



مبادئ الإحصاء





منشورات جامعة دمشق
كلية الاقتصاد

مبادئ الإحصاء

تأليف

د. فريد الجاعوني
أستاذ مساعد في قسم الإحصاء التطبيقي

د. منذر العواد
مدرس في قسم الإحصاء التطبيقي

د. عدنان عباس حميدان
أستاذ مساعد في قسم الإحصاء التطبيقي

د. عمار ناصر آغا
مدرس في قسم الإحصاء التطبيقي

منشورات جامعة دمشق



الفهرس

الصفحة

٣	 المقدمة
١٧	١- الفصل الأول : الإحصاء واستخداماته
١٩	١-١ مقدمة
٢٣	١-٢ استخدام علم الإحصاء في التحليل الاقتصادي
٢٣	١-٣ استخدام علم الإحصاء في إدارة الأعمال
٢٤	١-٤ استخدام علم الإحصاء في الدراسات المحاسبية
٢٤	١-٥ استخدام علم الإحصاء في مجالات التأمين وأسواق المال
٢٥	١-٦ استخدام علم الإحصاء في الدراسات السكانية (الديمغرافية)
٢٦	١-٧ استخدام علم الإحصاء في العلوم الإنسانية المختلفة
٢٩	أسئلة الفصل الأول
٣١	٢- الفصل الثاني : جمع البيانات والمخالفات الإحصائية
٣٣	٢-١ مقدمة
٣٤	٢-٢ جمع البيانات العددية
٣٤	٢-٢-١ المتغيرات
٣٥	٢-٢-١-١ متغيرات نوعية أو وصفية
٣٥	٢-٢-١-٢ متغيرات كمية
٣٦	٢-٢-٢ المجتمع الإحصائي
٣٧	٢-٢-٣ الدراسة التحضيرية
٣٧	٢-٢-٤ مصادر البيانات

٣٧	٢-٢-٤-١ البيانات الثانوية
٣٧	٢-٢-٤-٢ البيانات الأولية
٣٨	٢-٢-٤-٢-١ شكل الاستمارة وترتيبها
٣٨	٢-٢-٤-٢-٢ القواعد الواجب مراعاتها عند صياغة الأسئلة
٤١	٢-٢-٤-٢-٣ ملء الاستمارة الإحصائية
٤٤	٢-٢-٥ أسلوب جمع البيانات
٤٤	٢-٢-٥-١ أسلوب الحصر الشامل
٤٤	٢-٢-٥-٢ أسلوب العينة
٤٥	٢-٢-٥-٣ المفاضلة بين أسلوب الحصر الشامل وأسلوب العينة
٤٥	٢-٢-٥-٤ العينة وأنواعها
٤٨	٢-٢-٥-٤-١ أنواع العينات العشوائية
٥٢	٢-٢-٥-٤-٢ العينات الشخصية (غير العشوائية)
٥٢	٢-٢-٥-٥ أخطاء العينة
٥٧	٣- الفصل الثالث : عرض البيانات الإحصائية
٥٩	٣-١ مقدمة
٦٠	٣-٢ عرض البيانات النوعية (الوصفية)
٦٠	٣-٢-١ العرض الجدولي للمتغيرات النوعية
٦٤	٣-٢-٢ العرض البياني للمتغيرات النوعية
٦٤	٣-٢-٢-١ الأعمدة البيانية
٧٢	٣-٢-٢-٢ الرسوم الدائرية
٧٥	٣-٣ عرض البيانات الكمية
٧٥	٣-٣-١ العرض الجدولي للمتغيرات الكمية
٨٦	٣-٣-٢ العرض البياني للمتغيرات الكمية
٩٤	مسائل وتمارين الفصل الثاني والثالث

٩٧	٤- الفصل الرابع : مقاييس النزعة المركزية (المتوسطات)
٩٩	٤-١ مقدمة
١٠١	٤-٢ الوسط الحسابي
١٠١	٤-٢-١ حساب الوسط الحسابي لبيانات إحصائية غير مبوبة
١٠٦	٤-٢-٢ حساب الوسط الحسابي لبيانات إحصائية مبوبة في فئات
١١٠	٤-٢-٣ خصائص الوسط الحسابي
	٤-٢-٣-١ بعض الأساليب غير المباشرة في إيجاد قيمة
١١٠	الوسط الحسابي
١٢٨	٤-٢-٣-٢ بعض الخصائص الجبرية الأخرى للوسط الحسابي
١٣٢	٤-٢-٤ محاسن ومساوئ الوسط الحسابي
١٣٤	٤-٣ الوسيط
١٣٤	٤-٣-١ حساب الوسيط للبيانات غير المبوبة
١٣٦	٤-٣-٢ حساب الوسيط للبيانات المبوبة في فئات
١٤٢	٤-٣-٣ العشرية والربيعات والمئينات
١٤٧	٤-٣-٤ خصائص الوسيط ومحاسنه
١٥٥	٤-٤ المنوال
٢٥٠	٤-٤-١ حساب المنوال للبيانات غير المبوبة
١٥٠	٤-٤-٢ حساب المنوال للبيانات المبوبة في فئات
١٥٦	٤-٤-٣ محاسن ومساوئ المنوال
١٥٧	٤-٤-٤ العلاقة التقريبية بين الوسط الحسابي والوسيط والمنوال
١٦٢	٤-٥ الوسط الهندسي
١٦٣	٤-٥-١ حساب الوسط الهندسي للبيانات غير المبوبة

١٦٧	٤-٥-٢ حساب الوسط الهندسي للبيانات المبوبة في فئات
١٦٩	٤-٥-٣ محاسن ومساوئ الوسط الهندسي
١٧٠	٤-٦ الوسط التوافقي
١٧٠	٤-٦-١ حساب الوسط التوافقي للبيانات غير المبوبة
١٧٥	٤-٦-٢ حساب الوسط التوافقي للبيانات المبوبة في فئات
١٧٧	٤-٦-٣ محاسن ومساوئ الوسط التوافقي
١٧٩	٥- الفصل الخامس: مقاييس التشتت (التباين)
١٨١	٥-١ مقدمة
١٨٤	٥-٢ المدى المطلق
١٨٥	٥-٢-١ المدى النسبي
١٦٧	٥-٣ الانحراف الربيعي (نصف المدى الربيعي)
١٩١	٥-٣-١ الانحراف الربيعي النهائي
١٩٣	٥-٣-٢ محاسن ومساوئ الانحراف الربيعي
١٩٣	٥-٤ الانحراف المتوسط
١٩٤	٥-٤-١ حساب الانحراف المتوسط للبيانات المبوبة في فئات
١٩٦	٥-٤-٢ حساب الانحراف المتوسط للبيانات المبوبة في فئات
١٩٨	٥-٤-٣ الانحراف المتوسط النسبي
١٩٩	٥-٤-٤ محاسن ومساوئ الانحراف المتوسط
٢٠٠	٥-٥ الانحراف المعياري
٢٠١	٥-٥-١ حساب الانحراف المعياري للبيانات غير المبوبة
٢٠٧	٥-٥-٢ حساب الانحراف المعياري للبيانات المبوبة في فئات
٢١٣	٥-٥-٣ الانحراف المعياري النسبي (معامل الاختلاف)
٢١٥	٥-٥-٤ الدرجة المعيارية

٢١٦	٥-٥-٥ أهم خصائص الانحراف المعياري
٢١٨	٥-٥-٦ محاسن ومساوئ الانحراف المعياري
٢٢٠	تمارين ومسائل الفصل الرابع والخامس
٢٢٩	٦- الفصل السادس: مقاييس الالتواء والتفطح
٢٣١	٦-١ مفهوم الالتواء وأساليب قياسه
٢٣٧	٦-٢ فوائد مقاييس الالتواء
٢٣٧	٦-٣ مفهوم التفطح وأساليب قياسه
٢٣٩	٦-٤ فوائد مقاييس التفطح
٢٤١	تمارين ومسائل الفصل السادس
٢٢١	٧- الفصل السابع: الاحتمالات وبعض التوزيعات الاحتمالية
٢٤٥	٧-١ الاحتمالات
٢٤٥	٧-١-١ مقدمة
٢٤٥	٧-١-٢ بعض المفاهيم الأساسية في نظرية الاحتمالات
٢٥٢	٧-١-٣ العمليات على الحوادث
٢٥٨	٧-١-٤ التعاريف الأساسية للاحتمال
٢٥٨	٧-١-٤-١ التعريف الكلاسيكي أو التقليدي للاحتمال
٢٥٩	٧-١-٤-٢ التعريف الاحصائي أو التكرار النسبي للاحتمال
٢٦٠	٧-١-٤-٣ الاحتمال حسب مسلمات كولموروف
٢٦٠	٧-١-٥ القوانين الأساسية في الاحتمالات
٢٨٢	٧-٢ التوزيع الطبيعي
٢٨٢	٧-٢-١ مقدمة
٢٨٣	٧-٢-٢ تابع الكثافة الاحتمالي للتوزيع الطبيعي
٢٨٧	٧-٣ التوزيع الطبيعي المعياري

٢٨٩	٧-٣-١ تابع الكثافة الاحتمالي للتوزيع الطبيعي المعياري
٢٩٤	٧-٣-٢ تحديد قيم Z باحتمال معلوم
٢٩٦	مسائل وتمارين
٣٠٥	٨- الفصل الثامن : تحليل الارتباط والانحدار
٣٠٧	٨-١ الغاية من تحليل الارتباط والانحدار
٣٠٩	٨-٢ تحليل الارتباط
٣٠٩	٨-٢-١ الارتباط البسيط للمتغيرات الكمية
٣٠٩	٨-٢-٢ الارتباط للمتغيرات النوعية
٣٢٩	٨-٣ تحليل الانحدار
٣٣٠	٨-٣-١ الانحدار الخطي البسيط
٣٣٠	٨-٣-١-١ إيجاد معادلة مستقيم الانحدار
٣٣٧	٨-٣-١-٢ أهمية معادلة مستقيم الانحدار
٣٤١	٨-٣-١-٣ دراسة التباينات
٣٤٧	٨-٣-١-٤ التقدير والتنبؤ
٣٥١	مسائل وتمارين
٣٥٩	٩- الفصل التاسع : الأرقام القياسية
٣٦١	٩-١ مفهوم الرقم القياسي وأهميته
٣٧٢	٩-٢ تركيب الأرقام القياسية
٣٨٦	٩-٣ طرائق تركيب الأرقام القياسية
٣٨٧	رموز ومصطلحات
٣٨٩	٩-٣-١ طريقة المناسيب في تركيب الأرقام القياسية
٣٨٩	٩-٣-١-١ المناسيب البسيطة

٤٠٠	٩-٣-١-٢ المناسب المرجحة
٤١١	٩-٣-١-٣ خصائص المناسب
٤١٨	٩-٣-٢ طريقة السلسلة في تركيب الأرقام القياسية
٤٢٤	٩-٣-٣ الطريقة التجميعية في تركيب الأرقام القياسية
٤٣٨	٩-٣-٣-١ الطريقة التجميعية في تركيب الأرقام القياسية للأسعار
٤٣٨	٩-٣-٣-٢ الطريقة التجميعية في تركيب الأرقام القياسية للكميات
٤٤٦	٩-٤ اختبارات الأرقام القياسية
٤٤٧	٩-٤-١ اختبار الانعكاس في الزمن
٤٥٠	٩-٤-٢ اختبار الانعكاس في المعامل
٤٥٣	٩-٤-٣ الاختبار الدائري
٤٥٨	٩-٥ تعديل الأرقام القياسية
٤٦٤	٩-٦ الأرقام القياسية الشائعة الاستعمال
٤٦٩	تمارين ومسائل
٤٧٥	— بعض الجداول الإحصائية
٤٧٥	١- جدول التوزيع الطبيعي $\phi(z)$
٤٧٦	٢- جدول الأرقام العشوائية
٤٧٨	٣- جدول توزيع ستودنت (t)
٤٧٩	— بعض المصطلحات الإحصائية
٥٠٥	— المراجع العربية والأجنبية



المقدمة

أبرز التطور العلمي الحديث أهمية استخدام الأساليب الإحصائية في جميع المجالات التطبيقية، وجميع فروع المعرفة والعلوم مثل الفيزياء والكيمياء والزراعة والاقتصاد والمحاسبة والإدارة والإلكترونيات والاتصالات وعلم النفس والاجتماع والعلوم الطبية والصيدلانية وغيرها من العلوم الهندسية، وربما اعتمدت هذه العلوم في تطور مادتها المعرفية على تطور مادة الإحصاء واستخدام الطرائق الإحصائية. والهدف من هذا الكتاب تقديم بعض الأسس العامة لمبادئ علم الإحصاء حيث يبدأ كل فصل بمقدمة عن الهدف، إضافة إلى التعاريف والنظريات والأمثلة التي توضح الموضوعات المستهدفة لكل فصل.

إن الموضوعات التي تضمنتها فصول الكتاب تقع ضمن المنهاج الموضوع لهذا المقرر الموجه إلى طلاب السنة الأولى بكلية الاقتصاد بجامعة دمشق، كما أنه موجه إلى جميع المهتمين بمختلف المواضيع الهامة المتعلقة بالتحليل الإحصائي.

ويمتاز هذا الكتاب بطريقة عرضه وبتعدد الأمثلة المخلولة لجميع الفصول وتنوع التمارين والمسائل التي جاءت في نهاية كل فصل. أما مادة الكتاب فهي شاملة وتغطي موضوعات الإحصاء الوصفي وتعتبر مرجعاً أساسياً وهاماً للطلاب ولجميع المختصين والمهتمين باستخدام الطرائق الإحصائية والوسائل والأدوات الحديثة التي تساعد الباحث على الوصول إلى القرار السليم إزاء مشكلات تتحكم فيها العشوائية، من خلال استخدام طرائق جمع وتطبيق وعرض البيانات واستخدام مقاييس النزعة المركزية ومقاييس التشتت والالتواء والتفرطح، بالإضافة إلى

الاحتمالات وبعض التوزيعات الاحتمالية الهامة ودراسة الارتباط والانحدار والأرقام القياسية وغيرها.

وقد قسمنا موضوعات الكتاب إلى تسعة فصول :

تناول الفصل الأول منها التعريف بالإحصاء واستخداماته في مختلف المجالات العلمي والحياتية، ويتناول الفصل الثاني أساليب جمع البيانات وتطبيقها في حين الفصل الثالث تضمن موضوع عرض البيانات من خلال عرض المتغيرات الكمية والمتغيرات وفق أساليب العرض الجدولي والعرض البياني، ويبحث الفصل الرابع في مقاييس النزعة المركزية، أما الفصل الخامس فيعرض مقاييس التشتت وأهميتها في التحليل الإحصائي، بالإضافة إلى دراسة موجزة لمقاييس الالتواء والتفلطح في الفصل السادس، أما في الفصل السابع قد تناولنا موضوع الاحتمالات وبعض التوزيعات الاحتمالية الهامة .

وتعرض الفصل الثامن بشكل واسع إلى دراسة الارتباط والانحدار، والفصل التاسع ناقش بصورة شاملة ومفصلة أهمية الأرقام القياسية في التحليل الإحصائي والاقتصادي.

وفي نهاية الكتاب تم وضع الجداول والمصطلحات الإحصائية والمراجع العربية والأجنبية.

إن تأليف هذا الكتاب على الرغم من أنه عمل جماعي ومسؤولية مشتركة، إلا أن أمانة التأليف تقتضي تخصيص دور كل مؤلف على حده، وعلى هذا الأساس فقد قام الدكتور عدنان عباس حميدان بكتابة الفصل الأول والسادس والثامن والتاسع، في حين قام بكتابة الفصل السابع الدكتور فريد الجاعوني، بينما اختص الدكتور عمار ناصر آغا بكتابة الفصل الرابع والخامس، أما الدكتور منذر العواد فقد قام بكتابة الفصل الثاني والثالث.

جاء هذا الكتاب باللغة العربية في كل محتوياته ما عدا المعادلات والمصطلحات فقد وضعناها بالرموز اللاتينية، والسبب الأساسي في ذلك يعود إلى كثرة عدد وأشكال الرموز اللاتينية وكفايتها لترميز مختلف متطلبات تدريس هذا المقرر من جهة، ولتسهيل الاطلاع على المراجع الأجنبية في الموضوعات المتصلة بالإحصاء والعلوم الاقتصادية الأخرى.

لقد بذلنا كل الجهد لتقديم كل ما هو ضروري في علم الإحصاء لطلبة السنة الأولى بكلية الاقتصاد بجامعة دمشق، ولجميع الباحثين والمهتمين في هذا المجال والواقين بأن الرقم الإحصائي أصبح هذه الأيام بمثابة المرآة التي تعكس الواقع والحقائق، وطبيعة الظواهر والعمليات، وآلية تطورها. وهو يعالج كلاً من الطرائق الإحصائية الأساسية والجواب الإحصائية التي تهم الاقتصاديين والإداريين والمحاسبين، والفئات الأخرى التي لها علاقة وطيدة بالطرائق الإحصائية كالمصارف ومؤسسات التأمين والهيئات والمنظمات المتخصصة بالدراسات السكانية والاجتماعية وغيرها. وهو بذلك يعدّ كتاباً جامعاً لطلبة الاقتصاد والتجارة من جهة، ومن جهة أخرى يعتبر مرجعاً كاملاً للمختصين والمهتمين والعاملين في مجالات الإحصاء والتحليل الإحصائي.

آملين أن نكون قد وفقنا في إخراج هذا الكتاب بالشكل المناسب.

المؤلفون



الفصل الأول

الإحصاء واستخداماته

- مقدمة
- استخدام علم الإحصاء في التحليل الاقتصادي
- استخدام علم الإحصاء في إدارة الأعمال
- استخدام علم الإحصاء في الدراسات الخشبية
- استخدام علم الإحصاء في مجالات التأمين وأسواق المال
- استخدام علم الإحصاء في الدراسات السكانية (الديمغرافية)
- استخدام علم الإحصاء في العلوم الإنسانية المختلفة



1-1- مقدمة :

إن عملية رصد تطور الوعي الإنساني وملاحظة وإدراك جميع العمليات والملاحظات التي يتم جمع المعلومات والبيانات عنها بهدف دراستها وتحليلها تعتبر خطوة هامة وأساسية على صعيد التفكير العلمي السليم، حيث تعتبر الملاحظة أو المشاهدة بمثابة الخطوة الأولى للبحوث العلمية والتطبيقية التي يجب أن تتميز ببيانها بالموضوعية والثقة والشمولية، والمعلومات التي يتم جمعها تعتبر في كثير من الأحيان، كمواد البناء كما قال الإحصائي الروسي (كاوفمان) (إن اليد الإحصائية اللطيفة والناعمة يجب أن تشيد بناءً متيناً وقوياً)، وهذا الشكل يمكننا القول إن المشاهدة (الملاحظة) تعتبر أساس الإدراك العلمي من جهة، ومن جهة ثانية تعتبر من الشروط الضرورية للتأثير الفعال لماهية تطور الظواهر تحت تأثير القوانين الطبيعية النازمة التي تخضع لها، والتي تعتبر مقدمات هامة لبلوغ العلوم إلى أهدافها السامية.

إن الأساس الشامل للمشاهدة يمكن أن يتفق مع التأمل الحي في عملية وعي إنسانية، حيث تسمى المشاهدة الإحصائية بعملية جمع المعلومات الضرورية حول ظواهر وعمليات الحياة الاقتصادية والاجتماعية التي تبدو دائمة الحركة وسريعة التغير والتطور. فقد شهد العالم ومنه بلدنا سورية تطوراً ملحوظاً في الحياة الاقتصادية والاجتماعية نتيجةً للتقدم التقني واستخدام وسائل التكنولوجيا والتقنية الحديثة، حيث اعتلى الحاسوب والانترنت عرش مجالات الحياة جميعها. لذلك كان لا بُدَّ من الاعتماد على علم الإحصاء بهدف الرقابة على العمل وتوعّي الدقة في جميع المعلومات وتبويبها وتصنيفها وتحليلها وصولاً إلى النتائج المرجوة .

فالحاجة إلى الإحصاء تزداد يوماً بعد يوم لمواكبة التطور الاجتماعي والاقتصادي والتقني، ولاستعماله في رسم السياسات واتخاذ القرارات المبنية على أسس علمية سليمة. وهذا ما يستدعي إلى المعرفة الدقيقة لجميع المتغيرات العشوائية التي تؤثر سلباً أو إيجاباً في العملية الاقتصادية، وبالتالي التنبؤ بحجم الظواهر المدروسة في المستقبل. ومن هنا نرى أن علم الإحصاء ببياناته وقوانينه لا يعتبر علماً حديثاً، إنه وجد بعد أن بدأ الإنسان يمتلك ويحصى أملاكه ويخاف عليها، وتطور مع تطوره، فقد استعان العرب بالأرقام الحسابية لإحصاء ما ينتجون وما يملكون. وهذا يعني أن ارتباط علم الإحصاء بالمظهر الاقتصادي الاجتماعي وثيق لأن دراسة الإنتاج بشكله ووسائله وأسلوبه والقوى المنتجة والعلاقات الاقتصادية والاجتماعية الناشئة لم يكن وليد عصر قريب لكنه ولد مع ولادة الحضارات والمجتمعات الإنسانية.

هذا التطور وتسارع هذا التقدم جعل الإحصاء يغدو أداة بحث علمي يخدم جميع الباحثين في شتى الميادين العلمية والاقتصادية والاجتماعية... إلخ، ولا غنى لباحث عنه لأن التعليم لا يهدف إلى تخريج موظفين وفنيين فحسب، بل يهدف إلى خدمة التنمية الاقتصادية والاجتماعية بالاعتماد على العلم والتكنولوجيا وعلم الإحصاء وفق هذا المفهوم هو ذلك العلم الذي يختص بالطرائق العلمية لجمع البيانات وتنظيمها، وعرضها وتحليلها، بغية الوصول إلى النتائج المتوخاة التي تؤدي إلى قرارات سليمة، فهو علم ابتكار وتطوير، يستند إلى منطق سليم مدعوم بأساليب رياضية متعددة، وهو بذلك قد أصبح علماً مستقلاً له تفرعات وتشعبات هامة كالإحصاء الرياضي والإحصاء الاقتصادي والإحصاء التطبيقي والسكاني والصناعي والزراعي والتجاري... وقد أضحت هذا العلم المستقل بذاته الأداة الفعالة في كل التخصصات بعد الذي شهدته العالم من تطورات الظواهر والعمليات المختلفة التي يمكن دراستها باستخدام المعلومات التاريخية بأساليب وأدوات إحصائية توضح مسار كل ظاهرة، وتسهل التنبؤ المستقبلي لها، وبذلك يكون هذا العلم بمثابة اليد اللطيفة والأمانة والناعمة المهمة برسم وتنفيذ السياسة الإغاثية للمجتمع، التي تقول بأنه يجب النظر إلى النمو الاقتصادي من باب

علاقته الوثيقة بالتحول الاجتماعي وتحويل الكم إلى كيف ونوع في المجتمع. لذلك لا بُدَّ للإنسان قبل قيامه بأي مشروع تنموي، من أن يخطط له، ويتوصل إلى معرفة تأثيره في البنيان الاجتماعي، ثم الاقتصادي، لأنَّ المقاييس الاجتماعية تساهم في الحكم على مشاريع اقتصادية.. والسياسة الإنمائية تعتمد على المشاريع المتوجهة وجهة تقنية واقتصادية، لينمو المجتمع نمواً اقتصادياً، ويتطور تطوراً اجتماعياً، وليتحقق ذلك لا بُدَّ من البحث الدقيق، والاعتماد على العلوم الاجتماعية والاقتصادية التي تُوضِّح نتائجها في بيانات دقيقة وجداول مفصلة تعتمد على علم الإحصاء الذي يعتبر مرتكزاً لعلم الاقتصاد في دراسة جوانبه، وتكامله، وإيجاد نتائجه، وبخاصة بعد أن تطور كثيراً من اقتصاد زراعي حربي بدائي إلى اقتصاد صناعي. وما نراه اليوم من اعتماد على تقنيات حديثة في الزراعة والصناعة لم يُن على فراغ، بل اعتمد على أسس علمية من تباين وارتباط والتحدار وتحليل التباين... قوامها علم الإحصاء الذي تجاوز مرحلة بسيطة في تشكيلها. فالمشاريع التنموية للري والزراعة والاستثمار لا يمكن أن تكون ذات جدوى اقتصادية ما لم تعتمد الإحصاء وبياناته وقوانينه مرتكزاً لها وأساساً للتخطيط ليصار إلى تحسين المحاصيل والإنتاج الزراعي والحيواني اللذين يشكلان مع جوانب اقتصادية أخرى ما يسمى بالدخل الوطني أو الدخل القومي، الذي يعتمد في دراسته خلال فترات زمنية متعاقبة على مؤشرات اقتصادية هامة تستند بدورها على مؤشرات إحصائية تدرس الهيكل الإحصائي وتطوره في بلد ما، لذلك لا تقتصر مهمة رجل الإحصاء الاقتصادي في أي بلد من البلدان على جمع المعلومات الإحصائية عن تطور وتركيب الدخل الوطني والنتائج الاجتماعي فحسب، وإنما تتناول أيضاً معرفة طرائق الحصول على هذه المؤشرات والأسس النظرية التي بُنيت عليها والغاية المرجوة من جمع هذه المؤشرات التي ستستخدم في البحث والتحليل الاقتصاديين، فالإحصاء في هذا المجال يقوم بتصنيف النشاطات الاقتصادية، وعلاقتها مع الاقتصاد الوطني. ومثال ذلك: «دراسة العمليات الاقتصادية في بلد ما ودراسة تركيب الاقتصاد الوطني والعلاقات المتبادلة بين عناصره، والمقارنة مع المؤشرات المقابلة في بلدان أخرى، ووضع نظام الموازين الاقتصادية

لاستخدامها في التخطيط والتحليل الاقتصاديين، ودراسة المنتج الاجتماعي والدخل الوطني... » .

وللوصول إلى تصنيف علمي للاقتصاد لا بُدَّ من الاعتماد على تصنيف إحصائي فيه تفصيل للقطاعات الاقتصادية بشكلٍ دقيقٍ حسب التقسيم الاجتماعي للعمل الذي نعتبره أساساً لتركيب الاقتصاد الوطني، وتصنيف النشاطات الاقتصادية وفق التركيب القطاعي للدخل الوطني، كما هو الحال في الجمهورية العربية السورية.

فالإحصاء هو اللبنة الأساسية في التصنيف العلمي القطاعي للاقتصاد لدراسة الدخل الوطني الذي يتأثر تحديده بسياسة الأسعار والتسعير، والتي بدورها تؤثر في المتغيرات الاجتماعية والاقتصادية، والتي تتأثر وتؤثر في السياسات الإنمائية والتنمية للاقتصاد الوطني.

إن عملية التنمية الشاملة تتطلب توافقاً بين السياسات التنموية الاقتصادية والاجتماعية والسياسات السكانية، وعلاقتها بالموارد الاقتصادية وهذا يحتاج إلى التخطيط الاقتصادي والاجتماعي من أجل تحديد معدلات الطلب على خدمات التنمية الشاملة وتقديم مؤشرات ومقاييس هامة تعكس طبيعة هذا التوافق ومتطلبات التخطيط للتنمية الاقتصادية والاجتماعية.

مما تقدم يمكننا أن نستنتج أن الإحصاء علم قائم بذاته لأنه يتميز بمجموعة من الخصائص والميزات التي تجعله مستقلاً عن غيره من العلوم، ففي حين أن علم الرياضيات يبحث في القوانين المجردة بشكل عام، نجد أن علم الإحصاء يدرس الناحية الكمية للظواهر الاقتصادية والاجتماعية، فهو ذو صفة تطبيقية، وبالتالي يستخدم في جميع المجالات العلمية والحياتية، ويجعل من جملة المفاهيم الأساسية والمعرفية للعلوم المختلفة مؤشرات ومعايير علمية تعكس المادة المعرفية لهذه العلوم. بمعنى آخر يحول علم الإحصاء مفاهيم مختلف العلوم إلى قالب رقمي قياسي قابل للعد والملاحظة، ويعكس جوهر الظواهر والعمليات الخاصة بهذه العلوم ومدى تغيرها وتطورها.

لذلك نلاحظ أن مجال تطبيق علم الإحصاء يشمل الاقتصاد والإدارة والمحاسبة وعلم الحياة وعلم الوراثة وعلم الاجتماع والعلوم الزراعية والطبيعية والعلوم التربوية وغيرها. لهذا نجد أن مادة الإحصاء أصبحت مادة هامة تدرس في مختلف الكليات العملية والنظرية في الجامعات السورية، وسوف نتناول بإيجاز استخدامات الإحصاء في المجالات الاقتصادية والاجتماعية وغيرها.

2-1- استخدام علم الإحصاء في التحليل الاقتصادي

يستخدم علم الإحصاء في معظم الدراسات الاقتصادية التي تهدف عادة إلى التنبؤ والتخطيط، سواء أكانت هذه الدراسات على مستوى مشروع صناعي (Micro-Level) أم على مستوى الاقتصاد القومي، حتى أصبحت المؤشرات والمقاييس الإحصائية من أهم الوسائل العلمية اللازمة في التحليل الاقتصادي، فدراسة الأسعار والأجور والاستثمار والادخار والاستهلاك والتصدير والاستيراد، وأي متغير آخر من المتغيرات الاقتصادية، أصبحت تعتمد إلى درجة كبيرة على الأسلوب الاستقرائي أو الاستدلالي الإحصائي والأسلوب الكمي في التحليل الاقتصادي. أما بالنسبة للأسلوب الأول فإننا سوف نناقشه في السنة القادمة في مقرر الإحصاء التطبيقي، في حين أن الأسلوب الثاني الذي يستخدم المعلومات العددية والمعطيات الإحصائية في الدراسات الاقتصادية فإننا سوف نناقشه في هذا الكتاب .

3-1- استخدام علم الإحصاء في إدارة الأعمال

يلعب علم الإحصاء دوراً هاماً في مختلف النواحي الخاصة بإدارة الأعمال لما يقدمه من أساليب إحصائية تعتبر بمثابة منهج علمي، وطريقة للبحث في إطار إدارة الأعمال، حيث يساهم الإحصاء في دراسة اتجاهات المبيعات وبيان الآثار الموسمية والدورية لها بهدف إعداد الخطط المستقبلية للمؤسسات والشركات المدروسة. بالإضافة إلى تتبع حجم المخزون من المنتجات والمواد الخام والوقود، وذلك بهدف المحافظة على حجم معين من المخزون. كما أن الإحصاء يستخدم بشكل كبير في دراسة احتياجات

المستهلكين ورغباتهم وأذواقهم بهدف توافق الرغبات مع الأذواق وبالتالي دراسة الأسواق عن طريق بحوث التسويق كما يستخدم الإحصاء في تحديد درجة الجودة والكمية والرقابة على الجودة، حيث تلعب نظرية العينات والاعتبارات المرتبطة بها دوراً هاماً في تنظيم الرقابة على الجودة وبالتالي تقدير المبيعات المستقبلية في إعداد الموازنات الخاصة بالإنتاج والتكاليف والمخزون من المواد المختلفة.

4-1- استخدام علم الإحصاء في الدراسات المحاسبية

لاحظنا أن جميع العلوم تستخدم بشكل أو بآخر الأساليب الإحصائية التي تجعل من علم الإحصاء قاسماً مشتركاً بين العلوم المختلفة، وإن اختلفت فروعها تختلف في مدى التعمق باستخدام علم الإحصاء الذي يرتبط بصفة وثيقة بالحياة اليومية والعملية، ودراسة البيانات وتحليلها واتخاذ القرارات المختلفة. لهذا نجد أن علم الإحصاء يستخدم بشكل واسع في مجال الأعمال والدراسات المحاسبية مثل تحليل القوائم المالية، والبحث عن مؤشرات ومعايير محاسبية، والتنبؤ بالأرباح التجارية أو الصناعية، وتحديد القيمة المضافة على صعيد المؤسسة أو القطاع الصناعي بشكل عام، بالإضافة إلى ذلك تستخدم الطرائق الإحصائية وبشكل خاص نظرية العينات في بحوث المراجعة التي تعتمد اعتماداً أساسياً على تطبيق طرائق المعاينة في حالات التدقيق والبحث والاكتشاف والتقدير والمعالجة.

5-1- استخدام علم الإحصاء في مجالات التأمين ودراسات أسواق المال

هناك عدد من الاعتبارات التي تجعل الإحصاء ذا أهمية كبيرة في مجال العلوم المالية والمصرفية وهي أن الظواهر والعمليات الخاصة بهذه العلوم تتميز بالتغير الدائم الذي يجعل التنبؤ بها يتطلب أساليب إحصائية متطورة وعلمية بالإضافة إلى السلوك البشري المعقد الذي يشكل أحد العوامل الهامة في أثناء دراسة الأسواق المالية والتغيرات التي تتعرض لها أو خلال دراسة القروض المصرفية التي تمنحها المصارف في دولة ما في فترة زمنية معينة تبعاً لفترات آجالها وللأغراض التي منحت من أجلها. وهناك العديد من الظواهر التي يمكننا دراستها إحصائياً مثل : النقد المتداول - المبيعات - المخزون -

الودائع - القروض - رؤوس الأموال - أسعار الفائدة - أسعار الصرف - الصادرات - الواردات - الدخيل القومي، وفي أسواق البورصة : حركة الأوراق المالية - أسعار الأسهم - أسعار السندات - أسعار فائدة الأسهم أو السندات - معدل العائد على الاستثمار في محافظة الاستثمار - حجم الأموال المستثمرة في المحافظ المالية... إلخ، حيث تحتاج كل ظاهرة من هذه الظواهر إلى البيانات والتحليل باستخدام الأساليب الإحصائية والرياضيات الاكتوارية في دراسات التأمين.

6-1- استخدام علم الإحصاء في الدراسات السكانية (الديمغرافية)

يستخدم علم الإحصاء بشكل أساسي في مجالات علم السكان والدراسات السكانية، بهدف التعبير عن كثير من المؤشرات السكانية كالولادات والوفيات والهجرة، وتطورها عبر الزمن، فقد أصبحت الأبحاث السكانية تعتمد بشكل رئيسي على الدراسات الكمية سواء أكانت تتعلق بالتعدادات السكانية والواقعات التي تحدث خلال فترة حياة الإنسان أم بمجال الدراسات الحيوية، وتهدف إلى دراسة العناصر التي تتألف منها المجتمعات السكانية والتعرف على خصائصها العديدة من حيث توزيع السكان حسب العمر والجنس والحالة الزوجية ومكان الولادة ودرجة التعليم والنشاط الاقتصادي والهجرة... إلخ، بالإضافة إلى دراسة آلية تحديد هذه المجتمعات السكانية عن طريق التحليل الإحصائي للعلاقات القائمة بين حوادث الولادة والوفاة والخصوبة والطلاق والهجرة وعدد السكان... إلخ. وما الإحصاءات الحيوية سوى الإحصاءات الخاصة بالأمور المهمة من حياة الإنسان من حيث أنه كائن حي منذ ولادته إلى وفاته، وبذلك فهي تبحث في حالة السكان وتكوينهم وحركتهم من حيث الزيادة والنقصان والحوادث الهامة التي تقع لهم والمذكورة أعلاه.

إن القياس الإحصائي يساعد في فهم الظواهر السكانية وتحضيرها للتحليل العددي والبحث عن العلاقات السببية وتحديد العوامل التي تؤثر فيها وتتأثر بها. وتعتبر المقاييس والمؤشرات الإحصائية أساس الأبحاث السكانية المختلفة، لهذا نرى أن الديمغرافي يسمى باستخدام الطرائق الإحصائية إلى دراسة الحوادث السكانية من

ولادات وزواج وطلاق ووفيات وأفراد، أعداً بالاعتبار الزمن كمستفرد أساسي في قياس العمر وفترة الزواج وفترة الترميل ومدة الدراسة وتوقيت الهجرة وسن الحمل وعدد وفيات الأجنة وعدد المواليد... إلخ عن طريق استخدام جداول المواليد أو جداول الحياة التي تعتبر من الجداول الإحصائية الهامة صعبة التركيب، وهكذا فإن الإحصاء يستخدم في جميع مراحل البحث السكاني، ولا يمكن التوصل إلى نتائج هامة ومفيدة ما لم يتمكن الديمغرافي من استخدام الأساليب الإحصائية بكفاءة عالية.

ويسمى علم الإحصاء في الدراسات الديمغرافية⁽¹⁾ بالإحصاء الديمغرافي (السكاني) الذي يشمل بيانات تعدادات السكان Concusses لمعرفة عدد السكان في لحظة معينة، ومعرفة توزيعهم في تلك اللحظة ذاتها حسب الخصائص والأوصاف المختلفة، من حيث العمر والنوع والخصائص الأخرى الاجتماعية والاقتصادية والعائلية وغيرها من الخواص، كما يشمل الإحصاءات الحيوية Vital Statistics التي تصف مجموعة الأحداث والحوادث التي تصيب الإنسان منذ ولادته حتى وفاته، بالإضافة إلى إحصاءات الهجرة Emigration Statistics كمصدر ثالث للإحصاء الديمغرافي.

7-1- استخدام علم الإحصاء في العلوم الإنسانية المختلفة

نظراً للاستخدامات الكبيرة والهامة لعلم الإحصاء في مجال العلوم الإنسانية أصبح فرعاً من فروع هذه العلوم فيسمى في علم الاجتماع بالإحصاء الاجتماعي، وفي علم الجغرافيا بالتحليل الكمي، وفي علم السكان بالإحصاء السكاني، وفي علم المكتبات والإعلان بالإحصاء التطبيقي، وفي علم التربية وعلم النفس بالإحصاء التطبيقي وأحياناً بالإحصاء اللا معلمي وذلك انطلاقاً من أهمية الوسائل الإحصائية في تحليل ودراسة ظواهر هذه العلوم التي تعتمد على وضع الفروض والاختبارات الإحصائية، وهذه سمة

(1) كلمة الديمغرافيا Demography هي من أصل يوناني وتتألف من شقين الأول Demos (البشر أو الناس) ، والثاني Graphia بمعنى (وصف) وعلى ذلك فالإحصاء الديمغرافي يهدف إلى وصف البشر مقيمين كان هؤلاء البشر أم متنقلين .

مميزة لأبحاث هذه العلوم، حيث أصبح تطور العلوم الإنسانية مرتبطاً بتطبيق هذه الأساليب والأدوات الإحصائية المختلفة. فلاحظ أن الرقم الإحصائي الذي يعبر عن مؤشرات سكانية معينة كالولادات والوفيات والهجرة تساعد الدراسات السكانية ليكون لها مكان متقدم بين العلوم الاجتماعية الأخرى، حتى أصبحت الأبحاث السكانية تعتمد على الدراسات الكمية من خلال دراسة العلاقة بين مختلف المتغيرات وتأثير بعضها على البعض الآخر بدراسة معادلات التعداد المتغيرات التابعة على المتغيرات المستقلة. بالإضافة إلى ذلك هناك التخطيط التربوي بوظائفه المتعددة التي تلخص بتوفير التعليم المناسب لسكان المجتمع وعصائص هؤلاء السكان من خلال أعمارهم وجنسهم، وسكنهم في الريف أو المدينة. فالباحث التربوي يحتاج في موضوع بحثه إلى تنظيم القيم ليستطيع استخراج دلالتها، أو يحتاج للمقارنة بين الأفراد والفئات أو يتعرف على التغير الحاصل بين القيم في جدول القيم، أو معرفة تشتت هذه القيم أو العلاقة التي تربط بين المتغيرات، أو إذا أراد دراسة ظاهرة تربوية ما ليصل إلى ما هي عليه بين أفراد مجموعة واسعة من الطلبة، وأراد تعميم نتائجه، فإنه يعتمد مجموعة صغيرة تمثل عينة ليدرسها ويعمم نتائجها، وإن كل ذلك لا يتم إلا باستخدام مفاهيم الإحصاء من جداول تمثل توزيع القيم وتنظيمها، ورسوم بيانية تمثل المدرجات التكرارية أو المضلعات أو المنحنيات التكرارية، ولا يكفي بالجدول بل يعتمد مقاييس النزعة المركزية للمقارنة بين أفراد مجموعة غير معينة أو محددة، وذلك باعتماد متوسط القيم وتكرارها المحسوب أو باعتماد حساب الوسيط أو المنوال، هذه الحسابات للقيم تمكن الباحث من القيام بالمقارنة بين الأفراد والفئات، فإذا ما أراد الباحث معرفة التغير الحاصل بين القيم أي معرفة تشتت القيم من خلال النظر إلى مركزها، فإنه يستطيع استخدام مقاييس التشتت، وإذا أراد معرفة العلاقة بين متحولين، فإنه يقوم بحساب معامل الارتباط أو إيجاد الانحدار لاستخراج قيمه، بمتغير من قيمة في متغير آخر بينهما ترابط.

إنها مقاييس إحصائية تسهل مهمة الباحث في العلوم الإنسانية، وتزيل العقبات التي تعترضه في أثناء قيامه بالبحث، وفي إيجاد الحلول العلمية والمنطقية السليمة للمسائل

الأساسية المنبثقة من طبيعة وجوهر هذه العلوم من جهة ومن قانونية تطورها من جهة أخرى.

بالإضافة إلى ما سبق يستخدم الإحصاء بشكل واسع في فروع العلم المتصلة بالعلوم البيولوجية «Biological Sciences» مثل الطب والزراعة والصيدلة وعلم الوراثة وأخيراً في العلوم الطبيعية وفي الهندسة وغيرها، لدرجة يصعب معها وجود أي مجال من مجالات البحث والدراسة لم يستخدم فيها علم الإحصاء في الوقت الذي لم يكن فيه الإحصاء في مراحله الأولى أكثر من مجرد جمع للبيانات وتصنيفها وعرضها، حيث يسمى هذا الأسلوب في التعامل مع البيانات بالإحصاء الوصفي الذي سوف نناقشه في هذا الكتاب، لأن هدفه الأساسي هو وصف خصائص البيانات ووصف ملاحظاتها الأساسية، أي إن الإحصاء الوصفي هو ذلك النوع من الدراسة الذي يتضمن جمع وتنظيم وعرض ووصف البيانات المتاحة، في حين أن النوع الآخر الذي يتضمن اتخاذ القرارات وتعميم النتائج بناءً على المعلومات الجزئية التي نحصل عليها باستخدام الأساليب الوصفية يسمى بالإحصاء الاستدلالي الذي سوف نناقشه في السنة الثانية بمقرر الإحصاء التطبيقي.

وهنا تجب الإشارة إلى أن الإحصاء الوصفي والإحصاء الاستدلالي هما فرعا علم الإحصاء الحديث، وإن كليهما ضروري للأبحاث واتخاذ القرارات السليمة، على الرغم من تزايد أهمية الإحصاء الاستدلالي، وعلى وجه التحديد فإن الإحصاء يشير إلى طرائق جمع وتنظيم وعرض وتحليل البيانات، بالإضافة إلى اتخاذ القرارات وإجراء التعميمات عن خصائص جميع مفردات الظاهرة الخاضعة للدراسة، وبالتالي يمكننا تحديد مفهوم علم الإحصاء بأنه العلم العملي الذي يسجل بطرائق مختلفة واقع الأمر. لظاهرة ما أو عملية أو شيء، ويجسد هذه الوقائع في قالب رقمي قياسي يبرز اتجاهاتها وعلاقة بعضها ببعض وذلك بجمعها وتصنيفها وتبويبها وعرضها وتحليلها واتخاذ القرارات السليمة وتعميم النتائج وتفسيرها، أي إنه يهتم بجمع وتنظيم وتلخيص وعرض البيانات العددية والرقمية، ثم تحليل النتائج واتخاذ القرارات في ضوء التحليل.

أسئلة الفصل الأول

- 1-1- عرف كلاً من المقولات الإحصائية الآتية :
المشاهدة الإحصائية، علم الإحصاء، النمو الاقتصادي، السياسة الإنمائية، التصنيف الإحصائي، الدخل الوطني، التنمية الشاملة، الرقم الإحصائي.
- 1-2- بين أهمية استخدام علم الإحصاء في الدراسات الاقتصادية والتحليل الاقتصادي.
- 1-3- وضح العلاقة العلمية التي تربط بين علم الإحصاء والدراسات المحاسبية.
- 1-4- اذكر الاعتبارات العلمية التي تجعل الإحصاء ذا أهمية كبيرة في مجال العلوم المالية والمصرفية وعلوم التأمين .
- 1-5- بين أهمية استخدام علم الإحصاء في مختلف العلوم الإنسانية والاجتماعية مدعماً بإحاثتك ببعض الأمثلة.
- 1-6- وضح أهمية الإحصاء والطرائق الإحصائية في دراسة الظواهر والعمليات الخاصة بالعلوم البيولوجية مثل الطب والصيدلة والزراعة وغيرها.
- 1-7- وضح باختصار دور المقاييس الإحصائية، كمقاييس النسبة المركبة، ومقاييس التشتت والمقاييس الخاصة بالارتباط في وصف خصائص الظواهر المدروسة وتحديد ملامحها الأساسية.



الفصل الثاني

جمع البيانات والحقائق الإحصائية

- مقدمة
- جمع البيانات العددية
- المتغيرات
- متغيرات نوعية أو وصفية
- متغيرات كمية
- المجتمع الإحصائي
- الدراسة التحضيرية
- مصادر البيانات
- البيانات الثانوية
- البيانات الأولية
- شكل الاستمارة وترتيبها
- القواعد الواجب مراعاتها عند صياغة الأسئلة
- ملء الاستمارة الإحصائية
- أسلوب جمع البيانات
- أسلوب الحصر الشامل
- أسلوب العينة
- المفاضلة بين أسلوب الحصر الشامل وأسلوب العينة
- العينة وأنواعها
- أنواع العينات العشوائية
- العينات الشخصية (غير العشوائية)
- أخطاء العينة



1-2- مقدمة :

تعتبر عملية اتخاذ القرار وعلى جميع المستويات والمجالات سواء أكانت سياسية أم اقتصادية أم إدارية أم عسكرية... من العمليات متكررة الحدوث كثيراً، وفي أغلب الأحيان تكون هذه القرارات هامة جداً سواء أكانت على مستوى فرد أم مؤسسة أم دولة، حيث سيكون لهذه القرارات تأثير على قرارات أخرى وأعمال أخرى. من هنا يجب توخي الحذر وبذل قصارى جهد ممكن قبل اتخاذ مثل هذه القرارات. وبرأينا إن القرار الصحيح هو ذلك القرار المبني على دراسة وتحليل الواقع بشكل علمي ومنطقي صحيحين، كما أن هذا التحليل يحتاج إلى معلومات تعتمد بدورها على بيانات تشكل المادة الخام في تصنيع وتوفير هذه المعلومات، حيث نحصل على المعلومات لاتخاذ قرار ما أو لإجراء بحث ما من معارفنا المتراكمة، ولكن نادراً ما يلم شخص ما بجميع المعلومات، فقد يكون بحاجة إلى جمع البيانات عن الموضوع أو المشكلة قيد البحث ليحصل على معلومات، يتضح أن الغاية الأساسية هي ليست في جمع البيانات، ولكن في الحصول على معلومات لاتخاذ القرار، وهذا يتم عن طريق تحليل البيانات التي تم جمعها وترتيبها بشكل يسهل تحليلها. من هنا نرى أن تلك العملية تمر بعدة خطوات هي بعد ذاتها خطوات البحث العلمي :

- 1 - تحديد المشكلة .
- 2 - جمع البيانات الإحصائية عن المشكلة.
- 3 - عرض البيانات الإحصائية .
- 4 - تحليل البيانات .

وبالاعتماد على هذه الخطوات يمكننا تعريف الإحصاء بأنه العلم الذي يقدم البيانات لاتخاذ القرار وذلك من خلال جمع البيانات العددية وعرضها وتحليلها، وإن هذا التعريف لعلم الإحصاء كان نتيجة إدخال الرياضيات ونظرية الاحتمالات واستخدام الحاسب الإلكتروني في جميع مراحل البحث الإحصائي. وإننا سوف نتناول في هذا الكتاب هذه الخطوات بالتفصيل بالنسبة لعملية جمع البيانات وعرضها، أما التحليل فسوف نتناول جزءاً منه في هذا الكتاب ونترك بعض أساليب التحليل للكتب الأخرى.

2-2- جمع البيانات العددية : Collection of Numerical data

تعدّ عملية جمع البيانات العددية من أهم مراحل البحث الإحصائي، ونحتاج إلى الكثير من المال والوقت بالمقارنة مع المراحل التالية التي يمكن إنجازها بوقت قصير جداً، باستخدام الحاسب الإلكتروني، لذلك يجب في هذه المرحلة بذل قصارى الجهد والعناية لتكون البيانات التي سنقوم بجمعها صحيحة، وبالتالي تكون مراحل البحث التالية والقرار المتخذ صحيحة أيضاً، وأنا نرى أنه للحصول على هذه البيانات - والتي يمكن اعتبارها صحيحة - من الضروري الاستئناس ببعض المبادئ هنا والتي سنتعرف عليها بالتفصيل :

2-2-1 المتغيرات : Variables

يعرف المتغير بأنه الصفة أو الخاصية التي يمكن أن تتغير من مفردة إلى أخرى سواء أكان التغير كمياً أم نوعياً ، وأن المفردات يمكن أن تشترك في صفة ما أو عدة صفات بالوقت نفسه ولكنها تختلف هذه الصفات من مفردة لأخرى.

فأية مجموعة من الناس تحكمها المتغيرات التالية: أهم بشر - يختلفون من حيث الطول والوزن والجنس ومستوى المعيشة ومكان الإقامة ومتوسط الإنفاق ومتوسط الدخل. وهذا ما يؤكد أهمية التعرف على المتغيرات من أجل جمع البيانات، فبعد التعرف على المتغير يقوم الباحث بجمع البيانات عن المفردات التي تتمتع بهذه الصفة ويستبعد المفردات الأخرى، وبالتالي يكون قد اختصر البيانات إلى البيانات الضرورية. وتقسّم المتغيرات إلى قسمين :

1-1-2-2 - متغيرات نوعية أو وصفية Qualitative Variables :

وهي تلك المتغيرات التي تتمتع بها المفردة بصفة معينة، ولا يمكن لها أن تتمتع بما بدرجات مختلفة، ومعنى آخر هي تلك المتغيرات التي يعبر عنها بكلمة أو جملة مثل: ذكر - أنثى - طاولة - مدرسة - جامعة - إلخ .

2-1-2-2 - متغيرات كمية Quantitative Variables :

وهي تلك المتغيرات التي تتمتع بها المفردات بدرجات مختلفة، أو هي المتغيرات التي يمكن قياسها بوحدات القياس - الوزن - الطول - الحجم - الزمن - القيم النقدية إلخ. وتقسّم إلى قسمين:

- متغيرات مستمرة Continuous Variables :

وذلك عندما يأخذ المتغير أية قيمة ضمن مجال معين فطول إنسان ما يمكن أن يأخذ أية قيمة ضمن المجال 150 سم و 200 سم .

- متغيرات منقطعة Discrete Variables :

وذلك عندما يأخذ المتغير قيمة صحيحة، فعدد أفراد الأسرة هو 2 أو 3 ... وعدد العمال هو 100 أو 200 إلخ .

وأخيراً نقول يمكن أن تكون المتغيرات بسيطة عندما نعبر عنها بصفة واحدة وهنا يمكن أن نستخدم الوحدات الآتية :

1 - وحدات العد .

2 - وحدات القياس .

3 - الوحدات النقدية .

ويمكن أن تكون المتغيرات مركبة، وذلك عندما نعبّر عنها بأكثر من صفة مثل سرعة السيارة في وحدة الزمن، واستهلاك السيارة في وحدة المسافة ... إلخ .

2-2-2 - المجتمع الإحصائي Statistical Population :

المجتمع الإحصائي هو جميع المفردات التي تتمتع بصفة ما أو خاصية ما (متغير) وهذه المفردات قد تكون بشراً أو أشياء أو ظواهر، وهذا ما يميز المجتمع الإحصائي عن غيره بالمعنى الاجتماعي والذي هو مجموعة من الناس تربط بينهم علاقة ما، لذلك لا بد من تحديد المجتمع الإحصائي بهدف جمع البيانات العددية الضرورية عنه، ويتحدد المجتمع الإحصائي بتحديد المكان والزمان معاً، وهذا نرى أن مفهوم المجتمع الإحصائي مفهوم مرن حيث يمكن أن يوسع المجتمع، أو العكس بتوسيع المكان أو الزمان أو كليهما .

وكأمثلة على المجتمع الإحصائي: أسر مدينة دمشق عام 2000 ، وعمال شركة بردي عام 2001 ، وطلاب جامعة دمشق عام 2001 ، وإنتاج معمل ألبيسة الشرق عام 2000، نرى من جميع هذه الأمثلة أن هناك تحديداً للمجتمع من حيث الزمان والمكان.

- وقد يكون المجتمع الإحصائي محدوداً Finite وذلك عندما نستطيع حصر عدد مفردات المجتمع، وبالتالي جمع بيانات عنها مثل طلاب جامعة دمشق - سكان مدينة دمشق .

- وقد يكون المجتمع الإحصائي غير محدود Infinite، وذلك عندما لا نستطيع حصر عدد مفرداته: أسماك البحر الأبيض المتوسط أو الاحتياطي من النفط ، وهذه المجتمعات يطلق عليها عادة تسمية مجتمعات افتراضية Hypothetical Populations .

3-2-2-2- الدراسة التحضيرية⁽¹⁾ :

- يقصد بالدراسة التحضيرية الاطلاع المسبق على كل جوانب المشكلة - موضوع الدراسة - للتعرف على: جوانبها الفنية، نشأتها التاريخية، وتعتبر هذه العملية هامة حيث تساعد في اختصار البيانات الضرورية والهامة وذلك على النحو الآتي :
- قد يجد الباحث أن هناك من قام بهذه الدراسة سابقاً وتوصل إلى نتائج ما، وعليه تطوير هذه النتائج .
 - قد يجد الباحث أن هناك من أكد استحالة هذا البحث وبالتالي يمكن الاستفادة من هذه النتيجة .
 - بالإضافة لذلك تساعد الدراسة التحضيرية بالتعرف على المشكلة بشكل أكثر والتعرف على المصطلحات الفنية الخاصة بها.

4-2-2-2- مصادر البيانات :

يمكن تقسيم البيانات حسب مصادرها إلى قسمين: بيانات ثانوية يمكن الحصول عليها من مصادر غير مباشرة، وبيانات أولية يمكن الحصول عليها من مصادر الميدان. وعلى الباحث أن يحدد مسبقاً نوع البيانات التي سوف يعتمد عليها، ويجب تفضيل بعضها على الآخر من خلال مقارنة مزايا ومساوئ كل أسلوب، وهذا ما سنتعرف عليه من خلال دراسة كل مصدر .

1-4-2-2- البيانات الثانوية Secondary data :

هي كل ما يمكن الحصول عليه من بيانات قد جمعت من قبل الباحثين بشكل مسبق سواء أكانت منشورة أم غير منشورة. ومن الأمثلة عليها البيانات الصادرة عن جهات حكومية كالبيانات الإحصائية للوزارات المختلفة أو المجموعات الإحصائية عن الدول أو المنظمات الدولية ، الأمم المتحدة، جامعة الدول العربية... إلخ.

(1) د، حميد ناظم، نهدي الإحصاء . منشورات جامعة دمشق 1981 .

وإن أهم ما تتميز به هذه البيانات هي أنها توفر الوقت والجهد والكلفة اللازمة لجمع البيانات، ولكن إلى جانب هذه الميزات هناك مساوئ لهذه البيانات قد تجعل الباحث يتعد عنها، وأهم هذه المساوئ هو مقدار الثقة التي تتمتع به هذه البيانات، وأحياناً اختلاف المتغيرات لهذه البيانات من المتغيرات التي حددها الباحث .

2-2-4-2- Primary data البيانات الأولية :

البيانات الأولية هي تلك البيانات التي يتم جمعها من الميدان وذلك بجمع البيانات عن المفردات موضوع البحث مباشرة من قبل الباحث عن طريق العدادين، أو من خلال استخدام الاستبيان. ويقوم الباحث بجمع البيانات في استبيان أو استمارة إحصائية Questionnaire أو استبيان مصممة خصيصاً لهذه الغاية بذاتها من قبل مختصين، ويجب أن تصاغ هذه الاستمارة بشكل صحيح للحصول على النتائج المرجوة منها، ويتم ذلك بالاستئناس ببعض القواعد العامة الواجب مراعاتها عند صياغة الاستمارة، وهذه القواعد تقسم إلى قسمين: قسم يتعلق بشكل الاستمارة والآخر بطريقة صياغة الأسئلة.

2-2-4-2-1 شكل الاستمارة وترتيبها :

يمكن وضع القواعد والأسس العامة الواجب مراعاتها عند تصميم الاستمارة كالآتي:

- 1- يجب اختيار نوع وشكل ولون الورق المناسب بحيث نستطيع الكتابة والطباعة عليه، ويمكن استخدام أكثر من ورقة ترقيم وتجمع معاً.
- 2- يكتب عنوان البحث على الصفحة الأولى والهدف من البحث واسم الجهة المشرفة عليه.
- 3- يجب الإشارة إلى سرية البيانات والإشارة إلى عدم استخدامها لأغراض أخرى .
- 4- يجب إعطاء الأسئلة أرقاماً متسلسلة حتى يسهل التعرف عليها عند تفريغ الإجابات ، ويمكن أن يعطى لكل سؤال رمز Code في حال كان بالإمكان استخدام الحاسبات في تفريغ النتائج .

- 5- ترتب الأسئلة بشكل منطقي، ويجب أن تبدأ بالأسئلة السهلة مثل الأسئلة الخاصة:
الاسم، الجنس، السن....، ويمكن تقسيم الأسئلة إلى مجموعات تكون الأسئلة
ضمن كل مجموعة مشتركة بعامل ما .
- 6- عدم كتابة أكثر من سؤال في السطر الواحد .
- 7- يمكن تضمين الاستمارة جداول، إذا كانت الاستمارة ستشمل عدة مبحوثين
وكل جدول خاص بمبحوث .

2-2-4-2- القواعد الواجب مراعاتها عند صياغة الأسئلة :

- 1- يجب طرح الأسئلة الهامة ذات العلاقة بموضوع البحث حيث لا نحصل على
بيانات زائدة تعرقل العمليات التالية.
- 2- يجب أن تصاغ الأسئلة بشكل سهل وواضح ولا تقبل اللبس أو إساءة الفهم أو
التأويل، مثلاً السؤال الآتي « الفرع ؟ » فهل المقصود بفرع الدراسة أدبي أو
علمي أم المقصود فرع المنظمة الشعبية أو فرع هيئة سياسية
- 3- تجنب صياغة الأسئلة بحيث تكون الإجابة عليها قاطعة مانعة، ويفضل أن تكون
الإجابة بعدد أو بكلمة صح أو خطأ أو الإشارة إليها (✓ أو x) .
- 4- في حال كانت هناك خيارات، للإجابة أكثر من خيارين صح أو خطأ يجب
هذه الإجابات ويترك للقارئ الاختيار بينها .

مثال : الحالة الاجتماعية ؟

- 1- أعزب. 2- متزوج. 3- مطلق.
- 4- أرمل. 5- مخطوب (مخطوبة) .
- 5- إذا كانت هناك خيارات كثيرة ومتعددة للإجابة، وكان اهتمامنا ينصب على
بعضها فقط فهنا نذكر بعض الإجابات الهامة ونفرد للباقي إجابة عدا ذلك .

مثال: ما هي الهواية المفضلة ؟

- 1- الرياضة. 2- السياحة. 3- السفر. 4- عدا ذلك ...

6- يجب ألا تتطلب الأسئلة عمليات حسابية مطولة فلا نسأل كم كان عمرك في 21-03-2000 مثلاً، يكفي هنا أن نسأل عن تاريخ الولادة .

7- يجب ألا نسأل أكثر من سؤال في السؤال الواحد. فمثلاً نسأل: «متزوج وعدد الأولاد الذكور والإناث؟» يجب هنا توجيه ثلاثة أو أربعة أسئلة.

1- الحالة الاجتماعية :

- ☐ معازب. ☐ متزوج.
☐ مطلق أو مطلقة. ☐ أرمل أو أرملة.

2- عدد الأولاد ؟

3- عدد الذكور ؟

4- عدد الإناث ؟

8- يجب قدر الإمكان تجنب الأسئلة التي تخلق بال المستحوب أو تثير حساسيته :

ما دخلك ؟ ما رصيدك في البنوك ؟

هذه الأسئلة يجب تجنبها لاتصالها بالضرائب أو بأمور أخرى قد تؤذي مصالحه.

9- ألا نسأل أسئلة قد تثير بال المستحوب، كأن نسأله:

- هل تتأخر عن العمل أحياناً بسبب المواصلات ؟

- هل الدخل قليل ؟

- هل تقرأ الجرائد يومياً ؟

- هل رسبت بسبب صعوبة الأسئلة ؟

10- يجب ذكر الوحدات المستخدمة وتوضيحها، فعند السؤال عن الدخل يجب

توضيح المقصود بالدخل، هل هو الراتب الشهري لرب الأسرة أو لأفراد الأسرة؟

أو هو مجموع الدخول لهذه الأسرة من الراتب والعقارات والأموال إلخ،

ومن ثم يجب توضيح الوحدة المستخدمة في قياس الدخل، وهل هو الدخل

الشهري أو السنوي... إلخ .

11- وأخيراً يستحسن إضافة أسئلة راجعة، وهي توجيه أسئلة طرحت سابقاً بطريقة

أخرى لكشف صحة الإجابة من عدمها.

مثلاً : السؤال : هل تقرأ الجرائد يومياً ؟

السؤال الراجع: كم تصبر من الوقت يومياً لقراءة الجرائد ؟

3-2-4-2-2- ملء الاستمارة الإحصائية :

يتم ملء الاستمارة الإحصائية Questionnaire بإحدى الطرق الآتية :

1 - المقابلة الشخصية .

2 - المراسلة (البريد) .

3 - الهاتف .

4 - الإنترنت .

ولكل من هذه الطرق مزاياه ومساوئه بحيث تجعله ممكناً في بعض الحالات، وغير ممكن في حالات أخرى، لذلك تتطلب كل طريقة بعض الشرح بشيء من التفصيل.

1 - المقابلة الشخصية :

يقوم الباحث بحسب هذه الطريقة بنفسه، أو بواسطة أشخاص مدربين على البحث بشكل جيد. بمقابلة المبحوث، وتوجيه الأسئلة إليه بشكل مباشر سواءً سؤالا، حسب ورودها في الاستمارة أو ملاحظة المبحوث بشكل مباشر وتسجيل الإجابات. وتعتبر هذه الطريقة من أكثر الطرق شيوعاً وفائدة في الحالات الآتية :

- في حال كثرة عدد الأميين بين أفراد البحث .

- في حال الأسئلة التي تتطلب الإجابة عليها ملاحظة مباشرة، مثلاً ملاحظة حركات المستجوب (المبحوث)، نظافة البيت..... إلخ .

- بالإضافة لما ورد أعلاه فإن لهذا الأسلوب مزايا أخرى هي إلمام الباحث بالبحث، فسيكون لأسلوب توجيه السؤال وتوضيحه تأثير في الحصول على الجواب المطلوب والتأكد من الحصول على كل الأجوبة.

- بالإضافة إلى تلك المزايا هناك مساوئ لهذا الأسلوب هي :
- قد لا يصلح هذا الأسلوب في الحصول على أجوبة في حالة الأسئلة المخرجة كأن نسأل عن الأمراض الجلدية أو العلاقات الزوجية .
 - تخضع لتحيز الباحث حيث يستطيع توجيه المبحوث للإجابة باتجاه معين عن طريق الإيحاء .
 - تحتاج إلى وقت طويل وكلفة عالية إذا كان البحث يتطلب عدداً كبيراً من الباحثين .

2 - المراسلة (الاستبيان البريدي) :

ترسل الاستمارة بالبريد أو توزع باليد على المبحوثين، أو يمكن أن تنشر بالصحف ويقوم هنا المبحوث بعملء الاستمارة بنفسه، بالإجابة على الأسئلة الواردة فيها بالترتيب.

ولهذا الأسلوب مزاياه ومساوئه والتي نذكرها بما يأتي :

● المزايا :

- 1- سهولة الوصول إلى المبحوثين وقلة التكاليف.
- 2- تصلح في الحالات التي تتضمن أسئلة مخرجة .
- 3- التخلص من تأثير الباحث على المبحوث عن طريق الإيحاء وبالتالي التخلص من تحيز الباحث .
- 4- تسمح للمبحوث بالتفكير بالأجوبة.

● المساوئ :

- 1- لا تصلح في حال كثر عدد الأميين بين أفراد البحث.
- 2- قد يؤدي هذا الأسلوب إلى الحصول على أجوبة متحيزة، وذلك لأن الأجوبة ستعود من بعض فئات الناس، وهم الفئة الذكية التي تريد توجيه البحث باتجاه معين وغياب بعض الإجابات من الناس المهملين.

لذلك عند استخدام هذا الأسلوب ننصح بإضافة بعض الإجراءات إلى صياغة الاستمارة، والتي قد تحفز المهملين إلى الإجابة عنها :

- إرسال طابع بريدي ومغلف كتب عليه العنوان أو إجراء اتفاق مع مؤسسة البريد بالدفع عند وصول الرسائل وذكر ذلك على الاستمارة .
- ذكر بعض العبارات الكفيلة بتحريض المبحوث على الإجابة :
- « أضي المواطن نحن نثق بوعيك وحبك للتطوير والتقدم إلى الأمام، نرجو المساعدة بالإجابة على هذه الاستمارة » .
- من الممكن استخدام أساليب محرضة أخرى، كأن تعطى للاستمارات أرقام، ويجرى سحب عليها للحصول على جوائز بعد الإجابة عليها.

3- استخدام الهاتف للـ استمارة :

هذه الحالة يقوم الباحث بتوجيه الأسئلة الواردة بالاستمارة للمبحوث عن طريق الهاتف .

● مزاياها :

- السرعة في الحصول على البيانات .
- تعطي هذه الطريقة المجال لتوضيح الأسئلة للمبحوث .

● مساوئها :

- لا تصلح إلا للأشخاص الذين يملكون هاتفاً .
- تحتاج إلى تكاليف عالية أحياناً .
- لا تصلح في الحالات التي تتطلب ملاحظة مباشرة .

4- استخدام الإنترنت للـ استمارة «Internet» :

يتم نشر الاستمارة البريدية على شبكة الإنترنت، ويطلب من المبحوثين ملء هذه الاستمارة، ويستخدم هذا الأسلوب بشكل واسع من قبل المخططات التلفازية والصحف والمجلات لاستطلاع الرأي. مزايا هذا الأسلوب السرعة الفائقة في الحصول على

البيانات وقلة التكاليف، أما مساوئه فهو عدم امتلاك كل الناس للحاسب ولخطوط الإنترنت، وأمية بعض الأفراد باستخدام الحاسب. كما أن استخدام هذا الأسلوب يتطلب الإعلان عن وجود هذه الاستمارة عن طريق الصحف أو الإذاعة أو التلفاز.

5-2-2- أسلوب جمع البيانات :

بعد أن يحدد الباحث أنه سيستخدم مصادر الميدان للحصول على البيانات يجب عليه أن يحدد الأسلوب الواجب استخدامه لجمع هذه البيانات، وتجمع البيانات من الميدان بأحد أسلوبين :

1 - أسلوب الحصر الشامل .

2 - أسلوب العينة .

واعتماداً على مزايا ومساوئ كل أسلوب وظروف البحث الخاصة يحدد الباحث الأسلوب الواجب استخدامه :

1-5-2-2- أسلوب الحصر الشامل « Census » :

يقوم هذا الأسلوب على جمع البيانات العددية لجميع مفردات المجتمع الإحصائي، ويستخدم هذا الأسلوب عادة في التعدادات العامة مثل تعداد السكان، وإن مزايا هذا الأسلوب هي دقة النتائج، فإننا سوف نحصل على نتائج عن المجتمع الإحصائي، كما هو فعلاً باستثناء بعض الأخطاء الحسابية، بمقابل هذه المزية هناك مساوئ هي أن هذا الأسلوب يحتاج لوقت طويل وتكاليف عالية عندما يكون حجم المجتمع كبيراً .

2-5-2-2- أسلوب العينة Sample Method :

العينة هي جزء يحوي عدة مفردات عن المجتمع، ويحدد عدد هذه المفردات حسب قواعد علمية سوف نذكرها في فصول أخرى، ونقوم بدراسة العينة للحصول منها على معلومات عن المجتمع الإحصائي الذي سميت منه، ومن الضروري - حتى تكون هذه المعلومات صحيحة في المجتمع- أن تكون العينة ممثلة لمجتمعها تمثيلاً صحيحاً، أي

إنها تحوي الصفات الموجودة بالمجتمع الإحصائي نفسها وبالنسب نفسها، ويمكن الحصول على هذه العينة باستخدام الأساليب العلمية عند سحبها وهذا ما سنعالجه أدناه.

3-5-2-2- المفاضلة بين أسلوب الحصر الشامل وأسلوب العينة :

- هناك عدة ملاحظات يجب أخذها بعين الاعتبار لاختيار أسلوب الجمع :
- لكل بحث وقت لإنجازه وتكاليف مخصصة له، فاعتماداً على الوقت المخصص والتكاليف يمكن للباحث أن يقرر أو يختار الأسلوب المناسب، مع الأخذ بعين الاعتبار أن أسلوب الحصر الشامل يعطي نتائج أكثر دقة .
 - عندما يكون المجتمع الإحصائي غير محدود، فمن الصعوبة الوصول إلى كل مفرداته لدراستها، مثل تحليل ماء البحر أو تعداد الأسماك في البحر فهنا يجب استخدام أسلوب العينة في الدراسة .
 - إذا كانت الدراسة ستؤدي إلى تلف الوحدات موضوع الدراسة يجب استخدام أسلوب العينة، مثل تحليل دم الإنسان، أو فحص شحنة مستوردة من العبوات المعلبة، فإن الدراسة بأسلوب الحصر الشامل معناه سحب الشخص بشكل كامل، وفتح كل علب الشحنة المستوردة من المعلبات وبالتالي نكون في كلتا الحالتين قد أتلفنا وحدات المجتمع .

4-5-2-2- العينات :

عرف استخدام العينات منذ القلم ويؤدي هذا الأسلوب الفرض بنجاح ودون العودة إلى أسلوب الحصر الشامل، فمن المشاهد اليومية لاستخدام أسلوب العينة فحص الزمرة الدموية وتناول الأدوية، وكل الفحوصات المخبرية والنماذج التسويقية .

وتقسم العينات إلى نوعين :

- عينات عشوائية Random Samples .
- عينات غير عشوائية (عمدية) Purposive Sample .

وحتى تؤدي العينات الغاية المرجوة منها وهي التعرف على المجتمع الإحصائي من خلالها دون العودة إلى أسلوب الحصر الشامل يجب أن تحوي على صفات المجتمع، بالنسبة نفسها وهذا يمكن أن يتحقق إذا كانت العينة عشوائية. والمقصود بالعشوائية هنا أن نعطي لكل مفردة من مفردات المجتمع الإحصائي حظاً معيناً للظهور بالعينة، ويتم هذا باستخدام أساليب علمية عند السحب تدعى بأساليب السحب العشوائي أو الاختيار العشوائي Random Selection :

قبل أن نستعرض كيفية السحب بهذه الأساليب لنلاحظ الفرق بين الأمثلة الآتية: أخذ عينة صغيرة من دم الإنسان تعتبر ممثلة لمجتمعها من خلال هذه العينة، مثال آخر: هناك بعض الأدوية كتب عليها: (يرجى رج الزجاجة قبل تناول الدواء). والغاية من هذه العملية تحقيق التجانس بالخلول ككل وعندها فإن أخذ جزء منه يكون ممثلاً للمجتمع. لو أردنا فحص نسبة المعدن في فلز ما فيتوجب علينا طحن هذا الفلز جيداً، وخلطه جيداً حتى نحصل على عينة ممثلة لمجتمعها. لكن في الحياة العملية نصادف حالات لا يمكن إجراء عمليات الرج والطحن والخلط عليها لتحقيق التجانس في المجتمع، مثل سحب عينة من الأشخاص أو الأسر أو السيارات، فهنا لا بد من استخدام أساليب علمية تحقق هذا التجانس أولاً، وتحقق عشوائية العينة ثانياً. وهذه الأساليب هي :

1 - أسلوب القرعة :

نقوم بترقيم مفردات المجتمع $N - 1$ حيث N هي حجم المجتمع، وبعدها نقوم بكتابة هذه الأرقام أو أسماء المفردات على قصاصات من الورق وتوضع في كبسولات متماثلة على ورقة في كبسولة، وقبل البدء بالسحب نخلط الكبسولات في وعاء بشكل جيد، ونبدأ بسحب الكبسولات واحدة تلو الأخرى مع تكرار الخلط قبل كل سحب، وهكذا حتى نحصل على عدد من الكبسولات والتي هي ممثلة لعدد من المفردات n حجم العينة.

يمكن استخدام هذا الأسلوب، ونجاح، إذا كان حجم المجتمع الإحصائي صغيراً، وهذا ما نشاهده بشكل متكرر على شاشة التلفاز عند اختيار مجموعات كرة القدم بكأس العالم. أما عندما يكون حجم المجتمع الإحصائي كبيراً، فمن الصعوبة تطبيق هذا الأسلوب.

2 - دواليب الحظ :

لا تختلف هذه الطريقة بجوهرها عن الطريقة السابقة ولكنها أسهل بكثير . نقوم بترقيم مفردات المجتمع $1 - N$ ونبدأ بالسحب بتدوير الدواليب، وعندما تتوقف الدواليب نقرأ الرقم فتكون الوحدة في المجتمع الإحصائي التي تحمل ذلك الرقم هي من ستدخل في العينة، وهكذا نكرر العملية لسحب كل مفردة حتى نحصل على n مفردة حجم العينة .

كما أن هذا الأسلوب أصبح أكثر سهولة في وقتنا الحاضر، إذ لا نحتاج إلى دواليب خاصة لهذه الغاية حيث توجد هذه الدواليب على أغلب برامج الحاسب.

3 - جداول الأرقام العشوائية Random numbers :

الجدول العشوائية هي عبارة عن الأعداد 0, 1, 2,, 9 مرتبة عشوائياً على شكل أعمدة وسطور بحيث نستطيع من هذه الأعمدة تشكيل أي رقم بأخذ عدد منها عموداً أو اثنين ويتم استخدامها كما يأتي :

نرقم مفردات المجتمع $1 - N$ ثم نحدد عدد الأعمدة الواجب أخذها، وذلك حسب عدد مفردات المجتمع N فإذا كانت $N = 60$ يجب أخذ عمودين (خانتين) نختارهما بشكل عشوائي من الجدول، ونبدأ بقراءة الأرقام وكل رقم أكبر من 60 لا نأخذه، لأن حدود المجتمع حتى المفردة ذات الرقم 60 وهكذا حتى نحصل على مفردات العينة N .

مثال (1 - 2) :

قم بسحب عينة بحجم 10 مفردات من مجتمع حجمه 80 مفردة مستخدماً جدول الأرقام العشوائية في الملحق .

ليختار بشكل عشوائي عموداً وليكن العمود الثالث من اليسار، وكذلك سطرأ بشكل عشوائي وليكن الرابع، ونبدأ بقراءة الأرقام بأعداد ثنائيتين لأن أكبر مفردة في المجتمع هي 80 فتكون مفردات العينة هي :

58, 63, 44, 59, 53, 13, 18, 39, 54, 26

لاحظ أننا أهملنا المفردة ذات القيمة 93 لأنها أكبر من 80.

1-4-2- أنواع العينات العشوائية :

1- العينة العشوائية البسيطة Simple Random Sample :

تعّد العينة العشوائية البسيطة من أهم أنواع العينات العشوائية التي يتم سحبها بإحدى طرائق السحب العشوائي السابقة دون أية إجراءات أخرى، ولكن إلى جانب هذه البساطة والتي من أهم مزاياها بعض المساوئ هي :

- لا يمكن سحب عينة بهذا الأسلوب إلا إذا كان حجم المجتمع محدوداً ومعلومأ كل مفرداته.

- في بعض الأحيان قد لا نحصل على نسب الصفات الموجودة نفسها في المجتمع، مثلاً سحب عينة من طلاب الجامعة الموزعين كما يلي 60 % ذكور و 40 % إناث، فإن العينة البسيطة قد تعطي نسباً تختلف عن الموجودة في المجتمع كأن نحصل على عينة تحوي 20 % إناث وذلك بسبب العشوائية .

2- العينة المنتظمة Systematic Sample :

نقوم بتقسيم مفردات المجتمع المرتبة والرقمة N-1 إلى مجموعات تسمى كل مجموعة بفترة السحب وعدد مفرداتها S، حيث نحصل على طول فترة السحب بتقسيم حجم المجتمع N على حجم العينة n :

$$S = \frac{N}{n}$$

ثم نختار مفردة بشكل عشوائي من المفردات ولتكن المفردة Z فتكون المفردة Z أول مفردة في العينة، وللحصول على باقي مفردات العينة نضيف إليها على التتالي فترة السحب S فتكون بهذا الشكل مفردات العينة هي المفردات ذات الأرقام :

$$Z, Z + S, Z + 2S, Z + 3S, \dots, Z + (n-1)S$$

مثال (2-2) :

بمجمع حجم مفرداته 200 مفردة نريد سحب عينة منتظمة بحجم 10 مفردة :

$$S = \frac{N}{n} = \frac{200}{10} = 20$$

نختار من ضمن فترة السحب الأولى، والبالغ عددها مفرد بشكل عشوائي، بإحدى الطرائق التي تعرفنا عليها ولتكن هذه المفردة ذات الرقم 8 وعليه تكون مفردات العينة هي :

$$8, 28, 48, 68, 88, 108, 128, 148, 168, 188$$

مزايا هذه العينة هي قلة التكاليف وسهولة سحبها، وإنما قد تعطي نسب الصفات الموجودة نفسها في المجتمع أحياناً، ولكن إلى جانب هذه المزايا هناك بعض المساوئ هي أنه لا يمكن سحبها إلا إذا كان المجتمع محدوداً ومرتباً ومرقماً ومعلوم كل مفرداته، كما أنه في بعض الأحيان قد تعطي عينة متحيزة إذا كانت البيانات في المجتمع مرتبة بشكل دوري يتوافق مع فترة السحب، مثلاً عند اختيار عينة من البيوت في منطقة مخططة بطريقة الشطرنج، فإذا كانت وحدة البدء منزلاً يقع على الشارع، وكان طول فترة السحب يساوي عدد البيوت في رتل البيت الذي يقع على الشارع، فستكون البيوت الداخلة في العينة كلها تقع على الشارع، وبذلك نكون قد أغفلنا تمثيل البيوت غير الواقعة على الشارع.

3 - العينة الطبقية Stratified Sample :

عندما يتألف المجتمع الإحصائي من مفردات غير متجانسة نقوم بتقسيمه إلى مجموعات مؤلفة من مفردات متجانسة تدعى كل مجموعة طبقة (Stratum) ، ثم تؤخذ

من كل طبقة عينة عشوائية مجموع هذه العينات يدعى بالعينة الطبقية أما نسبة كل طبقة في العينة فيمكن تمثيل الطبقة بالعينة كما يأتي :

- تمثيل الطبقة بالعينة حسب نسبة الطبقة في المجتمع :

إذا كان N حجم المجتمع مولف من 3 طبقات عدد أفرادها N_1, N_2, N_3 .
حجم العينة n .

ومنه :

$$N_1 + N_2 + N_3 = n$$

ويجب أن يكون :

$$n_1 + n_2 + n_3 = n$$

حيث n_3, n_2, n_1 حجم كل عينة من كل طبقة .

وللحصول على قيم n_3, n_2, n_1 نطبق نسب الطبقات في المجتمع :

$$\frac{n}{N} = \frac{n_1}{N_1} = \frac{n_2}{N_2} = \frac{n_3}{N_3} = \dots$$

مثال (2-3) :

نود سحب عينة من طلاب الجامعة لمعرفة رأيهم بدراسة وتدريس اللغة الأجنبية، ومن المعروف بشكل مسبق هناك اختلاف بصعوبة وممارسة هذه اللغات، فإذا علمنا أن عدد الطلاب 20000 طالب، 10000 طالب لغتهم الإنكليزية و 7000 طالب لغتهم الفرنسية، والباقي لغات مختلفة، ونود سحب عينة بحجم 100 طالب، فما عدد الطلاب من كل طبقة ؟

$$\frac{n}{N} = \frac{n_1}{N_1} = \frac{n_2}{N_2} = \frac{n_3}{N_3}$$

$$\frac{100}{20000} = \frac{n_1}{10000} = \frac{n_2}{7000} = \frac{n_3}{3000}$$

$$n_1 = 50$$

$$n_2 = 35$$

$$n_3 = 15$$

ومنه

- يمكن أن تمثل الطبقات في العينة حسب التباين ضمن كل طبقة كلما زاد التباين
'يريد من ممثليها' 'معيه' 'وكمسح' .

تتاز هذه العينة الطبقية بأنها تمثل كل الصفات بالنسب نفسها كما أنها يمكن من
دراسة الطبقات على انفراد إذا كان ذلك مفيداً بالإضافة إلى دراسة المجتمع. ولكن من
مساوئها أنه لا يمكن سحبها إلا إذا كان معلوماً مفردات المجتمع ككل وعددها،
وكذلك مفردات كل طبقة وعددها.

4 - العينة العنقودية Cluster Sample :

رأينا أعلاه أن تلك العينات لا يمكن سحبها إلا إذا توافرت قوائم بمفردات
المجتمع ككل، ولكن هناك حالات قد لا يتوفر لدينا مثل هذه القوائم بمفردات المجتمع،
فيتوجب علينا إعداد هذه القوائم عن طريق استخدام أساليب أخرى لسحب العينات
منها بأسلوب العينة العنقودية. وحسب هذا الأسلوب نقوم بتقسيم المجتمع إلى
مجموعات أو أجزاء حسب خاصية ما أو صفة ما، ونعتبر هذه الأجزاء هي وحدات
المجتمع، والتي تدعى كل منها عنقوداً، ثم نختار بشكل عشوائي عدداً من هذه العناقيد
لنجمع البيانات منها.

5 - العينة المتعددة المراحل Multi-Stage Sample :

تستخدم هذه العينة في تلك الحالات التي لا تتوفر قائمة لدينا عن مفردات
المجتمع، وكذلك عندما يكون الوقت المتاح قصيراً والإمكانات المادية المتاحة قليلة.
يتم حسب هذه الطريقة سحب العينة على عدة مراحل، وكمثال عليها لدراسة
متوسط إنفاق الأسرة في الجمهورية العربية السورية نختار عدة محافظات بشكل عشوائي
(كمرحلة أولى)، بعد ذلك من ضمن المحافظات المختارة في المرحلة الأولى نختار عدة
مناطق (كمرحلة ثانية)، ثم نختار عدة قرى وأحياء بشكل عشوائي من المناطق المختارة
بالمرحلة الثانية (كمرحلة ثالثة)، وأخيراً كمرحلة رابعة نختار عدداً من الأسر من القرى
والأحياء المختارة في المرحلة الثالثة ونجمع البيانات من هذه الأسر فقط المختارة بالمرحلة الرابعة .

2-4-5-2-2 - العينات الشخصية (غير العشوائية) Non Random Samples :

يتم اختيار بعض العينات بشكل غير عشوائي أي عمدي، فهذا النوع من العينات هو غير عشوائي، وبالتالي لا يمكن حساب احتمال الخطأ باستخدام نظرية الاحتمالات في هذا النوع من العينات :

1 - عينة الحصص Quota Sample :

يحدد الباحث عدداً من المفردات لكل صفة، وهي بهذا الشكل تتشابه مع العينة الطبقية وتختلف معها بأنها ليست عشوائية .

2 - العينة المتقاة Selective Sample :

يستخدم هذا النوع بشكل أوسع من العينة السابقة، وذلك لما حققه من نجاح ناتج عن مهارة الباحث في اختيار المفردات التي ستدخل في العينة، ومن الأمثلة على هذا النوع السلع الداخلة في حساب الرقم القياسي للأسعار.

3 - العينات كبيرة الحجم Extenovie Sample :

وهي عينات تحوي عدداً كبيراً من المفردات مختارة بشكل غير عشوائي، ويدعي مؤيدو هذا النوع من العينات أنه كلما كبر حجم العينة كانت أكثر تمثيلاً للمجتمع، ولكن هذا غير صحيح إلا إذا كانت العينة عشوائية، أما في حال العينات غير العشوائية فإن التحيز سوف يبقى أو يمكن أن يزداد ما دام سحبها يتم بشكل غير عشوائي.

5-2-2-2 - أخطاء العينة Errors of Sample :

أشرنا أعلاه بأننا سوف نقوم بسحب عينة للتعرف منها على صفات المجتمعات الإحصائية، ولكن على الغالب نحصل على نتائج لها تماثل النتائج الممكن الحصول عليها، لو قمنا بالحصر الشامل وذلك لأن العينة عرضة لنوعين من الخطأ هما:

1 - خطأ الحظ والمصادفة Chance أو الخطأ العشوائي..

2 - خطأ التحيز Bias Error .

1 - الخطأ العشوائي :

ينتج هذا الخطأ بسبب العشوائية ذاتها حيث تختلف نتائج العينة عن القيم الموجودة في المجتمع ولتبيان ذلك نضرب المثال التالي:

مثال (2-4) :

لدينا أربعة طلاب A, B, C, D أوزانهم على التوالي 60, 66, 72, 78 كغ فإذا اعتبرنا أن هؤلاء طلاب يولفون مجتمعاً متوسط الوزن فيه

$$69 = \frac{60 + 66 + 72 + 78}{4} \text{ كغ}$$

الآن إذا قمنا بسحب عينة عشوائية بحجم ثلاثة أشخاص، فإنها ستكون إحدى العينات التالية في الجدول رقم (2 - 1) :

جدول رقم (2 - 1)

الخطأ	متوسط الوزن	مجموع أوزان العينة	مفردات العينة	العينة
3 - 66 = 69	66	60 + 66 + 72 = 198	A B C	1
1 - 68 = 69	68	60 + 66 + 78 = 204	A B D	2
1 - 70 = 69	70	60 + 78 + 72 = 210	A D C	3
3 - 72 = 69	72	66 + 72 + 78 = 216	B C D	4
0				

لدى تدقيق متوسط الأوزان التي أعطتها العينات نلاحظ أن هناك فروقاً بين متوسط العينة ومتوسط المجتمع لكل العينات الممكنة ، وهذا الفرق ناتج عن عملية الاختيار، ونطلق عليه خطأ الحظ والمصادفة، ولكن لاحظ أن هذا الخطأ الناتج هو لبعض العينات موجب ولبعضها سالب، أي يحدث في الاتجاهين، وأن مجموع الخطأ يساوي الصفر. من هنا يمكن القول إن القوى التي أدت إلى ظهور هذا الخطأ متساوية، وإلّا تلغى بعضها بعضاً عندما يزداد حجم العينة، ففي مثالنا السابق أكبر خطأ كان 3 كغ

أما لو سحبنا عينة بحجم أصغر ولتكن مفردتين، هما D و C ومتوسطهما 75 كغ والخطأ سيكون « $6 = 75 - 69$ » أي أن الخطأ أكبر عندما كان حجم العينة صغيراً. وهذا يؤدي بنا إلى القول إنه كلما كبر حجم العينة قل مقدار خطأ الحظ والمصادفة. وأخيراً نقول إن هذا الخطأ لا بد وأن يقع لأنه ناتج عن العشوائية نستطيع تقليل قيمته بزيادة حجم العينة كما نستطيع حساب احتمال الوقوع به كما سنرى لاحقاً.

2 - خطأ التحيز :

يمكن وصف التحيز بأنه الحصول على عينة كل مفرداتها أو أغلب مفرداتها، تحمل الصفة نفسها مثلاً سحب عينة لقياس متوسط الطول، وكل مفرداتها أشخاص طوال القامة، أو سحب عينة لتحديد نسبة الأمية في بلد ما من طلاب الجامعة.... إلخ. والتحيز دوماً يحدث باتجاه واحد سالب أو موجب، ويعتبر التحيز من أخطر الأخطاء التي يمكن أن تؤثر على البحث، وتكمن خطورته بأنه لا يمكن التنبؤ به وقياسه. ونحمل أسباب التحيز بما يأتي :

- عدم السحب العشوائي أو استخدام إطار معيب عند سحب العينة، أي السحب ليس من كامل المجتمع.
- عدم تنفيذ العينة بشكل كامل - إهمال بعض وحدات العينة.
- الاستعاضة بوحدات مكان عند سحب العينة .

ويمكن تصحيح خطأ التحيز بإدخال بعض التعديلات على مفردات العينة عندما يكون معلوماً اتجاه التحيز وحجمه، وإلا وجب إعادة سحب العينة.

عرضنا في هذا الفصل بعض الوسائل الفنية والإرشادات التي يمكن استخدامها في عملية جمع البيانات، والتي ستساعد في الحصول على البيانات الصحيحة والضرورية للبحث، مما يسهل عمل الباحث في المراحل اللاحقة، حيث إن الهدف من عملية جمع البيانات هو ليس غاية بحد ذاته، وإنما وسيلة وخطة للانتقال إلى المراحل الأخرى: عرض البيانات وتحليلها واتخاذ القرار، وتنتهي عملية جمع البيانات بالقيام بمراجعة الاستثمارات (Revision) واحدة واحدة، والغاية من هذه العملية هو اكتشاف الاستثمارات التي

تحتوي على نقص في الإجابات أو تناقض. وهنا يجب استكمال الأجوبة من الميدان أو إلغاؤها ، كما يمكن تحليل عملية المراجعة إجراء بعض الحسابات اللازمة التي وفرنا على المبحوث القيام بها.

إن نتيجة مرحلة جمع البيانات هي الحصول على عدد كبير من البيانات ندعوها بشكلها هذا دون أي إجراءات عليها بيانات خام Raw data ، فإذا ما عرضت بهذا الشكل ضمن بحث ما أو على الصحف أو على الشاشة لن تقدم أي فائدة للباحث أو للقارئ، ولذا كان من الضروري عرض هذه البيانات بشكل علمي وفني وسهل، وهذا ما سنناقشه في الفصل القادم.



الفصل الثالث

عرض البيانات الإحصائية

(Presentation of Statistical data)

- مقدمة

- عرض البيانات النوعية (الوصفية)

العرض الجدولي للمتغيرات النوعية

العرض البياني للمتغيرات النوعية

الأعمدة البيانية

الرسوم الدائرية

- عرض البيانات الكمية

العرض الجدولي للمتغيرات الكمية

العرض البياني للمتغيرات الكمية



1-3- مقدمة :

تستخدم أساليب عرض البيانات الإحصائية بشكل واسع في كل المجالات والعلوم، حيث تبلغ أهميتها أهمية علم الإحصاء من هذه العلوم، وأحياناً أهمية أكبر منه وذلك بسبب شعبية هذه الأساليب، وعدم غلو بحث أو كتاب أو صحيفة أو نشرة أخبار تلفزيونية أو إعلان تسويقي من أحد هذه الأساليب، واكتسبت أساليب العرض أهمية أعظم في وقتنا الحاضر وذلك لقدرة الحاسبات الإلكترونية على القيام بالحسابات الإحصائية اللازمة ووضع أساليب العرض المختلفة، كل هذا جعل معرفة أساليب العرض من المعارف الضرورية لكل إنسان معاصر، مما دعا إلى إعطاء هذا الفصل أهمية خاصة، وبالتالي الإسهاب في عرض وسائل وأساليب عرض البيانات الإحصائية عرضاً جدولياً وبيانياً، حيث لكل منها عدة أساليب عرض نتعرف عليها من خلال تقسيمها حسب طبيعة البيانات، ليتمكن الطالب بسهولة من معرفة حالات استخدام كل أسلوب وهي كما يلي :

- عرض البيانات النوعية .

- عرض البيانات الكمية .

2-3- عرض البيانات النوعية (الوصفية) :

تعرفنا في الفصل السابق على المتغيرات النوعية، ورأينا أنها بيانات تعود لمفردات تتصف أو لا تتصف بصفة ما، أو هي التي يمكن التعبير عنها بكلمة أو جملة: شجرة زيتون، طالب، طالب اقتصاد، حميد إلخ .

نعرض المتغيرات النوعية بطريقتين :

1- العرض الجدولي للنوعيات .

2- العرض البياني للنوعيات .

وستتعرف بالتفصيل على أساليب عرض المتغيرات النوعية فيما يأتي :

1-2-3 العرض الجدولي للمتغيرات النوعية :

رأينا أعلاه أن نتيجة مرحلة جمع البيانات العددية هي الحصول على عدد كبير من البيانات، تعذر الاستفادة من هذه البيانات، لذلك لا بد من عرضها بشكل منظم، ويتم ذلك عن طريق تبويبها .

- التبويب Tabulation :

وهو وضع البيانات الإحصائية في جداول، وذلك بعد تقسيمها حسب صفاتها المشتركة، والغاية من هذه العملية هو اختصار البيانات إلى أصغر حيز يمكن أن يستوعبها، وللمر عملية التبويب بثلاث خطوات :

1 - تقسيم البيانات إلى مجموعات حسب كمياتها أو صفاتها المشتركة تقسيماً

جامعاً مانعاً، حيث إن كل مفردة تنتمي لإحدى المجموعات .

2 - تعداد مفردات كل مجموعة .

3 - إنشاء الجدول الإحصائي ونقل البيانات إليه، بحيث يظهر عدد مفردات كل

مجموعة .

نرى أن نتيجة عملية التبويب هو الحصول على الجدول الإحصائي الذي يعرف

بأنه ترتيب البيانات على أسطر وأعمدة حسب كمياتها أو صفاتها المشتركة، والغاية هو

اختصار البيانات وملاحظة صفاتها العامة، ولكي نحصل من الجدول الإحصائي على الأغراض المرجوة منه لا بد وأن تتوفر فيه القواعد التالية :

- 1 - يجب أن يكون للجدول عنوان واضح يعبر عن محتوى الجدول وزمن الحصول على البيانات ومكان الحصول عليها، وبشكل آخر يجب أن يجب عنوان الجدول على الأسئلة الآتية: ماذا يعبر الجدول ؟ ومن أين تم الحصول على محتوياته وزمن الحصول عليها ؟
 - 2 - يجب ذكر وحدات القياس المستخدمة في الجدول .
 - 3 - يجب وضع عنوان لكل عمود ولكل سطر، وفي حال وضع الجدول بأكثر من صفحة يجب إعادة العناوين.
 - 4 - يمكن ترقيم الأسطر والأعمدة لتسهيل العودة إليها إذا كان لا بد من العودة إليها عند شرح محتوياتها .
 - 5 - يجب ذكر مصدر البيانات في الجدول وذلك بالإشارة إلى اسم المصدر الذي أخذت منه .
 - 6 - يجب وضع رقم للجدول إذا كان سيوضع ضمن كتاب أو كتيب، وذلك للعودة إليه عند شرح البيانات الموجودة ضمنه .
 - 7 - يجب أن يكون شكل الجدول متناسباً مع شكل الصفحة التي سيعرض عليها، وذلك حتى لا يضطر القارئ لتغيير وضع الكتاب عند القراءة .
 - 8 - يجب عرض البيانات في الجدول الإحصائي حسب ترتيب ما، حيث يمكن أن تعرض حسب الحجم تصاعدياً أو تنازلياً، أو حسب التاريخ، أو يمكن أن تعرض جغرافياً، والغاية من استخدام ترتيب ما هو سهولة الحصول على البيانات عند البحث في من الجدول، وقد ترتب البيانات ضمن الجدول على أكثر من أساس واحد، فقد يكون الترتيب زمنياً وجغرافياً معاً.
- والجدول رقم (1 - 3) هو جدول إحصائي تتوفر فيه بعض القواعد المذكورة أعلاه .

الجدول رقم (1 - 3) يمثل توزيع الأراضي في الجمهورية العربية السورية
حسب الاستعمال والحافظات لعام 2000 (الف هكتار)

المحافظة	الأراضي القابلة للزراعة	الأراضي غير القابلة للزراعة	مروج ومراعي	حراج	المجموع
دمشق	129	246	1331	45	1814
حلب	1227	351	228	43	1849
حمص	419	1009	2617	49	4094
حماة	489	158	294	75	1016
اللاذقية	113	27	5	85	230
دير الزور	233	1213	1854	6	3306
إدلب	343	140	49	78	610
الحسكة	1393	84	744	112	2333
الرقة	780	153	977	15	1925
السويداء	189	143	215	8	555
درعا	228	110	27	8	373
طرطوس	120	36	4	30	190
القنيطرة	154	15	14	3	186
حوض الفرات	25	12	0	0	37
المجموع	2905	3697	8359	557	18518

المصدر : المجموعة الإحصائية السورية لعام 2001 .

- أنواع الجداول الإحصائية :

يمكن تقسيم الجداول الإحصائية حسب الغاية من إعدادها، وكذلك حسب عدد الصفات (المتغيرات) في الجدول .
حسب الغاية من الجدول الإحصائي تقسم الجداول إلى جداول عامة وخاصة .

الجداول العامة :

هي تلك الجداول التي تعدّ دون أي هدف سوى العد والاختصار وعادة ما تنشأ هذه الجداول من قبل الحكومات لتوفير البيانات اللازمة عن النشاطات المختلفة في البلد وكمثال عليها المجموعة الإحصائية السورية الصادرة من المكتب المركزي للإحصاء، وتستخدم هذه البيانات في الأبحاث كبيانات ثانوية .

الجداول الخاصة :

هي تلك الجداول التي يضعها الباحث بنفسه عند دراسة مشكلة ما، ويمكن أن تكون مستمدة من الجداول العامة .

أما حسب عدد الصفات في الجدول، فيقسم إلى جداول بسيطة ومزدوجة ومركبة. في الجداول البسيطة تعرض البيانات حسب صفة واحدة الجدول رقم (2-3) أما الجدول المزدوج هو الذي تعرض فيه البيانات حسب صفتين، تمثل الأسطر تقسيمات إحدى الصفات والأعمدة تقسيمات الصفة الأخرى، انظر الجدول رقم (1-3) أعلاه حيث وضعت الصفة الأولى المحافظة في الأسطر، والصفة الثانية استعمال الأراضي في الأعمدة.

الجدول رقم (2-3) يمثل عدد أشجار الزيتون والتفاح والمشمش

في الجمهورية العربية السورية لعام 2000

(عدد الأشجار بالآلاف)

نوع الشجرة	عدد الأشجار
الزيتون	64344
التفاح	16204
المشمش	3222

المصدر : المجموعة الإحصائية السورية لعام 2001 .

ويدعى الجدول الإحصائي جدولاً مركباً عندما تعرض البيانات فيه لأكثر من صفتين كما في الجدول رقم (3-3) ، حيث تم عرض صفة واحدة في الأسطر وعدة صفات في الأعمدة.

جدول رقم (3-3)

يمثل عدد الأشجار والمساحة والإنتاج للزيتون والمشمش والتين والعنب

في الجمهورية العربية السورية لعام 2000

المساحة الأرض (ألف هكتار)	الإنتاج واستعمالاته (طن)			عدد الأشجار (ألف شجرة)			المحصول
	المجموع	استعمالات أخرى	للأكل طازج	المجموع	شجر مفمرة	مفمرة	
478	866052	755942	110110	64344	24038	40306	الزيتون
12.4	78873	11377	67496	3222	692	2530	المشمش
10.7	44071	12364	31707	2665	219	2446	التين
69.3	409450	148855	260595	49353	5431	43922	العنب

المصدر : المجموعة الإحصائية السورية لعام 2001 .

3-2-2 العرض البياني للمتغيرات النوعية :

يقصد بالعرض البياني استخدام الرسوم للتعبير عن البيانات العددية، وتفيد هذه الطريقة في إظهار البيانات وملاحظة التغيرات فيها بشكل يجذب الانتباه ويساعد على تذكرها، كما تبين هذه الطريقة العلاقة بين الظواهر والمتغيرات التي ندرسها، وهناك عدة أساليب لعرض المتغيرات النوعية بيانياً أهمها :

- الأعمدة البيانية البسيطة والمهزأة .
- الرسوم الدائرية .

3-2-2-1 الأعمدة البيانية (Bar Charts) :

وهي التعبير عن البيانات العددية بمستطيلات على المحاور الإحداثية، قواعد المستطيلات على محور السينات، وارتفاعاتها، تتناسب مع قيمة الظاهرة التي ندرسها. وتستخدم الأعمدة البيانية لعرض ظاهرتين أو أكثر الفترة الزمنية نفسها مثلاً مقارنة عدد طلاب كلية الاقتصاد وعدد طلاب كلية الحقوق لعام 2000، أو عدد أشجار الزيتون

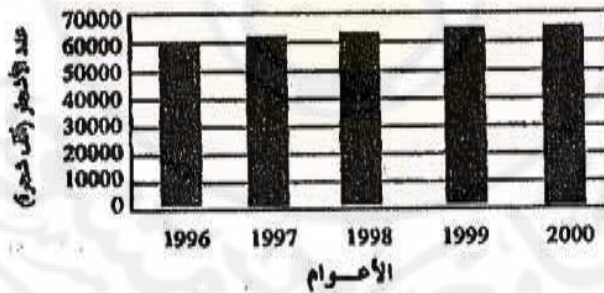
وعدد أشجار العنب في الجمهورية العربية السورية لعام 2000 كما في الشكل رقم (3-1) ، وكذلك تستخدم لعرض بيانات لظاهرة واحدة لفترات زمنية مختلفة، كأن نقارن عدد أشجار الزيتون لعدة أعوام كما في الشكل رقم (3-2) والفائدة من هذا الأسلوب هي مقارنة عدة ظواهر مع بعضها، أو ملاحظة تطور ظاهرة أو عدة ظواهر خلال فترات زمنية مختلفة.



الشكل رقم (3-1)

يحتل مقارنة عدد أشجار الزيتون والعنب في الجمهورية العربية السورية لعام 2000

المصدر : المجموعة الإحصائية السورية لعام 2001 .



الشكل رقم (3-2) تطور عدد أشجار الزيتون

في الجمهورية العربية السورية لأعوام 1996 - 2000

المصدر : المجموعة الإحصائية السورية لعام 2001 .

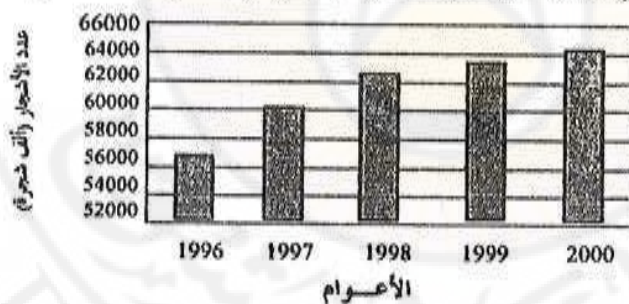
أهم القواعد والملاحظات الواجب مراعاتها عند رسم الأعمدة البيانية :

1 - يجب أن تكون قاعدة المستطيلات متساوية وتفصل بين المستطيلات مسافة لا تقل عن قاعدة المستطيل.

2 - يفضل تمييز المستطيلات عن بعضها بالوان مختلفة أو بالتظليل إذا كانت تعود لظواهر مختلفة كما في الشكل (3-1) .

3 - يستحسن إظهار تقسيمات محاور العينات على المستوى الإحداثي وذلك لبيان قيمة الظاهرة بشكل أكثر دقة .

4 - يجب أن تبدأ تقسيمات محاور العينات من الصفر وإلا سوف نحصل على مقارنة خاطئة، فنلاحظ في الشكل (3-2) أن محور العينات قد بدأ من الصفر ومقارنة عدد أشجار عام 2000 مع عدد أشجار عام 1996 تبين أن هناك زيادة في عدد الأشجار في عام 2000. أما لو بدأنا التقسيم من قيمة أخرى كما في الشكل رقم (3-3)، فإن المقارنة من خلال ارتفاع العمود تبين أن عدد الأشجار في عام 2000 هو ضعف عدد الأشجار عام 1996 وهذا غير صحيح، أما إذا اضطررنا لبدء التقسيم من قيمة غير الصفر لسبب ما، كأن تكون بعض الأعمدة طويلة يجب تبين ذلك على الرسم بكسر محاور العينات من الأسفل أو بكسر الأعمدة من الأعلى.



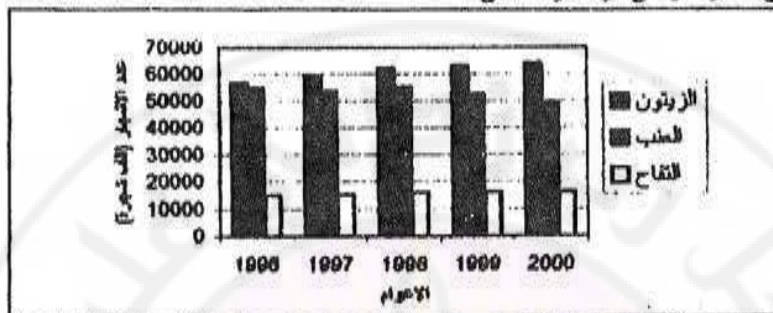
الشكل رقم (3-3) تطور عدد أشجار الزيتون في الجمهورية العربية السورية

للأعوام 2000 - 1996

المصدر : المجموعة الإحصائية السورية لعام 2001 .

5 - يمكن باستخدام الأعمدة البيانية عرض بيانات لعدة ظواهر لفترات زمنية مختلفة على الشكل نفسه وهنا تمثل الظواهر العائلة الفترة الزمنية نفسها بأعمدة متلاصقة

ويعطى لكل ظاهرة لون مختلف. مثلاً مقارنة عدد أشجار الزيتون والعنب والتفاح للأعوام 1996 حتى عام 2000 في سورية كما في الشكل رقم (3-4)، ويجب ألا ننسى هذه الحالة وضع وسيلة إيضاح على الشكل ليشير إلى ماذا يرمز كل عمود وكل لون أو تظليل.

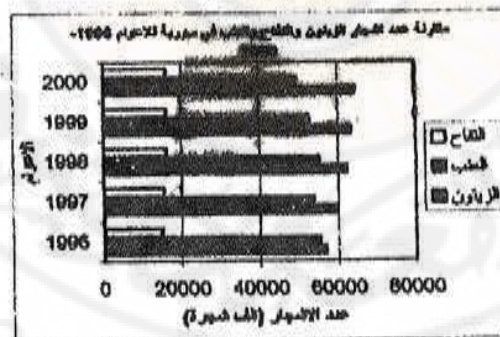


الشكل رقم (3-4)

مقارنة عدد أشجار الزيتون والعنب والتفاح للأعوام 2000-1996 في سورية

المصدر : المجموعة الإحصائية السورية لعام 2001 .

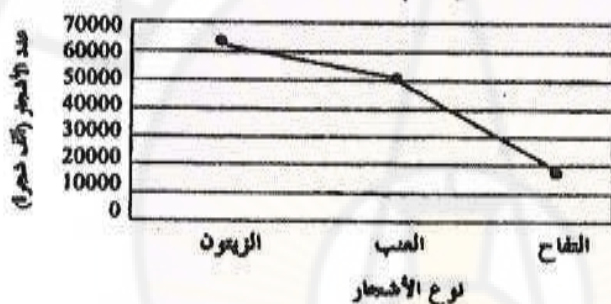
6 ... إن أسلوب **الأعمدة الأفقية** **بالشكل** الواردة أعلاه هو الأكثر شيوعاً، ولكن يمكن بالحالات السابقة استخدام المستطيلات الأفقية، وذلك يجعل محور السينات عينات والعينات مبنات، الشكل رقم (3-5) .



الشكل رقم (3-5)

المصدر : المجموعة الإحصائية السورية لعام 2001 .

7 - يستخلص بعض المؤلفين المنحنيات البيانية لعرض النوعيات بدلاً من الأعمدة البيانية. ونحن نفضل ترك المنحنيات البيانية لعرض السلاسل الزمنية والمتغيرات الكمية المستمرة فقط، لأنه في هذه الحالات يوجد تسلسل مستمر سواء أكان للزمن أم للمتغير المستمر، أما للنوعيات: لا يوجد هذا الاستمرار، هذا من جهة، ومن جهة ثانية عند استخدام المنحنيات في حال النوعيات فإن المنحني سيتغير بمجرد تغير موقع الظاهرة في الشكل، لأنه لا يوجد تسلسل أكيد أو مفترض لهذه الظواهر. وكمثال على استخدام المنحنيات البيانية في حال النوعيات نورد المنحني البياني لعدد أشجار الزيتون والتفاح والعنب في سورية لعام 2000 وبشكلين مختلفين بالترتيب، الشكل رقم (3-6) .



عدد أشجار العنب والزيتون والتفاح في سورية للأعوام 1996 - 2000

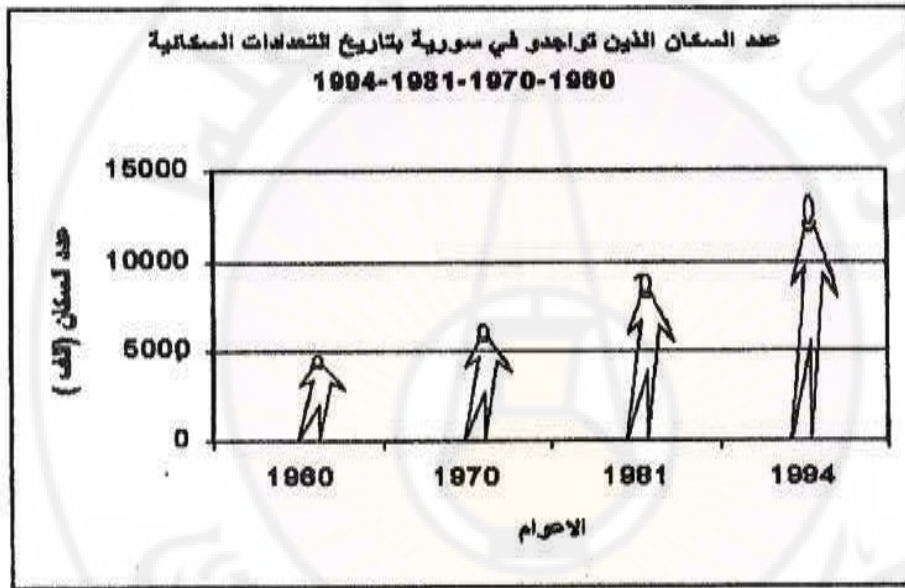


مقارنة عدد أشجار العنب والزيتون والتفاح في سورية للأعوام 1996 - 2000

الشكل رقم (3-6)

مصادر البيانات : المجموعة الإحصائية السورية لعام 2001 .

8 - يمكن استبدال الأعمدة بشكل مستمد من الظاهرة نفسها، وضمن الشروط والقواعد نفسها التي تستخدم بها الأعمدة، وذلك كأن نعرض عدد الأشجار بشكل شجرة بدلاً من العمود وعدد السكان بشكل آدمي وحجم المبيعات برزمة من القطع النقدية.... إلخ، وهذا الشكل هو الأكثر استخداماً في وقتنا الحاضر سواء أكان في أعمال الدعاية والإعلان أم في المقالات أم البرامج التلفازية، كما في الشكل رقم (3-7) .



الشكل رقم (3-7)

مصدر البيانات: المجموعة الإحصائية السورية لعام 2001 .

9 - في حال كان اهتمامنا ينصب على المكونات الجزئية للظاهرة بالإضافة إلى الظاهرة ككل يجب استخدام الأعمدة البيانية المقسمة (Durded Bar Diagrams) حيث يقسم العمود إلى أجزاء، وارتفاع كل جزء يتناسب مع حجم الجزء المكون للظاهرة. وإن الجدول الآتي رقم (3-4) يحوي بيانات عن النتائج الإجمالي بسعر

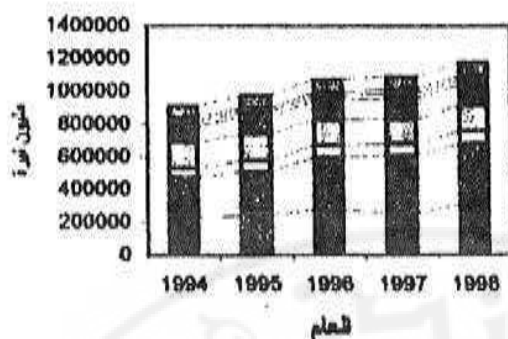
المنتج في سورية حسب القطاعات للأعوام 1994-1998 بأسعار 1995 الثابتة وملايين الليرات . إن عرض بيانات هذا الجدول بالأعمدة المقسمة كما في الشكل رقم (3-8) تساعد على مقارنة تطور الناتج الإجمالي ككل، وكذلك تطور مساهمة كل قطاع في الناتج الإجمالي. لكن من مساوئ هذا الأسلوب هو صعوبة مقارنة تطور الأجزاء، باستثناء الجزء المجاور لمحور السينات، وللتخلص من هذه السببية يفضل رسم خطوط متقطعة تصل بين نهايات الأجزاء، وعندها تكون المقارنة اعتماداً على هذه الخطوط، فإذا تباعدت دل ذلك على زيادة حجم الجزء، أما في حال تقاربها دل على التناقص، وإذا ما بقيت متوازية دل على ثبات قيمة هذا الجزء كما هو مبين في الشكل رقم (3-8) .

جدول رقم (3-4)

يمثل الإنتاج الإجمالي بسعر المنتج حسب القطاعات للأعوام 1994-1998
(بأسعار 1995 بملايين الليرات السورية)

الأعوام					القطاع
1998	1997	1996	1995	1994	
308182	255932	276347	241508	225902	الزراعة
373630	349685	320714	266524	247403	الصناعة والتعدين
76842	74893	73748	68101	58506	البناء والتشييد
156345	152197	158341	164526	159882	تجارة الجملة والمفرق
122444	123016	113453	107366	95995	النقل والمواصلات والتخزين
30774	30087	28093	28939	27981	المال والتأمين والعقارات
26814	24485	21592	20759	22146	خدمات المجتمع والشخصية
81127	78812	78277	77865	75046	الخدمات الحكومية
369	321	292	266	242	المبيعات التي لا تهدف إلى الربح
1176527	1089428	1070857	975854	913121	المجموع

المصادر : المجموعة الإحصائية السورية لعام 2001 .

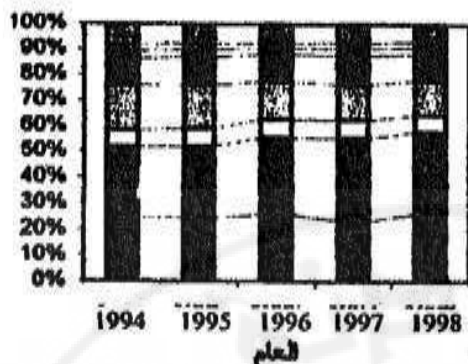


الزراعة	الصناعة والتعدين	البناء والتشييد
تجارة الجملة والمغزى	النقل والمواصلات والتخزين	السكن والتأمين والمعارف
خدمات المجتمع والشخصية	الخدمات الحكومية	الهيئات التي لا تهدف إلى الربح

الشكل رقم (3-8) الناتج الإجمالي حسب القطاعات
للأعوام 1994 - 1998 (بأسعار 1995 الثابتة ، مليون ليرة)

مصدر البيانات : المجموعة الإحصائية السورية لعام 2001 .

10- تعتبر الأعمدة البيانية المقسمة أكثر فائدة في تلك الحالة التي نريد التركيز فيها على مقارنة التغيرات التي تطرأ على تركيب الظاهرة، أو ما يطلق عليه البنية النوعية للظاهرة، وفي هذه الحالة يفضل استخدام الأعمدة البيانية المقسمة على أساس مئوي (%). وكمثال على هذه الحالة نعرض بيانات الجدول السابق رقم (3-4) بأعمدة بيانية مقسمة ولكن على أساس مئوي. لاحظ في الشكل رقم (3-9) أن الأعمدة البيانية متساوية في ارتفاعاتها حيث ارتفاع كل عمود 100 % .



الزراعة	الصناعة والبناء	الخدمات والتجارة
الغابات والصيد	النقل والاتصالات	الضرائب على المنتجات

الشكل رقم (3-9)

تركيب الإنتاج الإجمالي في سورية حسب القطاعات للأعوام 1994 - 1998

مصدر البيانات : المجموعة الإحصائية السورية لعام 2001 .

2-2-2-3 - الرسوم الدائرية Circle Charts :

الرسوم الدائرية شكل من أشكال العرض البياني، وفيه نعر عن الظواهر بمساحات

دائرية وتستخدم في الحالات الآتية :

1 - مقارنة عدة ظواهر مع بعضها، وهنا نعر عن كل ظاهرة بدائرة مساحتها تتناسب مع حجم الظاهرة، وتتم معرفة نصف قطر كل دائرة بأحد نسبة أحد نصفي الأقطار إلى الآخر الذي يجب أن يساوي جذر مساحة إحدى الدوائر إلى الأخرى على الترتيب، وذلك لأن مساحة الدائرة تتناسب مع مربع نصف قطرها.

مثال: لمقارنة إنتاج القطن في سورية في عام 1996 والبالغ 760 ألف طن، مع إنتاج القطن في عام 2000 والبالغ 1082 ألف طن (مصدر البيانات: المجموعة الإحصائية لعام 2001). تمثل إنتاج عام 1996 ودائرة تمثل إنتاج 2000، وهنا نحري حساب نصف قطر الدائرة الواجب رسمها:

$$\frac{R_2}{R_1} = \sqrt{\frac{1082}{760}} \approx 1.2$$

أي يجب أن تكون نسبة قطري الدائرتين هي 1.2 ، فلو اخترنا نصف قطر دائرة 2 سم لإنتاج عام 1996 يجب أن يكون نصف قطر الدائرة لعام 2000 هو 2.4 سم.

ما يعاب على هذا الأسلوب في هذه الحالة هو صعوبة مقارنة مساحة الدوائر مع بعضها عندما تعرض جاهزة، ولهذا يندر استخدام الدوائر في هذه الحالة.

2 - تستخدم الدوائر لبيان المكونات الجبرية لظاهرة ما، حيث تمثل للظاهرة بدائرة وتقسّم إلى قطاعات تتناسب مساحاتها مع حجم الأجزاء المكونة للظاهرة وتلون القطاعات بألوان مختلفة لتمييزها عن بعضها مع الانتباه إلى وضع وسيلة إيضاح تبين ماذا يمثل كل لون. الشكل رقم (3-10). ويتم تحديد مساحة كل جزء عن طريق تقسيم الزوايا المركزية للدائرة والبالغة 360° إلى أجزاء، حيث مجموع الأجزاء الظاهرة يقابل 360°، وبالتناسب نستطيع حساب زاوية كل قطاع من بيانات الجدول (3-4) نرى أن حصة قطاع الزراعة لعام 1998 تبلغ 308182 مليون ليرة في حين تبلغ قيمة الناتج الإجمالي لذلك العام 1176527 مليون ليرة، وعليه تكون زاوية القطاع المثلثة لحصة الزراعة هي :

$$\begin{array}{ccc} 1176527 & \text{تقابل} & 360 \\ \times & \text{تقابل} & \times \\ 308182 & & \\ \hline x = \frac{308182}{1176527} \times 360^\circ = 94.3^\circ \end{array}$$

أي زاوية القطاع المثلث لحصة الزراعة تبلغ 94.3° انظر الشكل رقم (3-10) و (3-11) وهكذا نحسب الزوايا كلها .

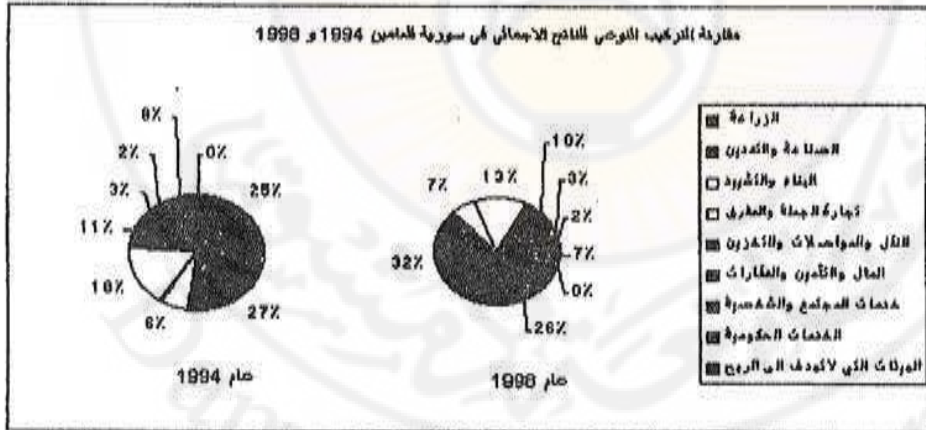


الزراعة	الصناعة والتعدين
بناء وتشهيد	تجارة الجملة والمفرق
النقل والمواصلات والتخزين	المال والتأمين والمعارف
خدمات المجهز والشخصية	الخدمات الحكومية
المنتجات التي لا تهدف الى الربح	

الشكل رقم (3-10) الناتج الإجمالي حسب القطاعات لعام 1998 (باسعار 1995 بملايين الليرات السورية)

مصدر البيانات : المجموعة الإحصائية السورية لعام 2001 .

3 - تعتبر الرسوم الدائرية أكثر فائدة عندما ينصب اهتمامنا على البنية النوعية للظاهرة وتطورها حيث تمثل الأجزاء بشكل متوي (%) كما في الشكل (3-11) .



الشكل رقم (3-11)

مصدر البيانات : المجموعة الإحصائية السورية لعام 2001 .

الجدول رقم (3-9) الجدول التكراري للدرجات التي حصل عليها
طلاب عينة عشوائية حجمها 300 طالب في مادتي الرياضيات والإحصاء

الإحصاء	الرياضيات										المجموع
	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100	
0-10	2	-	-	-	-	-	-	1	-	-	3
10-20	3	5	-	-	-	-	-	5	-	-	13
20-30	-	1	10	20	-	-	-	-	-	-	31
30-40	6	-	5	10	20	6	-	-	-	-	47
40-50	-	-	-	5	8	21	-	-	-	-	34
50-60	2	-	-	-	30	20	10	-	-	-	62
60-70	-	-	-	-	-	30	20	8	7	-	65
70-80	-	2	-	5	-	-	-	15	6	-	28
80-90	-	-	-	-	-	4	-	4	4	1	13
90-100	-	-	-	-	-	-	-	-	2	2	4
300	13	8	15	40	58	81	30	33	19	3	300

لاحظنا من الجدول السابق بأنه إذا أخذنا العمود الأول من اليسار مع العمود الأخير فإننا نحصل على الجدول التكراري للدرجات التي حصل عليها الطلاب في مادة الإحصاء، أما إذا أخذنا السطر الأول مع السطر الأخير فإننا نحصل على الجدول التكراري للدرجات التي حصل عليها الطلاب في مادة الرياضيات.

مزايا ومساوئ الجدول التكراري :

يساعد الجدول التكراري على اختصار البيانات بشكل كبير، حيث نستطيع حصر عدد غير متناه من المفردات في جدول واحد وبصفحة واحدة، كما يساعدنا الجدول على ملاحظة مركز القيم في الجدول وملاحظة انتشارها. ولكن بالإضافة إلى هذه المزايا هناك سيطرة للجدول التكراري يجب الانتباه إليها حين إعداده ، وكذلك عند تحليل البيانات، وهي أنه تضعيع معالم المفردات في الجدول التكراري وتعطى قيمة جديدة تمثل وسط الفئة التي تنتمي إليها المفردة، ونتيجة ذلك تختلف المقاييس

من مراجعة البيانات نرى أن عدد أفراد الأسر هو بين 2 و 8 فنقول إن لدينا 7 مجموعات، ننشئ جدولاً من ثلاثة أعمدة نضع في العمود الأول قيم المتغير .

جدول رقم (3-5)

عدد الأسر f	التفريع	المتغير عدد أفراد الأسرة
1	I	2
2	II	3
3	III	4
6	I II II I	5
1	I	6
4	II II I	7
3	III	8
20		المجموع

بعد رسم الجدول ووضع قيم المتغير فيه، نبدأ بتفريع البيانات الختام، وذلك بقراءتها مفردة مفردة ووضع إشارة في العمود الثاني مقابل القيمة المساوية لقيمة المفردة كما في الجدول رقم (3-5)، وعندما يبلغ عدد الإشارات 5 نشكل منها حزمة بوضع الإشارة الخامسة بشكل مائل. أخيراً نعد عدد الإشارات المقابلة لكل قيمة، ونكتبها بشكل عشري في العمود الأخير عمود التكرارات لكل قيمة Frequency، والذي سنرمز له من الآن فصاعداً بالرمز F أو (f) ، وبذلك نكون قد حصلنا على الجدول التكراري، والذي سوف يعرض فيما بعد بعمودين فقط الأول والثالث ويحذف عمود التفريع .

- توبيب المتغيرات الكمية المستمرة :

يعتبر توبيب المتغيرات الكمية المستمرة أكثر تعقيداً من المنقطعة، حيث في المتغيرات المستمرة عدد القيم التي سوف يأخذها المتغير غير محدود، وبالتالي ستكون هناك صعوبة بتقسيم البيانات إلى مجموعات، وسنورد فيما يلي خطوات إنشاء الجدول التكراري:

1 - ترتيب البيانات الخام بشكل تصاعدي أو تنازلي :

ويفيد ذلك في تسهيل التعرف على أكبر مفردة وأصغر مفردة وبالتالي المدى العددي للبيانات حيث إن المدى (Range) يساوي :

$$\text{المدى} = \text{أكبر مفردة} - \text{أصغر مفردة}$$

عدا ذلك سوف يساعد الترتيب في معرفة مدى اختلاف البيانات أو تجانسها، وبالتالي يساعد على تقسيمها إلى مجموعات، وكذلك عند تفرغ البيانات، هذا بالإضافة إلى أن الترتيب يعتبر شكلاً مفيداً من أشكال العرض يكون ذا فائدة كبيرة إذا كان عدد المفردات قليلاً، أما إذا كان عدد المفردات كبيراً فهو عديم الفائدة ، لأنه لا يختصر البيانات .

2 - حساب عدد الفئات وطول كل فئة :

عما أن عدد القيم التي يأخذها المتغير غير محدودة أو كبيرة نقوم بتقسيم المدى الذي أوجدناه أعلاه إلى فئات Classes، أو فترات Intervals، وليس هناك من قاعدة ثابتة لتحديد طول الفئة، وبالتالي عدد الفئات، ولكن نقول يجب ألا يكون عدد الفئات كبيراً وبالتالي تضعيف الفائدة المرجوة في الجدول وهي اختصار البيانات، كما أنه إذا كان عدد الفئات صغيرة تضعيف معالم التوزيع.

يقترح بعض المؤلفين أن يكون عدد الفئات بين 10 و 25 فئة، حيث يتم حساب طول الفئة بتقسيم المدى على عدد الفئات. ولكن يفضل استخدام طريقة سترجنس

(Sturges) في تحديد طول الفئة، نظراً لأنها تعتبر أكثر ملائمة، وتأخذ المدى، وعدد المفردات بعين الاعتبار، وبحسب طول الفئة وفق العلاقة الآتية :

$$C = \frac{\text{Range}}{1 + 3.32 (\text{Log } n)}$$

حيث n عدد المفردات الخام الواجب تبويبها و C طول الفئة و Range المدى وباقي المعادلة هي ثوابت، وبعد حساب طول الفئة يمكن استنتاج عدد الفئات، وذلك بتقسيم المدى على طول الفئة، ويجب مراعاة أن يكون طول الفئة عدداً صحيحاً، وفي حال استخدام علاقة سترجس يمكن تقريب طول الفئة.

3 - اختيار حدود الفئات وطريقة كتابتها :

يؤثر اختيار حدود الفئات كثيراً على نتائج تحليل البيانات، والجدول الصحيح أو الجدول السوي هو ذلك الجدول الذي يعطي نتائج مطابقة أو قريبة جداً من النتائج التي يمكن حسابها من البيانات قبل التبويب، ويمكن الحصول على جدول سوي باختيار حدود الفئات بشكل صحيح، وهذه العملية معقدة وتتطلب عدداً كبيراً من الحسابات والمقارنات وذلك لأن القيم في كل فئة ستعامل وكأن لها قيمة منتصف الفئة، لذلك لا بد من اختيار حدود الفئات بحيث يكون الوسط الحسابي لمفردات كل فئة يساوي قيمة منتصفها. أو يمكن بطريقة أخرى مراجعة الترتيب واختيار حدود الفئات بحيث يكون منتصفها يساوي القيم الأكثر تكراراً .

بعد اختيار حدود الفئات نختار إحدى الطرائق الآتية لكتابة حدود الفئات. والجدول رقم (3-6) يبين كيفية كتابة حدود الفئات .

جدول رقم (3-6)

1	2	3	4	5
10 إلى أقل من 5	9 إلى 5	5 -	5 - 10	5 - 9
15 إلى أقل من 10	14 إلى 10	10 -	10 - 15	10 - 14
20 إلى أقل من 15	19 إلى 15	15 -	15 - 20	15 - 19
أو				
10 إلى ولكن لا تشمل 5				
15 إلى ولكن لا تشمل 10				
20 إلى ولكن لا تشمل 15				

يتضح من الأشكال السابقة لطرائق كتابة حدود الفئات بأن الطريقة الأولى والثانية القيم المنتمة إلى هذه الفئات واضحة كل الوضوح، ففي الطريقة الأولى تدخل المفردات التي قيمها تساوي الحد الأدنى، وأكبر منه إلى أقل من الحد الأعلى، أي هي مجال من الشكل [5 و 10]. أما في الطريقة الثانية فالمجال من الشكل [5 و 9] يدخل الحد الأدنى والأعلى، وتستخدم هذه الطريقة في حال المتغيرات المنقطعة، ولتسهيل كتابة حدود الفئات نفضل عدم ذكر العبارات الدالة على القيم الداخلة في الفئة كما في الطريقة الأولى والثانية، ونكتفي بكتابة حدود الفئات وبينها إشارة معترضة. انظر العمودين الرابع والخامس في الجدول (3-6) .

ونجرب العرف على قراءة حدود الفئات كما يأتي : إذا تكررت كتابة الحد الأعلى للفئة كمحد أدنى للفئة التالية، فهذا يدل أن المجال من الشكل [5 و 10]، وإذا لم يتكرر كتابة الحد الأعلى في الفئة التالية هذا يدل أن المجال في الشكل [5 و 9] ويمكن التخلص من هذه الصعوبات بكتابة الحد الأدنى للفئة فقط، كما في العمود الثالث الجدول رقم (3-6) وهنا نقرأ الفئة بأنها من 5 فما فوق إلى أقل من 10، والثانية 10 فما فوق إلى أقل من 15 وهكذا ...

4 - التفريغ :

بعد تحديد عدد الفئات واختيار طريقة كتابتها نرسم جدولاً من ثلاثة أعمدة، في العمود الأول نكتب حدود الفئات، وفي العمود الثاني نفرغ البيانات بوضع إشارات لكل مفردة إشارة كما مر معنا أعلاه، وفي العمود الثالث نترجم الإشارات إلى أعداد عشرية ونضع له عنوان: عدد التكرارات f أو F .

ولتبيان كيف تتم عملياً الخطوات السابقة نورد المثال الآتي :

مثال (3-2) :

لدينا البيانات التالية عن أطوال 50 طالب من طلاب كلية الاقتصاد المذكور :

176	173	171	165	167	181	158	179	150	171
178	177	170	165	171	185	157	183	155	171
185	177	171	163	171	190	165	183	165	171
175	171	172	164	183	187	161	171	159	177
182	171	169	166	177	190	165	171	160	177

الحل :

1 - نرتب البيانات ترتيباً تصاعدياً لمعرفة مدى البيانات، ولتساعدنا الترتيب في

عملية التفريغ :

185	181	177	173	171	171	167	165	160	150
185	182	177	175	171	171	169	165	161	155
187	183	177	176	171	171	170	165	163	157
190	183	178	177	171	171	171	165	164	158
190	183	179	177	172	171	171	166	165	159

$$\text{سم } \text{Rang} = 190 - 150 = 40$$

2 - نحسب طول الفئة من علاقة سترجس :

$$C = \frac{\text{Range}}{1 + 3.32 (\text{Log } n)}$$

$$= \frac{40}{1 + 3.32 (\text{Log } 50)} = \frac{40}{1 + 3.32 (1.689)} = \frac{40}{6.644} = 6.02$$

وندوره إلى العدد 6 فيكون هو طول الفئة .

3 - نكتب حدود الفئات حيث سنختار أصغر مفردة كحد أدنى للفئة الأولى،

ونضيف إليها طول الفئة 6 على التوالي فنحصل على حدود الفئات.

وسوف نختار كتابة الحدود بالشكل :

$$150 - 156$$

$$156 - 162$$

ويتضح هنا أن المجال هو [150 - 156] كما بينا أعلاه .

جدول رقم (7-3)

فئات الطول سم	التفريع	التكرار f
150 - 156	11	2
156 - 162	1111	5
162 - 168	1111 1111	9
168 - 174	1111 1111 1111	15
174 - 180	1111 1111	9
180 - 186	1111 11	7
186 - 192	111	3
المجموع		50

أوساط الفئات :

عند إعداد الجدول التكراري ونقل المفردات تضيع المعالم الأصلية للبيانات، وكل ما يفيدنا به الجدول هو أن المفردة تنتمي إلى هذه الفئة أو تلك، كذلك سوف نفترض أن قيمة كل المفردات في فئة ما متساوية وتساوي وسط تلك الفئة، ويتم حساب وسط الفئة كما يأتي :

$$\text{وسط الفئة } x_i = \frac{\text{الحد الأدنى للفئة} + \text{طول الفئة}}{2}$$

وهذه القاعدة تصلح لكل أشكال كتابة حدود الفئات وبحسب طول الفئة بأنه الفرق بين حدين متتاليين: الحد الأدنى للفئة الثانية - الحد الأدنى للفئة الأولى = طول الفئة الأولى، وكذلك الحد الأعلى للفئة الثانية - الحد الأعلى للفئة الأولى = طول الفئة الثانية وهكذا .

$$\frac{\text{الحد الأعلى} + \text{الحد الأدنى}}{2} = x_i \text{ ويمكن ببساطة حساب طول الفئة :}$$

إذا كانت الفئات هي مجالات من الشكل [150 - 156] أي مفتوحة من الأعلى وطول الفئة عندها يكون: الحد الأعلى - الحد الأدنى .

التوزيع المنتظم وغير المنتظم :

إن حساب طول الفئة حسب علاقة سترجس أو بطريقة أخرى تحدد عدد الفئات بين 10 - 25 ، يؤدي إلى الحصول على فئات متساوية الطول للجدول كما في الجدول (3-7)، ويدعى التوزيع التكراري في هذه الحالة توزيعاً منتظماً، ولكن في بعض الحالات نرى أن هناك فئات فارغة من المفردات أو قليلة المفردات فندمجها مع بعضها، فنحصل على جدول بأطوال فئات غير متساوية الطول، وهذه الحالة يدعى التوزيع غير المنتظم، ويفضل أن يكون الجدول بأطوال فئات متساوية لتسهيل الحسابات، وتسهيل التمثيل البياني للتوزيع كما سنرى لاحقاً .

الجدول المغلق والجدول المفتوح :

يدعى الجدول التكراري مغلقاً إذا عُلمت حدود كل الفئات، أما إذا وجد في الجدول التكراري أن الفئة الأولى بدون حد أدنى دعي الجدول مفتوحاً من الأدنى، ويدعى الجدول مفتوحاً من الأعلى إذا كانت آخر الفئات بدون حد أعلى، ويدعى الجدول مفتوحاً من الطرفين إذا كانت الفئة الأولى والأخيرة مفتوحتين، وجعل أحد

الحدود غير معلوم قد يكون بسبب تباعد البيانات في الفئة الأولى والأخيرة، مثلاً: توزيع الدخل، حيث نعلم أن أكثر الدخل في حدود معينة، وهناك بعض الدخل تكون كبيرة فيفضل في هذه الحالة ترك الحد الأعلى للفئة الأخيرة مفتوحاً، وقد يكون سبب جعل أحد الحدود مفتوحاً هو سرية البيانات.

من عيوب الجداول التكرارية المفتوحة هو أننا لا نستطيع إيجاد بعض المقاييس الإحصائية، كما أننا لا نستطيع إيجاد التمثيل البياني للتوزيع كما سنرى لاحقاً.

التكرار التجميعي Cumulative Frequency :

يبين الجدول التكراري عدد المفردات في كل فئة، ولكن قد نرغب بمعرفة عدد المفردات التي تقل عن حد معين، أو عدد المفردات التي تزيد عن حد معين، فإننا نحصل على هذا بإيجاد التكرار التجميعي، والتكرار التجميعي صاعد وهابط، نعرف التكرار التجميعي الصاعد بأنه عدد المفردات في التوزيع التي تقل عن الحد الأعلى لكل فئة، ونحصل عليه بإضافة تكرار الفئات إلى بعضها على التوالي.

أما التكرار التجميعي الهابط فيعرف بأنه عدد المفردات التي قيمها أكبر أو تساوي الحد الأدنى لكل فئة، ونحصل عليه بطرح تكرار كل فئة في المجموع بدءاً من أول فئة، انظر الجدول رقم (3-8) :

الجدول رقم (3-8)

يثل الدخل الشهري لـ 200 أسرة من إحدى الدول العربية

فئات الدخل ألف وحدة نقدية	التكرار f	وسط الفئة x	التكرار التجميعي الصاعد	التكرار التجميعي الهابط	التكرار النسبي fr	التكرار النسبي % fr. 100	التكرار التجميعي الصاعد النسبي	التكرار التجميعي الهابط النسبي
1 - 3	20	2	20	200	0.1	10	0.1	1
3 - 5	25	4	45	180	0.125	12.5	0.225	0.9
5 - 7	30	6	75	155	0.15	15	0.375	0.775
7 - 9	50	8	125	125	0.25	25	0.625	0.625
9 - 11	40	10	165	75	0.2	20	0.852	0.375
11 - 13	20	12	185	35	0.1	10	0.925	0.175
13 - 15	15	14	200	15	0.075	7.5	1	0.075
المجموع	200					100		

كما يمكن أن يكون التكرار التجميعي بشكل نسبي، ونحصل عليه بتقسيم التكرار التجميعي على مجموع التكرارات، فالتكرار التجميعي الصاعد النسبي يعطي نسبة المفردات التي تقل قيمتها عن الحد الأعلى لكل فئة، أما التكرار التجميعي الهابط النسبي فيعطي نسبة المفردات التي قيمتها تساوي أو أكبر من الحد الأدنى لكل فئة.

التكرار النسبي Relative Frequency :

نحصل على التكرار النسبي بتقسيم تكرار كل فئة على مجموع التكرارات وهو يعطي نسبة المفردات التي قيمتها تنتمي لكل فئة، ويمكن التعبير عنه بشكل عشري أو بشكل مئوي % وذلك بضرب التكرار النسبي بـ 100. انظر الجدول السابق رقم (8-3). وللتكرار النسبي تطبيقات هامة سنراها عند دراسة الاحتمال والتوزيعات الاحتمالية.

الجدول التكراري المزدوج Double Frequency Table :

بينما أعلاه كيف يتم تبويب المتغيرات الكمية حسب متغير واحد فقط، ولكن قد نحتاج في بعض الأحيان إلى تبويب المتغيرات حسب متغيرين (صفتين) كأن نبوب بيانات تمثل أطوال مجموعة من الأشخاص وأوزانهم، أو بيانات تمثل الدرجات التي حصل عليها مجموعة من الطلاب في مادتي الإحصاء والرياضيات...، وواضح هنا أنه يجب أن يكون لكل مفردة صفتان وتعرض في جدول يدعى جدول تكراري مزدوج .

يتم إعداد الجدول المزدوج باتباع الخطوات نفسها التي تمر بها عملية الجدول البسيط ، ولكن لكل صفة على حده، حيث نقوم بوضع فئات إحدى الصفات في الأسطر والصفة الأخرى بالأعمدة ، ثم نفرغ البيانات اعتماداً على الصفتين، انظر الجدول رقم (9-3) :

الجدول رقم (3-9) الجدول التكراري للدرجات التي حصل عليها
طلاب عينة عشوائية حجمها 300 طالب في مادتي الرياضيات والإحصاء

الإحصاء	الرياضيات										المجموع
	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100	
0-10	2	-	-	-	-	-	-	1	-	-	3
10-20	3	5	-	-	-	-	-	5	-	-	13
20-30	-	1	10	20	-	-	-	-	-	-	31
30-40	6	-	5	10	20	6	-	-	-	-	47
40-50	-	-	-	5	8	21	-	-	-	-	34
50-60	2	-	-	-	30	20	10	-	-	-	62
60-70	-	-	-	-	-	30	20	8	7	-	65
70-80	-	2	-	5	-	-	-	15	6	-	28
80-90	-	-	-	-	-	4	-	4	4	1	13
90-100	-	-	-	-	-	-	-	-	2	2	4
300	13	8	15	40	58	81	30	33	19	3	300

لاحظنا من الجدول السابق بأنه إذا أخذنا العمود الأول من اليسار مع العمود الأخير فإننا نحصل على الجدول التكراري للدرجات التي حصل عليها الطلاب في مادة الإحصاء، أما إذا أخذنا السطر الأول مع السطر الأخير فإننا نحصل على الجدول التكراري للدرجات التي حصل عليها الطلاب في مادة الرياضيات.

مزايا ومساوئ الجدول التكراري :

يساعد الجدول التكراري على اختصار البيانات بشكل كبير، حيث نستطيع حصر عدد غير متناه من المفردات في جدول واحد وبصفحة واحدة، كما يساعدنا الجدول على ملاحظة متركز القيم في الجدول وملاحظة انتشارها. ولكن بالإضافة إلى هذه المزايا هناك سمة خطيرة للجدول التكراري يجب الانتباه إليها حين إعداده ، وكذلك عند تحليل البيانات، وهي أنه تضعيع معالم المفردات في الجدول التكراري وتعطى قيمة جديدة تمثل وسط الفئة التي تنتمي إليها المفردة، ونتيجة ذلك تختلف المقاييس

المحسوبة من البيانات قبل التوبيخ عن المقاييس المحسوبة بعد توبيخ البيانات وللتخلص من هذه السببية رأينا سابقاً كيف يتم إعداد جدول تكراري صحيح أو سوي باختيار حدود الفئات بحيث أن وسط الفئة يساوي وسط القيم المنتمة إليها.

ويمكن التأكد من أن الجدول سوي باختيار الفرق بين الوسط الحسابي قبل التوبيخ والوسط الحسابي بعد التوبيخ، فإذا كان الفرق ظاهرياً أو بمعنى آخر قليلاً جداً اعتبرنا أن الجدول سوي .

2-3-3- العرض البياني للمتغيرات الكمية :

تعرض المتغيرات الكمية بيانياً بأحد أنواع العرض الآتية :

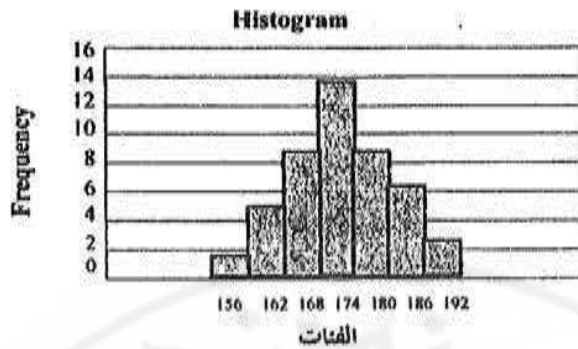
- المدرج التكراري .
- المضلع التكراري .
- المنحنى التكراري .

وسوف نرى كيف يمكن إنشاء كل نوع من أنواع العرض هذه .

- المدرج التكراري Histogram :

المدرج التكراري هو عرض بيانات الجدول التكراري على المحاور الإحداثية، حيث تمثل الفئات على محور السينات والتكرارات f على محور العيّنات، ونعبر عن تكرار كل فئة بمسقط عمودي قاعدته مساوية لطول الفئة، وارتفاعه يشير إلى عدد التكرارات وبذلك نحصل على مجموعة مستطيلة متلاصقة تشبه المدرج (الشكل رقم 12-3)، حيث مثلنا بيانات الجدول رقم (7-3) بيانياً بمدرج تكراري .

يساعد المدرج التكراري في معرفة الفئات التي تتمركز فيها المفردات ومدى انتشار أو اختلاف قيم المفردات في التوزيع، ولكن الفائدة الأهم التي يقدمها المدرج التكراري كما هو الحال في أساليب عرض المتغيرات الكمية الأخرى، هي أنه يبين شكل التوزيع التكراري عندما تكون أطوال الفئات متساوية ، أما عندما تكون أطوال



الشكل رقم (3-12) يمثل المدرج التكراري

الفئات غير متساوية فإن الشكل الناتج لا يعبر عن طبيعة أو شكل التوزيع، ولكي نحصل على الشكل الحقيقي للتوزيع يجب تعديل التكرارات f ، وذلك بتقسيم تكرار كل فئة على طولها فنحصل على عدد التكرارات لكل وحدة طول من الفئة، وهو ما ندعوه التكرار المعدل وسوف نرمز له بالرمز fc . لاحظ الفرق من خلال تمثيل بيانات الجدول (3-10) بالشكل (3-13). إن التمثيل البياني للتكرار f أو التكرار النسبي fr أو التكرار المعدل fc يعطي الشكل نفسه عندما تكون أطوال الفئات متساوية، والشكل الناتج هو الشكل الحقيقي للتوزيع التكراري، أما إذا كانت أطوال الفئات مختلفة فالشكل الناتج عن تمثيل التكرار المعدل fc يختلف عن الشكل الناتج عن تمثيل التكرار f أو التكرار النسبي fr .

بعد أن تعرفنا على التكرار المعدل نستطيع القول إن مساحة كل مستطيل من المستطيلات المشكلة للمدرج التكراري تساوي عدد تكرارات الفئة المقام عليها حيث Δ مساحة المستطيل تساوي القاعدة (طول الفئة) ضرب الارتفاع (التكرار المعدل):

التكرار المعدل للفئة i :

$$fc_i = \frac{f_i}{C_i}$$

مساحة المستطيل المقام على الفئة i :

$$fc_i \cdot C_i = f_i$$

ومساحة الشكل الناتج تساوي مجموع مساحات المستطيلات، وبالتالي تساوي مجموع التكرارات $\sum f_i$.

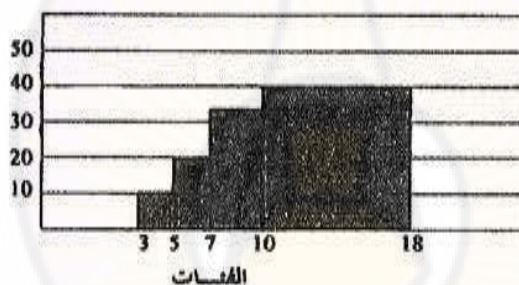
حيث $\sum$ تقرأ سيجما وترمز إلى المجموع أو يمكن أن تقرأ مباشرة مجموع f_i .

الجدول رقم (3-10)

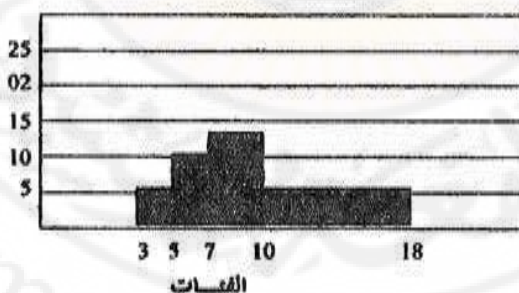
الدخل الشهري لعينة حجمها 106 أسرة من إحدى المدن العربية (ألف وحدة نقدية)

فئات الدخل	التكرار f	التكرار المعدل
3 - 5	10	5
5 - 7	20	10
7 - 10	36	12
10 - 18	40	5
المجموع	106	

التكرار f



التكرار المعدل f_c



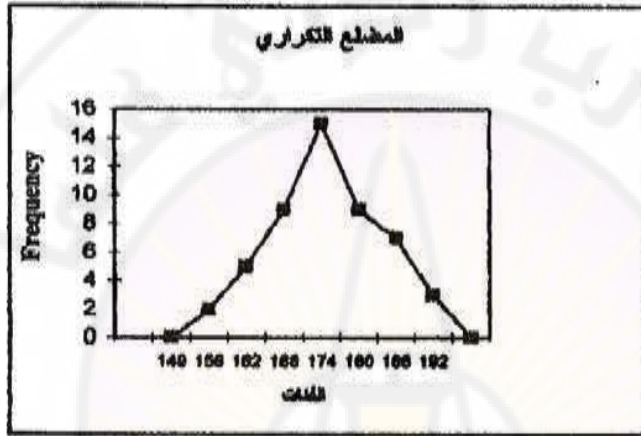
الشكل رقم (3-13)

يمثل مقارنة بين التمثيل البياني للتكرار f والتكرار المعدل f_c

- المضلع التكراري Frequency Polygon :

نحصل على المضلع التكراري بإحدى الطريقتين الآتيتين :

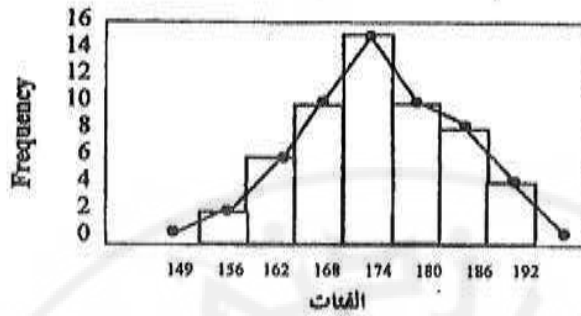
- 1- نمثل أوساط الفئات وفئات الجدول التكراري على المحاور الإحداثية، حيث نضع الفئات على محور السينات، والتكرارات على محور العيّنات، كما في الشكل (3-14) قمنا بتمثيل بيانات الجدول رقم (3-7) .



الشكل رقم (3-14)

- 2- الطريقة الثانية للحصول على المضلع التكراري هي كما يأتي: بعد رسم المدرج التكراري نصل بين نقاط أعلي ومتنصف المستطيلات كما في الشكل رقم (3-15)، فقد قمنا بالحصول على المضلع التكراري من المدرج التكراري في الشكل رقم (3-12)، لاحظ أنه في كلتا الطريقتين قمنا بإضافة فئتين وهميتين تكرر كل منهما يساوي الصفر، إحداهما قبل أول فئة والأخرى بعد آخر فئة، وذلك لإغلاق الشكل الناتج. كما يمكن من الشكل (3-15) ملاحظة أن مساحة الشكل الناتج (المضلع) تساوي تماماً مساحة المدرج، وتساوي مجموع التكرارات لأن مساحة المناطق التي تم ضمها إلى المدرج تساوي مساحة المناطق التي تم حذفها.

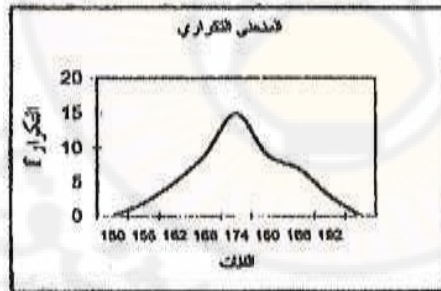
المضلع التكراري



الشكل رقم (3-15)

المنحنى التكراري Frequency Curve :

بعد رسم المضلع التكراري نقوم بتمهيد الخط المنكسر ليصبح منحنى الشكل رقم (3-16)، ويلاحظ أنه كلما كان عدد الفئات كبيراً وطولها قصيراً اقترب المضلع التكراري من المنحنى التكراري.



الشكل رقم (3-16)

ما يجب الانتباه إليه عند النظر إلى المنحنى التكراري، هو أن ارتفاعه يشير إلى التكرار دوماً، فابتعاد المنحنى عن محور السينات يدل على تكرارات أكثر واقتربه على تكرارات أقل.

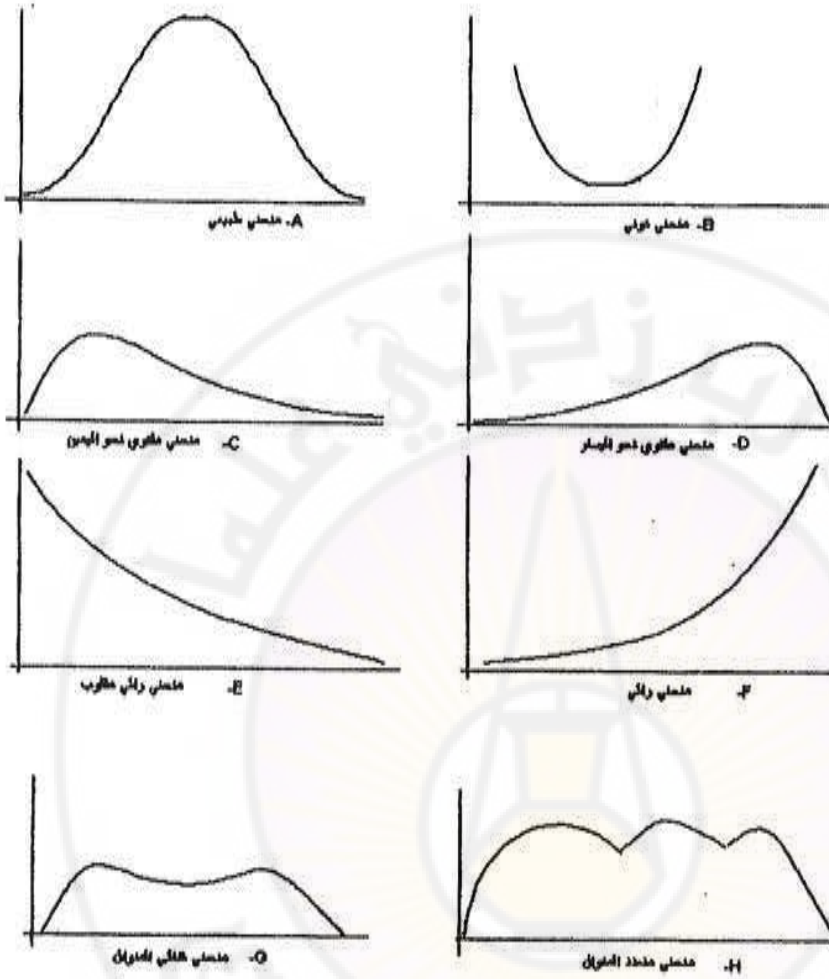
يساعد المنحنى التكراري في معرفة شكل التوزيع التكراري، وبالتالي نوع المعادلة الرياضية أو النموذج الرياضي الذي ينضج له التوزيع.

- أشكال المنحنيات التكرارية :

يُحوي الشكل رقم (3-17) مجموعة من الرسوم التي تعبر عن أهم أشكال المنحنيات التكرارية، بعض التوزيعات لها شكل الجرس المقلوب ويدعى بالمنحنى العظمى (Normal Curve) وهو منحنى متماثل حول العمود الذي يمر من نقطة لهافته العظيمة وهذا العمود يقسمه إلى قسمين متساويين، ومساحة كل قسم تساوي نصف مساحة المنحنى. الرسم A من الشكل (3-17). ويعتبر هذا الشكل من أهم الأشكال حيث ينضج له معظم الظواهر الطبيعية، لذلك دعي بالمنحنى الطبيعي، وكأمثلة على الظواهر التي تتوزع وفق التوزيع الطبيعي نذكر أطوال مجموعة من الناس وأوزانهم.... إلخ، لذلك سوف ندرسه بالتفصيل فيما بعد .

بعض التوزيعات لها شكل الحرف U ندعوه بالمنحنى النوبي، وذلك عندما تكون التكرارات العظمى على طرفي المنحنى والتكرارات الدنيا في الوسط مثل توزع الوفيات على الفئات العمودية، حيث تكثر الوفيات في الفئات العمرية الدنيا والعليا وتقل في الفئات المتوسطة الرسم B.

هناك بعد التوزيعات يكون منحناها غير متماثل تماماً كما في المنحنى الطبيعي، وإنما نجد لأحد نصفيه امتداد ذيل فندعوه (ملتوي) (Stewness). والالتواء إما موجب نحو اليمين عندما تكون القمة إلى اليسار والذيل يمتد نحو اليمين C، وهذه الحالة تكون المفردات ذات القيم الدنيا أكثر تكراراً، والقيم ذات القيم العليا أقل تكراراً. وإما يكون التوزيع ملتوي نحو اليسار (التواء سالب) وذلك عندما تكون القمة نحو اليمين والذيل يمتد نحو اليسار، وهذه الحالة تكون المفردات ذات القيم الدنيا أقل تكراراً، والمفردات ذات القيم العليا أكثر تكراراً الرسم D .



الشكل رقم (3-17)

بعض التوزيعات تكون شديدة الالتواء وهذه الحالة يدعى التوزيع الرائي، وفي التوزيع الرائي تكون أكثر التكرارات على أحد طرفيه مثل الرسم B حيث التكرارات العظمى على اليسار. ويدعى: المنحنى رأسي مقلوب، لأنه يشبه حرف الراء مقلوباً، وكمثال عليه توزيع الدخل وتوزيع الأراضي. أما إذا كانت التكرارات العظمى على اليمين دعى التوزيع بالتوزيع الرائي مثل توزيع مجموعة من الناس من حيث قوة البصر .

عندما تكون البيانات غير متجانسة مأخوذة من مجتمعين مختلفين أو أكثر، فإننا نحصل على منحنى ذي قيمتين (ثنائي المنوال Bimodal) أو (متعدد المنوال Multimode) الرسم G و H على التوالي، مع العلم أن التجانس يعتبر شيئاً مهماً لتطبيق الأساليب الإحصائية على البيانات، وعندما يكون لدينا منحنى بأكثر من منوال يمكن فصل البيانات لنحصل على عدة منحنيات، لكل منها قمة واحدة أو منوال واحد .

- خصائص وصفات التوزيعات التكرارية :

عندما تكون البيانات الإحصائية متجانسة ومأخوذة من مجتمع إحصائي واحد، فإن التوزيع يستع بالخصائص الآتية :

1 - هناك قيمة في التوزيع تتجمع عندها القيم، وهذا ما يدعى بالنزعة المركزية (Central Tendency) وتقاس النزعة المركزية بمقاييس النزعة المركزية، والتي سندرسها بالتفصيل في الفصل القادم.

2 - بالإضافة إلى تجمع القيم حول قيمة معينة يبدى التوزيع اختلافاً في القيم، ويبدو ذلك باعتداد طرفي المنحنى عن الوسط، وهذا ما يدعى بالتشتت Dispersion ويقاس بمقاييس التشتت وسندرسها لاحقاً أيضاً .

3 - بعض التوزيعات لها شكل التوزيع الطبيعي وبعضها ملتوي الرسوم A و C و D شكل رقم (17-3).

4 - تختلف التوزيعات من حيث شدة التمرکز بعض التوزيعات شديدة التمرکز وقليلة التشتت يدعى توزيعاً متطاولاً (Leptokurtic) ، وبعضها على العكس قليل التمرکز وكثير التشتت يدعى توزيعاً مفرطحاً (Platykurtic)، وهذا ما سوف نناقشه في الفصل السادس الخاص بمقاييس الالتواء والتفرطح .

تمارين ومسائل الفصل الثاني والثالث

1 - بين ما هو نوع الجدول الواجب استخدامه لعرض المتغيرات الآتية :

- عدد أشجار الزيتون في محافظات القطر 2000 .
- إنتاج عينة حجمها 500 شجرة زيتون .
- الدخل الشهري لـ 200 عامل .
- عدد الأشجار المثمرة حسب كل صنف في محافظات القطر .
- عدد الأشجار المثمرة حسب كل صنف وحجم الإنتاج لها في محافظات القطر .

2 - بين ما هو أسلوب العرض البياني الممكن استخدامه لعرض المتغيرات الآتية :

- عدد أشجار الزيتون حسب محافظات القطر لعام 2000 .
- إنتاج عينة حجمها 500 شجرة زيتون .
- الدخل الشهري لـ 200 عامل .
- عدد الأشجار المثمرة حسب أنواعها في محافظة حلب لعام 1995 .
- عدد طلاب كلية الاقتصاد للأعوام 1990 حتى 2000 .
- عدد الأشجار المثمرة في محافظة حلب للأعوام 1990 حتى 2000 .
- مقارنة عدد الطلاب في السنة الثانية حسب اللغة الأجنبية التي يدرسوها لعام 1995 .
- مقارنة عدد الطلاب في السنة الثانية حسب اللغة الأجنبية التي يدرسوها للأعوام 1990 حتى 2000 .

3 - قم بعرض البيانات الآتية بيانياً باستخدام الأساليب الممكنة :

- حجم إنتاج الزيتون والبالغ 423.3 ألف طن وحجم إنتاج التفاح والبالغ 224 ألف طن في الجمهورية العربية السورية للعام 1995 .

- حجم إنتاج الزيتون للأعوام 1995 - 2000 (ألف طن) .

العام	1995	1996	1997	1998	1999	2000
الإنتاج	423.3	647.6	403	785	400.5	866.1

المصدر : المجموعة الإحصائية لعام 2001.

- عدد أشجار الزيتون وحجم الإنتاج للأعوام 1995 - 2000

العام	1995	1996	1997	1998	1999	2000
الإنتاج ألف طن	423.3	647.6	403	785	400.5	866.1
عدد الأشجار ألف شجرة	54215	56903	59739	62300	63379	64344

المصدر : المجموعة الإحصائية لعام 2001.

4 - قم بإنشاء الجدول التكراري للدخل الشهري لعينة حجمها 50 عاملاً حيث كانت
دعوتهم كما يلي (وحدة نقدية) :

4730	6230	4715	5230	5240	4220	5230	4240	4730	5230
5220	6230	4715	5220	3480	5240	5240	5230	5000	4730
5220	4230	5460	5730	4220	6660	3730	5700	4730	5720
5230	5730	5730	4230	4740	5000	5760	5460	6260	6890
4720	6230	4720	4745	6730	4245	6230	4730	6200	5740

5 - أوجد التكرار التجميعي الصاعد والهابط والتكرار النسبي، وكذلك التكرار
التجميعي الصاعد والهابط النسبيين للجدول التكراري الآتي والممثل للمبيعات
الشهرية لـ 200 متجر من متاجر التجزئة، ومن ثم أوجد المدرج التكراري
والمضلع التكراري والمنحنى التكراري .

عدد المتاجر	المبيعات الشهرية ألف وحدة نقدية
5	20 - 30
10	30 - 40
30	40 - 50
40	50 - 60
65	60 - 70
30	70 - 80
20	80 - 90
200	المجموع

6 - حدد طبيعة التوزيع للبيانات الآتية من خلال المنحنى التكراري :

عدد الأسر	الدخل الشهري ألف وحدة نقدية
10	5 - 9
20	9 - 13
30	13 - 17
40	17 - 25
50	25 - 37
150	المجموع

الفصل الرابع

مقاييس النزعة المركزية . المتوسطات .

Central Tendency Measures «Introduction»

- مقدمة .

- الوسط الحسابي .

حساب الوسط الحسابي لبيانات إحصائية غير مبوبة .

حساب الوسط الحسابي لبيانات إحصائية مبوبة في فئات .

خصائص الوسط الحسابي .

بعض الأساليب غير المباشرة في إيجاد قيمة بعض الخصائص الجبرية الأخرى

للوسط الحسابي .

محاسن ومساوئ الوسط الحسابي .

- الوسيط .

حساب الوسيط للبيانات غير المبوبة .

حساب الوسيط للبيانات المبوبة في فئات .

العشريات والربيعات والمئينات

خصائص الوسيط ومحاسنه

- المتوال

حساب المتوال للبيانات غير الموزونة
حساب المتوال للبيانات الموزونة في فئات
محاسن ومساوئ المتوال
العلاقة التقريبية بين الوسط الحسابي والوسط والمتوال

- الوسط الهندسي

حساب الوسط الهندسي للبيانات غير الموزونة
حساب الوسط الهندسي للبيانات الموزونة في فئات
محاسن ومساوئ الوسط الهندسي

- الوسط التوافقي

حساب الوسط التوافقي للبيانات غير الموزونة
حساب الوسط التوافقي للبيانات الموزونة في فئات
محاسن ومساوئ الوسط التوافقي

1-4- مقدمة :

درسنا سابقاً الطرائق المتعددة لجمع البيانات وتلخيصها في جداول تكرارية، وكيفية عرضها بيانياً لتصبح أسهل للقارئ، لهذا سوف نتناول في هذا الفصل موضوع أساليب التحليل الإحصائي وإيجاد المقياس المناسب الذي يمثل الظاهرة موضوع الدراسة، واستخدامه للمقارنة بين تلك الظاهرة والظواهر الأخرى .

فإذا نظرنا إلى مفردات أي ظاهرة مدروسة، نجد أن أكثر هذه المفردات تتجمع حول قيمة ما تمثل متوسط الظاهرة، ونجد أيضاً بأن هناك تناقصاً تدريجياً في عدد المفردات كلما بعدت هذه القيم عن المتوسط، ومن الجانبين، وإن هذا التجمع حول القيمة المتوسطة هو ما يسمى بـ (النزعة المركزية) أي ميل المفردات إلى الاقتراب والتجمع حول قيمة معينة هي القيمة المتوسطة، وكان القيمة المتوسطة تحاول شد المفردات لكي تتراكم حولها. لكن لا تحدث هذه الظاهرة في جميع التوزيعات التكرارية فمن الطبيعي ألا تنجح جميع المفردات في الاقتراب من القيمة المتوسطة، فهناك مفردات تبتعد بشكل بسيط عن المتوسط، وهناك مفردات عددها أقل تبتعد بشكل أكبر عن المتوسط، وهكذا نجد أن العدد الذي يختلف عن المتوسط اختلافاً كبيراً جداً (بالزيادة أو النقصان) يكاد يتلاشى، فمثلاً الأقزام والعمالقة لا نشاهدهم إلا نادراً في حياتنا اليومية،

ولكننا نجد أن معظم الناس متوسطو الطول، وعدد قليل من الناس يختلف طولهم عن متوسط الطول بالزيادة أو النقصان.

وبشكل آخر نجد أن هناك خاصيتين أساسيتين للمعطيات الإحصائية تساعد على إعطاء تصور ومدلول يصف تلك المعطيات هي : النزعة المركزية ومقاييسها متمثلة بالمتوسطات التي بواسطتها يتمكن من إيجاد تلك القيمة التي تتمحور حولها كل القيم. أما الخاصية الثانية فهي مقاييس التشتت والتي توجد حالة الانتشار أو البعد للمعطيات عن المتوسط .

وهناك عدة أنواع من المتوسطات لكل منها طريقته الخاصة في الحساب وأهم هذه الأنواع: الوسط الحسابي - الوسيط - المنوال - الوسط الهندسي - الوسط التوافقي - الوسط التربيعي ، إلا أن المتوسط الأخير نادر الاستعمال .

«ومع الأسف، فإن مجرد القيام بحساب أحد المتوسطات الإحصائية قد لا يؤمن نتيجة ذات مدلول كما أن مختلف أنواع المتوسطات ليست ناجحة بصورة متساوية في وصف الحالة النموذجية لمختلف أنواع التوزيعات التكراري. وهكذا يقيس كل متوسط، أو يحدد، نقطة النزعة المركزية بطريقة مختلفة عن غيرها. وإذا لم يكن التوزيع متماثلاً بشكل تام فإن شكل متوسط من المتوسطات يعطي نتيجة مختلفة عن غيرها، ولذلك على الإحصائي أن يقوم باختيار المتوسط الملائم الذي يمكنه من تأمين المقياس الذي يصف بأفضل شكل النزعة المركزية لكل نوع من أنواع البيانات الإحصائية، علماً بأنه لا يوجد متوسط تام بل إن كل متوسط من المتوسطات يستعمل في مجالات معينة، وله محاسنه ومخاذهره، وبالتالي فإن اختيار المتوسط الصحيح لحالة ما يجب أن يكون بعد دراسة جدية لميزات كل واحد منها»⁽¹⁾. وأن استخراج قيمة المتوسطات يكون إما من قيم غير مبوبة (Ungrouped data) أي بشكل وحدات لكل منها قيمتها الخاصة أو من قيم مبوبة (Grouped data) في جداول تكرارية بحيث كل تكرار يمثل عدد الوحدات التي تقع ضمن فئة معينة، لها لهاتان (أو حدان) دنيا وعليا دون تحديد القيم الفعلية لتلك الوحدات .

(1) مبادئ الإحصاء - د. ناظم حيدر : 1987 - 1988 - ص 65-66 .

2-4- الوسيط الحسابي (The Arithmetic Mean) :

الوسيط الحسابي لقيم متغير ما هو القيمة الناتجة من قسمة مجموع تلك القيم على عددها، ويرمز له بـ $\bar{X}$ (اكس بار) وتعتمد طريقة حساب الوسيط الحسابي على نوعية البيانات من حيث كونها غير مبوبة أو مبوبة :

1-2-4- حساب الوسيط الحسابي لبيانات إحصائية غير مبوبة :

إذا كان لدينا n من القيم :

$$x_1, x_2, \dots, x_n$$

فإن الوسيط الحسابي لهذه القيم هو :

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\text{مجموع قيم المفردات}}{\text{عدد المفردات}}$$

حيث $\sum_{i=1}^n x_i$ هي مجموع قيم المفردات و n هي عدد المفردات .

يدعى هذا الأسلوب في حساب الوسيط الحسابي بالأسلوب المباشر للبيانات غير المبوبة.

مثال (4-1) :

البيانات الآتية هي عبارة عن علامات أحد الطلاب في خمس مواد :

$$60, 52, 70, 63, 55,$$

فما هو متوسط علامات الطالب في مواد الخمس (أي احسب معدل هذا الطالب في مواد الخمس) نجد أن :

$$x_1 = 60, x_2 = 52, x_3 = 70, x_4 = 63, x_5 = 55$$

وكذلك نجد أن عدد المواد $n = 5$ ، وباستخدام الأسلوب المباشر في حساب متوسط العلامات لهذا الطالب :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{60 + 52 + 70 + 63 + 55}{5} = \frac{300}{5} = 60 \text{ علامة}$$

وهي متوسط علامات هذا الطالب في مواد الخمس، بمعنى آخر إن هذه القيمة 60 علامة لو حصل عليها الطالب في كل مادة من مواد الخمس، لكان مجموع علاماته في مواد الخمس يساوي أيضاً 300 علامة أي :

$$\sum x = \bar{x} \cdot n$$

$$300 = (60) (5)$$

في بعض الحالات تكون القيم غير متساوية من حيث أهميتها، عندها سوف يتطلب ترجيح القيم بما يتناسب وأهمية ووزن كل منها، عندها يكون الوسط الحسابي وسطاً مثنقاً أو مرجحاً (weighted mean) ونرمز له بـ $\bar{x}_w$ وتصبح صيغته :

$$\bar{x}_w = \frac{\sum x_i w_i}{\sum w_i}$$

(حيث w_i هي وزن القيمة x_i)

مثال (4-2) :

بائع أحد تجار الجملة نوعاً من البضاعة بأربعة أسعار مختلفة حسب الكمية حيث باع 5 قطع بسعر 90 ل.س. للقطعة الواحدة و 60 قطعة بسعر 60 ل.س. للقطعة و 300 قطعة بسعر 20 ل.س. للقطعة و 310 قطعة بسعر 15 ل.س. للقطعة الواحدة. احسب متوسط السعر الذي باع فيه (الوسط الحسابي المرجح أو المثنق لسعر البيع).

نحدد أن الكميات التي تم بيعها هي أوزان لترجيح الأسعار:

$$x_1 = 90, x_2 = 60, x_3 = 20, x_4 = 15$$

$$w_1 = 5, w_2 = 60, w_3 = 300, w_4 = 310$$

و بتطبيق صيغة الوسط الحسابي الموزون نجد :

$$- ١٠٢ -$$

$$\bar{x}_w = \frac{\sum x_i w_i}{\sum w_i} = \frac{(90)(5) + (60)(60) + (20)(300) + (15)(310)}{5 + 60 + 300 + 310} = \frac{14700}{675} = 21,78 \text{ L.S}$$

وهو متوسط السعر للقطعة الواحدة.

في بعض الحالات أيضاً تتغير القيم عن بعضها البعض بعدد ثابت تولف سلسلة حسابية، فإذا كان لدينا الأرقام الآتية : 5, 7, 9, 11, 13, 15 فالوسط الحسابي يمكن حسابه بطريقة سريعة في هذه الحالة الخاصة :

$$\frac{\text{القيمة الكبرى} + \text{القيمة الصغرى}}{2}$$

$$\bar{x} = \frac{5 + 15}{2} = 10$$

وهي النتيجة نفسها إذا طبقنا الأسلوب المباشر :

$$\bar{x} = \frac{\sum x}{n} = \frac{5 + 7 + 9 + 11 + 13 + 15}{6} = 10$$

ويمكننا أيضاً استعمال طريقة أخرى لحساب الوسط الحسابي تدعى طريقة الانحرافات وهو أسلوب غير مباشر يوجز العمليات الحسابية في حساب الوسط الحسابي باعتماده على فرض وسط حسابي من بين الأعداد المراد إيجاد وسطها الحسابي، وإيجاد مجموع انحرافات القيم عن الوسط الحسابي المفترض وتطبيق الصيغة الآتية :

$$\bar{x} = \bar{x}_0 + \frac{\sum (x - \bar{x}_0)}{n} = \bar{x}_0 + \frac{\sum d}{n}$$

حيث $\bar{x}_0$ = الوسط الفرضي ، ويفضل أن يكون من ضمن البيانات المراد حساب وسطها الحسابي .

n = عدد القيم (المفردات) و $\sum d = \sum (x - \bar{x}_0)$ مجموع الانحرافات القيم عن الوسط الحسابي الفرضي $\bar{x}_0$.

مثال (4-3) :

احسب الوسط الحسابي بطريقة الانحرافات للبيانات الآتية :

x	$x - \bar{x}_0 = d$
200	$200 - 270 = - 70$
210	$210 - 270 = - 60$
240	$= - 30$
$\bar{x}_0$ ← 270	$= 0$
275	$= + 5$
290	$= + 20$
300	$= + 30$
1785	$- 105$

نفرض الوسط الفرضي $\bar{x}_0 = 270$ وتطبيق أسلوب الانحرافات للبيانات غير

المبوبة نجد :

$$\bar{x} = \bar{x}_0 + \frac{\sum d}{n} = 270 + \frac{- 105}{7}$$

$$= 270 - 15 = 255$$

وبالطريقة المباشرة :

$$\bar{x} = \frac{\sum x}{n} = \frac{1785}{7} = 255$$

وهي النتيجة نفسها وفق الأسلوب المباشر .

وتجيب الإشارة هنا إلى أنه لا تختلف قيمة الوسط الحسابي بأسلوب الانحرافات عن قيمته بالأسلوب المباشر، وذلك مهما كان الوسط الفرضي $\bar{x}_0$ (إن كان من داخل البيانات أو من خارجها) غير أنه يفضل أن يكون $\bar{x}_0$ من ضمن البيانات وأن يكون من منتصف البيانات المرتبة، وذلك بهدف تصغير الانحرافات وبالتالي سهولة الحساب، وسوف نتطرق لهذا الأسلوب بتفصيل أكبر عندما نقوم بدراسة الوسط الحسابي لبيانات مبنية في فئات .

أما إذا تكررت بعض قيم المتغير أو كلها، فإنه في هذه الحالة يمكننا حصر القيم الممكنة للمتغير وعدد مرات تكرار كل منها حيث تكون :

قيم المتغير : $x_1, x_2, \dots, x_n$

التكرار : $f_1, f_2, \dots, f_n$

ويكون الوسط الحسابي (المثقل) :

$$\bar{x} = \frac{\sum_{i=1}^n x f_i}{\sum f_i}$$

حيث : n هي حجم العينة وتساوي $\sum f$.

x = قيم المتغير و f = عدد مرات تكرار المتغير .

مثال (4-4) :

فيما يلي عدد أفراد عشر أسر تم اختيارهم بشكل عشوائي من إحدى أحياء مدينة دمشق. أوجد الوسط الحسابي لعدد أفراد الأسرة الواحدة .

جدول رقم (4-1)

رقم الأسرة	1	2	3	4	5	6	7	8	9	10
عدد الأفراد	5	4	3	7	4	4	5	6	3	5

$$\bar{x} = \frac{\sum x}{n} = \frac{5+3+6+5+4+4+7+3+4+5}{10}$$

$$= \frac{46}{10} = 4.6 \quad \text{فرداً}$$

ويمكن أن نحصر تكرار قيم المتغير بحيث ننشئ الجدول المساعد الآتي :

جدول رقم (4-2)

عدد أفراد الأسرة (x)	عدد الأسر (f)	x . f
3	2	6
4	3	12
5	3	15
6	1	6
7	1	7
Σ	10	46

وهكذا، فإن الوسط الحسابي (المثقل) لعدد أفراد الأسرة الواحدة يساوي :

$$\bar{x} = \frac{\sum x.f}{\sum f} = \frac{46}{10} = 4.6 \quad \text{فرداً}$$

وهي النتيجة نفسها حسب الأسلوب المباشر .

2-2-4- حساب الوسط الحسابي لبيانات إحصائية مبوبة في فئات :

عند التعامل مع بيانات إحصائية مبوبة في فئات (وذلك بعد توزيع القيم الأصلية للظاهرة على الفئات في الجدول التكراري) تختفي هذه القيم وتضيق معالمها الأصلية بحيث إن لكل فئة مدى، له حد أدنى وحد أعلى، وسوف نعتبر أن كل المفردات التي في فئة تكرارية واحدة متساوية، وقيمتها تساوي مركز الفئة التي تناظرها، وبالتالي فإن وسط الفئة سيعتبر ممثلاً لقيم هذه الفئة. إن هذا المنطق صحيح من الناحية العملية

والنظرية وإن استعمال وسطي الفئة الذي يعبر عن قيمة كل تكرارات هذه الفئة يفترض أن التكرارات المقابلة لكل فئة تتوزع بشكل طبيعي بين حدي الفئة الواحدة بحيث تتمحور حول قيمة واحدة هي متوسط الفئة.

وبصورة أخرى لو تم حساب الوسط الحسابي الحقيقي للقيم الحقيقية لتكرارات فئة ما لوجدنا بشكل قريب جداً أن هذا الوسط الحقيقي يساوي وسطياً الفئة، وقد يزيد هذا الوسط الحقيقي أو ينقص عن متوسط الفئة هذا القدر أو ذاك، بحيث نجد أن في فئة يكون الوسط الحسابي الحقيقي أقل من متوسط الفئة، وفي فئة أخرى أكبر من متوسط فئتها بحيث تلغي بعضها البعض، وبالنسبة سوف نحصل على أوساط فئات تمثل البيانات الحقيقية جميعها وبشكل قريب جداً. فإذا رمزنا لمراكز الفئات بـ x_i وللتكرارات المقابلة لكل فئة بـ f_i فإن الأسلوب المباشر لحساب الوسط الحسابي المرجح لبيانات إحصائية مبوبة في فئات يساوي :

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

وخطوات إيجاد الوسط الحسابي لتلك البيانات المبوبة تتلخص بما يأتي :

- (1) نحسب مراكز الفئات f_i .
- (2) نضرب مركز كل فئة بتكرارها المقابل $(x_i f_i)$ ومن ثم نوجد $(\sum x_i f_i)$.
- (3) نقسم مجموع حاصل ضرب مركز كل فئة بتكرارها $(\sum x_i f_i)$ على مجموع التكرارات $\sum f_i$ فنحصل عندها على الوسط الحسابي المرجح للبيانات المبوبة.

مثال (4-5) :

يبين الجدول الآتي توزيع أعمار المصاييح الكهربائية من إنتاج شركة ما، والمطلوب حساب الوسط الحسابي لأعمار هذه المصاييح.

جدول رقم (4-3)

فئات العمر (بالساعات)	عدد المصاييح
100 - 200	14
200 - 300	46
300 - 400	58
400 - 500	76
500 - 600	68
600 - 700	62
700 - 800	48
800 - 900	22
900 - 1000	6
	400

ننشئ الجدول المساعد لإيجاد الوسط الحسابي :

جدول رقم (4-4)

فئات العمر (بالساعات)	التكرارات عدد المصاييح (f)	وسطي الفئة x	x . f
100 - 200	14	150	2100
200 - 300	46	250	11500
300 - 400	58	350	20300
400 - 500	76	450	34200
500 - 600	68	550	37400
600 - 700	62	650	40300
700 - 800	48	750	36000
800 - 900	22	850	18700
900 - 1000	6	950	5700
Σ	400		206200

و بتطبيق الأسلوب المباشر :

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{206200}{400} = 515,5 \text{ ساعة}$$

أي إن عمر المصباح الواحد من هذه المصابيح هو بالمتوسط 515,5 ساعة .

وفي بعض الحالات قد لا يتوفر لدينا معطيات إحصائية بشكل تكرارات، بل بشكل نسب هذه التكرارات (تكرارات نسبية)، وهي قد وجدت أصلاً من تقسيم تكرار كل فئة على التكرار الإجمالي مضروباً بقيمة معينة، بمعنى أن التكرار النسبي وجد من خلال $\frac{f}{\sum f} \cdot K$ حيث K هي 1 أو 10 أو 100 أو 1000 أو 10000 ... إلخ، فيتحول عندها التكرار المطلق للفئة إلى تكرار نسبي.

وبالتالي تصبح صيغة الوسط الحسابي لبيانات جدول تكراري نسبي :

$$\bar{x} = \frac{\sum \left(x_i \frac{f_i}{\sum f_i} \cdot K \right)}{\sum \left(\frac{f_i}{\sum f_i} \cdot K \right)} = \frac{\sum x \cdot f_r}{\sum f_r}$$

حيث f_r التكرار النسبي و $\sum f_r$ يساوي إما 1 أو 100 أو 1000 .

بالعودة للمثال رقم (4-5) وبفرض أنه تم تحويل التكرارات المطلقة إلى تكرارات نسبية مئوية أو تكرارات نسبية ألفية أي $K = 100$ ومرة أخرى $K = 1000$ نجد ما يأتي :

جدول رقم (4-5)

فئات العمر (بالساعات)	$\frac{f}{\sum f} 100 = f_{r1}$	$\frac{f}{\sum f} 1000 = f_{r2}$	$x \cdot f_{r1}$	$x \cdot f_{r2}$
100 - 200	3,5	35	525	5250
200 - 300	11,5	115	2875	28750
300 - 400	14,5	145	5075	50750
400 - 500	19	190	8550	85500
500 - 600	17	170	9350	93500
600 - 700	15,5	155	10075	100750
700 - 800	12	120	9000	90000
800 - 900	5,5	55	4675	46750
900 - 1000	1,5	15	1425	14250
Σ	100	1000	51550	515500

وبتطبيق الصيغة السابقة في حساب الوسط الحسابي المرجح لبيانات تحتوي على نسب:

$$\bar{x} = \frac{\sum x \cdot f_{r1}}{\sum f_1} = \frac{51550}{100} = 515,5 \text{ ساعة}$$

$$\bar{x} = \frac{\sum x \cdot f_{r2}}{\sum f_2} = \frac{515500}{1000} = 515,5 \text{ ساعة}$$

وهي النتيجة نفسها التي حصلنا عليها سابقاً باستخدام التكرارات الفعلية المطلقة.

4-2-3- خصائص الوسط الحسابي:

4-2-3-1- بعض الأساليب غير المباشرة في إيجاد قيمة الوسط الحسابي:

يُعدّ الوسط الحسابي من أهم مقاييس النزعة المركزية لما يتميز به من خصائص جبرية هامة تجعل إمكانية حسابه بأساليب متعددة وغير مباشرة تؤدي إلى النتيجة

نفسها التي يمكن التوصل إليها بالأسلوب المباشر المذكور سابقاً ومن أهم هذه الخصائص التي تؤدي إلى أساليب غير مباشرة في حسابه نذكر :

1- إن قيمة الوسط الحسابي لن تتغير إذا تم ضرب كل التكرارات بقيمة ما أو إذا قسمنا جميع التكرارات على قيمة ما (أسلوب تعديل التكرارات) .
من الصيغة المباشرة للوسط الحسابي :

$$\bar{x} = \frac{\sum x f}{\sum f}$$

وبضرب جميع التكرارات بالقيمة K حيث $0 < K$ نجد :

$$\bar{x}' = \frac{\sum x \cdot f \cdot K}{\sum f \cdot K}$$

وبما أن K هو رقم ثابت أكبر من الصفر وبالتالي تصبح الصيغة :

$$\bar{x}' = \frac{K \sum x f}{K \sum f} = \bar{x}$$

أي :

$$\bar{x}' = \bar{x}$$

وأيضاً بتقسيم جميع التكرارات على القيمة $0 < K$ نجد :

$$\bar{x}' = \frac{\sum x \cdot \frac{f}{K}}{\sum \frac{f}{K}}$$

$$\bar{x}' = \frac{\frac{1}{K} \sum x \cdot f}{\frac{1}{K} \sum f} = \bar{x}$$

وبالتالي :

$$\bar{x}' = \bar{x}$$

مثال (4-6) :

قمنا بدراسة لعلامات 300 طالباً تم امتحانهم بمادة مبادئ الإحصاء فحصلنا على المعلومات الواردة في الجدول (4-6) :

جدول رقم (4-6)

فئات العلامات	عدد الطلاب
30 - 40	20
40 - 50	40
50 - 60	60
60 - 70	100
70 - 80	40
80 - 90	30
90 - 100	10

احسب الوسط الحسابي بطريقة تعديل التكرارات .

الحل :

لتبسيط أرقام التكرارات يمكن أن نقسم كل تكرار على $K = 10$ أو أي رقم آخر مناسب فننشئ الجدول المساعد الآتي :

جدول رقم (4-7)

وسطي الفئة x	f	$f/10$	$x f/10$	$x \cdot f$
35	20	2	70	700
45	40	4	180	1800
55	60	6	330	3300
65	100	10	650	6500
75	40	4	300	3000
85	30	3	255	2550
95	10	1	95	950
Σ	300	30	1880	18800

وبتطبيق صيغة تعديل التكرارات نجد أن :

$$\bar{x} = \frac{\sum x \frac{f}{K}}{\sum \frac{f}{K}}$$

حيث $K = 10$ نجد :

$$\bar{x} = \frac{1880}{30} = 62,67 \text{ علامة}$$

أما استعمال الأسلوب المباشر فنجد :

$$\bar{x} = \frac{\sum x f}{\sum f} = \frac{18800}{300} = 62,67 \text{ علامة}$$

وهي النتيجة السابقة نفسها، حيث يلاحظ كيف تمت الاستفادة من خاصية تعديل التكرارات في تسهيل عملية حساب الوسط الحسابي.

كما يمكننا أيضاً أن نضرب جميع التكرارات ونقسمها في آن واحد على عددين ثابتين كأن نضرب التكرارات بـ 100 ونقسمها على 500 بمعنى ضرب كل تكرار بـ 0,2 وبعد ذلك نطبق الصيغة السابقة (صيغة تعديل التكرارات) حيث K هنا تساوي :

$$\frac{100}{500} = (0,2)$$

وبالتالي سوف تؤدي إلى النتائج السابقة نفسها.

2- إذا ضربت كل قيمة من قيم التوزيع التكراري بقيمة ثابتة (K) فإن الوسط الحسابي للقيم الجديدة يساوي الوسط الحسابي للقيم الحقيقية مضروباً بالقيمة الثابتة (K)، أي الوسط الحسابي للقيم الحقيقية يساوي الوسط الحسابي للقيم الجديدة مقسوماً على القيمة الثابتة (K).

وإذا قسمت كل قيمة من قيم التوزيع التكراري على القيمة الثابتة (K) فإن

الوسط الحسابي للقيم الجديدة يساوي الوسط الحسابي للقيم الحقيقية مقسوماً على القيمة الثابتة (K) أي الوسط الحسابي للقيم الحقيقية يساوي الوسط الحسابي للقيم الجديدة مضروباً بالقيمة الثابتة (K) (أسلوب تعديل أوساط الفئات) .

بتطبيق الصيغة المباشرة للوسط الحسابي المرجح :

$$\bar{x} = \frac{\sum x \cdot f}{\sum f}$$

وبتقسيم جميع القيم (أوساط الفئات) على K فتصبح المعادلة السابقة، حيث (K

عدد ثابت) :

$$\bar{x}' = \frac{\sum \frac{x}{K} \cdot f}{\sum f} \quad (\text{الوسط الحسابي الجديد})$$

$$\bar{x}' = \frac{\frac{1}{K} \sum x \cdot f}{\sum f}$$

$$\bar{x}' = \frac{1}{K} \cdot \frac{\sum x \cdot f}{\sum f}$$

$$\bar{x}' = \frac{1}{K} \cdot \bar{x} \Rightarrow$$

$$\bar{x}' = \frac{\bar{x}}{K} \Rightarrow \boxed{\bar{x} = \bar{x}' \cdot K}$$

ولتوضيح ذلك نورد المثال الآتي :

مثال (4-7) :

إذا كان لدينا القيم الآتية :

$$x = 15, 3, 6, 6, 3, 6, 3$$

$$\text{الوسط الحسابي الحقيقي } \bar{x} = \frac{42}{7} = 6$$

وبتقسيم كل قيمة على 3 فنتحول قيم x إلى قيم جديدة x' :

$$x' = 5, 1, 2, 2, 1, 2, 1$$

$$\text{الوسط الحسابي الجديد } \bar{x}' = \frac{14}{7} = 2$$

وإذا ضربنا الوسط الحسابي الجديد بالقيمة الثابتة $K = 3$ ينتج لدينا الوسط الحسابي الحقيقي:

$$(2)(3) = 6$$

$$(\bar{x}')(K) = \bar{x}$$

كذلك من الصيغة المباشرة:

$$\bar{x} = \frac{\sum x \cdot f}{\sum f}$$

وبضرب جميع القيم (أوساط الفئات) بالقيمة K فتصبح المعادلة السابقة:

$$(\text{الوسط الحسابي الجديد}) \bar{x}' = \frac{\sum (x \cdot K) f}{\sum f}$$

$$\bar{x}' = \frac{K \sum x \cdot f}{\sum f}$$

حيث K عدد ثابت يمكن إخراجها خارج المجموع

$$\bar{x}' = K \frac{\sum x \cdot f}{\sum f}$$

$$\bar{x}' = K \cdot \bar{x} \Rightarrow \boxed{\bar{x}' = \frac{\bar{x}}{K}}$$

وتعني المعادلة الأخيرة بأنه لو تم ضرب كل القيم بعدد ثابت فإن الوسط الحسابي الحقيقي يساوي الوسط الحسابي الجديد مقسوماً على العدد الثابت .

مثال (4-8) :

ليكن لدينا القيم الآتية :

$$x = 3, 10, 12, 2, 6$$

$$\bar{x} = \frac{33}{5} = 6,6$$

الوسط الحقيقي

بضرب جميع القيم بالعدد $K = 4$ فتصبح القيم الجديدة :

$$x' = 12, 40, 48, 8, 24$$

$$\bar{x}' = \frac{132}{5} = 26,4$$

والوسط الجديد

وبالتالي يمكن إيجاد الوسط الحسابي الحقيقي من الوسط الحسابي الجديد بحسب

المعادلة السابقة :

$$\bar{x} = \frac{\bar{x}'}{K}$$

$$\bar{x} = \frac{26,4}{4} = 6,6$$

ولبيان كيفية إجراء الحسابات لما تم ذكره سابقاً نعود إلى المثال رقم (4-6) عن توزيع علامات 300 طالب ولنقسم أوساط الفئات في هذا التوزيع على القيمة 5 مثلاً .

جدول رقم (4-8)

فئات الدرجات	عدد الطلاب (f)	وسطي الفئة x	'x = x/5	'x . f
30 - 40	20	35	7	140
40 - 50	40	45	9	360
50 - 60	60	55	11	660
60 - 70	100	65	13	1300
70 - 80	40	75	15	600
80 - 90	30	85	17	510
90 - 100	10	95	19	190
Σ	300		91	3760

فالوسط الحسابي للقيم الجديدة (تعديل أوساط الفئات) :

$$\bar{x}' = \frac{\sum x' f}{\sum f} = \frac{3760}{300} = 12,533$$

أما الوسط الحسابي الحقيقي فيساوي :

$$\bar{x} = \bar{x}' . 5 = 62,67 \quad \text{علامة}$$

3- عند إضافة عدد ثابت إلى كل قيمة من قيم التوزيع، فالوسط الحسابي الجديد سوف يزيد بمقدار العدد الثابت الذي تمت إضافته، وبالمقابل عند طرح العدد الثابت من كل قيمة من قيم التوزيع فإن الوسط الحسابي الجديد سوف يقل بمقدار العدد الثابت الذي تم طرحه (أسلوب الانحرافات).
من الصيغة المباشرة :

$$\bar{x} = \frac{\sum x f}{\sum f}$$

وعند إضافة العدد الثابت $\bar{x}_0$ إلى جميع قيم التوزيع، حيث يدعى $\bar{x}_0$ بالوسط الحسابي الفرضي الذي يمكن أخذه من داخل أو خارج البيانات، ولكن يفضل أن يكون من داخل ومتنصف البيانات والمقابلة الأكبر تكرار، عندها تصبح الصيغة السابقة :

$$\bar{x}' = \frac{\sum (x + \bar{x}_0) f}{\sum f} \quad (\text{الوسط الحسابي الجديد})$$

$$\bar{x}' = \frac{\sum x f + \sum \bar{x}_0 f}{\sum f}$$

وبما أن $\bar{x}_0$ عدد ثابت :

$$\bar{x}' = \frac{\sum x f}{\sum f} + \frac{\bar{x}_0 \sum f}{\sum f}$$

$$\bar{x}' = \bar{x} + \bar{x}_0 \Rightarrow$$

$$\boxed{\bar{x} = \bar{x}_0 - \bar{x}'} \Rightarrow \bar{x} = \bar{x}_0 - \frac{\sum (x + \bar{x}_0) f}{\sum f}$$

الوسط الحسابي الحقيقي

وعند طرح العدد الثابت $\bar{x}_0$ من جميع قيم التوزيع، نجد عندها أن الصيغة المباشرة للوسط الحسابي الحقيقي تتحول إلى :

$$\bar{x}' = \frac{\sum (x - \bar{x}_0) f}{\sum f} = \frac{\sum x f - \bar{x}_0 \sum f}{\sum f}$$

$$\bar{x}' = \frac{\sum x f}{\sum f} - \frac{\bar{x}_0 \sum f}{\sum f} = \bar{x} - \bar{x}_0$$

$$\bar{x}' = \bar{x} - \bar{x}_0 \Rightarrow \bar{x} = \bar{x}_0 + \bar{x}'$$

$$\bar{x} = \bar{x}_0 + \frac{\sum (x - \bar{x}_0) f}{\sum f}$$

نرمز لـ $(x - \bar{x}_0)$ بـ d وبالتالي فالوسط الحسابي بأسلوب الانحرافات لبيانات

مبوبة في فئات يساوي :

$$\bar{x} = \bar{x}_0 + \frac{\sum d f}{\sum f}$$

لنطبق أسلوب الانحرافات المذكورة سابقاً على المثال رقم (4-6) جدول رقم (4-6) حيث نحدد في بادئ الأمر قيمة الوسط الحسابي الفرضي $\bar{x}_0$ ، ونأخذه من ضمن البيانات وليكن $\bar{x}_0 = 65$ ثم نحسب الانحرافات قيم التوزيع (متوسط الفئات) عن الوسط الحسابي الفرضي، وبعدها نقوم بضرب التكرارات المقابلة للانحرافات الناتجة، ومن ثم يتم تقسيم مجموع انحرافات القيم عن الوسط الحسابي الفرضي والمضروب بالتكرارات المقابلة على إجمالي التكرارات، فنحصل بذلك على الوسط الحسابي لانحرافات قيم التوزيع عن الوسط الحسابي الفرضي، وهو يساوي $\bar{x}$ (الوسط الحسابي الجديد) ، وبالتالي وللحصول على الوسط الحقيقي نقوم بإضافة الوسط الحسابي الفرضي $\bar{x}_0$.

حساب الوسط الحسابي بطريقة الانحرافات عن الوسط الحسابي الفرضي $\bar{x}_0 = 65$

جدول رقم (4-9)

x	f	$x - \bar{x}_0 = d$	$(x - \bar{x}_0) f = df$
35	20	- 30	- 600
45	40	- 20	- 800
55	60	- 10	- 600
$\bar{x}_0 \Leftarrow 65$	100	0	0
75	40	10	400
85	30	20	600
95	10	30	300
	300	0	- 2000
			+ 1300
			- 700

و بتطبيق أسلوب الانحرافات وبمساعدة الجدول رقم (4-9) نجد :

$$\begin{aligned}\bar{x} &= \bar{x}_0 + \frac{\sum d f}{\sum f} \\ &= 65 + \frac{-700}{300} \\ &= 65 - 2,33 = 62,67 \quad \text{علامة}\end{aligned}$$

4- إذا طرحنا من كل قيمة من قيم التوزيع المقدار $\bar{x}_0$ (الوسط الحسابي الفرضي) وقسمنا الناتج على قيمة محددة وتساوي طول الفئة C عندها يكون الوسط الحسابي الحقيقي يساوي طول الفئة مضروباً بـ وسطي الانحرافات القيم عن الوسط الحسابي الفرضي والمعدل بطول الفئة مضافاً إليه الوسط الحالي الفرضي $\bar{x}_0$ (أسلوب الانحرافات المختصرة)، لننتقل أيضاً من الصيغة المباشرة للوسط الحسابي الحقيقي .

$$\bar{x} = \frac{\sum x \cdot f}{\sum f}$$

ولنطرح من كل قيمة من قيم التوزيع الوسط الحسابي الفرضي $\bar{x}_0$ ولنقسم الفارق على طول الفئة C عندها يصبح الوسط الحسابي الجديد :

$$\begin{aligned}\bar{x}' &= \frac{\sum \left(\frac{x - \bar{x}_0}{C} \right) \cdot f}{\sum f} \\ \bar{x}' &= \frac{\frac{1}{C} \sum (x - \bar{x}_0) \cdot f}{\sum f} \Rightarrow \\ \frac{\bar{x}'}{\frac{1}{C}} &= \frac{\sum (x - \bar{x}_0) f}{\sum f}\end{aligned}$$

$$C. ' \bar{x} = \frac{\sum x f - \bar{x}_0 \sum f}{\sum f}$$

$$C. ' \bar{x} = \bar{x} - \bar{x}_0 \Rightarrow$$

$$\bar{x} = \bar{x}_0 + ' \bar{x} . C = \bar{x}_0 + \frac{\sum \left(\frac{x - \bar{x}_0}{C} \right) f}{\sum f} . C$$

لنرمز لـ $\frac{x - \bar{x}_0}{C}$ بـ D وبالتالي فالوسط الحسابي بأسلوب الانحرافات المختصرة:

$$\bar{x} = \bar{x}_0 + \frac{\sum D f}{\sum f} . C$$

وبتطبيق هذا الأسلوب على المثال رقم (4-6) جدول رقم (4-6) وبفرض أن الوسط الحسابي الفرضي $\bar{x}_0 = 65$ وطول الفئة يساوي $C = 10$.

حساب الوسط الحسابي بأسلوب الانحرافات المختصرة حيث $\bar{x}_0 = 65$ و $C = 10$

جدول رقم (4-10)

x	f	$x - \bar{x}_0$	$\frac{x - \bar{x}_0}{C} = D$	$\left(\frac{x - \bar{x}_0}{C} \right) . f = Df$
35	20	- 30	- 3	- 60
45	40	- 20	- 2	- 80
55	60	- 10	- 1	- 60
$\bar{x}_0 \Leftarrow 65$	100	0	0	0
75	40	10	1	40
85	30	20	2	60
95	10	30	3	30
	300	0	0	- 200
				+ 130
				- 700

بمساعدة الجدول (4-10) وتطبيق أسلوب الانحرافات المختصرة نجد :

$$\bar{x} = \bar{x}_0 + \frac{\sum D f}{\sum f} \cdot C$$

$$\bar{x} = 65 + \frac{-70}{300} \cdot 10 = 65 - 0,23 (10) = 62,67 \quad \text{علامة}$$

إن أسلوب الانحرافات المختصرة يشترط أن تكون أطوال الفئات متساوية في الجدول التكراري، فهو لا يصلح لجدول تكراري ذي فئات مختلفة الأطوال.

5 - وأيضاً عندما تكون أطوال الفئات متساوية في الجدول التكراري بحيث أن مراكز الفئات لهذا الجدول تأخذ فيه شكل سلسلة حسابية بمعنى أن الفارق بين كل مركز فئة مع الفئة السابقة لها أو اللاحقة لها تأخذ قيمة ثابتة، فيمكن عندها إيجاد الوسط الحسابي بالأسلوب الآتي (أسلوب السلسلة الحسابية) نطلق أيضاً من الصيغة المباشرة للوسط الحسابي :

$$\bar{x} = \frac{\sum x_i f_i}{\sum f}$$

حيث : $x_i = x_1, x_2, \dots, x_n$ و $i = 1, 2, \dots, n$

وحيث : $f_i = f_1, f_2, \dots, f_n$ وهي التكرارات المقابلة لكل قيمة من قيم x .

وبما أن أطوال الفئات متساوية فإن كل قيمة تزيد عن القيمة التي تسبقها بمقدار طول الفئة C :

$$x_1 = x_0 + C$$

بمعنى أن متوسط الفئة الأولى x_1 تساوي متوسط الفئة التي تسبقها مباشرة x_0 مضافاً إليها طول الفئة C ، وبالتالي يمكن تحويل جميع أوساط الفئات بدلالة متوسط الفئة، والتي هي قبل الفئة الأولى x_0 :

$$\begin{aligned}
\bar{x} &= \frac{\sum x_i f_i}{\sum f_i} = \frac{(x_0 + C) f_1 + (x_0 + 2C) f_2 + \dots + (x_0 + nC) f_n}{f_1 + f_2 + \dots + f_n} \\
&= \frac{(x_0 f_1 + C f_1) + (x_0 f_2 + 2C f_2) + \dots + (x_0 f_n + nC f_n)}{f_1 + f_2 + \dots + f_n} \\
&= \frac{(x_0 f_1 + x_0 f_2 + \dots + x_0 f_n) + (C f_1 + 2C f_2 + \dots + nC f_n)}{f_1 + f_2 + \dots + f_n} \\
&= \frac{x_0 (f_1 + f_2 + \dots + f_n) + C (f_1 + 2 f_2 + \dots + n f_n)}{f_1 + f_2 + \dots + f_n} \\
&= \frac{x_0 (f_1 + f_2 + \dots + f_n)}{f_1 + f_2 + \dots + f_n} + \frac{C (f_1 + 2 f_2 + \dots + n f_n)}{f_1 + f_2 + \dots + f_n}
\end{aligned}$$

$$\boxed{\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = x_0 + C \frac{\sum i f_i}{\sum f_i}}$$

حيث x_0 = هي مركز الفئة التي تسبق الفئة الأولى وتساوي $x_0 = x_1 - C$.

C = طول الفئة (أطوال الفئات جميعها متساوية) .

i = تأخذ القيم من الواحد وحتى n (عدد الفئات) .

وبالعودة للمثال رقم (4-6) والجدول رقم (4-6) لحساب الوسط الحسابي المرجح

بأسلوب السلسلة الحسابية نقوم بإنشاء الجدول المساعد رقم (4-11) :

جدول رقم (4-11)

وسط العينة x_i	f_i	i	$i f_i$
35	20	1	20
45	40	2	80
55	60	3	180
65	100	4	400
75	40	5	200
85	30	6	180
95	10	7	70
Σ	300		1130

$$\bar{x} = x_0 + C \frac{\Sigma i f_i}{\Sigma f_i}$$

$$x_0 = x_1 - C$$

$$x_0 = 35 - 10 = 25$$

وبالتالي :

$$\bar{x} = 25 + 10 \frac{1130}{300}$$

$$= 62,67 \text{ علامة}$$

وتجيب الإشارة إلى أن جميع الأساليب المستخدمة في حساب الوسط الحسابي، والتي تعكس خصائصه تؤدي جميعها إلى النتيجة نفسها التي حصلنا عليها وفق الأسلوب المباشر التي اشتقت منه هذه الأساليب، بهدف اختزال العمليات الحسابية وتبسيطها إلى أكبر قدر ممكن، إلا أن الطريقتين السابقتين (طريقة الانحرافات المختصرة وطريقة السلسلة الحسابية) لا يمكن استخدامها في التوزيعات التكرارية التي تكون أطوال فئاتها غير متساوية، إلا بعد إجراء التعديلات التي تتطلب العودة مرة أخرى إلى المعطيات،

ووضع فرضيات إضافية، وإعادة توزيعها من جديد، الأمر الذي لا يؤدي إلى تحقيق الاختصار المستهدف من تطبيق هاتين الطريقتين .

وأخيراً إذا كان لدينا توزيع تكراري ما يحتوي على نوعين من التكرارات، فلايجاد الوسط الحسابي للنوع الأول والوسط الحسابي للنوع الثاني، يمكن إيجاد الوسط الحسابي للوسطين الحسابيين بالعلاقة الآتية في حال تساوي مجموع التكرارين :

$$\bar{x} = \frac{\bar{x}_1 + \bar{x}_2}{2}$$

حيث :

$$\sum f_2 = \sum f_1$$

أما في حالة عدم التساوي لـ $\sum f_2 \neq \sum f_1$ فالوسط الحسابي للوسطين الحسابيين يساوي :

$$\bar{x} = \frac{\bar{x}_1 \sum f_1 + \bar{x}_2 \sum f_2}{\sum f_1 + \sum f_2}$$

ولتوضيح ذلك نورد المثال الآتي :

مثال (4-9) :

ليكن لدينا الجدول الآتي الذي يعبر عن توزيع عمال عهتين عشوائيتين من عمال القطاع الخدمي في كل من دمشق وحلب وحسب فئات العمر :

جدول رقم (4-12)

فئات العمر (سنة)	عدد العمال في دمشق	عدد العمال في حلب
20 - 26	20	10
26 - 32	60	20
32 - 38	30	25
38 - 44	25	30
44 - 50	10	10
50 - 56	5	5

- 1- احسب الوسط الحسابي لأعمار عمال عينة القطاع الخدمي في دمشق ولأعمار عينة عمال القطاع الخدمي في حلب .
- 2- احسب الوسط الحسابي لأعمار عمال القطاع الخدمي في كل من دمشق وحلب معاً (وذلك بدمج التكرارات) .
- 3- احسب الوسط الحسابي للوسطين الحسابيين ، ماذا تستنتج ؟

الجدول المساعد رقم (13 - 4)

(1)

وسطي الفئة x	f ₁	f ₂	D	Df ₁	Df ₂
23	20	10	- 3	- 60	- 30
29	60	20	- 2	- 120	- 40
35	30	25	- 1	- 30	- 25
41	25	30	0	0	0
47	10	10	1	10	10
53	5	5	2	10	10
$\bar{x}_0 = 41$ C = 6	150	100		- 190	- 75

$$\begin{aligned} \bar{x}_1 &= \bar{x}_0 + \frac{\sum D f_1}{\sum f_1} \cdot C \\ &= 41 + \frac{- 190}{150} \cdot 6 = 33,4 \text{ سنة} \end{aligned}$$

$$\begin{aligned} \bar{x}_2 &= \bar{x}_0 + \frac{\sum D f_2}{\sum f_2} \cdot C \\ &= 41 + \frac{- 75}{100} \cdot 6 = 36,4 \text{ سنة} \end{aligned}$$

(2) إذا تم دمج التكرارات $f = f_2 + f_1$ عندها يمكن حساب الوسط الحسابي لأعمار عمال القطاع الخدمي في عيني دمشق وحلب معاً كما يأتي :

جدول رقم (14 - 4)

x	f	D	Df
23	30	- 3	- 90
29	80	- 2	- 16
35	55	- 1	- 55
41	55	0	0
47	20	1	20
53	10	2	20
$\bar{x}_0 = 41$ C = 6	250		- 265

$$\begin{aligned}\bar{x} &= \bar{x}_0 + \frac{\sum D f}{\sum f} \cdot C \\ &= 41 + \frac{- 265}{250} \cdot 6 \\ &= 34,64 \text{ سنة}\end{aligned}$$

متوسط أعمار عيني القطاع الخدمي في كل من دمشق وحلب معاً .
(3) ويمكن إيجاد الوسط الحسابي للوسطين الحسابيين للمثال السابق بدون دمج التكرارات وباستخدام :

$$\begin{aligned}\bar{x} &= \frac{\bar{x}_1 \sum f_1 + \bar{x}_2 \sum f_2}{\sum f_1 + \sum f_2} \\ \bar{x} &= \frac{(33,4) (150) + (36,5) (100)}{150 + 100} = 34,64 \text{ سنة}\end{aligned}$$

وهو الوسط الحسابي للوسطين الحسابيين (الوسط الحسابي لمتوسط أعمار عيني دمشق وحلب) نستنتج أنه سواء أتم التكرارات أم لا فإن كلتا الطريقتين تؤدي إلى

نتيجة واحدة في حساب الوسط الحسابي للمجموعتين، أو بإيجاد متوسط الأوساط الحسابية المرجحة لإجمالي التكرارات للعينتين .

2-3-2-4 بعض الخصائص الجبرية الأخرى للوسط الحسابي :

عندما تطرقنا إلى الطرق غير المباشرة لحساب الوسط الحسابي انطلقنا من بعض خصائصه لإيجاد هذه الطرق، فالوسط الحسابي - كما ذكرنا سابقاً - يتمتع بسميزات تجعله أفضل مقياس للنزعة المركزية لأساسه النظري الذي يسمح لاستخدامه في التحليلات الإحصائية المتقدمة. ومن خصائصه الجبرية أيضاً :

1- مجموع انحرافات القيم عن الوسط الحسابي يساوي الصفر .

للبينات غير المبوبة :

$$\sum (x - \bar{x}) = 0$$

$$\sum (x - \bar{x}) = \sum x - \sum \bar{x}$$

وبما أن $\bar{x}$ عدد ثابت فـ :

$$\sum (x - \bar{x}) = \sum x - n \bar{x}$$

$$= \sum x - n \frac{\sum x}{n}$$

$$= \sum x - \sum x$$

$$= 0$$

وللبينات المبوبة في فئات :

$$\sum (x - \bar{x}) f = 0$$

$$\sum (x - \bar{x}) f = \sum x f - \sum \bar{x} f$$

وبما أن $\bar{x}$ عدد ثابت، فيمكن إخراج $\bar{x}$ خارج المجموع :

$$\sum (x - \bar{x}) f = \sum x f - \sum \bar{x} f$$

$$= \sum x f - \frac{\sum x f}{\sum f} \cdot \sum f$$

$$= \sum x f - \sum x f$$

$$= 0$$

مثال (4-10) :

ليكن لدينا القيم الآتية :

$$x_i = 2, 6, 4, 9, 10, 5$$

برهن أن مجموع الانحرافات القيم عن وسطها الحسابي يساوي الصفر :

$$\bar{x} = \frac{36}{6} = 6$$

$$\sum (x - \bar{x}) = (2-6) + (6-6) + (4-6) + (9-6) + (10-6) + (5-6)$$

$$= (-4) + (0) + (-2) + (3) + (4) + (-1)$$

$$= 0$$

مثال (4-11) :

ليكن لدينا الجدول الآتي : برهن أن مجموع الانحرافات القيم عن وسطها الحسابي يساوي الصفر .

جدول رقم (4-15)

الفئات	f	x	x . f	x - $\bar{x}$	(x - $\bar{x}$) f
10 - 14	5	12	60	- 8,6	- 43
14 - 18	18	16	288	- 4,6	- 82,8
18 - 22	42	20	840	- 0,6	- 25,2
22 - 26	27	24	648	3,4	91,8
26 - 30	8	28	224	7,4	59,2
Σ	100		2060		0

$$\bar{x} = \frac{\sum x f}{\sum f} = \frac{2060}{100} = 20,6$$

وبالتالي من الجدول (4-15) نجد :

$$\sum (x - \bar{x}) f = 0$$

(2) مجموع مربع الانحرافات القيم عن الوسط الحسابي أقل من مجموع مربع

انحرافات القيم عن أية قيمة أخرى غير الوسط الحسابي :

$$\sum (x - \bar{x})^2 \longrightarrow \text{أقل ما يمكن}$$

نفرض أن K هي قيمة غير قيمة الوسط الحسابي فإن :

$$\sum (x - \bar{x})^2 < \sum (x - K)^2$$

مهما تكن قيمة K وبشرط أن $\bar{x} \neq K$

$$\sum (x - K)^2 = \sum (x^2 - 2xK + K^2)$$

$$= \sum x^2 - \sum 2xK + \sum K^2$$

وبما أن K عدد ثابت فإن :

$$= \sum x^2 - 2K \sum x + nK^2$$

$$= \sum x^2 - 2Kn\bar{x} + nK^2$$

وبطرح وإضافة $n\bar{x}^2$ نجد :

$$= \sum x^2 - 2Kn\bar{x} + nK^2 + n\bar{x}^2 - n\bar{x}^2$$

$$\sum (x - K)^2 = (\sum x^2 - n\bar{x}^2) + n(K^2 - 2K\bar{x} + \bar{x}^2)$$

ونجد أن :

$$n(K^2 - 2K\bar{x} + \bar{x}^2) = n(K - \bar{x})^2$$

$$(\sum x^2 - n\bar{x}^2) = \sum (x - \bar{x})^2$$

لأن :

$$\begin{aligned}\sum (x - \bar{x})^2 &= \sum (x^2 - 2x\bar{x} + \bar{x}^2) \\ &= \sum x^2 - 2\bar{x} \sum x + n\bar{x}^2 \\ &= \sum x^2 - 2\bar{x}n\bar{x} + n\bar{x}^2 \\ &= \sum x^2 - 2n\bar{x}^2 + n\bar{x}^2\end{aligned}$$

$$\sum (x - \bar{x})^2 = (\sum x^2 - n\bar{x}^2)$$

وبالتالي :

$$\sum (x - K)^2 = \sum (x - \bar{x})^2 + n(K - \bar{x})^2$$

مما تقدم يتضح أن مجموع مربع انحراف القيم عن أية قيمة غير الوسط الحسابي، هي أكبر من مجموع مربع الانحرافات القيم عن الوسط الحسابي بمقدار $n(K - \bar{x})^2$ وهي قيمة موجبة دائماً. ولتوضيح الفكرة السابقة نورد المثال الآتي :

مثال (4-12) :

ليكن لدينا القيم الآتية :

$$x_i = 8, 7, 5, 4, 6$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{30}{5} = 6$$

$$\begin{aligned}\sum (x - \bar{x})^2 &= (8 - 6)^2 + (7 - 6)^2 + (5 - 6)^2 + (4 - 6)^2 + (6 - 6)^2 \\ &= 4 + 1 + 1 + 4 + 0 \\ &= 10\end{aligned}$$

إذا طرحنا من كل قيمة من هذه القيم قيمة معينة ما غير الوسط الحسابي (مثلاً :

$K = 8$) :

$$\begin{aligned}\sum (x - K)^2 &= \sum (x_i - 8)^2 \\ &= (8 - 8)^2 + (7 - 8)^2 + (5 - 8)^2 + (4 - 8)^2 + (6 - 8)^2 \\ &= 0 + 1 + 9 + 16 + 4 \\ &= 30\end{aligned}$$

وبالتالي فإن $10 < 30$ ونلاحظ أيضاً أن الفارق بينهما $20 = 30 - 10$ ما هو

$$20 = 5(8-6)^2 \Leftrightarrow n(K - \bar{x})^2$$

من المثال رقم (4-11) جدول رقم (4-12) احسب مجموع مربع الانحرافات القيم عن الوسط الحسابي، ثم احسب مجموع مربع الانحرافات القيم عن أية قيمة أخرى ولتكن $K = 16$.

جدول رقم (4-16)

x	f	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$	$(x - K)^2$	$(x - K)^2 f$
12	5	-8,4	70,56	352,8	16	80
16	18	-4,6	21,16	380,88	0	0
20	42	-0,6	0,36	15,12	16	672
24	27	3,4	11,56	312,12	64	1728
28	8	7,4	54,76	438,08	144	1152
Σ	100			1499		3632

بمساعدة الجدول (4-16) نلاحظ أن :

$$\Sigma (x - \bar{x})^2 f < \Sigma (x - K)^2 f$$

$$1499 < 3632$$

4-2-4- محاسن ومساوئ الوسط الحسابي :

محاسن الوسط الحسابي :

- 1- وضوح معناه وتعريفه .
- 2- سهولة حسابه ومعالجته إحصائياً .
- 3- يتصف بميزات جبرية لا يتصف بها أكثر المتوسطات .
- 4- إذا ما تم سحب عدة عينات من مجتمع ما، نجد أن الوسط الحسابي هو أقل تذبذباً عما عليه من مقاييس النزعة المركزية الأخرى، وهو بذلك يعتبر أفضل تقدير لمتوسط المجتمع .
- 5- إمكانية توظيفه لإيجاد مجموعة القيم Σx عند معرفة حجم العينة حيث:

$$\bar{x} = \frac{\sum x}{n} \Rightarrow \sum x = n \bar{x}$$

6- عملية حسابه تشمل جميع وحدات التوزيع .

مساوي الوسط الحسابي :

1- الوسط الحسابي قد يكون مقياساً مضللاً لوصف مجموعة من البيانات والسبب في ذلك يرجع إلى تأثيره بالقيم الشاذة. بمعنى وجود عدد قليل من القيم الكبيرة في مجموعة من القيم الصغيرة جداً. فمثلاً إذا كان لدينا 1000 شخص تبرع 997 منهم بمبلغ 1000 ل.س لكل شخص لبناء مبنى ماء، وتبرع الثلاثة الباقون ما قيمته 2000000 ل.س لكل شخص، فالوسط الحسابي لمقدار تبرع الشخص الواحد يساوي:

$$\bar{x} = \frac{(997)(1000) + (3)(2000000)}{1000} = 6997 \text{ ل.س}$$

نجد أن مقدار التبرع لأغلب المتبرعين (99,7% منهم) هو 1000 ل.س، وأن هذا المقدار يساوي تقريباً $\frac{1}{7}$ من الوسط الحسابي، فالوسط الحسابي ينحذب ويتجه باتجاه القيم الشاذة، وبالتالي نجد أن الوسط الحسابي في التوزيعات التكرارية المتماثلة أو القريبة من التماثل ممثلاً جيداً لهذه التوزيعات، أما في التوزيعات التكرارية المتوترة بشكل كبير، فإن الوسط الحسابي يفقد قيمته التمثيلية .

2- لا يمكن قياسه والتأكد منه بالطرق البيانية .

3- عدم إمكانية حسابه من جدول تكراري مفتوح «لأنه من المستحيل أن نعرف نقطة الوسط لفئة أحد حديها مفتوح، فحساب الوسط الحسابي يتطلب معرفة نقاط مراكز الفئات لجميع فئات الجدول التكراري، لذا إما أن يتم إهمال الفئات المفتوحة من الحساب أو إغلاق الفئة بدلالة أطوال الفئات الأخرى»، إذا كان الجدول التكراري ذا أطوال فئات واحدة ، لكن كلتا هاتين الطريقتين ليس مرغوباً فيها إلا إذا كنا بحاجة إلى الوسط الحسابي فقط.

فإذا كنا بحاجة لمقياس يعبر عن النزعة المركزية بشكل جيد في حال وجود فئات مفتوحة فيفضل استعمال مقياس آخر غير الوسط الحسابي .
4- لا يصلح في حال المتغيرات النوعية .

3-4- الوسيط (The Median) :

يُعدّ الوسيط مقياساً ترتيبياً Ordinal أو (مقياس موقع Position measure) لأن قيمته تعتمد على موقعه في البيانات، وهو بذلك عكس الوسط الحسابي الذي لا يستخدم إلا كمقياس كمي Interval Measurement .
ويعرف الوسيط بأنه القيمة التي تقسم القيم إلى جزأين بحيث يكون عدد القيم التي أقل منها يساوي عدد القيم التي أكبر منها، وذلك بعد ترتيب تلك القيم جميعها إما تصاعدياً أو تنازلياً، حيث تعتمد طريقة حساب القيمة التي تقسم البيانات إلى نصفين متساويين على موقع الوسيط سواء أكانت البيانات المبوبة أم غير المبوبة .

1-3-4- حساب الوسيط للبيانات غير المبوبة :

لايجاد قيمة الوسيط للبيانات غير المبوبة (الخام) يجب أولاً ترتيب تلك البيانات تصاعدياً أو تنازلياً، فإذا كان لدينا n من القيم أو المشاهدات $x_1, x_2, \dots, x_n$ المرتبة ترتيباً تصاعدياً أو تنازلياً وكانت n عدداً فردياً، فإن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$ ، أما إذا كانت n عدداً زوجياً فإن الوسيط هو الوسط الحسابي للقيمتين اللتين ترتيبهما $\frac{n}{2}$ و $\left(\frac{n}{2}\right) + 1$ ، والموجودتين في منتصف البيانات المرتبة، وذلك على الرغم من إمكانية اعتبار الوسيط أي قيمة بين هاتين القيمتين .

مثال (13-4) :

حصل طالب على العلامات الآتية في ثلاثة عشر مقررًا من مقررات السنة الأولى. احسب الوسيط لعلامات هذا الطالب .

100 - 95 - 50 - 55 - 51 - 54 - 52 - 51 - 52 - 53 - 55 - 50 - 55

نرتب البيانات ترتيباً تصاعدياً أو تنازلياً وفقاً لقيمتها، ولننحتر مثلاً الترتيب التصاعدي:

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13)

50 - 50 - 51 - 51 - 52 - 52 - 53 - 54 - 55 - 55 - 55 - 95 - 100

وموقع أو ترتيب الوسيط هو :

$$\frac{n+1}{2} = \frac{13+1}{2} = 7$$

حيث n عدد فردي فإن الوسيط هو قيمة المفردة السابعة، وفي هذه المسألة فإن المفردة السابقة تساوي 53 وهي قيمة الوسيط فالقيمة 53 تقع في منتصف البيانات حيث توجد ست مفردات قبلها وست مفردات بعدها.

ولإيجاد الوسيط للمفردات الزوجية الآتية :

↓ ↓
(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

3 - 4 - 4 - 4 - 4 - 5 - 7 - 7 - 9 - 11

نجد أن : $n = 10$ لذلك ترتيب الوسيط الأول =

$$\frac{n}{2} = \frac{10}{2} = (5)$$

أي القيمة التي ترتيبها الخامس وتساوي 4 .

وترتيب الوسيط الثاني =

$$\frac{n}{2} + 1 = \frac{10}{2} + 1 = (6)$$

أي القيمة التي ترتيبها السادس وتساوي 5 .

فالوسيط هو :

$$\frac{4+5}{2} = \frac{9}{2} = 4,5$$

وبالتالي فالوسيط في البيانات الزوجية يساوي الوسط الحسابي لقيمة المفردتين التي تنصف البيانات إلى نصفين متساويين، وعلى الرغم من ذلك يمكن اعتبار قيمة الوسيط أية قيمة بين هاتين المفردتين، فيمكن اعتبار قيمة الوسيط للمثال السابق هو 4.1 أو 4.9 أو أية قيمة أخرى بين القيمتين 5 و 4 لأن أيًا من هذه القيم 4.3 أو 4.6 أو ... يقسم البيانات إلى قسمين متساويين .

2-3-4 - حساب الوسيط للبيانات المبوبة في فئات :

يتم حساب الوسيط في حالة البيانات المبوبة بالحساب وبالرسم .

(أ) الطريقة الحسابية :

إذا كانت $x_1, x_2, \dots, x_n$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكرارها $f_1, f_2, \dots, f_n$ على التوالي، فإن قيمة الوسيط لهذه البيانات هي :

$$M_{ed} = L_m + \left(\frac{\frac{\sum f}{2} - f_{m-1}}{f_m} C_m \right)$$

حيث :

M_{ed} = قيمة الوسيط .

$\sum f$ = مجموع التكرارات .

f_{m-1} = التكرار التجميعي الصاعد السابق لفئة الوسيط (مجموع كل التكرارات

قبل الفئة الوسيطة) .

f_m = تكرار الفئة الوسيطة .

C_m = طول الفئة الوسيطة .

وخطوات إيجاد الوسيط للبيانات المبوبة هي :

أ - إيجاد التكرار التجميعي الصاعد (أو الهابط) .

ب- نوجد ترتيب الوسيط وهو $\frac{\sum f}{2} = \frac{n}{2}$ (سواء أكان هذا المجموع زوجياً أم فردياً).

ج- نبحث في التكرار التجميعي الصاعد عن الرقم $\frac{\sum f}{2}$ أو على رقم أكبر منه مباشرة.

د - نتعرف على الفئة الوسطية، والتي تكون مقابلة للرقم الذي أوجدناه من تدقيق التكرار التجميعي .

هـ - نطبق المعادلة :

$$M_{ed} = L_m + \left(\frac{\frac{\sum f}{2} - f_{m-1}}{f_m} \cdot C_m \right)$$

مثال (4-14) :

بالعودة إلى الجدول (4-6) مثال رقم (4-6) عن علامات 300 طالب نقوم بحساب الوسيط لهذا الجدول التكراري :

عدد الطلاب	فئات الدرجات
20	30 - 40
40	40 - 50
60	50 - 60
100	60 - 70
40	70 - 80
30	80 - 90
10	90 - 100

نوجد التكرار التجميعي الصاعد أو الهابط :

جدول رقم (4-17)

	x	f	التكرار التجميعي الصاعد	التكرار التجميعي الحابط
30 - 40	35	20	20	300
40 - 50	45	40	60	280
50 - 60	55	60	120	240
[60 - 70]	65	100	220	180
70 - 80	75	40	260	80
80 - 90	85	30	290	40
90 - 100	95	10	300	10
		300		

من تدقيق التكرار التجميعي الصاعد، نجد أن المفردة :

$$\frac{\sum f}{2} = \frac{300}{2} = 150$$

تقع في الفئة التي مركزها 65 وهي الفئة 60 - 70 (الفئة الوسيطة) وبالتالي :

$$L_m = 60$$

$$f_m = 100$$

$$f_{m-1} = 120$$

$$C_m = 10$$

فالوسيط يساوي :

$$M_{ed} = L_m + \left(\frac{\frac{\sum f}{2} - f_{m-1}}{f_m} \cdot C_m \right)$$

$$= 60 + \left(\frac{150 - 120}{100} \cdot 10 \right)$$

$$= 60 + 3$$

$$= 63 \quad \text{علامة}$$

نلاحظ (*) :

من التكرار التجميعي الصاعد والقيمة السابقة لفئة الوسيط هي 120، وهذا يسبق المفردة الوسيطة (المفردة الوسيطة ترتيبها 150) بثلاثين مفردة، أي إن الوسيط هو المفردة 3 داخل الفئة الوسيطة لأنه في التوزيعات التكرارية يفترض أن مفردات الفئة الوسيطة تتوزع بانتظام داخل الفئة، وبالتالي تبعد كل مفردة عن الأخرى بمقدار يساوي:

$$\frac{\text{طول الفئة}}{\text{تكرار الفئة}}$$

فإذا علمنا أن المفردة الوسيطة هي المفردة رقم $30 = (150 - 120)$ داخل الفئة الوسيطة فإنها تبعد عن الحد الأدنى بمقدار :

$$3 = 30 \times \frac{10}{100} = 30 \times \left(\frac{\text{طول الفئة}}{\text{تكرار الفئة}} \right)$$

ونلاحظ وجود 10 مفردات داخل الفئة الوسيطة، وبالتالي فالمفردة الوسيطة سوف تزيد عن الحد الأدنى لفئة الوسيط بمقدار $\frac{3}{10}$ من طول هذه الفئة، لذا يجب إضافة $10 \left(\frac{3}{10} \right) = 3$ إلى الحد الأدنى لفئة الوسيط لنحصل على قيمة الوسيط وهي 63.

(ب) الطريقة البيانية :

يمكن أيضاً إيجاد قيمة الوسيط عن طريق رسم المنحنى التجميعي الصاعد (أو الهابط) ومن ثم :

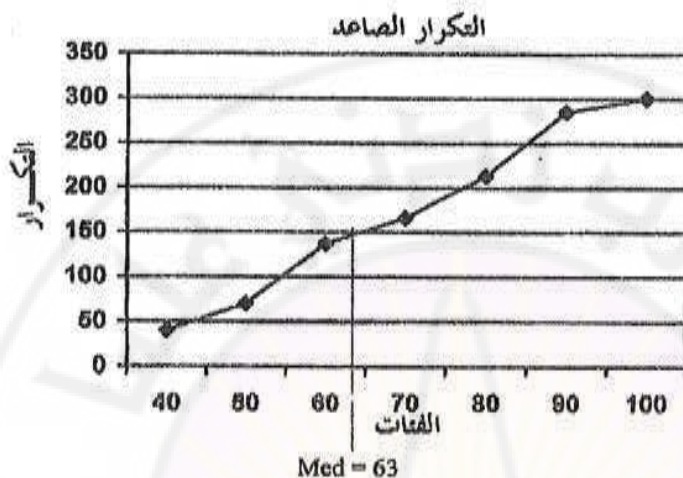
1- نعين ترتيب الوسيط $\frac{\sum f}{2}$ على المحور العامودي $150 = \frac{300}{2}$.

2- ثم نرسم مستقيماً أفقياً من نقطة ترتيب الوسيط (150) يوازي المحور الأفقي ويقطع المنحنى التجميعي الصاعد (أو الهابط) في نقطة وتكن A.

(٥) الإحصاء في الإدارة : لنكولن تشاو - دار المريخ ص 97.

3- نسقط من النقطة A عموداً يوازي المحور العمودي، ويتقاطع مع المحور الأفقي في نقطة Med تكون هي قيمة الوسيط .

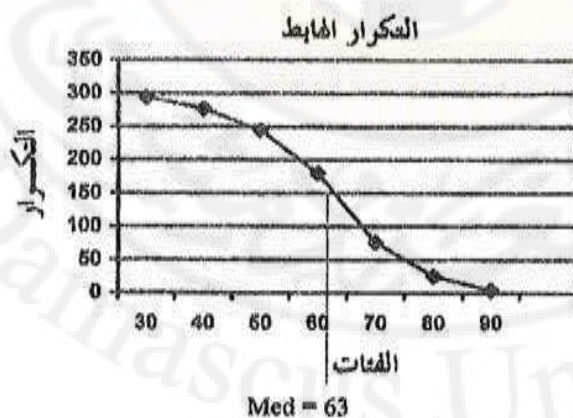
إيجاد قيمة الوسيط برسم المنحنى التجميعي الصاعد لعلامات 300 طالب



الشكل (4-1)

وكلما كان الرسم دقيقاً نحصل على قيمة للوسيط بدقة أكبر .

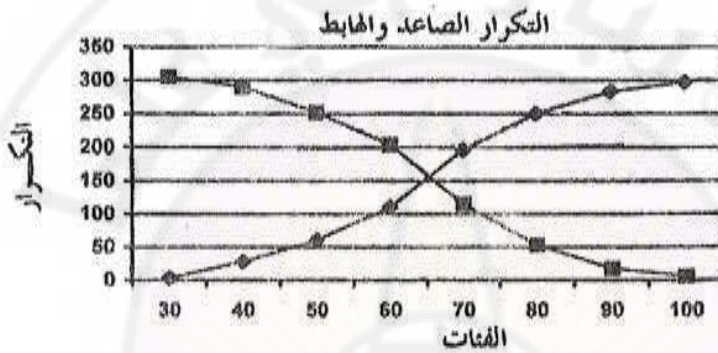
إيجاد قيمة الوسيط برسم المنحنى التجميعي الهابط لعلامات 300 طالب



الشكل (4-2)

ويمكن أيضاً إيجاد قيمة الوسيط وبدقة أكبر برسم المنحنيين التجميعي الصاعد والتجميعي الهابط معاً، حيث إن نقطة التقاطع للمنحنيين تقع عند منتصف مجموع التكرارات (أي عند نقطة ترتيب الوسيط)، وبذلك يكون الخط النازل من نقطة التقاطع الموازي للمحور العمودي يتقاطع مع المحور الأفقي في نقطة Med (قيمة الوسيط).

إيجاد قيمة الوسيط بالرسم للمنحنيين التجميعيين الصاعد والهابط معاً
لعلامات 300 طالب



الشكل (4-3)

وكما أشرنا سابقاً بأن الوسيط يعتبر من مقاييس الموقع، وقيمه تكون في منتصف البيانات، بمعنى أن 50% من البيانات أكبر من قيمة الوسيط، و 50% من البيانات هي أقل من قيمة الوسيط بالإضافة إلى مقياس الوسيط. توجد أيضاً بعض المقاييس الأخرى، والتي تعتبر أيضاً من مقاييس الموقع مثل العشرييات والربيعات والمئينات، والتي تحدد لنا مواقع غير مركزية وإن مهمتها هي إيجاد القيم التي تزيد عن 10% أو 25% أو 40% من المعطيات، وإن هذه التقسيمات غير ضرورية مع البيانات القليلة وغير المبوبة، ولذلك سوف نتناول تطبيقاتها في البيانات المبوبة.

(أ) العشير Decile :

يمكن أن نعرف العشير الأول بأنه القيمة التي تقع عند العشير $\left(\frac{1}{10}\right)$ الأول من البيانات، بمعنى أن $\frac{1}{10}$ من القيم أقل منه و $\frac{9}{10}$ من القيم أكبر منه. أما العشير الثاني فهو القيمة التي تقع عند $\frac{2}{10}$ الأولى من البيانات، بمعنى أن $\frac{2}{10}$ من القيم أقل منه و $\frac{8}{10}$ من القيم أكبر منه، وهكذا بالنسبة للأعشار الأخرى، وإن تحديد موقع العشير يتحدد بالصيغة الآتية :

$$\frac{(i) \sum f_i}{100}$$

حيث i ترمز إلى العشير ($i = 10, 20, \dots$) .

فموقع العشير الأول هو $\frac{10 \sum f_i}{100}$ وهو يساوي حسب المثال السابق :

$$30 = \frac{(10) \cdot (300)}{100}$$

و بتطبيق العلاقة التالية :

$$D_1 = L_{D_1} + \frac{\frac{(i) \sum f_i}{100} - f_{D_1-1}}{f_{D_1}} \cdot C_{D_1}$$

حيث : D_1 = قيمة العشير الأول .

L_{D_1} = الحد الأدنى لفئة العشير الأول .

f_{D_1-1} = التكرار التجميعي الصاعد السابق لفئة العشير الأول (مجموع

كل التكرارات قبل فئة العشير الأول) .

f_{D_1} = التكرارات المقابلة لفئة العشير الأول .

C_{D_1} = طول الفئة العشرية .

مثال (4-15) :

وبالاستفادة من الجدول رقم (4-6) مثال رقم (4-6) نجد :

تكرار تجميعي صاعد	عدد الطلاب (n)	فئات العلامات
20	20	30 - 40
60	40	40 - 50
120	60	50 - 60
220	100	60 - 70
260	40	70 - 80
290	30	80 - 90
300	10	90 - 100
	300	

من التكرار التجميعي الصاعد نبحث عن القيمة (30) موقع العشير الأول أو القيمة التي أكبر منها فنجد بأن فئة العشير الأول هي 40-50 وبالتالي نطبق الدستور السابق:

$$D_1 = 40 + \frac{\frac{(10)(300)}{100} - 20}{40} \cdot 10$$

$$= 40 + 2,5 = 42,5 \quad \text{علامة}$$

بمعنى أن هناك 90% من الطلاب قد حصلوا على علامات أكبر من 42,5 علامة، وأن هناك 10% من الطلاب قد حصلوا على علامات أقل من 42,5 علامة .
وبالمثل للحصول على العشير الثاني نجد أن موقعه :

$$60 = \frac{(20)(300)}{100}$$

يقابل الفئة السابقة نفسها [40 - 50]

$$D_2 = 40 + \frac{60 - 20}{40} \cdot 10 =$$

$$= 40 + 10 = 50 \text{ علامة}$$

بمعنى أن 80% من الطلاب حصلوا على علامات أكبر من 50 علامة، وأن هناك 20% منهم حصلوا على علامات أقل من 50 علامة .

(ب) الربع Quartile :

يمكن أيضاً تحديد مواقع الأرباع التي تنقسم البيانات إلى أربعة أرباع، فالربع الأول Q_1 ويقصد به بأنه القيمة التي تكون 25% من البيانات. أقل منها و 75% من البيانات أكبر منها، و Q_2 (الربع الثاني) هي القيمة التي تكون 50% من البيانات أقل منها و 50% من البيانات أكبر منها ، أما الربع الثالث Q_3 فإن قيمته تقع عند البيانات بحيث أن 75% أقل من قيمة الربع الثالث و 25% أكبر من الربع الثالث .
والجددير بالذكر بأن الربع الثاني Q_2 يفترض أن تتطابق قيمته مع الوسيط Med في تقسيم المساحة تحت المنحنى. وأن موقع الربع Q_i حيث $i = 1, 2, 3$ يتحدد بالصيغة الآتية:

$$\frac{(i) (\sum f)}{4}$$

فمن التكرار التجميعي الصاعد نوجد موقع الربع الثالث والأول، ومن ثم نوجه الفئة الربعية المقابلة، وبعد ذلك نطبق العلاقة الآتية لتحديد قيمة الربع :

$$Q_i = L_{Q_i} + \frac{\frac{(i) \sum f}{4} - f_{Q_{i-1}}}{f_{Q_i}} \cdot C_{Q_i}$$

حيث : $f_{Q_{i-1}}$ = التكرار التجميعي الصاعد السابق لفئة الربع i .

f_{Q_i} = تكرار الفئة الربعية i .

C_{Q_i} = طول الفئة الربعية i .

L_{Q_i} = الحد الأدنى لفئة الربع i .

Q_i = قيمة الربع .

مثال (4-16) :

ولإيجاد الربع الثالث والأول للجدول التكراري جدول رقم (4-6) نحدد موقع

الربع الثالث Q_3 :

$$225 = \frac{(3)(300)}{4}$$

نطبق علاقة قيمة الربع الثالث :

$$Q_3 = L_{Q_3} + \frac{\frac{(3)(300)}{4} - f_{Q_3-1}}{f_{Q_3}} \cdot C_{Q_3}$$

نجد من الجدول أن موقع الربع الثالث هو 225 يتحدد من التكرار التجميعي الصاعد، ويقابل الفئة (70-80) :

$$\begin{aligned} Q_3 &= 70 + \frac{225 - 220}{40} \cdot 10 \\ &= 70 + 1,25 = 71,25 \end{aligned} \quad \text{علامة}$$

71,25 علامة تقسم الطلاب إلى 75% منهم حصلوا على علامات أقل من 71,25 و 25% منهم حصلوا على علامات أكبر من 71,25 علامة .
ولنحدد موقع الربع الأول Q_1 :

$$75 = \frac{(1)(300)}{4}$$

ومن التكرار التجميعي الصاعد، يتحدد موقع الربع الأول عند الـ 75 أو أكبر منها مباشرة، حيث نجد أن الفئة الربعية (50-60) وقيمة الربع الأول :

$$\begin{aligned} Q_1 &= 50 + \frac{75 - 60}{60} \cdot 10 \\ &= 50 + 0,83 \\ &= 50,83 \end{aligned} \quad \text{علامة}$$

25% من الطلاب حصلوا على علامات أقل من 50.83 علامة و 75% منهم حصلوا على أكبر من 50.83 علامة .

جـ) المئين Percentile :

يقصد بالمئين هي القيمة التي تقسم البيانات لأجزاء صغيرة كواحدات من المائة
فمثلاً المئين الأول هو القيمة التي تقع عند $\frac{1}{100}$ من البيانات. ولنرمز له بـ W_1 والمئين السابع هو القيمة التي تقع عند $\frac{7}{100}$ من البيانات W_7 ولإيجاد موقع W_i :

$$\frac{(i) \sum f}{100}$$

حيث $i = 1, 2, 3, \dots$

وإن قيمة المئين تحسب بالعلاقة الآتية :

$$W_i = L_{w_i} + \frac{\frac{(i) \sum f}{100} - f_{w_i-1}}{f_{w_i}} \cdot C_{w_i}$$

مثال (4-17) :

ولتحديد فئة المئين للمثال رقم (4-6) نبحث في جدول التكرار التجميعي
الصاعد عن موقع المئين، فلإيجاد المئين الـ 60 مثلاً نوجد موقعه :

$$180 = \frac{(60)(300)}{100}$$

يقابل الفئة (60 - 70) وبتطبيق علاقة قيمة المئين :

$$W_{60} = 60 + \frac{180 - 120}{100} \cdot 10$$

$$= 60 + 6 = 66 \text{ علامة}$$

فإن 66 علامة تقسم البيانات إلى أن 40% حصلوا على علامات أكبر من 66
وأن 60% حصلوا على علامات أقل من 66 علامة .

كما يمكن تحديد قيمة العشرية والربعية والمئينية عن طريق رسم منحني التكرار التجميعي الصاعد (أو الهابط)، كما هو الحال بالنسبة إلى الوسيط، وذلك باستخدام صيغ موقع كل من العشرية والربعية والمئينية، وتحديد ذلك الموقع على الخط العامودي لمحوري الإحداثيات، ومن ثم رسم خط أفقي من ذلك الموقع يوازي المحور الأفقي، ويتقاطع مع منحني التكرار التجميعي الصاعد (أو الهابط) في نقطة ما وبإسقاط تلك النقطة على المحور الأفقي (محور الفئات) تتحدد كل من قيم المواقع المراد حسابها.

4-3-4- خصائص الوسيط ومحاسنه :

من أهم خواص الوسيط :

- 1- إن مجموع انحرافات القيم عن وسيطها بالقيمة المطلقة، هي دائماً أصغر من مجموع انحرافات تلك القيم عن أية قيمة أخرى بالقيمة المطلقة بمعنى :

$$\sum |X - \text{Med}| < \sum |X - K|$$

حيث $K \neq \text{Med}$.

مثال (4-18) :

ليكن لدينا البيانات الآتية التي تمثل أجور تسعة عمال في إحدى الشركات:
احسب القيمة المتوسطة لأجور هؤلاء العمال باستعمال الوسيط، ثم احسب مجموع انحرافات القيم عن هذا الوسيط بالقيمة المطلقة، وأثبت أن هذا المجموع هو أصغر من أي مجموع ممكن لانحرافات القيم عن أية قيمة كانت (بالقيمة المطلقة) :

$$[2500, 5000, 5000, 5500, 6000, 6000, 7000, 7000, 50000]$$

$$\frac{n+1}{2} = \frac{10}{2} = 5 \quad \text{نوجد موقع أو ترتيب الوسيط}$$

وبالتالي الوسيط يساوي 6000 ل.س

لنوجد مجموع انحرافات القيم عن الوسيط (بالقيمة المطلقة) ويساوي :

$$\begin{aligned}\Sigma |X - \text{Med}| &= 3500 + 1000 + 1000 + 500 + 0 + 0 + 1000 + 1000 + 44000 \\ &= 52000\end{aligned}$$

وبفرض $K = 5000$

بمجموع انحرافات القيم عن تلك القيمة (بالقيمة المطلقة) يساوي :

$$\begin{aligned}\Sigma |X - K| &= 2500 + 0 + 0 + 500 + 1000 + 1000 + 2000 + 2000 + 45000 \\ &= 54000\end{aligned}$$

فنجد أنه دوماً :

$$\Sigma |X - \text{Med}| < \Sigma |X - K|$$

$$52000 < 54000$$

2 - عدم تأثره بشكل مباشر بالقيم المتطرفة في تمثيله للبيانات فهو يتحدد بموقعه، ففي الحالات التي توجد فيها قيم متطرفة (شاذة) نتأجه أفضل مما لو استعملنا الوسط الحسابي، والذي يتأثر كما مر معنا سابقاً بالقيم الشاذة ، فقيمة الوسيط موضعية، ويفضل استعماله إذا ازدادت درجة الالتواء في التوزيعات التكرارية .

3 - بالإمكان استخدام الوسيط في الجداول التكرارية، والتي تحتوي على فئات مفتوحة فمعادلة الوسيط لا تتطلب حساب مراكز الفئات بعكس الوسط الحسابي الذي يشترط معرفة مركز كل فئات التوزيع، وكذلك يمكن استخدامه مع أطوال فئات غير متساوية، فلا يهمنا إلا طول الفئة الوسيطة .

أما مساوئ الوسيط فهي :

1 - طرق حسابه تقريبية .

2 - في حال المفردات القليلة، لا يعبر الوسيط عن مركز تجمع المفردات بصورة صحيحة .

3 - يعتمد الوسيط على قيمة واحدة أو قيمتين في البيانات غير المبوبة، وعلى فئة واحدة فقط في حال البيانات المبوبة، وبالتالي لا يأخذ جميع قيم التوزيع بعين الاعتبار، فهو ذو حساسية كبيرة للقيم الوسيطة .

4 - قد يتطلب لعمليات غير حيزية مطولة، وبالأخص في حالة البيانات غير المبوبة كترتيب البيانات بشكل تصاعدي أو تنازلي .

5 - لا يمكن حساب الوسيط العام لعدة مجموعات من البيانات .
في كثير من الأحيان يتم تقسيم البيانات إلى مجموعات وفقاً لصفات معينة، ففي مثل هذه الحالات سوف نحتاج إلى مقياس آخر من مقاييس النزعة المركزية .

4-4- المنوال «The Mode» :

يعتبر المنوال أحد مقاييس النزعة المركزية الهامة، لأنه المقياس الوحيد الذي يستخدم لحساب المتوسط لظاهرة ما لا يمكن قياسها بالمقياس الكمي (المتغيرات النوعية)، فهذا المقياس يمكن استخدامه للقيم الكمية والنوعية (الوصفية)، ويعرف المنوال بمجموعة من القيم بأنه القيمة أو الصفة الأكثر تكراراً أو الأكثر شيوعاً .

مثال (4-19) :

البيانات الآتية تمثل التوزيع التكراري لـ 128 زوجاً من الأحذية مباعه في إحدى المحلات خلال شهر ما وذلك حسب القياس :

القياس	40	41	42	43	44	45
عدد الأحذية	20	25	35	20	18	10

في هذا المثال نجد أن القياس رقم 42 هو الأكثر مبيعاً والأكثر طلباً، وبالتالي فالمنوال هو 42 ونجد في هذه الحالة بأن الوسيط الحسابي هو قيمة غير موجودة عملياً، وكذلك الوسيط باعتبار أن هذه البيانات هي بيانات وصفية .

مثال (4-20) :

ليكن لدينا مبيعات 128 زوجاً من الأحذية مباعه في إحدى المحلات خلال شهر ما وذلك بحسب اللون :

اللون	أسود	بني غامق	بني فاتح	عسلي
عدد الأحذية	60	40	14	14

نجد أن المقياس المناسب لإيجاد اللون كممثل للظاهرة هو المتوال، حيث نجد أن اللون الأسود هو الأكثر شيوعاً والأكثر تكراراً والأكثر مبيعاً. فمتوال لون الحذاء هو اللون الأسود، فلا يمكننا والحالة هكذا إيجاد الوسط الحسابي أو الوسيط للون . والجدير بالذكر فإن قيمة المتوال قد لا تكون الوحيدة، فيمكن أن تكون هناك أكثر من قيمة متوالية واحدة وقد لا توجد قيمة متوالية بين القيم، ويحصل ذلك عندما تكون جميع القيم لها العدد نفسه من التكرارات .

مثال (21-4) :

إذا كان لدينا القيم الآتية : 4, 3, 5, 3, 6, 4, 3, 2

فالمتوال لهذه المجموعة هو القيمة (3) لتكرارها أكثر من غيرها .

وإذا كان لدينا القيم الآتية : 3, 3, 3, 2, 2, 2, 6, 6, 7, 7, 5

فالمتوال لهذه المجموعة هو القيمتين (2) و (3) (ثنائية المتوال) .

وإذا كان لدينا القيم الآتية : 80, 60, 40, 70, 90, 100

فنلاحظ أنه لا توجد قيمة متوالية بين القيم السابقة، حيث إن هذه القيم لها العدد نفسه من التكرارات .

1-4-4- حساب المتوال للبيانات غير المبوبة :

إن طريقة حساب المتوال للبيانات الخام (غير المبوبة) بسيطة جداً حيث يتم حسابه لمثل هذه البيانات من واقع تعريفه مباشرة (كما مر معنا سابقاً) .

2-4-4- حساب المتوال للبيانات المبوبة في فئات :

إن حساب المتوال للبيانات المبوبة في فئات يتطلب ما يسمى بتوفيق أفضل منحنى يمثل التوزيع، حيث -بعد هذا التوفيق- نوجد القيمة التي تكون الأكثر تكراراً،

والمناظرة لقيمة المنحنى الموفق فتكون هي قيمة المنوال، وإن عملية توفيق أفضل منحني يتطلب استنتاج معادلة لهذا المنحنى تمكننا من الحصول على قيمة المنوال بشكل دقيق، وإن عملية توفيق ذلك المنحنى سوف يترك للعام القادم، فهو ليس مجال دراستنا في هذا الكتاب .

إلا أنه يمكن إيجاد قيمة المنوال لجدول تكراري بطريقتين بالحساب وبالرسم:

أ) الطريقة الحسابية :

هناك عدة طرائق حسابية لتقدير قيمة المنوال منها :

1- طريقة الرافعة .

2- طريقة بيرسون (الفروق) .

■ طريقة الرافعة :

إن قيمة المنوال سوف تقع ضمن الفئة المنوالية (التي تقابل أكبر تكرار في التوزيع)، وعادة ما تتحاذر قيمة المنوال إما باتجاه بداية الفئة المنوالية، أو باتجاه نهايتها تبعاً لتكرارات الفئتين السابفة واللاحقة للفئة المنوالية، وتتطابق قيمة المنوال مع مركز الفئة المنوالية إذا تساوى تكراري الفئتين السابفة واللاحقة للفئة المنوالية، فإذا كان تكرار الفئة السابفة للفئة المنوالية أكبر من تكرار الفئة اللاحقة للفئة المنوالية، فإن المنوال سوف يميل نحو بداية الفئة المنوالية، والعكس صحيح، وعلى ذلك سوف تكون قيمة المنوال داخل الفئة المنوالية بنسبة التكرارين السابق واللاحق لهذه الفئة المنوالية، ففي طريقة حساب المنوال باستخدام تكراري الفئتين السابفة واللاحقة للفئة المنوالية (طريقة الرافعة) يتم تقسيم طول الفئة المنوالية تقسيماً يتناسب مع التكرارين السابق واللاحق، وبالتالي يمكن حساب المنوال بحسب هذه الطريقة التقريبية كما يأتي :

$$\text{Mod} = \text{Lo} + \frac{f_{o_2}}{f_{o_2} + f_{o_1}} \cdot \text{Co}$$

حيث :

Mod = قيمة المتوال .

Lo = الحد الأدنى للفترة المتوالية .

fo₂ = التكرار اللاحق للفترة المتوالية .

fo₁ = التكرار السابق للفترة المتوالية .

Co = طول الفترة المتوالية .

مثال (4-22) :

وبالعودة إلى الجدول التكراري السابق عن علامات 300 طالب من الجدول رقم

(4-6) نجد :

فئات العلامات	f
30 - 40	20
40 - 50	40
50 - 60	60
الفترة المتوالية المقابلة لأكثر تكرار [60 - 70]	(100)
70 - 80	40
80 - 90	30
90 - 100	10
	300

$$\text{Mod} = 60 + \frac{40}{40 + 60} \cdot 10$$

$$= 60 + 4 = 64$$

إن العلامة الأكثر تردداً والأكثر تكراراً وشيوعاً هي 64 علامة .

■ طريقة بيرسون (الفروق) :

إن الطريقة السابقة (طريقة الرافعة) في حساب القيمة المتوالية تهمل تكرار الفترة

المتوالية نجد أن طريقة بيرسون في حساب قيمة المتوال تتلافى هذا النقص الموجود في

طريقة الرافعة، حيث تعتمد طريقة بيرسون على تكرار الفئة المنوالية، وتكرار الفئة السابقة لها واللاحقة لها، ولتحديد قيمة المنوال داخل الفئة المنوالية يؤخذ الفرق بين تكراري الفئة المنوالية والفئة السابقة لها، وكذلك الفرق بين تكراري الفئة المنوالية والفئة اللاحقة لها كعاملين مؤثرين في تحديد قيمة المنوال، وقيمة المنوال حسب طريقة بيرسون تحسب وفق العلاقة الآتية :

$$\text{Mod} = L_o + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot C_o$$

حيث : Δ_1 = الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.

Δ_2 = الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.

C_o = طول الفئة المنوالية .

L_o = الحد الأدنى للفئة المنوالية .

مثال (4-23) :

وبالرجوع إلى الجدول التكراري السابق جدول رقم (4-6) عن علامات 300 طالب نجد :

$$\begin{aligned} \text{Mod} &= 60 + \frac{(100 - 60)}{(100 - 60) + (100 - 40)} \cdot 10 \\ &= 60 + \frac{40}{40 + 60} \cdot 10 \\ &= 60 + 4 = 64 \quad \text{علامة} \end{aligned}$$

نلاحظ في هذا المثال أن الطريقتين توديان لنتيجة واحدة للمنوال، لأن الفارق بين تكرار الفئة المنوالية والتكرار السابق لها يساوي تماماً تكرار الفئة اللاحقة للفئة المنوالية، وإن الفارق بين تكرار الفئة المنوالية والتكرار اللاحق لها يساوي تماماً تكرار الفئة السابقة للفئة المنوالية .

فمثلاً : إذا كانت الفئة المنوالية تساوي [5-15] وتكرارها يساوي 120 والتكرار السابق لهذه الفئة يساوي 30 والتكرار اللاحق لهذه الفئة المنوالية يساوي 70 وإن الجدول التكراري ذو فئات متساوية الطول ، فالمنوال بطريقة الرافعة يساوي :

$$\text{Mod} = 5 + \frac{70}{70 + 30} \cdot 10$$

$$= 5 + 7 = 11$$

وبطريقة بيرسون :

$$\text{Mod} = 5 + \frac{(120 - 30)}{(120 - 30) + (120 - 70)} \cdot 10$$

$$= 5 + \frac{90}{90 + 50} \cdot 10$$

$$= 5 + 6,428$$

$$= 11,428$$

نلاحظ أن هناك بعض الاختلاف البسيط بين قيمتي المنوال، غير أن ما يؤخذ عليه (المنوال) أنه مقياس غير مستقر، وإن قيمة المنوال تتوقف على كيفية تبويب البيانات للتوزيعات التكرارية، فإذا كان التبويب في فئات ذي أطوال مختلفة فإن ذلك سوف يؤدي إلى اختلاف في قيمة المنوال بالنسبة للبيانات فيما إذا تم تبويبها في فئات ذات أطوال متساوية. في مثال علامات 300 طالب كانت أطوال الفئات متساوية، أما إذا كانت أطوال الفئات غير متساوية، فعندها يجب على الباحث إجراء بعض التعديلات على التكرارات بأن يعدل هذه التكرارات لكل فئة بالصيغة :

$$\frac{\text{التكرار الفعلي للفئة}}{\text{طول الفئة}}$$

والفئة المنوالية هي الفئة التي يوجد أمامها أكبر تكرار معدل. وفي هذه الحال يمكن للباحث أن يستخدم طريقة بيرسون المذكورة سابقاً. والجدير بالذكر أن المدرج التكراري يرسم من التكرارات المعدلة إذا كانت أطوال الفئات غير متساوية.

أما^(*) إذا أراد الباحث استخدام التكرارات الأصلية لجدول تكراري أطوال فئاته مختلفة دون إدخال أي تعديلات، فيمكن أن يحسب قيمة المنوال في هذه الحالة بالصيغة الآتية :

$$\text{Mod} = L_o + \frac{(Co_1) (fo)}{(Co_2) (fo_1) + (Co_1) (fo_2)} \cdot Co$$

حيث : L_o = الحد الأدنى للفترة المنوالية .

Co_1 = طول الفترة قبل المنوالية .

Co_2 = طول الفترة بعد المنوالية .

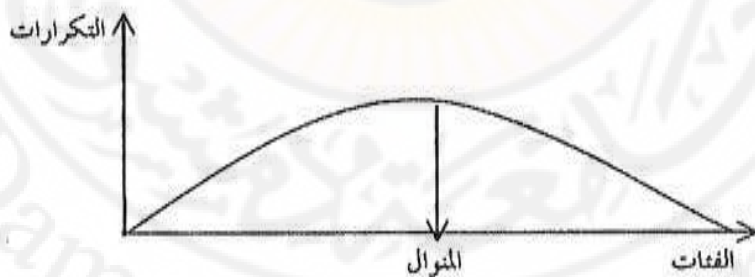
fo_1 = تكرار الفترة قبل المنوالية (السابق لها) .

fo_2 = تكرار الفترة بعد المنوالية (اللاحق لها) .

Co = طول الفترة المنوالية .

ب) الطريقة البيانية :

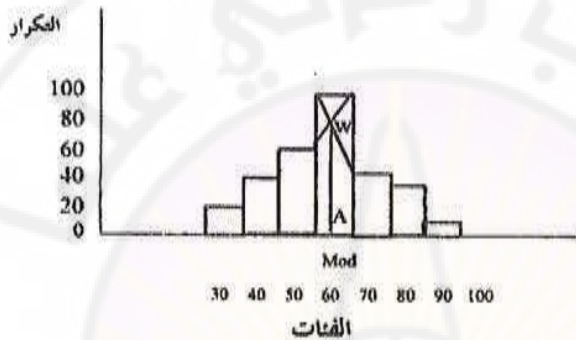
1- المنحنى التكراري: من تعريف المنوال والذي هو القيمة التي تقابل أكبر تكرار، فإن تم رسم المنحنى التكراري بإسقاط عامود من أعلى نقطة من المنحنى التكراري على المحور الأفقي (محور الفئات)، فإنه سوف يقطع محور الفئات في نقطة هي قيمة المنوال كما في الشكل (4-4) :



الشكل (4-4)

(*) أساسيات الإحصاء الاجتماعي : د. اعتماد علام ص 111 .

2- المدرج التكراري: نقوم برسم المدرج التكراري وذلك باستعمال مستطيل الفئة المتوالية (وهو أعلى مستطيل لأنه يمثل أكثر التكرارات) والمستطيلان المجاوران له وبعد ذلك نصل الركن الأيسر العلوي لفئة المتوال بالركن الأيسر العلوي للفئة التالية بخط مستقيم، ثم نصل الركن الأيمن العلوي لفئة المتوال بالركن الأيمن العلوي للفئة السابقة بخط مستقيم فيتقاطع الخطان في نقطة (w)، نسقط عموداً على محور الفئات ليلتقي به عند قيمة المتوال كما في الشكل (4-5) :



الشكل (4-5)

نجد من الرسم أن المتوال يقابل تقريباً القيمة 64 .

3-4-4- محاسن ومساوي المتوال :

محاسن المتوال :

- 1- عدم تأثره بالقيم المتطرفة (الشاذة) .
- 2- يمثل غالبية المشاهدات مع أن حسابه لا يحتاج لجميع قيم التوزيع.
- 3- إمكانية حسابه في حال الجداول التكرارية ذات الفئات المفتوحة.
- 4- له فائدة كبيرة للبيانات الوصفية إذا أردنا الحصول على مقياس للنزعة المركزية لهذه البيانات مثل اللون والحالة الاجتماعية ، الدرجة العلمية، الوظيفة، إلخ. ففي مثل هذه الأحوال لا نستطيع إيجاد الوسط الحسابي أو الوسط كمتقاس يعبر عن المتوسط، عندها يكون المتوال هو المقياس المناسب لمثل هذه الأنواع من

البيانات، حيث نستطيع أن نتكلم عن لون العيون الأكثر شيوعاً في بلد ما أو مستوى التعليم الأكثر شيوعاً في مؤسسات القطاع العام، ففي الحالة الأولى نتكلم عن منوال لون العيون، وفي الحالة الثانية نتكلم عن منوال الحالة التعليمية وهكذا...

مساوي المنوال :

1- غير دقيق حيث يتم حسابه بطرق كلها تقريبية تعطي نتائج مختلفة عندما تطبق على الجدول التكراري الواحد، وإن الأسلوب الواحد إذا ما طبق على البيانات الإحصائية نفسها، فإنه قد يعطي تقديرات مختلفة للمنوال باختلاف أطوال الفئات.

2- عندما تكون القيم منتشرة على مدى واسع، عندها يصبح المنوال أقل تمثيلاً كموسط .

3- لا يمكن حسابه في حالة التوزيعات التكرارية ذات المنحنى النوني كما في الشكل الآتي :



الشكل (4-6)

4-4-4- العلاقة التقريبية بين الوسط الحسابي والوسيط والمنوال :

Approximate Relation the Mean, Median, and Mode

لمعرفة مدى العلاقة بين المتوسطات الثلاث نورد المثال التالي :

مثال (4-24) :

الجدول الآتي يتضمن الأجر اليومي بالليرة لـ 140 عاملاً من عمال المؤسسة العامة للتبغ. المطلوب: احسب المتوسطات الثلاثة :

جدول رقم (4-18)

التكرار التجميعي الصاعد	$x \cdot f$	x	عدد العمال (f)	فئات الأجور
20	1900	95	20	80 - 110
50	3750	125	30	110 - 140
90	6200	155	40	140 - 170
120	5550	185	30	170 - 200
140	4300	215	20	200 - 230
	21700		140	Σ

آ الوسط الحسابي :

$$\bar{X} = \frac{21700}{140} = 155 \text{ ليرة}$$

ب) الوسط : $\frac{140}{2} = 70$ ترتيب الوسيط

$$\text{Med} = 140 + \frac{70 - 50}{40} \cdot 30$$

$$= 140 + 15 = 155 \text{ ليرة}$$

ج) المنوال :

$$\text{Med} = 140 + \frac{10}{10 + 10} \cdot 30$$

$$= 140 + 15 = 155 \text{ ليرة}$$

نلاحظ بأن المتوسطات الثلاثة منطبقة على بعضها البعض، والسبب أن هذا التوزيع التكراري متماثل تماماً، فتكرارات الفئات المتوسطة متدرجة، والآن لنضيف إلى هذا التوزيع فئتين جديدتين على نهايته كما يأتي :

جدول رقم (4-19)

التكرار التجميعي الصاعد	$x \cdot f$	x	عدد العمال (f)	فئات الأجور
20	1900	95	20	80 - 110
50	3750	125	30	110 - 140
90	6200	155	40	140 - 170
120	5550	185	30	170 - 200
140	4300	215	20	200 - 230
156	3920	245	16	230 - 260
163	1925	275	7	260 - 290
	27545		163	Σ

فالموسط الحسابي بعد إضافة فئات إلى نهاية التوزيع يساوي :

$$\bar{x} = \frac{27545}{163} = \boxed{168,99}$$

أما الوسيط فإنه يساوي :

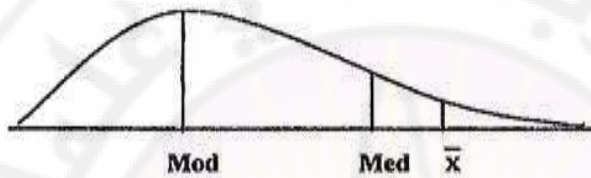
$$\begin{aligned} \text{Med} &= 140 + \frac{81,5 - 50}{40} \cdot 30 \\ &= 140 + 23,625 = \\ &= \boxed{163,625} \end{aligned}$$

والمتوال يساوي :

$$\begin{aligned} \text{Mod} &= 140 + \frac{10}{10 + 10} \cdot 30 \\ &= 140 + 15 \\ &= \boxed{155} \end{aligned}$$

نشاهد من النتائج الجديدة (أي بعد إضافة فئات إلى نهاية التوزيع) أن الوسط الحسابي والوسيط قد انحرفا نحو الفئات الجديدة (نحو اليمين)، علماً بأن قيمة الوسط الحسابي قد انحرفت بشكل أكبر من قيمة الوسيط، أما المنوال فلم تتغير قيمته مطلقاً لأن المنوال لا يتأثر بالقيم الشاذة، في حين أن الوسط الحسابي يتأثر بالقيم وبشكل كبير جداً، وبالتالي يصبح ترتيب قيم المتوسطات الثلاثة لتوزيع تكراري مائلاً نحو اليمين بالشكل الآتي :

$$\text{الوسط الحسابي} < \text{الوسيط} < \text{المنوال}$$



شكل (4-7)

والجدير بالذكر أن المنحنى التكراري المائل نحو اليمين (موجب الالتواء) (Positive Skewness) هو التوزيع التكراري الذي تتزايد تكرارات فئاته الأولى بسرعة حتى تصل إلى القمة لتبدأ بعدها بالتناقص ببطء. ولنضيف الآن إلى التوزيع المتماثل تماماً مثال (4-24) جدول رقم (4-18) بعض الفئات إلى بدايته كما يأتي :

جدول رقم (4-20)

فئات الأجور	عدد العمال (f)	x	x . f	التكرار التجميعي الصاعد
20 - 50	8	35	280	8
50 - 80	10	65	650	18
80 - 110	20	95	1900	38
110 - 140	30	125	3750	68
140 - 170	40	155	6200	108
170 - 200	30	185	5550	138
200 - 230	20	215	4300	158
Σ	158		22630	

فالوسط الحسابي بعد إضافة فئات جديدة إلى بداية التوزيع تساوي:

$$\bar{x} = \frac{22630}{158} = \boxed{143,228}$$

والوسيط يساوي :

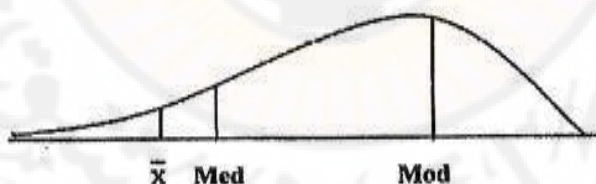
$$\begin{aligned} \text{Med} &= 140 + \frac{79 - 68}{40} \cdot 30 \\ &= 140 + 8,25 \\ &= \boxed{148,25} \end{aligned}$$

أما المنوال فيساوي :

$$\begin{aligned} \text{Mod} &= 140 + \frac{10}{10 + 10} \cdot 30 = \\ &= \boxed{155} \end{aligned}$$

نلاحظ من النتائج الجديدة (أي بعد إضافة فئات إلى بداية التوزيع) أن الوسط الحسابي والوسيط قد انحرفا إلى بداية التوزيع، حيث يكون الوسط الحسابي أصغر من الوسيط، والوسيط أصغر من المنوال الذي لم تتغير قيمته مطلقاً، وبالتالي يصبح ترتيب قيم المتوسطات الثلاثة لتوزيع تكراري مائلاً نحو اليسار على الشكل الآتي :

المنوال < الوسيط < الوسط الحسابي



شكل (4-8)

والجدير بالذكر أيضاً أن المنحنى التكراري المائل نحو اليسار (سالب الالتواء) (Negative Skew ness) هو التوزيع التكراري الذي تتزايد تكرارات الفئات الأولى ببطء حتى تصل إلى القمة لتتحدّر بعدها تكرارات الفئات الأخرى بسرعة.

من الأمثلة السابقة نلاحظ :

1- الوسيط يقع بصورة تقريبية في ثلث المسافة بين الوسط الحسابي والمنوال في التوزيعات القريبة من التماثل، ويكون بشكل عام أقرب إلى الوسط الحسابي منه إلى المنوال.

2- في التوزيعات المتماثلة تتساوى قيم المتوسطات الثلاثة.

3- عندما يكون التوزيع قريباً من التماثل ومائلاً نحو اليمين يكون $\bar{x} < Med < Mod$.

4- عندما يكون التوزيع قريباً من التماثل ومائلاً نحو اليسار $\bar{x} < Med < Mod$.

وقد وجد بالتجربة أنه في التوزيعات التكرارية القريبة من التماثل تكون العلاقة التقريبية بين المتوسطات الثلاثة وهي :

$$(\bar{x} - Mod) = 3 (\bar{x} - Med)$$

إن هذه العلاقة تعطينا قيمة أحد المقاييس إذا عرفنا المقياسين الآخرين، وذلك بشكل تقريبي فقط في التوزيعات القريبة من التماثل، كما تساعدنا على تقدير قيمة الوسط الحسابي في حال التوزيعات التكرارية المفتوحة باستعمال العلاقة السابقة .

4-5- الوسط الهندسي The Geometric Mean :

الوسط الهندسي لمجموعة من القيم هو حاصل ضرب هذه القيم بعضها ببعض مجزئاً إلى عددها، فإذا كان لدينا n من القيم $x_1, x_2, \dots, x_n$ فالوسط الهندسي لهذه القيم يساوي :

$$G = \sqrt[n]{(x_1)(x_2) \dots (x_n)}$$

ويمكن حل هذه المعادلة للحصول على قيمة الوسط الهندسي (G) بطريقتين :

أ) باستخدام الآلة الحاسبة بصورة مباشرة .

ب) باستخدام اللوغاريتم، ويتم ذلك برفع الطرفين إلى القوة n ، ثم نأخذ

لوغاريتم الطرفين :

$$G^n = (x_1)(x_2) \dots (x_n)$$

$$n \log G = \text{Log} [(x_1)(x_2) \dots (x_n)]$$

$$\text{Log } G = \frac{\text{Log } x_1 + \text{Log } x_2 + \dots + \text{Log } x_n}{n}$$

$$\text{Log } G = \frac{\sum \text{Log } x_i}{n}$$

ويجب الانتباه إلى أن المعادلة الأخيرة تعطي لوغاريتم الوسط الهندسي وبأخذ
مقابل لوغاريتم (Anti Logarithm) نحصل على قيمة الوسط الهندسي .

1-5-4- حساب الوسط الهندسي لبيانات غير مبوبة :

مثال (4-25) :

لدينا القيم التالية التي تمثل الأجور الساعية لمجموعة من العمال :

4, 6, 10, 15

لنحسب الوسط الهندسي :

$$G = \sqrt[4]{(4)(6)(10)(15)} = 7,746 \text{ ليرة}$$

$$\text{Log } G = \frac{\text{Log } 4 + \text{Log } 6 + \text{Log } 10 + \text{Log } 15}{4} = 7,746 \text{ ليرة}$$

لنتساءل عن معنى الوسط الهندسي وما هو تفسيره .

قبل الإجابة عن هذا التساؤل نحسب الوسط الحسابي للأجور الساعية للعمال في

المثال السابق، ثم نقارن ذلك الوسط الحسابي للأجور مع الوسط الهندسي للأجور نفسها:

$$\bar{x} = \frac{4 + 6 + 10 + 15}{4} = \frac{35}{4} = 8,75 \text{ ليرة}$$

يفسر الوسط الحسابي: لو تساوت الأجور لثلاثة عمال وسأوى كل منها 8,75 ليرة لكان مجموع الأجور الساعية للعمال الأربع يساوي 35 ليرة ما يدفع لثلاثة عمال، وهذا منطقي بحيث إننا نستخدم هذا المجموع في الحياة العملية .

ولنفسر الآن حسب المثال السابق (25-4) معنى قيمة الوسط الهندسي (7,746): لو تساوت الأجور الساعية لثلاثة عمال وكان كل منها يساوي (7,746) ليرة لكان جداءها يساوي جداء الأجور الساعية للعمال الأربعة (رياضياً صحيح) لكن من الناحية العملية ليس له أي معنى (جداء الأجور) .

ولنتساءل مرة أخرى متى يجب استخدام الوسط الهندسي ؟
للإجابة على هذا التساؤل نورد المثال الآتي :

مثال (26-4) :

خزان ماء له شكل متوازي مستطيلات أبعاده 1، 2، 4 متر، بمعنى أنه يسع ما مقداره (1) (2) (4) = 8 متر مكعب من الماء، نريد استبدال متوازي المستطيلات بمكعب بحيث نحصل على السعة نفسها 8 أمتار مكعبة من الماء أيضاً، فإذا استخدمنا الوسط الحسابي لإيجاد متوسط طول ضلع المكعب بحيث :

$$\bar{x} = \frac{4 + 2 + 1}{3} = 2,333 \text{ متر}$$

فإن سعة الخزان الجديد (المكعب) يساوي : (2,333) (2,333) (2,333) = 12,7 أمتار مكعب، وهذا غير صحيح لأن سعته المطلوبة هي 8 أمتار مكعبة .
أما إذا استخدمنا الوسط الهندسي لإيجاد متوسط طول ضلع المكعب بحيث :

$$G = \sqrt[3]{(4)(2)(1)} = 2$$

لكان سعة الخزان الجديد (المكعب) يساوي :

$$(2) (2) (2) = 8 \text{ متر مكعب وهي السعة المطلوبة .}$$

نستنتج أن المقياس الصحيح في هذه الحالة هو الوسط الهندسي، وبالتالي فهناك حالات في الحياة العملية يجب عندها استخدام الوسط الهندسي لإيجاد القيمة التمثيلية للبيانات، ويستخدم الوسط الهندسي في تلك الحالات التي يكون للجداء فيها معنى مثل الحجم والمساحة كما المثال السابق، ويستخدم أيضاً في الأرقام القياسية للأسعار والكميات أو إيجاد متوسط عدد من النسب أو إيجاد معدلات التغير في المبيعات أو السكان، فعند وجود سلاسل زمنية تعبر عن تطور ظاهرة ما خلال مدة من الزمن، عندها يستعاض في بعض الأحيان عن القيم الحقيقية المطلقة بالنسب المئوية والتي يكون أساسها متحركاً (نسبة كل قيمة إلى سابقتها) عندها نوجد الوسط الهندسي لهذه النسب (كمعدل للنمو) .

مثال (4-27) :

إذا أردنا معرفة متوسط نسبة تغير سرعة سيارة ما (متوسط معدل التغير) إذا علمنا أنه في الساعة الأولى كانت سرعتها تساوي 200 كم، في الساعة الثانية تساوي 100 كم، وفي الساعة الثالثة سرعتها ساوت 200 كم، نجد أن السرعة 200 تهبط إلى السرعة 100 ومن ثم ترتفع إلى 200 من جديد، فإن نسب التغير في الساعة الثانية بالنسبة للأولى $0,5 = \frac{100}{200}$ وإن نسبة التغير في الساعة الثالثة بالنسبة للساعة الثانية يساوي $2 = \frac{200}{100}$ وبالتالي فالوسط الهندسي لتغير السرعة هو :

$$G = \sqrt{(0,5)(2)} = 1$$

أي إن متوسط نسبة تغير السرعة (معدل التغير) خلال الزمن المدروس لهذه السيارة هو $1 - 1 = 0$ لكن عند استخدام الوسط الحسابي لنسب تغير السرعة نجد :

$$\bar{x} = \frac{0,5 + 2}{2} = 1,25$$

بمعنى أن معدل تغير السرعة خلال الزمن المدروس لهذه السيارة هو $1,25 - 1 = 0,25$ (أي 25%) وهي نتيجة غير منطقية والسبب الرئيسي في هذا الاختلاف بين

الوسطين يعود إلى تأثير الوسط الحسابي بتباين حجم القيم، ولتوضيح ذلك نعود إلى المثال السابق، وبفرض أن السرعة في الساعة الأولى تساوي 200 كم، وفي الساعة الثانية 195 كم، ومن ثم عادت في الساعة الثالثة إلى 200 فإن نسب تغير السرعة على التوالي:

$$1,0256 = \frac{200}{195} \text{ و } 0,975 = \frac{195}{200}$$

فالوسط الهندسي لتغير السرعة خلال تلك الفترة هو :

$$G = \sqrt{(0,975)(1,0256)} = 1$$

وبدلت التغير يساوي $1 - 1 = 0$.

أما الوسط الحسابي لتغير السرعة خلال تلك الفترة هو :

$$\bar{x} = \frac{0,975 + 1,0256}{2} = 1,003$$

وبالتالي فمعدل التغير يساوي تقريباً الصفر $1,0003 - 1 \approx 0$.

نستنتج من ذلك أن الوسطين الحسابي والهندسي متساويان تقريباً، وذلك كلما اقتربت أحجام القيم من بعضها البعض.

(سوف نتطرق لمعدلات النمو للأسعار والكميات وبشكل مفصل عند بحث الأرقام القياسية).

مثال (4-28) :

ليكن لدينا الجدول الآتي والمتضمن على الأرقام القياسية لأسعار 8 سلع استهلاكية لعام 2002 بالنسبة لعام 2000 :

السلعة	1	2	3	4	5	6	7	8
الرقم القياسي للأسعار في عام 2002 بالنسبة لعام 2000	101%	90%	108%	96%	103%	79%	85%	100%

بمعنى أن السلعة الأولى زادت أسعارها في عام 2002 بالمقارنة مع عام 2000 بنسبة $1\% = 100 - 101$. والسلعة الثانية قد انخفضت أسعارها في عام 2002 بالمقارنة مع عام 2000 بنسبة $10\% = 100 - 90$... وهكذا .
أوجد الوسط الهندسي لمعدل نمو الأسعار .

الحل :

$$\text{Log } G = \frac{\text{Log } 101 + \text{Log } 90 + \dots + \text{Log } 100}{8}$$

$$\text{Log } G = \frac{2,0043 + 1,9556 + \dots + 2,000}{8} = \frac{15,8554}{8}$$

$$\text{Log } G = 1,9819$$

وباستخراج القيمة العشرية لـ $\log G$ نجد أن الوسط الهندسي يساوي $G = 0,96$.
وبمعنى هذا الرقم 0,96 أن معدل نمو أسعار السلعة في عام 2002 بالنسبة لعام 2000 هو سلبى $0,04 = 1 - 0,96$ أي انخفضت الأسعار في عام 2002 عن عام 2000 بنسبة 4% .

2-5-4- حساب الوسط الهندسي للبيانات المبوبة في فئات (Grouped data):

إذا كانت $x_1, x_2, \dots, x_n$ هي مراكز الفئات في الجدول التكراري، وكانت التكرارات المقابلة لمراكز الفئات على التوالي $f_1, f_2, \dots, f_n$ فالوسط الهندسي يساوي:

$$G = \sqrt[n]{(x_1)^{f_1} (x_2)^{f_2} \dots (x_n)^{f_n}}$$

$$G = \sqrt[n]{\prod (x_i)^{f_i}}$$

حيث Π ترمز لعملية ضرب القيم. وباستخدام اللوغاريتمات نجد أن :

$$\log G = \frac{\sum_{i=1}^n \log x_i f_i}{\sum f_i}$$

للحصول على الوسط الهندسي في الجداول التكرارية نتبع المراحل الآتية :

- 1- نوجد لوغاريتم أوساط الفئات .
- 2- نضرب لوغاريتم أوساط الفئات بالتكرارات المقابلة .
- 3- نجمع النتائج ويتم تقسيمه على إجمالي التكرارات، فنحصل على لوغاريتم الوسط الهندسي .
- 4- نوجد العدد المقابل للوغاريتم فنحصل على الوسط الهندسي .

مثال (4-29) :

لدينا البيانات الآتية الممثلة لعلامات 150 طالباً في مادة الإحصاء :

جدول رقم (4-21)

فئات العلامات	f	x_i	$\log x_i$	$\log x_i f_i$
10 - 20	7	15	1.17	8.23
20 - 30	8	25	1.39	11.1
30 - 40	10	35	1.54	15.44
40 - 50	25	45	1.65	41.33
50 - 60	50	55	1.74	87.82
60 - 70	25	65	1.81	45.622
70 - 80	10	75	1.87	18.75
80 - 90	8	85	1.92	15.435
90 - 100	7	95	1.97	13.844
Σ	150			256.55

$$\text{Log } G = \frac{\sum \text{Log } x_i f_i}{\sum f_i} = \frac{256.55}{150} = 1.71$$

وبإيجاد مقابل لوغاريتم نجد أن $G = 51.33$.

لاحظ أن التوزيع طبيعي وسطه الحسابي 55 درجة، وكذلك وسيطه ومنواله ولكن الوسط الهندسي لا يساويهما. لماذا ؟

3-5-4- محاسن ومساوئ الوسط الهندسي :

(أ) محاسنه :

1- يعتبر الوسط الهندسي مقياساً محدداً رياضياً كالوسط الحسابي، فهو يتمتع ببعض الميزات الجبرية الهامة (لوغاريتم الوسط الهندسي = الوسط الحسابي للوغاريتميات القيم) :

$$\text{Log } G = \frac{\sum f_i \text{Log } x}{\sum f} \Rightarrow$$

$$\sum f_i \text{Log } G = \sum f_i \text{Log } x$$

2- يأخذ هذا المقياس جميع القيم بعين الاعتبار، بالإضافة إلى أنه قليل التأثير بالقيم المتطرفة نحو اليمين .

3- إن قيمة الوسط الهندسي أكثر تمثيلاً من باقي المتوسطات عند التعامل مع توزيعات تكرارية شديدة الالتواء نحو اليمين .

4- لا يتأثر الوسط الهندسي بأطوال الفئات مثله مثل الوسط الحسابي .

(ب) مساوئه :

1- لا يمكن حسابه إلا إذا كان مجموع القيم موجبة .

2 - محدودية تطبيقاته فأكثر استعمالاته تنحصر في حساب معدلات النمو والأرقام القياسية ومعدلات الفائدة.

3- التعقيدات في حسابه .

4- لا يمكن حسابه إذا كان الجدول التكراري له فئات مفتوحة .

4-6- الوسيط التوافقي Harmonic Mean :

عندما نريد إيجاد مقياس مناسب من مقاييس النزعة المركزية الذي يعبر عن إيجاد المتوسط وفقاً لوحدة قياسية معينة، فإن الوسيط التوافقي يعتبر عندها أفضل مقياس. فالوسيط التوافقي لمجموعة من القيم هو مقلوب الوسيط الحسابي لمقلوبات تلك القيم، بمعنى آخر هو مقلوب (reciprocals) الوسيط الحسابي .

4-6-1- حساب الوسيط التوافقي لبيانات غير مبهمة :

ليكن لدينا القيم الآتية: x_1, x_2, x_3, x_n فالوسيط الحسابي لتلك القيم يساوي :

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

وبإيجاد الوسيط الحسابي لمقلوب القيم السابقة :

$$\frac{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}{n}$$

ثم نوجد مقلوب هذا الوسيط الحسابي :

$$\frac{1}{\frac{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}{n}}$$

والقيمة الأخيرة تعتبر هي قيمة الوسيط التوافقي، إذاً :

$$H = \frac{1}{\frac{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}{n}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

$$H = \frac{n}{\sum \frac{1}{x_i}}$$

حيث H هو الوسط التوافقي .

والمثال الآتي يوضح فكرة الوسط التوافقي بشكل عملي :

مثال (30-4) :

نرغب بدراسة متوسط إنتاجية بعض الشركات الخاصة بتنظيف المنازل، إذا علمت أن الشركة A تنظف بيتاً من مساحة 200 متر خلال $\frac{1}{3}$ ساعة، وإن الشركة B تنظف بيتاً من المساحة السابقة نفسها خلال $\frac{1}{2}$ ساعة، والشركة C تنظفه بـ $\frac{1}{4}$ ساعة، والشركة D تنظفه بـ $\frac{1}{5}$ ساعة، والشركة E تنظفه بـ $\frac{1}{6}$ ساعة .

كم بيتاً بالمتوسط تستطيع هذه الشركات تنظيفها خلال ساعة واحدة ؟
أو كم من الوقت تستغرق هذه الشركات في المتوسط لتنظف بيتاً واحداً ؟

الحل :

خلال ساعة واحدة تنظف الشركة A ثلاثة بيوت، والشركة B تنظف بيتين
والشركة C تنظف أربعة بيوت، والشركة D تنظف خمسة بيوت، والشركة E تنظف
سنة بيوت في الساعة الواحدة، وعندها يكون متوسط ما تقوم به هذه الشركات في
الساعة الواحدة :

$$\frac{3 + 2 + 4 + 5 + 6}{5} = 4 \text{ بيوت في الساعة}$$

بمعنى أن الشركات الخمس للتنظيفات تستطيع بالمتوسط تنظيف أربعة بيوت
(ذات مساحة 200 متر) في الساعة الواحدة . بمعنى ثانٍ وسطياً تنظف الشركة الواحدة
للبيت الواحد هو $\frac{1}{4}$ ساعة (0,25) نجد أننا استعملنا مقلوب الوسط الحسابي $\frac{1}{4}$ لإيجاد
متوسط الوقت لإنتاجية هذه الشركات في تنظيف البيت الواحد .

نشاهد هنا أنه تم حساب المتوسط بطريقتين :

- الطريقة الأولى: وهي حساب متوسط إنتاجية الشركات الخمس في الساعة
الواحدة .

- الطريقة الثالثة: وهي حساب متوسط الوقت لإنتاجية الشركات الخمس للبيت الواحد.

فإذا استعملنا الوسط الحسابي البسيط لإيجاد متوسط الوقت لهذه الشركات في تنظيف بيت واحد لكان :

$$\frac{\frac{1}{3} + \frac{1}{2} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}}{5} = \frac{1,45}{5} = 0,29 \text{ ساعة}$$

وهذا لا يساوي 0,25 ساعة .

∴ يجب استعمال طريقة صحيحة في حساب المتوسط وفقاً لوحدة قياسية معينة.

إن طريقة إيجاد ذلك المتوسط (الذي يدعى بالوسط التوافقي) تمر بالمراحل الآتية:

نوجد المقلوب للزمن اللازم لتنظيف البيت الواحد لكل شركة، ومن ثم نحسب الوسط الحسابي لمقلوب الزمن اللازم لتنظيف البيت الواحد، وأخيراً نوجد مقلوب الوسط الحسابي لمقلوب الزمن اللازم لتنظيف البيت الواحد أي :

$$\frac{\frac{1}{\frac{1}{3}} + \frac{1}{\frac{1}{2}} + \frac{1}{\frac{1}{4}} + \frac{1}{\frac{1}{5}} + \frac{1}{\frac{1}{6}}}{5}$$

ثم :

$$\frac{1}{\frac{\frac{1}{3} + \frac{1}{2} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}}{5}} = \frac{5}{\frac{1}{3} + \frac{1}{2} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}} = \frac{5}{20}$$

ساعة $\frac{1}{4}$

بمعنى تنظيف البيت الواحد تحتاج الشركة بالمتوسط $\frac{1}{4}$ ساعة، وبشكل آخر

تستطيع الشركة أن تنظف أربعة بيوت في الساعة الواحدة .

مثال (4-31) :

يستهلك شخص في الصيف 50 كغ من اللحم حيث إن سعر الكيلو غرام الواحد يساوي 300 ليرة، أما استهلاكه في الشتاء فهو 60 كغ من اللحم حيث كان سعر الكيلو غرام الواحد يساوي 250 ليرة. احسب متوسط سعر الكيلو غرام الواحد بطريقتين بطريقة الوسط الحسابي وبطريقة الوسط التوافقي .

الحل :

بما أن الكمية المستهلكة من اللحم هي أوزان لترجيح الأسعار فالوسط الحسابي الموزون للأسعار يساوي :

$$\begin{aligned}\bar{x} &= \frac{x_1 \cdot w_1 + x_2 \cdot w_2}{w_1 + w_2} = \frac{(300)(50) + (250)(60)}{50 + 60} \\ &= \frac{15000 + 15000}{110} = \frac{30000}{110} = 272,73 \text{ ليرة}\end{aligned}$$

متوسط سعر الكغ الواحد من اللحم بالنسبة لمقدار استهلاك هذا الشخص في الشتاء والصيف .

أما إذا أردنا استعمال صيغة الوسط التوافقي لحساب متوسط الأسعار للكغ الواحد من اللحم، فنوجد أولاً الوسط الحسابي البسيط لمقلوب أسعار الكغ من اللحم، ومن ثم نحسب مقلوب الوسط الحسابي البسيط لمقلوب الأسعار كما يأتي :

$$\begin{aligned}&\frac{\frac{1}{300} + \frac{1}{250}}{2} \\ H &= \frac{1}{\frac{\frac{1}{300} + \frac{1}{250}}{2}} = \frac{2}{\frac{1}{300} + \frac{1}{250}} = 272,73 \text{ ليرة}\end{aligned}$$

نجد من هذا المثال ما يأتي :

1- تتساوى قيمة كل من الوسط الحسابي والوسط التوافقي لوسطي أسعار الكغ الواحد من اللحم $\bar{x}$ $H =$.

2- تم استعمال الوسط الحسابي للأسعار الموزونة بالكميات المستهلكة، بينما الوسط التوافقي البسيط استخدم فقط الأسعار، ولم يرجح (يوزن) بالكميات المستهلكة، وذلك لتساوي قيم الاستهلاك لهذا الشخص، فالقيمة والتي تساوي (الكمية $\times$ السعر) كانت متساوية في كل من الصيف والشتاء لاستهلاك هذا الشخص $15000 = 250 \times 60$ و $15000 = 300 \times 50$ وبذلك لم تتغير قيمة الوسط الحسابي الموزون عن قيمة الوسط التوافقي البسيط.

مثال (4-32) :

بالعودة إلى المثال رقم (4-31) وبفرض أن هذا الشخص يستهلك 65 كغ من اللحم شتاءً وبالسعر السابق نفسه للكغ الواحد، فإن القيمة والتي تساوي الكمية المستهلكة $\times$ السعر تساوي $16250 = 250 \times 65$ ل.س لاستهلاكه شتاءً عندها سوف نجد أنه يطرأ اختلاف على كل من الوسطين الحسابي والتوافقي، ولكنهما يكونان متساويين أيضاً :

$$\bar{x} = \frac{(300)(50) + (250)(65)}{50 + 65} = \frac{1500 + 16250}{115} = \frac{31250}{115} = 271,74$$

أما الوسط التوافقي فلا يحسب في هذه الحالة كما تم حسابه سابقاً، بسبب عدم تساوي قيم الاستهلاك في كل من الصيف والشتاء عندها يجب استعمال صيغة ثانية للوسط التوافقي في حال وجود قيم غير متساوية، وهذا ما ينطبق على التعديل الذي تم إجرأه للمثال السابق، والجدير بالذكر أن الوسط التوافقي الموزون قد يوزن (يرجح) بالأسعار أو الكميات أو القيم .

وبالتالي فالوسط التوافقي الموزون (بالقيم الجديدة) سوف يساوي :

$$H = \frac{1}{\frac{1}{300}(15000) + \frac{1}{250}(16250)} = \frac{31250}{114,99} = 271,74$$

نلاحظ أن الوسط التوافقي البسيط والذي يساوي 272,73 يختلف عن الوسط التوافقي الموزون بالقيم المختلفة 271,74 والسبب في ذلك هو اختلاف تلك القيم في الشتاء عنه في الصيف، أما إذا تساوت القيم فإن الوسطين التوافقي البسيط والتوافقي الموزون سوف يتساويان تماماً، ونلاحظ أيضاً أن الوسط الحسابي الموزون بالكميات يساوي الوسط التوافقي الموزون بالقيم، وإن هذين الوسطين (الحسابي الموزون والتوافقي الموزون) لن يتساويا في حال تم ترجيح الوسطين بالكميات، أو تم ترجيح الوسطين بالقيم ودائماً يكون الوسط الحسابي أكبر من الوسط التوافقي في الحالة الأخيرة. فمثلاً لو تم ترجيح الوسط التوافقي بالكميات نجد أنه يساوي :

$$H = \frac{1}{\frac{1}{300} \cdot (50) + \frac{1}{250} (65)} = 269,53 \text{ ليرة}$$

بمعنى أن قيمة الوسط التوافقي الموزون بالكميات (269,53) هو أصغر من الوسط الحسابي الموزون بالكميات أيضاً (271,74) (وسوف نشاهد بعض التطبيقات الهامة للوسط التوافقي في بحث الأرقام القياسية).

2-6-4- الوسط التوافقي لبيانات مبوبة في فئات :

عند التعامل مع بيانات مبوبة في فئات سنفترض بأن المعطيات سوف تقع في مراكز الفئات كما تعلمنا سابقاً $x_1, x_2, \dots$ وإن التكرارات المقابلة على التوالي هي $f_1, f_2, \dots$ فإن صيغة الوسط التوافقي المرجح يساوي :

$$H = \frac{\sum f}{\sum \frac{f}{x}}$$

حيث x_i = هي مراكز الفئات للجدول التكراري .

مثال (4-33) :

ليكن لدينا الجدول الآتي المتضمن فئات الأسعار التي باع بها أحد مندوبي المبيعات عدة سلع حيث الكميات المباعة من هذه السلع بحسب الجدول (4-22):

جدول رقم (4-22)

الكميات المباعة	فئات الأسعار (وحدة نقدية)
10	2 - 4
25	4 - 6
30	6 - 8
20	8 - 10
15	10 - 12
100	Σ

احسب الوسط الحسابي المرجح بالكميات للأسعار التي باع فيها هذا المندوب.

ثم احسب الوسط التوافقي المرجح بالكميات للأسعار التي باع فيها هذا المندوب.

واحسب الوسط التوافقي المرجح بالقيم (السعر $\times$ الكمية) ماذا تستنتج ؟

الحل :

x	f	$x f$	$\frac{f}{x}$
3	10	30	3,33
5	25	125	5
7	30	210	4,28
9	20	180	2,22
11	15	165	1,36
Σ	100	710	16,19

$$I) \bar{x} = \frac{\sum x f}{\sum f} = \frac{710}{100} = 7,10 \quad \text{وحدة نقدية}$$

$$II) H = \frac{\sum f}{\sum \frac{f}{x}} = \frac{100}{16,19} = 6,18 \quad \text{وحدة نقدية}$$

$$III) H = \frac{\sum x f}{\sum \frac{x f}{x}} = \frac{\sum x f}{\sum f} = \frac{710}{100} = 7,10 \quad \text{وحدة نقدية}$$

حيث إن القيمة $(f) (x)$.

نستنتج أنه إذا تم الترجيح بالكميات لكل من الوسط الحسابي والوسط التوافقي، فإن الوسط الحسابي يكون أكبر من الوسط التوافقي، ونلاحظ أيضاً أنه إذا تم ترجيح الوسط الحسابي بالكميات والوسط التوافقي بالقيم فإن الوسطين سوف يتساويان.

3-6-4 محاسن ومساوئ الوسط التوافقي :

محاسنه :

- 1- يأخذ جميع القيم بعين الاعتبار عند حسابه .
- 2- جيد الاستخدام ويعطي نتائج دقيقة عند إيجاد متوسط التغير عبر الزمن مثل إيجاد إنتاجية العمل كنسبة الإنتاج للزمن الذي أنتج به الإنتاج، الأسعار التي تظهر النسبة بين قيمة البضاعة من نوع ما وكميتها، وقيمة البضاعة من نوع آخر وكميتها، ويمكن عن طريقة إيجاد وسطى الأرقام القياسية التجميعية لمناسيب الأسعار (أو الكميات) المرجحة بالقيم أيضاً في سنة الأساس أو القيم في سنة المقارنة .

مساوئه :

- 1- يتأثر الوسط التوافقي بالقيم المتطرفة في الصفر حيث ليس له مدلول مع وجود قيم مساوية للصفر.
- 2- لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة .

وأخيراً يمكن أن يطرح السؤال الآتي :

أي مقياس من مقاييس النزعة المركزية هو أفضل المقاييس؟

وللجواب على هذا السؤال هو أن تتوفر في مقياس النزعة المركزية الصفات الآتية لكي يكون مقياساً جيداً :

1- أن يعتمد المقياس في حسابه على كل القيم.

2- أن يكون المقياس سهل الفهم والحساب.

3- أن تتوفر في المقياس الصفات الرياضية والجبرية .

4- أن لا يتأثر المقياس بوجود قيم متطرفة وشاذة .

إن هذه الصفات المشار إليها أعلاه لا تتوفر كلها في مقياس واحد، ولكن كل مقياس من مقاييس النزعة المركزية يفضل استخدامه في حالات معينة، ولا يفضل استخدامه في حالات أخرى، فعلى الباحث أن يختار المقياس الملائم والمناسب لمعطياته الإحصائية .

وبشكل عام إذا أردنا إعطاء صورة واضحة عن ظاهرة ما، فإنه من الناحية العملية يجب علينا حساب عدة متوسطات، لأنه - كما ذكرنا سابقاً - لكل متوسط غاية وهدف في وصف التوزيع المدروس .

وإن وصف التوزيع عن طريق المتوسطات لا يكون كافياً، فيجب إضافة إلى حساب المتوسطات حساب مقاييس التشتت والالتواء .

الفصل الخامس

مقاييس التشتت (التباين)

Measures of Dispersion (Variation)

- مقدمة .
- المدى المطلق .
- المدى النسبي .
- الانحراف الربيعي (نصف المدى الربيعي) .
- الانحراف الربيعي النسبي .
- محاسن ومساوى الانحراف الربيعي .
- الانحراف المتوسط .
- حساب الانحراف المتوسط للبيانات المبوبة في فئات .
- حساب الانحراف المتوسط للبيانات المبوبة في فئات .
- الانحراف المتوسط النسبي .
- محاسن ومساوى الانحراف المتوسط .
- الانحراف المعياري .
- حساب الانحراف المعياري للبيانات غير المبوبة .
- حساب الانحراف المعياري للبيانات المبوبة في فئات .
- الانحراف المعياري النسبي (معامل الاختلاف) .
- الدرجة المعيارية .
- أهم خصائص الانحراف المعياري .
- محاسن ومساوى الانحراف المعياري .



1-5- مقدمة :

إن مفردات أي ظاهرة تحاول أن تتمركز حول قيمة معينة من قيم الظاهرة المدروسة، وقد تكلمنا سابقاً عن أحد خصائص التوزيعات التكرارية، وهي النزعة المركزية لكن قد تتماثل التوزيعات التكرارية (البينانية) فيما بينها من حيث النزعة المركزية، بمعنى آخر قد نجد توزيعين أو أكثر يحملان الوسط الحسابي نفسه أو المنوال أو الوسيط، وهذا ليس مؤشراً على أن هناك تماثلاً في توزيع المفردات لتلك التوزيعات، فمقاييس النزعة المركزية تمدنا بمؤشرات إحصائية ذات دلالة وصفية لتوضيح الشكل العام للتوزيع البياني دون الإشارة عن ماهية التوزيع عن ما يجري داخل هذا التوزيع، عن التباعد أو التناثر أو الاقتراب لمفردات التوزيع بعضها مع بعض، يمكن أن تتساوى قيم الوسط الحسابي والوسيط والمنوال ويمكن أن تقترب من بعضها إلى البعض إلى حد كبير ويعني ذلك أن التوزيع البياني في الحالة الأولى متماثل، وفي الحالة الثانية قريباً من التماثل، أما إن اختلفت تلك المقاييس الثلاثة اختلافاً كبيراً دل ذلك على عدم تماثل التوزيع، ولكن السؤال المطروح : إلى أي مدى تتناثر القيم داخل التوزيع حول المتوسط الحسابي؟ وإلى أي مدى تختلف القيم داخل التوزيع فيما بينها؟ وهل تكفي مقاييس النزعة المركزية للمقارنة بين توزيع وآخر؟

للإجابة على هذه التساؤلات ينبغي الاستعانة بمقاييس مكتملة لمقاييس النزعة المركزية، وهي مقاييس التشتت التي تقيس مدى التباعد أو التقارب الموجود بين قيم المشاهدات التابعة لمتغير ما، وتقيس مدى تشتت قيم المشاهدات عن وسطها .

مثال (5-1) :

يفترض أننا اخترنا وبشكل عشوائي مجموعتين من الطلبة الذكور والإناث قد امتحنوا في المادة نفسها فحصلت مجموعة الذكور - وهم عشرة طلاب - على العلامات الآتية :

65, 66, 67, 68, 71, 73, 74, 77, 77, 77

كما حصلت مجموعة الإناث وهن عشر طالبات على العلامات الآتية :

41, 55, 57, 63, 67, 77, 77, 84, 94, 100

وبحساب كل من المتوسطات الثلاثة لكل مجموعة نجد :

$\bar{x} = 71,5$ للذكور	$Med = 72$ للذكور	$Mod = 77$ للذكور
$\bar{x} = 71,5$ للإناث	$Med = 72$ للإناث	$Mod = 77$ للإناث

وبالنظر إلى علامات كل من الطلاب والطالبات، نلاحظ للوهلة الأولى أن علاماتهم متماثلة ولكن لو تفحصنا علامات كل طالب وكل طالبة لوجدنا بأن هناك فرقاً بين الطلاب والطالبات، وأن علامات الطالبات تنتشر وتباعد عن بعضها البعض بشكل أكبر ويمدى أوسع من علامات الطلاب، وبالتالي فالاختلاف بين المشاهدات صفة لا يمكن أن تظهرها مقاييس النزعة المركزية ، بل يظهرها ما دعوناها سابقاً بمقاييس التشتت. ومقاييس التشتت تفيدنا في :

1- تزودنا بمعلومات حول تبعثر (تشتت)، أو تجمع البيانات داخل التوزيع، وحول الوسط الحسابي لهذا التوزيع.

2- تقيّم تلك المقاييس مدى فعالية مقاييس النزعة المركزية (الصفة التمثيلية لمقاييس النزعة المركزية)، فمثلاً كلما تجمعت القيم حول وسطها الحسابي كان هذا

الوسط الحسابي يمثل بشكل جيد هذه القيم والعكس صحيح، بمعنى أنه كلما كان مقياس التشتت صغيراً كان هذا مؤشراً على أن مقياس النزعة المركزية يمثل بياناته أصدق تمثيل.

3- لمقاييس التشتت دور هام من الناحية التحليلية فيما يخص الاستدلال الإحصائي، فالعينات تتم دراستها للاستدلال على المجتمعات التي سحبت منها، ففي كل الأحوال يستخدم الوسط الحسابي للعينات للاستدلال على الوسط الحسابي للمجتمع الإحصائي الذي سحبت منه تلك العينة، ولكن في أغلب الأحيان لا ينطبق الوسط الحسابي للعينات مع الوسط الحسابي للمجتمع الإحصائي، فهناك خطأ متوقع نتيجة لاستخدام أسلوب العينات، وهنا يأتي دور مقاييس التشتت في المساهمة بمعلومات توضح حجم هذا الخطأ، وسوف نتطرق لهذا الموضوع بشكل مفصل في العام القادم.

4- إن مقاييس التشتت لها الدور التكميلي لمقاييس النزعة المركزية في وصف التوزيع البياني، لها دور هام أيضاً حيث تتم بمساعدتها مقارنة توزيع بياني مع آخر .
وتصنف مقاييس التشتت إلى صنفين :

أ) الصنف الأول: هي المقاييس التي توضح مدى وكيفية اختلاف القيم فيما بينها.
ب) الصنف الثاني: هي المقاييس التي توضح مدى تناثر ، أو تجمع القيم حول أحد المتوسطات.

وسوف تتم دراسة مقياس المدى ومقياس الانحراف الربيعي «الصنف الأول»، والانحراف المتوسط والانحراف المعياري (التباين) «الصنف الثاني» .
وتجدر الإشارة إلى أن هناك نوعين من مقاييس التشتت :

أ) مقاييس العشنت المطلقة: والتي تكون وحداتها وحدات القيم الأصلية نفسها.
ب) مقاييس العشنت النسبية: وهي التي تكون خالية من وحدات القياس (كنسبة مئوية).

2-5- المدى المطلق The Range :

وهو أبسط مقياس من مقاييس التشتت ويعبر عن الفرق بين أكبر مشاهدته، وأصغر مشاهدته، فمثلاً بالنسبة لمثال علامات الطلاب وعلامات الطالبات من المثال السابق (5-1) نجد :

$$\text{علامة } R_1 = 77 - 65 = 12 \text{ مدى علامات الطلاب}$$

$$\text{علامة } R_2 = 100 - 41 = 59 \text{ مدى علامات الطالبات.}$$

نجد أن المدى المطلق لعلامات الطلاب أصغر من المدى المطلق لعلامات الطالبات، وهذا يعني أن تشتت علامات الطلاب أصغر من تشتت علامات الطالبات، ولذلك يمكن أن نقرر أن علامات الطلاب أكثر تجانساً من علامات الطالبات، بمعنى أوضح أن الطلاب كانوا أكثر تقارباً مع بعضهم البعض من حيث علاماتهم في تلك المادة، بعكس الطالبات فإن هناك تباعداً وتناثراً فيما بينهن من حيث علاماتهم في هذه المادة.

مثال (5-2) :

أوجد المدى المطلق لكل من المجموعتين الآتيتين :

$$(A) : 8, 2, 9, 9, 8, 9, 8, 19$$

$$(B) : 11, 7, 6, 2, 17, 9, 19, 5$$

$$R_A = 19 - 2 = 17$$

$$R_B = 19 - 2 = 17$$

نشاهد أن المدى المطلق في كلتا المجموعتين متساوٍ، ولكن نلاحظ من بيانات المجموعة (B) أن الاختلاف بين مشاهداتها أكبر من الاختلاف بين مشاهدات المجموعة (A) لأن أغلب مشاهدات المجموعة A متقاربة، فهي تتألف في معظمها من الرقمين 8 و 9، لذلك نجد أن المدى يكون في معظم الأحيان مقياساً مفضلاً لاعتماده على القيمتين المتطرفتين (بعد ترتيب القيم تصاعدياً أو تنازلياً) اللتين غالباً ما تكونان شاذتين، وبالأخص إذا كان حجم العينة أو المجتمع كبيراً نسبياً .

أما عن كيفية حساب المدى المطلق للبيانات المبوبة في فئات هناك طريقتان :

1- الحد الأعلى للفئة الأخيرة - الحد الأدنى للفئة الأولى .

2- مركز الفئة الأخيرة - مركز الفئة الأولى .

ونظراً لعدم معرفة قيمة المفردة الكبرى وقيمة المفردة الصغرى، ولوجود بعض التوزيعات التكرارية ذات الفئات المفتوحة، أي عندما تترك نهاية الفئة الأخيرة في التوزيع التكراري بدون حد، فإن قيمة المدة لمثل هذه الجداول تكون تقريبية وليست حقيقية .

يتميز مقياس المدى المطلق بأن عملياته الحسابية سهلة وسريعة، كما أن أغلب استخداماته في مجال السيطرة النوعية للإنتاج، وذلك لضمان أن تكون الوحدات المنتجة ضمن حدي أدنى وأعلى للقبول، كما يستخدم في قياس التغير بدرجات الحرارة وتشتت الأسعار اليومية للسندات والأوراق المالية (افتتاح سعر السندات في بداية اليوم وإقفال سعر السندات في نهاية اليوم) ويستخدم في بعض الدراسات المخبرية الطبية (سرعة تدفق الدم في الساعة الأولى والساعة الثانية....)، غير أنه لا يصلح أن يكون مقياساً دقيقاً لمدى تشتت البيانات، فهو يهمل بقية القيم الواقعة بين هاتين القيمتين المتطرفتين وبالتالي فهو شديد التأثير بالقيم المتطرفة كالوسط الحسابي.. كما أنه لا يصلح استخدامه إذا كان التوزيع شديد الالتواء، وأن المدى المطلق لا يصلح إلا للمقارنة بين توزيعين لمتغير واحد ولهما وحدة القياس نفسها .

1-2-5- المدى النسبي :

إن المقاييس النسبية تقيس التشتت على شكل نسبة مئوية، ولها مدلول أكبر من المقاييس المطلقة، وبالأخص إذا كانت المقارنة بين بيانات أو مجموعات من وحدات إحصائية مختلفة في مقاييسها أو يختلف متوسط كل منها عن الآخر اختلافاً كبيراً، وإن كان لها وحدات القياس نفسها، ولإجراء مثل هذه المقارنة لا بد من التخلص من وحدات القياس باستخدام مقياس نسبي يخلصنا من هذه الوحدات المختلفة، وهذه المقاييس تدعى بالمقاييس النسبية للتشتت، فمقياس التشتت النسبي يساوي

$$100 \times \frac{\text{مقياس التشتت المطلق}}{\text{مقياس النزعة المركزية المقابل}}$$

مثال (5-3) :

إذا كان لدينا نوعان من السندات المالية تتراوح أسعارها بين (20-28) للنوع الأول و (110-118) للنوع الثاني. نجد أن كلا النوعين لهما المدى نفسه غير أن الاختلاف في أسعار النوع الأول أكثر أهمية من الاختلاف في أسعار النوع الثاني، ففي هذه الحالة يعبر عن المدى كنسبة مئوية من الوسط الحسابي الحقيقي :

$$\left(\bar{X} = \frac{\sum X}{n} \right)$$

أو إذا كان الوسط الحسابي مجهولاً فيعبر عنه بـ :

$$\frac{\text{أكبر مفردة} - \text{أصغر مفردة}}{2}$$

$$\text{المدى النسبي} = 100 \times \frac{\text{أكبر مفردة} - \text{أصغر مفردة}}{\text{الوسط الحسابي}}$$

$$\text{المدى النسبي للنوع الأول} \quad R_1\% = \frac{28 - 20}{\frac{28 + 20}{2}} \cdot 100$$

$$= \frac{8}{24} \cdot 100 = 33,33 \%$$

$$\text{المدى النسبي للنوع الثاني} \quad R_2\% = \frac{118 - 110}{\frac{118 + 110}{2}} \cdot 100$$

$$= \frac{8}{114} \cdot 100 = 7,02 \%$$

- ١٨٦ -

وهكذا ، نجد أولاً أن الوسط الحسابي للنوع الأول يساوي 24 والوسط الحسابي للنوع الثاني يساوي 114 والفارق بينهما كبير جداً، فيجب في هذه الحالة أن تتم المقارنات بين النوعين من السندات كنسبة مئوية، حتى وإن كانت وحدات القياس موحدة بين النوعين ، نشاهد أن المدى النسبي للنوع الأول هو أكبر من المدى النسبي للنوع الثاني، مع العلم أن المدى المطلق يساوي 8 لكلا النوعين، وبالتالي فإن التشتت في النوع الأول هو أكبر من النوع الثاني وذلك كنسبة مئوية من الوسط الحسابي.

والجدير بالذكر أنه ليس لمقياس التشتت المطلق ومقياس التشتت النسبي أي مدلول إلا إذا ما قورنت مع مقياس التشتت المماثلة، فمثلاً إذا كان لدينا نوع واحد من السندات (النوع الأول فقط) فإن المدى النسبي كما رأينا يساوي 33,33% فإن هذه النسبة ليس لها أي مدلول إلا إذا قورنت بمدى نسبي آخر لمعرفة التغيير أكان كبيراً أم صغيراً .

3-5- الانحراف الربيعي (لصف المدى الربيعي)

The Quartile Deviation (Semi-interquartile Range)

إن تركيز المدى المطلق على أصغر قيمة وأكبر قيمة في التوزيع يعتبر من أهم عيوبه لإهماله بقية القيم الواقعة بين هاتين القيمتين، وللتغلب بعض الشيء على هذا العيب ولتجنب القيم المتطرفة والشاذة يجب الاستناد إلى قيمتين من داخل التوزيع، فالمدى الربيعي لا يأخذ القيم المتطرفة بعين الاعتبار بل يأخذ قيمتين من داخل التوزيع، وهما قيمة الربع الأول (الأدنى) وقيمة الربع الثالث (الأعلى) للحصول على بعض المعلومات عن بقية قيم التوزيع، بمعنى أنه يتم استبعاد نسبة من القيم من الطرفين بعد أن يجري ترتيبها تصاعدياً باعتبارها قيماً متطرفة، فإذا ما تم استبعاد الربع الأول من البيانات والربع الرابع منها (الأخير) طبعاً بعد ترتيبها تصاعدياً سمي المقياس بمدى الربيعي، وفي بعض الأحيان يتم استبعاد العشر الأول والعشر الأخير من البيانات عندها يدعى المقياس بالمدى العشري.

والانحراف الربيعي يعرف على أنه نصف المدى الربيعي، والمدى الربيعي هو الفرق بين الربع الثالث والربع الأول.

وغالباً ما يستخدم الانحراف الربيعي كمقياس للتشتت عندما يكون الوسيط هو مقياس النزعة المركزية للبيانات نفسها.
وعن كيفية إيجاد قيمة الربع الأول وقيمة الربع الثالث كنا قد تطرقنا لطريقة إيجادهم في الفصل السابق.

ولنرمز للانحراف الربيعي بـ Q ويساوي :

$$Q = \frac{\text{المدى الربيعي}}{2}$$

$$= \frac{Q_3 - Q_1}{2}$$

مثال (5-4) :

يبين الجدول الآتي علامات ست مواد لطالبتين (A, B) من طلاب السنة الأولى في كلية الاقتصاد، والمطلوب حساب الانحراف الربيعي .

جدول رقم (5-1)

المادة \ الطالبة	1	2	3	4	5	6	Σ
A	50	56	49	56	59	48	318
B	68	46	47	49	41	67	318

نرتب البيانات ترتيباً تصاعدياً :

	1	2	3	4	5	6
A	48	49	50	56	56	59
B	41	46	47	49	67	68

- موقع الربع الأول لـ A يساوي :

$$\frac{n}{4} = \frac{6}{4} = 1,5$$

أي إن موقع الربع الأول لعلامات الطالبة A يقع بين القيمتين الأولى والثانية (48, 49)، وبالتالي قيمة الربع الأول تساوي :

$$\frac{48 + 49}{2} = 48,5$$

- موقع الربيع الثالث لـ A يساوي :

$$\frac{3n}{4} = \frac{(3)(6)}{4} = 4,5$$

وبالتالي قيمة الربيع الثالث تساوي :

$$\frac{56 + 56}{2} = 56$$

(الربيع الثالث يقع بين القيمتين الرابعة والخامسة).

- موقع الربيع الأول للطالبة B يساوي :

$$\frac{n}{4} = \frac{6}{4} = 1,5$$

وقيمة هذا الربيع تساوي :

$$\frac{41 + 46}{2} = 43,5$$

- موقع الربيع الثالث للطالبة B يساوي :

$$\frac{3n}{4} = \frac{(3)(6)}{4} = 4,5$$

وقيمة هذا الربيع تساوي :

$$\frac{49 + 67}{2} = 58$$

فالانحراف الربيعي للطالبة A هو :

$$Q_A = \frac{Q_{3A} - Q_{1A}}{2} = \frac{n}{4}$$

$$= \frac{56 - 48,5}{2} = 3,75 \text{ علامة}$$

والانحراف الربيعي للطالبة B يساوي :

$$Q_B = \frac{Q_{3B} - Q_{1B}}{2}$$

$$= \frac{58 - 43,5}{2} = 3,75 \text{ علامة}$$

نستنتج أن تشتت علامات الطالبة A هو أقل من تشتت علامات الطالبة B.
ولمعرفة كيفية حساب الانحراف الربيعي لبيانات مهوبة في فئات نطرح المثال الآتي:

مثال (5-5) :

يبين الجدول الآتي النفقات اليومية لعينة عشوائية مكونة من 200 أسرة عمالية تعمل في القطاع العام .

جدول رقم (5-2)

نفقات النفقات (وحدة نقدية)	عدد العمال	التكرار التجميعي الصاعد
100 - 120	18	18
120 - 150	44	62
150 - 200	100	162
200 - 240	24	186
240 فما فوق	14	200
Σ	200	

موقع الربيع الأول :

$$\frac{(i)(\Sigma f)}{4} = \frac{(1)(200)}{4} = 50$$

ومن التكرار التجميعي الصاعد نحدد الفئة الربيعية 120-150 (فئة الربيع الأول).

فقيمة الربيع الأول تساوي :

$$Q_1 = L_{Q_1} + \frac{\frac{\Sigma f}{4} - f_{Q_1-1}}{f_{Q_1}} \cdot C_{Q_1}$$

$$= 120 + \frac{50 - 18}{44} \cdot 30$$

$$= 120 + 21,82 = 141,82 \quad \text{وحدة نقدية}$$

وموقع الربيع الثالث :

$$\frac{(3)(200)}{4} = 150$$

ومن التكرار التجميعي الصاعد لتحديد الفئة الربيعية وهي 150-200 (فئة الربيع الثالث).
فقيمة الربيع الثالث تساوي :

$$Q_3 = L_{Q_3} + \frac{\frac{(3)\Sigma f}{4} - f_{Q_3-1}}{f_{Q_3}} \cdot C_{Q_3}$$

$$= 150 + \frac{150 - 62}{100} \cdot 50$$

$$= 150 + 44 = 194 \quad \text{وحدة نقدية}$$

فالانحراف الربيعي (نصف المدى الربيعي) يساوي :

$$Q = \frac{Q_3 - Q_1}{2}$$

$$= \frac{194 - 141,82}{2} = 26,09 \quad \text{وحدة نقدية}$$

1-3-5- الانحراف الربيعي النسبي :

في أغلب الأحيان يستعمل الانحراف الربيعي للمقارنة، ولذلك يجب أن يعبر عنه

كنسبة مئوية من الوسيط فيجب تحويل الانحراف الربيعي إلى انحراف ربيعي نسبي كما يأتي :

$$Q\% = \frac{Q}{Med} \cdot 100$$

$$Q\% = \frac{Q_3 - Q_1}{2 Med} \cdot 100$$

ولحساب الوسيط لبيانات الجدول السابق رقم (2-5) نجد أن موقع الوسيط

يساوي :

$$\frac{\Sigma f}{2} = \frac{200}{2} = 100$$

ومن التكرار التجميعي الصاعد نحدد الفئة الوسيطة فنجدها تقابل الفئة

150-200 فقيمة الوسيط تساوي :

$$Med = 150 + \frac{100 - 62}{100} \cdot 50$$

$$= 150 + 19 = 169 \text{ وحدة نقدية}$$

والانحراف الربيعي النسبي لانفاق 200 عامل من عمال القطاع العام يساوي:

$$Q\% = \frac{26,09}{169} \cdot 100$$

$$= 15,44 \%$$

وبفرض أننا سحبنا عينة عشوائية أخرى من 200 عامل من عمال القطاع

الخاص، لدراسة نفقاتهم اليومية، فتبين لنا أن الانحراف الربيعي النسبي يساوي 20%.

والسؤال أي النفقات ذات انتشار أكبر؟ هل نفقات عمال القطاع العام أم نفقات

عمال القطاع الخاص؟

مقارنة الانحراف الربيعي النسبي لنفقات عمال القطاع العام مع الانحراف الربيعي

النسبي لنفقات عمال القطاع الخاص نجد أن $Q\%$ لعمال القطاع العام $> Q\%$ لعمال

القطاع الخاص $15,44\% > 20\%$.

وبالتالي فإن نفقات عمال القطاع العام تعرض انتشاراً أقل من نفقات عمال القطاع الخاص، وإن الوسيط لنفقات عمال القطاع العام يعبر عن بياناته بشكل أفضل من وسيط نفقات عمال القطاع الخاص .

2-3-5- محاسن ومساوئ الانحراف الربيعي :

آ محاسنه :

1- إنه المقياس الوحيد للتشتت إذا كانت بعض فئات السندول التكراري مفتوحة.

2- إنه أقل المقاييس تأثراً بالقيم الشاذة والمتطرفة على جانبي التوزيع.

ب) مساوئه :

1- إنه يأخذ بالحسبان فقط 50% من المفردات، وليهمل الـ 50% الباقية (ربع المفردات أقل من الربع الأول وربع المفردات أكبر من الربع الثالث).

2- لا يظهر لنا بشكل واضح تباعد أو تقارب المفردات الواقعة بين الربع الأول والربع الثالث، فهو يعطي فكرة عن هذا التشتت وليس عن القيمة بدقة.

3- لا يتصف بخواص رياضية وحبرية، الأمر الذي لا يساعد على استخدامه في الدراسات الإحصائية المتقدمة.

4-5- الانحراف المتوسط (The Average Deviation) :

من الممكن معرفة تشتت القيم باستخدام القيمة الوسط الحسابي لتلك القيم، أي باستخدام $(x - \bar{x})$ فإذا اتضح أن هذه الانحرافات كبيرة، يعني ذلك أن تشتت البيانات، هو أكبر من تلك الحالة التي تكون الـ $(x - \bar{x})$ إشارات عن الوسط الحسابي تشكل قيمة صغرى، وبالتالي يمكن استخدام انحرافات القيم عن وسطها الحسابي $(x - \bar{x})$ لمعرفة مقدار تباعد تشتت البيانات المدروسة، وقد يتوقع البعض أنه يمكن استخدام الوسط الحسابي لهذه الانحرافات كمقياس للتشتت، ولكن ذلك سوف يكون مستحيلاً بسبب

أن بعض هذه الانحرافات سوف تكون سالبة والبعض الآخر سوف تكون موجبة، وأن مجموعها سوف يكون معدوماً (كما ذكرنا سابقاً في خواص الوسط الحسابي) مجموع الانحرافات القيم عن وسطها الحسابي يساوي الصفر :

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

وللتخلص من أن $\sum (x - \bar{x}) = 0$ تم حساب ما يسمى بالانحراف المتوسط، وهو الوسط الحسابي لمجموع انحرافات القيم عن وسطها الحسابي بالقيمة المطلقة، وهذا يعني إهمال الإشارة السالبة .

1-4-5- حساب الانحراف المتوسط للبيانات غير المبوبة :

إذا كانت $x_1, x_2, \dots, x_n$ قيم المشاهدات وكان $\bar{x} = \frac{\sum x}{n}$ فإن الانحراف المتوسط للبيانات غير المبوبة يساوي :

$$D_m = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

حيث : $|x_i - \bar{x}|$ هي القيمة المطلقة أي تجاهل الإشارات السالبة .

x_i = قيم المشاهدات .

$\bar{x}$ = الوسط الحسابي لتلك القيم .

D_m = الانحراف المتوسط .

مثال (5-6) :

احسب الانحراف المتوسط لعلامات الطالبين (A, B) من المثال رقم (5-4) .

جدول رقم (3-5)

المادة	علامات الطالبة A	علامات الطالبة B	الانحرافات المطلقة عن الوسط الحسابي للطالبة A $ x_A - \bar{x}_A $	الانحرافات المطلقة عن الوسط الحسابي للطالبة B $ x_B - \bar{x}_B $
1	50	68	3	15
2	56	46	3	7
3	49	47	4	6
4	56	49	3	4
5	59	41	6	12
6	48	67	5	14
Σ	318	318	24	58

$$\bar{x}_A = \frac{318}{6} = 53 \text{ علامة}$$

$$\bar{x}_B = \frac{318}{6} = 53 \text{ علامة}$$

والانحراف المتوسط لعلامات الطالبة A يساوي :

$$D_m = \frac{\Sigma |x_A - \bar{x}_A|}{n} = 53 \text{ علامة}$$

$$= \frac{24}{6} = 4 \text{ علامات}$$

والانحراف المتوسط لعلامات الطالبة B يساوي :

$$D_m = \frac{\Sigma |x_B - \bar{x}_B|}{n} = 53 \text{ علامة}$$

$$= \frac{58}{6} = 9,67 \text{ علامة}$$

إن متوسط انحرافات علامات الطالبة A عن الوسط الحسابي لعلاماتها يعادل 4 علامات ، وإن متوسط انحرافات علامات الطالبة B عن الوسط الحسابي لعلاماتها يعادل 9,67 علامة.

2-4-5- حساب الانحراف المتوسط للبيانات المبوبة Grouped data في فئات:

بفرض أن القيم الواقعة ضمن الفئات هي عند مركز هذه الفئات، وبالتالي فإن الانحراف المتوسط للبيانات المبوبة في فئات هو عبارة عن مجموع انحرافات قيم مراكز الفئات عن الوسط الحسابي بالقيمة المطلقة مضروبة بتكراراتها، وتقسيم الناتج على مجموع التكرارات، وبالتالي فالانحراف المتوسط للبيانات المبوبة يساوي:

$$D_m = \frac{\sum |x_i - \bar{x}| f}{\sum f}$$

مثال (5-7) :

ليكن لدينا الجدول الآتي المتضمن توزيع النفقات الأسبوعية لعينتين عشوائيتين من الأسر العمالية والتي تعمل في القطاع العام والقطاع الخاص (عمات الليرات) .

جدول رقم (5-4)

فئات النفقات (عمات الليرات)	عدد عمال القطاع العام f_1	عدد عمال القطاع الخاص f_2
1,5 - 2,5	18	34
2,5 - 3,5	42	38
3,5 - 4,5	80	62
4,5 - 5,5	46	36
5,5 - 6,5	10	20
6,5 - 7,5	4	10
Σ	200	200

احسب الوسط الحسابي للنفقات الأسبوعية لهاتين العينتين في كل قطاع على حده ومن ثم أوجد الانحراف المتوسط لنفقات عمال القطاع العام والقطاع الخاص لهاتين العينتين ماذا تستنتج ؟

الحل :

الجدول الآتي هو جدول مساعد لحساب كل من الوسط الحسابي والانحراف المتوسط :

جدول رقم (5-5)

وسطي الفئة (x) (بمئات الليرات)	f ₁	f ₂	x f ₁	x f ₂	x - $\bar{x}$	x - $\bar{x}$ f ₁	x - $\bar{x}$ f ₂
2	18	34	36	68	2	36	68
3	42	38	126	114	1	42	38
4	80	62	320	248	0	0	0
5	46	36	230	180	1	46	36
6	10	20	60	120	2	20	40
7	4	10	28	70	3	12	30
Σ	200	200	800	800		156	212

$$\bar{x}_1 = \frac{\Sigma x f_1}{\Sigma f_1} = \frac{800}{200} = 4$$

$$\bar{x}_2 = \frac{\Sigma x f_2}{\Sigma f_2} = \frac{800}{200} = 4$$

$$\bar{x} = \bar{x}_1 = \bar{x}_2$$

هذا يعني أن متوسط الإنفاق لكل من أسر القطاع العام وأسرة القطاع الخاص يساوي 400 ليرة لهاتين العينتين.

أما الانحراف المتوسط لنفقات أسر عمال القطاع العام يساوي :

$$D_{m_1} = \frac{\Sigma |x - \bar{x}| f_1}{\Sigma f_1}$$

$$= \frac{156}{200} = 0,78$$

بمعنى أن متوسط الانحرافات نفقات أسر عمال القطاع العام عن الوسط الحسابي لنفقاتها يساوي 78 ليرة لهذه العينة .

الانحراف المتوسط لنفقات أسر عمال القطاع الخاص :

$$D_{m2} = \frac{\sum |x - \bar{x}| f_2}{\sum f_2}$$

$$= \frac{212}{200} = 1,06$$

أي انحرافات نفقات أسر عمال القطاع الخاص عن الوسط الحسابي لنفقاتها يساوي 106 ليرة لهذه العينة، وبالتالي فإن نفقات أسر عمال القطاع الخاص تظهر تشتتاً أكبر من تشتت نفقات أسر عمال القطاع العام. وبمعنى آخر فالوسط الحسابي لنفقات أسر عمال القطاع العام يمثل متوسط نفقاتهم بشكل أفضل من تمثيل متوسط نفقات أسر القطاع الخاص، فكلما قل مقياس التشتت أدى ذلك إلى تمثيل مقياس النزعة المركزية لبياناته بشكل أدق.

3-4-5- الانحراف المتوسط النسبي :

في المثال السابق رقم (5-7) كان الوسطان الحسابيان لأسر القطاعين متساويين وكانت وحدات القياس (الليرة) واحدة، ولكن في أغلب الأحيان لن تكون الشروط السابقة متوفرة للمقارنة، بمعنى أنه قد تختلف وحدات القياس في القطاعين (ليرة) و (دولار)، وبالتالي سوف يختلف المتوسطان الحسابيان. عندها لن نستطيع المقارنة بين القطاعين إلا إذا تم تحويل مقياس التشتت إلى مقياس نسبي (كما تم ذكره سابقاً) فالانحراف المتوسط النسبي سوف يعبر عنه كنسبة مئوية من الوسط الحسابي :

$$\text{الانحراف المتوسط النسبي} = 100 \times \frac{\text{الانحراف المتوسط}}{\text{الوسط الحسابي}}$$

من المثال السابق رقم (5-7) نجد :

$$D_m \% = \frac{D_m}{\bar{x}} \cdot 100$$

إن الانحراف المتوسط النسبي لنفقات أسر عمال القطاع العام في هذه العينة يساوي :

$$D_{m_1} \% = \frac{0,78}{4} \cdot 100 = 19,5 \%$$

والانحراف المتوسط النسبي لنفقات أسر عمال القطاع الخاص في هذه العينة يساوي:

$$D_{m_2} \% = \frac{1,06}{4} \cdot 100 = 26,5 \%$$

نجد أن $D_{m_2} \% > D_{m_1} \%$ وبالتالي تشتت نفقات أسر عمال القطاع العام أصغر من تشتت نفقات أسر عمال القطاع الخاص في هاتين العينتين، بمعنى أن نسبة انتشار النفقات لأسر عمال القطاع العام حول الوسط الحسابي لنفقاتها يعادل 19,5% من قيمة وسطي إنفاقها. وإن نسبة انتشار النفقات لأسر عمال القطاع الخاص حول الوسط الحسابي لنفقاتها يعادل 26,5% من قيمة وسطي إنفاقها.

4-4-5- محاسن ومساوى الانحراف المتوسط :

آ) محاسنه :

- 1- إنه يأخذ بعين الاعتبار جميع القيم (انحرافات كل قيمة عن الوسط الحسابي) فهو يعتبر مقياساً جيداً للتشتت بالمقارنة مع مقياسي المدى والانحراف الربيعي.
- 2- يمكن حسابه من كل من الوسط الحسابي والوسيط، إلا أن قيمته عند حسابه من الوسيط تكون دائماً أقل من قيمته عند حسابه من الوسط، ففي التوزيعات المتماثلة والقريبة من التماثل يفضل حسابه من الوسط الحسابي، أما في التوزيعات غير المتماثلة فيفضل حسابه من الوسيط.

3- سهل الحساب والفهم فهو أسهل حسابياً من الانحراف المعياري، فبساطة العمليات الحسابية اللازمة لإيجاده تشجع غير المتخصصين في الإحصاء على استخدامه .
(آ) مساوئه :

- 1- نادر الاستعمال فاستعماله يقتصر على سرعته في قياس التنسب. وسبب قلة استعماله كون عملية احتسابه تعتمد على القيم المطلقة، وبالتالي إهمال الإشارات الجبرية الأمر الذي يتعذر التعامل معه رياضياً، فمثلاً إذا كانت نتيجة جميع البدائل تؤدي إلى حالة واحدة $|x| + |L| = 1$ سواء أكانت $x + L$ أم $x - L$ أو $x - L$ أو $x + L$.
- 2- لا يمكن حسابه من الجداول التكرارية ذات الفئات المفتوحة .

5-5- الانحراف المعياري «Standard Deviation» :

في سبيل التغلب على المشكلة التي تقيس مجموع الانحرافات الم من عن وسط الحساب والتي تساوي الصفر أي على مشحلة الاشارات الجبرية عند جميع الانحرافات، فبدلاً من أخذ القيم المطلقة للانحرافات وإهمال الإشارات الجبرية، فإننا نستطيع أن تغلب على هذه الخاصة التي يتسبب 4 أن متوسط انحرافات، بتربيع قيم الانحرافات، ولذلك تصبح جميعها موجبة، وبذلك نحصل على مجموع مربعات الانحرافات، Sum of Square variations بدلاً من مجموع الانحرافات، بمعنى أنه ليس من الضروري إهمال الإشارات الجبرية عند إيجاد مقياس مناسب للتشتت، فمجموع مربعات الانحرافات $\sum (x_i - \bar{x})^2$ يتم تقسيمها على n وبذلك نحصل على ما يدعى بالتباين (S^2) .
إذا كان لدينا عينة من حجم n: $x_1, x_2, \dots, x_n$ و x_i فالتباين للمجتمع يساوي :

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

أما إذا كان لدينا مجتمع إحصائي حجمه N و $x_1, x_2, \dots, x_N$ فالمجتمع يساوي :

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

حيث : μ = الوسط الحسابي للمجتمع .

σ^2 = تباين المجتمع (Sigma Square).

وبما أننا في حساب التباين تم تربيع الانحرافات، فإن المقياس محدد على أساس مربع القيم المستخدمة، فإذا كانت قيم المشاهدات مقاسة بالتر، فالتباين سيكون مقاساً بالتر المربع، وهنا لن توجد مشكلة في ذلك، ولكن إذا كانت قيم المشاهدات تعبر عن مبالغ بالليرة أو إعداد طلاب أو أوزان بالطن مثلاً فإن التباين يكون مقاساً بالليرة المربعة، أو الطالب المربع أو الطن المربع، وكلها ليس لها معنى.. ونتيجة لذلك ولكي نعيد وحدات القياس إلى أصلها، فإنه يتم أخذ الجذر التربيعي للتباين لنحصل على ما يدعى بالانحراف المعياري S وهو مقياس بالوحدات الأصلية وليست بالوحدات المربعة، وبالتالي $S = \sqrt{S^2}$ وحيث إن أي مقياس للتشتت يجب أن لا يقل عن الصفر لذا لزم أخذ الجذر التربيعي الموجب للتباين لإيجاد الانحراف المعياري .

5-5-1- حساب الانحراف المعياري للبيانات غير المبوبة :

(1) - الطريقة المباشرة :

الانحراف المعياري لعينة ما هو الجذر التربيعي الموجب لتباين تلك العينة :

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

وتلخص العمليات الحسابية لإيجاد الانحراف المعياري بالطريقة المباشرة كما يأتي:

- 1- نوجد الوسط الحسابي للبيانات .
- 2- نوجد انحرافات القيم x_i عن وسطها الحسابي $\bar{x}$.
- 3- نربع كل انحراف من الانحرافات .
- 4- نحسب مجموع مربع الانحرافات .

5- نقسم مجموع مربيع الانحرافات على حجم العينة n فنحصل على التباين S^2 .

6- نأخذ الجذر التربيعي الموجب للتباين فنحصل على الانحراف المعياري S .

مثال (5-8) :

لدينا المعلومات الآتية والتي تمثل عدد الساعات الأسبوعية لدراسة أحد الطلاب

خلال سبعة أسابيع : احسب الانحراف المعياري لتشتت ساعات الدراسة لهذا الطالب بالطريقة المباشرة .

عدد الساعات (x)	الأسبوع
15	1
9	2
14	3
6	4
23	5
19	6
33	7
119	Σ

$$\bar{x}_1 = \frac{\Sigma x}{n} = \frac{119}{7} = 17 \text{ ساعة}$$

متوسط الدراسة الأسبوعية للطلاب .

جدول مساعد لحساب الانحراف المعياري بالطريقة المباشرة.

جدول رقم (5-6)

x	$x - \bar{x}$	$(x - \bar{x})^2$
15	- 2	4
9	- 8	64
14	- 3	9
6	- 11	121
23	6	36
19	2	4
33	16	256
119	0	494

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{494}{7}}$$

$$S = 8,40 \text{ ساعة}$$

(2) طريقة المربعات :

من الطريقة المباشرة نجد :

$$S^2 = \frac{\sum (x - \bar{x})^2}{n} \Rightarrow$$

$$\sum (x - \bar{x})^2 = \sum (x^2 - 2x\bar{x} + \bar{x}^2)$$

$$= \sum x^2 - 2\bar{x} \sum x + n\bar{x}^2$$

$$= \sum x^2 - 2 \left(\frac{\sum x}{n} \right) (\sum x) + n \frac{(\sum x)^2}{n^2}$$

$$= \sum x^2 - 2 \frac{(\sum x)^2}{n} + \frac{(\sum x)^2}{n}$$

$$= \sum x^2 - \frac{(\sum x)^2}{n}$$

فالتباين يساوي :

$$S^2 = \frac{\sum (x - \bar{x})^2}{n} = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}$$

والانحراف المعياري يساوي:

$$S = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}}$$

أو :

$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2} = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$$

و بتطبيق الصيغ الأخيرة على المثال السابق رقم (5-8) بعد إيجاد ΣX و ΣX^2
(انظر الجدول المساعد رقم 5-7) :

$$S = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n}} = \sqrt{\frac{2517 - \frac{(119)^2}{7}}{7}} = 8,40 \text{ ساعة}$$

أو :

$$S = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} = \sqrt{\frac{2517}{7} - \left(\frac{119}{7}\right)^2} = 8,40 \text{ ساعة}$$

نجد بأن هذه الصيغ تختصر الكثير من العمليات الحسابية بالمقارنة مع الطريقة المباشرة .
جدول مساعد لحساب الانحراف المعياري بطريقة المربعات :

جدول رقم (5-7)

x	x ²
15	225
9	81
14	196
6	36
23	529
19	361
33	1089
119	2517

(3) طريقة الانحرافات :

تقوم هذه الطريقة بالاعتماد على أحد خواص الانحراف المعياري أو التباين فعند طرح أو إضافة عدد ثابت، وليكن $\bar{x}_0$ إلى كل قيمة من قيم المشاهدات x فإن قيمة الانحراف المعياري أو التباين لا يتغيران، بمعنى أن الانحراف المعياري للقيم الجديدة يساوي الانحراف المعياري للقيم الحقيقية .

بالانطلاق من الصيغة الأخيرة وبطرح القيمة $\bar{x}_0$ من كل قيمة من قيم المشاهدات

x نجد :

$$S = \sqrt{\frac{\sum (x - \bar{x}_0)^2}{n} - \left(\frac{\sum (x - \bar{x}_0)}{n} \right)^2}$$

ولنرمز لـ $\sum (x - \bar{x}_0)$ بـ d. فالانحراف المعياري بطريقة الانحرافات للبيانات

غير المبوبة :

$$S = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n} \right)^2}$$

وبتطبيق المعادلة السابقة على المثال السابق بعد فرض $\bar{x}_0 = 5$ وإنشاء الجدول

المساعد، جدول رقم (5-8) نجد: $\sum d = \sum (x - \bar{x}_0)$ و $d^2 = \sum (x - \bar{x}_0)^2$ (يفضل أن

تكون قيمة $\bar{x}_0$ (الوسط الحسابي الفرضي) من ضمن البيانات لتسهيل الحسابات).

جدول مساعد لحساب الانحراف المعياري بطريقة الانحرافات:

جدول رقم (5-8)

x	d	d ²
$\bar{x}_0 = (15)$	0	0
9	- 6	36
14	- 1	1
6	- 9	81
23	8	64
19	4	16
33	18	324
Σ	14	522

$$S = \sqrt{\frac{522}{7} - \left(\frac{14}{7} \right)^2} = 8,40 \text{ ساعة}$$

وهي النتيجة السابقة نفسها وبالتالي فإن الانحراف المعياري لم يتأثر بطرح العدد الثابت إلى كل قيمة من القيم الأصلية، ويمكن استخدام طريقة الانحرافات هذه في حالة القيم الكبيرة أو العديدة اختصاراً للوقت .

مثال (5-8) :

لدينا البيانات الآتية $x_i = 9, 2, 3, 11, 10$

احسب الانحراف المعياري بالطرق الثلاث :

1- الطريقة المباشرة :

$$\bar{x} = \frac{\sum x}{n} = \frac{35}{5} = 7$$

$$S^2 = \frac{\sum (x - \bar{x})^2}{n} = \frac{(9-7)^2 + (2-7)^2 + (3-7)^2 + (11-7)^2 + (10-7)^2}{5}$$

$$S^2 = \frac{4 + 25 + 16 + 16 + 9}{5} = 14$$

$$S = \sqrt{14} = 3,74$$

2- طريقة المربعات :

نجد أن :

$$\sum x = 9 + 2 + 3 + 11 + 10 = 35$$

$$\sum x^2 = 81 + 4 + 9 + 121 + 100 = 315$$

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2 = \frac{315}{5} - \left(\frac{35}{5} \right)^2 = 14$$

$$S = \sqrt{14} = 3,74$$

3- طريقة الانحرافات :

نفرض $\bar{x}_0 = 5$ ونطرح من كل قيمة لـ x القيمة $\bar{x}_0 = 5$ مثلاً :

$$\Sigma d = \Sigma (x - \bar{x}_0) = (4) + (-3) + (-2) + (6) + (5) = 10$$

$$\Sigma d^2 = \Sigma (x - \bar{x}_0)^2 = (16) + (9) + (4) + (36) + (25) = 90$$

$$S^2 = \frac{\Sigma (x - \bar{x}_0)^2}{n} - \left(\frac{\Sigma (x - \bar{x}_0)}{n} \right)^2 = \frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n} \right)^2$$

$$S^2 = \frac{90}{5} - \left(\frac{10}{5} \right)^2 = 14$$

$$S = \sqrt{14} \approx 3,74$$

ولنحسب لهذا المثال الانحراف المتوسط : حيث

$$D_m = \frac{\Sigma |x - \bar{x}|}{n} = \frac{18}{5} = 3,6$$

نلاحظ هنا أن قيمة الانحراف المعياري هي قيمة أكبر من الانحراف المتوسط، فعندما يكون التوزيع شبه متماثل، فإن قيمة الانحراف المتوسط تساوي $\frac{4}{5}$ من قيمة الانحراف المعياري، بمعنى :

$$D_m = \frac{4}{5} \cdot S$$

2-5-5- حساب الانحراف المعياري للبيانات المبوبة في فئات :

1- الطريقة المباشرة :

إذا كانت x_i تمثل مراكز الفئات في الجدول التكراري وإن تكرارها المقابلة f_i فالانحراف المعياري بالطريقة المباشرة هو :

$$S = \sqrt{\frac{\Sigma (x - \bar{x})^2 \cdot f}{\Sigma f}}$$

2- طريقة المربعات .

تقوم على إيجاد مراكز الفئات واحتساب مربعاتها ومن ثم استخدام العلاقة الآتية:

$$S = \sqrt{\frac{\sum x^2 f - \frac{(\sum x f)^2}{\sum f}}{\sum f}}$$

هذا ويمكن الوصول إلى العلاقة الأخيرة بالطريقة نفسها التي حصلنا فيها على قيمة الانحراف المعياري في البيانات غير المبوبة (طريقة المربعات) .
أو يمكن استخدام العلاقة الآتية :

$$S = \sqrt{\frac{\sum x^2 f}{\sum f} - \left(\frac{\sum x f}{\sum f}\right)^2}$$

3- طريقة الانحرافات :

وكما هو الحال في المعطيات غير المبوبة هناك إمكانية اختصار العمليات الحسابية أيضاً مع البيانات المبوبة في فئات، وذلك بتخفيض قيم المشاهدات من خلال استخدام قيمة فرضية $\bar{x}_0$ وطرحها من قيم المشاهدات الأصلية (مراكز الفئات) ومن ثم نطبق العلاقة :

$$S = \sqrt{\frac{\sum (x - \bar{x}_0)^2 f}{\sum f} - \left(\frac{\sum (x - \bar{x}_0) f}{\sum f}\right)^2}$$

$$= \sqrt{\frac{\sum d^2 f}{\sum f} - \left(\frac{\sum d f}{\sum f}\right)^2}$$

أو :

$$S = \sqrt{\frac{\sum (x - \bar{x}_0)^2 f - \frac{[\sum (x - \bar{x}_0) f]^2}{\sum f}}{\sum f}} = \sqrt{\frac{\sum d^2 f - \frac{(\sum d f)^2}{\sum f}}{\sum f}}$$

ويفضل أن تكون قيمة $\bar{x}_0$ من أحد قيم أوساط الفئات في الجدول التكراري والواقعة في منتصف الجدول .

4- طريقة الانحرافات المختصرة :

يمكن الزيادة في اختصار العمليات الحسابية (بشرط أن تكون أطوال الفئات في الجدول التكراري متساوية)، وذلك بتقسيم انحرافات القيمة الفرضية $\bar{x}_0$ عن أوساط الفئات ولكل فئة على طول الفئة C حيث تصبح العلاقة على الشكل الآتي :

$$S^2 = C \cdot \frac{\sum \left(\frac{x - \bar{x}_0}{C} \right)^2 f - \left[\frac{\sum \left(\frac{x - \bar{x}_0}{C} \right) f}{\sum f} \right]^2}{\sum f}$$

وإذا رمزنا لـ $\frac{x - \bar{x}_0}{C}$ بـ D فيصبح دستور الانحراف المعياري بطريقة الانحرافات المختصرة:

$$S = C \cdot \sqrt{\frac{\sum D^2 f - \frac{(\sum D f)^2}{\sum f}}{\sum f}}$$

أو :

$$S = C \cdot \sqrt{\frac{\sum D^2 f}{\sum f} - \left(\frac{\sum D f}{\sum f} \right)^2}$$

حيث : C = قيمة ثابتة (طول الفئة في الجدول التكراري) .

$\bar{x}_0$ = الوسط الحسابي الفرضي وهو قيمة ثابتة، وغالباً ما يكون من وسطى الفئة الواقعة في منتصف الجدول .

مثال (5-9) :

الجدول الآتي يمثل التوزيع التكراري لمدة الخدمة بالسنوات لعينة من 100 عامل من عمال إحدى الشركات :

جدول رقم (5-9)

عدد العمال	مدة الخدمة بالسنوات
6	0,5 – 5,5
25	5,5 – 10,5
29	10,5 – 15,5
20	15,5 – 20,5
10	20,5 – 25,5
5	25,5 – 30,5
3	30,5 – 35,5
2	35,5 – 40,5
100	Σ

احسب الانحراف المعياري لمدة الخدمة لهؤلاء العمال بالطرائق الأربع.

ننشي الجدول المساعد الآتي :

جدول مساعد لإيجاد قيمة الوسط الحسابي والانحراف المعياري بالطريقة المباشرة

وطريقة المربعات

جدول رقم (5-10)

وسطي الفئة x	عدد العمال f	$x \cdot f$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$	x^2	$x^2 \cdot f$
3	6	18	144	864	9	54
8	25	200	49	1225	64	1600
13	29	377	4	116	169	4901
18	20	360	9	180	324	6480
23	10	230	64	640	529	5290
28	5	140	169	845	784	3920
33	3	99	324	972	1089	3267
38	2	76	529	1058	1444	2888
Σ	100	1500		5900		28400

أولاً - الطريقة المباشرة :

نوجد أولاً الوسط الحسابي لمدة الخدمة طويلاً العمال :

$$\bar{x} = \frac{\sum x f}{\sum f} = \frac{1500}{100} = 15 \text{ سنة}$$

ومن ثم نجد الانحراف المعياري بمساعدة الجدول (5-10) نجد :

$$S = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}} = 7,68 \text{ سنة}$$

ثانياً - طريقة المربعات :

وبمساعدة الجدول (5-10) نجد :

$$S = \sqrt{\frac{\sum x^2 f - \frac{(\sum x f)^2}{\sum f}}{\sum f}} = \sqrt{\frac{28400 - \frac{(1500)^2}{100}}{100}}$$

$$S = 7,68 \text{ سنة}$$

ثالثاً - طريقة الانحرافات :

نفرض $\bar{x}_0 = 23$ فالانحراف المعياري يساوي وبحسب هذه الطريقة وبمساعدة

الجدول رقم (5 - 11) :

$$S = \sqrt{\frac{\sum d^2 f - \frac{(\sum d f)^2}{\sum f}}{\sum f}} = \sqrt{\frac{12300 - \frac{(-800)^2}{100}}{100}}$$

$$S = 7,68 \text{ سنة}$$

رابعاً - طريقة الانحرافات المختصرة :

كما أشرنا يجب أن تكون أطوال الفئات متساوية.

حيث : $C = 5$ (طول الفئة) و $\bar{x}_0 = 23$

وبمساعدة الجدول (5-11) نجد :

$$S = C \cdot \sqrt{\frac{\sum D^2 f - \frac{(\sum D f)^2}{\sum f}}{\sum f}}$$

$$S = (5) \cdot \sqrt{\frac{492 - \frac{(-160)^2}{100}}{100}}$$

$$= (5) \cdot \sqrt{2,36} = 7,68 \text{ سنة}$$

الجدول المساعد لإيجاد قيمة الانحراف المعياري بطريقة الانحرافات والانحرافات

المختصرة حيث : $\bar{x}_0 = 23$ و $C = 5$.

جدول رقم (5-11)

x	f	طريقة الانحرافات حيث $23 = \bar{x}_0$			طريقة الانحرافات المختصرة حيث $\bar{x}_0 = 23$ و $C = 5$ (طول الفئة)		
		$(x - \bar{x}_0)$ (d)	$(x - \bar{x}_0)f$ (d f)	$(x - \bar{x}_0)^2 f$ (d ² f)	$\left(\frac{x - \bar{x}_0}{C}\right)$ (D)	$\left(\frac{x - \bar{x}_0}{C}\right) f$ (D f)	$\left(\frac{x - \bar{x}_0}{C}\right)^2 f$ D ² f
3	6	- 20	- 120	2400	- 4	- 24	96
8	25	- 15	- 375	5625	- 3	- 75	225
13	29	- 10	- 290	2900	- 2	- 58	116
18	20	- 5	- 100	500	- 1	- 20	20
$\bar{x}_0 = 23$	10	0	0	0	0	0	0
28	5	5	25	125	1	5	5
33	3	10	30	300	2	6	12
38	2	15	30	450	3	6	18
	100		- 885 + 85 - 800	12300		- 177 + 17 - 160	492

نجد من المثال رقم (9-5) بأن كل الأساليب تؤدي إلى نتيجة واحدة، إلا أن أسلوب الانحرافات المختصرة هو الأسلوب الأقل عناء في حسابه، لاختصاره لبعض العمليات الحسابية، وبالأخص إذا كانت القيم كثيرة وكبيرة فمن الأفضل استعماله كلما كان ذلك ممكناً .

3-5-5- الانحراف المعياري النسبي (معامل الاختلاف) :(Coefficient of Variation)

يعتبر معامل الاختلاف (C.V) من أهم مقاييس التشتت النسبية، ويستعمل بشكل واسع في التحليل الإحصائي لمقارنة مقدار التشتت أو الاختلاف لمجموعتين أو أكثر من البيانات في حالة اختلاف الوحدات القياسية المستخدمة مع وحدات كل مجموعة، وكذلك عند اختلاف قيم الوسط الحسابي ، ولو كانت الوحدات القياسية غير مختلفة، فمعامل الاختلاف يعبر عن الانحراف المعياري كنسبة مئوية من الوسط الحسابي وصيغته للعينة :

$$C.V = \frac{S}{\bar{x}} \cdot 100$$

وصيغته للمجتمع الإحصائي هو :

$$C.V' = \frac{\sigma}{\mu} \cdot 100$$

حيث : σ = الانحراف المعياري للمجتمع .

μ = الوسط الحسابي للمجتمع .

مثال (10-5) :

بفرض أن المعدل العام للخرابين من كلية الاقتصاد في جامعة دمشق عام (2000) يساوي 62 علامة بانحراف معياري قدره 6,5 علامة، وإن المعدل العام للخرابين من كلية الاقتصاد في جامعة حلب يساوي 70 علامة بانحراف معياري 7 علامات العام نفسه، فهل نستدل من هذه البيانات أن مقدار التشتت لعلامات الخرابين من جامعة

حلب كلية الاقتصاد هي أكبر من تشتت علامات الخريجين من جامعة دمشق في كلية الاقتصاد ؟

$$C.V' = \frac{\sigma}{\mu} \cdot 100$$

$$C.V'_1 = \frac{6,5}{62} \cdot 100 = 10,48 \%$$

$$C.V'_2 = \frac{7}{70} \cdot 100 = 10 \%$$

فنجد أنه بالرغم من ارتفاع مقدار التشتت للخريجين من كلية الاقتصاد في جامعة حلب، إلا أن معامل الاختلاف يؤكد لنا أن التشتت عن متوسط العلامات لخريجي كلية الاقتصاد في جامعة دمشق هو أكبر مما عليه في حلب .

مثال (5-11) :

أوجد أيًا من الدولتين توجد فيها عدالة اجتماعية أكثر في توزيع الدخل على أفراد مجتمعهما .

جدول رقم (5-12)

الدولة	متوسط الدخل	الانحراف المعياري	نوعية العملة
I	8000	360	ليرة
II	2000	160	دولار

$$C.V'_1 = \frac{360}{8000} \cdot 100 = 4,5 \%$$

$$C.V'_2 = \frac{160}{2000} \cdot 100 = 8 \%$$

نلاحظ أن مقدار التشتت النسبي في توزيع الدخل للدولة I هو أقل من الدولة II، وبذلك فإن الدولة I هي أكثر عدالة في توزيع الدخل على أفرادها من الدولة II .

نلاحظ أنه لو تمت المقارنة بمقاييس الانحراف المعياري لكان التشتت أكبر في الدولة I عنه في الدولة II .

4-5-5- الدرجة المعيارية (Standardized Scores) :

في بعض الأحيان نضطر إلى مقارنات من نوع آخر، فمثلاً يمكن أن نقارن مشاهدين من مجموعتين مختلفتين، في هذه الحالة يتم تحويل وحدات كل مفردة إلى وحدات معيارية حتى نتمكن من المقارنة، وذلك باستخدام كل من الوسط الحسابي والانحراف المعياري. وعملية التحويل إلى درجات معيارية تتم وفق الصيغة الآتية:

$$Z = \frac{x_i - \bar{x}}{S}$$

أو

$$Z' = \frac{x_i - \mu}{\sigma}$$

حيث القيمة Z هي الدرجة المعيارية.
وبذلك يتضح أن الدرجات المعيارية نحالية من الوحدات المستخدمة في القياس.
ولمعرفة كيفية إجراء هذا التحويل نأخذ المثال الآتي :

مثال (5-12) :

في امتحان مقرر مبادئ الإحصاء حصل أحد الطلاب على 75 علامة، علماً بأن متوسط علامات جميع الطلاب في مبادئ الإحصاء يساوي 67 علامة بالانحراف المعياري يساوي 10 علامات، وإن هذا الطالب حصل في امتحان المداخل إلى المحاسبة على 85 علامة ، علماً بأن متوسط علامات جميع الطلاب في مقرر مداخل إلى المحاسبة يساوي 77 علامة بالانحراف المعياري 16 علامة.

السؤال: في أي المقرر كانت علامات الطالب أعلى بالمقارنة مع زملائه ؟

الحل :

عند مقارنة علامات الامتحانين لهذا الطالب نجد أن علامته في مدخل المحاسبة أعلى من علامته في مبادئ الإحصاء .

لكن عند تحويل علامات الطالب إلى درجات معيارية نجد أن :

$$Z'_1 = \frac{x_1 - \mu_1}{\sigma_1} \quad \text{لمبادئ الإحصاء}$$
$$= \frac{75 - 67}{10} = 0,8 \quad \text{درجة معيارية}$$

$$Z'_2 = \frac{x_2 - \mu_2}{\sigma_2} \quad \text{المدخل إلى المحاسبة}$$
$$= \frac{85 - 77}{16} = 0,5 \quad \text{درجة معيارية}$$

يتضح من مقارنات الدرجات المعيارية لهذا الطالب بأن تفوقه في مقرر مبادئ الإحصاء هو أعلى من تفوقه في مقرر المدخل إلى المحاسبة، لأن الدرجة المعيارية لمبادئ الإحصاء هي أكبر من الدرجة المعيارية لمدخل المحاسبة مقارنة مع جميع الطلاب المتشحين في تلك المقررات .

5-5-5- أهم خصائص الانحراف المعياري :

- 1- عند إضافة أو طرح عدد ثابت K إلى كل قيمة من قيم المشاهدات فإن قيمة الانحراف المعياري أو التباين لا تتغير. وهذا ما شاهدناه عند استخدام طريقة الانحرافات في إيجاد الانحراف المعياري أو التباين .
- 2- إذا ضربت كل قيمة من قيم المشاهدات بعدد ثابت (K) فإن الانحراف المعياري للقيم الجديدة = الانحراف المعياري للقيم الأصلية ضرب العدد الثابت (K) .

مثال (5-13) :

إذا كان لدينا القيم الآتية :

$$x_i = 8, 3, 2, 12, 10$$

فالوسط الحسابي = 7 فالانحراف المعياري لهذه القيم تساوي :

$$S_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{76}{5}} = 3,9$$

فإذا ضربنا كل قيمة بـ $K = 2$ فتصبح القيم الجديدة :

$$Y_i = 16, 6, 4, 24, 20$$

فالوسط الحسابي الجديد = 14 والانحراف المعياري الجديد :

$$S_y = \sqrt{\frac{\sum (y - \bar{y})^2}{n}} = \sqrt{\frac{304}{5}} = 7,8$$

وهكذا نستنتج من الانحراف المعياري للقيم الأصلية (3,9) ومن الانحراف

المعيار ، التيم الجديد (7,8) أن :

$$S_y = S_x \cdot K$$

$$7,8 = (3,9) (2) = 7,8$$

3- تبين قيم التوزيع حول الوسط الحسابي أصغر من تبين هذه القيم حول أية قيمة

أخرى بمقدار يعادل مربع الفرق بين الوسط الحسابي وهذه القيمة .

معنى :

$$S^2 = \frac{\sum (x - \bar{x})^2}{n} = \frac{\sum (x - K)^2}{n} = (\bar{x} \cdot K)^2$$

حيث K هي قيمة ثابتة .

مثال (5-14) :

إذا كانت لدينا القيم الآتية : 5, 10, 15
فتباين هذه القيم حول الوسط الحسابي ($\bar{x} = 10$) :

$$S^2 = \frac{\sum (x - \bar{x})^2}{n} = \frac{50}{3} = 16,67$$

ولنحسب تباين هذه القيم حول قيمة أخرى ($K = 5$ مثلاً) :

$$S^2 = \frac{\sum (x - K)^2}{n} = \frac{0 + 25 + 100}{3} = 41,67$$

ومربع الفرق بين القيمة K والوسط الحسابي $\bar{x}$ يساوي :

$$(\bar{x} - K)^2 = (10 - 5)^2 = 25$$

فإذا ما أضفنا مربع الفرق بين K و $\bar{x}$ ويساوي 25 إلى تباين التوزيع حول الوسط الحسابي، فنجد أنه سوف يساوي تباين التوزيع حول القيمة الفرضية K :

$$16,67 + 25 = 41,67$$

$$\frac{\sum (x - \bar{x})^2}{n} + (\bar{x} - K)^2 = \frac{\sum (x - K)^2}{n}$$

5-5-6- محاسن ومساوي الانحراف المعياري :

آ محاسنه :

- 1- يعرف بشكل جيد مفهوم التشتت، إذ يقوم على أساس متوسط انحرافات القيم عن الوسط الحسابي .
- 2- يأخذ جميع القيم بعين الاعتبار عند حسابه وبذلك لا يستبعد أي قيمة .
- 3- يتمتع بالكثير من الخواص الجبرية، ولذلك فاستعمالاته كثيرة أكثر من أي مقياس آخر للتشتت.

- 4- تسمح قيمته في معرفة القدرة التمثيلية للوسط الحسابي .
5- يمكن عن طريق الانحراف المعياري والوسط الحسابي أن تعرف نسب التوزيع تحت المنحني الطبيعي .

آ) مساوئه :

- 1- تأثره بالقيم المتطرفة أكثر من الانحراف المتوسط .
2- ليس له مدلول إلا عند مقارنته بانحرافات معيارية لمجموعات أخرى من القيم، وبفضل أن تكون المقارنات بشكل نسبي (معامل الاختلاف) .
3- قيمته دوماً أكبر من أي مقياس آخر للتشتت ما عدا المدى.
4- قيمته دوماً أصغر في العينة عنه في المجتمع الذي سحبت منه هذه العينة، فالانحراف المعياري المحسوب من عينة عشوائية سحبت من مجتمع إحصائي تكون أصغر من الانحراف المعياري لهذا المجتمع الذي سحبت منه هذه العينة، ويزداد هذا الفرق كلما صغرت حجم العينة.
5- لا يمكن حسابه في الجداول التكرارية إذا كان أحد فئاتها مفتوحاً، لأن صيغته تشترط معرفة مراكز كل الفئات في التوزيع التكراري .

مسائل وتقارين الفصل الرابع والخامس

1 - حدد مقياس النزرعة المناسب للحالات الآتية :

- I - معرفة الحالة الاجتماعية لأغلبية عمال القطاع الخاص في دولة ما .
- II - معرفة معدل تخرج أحد الطلاب خلال السنوات الأربع التي قضاها في كليته .
- III - معرفة المعدل العام للتخرج بحيث أن 50% من المتخرجين تكون معدلهم أكبر من هذا المعدل العام .
- IV - وسطي معدل النمو لإنتاج القطن في سورية من عام 1990 وحتى عام 2002 .
- V - معرفة متوسط أسعار الكغ الواحد من البطاطا خلال ثلاث سنوات متعاقبة لشخص يستهلك كميات متفاوتة منها كل عام .
- VI - حساب كم من الوقت ينفق ثلاثة عمال في المتوسط لإنتاج سلعة واحدة إذا علمت أن كل عامل ينتج السلعة نفسها بوقت مختلف عن العامل الآخر .
- VII - معرفة مقدار الطول بحيث أن 10% من الأطوال هي أقل منه .
- VIII - معرفة مقدار الراتب بحيث أن 35% من الرواتب تكون أكبر منه .
- IX - تحديد سعر السلعة في محل تجاري ما بحيث 25% من السلع الموجودة في هذا المحل أسعارها أقل .

جاءت البيانات الآتية تمثل علامات عشرة طلاب في مادة ما :

الطالب :	1	2	3	4	5	6	7	8	9	10
العلامة :	10	25	30	20	37	85	65	83	75	55

والمطلوب : احسب المدى المطلق، الانحراف الربيعي، الانحراف المتوسط والانحراف المعياري مقارناً بينهم مع التعقيب .

3 - إذا علمت أن مجموع مربع انحرافات قيم X عن وسطه الحسابي يساوي 360 وأن مجموع مربع قيم X يساوي 2060 وعدد القيم يساوي عشرين.
احسب الوسط الحسابي.

4 - ليكن لدينا المعلومات الآتية التي تمثل إنتاج النفط الخام في إحدى الدول العربية خلال الفترة الزمنية 1995 - 1999 مقدراً بملايين الأطنان :

السنوات	كميات الإنتاج من النفط
1995	86,8
1996	97,2
1997	104,2
1998	114,56
1999	117,1

احسب متوسط معدل النمو لإنتاج النفط في هذه الدولة ؟

5 - ثلاث صيدليات تشتري المضادات الحيوية بأسعار مختلفة، والجدول الآتي يوضح الأسعار والكميات التي تشتريها كل صيدلية من ذلك الدواء:

الصيدلية	سعر العبوة (x) (ل.س)	كمية العبوات المشتراة (f)
1	100	50
2	125	40
3	200	25

احسب :

1- الوسط الحسابي الموزون بالكميات لمتوسط سعر العبوة الواحدة.

2- احسب الوسط التوافقي الموزون بالقيم (الكمية $\times$ السعر) لمتوسط سعر العبوة الواحدة .

3- احسب الوسط التوافقي الموزون بالكميات لمتوسط سعر العبوة الواحدة.

4- احسب الوسط التوافقي البسيط ، فسر النتائج ، وماذا تستنتج ؟

6 - البيانات الآتية تمثل أسعار الإقفال لأسهم شركتين من الشركات الكبرى في بورصة الأوراق المالية خلال أسبوع معين :

أيام الأسبوع	أسعار أسهم الشركة A	أسعار أسهم الشركة B
السبت	61	22
الأحد	13	25
الاثنين	32	26
الثلاثاء	25	24
الأربعاء	29	28
الخميس	40	25
Σ	200	150

والمطلوب :

1- قارن أسعار أسهم الشركتين مستخدماً الوسط الحسابي .

2- أي الشركتين لديها تجمانس أكثر في أسعار أسهمها الشركة A أم الشركة B ؟

3- إذا أردت شراء أسهم من إحدى الشركتين، أي الشركتين تحقق لك الطمأنينة المالية الشركة A أم B ؟

7 - البيانات الآتية تمثل مراكز الفئات لأجرة ساعة العمل بالليرة والتكرارات المناظرة في عينة من 25 عاملاً .

عدد العمال	أجرة ساعة العمل (x)
2	25
4	30
8	35
5	40
3	45
3	50
25	Σ

والمطلوب :

1- احسب الوسط الحسابي والوسيط والنوال لأجرة ساعة العمل لهذه العينة، ماذا تستنتج ؟

2- استخدم التكرار التجميعي الصاعد لإيجاد قيمة الوسيط .

3- هل يمكن استخدام قيمة الانحراف المعياري لإيجاد قيمة الانحراف المتوسط للبيانات السابقة ؟

8 - البيانات الآتية تبين توزيع عدد من شركات التجزئة بحسب المبيعات الشهرية بآلاف الليرات :

عدد الشركات	المبيعات الشهرية
25	5 - 10
30	10 - 15
41	15 - 20
49	20 - 25
35	25 - 30
30	30 - 35
20	35 - 40
10	40 - 45

المطلوب :

1- احسب معامل الاختلاف والانحراف الربيعي النسبي للمبيعات الشهرية. وماذا تعلق الاختلاف بين الرقمين ؟

2- استنتج قيمة المبيعات الأكثر شيوعاً .

3- أوجد من خلال رسم منحني التكرار التجميعي الصاعد كل من قيمة الوسيط، وعدد الشركات التي تقل مبيعاتها عن 32 ألف ليرة شهرياً .

9 - ليكن لدينا الجدول الآتي الذي يظهر توزيع عينة عشوائية من عمال معمل (أوبري) للأدوية إلى فئات بحسب الدخل اليومي لهم :

عدد العمال	فئات الدخل (ل.س)
8	200 - 250
24	250 - 300
30	300 - 350
40	350 - 400
30	400 - 450
20	450 - 500
8	500 - وأكثر

والمطلوب :

1- احسب معامل الاختلاف للسرعة المركزية والتشتت لبيانات الجدول السابق، ثم احسبهما .

2- بفرض أن : سحب عينة عشوائية من الحجم السابق نفسه من معدل (ابن زهر) للأدوية، لدراسة توزيع الدخل اليومي، فوجد أن الانحراف الربيعي النسبي يساوي 10%، أي العاملين توجد فيه عدالة أكبر في توزيع الدخل على عماله: معمل أوبري أو معمل ابن زهر؟ أثبت ذلك إحصائياً .

- 3- أوجد قيمة الوسط الحسابي التقريبية لمتوسط الدخل اليومي لعينة عمال معمل أوبري للأدوية باستعمال العلاقة بين المتوسطات الثلاثة، ثم أغلق الفقة الأخيرة بدلالة الفقة السابقة لها للجدول السابق، واحسب الوسط الحسابي لذلك الجدول، قارن بين الوسط الحسابي والوسط الحسابي التقريبي، ماذا تستنتج ؟
- 4- حدد طبيعة التوزيع الذي تخضع له بيانات الجدول السابق .
- 5- أوجد قيمة الدخل لعينة معمل أوبري بحيث 30% من العمال يتقاضون دخلاً يومياً أكبر منه.

- 10- بغية معرفة متوسط الزمن الذي يستغرقه الطالب في الإجابة على أسئلة الامتحان الموثمة تمت مراقبة أربع طلاب من طلاب السنة الأولى (أدخلوا بشكل عشوائي) في أثناء امتحانهم في إحدى المواد الموثمة فشكل الجدول الآتي :

الطالب	زمن الإجابة على سؤال واحد (بالدقائق)	عدد الأسئلة الجواب عليها في الساعة
A	1	60
B	2	30
C	1,5	40
D	5	12
Σ	9,5	142

والمطلوب :

احسب متوسط الزمن الذي يستغرقه الطالب الواحد للإجابة على سؤال واحد موثمت باستخدام متوسط مناسب. (الجواب) = (1,688) دقيقة .

- 11- اخضع مجموعتين مستقلتين من الطلاب لامتحانين في المقرر نفسه للامتحان الأول بطريقة الأتمتة، والامتحان الثاني بالطريقة التقليدية فتم الحصول على الجدول الآتي:

فئات العلامات	عدد الطلاب بالطريقة التقليدية	عدد الطلاب بالطريقة الموثمة
10 - 20	1	-
20 - 30	2	-
30 - 40	3	8
40 - 45	11	21
45 - 50	29	35
50 - 60	32	25
60 - 70	20	7
70 - 80	10	4
80 - 90	2	-
Σ	110	100

المطلوب :

- 1- احسب الوسط الحسابي لعلامات الطلاب المتحنيين بالطريقة التقليدية، وعلامات الطلاب بالطريقة الموثمة .
- 2- احسب الوسط الحسابي لعلامات الطلاب في هذا المقرر .
- 3- أنشئ جدولاً تكرارياً جديداً دون التفرقة بين الطلاب المتحنيين بالطريقة الموثمة والمتحنيين بالطريقة التقليدية، ثم احسب الوسط الحسابي والمنوال لعلاماتهم، ماذا تستنتج ؟
- 4- لو جعلنا أطوال الفئات في الجدول التكراري متساوية، احسب قيمة المنوال وقيمة الوسط الحسابي في الحالة الجديدة، ثم قارن قيمتهما في حال كانت أطوال الفئات غير متساوية. ماذا تستنتج ؟

5- أي الطلاب مستواهم متقارب أكثر؟ هل الطلاب المتبحرون بالطريقة التقليدية، أو المتبحرون بطريقة الأتمتة ؟ برهن ذلك إحصائياً .

12- الجدول الآتي يمثل توزيع مجموعة من الطلبة الناجحين في مقرر مدخل الإدارة تبعاً لعلاماتهم في ذلك المقرر :

أقل من 50 علامة	0	أقل من 56 علامة	289
أقل من 51 علامة	5	أقل من 57 علامة	333
أقل من 52 علامة	26	أقل من 58 علامة	361
أقل من 53 علامة	75	أقل من 59 علامة	383
أقل من 54 علامة	152	أقل من 60 علامة	400
أقل من 55 علامة	224		

المطلوب :

(أجراء الحسابات على أساس طول الفئة يساوي علامتين)

- 1- احسب معامل الاختلاف .
- 2- احسب الانحراف الربيعي النسبي .
- 3- احسب العلامة الأكثر شيوعاً لولاء الناجحين .



الفصل السادس

مقاييس الالتواء والتفلطح

Measures of Skewness and Kurtosis

- مفهوم الالتواء وأساليب قياسه .
- فوائد مقاييس الالتواء .
- مفهوم التفلطح وأساليب قياسه .
- فوائد مقاييس التفلطح .



6-1 مفهوم الالتواء وأساليب قياسه:

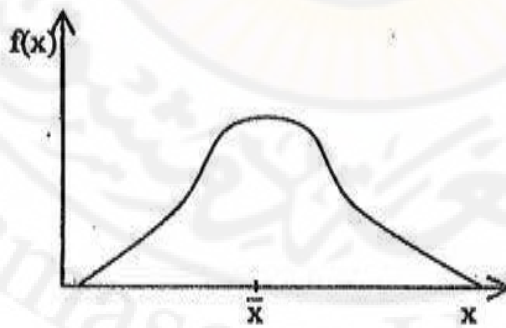
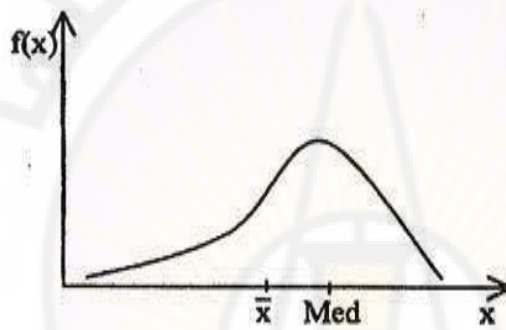
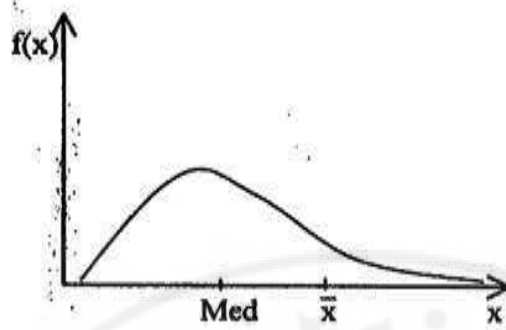
لقد درسنا المتوسطات ومقاييس التشتت، واستعرضنا طرائق العرض البياني جدولياً وبياناً، ولكن لم نتطرق عند دراسة المنحنيات التكرارية إلى مقدار التواء هذه المنحنيات، وإلى أية جهة ملتوية، علماً أن الالتواء هو بُعد المنحنى عن التماثل، فقد يكون الالتواء موجباً أي إلى اليمين أو سالباً إلى اليسار، فعندما يكون التوزيع ملتوياً نحو اليمين، فإن القيم المتطرفة نحو اليمين تؤثر في الوسط الحسابي وتسحبه نحو اليمين، وبذلك يكون الوسط الحسابي أكبر من الوسيط، أما إذا كان التوزيع ملتوياً نحو اليسار فإن الوسط يكون أصغر من الوسيط كما في الشكل (6-1) .

وفي حالة المنحنيات المتماثلة ينطبق الوسط الحسابي والوسيط والمنوال على بعضها تماماً، وكلما بعد المنحنى عن التماثل بعدت هذه القيم عن بعضها، وأصبحت الفروق بين هذه المقاييس الثلاثة تصلح لقياس التواء التوزيع، لذلك يفضل قياس الالتواء بمعامل يشمل هذا الفرق منسوباً إلى أحد مقاييس التشتت، وأهم هذه المقاييس أو المعاملات «معامل بيرسون» للالتواء وله الصيغتان الآتيتان :

$$SK = \frac{(\bar{x} - \text{Mod})}{S} \quad \text{الأولى}$$

$$SK = \frac{3(\bar{x} - \text{Med})}{S} \quad \text{والثانية}$$

حيث SK معامل الالتواء .



مجموعات الأشكال رقم (6-1)

يؤخذ على هذه الطريقة في حساب معامل الالتواء أنها تعتمد على المنوال، وهو مقياس غير دقيق لا يمكن الاعتماد عليه في قياس درجة الالتواء، لذلك يفضل استخدام الطريقة الثانية للتعبير عن معامل الالتواء، بأن يحسب الفرق بين الوسط الحسابي والمنوال بدلالة الوسيط بدلاً من المنوال، وذلك باستعمال العلاقة التقريبية التي مرت معنا سابقاً، والتي تربط بين المتوسطات الثلاثة ($\bar{x}$, med, Mod) في حالة التوزيعات المعتدلة الالتواء (القريبة من التماثل) وهي :

$$\bar{x} - \text{Mod} = 3 (\bar{x} - \text{Med})$$

إن معامل الالتواء وفق الطريقة الأولى يختلف بعض الشيء عن معامل الالتواء وفق الطريقة الثانية، وذلك عائد لكون العلاقة بين المنوال والوسيط هي علاقة تقريبية، بالإضافة إلى أن الطريقة الثانية لا تستخدم إذا كان المنحنى شديد الالتواء .

إن درجة الالتواء تتوقف على قيمة المعامل وإشارته الجبرية. فحينما يكون معامل الالتواء صغيراً فإن ذلك يدل على ضعف الالتواء، والعكس إذا كبر المعامل فإنه يدل على شدة الالتواء، ونشير إلى أنه حينما يكون الوسط الحسابي أكبر من المنوال، فإن معامل الالتواء يكون موجباً، والتواء التوزيع يكون نحو اليمين، أما حينما يكون الوسط الحسابي أصغر من المنوال، فإن معامل الالتواء يكون سالباً والتواء التوزيع يكون نحو اليسار، وإذا كان التوزيع متماثلاً، فإن الوسط الحسابي يتساوى مع المنوال، وحينئذ يكون مقياس الالتواء مساوياً للصفر، وكذلك معامل الالتواء، أما إذا كانت النتيجة أكثر من صفر، فإن ذلك يدل على وجود التواء في التوزيع .

بالإضافة إلى ما سبق يمكن استخدام طريقة العزم في إيجاد معامل الالتواء، عن طريق العزم المركزي من المرتبة الثالثة، حيث يعرف العزم المركزي من المرتبة k بشكل عام لمتغير إحصائي بالعلاقة :

$$\mu_{k(x)} = \frac{\sum (x_i - \bar{x})^k}{n}$$

أو في حالة وجود توزيع تكراري :

$$\mu_{k(x)} = \frac{\sum f_i (x_i - \bar{x})^k}{\sum f_i}$$

حيث μ_K العزم المركزي من المرتبة K .
ويقصد بالعزم متوسط الانحرافات عن قيمة ما، فإذا أخذنا الانحرافات عن
الوسط الحسابي أمكننا الحصول على العزم أو القوة الأولى كما يأتي :

$$\mu_{1(x)} = \frac{\sum f_i (x_i - \bar{x})}{\sum f_i} = 0$$

أي إذا كان $K = 1$ فإن العزم المركزي يساوي الصفر، وذلك حسب الخاصية
الأولى من خواص الوسط الحسابي .

أما إذا كان $K = 2$ فإن العزم المركزي يساوي التباين، ويسمى العزم أو القوة
التي نحصل عليها العزم أو القوة الثانية وهي الوسط الحسابي لمربع الانحرافات أي :

$$\mu_{2(x)} = S^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i} = 0$$

وإذا كان $K = 3$ فإن العزم المركزي من المرتبة الثالثة يساوي :

$$\mu_{3(x)} = \frac{\sum f_i (x_i - \bar{x})^3}{\sum f_i}$$

وإذا كان $K = 4$ فإن العزم المركزي من المرتبة الرابعة يكون :

$$\mu_{4(x)} = \frac{\sum f_i (x_i - \bar{x})^4}{\sum f_i}$$

وبصورة عامة فإن العزوم المركزية من المراتب الفردية للمتغير الإحصائي تكون
مساوية للصفر في حالة التوزيع المتناظر (بعد التوزيع متناظراً عندما تكون قيم التوزيع
تبتعد بشكل متناظر عن القيمة المتوسطة)، وذلك لوجود الإشارات الجبرية، لهذا فمن
المنطقي أن نختار أصغر العزوم المركزية الفردية المرتبة لقياس عدم التناظر في التوزيع أو
الالتواء بالنسبة للمنحنى التكراري. وبما أن $\mu_1(x) = 0$ لذلك اختير العزم المركزي من

المرتبة الثالثة كمقياس لواقع التناظر حول الوسط الحسابي معبراً عنه بوحدات معيارية،
ومنه :

$$\beta_1 = SK = \frac{\mu_{3(x)}}{\sqrt{\mu_{2(x)}^3}} = \frac{\mu_{3(x)}}{S_{(x)}^3}$$

$$\beta_1 = \frac{\frac{\sum f_i (x_i - \bar{x})^3}{\sum f_i}}{\left[\sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}} \right]^3}$$

- فإذا كانت $\beta_1 = 0$ يكون التوزيع التكراري متناظراً .
 $\beta_1 > 0$ يكون التوزيع التكراري ملتوياً نحو اليمين .
 $\beta_1 < 0$ يكون التوزيع التكراري ملتوياً نحو اليسار .

مثال (6-1) :

ليكن لدينا الجدول الآتي الذي يعطينا توزيع مجموعة من عمال شركة الصناعات
الكيمياوية حسب الأجر اليومي لهم .

جدول (6-1)

فئات الأجر	عدد العمال
[20 - 30]	7
[30 - 40]	20
[40 - 50]	38
[50 - 60]	25
[60 - 70]	10
Σ	100

والمطلوب : حساب معامل الالتواء بطريقة العزوم .

الحل : لحساب معامل الالتواء نقوم بإعداد الجدول المساعد الآتي :

حساب معامل الالتواء لتوزيع مجموعة من عمال شركة الصناعات الكيماوية
حسب الأجر اليومي لهم

جدول (6-2)

الفئات	f_i	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$f_i(x_i - \bar{x})^3$
[20 - 30]	7	25	- 21.1	445.21	3116.47	- 9393.931	- 65757.517
[30 - 40]	20	35	- 11.1	123.21	2464.20	- 1367.631	- 27352.620
[40 - 50]	38	45	- 1.1	1.21	45.98	- 1.331	- 50.578
[50 - 60]	25	55	+ 8.9	79.21	1980.25	+ 704.969	+ 17624.225
[60 - 70]	10	65	+ 18.9	357.21	3572.10	+ 6751.269	+ 67512.690
	100			1006.05	11179		- 8023.8

الوسط الحسابي :

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{4610}{100} = 46.1$$

العزم المركزي من المرتبة الثالثة :

$$\mu_{3(x)} = \frac{\sum f_i (x_i - \bar{x})^3}{\sum f_i} = \frac{-8023.8}{100} = -80.238$$

الانحراف المعياري :

$$S_x = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}} = \sqrt{\frac{11179}{100}} = 10.57$$

معامل الالتواء :

$$\beta_1 = \frac{\mu_{3(x)}}{S_x^3} = \frac{-80.238}{(10.57)^3} = -0.07$$

نلاحظ أن معامل الالتواء قريباً جداً من الصفر وهذا يعني أنه شبه متناظر .

2-6 فوائده مقياس الالتواء :

يمكن تلخيص فوائده معامل الالتواء كما يأتي :

- 1- معرفة نوعية التواء التوزيع التكراري، فإذا كان موجباً فيكون الالتواء نحو اليمين وإذا كان سالباً فإن الالتواء نحو اليسار وبالتالي $\bar{x} > Med$.
- 2- يستخدم بهدف المقارنة بين التواء توزيعين تكراريين، أو مجموعتين من البيانات حيث إن المجموعة التي يكون فيها معامل الالتواء أكبر يكون توزيعها ملتوياً أكثر من المجموعة الأخرى، والعكس صحيح أيضاً .

3-6 مفهوم التفلطح وأساليبه قياسه :

إن ما يميز التوزيعات التكرارية ومنحنياها هو مقدار التفلطح، وهو مقياس يحدد درجة علو أي منحنى تكراري أو انخفاضه بالنسبة للمنحنى الطبيعي، فإذا كان معامل التفلطح $(K = 3)$ فإن ذلك يعني أن المنحنى طبيعي، وإذا كان أقل من ثلاثة فإن ذلك يعدّ دليلاً على وجود انبساط أو تفلطح، أما إذا كان معامل التفلطح أكبر من ثلاثة فإن هذا يعني وجود حده في قمة المنحنى .

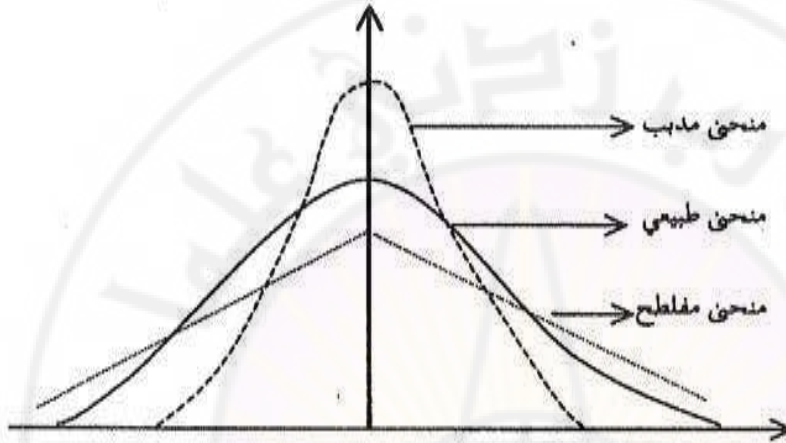
إن مقياس التفلطح يتعلق بمقدار تشتت القيم حول الوسط الحسابي، وتتخذ العزوم المركزية من المراتب الزوجية أساساً لهذا المقياس، وبما أن العزم المركزي من المرتبة الثانية يعطينا التباين، لذا نلجأ إلى العزم المركزي من المرتبة الرابعة، حيث نعرف معامل التفلطح بالعلاقة :

$$\beta_2 = K = \frac{\mu_{4(x)}}{\mu_{2(x)}^2} = \frac{\mu_{4(x)}}{S_{(x)}^4} = \frac{\frac{\sum f_i (x_i - \bar{x})^4}{\sum f_i}}{\left[\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i} \right]^2} = \frac{\frac{\sum f_i (x_i - \bar{x})^4}{\sum f_i}}{\left[\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i} \right]^2}$$

حيث β_2 معامل التفلطح وتأخذ القيم الآتية :

- $\beta_2 = 3$ إذا كان التوزيع التكراري طبيعياً .
- $\beta_2 > 3$ إذا كان التوزيع التكراري مديباً .
- $\beta_2 < 3$ إذا كان التوزيع التكراري مفلطحاً .

والرسم البياني (6-2) يوضح أشكال هذه المنحنيات .



الشكل (6-2) يبين بعض أشكال المنحنيات

بالإضافة إلى أن التفلطح لا يرتبط بالمتوسط الحسابي للتوزيع، فقد يكون التوزيع مسطحاً أو معتدلاً أو رفيعاً ويكون له متوسط حسابي محدد كما أن زيادة التفلطح أو نقصانه لا يتعارض مع تماثل التوزيع .

مثال (6-2) :

احسب معامل التفلطح من التوزيع التكراري الوارد في الجدول رقم (6-1) والمتعلق بتوزيع مجموعة من عمال شركة الصناعات الكيماوية حسب الأجر اليومي لهم.

الحل : نعلم أن الوسط الحسابي $(\bar{x} = 46.1)$ ، لذلك نقوم بإعداد الجدول المساعد لإيجاد القيم اللازمة .

جدول (6-3)

فئات الأجر	x_i	f_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$	$(x_i - \bar{x})^4$	$f_i(x_i - \bar{x})^4$
[20 - 30]	25	7	- 21.1	445.21	3116.47	198211.94	1387483.58
[30 - 40]	35	20	- 11.1	123.21	2464.20	15180.704	303614.08
[40 - 50]	45	38	- 1.1	1.21	45.98	1.4641	55.6358
[50 - 60]	55	25	+ 8.9	79.21	1980.25	6274.2241	156855.6025
[60 - 70]	65	10	+ 18.9	357.21	3572.10	127598.98	1275989.8
Σ					11179		3123998.697

لايجاد معامل التفلطح نحسب العزم المركزي من المرتبة الرابعة :

$$\mu_{4(x)} = \frac{\Sigma f_i (x_i - \bar{x})^4}{\Sigma f_i} = \frac{3123998.697}{100} = 31239.98697$$

$$\mu_{2(x)}^2 = S_{(x)}^4 \left[\frac{\Sigma f_i (x_i - \bar{x})^4}{\Sigma f_i} \right]^4 = \left[\frac{11179}{100} \right]^4 = 12497.004$$

أو :

$$\mu_{2(x)}^2 = S_{(x)}^4 \left[\frac{\Sigma f_i (x_i - \bar{x})^2}{\Sigma f_i} \right]^2 = \left[\frac{11179}{100} \right]^2 = 12497.004$$

$$\beta_2 = \frac{\mu_{4(x)}}{\mu_{2(x)}^2} = \frac{31239.98694}{12497.004} = 2.49979$$

أي أن التوزيع مفلطح لأن $\beta_2 < 3$.

6-4 فوائد مقياس التفلطح :

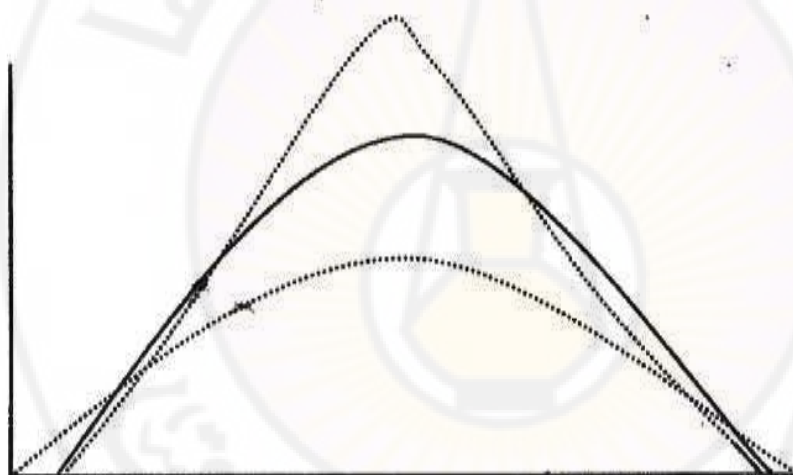
يمكن تلخيص فوائد معامل التفلطح كما يأتي :

- 1- إن معامل التفلطح يبين فيما إذا كان للتوزيع قمة عريضة مسطحة، أو قمة حادة رفيعة، حيث يطلق على التوزيع ذي القمة العريضة المسطحة اسم التوزيع

مسطح التفلطح، أما إذا كان التوزيع ذا قمة حادة رفيعة فيطلق عليه اسم التوزيع مذهب التفلطح .

2- بمعرفة قيمة معامل التفلطح يمكننا التعرف على شكل التوزيع سواء أكان مسطحاً أم معتدلاً أم رفيعاً، حيث إن زيادة التفلطح أو نقصانه لا يتعارض مع تماثل التوزيع.

3- بالإضافة لما سبق نلاحظ أن التفلطح ليست له علاقة بالمتوسط الحسابي للتوزيع، حيث قد يكون التوزيع مفلطحاً أو رفيعاً أو مسطحاً، ويكون له متوسط حسابي محدد كما هو مبين بالشكل (6-3) .



الشكل (6-3) يبين اختلاف التفلطح ومتوسط حسابي موحد

تمارين ومسائل الفصل السادس

- 1- احسب مقاييس الالتواء ومقاييس التفلطح للبيانات الآتية التي تمثل توزيع أوزان 50 طالباً من طلاب كلية الاقتصاد بجامعة دمشق .

فئات الوزن	50-60	61-63	64-66	67-69	70-72	73-75
عدد الطلاب	2	7	14	15	8	4

- 2- الجدول الآتي يمثل دخل مجموعة من الأسر بمئات الدولارات :

فئات الدخل	أقل من 10	10-14	15-19	20-24	25-29	أكثر من 30
عدد الأسر	5	20	35	19	13	8

والمطلوب : حساب معاملي الالتواء والتفلطح لهذه البيانات وتفسير النتائج .

- 3- ليكن لدينا التوزيع التكراري لعلامات 250 طالباً في مادة مبادئ الإحصاء في كلية الاقتصاد بجامعة دمشق :

فئات العلامات	العدد الفعلي أو التكرار المطلق f_i
[10 - 20 [	14
[20 - 30 [	16
[30 - 40 [	20
[40 - 50 [	29
[50 - 60 [	40
[60 - 70 [	56
[70 - 80 [	35
[80 - 90 [	23
[90 - 100 [	17
Σ	250

والمطلوب :

حساب معامل الالتواء ومعامل التفلطح لتوزيع علامات الطلاب في مادة مبادئ الإحصاء وتفسير النتائج .

4- سحبت عينة عشوائية من مرضى إحدى المستشفيات حجمها (40) مريضاً لدراسة تعداد الكريات الحمراء في الدم، فكانت النتائج كما هو موضح في الجدول الآتي:

عدد الحالات	فئات تعداد الكريات الحمراء (بالملايين)
6	[1.5 – 2.5]
10	[2.5 – 3.5]
15	[3.5 – 4.5]
9	[4.5 – 5.5]

والمطلوب :

1- حساب معامل الالتواء وتفسير النتائج، ورسم منحني التوزيع وتحديد جهة الالتواء .

2- حساب معامل التفلطح وتفسيره، وبيان طبيعة قمة التفلطح .

الفصل السابع

الاحتمالات وبعض التوزيعات الاحتمالية

Probability and Probability Distribution

- الاحتمالات .

مقدمة .

بعض المفاهيم الأساسية في نظرية الاحتمالات.

العمليات على الحوادث .

التعاريف الأساسية للاحتمال .

التعريف الكلاسيكي أو التقليدي للاحتمال.

التعريف الاحصائي أو التكرار النسبي للاحتمال.

الاحتمال حسب مسلمات كوجوروف .

القوانين الأساسية في الاحتمالات.

- التوزيع الطبيعي

مقدمة.

تابع الكثافة الاحتمالي للتوزيع الطبيعي .

- التوزيع الطبيعي المعياري.

تابع الكثافة الاحتمالي للتوزيع الطبيعي المعياري .

تحديد قيم Z باحتمال معلوم.



7-1 الاحتمالات

7-1-1 مقدمة :

تعد نظرية الاحتمال أو قوانين المصادفة (Laws of chances) من النظريات الرياضية المهمة التي تخصص في دراسة الحوادث، أو الظواهر التي تتميز بعدم تأكيد حدوثها.

ولا تبرز أهمية نظرية الاحتمال لتطبيقاتها المهمة في شتى المجالات فحسب، بل إن هناك علوماً ونظريات أصبحت تركز أساساً على نظرية الاحتمال، ومنها علم الإحصاء (Statistics) ونظرية القرار (Decision theory) ونظرية الطوابير (Queuing theory) ونظرية المعلومات (Information theory)، ولقد دخلت نظرية الاحتمال بتطبيقاتها المختلفة في شتى المجالات التقنية، والهندسية، والاجتماعية، والاقتصادية، والعسكرية.

7-1-2 بعض المفاهيم الأساسية في نظرية الاحتمال

1- التجربة العشوائية (Random experiment) :

هي كل عملية تؤدي إلى تسجيل ملاحظة، أو قياس ظاهرة ما، علماً بأن جميع النتائج الممكنة للتجربة العشوائية هي معروفة مسبقاً (أي قبل إجرائها) ، وإن كنا لا

نستطيع أن نحدد أي هذه النتائج سيتحقق في محاولة معينة، وهي تخضع لعامل المصادفة فقط.

أمثلة على التجارب العشوائية :

أ- إلقاء قطعة نقود عادية وملاحظة أي وجه سوف يظهر، ومن الواضح أننا نعلم مسبقاً جميع النتائج الممكنة وهي: صورة (H) وكتابة (T) .

ولكن لا يمكن معرفة نتيجة رمية واحدة قبل إجرائها فهي تجربة عشوائية.

ب- إلقاء حجر النرد (زهرة الطاولة) مرة واحدة، وملاحظة العدد على الوجه الذي يظهر... نعلم أن جميع النتائج الممكنة هي الأرقام (1. 2. 3. 4. 5. 6) ولكن لا يمكن تحديد أي هذه الأرقام سيظهر بالفعل من خلال رمية واحدة وبالتالي فهي تجربة عشوائية.

ج- فصيلة الدم عند شخص ما .

د- نتيجة امتحان طالب بمادة الإحصاء .

وهذه التجارب جميعها تجارب عشوائية، وهي على العكس من التجارب التي يمكن التحكم بنتائجها، مثل وضع ذرة من الأكسجين وذرتين من الهيدروجين في ظروف معينة، فالنتائج هو الماء، ومثل هذه التجربة تدعى بالتجربة العملية وهي غير عشوائية .

2- فضاء العينة (Sample Space) :

يعرف فضاء العينة المرتبط بتجربة عشوائية، بأنه مجموعة جميع النتائج الممكنة لهذه التجربة ويرمز له عادة بـ (S) أو (Ω) وسوف نستخدم الرمز $N(S)$ للتعبير عن عناصر فضاء العينة (S) .

أمثلة على فضاءات العينة لتجارب مختلفة :

أ- التجربة العشوائية «إلقاء حجر النرد»

الفضاء : $N(S) = \{ 1, 2, 3, 4, 5, 6 \}$

ب- التجربة العشوائية «إلقاء قطعة نقد في الهواء»

الفضاء : $N(S) = \{ H, T \}$

ج- التجربة العشوائية «إلقاء قطعتي نقود الواحدة بعد الأخرى»

الفضاء : $N(S) = \{ HH, HT, TH, TT \}$

د - التجربة العشوائية «تحديد جنس المولود عند الولادة»

الفضاء : $N(S) = \{ \text{أنثى} , \text{ذكر} \}$

3- الحوادث (Events) :

الحادثة : هي أي مجموعة جزئية من فضاء العينة (S) ورياضياً يقال A حدث من S ويكتب S A ، ونرمز عادة للحوادث بالأحرف اللاتينية الكبيرة A, B, C,

أمثلة على الحوادث :

أ- إذا كانت A هي حادثة حصولنا على رقم زوجي عند رمي حجر النرد فإن:

$$A = \{ 2, 4, 6 \}$$

وهي مجموعة جزئية من (S) أي: $A \subset S$ وبالتالي فإن A حادثة معرفة على (S) .

ب- إذا كانت B هي حادثة حصولنا على الوجه صورة عند رمي قطعة نقد فإن:

$$B = \{ H \}$$

وهي حادثة في (S) .

ومن الضروري الإشارة أن الحوادث تقسم إلى الأنواع الآتية :

أ- الحوادث البسيط (Elementary event) ويعرف: بأنه الحادث الذي يحتوي على عنصر واحد من عناصر فضاء العينة .

ب- الحوادث المركب (Compound event) ويعرف: بأنه الحادث الذي يحتوي على عنصرين أو أكثر من عناصر فضاء العينة (S) .

جـ - الحادث الأكيد (Sure event) ويعرف: بأنه الحادث الذي يظهر دائماً عند إجراء التجربة.

مثال: حادثة ظهور رقم أقل من (7) عند رمي حجر النرد هي حادثة أكيدة.
د - الحادث المستحيل (Impossible event): وهو الحادث الذي لا يمكن أن يظهر أبداً ، ويرمز له بـ (ϕ) .

مثال: حادثة حصولنا على الوجه (7) عند رمينا حجر النرد عدداً كبيراً من المرات.

هـ - الحادث العشوائي (Random event): وهو الحادث الذي لا نستطيع أن نؤكد حدوثه أو ننفيه .

مثال: حصولنا على الوجه 5 عند رمينا لحجر النرد، فلا نستطيع الجزم بأننا سنحصل على الوجه (5) ولا نستطيع أن نؤكد استحالة حصولنا على الوجه 5 فهو حادث عشوائي.

ولما كنا نعرف سلفاً نتائج الحوادث (الأكيدة والمستحيلة)، فإننا لا نلتم بدراستها، أما الحوادث العشوائية، فإنه لا يمكن معرفة نتائجها معرفة قاطعة قبل حدوثها، لهذا فإننا في نظرية الاحتمال نلتم بدراسة الحوادث العشوائية، لكي نتمكن بشكل احتمالي من التكهّن بنتائجها ومعرفة القوانين المتحكممة بحدوثها.
وبعبارة أخرى فإن نظرية الاحتمال، هي نظرية هدفها دراسة الحوادث العشوائية وتقدير احتمال حدوثها.

و - الحادثة الجزئية (Sub event): يقال للحادثة (A) بأنها حادثة جزئية إذا كانت (A) محتواة بكامل عناصرها في (B) وليس بعض عناصرها وتكتب $A \subset B$ ، وهذا يعني أن حدوث B يعني ضمناً حدوث A والعكس غير صحيح.
مثال: إذا عرفنا (A) بأنها حادثة حصولنا على رقم فردي في أثناء رمينا لحجر النرد أي:

$$A = \{ 1, 3, 5 \}$$

وعرفنا (B) بأنها حادثة حصولنا على رقم أصغر من 6 في أثناء رمينا لحجر النرد

أي:

$$B = \{ 1, 2, 3, 4, 5, \}$$

نلاحظ هنا أن كل عناصر A هي موجودة في B ، إذا نقول إن A هي مجموعة

جزئية من B وإن $A \subset B$.

وإذا وجد عنصر واحد على الأقل من A غير موجود أو لا ينتمي إلى B فإن A

تكون غير محتواة في B أي : $A \not\subset B$ ، وبالتالي ليست مجموعة جزئية منها، وذلك بغض

النظر عن بعض العناصر المشتركة .

مثال (7-1) :

إذا رمينا قطعة نقد سليمة في الهواء وعرفنا على فضاء العينة (S) الحرف (H)

حادثة حصولنا على الوجه صورة والحرف (T) حادثة حصولنا على الوجه كتابة .

والمطلوب: عرّف الحوادث الآتية :

أ - حصولنا على صورة.

ب - حصولنا على كتابة .

ج- حصولنا على صورة أو كتابة .

الحل :

$$A_1 = \{ H \}$$

$$A_2 = \{ T \} , \quad S = \{ H, T \}$$

مثال (7-2) :

في تجربة رمي حجر النرد عرّف الحوادث الآتية على فضاء العينة (S) :

أ - حصولنا على عدد فردي .

ب - حصولنا على عدد زوجي .

ج- حصولنا على عدد أقل من (3).

د - حصولنا على عدد يقبل القسمة على (3) .

الحل :

$$A_1 = \{ 1, 3, 5 \}$$

$$A_2 = \{ 2, 4, 6 \}$$

$$A_3 = \{ 1, 2 \}$$

$$A_4 = \{ 3, 6 \}$$

مثال (7-3) :

في تجربة قياس عمر جهاز إلكتروني فضاء عينة $S = \{t : t \geq 0\}$ والمطلوب: عرف الحوادث الآتية على فضاء العينة :

أ - أن لا يقل عمر الجهاز الإلكتروني عن 10 سنوات .

ب - أن يقل عمر الجهاز عن 3 سنوات .

ج - أن يتراوح عمر الجهاز بين (8 - 12) سنة .

الحل :

$$A_1 = \{t : t \geq 10\}$$

$$A_2 = \{t : t < 3\}$$

$$A_3 = \{t : 8 \leq t \leq 12\}$$

مثال (7-4) :

ألقى قطعتي نقود الواحدة بعد الأخرى والمطلوب :

أولاً: اكتب فضاء العينة .

ثانياً: اكتب كل من الأحداث الآتية :

(1) حادث ظهور صورة على الأقل .

(2) حادث ظهور صورة واحدة فقط .

(3) حادث ظهور صورة أو كتابة وليس كلاهما معاً .

الحل :

$$S = \{ HH, HT, TH, TT \}$$

أولاً :

ثانياً :

$$A_1 = \{ (HH), (HT), (TH) \} \quad (1)$$

$$A_2 = \{ (HT), (TH) \} \quad (2)$$

$$A_3 = \{ (HH), (TT) \} \quad (3)$$

مثال (5-7) :

ألقيت ثلاث قطع من النقود مميزة بأن واحد والمطلوب :

أولاً: اكتب فراغ العينة .

ثانياً : اكتب كلاً من الحوادث الآتية :

(1) حادث ظهور صور متتالية .

(2) حادث ظهور نتائج متشابهة .

(3) حادث ظهور صورة واحدة بالضبط .

الحل :

$$S = \{ (HHH), (HHT), (HTH), (THT), (TTH), (TTH), (THT) \} - \text{أولاً :}$$

$$(TTH), (TTH), (THT) \}$$

ثانياً :

$$A = \{ (HHH), (HHT), (THH) \} \quad (1)$$

$$B = \{ (HHH), (T.T.T) \} \quad (2)$$

$$C = \{ (HTT), ((TTH), (THT)) \} \quad (3)$$

مثال (6-7) :

ألقيت حجرين نرد مميزين بأن واحد والمطلوب :

أولاً: اكتب فضاء العينة .

ثانياً : اكتب كلاً من الحوادث الآتية :

(1) ظهور العدد (5) على أحد الترددين .

(2) حادث مجموع الوجهين ≤ 10 .

(3) حادث مجموع العددين ≥ 4 مع اختلاف العددين .

الحل :

أولاً :

$$S = \left\{ \begin{array}{l} (1, 1), (1, 2), \dots, (1, 6) \\ (2, 1), (2, 2), \dots, (2, 6) \\ (3, 1), \dots, (3, 6) \\ (4, 1), \dots, (4, 6) \\ (5, 1), \dots, (5, 6) \\ (6, 1), \dots, (6, 6) \end{array} \right\}$$

ثانياً :

$$A = \{ (5, 1), (5, 2), (5, 3), (5, 4), (5, 6), (1, 5), (2, 5), (3, 5), (6, 5) \} \quad (1)$$

$$B = \{ (4, 6), (6, 4), (5, 5), (5, 6), (6, 5), (6, 6) \} \quad (2)$$

$$C = \{ (1, 2), (1, 3), (2, 1), (3, 1) \}$$

3-1-7 العمليات على الحوادث

1- تكافؤ حادثين (Equivalence) :

نقول عن حادثين A و B بأنهما متكافئان، إذا تحققت العلاقة الآتية :

$$A \subset B \text{ و } B \subset A$$

وتكتب على الصورة $A = B$

2- تقاطع حادثين (Intersection) :

نرمز لتقاطع الحادثين A و B بالرمز $(A \cap B)$ وهو الحادث الذي يتحقق إذا تحقق كل من A و B معاً، أو هو حادث وقوع A و B في الوقت نفسه .

مثال (7-7) :

إذا قمنا بتجربة اختيار طالب من طلبة الكلية وعرفنا الحوادث الآتية :

$A = \{ \text{أن يكون الطالب من قسم المحاسبة} \}$

$B = \{ \text{أن يكون الطالب من طلاب السنة الثالثة} \}$

فإن الحادث $(A \cap B)$ يقع إذا كان الطالب المختار هو من قسم المحاسبة ومن طلاب السنة الثالثة .

مثال (7-8) :

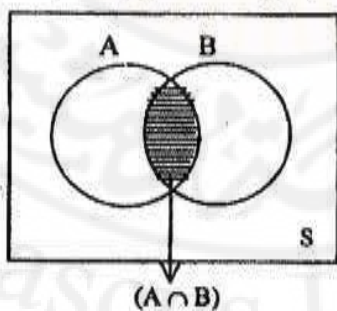
إذا عرفنا الحادثين A و B على النحو التالي :

$A = \{ \text{حصولنا على الورقة التي تحمل الرقم 10 من ورق اللعب} \}$

$B = \{ \text{حصولنا على ورقة ديناري من ورق اللعب} \}$

فإن الحادث $(A \cap B)$ يقع إذا كانت الورقة المختارة تحمل الرقم 10 ومن نوع الديناري حصراً .

ويوضح الشكل رقم (7-1) من أشكال فن (Venn) منطقة التقاطع .



الشكل رقم (7-1)

3- اتحاد حادثين (Union) :

يرمز لاجتماع حادثين A و B أو اتحادهما بالرمز $(A \cup B)$ وهو الحادث الذي يتحقق إذا تحقق أو حدث A أو B أو كلاهما معاً .

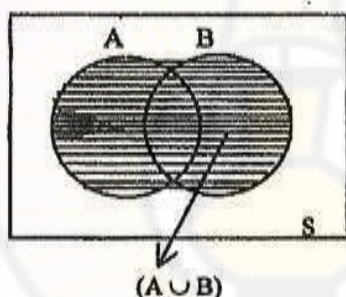
مثال (7-9) :

في تجربة اختبار طالب من طلاب كلية التجارة نعرف الحوادث الآتية :

$A = \{ \text{أن يكون الطالب المختار من طلاب قسم المحاسبة} \}$

$B = \{ \text{أن يكون الطالب المختار من طلاب السنة الثالثة} \}$

فإن الحادث $(A \cup B)$ يقع إذا كان الطالب المختار من طلاب قسم المحاسبة أو من طلاب السنة الثالثة ، أو من طلاب السنة الثالثة ومن قسم المحاسبة .
ويوضح الشكل (7-2) من أشكال فن منطقة الاتحاد .



الشكل رقم (7-2)

مثال (7-10) :

إذا عرفنا الحوادث الآتية أثناء سحبنا لورقة من ورق اللعب :

$A = \{ \text{الورقة المختارة تحمل الرقم 10} \}$

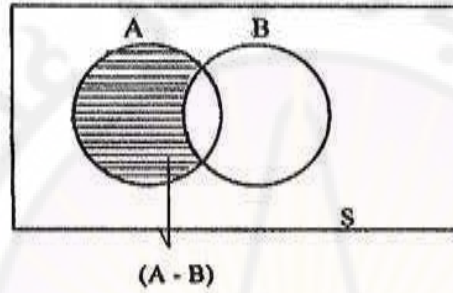
$B = \{ \text{الورقة المختارة من نوع الديناري} \}$

فإن الحادث $(A \cup B)$ يقع إذا كانت الورقة المختارة تحمل الرقم 10 أو من نوع الديناري أو تحمل الرقم 10 ومن نوع الديناري .

4- الفرق بين حادثين (Difference) :

نرمز للفرق بين حادثين A و B بالرمز $(A - B)$ أو (A/B) وهو الحادث الذي يتحقق إذا تحقق A ولم يتحقق B ويشير الشكل رقم (7-3) من أشكال فن إلى منطقة الفرق بين A و B وإن :

$$\begin{aligned} A/B &= A - (A \cap B) = A \cap \bar{B} \\ B/A &= B - (B \cap A) = B \cap \bar{A} \end{aligned}$$



الشكل رقم (7-3)

مثال (7-11) :

إذا عرفنا الحوادث الآتية على مجموعة الأعداد الطبيعية N :

$$A = \{ 1, 3, 7, 9 \}$$

$$B = \{ 5, 4, 3, 2 \}$$

$$A/B = \{ 1, 7, 9 \}$$

فإن الحادث A/B

$$B/A = \{ 5, 4, 2 \}$$

والحادث B/A

مثال (7-12) :

إذا عرفنا الحوادث الآتية في أثناء اختيارنا لطلاب من طلاب الكلية :

$$A = \{ \text{أن يكون الطالب المختار من قسم المحاسبة} \}$$

$$B = \{ \text{أن يكون الطالب المختار ذكراً} \}$$

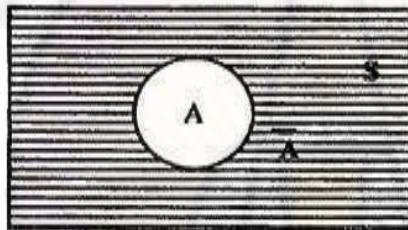
فإن الحادث A/B يتحقق إذا كان الطالب المختار من قسم المحاسبة وأنتى .

5- الحادث المتمم أو المكمل (Complement) :

الحادثة المتممة أو المكمل للحادثة A هي الحادثة التي تتكون من عناصر (S) التي لا تنتمي إلى A ونرمز لها بالرمز $\bar{A}$ أو هي حادثة عدم وقوع A المعرفة على (S) وإن:

$A \cup \bar{A} = S$
$A \cap \bar{A} = \phi$

وتشير المنطقة المظلمة في الشكل (7-4) إلى متممة أو مكمل الحادثة A المعرفة على (S) .



الشكل رقم (7-4)

مثال (7-13) :

إذا كانت الحادثة A في رمي حجر النرد ممثلة بالعلاقة :

$$A = \{ X : X \geq 3 \}$$

فإن متمم الحادثة A مساو :

$$\bar{A} = \{ X : X < 3 \}$$

6- الحوادث المتنافية (Mutually exclusive) :

نقول عن حادثين A و B أنهما متنافيان إذا كان احتمال وقوع أحدهما ينفي احتمال وقوع الآخر، أي بمعنى استحالة وقوعهما معاً أي :

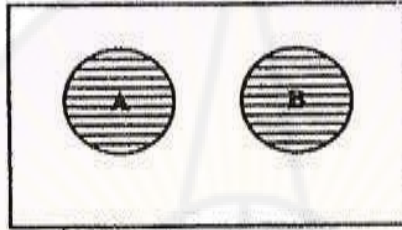
$$A \cap B = \phi$$

مثال (7-14) :

إذا رمزنا بـ A لحادثة ظهور الصورة في أثناء رمينا لقطعة النقد، و B لحادثة ظهور الكتابة وبالتالي فإن ظهور الصورة ينفي ظهور الكتابة في رمية واحدة أي :

$$A \cap B = \phi$$

وإذا رمزنا أيضاً بـ A لحادثة أن يكون لشخص ما الزمرة الدموية O^+ وللحادثة B أن يكون لشخص ما الزمرة الدموية O^- وبالتالي فإن $A \cap B = \phi$.



الشكل رقم (7-4)

7- الحوادث الشاملة (Exhaustive events) :

نقول عن الحوادث $A, B, C, \dots$ بأنها حوادث شاملة إذا حققت الشروط الآتية:

أ- إذا كانت إحداها واجبة الحدوث ولا تساوي ϕ .

ب- إذا كان اتحادها $A \cup B \cup C = S$.

ج- إذا كانت الأحداث متنافية مثنى مثنى.

مثال (7-15) :

إن الحوادث A و $\bar{A}$ حادثتان شاملتان لأن :

أ - أحدهما ستقع حتماً .

ب - $A \cap \bar{A} = \phi$

ج - $A \cup \bar{A} = S$

مثال: إذا عرفنا على فراغ العينة في رمية حجر النرد الحادث A الذي يشير إلى الأعداد الزوجية والحادث B الذي يشير إلى الأعداد الفردية، فإن الحوادث A و B هي حوادث شاملة.

7-1-4 التعاريف الأساسية للاحتمال :

7-1-4-1 التعريف الكلاسيكي أو التقليدي للاحتمال :

لنفرض أننا قمنا برمي قطعة نقود ثلاث مرات متتالية الواحدة بعد الأخرى، فإن فضاء العينة لهذا المثال هو :

$$S = \{ (HHH), (HHT), (HTH), (HTT), (THH), (THT), (TTH), (TTT) \}$$

نلاحظ من خلال هذا المثال ما يأتي:

أ- إن فضاء العينة هو منتهى ومعروف مسبقاً، ويساوي إلى 8 نتائج مختلفة الواحدة عن الأخرى.

ب- إن احتمال تحقق أي نتيجة أو حالة هو متساوٍ بالنسبة لجميع الحالات ويساوي $\frac{1}{8}$ أي أن لكل عنصر من عناصر فضاء العينة نفس احتمال الحدوث .

ج- إن هذا النموذج، هو نموذج نظري، أو مثل أعلى، قد لا يكون له ما يناظره في الحياة الواقعية، لكنه يوفر معياراً مفيداً للمقارنة مع تعاريف أخرى للاحتمال.

- إذا عرفنا الحادث (A) بالنسبة لمثالنا السابق بأنه حادثة حصولنا على وجه الصبورة في الرميات الثلاث (الحالة المرغوبة أو الحالة الملائمة)، فإن احتمال حدوث A ونرمز له بـ $P(A)$ حسب التعريف الكلاسيكي أو التقليدي يساوي إلى النسبة بين عدد الحالات الملائمة أو المرغوبة إلى عدد الحالات الممكنة، أو النسبة بين عدد عناصر الحادث A وعدد عناصر فضاء العينة (S) بشرط أن يكون A و B محققين أي :

$$P(A) = \frac{\text{عدد الحالات المرغوبة أو الملائمة}}{\text{عدد الحالات الممكنة}} = \frac{1}{8}$$

مواصفات الاحتمال النظري أو الكلاسيكي :

$$0 \leq P(A) \leq 1 \quad (1)$$

(2) احتمال تحقق الحادث (A) ندعوه باحتمال النجاح ونرمز له بـ $P(A)$.

(3) احتمال عدم تحقق الحادث المرغوب A ندعوه باحتمال الفشل ونرمز له بـ

$P(\bar{A})$ وأن :

$$P(\bar{A}) = 1 - P(A)$$

$$P(A) = 1 - P(\bar{A}) \quad (4)$$

$$P(A) + P(\bar{A}) = P(S) = 1 \quad (5)$$

عيوب التعريف الكلاسيكي :

1- يشترط تطبيق التعريف الكلاسيكي للاحتمال أن يكون فضاء العينة منتهياً.

2- يشترط أن يكون لجميع عناصر فضاء العينة احتمال الحدوث نفسه.

وقد يصادف أن لا يتحقق أحد هذين الشرطين أو كلاهما، فقد يكون فضاء العينة غير منته، كما في تجربة قياس عمر جهاز إلكتروني، أو لا يكون لجميع عناصر فضاء العينة احتمال الحدوث نفسه.

7-1-4-2 التعريف الإحصائي أو التكرار النسبي للاحتمال :

إذا كانت A حادثة معرفة على (S) المتعلق بتجربة عشوائية لمجتمع إحصائي فضاء عينته غير منته ، أو أن حجم المجتمع كبير ، أو إذا كانت عناصر فضاء العينة غير متساوية احتمال الحدوث، ولحساب الاحتمال بحسب التعريف الإحصائي نقوم بإجراء التجربة N من المرات، ولنفرض أن الحادثة المرغوبة A قد وقعت n مرة، فإن الاحتمال التجريبي أو الإحصائي لحدوث الحادثة A يعرف بأنه عدد مرات حدوث الحادثة A منسوبة إلى العدد الكلي من المرات التي أعيدت فيها التجربة أي أن :

$$P(A) = \frac{\text{عدد الحالات التي تحققت فيه الحادثة المرغوبة}}{\text{عدد مرات إجراء التجربة}} = \frac{n}{N}$$

عيوب التعريف الإحصائي (التكرار النسبي) :

1- إن قيمة $P(A)$ هي قيمة تقريبية ومتغيرة ومرتبطة بعدد مرات إجراء التجربة N ، وأنه لا توجد قيمة لـ N (مثلاً) بحيث عندما تصلها N تكون النسبة $\frac{n}{N}$ ثابتة لجميع قيم N التي هي أكبر من N_0 وتعد الإشارة إلى أنه لا يمكننا معرفة قيمة الاحتمال حسب هذا التعريف إلا بعد إجراء التجربة .

7-1-4-3 الاحتمال حسب مسلمات كولجوروف :

تعتبر مسلمات كولجوروف من أكثر تعاريف الاحتمال رصانة وهي تنص على ما يأتي :

يقابل كل حادثة (A) في فضاء العينة (S) ، عدد حقيقي R يسمى احتمال حدوث A ويرمز له بـ $P(A)$ بحيث يحقق المسلمات التالية :

$$P(S) = 1 \quad \text{المسلمة الأولى :}$$

$$P(\phi) = 0$$

المسلمة الثانية : من أجل كل حادث (A) ينتمي إلى S يكون لدينا :

$$0 \leq P(A) \leq 1$$

المسلمة الثالثة : لأي حادثين B و A المتنافيتين فإن :

$$P(A \cup B) = P(A) + P(B)$$

ويمكن تعميم المسلمة الثالثة لأكثر من حادثين، فإذا كانت لدينا الحوادث المتنافية:

$$A_1, A_2, \dots, A_k, \dots$$

فإن :

$$P(A_1 \cup A_2 \cup \dots \cup A_k \cup \dots) = P(A_1) + P(A_2) + \dots + P(A_k) + \dots$$

7-1-5 القوانين الأساسية في الاحتمالات :

1- احتمال الحادثة المستحيلة هو الصفر:

$$A(\phi) = 0$$

$$S = S \cup \phi$$

حيث ϕ و S حادثتان متنافيتان ، وحسب المسلمة الثالثة من مسلمات

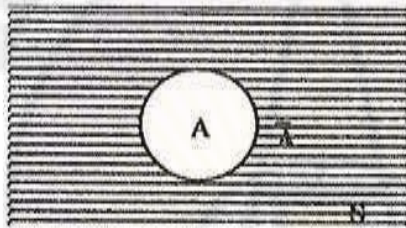
كولموغوروف فإن :

$$P(S) = P(S) + P(\phi)$$

$$P(\phi) = P(S) - P(S) = 0$$

2- لأي حادثة (A) فإن :

$$P(\bar{A}) = 1 - P(A)$$



الشكل رقم (7-6)

البرهان: من شكل فن نجد أن :

$$S = A \cup \bar{A}$$

وحيث إن A و $\bar{A}$ حادثتان متنافيتان ونحسب المسلمة الثالثة نجد :

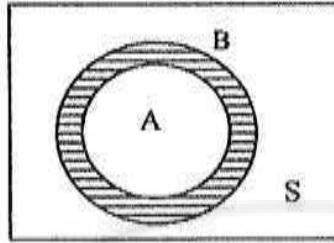
$$P(S) = P(A) + P(\bar{A})$$

$$P(\bar{A}) = 1 - P(A)$$

3- لأي حادثتين A و B فإن :

$$P(A) \leq P(B)$$

إذا كانت $A \subset B$



الشكل رقم (7-7)

البرهان: من الشكل رقم (7-7) نجد أن :

$$B = A \cup (B - A)$$

وبما أن A و B-A حادثتان متنافيتان، وحسب المسلمة الثالثة فإن :

$$P(B) = P(A) + P(B - A)$$

وبما أن $P(B - A) \neq \phi$

$$P(A) \leq P(B)$$

إذاً

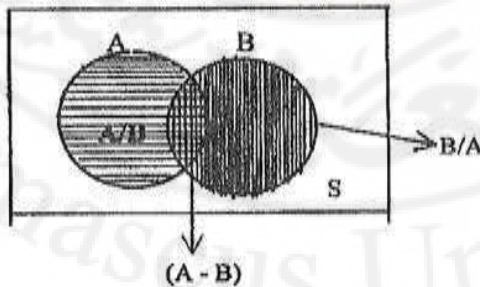
4- قانون الجمع للحوادث غير المتنافية لأي حادثتين A و B فإن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

من شكل فن (شكل رقم 7-8) نجد :

$$A \cup B = A \cup (B - A)$$

$$P(A \cup B) = P[A \cup (B - A)]$$



الشكل رقم (7-8)

وبما أن A و $B-A$ حادثتان متنافيتان، وحسب المسلمة الثالثة فإن :

$$P(A \cup B) = P(A) + P(B - A) \quad (1)$$

كذلك فإن :

$$P(B) = P[(A \cap B) \cup (B - A)]$$

وبما أن $(B-A)$ و $A \cap B$ حادثتان متنافيتان وحسب المسلمة الثالثة فإن :

$$P(B) = P(A \cap B) + P(B - A) \quad (1)$$

$$P(B - A) = P(B) - P(A \cap B) \quad (2)$$

وبتعويض (2) في (1) نجد أن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ويمكن تعميم قانون الجمع لثلاث حوادث A, B, C حيث :

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

مثال (7-16) :

في تجربة رمي ثلاث قطع نقود سوية ثلاث مرات متتالية عرفنا على فراغ العينة

(S) الحوادث الآتية :

$$A = \{ \text{جميع الأوجه صور} \}$$

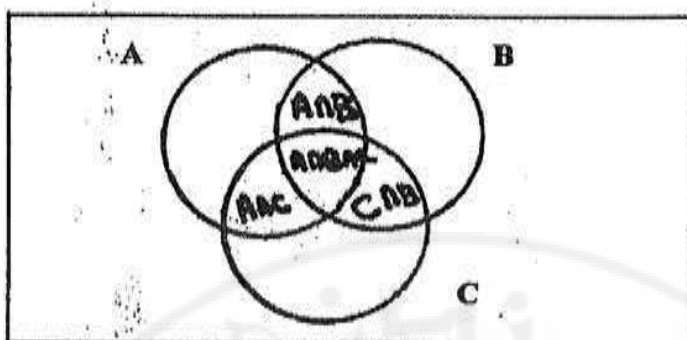
$$B = \{ \text{جميع الأوجه متشابهة} \}$$

$$C = \{ \text{ظهور صورة واحدة فقط} \}$$

والمطلوب : أوجد ما يأتي :

$$(A \cap B), (A \cap C), (B \cap C), (A \cap B \cap C), (A \cup B)$$

$$(B \cup C), (A \cup B \cup C), (A \cup C)$$



الشكل رقم (7-9)

الحل :

$$S = \{ (H, H, H), (H, H, T), (H, T, H), (H, T, T), (T, H, H), (T, H, T), (T, T, H), (T, T, T) \}$$

$$A = \{ (H, H, H) \}$$

$$B = \{ (H, H, H), (T, T, T) \}$$

$$C = \{ (H, T, T), (T, H, T), (T, T, H) \}$$

$$1 - (A \cap B) = A$$

$$3 - (B \cap C) = \phi$$

$$2 - (A \cap C) = \phi$$

$$4 - (A \cap B \cap C) = \phi$$

$$5 - (A \cup C) = B$$

$$6 - (A \cup C) = A + C$$

$$7 - (B \cup C) = B + C$$

$$8 - (A \cup B \cup C) = B + C$$

ولحساب الاحتمالات السابقة يجب أن تحقق التجربة العشوائية الشرطين الآتيين:

1- عدد النتائج الممكنة للتجربة محدود أو منته وهو مساوٍ إلى (8) في مثالنا.

2- كل نتيجة ممكنة لها الفرصة أو الاحتمال نفسه في التحقق .

عندها يمكن حساب احتمال أية حادثة A معرفة على فضاء العينة S لهذه

التجربة من العلاقة :

$$P(A) = \frac{\text{عدد الحالات الملائمة أو المرغوبة}}{\text{عدد الحالات الممكنة}}$$

وذلك حسب التعريف الكلاسيكي للاحتمال .

ولحساب الاحتمالات :

$$1 - P(A \cap B) = P(A) = 1/8$$

وذلك لأن $A \subset B$

$$2 - P(A \cap C) = 0$$

وذلك لأن A و C حوادث متنافية

$$3 - P(B \cap C) = 0$$

وذلك لأن الحوادث متنافية

$$4 - P(A \cap B \cap C) = 0$$

وذلك لأن الحوادث متنافية

$$5 - P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= 1/8 + 3/8 - 1/8 = 2/8$$

$$= P(B)$$

لأن : $A \subset B$

$$6 - P(A \cup C) = P(A) + P(C)$$

$$= 1/8 + 3/8 = 1/2$$

الحوادث متنافية

$$7 - P(B \cup C) = P(B) + P(C)$$

$$= 2/8 + 3/8 = 5/8$$

وإن الحوادث متنافية

$$8 - P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) -$$

$$P(B \cap C) + P(B \cap A \cap C)$$

$$= 1/8 + 2/8 + 3/8 + 1/8 - 0 - 0 + 0 = 5/8$$

طريقة أخرى للحل :

بما أن $A \subset B$ فإن :

$$P(A \cup B \cup C) = P(B \cup C) = P(B) + P(C)$$

وذلك لأن B و C حوادث متنافية .

$$= 2/8 + 3/8 = 5/8$$

وهي النتيجة السابقة نفسها.

مثال (7-17) :

في أحد المختبرات عشر حاسبات (6) منها تعمل بلغة البيسك فقط و (3) تعمل بلغة فورتران فقط و (1) يعمل بكلتا اللغتين، فإذا اختيرت إحدى هذه الحاسبات بصورة عشوائية وكانت الحادثة (B) ترمز إلى أن الحاسبة تعمل بلغة البيسك والحادثة (F) إلى أن الحاسبة تعمل بلغة فورتران والمطلوب :

احسب الاحتمالات الآتية مبيناً معنى كل منها :

$$P(B), P(F), P(B \cap \bar{F}), P(B \cap F), P(\bar{B} \cap F), P[(B \cap \bar{F}) \cup (B \cap F)]$$

$$P[(\bar{B} \cap F) \cup P(\bar{B} \cap F)], P(B \cup F)$$

الحل :

$$S = \{ B, B, B, B, B, B, F, F, F, B \cap F \}$$

$$B = \{ \text{الحاسبة تعمل بلغة البيسك} \}$$

$$F = \{ \text{الحاسبة تعمل بلغة فورتران} \}$$

$$1 - P(B) = \frac{7}{10}$$

الحاسبة تعمل بلغة البيسك

$$2 - P(F) = \frac{4}{10}$$

الحاسبة تعمل بلغة فورتران

$$3 - P(B \cap \bar{F}) = \frac{6}{10}$$

الحاسبة تعمل بلغة البيسك، ولا تعمل بلغة الفورتران

وهي مساوية F/B

$$4 - P(B \cap F) = \frac{1}{10}$$

الحاسبة تعمل بلغة البيسك والفورتران

$$5 - P(\bar{B} \cap F) = \frac{3}{10}$$

الحاسبة تعمل بلغة فورتران ولا تعمل

بلغة البيسك وهي مساوية F/B

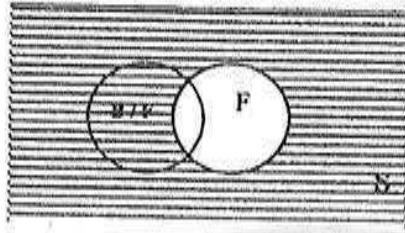
$$6 - P[(B \cap \bar{F}) \cup (B \cap F)]$$

ومن خلال خاصية مورغان الآتية :

$$(B \cap \bar{F}) \cup (B \cap F) = B \cap (F \cup \bar{F}) = B \cap S = B$$

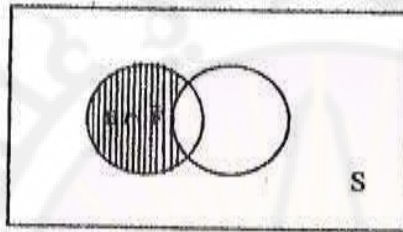
أي أن :

$$P[(B \cap \bar{F}) \cup (B \cap F)] = P(B) = \frac{7}{10}$$



الشكل رقم (7-10)

ومعناه أن تعمل الحاسبة بلغة بيسك .



الشكل رقم (7-11)

$$7 - P[(B \cap \bar{F}) \cup (\bar{B} \cap F)] = P(B \cap \bar{F}) + P(\bar{B} \cap F)$$

على اعتبار أن الحوادث متنافية .

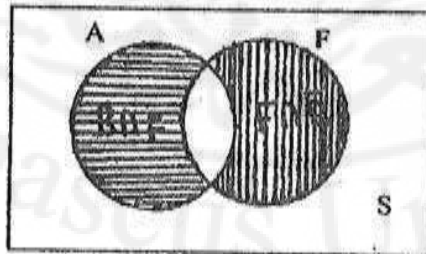
$$= 6/10 + 3/10 = 9/10$$

ومعناه أن تعمل الحاسبة إما بلغة فورتران فقط أو لغة بيسك فقط .

$$8 - P(B \cup F) = P(B) + P(F) - P(B \cap F)$$

$$= 7/10 + 4/10 - 1/10 = 1$$

ومعناه أن تعمل الحاسبة بلغة بيسك أو فورتران أو كلتا اللغتين .



الشكل رقم (7-12)

مثال (7-18) :

تتكون إحدى المستويات الدراسية من 180 طالباً كانت نتائجهم في الاختبارات كما يأتي :

18 رسبوا في الرياضيات .

24 رسبوا في الإحصاء .

12 رسبوا في كلتا المادتين .

تم اختيار طالب عشوائي من هذه الفرقة .

والمطلوب: احسب الاحتمالات الآتية :

1- رسوب الطالب في الرياضيات .

2- رسوب الطالب في الإحصاء .

3- رسوبه في الرياضيات والإحصاء .

4- رسوبه على الأقل في مادة واحدة .

5- نجاحه في الرياضيات .

6- نجاحه في الإحصاء .

7- رسوبه في الرياضيات ونجاحه في الإحصاء .

8- رسوبه في الإحصاء ونجاحه في الرياضيات .

الحل :

سوف نستخدم الرموز الآتية للتعبير عن الحوادث الأساسية في المسألة :

$A = \{ \text{رسوب الطالب في الرياضيات} \}$

$B = \{ \text{رسوب الطالب في الإحصاء} \}$

$C = \{ \text{رسوب الطالب في كلتا المادتين} \}$

$$P(A) = \frac{18}{180}$$

-1

$$P(B) = \frac{24}{180}$$

-2

$$P(A \cap B) = \frac{12}{180}$$

-3

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{18}{180} + \frac{24}{180} - \frac{12}{180} = \frac{30}{180}$$

-4

ومن تعريف العلاقات بين الحوادث يمكن التعبير عن الحوادث المراد إيجاد احتمالاتها باستخدام الرموز كما يأتي :

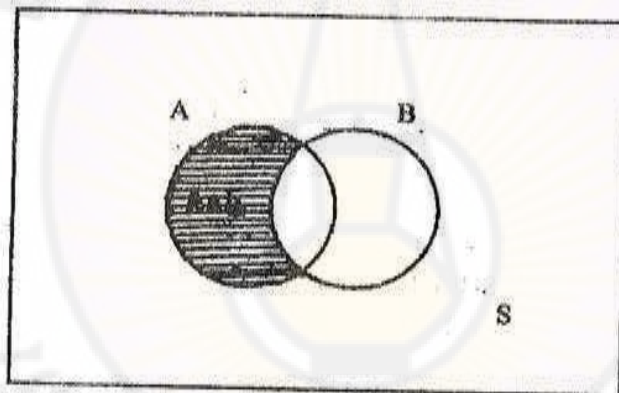
$$P(\bar{A}) = 1 - P(A) = 1 - \frac{18}{180} = \frac{162}{180}$$

-5

$$P(\bar{B}) = 1 - P(B) = 1 - \frac{24}{180} = \frac{156}{180}$$

-6

7- رسوب في الرياضيات ونجاحه في الإحصاء، أي أن المطلوب حسابه هو $A \cap \bar{B}$.



الشكل رقم (7-13)

ومن شكل فن المقابل نجد أن :

$$P(A \cap \bar{B}) = P(A) - P(A \cap B)$$

$$= \frac{18}{180} - \frac{12}{180} = \frac{6}{180}$$

8- يمكن التعبير عن حادث رسوب الطالب في الإحصاء، ونجاحه في الرياضيات

$$A - B \cap \bar{A} = B$$

$$P(B \cap \bar{B}) = P(B) - P(B \cap A)$$

$$= \frac{29}{180} - \frac{12}{180} = \frac{12}{180}$$

5- قانون الضرب للحوادث المستقلة والشرطية :

آ- الحوادث المستقلة (Independent events) :

نقول عن حادثين A و B أنهما مستقلان، إذا كان حدوث أحدهما مستقلاً عن حدوث الآخر، فالنتائج التي نحصل عليها عند رمي حجر نرد معاً، تمثل حوادث مستقلة عن بعضها البعض، لأن حدوث أحدهما مستقل عن حدوث الآخر، وكذلك الأمر في سحب كرات من صندوق (مع الإعادة) تمثل حوادث مستقلة، وعلى العكس إذا كان السحب بدون إعادة.

ولمزيد من التوضيح نقول: إذا كانت A و B حادثتين، وكان حدوث أو عدم حدوث (A)، لا يتأثر بحدوث أو عدم حدوث (B) نقول عندها: إن A و B حادثتان مستقلتان.

ونقول عن الحادثة (A) إنها مستقلة عن الحادثة (B) إذا تحقق الشرط الآتي :

$$P(A/B) = P(A) \quad \text{و} \quad P(B/A) = P(B)$$

إن الرمز أو التعبير $P(A/B)$ تعني احتمال حدوث A علماً بأن B قد حدثت ويكون الحادثان A و B غير مستقلين إذا كان :

$$P(A/B) \neq P(A)$$

$$P(B/A) \neq P(B)$$

وإن قانون الضرب للحوادث المستقلة يعطى بالعلاقة :

$$P(A \cap B) = P(A) \cdot P(B)$$

ونشير هنا إلى أن الحوادث المستقلة هي ليست حوادث متنافية، وإن وقوع (A) لا يمنع من وقوع (B)، وإن الحوادث غير المتنافية قد تكون مستقلة أو غير مستقلة. وبالنسبة للحوادث المتنافية فإن :

$$P(A/B) = \phi$$

مثال (7-19) على الحوادث المستقلة:

1- لنفرض أن لدينا باحثين A و B يعمل كل منهما على حل لمشكلة في الاقتصاد بشكل مستقل، نلاحظ هنا أن الحوادث A و B مستقلة لأن :

$$P(A/B) = P(A)$$

$$P(B/A) = P(B)$$

نلاحظ أن الحوادث غير متنافية إذ من الممكن أن يتوصل كلا الباحثين إلى حل القضية بالوقت نفسه.

2- إن نجاح محمد بالرياضيات وسفر عبد الله إلى الخارج هي حوادث مستقلة وغير متنافية.

3- إن تخرج محمد من قسم الإحصاء بمرتبة الشرف، وسفره للدراسة في الخارج هما حدثان غير متنافيان وبالوقت نفسه غير مستقلين.

مثال (7-20) :

صندوق يحتوي على (20) كرة (12) منها بيضاء و (8) كرات حمراء، سحبت من الصندوق كرتان الواحدة تلو الأخرى، علماً بأن الكرة المسحوبة الأولى قد أعيدت إلى الصندوق قبل سحب الكرة الثانية .

والمطلوب :

- 1- ما احتمال أن تكون الكرتان المسحوبتان بيضاوين .
- 2- ما احتمال أن تكون الكرتان المسحوبتان حمراوين .
- 3- ما احتمال أن تكون الكرة المسحوبة الأولى حمراء والثانية بيضاء .

الحل :

نرمز إلى سحب كرة بيضاء بالرمز $W = \{ \text{كرة بيضاء} \}$ وإن :

$$P(W) = \frac{12}{20}$$

ونرمز إلى سحب كرة حمراء بالرمز $R = \{ \text{كرة حمراء} \}$ وإن :

$$P(R) = \frac{8}{20}$$

وعلينا هنا أن نميز بين طرق السحب، فإذا أعيدت الكرة المسحوبة الأولى قبل سحب الكرة الثانية (السحب مع الإعادة) فإن الحوادث هنا مستقلة .

$$P(W \cap W) = P(W) \cdot P(W) = \frac{12}{20} \cdot \frac{12}{20} = \frac{9}{25} \quad - 1$$

$$P(R \cap R) = P(R) \cdot P(R) = \frac{8}{20} \cdot \frac{8}{20} = \frac{4}{25} \quad - 2$$

$$P(R \cap W) = P(R) \cdot P(W) + P(W) \cdot P(R) \quad - 3$$

$$= 2 \left(\frac{8}{20} \cdot \frac{12}{20} \right) = \frac{12}{25}$$

ب- الحوادث غير المستقلة (الاحتمال الشرطي) :

يعطى قانون الاحتمال الشرطي لحدثين A و B غير مستقلتين ومعرفتين على فضاء العينة S من العلاقة :

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \text{حيث} \quad P(B) > 0$$

$$P(B/A) = \frac{P(B \cap A)}{P(A)} \quad \text{حيث} \quad P(A) > 0$$

مثال (7-21) :

قمنا برمي حجر نرد متوازن وعرفنا على فضاء العينة S الحوادث الآتية :

$$A_1 = \{ \text{حادثة الحصول على الوجه الذي يحمل العدد 2} \}$$

$$A_2 = \{ \text{حادثة الحصول على الوجه الذي يحمل الرقم أقل من 4} \}$$

$$A_3 = \{ \text{حادثة الحصول على الوجه الذي يحمل الرقم أقل من 5} \}$$

$$B = \{ \text{حادثة الحصول على عدد زوجي} \}$$

والمطلوب : احسب الاحتمالات الآتية :

$$1 - P(A_1)$$

$$2 - P(A_2)$$

$$3 - P(A_3)$$

$$4 - P(B)$$

$$5 - P(A_1 \cap B)$$

$$6 - P(A_2 \cap B)$$

$$7 - P(A_3 \cap B)$$

$$8 - P(A_1/B)$$

$$9 - P(A_2/B)$$

$$10 - P(B/A_3)$$

الحل :

$$1 - P(A_1) = 1/6$$

$$2 - P(A_2) = 3/6 = 1/2$$

$$3 - P(A_3) = 4/6 = 2/3$$

$$4 - P(B) = 3/6 = 1/2$$

$$5 - P(A_1 \cap B) = 1/6$$

$$6 - P(A_2 \cap B) = 1/6$$

$$7 - P(A_3 \cap B) = 2/6$$

$$8 - P(A_1/B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{1/6}{1/2} = 1/3$$

$$9 - P(A_2/B) = \frac{P(A_2 \cap B)}{P(B)} = \frac{1/6}{1/2} = 1/3$$

$$10 - P(B/A_3) = \frac{P(B \cap A_3)}{P(A_3)} = \frac{2/6}{2/3} = 1/2$$

$$P(B \cap A_3) = 2/6$$

وبصورة عامة فإن الاحتمال المشروط بقيس احتمال تحقق الحادث A ، إذا علم

أن الحادث (B) قد وقع .

وعند التعامل مع الاحتمال الشرطي، فإننا لا نهتم من فضاء العينة (S) إلا بالفضاء المختزل للحادثة (B)، مما يعني أنه قد تم اختزال فضاء العينة (S) إلى الجزء الممثل للحادثة (B) فقط. ففي تجربة رمي حجر النرد فإن فضاء العينة هو :

$$S = \{ 1, 2, 3, \dots, 6 \}$$

فإذا عرفنا على فضاء العينة الحادث (B) بأنه ظهور رقم زوجي أي :

$$B = \{ 2, 4, 6 \}$$

وعرفنا أيضاً على فراغ العينة الحادث (A) بأنه ظهور الوجه الذي يحمل رقم {2}، وعند علمنا أن الحادث B قد وقع فإن :

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{3}$$

وإن وقوع الحادث B قد أدى إلى اختزال فضاء العينة من (S) إلى (B) .

مثال (7-22) :

في تجربة إلقاء حجري نرد متميزين، عرفنا الحوادث التالية على فضاء العينة الحوادث (A) والمتمثل في الحصول على المجموع (8) من النردين، والحادثة (B) والمتمثل في الحصول على مجموع زوجي .

والمطلوب : احسب الاحتمالات الآتية :

$$P(A) , P(B) , P(A \cap B) , P(A/B) \quad -1$$

$$P(A/B) \approx P(A) \quad -2 \text{ تحقق من أن}$$

الحل :

إن فراغ العينة هو على النحو الآتي :

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,3)	(3,5)	(3,6)
(4,1)	(4,2)	(4,2)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

(الحصول على المجموع 8 من
النردين) $A = \{ \dots \}$

(الحصول على مجموع زوجي من
النردين) $B = \{ \dots \}$

$$P(A) = \frac{5}{36}$$

-1

$$P(B) = \frac{18}{36}$$

$$P(A \cap B) = \frac{5}{36}$$

$A \subset B$ لأن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{5}{36} + \frac{18}{36} - \frac{5}{36} = \frac{18}{36}$$

وبما أن $A \subset B$ بكاملها فإن :

$$P(A \cup B) = P(B) = \frac{18}{36}$$

$$P(A / B) = \frac{P(A \cap B)}{P(B)} = \frac{5}{18}$$

- 2

$$P(A / B) = \frac{5}{18} \neq P(A) = \frac{5}{36}$$

نتيجة (1) :

لتكن لدينا الحادثتان A و B غير المتنافيتين والمعرفتين على فضاء العينة (S) فإن :

$$P(\bar{A} \cap \bar{B}) = 1 - [P(A) + P(B) - P(A \cap B)]$$

وهذا ما يعرف بقانون دي مورغان الأول.

ويمكن تعميم هذه النتيجة على مجموعة حوادث $A_1, A_2, A_3, \dots$:

$$P(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3 \dots \cap \bar{A}_n) = 1 - P(A_1 \cup A_2 \cup A_3 \dots \cup A_n)$$

وإذا كانت الحوادث A و B متنافية فإن :

$$P(\bar{A} \cap \bar{B}) = 1 - [P(A) + P(B)]$$

وعندما تكون $A \subset B$

تصبح العلاقة السابقة على النحو :

$$P(\bar{A} \cap \bar{B}) = 1 - P(B)$$

لنتيجة (2) :

لتكن لدينا الحوادث A و B غير المتنافية والمعرفة على فضاء العينة (S) فإن :

$$P(\bar{A} \cup \bar{B}) = 1 - [P(A \cap B)]$$

وهذا ما يدعى بقانون دي مورغان الثاني .

ويمكن التعميم على مجموعة حوادث $A_1, A_2, \dots, A_n$:

$$P(\bar{A}_1 \cup \bar{A}_2 \cup \dots \cup \bar{A}_n) = 1 - P(A_1 \cap A_2 \dots \cap A_n)$$

وإذا كانت الحوادث A و B متنافية فإن :

$$P(\bar{A} \cup \bar{B}) = 1$$

وإذا كانت $A \subset B$ فإن :

$$P(\bar{A} \cup \bar{B}) = 1 - P(A)$$

مثال (7-23) :

إذا كان A و B حادثين في (S) بحيث أن $P(B) = 0.6$ و $P(A) = 0.5$ و $P(A \cup B) = 0.8$ فهل A و B حادثان مستقلان ؟

الحل :

إن شرط الاستقلال الثاني هو :

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B) \\ &= (0.5) (0.6) = 0.3 \end{aligned}$$

وللتحقق وبما أن A و B هما حادثان غير متنافيين فإن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

$$P(A \cap B) = 0.5 + 0.6 - 0.8 = 0.3$$

إذن الحادثان A و B مستقلان .

مثال (7-24) :

ثلاثة باحثين a, b, c يقوم كل منهم وبشكل مستقل بدراسة مشكلة في الاقتصاد بغرض الوصول إلى حل لها، فإذا كان احتمال أن يصل الباحث a إلى حل هو $2/3$ ، واحتمال أن يصل الباحث b إلى حل هو $3/4$ ، واحتمال أن يصل الباحث c إلى حل هو $4/5$.

والمطلوب : احسب احتمال الحوادث الآتية :

- 1- التوصل إلى حل من قبل باحث واحد على الأقل .
- 2- عدم التوصل إلى حل.
- 3- توصل باحث واحد فقط للحل .

4- توصل باحثين فقط للحل .

5- توصل ثلاثة باحثين للحل .

6- توصل الباحث A إلى حل وعدم توصل الباحث B والباحث C .

الحل :

سوف نستخدم الرموز الآتية للإشارة إلى الحوادث :

A = { الباحث A يتوصل إلى حل }

B = { الباحث B يتوصل إلى حل }

C = { الباحث C يتوصل إلى حل }

1- التوصل إلى حل أي أن يتوصل الباحث A أو B أو C أو $A \cap B$ ، أو

$A \cap B \cap C$ ، أو $B \cap C$ ، أو $A \cap C$:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

وحيث إن الحوادث مستقلة :

$$P(A \cup B \cup C) = (2/3) + (3/4) + (4/5) - (2/3)(3/4) - (2/3)(4/5) - (3/4)(4/5)$$

$$+ (2/3)(3/4)(4/5) = \frac{59}{60}$$

طريقة أخرى للحل :

كون A و B و C حوادث مستقلة تكون متمماتها أيضاً حوادث مستقلة،
وبالاعتماد على قانون دي مورغان الأول :

$$P(A \cup B \cup C) = 1 - P(\bar{A} \cap \bar{B} \cap \bar{C}) = 1 - P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C}) =$$

$$= 1 - [1 - 2/3][1 - 3/4][1 - 4/5] = 59/60$$

وهذا الاحتمال هو عبارة عن توصل باحث واحد فقط للحل + توصل باحثين

فقط للحل + توصل ثلاثة باحثين للحل .

2- عدم التوصل إلى حل :

$$P(\bar{A} \cap \bar{B} \cap \bar{C}) = P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C}) = [1 - \frac{2}{3}] [1 - \frac{3}{4}] [1 - \frac{4}{5}] = \frac{1}{60}$$

أو :

$$P(\bar{A} \cap \bar{B} \cap \bar{C}) = 1 - P(A \cup B \cup C) = 1 - \frac{59}{60} = \frac{1}{60}$$

$$P[(A \cap \bar{B} \cap \bar{C}) \cup (\bar{A} \cap B \cap \bar{C}) \cup (\bar{A} \cap \bar{B} \cap C)] \quad -3$$

$$= P(A \cap \bar{B} \cap \bar{C}) + P(\bar{A} \cap B \cap \bar{C}) + P(\bar{A} \cap \bar{B} \cap C)$$

على اعتبار أن الحوادث الثلاث متنافية .

وبما أن الحوادث مستقلة :

$$= P(A) P(\bar{B}) P(\bar{C}) + P(\bar{A}) P(B) P(\bar{C}) + P(\bar{A}) P(\bar{B}) P(C)$$

$$= (\frac{2}{3}) (\frac{1}{4}) (\frac{1}{5}) + (\frac{1}{3}) (\frac{3}{4}) (\frac{1}{5}) + (\frac{1}{3}) (\frac{1}{4}) (\frac{4}{5}) = \frac{9}{60}$$

4- توصل باحثين فقط للحل :

$$P[(A \cap B \cap \bar{C}) \cup (A \cap \bar{B} \cap C) \cup (\bar{A} \cap B \cap C)]$$

$$= (\frac{2}{3}) (\frac{3}{4}) (\frac{1}{5}) + (\frac{2}{3}) (\frac{1}{4}) (\frac{4}{5}) + (\frac{1}{3}) (\frac{3}{4}) (\frac{4}{5}) = \frac{26}{60}$$

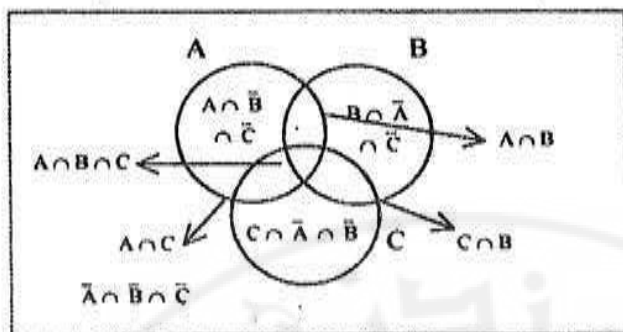
5- توصل ثلاثة باحثين للحل :

$$P(A \cap B \cap C) = P(A) P(B) P(C) = (\frac{2}{3}) (\frac{3}{4}) (\frac{4}{5}) = \frac{24}{60}$$

$$P(A \cap \bar{B} \cap \bar{C}) = P(A) P(\bar{B}) P(\bar{C})$$

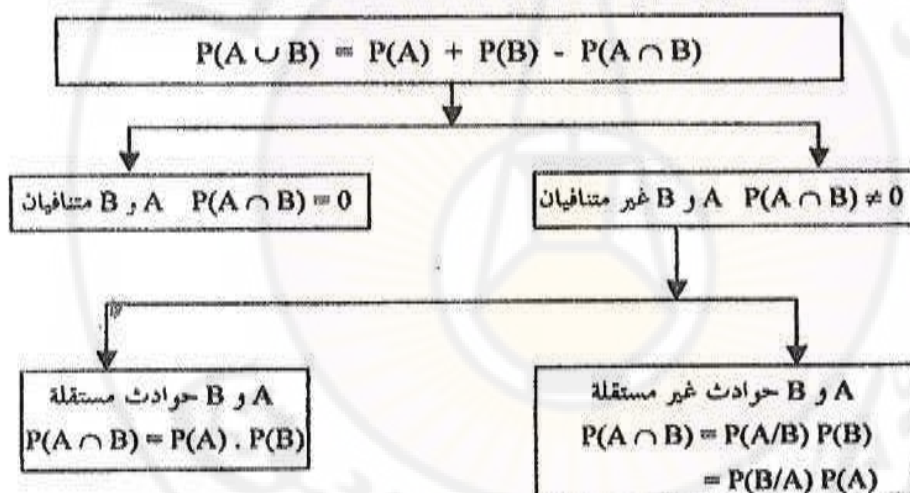
$$= (\frac{2}{3}) (1 - \frac{3}{4}) (1 - \frac{4}{5}) = 0.0333$$

- 6



الشكل رقم (7-16)

يمكن تصنيف عمليات الضرب على الحوادث وفق المخطط الآتي، علماً بأن المخطط قد وضع لحادثتين A و B.



6 - قانون الاحتمال الكلي (Law of total probability):

لتكن الحوادث $A_1, A_2, \dots, A_n$ مجموعة حوادث متنافية وشاملة معرفة على فضاء العينة S لتجربة عشوائية أي:

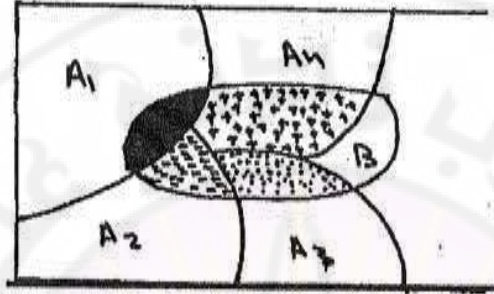
$$A_i \cap A_j = \phi \text{ for } i \neq j$$

$$\bigcup_{i=1}^n A_i = S$$

وعرفنا على S الحادثة B أي أن $B \subset S$ فإن :

$$P(B) = P(A_1) P(B/A_1) + P(A_2) P(B/A_2) + P(A_3) P(B/A_3) \\ + \dots + P(A_n) P(B/A_n)$$

$$P(B) = \sum_{i=1}^n P(A_i \cap B) = \sum_{i=1}^n P(A_i) P(B/A_i)$$



الشكل رقم (7-17)

مثال (7-25) :

ثلاثة صناديق متشابهة تحوي الكرات الآتية :

الصندوق الأول : (10 كرات) منها (7) بيضاء و (3) سوداء .

الصندوق الثاني : (5 كرات) منها (4) بيضاء و (1) سوداء .

الصندوق الثالث : (5 كرات) منها (2) بيضاء و (3) سوداء .

اختير صندوق من الصناديق الثلاثة بشكل عشوائي، ثم سحبت منه كرة بشكل

عشوائي.

والمطلوب : احسب احتمال سحب كرة بيضاء .

الحل :

نرمز للحوادث بما يأتي :

$A_1 = \{ \text{احتمال سحب كرة من الصندوق الأول} \}$

$A_2 = \{ \text{احتمال سحب كرة من الصندوق الثاني} \}$

$A_3 = \{ \text{احتمال سحب كرة من الصندوق الثالث} \}$

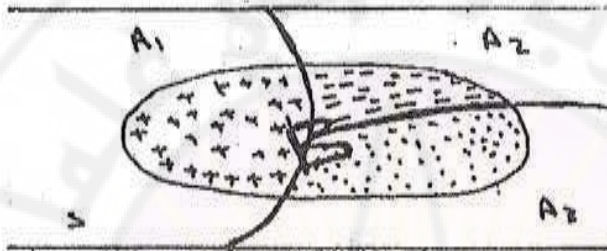
$B = \{ \text{احتمال سحب كرة بيضاء} \}$

$$P(B) = P(A_1) P(B/A_1) + P(A_2) P(B/A_2) + P(A_3) P(B/A_3)$$

وكما نلاحظ فإن الأحداث الثلاث A_3, A_2, A_1 تولف مجموعة شاملة ولها

متنافية متني متني .

$$P(B) = (1/3) \cdot (7/10) + (1/3) (4/5) + (1/3) (2/5) = 19/30$$



الشكل رقم (7-18)

7-2 التوزيع الطبيعي Normal Distribution :

7-2-1 مقدمة :

إن التوزيع الأكثر أهمية في الإحصاء هو التوزيع الطبيعي، ويعود السبب في ذلك إلى أن كثيراً من القياسات الفيزيائية، والحوادث الطبيعية، تقارب توزيعاتها الاحتمالية هذا التوزيع، مثل الطول، والوزن، وعدد الكريات في جسم الإنسان وخضاب الدم... تنتظم وفق هذا التوزيع، والسبب الآخر لأهمية هذا التوزيع أن توزيعات المعاينة وأنحطاءها تتبع بدورها للتوزيع الطبيعي، فعندما يكون المتحول العشوائي (X) مستمراً، فإن عدد قيمه الممكنة يصبح لانهائياً، وعندها يتعذر علينا إعداد جدول بقيم المتحول (X) واحتمالاتها المقابلة وتصبح العلاقة $P(X = x_i) = 0$.

لذلك فإن توزيع الاحتمالات على قيم المتحولات يعطي بدلالة تابع الكثافة الاحتمالية لتغير عشوائي مستمر وهي من الشكل :

7-2-2 تابع الكثافة للتوزيع الطبيعي :

$$f(x) = \frac{1}{\delta \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-168}{\delta}\right)^2}$$

ونلاحظ أن لهذا التوزيع وسيطين هما: الوسط الحسابي (μ) والانحراف المعياري (δ) وإن مجال تحوله $-\infty < X < +\infty$.

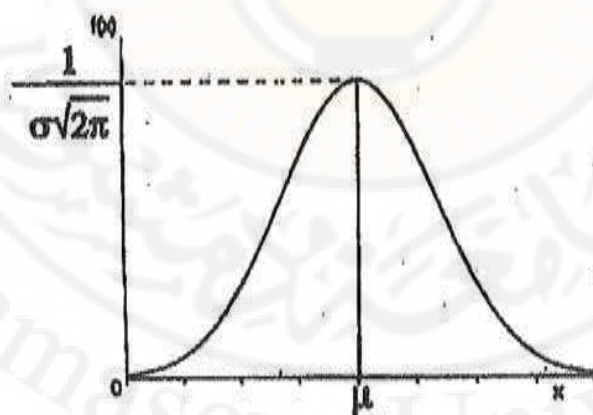
مثال (7-26) :

أنشئ نموذجاً رياضياً لتوزيع أطوال مجتمع الطلاب الجامعيين إذا علمت أن :

$$\delta = 12 \text{ cm} \quad \mu = 168 \text{ c.m}$$

$$f(x) = \frac{1}{12 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-168}{12}\right)^2}$$

ونرمز عادة لهذا التوزيع بالرمز $N \sim (\mu, \delta^2)$ حيث يمثل (μ) الوسط الحسابي و(δ) الانحراف المعياري وهو يشكل منحنياً متناظراً حول (μ) كما هو مبين في الشكل رقم (7-19).



الشكل رقم (7-19)

خواص التوزيع الطبيعي :

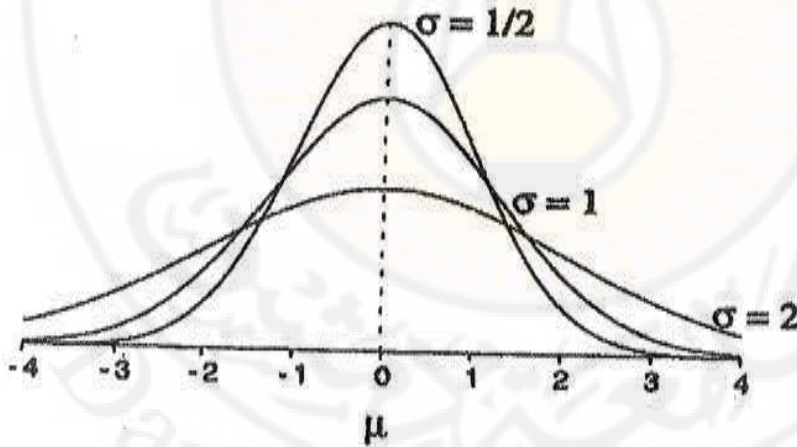
1- إن المساحة التي يحددها تابع الكثافة $f(x)$ على المحور $(0x)$ تساوي الواحد الصحيح :

$$\int_{-\infty}^{+\infty} f(x) \cdot dx = 1$$

2- إن منحنى هذا التابع متناظر حول النقطة $x = \mu$ وهي النقطة التي يتمركز عندها التوزيع أو يبلغ لهايته العظمى، ثم ينتشر على جانبيها بصورة متناظرة، وله نقطتا انعطاف في الموضعين $x = \mu - \delta$ والنقطة $x = \mu + \delta$.

ونلاحظ أنه من أجل قيم صغيرة لـ δ يكون انتشار المنحنى على جانبي μ ضيقاً، فيما يمتد إلى مسافات أبعد ويأخذ شكلاً أكثر انبساطاً كلما ازدادت δ .

الشكل رقم (7-20) وإن المساحة التي يحددها المنحنى على المحور $(0x)$ مساوية الواحد في جميع الأحوال .

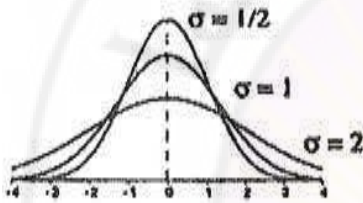


الشكل رقم (7-20)

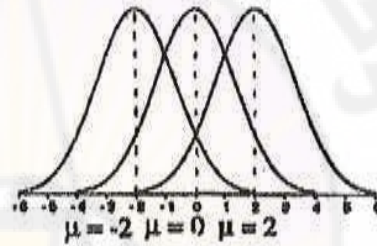
3- الأمل الرياضي :

$$E(x) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx$$

4- إن تابع الكثافة لمنحني التوزيع الطبيعي يعوي -كما رأينا- وسيطون (μ, δ) ومن أجل قيم مختلفة لهما، نحصل على منحنيات متعددة لها جميعاً شكل الجرس المقلوب، ويمثل الشكلان (7-21)، (7-22) مجموعتين من المنحنيات الطبيعية، تختلف منحنيات المجموعة الأولى عن بعضها (7-21) باختلاف متوسطاتها، بينما تختلف المجموعة الثانية (7-22) باختلاف انحرافاتها المعيارية .

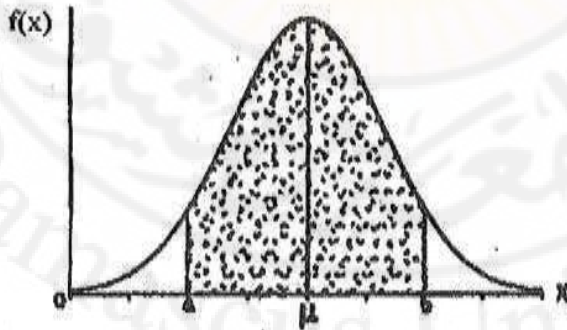


الشكل رقم (7-22)



الشكل رقم (7-21)

وإن احتمال أن يقع x في المجال $[a, b]$ هي عبارة عن المساحة المحدودة بهذا المنحني $f(x)$ والمحور $(0x)$ والمستقيمين $x = a$ و $x = b$ حيث $a < b$ الشكل (7-23) .



الشكل رقم (7-23)

ولحساب الاحتمال :

$$P(a \leq x \leq b) = \int_a^b \frac{1}{\delta \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2} \cdot dx$$

إذا كانت الظاهرة تخضع لمنحنى توزيع طبيعي في المجتمع. فإن قيم المتغيرات العشوائية تتوزع وفق النسب الآتية :

$$P(\mu - 1.8 \leq x \leq \mu + 1.8) = 68.27 \%$$

$$P(\mu - 1.96 \delta \leq x \leq \mu + 1.96 \delta) = 95 \%$$

$$P(\mu - 3 \delta \leq x \leq \mu + 3 \delta) = 99.73 \%$$

مثال (7-27) :

إذا كانت علامات الطلاب الناجحين، بمقرر مبادئ الإحصاء في إحدى الكليات في الجامعة تتوزع طبيعياً حول متوسط $\mu = 70$ درجة وانحراف معياري $\delta = 10$ درجات، سحبنا بشكل عشوائي طالباً من الطلاب الناجحين .

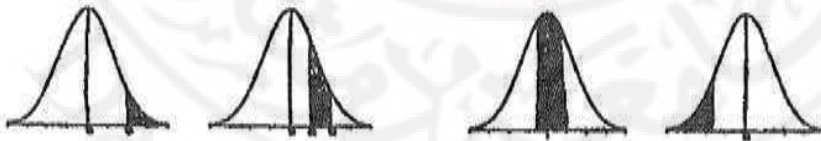
والمطلوب : احسب الاحتمالات الآتية :

$$P(x \geq 90) , P(80 \leq x \leq 90) , P(60 \leq x \leq 80) , P(x \leq 55)$$

الحل :

ولحساب هذه الاحتمالات نقوم بتمثيلها على المنحنى الطبيعي القياسي شكل

(7-24) .



الشكل رقم (7-24)

$$P(x \geq 90) = \int_{90}^{+\infty} \frac{1}{10 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2} \cdot dx$$

$$P(80 \leq x \leq 90) = \frac{1}{10\sqrt{2\pi}} \int_{80}^{90} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2} dx$$

$$P(60 \leq x \leq 80) = \frac{1}{10\sqrt{2\pi}} \int_{60}^{80} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2} dx$$

$$P(x \leq 55) = \frac{1}{10\sqrt{2\pi}} \int_{-\infty}^{55} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2} dx$$

وفي كل مرة نحتاج فيها إلى حساب المساحة أو الاحتمال لا بد من إجراء هذه التكاملات الشاقة، والسبب في ذلك يعود إلى اختلاف الأوساط الحسابية للمنحنيات وانحرافاتها المعيارية، من هنا جاءت فكرة البحث عن منحني طبيعي له شكل واحد، مهما اختلفت قيم المتحول العشوائي x ، وهو ما يعرف بالتوزيع الطبيعي والقياسي . وقبل البحث في مفهوم هذا التوزيع، لا بد من التعرض لمفهوم الدرجات المعيارية، والتي هي عبارة عن قيم المتغيرات العشوائية المستمرة (x) غير المقاسة بوحدات القياس المعروفة، وإنما مقاسة بالدرجات المعيارية، وتعرف الدرجة المعيارية (Z) بأنها وحدة قياس مستقلة تماماً عن وحدة القياس الأساسية، ويتم ذلك بتحويل المتغير العشوائي (x) إلى متغير عشوائي (Z) بالعلاقة الآتية :

$$Z = \frac{x - \mu}{\sigma}$$

7-3 التوزيع الطبيعي المعياري (القياسي) Standard Normal Distribution :

إن عملية تحويل المتغير العشوائي (x) إلى متغير (Z) تقودنا إلى توزيع طبيعي قياسي متغيره العشوائي الدرجات المعيارية (Z) ووسطه الحسابي ($\mu = 0$) وانحرافه المعياري $\sigma = 1$ ونرمز له بالرمز $N(0,1)$ ويمكن برهان ذلك رياضياً .

$$\begin{aligned}
 E(Z) &= E \left(\frac{x - \mu}{\delta} \right) \\
 &= \frac{1}{\delta} E(x - \mu) \\
 &= \frac{1}{\delta} (E(x) - \mu)
 \end{aligned}$$

ونحن نعلم أن: $E(x) = \mu$

$$E(Z) = \frac{1}{\delta} (\mu - \mu) = 0 \quad \text{إذا:}$$

أي أن μ بالنسبة لمنحنى التوزيع الطبيعي القياسي - 0 =

$$\text{var}(Z) = \delta^2 = E(Z^2) - \mu^2$$

ولكن $\mu = 0$

$$E(Z^2) = E \left(\frac{x - \mu}{\delta} \right)^2$$

$$E \left(\frac{x - \mu}{\delta} \right)^2 = \frac{1}{\delta^2} (x - \mu)^2 = \frac{1}{\delta^2} \cdot \delta^2 = 1$$

ولإيجاد تابع الكثافة لمنحنى التوزيع الطبيعي القياسي $N \sim (0, 1)$ تجري عملية

التحويل لتابع الكثافة لمنحنى التوزيع الطبيعي $N \sim (\mu, \delta^2)$.

$$Z^2 = \left(\frac{x - \mu}{\delta} \right)^2$$

والانحراف المعياري $\delta = 1$ إذا أصبح تابع الكثافة لمنحنى التوزيع الطبيعي من

الشكل:

$$\phi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} Z^2}$$

7-3-1 تابع الكثافة للتوزيع الطبيعي المعياري :

وهو ما يعرف بالتوزيع الطبيعي القياسي .

حيث إن :

$$Z = \frac{x - \mu}{\delta}$$

$$-\infty < Z < +\infty$$

إن هذا التابع متناظر حول وسطه الحسابي حيث $(\mu = 0)$ ، كما أن انحرافه المعياري $\delta = 1$.

نسمي المتغير العشوائي الجديد (Z) ، المتغير المعياري أو الدرجة المعيارية. ولحساب قيمة الاحتمال أو المساحة :

$$P(a \leq x \leq b) = \int_a^b \frac{1}{\delta \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - \mu}{\delta} \right)^2} . dx$$

نحول المتغير العشوائي (x) إلى المتغير المعياري (Z) في تابع الكثافة الاحتمالية $N(\mu, \delta^2)$ ليصبح $N(0, 1)$.

ولإيجاد التكامل السابق لا بد أيضاً من تحويل حدود التكامل إلى الدرجات المعيارية أو المتغير المعياري .

$$Z = \frac{a - \mu}{\delta} \text{ عندما } x = a \text{ يصبح الحد الأدنى للتكامل}$$

$$Z = \frac{b - \mu}{\delta} \text{ عندما } x = b \text{ يصبح الحد الأعلى للتكامل}$$

$$dz = \frac{dx}{\delta} \Rightarrow dx = dz \cdot \delta \quad \text{وإن :}$$

وبالتعويض في علاقة منحنى التوزيع الطبيعي $N(\mu, \delta^2)$:

$$P(a \leq x \leq b) = \int_{\frac{a - \mu}{\delta}}^{\frac{b - \mu}{\delta}} \frac{1}{\delta \sqrt{2\pi}} e^{-\frac{1}{2} z^2} \delta . dz$$

وبما أن $\delta = 1$ تصبح العلاقة :

$$P(a \leq x \leq b) = \frac{1}{\delta \sqrt{2\pi}} \int_{\frac{a-\mu}{\delta}}^{\frac{b-\mu}{\delta}} e^{-\frac{1}{2}z^2} \cdot dZ$$

مثال (7-28) :

احسب المساحة المحصورة $P(0 \leq x \leq 2)$

الحل :

$$P(0 \leq x \leq 2) = \frac{1}{\sqrt{2\pi}} \int_0^2 e^{-\frac{1}{2}z^2} \cdot dZ$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^2 \left(1 - \frac{Z^2}{2} + \frac{Z^4}{4 \times 2} - \frac{Z^6}{8 \times 2 \times 3} + \frac{Z^8}{16 \times 2 \times 3 \times 4} - \frac{Z^{10}}{32 \times 2 \times 3 \times 4 \times 5} \right) \cdot dZ$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^2 \left(1 - \frac{Z^2}{2} + \frac{Z^4}{8} - \frac{Z^6}{48} + \frac{Z^8}{384} - \frac{Z^{10}}{3840} \right) \cdot dZ$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^2 \left(\frac{3840 - 1920 Z^2 + 480 Z^4 - 80 Z^6 + 10 Z^8 - Z^{10}}{3840} \right) \cdot dZ$$

$$= \frac{1}{3840 \sqrt{2\pi}} \left[3840 Z - 1920 \frac{Z^3}{3} + 480 \frac{Z^5}{5} - 80 \frac{Z^7}{7} + 10 \frac{Z^9}{9} - \frac{Z^{11}}{11} \right]_0^2$$

$$= \frac{1}{3840 \times 2.5} \left(3840 \times 2 - 1920 \cdot \frac{8}{3} + 480 \cdot \frac{32}{5} - 80 \cdot \frac{128}{7} + 10 \cdot \frac{512}{9} - \frac{2048}{11} \right)$$

$$= \frac{4551.85}{6900} = 0.474151$$

وكما لاحظنا من خلال التمرين السابق أن إجراء هذا التكامل شاق، ويتطلب جهداً كبيراً، لذلك تم وضع جدول خاص لحساب هذه التكاملات، ويدعى بجدول المساحات تحت منحنى التوزيع الطبيعي القياسي $N(0,1)$ ، وقد حسبت المساحات في

هذا الجدول لنصف المنحني على اعتبار أن منحني التوزيع القياسي متناظر حول النقطة $\mu = 0$ ، وهي عبارة عن الوسط الحسابي ، وتقسم المنحني إلى قسمين متساويين ومتناظرين تماماً، وقد حسبت المساحات بدءاً من النقطة $Z = 0$ وبتزايد مقداره 0.01. ويقسم جدول المساحات، إلى سطر وعمود، وتبدأ قيم (Z) في السطر بالقيمة $Z = 0$ وتزايد مقداره 0.01 حتى القيمة 0.09 وتبدأ قيم العمود من القيمة $Z = 0$ وتزايد مقداره 0.1 حتى القيمة $Z = 3$ حيث المساحة المحصورة $P(Z \geq 0) \approx 0.5$.

مثال (7-29) :

احسب المساحة المحصورة بين القيمتين $(Z = 0, Z = 1.96)$ درجة معيارية أي $P(0 \leq Z \leq 1.96)$ باستخدام جدول المساحات .

الحل :

لإيجاد المساحة المطلوبة نبحث في عمود (Z) عن القيمة 1.9 وفي سطر (Z) عن القيمة 0.06 ونقطة تقاطع السطر والعمود هي عبارة عن المساحة المطلوبة .

$$P(0 \leq Z \leq 1.96) = 0.475$$

مثال (7-30) :

احسب احتمال وقوع (Z) بين $(0, 0.98)$ أي حساب الاحتمال $P(0 \leq Z \leq 0.98)$. لهذه الغاية نبحث في عمود (Z) عن القيمة (0.9)، وفي السطر عن القيمة (0.08) ونقطة التقاطع بينهما هي المساحة المطلوبة :

$$P(0 \leq Z \leq 0.98) = 0.3365$$

مثال (7-31) :

إذا علمت أن كمية الكوليسترول في لتر من الدم تتوزع طبيعياً بمتوسط $\mu = 2.20$ والانحراف المعياري $\delta = 0.36$.

والمطلوب :

- ما هي نسبة الأشخاص الذين تتراوح كمية الكوليسترول في دمائهم بين القيمتين 2.38 , 2.74 في المجتمع .

الحل :

$$P(2.38 \leq x \leq 2.74) = P\left(\frac{2.38 - 2.20}{0.36} \leq Z \leq \frac{2.74 - 2.20}{0.36}\right)$$

$$P(2.38 \leq x \leq 2.74) = P(0.5 \leq Z \leq 1.5)$$

$$\begin{aligned} P(0.5 \leq Z \leq 1.5) &= P(0 \leq Z \leq 1.5) - P(0 \leq Z \leq 0.5) \\ &= 0.4332 - 0.1915 = 0.2417 \end{aligned}$$

مثال (7-32) :

إن طول الشخص هو متحول عشوائي مستمر (x) يخضع للتوزيع الطبيعي بمتوسط $\mu = 170$ وانحراف معياري $\sigma = 5$.

1- ما هو احتمال أن يقع x في المجال $P(160 \leq x \leq 185)$.

2- ما هو احتمال أن يقع x في المجال $P(x \geq 170)$.

3- ما هو احتمال أن يقع x في المجال $P(x \leq 160)$.

الحل :

$$1 - P(160 \leq x \leq 185) = P\left(\frac{160 - 170}{5} \leq Z \leq \frac{185 - 170}{5}\right)$$

$$= P(-2 \leq Z \leq 3)$$

$$= P(0 \leq Z \leq 2) + P(0 \leq Z \leq 3)$$

$$= 0.4772 + 0.49865 = 0.975$$

$$2 - P(x \geq 170) = P\left(Z \geq \frac{170-170}{5}\right)$$

$$P(Z \geq 0) = P(0 \leq Z \leq 3) \approx 0.5$$

$$3 - P(x \leq 160) = P\left(Z \leq \frac{160-170}{5}\right) = P(Z \leq -2)$$

$$P(Z \geq 2) = P(0 \leq Z < 2) - P(Z \leq 2)$$

$$= 0.5 - 0.4772 = 0.0228$$

مثال (7-33) :

إذا كان (Z) متغيراً عشوائياً مستمراً، يتبع التوزيع الطبيعي المعياري $N \sim (0,1)$.

والمطلوب :

احسب الاحتمالات الآتية :

$$1- P(0 \leq Z \leq 1.45)$$

$$2- P(1 \leq Z \leq 2)$$

$$3- P(Z \leq -1.2)$$

$$4- P(1 \leq Z \leq 1.28)$$

$$5- P(Z \geq 1.25)$$

$$6- P(-1.8 \leq Z \leq 3)$$

$$7- P(-2 \leq Z \leq -1)$$

الحل :

$$1 - P(0 \leq Z \leq 1.54)$$

$$P(Z \leq 1.54) = 0.4382$$

$$2 - P(1 \leq Z \leq 2) = P(Z \leq 2) - P(Z \leq 1)$$

$$= 0.4712 - 0.3413 = 0.1359$$

$$3 - P(Z \leq 1.2)$$

نحن نعلم أن جدول المساحات لا يعطي المساحة للقيم السالبة ونظراً لتماثل المنحني في كل من طرفيه الموجب والسالب فإن :

$$\begin{aligned} P(Z \leq -1.2) &= P(Z \geq 1.2) \\ &= P(Z \leq 0) - P(Z \leq 1.2) \\ &= P(0.5) - 0.3849 = 0.1151 \end{aligned}$$

ويمكن أن نقوم بحساب بقية الاحتمالات بالطريقة نفسها، وتترك واجباً للطالب.

7-3-2 تحديد قيم Z باحتمال معلوم $(1 - \alpha)$:

نرغب أحياناً في تحديد مجالات تحوي (Z) باحتمال معلوم، فإذا كان لدينا توزيع $N(0,1)$ وكنا نرغب في تعيين مجال متناظر حول $\mu = 0$ كالمجال $(-Z_1, Z_1)$ والذي يحتوي على 98% من قيم التوزيع مثلاً :

$$P(Z_1 \leq Z \leq Z_1) = 0.98$$

أو :

$$P(0 \leq Z \leq Z_1) = 0.49$$

وبالرجوع إلى جدول المساحات نجد أن قيمة $Z_1 = 2.33$ أي أن المجال $[-2.33, 2.33]$ يحتوي على (98%) من قيم التوزيع وبالطريقة ذاتها نجد أن المجال $[-1.96, 1.96]$ يحتوي على 95% من قيم التوزيع .

وبصورة عامة إذا رغبنا في تعيين مجال يحوي $(1 - \alpha)$ من قيم التوزيع نطبق العلاقة:

$$P\left(-Z_{\frac{1-\alpha}{2}} \leq Z \leq Z_{\frac{1-\alpha}{2}}\right) = 1 - \alpha$$

حيث تمثل $Z_{\frac{1-\alpha}{2}}$ فاصلة النقطة (OZ) التي يقع على يمينها $\alpha/2$ من قيم التوزيع

ويمكن تعميم ذلك من أجل مجتمع طبيعي عادي $N(\mu, \sigma^2)$.

فإذا رغبتنا في تعيين مجالات متناظرة حول μ تحوي $(1 - \alpha)$ من قيم التوزيع نبذل (Z) بقيمتها بدلالة (x) من العلاقة السابقة فتصبح :

$$P\left(-Z_{\frac{1-\alpha}{2}} \leq \frac{x-\mu}{\delta} \leq Z_{\frac{1-\alpha}{2}}\right) = 1-\alpha$$

$$P\left(\mu - \delta_{\frac{1-\alpha}{2}} \leq x \leq \mu + \delta_{\frac{1-\alpha}{2}}\right) = 1-\alpha$$

ويمكننا تعيين القيم المختلفة لـ $Z_{\frac{1-\alpha}{2}}$ حسب مستويات الثقة $(1 - \alpha)$ المطلوبة. والجدول الآتي يعطينا بعض هذه القيم :

α	31.75 %	10 %	5 %	2 %	1 %	0.27 %
$1 - \alpha$	68.27 %	90 %	95 %	98 %	99 %	99.73 %
$Z_{\frac{1-\alpha}{2}}$	± 1	± 1.64	± 1.96	2.33	± 2.58	± 3

مسائل وتمارين

1- يحتوي صندوق على 20 وحدة من منتج معين منها 5 تالفة. سحبت وحدتان بالتتالي وبدون إرجاع.

والمطلوب :

ما احتمال أن تكون كل من الوحدتين المسحوبتين تالفتين ؟

ما احتمال أن تكون الوحدة الأولى تالفة والثانية غير تالفة ؟

ما احتمال أن تكون الوحدة الأولى غير تالفة والثانية غير تالفة أيضاً ؟

ما احتمال أن تكون الوحدة الأولى غير تالفة والثانية تالفة ؟

2- ثلاثة متبارين (I, II, III) يصوبون بنادقهم إلى هدف واحتمال إصابتهم للهدف على التوالي (I- 0.7, II- 0.5, III- 0.4) علماً بأن كل واحد منهم أطلق طلقة واحدة.

والمطلوب:

1- إيجاد احتمال إصابة الهدف بطلقة واحدة فقط .

2- إيجاد احتمال إصابة الهدف بطلقتين فقط .

3- إيجاد احتمال إصابة الهدف بثلاث طلقات .

4- إيجاد احتمال إصابة الهدف بطلقة واحدة على الأقل .

5- إيجاد احتمال إصابة الرامي I وعدم إصابة II, III .

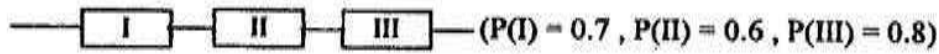
6- إيجاد احتمال عدم إصابة الهدف من قبل أي رامي.

7- إيجاد احتمال إصابة الهدف بطلقتين على الأقل.

3- دائرة كهربائية تم توصيل مقاومتها الثلاثة على التوالي I, II, III وكان معلوماً أن

المقاومات الثلاثة يمكن أن تعمل بكفاءة لمدة 500 ساعة بشرط أن لا تتعطل أي

واحدة منها، علماً بأن المقاومات تعمل بشكل مستقل الواحدة عند الأخرى، وإن احتمال عمل كل مقاومة لمدة 4000 ساعة هو على النحو الآتي :



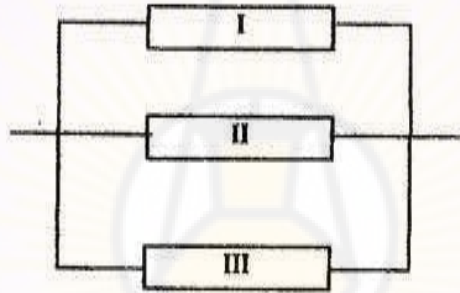
والمطلوب :

ما احتمال أن تظل الدارة تعمل لمدة 4000 ساعة بدون تعطل ؟

4- في التمرين السابق (3) إذا كانت المقاومات الثلاث في الدارة مربوطة على التوازي (أي أن الدارة تبقى تعمل طالما أن مقاومة واحدة على الأقل تعمل)

والمطلوب :

ما احتمال أن تظل الدارة تعمل بكفاءة لمدة 4000 ساعة ؟



5- تم سحب ورقتين من ورق اللعب الكوتشينه على التوالي وبدون إعادة .

والمطلوب :

احسب احتمال أن تكون الورقتان المسحورتان صورتين .

6- في دراسة فرضية أجريت حول العلاقة بين مكان الإقامة (حضر - ريف) والتعليم

(أمي - متعلم)، سحبت عينة عشوائية مؤلفة من 2000 شخص من الذكور

والإناث بعمر 10 سنوات فما فوق، فأعطت العينة النتائج الآتية :

--- إن 60% من أفراد العينة كانوا يقيمون في الريف و 40% في المدينة .

- تبين أن هناك 450 شخصاً من المقيمين في المدينة هم أميون و 350 هم متعلمون.

- تبين أن هناك 700 شخص من المقيمين في الريف هم أميون و 500 هم متعلمون.
والمطلوب :

- 1- احسب احتمال أن يكون شخص ما متعلماً إذا علمت أنه يقطن في الريف.
 - 2- احسب احتمال أن يكون شخص ما أمياً إذا علمت أنه يقطن في المدينة.
 - 3- احسب احتمال أن يكون شخص ما أمياً إذا علمت أنه يقطن في الريف.
 - 4- احسب احتمال أن يكون شخص ما متعلماً إذا علمت أنه يقطن في المدينة.
 - 5- احسب احتمال أن يكون شخص ما يقطن في الريف إذا علمت أنه متعلم.
 - 6- احسب احتمال أن يكون شخص ما يقطن في المدينة إذا علمت أنه أمي.
 - 7- احسب احتمال أن يكون شخص ما يقطن في المدينة إذا علمت أنه متعلم.
 - 8- احسب احتمال أن يكون شخص ما يقطن في الريف إذا علمت أنه أمي .
- 7- شركة تأمين تقدر لعائلة مكونة من زوج وزوجته احتمال أن يعيش الزوج مدة أكثر من (15) عاماً هو (0.2) واحتمال أن تعيش زوجته أكثر من المدة نفسها هو (0.25) .

والمطلوب : احسب الاحتمالات الآتية :

- 1- أن يعيشا معاً أكثر من 15 عاماً .
 - 2- احتمال أن يعيش أحدهما على الأقل أكثر من 15 عاماً .
 - 3- احتمال أن تعيش الزوجة فقط أكثر من 15 عاماً .
 - 4- احتمال أن يعيش الزوج فقط أكثر من 15 عاماً .
 - 5- احتمال أن يعيش أحدهما فقط أكثر من 15 عاماً .
- 8- صندوق يحتوي على عدد w من الكرات البيضاء، و b من الكرات السوداء .
سحبت من الصندوق ثلاث كرات على التوالي دون إعادة .
والمطلوب : احسب الاحتمالات الآتية :

1- الكرة الأولى سوداء والثانية بيضاء والثالثة بيضاء .

2- سوداء سوداء سوداء .

3- سوداء سوداء بيضاء .

4- سوداء بيضاء سوداء .

5- احسب الاحتمالات لبقية المسارات مستخدماً الشجرة البهائية .

9- ألقيت قطعة نقود سورية ثلاث مرات متتالية :

بفرض أننا عرفنا على فراغ العينة الحوادث الآتية:

الحادثة A = { الرمية الأولى صورة } .

الحادثة B = { الرمية الثانية صورة } .

الحادثة C = { وقوع الصورة مرتان متتاليتين فقط } .

والمطلوب : تحقق من استقلالية الحوادث A, B, C .

10- تمت معايرة سكر الدم عند شخص ثلاث مرات في ظروف مختلفة وكانت النتائج

التي نحصل عليها في كل مرة طبيعي (N) غير طبيعي (O)، وعرفنا على فضاء العينة الحوادث الآتية :

الحادث الأول: هو الحادث الدال على التسجيل طبيعي (N) في المرة الأولى، ونرمز له بـ (A).

الحادث الثاني: هو الحادث الدال على التسجيل طبيعي في المرة الثانية، ونرمز له بـ (B).

الحادث الثالث: هو الحادث الدال على التسجيل طبيعي مرتين متتاليتين فقط ونرمز له بـ (C).

والمطلوب : احسب الاحتمالات الآتية :

1- $P(A \cap B)$, $P(A \cap C)$, $P(B \cap C)$, $P(A \cap B \cap C)$

2- تحقق من استقلال الحوادث A, B, C .

3- تحقق من استقلال الحوادث $\bar{A}$, $\bar{B}$, $\bar{C}$.

4- تحقق من استقلال الحادثين $\bar{A}$, $\bar{B}$.

11- شركة أحذية لها ثلاثة مصانع إنتاجية، المصنع الأول ينتج 40% من إنتاج الشركة

وفيه 2% من البضاعة معيبة، المصنع الثاني ينتج 35% من إنتاج الشركة وفيه 1%

من البضاعة معيبة. المصنع الثالث ينتج 25% من إنتاج الشركة ونسبة الإنتاج

المعيب فيه 4%، إذا اشترت حذاء وبشكل عشوائي.

والمطلوب :

احسب الاحتمالات الآتية :

1- أن يكون الحذاء معيباً .

2- أن يكون الحذاء غير معيب .

3- إذا اشتريت حذاء وكان معيباً، ما احتمال أن يكون من إنتاج المصنع الأول؟

4- إذا اشتريت حذاء وكان معيباً، ما احتمال أن يكون من إنتاج المصنع الثاني؟

5- إذا اشتريت حذاء وكان معيباً، ما احتمال أن يكون من إنتاج المصنع الثالث ؟

6- ما احتمال أن يكون من إنتاج المصنع الأول أو الثاني ؟

7- ما احتمال أن يكون من إنتاج المصنع الأول أو الثالث ؟

8- ما احتمال أن يكون من إنتاج المصنع الثاني أو الثالث ؟

9- ما احتمال أن يكون من إنتاج المصنع الأول أو الثاني أو الثالث ؟

10- اشترت حذاء ووجدته غير معيب، ما احتمال أن يكون من إنتاج المصنع الأول ؟

11- اشترت حذاء ووجدته غير معيب، ما احتمال أن يكون من إنتاج المصنع الأول أو الثاني ؟

12- صندوقان يحتويان على كرات حمراء وكرات بيضاء :

الصندوق الثاني	الصندوق الأول
(II)	(I)
5	3 حمراء
2	6 بيضاء

وتلقى زهرة نرد بحيث إذا كانت النتيجة 2 أو 4 يتم سحب كرة من الصندوق الثاني وبخلاف ذلك فإن الكرة تسحب من الصندوق I .
ألقيت زهرة النرد وبناء على النتيجة تم سحب كرة من الصندوق الموافق لنتيجة زهرة النرد .
والمطلوب :

- 1- احسب احتمال أن تكون الكرة المسحوبة حمراء .
- 2- إذا علمت أن الكرة المسحوبة حمراء ما احتمال أن تكون قد سحبت من الصندوق I ؟

13- أُلقيت قطعة نقود مرتين متتاليتين وقد تم تعريف الحوادث الآتية على فراغ العينة:

- $A = \{ \text{ظهور صورة في الرمية الأولى} \}$
 $B = \{ \text{ظهور صورة في الرمية الثانية} \}$
 $C = \{ \text{الرميتان تعطيان النتيجة نفسها} \}$

والمطلوب :

- تحقق من الحوادث A, B, C هي حوادث مستقلة .

14- صف مكون من 40 فتاة، 30 منهن ذوات عيون عسلية، و 25 ذوات شعر أسود و 20 من ذوات العيون العسلية والشعر الأسود، اختيرت فتاة بشكل عشوائي .

والمطلوب : احسب الاحتمالات الآتية :

- 1- أن تكون الفتاة المختارة ذات عيون عسلية أو شعر أسود .
- 2- أن تكون الفتاة المختارة ذات شعر أسود وعيون ليست عسلية.
- 3- أن تكون الفتاة المختارة ذات شعر ليس أسوداً وذات عيون ليست عسلية.
- 4- أن تكون الفتاة المختارة ذات عيون عسلية وشعر ليس أسوداً.
- 5- أن تكون الفتاة المختارة ذات عيون عسلية وشعر أسود أو ذات شعر أسود وعيون ليست عسلية .

15- كان الاعتقاد السائد لدى المسؤولين عن الصناعات البترولية أن متوسط الأجر الأسبوعي للعامل 3100 وحدة نقدية بانحراف معياري 250 وحدة، فإذا علمت أن الأجور كانت تتبع توزيعاً طبيعياً في هذه الصناعة، وسحبنا بصورة عشوائية عاملاً من عمال هذه الصناعة فما هو احتمال أن يكون أجره :

$$P(2780 \leq x \leq 3420) \quad -1$$

$$P(3590 \leq x \leq 3350) \quad -2$$

$$P(x \leq 2455) \quad -3$$

على اعتبار أن الأجر هو متغير عشوائي مستمر ونرمز له بـ (x) .

16- رغبت إحدى المنشآت الصناعية المتخصصة لصناعة البطاريات الصغيرة طرح نوع جديد من هذه البطاريات، بحيث يكون العمر الوسطي للبطارية الواحدة $\mu = 200$ ساعة، وقد تبين من الدراسات التي أجريت على هذه البطاريات أن أعمارها تتبع التوزيع الطبيعي، انحرافه المعياري $\sigma = 15$ ساعة ، كما تقرر تسويق البطاريات بعلب تحتوي الواحدة منها على (100) بطارية .

1- إذا سحبنا من إحدى ألعاب بطارية واحدة، فما احتمال أن يزيد عمرها على (230) ساعة ؟

2- ما عدد البطاريات التي يتراوح عمرها بين 170 ساعة ، و 215 ساعة في علبة ما سمحت عشوائياً ؟

17- يتوزع المتغير العشوائي (x) توزيعاً طبيعياً (120.400) ~ N ، والمطلوب تشكيل مجال متناظر حول (μ) بحيث يكون احتمال وقوع (x) فيه 94% .
الحل :

إذا رمزنا بـ (Z) للمتغير المعياري الموافق لـ (x) فإن المجال المطلوب حول (μ) يصبح مجالاً متناظراً حول (0) بالنسبة للمتغير (Z) .

$$P \left(-Z_{\frac{1-\alpha}{2}} \leq Z \leq Z_{\frac{1-\alpha}{2}} \right) = 0.94$$

نبحث عن $Z_{\frac{1-\alpha}{2}}$ الموافقة للاحتمال 0.94 فنجد :

$$Z_{\frac{1-\alpha}{2}} = 1.88$$

$$Z = \frac{x - 120}{20}$$

لكن

$$P \left(-1.88 \leq \frac{x - 120}{20} \leq 1.88 \right)$$

$$(120 - 1.88 \times 20 \leq x \leq 120 + 1.88 \times 20) = 0.94$$

$$(82.4 \leq x \leq 157.6) = 0.94$$

والجواب المطلوب هو : (82.4 , 157.6) .



الفصل الثامن

تحليل الارتباط والانحدار

Correlation Regression Analysis

- الغاية من تحليل الارتباط والانحدار.

- تحليل الارتباط .

الارتباط البسيط للمتغيرات الكمية .

الارتباط للمتغيرات النوعية .

- تحليل الانحدار

الانحدار الخطي البسيط .

إيجاد معادلة مستقيم الانحدار.

أهمية معادلة مستقيم الانحدار .

دراسة التباينات.

التقدير والتنبؤ.



8-1 الغاية من تحليل الارتباط والانحدار

تعرفنا في الفصول السابقة على بعض أساليب التحليل الإحصائي، التي يمكن استخدامها لتحليل متغير واحد مثل أطوال مجموعة من الأشخاص، أو أوزانهم أو دخولهم.... الخ، فرأينا كيف يمكن حساب النسبة المركزية والتشتت واحتمال أن يأخذ المتغير قيمة ما أو تكون قيمته في مجال ما .

سوف نتعرف في هذا الفصل على أساليب تحليل جديدة نستطيع باستخدامها تحليل متغير ما، من خلال علاقته بمتغير آخر أو بعدة متغيرات، حيث نستطيع أن نذكر عدداً كبيراً من الأمثلة على الظواهر أو المتغيرات التي يوجد علاقة بينها فمثلاً: دخل الأسرة وإنفاقها، دخل الأسرة ومداخيلها، الغلة الزراعية وكمية السماد المستخدم، لون بشرة الإنسان وعدد شعر الرأس، الخ. ومن خلال دراسة وتحليل متغير ما بعلاقته مع متغير آخر، أو مجموعة من المتغيرات يمكننا التأكد من وجود هذه العلاقة وتحديد متانتها بين المتغيرات، وهذا الأسلوب يسمى بتحليل الارتباط (Correlation analysis)، كما نستطيع معرفة تأثير أحد المتغيرين على متغير آخر، أو تأثير عدة متغيرات على متغير واحد من خلال دراسة الانحدار (Regression).

إن تحليل الارتباط والانحدار يتم من خلال إجراء عدد من الحسابات، يسبقها التأكد من وجود علاقة بين المتغيرات عن طريق التحليل المنطقي بالاعتماد على معارفنا

العامة كما يمكن الاستعانة بأصحاب الخبرات للتأكد من منطقية العلاقة بين المتغيرات، وإلا سوف نحصل على مؤشرات تؤكد وجود العلاقة وهي غير موجودة منطقياً، وهذا ما يدعى بالارتباط الزائف، مثلاً تزايد المبيعات السنوية لمتجرين مختلفين من مدينتين مختلفتين، والتأكيد المنطقي للعلاقة يتطلب بالإضافة إلى وجود العلاقة أن تكون الفترة الزمنية للعلاقة منطقية مثلاً: الغلة وكمية السماد الكيماوي للعام نفسه، أما الغلة وكمية السماد البلدي تكون الفترة الزمنية أكثر من عام، كما يجب ألا يكون هناك فاصل مكاني يلغي منطقية العلاقة حيث نعلم أن هناك علاقة بين دخل الأسرة ونفقاتها، لكن إذا أخذنا دخل عدة أسر في دمشق وإنفاق عدة أسر في عمان، فإن الفاصل المكاني ينفي هذه العلاقة. ونشير إلى أن أهمية تأكيد وجود العلاقة المنطقية لا يعني بالضرورة وجود علاقة سببية بين المتغيرين، أي أن أحدهما يؤثر بالآخر، فقد يكون أحدهما يؤثر بالآخر من خلال متغير آخر مثلاً: علاقة الزمن بتزايد عدد السكان في بلد ما، فالزمن كمتغير يعكس تأثير متغيرات أخرى كثيرة، ومن الممكن أن تكون كلتا الظاهرتين تتحركان بسبب متغيرات أخرى تؤثر عليهما مثلاً: تطور عدد الحيوانات الداجنة في بلد ما، وإنتاجها النباتي فإنه يؤثر على كليهما عدد كبير من المتغيرات الأخرى.

بعد التأكد من وجود العلاقة بين الظواهر (المتغيرات) يجب تحديد أي المتغيرات يؤثر بالآخر، وهو ما سوف ندعوه بالمتغير المستقل (Independent variable) ونرمز له بالرمز x ، أما المتغير الذي يتأثر بالآخر ندعوه بالمتغير التابع (Dependent variable) ونرمز له بالرمز y ، وإن تحديد طبيعة المتغيرات هو امتداد للخطوة السابقة، التأكد من وجود العلاقة، حيث يتم تحديد طبيعة المتغيرات في أكثر الأحيان من خلال التحليل المنطقي، فالدخل يؤثر باستهلاك الأسرة، فهو المستقل، والاستهلاك تابع والسماد يؤثر في الغلة فهو المستقل والغلة تابع... أما إذا تعذر تحديد طبيعة المتغيرات بالطريقة السابقة فعندها من المنطق القول إن من يحدث أولاً هو الذي يؤثر بالآخر، وبالتالي هو المستقل، وإذا تعذر حسب الطريقتين السابقتين تحديد طبيعة المتغيرات يكون للباحث الخيار في اعتبار أحدهما مستقلاً والآخر تابعاً، وأن الخطأ المرتكب سيكون ضئيلاً كما سنرى فيما بعد.

بعد القيام بالإجراءات السابقة يتم حساب الارتباط والانحدار للظواهر، مع العلم بأنه يمكن دراسة علاقة متغير مستقل واحد مع متغير تابع واحد، وهنا يدعى الارتباط البسيط والانحدار البسيط، أما إذا درسنا علاقة عدة متغيرات مستقلة بمتغير تابع واحد، فالارتباط يدعى متعددًا والانحدار يدعى متعددًا مثل علاقة إنفاق الأسرة كمتغير تابع بدخلها، وعدد أفرادها ومكان إقامتها ودرجة تعلم أفرادها... كمتغيرات مستقلة .

8-2 تحليل الارتباط (Correlation Analysis)

يمكن تقسيم العلاقات بين الظواهر إلى قسمين: علاقات تابعة عندما يكون لكل قيمة للمتغير المستقل قيمة بالذات للمتغير التابع، مثل علاقة مساحة الدائرة بنصف القطر، ومساحة المربع بالضلع، واستهلاك معمل ما في أحد المواد الأولية بحجم الإنتاج، وعلاقات ارتباطية عندما يكون مقابل كل قيمة للمتغير المستقل قيمة تقريبية أو احتمالية للمتغير التابع، وإننا هنا سوف ندرس النوع الثاني فقط، حيث تمت دراسة النوع الأول من العلاقات في المرحلة الثانوية.

والعلاقة بين الظواهر إما أن تكون خطية (Linear) يمكن تمثيلها بخط مستقيم، أو منحنية (Non-Linear) تمثل بمنحنى غير المستقيم .

وسوف ندرس هنا الارتباط، ولأسباب دراسية كما يأتي :

8-2-1 الارتباط البسيط للمتغيرات الكمية

8-2-2 الارتباط للمتغيرات النوعية

8-2-1 الارتباط البسيط للمتغيرات الكمية

بعد أن يتأكد الباحث من وجود علاقة بين المتغيرات قيد الدراسة يحدد المتغير المستقل والمتغير التابع ويقوم بجمع البيانات عن الظاهرتين، حيث من خلال هذه البيانات يستطيع أن يحدد متانة هذه العلاقة وطبيعتها بالطرائق الآتية :

8-2-1-1 التحليل المباشر :

إذا تبين من خلال قراءة بيانات المتغيرين أن القيم الدنيا للمتغير المستقل تتقابل مع القيم الدنيا للمتغير التابع والعليا مع العليا، فالعلاقة بين المتغيرين هي علاقة طردية، أما إذا كانت القيم الدنيا للمتغير المستقل تتقابل مع القيم العليا للمتغير التابع وبالعكس، فالعلاقة عكسية أي الزيادة في المستقل يقابلها تناقص بالمتغير التابع، ولمعرفة نوع العلاقة فيما إذا كانت خطية أو منحنية، نقارن الفروق بين كل قيمة للمتغير المستقل والقيمة السابقة لها مع مثيلاتها للمتغير التابع، فإذا كانت متقاربة فالعلاقة خطية، أما إذا كانت متباعدة فالعلاقة غير خطية .

8-2-1-2 شكل الانتشار (Scatter diagram)

يقصد بشكل الانتشار تمثيل البيانات عن الظاهرتين على المحاور الإحداثية، حيث يمثل المتغير المستقل X على محور السينات والمتغير التابع Y على محور العيّنات، وتمثل كل ثنائية (X, Y) بنقطة على المستوى الإحداثي، والنتيجة تكون شكلاً مؤلفاً من عدد كبير من النقاط يساوي عدد المشاهدات المأخوذة عن متغيرين، حيث يدعى هذا الشكل بشكل الانتشار .

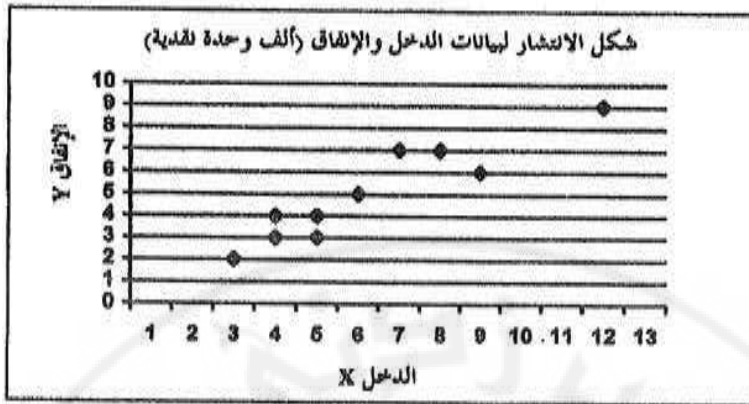
مثال (8-1) :

الجدول الآتي يمثل بيانات الدخل الشهري، والإنفاق الشهري لعينة مؤلفة من 10 أسر .

جدول رقم (8-1)

1	2	3	4	5	6	7	8	9	10	الأسرة
4	5	3	6	7	8	9	4	5	12	الدخل (بالآلاف الوحدات)
3	4	2	5	7	7	6	4	3	9	الإنفاق (بالآلاف الوحدات)

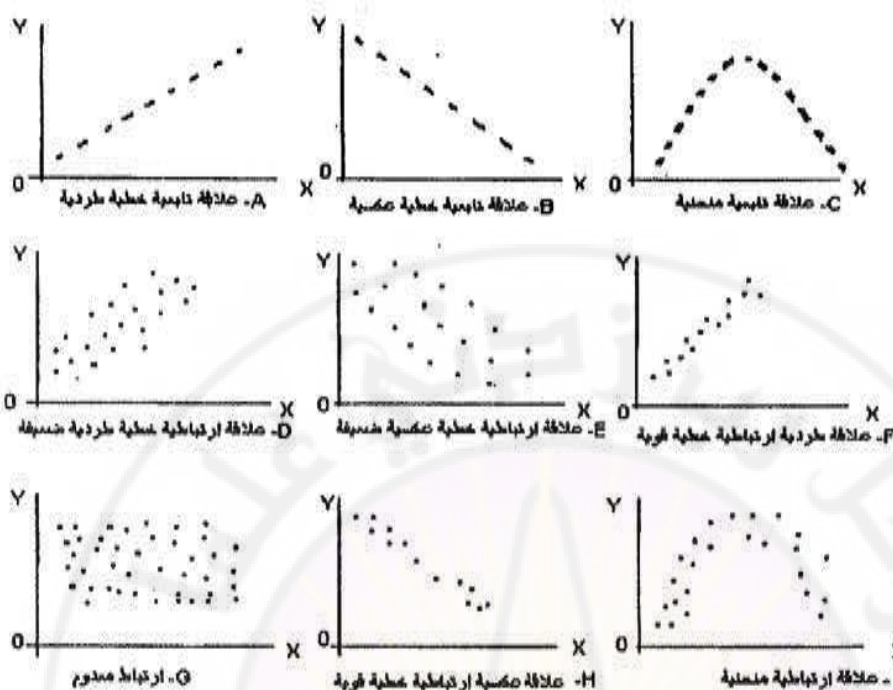
إن رسم شكل الانتشار للبيانات السابقة يعطي الشكل الآتي رقم (8-1).



الشكل رقم (8-1)

يمكننا من خلال هذا الشكل ملاحظة طبيعة العلاقة، حيث القيم الدنيا للدخل تتوافق مع القيم الدنيا للإنفاق والعليا مع العليا مما يبين أن العلاقة طردية. كما يمكن من الشكل معرفة متانة العلاقة، حيث تكون العلاقة تابعة عندما تقع النقاط على منحنى واحد، وتكون ارتباطية وقوية كلما اقتربت النقاط، من المنحنى، ففي الشكل رقم (8-1) النقاط تنتشر حول مستقيم يمكن تمريره من بين النقاط مما يسمح لنا بالقول إن العلاقة خطية وقوية. وبشكل عام يمكن أن يأخذ شكل الانتشار أحد الأشكال الآتية الواردة في الشكل رقم (8-2)، ومن خلال قراءة هذه الأشكال يمكن تحديد طبيعة ونوع ومتانة العلاقة بين المتغيرات، فإذا كانت نقاط الانتشار تقع على مستقيم واحد تكون العلاقة قوية جداً، وبالتالي يكون الارتباط تاماً. ويمكن أن تكون العلاقة عكسية أو طردية حسب موقع المستقيم الأشكال: A, B (الشكل (8-2))، وإذا كانت تقع على منحنى واحد كانت العلاقة تابعة منحنية الشكل ((8-2) - C). وعندما تكون نقاط الانتشار تشكل حزمة لا تقع على منحنى واحد تكون العلاقة طردية، وتكون العلاقة أقوى كلما قل عرض الحزمة، لاحظ الأشكال D, E, F, H, I في الشكل رقم (8-2) وعندما تكون النقاط مبعثرة دون أن يكون لها اتجاه معين يكون الارتباط معدوماً بين المتغيرين الشكل ((8-2) G)⁽¹⁾.

(1) لمزيد من المعلومات انظر كتاب: مبادئ الإحصاء، للأستاذ الدكتور ناظم حيدر، الصادر في دمشق عن مطبعة ابن حيان للعام 1982-1983.



الشكل (8-2)

8-2-1-3 معامل الارتباط (Coefficient of correlation)

يساعد معامل ارتباط بيرسون (Pearson) على تحديد متانة وطبيعة العلاقة بين المتغيرين X و Y وتعطى قيمة هذا المعامل بالعلاقة :

$$r_{xy} = \frac{\sum (x - \bar{x}) (y - \bar{y})}{n S_x S_y} \quad (1)$$

حيث : $\bar{x}$ و $\bar{y}$ الأوساط الحسابية للمتغير المستقل والمتغير التابع على التوالي .

S_x و S_y الانحرافات المعيارية للمتغير المستقل والمتغير التابع.

n عدد المشاهدات .

ويمكن إصلاح العلاقة السابقة وكتابتها بأشكال أخرى أكثر سهولة لحساب معامل الارتباط :

$$r_{xy} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (2)$$

$$r_{xy} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{[n\sum x^2 - n\bar{x}^2][\sum y^2 - n\bar{y}^2]}} \quad (3)$$

وسوف نوضح فيما يأتي كيفية حساب معامل الارتباط بالعلاقات الثلاث السابقة من خلال العودة إلى المثال رقم (8-1) .

جدول رقم (8-2)

الأسرة	الإنفاق (الف) Y	الإنفاق (الف) Y	xy	x*x	y*y	(x-x)	(x-y)	(x-x) * (y-y)
1	4	3	12	16	9	-2.3	-2	4.6
2	5	4	20	25	16	-1.3	-1	1.3
3	3	2	6	9	4	-3.3	-3	9.9
4	6	5	30	36	25	-0.3	0	0
5	7	7	49	49	49	0.7	2	1.4
6	8	7	56	64	49	1.7	2	3.4
7	9	6	54	81	36	2.7	1	2.7
8	4	4	16	16	16	-2.3	-1	2.3
9	5	3	15	25	9	-1.3	-2	2.6
10	12	9	108	144	81	5.7	4	22.8
المجموع	63	50	366	465	294	0	0	51

الحل حسب الصيغة رقم (1) .

نوجد الانحراف المعياري لـ x وكذلك لـ y :

$$S_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$$

$$= \sqrt{\frac{465}{10} - (6.3)^2} = 2.6096$$

$$S_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2}$$

$$= \sqrt{\frac{294}{10} - (5)^2} = 2.0976$$

ومنه :

$$r_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n S_x S_y}$$

$$= \frac{51}{(10)(2.6096)(2.0976)} = + 0.9317$$

الحل حسب الصيغة رقم (2) :

$$r_{xy} = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

$$= \frac{10(366) - (63)(50)}{\sqrt{[10(465) - (63)^2][10(294) - (50)^2]}} = + 0.9317$$

الحل حسب الصيغة رقم (3) :

$$r_{xy} = \frac{\sum xy - n \bar{x} \bar{y}}{\sqrt{[\sum x^2 - n \bar{x}^2][\sum y^2 - n \bar{y}^2]}}$$

$$= \frac{366 - 10(6.3)(5)}{\sqrt{[465 - 10(6.3)^2][294 - 10(5)^2]}} = + 0.9317$$

8-2-1-3-1 مزايا ومساوئ معامل ارتباط بيرسون :

- 1- يأخذ معامل ارتباط r قيمة تتراوح بين -1 إلى $+1$ ويمكن إثبات ذلك رياضياً.
- 2- تكون قيمة r تساوي الواحد عندما تكون العلاقة تامة، وتبلغ قيمته الصفر عندما يكون الارتباط معدوماً، وتقترب من الواحد عندما تكون العلاقة قوية جداً.
- 3- يأخذ معامل الارتباط قيمة موجبة عندما تكون العلاقة طردية، وسالبة عندما تكون عكسية.
- 4- قيمة معامل الارتباط مجردة من وحدات القياس، وهذه من أهم مزاياه حيث يتم التخلص من وحدات القياس عند التقسيم على الانحرافات المعيارية (لاحظ الصيغة 1) في حين أن أهم مساوئ معامل ارتباط بيرسون أنه لا يعبر عن متانة العلاقة بشكل صحيح إلا إذا كانت العلاقة خطية، أما إذا كانت العلاقة منحنية، فإن قيمته لا تعبر عن مكانة العلاقة، وفي هذه الحالة يجب استخدام مقاييس أخرى، كما أنه لا يمكن استخدام معامل بيرسون في حالة المتغيرات النوعية.

8-2-1-4 معامل ارتباط بيرسون للجداول التكرارية المزدوجة :

تبقى الصيغة السابقة لمعامل ارتباط بيرسون هي نفسها، مع الأخذ بالاعتبار عدد تكرارات كل من المتغير التابع والمتغير المستقل، حيث نعتبر أوساط الفئات في الجداول التكرارية المزدوجة في السطر والعمود للمتغير المستقل x ، والمتغير التابع y ، وعليه تصبح الصيغة السابقة لمعامل الارتباط كما يأتي :

$$r_{xy} = \frac{\sum \sum f_{ij} x_i y_j - n \bar{x} \bar{y}}{n S_x S_y}$$

$$r_{xy} = \frac{n \sum f_{ij} x_i y_j - \sum f_i x_i \sum f_j y_j}{\sqrt{[n \sum f_i x_i^2 - (\sum f_i x_i)^2][n \sum f_j y_j^2 - (\sum f_j y_j)^2]}}$$

حيث f_{ij} : عدد التكرارات عندما x يأخذ القيمة x_i و y يأخذ القيمة y_j .

ولتوضيح كيفية حساب معامل الارتباط للجدول التكراري المزوج نورد المثال الآتي:

مثال (8-2) :

يمثل الجدول رقم (8-3) بيانات عن الغلة الزراعية من القمح لمدة قطعة زراعية استخدمت فيها كميات مختلفة من السماد.

والمطلوب: حساب معامل ارتباط بيرسون بين كمية السماد والغلة مستخدماً الصيغتين السابقتين لمعامل الارتباط بيرسون .

جدول رقم (8-3)

فئات السماد (بالكغ)		2-4	4-6	6-8	8-10	10-12	المجموع			
فئات الغلة 10 كغ	أوساط الفئات x_i	3	5	7	9	11	f_j	$f_j y_j$	$f_j y_j^2$	$f_{ij} x_i y_j$
10-12	11	5	-	-	-	-	5	55	605	165
12-14	13	3	14	-	-	-	17	221	2873	1027
14-16	15	-	5	15	6	-	26	390	5850	2760
16-18	17	-	1	10	20	2	33	561	9537	4709
18-20	19	-	-	-	4	15	19	361	6859	3819
المجموع	f_i	8	20	25	30	17	100	1588	25724	12480
	$f_j x_i$	24	100	175	270	187	756			
	$f_j x_i^2$	72	500	1225	2430	2057	6282			
	$f_{ij} x_i y_j$	282	1370	2765	4554	3509	12480			

الحل حسب الصيغة الأولى :

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{756}{100} = 7.56$$

$$\bar{y} = \frac{\sum f_j y_j}{\sum f_j} = \frac{1588}{100} = 15.88$$

$$s_x = \sqrt{\frac{\sum f_i y_i^2}{\sum f_i} - \left(\frac{\sum f_i x_i}{\sum f_i} \right)^2}$$

$$= \sqrt{\frac{6282}{100} - \left(\frac{756}{100} \right)^2} = 2.25$$

$$r_{xy} = \frac{\sum \sum f_{ij} x_i y_j - n \bar{x} \bar{y}}{n S_x S_y}$$

$$= \frac{12480 - 100(7.56)(15.88)}{100(2.38)(2.25)}$$

$$= + 0.88$$

$$r_{xy} = \frac{n \sum \sum f_{ij} x_i y_j - (\sum f_i x_i)(\sum f_j y_j)}{\sqrt{[n \sum f_i x_i^2 - (\sum f_i x_i)^2][n \sum f_j y_j^2 - (\sum f_j y_j)^2]}}$$

$$r_{xy} = \frac{100(12480) - (756)(1588)}{\sqrt{[100(6282) - (756)^2][100(25724) - (1588)^2]}} = + 0.88$$

وتحب الإشارة هنا إلى أنه يمكن استخدام أسلوب الانحرافات المختصرة في الصيغة الأولى أو الثانية لحساب معامل الارتباط .

8-2-1-5 الدلالة الإحصائية لمعامل الارتباط

عند دراسة الارتباط بين متغيرين فإننا في أغلب الأحيان نستخدم عدداً من المشاهدات (n) وهي عبارة عن عينة عشوائية، حيث نحصل من هذه العينة على قيمة لمعامل الارتباط ولكن السؤال الذي يطرح هو: هل هذه القيمة لمعامل الارتباط هي فعلاً موجودة في المجتمع الإحصائي أم لا ؟ وللتأكد من ذلك نقوم عادةً باختبار دلالة أو معنوية معامل الارتباط، وجوهر هذا الاختبار يقوم على أساس اختبار الفرضية القائلة بأن قيمة معامل الارتباط في المجتمع الإحصائي تساوي الصفر، فإذا ما رفضنا هذه الفرضية نكون عندها قد تأكدنا من وجود ارتباط في المجتمع، والمعامل الناتج يعبر عنه أما إذا قبلنا الفرضية فهذا يعني أن قيمة معامل الارتباط لا تعبر عن الارتباط في المجتمع الإحصائي، وإن قيمته يمكن أن تساوي الصفر، وبالتالي ليس لمعامل الارتباط أية دلالة أو قيمة غير معنوية .

ولإجراء الاختبار نستخدم التحويل t حيث :

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

وندعو هذه القيمة بالقيمة المحسوبة لـ t ، نقارن هذه القيمة مع القيمة النظرية لـ t المستخرجة من جداول ستودنت عند مستوى دلالة α محدد وعدد درجات حرية $n-2$ حيث n عدد المشاهدات .

لنتأكد الآن من الدلالة الإحصائية لمعامل الارتباط الذي أوجدناه سابقاً من بيانات الجدول رقم (8-2) حيث وجدنا أن : $r = 0.93$ وإن $n = 10$:

$$t = 0.93 \sqrt{\frac{10-2}{1-(0.93)^2}} = 7.15$$

أما قيمة t النظرية عند مستوى معنوية 5% و 8 درجات حرية :

$$t_{(0.05)(8)} = 2.3$$

نلاحظ أن قيمة t المحسوبة أكبر من t النظرية، وبالتالي نرفض الفرضية من أن قيمة معامل الارتباط بالمجتمع الإحصائي تساوي الصفر، وبالتالي قيمة r التي حصلنا عليها هي قيمة معنوية .

أو بطريقة أخرى نحسب الخطأ المعياري لمعامل الارتباط :

$$S_r = \sqrt{\frac{1-r^2}{n-2}}$$

فإذا كانت قيمة معامل الارتباط r أكبر من ثلاثة أمثال خطئه المعياري، يكون لمعامل الارتباط أهمية إحصائية أو دلالة إحصائية، وذلك باحتمال 99.7% ، وبتطبيق هذه الطريقة على بيانات المثال رقم (1) نجد :

$$S_r = \sqrt{\frac{1-(0.93)^2}{10-2}} = 0.129$$

وهنا نلاحظ أن $r > 3 S_r$ وبالتالي نقول إن لمعامل الارتباط دلالة إحصائية، وإن قيمته في المجتمع تختلف عن الصفر .

ونوجه العناية إلى أننا سوف نقوم بدراسة هذا الموضوع بمزيد من التفصيل في مقرر الإحصاء التطبيقي للسنة الثانية ضمن فقرة الاختبارات الإحصائية لمعاملات الارتباط والانحدار .

8-2-2 الارتباط للمتغيرات النوعية

شاهدنا سابقاً بعض الأساليب التي نستطيع بمساعدتها الكشف عن مكانة العلاقة الارتباطية للمتغيرات الكمية، ولكن أياً من هذه الأساليب لا يصلح لحساب مكانة العلاقة الارتباطية للمتغيرات النوعية، حيث تستخدم في هذا المجال عدة أساليب مختلفة تبعاً للحالة المدروسة، وسوف نقوم فيما يأتي بدراسة هذه الأساليب بالتفصيل .

8-2-2-1 معامل الاقتران

يستخدم معامل الاقتران عندما يكون لدينا متغيران نوعيان، ولكل متغير قيمتان أو صفتان كأن ندرس العلاقة بين الإصابة بالسرطان لمجموعة من الأشخاص مدخنين وغير مدخنين، فالمتغيران هنا المرضى (مصاب بالسرطان وغير مصاب) ، والتدخين (مدخن وغير مدخن) ، أو دراسة العلاقة بين لون الزهرة ورائحتها ..
ولحساب معامل الاقتران نقسم المفردات حسب المتغيرات والصفات كما في الجدول (8-4) .

جدول رقم (8-4)

المتغير A	الصفة 1	الصفة 2	المجموع
المتغير B			
الصفة 1	n_{11}	n_{12}	$'n_2$
الصفة 2	n_{21}	n_{22}	$'n_2$
المجموع	n_1	n_2	N

$$n_{11} + n_{12} + n_{21} + n_{22} = N \quad \text{حيث :}$$

ويحسب معامل الاقتران كما يأتي :

$$r_0 = \frac{n_{11} * n_{22} - n_{12} * n_{21}}{n_{11} * n_{22} + n_{12} * n_{21}}$$

وتكون العلاقة بين المتغيرات قوية جداً كلما اقتربت قيمة معامل الاقتران من الواحد وبشكل عام نقول إن هناك علاقة عندما تكون قيمة r_0 أكبر من 0.5.

مثال (8-3) :

مهدف دراسة العلاقة بين التدخين والإصابة بالتهاب اللثة قام أحد الباحثين بجمع بيانات عن 200 شخص كان توزيعهم حسب المتغيرات والصفات كما في الجدول (8-5).

جدول رقم (8-5)

المتغير A \ المتغير B	مدخن	غير مدخن	المجموع
مصاب	70	10	80
غير مصاب	20	100	120
المجموع	90	110	200

$$r_0 = \frac{70 * 100 - 20 * 10}{70 * 100 + 20 * 10} = \frac{6800}{7200} = 0.944$$

وهذا يعني أن هناك علاقة قوية بين الإصابة بالتهاب اللثة والتدخين. وتجب الإشارة هنا إلى أنه ليس لإشارة r_0 أي مدلول، لأن إشارته تختلف حسب ترتيب الصفات في الجدول، لذلك يفضل حسابه بالقيمة المطلقة للدلالة على وجود علاقة أو لا. (لاحظ لو بدلنا السطر الأول بالثاني في الجدول السابق لأصبحت قيمة معامل الاقتران -0.944).

8-2-2-2 معامل التوافق

رأينا أن معامل الاقتران يستخدم لقياس متانة العلاقة بين متغيرين نوعيين، لكل منهما صفتان (أي الجدول يحوي أربع خلايا)، في حين لا يمكن استخدامه إذا كان لأحد المتغيرين على الأقل أكثر من صفتين حيث يصبح الجدول في هذه الحالة مركباً في أكثر من أربع خلايا، والمعامل الذي يمكن استخدامه هو معامل التوافق وصيغته كما يأتي:

$$r_A = \sqrt{\frac{G-1}{G}}$$

حيث :

$$G = \sum \left(\frac{f_{ij}^2}{f_i * f_j} \right)$$

f_{ij} : عدد التكرارات في الخلية الواقعة في السطر i والعمود j .

f_i : مجموع السطر i .

f_j : مجموع العمود j .

مثال (8-4) :

مهدف دراسة تأثير الحالة التعليمية للأُم على مستوى الابن في المدرسة الابتدائية
تم جمع البيانات الآتية عن 200 تلميذ في إحدى المدارس .

جدول رقم (8-6)

الحالة التعليمية للأُم مستوى الابن في المدرسة	أمية وابتدائية	ثانوية وإعدادية	فوق الثانوية	المجموع
متفوق	5	15	30	50
وسط	5	40	15	60
ضعيف	70	15	5	90
المجموع	80	70	50	100

$$G = \frac{5^2}{80 \times 50} + \frac{5^2}{80 \times 60} + \frac{70^2}{80 \times 90} + \frac{15^2}{70 \times 50} + \frac{40^2}{70 \times 60} \\ + \frac{15^2}{70 \times 90} + \frac{30^2}{50 \times 50} + \frac{15^2}{60 \times 50} + \frac{5^2}{50 \times 90} = 1.613$$

$$r_A = \sqrt{\frac{G - 1}{G}} = \sqrt{\frac{1.613 - 1}{1.613}} = 0.616$$

وهذا يعني أن هناك علاقة قوية بين حالة الأم التعليمية ومستوى الابن في الصف. وبشكل عام نقول إن هناك علاقة قوية أو جيدة بين المتغيرين، كلما كانت قيمة معامل التوافق أكبر من 0.5 .

8-2-2-3 معامل ارتباط الرتب (Coefficient of Rank Correlation)

يستخدم معامل ارتباط الرتب أو معامل ارتباط سبيرمان (Sperman) نسبة إلى واضعه، لقياس متانة العلاقة بين متغيرين نوعيين، لكل منهما عدد كبير من الصفات، أو إذا كان أحد المتغيرين نوعياً والأخر كمية. وكذلك الأمر يمكن استخدامه في حالة المتغيرات الكمية أحياناً لاختصار الحسابات .

يتم حساب ارتباط الرتب باستبدال الصفات النوعية بقيم عددية تعبر عنها وتسمى الرتب Rank، حيث يتم وضعها عن طريق إيجاد العلاقة الداخلية للصفات ولكل متغير على حده، فنبدأ بأضعف الصفات مثلاً ونعطيها الرتبة 1 والأقوى منها رقم 2 وهكذا، بحيث يتولد لدينا متوالية عددية لكل متغير يكون مجموعها متساوياً، لأن عدد الصفات لكل متغير متساوية، وبعد ذلك نحسب معامل ارتباط الرتب بالعلاقة:

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

حيث D : فرق رتبة X من رتبة Y لكل ثنائية .

n : عدد المشاهدات .

مثال (8-5) :

ليكن لدينا الجدول الآتي الذي يمثل البيانات الخاصة بشكل علبه الحلوى والطلب على هذا النوع من الحلوى .

جدول رقم (8-7)

شكل علبه الحلوى (x)	نوع الطلب (y)	رتب x	رتب y	D	D ²
علبه بسيطة دون رسوم	ضعيف	1	1	0	0
علبه بسيطة برسوم	وسط	2	2	0	0
علبه بسيطة برسوم وغللاف	جيد جداً	3	4	1	1
علبه مزخرفة	جيد	4	3	1	1
علبه حديدية تقليدية	عالي جداً	5	5	0	0
المجموع					2

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6 \times 2}{5(5^2 - 1)} = 1 - \frac{12}{120} = 0.9$$

أي أن هناك علاقة قوية جداً بين شكل العلبه ونوع الطلب عليها. ونحب الإشارة هنا إلى أن معامل ارتباط الرتب يتمتع بالمزايا نفسها التي يتمتع بها معامل ارتباط بيرسون، باعتبار أنه مستنتج من معامل ارتباط بيرسون بعد تعويض $\sum X$ و $\sum y$ و $\sum X^2$ و $\sum Y^2$ بما يقابلها باعتبارها متواليات عددية بعد إعطاء الرتب للمتغيرين .

8-2-2-3-1 معامل ارتباط الرتب في حالة تكرار عدد من الصفات:

مثال (8-6) :

تبين البيانات الآتية غزارة الأمطار ولون الغيم في إحدى المدن في فصل الشتاء .

جدول رقم (8-8)

لون الغيم (x)	كمية المطر ميليمتر/الدقيقة (y)	رتب x	رتب y	D	D ²
أبيض متفرق	0	1	1	0	0
أبيض كثيف	2	2	2	0	0
رمادي	3	3	3.5	0.5 -	0.25
رمادي غامق	3	4	3.5	0.5	0.25
أسود	4	5	5	0	0
المجموع		15	15		0.5

نلاحظ من هذه البيانات أن أحد المتغيرين نوعي وهو لون الغيوم والآخر كمي. ولقد تم وضع رتب x مبتدئين بالأضعف حسب معرفتنا بحالات هطول المطر، والأضعف في هذه الحالة إذا كان لون الغيم أبيض متفرقاً، وأعطيناه الرتبة 1، والأفضل هو أبيض كثيف 2 وهكذا... أما المتغير الآخر هو كمية المطر وبياناته كمية، وهنا لن تكون هناك صعوبة في وضع الرتب، حيث أقل هذه البيانات سوف يأخذ الرتبة 1 وهكذا، ولكن هناك مشكلة تكرار إحدى صفات الرقم 3، في هذه الحالة تعطى الرتب كما يأتي : حسب الترتيب أول صفة يجب أن تعطى الرتبة 3 في البيانات، والصفة المكررة الأخرى يجب أن تعطى الرتبة 4 ، ولكن بما أن الصفات متساوية فيجب أن تساوي بين الرتبتين، ويتم ذلك من خلال إيجاد الوسط الحسابي للرتبتين 3 و 4 وهو 3.5، فيعطى كل منها الرتبة 3.5. بهذه الطريقة نكون قد حافظنا على مجموع الرتب لكل من المتغيرين بحيث يكون متساوياً، انظر الجدول رقم (8-8) من المثال (8-6) .

8-2-2-3-2 معامل ارتباط الرتب للجدول التكرارية :

يحسب معامل ارتباط الرتب للجدول التكرارية بالعلاقة :

$$r_s = \frac{S_D^2 - S_x^2 - S_y^2}{2 * S_x * S_y}$$

- ٣٢٤ -

حيث S_x و S_y : الانحراف المعياري لرتب x ورتب y على التوالي ومربعهما هو تباينهما .

S_D^2 : تباين مجموع الرتب لكل سطر وعمود .

وسوف نبين كيفية حساب معامل ارتباط الرتب للجدول التكراري من خلال

المثال الآتي:

مثال (7-8) :

ليكن لدينا البيانات الآتية عن حالة الأم التعليمية وعدد الولادات خلال حياتها .
والمطلوب: حساب معامل ارتباط الرتب لهذه البيانات .

جدول رقم (8-9)

عدد الولادات حالة الأم التعليمية	2	3	4	5	6	المجموع
أمية	-	-	5	5	30	40
ابتدائية	-	-	5	15	10	30
إعدادية وثانوية	5	8	7	-	-	20
فوق الثانوية	10	-	-	-	-	10
المجموع	15	8	17	20	40	100

الحل :

1- نضع رتب x و y ونحسب تباينها وانحرافها المعياري .

جدول رقم (8-10)

عدد الولادات y	رتب y R_y	f_y	$R_y f_y$	$f_y R_y^2$
2	1	15	15	15
3	2	8	16	32
4	3	17	51	153
5	4	20	80	320
6	5	40	200	1000
المجموع		100	362	1520

$$S_y^2 = \frac{\sum f_y R_y^2}{\sum f_y} - \left(\frac{\sum f_y R_y}{\sum f_y} \right)^2$$

$$= \frac{1520}{100} - \left(\frac{362}{100} \right)^2 = 2.09$$

$$S_y = \sqrt{2.09} = 1.447$$

جدول رقم (8-11)

حالة الأم التعليمية x	رتبة x R_x	f_x	$f_x R_x$	$f_x R_x^2$
أمية	1	40	40	40
ابتدائية	2	30	60	120
إعدادية وثانوية	3	20	60	180
فوق الثانوية	4	10	40	160
المجموع		100	200	500

$$S_x^2 = \frac{\sum f_x R_x^2}{\sum f_x} - \left(\frac{\sum f_x R_x}{\sum f_x} \right)^2$$

$$S_x^2 = \frac{500}{100} - \left(\frac{200}{100} \right)^2 = 1$$

$$S_x = \sqrt{1} = 1$$

2- نحسب تباين مجموع الرتب لكل سطر وعمود، مع أخذ التكرارات لكل مجموع بالاعتبار S_D^2 ، ولذلك نعد الجدول الآتي الذي يبين الرتب للأسطر والأعمدة والتكرارات :

جدول رقم (8-12)

Ry \ Rx	1	2	3	4	5
1	-	-	5	5	30
2	-	-	5	15	10
3	5	8	7	-	-
4	10	-	-	-	-

من الجدول رقم (8-12) نحسب بمجموع الرتب للعمود والسطر مثلاً $1 + 1 = 2$ ولكن لا يقابلها تكرارات، ثم $1 + 2 = 3$ أو $2 + 1 = 3$ كذلك لا يقابلها تكرارات في مثالنا. وهكذا، ومن ثم نعد الجدول الآتي رقم (8-13) للمجموع D. أما المجموع 4 فهو يحدث $2 + 2$ أو $3 + 1$ أو $1 + 3$ والتكرارات F هي $0 + 5 + 5 = 10$ انظر الجدول رقم (8-13).

جدول رقم (8-13)

مجموع الرتب D	f	Df	fD ²
2	-	-	-
3	-	-	-
4	10	40	160
5	28	140	700
6	52	312	1872
7	10	70	490
8	-	-	-
9	-	-	-
المجموع	10	562	3222

نحسب S_D^2 من العلاقة :

$$S_D^2 = \frac{\sum f D^2}{\sum f} - \left(\frac{\sum f D}{\sum f} \right)^2$$

$$= \frac{3222}{100} - \left(\frac{562}{100} \right)^2 = 0.6356$$

نحسب معامل الارتباط من العلاقة :

$$r_s = \frac{S_D^2 - S_x^2 - S_y^2}{2 * S_x * S_y}$$

$$= \frac{0.6356 - 1 - 2.09}{2 * 1 * 1.447} = - 0.849$$

وهذا يعني أن العلاقة قوية جداً ، وعكسية بين المستوى التعليمي للأم وعدد الولادات.

8-3 تحليل الانحدار Regression Analysis :

رأينا سابقاً كيف يمكننا تحديد وجود أو عدم وجود علاقة بين المتغيرات، وكيف نستطيع تحديد طبيعة العلاقة ومكانتها، ولكن هل يمكننا بالإضافة إلى ذلك تحديد مدى تأثير المتغير المستقل X على المتغير التابع Y ؟ وحساب مقدار التغير في المتغير التابع عندما يزداد المتغير المستقل وحدة واحدة؟ وتقدير قيمة المتغير التابع عند قيمة محددة للمتغير المستقل؟ إن الإجابة على هذه الأسئلة يكون من خلال دراسة الانحدار Regression. والانحدار هو تمثيل العلاقة بين المتغيرات بمعادلة رياضية، حيث يمكن تقسيم الانحدار حسب نوع المعادلة وعدد المتغيرات إلى :

1- الانحدار البسيط: وهو دراسة العلاقة بين متغيرين أحدهما مستقل والآخر تابع، وهو بدوره يقسم إلى قسمين :

- الانحدار الخطي البسيط Simple linear Regression :

وذلك عندما يمكن تمثيل العلاقة بخط مستقيم .

- الانحدار غير الخطي البسيط Simple non-linear Regression :

وذلك عندما تكون العلاقة بين المتغيرين تمثل بعلاقة غير معادلة المستقيم: درجة ثانية - ثلاثة - أسية - لوغاريتمية .

2- الانحدار المتعدد Multiple Regression : وهو دراسة الانحدار بين عدة متغيرات مستقلة ومتغير واحد تابع، ويقسم بدوره إلى قسمين:

- الانحدار المتعدد الخطي Multiple non-Linear Regression :

وتكون بتمثيل العلاقة بين المتغيرات المستقلة والمتغير التابع بعلاقة غير خطية . وسوف نقوم بدراسة الانحدار الخطي البسيط بشكل مفصل، وذلك لاستعمالاته المتعددة ، ولكي يكون مدخلاً لفهم علاقات الانحدار الأخرى ، ونترك

الانحدار المتعدد الخطي أو اللاخطي بالإضافة إلى الارتباط والانحدار المتعدد للسنوات القادمة .

8-3-1 الانحدار الخطي البسيط :

إن دراسة الانحدار بين متغيرين x و y تبدأ بتمثيل العلاقة بينهما بخط مستقيم يدعى مستقيم الانحدار، والمعادلة الممثلة لهذا المستقيم تسمى بمعادلة الانحدار وهي من الشكل :

$$\hat{y}_i = a + b x_i$$

حيث : $\hat{y}_i$ - القيمة المقدرة أو المحسوبة للمتغير التابع.

a - ثابت الانحدار .

b - معامل الانحدار .

ومن الممكن إيجاد انحدار y على x ، أو انحدار x على y ، غير أنه يفضل التركيز على انحدار y على x نظراً لأنه أكثر أهمية من انحدار x على y من جهة، وباعتبار أننا حددنا مسبقاً المتغير المستقل والمتغير التابع من جهة أخرى.

بعد إيجاد معادلة الانحدار الممثلة للعلاقة يمكن الاستمرار بدراسة الانحدار باستخدام أساليب تحليل أخرى سوف نتعرض لها لاحقاً ولكن من الضروري أولاً بيان كيفية إيجاد معادلة الانحدار .

8-3-1-1 إيجاد معادلة مستقيم الانحدار :

إن التدقيق في الشكل رقم (8-1) أو في الشكل رقم (8-3) الذي يمثل شكل الانتشار لبيانات الدخل والإنفاق لـ 10 أسر كما هو وارد في الجدول رقم (8-14).

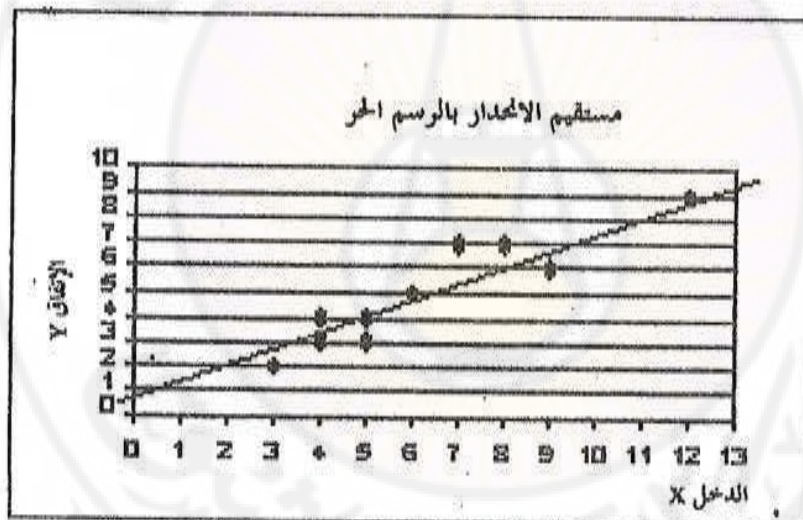
جدول رقم (8-14)

رقم الأسرة	1	2	3	4	5	6	7	8	9	10
دخل الأسرة (1000) وحدة	4	5	3	6	7	8	9	4	5	12
إنفاق الأسرة (1000) وحدة	3	4	2	5	7	7	6	4	3	9

نرى أنه يمكن من خلال نقاط الانتشار رسم عدد لا نهائي من المستقيمات، وذلك لتعدد الحصول على مستقيم يمر من كل هذه النقاط، وبالتالي يمكن اعتبار أي مستقيم يمر خلال الشكل ممثلاً للعلاقة، ويكون هذا المستقيم ممثلاً بشكل أفضل، بالاعتماد على طريقة علمية ما في إيجادها، حيث توجد عدة طرائق تستخدم لإيجاد أفضل مستقيم، وسوف نناقش بعضها أدناه .

- طريقة الرسم الحر في إيجاد مستقيم الانحدار Freehand Method :

تعتمد هذه الطريقة في إيجاد مستقيم الانحدار على رسم مستقيم يمر من جميع نقاط الانتشار أو معظمها، وإن عملية الحصول على المستقيم الأفضل يعتمد على الخبرة في عملية رسم . انظر الشكل (8-3) .



الشكل رقم (8-3)

لإيجاد معادلة الانحدار $\hat{y}_i = a + bX_i$ لا بد من إيجاد الثوابت a , b ، حيث يتم ذلك كما يأتي : إن الثابت b هو عبارة عن ميل المستقيم، ويعطى بالعلاقة :

$$b = \frac{\Delta Y}{\Delta X}$$

نفتش عن نقطتين يمر منها المستقيم ونحسب ΔY و ΔX حيث نرى في الشكل (8-3) أن المستقيم يمر من النقطتين (5, 4) و (12, 9)، وبالتالي يكون :

$$\Delta Y = 9 - 4 = 5$$

$$\Delta X = 12 - 5 = 7$$

وبالتالي :

$$b = \frac{5}{7} = 0.714$$

أما قيمة الثابت a فهي عبارة عن نقطة التقاطع مع محور العينات وهي في الشكل (8-3) تساوي تقريباً 0.7. وعليه تكون معادلة الانحدار من الشكل :

$$\hat{y}_i = 0.7 + 0.714 x_i$$

تتميز هذه الطريقة بالسهولة والسرعة، ولكن من سيئاتها أنها قد تعطي نتائج غير دقيقة ومختلفة تبعاً للخبرة الشخصية، واختلاف أسلوب الرسم، حيث يمكن أن نحصل على عدد من الخطوط المستقيمة تساوي عدد القائمين على عملية الرسم .

– طريقة المربعات الصغرى The Method of Least Squares :

إن جوهر هذه الطريقة يتلخص بجعل الانحرافات بين القيم الفعلية y_i للمتغير التابع عن مستقيم الانحدار y_i أصغر ما يمكن أي :

$$\Sigma (y_i - \hat{y}_i) \rightarrow \min$$

ولكن $\sum (y_i - \hat{y}_i) = 0$ حيث $\hat{y}$ يمثل القيمة المتوسطة لـ y_i عند قيمة معينة لـ x كما سنرى فيما بعد، ولكي نتخلص من هذه النتيجة نأخذ مربعات الفروق وعليه نشكل التابع f :

$$f = \sum (y_i - \hat{y}_i)^2 \rightarrow \min$$

أو :

$$f = \sum (y_i - a - b x_i)^2 \rightarrow \min$$

نلاحظ أن قيمة هذا التابع ترتبط بالمتغيرين a و b ، حيث تزداد بازديادهما وتنقص بنقصانهما، وإذا ما نظرنا إلى شكل الانتشار لوحدنا أن عدد الخطوط المستقيمة التي يمكن أن تمر عبر نقاط الشكل هو لا نهائي، وتختلف عن بعضها باختلاف أحد المتغيرين على الأقل، فإذا ما بلغ التابع F النهاية الصغرى تكون الانحرافات أقل ما يمكن، وبالتالي نشق التابع اشتقاقاً جزئياً بالنسبة لـ a أولاً، ومن ثم لـ b ونعدم بعد ذلك المشتق :

$$\frac{\partial F}{\partial a} = 2 \sum (y_i - a - b x_i) (-1) = 0$$

وبإصلاح العلاقة نحصل على المعادلة الآتية التي تدعى المعادلة الطبيعية الأولى:

$$\sum Y_i = na + b \sum X_i$$

$$\frac{\partial F}{\partial b} = 2 \sum (y_i - a - b x_i) (-x_i) = 0$$

ومنه :

$$\sum x_i y_i = a \sum x_i + b \sum x_i^2$$

وهي المعادلة الطبيعية الثانية .

وبالحل المشترك لهاتين المعادلتين الطبيعيتين :

$$\sum y_i = na + b \sum x_i$$

$$\sum x_i y_i = a \sum x_i + b \sum x_i^2$$

عن طريق استخدام إحدى الطرائق المعروفة لحل معادلتين مجهولين نحصل على قيمة الثوابت a و b المجهولين، ومن هذه الطرائق نذكر :

- الحذف بالتعويض .
- الحذف بالطرح .
- الحل المصفوفي .
- المعينات .

ونلاحظ من المعادلتين السابقتين أن جميع القيم الأخرى يمكن الحصول عليها من البيانات الفعلية .

سوف نقوم فيما يلي بإيجاد معادلة مستقيم الانحدار للبيانات الواردة في الجدول رقم (8-14)، والتي أعدنا كتابتها بالجدول رقم (8-15) لإيجاد القيم المطلوبة لحل المعادلتين :

جدول رقم (8-15)

رقم الأسرة	x	y	$x*y$	x^2	y^2
1	4	3	12	16	9
2	5	4	20	25	16
3	3	2	6	9	4
4	6	5	30	36	25
5	7	7	49	49	49
6	8	7	56	64	49
7	9	6	54	81	36
8	4	4	16	16	16
9	5	3	15	25	9
10	12	9	108	144	81
المجموع	63	50	366	465	294

— الحل باستخدام طريقة الحذف بالطرح :

نعوض في المعادلتين بالقيم التي حصلنا عليها من الجدول رقم (8-15) :

$$50 = 10 * a + 63 b \quad I$$

$$366 = 63 * a + 465 b \quad II$$

نلاحظ أنه لحذف أحد المتغيرين من المعادلتين بالطرح يجب ضرب المعادلة الأولى

— 6.3، وبالتالي نحصل على المعادلة الآتية :

$$315 = 63 * a + 396.9 b \quad III$$

وبطرح المعادلة III من II نحصل على المعادلة IV .

$$15 = 68.1 b \quad IV$$

ومنه :

$$b = \frac{51}{68.1} = 0.7489$$

وبتعويض قيمة b في المعادلة I نحصل على قيمة a = 0.2819 .

— الحل باستخدام طريقة المعينات (الطريقة المباشرة) :

المعين الأساسي Δ :

$$\Delta = \begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix} = n \sum x_i^2 - (\sum x_i)(\sum x_i) = n \sum x_i^2 - (\sum x_i)^2$$

المعين المساعد Δ_a :

$$\Delta_a = \begin{vmatrix} \sum y_i & \sum x_i \\ \sum x_i y_i & \sum x_i^2 \end{vmatrix} = (\sum y_i)(\sum x_i^2) - (\sum x_i y_i)(\sum x_i)$$

المعين المساعد Δ_b :

$$\Delta_b = \begin{vmatrix} n & \sum y_i \\ \sum x_i & \sum x_i y_i \end{vmatrix} = n \sum x_i y_i - \sum x_i \sum y_i$$

$$b = \frac{\Delta b}{\Delta} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)(\sum x_i)}$$

ويتقسم البسط والمقام على n لحصل على :

$$= \frac{\sum x_i y_i - \bar{x} \sum y_i}{\sum x_i^2 - \bar{x} \sum x_i}$$

وهذا يساوي أيضاً :

$$= \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2}$$

أما الثابت a :

$$a = \frac{\Delta a}{\Delta} = \frac{(\sum y_i)(\sum x_i^2) - (\sum x_i y_i)(\sum x_i)}{n \sum x_i^2 - (\sum x_i)(\sum x_i)}$$

ويمكننا بطريقة أخرى أسهل إيجاد قيمة الثابت a بعد الحصول على قيمة الثابت

b من المعادلة الطبيعية الأولى إذا قسمنا هذه المعادلة على n فإننا نحصل على :

$$\bar{y} = a + b \bar{x}$$

ومنه :

$$a = \bar{y} - b \bar{x}$$

وبتطبيق هذه العلاقات بشكل مباشر على المثال السابق نستطيع تحديد قيمة ثابتي

المعادلة a و b :

$$b = \frac{366 - 6.3 * 50}{465 - 6.3 * 63} = \frac{51}{68.1} = 0.7489$$

$$= \frac{366 - 10(6.3 * 5)}{465 - 10(6.3)^2} = \frac{51}{68.1} = 0.7489$$

$$a = \frac{50(465) - 366(63)}{10(465) - (63)(63)} = \frac{192}{681} = 0.2819$$

$$= 5 - 0.7489 \cdot 6.03 = 0.2819$$

حيث :

$$\bar{x} = \frac{\sum x}{n} = \frac{63}{10} = 6.3$$

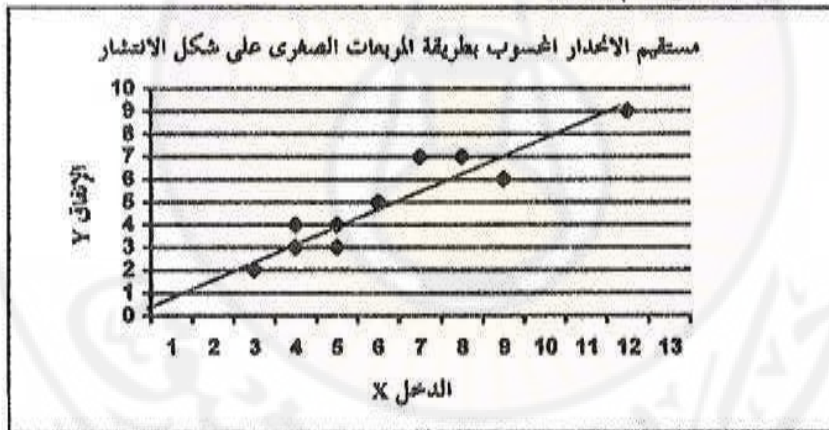
و

$$\bar{y} = \frac{\sum Y}{n} = \frac{50}{10} = 5$$

وهكذا، فإن معادلة الانحدار تكون من الشكل الآتي :

$$\hat{y}_i = 0.2819 + 0.7489 x_i$$

انظر الشكل رقم (8-4) .



الشكل رقم (8-4)

8-3-1-2 أهمية معادلة مستقيم الانحدار :

تقدم معادلة مستقيم الانحدار التي يتم إيجادها بطريقة المربعات الصغرى مجموعة من المعلومات الهامة عند تحليل البيانات وهي :

A - تعطي معادلة الانحدار القيمة التي يأخذها المتحول التابع بالمتوسط عندما يأخذ المتحول المستقل قيمة محددة. ولتوضيح ذلك نورد المثال الآتي :

مثال (8-8) :

يبين الجدول رقم (8-16) كمية السماد المستخدم وغلة القطعة من القمح لـ 12 قطعة أرض استخدمت في تجربة قام بها أحد الباحثين لمعرفة مدى تأثير السماد على الغلة فكانت النتائج :

جدول رقم (8-16)

رقم قطعة الأرض	كمية السماد (1) كيلو غرام x	الغلة (10) كيلو غرام y	x y	x ²
1	0	19	0	0
2	0	20	0	0
3	0	21	0	0
4	1	21	21	1
5	1	22	22	1
6	1	23	23	1
7	2	23	46	4
8	2	24	48	4
9	2	25	50	4
10	3	25	75	9
11	3	26	78	9
12	3	27	81	9
المجموع	18	276	444	42

لايجاد معادلة مستقيم الانحدار الممثلة للعلاقة بين كمية السماد والغلة يجب تحديد قيمة الثوابت a و b :

$$b = \frac{\sum x_i y_i - \bar{x} \sum y_i}{\sum x_i^2 - \bar{x} \sum x_i}$$

$$= \frac{444 - 1.5 * 276}{42 - 1.5 * 18} = \frac{30}{15} = 2$$

$$a = \bar{y} - b \bar{x}$$

$$= 23 - 2 * 1.5 = 20$$

$$\hat{y}_i = 20 + 2 x_i$$

إن القيمة التي تعطىها هذه المعادلة تتوضع من خلال بيانات الجدول رقم (8-16)، فعندما تكون كمية السماد 0 كغ يكون متوسط الغلة 20 وحدة (200 كغ)، وكذلك يكون متوسط الغلة 22 وحدة عند استخدام 1 كغ سماد للغلة وهكذا... فإذا عوضنا في المعادلة المشحول المستقل X بقيمة محدد ولتكن 2 كغ، فلها تأخذ الشكل الآتي :

$$\hat{y}_i = 20 + 2 * 2 = 24$$

أي عند استخدام 2 كغ سماد للقطعة، فإن الغلة سوف تكون 24 وحدة، وهي تعبر عن متوسط الغلة عند استخدام 2 كغ من السماد، انظر الجدول رقم (8-16). وهكذا فإن معادلة مستقيم الانحدار تعطي دوماً قيمة التابع بالمتوسط، حيث تبقى هذه النتيجة نفسها دون تغيير، حتى لو كانت البيانات موضع الدراسة لا تحتوي على قيمة الأوساط الحسابية للتابع كما هو الحال في مثالنا. وللتأكد من ذلك يمكننا حذف المشاهدات التي تحوي أوساطاً حسابية لـ y، وأن المعادلة سوف تبقى نفسها والقيم التي تعطىها هي القيم المتوسطة نفسها. وأن مجموع القيم الفعلية للتابع يساوي مجموع القيم المحسوبة من المعادلة .

B - نستطيع من خلال معادلة مستقيم الانحدار التنبؤ أو تقدير القيمة التي يأخذها المتغير التابع بالمتوسط، عندما تكون قيمة المتغير المستقل معدومة، وهذه القيمة هي قيمة ثابت معادلة الانحدار a ، وللتأكد من ذلك نعوض قيمة x بالصفر في معادلة الانحدار للمثال السابق فنجد :

$$\hat{y}_1 = 20 + 2 * 0 = 20$$

وهذا يعني أنه عندما تكون كمية السماد معدومة أي تساوي الصفر، فإن الغلة بالمتوسط تساوي 20 وحدة، وهي متوسط الغلة في حالة عدم استخدام السماد. انظر الجدول رقم (8-16) .

C - تساعد معادلة مستقيم الانحدار على معرفة مقدار التغير في المتغير التابع (زيادة أو نقصان)، عندما يزداد المتغير المستقل وحدة واحدة، مع العلم أن التغير سيكون بالمتوسط، وللتأكد من ذلك انظر المثال السابق: فإذا استخدمنا 1 كغ سماد فإن متوسط الغلة سوف يعادل 22 وحدة، وتصبح الغلة بالمتوسط 24 وحدة إذا استخدمنا 2 كغ سماد، أي عندما يزداد المتغير المستقل وحدة واحدة، فإن المتغير التابع سوف يزداد بالمتوسط بمقدار 2 وحدة وهي قيمة b معامل الانحدار.

D - تبلغ القيمة المقدرة للمتغير التابع $\hat{y}_1$ قيمة الوسط الحسابي للمتغير التابع عندما تكون قيمة المتغير المستقل تساوي وسطه الحسابي $\bar{x}$ أي بمعنى أن مستقيم الانحدار يمر بالنقطة $(\bar{x}, \bar{y})$ ، حيث يمكننا التأكد من ذلك من خلال المعادلة الآتية :

$$\bar{y} = a + b \bar{x}$$

E - تستخدم معادلة مستقيم الانحدار للتنبؤ بقيم المتغير التابع عندما يأخذ المتغير المستقل قيمة محددة، ولكن هنا يجب الانتباه إلى أي مدى يمكن أن تحافظ هذه العلاقة على نوعها وطبيعتها خارج حدود هذه البيانات، وبشكل عام يمكننا القول إنه يمكن استخدام المعادلة في حدود معينة من الثقة على طرفي مجال قيم المتغير المستقل .

F - تقدم معادلة مستقيم الانحدار معلومات بشكل غير مباشر عن مدى تأثير المتغير المستقل على المتغير التابع، وهذا سوف نتعرف عليه بالتفصيل من خلال دراسة التباينات .

8-3-1-3 دراسة التباينات : Study of variances

تعرفنا عند دراسة التشتت على كيفية حساب التباين لمتغير واحد وعليه نستطيع أن نكتب :

$$S_x^2 = \frac{\sum (x - \bar{x})^2}{n}$$

أو :

$$S_x^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

وهي قيمة تباين المتغير المستقل x .

$$S_y^2 = \frac{\sum (y - \bar{y})^2}{n}$$

أو :

$$S_y^2 = \frac{\sum y^2}{n} - \left(\frac{\sum y}{n} \right)^2$$

وهي قيمة تباين المتغير التابع دون النظر إلى علاقته بالمتغير المستقل x . ولكن الآن بعد أن أوجدنا المستقيم الممثل للعلاقة بين المتغيرين، لتساءل ماذا حل بتباين المتغير التابع S_y^2 ؟

يمكن كتابة علاقة التباين الآتية :

$$S_y^2 = \frac{1}{n} \sum (y - \bar{y})^2$$

بالشكل الآتي :

$$S_y^2 = \frac{1}{n} \sum [(y - \hat{y}) + (\hat{y} - \bar{y})]^2$$

أي أضفنا وطرحنا y دون أن تتغير قيمة التباين، وبفك المطابقة الشهيرة نحصل

على :

$$S_y^2 = \frac{1}{n} \sum (y - \hat{y})^2 + \frac{1}{n} \sum (\hat{y} - \bar{y})^2 + 2 \sum (y - \hat{y})(\hat{y} - \bar{y})$$

إن قيمة الحد الثالث تساوي الصفر عندما تكون العلاقة مستقيمة بين المتغيرين، حيث يمكن إثبات ذلك ببساطة إذا بدلنا $\hat{y}$ بـ $(a + bx)$ وأجرينا بعض الإصلاحات. نستطيع أن نكتب :

$$S_y^2 = \frac{1}{n} \sum (y - \hat{y})^2 + \frac{1}{n} \sum (\hat{y} - \bar{y})^2$$

أي أن التباين S_y^2 الذي يطلق عليه التباين الإجمالي أو الكلي لقيم y ينقسم إلى

قسمين :

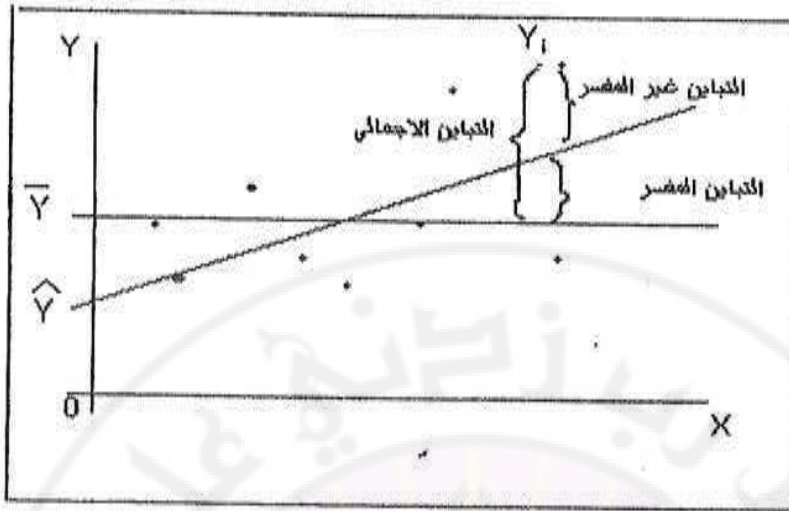
1- تباين القيم الفعلية عن النظرية المحسوبة، ونرمز له بالرمز $S_{y\hat{y}}^2$ ونطلق عليه التباين غير المفسر وسوف نرى سبب التسمية أدناه :

$$S_{y\hat{y}}^2 = \frac{1}{n} \sum (y - \hat{y})^2$$

2- تباين القيم النظرية المحسوبة عن الوسط الحسابي، ونرمز له بالرمز $S_{\hat{y}}^2$ ونطلق عليه التباين المفسر :

$$S_{\hat{y}}^2 = \frac{1}{n} \sum (\hat{y} - \bar{y})^2$$

ولتبسيط وصف هذه التباينات يمكننا بسهولة عرضها بيانياً في الشكل رقم (8-5) .



الشكل رقم (8-5)

حساب التباين أو التباين الكلي يتم من خلال العلاقة :

$$S_y^2 = \frac{1}{n} \sum (y - \bar{y})$$

أو :

$$S_y^2 = \frac{\sum y^2}{n} - \bar{y}^2$$

- التباين المفسر S_y^2 يحسب من العلاقة :

$$S_y^2 = \frac{1}{n} \sum (\hat{y} - \bar{y})^2$$

ويمكن تبديل $\hat{y} = a + bx$ في العلاقة السابقة وإجراء بعض الإصلاحات فنحصل

على العلاقة الآتية :

$$S_y^2 = \frac{b(\sum xy - \bar{x} \sum y)}{n}$$

- التباين غير المفسر $S_{y\hat{y}}^2$ يمكن حسابه أيضاً بطريقة الفروق :

$$S_{y\hat{y}}^2 = S_y^2 - S_{\hat{y}}^2$$

باعتبار أن التباين الإجمالي أو الكلي يساوي مجموع التباين المفسر وغير المفسر.
أو يحسب من العلاقة :

$$S_{y\hat{y}}^2 = \frac{1}{n} \sum (y - \hat{y})^2$$

التي يمكن أن تكتب بعد تبديل $\hat{y} = a + bx$ وإجراء الإصلاحات كما يأتي :

$$S_{y\hat{y}}^2 = \frac{\sum y^2 - a \sum y - b \sum xy}{n}$$

مثال (8-9) :

بالعودة إلى المثال الوارد في الجدولين رقم (8-14) و (8-15) نقوم فيما يأتي بحساب التباين بالطرائق الممكنة، حيث كانت معادلة مستقيم الانحدار لتلك البيانات من الشكل :

$$\hat{y}_i = 0.2819 + 0.7489 x_i$$

ولمتابعة الحل يلزمنا حساب القيم :

$$\sum (y - \hat{y})^2 , \sum (\hat{y} - \bar{y})^2 , \sum (y - \bar{y})^2$$

وهذا ما سوف نفعله بعد إيجاد القيم المقدرة (المحسوبة) من خلال تبديل قيم x بمعادلة مستقيم الانحدار. انظر الجدول رقم (8-17). كما أننا قمنا بعرض حساب التباين بالطرائق الممكنة المذكورة أعلاه في الجدول رقم (8-18).

جدول رقم (8-17)

الأسرة	الدخل (الف) x	الإنفاق (الف) y	xy	x*x	y*y	$\hat{y}_i$	$(y-\bar{y})^2$	$(\hat{y}-\bar{y})^2$	$(y-\hat{y})^2$
1	4	3	12	16	9	3.28	4	2.967	0.077
2	5	4	20	25	16	4.03	1	0.948	0.0007
3	3	2	6	9	4	2.53	9	6.108	0.2794
4	6	5	30	36	25	4.78	0	0.05	0.0505
5	7	7	49	49	49	5.52	4	0.275	2.178
6	8	7	56	64	49	6.27	4	1.621	0.5284
7	9	6	54	81	36	7.02	1	4.088	1.0445
8	4	4	16	16	16	3.28	1	2.967	0.522
9	5	3	15	25	9	4.03	4	2.948	1.0535
10	12	9	108	144	81	9.27	16	18.22	0.0722
المجموع	63	50	366	465	294	50	44	38.19	5.8062

جدول رقم (8-18)

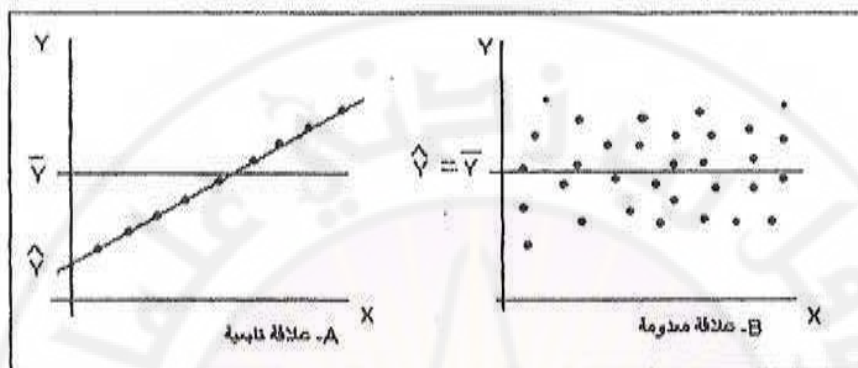
التباين	حسب التعريف	بالطريقة المباشرة
S_y^2	$\frac{\sum (y - \bar{y})^2}{n} = \frac{44}{10} = 4.4$	$\frac{\sum y^2}{n} - \bar{y}^2 = \frac{204}{10} - 5^2 = 4.4$
$S_{\hat{y}}^2$	$\frac{\sum (\hat{y} - \bar{y})^2}{n} = \frac{38.19}{10} = 3.819$	$\frac{b(\sum xy - \bar{x} \sum y)}{n} =$ $= \frac{0.75(366 - 6.3 * 50)}{10} = 3.819$
S_{yy}^2	$\frac{\sum (y - \hat{y})^2}{n} = \frac{5.806}{10} = 0.5806$ أو $S_y^2 - S_{\hat{y}}^2 = 4.4 - 3.819 = 0.58$	$\frac{\sum y^2 - a \sum y - b \sum xy}{n} =$ $= \frac{294 - 0.2819 * 50 - 0.75 * 366}{10} = 0.5806$

Coefficient of Determination

قبل دراسة العلاقة بين التباينات يمكننا أن نتساءل عن الفوائد التي تعطيها هذه المقارنات ولذلك لا بد من توضيح مفهوم كل تباين من التباينات، والأسباب التي أدت إلى حدوثه، فقد رأينا سابقاً أن التباين لمجموعة من القيم هو عبارة عن الوسط الحسابي لمربع فروق القيم عن وسطها الحسابي، فإذا كان هذا المتغير هو المتغير التابع فعندها نستطيع القول: إن بعض قيم المتغير التابع أكبر من الوسط الحسابي والباقي أقل منه، أي أن هناك تغيراً في قيم التابع، وسبب هذا التغير يعود إلى جميع العوامل التي يمكن أن يكون لها علاقة مع المتغير التابع، فإذا كان المتغير التابع هو الإنفاق لمجموعة من الأسر، ولاحظنا أن هناك بعض الأسر تنفق أقل وبعضها أكثر، فما هو سبب هذا الاختلاف في إنفاق هذه الأسر؟ في هذا المجال يمكن أن نذكر كل العوامل التي تؤثر بالإنفاق: الدخل - عدد أفراد الأسرة - مكان إقامة الأسرة - ثقافة أفراد الأسرة - العادات والتقاليد من جهة أخرى إذا لم يكن هناك أي تغير في قيم تابع ما، فهذا يعني أنه ليس هناك متغيرات تؤثر عليه وقيمه ثابتة وتباينه يساوي الصفر، من هنا نستطيع القول إن سبب التباين الكلي للمتغير التابع S_y^2 هو جميع المتغيرات المستقلة التي يمكن أن يكون لها علاقة بالمتغير التابع.

إن التباين الكلي - كما ذكرنا سابقاً - هو عبارة عن مجموع التباينين المفسر وغير المفسر، أما إذا كانت العلاقة تابعة بين المتغير المستقل والتابع أي أنه مقابل كل قيمة للمتغير المستقل يوجد قيمة محددة للمتغير التابع، بمعنى أنه لا يؤثر على التابع سوى متغير مستقل واحد، وفي هذه الحالة سوف تكون نقاط الانتشار واقعة على مستقيم الانحدار، ولن يكون هناك تباين غير مفسر، والتباين الإجمالي سوف يساوي التباين المفسر انظر الشكل رقم (8-6) الرسم A، وهنا نستطيع القول بأن التباين هو ناتج عن المتغير المستقل الداخلي في معادلة الانحدار وهو المتغير الوحيد المؤثر، ولا يوجد متغيرات مستقلة أخرى تؤثر بالتابع، وبالتالي لا يوجد تباين غير مفسر. من جهة أخرى إذا كان

لدينا متغيرين العلاقة بينهما معدومة الشكل رقم (8-6) الرسم B، فإن مستقيم الانحدار سوف ينطبق على الوسط الحسابي، مهما ازداد المتغير المستقل يبق المتغير التابع مساوياً للوسط الحسابي ولن يكون هناك تباين مفسر، والتباين غير المفسر سوف يساوي التباين الكلي .



الشكل رقم (8-6)

في مثل هذه الحالة الأخيرة لا توجد علاقة بين المتغير الداعل في المعادلة والمتغير التابع، أما إذا أدخلنا متغيراً مستقلاً آخر في المعادلة توجد بينه وبين المتغير التابع، فإن التباين غير المفسر سوف يتناقص ويزداد التباين المفسر وإذا أضفنا متغيراً مستقلاً آخر نلاحظ أن التباين المفسر سوف يزداد ويتناقص التباين غير المفسر، وهكذا إلى أن ندخل جميع المتغيرات المستقلة وعندها سوف تصبح العلاقة تابعة والتباين غير المفسر مساوياً للصفر، وكذلك الأمر لو أدخلنا متغيراً مستقلاً واحداً علاقته بالمتغير التابع علاقة تابعة نستطيع ببساطة فهم جوهر تسمية التباين المفسر. ومن خلال مقارنة التباين المفسر مع التباين الكلي يمكننا تحديد مقدار تأثير المتغير المستقل على المتغير التابع، فإذا كانت العلاقة تابعة فإن نسبة التباين المفسر إلى التباين الكلي تساوي الواحد، وندعو هذه النسبة بمعامل التحديد ونرمز له بالرمز r^2 :

$$r^2 = \frac{S_y^2}{S_y^2} = 1 - \frac{S_{y\hat{y}}^2}{S_y^2}$$

فإذا كانت العلاقة معدومة فمعامل التحديد r^2 يساوي الصفر، ولم يفسر شيء من هذه الحالة من التباين، والتباين الإجمالي يكون ناتجاً عن المتغيرات المستقلة غير الداخلة في معادلة الانحدار .

8-3-1-3-2 خصائص معامل التحديد :

1- تقع قيمة معامل التحديد بين الصفر والواحد، حيث تكون قيمته مساوية الصفر إذا كانت العلاقة معدومة وتساوي الواحد إذا كانت علاقة تابعة، وكلما اقتربت من الواحد كانت العلاقة أقوى .

2- مربع معامل ارتباط بيرسون يساوي معامل التحديد إذا كانت العلاقة مستقيمة فقط ويمكن إثبات ذلك رياضياً كما يأتي :

$$r = \sqrt{r^2}$$

8-3-1-4 التقدير والتنبؤ Estimation and forecast :

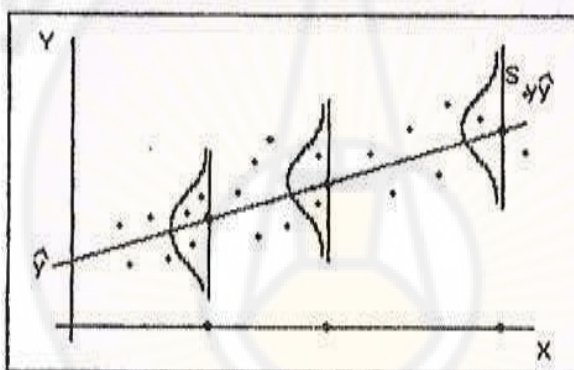
تعرفنا سابقاً على التوزيعات الاحتمالية وأهمها التوزيع الطبيعي، ورأينا أنه باستخدام هذا التوزيع يمكننا التنبؤ بقيمة متغير ما دون النظر إلى علاقته بمتغيرات أخرى، وذلك باستخدام الوسط الحسابي $\bar{x}$ والانحراف المعياري للقيم S_x والنسب تحت المنحنى الطبيعي، حيث نستطيع تقدير قيمة المتغير بالوسط الحسابي، إذا أخذنا تقدير نقطة أو تقدير مجال بإنشاء حدي ثقة باحتمال معين، كان يكون هذا الاحتمال 95.5% أو 95% أو 99%... والجمل يعطى بالعلاقة :

$$\bar{x} \pm Z S_x$$

حيث Z تمثل عدداً من الانحرافات المعيارية المقابلة لاحتمال الثقة .
أما بالنسبة لعملية التنبؤ بقيمة المتغير التابع بعد معرفة علاقته بالمتغير المستقل، فيمكن أن تتم عن طريق التقدير بنقطة، وهي هنا تساوي القيمة المتوسطة للتابع عندما يأخذ المتغير المستقل قيمة محددة، ويكون ذلك باستبدال x بالمعادلة بقيمتها المعطاة.

كما يمكن التنبؤ بقيمة المتغير التابع عن طريق إنشاء مجال ثقة، حيث نعتبر النقاط موزعة توزيعاً طبيعياً حول مستقيم الانحدار، فكما رأينا عند كل قيمة للمتغير المستقل x هناك قيمة متوسطة للمتغير التابع $\hat{y}$ ، وللتعرف على هذه القيمة يمكن تصور المنحنى الطبيعي (انظر الشكل (8-7)) وإنشاء مجال حولها بعدد من مقاييس التشتت حسب احتمال الثقة، ولكن التشتت هنا هو ليس عن الوسط الحسابي $\bar{y}$ (الانحراف المعياري) بل تشتت القيم الفعلية y عن القيمة المتوسطة $\hat{y}$ عند قيمة محددة لـ x ، وهذا هو التباين غير المفسر وندعوه بالخطأ المعياري للتقدير (Standard Error of Estimates).

$$S_{y\hat{y}} = \sqrt{S_{y\hat{y}}^2} = \sqrt{\frac{\sum (y - \hat{y})^2}{n}}$$



الشكل رقم (8-7)

وفي حالة العينات صغيرة الحجم، فإنه لحساب الخطأ المعياري للتقدير يجب أن يصحح بطرح 2 (درجتين ثقة)، كما هو الحال في تصحيح التباين بشكل عام.

مثال (8-10) :

بالعودة إلى المثال الوارد في الجدولين (8-14) و (8-15) يطلب حساب ما يأتي :

1- معامل التحديد وتفسيره.

2- التأكد من أن جذر معامل التحديد يساوي معامل الارتباط الذي أوجدته سابقاً .

3- تقدير إنفاق أسرة ما من هذه المدينة باحتمال 95.5% .

4- تقدير إنفاق أسرة يبلغ دخلها 10 ألف وحدة باحتمال 95.5% .

الحل : (1)

$$r^2 = \frac{S_y^2}{S_y^2} = \frac{3.819}{4.4} = 0.868$$

$$r^2 = 1 - \frac{S_{y\hat{y}}^2}{S_y^2} = 1 - \frac{0.58}{4.4} = 0.868$$

أي أن 86.8% من تباين المتغير التابع (الإنفاق) كان مرافقاً لتغيرات المتغير المستقل (الدخل)، أو يمكن القول إن 86.8 % من تباين الإنفاق يعود سببه لتباين الدخل (انظر الجدول رقم (8-18) لحساب التباين) .

$$r = \sqrt{r^2} = \sqrt{0.868} = 0.9317 \quad (2)$$

وهي النتيجة نفسها التي حصلنا عليها عند حساب معامل الارتباط سابقاً .

3- حدا الثقة لإنفاق أسرة ما باحتمال 95.5% .

$$\bar{y} \pm Z S_y$$

$$5 \pm 2 \sqrt{4.4}$$

$$[0.805 , 9.195]$$

أي باحتمال قدره 95.5% فإن إنفاق أسرة ما لن يقل عن 805 وحدة، ولن يزيد عن 9195 وحدة.

4- حدا الثقة لإنفاق أسرة يبلغ دخلها 10 ألف وحدة باحتمال 95.5% هو:

$$\hat{y} \pm Z S_{y\hat{y}}$$

$$\hat{y} = 0.2819 + 0.7489 * 10 = 7.77$$

$$S_{y\hat{y}} = \sqrt{S_{y\hat{y}}^2} = \sqrt{0.5806} = 0.76$$

$$7.77 \pm 2 * 0.76$$

$$[6.246, 9.294]$$

وهذا يعني أنه باحتمال قدره 95.5% فإن إنفاق أسرة يبلغ دخلها 10 آلاف وحدة سوف لن يقل عن 6.246 ألف وحدة ولن يزيد عن 9.294 ألف وحدة.

مسائل وتبايرين

1- بين ما هي طبيعة العلاقة ونوع المتغيرات في العلاقات الآتية :

- دخل ونفقات المستهلك .
- دخل ونفقات المنتج .
- دخل ومدخرات الأسرة .
- نفقات الأسرة وعدد أفرادها .
- حجم الإنتاج والنفقات الكلية .
- حجم الإنتاج وكلفة الوحدة الواحدة من الإنتاج .
- حجم الطلب على سلعة ما وسعرها .
- حجم عرض سلعة ما وسعرها .
- مدخرات الأسر والفوائد عليها .
- حجم الاستثمار وعوائد الاستثمار .

2- قُدِّر متانة العلاقة ونوعها من خلال رسم شكل الانتشار للحالات الآتية :

-A-

X	2	3	5	8	12	13	15	16	17	20
Y	5	6	8	11	15	16	18	19	20	23

-B-

X	1	2	4	5	7	8	10	11	13	15
Y	3	6	18	27	51	66	102	123	171	227

-C-

X	4	5	7	8	10	11	12	15	17	20
Y	29	36	44	51	58	66	68	87	93	108

3- ليكن لدينا البيانات الآتية التي تمثل السماد المستخدم والغلة لـ 100 قطعة متشابهة من الأرض :

الغلة كمية السماد	غير وفيرة	وفيرة	المجموع
عالية	15	55	70
قليلة	25	5	30
المجموع	40	60	100

والمطلوب: إيجاد متانة العلاقة بين كمية السماد والغلة مستخدماً المقياس المناسب .

4- أراد أحد الباحثين دراسة العلاقة بين لون بشرة الإنسان وعدد شعر الرأس فقام بجمع البيانات الآتية عن 60 شخص فتبين أن هناك 40 شخصاً لهم شعر كثيف من بينهم 30 شخصاً لون بشرتهم فاتح، والباقي لهم لون بشرة غامق، أما العشرون الباقون فلهم شعر خفيف بينهم 10 أشخاص لون بشرتهم فاتح، والباقيون غامق.
والمطلوب: مساعدة ~~في~~ تأكيد هذه العلاقة أو نفيها.

5- البيانات الآتية تمثل ~~200~~ طالب في امتحانات الفصل الأول والثاني :

المعدل	امتحانات الفصل الأول	امتحانات الفصل الثاني
50 - 60	60	55
60 - 80	90	80
فما فوق - 80	50	65
المجموع	200	200

المطلوب: هل توجد علاقة بين معدل هؤلاء الطلاب في كلا الامتحانين.

6- احسب متانة العلاقة بين درجة نضج العنب، والمسافة بين الأشجار في البيانات الآتية التي تم الحصول عليها من إحدى التجارب.

المسافة (بالمتر)	3	4	4.5	5	6	7
درجة النضج	أخضر صغير (محصرم)	أخضر كبير (محصرم)	أخضر كبير منقط بالأحمر	أخضر كبير منقط بالأحمر	أخضر كبير محمر	أحمر كبير

7- الجدول الآتي يبين النفقات الشهرية والدخل الشهري لـ 10 مؤسسات إنتاجية متشابهة :

المؤسسة	1	2	3	4	5	6	7	8	9	10
الدخل (بالآلاف)	8	10	10	12	15	16	17	19	20	25
الإنفاق (بالآلاف)	20	22	24	25	26	30	30	32	35	36

المطلوب :

1- إيجاد معادلة مستقيم الانحدار المثلثة للعلاقة وتفسير ثوابتها.

2- حساب معامل التحديد ومعامل الارتباط وتفسيرهما.

3- حساب الخطأ المعياري للتقدير .

4- قدر دخل مؤسسة تبلغ نفقاتها 23 ألف وحدة .

5- قدر باحتمال 95.5% دخل مؤسسة تبلغ نفقاتها 23 ألف وحدة .

8- تبين من سجلات مؤسسة المياه في إحدى الأحياء ما يأتي : إن عدد المرات التي تغير

به سعر المتر المكعب هو 14 مرة وأن حجم الاستهلاك اليومي عندما كانت

المؤسسة توزع المياه مجاناً للمشاركين هو 20 ألف متر مكعب وأن :

$$\sum x = 252 , \sum y = 154 , \sum xy = 1596$$

$$\sum x^2 = 6888 , \sum y^2 = 2290$$

والمطلوب:

- 1- بفرض أن العلاقة مستقيمة بين سعر المتر المكعب، وحجم الاستهلاك، أوجد معادلة مستقيم الانحدار .
- 2- حساب التباين المفسر وغير المفسر وما مدلولهما ؟
- 3- حساب معامل التحديد وتفسيره واحسب معامل الارتباط واختبر معنويته .
- 4- قدر باحتمال 95.5% حجم الاستهلاك لهذا الحي عندما يكون سعر المتر المكعب 12 وحدة.
- 5- ارسم مستقيم الانحدار الناتج وبيّن على الشكل التباين المفسر.
- 9- بهدف دراسة العلاقة بين عدد سنوات الخبرة والأجر الشهري للعامل في ورشات صيانة السيارات تم جمع البيانات الآتية عن 12 عاملاً:

العامل	1	2	3	4	5	6	7	8	9	10	11	12
عدد سنوات الخبرة	1	2	4	5	7	9	10	12	15	20	23	25
الأجر الشهري (ألف ليرة)	5	7	10	12	16	20	25	30	40	50	55	60

المطلوب:

- 1- أوجد متانة العلاقة بين عدد سنوات الخبرة والأجر الشهري.
- 2- أوجد معادلة مستقيم الانحدار وفسر ثوابتها .
- 3- أوجد التباين المفسر وغير المفسر والتباين الكلي .
- 4- احسب معامل التحديد ومعامل الارتباط، وفسرها وتأكد من معنوية معامل الارتباط .
- 5- قدر باحتمال 95.5% الأجر الشهري لعامل عدد سنوات الخبرة لديه 18 عاماً .

6- ارسم مستقيم الانحدار الناتج وارسم حوله مدى من خطاين معياريين .

10- البيانات الآتية تمثل دخل الأسرة ونفقاتها لعينة عشوائية حجمها 12 أسرة من إحدى المدن :

الدخل (ألف وحدة)	الإنفاق (ألف وحدة)	الأسرة
6	5	1
7	6	2
9	9	3
12	10	4
10	8	5
9	7	6
8	7	7
6	5	8
18	16	9
15	11	10
12	10	11
8	6	12
120	100	المجموع

والمطلوب :

- 1- احسب متانة العلاقة بين الدخل والإنفاق لهذه الأسر.
- 2- هل يمكن الاعتماد على قيمة معامل الارتباط الذي حصلت عليه في الطلب السابق في تقدير متانة العلاقة في المجتمع؟ برهن ذلك.
- 3- أوجد معادلة الانحدار الممثلة للعلاقة بين الإنفاق والدخل وفسر ثوابتها.
- 4- قدر باحتمال 95.5% إنفاق أسرة يبلغ دخلها 17 ألف وحدة .
- 11- بهدف دراسة تأثير نوع من السماد على إنتاج القمح تم استخدام كميات مختلفة من ذلك السماد في 7 قطع متشابهة الظروف :

المجموع	7	6	5	4	3	2	1	القطعة
	9	7	5	4	2	1	0	كمية السماد (كغ)
	325	325	260	265	235	210	200	الغلة (كغ)

والمطلوب :

1- حساب معامل الارتباط بين كمية السماد والغلة وتفسيره وهل قيمته

معنوية ؟

2- إيجاد معادلة مستقيم الانحدار بطريقة المربعات الصغرى وتفسير ثوابتها.

3- حاول إيجاد معادلة مستقيم الانحدار بطرائق أخرى معتمداً على خصائص

مستقيم الانحدار.

4- احسب معامل التحديد وفسره .

5- قدر باحتمال 95.5% غلة القطعة إذا تم رش 4 كغ من السماد في القطعة.

12- (مسألة دورة امتحانية 2001-2000 الفصل الأول) :

يهدف دراسة إنفاق الفرد الشهري في الأسر ذات الدخل المحدود قام أحد الباحثين

بجمع بيانات عن عدد الأفراد في 100 أسرة، وإنفاق الفرد الشهري في كل أسرة

فحصل على الجدول الآتي :

عدد الأسر	عدد أفراد الأسرة	نفقات الفرد الشهرية
12	7	750 - 850
18	6	850 - 950
40	5	950 - 1050
18	4	1050 - 1150
12	3	1150 - 1250

والمطلوب :

1- احسب متوسط إنفاق الفرد شهرياً في تلك العينة واحسب متوسط عدد أفراد الأسرة .

2- أيهما أكثر تمثيلاً لبياناته: متوسط عدد أفراد الأسرة أو متوسط إنفاق الفرد شهرياً ؟ دعم إجابتك بالحسابات اللازمة.

3- حدد طبيعة التوزيع الذي تخضع له بيانات إنفاق الفرد الشهري وطبيعة التوزيع الذي تخضع له بيانات عدد أفراد الأسرة .

كما رغب الباحث أيضاً بدراسة العلاقة بين الدخل الشهري لهذه الأسر وإنفاقها الشهري على الغذاء بصرف النظر عن عدد أفرادها، وجمع لذلك البيانات اللازمة وحصل على المجاميع الآتية :

$$\sum x = 6000 \quad \sum x^2 = 364000 \quad \sum xy = 297000$$

$$\sum y = 4900 \quad \sum y^2 = 242575$$

حيث x تشير إلى الدخل وتتراوح قيمته بين 40 و 70 وحدة و y الإنفاق، وتتراوح قيمته بين 36 و 56 وحدة، وكانت وحدة القياس لهذه البيانات هي مئة الليرة السورية، والمطلوب:

4- احسب معادلة مستقيم الانحدار وفسر ثوابتها.

5 - احسب الخطأ المعياري للتقدير و اشرح مدلوله.

6- ارسم مستقيم الانحدار وارسم حوله مدى من خطأ معياري واحد.

7- احسب معامل التحديد وفسره وبين الأهمية الإحصائية لمعامل الارتباط.

8- تبين أن أسرة يبلغ دخلها الشهري 6600 ليرة تنفق منها على الغذاء 5850 ليرة شهرياً، فهل تعتقد أن إنفاقها على الغذاء اعتيادياً ؟ برر إجابتك إحصائياً .

الفصل التاسع

الأرقام القياسية

Index Numbers

- مفهوم الرقم القياسي وأهميته .
- تركيب الأرقام القياسية .
- طرائق تركيب الأرقام القياسية .
- رموز ومصطلحات .
- طريقة المناسيب في تركيب الأرقام القياسية .
- المناسيب البسيطة .
- المناسيب المرجحة .
- خصائص المناسيب .
- طريقة السلسلة في تركيب الأرقام القياسية .
- الطريقة التجميعية في تركيب الأرقام القياسية .
- الطريقة التجميعية في تركيب الأرقام القياسية للأسعار .
- الطريقة التجميعية في تركيب الأرقام القياسية للكميات .
- اختبارات الأرقام القياسية .
- اختبار الانعكاس في الزمن .
- اختبار الانعكاس في المعامل .
- الاختبار الدائري .
- تعديل الأرقام القياسية .
- الأرقام القياسية الشائعة الاستعمال .



9-1 مفهوم الرقم القياسي وأهميته Concept of Index Numbers :

استعرضنا في البحوث السابقة مجموعة كبيرة من المقاييس الإحصائية الهامة (مقاييس النزعة المركزية - التشتت... الخ) التي يمكن استخدامها في قياس التطور والتغير في الظواهر والعمليات الاقتصادية والاجتماعية... الخ، فقد رأينا أن الوسط الحسابي لعدة قيم خاصة بظاهرة معينة هو رقم أو قيمة تعبر عن المستوى العام لهذه الظاهرة، كما رأينا أنه حتى يتسنى لنا أخذ الأهمية النسبية لمفردات هذه الظاهرة في الاعتبار، فإن حساب الوسط الحسابي من الممكن أن نحصل عليه بطريقة الترجيح بهذه الأهمية النسبية من خلال حساب الوسط الحسابي الموزون (المثقل Weighted Mean) والذي يمكن أن يستخدم لأغراض التنبؤ والاستدلال عن المستقبل، حيث إن معظم المؤسسات التجارية والحكومية تعبر عن موقفها المالي في صورة أرقام قياسية تستخدم كأداة في عمليات التحليل الإحصائي، مع الإشارة إلى أن الأرقام القياسية لا ترتبط فقط بالأنشطة التجارية والاقتصادية بل ترتبط أيضاً بالعديد من المجالات الأخرى كالتعليم والتخطيط وعلم النفس... الخ فمثلاً يستخدم علماء النفس حاصل الذكاء - وهو حاصل قسمة سن الشخص الفعلي على سنه الزمني مضروباً في مئة - وهذا في الحقيقة يعتبر نوعاً من الأرقام القياسية. كما يستخدم خبراء التعليم الأرقام القياسية لمعرفة

اتجاهات التسجيل بالمدارس والجامعات كأساس للتخطيط للمستقبل. وباختصار شديد تستخدم الأرقام القياسية تقريباً في جميع أوجه النشاط. وتستخدم الأرقام القياسية عموماً كمقياس للمتوسط التغير في قيم أو كميات أسعار السلع والخدمات.

وتستخدم الفترة التي يقاس بها التغير كفترة أساس (base period) للرقم القياسي، بينما تستخدم الفترة التي يقاس بها التغير كفترة مقارنة (given period) وتأخذ الظاهرة في فترة الأساس القيمة 100 وزيادة أو نقصان قيمة الظاهرة في فترة المقارنة عن قيمتها في فترة الأساس يؤدي إلى الحصول على نسبة مئوية أكبر أو أقل من 100، وبالتالي فإن دراسة التغير الذي يطرأ على الظواهر والعمليات والأشياء يجعلنا دوماً بحاجة إلى مقارنة ظاهرة معينة في زمان أو مكان ما بمثلها في زمان أو مكان آخر أو كليهما معاً، وذلك بهدف قياس درجة التغير الحاصل وطبيعته ونتيجة اختلاف المكان أو مرور الزمان. ولهذا المهدف نلجأ إلى دراسة الرقم القياسي كأداة إحصائية لقياس تغيرات مجموعة من البيانات بالمقارنة مع أساس معين قد تكون هذه البيانات أسعار منتج ما، أو مجموعة منتجات، أو قد تكون كميات، أو ألها تمثل حجم إنتاج زراعي أو صناعي، أو قد تكون صادرات أو واردات، أو تنطلق بالعمالة أو البطالة أو الصحة والتعليم،... إلى ما هنالك من بيانات يمكن قياس تغيراتها، من خلال مقارنتها مع أساس معين، قد يكون هذا الأساس فترة زمنية معينة أو قد تتم المقارنة بين بلدين، أو حتى بين قطاعين اقتصاديين ضمن البلد الواحد، أو ربما بين نشاطين اقتصاديين ضمن قطاع واحد. وهكذا يمكن لنا من خلال هذا المفهوم أن نعرف الرقم القياسي على أنه مؤشر أو قيمة تعبر عن المستوى العام للتغير في قيم ظاهرة معينة خلال فترات أو أماكن معينة ويستخدم لمقارنة التغير في ظاهرة واحدة، ولمقارنة التغير في المستوى العام لمجموعة من المتغيرات أو الظواهر المختلفة فيما بينها، والمرتبطة بشكل أو بآخر لتكون مجموعة متجانسة، فمثلاً مقارنة إنتاج مجموعة متجانسة مكونة من الثلاجات والغسالات والمكيفات في عام 2002 مع نظيره في عام 1992، بالرغم من وجود اختلافات كثيرة فيما بين هذه المفردات، أي أن كل سلعة تحكمها ظروف مختلفة، وبالتالي يتغير إنتاجها بنسب مختلفة بالزيادة للبعض،

وبالنقص للبعض الآخر، وبناء على ذلك يكون من المفيد إيجاد وسيلة أو مقياس يحتوي على كل هذه العوامل الثابتة والمتغيرة، وبالتالي فإن الرقم القياسي هو مقياس إحصائي يبين التغيرات التي تطرأ على ظاهرة ما، أو على مجموعة من الظواهر المترابطة خلال فترة زمنية أو منطقة جغرافية .

إذن من أجل استخدام الأرقام القياسية لا بد أن تأخذ دوماً قيم الظاهرة، أو كمياتها، أو أسعارها في فترة المقارنة أو الدراسة ونقارنها مع مثيلتها في فترة الأساس، أو مكان الأساس سواء أكان ثابتاً أم متحركاً، حيث نقول عن الأساس إنه ثابت عندما ننسب وبشكل متتالي كل كمية أو قيمة من قيم السلسلة إلى قيمة ثابتة، أي بمعنى أن قيمة المقام تبقى ثابتة دوماً بالنسبة لجميع الكميات أو القيم .

مثال (9-1) :

إذا كانت كميات إنتاج القمح (بآلاف الأطنان) في منطقة ما في الفترة (1999 - 2000) كما يأتي :

1999	2000	2001	2002
40	50	60	70

باعتبار عام 1999 عام الأساس يكون الأساس الثابت هو 40 وبالتالي يكون :

$$\frac{40}{40} , \frac{50}{40} , \frac{60}{40} , \frac{70}{40}$$

ونقول عن الأساس إنه متغير عندما ننسب كل كمية أو قيمة من قيم السلسلة إلى القيمة التي تسبقها مباشرة، أي بمعنى أن قيمة المقام في هذه الحالة سوف تتغير باستمرار بالنسبة لجميع الكميات أو القيم كما نلاحظ في المثال السابق :

$$\frac{40}{40} , \frac{50}{40} , \frac{60}{50} , \frac{70}{60}$$

إن السبب في تحويل مجموعة من القيم من أرقام مطلقة إلى أرقام نسبية يعود لعاملين اثنين :

1- إمكانية إجراء المقارنة بين أرقام المجموعة نفسها من البيانات.

2- إمكانية مقارنة عدة مجموعات من القيم معبراً عنها بوحدة مختلفة.

وعندما يقتصر الأمر على قياس التغير النسبي لظاهرة واحدة من زمان إلى آخر، فمن الواضح أن هذا النوع من الأرقام القياسية النسبية تعرض من وجهة النظر الإحصائية مشكلات قليلة، ما عدا تلك المتعلقة بجميع البيانات العددية. وبالإضافة إلى ذلك، فإن تفسير هذه الأرقام البسيطة سهل غير معتمد، نظراً لأن العلاقة فيما بينها هي العلاقة نفسها لو بقيت بصورة مطلقة، وبالتالي فإن تحويلها إلى أرقام نسبية لا يضيف شيئاً جديداً إليها أكثر من وصف التغيرات المطلقة فيما بينها بصورة مثوية من أساس ثابت أكان زمانياً أم مكانياً .

وفي الحقيقة، فإن استعمال مثل هذه الأرقام أمر يتعلق بالملاءمة باعتبار أن فهم قيم قليلة يكون أكثر سهولة من قيم كبيرة، نظراً لأن الأرقام النسبية بشكل عام هي أصغر عددياً بالمقارنة فيما لو كانت مطلقة، وبالتالي غالباً ما يطلق عليه أرقام نسبية (Relative Numbers) في حين أن تعبير أرقام قياسية (Index Numbers) يطلق بصورة أكثر ملاءمة عندما يتعلق الأمر بقياس التغير النسبي سواء أكان من حيث الزمان أم المكان، لتلك الظواهر المركبة من عدة عوامل متغيرة في وقت واحد، ومثال على ذلك الأرقام القياسية لأسعار التجزئة أو الجملة. مثل هذه الأرقام تستخدم لقياس التغيرات ليس في سعر سلعة واحدة فحسب بل لقياس التغير الصافي في أسعار مجموعة كاملة من السلع، بعضها قد يكون متزايداً والآخر متناقصاً، وبالتالي فإن الرقم القياسي يعبر عن التغير النسبي سواء أكان من حيث الزمان أم المكان للظواهر المركبة من عدة عوامل مختلفة، أما الرقم النسبي فهو يصف التغيرات المطلقة للظواهر فيما

بينها بصورة مثوية سواء أكان هذا الأساس زمانياً أم مكانياً، أي أن هناك تطابقاً بين الرقم القياسي والرقم النسبي، وبالتالي فإن الرقم القياسي هو أداة لقياس التغير النسبي في كمية وأسعار وقيم ظاهرة ما أو مجموعة من الظواهر من زمان إلى آخر أو من مكان إلى آخر، وبما أن الرقم القياسي يعبر عن التغير النسبي للظاهرة زمانياً أو مكانياً وبين الأهمية النسبية لها، فهو مرتبط إذن بعلاقة أو شبه علاقة مع الوسط الحسابي المرجح (المثقل) الذي يقيس الأهمية النسبية للظاهرة المدروسة، ولما كان الرقم القياسي يصف تغير الظاهرة المدروسة بالزيادة أو النقصان فهو يرتبط أيضاً بمعدل النمو.

فإذا كان لدينا y_1, y_0 عبارة عن القيم التي تأخذها ظاهرة ما خلال الفترتين t_1, t_0 فإن معدل النمو يعطى بالعلاقة الآتية :

$$a(1/0) = \frac{\Delta y}{y_0} = \frac{y_1 - y_0}{y_0} = \frac{y_1}{y_0} - 1 = \text{معدل النمو}$$

أما الرقم القياسي فيمثل النسبة $\frac{y_1}{y_0}$ خلال الفترتين t_1, t_0 :

$$I(1/0) = \frac{y_1}{y_0}$$

وعادة نضرب العلاقة السابقة بـ 100 فيكون لدينا :

$$I(1/0) = \frac{y_1}{y_0} \cdot 100$$

ومنه يكون $a(1/0) = I(1/0) - 1$ أو $I(1/0) = [a(1/0) + 1] \cdot 100$.

أي أن معدل النمو للظاهرة (y) = الرقم القياسي - 1 .

والرقم القياسي هو 1 + معدل النمو .

من هذا المنطلق يمكن أن يستخدم الرقم القياسي لأغراض عديدة في المجالات الاقتصادية والاجتماعية كقياس ظاهرة الثراء النسبي أو الفقر النسبي لمنطقة معينة مقارنة

$$\frac{85}{125\%} = 68 \text{ وحدة نقدية}$$

وبالتالي يمكننا القول إن هذه الزيادة الحاصلة في دخل هذا الشخص من عام 2001 إلى عام 2002، وقدرها 25 وحدة نقدية يستطيع أن يشتري سلع وخدمات للاستهلاك الضروري بمقدار 8 وحدات نقدية بالمقارنة بإمكانياته في عام 2001 .

نسبة الزيادة الاسمية في دخل هذا الشخص هي :

$$\frac{25}{60} \cdot 100 = 41.67 \%$$

نسبة الزيادة الحقيقية في دخل هذا الشخص هي :

$$\frac{8}{60} \cdot 100 = 13.33 \%$$

يمكن أن نفسر ذلك اقتصادياً بأنه عندما ازداد الدخل بنسبة 41.67% فإن الأسعار قد زادت بنسبة 125%، أي أن هذه الزيادة في الدخل هي زيادة اسمية، في حين أن الزيادة الحقيقية هي بنسبة 13.33% أي بمعنى أن الأسعار امتصت هذه الزيادة في الدخل بما يعادل 11.67% (25% - 13.33%)، وهذا يعبر عنه على صعيد الاقتصاد الوطني مفهوم التضخم النقدي .

مثال (3-9) :

بفرض أن دخل أسرة ما في عامي 1992-2002 قد ارتفع من 10000 وحدة نقدية عام 1992 إلى 15000 وحدة نقدية عام 2002، وأن الرقم القياسي لتكاليف الحياة المعيشية قد ارتفع بنسبة 30% في عام 2002 عما كان عليه عام (1992 = 100%) فما هو مقدار الزيادة الحقيقية في دخل هذه الأسرة؟ وما مقدار الزيادة الاسمية ؟

الحل :

نلاحظ أن التغير في الدخل خلال العامين المدروسين 1992-2002 هو :

$$\frac{15000 - 10000}{10000} \times 100 = 50 \%$$

أي أن دخل هذه الأسرة قد زاد بنسبة 50% عام 2002 مقارنة، بدخلها عام 1992 .

غير أن التغير في الرقم القياسي للأسعار وتكاليف المعيشة قد زاد بنسبة 30% عام 2002 عما كان عليه عام 1992 .

مقدار الزيادة الحقيقية في دخل هذه الأسرة عام 2002 :

$$\frac{15000}{130} \times 100 = 11538.46 \text{ (وحدة نقدية)}$$

وبالتالي يكون لدينا الزيادة الحقيقية في دخل هذه الأسرة في عام 2002، مقارنة مع عام 1992 هي :

$$11538.46 - 10000 = 1538.46 \text{ (وحدة نقدية)}$$

أما الزيادة الاسمية في دخل هذه الأسرة هي :

$$15000 - 10000 = 5000 \text{ (وحدة نقدية)}$$

نلاحظ أن ارتفاع الرقم القياسي لتكاليف المعيشة أدى إلى انخفاض في القوة الشرائية، وهذا يعتبر مؤشراً من مؤشرات التضخم، بالإضافة إلى استخدامات الأرقام القياسية في مجال تحديد نوعية الزيادة في الدخل سواء أكانت حقيقية أم اسمية، كذلك يمكننا استخدامها بهدف معرفة القيمة الحقيقية للمشاريع الاستثمارية سواء أكانت زراعية أم سياحية أم صناعية أم خدمية... الخ، وذلك عندما نأخذ بعين الاعتبار الرقم القياسي الخاص بتكاليف المواد اللازمة لإنشاء المشروع. والمثال الآتي يوضح ذلك .

مثال (4-9) :

بفرض أنه تم تخصيص مبلغ 30 وحدة نقدية لإنشاء منشأة سياحية في عام 1995، حيث كان الرقم القياسي لتكاليف المواد اللازمة لهذا النوع من المشروعات هو

% 130 (100% = 1990) وبفرض أنه أريد إنشاء منشأة سياحية مماثلة في عام 2000،
 وتم تخصيص مبلغ 35 وحدة نقدية لهذا الغرض غير أن الرقم القياسي لتكاليف المواد
 اللازمة أصبح 180% (100% = 1990). الآن، هل تكفي هذه الأموال المخصصة
 لإنشاء المنشأة السياحية الجديدة ؟

الحل :

إن القيمة الحقيقية للأموال المخصصة لإنشاء المنشأة السياحية في عام 1995
 (100% = 1990)

$$= \frac{30}{130\%} = 23.07 \text{ مليون وحدة نقدية}$$

أما القيمة الحقيقية للأموال المخصصة لإنشاء المنشأة السياحية الجديدة في عام
 2000 (100% = 1990)

$$= \frac{35}{180\%} = 19.44 \text{ مليون وحدة نقدية}$$

إذن إن القيمة المخصصة لإنشاء هذه المنشأة غير كافية، بسبب ارتفاع الرقم
 القياسي لتكاليف المواد من 130% إلى 180% .

القيمة الحقيقية للأموال المخصصة عام 2000
 أما نسبة الإنجاز فتكون = $\frac{\text{القيمة الحقيقية للأموال المخصصة عام 2000}}{\text{القيمة الحقيقية للأموال المخصصة عام 1995}} \times 100$

$$\frac{19.44}{180\%} \times 100 = 84.26 \%$$

ويمكن حساب نسبة الإنجاز وفق الصيغة الآتية :

$$100 \times \frac{\text{القيمة المخصصة في سنة المقارنة} \times \text{الرقم القياسي لتكاليف في سنة الأساس}}{\text{القيمة المخصصة في سنة الأساس} \times \text{الرقم القياسي لتكاليف في سنة المقارنة}} = \text{نسبة الإنجاز الفعلية}$$

$$\frac{35 \times 130}{30 \times 180} = \frac{4550}{5400} \times 100 = 84.26 \%$$

نلاحظ عند مقارنة المبالغ المخصصة لإنشاء منشأة سياحية جديدة، أي مقارنة 35 مليون وحدة نقدية مع 30 مليون وحدة نقدية أن المبالغ قد زادت من 30 مليون وحدة نقدية إلى 35 مليون وحدة نقدية، أي أنه يمكن إنشاء منشأة سياحية جديدة، غير أنه عند استخدام الأرقام القياسية الخاصة بالتكاليف الصناعية تبين أنه في عام 2000 يمكن إنجاز حوالي 84.26% من هذا المشروع باستخدام الأموال المخصصة لإنشاء هذه المنشأة .

والخلاصة أن الرقم القياسي هو أداة قياسية تستخدم في قياس التغير في الأسعار وآثاره على النشاط الاقتصادي، وقياس التغيرات التي تصيب تكاليف الحياة المعيشية في مجتمع ما، أو طبقة معينة، وذلك بقسمة الرقم القياسي للدخول النقدية للمجتمع أو الطبقة على الرقم القياسي لنفقة المعيشة أو تكاليف المعيشة فنحصل على رقم قياسي يقيس التغير في الأجر الحقيقية بالنسبة لفترة زمنية معينة تعتبر أساساً للمقارنة. كما تستخدم الأرقام القياسية لقياس التغيرات التي تصيب حجم الإنتاج الصناعي أو الزراعي أو الموجودات في المخازن من السلع التامة الصنع أو النصف المصنعة، أو المواد الخام والوقود، أو التجارة الخارجية بشقيها الصادرات والواردات... الخ. وتستخدم الأرقام القياسية أيضاً في قياس التغيرات في عدد المشتغلين عامة وفي كل قطاع من قطاعات النشاط الاقتصادي على حده، وكذلك التغيرات في عدد العاطلين عن العمل وفي أسعار الأسهم والسندات التي تتداول في البورصات. كما يمكن بواسطة الرقم القياسي تحديد متوسط التغيرات في أسعار عدة أصناف لسلعة واحدة أو عدة سلع تقاس مبيعاتها بوحدات مختلفة. لهذا تعتمد الأبحاث في الاقتصاد القياسي على الأرقام القياسية عند تركيب النماذج الاقتصادية ، كما وتستخدم الأرقام القياسية للأسعار في استبعاد أثر تغيرات الأسعار من القيم التي تتأثر بهذه التغيرات للتعرف على التغيرات الحقيقية ، كما لاحظنا سابقاً.

9-2 تركيب الأرقام القياسية Composition of Index Numbers :

عند تركيب الأرقام القياسية يجب مراعاة مجموعة من العوامل التي تمكننا من الوصول إلى رقم قياسي جيد يُخدم الأهداف المنشودة منه، ويتحقق الغاية في قياس التغيرات ووصف الآثار المترتبة عليها، ومن بين هذه العوامل الهامة ندرس :

1 - تحديد الغرض من تركيب الرقم القياسي :

قبل أن نبدأ في تركيب الرقم القياسي لا بد لنا من تحديد الهدف أو الأهداف المرجوة من وضع هذا الرقم، حيث يبنى على ذلك اختيار عناصر الرقم، والطريقة التي سيتم العمل بها، فمثلاً الرقم القياسي لنفقة المعيشة يختلف عن الرقم القياسي للإنتاج الصناعي، وكذلك يختلف عن الرقم القياسي للصادرات والواردات، وإن تركيب رقم قياسي لنفقة المعيشة للطبقة العاملة يختلف من حيث مكوناته عن رقم قياسي لنفقة المعيشة لجميع طبقات المجتمع سوياً، كذلك يختلف تركيب رقم قياسي لأسعار الجملة عن رقم قياسي لأسعار الأسهم الصناعية التي يجري تبادلها في سوق الأوراق المالية، لهذا لا بد من تحديد الغرض من تركيب الرقم القياسي بكل وضوح ودقة وأن عملية استخدامه بعيداً عن هذا الغرض يجعل استخدامه لأي أغراض أخرى غير ذي دلالة دقيقة؛ فمثلاً تركيب رقم قياسي لقياس التغير في تكاليف البناء يجب أن لا يستخدم في تقييم المعدات.

2- اختيار العناصر التي يتركب منها الرقم القياسي :

يعتمد اختيار العناصر التي يتركب منها الرقم القياسي إلى مدى بعيد، على الغرض المنشود الذي يحدو الباحث للقيام بتركيبه، وعادة، فإن هدف الباحث يختلف بين أن يكون عاماً، كما هو الحال عند إنشاء أرقام قياسية لأسعار الجملة، حيث يكون الهدف من تركيبه إعطاء صورة واسعة عن تغيرات الأسعار الصافية لبلد من البلدان، دون أن يقتصر الأمر على منطقة معينة، أو أمور محددة تتعلق بطبقة ما من المستهلكين،

أو المنتجين، أو قد يكون الهدف خاصاً كما هو الحال عند تركيب أرقام قياسية لقياس نفقات المعيشة لنوع معين من الأفراد من طبقة محددة. وأياً كانت الحال، فإنه من غير العملي أن يتركب الرقم القياسي من جميع الأسعار التي تشكل المجتمع الإحصائي لهيكل الأسعار، نظراً لأن الضرورة العملية تقضي بالاعتصار على أخذ عينة من هذه الأسعار، تضم تلك الأسعار التي تعدّ ممثلة تمثيلاً صادقاً للمجتمع الإحصائي المأخوذة منه.

وحقّ تكون العينة التي يجري اختيارها ممثلة تمثيلاً صادقاً لموضوع الرقم القياسي، وتحقيقاً لأهدافه يجب أن تشمل عدداً لا بأس به من العناصر (السلع مثلاً)، وليس هناك قاعدة محددة يمكن أن نسترشد بها عند تحديد عدد العناصر، وإن كان يجب أن تلاحظ أن لا يزيد العدد كثيراً، لأن في ذلك تعقيداً للعمليات الحسابية التي يتطلبها تركيب الرقم، وأن لا يقل العدد كثيراً حيث تصبح العينة بذلك غير ممثلة تمثيلاً صادقاً لموضوع المقارنة والقياس. وعند تركيب الأرقام القياسية للأسعار يحسن تقسيم السلع -موضوع الدراسة- إلى مجموعات تشمل كل منها مفردات تتميز بصفات خاصة ذات أهمية فيما يتعلق بأهداف الرقم القياسي الذي نكون بصدد تركيبه، وهذا يتطلب من الباحث مهارة ودقة كبيرة، نظراً لأن العينة المأخوذة لهذه الغاية لا تؤخذ بصورة عشوائية، وبالتالي فليس هناك مقاييس إحصائية محددة لتقويم مقدار درجة صحة الأرقام القياسية ودقتها. فمثلاً في حالة تركيب أرقام قياسية لنفقات المعيشة، يتعين على الباحث أن يدرك أن نفقات المعيشة تختلف باختلاف مستويات الدخل، نظراً لأن طبيعة نفقات الأفراد وأنواع السلع التي يتناعوها تعتمد بصورة رئيسية على مستوى دخولهم. وتبعاً لذلك، فقد جرت العادة عند تركيب أرقام قياسية لنفقات المعيشة أن تقاس التغيرات في مستوى الأسعار التي يتأثر بها، أو لتلك الذين هم في مستويات الدخل الدنيا. إن قيمة الرقم القياسي تتوقف على حكم الباحث ومهارته التي تبذل في اختيار العناصر التي ستدخل في تركيب الرقم القياسي، وذلك تبعاً للهدف من وراء إنشائه، حيث يجب على الباحث وضع دراسة مفصلة حول أنواع السلع التي تشتري فعلياً من قبل أولئك الذين هم في مستويات الدخل الدنيا، وذلك عند تركيب الرقم القياسي لنفقات

المعيشة كأن يقوم بتقسيم السلع والخدمات موضوع الدراسة إلى مجموعة اللحوم والخضراوات والفواكه والبقول والزيوت والملابس والمواصلات والتعليم والعناية الصحية والطبية.. الخ، بحيث لا تشمل كل مجموعة جميع السلع والعناصر التي تتكون منها، إذ يكفي أن يأخذ بعضها بحيث يكون هذا البعض شاملاً للسلع الشائعة الاستعمال لدى الطبقة التي يهدف الرقم القياسي إلى قياس التغير في نفقة معيشتها، وبحيث تكون ممثلة تمثيلاً صادقاً للاتجاهات المختلفة في أسعار المجموعة الخاصة بها.

إن ما يقال عن الأرقام القياسية لنفقات المعيشة ينطبق على جميع الأرقام القياسية الأخرى، وهكذا، عندما يرغب الباحث في تركيب أرقام قياسية للإنتاج الصناعي، فإنه يلجأ عادة إلى تقسيمه إلى عدة أقسام مثل الأطعمة، والمنسوجات، والمنتجات الجلدية، والأخشاب والمصنوعات الخشبية، والحديد والفولاذ. ونظراً لأنه من الصعوبة أن يضمن الباحث العينة كل الأجزاء الصغيرة المكونة لكل قسم من هذه الأقسام، فإن عليه أن يختار بصورة عقلانية من كل قسم العناصر التي تعتبر ممثلة له. وهكذا الأمر بالنسبة للأنواع الأخرى من الأرقام القياسية. ويتعين على الباحث عند اختيار العناصر التي ستتركب منها العينة مراعاة الاعتبارات الآتية :

٢ - التكلفة (Cost) :

تعدّ التكلفة التي يتطلبها تركيب الأرقام القياسية من العوامل المهمة في اختيار المواد التي تدخل في تركيب الأرقام القياسية، كما أنها تعدّ من العوامل التي تضع حداً أعلى لعدد المواد التي تتضمنها العينة. وفي الحقيقة، هناك عناصر يمكن الحصول على معلومات عنها بأقل جهد ونفقة، نظراً لأنها لا تتطلب سوى نقلها من السجلات الرسمية المنشورة فيها، كما هو الحال بالنسبة للسلع ذات الأسعار العالمية، أو تلك السلع التي تشرف عليها أجهزة الدولة، كالقمح والأقطان وغيرها من السلع الأساسية، وعلى العكس من ذلك، لا يمكن تأمين أسعار سلع التجزئة إلا عن طريق الاتصال المباشر بالمنشآت التي تتعامل بها، حيث يشتري المستهلك الأخير عادة معظم مشترياته منها.

ب - التماثل (Symmetry) :

من الاعتبارات المهمة في اختيار العناصر التي تدخل في تركيب الأرقام القياسية هو بقاء العناصر متماثلة بين فترة زمنية وأخرى، حيث إن هناك عدداً كبيراً من السلع المصنوعة لا تحافظ على تماثلها مع مرور الزمن، فمن الصعوبة بمكان القيام بمقارنة أجهزة التلفاز لعام 2002 بتلك الأجهزة المنتجة عام 1940، نظراً لأن هناك فروقاً جوهرية بينهما، وبالتالي فإنه لا يمكن اعتبارها سلعة واحدة.

ج - الازدواجية (Double Digit) :

تعني الازدواجية أو العد المزدوج أن تدخل سلعة من السلع مرتين أو أكثر في العينة التي يتركب منها الرقم القياسي، كأن تدخل كمادة أولية، ثم كمادة مصنوعة جزئياً أو كمادة تامة الصنع، وهذه المشكلة قد تواجه الباحث عند قيامه بتركيب أرقام قياسية غير مرجحة، فمثلاً عند تركيب أرقام قياسية لأسعار الأطعمة، فقد يدخل الباحث في العينة التي تتركب منها أرقاماً قياسية لأسعار الأطعمة، وقد يدخل الباحث في العينة التي تتركب منها أرقامه القياسية أسعار الدقيق والخبز والمعجنات، إذ الدقيق هو المادة الرئيسية في جميع هذه السلع، فإذا لم يعامل المواد الأخرى الداخلة في هذه الأرقام القياسية المعاملة نفسها، فقد يحصل على أرقام قياسية يكون فيها لتقلبات أسعار الدقيق تأثير أكبر مما يجب، نظراً لأن قيمة كل سلعة من هذه السلع المتداخلة تتضمن سعر الدقيق، وتبعاً لذلك إذا ما أريد استعمال القيمة كأوزان للسلع المصنوعة، فإنها يفضل أن تكون القيمة المضافة⁽¹⁾، وذلك بهدف التخلص من الإعادة المتكررة لأسعار هذه المادة الأولية الأساسية.

(1) القيمة المضافة الصافية = قيمة الإنتاج - قيمة المستهلك من الوقود والمواد الأولية واهتلاك الأصول الثابتة.

3- اختيار مصادر جمع المعلومات التي يتطلبها تركيب الرقم القياسي:

إن عملية جمع المعلومات من مصادر متعددة تواجه بعض الصعوبات التي يجب مواجهتها وفق أسس موحدة لتقليل التفاوت الكبير في مقدار الثقة بهذه المعلومات التي تمثل العناصر التي يتركب منها، فمثلاً، إن أسعار السلع الاستهلاكية في أسواق الجملة هي أعلى من أسعار هذه السلع نفسها في تجارة الاستيراد والتصدير، نظراً لأن المصدر والمستورد لديه ميل ثابت للتصريح عن قيم السلع التي يتعامل بها بأقل من حقيقتها، وذلك لكي لا تترتب عليه رسوم باهظة، وبالإضافة إلى ذلك، فإن هناك أنواعاً من المعلومات يصعب تأمينها بصورة صحيحة، فأرباح مؤسسات الأعمال ونفقاتها وخسائرها ومبيعاتها... الخ كثيراً ما تكون المعلومات عنها مغلوبة ومشوهة بصورة معتمدة، ولذلك ليس من قبيل المصادفة إذا ما كانت الأرقام القياسية للمبيعات المحسوبة من عينات صغيرة مشكوك في صحتها وقيمتها التمثيلية.

وتجدر الإشارة إلى أن هناك مصدراً آخر لعدم صحة المعلومات اللازمة عن العناصر التي يتركب منها الرقم القياسي، هو اختلاف أسعار السلع باختلاف شروط البيع بالدفع نقداً، في حين أنها تباع بأسعار أكثر ارتفاعاً لأولئك الذين يحصلون على شروط ائتمانية دون إضافة أية أعباء عليهم لقاء حصولهم على هذه التسهيلات.

ولا شك أن الرقم القياسي يعاد تركيبه من وقت إلى آخر، كما أن حسابه يجري على فترات زمنية متتابعة، كل شهر مثلاً، لذلك يجب أن نلاحظ تغير استعمال السلع من وقت إلى آخر، وفقاً لتغير عادات المجتمع عامة وعادات الطبقة التي يقيس الرقم التغيرات في نفقة معيشتها، فهناك سلع يبطل استعمالها، كما أن هناك سلعاً ينتشر استعمالها، كما أن أنماط الإنفاق تتغير، ولهذا يجب تعديل تركيب الرقم القياسي كلما أدر كنا حدوث تغيرات جوهرية في عادات الاستهلاك وأنماط الإنفاق الخاصة بالطبقة، و المجتمع موضوع الدراسة، لذلك فإن مكونات الرقم القياسي لنفقة المعيشة تعتمد على قائمة السلع والخدمات التي تصدر عن مديرية الإحصاء والتي قد لا تشمل جميع بنود

الإنفاق شمولاً كاملاً، أو قد يكون هناك خطأ ما وقعت به مديرية الإحصاء يعود إلى أنها تجمع الأسعار من مصادر غير دقيقة، إذ قد تستخدم الأسعار الرسمية التي تحددها الحكومة للسلع والخدمات في الوقت الذي تكون فيه هذه الأسعار غير مطابقة، أي يتقيد بها التجار، وبذلك تكون بعيدة عن الواقع، الأمر الذي يجعل الرقم القياسي المحسوب على أساس لا يعبر تعبيراً دقيقاً عن التغيرات الفعلية في أسعار السلع والخدمات.

4- اختيار فترة الأساس (Base Period) :

تعدّ مشكلة تحديد الفترة الزمنية التي تتخذ أساساً للمقارنة عند تركيب الأرقام القياسية من المشاكل الهامة التي تواجه الباحث، حيث إن فترة زمنية ما قد تكون أساساً مناسباً في مرحلة تاريخية معينة، إلا أنها تصبح غير مناسبة بمرور الزمن، ولهذا يجب تغييرها حتى تصبح قريبة من فترات المقارنة (Given-period)، ويصبح تركيب الرقم القياسي على أساسها مفيداً من الناحية العملية، وتكون المقارنة بالنسبة لها غير مضللة. وفترة الأساس لأي رقم قياسي هي الفترة التي تعتبر القيم فيها أساساً أي تساوي 100، وبذلك يكون الفرق بين نتيجة الرقم في أي فترة أخرى وبين المئة هو مقياس التغير. ويكون التغير بالزيادة عندما تزيد النتيجة عن 100، ويكون بالنقص عندما تنقص النتيجة عن 100. لذلك عند اختيار فترة الأساس يجب مراعاة الشرطين الآتيين:

أ- أن تكون فترة الأساس حديثة وغير بعيدة كثيراً عن فترة المقارنة، وذلك لاحتمال تغير عادات المجتمع وأنماط إنفاقه، أو لتغير أهمية عناصر الرقم القياسي، أو ظهور عناصر جديدة ضمن الرقم، أو حتى اختفاء عناصر قديمة، فكل ذلك يمكن أن يؤثر على دقة الرقم .

ب- أن تتميز فترة الأساس بالاستقرار بحيث تكون الأسعار فيها عادية. بمعنى أن لا تكون شاذة اقتصادياً، فلا يجب أن تكون فترة رواج أو كساد، وكذلك غير شاذة طبيعياً فلا تكون فترة حرب أو كوارث طبيعية أو مجاعات أو ما شابه ذلك.

قد تكون فترة الأساس فترة زمنية معينة (سنة، أو شهراً أو أسبوعياً) لأنها تعدّ فترة سوية واعتيادية وتتصف بمميزات مهمة معينة، وقد تكون فترة الأساس متوسط عدة سنوات بدلاً من سنة واحدة، وذلك بسبب صعوبة إيجاد فترة زمنية مثالية خالية من التقلبات الاقتصادية، وقد تكون أيضاً فترة الأساس الفترة السابقة مباشرة لفترة المقارنة حيث يتم استخدام الأساس المتحرك (Changing Base) في تركيب الأرقام القياسية الذي يتغير بشكل دائم ومستمر.

فإذا كانت الأرقام القياسية شهرية، فإن الرقم القياسي للشهر الجاري يعبر عن نسبة مئوية من الشهر السابق له مباشرة، وكذلك الأمر إذا كانت الأرقام القياسية سنوية، فإنه يعبر عن الرقم القياسي للسنة الحالية كنسبة مئوية من السنة السابقة لها مباشرة، ولا شك أن هذه الطريقة تتميز بالمرونة، لأنها تأخذ بالاعتبار نتائج الأحوال الاقتصادية المتغيرة، وتأثير هذه الأحوال على السلع التي يتضمنها الرقم القياسي فيما يتعلق بأسعارها أو بأهميتها النسبية أي بأوزانها المختلفة وتأثيرها على المتوسط الذي يستخدم في حساب الرقم القياسي، فمثلاً يعبر الرقم القياسي ذو الأساس المتحرك عن إنتاج أو مبيعات الشهر الحالي لإحدى المؤسسات كنسبة مئوية من إنتاج أو مبيعات الشهر المقابل في السنة السابقة مباشرة، ومثل هذه الطريقة تعدّ مهمة خاصة بالنسبة للبيانات التي تتأثر تأثيراً كبيراً بالتغيرات الموسمية (Seasonal variations).

ونحب الإشارة إلى أن اختيار سنة واحدة كفترة أساس، تمكن الباحث من القيام بمقارنة أكثر واقعية، غير أنه كلما اتسعت الفترة الزمنية بين سنة المقارنة وسنة الأساس أصبحت المقارنة أقل تحديداً، وأكثر بعداً عن الحقيقة، وبالتالي يصبح من اللازم إعادة النظر في تركيب الأرقام القياسية واختيار فترة أساس تناسب وطبيعة الحياة الاقتصادية السائدة في فترة المقارنة، وذلك تبعاً لتغير أهمية العناصر المكونة للأرقام القياسية بين فترتي الأساس والمقارنة، وهذا يفسر لنا، ولو بصورة جزئية، إعادة النظر المتكررة للأرقام القياسية المنشورة في مختلف بلدان العالم من خلال تغيير السنة المعتمدة كأساس، أو عن طريق أخذ متوسط عدة سنوات كأساس، أو استخدام الأساس المتحرك في تركيب الأرقام القياسية.

إن اختيار متوسط عدة سنوات كأساس في تركيب الأرقام القياسية يعاني من بعض العيوب، بالمقارنة مع اختيار سنة واحدة كأساس، تلخص في ميل الأساس المتخذ لأن يصبح قديماً وميل الأوزان لأن تصبح بعيدة عن الواقع، كلما اتسع الفاصل الزمني، بالإضافة إلى ذلك فإن اختيارها ليس واقعياً، لأن المرء لا يحيل إلى التفكير والتحليل في متوسط فترة من الزمن، وإنما يسعى دوماً إلى تناول فترة زمنية محددة وثابتة تتميز بعناصر هامة يتركب منها الرقم القياسي. ومن أهم المحاسن التي يتصف بها اختيار متوسط عدة سنوات كأساس، هو أن المؤثرات أو التقلبات المؤقتة التي قد تطرأ على الوضع الاقتصادي في بلد ما يكون تأثيرها في متوسط عدة سنوات أقل بكثير منها على بيانات سنة واحدة، وفي الحقيقة، فإن الهزات التي تؤثر في توازن مختلف أوجه النشاط الاقتصادي تدوم فترة قصيرة، وبالتالي فإنها تؤثر في بيانات سنة واحدة تأثيراً كبيراً، في حين أنه يمكن التغلب على هذا التأثير إذا كانت فترة الأساس المتخذة عبارة عن متوسط عدة سنوات، حيث تكون قيم الأساس أقرب إلى التوازن بين مختلف أوجه النشاط الاقتصادي، مما لو استخدمنا سنة واحدة كأساس، ونظراً لميل الباحث لأن يعامل فترة الأساس المتخذة كفترة سوية أو طبيعية، فإن اختيار متوسط عدة سنوات يعطي إلى حد كبير أساساً يمكن عده سوية أو طبيعياً.

ونجيب الإشارة إلى أن طول الفترة الزمنية التي يتم اختيارها كأساس يجب أن تكون كافية، بحيث تحتوي على دورة اقتصادية كاملة أو أكثر، وفي الحقيقة، فإن استعمال فترة أساس تحتوي على دورة أو عدة دورات اقتصادية تعطي فرصة للعوامل الإيجابية والسلبية للدورات الاقتصادية لأن تلغي بعضها بعضاً عند حساب متوسط هذه الفترة. وعلى العكس من ذلك، فإن استعمال فترة زمنية أقل أو أكثر من دورة أو دورات اقتصادية تعطي قيماً تميل لأن تكون إما مرتفعة كثيراً أو منخفضة كثيراً، وهذا يعتمد على ما إذا كان الجزء الأكبر من الفترة الزمنية المنتقاة فترة ركود أو ازدهار اقتصادي.

أما اختيار الأساس المتحرك بصورة مستمرة كفترة أساس، فإنه ينطوي على بعض الصعوبات عند استخدام أساس ثابت غير متحرك، وبحسب هذه الطريقة فإن الرقم القياسي يعبر عن المقارنات المرغوب إجراؤها للفترة الجارية والقائمة بين أسبوع

وآخر أو شهر وآخر أو سنة وأخرى، حيث يسهل تغيير الأوزان كلما دعت الحاجة إلى ذلك، فالعناصر التي لم تعد مهمة يمكن استبعادها من الرقم القياسي، في حين أن عناصر جديدة يمكن إضافتها بدون إعادة حساب لكامل الأرقام القياسية، ومن الواضح أن هذا يعطي مرونة غير ممكنة التحقيق لأية أرقام قياسية ذات أساس ثابت (Fixed Base). وعلى الرغم من هذه المرونة التي تتصف بها الأرقام القياسية ذات الأساس المتحرك، فإنها تشكو من بعض المساوئ أو العيوب، لعل من أخطرها هو أن إجراء أية مقارنة فيما بين أرقام الفترات الزمنية غير المتتالية هي من الأمور الصعبة. وهكذا لنفترض أن الجدول رقم (9-1) يبين طريقة الأساس المتحرك للاستخدام الشهري في صناعة النسيج بدمشق لعام 2001.

الجدول رقم (9-1)

طريقة الأساس المتحرك للاستخدام الشهري في صناعة النسيج

بمدينة دمشق لعام 2001

الشهر	طريقة الأساس المتحرك
كانون الثاني	95
شباط	98
آذار	105
نيسان	104
ايار	100
حزيران	99
تموز	98
آب	96
أيلول	100
تشرين أول	105
تشرين ثاني	108
كانون أول	99

المصدر : أرقام فرضية.

يبين الجدول رقم (1-9) حالة الاستخدام بين شهر وآخر وهكذا، فإذا نظرنا إلى حالة الاستخدام في شهر حزيران، فإننا نرى أنها 99% بالنسبة لحالة الاستخدام في شهر أيار، وإن حالة الاستخدام في شهر آب كانت 96% بالنسبة لحالة الاستخدام في شهر تموز، غير أنه لا يمكن القيام بإجراء المقارنة بصورة مباشرة لحالة الاستخدام بين شهري تشرين الأول وآذار أو بين أي شهرين غير متتاليين.

إن كلاً من شهري آذار وتشرين الأول له رقم قياسي يساوي 105، غير أن الرقم القياسي لشهر آذار يعني أن حالة الاستخدام في شهر آذار كانت أكثر ارتفاعاً بـ 5% بالنسبة لشهر شباط، وأن الرقم القياسي لشهر تشرين الأول يعني أن حالة الاستخدام في شهر تشرين الأول كانت أكثر ارتفاعاً بـ 5% بالنسبة لشهر أيلول، غير أن هذا لا يدلنا ما إذا كانت حالة الاستخدام في شهر تشرين الأول أكثر ارتفاعاً أو انخفاضاً عما كان عليه الحال في شهر آذار ولا يمكن ملاحظته مباشرة من هذا الرقم القياسي.

وتحب الإشارة إلى أنه يمكننا تأمين المقارنة بين الأشهر غير المتتالية إذا استعملنا أساساً ثابتاً، وبالتالي ضحينا ببعض المرونة التي قد يتصف بها الرقم القياسي عن طريق تحويل الأرقام القياسية ذات الأساس المتحرك إلى أساس ثابت، وذلك باستعمال طريقة السلسلة (Chain Index)، ولتوضيح هذه الطريقة، لنفترض أننا نرغب في تحويل الأرقام القياسية للجدول رقم (1-9)، وأنها نرغب في اتخاذ شهر كانون الثاني كأساس، في هذه الحالة، فإننا نستطيع أن نعطي بصورة تحكيمية الرقم القياسي لشهر كانون الثاني القيمة 100، نظراً لأن أساس أي رقم قياسي هو دوماً 100%، وتبعاً لذلك، فإن المنسوب المتصل (The Link Relative) لشهر شباط يعني أنه كان 98% بالنسبة لشهر كانون الثاني، كما أننا نأخذ 97% من 100 حيث نحصل على 98 وهو الرقم القياسي لشهر شباط، وإن المنسوب المتصل لشهر آذار يدل على أنه كان 105% من شهر شباط، غير أن شهر شباط هو 98 إذا اتخذنا شهر كانون الثاني كأساس، وبالتالي فإننا نأخذ 105% من الـ 98 حيث نحصل على 102.9 وهو الرقم القياسي لشهر آذار إذا اتخذنا شهر

كانون الثاني كأساس، وإن المنسوب المتصل لشهر نيسان يبين أن شهر نيسان كان 104% من شهر آذار، غير أن شهر آذار إذا اتخذنا شهر كانون الثاني كأساس هو 102.9، لذلك فإننا نأخذ 104% من 102.9 حيث نحصل على 107، وهو الرقم القياسي لشهر نيسان إذا اتخذنا شهر كانون الثاني كأساس وهكذا...

ويمكن القول بشكل عام، إن إمكان تحويل الأرقام القياسية ذات الأساس المتحرك إلى أساس ثابت عن طريق استعمال طريقة السلسلة يمكن عده على جانب كبير من الأهمية بالنسبة لتركيب أنواع معينة من الأرقام القياسية، وبخاصة أرقام الاستخدام، والرواتب، والإنتاج، والمبيعات، وأن البيانات اللازمة لحساب أرقام قياسية من هذا النوع لا يمكن توفرها في كثير من البلدان إلا عن طريق التعاون الطوعي لأرباب العمل والمنتجين والموزعين الذين يوافقون على تقديم تقارير شهرية عن أحوال المؤسسات التي يشرفون عليها، وأن المؤسسات العاملة في مجال الصناعة أو التجارة والتي توافق على التعاون مع إحدى المؤسسات الحكومية في تقديم المعلومات اللازمة لتركيب الأرقام القياسية، هي عادة تلك المؤسسات الأكثر تقدمية والأكثر إحساساً بواجبها الاجتماعي، وبالتالي فإن تجاربها تكون عادة أحسن، بطريقة أو بأخرى، من غيرها من المؤسسات العادية العاملة في المجال نفسه، فعدد مستخدميها وجدول رواتبها مثلاً يثبت عادة بصورة أقل في فترة الركود الاقتصادي، ويزداد أكثر في فترة الازدهار الاقتصادي بالمقارنة بسائر المؤسسات الأخرى... وهكذا، فإن العينة التي تؤخذ في تركيب هذا النوع من الأرقام القياسية تميل لأن تمثل المؤسسات الأكثر نجاحاً، والأحسن تجربة في أي حقل من حقول الصناعة أو التجارة التي تؤخذ منه لتمثيله.

5- اختيار الأوزان المناسبة :

إن عملية اختيار الأوزان المناسبة للسلع الداخلة في تركيب الرقم القياسي يتوقف على البيانات المتوفرة لدينا من جهة، وعلى الغرض الذي من أجله يتم تركيب الرقم القياسي من جهة أخرى. ولكن هذه الأرقام القياسية يجب أن تعكس الأهمية النسبية

للسلع المختلفة الداخلة في تركيبه، فمثلاً الوزن المناسب للرقم القياسي لأسعار التجزئة هو كمية المستهلك من السلع الداخلة في تركيبه، وليس كمية الإنتاج منها، فإذا ما توفرت لدينا بيانات عن كميات الاستهلاك أو كميات الإنتاج من السلع المختلفة في سنة الأساس، نعمل على استخدامها كأوزان، بالرغم من أن هذه السنة قد تكون بعيدة نوعاً ما بحيث تكون الأهمية النسبية للسلع المختلفة قد تغيرت خلال المدة الطويلة بين سنة الأساس وسنة المقارنة، أما إذا وجد لدينا بيانات عن كميات الاستهلاك أو كميات الإنتاج من السلع المختلفة، موضوع دراستنا في سنة المقارنة، يكون من الأفضل استخدامها كأوزان حيث إنها تعبر تعبيراً صحيحاً عن الأهمية النسبية الحاضرة لهذه السلع، ومن الناحية النظرية يوجد خلاف جوهري حول هذا الموضوع، حيث يعتقد لاسبير أنه من الأفضل اتخاذ أوزان سنة الأساس للترجيح، لأن أوزان سنة المقارنة تكون في الغالب متأثرة بتغير أسعار السلع المختلفة، ولذلك يؤدي استخدامها إلى تضخيم تأثير الأسعار على نتيجة الرقم القياسي، بينما يهدف الرقم إلى قياس التغيرات الواقعية في الأسعار دون أي تضخم، وعكس ذلك يعتقد باش أنه من الأفضل استخدام أوزان سنة المقارنة، لأنها هي التي تعبر تعبيراً صحيحاً عن الأهمية النسبية الحاضرة للسلع الداخلة في تركيب الرقم القياسي.

وتجيب الإشارة إلى أنه إذا ما أريد إجراء المقارنة بين أرقام مختلف السنوات فلا بد من جعل الأوزان ثابتة وأثناء تركيب الأرقام القياسية للأسعار، وإذا بدلنا الأوزان في كل سنة، فإن التغيرات في الأرقام القياسية من سنة إلى أخرى لن تعكس التقلبات في الأسعار فحسب بل التبدلات في الأوزان المستعملة أيضاً، هذا من جهة ومن جهة أخرى إذا ما أريد للأرقام القياسية أن تكون قريبة من الواقع، فإن الأوزان المستعملة يجب أن تصف نموذج الحياة الاقتصادية أو نمطها، ومن الواضح فإن هناك تضارباً بين هذين الالتزامين بصورة متبادلة، نظراً لأن الأوزان المستعملة لا يمكن أن تكون قريبة إلى الواقع، وفي الوقت نفسه ثابتة، وفي الواقع، لو أخذنا مثلاً الأرقام القياسية لتكاليف المعيشة فإننا نجد أنها لا تقيس التغيرات التي تطرأ على تكاليف المعيشة أبداً، وإنما تقيس

التأثير الذي قد يكون لتغيرات الأسعار فيما لو بقي نموذج الاستهلاك ثابتاً، ولذلك يفضل عند تركيب الأرقام القياسية لتكاليف المعيشة اختيار الأوزان بناء على دراسة ميدانية بالعينة عن ميزانية الأسرة للتعرف على متوسط ما تنفقه الأسرة من دخلها الشهري على السلع والخدمات المختلفة التي تدخل في إطار معيشتها، هذه المتوسطات لنسب الإنفاق على السلع والخدمات المختلفة تشكل الأوزان المناسبة التي تستخدم في ترجيح أسعار هذه السلع والخدمات لإبراز أهميتها النسبية ضمن النتيجة التي يحققها الرقم القياسي لتكاليف المعيشة. وواضح أن هذه الدراسة الميدانية لميزانية الأسرة تجري في فترة زمنية معينة، ويعمل بنتائجها كأوزان طالما أن نمط إنفاق المجتمع لا يتغير، فإذا أدركنا أن هناك تغيراً جوهرياً في هذا النمط يعاد إجراء الدراسة من جديد، لتحديد التغير الذي حدث في الأهمية النسبية لنود الإنفاق المختلفة، ونظراً لأن الأوزان التي تستخدم في تركيب الأرقام القياسية تضرب بالأسعار في البسط، وكذلك في المقام كما سيتضح ذلك من مناقشات طرائق تركيب الأرقام القياسية، لذلك لا يؤثر الخطأ البسيط على النتيجة تأثيراً كبيراً. لهذا يفصل دائماً أن نأخذ أعداداً صحيحة، ونحمل الكسور الخاصة بالأوزان لتسهيل العمليات الحسابية الخاصة بحساب الأرقام القياسية.

6- اختيار الأسعار المناسبة :

يعتمد اختيار الأسعار على الهدف والغاية التي من أجلها يتم تركيب الرقم القياسي من جهة، وعلى اختيار مرحلة معينة من مراحل إنتاج السلع الداخلة في تركيب الرقم القياسي من جهة أخرى، لذلك يجب علينا أن نقرر في أي مرحلة من عمليات الإنتاج نريد التحقق من الأسعار، أي هل ستكون أسعار الجملة أو أسعار التجزئة أو أسعار الإنتاج هي التي سوف تستخدم في حساب الرقم القياسي؟ وكيف يمكننا الحصول على هذه الأسعار؟ حيث إن هناك أكثر من سعر نتيجة لتعدد جهات البيع والوكلاء المختلفين، بالإضافة إلى أن العديد من المصانع تنتج السلع نفسها الداخلة في تركيب الرقم، وعندما كانت هناك صعوبة كبيرة في الحصول على جميع الأسعار

المطلوبة للسلع من قبل محلات البيع المختلفة، نرى أنه من الضروري أخذ عينة من هذه الأسعار تكون ممثلة تمثيلاً صادقاً لمجتمع الأسعار المدروس، وبحيث تتوافق الأسعار المستخدمة مع الأسعار الحقيقية المطلوبة من قبل محلات البيع المختلفة، تجنب الإشارة إلى أن الأسعار تتغير بشكل دائم وفقاً لحجم السلع المعروضة، والطلب عليها في زمان ومكان محددين، فإن التغير في أسعار التجزئة مثلاً يعكس تأثير العوامل الفعلية المختلفة، وفي مقدمة هذه العوامل نذكر أثر التقدم العلمي والتقني في فترة سنة واحدة، ولذلك تعدّ أسعار التجزئة من المسائل الإحصائية الهامة التي يجب أن تبقى دوماً في إطار الملاحظة العلمية، وذلك نظراً لما يحدثه التغير في أسعار البضائع المشتراة من قبل السكان من أثر كبير على مستواهم المعاشي، بالإضافة إلى الآثار الكبيرة التي تتركها تغيرات الأسعار على السوق التي يتم دراستها من خلال تطبيق الأرقام القياسية كمؤشرات للحالة القائمة.

لذلك نجد عند تركيب الرقم القياسي لأسعار التجزئة الخاضعة لسيطرة الحكومة، أن من الخطأ أن تقتصر على إدراج الأسعار المطلوبة للسلع من قبل الشركات والمحلات التجارية، وإنما يترتب علينا عندما نريد الحصول على أسعار سلعة معينة أن نتأكد من أن هذه السلعة هي نفسها في جميع الأوقات التي نحصل منها على الأسعار، أي يجب علينا تثبيت مستوى معين من النوعية لكل سلعة والحفاظ عليه دائماً.

7- اختيار المعادلة التي سوف تستخدم في تصميم الرقم القياسي :

بعد الانتهاء من عملية اختيار العوامل السابقة المكونة للرقم القياسي ذي الغرض المحدد والهدف المنشود، يأتي دور اختيار المعادلة المناسبة التي ستستخدم في تركيب الرقم القياسي، وعملية الاختيار هذه تتوقف على التصنيف المسبق للمعادلات المختلفة المقترن بترتيب منطقي للأوزان اللازمة الترجيح ، متغيرة كانت أم ثابتة، وبناء على ذلك فإنه يجب استبعاد المعادلات البسيطة غير المرجحة مثل الرقم القياسي التجميعي البسيط، والرقم القياسي لمناسيب الأسعار من عملية الاختيار هذه، مع العلم أن اختيار المعادلات التي تعتمد على أوزان متغيرة يتطلب معلومات إحصائية عن الكميات، والأسعار بشكل دائم ومستمر أكثر بكثير من تلك المعادلات التي تعتمد على أوزان متغيرة يكون

في معظم الحالات غير عملي وصعب التفسير، في حين أن المعادلات التي تقوم على أساس الأوزان الثابتة تكون أكثر وضوحاً في الحياة العملية، على الرغم من ضرورة إعادة النظر في الأوزان الثابتة من فترة زمنية إلى أخرى، وتبديل فترة الأساس إلى فترة زمنية أحدث، وذلك لكي يكون للأرقام القياسية الأهمية التي من أجلها يتم حسابها.

إن هذا النوع من المعادلات يقدم دليلاً واضحاً عن التغيرات التي تحدث في كلفة مجموعة سلعية معينة من خلال المقارنات المباشرة بين فترتين زمنيتين، أو بواسطة الأرقام القياسية المتسلسلة. وهذا ما يجعل معادلة الرقم القياسي القائم على أساس الوزن الثابت من المعادلات الأكثر شيوعاً واستخداماً في الحياة العملية، مع العلم أنه تستخدم أحياناً معادلات مركبة في تكوين الرقم القياسي تحتاج إلى عمليات حسابية طويلة، كاستخدام الرقم القياسي الأمثل لفيشر في حساب الأرقام القياسية للتجارة الخارجية. ومن معادلات الأرقام القياسية الشائعة الاستخدام نذكر: الرقم القياسي التجميعي للأسعار المرجحة بكميات ثابتة، والرقم القياسي لمناسيب الأسعار المرجحة بأوزان ثابتة.

وبالتالي فإن عملية اختيار المعادلة المناسبة لتركيب الرقم القياسي تعتبر من المهام الصعبة، لأنه من الواضح والبدهي لا توجد أية معادلة لرقم قياسي يمكن أن نسميها «أفضل معادلة» (Best formula)، حيث إن لكل معادلة معنى دقيقاً خاصاً بها، وأغلبها يقيس الكلفة الإجمالية لمجموعة معينة من السلع، وهناك بالطبع معادلات متنوعة متاحة، يمكن اختيار المناسب منها عند تركيب الرقم القياسي، غير أنه يجب عدم استخدام أية معادلة لا تحقق شرط الترتب المنطقي للأوزان، ولا تسمح بوضع نظام محدد ومقبول للترجيح.

9-3 طرائق تركيب الأرقام القياسية :

يعتمد تركيب الرقم القياسي على مجموعة من الطرائق الهامة التي تستند من حيث الجوهر على اختيار فترة الأساس، بالإضافة إلى العوامل الأخرى التي سبق عرضها ودراستها.

ومن هذه الطرائق: طريقة التجميع وطريقة المناسيب، حيث يمكن تطبيق أي طريقة من هاتين الطريقتين وفق صيغ مختلفة يمكن اختيار المناسب منها، وفقاً للإمكانيات العملية المتوفرة، والتي تحقق الهدف من تركيب الرقم القياسي، فإذا كان الرقم القياسي خاصاً بالأسعار، فإن الصيغة التجميعية تقضي بقسمة متوسط أسعار السلع المدروسة في سنة المقارنة إلى متوسط أسعارها في سنة الأساس وضرب النتيجة بـ 100، أما إذا كان الرقم القياسي خاصاً بالكميات، فإن هذه الصيغة تقوم على أساس قسمة متوسط الكميات في سنة المقارنة إلى متوسط الكميات في سنة الأساس وضرب النتيجة بـ 100، وإذا كان الرقم القياسي خاصاً بالقيم، فإن هذه الصيغة تتمثل بقسمة متوسط قيم السلع المدروسة في سنة المقارنة إلى متوسط قيم السلع المدروسة في سنة الأساس، وضرب النتيجة بـ 100، لأن الرقم القياسي يعبر عنه دوماً في شكل نسبة مئوية، وهو بهذا يختلف عن المتوسط الحسابي الذي يمكن أن يكون بسيطاً، أي تعتبر السلع التي تدخل في تركيب الرقم القياسي متساوية في الأهمية، ويمكن أن يكون مرجحاً حيث تعطى كل سلعة وزناً يدل على أهميتها النسبية بالمقارنة مع السلع الأخرى الداخلة في تركيب الرقم القياسي موضوع الدراسة مثل الرقم القياسي التجميعي البسيط. وقبل البدء في دراسة الطرائق المتاحة لتركيب الرقم القياسي، نرى أنه من الضروري وضع دليل خاص بالرموز والمصطلحات التي سنستخدمها في هذا الفصل، وذلك نظراً لأهميتها واستخداماتها في أثناء مناقشة هذه الطرائق والفقرات الأخرى اللاحقة .

الرموز والمصطلحات :

- نرمز للسعر (Price) بالحرف P .
- نرمز للكمية (quantum) بالحرف q .
- نرمز للقيمة (Value) بالحرف V .
- الرمز (o) يشير إلى فترة الأساس .
- الرمز (n) يشير إلى فترة المقارنة .

الرمز (P_n) يشير إلى السعر في فترة المقارنة .

الرمز (q_n) يشير إلى الكمية في فترة المقارنة .

الرمز (V_n) يشير إلى القيمة في فترة المقارنة .

الرمز (P_o) يشير إلى السعر في فترة الأساس .

الرمز (q_o) يشير إلى الكمية في فترة الأساس .

الرمز (V_o) يشير إلى القيمة في فترة الأساس .

الرمز $(P_R = x)$ (Price Relative) يشير إلى منسوب السعر .

الرمز $(q_R = x)$ (Quantum Relative) يشير إلى منسوب الكمية .

الرمز $(V_R = x)$ (Value Relative) يشير إلى منسوب القيمة .

الرمز (I) (Index numbers) يشير إلى الرقم القياسي .

الرمز (P_I) (Price Index number) يشير إلى الرقم القياسي للأسعار .

الرمز (q_I) (Quantum Index number) يشير إلى الرقم القياسي للكميات .

الرمز (V_I) (Value Index number) يشير إلى الرقم القياسي للقيمة .

الرمز $(\sum P_n)$ يشير إلى مجموع أسعار السلع في فترة المقارنة .

الرمز $(\sum P_o)$ يشير إلى مجموع أسعار السلع في فترة الأساس .

الرمز $(\sum P_n q_n)$ يشير إلى كميات فترة المقارنة، مقومة بأسعار فترة المقارنة .

الرمز $(\sum P_o q_o)$ يشير إلى كميات فترة الأساس، مقومة بأسعار فترة الأساس .

الرمز $(\sum P_o q_n)$ يشير إلى كميات فترة المقارنة، مقومة بأسعار فترة الأساس .

الرمز $(\sum P_n q_o)$ يشير إلى كميات فترة الأساس، مقومة بأسعار فترة المقارنة .

الرمز $P_{I(L)}$ يشير إلى الرقم القياسي التجميعي المرجح للأسعار (لاسيير) .

الرمز $P_{I(P)}$ يشير إلى الرقم القياسي التجميعي المرجح للأسعار (باش) .

الرمز $P_{I(n)}$ يشير إلى الرقم القياسي التجميعي المرجح للأسعار (فيشر) .

الرمز $q_{I(L), (P), (n)}$ يشير إلى الرقم القياسي التجميعي المرجح للكميات (لاسيير -

باش - فيشر) .

الرمز $(V_n = P_n q_n)$ يشير إلى القيمة في سنة المقارنة .

الرمز $(V_n = P_n q_n)$ يشير إلى القيمة في سنة الأساس .

الرمز $(V_1 = P_1 q_1)$ يشير إلى القيمة في سنة مختارة .

الرمز $\sum \frac{P_n}{P_0}$ يشير إلى مجموع الأسعار النسبية لجميع السلع الداخلة في تركيب

الرقم القياسي.

الرمز $\sum \frac{q_n}{q_0}$ يشير إلى الكميات النسبية لجميع السلع الداخلة في تركيب الرقم

القياسي.

9-3-1 طريقة المناسيب في تركيب الأرقام القياسية (Relatives Method) :

تعدّ طريقة المناسيب من الطرائق الهامة في تركيب الرقم القياسي، وتقوم على أساس التركيز على حركات الأسعار أو الكميات أو القيم للسلع الفردية التي تدخل في تركيب الرقم القياسي بين نقطتي الأساس والمقارنة، وباستخدام هذه الطريقة تصبح نتيجة الرقم القياسي تعبر تعبيراً دقيقاً عن حقيقة تغيرات أسعار وكميات وقيم السلع التي يتضمنها الرقم القياسي في السوق موضوع الدراسة.

كما أنه يمكن التعبير عن الرقم القياسي لعدة سلع داخلة في تركيبه، عن طريق استخدام متوسطات المناسيب للأسعار، أو للكميات، أو لقيم هذه السلع، والمتوسط الذي نحسبه للمناسيب يكون إما وسطاً حسابياً أو هندسياً أو توافقياً حيث يتم حسابه على أساس أن يكون وسطاً بسيطاً أو وسطاً مرجحاً.

ولهذا نرى أنه من الضروري دراسة طريقة المناسيب من خلال التمييز بين المناسيب البسيطة والمناسيب المرجحة والمتوسطات المستخدمة في الحالتين.

9-3-1-1 المناسيب البسيطة :

تعدّ المناسيب البسيطة من أسهل وأبسط الأمثلة للرقم القياسي، وتقوم على أساس استخدام الأسعار، أو الكميات، أو القيم لكل سلعة على حدة تدخل في تركيب

الرقم القياسي، فإن حساب التغيرات في أسعار سلعة معينة بين نقطة الأساس ونقطة المقارنة يمكن الحصول عليها عن طريق النسبة المئوية التي تعرف بـ منسوب السعر :

$$P_R = \frac{P_n}{P_o}$$

وإن حساب التغير الكمي في سلعة معينة يتم من خلال استخدام ما يعرف بمنسوب الكمية :

$$q_R = \frac{q_n}{q_o}$$

وكذلك فإن حساب التغير في قيمة سلعة معينة بين فترتي الأساس والمقارنة يتم من خلال تطبيق منسوب القيمة :

$$V_R = \frac{\text{قيمة الظاهرة في فترة المقارنة}}{\text{قيمة الظاهرة في فترة الأساس}} = \frac{V_n}{V_o}$$

ويهدف تحديد طبيعة وحجم التغيرات التي تطرأ على الأسعار والكميات والقيم للسلع الداخلة في تركيب الرقم القياسي، سوف تميز بين مناسيب الأسعار ومناسيب الكميات ومناسيب القيمة .

— مناسيب الأسعار :

إن منسوب السعر لأي سلعة تدخل في تركيب الرقم القياسي، هو نسبة سعرها في فترة المقارنة إلى سعرها في فترة الأساس أو فترة الإسناد، ولحساب هذه النسبة نفترض أن الأسعار ثابتة لأي فترة زمنية، وإذا لم يكن هذا الفرض صحيحاً، فإنه يمكن استخدام متوسط ملائم لهذه الفترة حتى نجعل هذا الفرض صحيحاً، فإذا كانت P_o تمثل سعر السلعة خلال فترة الأساس و P_n سعرها خلال فترة المقارنة فإن منسوب السعر :

$$P_{R(o,n)} = x_{(o,n)} = \frac{P_n}{P_o} \cdot 100$$

وبصورة عامة إذا كانت P_b , P_a هي أسعار سلعة خلال الفترات a , b على الترتيب، فإن منسوب السعر في الفترة b بالنسبة للفترة a يعرف بأنه :

$$\frac{P_b}{P_a} \cdot 100$$

إذا فرضنا أن أسعار الكيلو غرام من مادة اللحم في عام 2000 هو 180 وحدة نقدية، وفي عام 1995 هو 150 وحدة نقدية .

فإن منسوب سعر مادة اللحم لعام 2000 بالنسبة لعام 1995 :

$$X_{(1995/2000)} = \frac{P_n}{P_o} \cdot 100 = \frac{180}{150} \cdot 100 = 120 \%$$

وهذا يعني أن أسعار مادة اللحم قد زادت في عام 2000 عنه في عام 1995 بنسبة 20% أي زيادة عن الرقم 100% .

مثال (6-9) :

يفرض أن أسعار السلعة a في السنوات 2001 , 1998 هي 90 ، 75 على الترتيب فإذا اعتمدنا 1998 كسنة أساس و 2001 سنة مقارنة، فإن منسوب السعر لهذه السلعة:

$$\begin{aligned} X_{(1998/2001)} &= \frac{\text{السعر في عام 2001}}{\text{السعر في عام 1998}} \cdot 100 \\ &= \frac{P_n}{P_o} \cdot 100 = \frac{90}{70} \cdot 100 = \frac{6}{5} \cdot 100 = 120 \% \end{aligned}$$

باختصار 120 بحذف علامة % كما هو متبع غالباً في المؤلفات الإحصائية، تعني هذه النتيجة ببساطة أن سعر السلعة a سنة 2001 أصبح 120% من سعرها في سنة 1998 أي أنه زاد بنسبة 20% عن سنة الأساس.

أما إذا أخذنا 2001 سنة أساس و 1998 سنة مقارنة في المثال فإن منسوب السعر:

$$X_{(2001/1998)} = \frac{\text{السعر في عام 1998}}{\text{السعر في عام 2001}} \cdot 100$$

$$= \frac{P_o}{P_n} \cdot 100 = \frac{5}{6} \cdot 100 = 83.33 \%$$

أي أن سعر السلعة a في عام 1998 كان يعادل 83.33 % من سعرها في عام 2001 أي أنه كان ينقص بنسبة 16.67 % .

مثال (7-9) :

بفرض أن سعر الكيلو غرام من البطاطا في السوق عام 2000 في شهر أيار 15 وحدة نقدية وفي شهر آب 30 وحدة نقدية K فإن منسوب سعر مادة البطاطا في شهر آب بالنسبة لشهر أيار يعادل :

$$X_{(أب/أيار)} = \frac{P_n}{P_o} \cdot 100 = \frac{30}{15} \cdot 100 = 200 \%$$

وهذا يعني أن سعر الكيلو غرام من البطاطا قد زاد بنسبة 100%، بالمقارنة مع سعر الكيلو غرام من البطاطا في شهر أيار من العام نفسه. وإذا بدلنا نقطة الأساس إلى شهر آب ونقطة المقارنة في شهر أيار، فإن منسوب السعر في هذه الحالة :

$$X_{(أب/أيار)} = \frac{15}{30} \cdot 100 = 50 \%$$

وتفسير ذلك أن سعر الكيلو غرام من البطاطا في شهر آب كان يعادل 50% من سعرها في شهر أيار أي بنقص قدره 50% .

مثال (8-9) :

ليكن لدينا البيانات الآتية عن أسعار وكميات أربع سلع في عام 2001 كسنة مقارنة باستخدام 1998 كسنة أساس، والمطلوب: تركيب الرقم القياسي للأسعار عام 2001، والجدول الآتي يبين أسعار وكميات الاستهلاك لهذه السلع موضوع الدراسة .

جدول رقم (9-2)

السلعة	أسعار عام 1998 (بالدولارات)	أسعار عام 2001 (بالدولارات)	كميات الاستهلاك عام 2001 (بالأطنان)
A	6	8	50
B	9	11	70
C	7	6	30
D	21	23	60

الحل:

الجدول الآتي يعطي مناسيب أسعار السلع الأربع .

جدول رقم (9-3)

السلعة	مناسيب الأسعار $X_{(o/n)} = P_n / P_o . 100$
A	$X_a = 8/6 . 100 = 133.33$
B	$X_b = 11/9 . 100 = 122.22$
C	$X_c = 6/7 . 100 = 85.71$
D	$X_D = 23/21 . 100 = 109.52$

نلاحظ أن كل منسوب من مناسيب الأسعار للسلع الأربع يقيس التغير النسبي في سعر السلعة الخاص به، فإن منسوب السعر للسلعة A يدل على ارتفاع سعرها بنسبة 33.33% عام 2001 بالمقارنة بعام الأساس 1998 وهكذا لبقية السلع B, C, D .
وحيث إن حساب الرقم القياسي يهدف إلى قياس التغير في أسعار وكميات عدة سلع سوياً، لذلك فإنه بعد حساب مناسيب الأسعار لكل سلعة من هذه السلع الأربع التي تدخل في تركيب الرقم القياسي يتوجب علينا اختيار المتوسط الملائم بهدف

حساب متوسط مناسب أسعار هذه السلع الأربع مجتمعة، ورغم أنه من الوجهة النظرية يمكن تطبيق أي مقياس من مقاييس النزعة المركزية لهذه الغاية، فإن الوسط الحسابي والوسط الهندسي والوسط التوافقي تعتبر من أهم المتوسطات من الناحية العملية، أما المتوسطات الأخرى كالوسط والمنوال، فإن أيًا منهما ليس مرضياً نظراً لعدم أخذ جميع القيم بعين الاعتبار، وعلى الرغم من أن هذا يعدّ من المزايا الرئيسية للوسط أو المنوال، فإنه من غير المرغوب فيه عند حساب مناسب الأسعار أن نقلل من تأثير تغيرات الأسعار.

1- الوسط الحسابي للمناسيب البسيطة للأسعار :

تعدّ طريقة حساب المتوسط الحسابي لمناسيب الأسعار من أسهل الطرائق المستخدمة في تركيب الرقم القياسي، فإذا ما أخذنا متوسط النسب السابقة في المثال (8-9) فإننا نحصل على الرقم القياسي المطلوب كما يأتي :

منسوب سعر السلعة A :

$$X_a = \frac{8}{6} \cdot 100 = 133.33$$

منسوب سعر السلعة B :

$$X_b = \frac{11}{9} \cdot 100 = 122.22$$

منسوب سعر السلعة C :

$$X_c = \frac{6}{7} \cdot 100 = 85.71$$

منسوب سعر السلعة D :

$$X_D = \frac{23}{21} \cdot 100 = 109.52$$

الوسط الحسابي للمناسيب البسيطة لأسعار السلع الأربع التي يتضمنها الرقم

القياسي :

$$\bar{X}_{(a, b, c, D)} = \frac{X_a + X_b + X_c + X_D}{n}$$

وبما أن أية قيمة X تساوي $\frac{P_n}{P_0}$ فإن الرقم القياسي المتوسط للمناسيب البسيطة

لأسعار n سلعة يصبح :

$$PI_{(\bar{x})} = \frac{\sum (P_n / P_0)}{n}$$

حيث : $X = \frac{P_n}{P_0}$ - منسوب السعر لكل سلعة من هذه السلع الأربع.

n - عدد السلع الداخلة في تركيب الرقم القياسي. وفي مثالنا

($n = 4$) وبالتالي:

$$PI_{(\bar{x})} = \frac{133.33 + 122.22 + 85.71 + 109.52}{4} = 112.70$$

أو :

$$PI_{(\bar{x})} = \frac{1.3333 + 1.2222 + 0.8571 + 1.0952}{4} \cdot 100 = 112.70$$

وتفسير ذلك أن أسعار هذه السلع الأربع التي يتضمنها الرقم القياسي قد

ارتفعت في المتوسط بنسبة 12.70% عام 2001 بالمقارنة مع عام 1998 .

يؤخذ على هذا الرقم أنه في بعض الأحيان تكون بعض السلع كبيرة الإنتاج

بالمقارنة مع غيرها من السلع المعطاة، وحينئذ لا بد من أخذ الكميات المنتجة بعين

الاعتبار في حساب الرقم القياسي المتوسط للمناسيب البسيطة للأسعار بدلاً من عدد

السلع، كما ويعاب على استخدام الوسط الحسابي في هذه الحالة أنه يعطي نتائج غير

دقيقة، فمثلاً لو أن منسوب سلعة قد تضاعف ومنسوب آخر قد نقص إلى النصف،

فالمنتطقي أن المتوسط هذين المنسوبين سوف لا يتغير، ولكن بحساب المتوسط الحسابي

سوف يتغير بنسبة 25% إذا كان لدينا :

$$\frac{2.0 + 0.50}{2} = 1.25$$

(راجع الوسط الهندسي في مقياس السرعة المركزية) لهذا السبب نرى أن المتوسط الهندسي في التعامل مع المناسيب لا يعاني من هذا القصور حيث إن :

$$\sqrt{0.50 \times 2.0} = 1$$

لهذا نرى أنه عندما يراد حساب متوسط مناسيب، فإنه يفضل استخدام الوسط الهندسي بدلاً من الوسط الحسابي.

2- الوسط الهندسي للمناسيب البسيطة للأسعار :

يعتمد هذا الرقم على حساب نسبة السعر في السنة الحالية إلى سنة الأساس ثم استخدام صيغة الوسط الهندسي في حساب الرقم القياسي المتوسط للمناسيب البسيطة للأسعار، حيث نعلم أن الوسط الهندسي لمجموعة من القيم هو حاصل ضرب هذه القيم بعضها ببعض محذورة إلى قوة تساوي عددها أي الوسط الهندسي :

$$G = \sqrt[n]{X_1 \cdot X_2 \cdot X_3 \cdot \dots \cdot X_n}$$

وفي مثالنا السابق الوسط الهندسي :

$$G = \sqrt[4]{X_a \cdot X_b \cdot X_c \cdot \dots \cdot X_D}$$

نستبدل X بالقيم التي تأخذها كمناسيب للسلع الأربع كما يأتي :

$$X_a = \frac{P_n^{(a)}}{P_0^{(a)}}, \quad X_b = \frac{P_n^{(b)}}{P_0^{(b)}}, \quad X_c = \frac{P_n^{(c)}}{P_0^{(c)}}, \quad X_D = \frac{P_n^{(D)}}{P_0^{(D)}}$$

ومنه يصبح الرقم القياسي الهندسي للمناسيب البسيطة لأسعار السلع الأربع في

المثال (9-8) كما يأتي :

$$PI_{(0)} = \sqrt[4]{\frac{P_n^{(a)}}{P_0^{(a)}} \cdot \frac{P_n^{(b)}}{P_0^{(b)}} \cdot \frac{P_n^{(c)}}{P_0^{(c)}} \cdot \frac{P_n^{(d)}}{P_0^{(d)}}} \cdot 100$$

$$PI_{(0)} = \sqrt[4]{\frac{8}{6} \cdot \frac{11}{9} \cdot \frac{6}{7} \cdot \frac{23}{21}} \cdot 100$$

$$PI_{(0)} = \sqrt[4]{(133.33) \cdot (122.22) \cdot (85.71) \cdot (109.52)}$$

$$PI_{(0)} = \sqrt[4]{152966062.2} = 111.21$$

وهذا يعني أن أسعار السلع الأربع قد زادت بالمتوسط بنسبة 11.21% في عام 2001 عما كانت عليه أسعار هذه السلع في عام 1998 .
إن العلاقة السابقة للوسط الهندسي يمكن استخدامها، إذا كانت السلع لدينا قليلة العدد، أما إذا كانت السلع كثيرة العدد، فعندئذٍ يجب الاستعانة باللوغاريتمات حيث يمكن أن تكتب المعادلة السابقة أعلاه كما يأتي :

$$\log PI_{(0)} = \frac{1}{n} \sum \log \left(\frac{P_n}{P_0} \right) \cdot 100$$

نأخذ لوغاريتم الطرفين أو تكتب :

$$\log PI_{(0)} = \frac{\sum \log \left(\frac{P_n}{P_0} \right)}{n} \cdot 100$$

حيث : $\log G$ - لوغاريتم الوسط الهندسي للمناسيب البسيطة للأسعار.

$\log \left(\frac{P_n}{P_0} \right)$ - لوغاريتم منسوب السعر لكل سلع داخلية في تركيب الرقم القياسي.

n - عدد السلع الداخلة في تركيب الرقم القياسي.

وفي مثالنا السابق يجب أن نحسب أولاً : $\log \left(\frac{P_n}{P_o} \right) \cdot 100$ لكل سلعة من السلع

الأربعة :

$$\log X_a = \log \frac{P_n^{(a)}}{P_o^{(a)}} \cdot 100 = \log \frac{8}{6} \cdot 100 = \log 133.33 = 2.1249$$

$$\log X_b = \log \frac{P_n^{(b)}}{P_o^{(b)}} \cdot 100 = \log \frac{11}{9} \cdot 100 = \log 122.22 = 2.0871$$

$$\log X_c = \log \frac{P_n^{(c)}}{P_o^{(c)}} \cdot 100 = \log \frac{6}{7} \cdot 100 = \log 85.71 = 1.9330$$

$$\log X_D = \log \frac{P_n^{(D)}}{P_o^{(D)}} \cdot 100 = \log \frac{23}{21} \cdot 100 = \log 109.52 = 2.0395$$

وحسب قاعدة لوغاريتم جداء = مجموع اللوغاريتمات يكون :

$$\log PI_{(G)} = \frac{1}{n} \left(\log \frac{P_n^{(a)}}{P_o^{(a)}} + \log \frac{P_n^{(b)}}{P_o^{(b)}} + \log \frac{P_n^{(c)}}{P_o^{(c)}} + \log \frac{P_n^{(D)}}{P_o^{(D)}} \right)$$

$$\log PI_{(G)} = \frac{1}{4} (2.1249 + 2.0871 + 1.9330 + 2.0395)$$

$$\log PI_{(G)} = \frac{1}{4} (8.1845) = 2.046125$$

$$\Rightarrow PI_{(G)} = \frac{2.046125}{10} = 111.21$$

وهو الرقم نفسه الذي حصلنا عليه سابقاً .

يؤخذ على هذا الرقم اعتماده على الجداول اللوغاريتمية، ولكن مع انتشار الحاسبات الإلكترونية قلل كثيراً من أهمية هذا العيب، ويمكن أن نشير إلى أن المتوسط الحسابي للمناسيب سوف يكون دائماً أكبر من المتوسط الهندسي أي :

$$PI(\bar{x}) > PI_{(G)}$$

3- الوسط التوافقي للمناسيب البسيطة للأسعار :

يقوم هذا الرقم على أساس حساب منسوب السعر، أو الكمية للسلع الداخلة في تركيب الرقم القياسي في فترة المقارنة بالاستناد إلى فترة الأساس ، ثم استخدام علاقة الوسط التوافقي في حساب الرقم القياسي المتوسط للمناسيب البسيطة للأسعار، حيث علمنا من دراسة مقاييس النزعة المركزية أن الوسط التوافقي لمجموعة من القيم يساوي إلى مقلوب الوسط الحسابي لمقلوب القيم، وأن علاقته تأخذ الشكل الآتي :

$$H = \frac{n}{\sum \frac{1}{X}}$$

وتطبق هذه العلاقة على مناسيب الأسعار للسلع الأربع الداخلة في تركيب الرقم القياسي في المثال (8-9) نحصل على الرقم القياسي المتوسط للمناسيب البسيطة للأسعار (طريقة الوسط التوافقي) أي :

$$PI_{(H)} = \frac{n}{\frac{1}{X_a} + \frac{1}{X_b} + \frac{1}{X_c} + \frac{1}{X_d}} \cdot 100$$

$$PI_{(H)} = \frac{4}{\frac{1}{\frac{P_n^a}{P_o^a}} + \frac{1}{\frac{P_n^b}{P_o^b}} + \frac{1}{\frac{P_n^c}{P_o^c}} + \frac{1}{\frac{P_n^d}{P_o^d}}} \cdot 100$$

$$PI_{(H)} = \frac{4}{\frac{P_o^a}{P_n^a} + \frac{P_o^b}{P_n^b} + \frac{P_o^c}{P_n^c} + \frac{P_o^d}{P_n^d}} \cdot 100$$

بالتعويض بقيم المناسيب نجد :

$$\begin{aligned}
&= \frac{4}{\frac{6}{8} + \frac{9}{11} + \frac{7}{6} + \frac{21}{23}} \cdot 100 \\
&= \frac{4}{\frac{6}{759} + \frac{9}{552} + \frac{7}{1012} + \frac{21}{264}} \cdot 100 \\
&= \frac{4}{\frac{4554 + 4968 + 7084 + 5544}{6072}} \cdot 100 \\
&= \frac{4}{\frac{22150}{6072}} \cdot 100 = \frac{24288}{22150} \cdot 100 = 109.65\%
\end{aligned}$$

وهذا يفسر على أن أسعار هذه السلع الأربع قد زادت في عام 2001 بنسبة 9.65% عما كانت عليه في سنة الأساس 1998.

وبنتيجة المقارنة بين المتوسطات الثلاثة المستخدمة في قياس المناسيب البسيطة لأسعار السلع الأربع بين عامي 1998 و 2001 .

إن اختيار أي من هذه المتوسطات يعتمد على المقارنة فيما بينها من حيث مزايا وعيوب كل منها من جهة، وما نتوخاه من سهولة في العمليات الحسابية من جهة أخرى ومن خلال المقارنة بين نتائج المتوسطات الثلاثة المستخدمة في قياس المناسيب البسيطة لأسعار السلع الأربع بين عام 1998 و 2001 نلاحظ أن :

$$PI_{(H)} > PI_{(\bar{x})} > PI_{(O)}$$

9-3-1-2 المناسيب المرجحة :

لاحظنا من خلال دراسة المتوسطات البسيطة أنها لم تعطِ أية أهمية لمسألة ترجيح المناسيب، وتعتبر أن التغيرات في أسعار السلع الداخلة في تركيب الرقم القياسي متساوية من حيث الأهمية، وهذا لا يمثل الواقع تمثيلاً دقيقاً، لأن الأهمية النسبية للسلع في السوق تختلف فيما بينها بشكل كبير، أي أن كل سلعة تدخل في تركيب الرقم القياسي لها الأهمية النسبية المختلفة عن السلع الأخرى، لهذا من المفضل الاعتماد على

المتوسطات المرجحة في جميع الحالات التي يمكن معها تأمين بيانات إحصائية يمكن استخدامها كأوزان للتعبير عن الأهمية النسبية للسلع موضوع الرقم القياسي، وقد تكون هذه الأوزان كميات الاستهلاك، أو كميات الإنتاج، أو إجمالي نفقات الأسر على السلع المختلفة، حيث إن أهمية أي سلع تتحدد بالنسبة إلى السلع الأخرى الداخلة في تركيب الرقم القياسي، وإن ترجيح كل منسوب سعر بالوزن الذي يتناسب مع أهمية السلعة المحسوب لها المنسوب، يمكننا من تركيب الأرقام القياسية للمناسيب المرجحة للأسعار، باستخدام المتوسطات الثلاثة السابقة: الوسط الحسابي، والوسط الهندسي، والوسط التوافقي .

1- الوسط الحسابي للمناسيب المرجحة للأسعار :

يعرف الوسط الحسابي المرحج للمناسيب على أنه حاصل قسمة مجموع جداء القيم في أوزانها على مجموع الأوزان، فإذا ما أخذنا بعين الاعتبار الكميات التي استهلكت من السلع الأربع الداخلة في تركيب الرقم القياسي عام 2001 في المثال (8-9)، فسوف نتوصل إلى حساب الرقم القياسي للمناسيب المرجحة للأسعار باستخدام الوسط الحسابي في سنة المقارنة 2001 .

منسوب سعر السلعة A مضروب بالكميات المستهلكة منها كأوزان للترجيح :

$$X_a = \frac{P_n^{(a)}}{P_0^{(a)}} \cdot q_n^{(a)} \cdot 100 = \frac{8}{6} \cdot (50) \cdot 100 = 6666.66$$

منسوب سعر السلعة B ضرب الكميات المستهلكة منها كأوزان للترجيح :

$$X_b = \frac{P_n^{(b)}}{P_0^{(b)}} \cdot q_n^{(b)} \cdot 100 = \frac{11}{9} \cdot (70) \cdot 100 = 8555.55$$

منسوب سعر السلعة C ضرب الكميات المستهلكة منها كأوزان للترجيح :

$$X_c = \frac{P_n^{(c)}}{P_0^{(c)}} \cdot q_n^{(c)} \cdot 100 = \frac{6}{7} \cdot (30) \cdot 100 = 2571.3$$

$(V_o = P_o q_o)$ - هي عبارة عن القيمة النقدية لما تم إنفاقه على مجموعة السلع الداخلة في تركيب الرقم القياسي في فترة الأساس مقومة بأسعار سنة الأساس.

$(V_n = P_n q_n)$ - هي عبارة عن القيمة النقدية لما تم إنفاقه على مجموعة السلع الداخلة في تركيب الرقم القياسي في فترة المقارنة، مقومة بأسعار سنة المقارنة.
 $(P_n q_o)$ - يقيس تكاليف السلع المشتراة في نقطة الأساس، مقومة بأسعار نقطة المقارنة.

$(P_o q_n)$ - يقيس تكاليف السلع المشتراة في نقطة المقارنة، مقومة بأسعار نقطة الأساس.

ويمكن بذلك أن يأخذ الوسط الحسابي المرجح للمناسيب أحد الصيغ الآتية :

1.1 - الرقم القياسي للمناسيب المرجحة بالقيمة في سنة الأساس (V_o) باستخدام الوسط الحسابي :

$$PI_{(X)} = \frac{\sum \left(\frac{P_n}{P_o} \cdot P_o q_o \right)}{\sum (P_o q_o)} \cdot 100$$

2.1 - الرقم القياسي للمناسيب المرجحة بالقيمة في سنة المقارنة (V_n) باستخدام الوسط الحسابي :

$$PI_{(X)} = \frac{\sum \left(\frac{P_n}{P_o} \cdot P_n q_n \right)}{\sum (P_n q_n)} \cdot 100$$

3.1 - الرقم القياسي للمناسيب المرجحة بتكلفة السلع المشتراة في نقطة الأساس، مقومة بأسعار سنة المقارنة $(P_n q_o)$:

$$PI_{(X)} = \frac{\sum \left(\frac{P_n}{P_o} \cdot P_n q_o \right)}{\sum (P_n q_o)} \cdot 100$$

4.1 - الرقم القياسي للمناسيب المرجحة بتكلفة السلع المشتراة في نقطة المقارنة مقومة بأسعار سنة الأساس $(P_0 q_n)$:

$$PI_{(N)} = \frac{\sum \left(\frac{P_n}{P_0} \cdot P_0 q_n \right)}{\sum (P_0 q_n)} \cdot 100$$

وسنبين طريقة حساب الوسط الحسابي للمناسيب المرجحة بالمجموعات الأربع السابقة من الأوزان من خلال المثال الآتي :

مثال (9-9) :

يبين الجدول الآتي أسعار وكميات مجموعة من المواد الغذائية المستهلكة في عامي 2000 و 2002 ، والمطلوب تركيب الرقم القياسي للمناسيب المرجحة بالمجموعات الأربعة السابقة من الأوزان باستخدام الوسط الحسابي باعتبار عام 2000 سنة أساس.

جدول رقم (9-4)

المواد الغذائية	الأسعار (بالليرات السورية)		الكميات (بالكيلو غرام)	
	عام 2000	عام 2002	عام 2000	عام 2002
A	5	6	30	34
B	40	50	12	15
C	12	16	22	26
D	9	11	14	16
F	80	120	5	10

الحل :

الجدول المساعد رقم (9-5)

المواد الغذائية	P_0	P_n	q_0	q_n	$V_0 = P_0 q_0$	$V_n = P_n q_n$	$P_n q_0$	$P_0 q_n$
A	5	6	30	34	150	204	180	170
B	40	50	12	15	480	750	600	600
C	12	16	22	26	264	416	352	312
D	9	11	14	16	126	176	154	144
E	80	120	5	10	400	1200	600	800
Σ				101	1420	2746	1886	2026

$X_{(o,n)} = \frac{P_n}{P_0} \cdot 100$	$\left(\frac{P_n}{P_0} \cdot 100 \right) \cdot V_0$	$\left(\frac{P_n}{P_0} \cdot 100 \right) \cdot V_n$	$\left(\frac{P_n}{P_0} \cdot 100 \right) \cdot P_n q_0$	$\left(\frac{P_n}{P_0} \cdot 100 \right) \cdot P_0 q_n$
120	18000	24480	21600	20400
125	60000	93750	75000	75000
133.33	32799.13	55465.28	46932.16	41598.96
122.22	15399.72	21510.72	18821.88	17599.68
150	60000	180000	90000	120000
Σ	186198.85	375206	252354	274598.64

$$(1) \quad PI_{(\bar{x})} = \frac{186198.85}{1420} = 131.125 \%$$

$$(2) \quad PI_{(\bar{x})} = \frac{375206}{2746} = 136.64 \%$$

$$(3) \quad PI_{(\bar{x})} = \frac{252354}{1886} = 133.80 \%$$

$$(4) \quad PI_{(\bar{x})} = \frac{274598.64}{2026} = 135.37 \%$$

نلاحظ أننا قد حصلنا على نتائج مختلفة للرقم القياسي للمناسيب المرجحة باستخدام الوسط الحسابي، إلا أن جميعها يدل على الاتجاه نفسه للتغير في أسعار المواد الغذائية موضوع الدراسة، وهذا الاختلاف في النتائج إنما يرجع إلى اختلاف أسلوب تركيب الرقم القياسي، أو إلى اختلاف الأوزان المستخدمة في الترجيح.

2- الوسط الهندسي للمناسيب المرجحة للأسعار :

نعلم أن الوسط الحسابي يقوم على أساس جمع القيم أما الوسط الهندسي يقوم على أساس ضرب القيم ولهذا فالمنسوب نرفعه للقوة أو الوزن المرجح به وقاعدة تطبيق الوسط الهندسي المرجح تقوم على أساس أخذ الجذر النوني (حيث n = مجموع الأوزان) لحاصل ضرب المناسيب مرفوعة إلى أوزانها. وبتطبيق ذلك على المثال (9-9) للسلع الأربع التي يتضمنها الرقم القياسي نجد أن الرقم القياسي للمناسيب المرجحة بكميات سنة المقارنة 2001 باستخدام الوسط الهندسي يتم حسابه كما يأتي :

$$PI_{(G)} = \sqrt[n]{\left(\frac{P_n^{(a)}}{P_0^{(a)}} \cdot 100\right)^{q_n^{(a)}} \cdot \left(\frac{P_n^{(b)}}{P_0^{(b)}} \cdot 100\right)^{q_n^{(b)}} \cdot \left(\frac{P_n^{(c)}}{P_0^{(c)}} \cdot 100\right)^{q_n^{(c)}} \cdot \left(\frac{P_n^{(D)}}{P_0^{(D)}} \cdot 100\right)^{q_n^{(D)}}}$$

$$PI_{(G)} = \sqrt[210]{(133.33)^{(50)} \cdot (122.22)^{(70)} \cdot (85.71)^{(30)} \cdot (109.52)^{(60)}}$$

وباستخدام اللوغاريتمات نجد :

$$\log PI_{(G)} = \frac{1}{\sum q_n} \sum q_n \log \left(\frac{P_n}{P_0} \cdot 100 \right)$$

$$\log PI_{(G)} = \frac{1}{210} (50 \cdot \log 133.33 + 70 \cdot \log 122.22 + 30 \cdot \log 85.71 + 60 \cdot \log 109.52)$$

$$= \frac{1}{210} (106.24 + 146.09 + 58 + 122.37)$$

$$= \frac{1}{210} (432.70) = 2.06$$

$$\Rightarrow PI_{(G)} = 10 \times 2.06 = 114.81$$

وهذا يدل على أن أسعار السلع الأربع المرححة بكميات سنة المقارنة 2002 قد زادت بالمتوسط الهندسي بنسبة 14.81% عما كانت عليه في سنة الأساس 1998، أما إذا تم ترجيح مناسيب الأسعار بإحدى المجموعات الأربع من الأوزان :

$$(P_0q_0) , (P_nq_n) , (P_nq_0) , (P_0q_n)$$

فسوف نجد أن الرقم القياسي للمناسيب المرححة للأسعار باستخدام الوسط الهندسي يعطى بالعلاقات الآتية حسب الأوزان المستخدمة في الترجيح :

$$(1) \log PI_{(G)} = \frac{1}{\sum P_0q_0} \cdot \sum P_0q_0 \log \left(\frac{P_n}{P_0} \cdot 100 \right)$$

$$(2) \log PI_{(G)} = \frac{1}{\sum P_nq_n} \cdot \sum P_nq_n \log \left(\frac{P_n}{P_0} \cdot 100 \right)$$

$$(3) \log PI_{(G)} = \frac{1}{\sum P_nq_0} \cdot \sum P_nq_0 \log \left(\frac{P_n}{P_0} \cdot 100 \right)$$

$$(4) \log PI_{(G)} = \frac{1}{\sum P_0q_n} \cdot \sum P_0q_n \log \left(\frac{P_n}{P_0} \cdot 100 \right)$$

وبالعودة إلى المثال (9-9) المتعلق بمجموعة المواد الغذائية المستهلكة في عامي

2000 و 2002 نجد أن قيمة الصيغ السابقة تعطى كما يأتي :

الجدول رقم (9-6)

المواد الغذائية	$P_o q_o \log \left(\frac{P_n}{P_o} \cdot 100 \right)$	$P_n q_n \log \left(\frac{P_n}{P_o} \cdot 100 \right)$	$P_n q_o \log \left(\frac{P_n}{P_o} \cdot 100 \right)$	$P_o q_n \log \left(\frac{P_n}{P_o} \cdot 100 \right)$
A	$150 \log 120 = 311.877$	424.152	374.252	352.460
B	$480 \log 125 = 1006.516$	1572.682	1258.146	1258.146
C	$264 \log 133.33 = 560.980$	883.970	747.974	662.977
D	$126 \log 122.22 = 262.979$	367.337	321.419	300.548
E	$400 \log 150 = 870.436$	2611.309	1305.654	1740.873
Σ	3012.788	5859.450	4007.445	4315

$$(1) \log PI_{(a)} = \frac{1}{1420} (150 \cdot \log 120 + 480 \cdot \log 125 + 264 \cdot \log 133.33 + 400 \cdot \log 150)$$

$$= \frac{1}{1420} (311.877 + 1006.516 + 560.980 + 870.436)$$

$$= \frac{1}{1420} (3012.788) = 2.12$$

$$\Rightarrow PI_{(a)} = \frac{(2.12)}{10} = 131.82$$

$$(2) \log PI_{(a)} = \frac{1}{2746} (5859.450) = 2.13$$

$$\Rightarrow PI_{(a)} = \frac{(2.13)}{10} = 134.89$$

$$(3) \log PI_{(G)} = \frac{1}{1886} (4007.445) = 2.12$$

$$(2.12) \\ \Rightarrow PI_{(G)} = 10^{2.12} = 131.82$$

$$(4) \log PI_{(G)} = \frac{1}{2026} (4315) = 2.13$$

$$(2.13) \\ \Rightarrow PI_{(G)} = 10^{2.13} = 134.89$$

نلاحظ أن النتائج التي توصلنا إليها باستخدام الوسط الهندسي في تركيب الرقم القياسي للمناسيب المرجحة لأسعار المواد الغذائية متباعدة بعض الشيء، غير أنها تدل على الاتجاه ذاته للتغير في أسعار هذه المواد بين فترتي المقارنة 2002 والأساس 2000 حيث إن جميعها يعبر عن زيادة في الأسعار تتراوح بين 31.82% و 34.89% ويعود هذا التباين في النتائج إلى اختلاف الأوزان المستخدمة في التجميع، أو إلى الأسلوب المستخدم في تركيب الرقم القياسي .

3- الوسط التوافقي للمناسيب المرجحة للأسعار :

بالإضافة إلى استخدام الوسطين السابقين الحسابي والهندسي في تركيب الرقم القياسي للمناسيب المرجحة للأسعار، يمكننا أيضاً إيجاد الوسط التوافقي مرجحاً إما بالكميات في سنة المقارنة، أو بأي صورة من المجموعات الأربع السابقة، فإذا ما تم ترجيح مناسيب الأسعار بالكميات في سنة المقارنة (q_n) فيمكننا عندها تركيب الرقم القياسي للمناسيب المرجحة بالكميات في سنة المقارنة لأسعار السلع الأربع موضوع الدراسة في المثال (9-9) كما يأتي :

نعلم أن صيغة الوسط التوافقي :

$$H = \frac{n}{\sum \frac{1}{x}}$$

وفي حالة الترجيح تصبح :

$$H = \frac{\sum q_n}{\sum \frac{1}{\frac{P_n}{P_o} \cdot 100} \cdot q_n} = \frac{\sum q_n}{\sum \frac{P_o}{P_n} \cdot q_n} \cdot 100$$

وبالتالي فإن الرقم القياسي للمناسيب المرحح لأسعار السلع الأربع باستخدام الوسط التوافقي يكون على النحو الآتي :

$$PI_{(H)} = \frac{210}{\frac{1}{133.33} \cdot 50 + \frac{1}{122.22} \cdot 70 + \frac{1}{85.71} \cdot 30 + \frac{1}{109.52} \cdot 60}$$

$$= \frac{210}{0.375 + 0.572 + 0.350 + 0.547} = \frac{210}{1.844} = 113.88$$

أما إذا تم تبني أية مجموعة من مجموعات الأوزان الأربعة الخاصة بالترجيح، فإن صيغة الرقم القياسي التوافقي للمناسيب المرححة لأسعار المواد الغذائية في المثال (9-9) تؤول إلى الشكل :

$$(1) \quad PI_{(H)} = \frac{\sum P_o q_o}{\sum \frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_o q_o}$$

$$(2) \quad PI_{(H)} = \frac{\sum P_n q_n}{\sum \frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_n q_n}$$

$$(3) \quad PI_{(H)} = \frac{\sum P_n q_o}{\sum \frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_n q_o}$$

$$(4) \quad PI_{(H)} = \frac{\sum P_o q_n}{\sum \frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_o q_n}$$

ولحساب نتائج هذه الصيغ حسب معطيات المثال (9-9) ننشئ الجدول المساعد رقم (9-7) .

الجدول رقم (9-7)

المواد الغذائية	$\frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_o q_o$	$\frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_n q_n$	$\frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_n q_o$	$\frac{1}{\left(\frac{P_n}{P_o} \cdot 100\right)} \cdot P_o q_n$
A	$\frac{1}{120} \cdot 150 = 1.24$	1.70	1.50	1.41
B	$\frac{1}{125} \cdot 480 = 3.84$	6	4.8	4.8
C	$\frac{1}{133.33} \cdot 264 = 1.98$	3.12	2.64	2.34
D	$\frac{1}{122.22} \cdot 126 = 1.03$	1.44	1.26	1.17
E	$\frac{1}{150} \cdot 400 = 2.67$	8	4	5.33
Σ	10.76	20.26	14.2	15.05

$$(1) \quad PI_{(H)} =$$

$$\frac{1420}{\left(\frac{1}{120} \cdot 150\right) + \left(\frac{1}{125} \cdot 480\right) + \left(\frac{1}{133.33} \cdot 264\right) + \left(\frac{1}{122.22} \cdot 126\right) + \left(\frac{1}{150} \cdot 400\right)} =$$

$$= \frac{1420}{1.24 + 3.84 + 1.98 + 1.03 + 2.67} = \frac{1420}{10.76} = 131.97$$

$$(2) \quad PI_{(H)} = \frac{2746}{20.26} = 135.53$$

$$(3) \quad PI_{(H)} = \frac{1886}{14.2} = 132.81$$

$$(4) \quad PI_{(H)} = \frac{2026}{1505} = 134.61$$

وتفسير نتائج هذه المتوسطات جميعاً هو نفسه في حالة الوسط الحسابي والوسط الهندسي والوسط التوافقي، مع ملاحظة أن اختلاف النتائج يعود إلى اختلاف المتوسط واختلاف طرائق الترجيح .

9-3-1-3 خصائص المناسيب

1- خصائص مناسيب الأسعار :

إذا كانت $P_{(a)}$, $P_{(b)}$, $P_{(c)}$ تعبر عن أسعار الفترات a , b , c على الترتيب فإن الخصائص الآتية تتحقق لمناسيب الأسعار المرتبطة بها :

$$A- \text{خاصية التطابق} \quad P_{a/a} = 1$$

وهذه ببساطة تعني أن منسوب السعر لفترة معينة بالنسبة للفترة نفسها تساوي 1 أو 100% .

$$B- \text{خاصية الانعكاس في الزمن} \quad P_{a/b} \cdot P_{b/a} = 1$$

أو

$$P_{a/b} = \frac{1}{P_{b/a}}$$

وهذه تعني أنه إذا أحصلنا فترتين كل واحدة محل الأخرى، فإن مناسيب الأسعار المقابلة تكون كل منها معكوس الأخرى .

جـ - خاصية الدورية أو الدائرية $P_{a/b} \cdot P_{b/c} \cdot P_{c/a} = 1$

وهكذا $P_{a/b} \cdot P_{b/c} \cdot P_{c/D} \cdot P_{D/a} = 1$

د - خاصية الدورية أو الدائرية المعدلة :

$$P_{a/b} \cdot P_{b/c} = P_{a/c}$$

وهكذا $P_{a/b} \cdot P_{b/c} \cdot P_{c/D} = P_{a/D}$

2- مناسب الكميات أو الحجم وخصائصها :

بدلاً من مقارنة أسعار السلعة، قد نلتم بمقارنة كميات أو أحجام السلعة، مثل كمية أو حجم الإنتاج، الاستهلاك، التصدير وغيرها، في مثل هذه الحالات نتكلم عن مناسب الكمية أو مناسب الحجم للتسهيل، كما في حالة الأسعار، نفترض أن الكميات ثابتة في أي فترة، وإذا لم يكن هذا صحيحاً، فإنه يمكن استخدام متوسط ملائم لجعل هذا الفرض ممكناً.

فإذا كانت q_0 تعبر عن كمية أو حجم السلعة المنتجة، المستهلكة، المصدرة، وغير ذلك خلال فترة الأساس بينما q_n تعبر عن كمية الإنتاج، الاستهلاك في فترة المقارنة فإن منسوب الكمية أو الحجم =

$$q_{R(0,n)} = X_{(0,n)} = \frac{q_n}{q_0} \cdot 100$$

وعندما كان تركيب الرقم القياسي بطريقة المناصب البسيطة أو المرجحة يهدف إلى قياس التغير في كميات السلع، أو المجموعة السلعية موضوع الدراسة، فإنه يمكننا استخدام المتوسطات التي درسناها سابقاً مع فارق جوهري هو أن الترجيح هنا يكون بالأسعار، أي أن الأوزان هي عبارة عن الأسعار، ولبيان كيفية استخدام هذه المتوسطات في تركيب الرقم القياسي البسيط أو المرجح للكميات نعرض المثال الآتي:

مثال (9-10) :

ليكن لدينا كميات وأسعار ست سلع في فترتي المقارنة عام 2002 والأساس عام 1999 والجدول الآتي يوضح البيانات الضرورية. والمطلوب :

- 1- إيجاد مناسيب الكميات للسلع الست .
- 2- إيجاد الوسط الحسابي البسيط لمناسيب الكميات .
- 3- إيجاد الوسط الهندسي البسيط لمناسيب الكميات .
- 4- إيجاد الوسط الحسابي المرجح لمناسيب الكميات .
- 5- إيجاد الوسط الهندسي المرجح لمناسيب الكميات .

الجدول رقم (9-8)

السلع	كميات عام 1999 (بالأطنان)	كميات عام 2002 (بالأطنان)	أسعار عام 2002 (بالوحدات النقدية)
A	40 ton	70	8
B	90 ton	120	7
C	130 ton	180	9
D	240 ton	270	5
E	70 ton	50	6
F	120 ton	150	4

1- مناسيب الكميات للسلع الست :

منسوب كمية السلعة A :

$$\frac{q_n^{(a)}}{q_o^{(a)}} \cdot 100 = \frac{70}{40} \cdot 100 = 175$$

منسوب كمية السلعة B :

$$\frac{q_n^{(b)}}{q_o^{(b)}} \cdot 100 = \frac{120}{90} \cdot 100 = 133.33$$

منسوب كمية السلعة C :

$$\frac{q_n^{(c)}}{q_o^{(c)}} \cdot 100 = \frac{180}{130} \cdot 100 = 138.46$$

منسوب كمية السلعة D :

$$\frac{q_n^{(D)}}{q_o^{(D)}} \cdot 100 = \frac{270}{240} \cdot 100 = 112.5$$

منسوب كمية السلعة E :

$$\frac{q_n^{(B)}}{q_o^{(B)}} \cdot 100 = \frac{50}{70} \cdot 100 = 71.42$$

منسوب كمية السلعة F :

$$\frac{q_n^{(F)}}{q_o^{(F)}} \cdot 100 = \frac{150}{120} \cdot 100 = 125$$

2- الوسط الحسابي البسيط لمناسيب الكميات للسلع الست :

$$PI_{(x)} = \frac{\sum \left(\frac{q_n}{q_o} \cdot 100 \right)}{n}$$

$$= \frac{175 + 133.33 + 138.46 + 112.5 + 71.42 + 125}{6}$$

$$= \frac{755.71}{6} = 125.95$$

ويمثل ذلك على أن الكميات المنتجة من هذه السلع الست قد زادت في عام 2002 بنسبة 25.95% بالمقارنة مع سنة الأساس 1999 .

3- الوسط الهندسي البسيط لمناسيب الكميات :

$$PI_{(G)} = \sqrt[n]{\frac{q_n^{(a)}}{q_o^{(a)}} \cdot \frac{q_n^{(b)}}{q_o^{(b)}} \cdot \frac{q_n^{(c)}}{q_o^{(c)}} \cdot \frac{q_n^{(D)}}{q_o^{(D)}} \cdot \frac{q_n^{(B)}}{q_o^{(B)}} \cdot \frac{q_n^{(F)}}{q_o^{(F)}}} \times 100$$

$$= \sqrt[6]{\frac{70}{40} \cdot \frac{120}{90} \cdot \frac{180}{130} \cdot \frac{270}{240} \cdot \frac{50}{70} \cdot \frac{150}{120} \cdot 100}$$

$$= \sqrt[6]{(175) \cdot (133.33) \cdot (138.46) \cdot (112.5) \cdot (71.42) \cdot (125)}$$

باستخدام اللوغاريتمات وفق الصيغة الآتية نحصل على قيمة الرقم القياسي

الهندسي البسيط للكميات :

$$\text{Log PI}_{(G)} = \frac{1}{n} \sum \log \left(\frac{q_n}{q_0} \right) \cdot 100$$

$$= \frac{1}{6} (\log 175 + \log 133.33 + \log 138.46 + \log 112.5 + \\ + \log 71.42 + \log 125)$$

$$= \frac{1}{6} (2.243 + 2.125 + 2.141 + 2.051 + 1.853 + 2.096)$$

$$= \frac{1}{6} (12.509) = 2.084$$

$$\Rightarrow \text{PI}_{(G)} = \frac{(2.084)}{10} = 121.33$$

4- الوسط الحسابي المرجح لنسب الكميات بأسعار 2002 :

$$\text{PI}_{(G)} = \frac{\sum (q_n/q_0) \cdot p_n}{\sum p_n} \cdot 100$$

$$= \frac{175(8) + 133.33(7) + 138.46(9) + 112.5(5) + 71.42(6) + 125(4)}{8 + 7 + 9 + 5 + 6 + 4}$$

$$= \frac{1400 + 933.31 + 1246.14 + 562.5 + 428.52 + 500}{39}$$

$$= \frac{5070.47}{39} = 130$$

يدل الرقم القياسي لمناسيب الكميات المرجحة بأسعار سنة المقارنة 2002 على أن الكميات المنتجة من هذه السلع الست قد زادت بنسبة 30%، بالمقارنة مع سنة الأساس 1999 .

5 - الوسط الهندسي المرجح لمناسيب الكميات للسلع الست بأسعار 2002 :

$$\text{Log PI}_{(0)} = \frac{1}{\sum p_n} \sum p_n \log \left(\frac{q_n}{q_0} \cdot 100 \right)$$

$$\begin{aligned} \text{Log PI}_{(0)} &= \frac{1}{39} (8 \cdot \log 175 + 7 \cdot \log 133.33 + 9 \cdot \log 138.46 \\ &\quad + 5 \cdot \log 112.5 + 6 \cdot \log 71.42 + 4 \cdot \log 125) \\ &= \frac{1}{39} (17.94 + 14.87 + 19.27 + 10.25 + 11.12 + 8.38) \end{aligned}$$

$$\text{Log PI}_{(0)} = \frac{1}{39} (81.83) = 2.09$$

$$\Rightarrow \text{PI}_{(0)} = \frac{(2.09)}{10} = 123$$

نلاحظ على الرغم من اختلاف نتائج المتوسطات، إلا أنها جميعاً تشير إلى الاتجاه نفسه بالزيادة في الكميات المنتجة في سنة المقارنة 2002 بالنسبة لسنة الأساس 1999. وتجب الإشارة هنا إلى أنه بالإمكان استخدام الوسط التوافقي في تركيب الأرقام القياسية لمناسيب الكميات البسيطة كانت أو المرجحة بالطريقة ذاتها التي تم استخدامها في تركيب الأرقام القياسية لمناسيب الأسعار البسيطة والمرجحة، آخذين بالاعتبار

منسوب الكمية $\frac{q_n}{q_0} \cdot 100$ في حين منسوب السعر $\frac{p_n}{p_0}$ كما أنه يمكن تطبيق المتوسطات الثلاثة في تركيب رقم قياسي مرجح بإحدى المجموعات الأربع لأوزان : $(p_0 q_n)$ ، $(p_n q_0)$ ، $(p_n q_n)$.

فإذا كان الترجيح بالقيمة في سنة الأساس $p_0 q_0$ ، فإن الرقم القياسي المناسب الكميات المرجحة بالقيمة في سنة الأساس بطريقة الوسط الحسابي تكون :

$$PI_{(q)} = \frac{\sum \left(\frac{q_n}{q_0} \cdot p_0 q_0 \right)}{\sum (p_0 q_0)} \cdot 100$$

أما إذا كان الترجيح بالقيمة في سنة المقارنة فإن :

$$PI_{(q)} = \frac{\sum \left(\frac{q_n}{q_0} \cdot p_n q_n \right)}{\sum (p_n q_n)} \cdot 100$$

وهكذا بالنسبة للأوزان الأخرى كما سبق دراستها في مناسيب الأسعار وفي حالة استخدام الوسط الهندسي، فإن الرقم القياسي المناسب الكميات المرجحة بالقيمة في سنة الأساس يكون على الشكل الآتي :

$$\log PI_{(q)} = \frac{1}{\sum p_0 q_0} \sum p_0 q_0 \log \left(\frac{q_n}{q_0} \cdot 100 \right)$$

وإذا كان الترجيح بالقيمة في سنة المقارنة، فإن الرقم القياسي الهندسي المناسب الكميات المرجحة بالقيمة في سنة المقارنة يأخذ الصيغة الآتية :

$$\log PI_{(q)} = \frac{1}{\sum p_n q_n} \sum p_n q_n \log \left(\frac{q_n}{q_0} \cdot 100 \right)$$

وهكذا بالنسبة للأوزان الأخرى بالإضافة إلى أنه يمكن استخدام الوسط التوافقي أيضاً في تركيب الرقم القياسي البسيط أو المرجح لمناسيب الكميات كما سبق دراسته

في مناسيب الأسعار، مع ملاحظة أن جميع الملاحظات والخصائص التي تتعلق بمناسيب الأسعار تعتبر خصائصاً لمناسيب الكميات .

3- مناسيب القيمة وخصائصها :

إذا كان p هو سعر السلعة خلال فترة ما و q هي الكمية أو الحجم المنتج، المباع وغير ذلك خلال الفترة نفسها. إذن pq تسمى القيمة الإجمالية، فإذا بيعت 1000 وحدة بسعر 30 قرشاً سورياً لكل وحدة فإن القيمة الإجمالية هي :

$$\text{ليرة سورية } 300 = (1000) \cdot (0.30)$$

إذا كانت P_0 تعبر عن السعر و q_0 عن الكمية لسلعة خلال فترة الأساس بينما P_n تعبر عن السعر المقابل و q_n الكمية المقابلة خلال الفترة المعطاة، كذلك فإن القيمة الإجمالية خلال هذه الفترات هي V_0 لفترة الأساس و V_n للفترة المعطاة، وبالتالي فإن منسوب القيمة :

$$V_{R(0,n)} = X_{(0,n)} = \frac{V_n}{V_0} \cdot 100 = \frac{P_n q_n}{P_0 q_0} \cdot 100$$

وعلى وجه الخصوص إذا كانت $P_{n/b}$ تعبر عن منسوب السعر و $q_{n/b}$ عن منسوب الكمية فإن $V_{n/b}$ تعبر عن منسوب القيمة للفترة (b) بالمقارنة بالفترة (a) كما أن :

$$V_{n/b} = P_{n/b} \cdot q_{n/b}$$

وتجب الإشارة إلى أن جميع الخصائص والملاحظات الخاصة بمناسيب الأسعار والكميات يمكن أن تكون خصائص هامة لمناسيب القيمة، وأن جميع الطرائق والأساليب التي استخدمناها في دراسة مناسيب الأسعار والكميات يمكن تطبيقها على مناسيب القيمة بهدف تركيب الرقم القياسي لمناسيب القيم البسيطة والمرجحة .

2-3-9 طريقة السلسلة في تركيب الأرقام القياسية :

تقوم فكرة تركيب الأرقام القياسية باستخدام طريقة السلسلة على أساس حساب أسعار أو كميات أو قيم كل سنة أو كل فترة زمنية كنسبة مئوية من أسعار،

أو كميات السنة السابقة، أو الفترة الزمنية السابقة لها، أي بالاعتماد على الأساس المتحرك لمناسيب الأسعار أو الكميات أو القيم، وذلك بسبب سلع حديثة أو اختفاء سلع قديمة وآثار ذلك على أذواق المستهلكين، وأنماط الاستهلاك بشكل عام، كما يؤثر على الأهمية النسبية للسلع المختلفة التي تدخل في تركيب الرقم القياسي، ولهذا نرى أنه من الضروري إجراء مراجعة مستمرة للسلع المنتجة، وللأسواق، لكي نكون على بينة ومعرفة دقيقة بالأوزان التي سوف تستخدم في عملية الترجيح من وقت إلى آخر، وهذا يتطلب تحريك فترة الأساس بالشكل الذي يماشى مع التغيرات الخاصة بالإنتاج والتوزيع والعادات الاستهلاكية، وما طريقة السلسلة في تركيب الرقم القياسي سوى الوعاء الذي يجب أن يحتوي على جميع هذه التبدلات والتغيرات، بحيث تصبح عملية مقارنة أكثر دقة وواقعية، وتكون الأرقام القياسية للمناسيب التي تعبر عن هذه الفترات الزمنية المتتالية مقياساً حقيقياً يحدد قيمة التغير وطبيعته خلال هذه الفترة الزمنية، مع نعلم أنه يمكننا الوصول بصيغة الأساس المتحرك التسلسلي إلى الأساس الثابت وخاصة عندما تكون المقارنة بين فترات ليست طويلة، حيث لا تتغير الكميات تغيراً كبيراً. ولإجراء هذه المقارنات نستخدم الرقم القياسي المتسلسل (Chained Index) وتكون صيغته على النحو الآتي :

$$X_{(0,t)} \text{ (المسلسل) } = X_{(0,1)} \cdot X_{(1,2)} \cdot X_{(2,3)} \cdot \dots \cdot X_{(t-1,t)}$$

حيث نعلم أن X — يدل على منسوب السعر .
فإذا كان لدينا :

$$X_{(0,1)} = 112 \%$$

$$X_{(1,2)} = 124 \%$$

$$X_{(2,3)} = 95 \%$$

فإن الرقم القياسي المتسلسل :

$$X_{(0,3)} = X_{(0,1)} \cdot X_{(1,2)} \cdot X_{(2,3)} \cdot 100 \%$$

$$= 1.12 \times 1.24 \cdot 0.95 \cdot 100 = 131.93 \%$$

بالرغم من أن هذا الرقم ليس له مدلول واضح، لأنه يعتمد **ن** تركيبة على مجموعات متغيرة من السلع، إلا أنه يبقى إلى حد بعيد بمثابة مؤشر **يعكس** تغيرات أوزان السلع خلال الفترة الزمنية المدروسة، حيث الكميات لا **تتغير** تغيراً كبيراً من نقطة ما إلى نقطة لاحقة مباشرة، بالإضافة إلى أنه لو حسبنا الرقم $X_{(0,3)}$ كعلاقة ثنائية سوف نجده مختلفاً عن نظيره $X_{(0,3)}$ المتسلسل .

مثال (9-11) :

إذا كانت أسعار سلعة ما خلال السنوات: 1998, 1999, 2000, 2001 هي 9, 13, 16, 19 وحدة نقدية على الترتيب، فإن الوصلات النسبية التي تعبر دوماً عن **مناسيب الأسعار لفترة زمنية معينة بالمقارنة مع فترة الأساس هي :**

$$P_{(1998/1999)} = \frac{13}{9} \cdot 100 = 144.44 \%$$

$$P_{(1999/2000)} = \frac{16}{13} \cdot 100 = 123 \%$$

$$P_{(2000/2001)} = \frac{19}{16} \cdot 100 = 118.75 \%$$

وبالتالي فإن الرقم القياسي المتسلسل لمناسيب الأسعار لسنة 2001 بالمقارنة مع سنة الأساس 1998 هو :

$$P_{(1998/2001)} = P_{(1998/1999)} \cdot P_{(1999/2000)} \cdot P_{(2000/2001)} =$$

$$= \frac{13}{9} \cdot \frac{16}{13} \cdot \frac{19}{16} \cdot 100 = 211 \%$$

ويطلق على هذا الرقم أحياناً المناسيب المسلسلة إلى أساس ثابت.

تستخدم طريقة السلسلة في تركيب الرقم القياسي بالاعتماد على سلسلة من المناسب أو وصلة المناسب حيث يجب أولاً تحديد مناسب الأسعار، أو مناسب الكميات، أو القيم في كل فترة زمنية بالنسبة للفترة السابقة كأساس، ثم تستخدم أحد المتوسطات الثلاثة الحسابي أو الهندسي أو التوافقي، فنحصل على أرقام قياسية متسلسلة حسابية أو هندسية.. الخ لقياس التغيرات في الأسعار أو الكميات أو القيمة .

مثال (9-12) :

بفرض أنه لدينا مناسب أسعار ثلاثة محاصيل في السنوات الثلاث الآتية 1999، 2000، 2001 بالنسبة لسنة 1998 كأساس كما في الجدول رقم (9-9) .

جدول رقم (9-9)

المحاصيل	2001	2000	1999
القمح	125	110	120
الشعير	115	112	107
العدس	108	104	93
متوسط المناسب (بأساس ثابت)	116	108.66	106.66

وبذلك فإن الأرقام القياسية التي تم تركيبها باستخدام المتوسط الحسابي في السنوات 1999-2001 بالنسبة إلى سنة 1998 كأساس ثابت هي على الترتيب 116، 108.66، 106.66 .

ولتركيب الأرقام القياسية وفق نظام السلسلة بأساس متحرك نحسب منسوب سعر السلعة في كل سنة بالنسبة إلى سابقتها مباشرة، كما في الجدول رقم (9-10) الذي يبين مناسب أسعار المحاصيل الثلاثة في السنوات 1999-2001 بأساس متحرك .

جدول رقم (9-10)

الخصائص	2001	2000	1999
القمح	113.63	91.66	120
الشعير	102.67	104.67	107
العدس	103.84	111.82	93
متوسط المناسيب (بأساس متحرك)	1.6.71	102.72	106.66

وعلى ذلك فالأرقام القياسية لمناسيب أسعار الخصائص الثلاثة في سنة 1999 إلى 2001 كل سنة بالنسبة إلى سابقتها هي بالمتوسط على الترتيب :

106.66 , 102.72 , 106.71

بالإضافة إلى ما سبق تستخدم طريقة السلسلة في تركيب الأرقام القياسية لمعرفة التغيرات التي تطرأ على مناسيب الأسعار، أو الكميات خلال فترة زمنية معينة تمثل عدداً من الأشهر. فإذا ما تم ضرب المناسيب بعدد من الأشهر أي نضرب نسبة الأسعار بين شهر وآخر بعضها ببعض، فنحصل على الرقم القياسي الموسمي الشهري، فإذا طلب حساب الرقم القياسي للأسعار في شهر حزيران بالنسبة لكانون ثاني، فيكون الرقم القياسي :

$$PI = \frac{\text{شباط}}{\text{كانون ثاني}} \cdot \frac{\text{آذار}}{\text{شباط}} \cdot \frac{\text{نيسان}}{\text{آذار}} \cdot \frac{\text{أيار}}{\text{نيسان}} \cdot \frac{\text{حزيران}}{\text{أيار}} \cdot 100$$

يستخدم هذا الرقم من أجل معرفة نسبة الزيادة بين شهر وآخر، وإجراء المقارنات بين الأسعار خلال أشهر السنة لقياس التغيرات التي طرأت عليها وبيان أسبابها وأساسها الاقتصادي والعوامل المسببة .

مثال (9-13) :

إذا علمت أن أسعار البطاطا في السوق من أيار ولغاية شهر آب قد تغيرت وفق الجدول الآتي :

جدول رقم (9-11)

البيان	أيار	حزيران	تموز	آب
سعر الكيلو غرام من البطاطا (بالليرات السورية)	30	25	20	15
الأرقام القياسية القاعدية أو ذات الأساس الثابت %	100.0	83.33	66.66	50.0
الأرقام القياسية المتسلسلة أو المتحركة %	-	83.33	80.0	75.0

نلاحظ أن أسعار السلع الإفرادية ذات الأساس الثابت والمتسلسل ترتبط فيما بينها من خلال علاقة مستوى السعر للفترة الأخيرة مع مستوى السعر للفترة الأولى أي:

$$PI = \frac{\text{السعر في شهر آب}}{\text{السعر في شهر أيار}} \cdot 100$$

$$= \frac{15}{30} \cdot 100 = 50 \%$$

حيث يتم حساب الرقم القياسي المتسلسل لأسعار البطاطا كما يأتي :

$$\frac{25}{30} \cdot \frac{20}{25} \cdot \frac{15}{20} \cdot 100 = 0.833 \times 0.80 \times 0.75 \times 100$$

$$= 50 \%$$

نلاحظ أن الرقم القياسي المتسلسل يعكس بعض التغيرات التنازلية البطيئة في أسعار البطاطا، بالمقارنة مع الرقم القياسي الثابت، ولكن أحياناً تظهر هناك ضرورة ماسة لدراسة ديناميكية السعر، ليس من خلال دراسة سعر بضاعة واحدة، كما هو

وارد في المثال، وإنما من خلال دراسة السعر وفق مجموعات من السلع، حيث إن الأرقام القياسية لأسعار المجموعات البضائية يتم حسابها وفق البضائع الموجودة في السوق. ومن أجل حساب التغير الحاصل في أسعار المجموعات البضائية المختلفة، يمكننا تطبيق متوسطات الأرقام القياسية، والذي يعتبر الوسط الحسابي من أهم هذه المتوسطات التي تستخدم كمؤشرات تعطي إمكانية قياس البضائع المختلفة .

9-3-3 الطريقة التجميعية في تركيب الأرقام القياسية Aggregative type

9-3-3-1 الطريقة التجميعية في تركيب الأرقام القياسية للأسعار

باستخدام هذه الطريقة سوف نستعرض عدداً من الطرائق التي تتبع في تركيب الأرقام القياسية للأسعار، والتي يمكن تصنيفها ضمن فئتين من حيث التركيب :

(9-3-3-1-1) الطريقة التجميعية البسيطة.

(9-3-3-1-2) الطريقة التجميعية المرجحة.

9-3-3-1-1 الطريقة التجميعية البسيطة (Simple Aggregate Method)

تعتمد هذه الطريقة من أبسط الطرائق المستخدمة في تركيب الرقم القياسي لأسعار مجموعة من السلع، وتتلخص في حساب النسبة بين مجموع الأسعار للسلع الداخلة في تركيب الرقم القياسي في فترتي المقارنة، والأساس، والرقم الذي يتم تركيبه وفق هذه الطريقة يطلق عليه الرقم القياسي التجميعي البسيط للأسعار (Simple Aggregate Price Index) ومعادلة حسابه كما يأتي :

$$PI_{(s)} = \frac{\sum P_n}{\sum P_0} \cdot 100$$

هذا إذا كان لدينا سلعة واحدة، أما إذا كان لدينا مجموعة سلع، فنرمز بالحرف $P_0^{(1)}$ للسلعة الأولى، $P_0^{(2)}$ للسلعة الثانية.... $P_0^{(n)}$ لأسعار سنة الأساس. وكما نرمز لأسعار سنة المقارنة بالحرف :

$P_0^{(1)}$ للسلعة الأولى ، $P_0^{(2)}$ للسلعة الثانية $P_0^{(n)}$

وتصبح معادلة الرقم القياسي التجميعي البسيط :

$$PI_{(s)} = \frac{P_n^{(1)} + P_n^{(2)} + P_n^{(3)} + \dots + P_n^{(n)}}{P_0^{(1)} + P_0^{(2)} + P_0^{(3)} + \dots + P_0^{(n)}} \cdot 100 = \frac{\sum P_n}{\sum P_0} \cdot 100$$

مثال (9-14) :

ليكن لدينا أسعار أربع سلع في عامي 2001 و 1998 والمطلوب حساب الرقم القياسي التجميعي البسيط لأسعار هذه السلع عام 2001 باستخدام أسعار 1998 كأساس .

جدول رقم (9-12)

السلع	أسعار عام 2001	أسعار عام 1998
A	7	6
B	5	4
C	9	8
D	3	2
Σ	24	20

الرقم القياسي التجميعي البسيط للسلع الأربع في عامي 2001 ، 1998 :

$$PI_{(s)} = \frac{\sum P_n}{\sum P_0} \cdot 100 = \frac{7 + 5 + 9 + 3}{6 + 4 + 8 + 2} \cdot 100$$

$$= \frac{24}{20} \cdot 100 = 120$$

وبدل ذلك على أن أسعار هذه السلع الأربع قد ارتفعت في عام 2001 بنسبة 20% عما كانت عليه الأسعار في عام 1996 .

على الرغم من اعتبار الصيغة السابقة من أبسط الصيغ التي يمكن استخدامها في تركيب الرقم القياسي، إلا أنها أكثر الصيغ بدائية ويطلب عليها أنها تساوي في الأهمية النسبية للسلع التي تدخل في تركيب الرقم القياسي في حين تكاد تلغى الحالات التي تتساوى فيها السلع في الأهمية. بالإضافة إلى استحالة استخدامها في تركيب الرقم القياسي للكميات إذا اختلفت وحدات قياس السلع الداخلة في تركيبه، ولذلك فمن النادر قياس التغير في الكمية بهذه الصيغة، والمثال الآتي يؤكد صحة ما ذكره .

مثال (9-15) :

نفرض أننا نرغب في تركيب أرقام قياسية لتغيرات الغذاء وأننا أخذنا من أجل ذلك خمسة أصناف من الأغذية هي الخبز ، البطاطا، الجبن، الشاي، في كل من عامي 2000 ، 1996 كما هو مبين في الجدول الآتي :

جدول رقم (9-13)

الأسعار (بالليرات) بعد تغير وحدة التسعير		الأسعار (بالليرات)		الوحدة التي يتم التسعير بها	السلعة
P_n	P_o	P_n	P_o		
8	7	8	7	1 كيلو غرام	خبز
55(11×5)	65(13×5)	11	13	1 كيلو غرام	بطاطا
26	24	26	24	250 غرام	الجبن
28 (4×7)	32 (4×8)	7	8	1/2 لتر	لبن
56 (28×2)	50 (25×2)	28	25	100 غرام	الشاي
173	178	80	77		

نجد أن الرقم القياسي التجميعي البسيط لأسعار الأصناف الخمسة من المواد

الغذائية:

$$PI_{(s)} = \frac{\sum P_n}{\sum P_0} \cdot 100 = \frac{80}{77} \cdot 100 = 103.89 \%$$

هذه النتيجة تعني أن هناك 3.89% زيادة في المستوى العام للأسعار في نقطة المقارنة 2001 بالنسبة لنقطة الأساس 1996 لتلك المجموعة من السلع الداخلة في تركيب الرقم القياسي، بالإضافة إلى أن هذا الرقم يقيس التغير في التكلفة المجمعة لشراء وحدة واحدة من مجموعة هذه السلع حسب الوحدات التي يتم على أساسها التسعير لهذه السلع، حيث يُحد في المثال 1 كيلو غرام بحز 250 غرام جبنة... الخ، أي أن تركيب هذا النوع من الأرقام يعتمد بشكل أساسي على وحدات التسعير، فإذا ما تغيرت وحدات التسعير، فسوف نلاحظ تغير في تكاليف هذه المجموعة السلعية، وبالتالي تغير في قيمة الرقم القياسي التجميعي البسيط، فمثلاً إذا أصبحت وحدة تسعير اللبن 2 لتر عوضاً عن 1/2 لتر و 200 غرام شاي عوضاً عن 100 غرام و 5 كيلو بطاطا عوضاً عن 1 كيلو غرام فسوف يُحد أن الرقم القياسي الجديد:

$$PI_{(s)} = \frac{173}{178} \cdot 100 = 97.19 \%$$

أي أن هناك انخفاضاً في مستوى الأسعار بحوالي 2.81% في حين كان هناك زيادة في مستوى الأسعار بنسبة 3.89%، ولتلافي عيوب هذا الرقم يُحد أنه من الضروري إعطاء السلع الداخلة في تركيب الرقم القياسي أوزاناً تتناسب مع أهميتها النسبية، وفي حالة تركيب الرقم القياسي أوزاناً تتناسب مع أهميتها النسبية، وفي حالة تركيب الرقم القياسي للأسعار تعتبر الكميات المنتجة من السلع موضوع الدراسة من أفضل الأوزان التي يمكن استخدامها في عملية الترجيح.

9-3-3-1 الطريقة التجميعية المرجحة (Weighted Aggregate Method)

من الواضح أن استخدام الطريقة التجميعية البسيطة في تركيب الأرقام القياسية للأسعار يعتمد على وحدات التسعير التي تحددها أساليب التسويق، وطبيعة التعامل في السوق في منطقة معينة، فإذا تغيرت هذه الوحدات تغير الرقم القياسي تغيراً ملحوظاً.

علاجاً لهذا نرى أنه من الضروري القيام بتركيب رقم قياسي تعطي فيه كل سلعة وزناً تبعاً لأهميتها، على أن يحدد هذا الوزن بأسلوب ما يعكس هذه الأهمية، هذا الرقم الذي يتمتع بهذه الخاصية يطلق عليه الرقم القياسي التجميعي المرجح (المثقل) للأسعار، حيث يقصد بالأوزان الترجيحية الكميات الخاصة بكل سلعة تدخل في تركيب الرقم القياسي، وهذه الكميات المنتجة من كل سلعة - كما أشرنا سابقاً - يرمز لها بالرمز q ، حيث q تختلف من سلعة إلى أخرى، والصيغة العامة للرقم القياسي التجميعي للأسعار المرجح بالكميات كأوزان هي :

$$PI_{(ng)} = \frac{\sum p_n q}{\sum p_0 q} \cdot 100$$

حيث : $p_n q$ - هي تكلفة مجموعة السلع في فترة المقارنة .

$p_0 q$ - هي تكلفة مجموعة السلع في فترة الأساس .

نلاحظ أن الرقم القياسي التجميعي البسيط هو حالة خاصة من هذا الأخير، إذا كانت قيمة $q = 1$ لكل السلع الداخلة في تركيب الرقم القياسي.

إن الأهمية النسبية للسلع التي تدخل في تركيب الرقم القياسي للأسعار تقوم على أساس مقارنة الكميات المختلفة بعضها البعض، فإذا بقيت هذه الكميات دون تغير من فترة إلى أخرى، فإن الترجيح بالكمية المشتراة يعطي النتيجة نفسها، غير أن الكميات تتغير دوماً من سلعة إلى أخرى ومن مدة زمنية إلى مدة زمنية أخرى، لذلك عند تركيب الرقم القياسي للأسعار وفق هذه الطريقة نجد أنه من الضروري استخدام الكميات المنتجة أو المشتراة في نقطة الأساس (q_0)، أو تلك المنتجة أو المشتراة في نقطة المقارنة (q_n) ، وذلك بهدف تركيب رقم قياسي مناسب يقيس هذه التغيرات سواء أكانت في الكميات أم في الأسعار. ومن الأرقام التي يمكن تركيبها لتحقيق هذا الغرض نعرض ما يأتي :

1- الرقم القياسي التجميعي للأسعار المرجح بكميات الأساس أو المقارنة (Weighted Aggregate Price Index Numbers)

عندما كانت الكميات المنتجة من كل سلعة تعتبر من أنسب الأوزان التي تستخدم في الترجيح، حيث $\sum Y$ هذه الكميات قد تكون منتجة في فترة الأساس (q_0) أو منتجة في فترة المقارنة (q_n) ، فإذا تم الترجيح باستخدام كميات فترة الأساس (q_0) ، فإنه يمكننا تركيب الرقم القياسي التجميعي للأسعار المرجح بكميات الأساس، والذي يعرف باسم رقم قياسي لاسبير للأسعار (Laspeyre's-Price Index) وصيغته هي :

$$PI_{(L)} = \frac{\sum P_n q_0}{\sum P_0 q_0} \cdot 100$$

أما إذا تم الترجيح باستخدام كميات فترة المقارنة (q_n) فإنه يمكننا تركيب الرقم القياسي التجميعي للأسعار المرجح بكميات المقارنة، والذي يعرف باسم رقم قياسي باش للأسعار (Paashe's-Price Index) وصيغته هي :

$$PI_{(P)} = \frac{\sum P_n q_n}{\sum P_0 q_n} \cdot 100$$

من الواضح أن رقم قياسي لاسبير للأسعار يفترض ثبات طلب المستهلكين على السلع التي تدخل في تركيب الرقم القياسي من فترة الأساس إلى فترة المقارنة دون النظر إلى التغيرات التي تطرأ على استهلاك كميات السلع الداخلة في تركيب الرقم، بسبب ارتفاع أو انخفاض أسعارها، أو إحلال سلع جديدة مكان السلع القديمة، وبمواصفات ذات جودة متميزة، في حين أن رقم قياسي باش يفترض أن الكميات المشتراة في فترة المقارنة هي ذات الكميات المشتراة في فترة الأساس، وهذا بطبيعة الحال يعتبر بعيداً عن الواقع. إن عملية المفاضلة بين هذين الرقمين تعتمد على خصائص كل منهما، والتي تعتبر سبباً منطقياً لتطبيق أحدهما دون الآخر من جهة، وعلى أسلوب تركيب الرقم القياسي من جهة أخرى، مع العلم أن طريقة لاسبير في الترجيح هي من أكثر الطرائق

استخداماً في العالم حيث يستخدمها أكثر من 65% من الدول رغم العيوب التي تشوبها، أما طريقة باش فيستخدمها 13% من دول العالم .

2- الرقم القياسي التجميعي للأسعار المرجح بكميات الأساس والمقارنة (Weighted Aggregate Price Index Numbers)

يهدف التلخيص من الجدل القائم حول نظام الترجيح بكميات الأساس، أو المقارنة، يمكننا استخدام كميات فترتي الأساس والمقارنة معاً، كأوزان للترجيح عند تركيب الرقم القياسي، وذلك بأخذ الوسط الحسابي أو الوسط الهندسي أو أي وسط آخر لكميات فترتي الأساس والمقارنة. فإذا اعتمدنا صيغة الوسط الحسابي في تركيب هذا النوع من الأرقام فسوف نحصل على الرقم القياسي التجميعي للأسعار المرجح بكميات الأساس والمقارنة معاً، والذي يعرف باسم رقم قياسي مارشال ادجورث للأسعار (Marshall and Edgeworth) نسبة إلى العالمين مارشال وادجورث والذي صيغته تعطى كما يأتي :

$$PI_{(ME)} = \frac{\sum P_n (q_0 + q_n)}{\sum P_0 (q_0 + q_n)} \cdot 100$$

$$= \frac{\sum (P_n q_0 + P_n q_n)}{\sum (P_0 q_0 + P_0 q_n)} \cdot 100$$

أما إذا استخدمنا صيغة الوسط الهندسي في تركيب هذا الرقم فإننا سوف نحصل على الصيغة الآتية للرقم القياسي التجميعي المرجح بمتوسط الكميات (مارشال ادجورث) :

$$PI_{(ME)} = \frac{\sum P_n \sqrt{q_0 q_n}}{\sum P_0 \sqrt{q_0 q_n}} \cdot 100$$

من الناحية العملية فإن رقم قياسي مارشال ادجورث بصيغته يعتبر قليل الاستخدام بالمقارنة مع رقمي لاسبير وباش، وذلك لاحتوائه على كميات فترتي الأساس والمقارنة معاً في قياس التغير الذي يطرأ على الأسعار، وهذا ما يقلل من أهمية الانتقادات الموجهة لرقمي لاسبير وباش للأسعار.

بالإضافة إلى رقم قياسي مارشال ادجورث بصيغته يعتبر قليل الاستخدام بالمقارنة معاً في قياس التغير الذي يطرأ على الأسعار، وهذا ما يقلل من أهمية الانتقادات الموجهة لرقمي لاسبير وباش للأسعار.

بالإضافة إلى رقم قياسي مارشال ادجورث للأسعار هناك أيضاً رقم قياسي فيشر (Fisher) الذي قام بتركيبه بالاعتماد على نظام الترجيح بكميات سنة الأساس، وكميات سنة المقارنة معاً، وذلك باستخدام صيغة الوسط الهندسي لكل من رقمي لاسبير وباش، حيث سمي هذا الرقم بالرقم القياسي الأمثل (Ideal Index Number) كونه يحقق اختباري الانعكاس في الزمن والانعكاس في المعامل كاختبارين أساسيين من اختبارات الأرقام القياسية التي سنتشاهدها لاحقاً، وهذا ما أعطى هذا الرقم بعض المزايا النظرية بالمقارنة مع الأرقام القياسية الأخرى، ويسمى هذا الرقم بالرقم القياسي التجميعي للأسعار المرجح بكميات فترتي الأساس والمقارنة معاً (فيشر)، أو الرقم القياسي التجميعي للأسعار باستخدام الوسط الهندسي لرقمي لاسبير وباش (فيشر) (Fisher Index Number) وصيغته هي :

$$PI(t) = \sqrt{I_{(L)} \times I_{(p)}} \cdot 100$$

$$PI(t) = \sqrt{\frac{\sum p_n q_o}{\sum p_o q_o} \times \frac{\sum p_n q_n}{\sum p_o q_n}} \cdot 100$$

مما تقدم نستنتج أن استخدام الطريقة التجميعية المرجحة في تركيب الأرقام القياسية للأسعار يعتمد على الأوزان المتغيرة كنظام للترجيح، فإذا ما تغيرت هذه الأوزان في الأساس، أو في المقارنة، فإن ذلك سوف يؤدي إلى تغير نتائج هذه الأرقام، وإن اعتماد الأوزان الثابتة كنظام للترجيح في فترتي الأساس والمقارنة يمكننا من تركيب الرقم القياسي التجميعي المرجح بأوزان ثابتة وصيغته هي :

$$PI_{(Fixed)} = \frac{\sum p_n q_{(FI)}}{\sum p_o q_{(FI)}} \cdot 100$$

مثال (9-16) :

ليكن لدينا الجدول الآتي الذي يمثل متوسط أسعار خمس سلع استهلاكية والكميات المستهلكة منها في عامي 1998 و 2002 .

جدول رقم (9-14)

السلع الاستهلاكية	متوسط الأسعار (بالليرات)		الكميات المستهلكة (بالآلاف الأطنان)	
	1998	2002	1998	2002
A	9	11	13	10
B	33	37	3	5
C	7	9	18	20
D	18	23	12	12
E	23	25	14	16

المطلوب : إيجاد الأرقام القياسية الآتية باعتبار 1998 سنة أساس :

- 1- الرقم القياسي التجميعي البسيط للأسعار .
- 2- الرقم القياسي التجميعي للأسعار المرجح بكميات الأساس (لاسيبر).
- 3- الرقم القياسي التجميعي للأسعار المرجح بكميات المقارنة (فيشر).
- 4- الرقم القياسي التجميعي للأسعار المرجح بمتوسط كميات الأساس، والمقارنة (مارشال ادجورث) .
- 5- الرقم القياسي التجميعي للأسعار المرجح بمتوسط الكميات في الأساس، والمقارنة (فيشر) .
- 6- الرقم القياسي للمناسيب البسيطة للأسعار، باستخدام الوسط الحسابي .
- 7- الرقم القياسي للمناسيب البسيطة للأسعار، باستخدام الوسط الهندسي.
- 8- الرقم القياسي للمناسيب المرجحة للأسعار، باستخدام المتوسط الحسابي للمناسيب المرجحة بالقيمة في سنة الأساس وسنة المقارنة .

الحل :

لتسهيل العمليات الحسابية نشكل الجدول الآتي :

جدول رقم (9-15)

السلع	p_o	p_n	q_o	q_n	$p_o q_o$ - v_o	$p_o q_n$	$p_n q_o$	$p_n q_n$ - v_n	$q_o + q_n$	$q_o \times q_n$
A	9	11	13	10	117	90	143	110	23	130
B	33	37	3	5	99	165	111	185	8	15
C	7	9	18	20	126	140	162	180	38	360
D	18	23	12	12	216	216	276	276	24	144
E	23	25	14	16	322	368	350	400	30	224
Σ	90	105	60	63	880	979	1042	1151		

تابع الجدول رقم (9-15) :

$p_n (q_o + q_n)$	$p_o (q_o + q_n)$	$\sqrt{q_o \cdot q_n}$	$p_n \sqrt{q_o \cdot q_n}$	$p_o \sqrt{q_o \cdot q_n}$
253	207	15.16	125.41	102.61
296	264	3.873	143.30	127.80
342	266	18.9736	170.73	132.81
552	432	12	276	216
750	690	14.9666	374.16	344.23
2193	1859		1089.6	923.45

تابع الجدول رقم (9-15) :

$X = \frac{P_n}{P_o} \cdot 100$	$\frac{P_n}{P_o} \cdot P_o q_o$	$\frac{P_n}{P_o} \cdot P_o q_n$	$\frac{P_n}{P_o} \cdot P_n q_o$	$\frac{P_n}{P_o} \cdot P_n q_n$
$\frac{11}{9} \cdot 100 = 122.22$	$\frac{11}{9} \cdot 117 = 143$	110	174.77	134.44
$\frac{37}{33} \cdot 100 = 112.12$	$\frac{37}{33} \cdot 99 = 111$	185	124.45	207.42
$\frac{9}{7} \cdot 100 = 128.57$	$\frac{9}{7} \cdot 126 = 162$	180	208.28	231.42
$\frac{23}{18} \cdot 100 = 127.77$	$\frac{23}{18} \cdot 216 = 276$	276	352.66	352.66
$\frac{25}{23} \cdot 100 = 108.96$	$\frac{25}{23} \cdot 322 = 350$	400	380.43	434.78
599.37	1042	1151	1240.59	1360.72

(1) الرقم القياسي التجميعي البسيط للأسعار :

$$PI_{(S)} = \frac{\sum P_n}{\sum P_o} \cdot 100$$

$$= \frac{105}{90} \cdot 100 = 116.66 \%$$

وهذا يعني أن الأسعار قد زادت في عام 2002 بنسبة 16.66%، عما كانت عليه

عام 1998 .

(2) رقم قياسي لاسبير للأسعار :

$$PI_{(L)} = \frac{\sum P_n q_o}{\sum P_o q_o} \cdot 100$$

$$= \frac{1042}{880} \cdot 100 = 118.40 \%$$

أي أن الأسعار وفق رقم قياسي لاسبير قد زادت بنسبة 18.40% في عام 2002،

بالمقارنة مع سنة الأساس 1998 .

(3) رقم قياسي باش للأسعار :

$$PI_{(p)} = \frac{\sum P_n q_n}{\sum P_o q_n} \cdot 100$$

$$= \frac{1151}{979} \cdot 100 = 117.56 \%$$

وهذا يفسر أن الأسعار قد زادت وفق رقم باش بنسبة 10.46% في عام 2002 بالمقارنة بسنة الأساس 1998 .

(4) الرقم القياسي الأمثل للأسعار (فيشر) :

$$PI_{(p)} = \sqrt{\frac{\sum p_n q_o}{\sum p_o q_o} \cdot \frac{\sum p_n q_n}{\sum p_o q_n}} \cdot 100 = \sqrt{I_{(L)} \cdot I_{(p)}}$$

$$= \sqrt{\frac{1042}{880} \cdot \frac{1151}{979}} \cdot 100 = \sqrt{118.40 \times 117.56}$$

$$= 117.97$$

ويبدل هذا الرقم أيضاً على زيادة في أسعار هذه السلع بنسبة 17.97% في عام 2002 ، بالمقارنة بسنة الأساس 1998 .

(5) رقم قياسي مارشال ادجورث للأسعار:

$$PI_{(ME)} = \frac{\sum P_n (q_o + q_n)}{\sum P_o (q_o + q_n)} \cdot 100 = \frac{\sum (P_n q_o + q_n q_o)}{\sum (P_o q_o + q_o q_n)} \cdot 100$$

$$PI_{(ME)} = \frac{2193}{1859} \cdot 100 = \frac{1042 + 1151}{880 + 979} \cdot 100 = 118 \%$$

ويمكن حساب هذا الرقم أيضاً بطريقة الوسط الهندسي أي :

$$PI_{(ME)} = \frac{\sum p_n \sqrt{q_o \cdot q_n}}{\sum p_o \sqrt{q_o \cdot q_n}} \cdot 100 = \frac{1089.6}{923.45} \cdot 100 = 118\%$$

نلاحظ أن نتائج الأرقام القياسية التجميعية لأسعار هذه السلع الخمسة مختلفة فيما بينها، غير أنها جميعاً تدل على الاتجاه نفسه بالزيادة في أسعار هذه السلع موضوع الدراسة وهذا يعود إلى اختلاف نظام الترجيح في تركيب هذه الأرقام .

(6) الرقم القياسي للمناسيب البسيطة لأسعار هذه السلع باستخدام المتوسط الحسابي:

$$PI_{(\bar{x})} = \frac{\sum \left(\frac{p_n}{p_o} \cdot 100 \right)}{n} = \frac{599.37}{5} = 119.87\%$$

(7) الرقم القياسي للمناسيب البسيطة لأسعار هذه السلع باستخدام الوسط الهندسي:

$$PI_{(G)} = \sqrt[5]{\frac{p_n^{(a)}}{p_o^{(a)}} \cdot \frac{p_n^{(b)}}{p_o^{(b)}} \cdot \frac{p_n^{(c)}}{p_o^{(c)}} \cdot \frac{p_n^{(d)}}{p_o^{(d)}} \cdot \frac{p_n^{(E)}}{p_o^{(E)}}} \cdot 100$$

$$= \sqrt[5]{122.22 \times 112.12 \times 128.27 \times 127.77 \times 108.96}$$

$$\log PI_G = \frac{1}{5} (\log 122.22 + \log 112.12 + \log 128.27 + \log 127.77 +$$

$$+ \log 108.96)$$

$$= \frac{1}{5} (2.08 + 2.04 + 2.10 + 2.10 + 2.03)$$

$$= \frac{1}{5} (10.35) = 2.07$$

$$PI_{(0)} = \frac{(2.07)}{10} = 117.48 \%$$

(8) الرقم القياسي للمناسيب المرجحة لأسعار هذه السلع باستخدام المتوسط الحسابي:

أ- الرقم القياسي لمناسيب الأسعار المرجحة بالقيمة في سنة الأساس $(P_0 Q_0)$:

$$PI_{(\bar{x})} = \frac{\sum \left(\frac{P_n}{P_0} \cdot P_0 Q_0 \right)}{\sum P_0 Q_0} \cdot 100 = \frac{1042}{880} \cdot 100 = 118.40$$

وهو نفسه رقم قياسي لاسير للأسعار .

ب- الرقم القياسي لمناسيب الأسعار المرجحة بالقيمة في سنة المقارنة $(P_n Q_n)$:

$$PI_{(\bar{x})} = \frac{\sum \left(\frac{P_n}{P_0} \cdot P_n Q_n \right)}{\sum P_n Q_n} \cdot 100 = \frac{1360.72}{1151} \cdot 100 = 118.22$$

وإذا حسبنا الوسط التوافقي لهذا الرقم فنحصل على رقم قياسي باش للأسعار .

ج- الرقم القياسي لمناسيب الأسعار المرجحة بالقيمة $(P_n Q_0)$:

$$PI_{(\bar{x})} = \frac{\sum \left(\frac{P_n}{P_0} \cdot P_n Q_0 \right)}{\sum P_n Q_0} \cdot 100 = \frac{1240.59}{1042} \cdot 100 = 119$$

د - الرقم القياسي لمناسيب الأسعار المرجحة بالقيمة $(P_0 Q_n)$:

$$PI_{(\bar{x})} = \frac{\sum \left(\frac{P_n}{P_0} \cdot P_0 Q_n \right)}{\sum P_0 Q_n} \cdot 100 = \frac{1151}{979} \cdot 100 = 117.56$$

وهو نفسه رقم قياسي باش للأسعار .

9-3-3-2 الطريقة التجميعية في تركيب الأرقام القياسية للكميات :

استعرضنا سابقاً عند دراسة مناسيب الكميات أهمية هذه المناسيب ودورها في تحديد التغيرات التي تطرأ على الكميات في فترتي الأساس والمقارنة، حيث تم تركيب بعض الأرقام القياسية باستخدام المتوسطات التي تعتبر جميعها مؤشرات هامة تعكس متوسط التغير في كميات الاستهلاك، أو كميات الإنتاج الخاصة بالسلع موضوع الدراسة، وما الطريقة التجميعية في تركيب الأرقام القياسية للكميات سوى إحدى الطرائق الهامة القائمة على أساس قسمة متوسط الكميات في سنة المقارنة على متوسط الكميات في سنة الأساس، حيث يمكن أن يكون هذا المتوسط حسابياً بسيطاً تكون جميع السلع الداخلة في تركيب الرقم القياسي متساوية في الأهمية، أو حسابياً مرجحاً تعطى فيه كل سلعة وزناً يدل على أهميتها النسبية بالمقارنة مع السلع الأخرى، ولذلك يمكن تصنيف الأرقام القياسية للكميات إلى مجموعتين حسب أسلوب التركيب :

9-3-3-2-1 الطريقة التجميعية البسيطة .

9-3-3-2-2 الطريقة التجميعية المرجحة .

9-3-3-2-1 الطريقة التجميعية البسيطة :

تستخدم الطريقة التجميعية البسيطة في تركيب الأرقام القياسية للكميات عندما تكون جميع وحدات قياس السلع الداخلة في تركيب الرقم متجانسة، أما إذا اختلفت وحدات التسعير أو القياس للسلع الداخلة في تركيب الرقم، فعندها لا يمكن استخدام هذه الطريقة بل ويجب البحث عن طريقة أخرى. إن الرقم الذي يمكن تركيبه وفق هذه الطريقة هو الرقم القياسي التجميعي البسيط للكميات وصيغته :

$$qI_{(s)} = \frac{\sum q_n}{\sum q_0} \cdot 100$$

غير أن هذه الصيغة تعاني من بعض العيوب كما هو الحال بالنسبة للرقم القياسي التجميعي البسيط للأسعار، حيث إنها لا تأخذ بالاعتبار الأهمية النسبية للسلع التي تدخل في تركيب الرقم، ولذلك يفضل الاعتماد على الطريقة التجميعية المرجحة في تركيب الأرقام القياسية للكميات بدلاً عنها .

2-2-3-9 الطريقة التجميعية المرجحة :

يمكن استخدام هذه الطريقة في تركيب الأرقام القياسية التجميعية للكميات بالأسلوب نفسه الذي تم استخدامه في تركيب الأرقام القياسية التجميعية للأسعار، والفارق الوحيد بين هاتين الطريقتين ينحصر في نوع الأرقام التي تستخدم كأوزان للترجيح، أي لإبراز الأهمية النسبية للسلع التي يتضمنها الرقم القياسي، ومن أهم الأرقام التي تستخدم في قياس التغير في الكميات حيث تستخدم الأسعار فيها للترجيح نذكر:

1- الرقم القياسي التجميعي للكميات المرجح بأسعار سنة الأساس (لاسيبر للكميات) وصيغته :

$$qI_{(L)} = \frac{\sum q_n P_0}{\sum q_0 P_0} \cdot 100$$

وباعتبار أن الأسعار هي عبارة عن أوزان ثابتة في الترجيح، فإن التغير في قيمة هذا الرقم سوف يعزى إلى التغير في الكميات بين فترتي الأساس والمقارنة.

2- الرقم القياسي التجميعي للكميات المرجح بأسعار سنة المقارنة (باش للكميات) وصيغته :

$$qI_{(P)} = \frac{\sum q_n P_n}{\sum q_0 P_n} \cdot 100$$

3- الرقم القياسي الأمثل للكميات (فيشر) :

$$qI(P) = \sqrt{I_{(L)} \cdot I_{(P)}} = \sqrt{\frac{\sum q_n P_o}{\sum q_o P_o} \cdot \frac{\sum q_n P_n}{\sum q_o P_n}} \cdot 100$$

4- الرقم القياسي التجميعي للكميات المرجح بأسعار سنة الأساس والمقارنة معاً وصيغته :

$$qI_{(ME)} = \frac{\sum q_n (P_o + P_n)}{\sum q_o (P_o + P_n)} \cdot 100$$

كذلك يمكن استخدام صيغ أخرى تقيس التغير في الكميات، بحيث يكون الترجيح بأسعار سنة الأساس، أو سنة المقارنة، أو سنة الأساس والمقارنة معاً، غير أن الصيغ التي تعتمد الترجيح بأسعار الأساس والمقارنة معاً في تركيب الأرقام القياسية للكميات ليست شائعة الاستعمال من الناحية العملية .

5- الرقم القياسي التجميعي للكميات المرجح بمتوسط أسعار عدة سنوات كأوزان ثابتة، بالإضافة إلى الصيغ السابقة التي تستخدم في تركيب الأرقام القياسية للكميات، يمكننا تركيب رقم قياسي للكميات مرجحاً بأوزان ثابتة، حيث تبقى الأسعار في فترة الأساس إلى فترة المقارنة ثابتة دون تغيير وصيغته :

$$qI_{(Pi)} = \frac{\sum q_n P_{(Pi)}}{\sum q_o P_{(Pi)}} \cdot 100$$

مثال (9-17) :

الجدول الآتي يعبر عن بيانات الإنتاج والأسعار لمجموعة من السلع في بلد ما خلال السنوات 2000، 2001، 2002 .

جدول رقم (9-16)

السلع	الكميات المتبعة			الأسعار		
	2000	2001	2002	2000	2001	2002
برادات	30	40	50	10	20	30
تلفزيونات	40	50	60	10	14	28
مكيفات	10	20	30	30	40	50
حسابات	60	70	80	20	23	26

باعتبار سنة 2000 سنة أساس، المطلوب :

- 1- إيجاد الرقم القياسي التجميعي البسيط للكميات في السنوات 2000، 2001، 2002.
- 2- إيجاد رقم قياسي لاسير للكميات في السنوات 2000، 2001، 2002 .
- 3- حساب رقم باس للكميات في السنوات 2000، 2001، 2002 .
- 4- إيجاد رقم قياسي فيشر للكميات في السنوات 2000، 2001، 2002 .
- 5- حساب رقم قياسي مارشال ادجورث للكميات في السنوات الثلاث .
- 6- إيجاد الرقم القياسي لمناسيب الكميات المرححة بنسب القيم لكل سلعة إلى مجموع القيم في نقطة الأساس كأوزان .

الحل :

نرمز للكميات في سنة الأساس 2000 بالرمز $q_{(0)}$ وللأسعار بالرمز $P_{(0)}$.
 وللكميات في سنة المقارنة 2001 بالرمز $q_{(1)}$ وللأسعار بالرمز $P_{(1)}$.
 وللكميات في سنة المقارنة 2002 بالرمز $q_{(2)}$ وللأسعار بالرمز $P_{(2)}$.
 رمزنا للكميات في سنة 2002 بالرمز $q_{(2)}$ للتمييز بينها وبين كميات 2001 ($q_{(1)}$)
 حيث تعتبر مثل $q_{(1)}$ كنقطة مقارنة بالنسبة لنقطة الأساس $q_{(0)}$ (2000) وهذا ينطبق
 أيضاً على الأسعار.

من أجل تسهيل العمليات الحسابية نقوم بتنظيم الجدول رقم (9-17) .

جدول رقم (17-9)

السلع	الكميات			الأسعار												
	2000	2001	2002	$P_{(0)}$	$P_{(1)}$	$P_{(2)}$	$q_{(0)}$	$q_{(1)}$	$q_{(2)}$	$q_{(0)}$	$q_{(1)}$	$q_{(2)}$	$q_{(0)}$	$q_{(1)}$	$q_{(2)}$	$q_1 \cdot p_2$
براديات	38	40	50	10	20	30	300	400	500	600	800	900	1500	40/30	50/30	1200
تلفزيونات	50	60	10	14	28	400	500	600	560	700	1120	1680	50/40	60/40	1400	
مكيفات	20	30	30	40	50	300	600	900	400	800	500	1500	2	3	1000	
خلاطات	70	80	20	23	26	1200	1400	1600	1380	1610	1560	2080	70/60	80/60	1820	
	180	220				2200	2900	3600	2940	3910	4080	6760			5420	
	$P_0 + P_{(2)}$	$q_{(0)} (P_{(0)} + P_{(1)})$	$q_1 (P_0 + P_{(1)})$	$q_2 (P_0 + P_{(2)})$	$q_{(0)} (P_{(0)} + P_{(2)})$	$\frac{q_{(1)}}{q_{(0)}} \cdot P_0 \cdot q_{(0)}$	$\frac{q_{(2)}}{q_{(0)}} \cdot P_0 \cdot q_{(0)}$									
	40	900	1200	1500	1200	400	500									
	38	960	1200	1440	1520	500	600									
	80	700	1400	2100	800	600	900									
	46	2580	3010	3440	2760	1400	1600									
		5140	6810	8480	6280	2900	3600									

(1) الرقم القياسي التجميعي البسيط للكميات في السنوات الثلاث :

في سنة الأساس 2000 :

$$qI_{(s)} = \frac{\sum (q_0)}{\sum (q_0)} \cdot 100 = \frac{140}{140} \cdot 100 = 100 \%$$

في سنة المقارنة 2001 :

$$qI_{(s)} = \frac{\sum (q_1)}{\sum (q_0)} \cdot 100 = \frac{180}{140} \cdot 100 = 128.57$$

في سنة المقارنة 2002 :

$$qI_{(s)} = \frac{\sum (q_2)}{\sum (q_0)} \cdot 100 = \frac{220}{140} \cdot 100 = 157.14$$

(2) رقم قياسي لاسير للكميات في السنوات الثلاث :

في سنة الأساس (2000) :

$$qI_{(L)} = \frac{\sum q_{(0)} P_{(0)}}{\sum q_{(0)} P_{(0)}} \cdot 100 = \frac{2200}{2200} \cdot 100 = 100 \%$$

في سنة المقارنة (2001) :

$$qI_{(L)} = \frac{\sum q_{(1)} P_{(0)}}{\sum q_{(0)} P_{(0)}} \cdot 100 = \frac{2900}{2200} \cdot 100 = 131.81$$

في سنة المقارنة (2002) :

$$qI_{(L)} = \frac{\sum q_{(2)} P_{(0)}}{\sum q_{(0)} P_{(0)}} \cdot 100 = \frac{3600}{2200} \cdot 100 = 163.63$$

وهذه النتائج تعني أنه في عام 2001 قد زادت الكميات بنسبة 31.81% وفي عام 2002 بنسبة 63.63% مقارنة بالكميات في سنة الأساس 2000 .

(3) رقم قياسي باش للكميات في السنوات الثلاث :

في سنة الأساس (2000) :

$$qI_{(P)} = \frac{\sum q_{(0)} P_{(0)}}{\sum q_{(0)} P_{(0)}} \cdot 100 = \frac{2200}{2200} \cdot 100 = 100 \%$$

في سنة المقارنة (2001) :

$$qI_{(P)} = \frac{\sum q_{(1)} P_{(1)}}{\sum q_{(0)} P_{(1)}} \cdot 100 = \frac{3910}{2940} \cdot 100 = 133$$

في سنة المقارنة (2002) :

$$qI_{(P)} = \frac{\sum q_{(2)} P_{(2)}}{\sum q_{(0)} P_{(2)}} \cdot 100 = \frac{6760}{4080} \cdot 100 = 165.68$$

نلاحظ من النتائج أن الكميات المنتجة من هذه السلع قد زادت في عام 2001 بنسبة 33%، بالمقارنة مع سنة الأساس 2000 ، في حين أن الزيادة سنة 2002 بلغت 65.68% ، وهذا يدل على تغير كبير في الكميات المنتجة في عامي 2001 و 2002 مقارنة بنسبة الأساس 2000 .

(4) رقم قياسي فيشر للكميات في السنوات الثلاث :

في سنة الأساس (2000) :

$$qI_{(0)} = \sqrt{I_{(L)} I_{(P)}} = \sqrt{100 \cdot 100} = 100 \%$$

في سنة المقارنة (2001) :

$$qI_{(0)} = \sqrt{\frac{\sum q_{(1)} p_{(0)}}{\sum q_{(0)} p_{(0)}} \cdot \frac{\sum q_{(1)} p_{(1)}}{\sum q_{(0)} p_{(1)}}} \cdot 100 = \sqrt{131.81 \times 133}$$

في سنة المقارنة (2002) :

$$qI_{(0)} = \sqrt{\frac{\sum q_{(2)} p_{(0)}}{\sum q_{(0)} p_{(0)}} \cdot \frac{\sum q_{(2)} p_{(2)}}{\sum q_{(0)} p_{(2)}}} \cdot 100 = \sqrt{163.63 \times 165.68}$$

$$= 164.65$$

(5) رقم قياسي مارشال ادجورث للكميات في السنوات الثلاث :

في سنة الأساس (2000) :

$$qI_{(ME)} = 100 \%$$

في سنة المقارنة (2001) :

$$qI_{(ME)} = \frac{\sum q_{(1)} (p_{(0)} + q_{(1)})}{\sum q_{(0)} (p_{(0)} + q_{(1)})} \cdot 100 = \frac{6810}{5140} \cdot 100 = 132.49$$

في سنة المقارنة (2002) :

$$qI_{(ME)} = \frac{\sum q_{(2)} (p_{(0)} + q_{(2)})}{\sum q_{(0)} (p_{(0)} + q_{(2)})} \cdot 100 = \frac{8480}{6280} \cdot 100 = 135$$

(6) الرقم القياسي لمناسيب الكميات المرجحة بالقيمة في سنة الأساس، وصيغة

هذا الرقم :

$$qI_{(\bar{x})} = \frac{\sum \left(\frac{q_n}{q_0} \cdot q_0 p_0 \right)}{\sum q_0 p_0} \cdot 100$$

في سنة المقارنة (2001) :

$$QI_{(x)} = \frac{\sum \left(\frac{q_{(1)}}{q_{(0)}} \cdot q_0 \cdot p_0 \right)}{\sum q_0 p_0} \cdot 100 = \frac{2900}{2200} \cdot 100 = 131.81$$

وهو نفسه رقم قياسي لاسير للكميات سنة (2001) .

في سنة المقارنة (2002) :

$$QI_{(x)} = \frac{\sum \left(\frac{q_{(2)}}{q_{(0)}} \cdot q_0 \cdot p_0 \right)}{\sum q_0 p_0} \cdot 100 = \frac{3600}{2200} \cdot 100 = 163.63$$

وهو نفسه رقم قياسي باش للكميات سنة 2002 .

9-4 اختبارات الأرقام القياسية :

تهدف اختبارات الأرقام القياسية إلى معرفة جودة الرقم القياسي الذي تم تركيبه، ومدى صلاحية استخدامه في دراسة التغيرات التي تطرأ على الظواهر الاقتصادية والاجتماعية موضع الدراسة، ولا سيما التغيرات التي تتعرض لها الأسعار والكميات المنتجة بالنسبة لسلعة واحدة أو مجموعة سلع معينة وسواء أكانت على مستوى الاقتصاد الجزئي أم الاقتصاد الكلي، بالارتفاع كانت أو بالانخفاض حيث يفضل من الناحية النظرية أن تحقق الأرقام القياسية لمجموعات من السلع الخواص التي تحققها المناسب، وأي رقم قياسي له خاصية معينة يمكن القول إنه يحقق الاختبار المرتبط بهذه الخاصية، فعلى سبيل المثال، الأرقام القياسية التي لها خاصية الانعكاس في الزمن يقال عنها أنها تحقق اختبار الانعكاس في الزمن، ... وهكذا .

لم يتم حتى الآن التوصل إلى تركيب رقم قياسي يحقق كل الاختبارات، على الرغم من أنه في بعض الحالات تتحقق هذه الاختبارات، فمثلاً رقم قياسي فيشر يحقق اختبار الانعكاس في الزمن واختبار الانعكاس في المعامل، ولهذا سُمي بالرقم القياسي

الأمثل، ومن الناحية العملية يمكن استخدام أرقام قياسية أخرى سوف نقوم باختبار بعضها وفقاً لثلاثة أنواع من الاختبارات :

- 9-4-1 اختبار الانعكاس في الزمن (Time Reversal Test) .
- 9-4-2 اختبار الانعكاس في المعامل (Factor Reversal Test) .
- 9-4-3 الاختبار الدائري (Circular Test) .

فعند اقتراح فيشر الاختبارين الأول والثاني، مؤكداً أن أي رقم قياسي لا يحقق هذين الاختبارين لا يعتبر رقماً قياسياً أمثلاً، وبناء على ذلك فقد قام بتركيب رقم قياسي يحقق هذين الاختبارين، أطلق عليه الرقم القياسي الأمثل .

9-4-1 اختبار الانعكاس في الزمن :

يعدّ اختبار الانعكاس في الزمن أحد الاختبارات التي تفرض على الأرقام القياسية نوعاً من التناسق الرياضي، الذي يقضي بأن يكون حاصل ضرب الرقمين القياسيين المتبادلين في الزمان أو المكان يساوي الواحد، أي أن الصيغة الجيدة للأرقام القياسية يجب أن تعمل من الناحية الزمنية في اتجاهين، ويعني ذلك أن الرقم القياسي في فترة المقارنة المحسوب بالنسبة لفترة الأساس يجب أن يعطي نتيجة عكسية بالنسبة نفسها للرقم القياسي في فترة الأساس بالنسبة لفترة المقارنة، بمعنى أن كل رقم منها يكون مقلوب الآخر، ويكون حاصل ضربها واحداً صحيحاً .

ولإجراء هذا الاختبار على الأرقام القياسية يجب إيجاد البديل الزمني للصيغة موضوع البحث، والبديل الزمني لأي رقم قياسي هو الرقم نفسه محسوباً بطريقة عكسية، وذلك بجعل سنة الأساس في الرقم الأصلي سنة مقارنة في البديل الزمني، وسنة المقارنة في الرقم الأصلي سنة أساس في البديل الزمني. وقاعدة هذا الاختبار هي :

$$\text{الرقم القياسي الأصلي} \times \text{البديل الزمني} = 1$$

وتطبق هذا الاختبار على الأرقام القياسية تبين أن هناك ثلاث صيغ فقط تحقق

هذا الاختبار وهي :

1- الرقم القياسي التجميعي البسيط للأسعار :

$$PI_{(S)} = \frac{\sum P_n}{\sum P_0}$$

بديله الزمني هو :

$$\frac{\sum P_0}{\sum P_n}$$

فيكون الاختبار :

$$\frac{\sum P_n}{\sum P_0} \cdot \frac{\sum P_0}{\sum P_n} = 1$$

2- الرقم القياسي التجميعي البسيط للكميات :

$$\frac{\sum q_n}{\sum q_0}$$

بديله الزمني هو :

$$\frac{\sum q_0}{\sum q_n}$$

الاختبار :

$$\frac{\sum q_n}{\sum q_0} \cdot \frac{\sum q_0}{\sum q_n} = 1$$

3- الرقم القياسي التجميعي للأسعار المرجح بالوسط الحسابي لكميات سنة

الأساس والمقارنة معاً (رقم قياسي مارشال ادجورث) :

$$PI_{(ME)} = \frac{\sum P_n (q_0 + q_n)}{\sum P_0 (q_0 + q_n)}$$

بديله الزمني :

$$= \frac{\sum P_0 (q_0 + q_n)}{\sum P_n (q_0 + q_n)}$$

اختبار الانعكاس في الزمن :

$$\frac{\sum P_n (q_0 + q_n)}{\sum P_0 (q_0 + q_n)} \cdot \frac{\sum P_0 (q_0 + q_n)}{\sum P_n (q_0 + q_n)} = 1$$

وكذلك بالنسبة للرقم القياسي التجميعي للكميات المرجح بمتوسط أسعار سنة الأساس والمقارنة معاً .

4- الرقم القياسي الأمثل (فيشر) للأسعار :

$$PI_{(f)} = \sqrt{\frac{\sum p_n q_o}{\sum p_o q_o} \cdot \frac{\sum p_n q_n}{\sum p_o q_n}} \cdot 100$$

بديله الزمني :

$$\sqrt{\frac{\sum p_o q_n}{\sum p_n q_n} \cdot \frac{\sum p_o q_o}{\sum p_n q_o}}$$

فإن الرقم القياسي الأصلي لفيشر $\times$ بديله الزمني = 1 .

أما بالنسبة إلى الأرقام القياسية الأخرى كجميعها لا يحقق هذا الاختبار رقم قياسي لاسبير مثلاً :

$$\frac{\sum p_n q_o}{\sum p_o q_o} \cdot \frac{\sum p_o q_n}{\sum p_n q_n} \neq 1$$

رقم باش :

$$\frac{\sum p_n q_n}{\sum p_o q_n} \cdot \frac{\sum p_o q_o}{\sum p_n q_o} \neq 1$$

وبتطبيق رقمي على هذه الأرقام نستخدم معطيات المثال (9-16) فنجد :

(1) اختبار الرقم القياسي التجميعي البسيط للأسعار :

$$\frac{105}{90} \cdot \frac{90}{105} = 1$$

(2) اختبار رقم مارشال ايجورث :

$$\frac{2193}{1859} \cdot \frac{1859}{2193} = 1$$

(3) اختبار رقم قياس فيشر :

$$\sqrt{\left[\frac{\sum p_n q_o}{\sum p_o q_o} \cdot \frac{\sum p_n q_n}{\sum p_o q_n} \right] \left[\frac{\sum p_o q_n}{\sum p_n q_n} \cdot \frac{\sum p_o q_o}{\sum p_n q_o} \right]}$$

$$\sqrt{\left[\frac{1042}{880} \cdot \frac{1151}{979} \right] \left[\frac{979}{1151} \cdot \frac{880}{1040} \right]} = 1.1797 \times 0.84835 = 1$$

بحق لقاعدة الاختبار، وهكذا بالنسبة لرقم قياس فيشر للكميات .

أما رقم قياسي لاسير للأسعار :

$$\frac{1042}{880} \cdot \frac{979}{1151} \approx 1$$

ورقم باش أيضاً لا يحقق هذا الاختبار :

$$\frac{1151}{979} \cdot \frac{880}{1042} \approx 1$$

وهكذا يمكن اختبار جميع الأرقام القياسية سواء أكانت للأسعار أم للكميات، والرقم الذي يحقق قاعدة هذا الاختبار تكون صلاحيته أفضل من غيره من الأرقام .

9-4-2 اختبار الانعكاس في المعامل :

يقوم هذا الاختبار على قاعدة التغيرات في الأسعار والكميات، وانعكاس ذلك على قيم السلع الداخلة في تركيب الرقم القياسي انطلاقاً من قاعدة (السعر × الكمية = القيمة) وعندما كانت الأرقام القياسية تقيس هذه التغيرات في الأسعار والكميات، فإنها تقيس في الوقت نفسه التغيرات التي تطرأ على القيم، فإذا ما ضربنا الرقم القياسي للأسعار بالرقم القياسي للكميات، فسوف نحصل على رقم قياسي للقيم، ولإجراء هذا الاختبار يجب أن نبحث عن البديل المعاملي وفق القاعدة الآتية لهذا الاختبار:

الرقم القياسي الأصلي × البديل المعاملي = الرقم القياسي للقيمة

أي إذا كان لدينا رقم قياسي للأسعار، فإن البديل المعاملي لهذا الرقم هو الرقم القياسي للكميات، أي أن البديل المعاملي لأي رقم قياسي هو الرقم نفسه من حيث فترة الأساس وفترة المقارنة، ولكن نستبدل الأسعار بالكميات والكميات بالأسعار، وبتطبيق هذا الاختبار على صيغ الأرقام القياسية التي سبق بحثها تبين أن الرقم القياسي الأمثل هو الرقم الوحيد الذي يحقق هذا الاختبار، وإن إقامة الدليل على أن الرقم القياسي الأمثل يحقق هذا الاختبار يمكن البرهان عليه عن طريق ضرب الرقم القياسي الأمثل للأسعار بالرقم القياسي الأمثل للكميات، أي بتطبيق قاعدة هذا الاختبار الرقم القياسي الأمثل للأسعار $\times$ الرقم القياسي الأمثل للكميات = الرقم القياسي للقيمة. ويساوي:

$$\begin{aligned} & \sqrt{\frac{\sum p_n q_o}{\sum p_o q_o} \times \frac{\sum p_n q_n}{\sum p_o q_n}} \times \sqrt{\frac{\sum q_n p_o}{\sum q_o p_o} \times \frac{\sum q_n p_n}{\sum q_o q_n}} \\ &= \sqrt{\frac{\sum (p_n q_n)^2}{\sum (p_o q_o)^2}} = \frac{\sum p_n q_n}{\sum p_o q_o} \end{aligned}$$

كما أنه يمكن البرهان عددياً على أن الرقم القياسي الأمثل يحقق هذا الاختبار، وذلك بالعودة إلى المثال (9-16) :

$$\begin{aligned} & \sqrt{\frac{1042}{880} \times \frac{1151}{979} \times \frac{979}{880} \times \frac{1151}{1042}} = \sqrt{\frac{(1151)^2}{(880)^2}} \\ &= \frac{1151}{880} \cdot 100 = 130.79 \% \end{aligned}$$

يُحقق اختبار الانعكاس في المعامل .

بالإضافة إلى رقم قياسي فيشر يمكننا اختبار بعض الأرقام القياسية، وفق قاعدة اختبار الانعكاس في المعامل للتأكد من أنها لا تحقق هذا الاختبار .

1- الرقم القياسي التجميعي البسيط للأسعار :

الاختبار :

$$\frac{\sum p_n}{\sum p_o} \cdot \frac{\sum q_n}{\sum q_o} \neq \frac{\sum p_n q_n}{\sum p_o q_o}$$

وبالتالي فإن هذا الرقم لا يحقق الاختبار ويمكن التأكد عددياً من خلال معطيات

المثال (9-16) :

$$\frac{105}{90} \cdot \frac{63}{60} \neq \frac{1151}{880}$$

على الرغم من أنه يحقق اختبار الانعكاس في الزمن .

2- رقم قياسي لاسبير للأسعار :

الاختبار :

$$\frac{\sum p_n q_o}{\sum p_o q_o} \cdot \frac{\sum q_n p_o}{\sum q_o p_o} \neq \frac{\sum p_n q_n}{\sum p_o q_o}$$

وبتطبيق ذلك عددياً على المثال (9-16) نجد :

$$\frac{1042}{880} \cdot \frac{979}{880} \neq \frac{1151}{880}$$

إذن هذا الرقم لا يحقق اختبار الانعكاس في المعامل

3- رقم قياسي باش للأسعار :

الاختبار :

$$\frac{\sum p_n q_n}{\sum p_o q_n} \cdot \frac{\sum q_n p_n}{\sum q_o p_n} \neq \frac{\sum p_n q_n}{\sum p_o q_o}$$

وبالتطبيق العددي على المثال (9-16) نجد :

$$\frac{1151}{979} \cdot \frac{1151}{1042} \neq \frac{1151}{880}$$

وبالتالي فإن هذا الرقم أيضاً لا يحقق قاعدة اختبار الانعكاس في المعامل .

4- رقم قياسي مارشال ادجورث للأسعار :

الاختبار :

$$\frac{\sum P_n (q_0 + q_n)}{\sum P_0 (q_0 + q_n)} \cdot \frac{\sum q_n (p_0 + p_n)}{\sum q_0 (p_0 + p_n)} \neq \frac{\sum P_n q_n}{\sum P_0 q_0}$$

أي أن هذا الرقم أيضاً لا يحقق اختبار الانعكاس في المعامل، في حين أنه حقق اختبار الانعكاس في الزمن، ولهذا يبقى الرقم القياسي الأمثل (فيشر) هو الرقم الوحيد من بين الأرقام القياسية الذي يحقق هذين الاختبارين .

9-4-3 الاختبار الدائري (Circular test)

يعدّ هذا الاختبار بمثابة امتداد لاختبار الانعكاس في الزمن، على الرغم من أنه يشترط توافر التناسق الرياضي التام بين أي عددين عائدين لسنتين مختلفتين لأية أرقام قياسية، سواء أكانت لأسعار التجزئة أم الجملة دون النظر عما إذا كان أحد العددين يعود إلى سنة الأساس أم لا .

يقوم هذا الاختبار في أبسط أشكاله على ضرورة توفر إمكانية تغير سنة الأساس لأي رقم قياسي، دون أن تكون هناك حاجة إلى إعادة الحساب من الأسعار الأصلية، وذلك عن طريق القيام بتقسيم كل رقم على الرقم القياسي المتخذ كأساس. وإن القيم التي نحصل عليها وفق الأساس الجديد يجب أن تكون مماثلة لتلك الأرقام التي يمكن توفيرها فيما لو قمنا بإعادة الحساب مباشرة من الأسعار الأصلية، وتعود أهمية هذا الاختبار إلى أن هناك عدداً من الأرقام القياسية للأسعار والإنتاج والاستخدام... الخ متوافرة ومنشورة في بلدان العالم، كل واحدة منها قد تختلف عن الأخرى من حيث سنة الأساس، وقد يكون من الملائم تحويل جميع هذه الأرقام القياسية إلى أساس واحد ومشترك بغية جعل المقارنة فيما بينها أكثر سهولة، وتجنب الإشارة هنا إلى أن تحويل

مختلف الأرقام القياسية إلى سنة أساس مشتركة لا يؤدي إلى جعلها قابلة للمقارنة فيما بينها إلا بصورة شكلية، حيث تجري المقارنات عادة بين عدة فترات 0, 1, 2, 3 في حين اقتصر المقارنات في دراستنا أعلاه للأرقام القياسية على المقارنات الثنائية (Binary comparisons) ومقارنة نقطة إلى نقطة (Point-to point comparisons)، ولهذا سوف نستخدم سلسلة من الأرقام القياسية :

$$P_{(0/1)} , P_{(0/2)} , P_{(0/3)}$$

وكل منها يعبر عن مستوى الأسعار في تلك الفترة بالقياس إلى فترة الأساس، وكل من هذه الأرقام هو عبارة عن مقارنات ثنائية بين فترة الأساس وفترة المقارنة، أما إذا رغبتنا في مقارنة الأسعار في فترة (2) لفترة (1) كأساس، فعندها نقوم بإجراء مقارنة ثنائية مباشرة، وحساب $P_{(1/2)}$ أو يمكن أن نتوصل إلى النتيجة ذاتها عن طريق قسمة $P_{(0/2)}$ إلى $P_{(0/1)}$ أي نستطيع في مثل هذه الحالة استخدام إما $P_{(1/2)}$ مباشرة أو $\frac{P_{(0/2)}}{P_{(0/1)}}$ وبالأسلوب ذاته نستطيع مقارنة الأسعار في فترة (2) مع الأسعار في فترة (0) باستخدام مقارنة ثنائية مباشرة $P_{(0/2)}$ أو عن طريق حاصل ضرب $P_{(0/2)} \times P_{(1/2)}$ وعلى ذلك فالصيغة العامة لهذه المقارنات بالنسبة لجميع الأرقام القياسية يمكن كتابتها كما يأتي :

$$PI_{(b/g)} = \frac{PI_{(a/g)}}{PI_{(a/b)}}$$

وبالتالي فإن :

$$PI_{(a/g)} = PI_{(a/b)} \cdot PI_{(b/g)}$$

حيث b = نرمز لفترة الأساس (base period) .

g = نرمز لفترة المقارنة (given period) .

مثال (9-18) :

إذا كانت :

$$PI_{(0/2)} = 190 , PI_{(0/1)} = 130$$

$$PI_{(1/2)} = \frac{PI_{(0/2)}}{PI_{(0/1)}} \cdot 100 = \frac{190}{130} \cdot 100 = 146.15 \%$$

أما إذا كانت لدينا :

$$PI_{(1/2)} = 150 \% , PI_{(0/1)} = 115 \%$$

$$\therefore PI_{(0/2)} = PI_{(0/1)} \cdot PI_{(1/2)} = 115 \% \times 150 \% = 172.50 \%$$

مثال (9-19) :

نفرض أنه توفرت لدينا المعلومات الآتية المنشورة عن الأرقام القياسية لعدة نقط مقارنة لها جميعاً نقطة أساس واحدة .

$PI_{(0/1)}$	رقم قياسي للأسعار سنة 1995 (أساس سنة 1990 = 100)
$PI_{(0/2)}$	رقم قياسي للأسعار سنة 1996 (أساس سنة 1990 = 100)
$PI_{(0/3)}$	رقم قياسي للأسعار سنة 1997 (أساس سنة 1990 = 100)
$PI_{(0/4)}$	رقم قياسي للأسعار سنة 1998 (أساس سنة 1990 = 100)
$PI_{(0/5)}$	رقم قياسي للأسعار سنة 1999 (أساس سنة 1990 = 100)
$PI_{(0/6)}$	رقم قياسي للأسعار سنة 2000 (أساس سنة 1990 = 100)
$PI_{(0/7)}$	رقم قياسي للأسعار سنة 2001 (أساس سنة 1990 = 100)
$PI_{(0/8)}$	رقم قياسي للأسعار سنة 2002 (أساس سنة 1990 = 100)

والمطلوب حساب الأرقام القياسية الآتية :

$$PI_{(1/2)} , PI_{(2/3)} , PI_{(3/4)} , PI_{(4/5)} , PI_{(5/6)} , PI_{(6/7)} , PI_{(7/8)}$$

يمكن حساب هذه الأرقام جميعاً باستخدام الصيغة العامة السابقة :

$$PI_{(b/g)} = \frac{PI_{(0/g)}}{PI_{(0/b)}}$$

فنجد :

$$PI_{(1/2)} = \frac{PI_{(0/2)}}{PI_{(0/1)}} \cdot 100 \quad PI_{(2/3)} = \frac{PI_{(0/3)}}{PI_{(0/2)}} \cdot 100$$

$$PI_{(3/4)} = \frac{PI_{(0/4)}}{PI_{(0/3)}} \cdot 100 \quad PI_{(4/5)} = \frac{PI_{(0/5)}}{PI_{(0/4)}} \cdot 100$$

$$PI_{(5/6)} = \frac{PI_{(0/6)}}{PI_{(0/5)}} \cdot 100 \quad PI_{(6/7)} = \frac{PI_{(0/7)}}{PI_{(0/6)}} \cdot 100$$

وهكذا :

$$PI_{(7/8)} = \frac{PI_{(0/8)}}{PI_{(0/7)}} \cdot 100$$

إن الأرقام القياسية وفق طريقة المقارنات المتعددة بين عدة نقاط تصبح أكثر مرونة من حيث التطبيق العملي وأفضل فائدة، ويمكن استخدامها جميعاً في دراسات الأسعار والإنتاج والاستهلاك والتوزيع و الخ، وبالتالي من الواضح أنه من المرغوب فيه أن تكون للأرقام القياسية الخاصية الدائرية أو الدورية أي :

$$PI_{(0/g)} = PI_{(0/B)} \cdot PI_{(B/g)}$$

وقاعدة الاختبار الدائري هي :

$$\frac{PI_{(0/B)} \cdot PI_{(B/g)}}{PI_{(0/g)}} = 1$$

ومن الناحية العملية فإن جميع الأرقام القياسية لا تحقق الاختبار الدائري باستثناء الرقمين الآتيين : الرقم القياسي الهندسي للمناسيب البسيطة للأسعار، والرقم القياسي التجميعي المرجح بأوزان ثابتة، ومهدف التأكد من هذه النتيجة يمكننا تطبيق هذا الاختبار على بعض الأرقام القياسية مثل :

(1) رقم قياسي لاسبير للأسعار خلال الفترات الزمنية 0, 1, 2 :

$$PI_{L(0/2)} = \frac{\sum p_{(2)} q_{(0)}}{\sum p_{(0)} q_{(0)}}$$

$$PI_{L(0/1)} = \frac{\sum p_{(1)} q_{(0)}}{\sum p_{(0)} q_{(0)}}$$

$$PI_{L(1/2)} = \frac{\sum p_{(2)} q_{(1)}}{\sum p_{(1)} q_{(1)}}$$

قاعدة الاختبار الدائري :

$$\frac{PI_{(0/1)} \cdot PI_{(1/2)}}{PI_{(0/2)}} = 1$$

وبتطبيق قاعدة الاختبار على رقم قياسي لاسير للأسعار نجد :

$$\frac{\left[\frac{\sum p_{(1)} q_{(0)}}{\sum p_{(0)} q_{(0)}} \cdot \frac{\sum p_{(2)} q_{(1)}}{\sum p_{(1)} q_{(1)}} \right]}{\frac{\sum p_{(2)} q_{(0)}}{\sum p_{(0)} q_{(0)}}} \neq 1$$

أي أن رقم قياسي لاسير لا يحقق قاعدة الاختبار الدائري .

(2) رقم قياسي هاش للأسعار :

$$PI_{P(0/2)} = \frac{\sum p_{(2)} q_{(2)}}{\sum p_{(0)} q_{(2)}}$$

$$PI_{P(0/1)} = \frac{\sum p_{(1)} q_{(1)}}{\sum p_{(0)} q_{(1)}}$$

$$PI_{P(1/2)} = \frac{\sum p_{(2)} q_{(2)}}{\sum p_{(1)} q_{(2)}}$$

قاعدة الاختبار الدائري :

$$\frac{\left[\frac{\sum p_{(1)} q_{(1)}}{\sum p_{(0)} q_{(1)}} \cdot \frac{\sum p_{(2)} q_{(2)}}{\sum p_{(1)} q_{(2)}} \right]}{\frac{\sum p_{(2)} q_{(2)}}{\sum p_{(0)} q_{(2)}}} \neq 1$$

وهذا الرقم أيضاً لا يحقق قاعدة الاختبار الدائري، وبالتالي لا يتصف بالدورية، وهذا الشكل يمكن تطبيق جميع الأرقام القياسية وفقاً لهذا الاختبار .

9-5 تعديل الأرقام القياسية :

لاحظنا من خلال دراسة اختبارات الأرقام القياسية أن معظم الأرقام لا تحقق هذه الاختبارات، حيث حقق اختبار الانعكاس في الزمن ثلاثة أرقام قياسية هي: الرقم القياسي التجميعي البسيط للأسعار والكميات، الرقم القياسي لمارشال ادجورث، والرقم القياسي الأمثل (فيشر) الذي حقق أيضاً اختبار الانعكاس في المعامل، واعتبر الرقم الوحيد الذي يحقق هذين الاختبارين، حيث سمي بالرقم القياسي الأمثل لشمته بهذه الخاصية، أما بالنسبة للاختبار الدائري فإن جمع الأرقام القياسية لا تحقق هذه الخاصية باستثناء الرقم القياسي الهندسي لمناسيب الأسعار، والرقم القياسي التجميعي المرجح بأوزان ثابتة.

من الواضح أن جميع هذه الاختبارات تستخدم بشكل أو بآخر للتحقق من صلاحية الأرقام القياسية المركبة وجودها، غير أن ذلك لا يعني أن أي رقم قياسي لا يحقق هذه الاختبارات يعتبر رقماً قياسياً غير صالح للاستخدام، وذلك لأن الرقم القياسي يكون له معنى فقط وفق نظام الترجيح المتبع، وبالتالي يمكن الحصول على عدة أرقام قياسية بعدة أساليب مختلفة للترجيح وللفترة الزمنية نفسها، فمثلاً نرى أن نظام الترجيح هو نفسه المتبع في تركيب رقم قياسي لاسبير $PI_{(0/2)}$ ، $PI_{(0/1)}$ في حين أنه يختلف عند تركيب $PI_{(1/2)}$. وهذا يعني أن الزيادة بين النقطتين (0, 1) و (1, 2) شيء مختلف

تماماً عن الزيادة التي يمكن أن يعكسها الرقم القياسي (2, 0) ، وحاصل ضرب هاتين الزيادةتين يمثل خليطاً ليس له معنى في تحليل الأرقام القياسية، ولهذا نرى أنه من الأفضل اللجوء إلى فكرة تعديل الأرقام القياسية ، وجعلها تتصف بالصفة الأمثلية أي تتمتع بخاصية الانعكاس في الزمن وخاصية الانعكاس في المعامل، وذلك بالاعتماد على الأسلوبين الآتيين :

1- تعديل البديل الزمني عن طريق قسمة الواحد الصحيح على البديل الزمني للرقم القياسي المطلوب تعديله، فنحصل على ما يسمى بالمقلوب الزمني أو المعكوس الزمني للرقم، وبأخذ المتوسط الهندسي للرقم للقياسي الأصلي مضروباً بمقلوبه الزمني نحصل على الرقم القياسي المعدل القابل لتحقيق اختبار الانعكاس في الزمن.

2- تعديل البديل المعاملي عن طريق قسمة منسوب القيم $(\sum P_n q_n / \sum P_0 q_0)$ على البديل المعاملي للرقم القياسي موضع التعديل، فنحصل على ما يسمى بالمقلوب المعاملي أو المعكوس المعاملي للرقم، وبأخذ المتوسط الهندسي للرقم القياسي الأصلي مضروباً بمقلوبه المعاملي نحصل على الرقم القياسي المعدل الذي يحقق اختبار الانعكاس في المعامل، بتطبيق هذين الأسلوبين على جميع الأرقام القياسية يمكننا الحصول على أرقام قياسية معدلة تحقق اختبار الانعكاس في الزمن واختبار الانعكاس في المعامل فمثلاً :

رقم قياسي ياش للأسعار :

$$PI_{(p)} = \frac{\sum P_n q_n}{\sum P_0 q_n}$$

بديله الزمني :

$$= \frac{\sum P_0 q_0}{\sum P_n q_0}$$

المقلوب أو المعكوس الزمني :

$$\frac{1}{\frac{\sum P_0 q_0}{\sum P_n q_0}} = \frac{\sum P_n q_0}{\sum P_0 q_0}$$

رقم قياسي باش المعدل :

$$= \sqrt{\frac{\sum p_n q_n}{\sum p_o q_n} \cdot \frac{\sum p_n q_o}{\sum p_o q_o}}$$

الذي يحقق خاصية الانعكاس في الزمن .

البديل المعاملي لرقم قياسي باش للأسعار (رقم الكميات) :

$$= \frac{\sum q_n p_n}{\sum q_o p_n}$$

$$= \frac{\text{الرقم القياسي للقيمة}}{\text{البديل المعاملي}} = \text{المقلوب (المعكوس) المعاملي}$$

$$= \frac{\frac{\sum p_n q_n}{\sum p_o q_o}}{\frac{\sum q_n p_n}{\sum q_o p_n}}$$

$$= \frac{\sum q_o p_n}{\sum p_o q_o}$$

رقم قياسي باش المعدل :

$$= \sqrt{\frac{\sum p_n q_n}{\sum p_o q_n} \cdot \frac{\sum q_o p_n}{\sum p_o q_o}}$$

نلاحظ أن الرقم القياسي المعدل لباش الذي تم الحصول عليه وفق الأسلوبين السابقين، هو نفسه الرقم القياسي الأمثل (فيشر) الذي يحقق الاختبارين معاً الانعكاس في الزمن والانعكاس في المعامل، وبالتالي فإن أية صيغة من صيغ الأرقام القياسية المركبة لا تحقق الانعكاس في الزمن، أو في المعامل يجب معالجتها مرتين: مرة بالنسبة للزمن وأخرى بالنسبة للمعامل، على الرغم من أن هذا التعديل للأرقام القياسية يتطلب

عمليات حسابية معقدة ويؤدي إلى الحصول على صيغ مبهمه وغير محددة، وأكثر صعوبة في التركيب، لذلك ينحصر هذا التعديل في الإطار النظري فقط بهدف تطوير نظرية الأرقام القياسية وجعلها تستند بصورة أفضل على أسس منطقية وعلمية دقيقة، تعكس واقع الحال، وتعمل على تفسير جميع الأرقام القياسية التي يتم الحصول عليها وتوحيدها في صيغ قائمة على أسس ثابتة، وبخاصة عندما يكون لدينا مجموعتان من الأرقام بنقاط أساس مختلفة ونرغب في تعديلها، حيث يحتاج ذلك إلى القيام بتعديل نقطة الأساس عن طريق وصل نقطة الأساس، أو نقلها إلى فترة زمنية أخرى تعكس الحالة موضع الاهتمام والدراسة .

إن تعديل نقطة الأساس لمجموعة من الأرقام القياسية يهدف إلى توحيد نقطة الأساس في مجموعتين من الأرقام القياسية يكون الأساس فيهما مختلفاً، وبالتالي إلى تعديل الأرقام القياسية التي تم تركيبها في هاتين المجموعتين من خلال وصل نقطة الأساس، أو نقلها إلى فترة زمنية أخرى .

مثال (9-20) :

ليكن لدينا الجدول الآتي الذي يعبر عن البيانات الضرورية لتركيب نوعين من الأرقام القياسية :

جدول (9-18)

السنوات	1998	1999	2000	2001	2002
النوع A	$\sum p_0 q_0 = 7$	$\sum p_n^{(1)} q_0 = 8$	$\sum p_n^{(2)} q_0 = 9$		
النوع B			$\sum p_n^{(2)} q_n^{(1)} = 13$	$\sum p_n^{(3)} q_n^{(1)} = 16$	$\sum p_n^{(4)} q_n^{(1)} = 18$

فإن الأرقام القياسية التي يمكن تركيبها لهذه البيانات تتضح من الجدول الآتي :

جدول (9-19)

السنوات	1998	1999	2000
النوع A	$PI_{(0/0)} = \frac{7}{7} \cdot 100 = 100$	$PI_{(0/1)} = \frac{8}{7} \cdot 100 = 114.28$	$PI_{(0/2)} = 128.85$
النوع B			$PI_{(1/2)} = \frac{13}{13} \cdot 100 = 100$

تابع الجدول (9-19)

السنوات	2000	2002
النوع A		
النوع B	$PI_{(1/3)} = \frac{16}{13} \cdot 100 = 123$	$PI_{(1/4)} = 138.46$

نلاحظ أن الأوزان في سنة 2000 تغيرت من النقطة q_0 إلى $q_n^{(1)}$ وبالتالي لتعديل الأرقام القياسية يجب تغير نقطة الأساس للأرقام في السنوات 1998 , 1999 , 2000 إلى سنة 2000 بدلاً من 1998، أو بتغير نقطة الأساس للأرقام في السنوات 2001 و 2002 إلى 1998، بدلاً من 2000 أي :

جدول (9-20)

السنوات	1998	1999	2000	2001	2002
النوع A	$PI_{(0/0)} = 100$	$PI_{(0/1)} = 114.28$	$PI_{(0/2)} = 128.85$	$PI_{(0/3)} = ?$	$PI_{(0/4)} = ?$
النوع B	$PI_{(2/0)} = ?$	$PI_{(2/1)} = ?$	$PI_{(2/2)} = 100$	$PI_{(2/3)} = 123$	$PI_{(2/4)} = 138.46$

وبحساب قيم الأرقام القياسية المجهولة بعد تعديل نقطة الأساس إلى عام 2000 نجد :

$$PI_{(0/3)} = \frac{PI_{(0/2)} \cdot PI_{(2/3)}}{PI_{(2/2)}} = \frac{128.85 \times 123}{100} = 158.48$$

$$PI_{(0/4)} = \frac{PI_{(0/2)} \cdot PI_{(2/4)}}{PI_{(2/2)}} = \frac{128.85 \times 138.46}{100} = 178.40$$

$$PI_{(2/0)} = \frac{PI_{(0/0)} \cdot PI_{(2/2)}}{PI_{(0/2)}} = \frac{100 \times 100}{128.85} = 77.60$$

$$PI_{(2/1)} = \frac{PI_{(0/1)} \cdot PI_{(2/2)}}{PI_{(0/2)}} = \frac{114.28 \times 100}{128.85} = 88.69$$

نلاحظ أن نسبة الزيادة بين الرقمين $PI_{(2/2)}$, $PI_{(2/3)} = 23\%$ وهي تساوي نسبة الزيادة بين $PI_{(0/2)}$, $PI_{(0/3)}$ التي تساوي :

$$= \frac{158.48 - 128.85}{128.85} \cdot 100 = 23 \%$$

وكذلك الزيادة بين الرقمين $PI_{(2/2)}$, $PI_{(2/4)} = 38.46 \%$ تساوي النسبة بين الرقمين : $PI_{(0/2)}$, $PI_{(0/4)}$

$$= \frac{178.40 - 128.85}{128.85} \cdot 100 = 38.46 \%$$

وهكذا بالنسبة لبقية الأرقام التي تم تعديلها بتغيير نقطة الأساس عن طريق وصلها بين المجموعتين المدروستين من الأرقام القياسية أما إذا تم نقل نقطة الأساس في مجموعتين من الأرقام القياسية لتكون سنة 2000 بدلاً من سنة 1998 كما هو واضح من الجدول رقم (9-21) :

جدول (9-21)

السنوات	1998	1999	2000
المجموعة A%	100	300	700
المجموعة B%	$PI_{(1)}$	$PI_{(2)}$	100

فيمكننا حساب أرقام 1998 (أساس سنة 2000) وأرقام 1999 (أساس سنة 2000 = 100) وذلك بالقسمة على 700 فنجد :

$$PI_{(1)} = \frac{100 \times 100}{700} = 14.28 \%$$

$$PI_{(2)} = \frac{300 \times 100}{700} = 42.85 \%$$

6-9 الأرقام القياسية الشائعة الاستعمال:

1- الرقم القياسي لأسعار الجملة :

الرقم القياسي لأسعار الجملة يقيس متوسط التغير في أسعار مجموعة معينة من مواد الجملة التي يتم تبادلها خلال فترة معينة من الزمن (شهر، فصل، سنة)، ولعل من أهم استعمالاته قياس المستوى العام للأسعار، وقياس القوة الشرائية للنقود، ويساعد على رسم السياسات المتعلقة بتحديد مستوى الأسعار والأرباح والضرائب والاستيراد والتصدير، بالإضافة إلى فائدته في الحسابات القومية، ويشتمل هذا الرقم عادةً على أهم الأصناف الموجودة في أسواق الجملة، فهو يتركب في سورية من (190) مادة مقسمة إلى ستة أقسام رئيسية هي :

- 1- المواد الغذائية .
- 2- المواد العلفية .
- 3- مواد أولية زراعية .
- 4- المواد المصنوعة .
- 5- المحروقات .
- 6- مواد البناء .

حيث يتألف كل قسم من عدة مواد مختلفة. فالمواد الغذائية مثلاً تتألف من ثمانين فئات هي: الحبوب والدقيق، والبقول الجافة، واللحوم، والزيتون والدهون، والخضار والفواكه، والمشروبات والتبغ، ومواد غذائية معلبة، ومواد غذائية أخرى. أما بالنسبة لفترة الأساس فقد تم اتخاذ سنة (1990) كسنة أساس، وذلك لأن الشروط الاقتصادية ووسطي الأسعار كانت ملائمة في هذه السنة وقد استعمل متوسط الكميات المباعة بالجملة في السنة نفسها كأوزان للترجيح، ونجيب الإشارة إلى أن الطريقة المستخدمة في تركيب هذا الرقم تم تبنيها بعد عام 1990 بالاعتماد على أسلوب لاسبير وباش في تركيب الأرقام القياسية.

2- الرقم القياسي لأسعار المستهلك :

يطلق على هذا الرقم أحياناً: «الرقم القياسي لأسعار التجزئة أو الرقم القياسي لنفقة المعيشة»⁽¹⁾ ويقيس هذا الرقم تكلفة السلع الضرورية اللازمة لسد حاجات المستهلكين خلال فترة زمنية محددة. ويستعمل هذا الرقم لتحقيق أغراض متعددة أهمها:

أ- تحديد أجور العمال والموظفين وحساب الأجر الحقيقي .

ب- تحديد السياسة الضريبية على الاستهلاك والأجور وسياسات الأسعار .

ج- تحديد معدل التضخم النقدي وحساب القوة الشرائية لوحدة النقد.

ويشتمل هذا الرقم عادةً على مواد تمثل أهم السلع والخدمات التي يستهلكها المستهلك في نقطة زمنية معينة، ويتكون هذا الرقم في سورية من (150) مادة موزعة على (14) قسمًا رئيسيًا، حيث اعتمد المكتب المركزي للإحصاء هذه المواد بناءً على نتائج المسوح بالعينة لتكاليف معيشة الأسر، والتي تسمى بحوث ميزانية الأسر التي قام بها المكتب المركزي للإحصاء، بالإضافة إلى بعض الجهات العامة مثل وزارة التموين والتجارة الداخلية، وشركات التجارة في القطاع العام، واتخذت سنة 1970 كسنة أساس، حيث حافظ المكتب المركزي على هذا الأسلوب في الحساب حتى عام 1993، وفي عام 1994 طور المكتب أسلوب تركيب هذا الرقم، واتخذ سنة 1990 سنة أساس، وذلك لأن الشروط الاقتصادية التي سادت خلالها اعتبرت طبيعية نسبياً، بالإضافة إلى أنه زاد عدد المواد التي يتركب منها الرقم إلى (183) مادة تشكل مجملها تكاليف معيشة الأسر. ويتم تركيب هذا الرقم باستخدام طريقة الوسط الحسابي المرجح لمناسيب الأسعار .

3- الرقم القياسي لحجم الصادرات والواردات:

يقيس الرقم القياسي لحجم الصادرات والواردات التغير الحاصل في قيم الواردات والصادرات بأسعار ثابتة، بهدف تحديد التغير الحقيقي في اتجاهاتها حيث إن التغير في

(1) نفقة المعيشة : تعني جملة ما ينفقه المستهلك على احتياجاته من سلع وخدمات، وهي التي تحدد مستوى المعيشة والدخل الحقيقي للمستهلكين .

حجم الصادرات والواردات، قد يكون نتيجة تغير كميات السلع المصدرة أو المستوردة أو نتيجة تغير أسعارها، ولذلك لكي نتمكن من معرفة حجم التجارة الخارجية للدولة ما يجب أن يكون الهدف من تركيب هذا الرقم هو استبعاد أثر تغيرات الأسعار. حيث يتم تركيب هذا الرقم باستخدام طريقة الوسط الحسابي المرجح لمناسيب الكميات، وتستخدم قيم السلع في فترة الأساس كأوزان حسب إحصاءات التجارة الخارجية. ومن الواضح أن إجراء مقارنة بين الرقم القياسي لأسعار الصادرات، والرقم القياسي لأسعار الواردات، تدلنا على معدل التبادل في التجارة الخارجية فيما إذا كان لصالح الدولة أم لا .

4- الرقم القياسي للإنتاج الصناعي :

يهدف الرقم القياسي للإنتاج الصناعي إلى إظهار التغيرات في حجم الإنتاج الصناعي، ويساعد على تتبع تطور الإنتاج الصناعي بعيداً عن أثر ارتفاع الأسعار وانخفاضها في حجم الإنتاج، وهو بذلك يحقق عدة أغراض من أهمها :

أ- قياس معدلات نمو العينة في قطاع الصناعة .

ب- تقويم الإنتاج في قطاع الصناعة بالأسعار الثابتة وإمكانية استخدامها في إعادة تسعير حسابات القيمة المضافة الصناعية وتحويلها من أسعار جارية إلى أسعار ثابتة بهدف المقارنة .

ج- التعرف على التغير الهيكلي في بنية القطاع الصناعي .

د - التنبؤ بالاتجاهات المستقبلية لحجم الإنتاج الصناعي، حيث فترة الأساس لهذا الرقم هي في الغالب سنة، والأوزان التي تستخدم في الترجيح هي القيمة الصافية لإنتاج كل سلعة في فترة الأساس، حيث يتم تركيب هذا الرقم باستخدام طريقة الوسط الحسابي المرجح لمناسيب الكميات.

5- الرقم القياسي للأوراق المالية في أسواق البورصة :

تساعد الأرقام القياسية للأوراق المالية في إظهار حركة النشاط الاقتصادي وما تعكسه من فوائد للمتعاملين في البورصات، وبشكل خاص للمشرفين على صناديق المعاشات والضمان الاجتماعي الذين يرغبون في تتبع أسعار الأسهم، حيث يُستثمر جزء كبير من أموال هذه الصناديق في الأسهم المختلفة التي يجري تداولها في سوق الأوراق المالية، فمثلاً يُجد الأرقام القياسية لجريدة الفاييننشال تايمز تقيس التغيرات التي تظهر على أسعار الأوراق المالية البريطانية التي تتداول في سوق (بورصة) الأوراق المالية، في لندن، بالإضافة إلى الأرقام القياسية الخاصة بقياس عائدات هذه الأوراق وكسبها، حيث يتم نشر هذه الأرقام يومياً مع بعض الشرح في الجريدة المذكورة، ويتم تركيب هذه الأرقام بالتعاون مع معهد خبراء التأمين في لندن وفي أدنبرة، وفقاً لطريقة السلسلة في تركيب الأرقام القياسية، وعامل الترجيح في هذه الأرقام هي الأوزان في سنة الأساس، بالإضافة إلى الأرقام القياسية لجريدة الفاييننشال تايمز هناك الرقم القياسي - داوجونز الذي يقيس تغيرات وتقلبات أسعار الأسهم في بورصة وول ستريت في الولايات المتحدة الأمريكية، حيث يعتبر مؤشراً هاماً يعكس حركة النشاط الاقتصادي في أمريكا، ونحب الإشارة إلى أننا يومياً نشاهد ونسمع ونقرأ عن الكثير من الأرقام القياسية التي تعبر عن بورصات مختلفة من العالم سواء أكانت للأوراق المالية أم لأسعار النفط أم لأسعار العملات أم لأسعار الذهب والمعادن الثمينة الخ .



تمارين ومسائل

- 1 - وضع مفهوم الرقم القياسي وبين أهميته .
- 2 - ما العلاقة الرابطة بين الرقم القياسي ومعدل النمو ؟
- 3 - وضع دور الرقم القياسي في بيان التغير الحقيقي في الظواهر، والعمليات الاقتصادية والاجتماعية وغيرها .
- 4 - بين العناصر الأساسية التي تدخل في تركيب الرقم القياسي.
- 5 - اذكر الاعتبارات التي يجب مراعاتها عند اختيار عناصر الرقم القياسي.
- 6 - ما الشروط التي يجب أن تتوفر عند اختيار نقطة الأساس ؟
- 7 - ما الأشكال التي يمكن أن تأخذها فترة الأساس ؟
- 8 - اذكر طرائق تركيب الأرقام القياسية، وشرح طريقة السلسلة في تركيب الأرقام.
- 9 - اذكر خصائص مناسبة الأسعار .
- 10 - اذكر خصائص مناسبة الكميات .
- 11 - اذكر خصائص مناسبة القيمة .
- 12 - برهن جبرياً أن رقمي لاسبير للأسعار وباش للكميات لا يحققان جميع الاختبارات الإحصائية للأرقام القياسية .
- 13 - في عام 1980 تم إنشاء المبنى الحالي لكلية الاقتصاد بمبلغ (20) مليون ليرة سورية. حيث كان الرقم القياسي لتكاليف المواد اللازمة للبناء هو 131% بالنسبة لعام

1970، وقد أريد الآن بناء مبنى مماثل للكلية عام 2002 وتم رصد مبلغ (40) مليون ليرة سورية، حيث بلغ الرقم القياسي لتكاليف المواد اللازمة للبناء 270% بالنسبة لعام 1970، هل تكفي المبالغ المخصصة لبناء المبنى الجديد عام 2002؟ وما نسبة الإنجاز؟

- 14- ما الفرق بين المقلوب الزمني والبديل الزمني لرقم قياسي ما .
- 15- بين كيف يمكنك تحويل رقم قياسي تجميعي مرجح لا يحقق أياً من اختبائي الانعكاس في الزمن والانعكاس في المعامل إلى رقم يحققهما معاً؟
- 16- لمة رقم قياسي يحقق اختبائي الانعكاس في الزمن والدائري معاً، اذكره ثم تأكد من صحة تحقيقه لاختبار الانعكاس في الزمن فقط .
- 17- ما هو الرقم القياسي المحقق لكل من اختبائي الانعكاس في الزمن، والانعكاس في المعامل (برهن ذلك جبرياً) .
- 18- في أي الحالات تصبح الأرقام القياسية التجميعية المرجحة محققة لجميع الاختبارات المطبقة عليها، ثم اكتب قاعدة كل اختبار من تلك الاختبارات؟
- 19- ما الفرق بين الأساس الثابت، والأساس المتغير في تركيب الرقم القياسي؟
- 20- اذكر الأرقام القياسية الأكثر استخداماً وبين مجالات استخدام كل منها.
- 21- الجدول الآتي يبين التغير في كميات الاستهلاك والأسعار لثلاث سلع غذائية :

السلعة	الأسعار (بالوحدات النقدية)		الكميات المستهلكة (بالكيلو غرام)	
	2000	2001	2000	2001
A	20	25	40	60
B	10	13	80	100
C	25	50	20	40

والمطلوب :

- 1- حساب الأرقام القياسية البسيطة لمناسيب الأسعار باستخدام الوسط الحسابي، الوسط الهندسي، الوسط التوافقي للسلع الثلاث.
 - 2- حساب الرقم القياسي التجميعي للأسعار في سنة المقارنة المرجع بكميات الأساس.
 - 3- حساب الرقم القياسي التجميعي للأسعار في سنة المقارنة المرجع بكميات المقارنة.
 - 4- حساب الرقم القياسي التجميعي للأسعار في سنة المقارنة المرجع بنقطتي الأساس والمقارنة .
 - 5- حساب الرقم القياسي التجميعي للأسعار في سنة المقارنة المرجع بمتوسط كميات نقطتي الأساس والمقارنة .
 - 6- حساب الرقم القياسي لمناسيب الكميات المرجح بنسب القيم لكل سلعة في سنة الأساس إلى مجموع القيم في نقطة الأساس كأوزان .
 - 7- حساب الرقم القياسي لمناسيب الأسعار المرجح بنسب القيم لكل سلعة في سنة الأساس إلى مجموع القيم في نقطة الأساس كأوزان .
 - 8- بين رقمياً الأسباب التي أدت إلى تسمية رقم فيشر بالرقم القياسي الأمثل .
- 22- الجدول الآتي يبين أسعار وقيم ثلاث سلع:

السلع	الأسعار		القيم	
	2001	2002	2001	2002
A	10	12	250	480
B	15	18	600	900
C	20	30	2000	3000

المطلوب :

- 1- حساب الرقم القياسي التجميعي للكميات المرحجة بأسعار سنة الأساس ومن ثم الرقم القياسي التجميعي للكميات المرحجة بأسعار سنة المقارنة .
- 2- هل يحقق رقم قياسي لاسبير الاختبار الدائري؟ برهن ذلك جبرياً ورقمياً.
- 3- حساب الرقم القياسي لمناسيب الكميات المرحجة بالقيمة في سنة المقارنة وفسر جميع النتائج .
- 4- حساب الرقم القياسي لمناسيب الكميات المرحجة بالقيمة في سنة الأساس .
- 23- ليكن لدينا أسعار ثلاث سلع خلال سنتين والكميات المشتراة خلال فترة المقارنة.

السلع	الأسعار		الكميات 2002
	2001	2002	
قمح	17	20	182
رز	16	21	3600
لحم	250	300	90

المطلوب :

- 1- تركيب جميع الأرقام القياسية التجميعية المرحجة والمنسجمة مع طبيعة هذه البيانات .
- 2- أثبت أن رقم قياسي لاسبير لا يحقق اختبار الانعكاس في المعامل إذا علمت أن الكميات المباعة في عام 1998 للسلع الثلاث على الترتيب 90 - 3500 - 150 .
- 24- الجدول الآتي يمثل سلسلة مناسيب الكميات المباعة من السلعة A والكميات المباعة من السلعة B :

السنة	2000	2001	2002
سلسلة مناسيب الكميات من السلعة (A)	120 %	130 %	135 %
الكميات المباعة من السلعة (B) بالطن	15	20	30

المطلوب :

- 1- إيجاد منسوب الكمية للسلعة (A) لعام 2002 بالنسبة لعام 1999 .
- 2- إذا علمت أن الكميات المباعة من السلعة (A) في السنوات الثلاث هي على التوالي 6, 7.8, 11.7 (طن)، وأن سنة الأساس عام 2000 وأسعار السلعتين A و B كانت ثابتة لا تتغير خلال السنوات الثلاث، وتساوي 10 وحدات نقدية. احسب الرقم القياسي التجميعي لمناسيب الكميات لعام 2002 فقط متخذاً الأسعار الثابتة كأوزان .
- 25- يمثل الجدول الآتي كميات وقهم صادرات إحدى الشركات الخاصة بحلال عامي 2000-2001 باعتبار عام 2000 سنة أساس، حيث قدرت الكميات بالطن والقيمة بملايين الوحدات النقدية :

المادة	الكمية		القيمة	
	2000	2001	2000	2001
A	100	120	2500	4800
B	150	180	6000	9000
C	200	300	20000	30000
D	50	100	3000	6000

المطلوب :

- 1- حساب الرقم القياسي للأسعار بصيغتي لاسبير وباش .
- 2- أثبت أن رقم قياسي فيشر ينعكس في المعامل.
- 3- حساب الرقم القياسي لمناسيب الأسعار المرجحة بكميات سنة الأساس كأوزان وفسر جميع النتائج .
- 4- هل يحقق الرقم القياسي للأسعار المركب وفق صيغتي لاسبير وباش اختباري الانعكاس في الزمن والانعكاس في المعامل؟ وما الإجراءات الممكنة اتباعها لجعلها تجتاز هذين الاختبارين ؟ (برهن وشرح ذلك) .

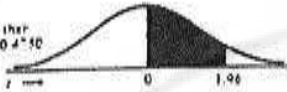


الجدول الإحصائية

1- جدول التوزيع الطبيعي $\Phi(z)$

TABLE V

For $\mu = 0$
 $\sigma = 1$ or there
 $\mu = 0, \sigma = 1$



AREAS UNDER THE NORMAL CURVE

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6951	0.6987	0.7023	0.7058	0.7093	0.7129	0.7164	0.7199	0.7234
0.6	0.7267	0.7301	0.7334	0.7367	0.7400	0.7433	0.7465	0.7497	0.7529	0.7561
0.7	0.7593	0.7625	0.7657	0.7688	0.7719	0.7749	0.7779	0.7809	0.7839	0.7869
0.8	0.7899	0.7929	0.7959	0.7989	0.8019	0.8049	0.8079	0.8109	0.8139	0.8169
0.9	0.8199	0.8229	0.8259	0.8289	0.8319	0.8349	0.8379	0.8409	0.8439	0.8469
1.0	0.8499	0.8529	0.8559	0.8589	0.8619	0.8649	0.8679	0.8709	0.8739	0.8769
1.1	0.8799	0.8829	0.8859	0.8889	0.8919	0.8949	0.8979	0.8999	0.9029	0.9059
1.2	0.9089	0.9119	0.9149	0.9179	0.9209	0.9239	0.9269	0.9299	0.9329	0.9359
1.3	0.9389	0.9419	0.9449	0.9479	0.9509	0.9539	0.9569	0.9599	0.9629	0.9659
1.4	0.9689	0.9719	0.9749	0.9779	0.9809	0.9839	0.9869	0.9899	0.9929	0.9959
1.5	0.9989	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

2- جدول الأرقام العشوائية

48611	62866	33963	14045	79451	04934	45376
78812	03509	78673	73181	29973	18664	04555
19472	63971	37271	31443	49019	49405	46925
51266	11569	08697	91120	64156	40365	74297
55906	96275	26130	47949	14877	69594	83641
77527	81360	18180	97421	53541	90275	18213
77640	58758	33016	61173	93049	04694	43534
15404	96554	88265	34537	38526	67924	40474
14045	22917	60718	66487	46346	30949	03173
68376	43918	77653	04127	69930	43283	35766
93385	13421	67957	20384	58731	53396	59723
09858	52104	32014	53113	03727	98624	84616
93307	34116	49516	42148	57740	31198	70336
04794	01534	92058	03157	91758	80611	45357
86265	49096	97021	92582	61422	75890	86442
65943	79232	45702	67055	39024	57383	44424
90038	94209	04055	27393	61517	23002	96560
97283	95943	78363	36498	40662	94188	18202
21913	72958	75637	99936	58715	07943	23748
41161	37341	81838	19389	80336	46346	91895
23777	98392	31417	98547	92058	02277	50315
59973	08144	61070	73094	27059	69181	55623
82690	74099	77885	23813	10054	11900	44653
83854	24715	48866	65745	31131	47636	45137
61980	34997	41825	11623	07320	15003	56774
99915	45821	97702	87125	44488	77613	56823
48293	86847	43186	42951	37804	85129	28993
33225	31280	41232	34750	91097	60752	69783
06846	32828	24425	30249	78801	26977	92074
32671	45587	79620	84831	38156	74211	82752
82096	21913	75544	53228	89796	05694	91552
51666	10433	10945	53306	78562	89630	41230
54044	67942	24145	42294	27427	84875	37022
66738	60184	75679	38120	17640	36242	99357
55064	17427	89180	74018	44865	53197	74810
69599	60264	84549	78007	88450	06488	72274
64756	87759	92354	78694	63638	80939	98644
80817	74533	68407	55862	32476	19326	95558
39847	96884	84657	33697	39578	90197	80532
90401	41700	95510	61166	33757	23279	85523
78227	90110	81378	96659	37008	04050	04228
87240	52716	87697	79433	16336	52862	69149
08486	10951	26832	39763	02485	71688	90936
39338	32169	03713	93510	61244	73774	01245
21188	01850	69689	49426	49128	14660	14143
13287	82531	04388	64693	11934	35051	68576
53609	04001	19648	14053	49623	10840	31915
87900	36194	31567	53506	34304	39910	79620
81641	00496	36058	75899	46620	70024	88753
19512	50277	71508	20116	79520	06269	74173

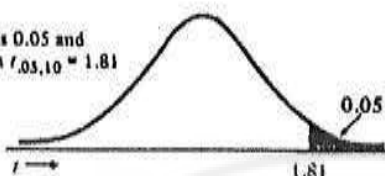
2- تابع جدول رقم (2) الأرقام العشوائية

04433	80674	24520	18222	10610	05794	37515
60298	47829	72648	37414	75755	04717	29879
67884	59651	67333	68123	17730	95862	08034
89512	32155	51906	61662	64130	16648	37275
32653	01495	12506	88535	36553	23757	54209
95913	15405	13772	76638	38423	25018	90041
55864	21694	13122	44115	01601	50541	00147
33334	49810	91601	40617	72876	33967	73830
57729	32196	76487	11622	96297	24160	09903
86648	13697	63677	70119	94739	25875	38829
30374	47609	07967	32422	76791	39725	53711
81307	43694	83580	79974	45929	83113	72268
02410	54905	79007	54939	21410	86980	91772
18969	75274	52333	62319	08598	09066	95288
87863	82384	66860	62297	80198	19347	73234
68397	71708	15438	62311	72844	60203	46412
28529	54447	58729	10854	99058	18260	38765
44285	06372	15867	70418	37012	72122	36634
86299	83430	33571	23309	57040	29285	67870
84842	68668	90894	61638	15001	94055	36308
56970	83609	52098	04184	54967	72938	56834
83125	71257	60490	44369	66130	72936	69848
55503	52423	02464	26141	68779	66388	75242
47019	76273	33203	29608	54553	25971	69573
84828	32592	79526	29554	84580	37859	28504
68921	08141	79227	05748	51276	57143	31926
36458	96045	30424	98420	72923	40729	22337
95752	59445	36847	87729	81679	59126	59437
26768	47323	58454	56958	20575	76746	49878
42613	37036	43636	58085	06766	60227	96414
95457	30566	65482	25596	02678	54592	63607
95276	17894	63564	95958	39730	64379	46059
66954	52324	64776	92345	95110	59448	77249
17457	18481	14113	62462	02798	54977	48349
03704	36872	83214	59337	01695	60666	97410
21538	86497	33210	60337	27976	70661	08250
57178	67619	98310	70348	11317	71623	55510
31048	97558	94953	53866	96383	46620	52087
69799	55380	16498	80733	96422	58078	99643
90595	61867	59231	17772	67831	33317	00520
33570	04981	98939	78784	09977	29398	93896
15340	93460	57477	13898	48431	72936	78160
64079	42483	36512	56186	99098	48850	72527
63491	05546	67118	62063	74958	20946	28147
92003	63868	41034	28260	79708	00770	88643
52360	46658	66511	04172	73085	11795	52594
74622	12142	68355	65635	21828	39539	18988
04157	50079	61343	64315	70836	82857	35335
86003	60070	66241	32836	27573	11479	94114
41268	80187	20351	09636	84668	42486	71303

* Based on parts of Tables of 101,000 Random Dermal Digits, Interstate Commerce Commission, Bureau of Transport Economics and Statistics, Washington D.C.

3- جدول توزیع استودیت t (ت)

Example:
If tail area is 0.05 and
 $\nu = 10$, then $t_{0.05,10} = 1.81$



THE STUDENT t DISTRIBUTION
(Entries are t values for selected tail areas)

Degrees of freedom ν	Tail area					
	0.25	0.10	0.05	0.025	0.01	0.005
1	1.00	3.08	6.31	12.71	31.82	63.66
2	0.82	1.89	2.92	4.30	6.98	9.92
3	0.76	1.64	2.35	3.18	4.54	5.84
4	0.74	1.53	2.13	2.78	3.75	4.60
5	0.73	1.48	2.02	2.57	3.38	4.03
6	0.72	1.44	1.94	2.45	3.14	3.71
7	0.71	1.41	1.89	2.36	3.00	3.50
8	0.71	1.40	1.86	2.31	2.90	3.36
9	0.70	1.38	1.83	2.26	2.82	3.25
10	0.70	1.37	1.81	2.23	2.76	3.17
11	0.70	1.36	1.80	2.20	2.72	3.11
12	0.70	1.36	1.78	2.18	2.68	3.05
13	0.69	1.35	1.77	2.16	2.65	3.01
14	0.69	1.35	1.76	2.14	2.62	2.98
15	0.69	1.34	1.75	2.13	2.60	2.95
16	0.69	1.34	1.75	2.12	2.58	2.92
17	0.69	1.33	1.74	2.11	2.57	2.90
18	0.69	1.33	1.73	2.10	2.55	2.88
19	0.69	1.33	1.73	2.09	2.54	2.86
20	0.69	1.33	1.72	2.09	2.53	2.85
21	0.69	1.32	1.72	2.08	2.52	2.83
22	0.69	1.32	1.72	2.07	2.51	2.82
23	0.69	1.32	1.71	2.07	2.50	2.81
24	0.68	1.32	1.71	2.06	2.49	2.80
25	0.68	1.32	1.71	2.06	2.49	2.79
26	0.68	1.31	1.71	2.06	2.48	2.78
27	0.68	1.31	1.70	2.05	2.47	2.77
28	0.68	1.31	1.70	2.05	2.47	2.76
29	0.68	1.31	1.70	2.05	2.46	2.76
30	0.68	1.31	1.70	2.04	2.46	2.75
40	0.68	1.30	1.68	2.02	2.42	2.70
60	0.68	1.30	1.67	2.00	2.39	2.66
∞	0.67	1.28	1.64	1.96	2.33	2.58

بعض المصطلحات الإحصائية
Some Statistical Terminology

الإنكليزية	العربية
A	
Absolute deviation	الانحراف المطلق
Absolute error	الخطأ المطلق
Absolute value	القيمة المطلقة
Accuracy	دقة
Additive	الجماعي
Adjustment of data	تعديل البيانات
Age distribution	توزيع الأعمار
Age-sex standardized rate	المعدل المنمط للعمر والجنس
Aggregate	تجميع
Aggregate value	القيمة التجميعية
Aging of the population	تغير السكان
Aliases	مترادفات
Allowance limits	حدود السماح
Amplitude	السعة
Amplitude ratio	نسبة السعة
Analysis of covariance	تحليل التغاير
Analysis of variance	تحليل التباين
Applied statistics	الإحصاء التطبيقي
Arbitrary constant	ثابت اختياري
Arbitrary origin	نقطة أصل اختيارية

Area sampling	المعاينة المساحية
Arithmetic mean	الوسط الحسابي
Arithmetic progression	متتالية حسابية
Array	منظومة ، ترتيب
Assembling	تجميع
Assumption	اقتران
Assumed origin	نقطة الأصل المفروضة
Assumption	فرض
Asymmetry	عدم التماثل
Asymmetrical	غير متماثل
Asymptotic	تقريبي
Attribute	صفة
Automatic	تلقائي
Average	متوسط
Average deviation	الانحراف المتوسط
Axis	محور
Axis of abscissas	المحور السيني أو الأفقي
Axis of ordinates	المحور الصادي أو الرأسي
Balanced	متوازن
Base	أساس أو قاعدة
Base line	الخط الأساسي
Base period	فترة الأساس
Basic statistics	إحصاءات أساسية
Battery of tests	مجموعة الاختبارات
Bell-shaped curve	منحنى جرس
Between groups	بين المجموعات
Bias	تحيز

Bimodal	ثنائي المتوال
Binominal distribution	التوزيع الثنائي الحدين
Birth rate	معدل المواليد
Bivariate	ذو متغيرين
Bivariate table	جدول ذو متغيرين
Breakdown	تقسيم
C	
Calculation	حساب
Calculus of probability	حساب الاحتمالات
Calendar month	شهر تقويمي
Caption	عنوان العمود في الجدول
Card counter	عداد للبطاقات
Cartesian coordinates	الأحداثيات الكارتيزية
Case fatality rate	معدل وفيات الحالات
Census	تعداد
Census enumerator	مندوب تعداد (عداد)
Census schedule	استمارة التعداد
Central tendency	النزعة المركزية
Chain indexes	الأرقام القياسية المسلسلة
Chain relatives	منسبات
Chart	لوحة أحصائية
Cumulative chart	لوحة تجميعية
Check	مراجعة
Chi-squared test	اختبار كاي مربع
Class boundaries	أطراف الفئة
Class frequencies	تكرار الفئة
Class interval	فترة الفئة
Class limits	حدود الفئة

Classification	تصنيف
Cluster sampling	المعاينة العنقودية
Code	رمز
Coder	مرمز
Coefficient	معامل
Column	عمود
Column diagram	شكل أعمدة
Completely randomized design	التصميم التام العشبية
Compute	يحسب
Confidence intervals	فترات الثقة
Confidence limits	حدود الثقة
Confounding (partial, complete)	إدماج (جزئي ، كلي)
Consistency	الاتساق
Consistent	متسق
Constant	ثابت
Contingency	توافق
Contingency table	جدول التوافق
Continuous	متصل
Continuous series	سلسلة متصلة
Control	مراقبة
Control chart	لوحة المراقبة
Coordinate	أحداثي
Corrected axis	محور الأحداثيات
Correction factor	معدل تصحيح
Correlation coefficient	عامل التصحيح
Correlation	ارتباط
Correlation coefficient	معامل الارتباط

Correlation index	دليل الارتباط
Correlation surface	نسبة الارتباط
Correlation table	سطح الارتباط
Cost of living	جداول الارتباط
Cost of living	لفتة المعيشة
Count	بعد
Covariable	متغير
Covariance	تغير
Coverage	شمول
Critical value	قيمة حرجية
Cross classifications	تصنيف متقاطع
Cross tabulation	تبويب متقاطع
Crude rate	معدل خام
Cumulant	متراكمة
Cumulative classification	تصنيف تراكمي
Cumulative chart	لوحة تراكمية
Cumulative frequencies	تكرارات تراكمية
Cumulative frequency curve	منحنى تكراري تراكمي
Current statistics	إحصاءات حالية
Curve fitting	توفيق المنحنيات
Curvilinear	منحني
Curvilinear correlation	ارتباط منحن أو غير مستقيم
Curvilinear regression	المحدار منحن
Curvilinear trend	اتجاه منحن
Cycle	دورة
Cyclical variations	تغيرات دورية

D

D ta

بيانات

Deduction	استنباط
De facto population	عدد السكان الفعلي
De jure population	عدد السكان القانوني
Death (S)	وفاة (وفيات)
Death rate	معدل الوفيات
Decile	عشر
Defining contrast	المقارنة المميزة
Degrees of freedom	درجات الحرية
Demography	ديموغرافيا
Dependent variable	متغير تابع
Derivative function	دالة مشتقة
Design of experiments	تصميم التجارب
Determinant	محددة
Deviation	المخالف
Diagram (figure)	شكل
Differential	تفاضلي
Differential calculus	الحساب التفاضلي
Differential equation	معادلة تفاضلية
Differential fertility	الخصوبة التمييزية
Differentiation (calculus)	مفاضلة
Digit	رقم
Discontinuous	غير متصل
Discrete	منفصل
Dispersion	تشتت
Distribution	توزيع أو توزيع
Distribution function	دالة التوزيع
Dot chart	لوحة بالنقط

Dot map

خريطة بالنقط

Double frequency table

جدول تكراري مزدوج

Double sampling

معاينة مزدوجة

or

أو

two-phase sampling

ذات وجهين

Downward bias

تحيز بالنقص أو نحو الأدنى

Downward trend

اتجاه نزولي

E

Econometrics

اقتصاد قياسي

Edit (check)

مراجع أو يدق

Effect

تأثير أو مفعول

Efficiency estimates

فعالية التقدير

Elimination

حذف

Enquiry

استعلام

Enumerator

عداد

Enumeration area or district

منطقة التعداد

Equation

معادلة

Error

خطأ

Error of estimate

خطأ التقدير

Estimate

تقدير

Estimated population

عدد السكان التقديري

Estimation equation

معادلة التقدير

Evaluate

يقوم

Event

واقعة

Expected value

القيمة المتوقعة

Expectation of life

توقع الحياة

Expectation error

الخطأ التجريبي

Experimental units

وحدات تجريبية

Exponential curve	منحنى أسّي
Extrapolation	استيفاء أو استكمال خارجي
F	
F table	جدول « ف »
Factor	عامل
Factor analysis	تحليل العوامل
Factor-reversal test	اختبار الانعكاس العائلي
Factorial analysis	تحليل عائلي
Factorial experiment	تجربة عائلية
Family budget	ميزانية الأسرة
Fertility	خصب أو خصوبة
Fertility	معدل الخصب أو الخصوبة
	حدود الاطمئنان
	الاحتمال الاطمئنان
	ميدان أو حقول
Field of enquiry	ميدان البحث أو مجال البحث
Field staff	موظفو الميدان
Field work	عمل الميدان
Finite	محدود
First quartile	الرابع الأول
Flow chart	لوحة السير
Fluctuations	تقلبات
Forecasting	التنبؤ
Forms	استمارات
Fourfold table	جدول رباعي
Fractional	كسري أو جزئي
Fractional replication	التكرير الجزئي
Frame (sampling)	إطار

Frequency	تكرار
Frequency curve	منحنى التكرار
Frequency density	كثافة التكرار
Frequency distribution	توزيع التكرار
Frequency polygon	مضلع التكرار
Frequency surface	سطح التكرار
Frequency table	جدول التكرار
Function	دالة
Functional relationship	علاقة دالة

G

Gainful occupation	سرفة مكسبة
Gang punch	ثقيب بالجملة
Gaps	ثغراب
Gaussian curve	منحنى جاوس
Geometric mean	وسط هندسي
Graduation	تدريج
Graph	رسم بياني
Graphic analysis	تحليل بياني
Graphic method	طريقة بيانية
Graphic presentation	عرض بياني
Graco-Latin square	مربع جريكو لاتيني
Grid	شبكة مستطيلات
Gross national product	الإنتاج القومي الإجمالي
Grouping	تجميع أو تصنيف
Grouping error	خطأ التجميع أو التصنيف
Growth curve	منحنى النمو

H

Hand tabulation	التبويب اليدوي أو المجدولة اليدوية
-----------------	------------------------------------

Harmonic equation	أدلة توافقية
Harmonic mean	وسط توافقي
Head of household	رب العائلة
Histogram	مدرج تكراري (هستغرام)
Histogram	سلسلة تاريخية
Homogeneous	متجانس
Household	عائلة
Housing census	تعداد السكان
Hypothesis	افتراض
I	
Ideal «index number»	الرقم القياسي «الأمثل»
Identity	متطابقة
Illegitimate birth	مواليد غير شرعية
Illiteracy	أمية
Immediate cause of death	السبب المباشر للوفاة
Implicit	ضمني
Imputed income	الدخل المحتسب
Imputed weight	الترجيح المضمن
Incidence rate	معدل الإصابة
Income in kind	دخل عيني
Incomplete blocks	القطاعات المنقوصة
Inconsistent	غير متسق
Independent variable	متغير مستقل
Index	دليل أو مقياس
Index number	رقم قياسي
Index of concentration	دليل التركيز
Index of correlation	دليل الارتباط
Index of prices	الرقم القياسي للأسعار

Index of quantity	الرقم القياسي للكمية
Index of seasonal variations	الرقم القياسي للتغيرات الموسمية
Indirect transit trade	تجارة عابرة غير مباشرة
Industry	صناعة
Inequality	متباينة أو لا متساوية
Infant mortality	وفيات الرضع
Infant mortality rate	معدل وفيات الرضع
Infectious disease	مرض معد
Infinite	لا نهائي
In-service training	تدريب أثناء الخدمة
Integration	تكامل
Intelligence quotient	نسبة الذكاء
Interaction	تفاعل
Intercensal	بين التعدادات
Interclass correlation	ارتباط بين الفئات
Interest	فائدة
Interest rate	سعر الفائدة
Interpolation	استكمال داخلي
Interquartile range	المدى بين الربيعين
Interval	فترة
Interview	مقابلة
Intraclass correlation	ارتباط داخل الفئات
Invariant	لزوم
Invariant	لازمة
Investigators	باحثون
Inverse correlation	ارتباط عكسي
Inverse function	دالة عكسية

Inverse probability	احتمال عكسي
Invisible trade	تجارة غير منظورة
Irregular variations	تغيرات غير منتظمة
Isometric chart	لوحة بخطوط التساوي

J

Joint causes of death	أسباب مشتركة للوفاة
Joint distribution	توزيع مشترك
Join regression	التحدر مشترك
J-curve	منحنى رالي

K

Kurtosis	تفرطح
----------	-------

L

Labels	علامات مميزة
Labour force	قوة العمل
Labour turn-over	تبدل العمال
Lag	تخلف
Lattice designs	التصميمات الشبكية
Lattice square	المربع الشبكي
Latin square	المربع اللاتيني
Law of large numbers	قانون الأعداد الكبيرة
Leas significant difference	أصغر فرق معنوي
Least squares	المربعات الصغرى
Least squares method	طريقة المربعات الصغرى
Legitimate births	مواليد شرعيين
Lethality rate	معدل الإماتة
Level of living	مستوى المعيشة
Levels of significance	مستويات المعنوية أو الدلالة
Life table	جدول الحياة

Likelihood	إمكان
Limit	نهاية أو حد
Line chart	لوحة خطية
Line graph	شكل خطي
Linear correlation	ارتباط مستقيم
Linear equation	معادلة خطية
Linear interpolation	استكمال مستقيم
Linear regression	المقدار مستقيم
Linear trend	اتجاه مستقيم
Linearity	استقامة
List	قائمة
Live births	مواليد أحياء
Living conditions	الظروف المعيشية
Logarithmic chart	لوحة لوغاريتمية
Logistic curve	منحنى لوجستي
Long cycles	دورات طويلة
Lorenz curve	منحنى لورنز
Lower quartile	الربيع الأدنى
Machine tabulation	التبويب الآلي
Mail investigation	التحري بالبريد
Man-days	يوم رجل
Man-hours	ساعة رجل
Margin of error	هامش الخطأ
Marital status	الحالة الزوجية
Market price	سعر السوق
Market research	بحث الأسواق
Marriage rate	معدل الزواج

Master sample	عينة رئيسية
Maternal mortality rate	معدل وفيات المواليدات الحضانة
Matrix	مصنوفة
Maximum variation	أقصى اختلاف
Mean	وسط
Mean absolute error	وسط الخطأ المطلق
Mean deviation	الانحراف الوسطي
Median	الوسيط
Mid-range	منتصف المدى
Missing Plots	القطع المفقودة
Modal class	الفئة المتوالية
Mode	المنوال
Model	نموذج
Moments	العزوم
Money income	الدخل النقدي
Morbid condition	حالة مرضية
Morbidity incidence rate	معدل الإصابة بالأمراض
Morbidity prevalence rate	معدل انتشار الأمراض
Mortality rate	معدل الوفيات
Moving average	متوسط متحرك
Multimodal	متعدد المنوال
Multi-phase sampling	معاينة متعددة الأوجه
Multiple axis chart	معاينة متعددة المحاور
Multiple correlation	ارتباط متعدد
Multiple factor analysis	تحليل عامل متعدد
Multiplicative	انضغابي ، تضاعفي
Multi-stage sampling	معاينة متعددة المراحل

Multivariate analysis	تحليل المتغيرات المتعددة
Mutually exclusive	متنافية
N	
Natality	ولادة
Natality rate	معدل الولادة
National accounts	الحسابات القومية
National income	الدخل القومي
Natural increase	زيادة طبيعية
Needle sorting	فرز بالإبرة
Negative correlation	ارتباط سالب
Non-determination coefficient	معامل عدم التحديد
Non-linear correlation	ارتباط غير مستقيم
Non-resident	غير مقيم
Non-sampling error	خطأ غير عيني
Normal deviate	المنحرف المعتدل
Normal distribution	التوزيع المعتاد
Normal equations	المعادلات المعتادة
Normal frequency curve	منحنى تكراري معتاد أو معتدل
Null hypothesis	افتراض العدم
Numerical	عددي
O	
Objective	موضوعي، هدف
Observation	مشاهدة
Occupation	حرفة
Ogive curve	منحنى تراكمي أو تجميعي
Open-end class	فئة مفتوحة الطرف
Optimum allocation of the sample	التوزيع الأفضل للعينة
Orthogonal	متعامد

Orthogomality	التعامد
Orthonormal	عادي التعامد
Own-account worker	يشتغل لحسابه
Parabola	قطع مكافئ
Parameter (S)	معلمة (معلم)، ثابت إحصائي
Partial correlation	ارتباط جزئي
Partial regression	المحدد جزئي
Per capita	للفرد
Percentage	نسبة مئوية
Percentile	مئين
Periodicity	دورية
Periodic movements	حركات دورية
Periodogram	شكل المدورات أو الشكل الدوري
Pictograph	رسم تصويري
Pie chart	لوحة دائرية
Pilot survey	بحث استرشادي أو استطلاعي
Platy-kurtic	مفرطح
Plot	قطعة
Polygon	مضلع
Polynomial series	متسلسلة متعددة الحدود
Pooled estimate	تقدير تجمعي
Population (statistical)	مجتمع (إحصائي)
Positive correlation	ارتباط موجب
Positive skewness	التواء موجب
Postcensal	بعد التعداد
Post-multiplication	ضرب لاحق
Precise	دقيق

Precision	أحكام
Preliminary figure	رقم مبدئي
Pre-multiplication	ضرب سابق
Prevalence rate	معدل انتشار (للأمراض)
Price changes	تغيرات الأسعار
Price index	رقم قياسي للأسعار
Price level	مستوى الأسعار
Price relative	منسوب السعر
Primary sampling unit	الوحدة الأولية للمعاينة
Basic source	مصدر أساسي
Primary trend	اتجاه أساسي
Probable error	الخطأ المتبادل
Probability	الاحتمال
Probability density	منحنى الاحتمال
Probability element	كثافة الاحتمال
Probability element	عنصر الاحتمال
Probability level	مستوى الاحتمال
Probability paper	ورق الاحتمالات
Probability ratio test	اختبار النسبة الاحتمالي
Probability sampling	معاينة احتمالية
Probability table	جدول الاحتمالات
Processing of data	تجهيز البيانات
Progression	متوالية
Progressive tax	ضريبة تصاعدية
Proportional allocation of the sample	التوزيع التناسبي للمعينة
Punch card	بطاقة للتنقيب
Purposive sample	عينة عمدية

Purposive selection

اختيار عمدي

Q

Quadratic mean

وسط تربيعي

Qualitative analysis

تحليل وصفي

Quality control

مراقبة الإتقان أو الجودة

Quantative analysis

تحليل كمي

Quantity index number

رقم قياسي للكمية

Quantity relative

منسوب الكمية

Quantum index number

رقم قياسي للكم

Quartile

ربيع

Quartile deviation

المخرف ربعي

Questionnaire

استبيات

Quintiles

خمسيات

Quota sampling

معاينة بالحصص

R

Random distribution

توزيع عشوائي

Random error

خطأ عشوائي

Random numbers

أعداد عشوائية

Random sample

عينة عشوائية

Random selection

اختيار عشوائي

Random variables

متغير عشوائي

Randomization

تعشية

Randomized

معشي

Range

مدى

Rank correlation coefficient

معامل ارتباط الرتب

Rate

معدل

Ratio chart

لوحة النسب

Ratio estimate

تقدير النسب

Ratio estimator	مقدر بالنسبة
Ratio test	اعتبار النسب
Record	سجل
Rectilinear	مستقيم
Regression	الانحدار
Regression coefficient	معامل الانحدار
Regression equation	معادلة الانحدار
Regression estimate	تقدير الانحداري
Regression line	خط الانحدار
Relative dispersion	تشتت نسبي
Relative efficiency	فعالية نسبية
Relative frequency	تكرار نسبي
Reliability	كفاءة أو مقدار الثقة
Replacement	إعادة
Replicate	يكرر - مكررة
Replication	تكرار
Representative sample	عينة ممثلة
Reproducible	قابل للتكرار المتشابه
Reproducibility	القابلية للتكرار المتشابه
Reproduction rate	معدل التغليف
Residual	المتبقي
Response	استجابة
Restricted sample	عينة مقيدة
Ratinal price index	رقم قياسي لأسعار التجزئة أو الفرق
Returns	إجابات
Root-mean square deviation	جذر متوسط مربع الانحراف
Round number	عدد مقرب أو مدور

Rounding error

خطأ التقريب

Row

صف

S

Sample

عينة

Sample census

تعداد بالعينة أو عيني

Sample design

تصميم العينة

Sampling

معاينة

Sampling error

خطأ المعاينة

Sampling fraction

كسر المعاينة

Sampling method

طريقة المعاينة

Sampling ratio

نسبة المعاينة

Sampling unit

وحدة المعاينة

Scale

مقياس الرسم

Scatter diagram

التشعار

Seasonal diagram

شكل الانتشار

Seasonal index

رقم قياسي موسمي

Seasonal variations

تغيرات موسمية

Secondary source

مصدر ثانوي

Secondary trend

اتجاه ثانوي

Second degree curve

منحنى من الدرجة الثانية

Second order coefficients

معاملات من الرتبة الثانية

Secular trend

الاتجاه العام

Semi-interquartile range

نصف المدى الربيعي

Sensitivity

حساسية

Sequence

تتابع

Sequential analysis

تحليل متابعي

Sequential sampling

المعاينة المتتابعة

Serial correlation

ارتباط متسلسل

Short-cuts	طرق مختصرة
Sight check	مراجعة بالنظر
Significance	معنوية
Significant digit	رقم معنوي
Similarity	تشابه
Size of sample	حجم العينة
Skew (in a distribution)	ملتوي (للتوزيع)
Skewness	التواء
Small samples	عينات صغيرة
Smoothed curve	منحنى ممدد
Sorter (machine)	فرازة
Source of error	مصدر الخطأ
Split-plots	القطع المنسقة
Spot check	مراجعة سريعة
Spurious correlation	ارتباط وهمي
Square	مربع
Sage (multi-stage)	مرحلة (متعددة المراحل)
Standard deviation	انحراف معياري
Standard error	خطأ معياري
Standard error of estimate	خطأ معياري للتقدير
Standard error of difference	خطأ معياري للفرق
Between means	بين الوسطين
Standard of living	مستوى المعيشة
Standard death rates	معدلات الوفاة المنمطة
Stationary population	مجتمع ثابت
Statistical induction	استنتاج إحصائي
Statistical inference	استدلال إحصائي

Statistical method	الطريقة الإحصائية
Statistical probability	احتمال إحصائي
Statistical processing	التجهيز الإحصائي
Statistics	إحصاء
Still birth	مولود ميت
Still birth rate	معدل المواليد أمواتاً
Straight line trend	خط اتجاه مستقيم
Stratified sample	عينة طبقية
Stratification	التقسيم إلى طبقات
Stratum	طبقة
Stub	عنوان السطر
Sub-grup	مجموعة فرعية
Subjective	ذاتي
Sub-plots	قطع فرعية
Sub-sample	عينة فرعية
Sub-sample unit	وحدة العينة الفرعية
Sub-title	عنوان فرعي
Summary punch	تخريم تلخيصي
Summation	جمع
Supplementary information	معلومات إضافية (متممة)
Survey	استقصاء (مسح)
Symmetrical	متماثل
System	نظام
Systematic	نظامي
Systematic sampling	معاينة منتظمة
Systematic selection	اختيار نظامي

T

Table

جدول

Tabulating equipment	معدات التبويب أو الجدولة
Tabulation	تبويب أو جدولة
Tabulator	مبوبة أو مجدولة
Tally sheet	صحيفة العلامات
Test of goodness of fit	اختبار جودة التوفيق أو حسن المطابقة
Test of homogeneity	اختبار التجانس
Test of reliability	اختبار الاعتمادية أو الكفاءة
Test of significance	اختبار المعنوية
Tetrad differences	فروق رباعية
Three dimensional chart	لوحة ثلاثية الأبعاد
Third degree curve	منحنى من الدرجة الثالثة
Third quartile	الربيع الثالث
Time-reversal test	اختبار الانعكاس الزمني
Time series	متسلسلة زمنية
Tolerance limits	حدول التسامح
Ton-kilometre	كيلو متر طن
Transformation	تحويل
Treatment	معاملة
Treatment effect	مفعول المعاملة
Trend analysis	تحليل الاتجاه
Trend line	خط الاتجاه
Trial	محاولة أو تجربة
Trial census	تعداد تجريبي
True mean	وسط حقيقي
Type	طراز ، نوع ، نمط
Typical value	قيمة نمطية
U-curve	منحنى ذو شعبتين

Unbiased	غير متحيز
Unemployment	بطالة
Uniform sampling fraction	كسر معاينة ثابت
Unimodal	وحيد المنوال
Unit of measurement	وحدة القياس
Unit value index	رقم قياسي لوحدة القيمة
Universe (statistical)	عالم (أحصائي)
Unweighted index	رقم قياسي غير مرجح أو غير مثقل
Upper quartile	الربيع الأعلى
Upward trend	اتجاه صعودي

V

Variable	متغير
Variance	تباين
Variance ratio	نسبة التباين
Variate	متغير احتمالي
Variation	اختلاف
Verification	مراجعة أو تحقيق
Verifier	مراجع أو محقق
Vital statistics	إحصاءات حيوية

W

Wage earner	عامل بأجر
Weights	أوزان، مرجحات أو ثقلات
Weighted average	متوسط مرجح أو مثقل
Weighted index number	رقم قياسي مرجح أو مثقل
Weighted mean	وسط مرجح أو مثقل
Weighted price index	رقم قياسي مرجح للسعر
Weighted relatives	مناسيب مرجحة
Weighting coefficient	معامل الترجيح

Weighting factor

عامل الترجيح

Wholesale price index

رقم قياسي لأسعار الجملة

Within groups

داخل المجموعات

X

X-Axis

محور السينات

Y

Y-Axis

محور الصادات

Z

Z-axis

محور العميات

Z-test

اختبار « ز »



المراجع العربية والأجنبية

٢ - المراجع العربية :

- 1 - د. صباح بقجة جي - مبادئ الإحصاء (الأمالى الجامعية) 1967 .
- 2 - د. صباح بقجة جي - تحليل الارتباط (الأمالى الجامعة) 1978 .
- 3 - د. ناظم حيدر - المدخل لدراسة الإحصاء (الكتب الجامعية) 1964 .
- 4 - د. ناظم حيدر - الوسيط فى الإحصاء التطبيقي (الكتب الجامعية) 1977 .
- 5 - د. ناظم حيدر - مبادئ الإحصاء (الكتب الجامعية) 1982-1983 .
- 6 - د. عبد القادر الأفندي - نظرية الإحصاء (الكتب الجامعية) 1975-1976 .
- 7 - د. إبراهيم محمد العلى - مدخل فى نظرية الارتباط (الكتب الجامعية) 1979-1980 .
- 8 - د. محمد مظلوم حمدي - طرق الإحصاء (دار المعارف - مصر) 1965 .
- 9 - أنيس كنجو - الإحصاء وتطبيقاته (جزآن) مؤسسة الرسالة - دمشق 1977 .
- 10 - د. محمد عادل العاقل - مبادئ الإحصاء، الجزء الأول، مديرية الكتب والمطبوعات الجامعية - الطبعة الثانية - حلب 1969 .
- 11 - د. مختار محمود الهانسي - مقدمة الطرق الإحصائية - دار النهضة العربية - بيروت 1982 .

- 12- د. مختار محمود الهانسي - مبادئ الإحصاء - بيروت - لبنان 1993 .
- 13- د. عبد العزيز فهمي هيكل - مبادئ في الإحصاء التطبيقي - جامعة بيروت العربية - بيروت 1986 .
- 14- د. عبد الله رمضان الكندري - مبادئ الإحصاء وأساليب التحليل الإحصائي - ذات السلاسل - الطبعة الأولى - الكويت 1985 .
- 15- د. عبد العزيز فهمي هيكل - مبادئ الإحصاء - جامعة الاسكندرية وبيروت العربية 1985 .
د. امثال محمد حسن
- 16- د. سمير كامل عاشور - مبادئ الإحصاء الوصفي والتحليلي . القاهرة - معهد الإحصاء 1977 .
- 17- د. مدني دسوقي مصطفى - مبادئ في علم الإحصاء . القاهرة: دار النهضة العربية 1979 .
- 18- د. جلال الصياد وآخرون - الإحصاء لطلاب الدراسات الاقتصادية والإدارية - الطبعة الأولى - جدة - المملكة العربية السعودية 1989 .
- 19- د. اعتماد علام - أساسيات الإحصاء الاجتماعي - الدوحة - قطر 1995 .
- 20- عبد الحميد عبد المجيد البلداوي - الإحصاء للعلوم الإدارية والتطبيقية - عمان - الأردن 1997 .
- 21- د. جلال مصطفى الصياد - مقدمة في الطرق الإحصائية - الرياض - المملكة العربية السعودية 1999 .
- 22- د. منير غانم - مبادئ الإحصاء - كتاب جامعي - جامعة حلب 1983 .

- 1 - Allen L. Edwards - «Statistical Methods», Second Edition, Holt, Rinehart & Winston, Inc, New York, 1967 .
- 2 - B. C. Ericker - «Advanced General Statistics». English Universities Press, London, 1972.
- 3 - George W. Sendecor - «Statistical Methods», Sixth edition, the Iowa & William Cochran State University Press, Iowa, U. S. A. 1974.
- 4 - Helen M. Welber, - «Elementary statistical Methods», John Wiley & Sons, Inc, New York 1966.
- 5 - N. R. Draper - «Applied Regression Analysis», John Wiley & Sons, Inc, New York 1966.
- 6 - William C. Guenther - «Concepts of Statistical Inference», Mc-Graw-Hill Book Company, U.S.A. 1973.
- 7 - Taro Yamane - «Statistics», (An Introductory Analysis) second Ed., U.S.A. 1970.
- 8 - Wilfrid J. Dixon - «Introduction to Statistical Analysis» Frank J. Massey, Jr. McGRAW - Hill Book Company 1969.
- 9 - Lincoln L. Chao - Statistics for management 1984 .

- 1 - А. П. Зинченко, В.К. Горкавый, Е. А. Гуртовник. Т. Ф. Хромова, И. Д. Политова. «статистика» учеб. Пособие - М. : Финсы и статистика, 1982 .
- 2 - Пасхавер И. С. Яблочник А. Л - « общая теория статистики » М. Финансы и статистика. 1983.
- 3 - Т. И. Козлов, В. Е. Овсиенко, В. И. Смирнский - « Курс общей Теории статистики » Москва 1965 .
- 4 - А. Я. Боярско. Г. Л. Громыко - « Общая Теория статистики » Московского университета 1985.
- 5 - Гмурман В.Е. - « Теория Вероятностей И Математическая статистики » М. « Выш. Школа », 1977.
- 6 - Б. В. Гнеденко. Г. Л. Громыко - « Курс Теории Вероятностей » Москва 1988 .
- 7 - Н. Дрейпер. Г. Смит - « Прикладной регрессионный Анализ » Книга 1, книга 2, М. «финансы и статистика» 1986.
- 8 - Н. И. Щедрин - Н. Н. Егоров - « Статистика торговли » - М. финансы и статистика, 1987 .
- 9 - И.К. Беляевский. Н. Н. Ряузов. Д. Н. Ряузов. - « Статистика торговли » Москва « финансы и статистики » 1989 .

- 10- А. М. Гольдъерга – В. С Козлова - « Общая Теория статистики »
Москва « финансы и статистики » 1985 .
- 11- Н. Н. Рязов - « Общая Теория статистики » Москва « финансы и
статистики » 1984 .
- 12- Ю. В. Прохоров - Ю. А. Розанов - « Теория Вероятностей »
Москва « Наука » 1987.
- 13- Г. Л. Громыко - « Статистика » издательство Московского
университета 1981.



• اللجنة العلمية :

- أ. د. ناظم حيدر .
- أ. د. صباح الدين بقجة جي .
- أ. م. د. خلف مطر الجراد .

• المدقق اللغوي :

- الدكتور ماجد أبو ماضي .

حقوق المطبع والترجمة والنشر محفوظة لمديرية الكتب والمطبوعات

