الإحصاء الوصفي


وقل رب زدني علما
جامعة دمشق
Damascus University

منشورات جامعة دمشق
كلية العلوم السياسية

الإحصاء الوصفي

الدكتور
نبيل علي
أستاذ مساعد في قسم الاقتصاد والإدارة العامة

الدكتور
عدنان غانم
أستاذ مساعد في قسم الإحصاء التطبيقي

جامعة دمشق
١٤٣٥ - ١٤٣٦ هـ
٢٠١٣ - ٢٠١٤ م


وقل رب زدني علما
جامعة دمشق
Damascus University

الفهرس

الموضوع	الصفحة
مقدمة	١٣
الفصل الأول	١٧
أهمية علم الإحصاء	
١ ــ ١ : مقدمة	١٩
١ ــ ٢ : تطور مفهوم الإحصاء	١٩
١ ــ ٣ : تعريف علم الإحصاء	٢٠
١ ــ ٤ : النماذج وطرق القياس والأدوات الإحصائية	٢٠
١ ــ ٥ : دور الإحصاء في التخطيط	٢١
١ ــ ٦ : مجالات استخدام علم الإحصاء :	٢١
ــ العلوم الاجتماعية	٢١
ــ العلوم الاقتصادية	٢٢
ــ التربية وعلم النفس	٢٢
ــ العلوم الطبيعية والحيوية	٢٢
ــ الصناعة والعلوم الهندسية	٢٢
١ ــ٧ : خطوات البحث الإحصائي	٢٢
ــ مثال تطبيقي عن مراحل البحث الإحصائي	٢٣
ــ شكل توضيحي لأهم مراحل البحث الإحصائي	٢٦

الفصل الثاني ٢٧

جمع البيانات الإحصائية

٢ ــ ١ : مقدمة ٢٩

٢ ــ ٢ : مصادر البيانات الإحصائية ٢٩

ــ المصادر الميدانية ٢٩

ــ المصادر التاريخية ٢٩

٢ ــ ٣ : أساليب جمع البيانات الإحصائية ٣٠

ــ أسلوب الحصر الشامل ٣١

ــ أسلوب العينة ٣١

٢ ــ ٤ : مميزات أسلوب العينة ٣٢

٢ ــ ٥ : عشوائية العينة ٣٣

ــ السحب مع الإعادة ٣٣

ــ السحب بدون إعادة ٣٣

٢ ــ ٦ : الطرق الاحتمالية لاختيار العينات والطرق التقديرية لاختيار العينات ٣٤

ــ (المعاينة العشوائية البسيطة، المنتظمة ، الطبقية ، العنقودية، المتعددة المراحل، المتدرجة ، المساحية ، الحصص، العمدية) ٣٥

٢ ــ ٧ : طرق جمع البيانات: ٤٠

ــ الطريقة المباشرة ٤٠

ــ الطريقة غير مباشرة ٤٠

٢ ــ ٨ : مكونات الاستمارة الإحصائية (الاستبيان) ٤١

الفصل الثالث ٤٣

أساليب عرض البيانات الإحصائية وتبويبها

٣ ــ ١ : العرض الجدولي: ٤٥

ــ تصنيف جغرافي ، كمي ، تاريخي ٤٦

ــ القواعد التي يجب مراعاتها عند إعداد الجداول الإحصائية ٤٧

٣ ــ ٢ : أنواع الجداول الإحصائية ٤٨

ــ الجدول البسيط، الجدول المركب ٤٨

٣ ــ ٣ : العرض البياني ٤٨

ــ الأشكال المعبرة، الرسوم الدائرية، الأعمدة البيانية، الخطوط البيانية ٤٩

٣ ــ ٤ : الجداول التكرارية ٥٤

ــ خطوات إنشاء الجدول التكراري ٥٥

٣ ــ ٥ : عرض التوزيعات التكرارية: ٥٨

ــ المدرج التكراري، المضلع التكراري، المنحني التكراري ٥٨

ــ المنحني التكراري المجتمع الصاعد، المنحني التكراري المجتمع الهابط ٦١

ــ أسئلة وتمارين ٦٥

الفصل الرابع ٧١

مقاييس النزعة المركزية

٤ ــ ١ : مقدمة ٧٣

٤ ــ ٢ : الوسط الحسابي ٧٣

ــ حساب الوسط الحسابي لبيانات غير مبوبة (الأسلوب ٧٤

المباشر، أسلوب الانحرافات)
ـــ حساب الوسط الحسابي لبيانات مبوبة (الأسلوب : المباشر، أسلوب الانحرافات، أسلوب الانحرافات المختصرة) ٧٦
٤ ـــ ٣ : الخصائص الرياضية والجبرية للوسط الحسابي : ٨٠
٤ ـــ ٤ : الوسيط ٨١
ـــ حساب الوسيط لبيانات غير مبوبة ٨١
ـــ حساب الوسيط لبيانات مبوبة ٨٢
٤ ـــ ٥ : تحديد الوسيط بيانياً ٨٥
٤ـــ ٦ : الربيعيات ٨٦
ـــ حساب الربيعيات لبيانات غير مبوبة ٨٧
ـــ حساب الربيعيات لبيانات مبوبة ٨٧
٤ ـــ ٧ : المنوال: ٩١
ـــ حساب المنوال لبيانات غير مبوبة ٩١
ـــ حساب المنوال لبيانات مبوبة (طريقة الفروق، طريقة العزوم) ٩٣
٤ ـــ ٨ : تحديد المنوال بيانياً ٩٦
٤ـــ ٩ : خصائص المنوال ٩٧
٤ ـــ ١٠ : العلاقة التقريبية بين المتوسطات ٩٨
٤ ـــ ١١ : الوسط الهندسي: ١٠٠
ـــ حساب الوسط الهندسي لبيانات غير مبوبة ١٠٠
ـــ حساب الوسط الهندسي لبيانات مبوبة ١٠٢
٤ ـــ ١٢ : خصائص الوسط الهندسي ١٠٣

الفصل الخامس ١٠٥

مقاييس التشتت ومقاييس الالتواء والتفلطح

٥ ــ ١ : مقدمة ١٠٧

٥ ــ ٢ : أهمية مقاييس التشتت (التشتت المطلق، التشتت النسبي) ١٠٧

٥ ــ ٣ : المدى : ١١١

ــ حساب المدى لبيانات غير مبوبة ١١١

ــ حساب المدى لبيانات مبوبة ١١١

٥ ــ ٤ : خصائص المدى ١١١

٥ ــ ٥ : الانحراف الربيعي: ١١٢

ــ حساب الانحراف الربيعي لبيانات غير مبوبة ١١٢

ــ حساب الانحراف الربيعي لبيانات مبوبة ١١٣

٥ ــ ٦ : خصائص الانحراف الربيعي ١١٤

٥ ــ ٧ : الانحراف المتوسط : ١١٤

ــ حساب الانحراف المتوسط لبيانات غير مبوبة ١١٤

ــ حساب الانحراف المتوسط لبيانات مبوبة ١١٥

٥ ــ ٨ : خصائص الانحراف المتوسط ١١٦

٥ ــ ٩ : الانحراف المعياري : ١١٧

ــ حساب الانحراف المعياري لبيانات غير مبوبة (الأسلوب المباشر، أسلوب الانحرافات) ١١٧

ــ حساب الانحراف المعياري لبيانات مبوبة (الأسلوب المباشر، أسلوب الانحرافات، أسلوب الانحرافات المختصرة). ١٢٠

٥ ــ ١٠ : الانحراف المعياري النسبي (معامل الاختلاف) ١٢٣
٥ ــ ١١ : خصائص الانحراف المعياري ١٢٤
٥ ــ ١٢ : أهمية الانحراف المعياري ١٢٥
٥ ــ ١٣ : الالتواء (معامل بيرسون للالتواء) ١٣٤
٥ ــ ١٤ : التفلطح (البيانات غير المبوبة، البيانات المبوبة) ١٣٦
أسئلة وتمارين ١٣٧

الفصل السادس ١٤٥

الارتباط والانحدار

٦ ــ ١ : مقدمة ١٤٧
٦ ــ ٢ : تحليل الارتباط: ١٤٨
ــ تحديد طبيعة المتغيرات ١٤٨
ــ رسم شكل الانتشار ١٤٩
٦ ــ ٣ : مقاييس الارتباط وأنواعه : ١٥١
ــ أسلوب الرسم الحر ١٥١
ــ أسلوب المربعات الصغرى ١٥٣
ــ خواص خط الانحدار ١٥٦
ــ استعمالات معادلة الانحدار ١٥٦
ــ رسم خط الانحدار ١٥٧
٦ ــ ٤ ــ معادلة الانحدار الخطي (الرسم الحر، المربعات الصغرى، الأسلوب المباشر). ١٥٨
٦ ــ ٥ : تحليل التباينات : ١٦١
- التباين غير المفسر ١٦٢

- التباين المفسر ١٦٢
معامل الارتباط بيرسون ١٦٣
٦ ــ ٦ : معامل التحديد ١٦٤
ــ معامل ارتباط الرتب – سبيرمان ١٦٦
٦ ــ ٧ : اختبار معنوية معامل الارتباط ١٦٧
- اختبار معامل الارتباط ١٦٨
ــ تمارين ١٧١
الفصل السابع ١٧٧
السلاسل الزمنية
٧ ــ ١ : تعريف السلسلة الزمنية ١٧٩
٧ ــ ٢ : مجالات استخدام السلاسل الزمنية (الاقتصادية، ١٧٩
الفيزيائية، للتسويق، السكانية، لمراقبة الإنتاج، الثنائية، النقطية)
٧ ــ ٣ : العوامل المؤثرة على السلسلة الزمنية ١٨٠
ــ تغيرات الاتجاه العام ١٨٠
ــ التغيرات الموسمية (الفصلية ، الشهرية) ١٨١
ــ التغيرات الدورية ١٨١
ــ التغيرات الفجائية أو العشوائية ١٨٢
٧ ــ ٤ : طرائق تحليل الاتجاه العام (طريقة الرسم الحر، ١٨٤
طريقة المتوسطات النصفية، طريقة المتوسطات المتحركة،
طريقة المربعات الصغرى)
٧ ــ ٥ : التذبذب في السلسلة الزمنية ٢٠١
٧ ــ ٦ : التغيرات الموسمية (الفصلية ــ الشهرية) ٢٠٢

٧ ــ ٧ : التغيرات الموسمية الشهرية ٢٠٦
٧ ــ ٨ : التغيرات الدورية والتغيرات العشوائية ٢١٧
ــ تمارين ٢٠٩
الفصل الثامن ٢٢٣
الأرقام القياسية
٨ ــ ١ : تعريف الأرقام القياسية ٢٢٥
٨ ــ ٢ : استخدامات الأرقام القياسية ٢٢٥
٨ ــ ٣ ــ الصيغ البسيطة للأرقام القياسية ٢٢٧
٨ ــ ٤ : الأرقام القياسية التجميعية البسيطة ٢٣٠
٨ ــ ٥ : الصيغ المرجحة للأرقام القياسية ٢٣٣
ــ الرقم القياسي الأمثل ٢٣٥
٨ ــ ٦ : الاختبارات المطبقة على الأرقام القياسية ٢٣٩
ــ اختبار الانعكاس في الزمن ٢٣٩
ــ اختبار الانعكاس المعاملي ٢٤١
٨ ــ ٧ : بعض الأرقام القياسية الهامة ٢٤٤
ــ الرقم القياسي لنفقة المعيشة ٢٤٥
ــ الرقم القياسي لكمية الصادرات والواردات ٢٤٦
ــ الرقم القياسي للقوة الشرائية للنقود ٢٤٦
ــ تمارين ٢٤٨
ــ المصطلحات الإحصائية ٢٥٣
ــ الجداول الإحصائية ٢٧١
ــ المصادر والمراجع العربية والأجنبية . ٢٧٩

مقدمة

إن عبارة " الإحصاء " تعني لغوياً كماً من المعطيات العددية. أما الإحصاء كمصطلح أكاديمي فيعني علم تجميع المعطيات العددية وتفسيرها، فعلم الإحصاء يهتم بتوفير المقاييس الكمية للظواهر الاقتصادية والدراسات الاجتماعية واستطلاع الرأي العام وغيرها من الظواهر حتى أصبح التفكير الإحصائي ضرورة ملحة لكل فرد، ومع تطور العلم الحديث تغيرت وظيفة الإحصائي واتخذ علم الإحصاء معنى آخر حتى أصبح علم اتخاذ القرارات.

لم يشذ علم السياسة عن بقية العلوم الاجتماعية الأخرى التي تستخدم الرياضيات في دراساتها المختلفة. فلغة الأرقام اقتحمت مجالات علم السياسة منذ فترة من الزمن على الرغم من أن هناك من انتقد الأرقام أو إدخال الرياضيات إلى حقل الدراسات السياسية غير أن الواقع وقف إلى جانب أنصار استخدام الرياضيات والإحصاء. وهكذا وجدنا علماء السياسة يستخدمون التوابع الرياضية والاحتمالات والإحصاءات في أبحاثهم المتعددة ولم يعد الإحصاء حكراً على الرياضيين وأصبح العديد من الباحثين لديهم المقدرة على تطبيق الإحصاء في المجالات التي يعملون بها، فالباحث السياسي اليوم يستخدم العديد من الموضوعات التي يمكن أن تخضع للقياس والتعبير عنها كمياً كالأوضاع في المجتمع الدولي الأكثر احتمالاً أو الأكثر ملائمة لإثارة الصراعات المسلحة؟ ودراسة العلاقة بين مستوى الدخل ومستوى المشاركة السياسية والعلاقة بين مستوى الدخل والانتماء الحزبي؟ وأثر وسائل الإعلام في السلوك السياسي؟ ودراسة العلاقة بين الوضع الداخلي والخارجي وتفاعلهما ؟.

يفتقد علم السياسة الإطار النظري الرياضي المنظور الذي يساعد على تفسير أغراض وأنشطة هذا المجال التخصصي وشرحها، لذا يعمد كثير

من المهتمين بالسياسة على تطوير النظريات والقوانين المتعلقة بسلوكيات هذا المجال التخصصي سواءً من حيث الكم فيها أم من حيث الكيف والاستفادة من تطور مناهج البحث العلمي في التخصصات الأخرى وذات الصلة بها.

لقد بات معروفاً، بأن تدريس الإحصاء لم يعد مقتصراً على تخصص بعينه، بل أصبح يشمل كل التخصصات، اقتصادية كانت، أم اجتماعية، ... الخ، لأن الإحصاء أصبح بالنسبة لهذه التخصصات الطابع المميز للمنهج العلمي المتبع فيها تخصص السياسة خاصة الذي أصبح بأمس الحاجة إلى الإحصاء لاستخدامه في مجال البحوث التي تؤدي إلى معارف جديدة، وفي تحليل البيانات التي تفسر الظواهر المعلوماتية لأن الباحث السياسي هو المستخدم لجميع أنواع البيانات التي يستفيد من دراستها في اتخاذ القرارات العقلانية .

نأمل في فصول هذا الكتاب، أن نكون قد قدمنا مفهوماً واضحاً لعلم الإحصاء والتحليل الإحصائي ومدى فائدتهما لتخصص السياسة، إذ تجلى هدفنا في تزويد طلاب العلوم السياسية والمهتمين في هذا التخصص بالمفاهيم الإحصائية الأكثر استخداماً والتي تفيدهم في مجال تخصصهم.

يتكون هذا الكتاب من ثمانية فصول، توزعت على النحو التالي :

الفصل الأول : أهمية علم الإحصاء

الفصل الثاني : جمع البيانات الإحصائية

الفصل الثالث : أساليب عرض البيانات الإحصائية وتبويبها

الفصل الرابع : مقاييس النزعة المركزية

الفصل الخامس : مقاييس التشتت

الفصل السادس : الارتباط والانحدار

الفصل السابع : السلاسل الزمنية

الفصل الثامن : الأرقام القياسية

وقد ضمّنا كل فصل من الفصول السابقة عدداً من التمارين المحلولة وختم كل منها بمجموعة من التمارين غير المحلولة لإتاحة الفرصة للطالب كي يختبر مدى تفهمه وإدراكه للمعلومات المعطاة وألحقت في نهاية الكتاب الجداول الإحصائية اللازمة، بالإضافة إلى قائمة بالمراجع العربية والأجنبية والمصطلحات باللغة العربية والأجنبية .

إن إعداد هذا الكتاب، هو إعداد جماعي ومسؤولية مشتركة، لكل من الدكتور عدنان غانم والدكتور نبيل علي .

إننا نضع هذا الكتاب بين أيديكم، نأمل منكم جميعاً طلاباً وباحثين أن يتفضلوا بتقديم أي نقد بنّاء يسهم في إغنائه وإخراجه مستقبلاً بشكل نجعل منه أكثر نفعاً وفائدة، وبما يسمح لنا في متابعة مسيرة البحث العلمي، وأخيراً نقدم جزيل الشكر والتقدير لكل من أسهم في إخراج هذا الكتاب والله ولي التوفيق .

والله وليُّ التوفيق.....

دمشق في ٢٠٠٦ ***المؤلفان***


وقل رب زدني علما
Damascus University

الفصل الأول

أهمية علم الإحصاء

- تطور مفهوم الإحصاء
- تعريف علم الإحصاء
- البحث الإحصائي
- النماذج وطرق القياس والأدوات الإحصائية
- دور الإحصاء في التخطيط
- مجالات استخدام علم الإحصاء
- خطوات البحث الإحصائي

Damascus University


وقل رب زدني علما
جامعة دمشق
Damascus University

١ ــ ١ : مقدمة :

أصبح الإحصاء بمفهومه العلمي الواسع، أداة للحصول على البيانات الإحصائية وتحليلها كما أصبح، يحتل مكاناً هاماً في مضمار العلم والمعرفة ويحظى باهتمام جميع الدول في رسم سياساتها ووضع خططها التنموية وذلك لوضع اقتصادياتها على المسار الذي يحقق الأهداف من هذه الخطط .

تطور في استخدامات الإحصاء ساعد على إدخال تغييرات جذرية على العملية الإنتاجية والإدارية وأصبح الأداة التي لا غنى عنها لأغلب العلوم في تفسير الظواهر الاجتماعية والتربوية والحيوية والاقتصادية وبناء التوقعات المستقبلية، وربط الظواهر بعلاقاتها الجدلية كمياً، واتخاذ القرارات العقلانية لتحقيق الأهداف المعنية.

١ ــ ٢ : تطور الإحصاء :

عُرف الإحصاء منذ أقدم العصور فقد عرفه المصريون القدماء، والرومان والصينيون منذ (٢٠) ق.م وكان الهدف منه معرفة المقدار العددي للقوة البشرية لمعرفة مقدرة الدولة على الدفاع ومعرفة مقدار المحاصيل الزراعية.

وقد اهتم المسلمون بعدد السكان منذ البدايات الأولى للإسلام لسببين: أحدهما معرفة أعداد المسلمين وأماكن تواجدهم. وثانيهما معرفة عدد سكان البلاد المفتوحة ومواردها ومساحة أراضيها وخصائص سكانها. ومع نهاية القرن التاسع عشر وبداية القرن العشرين، أصبح الإحصاء أداة عملية تشمل جميع المبادئ والأسس، في جمع وعرض وتحليل البيانات، وقد تطور الإحصاء في هذه المرحلة نتيجة لتطور نظرية الاحتمال، وفي الوقت الحاضر أصبح الإحصاء علم اتخاذ القرارات، وعلم كشف الحقائق باعتباره

يهتم بتوفير المقاييس الكمية للظواهر الاقتصادية والاجتماعية التي تمهد الطريق في تحليل هذه الظواهر.

١ ــ ٣ : تعريف علم الإحصاء :

هناك الكثير من التعاريف الشائعة والمألوفة للإحصاء وأن معظم هذه التعاريف تركز إلى ما يهدف إليه الإحصاء من أنه : أداة علمية تتناول الأسس والطرائق العلمية في جمع، وتبويب، وعرض، وتحليل البيانات العددية وتفسيرها، ومن ثم اتخاذ القرارات العقلانية في ضوء النتائج المستخلصة .

١ ــ ٤ : النماذج وطرق القياس في الأدوات الإحصائية :

تنبع أهمية الإحصاء في هذا المجال من استخدام أدواته في مختلف المجالات الاقتصادية والاجتماعية والطبيعية للتأكد من ظاهرة معينة، أو نفيها، أو قياسها أو عرض بعض النتائج في صورة مؤشرات تلخيصية وصفية أو استدلالية عامة، فكثيراً ما نقول، إن متوسط إنتاج النفط الخام، والغاز المسال، وأن متوسط الدخل في مجتمع ما كذا، ومعدل هطول الأمطار في مدينة ما كذا، وبذلك فقد عُبر عن ظاهرة إنتاج النفط أو الدخل في مجتمع ما، وهطول الأمطار، بمتوسط حسابي وحين نقول إن الارتباط بين مستوى الدخل ومستوى التعليم قوي فقد لخصنا العلاقة بين المتغيرات بمؤشر تلخيص عام يصف طبيعة العلاقة بينهما من حيث وجودها أو عدم وجودها ، وطبيعتها، وشدتها، وعندما نقول إن فرضية ما عن المجتمع صحيحة بدرجة ثقة معينة، نكون بذلك قد استخدمنا مفهوم الإحصاء الاستدلالي للتوصل إلى هذه النتيجة وتعميمها على المجتمع من خلال بيانات العينة.

١ ــ ٥ : دور الإحصاء في التخطيط :

يشكل الإحصاء أساساً يعتمد عليه، في رسم أية خطة مستقبلية تحددها السياسة العليا للدولة، لتحقيق أهداف معينة في مختلف المجالات والأنشطة، ومن المعروف أنه لا يمكن أن توضع أية خطة دون توفر البيانات الإحصائية وكلما كانت البيانات دقيقة ومفصلة كلما كانت الخطة أسلم وأقرب إلى الواقع وقد لا تفي البيانات المتوفرة أحياناً تلبية متطلبات الخطة لذلك يتم اللجوء إلى تنفيذ أبحاث خاصة بالعينة تخدم هذه الأهداف، وبما أن العملية الإحصائية عملية فنية وتخصصية لذلك فقد أكدت الدول على أحداث الأجهزة الإحصائية وربطها برئاسة مجلس الوزراء أو مكاتب الرئاسة نظراً لأهميتها ودورها المهم الذي تقوم به. وعلى هذا نجد أن معظم الدول في العالم تعتمد على الإحصاء اعتماداً تاماً وتدعمه مالياً وفنياً وبشرياً.

١ ــ ٦: مجالات استخدام علم الإحصاء :

العلوم الاجتماعية وتشمل :

تعداد السكان بهدف التعرف على خصائصهم وتركيباتهم المختلفة العمرية والنوعية والاقتصادي والتعليمي والزواجي والتوزيع الجغرافي .

ب ــ دراسة حركة السكان : (ولادات، وفيات، هجرة، زواج .. إلخ).

ج ــ دراسة الاتجاهات والآراء في المجتمع .

العلوم الاقتصادية وتشمل :

آ ــ دراسة التجارة الخارجية لمعرفة السلع المستوردة ومعرفة الأسواق التي تصدر إليها السلع المحلية .

ب ــ دراسة التقلبات في أسعار السلع في الأسواق الداخلية والخارجية.

ج ــ إحصاءات القوى العاملة والتوزيع النوعي للعاملين .

د ــ إحصاءات الإنتاج الزراعي والصناعي.

هـ ــ الحسابات القومية وتشمل الدخل القومي وتقسيماته واستخداماته في الاستهلاك أو الاستثمار.

التربية وعلم النفس وتشمل :

آ ــ دراسة القدرات العقلية للطلاب.

ب ــ دراسة النتائج التربوية لنظم وطرق التعليم المختلفة .

العلوم الطبيعية والحيوية وتشمل :

ــ دراسة التغيرات الجوية لمعرفة مدى تأثيرها في الحالة الصحية للسكان وعلى الإنتاج الزراعي والصناعي.

الصناعة والعلوم الهندسية وتشمل :

آ ــ تحديد مستويات التخزين .

ب ــ مراقبة جودة الإنتاج .

١ ــ ٧ : خطوات الدراسة الإحصائية :

تمر الدراسة الإحصائية بالخطوات أو المراحل الآتية :

١ ــ تحديد المشكلة والهدف من البحث.

٢ ــ تحديد تكلفة البحث.

٣ ــ جمع المشاهدات أو البيانات عن الظواهر محل الدراسة .

٤ ــ تحليل البيانات واستخلاص النتائج .

٥ ــ وضع الفروض المختلفة عن طبيعة هذه الظواهر أو المشكلات.

٦ ــ اختبار صحة الفروض .

٧ ــ دراسة احتمالات تعميم هذه النتائج التي تصل إليها عند اختبار الفروض.

من خلال ذلك نجد أن الإحصاء يخدم هدفين أساسيين :

- توفير الأسس العلمية لجمع البيانات وتبويبها وحساب المقاييس التلخيصية والوصفية منها.
- تحليل البيانات التي تم جمعها، بهدف بناء تقديرات إحصائية لبعض الظواهر المدروسة وتفسيرها واستخلاص الاستنتاجات وبناء على جزء أو (عينة) من المجتمع للتوصل إلى قرارات يمكن تعميمها على كامل أجزاء المجتمع، وهو ما يعرف بالإحصاء الاستدلالي.

- **مثال تطبيقي عن مراحل البحث الإحصائي :**

لنفرض أننا أمام ظاهرة تزايد عدد حوادث المرور في مدينة ما ونريد اتخاذ قرار بشأن الحلول الممكنة للحد من هذه الظواهر نظراً لآثارها السلبية على المجتمع، وتحليل أسباب الحوادث وملابساتها. مثل : (عدم التقيد بإرشادات المرور، السرعة، التجاوز غير القانوني، السائق تحت تأثير مخدر.. إلخ)، تتبع الخطوات التالية :

١ ــ تحديد المشكلة وهدفها : والظاهرة أو المشكلة هنا تزايد عدد حوادث المرور بالمقارنة مع السنوات الماضية مثلاً والهدف هو الحد من تزايد حدوث المرور.

٢ ــ الوحدة الإحصائية والمجتمع الإحصائي : إن الوحدة الإحصائية هنا في مثالنا سائق السيارة، والمجتمع هو مجتمع السائقين (مالكي السيارات).

٣ ــ تحديد نوع البيانات المطلوب جمعها : ويكون نوع البيانات المطلوب جمعها في هذه الحالة عدد الحوادث بالنسبة لكل نوع من أنواع السيارات (خاصة ، أجرة، شاحنات،..) وكذلك لكل نوع من أنواع الحوادث (تصادم سيارتين، تصادم مع جسم ثابت، تصادم سيارة بأحد المارة) ويمكن التقسيم أيضاً بالنسبة لآثار الحادث (إصابة سطحية، إصابة خطرة، إصابة قاتلة) ويمكن أيضاً (اصطدام ، دهس، انقلاب).

٤ ــ تحديد مصادر جمع البيانات : يمكن جمع البيانات في هذه الحالة بأسلوبين:

الأسلوب الأول : من سجلات إدارة المرور بالمدينة (المصادر الوثائقية أو التاريخية).

الأسلوب الثاني : عن طريق عمل استقصاء ميداني للسائقين لمختلف أنواع السيارات للسؤال عن الرأي في نظام المرور، والإشارات الضوئية، ومدى سلامة الطرق، (المصادر الميدانية).

٥ ــ تحديد كيفية جمع البيانات : بالنسبة للأسلوب الأول يتم الاتصال المباشر بإدارة المرور بالمدينة والحصول على البيانات المطلوبة كافة، أما بالنسبة للأسلوب الثاني فيتم عمل مجموعة مدربة من العدادين وتوزيعهم في مناطق مختلفة من المدينة لجمع البيانات ميدانياً بواسطة استمارات وجداول إحصائية أعدت خصيصاً لهذا الغرض .

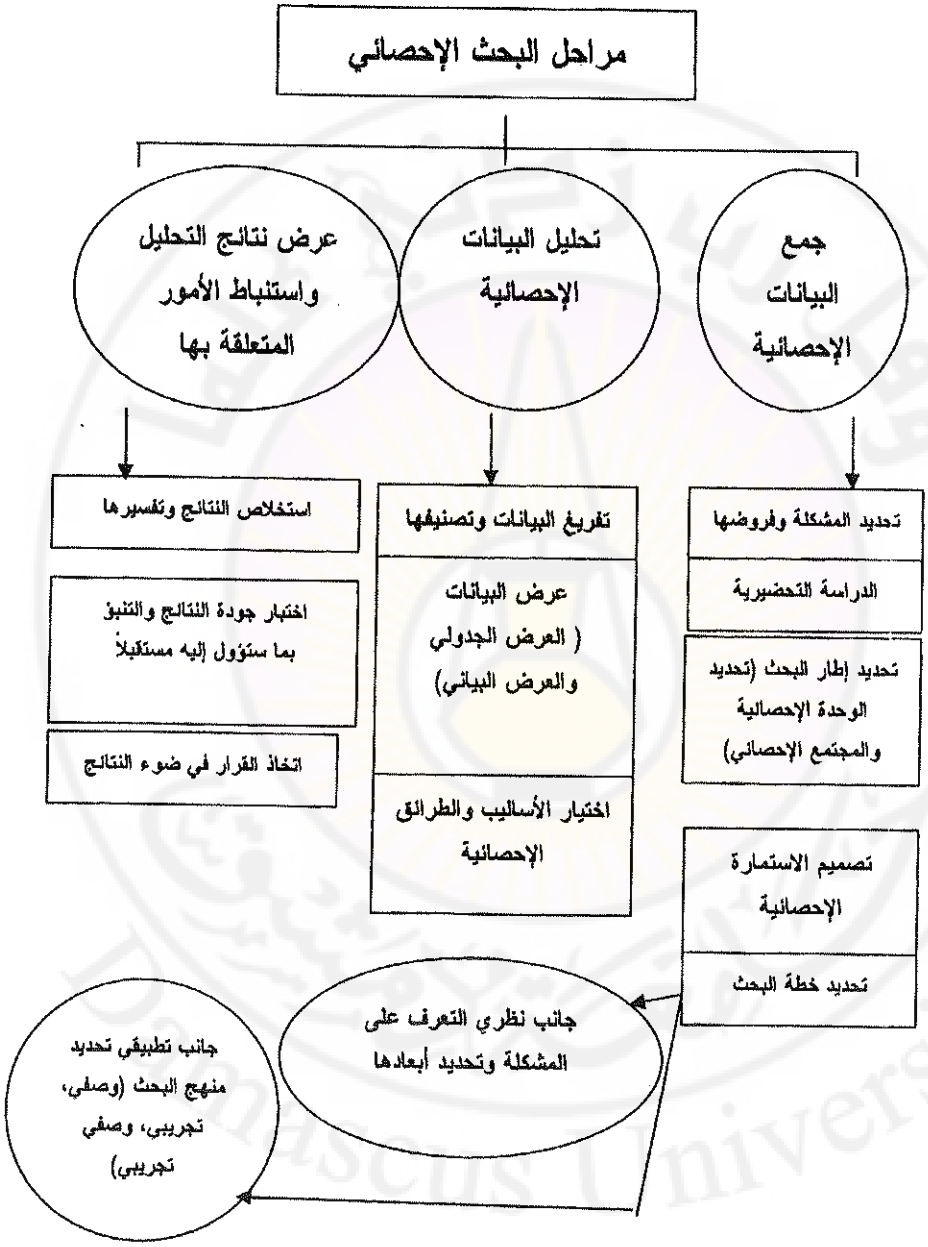
٦ ــ تحديد كلفة البحث : من خلال مكافأة القائمين على جمع البيانات من الميدان وتكلفة المطبوعات .

٧ ــ **تحديد الفترة الزمنية للدراسة : حيث يتم تحديد هذه الفترة بحد أقصى ثلاثة شهور مثلاً نظراً للأهمية في سرعة الإنجاز**، للوصول لجميع الجوانب الخاصة بالمشكلة، واتخاذ القرارات المناسبة في ضوء نتائج الدراسة .

٨ ــ استخدام المقاييس الإحصائية المناسبة : من أجل الحصول على المؤشرات التلخيصية والوصفية العامة، لجميع عناصر المشكلة (مثلاً : متوسط عدد حوادث المرور في منطقة معينة علاقة الارتباط بين عمر السائق وعدد الحوادث ، كثافة المرور على طريق معينة...إلخ).

٩ ــ اتخاذ القرار: بعد الحصول على البيانات وتحليلها إحصائياً يتم اتخاذ القرار والتوصيات المناسبة الكفيلة بالحد من هذه الظاهرة.

شكل توضيحي لأهم مراحل البحث الإحصائي



الفصل الثاني

جمع البيانات الإحصائية

- مصادر البيانات الإحصائية
- أساليب جمع البيانات
- مميزات أسلوب العينة
- عشوائية العينة
- الطرق الاحتمالية لاختيار العينات والطرق التقديرية لاختيار العينات
- طرق جمع البيانات
- أنواع أخطاء العينة

Damascus University

وقل رب زدني علما
جامعة دمشق
Damascus University

٢ ــ ١ : مقدمة :

تعد عملية جمع البيانات الإحصائية أهم مرحلة من مراحل الدراسة الإحصائية لأنها تعتبر بمثابة الأساس الذي تبنى عليه باقي مراحل الدراسة الإحصائية للظاهرة المدروسة، فليس لنتائج الدراسة أية قيمة إذا كانت البيانات التي قمنا بجمعها خاطئة أو متحيزة مهما استخدمنا من أساليب راقية في التحليل، وتبدو أهمية هذه المرحلة من خلال: التكلفة والزمن اللذين تتطلبهما هذه المرحلة.

٢ ــ ٢ : مصادر البيانات الإحصائية :

يمكن أن نميز بين نوعين أساسيين من مصادر البيانات الإحصائية :

١ ــ المصادر الميدانية (المباشرة) :

يتم جمع البيانات بهذا الأسلوب من الوحدات المدروسة مباشرة، فمثلاً يتم جمع البيانات الخاصة بالدخل والإنفاق، مباشرة من الأسر بواسطة استمارات وجداول إحصائية تعد لهذا الغرض.

وتمتاز عملية جمع البيانات بالطريقة المباشرة بإمكانية مراجعتها والتحقق من دقتها مع مصدر البيان نفسه، ويعاب عليها كثرة الجهد والتكاليف.

٢ ــ المصادر التاريخية (غير المباشرة) :

يتم الحصول على البيانات بهذا الأسلوب من السجلات، التي تقوم هيئات رسمية أو شبه رسمية بإصدارها بصفة دورية، فتقوم بحكم اختصاصها بجمع البيانات عن الظواهر المختلفة ووضعها في جداول عامة، يمكن الرجوع إليها للحصول على البيانات أو المعلومات الإحصائية دون الرجوع إلى الوحدات الأصلية ومن أمثلتها :

- سجلات السكان، والتي تستعمل في الأغراض الإدارية أو المدنية، ومصدرها الجهاز المركزي للإحصاء بالدولة، أو الإدارات المختصة بهذا الشأن كإدارة التسجيل المدني للمواطنين، بوزارة الداخلية، وكذلك إدارات السجل العام للموظفين في الوزارات والمؤسسات وسجلات الطلبة في المؤسسات التعليمية المختلفة.
- السجلات الحيوية التي تستوفي وتحفظ لأغراض قومية، مثل سجلات المواليد والوفيات والزواج والطلاق، وكذلك السجلات الخاصة بالقدوم والمغادرة بشكل مؤقت أو دائم .
- السجلات المالية الخاصة بالدخل، والرسوم الجمركية، والقروض، والودائع، التي تقوم باستيفائها الجهات الرسمية الحكومية، مثل وزارة المالية والتخطيط والبنوك المركزية.
- سجلات الأمن العام، والتي تضم بيانات عن الحوادث والجرائم وتقوم على استيفائها إدارات وزارة الداخلية كما تشمل أيضاً أية إحصاءات أو سجلات خاصة تعدها جهات معينة ليست معدة لغرض النشر.

وتمتاز طريقة اللجوء إلى المصادر التاريخية، بالسهولة، والتوفير في الجهد والتكاليف والوقت بينما يعاب عليها عدم إمكانية التحقق من صحة البيانات .

٢ ـ ٣ : أساليب جمع البيانات الإحصائية :

إذا ما تقرر جمع البيانات الإحصائية من مصادرها الأصلية أو المباشرة أي بالطريقة الميدانية فإن عملية جمع البيانات يمكن أن تتم بأحد الأسلوبين الأساسيين وهما :

أ ــ أسلوب الحصر الشامل : Complete Census

حيث يتم الحصول على البيانات المطلوبة، من جميع أفراد المجتمع، موضع البحث، دون استثناء، وفي هذه الحالة يجب دراسة مدى توفر جميع الإمكانيات اللازمة لجمع البيانات من كل وحدة من وحدات المجتمع، ويتبع هذا الأسلوب في حالات خاصة، تستدعي جمع البيانات عن ظاهرة أو مجموعة ظواهر على المستوى القومي، كما هو الحال في التعدادات العامة للسكان والمساكن التي تقوم بها الحكومات كل خمس أو عشر سنوات، كما يتبع هذا الأسلوب في حالة ما إذا كان مجتمع البحث صغيراً، ورغم أن جميع البيانات باتباع أسلوب الحصر الشامل يمتاز بتوفر معلومات عن جميع وحدات المجتمع محل الدراسة، وبالتالي تكون الدراسة شاملة، إلا أن الدراسة بهذا الأسلوب لها بعض العيوب منها:

١ ــ إن جمع البيانات وفق هذا الأسلوب يحتاج إلى إمكانيات مادية وبشرية كبيرة نسبياً قد لا يمكن توفرها في بعض الأحيان .

٢ ــ إن عملية الجمع تحتاج وقتاً طويلاً وبالتالي، تأخر صدور النتائج، الأمر الذي قد يفقدها أهميتها أو يعرضها للتقادم.

٣ ــ تتعرض عملية جمع البيانات وفق هذا الأسلوب إلى أخطاء ذات أسباب متعددة، كعدم الشمول لجميع وحدات المجتمع، مما يستدعي ضرورة فرض رقابة شديدة على عملية جمع البيانات .

٤ ــ الحاجة إلى الكوادر المتخصصة وارتفاع التكلفة .

ب ــ اسلوب العينة Sampling Approsh :

كلما ازداد عدد وحدات المجتمع محل الدراسة، ازدادت عملية جمع البيانات تعقيداً، خصوصاً إذا كانت وحدات المجتمع منتشرة في مساحة جغرافية واسعة يصعب الوصول إليها، هذا بالإضافة إلى التكاليف الباهظة،

وتدريب الكادر والزمن الكبير لمرحلتي التحضير والتنفيذ، عدا ذلك هناك بعض الدراسات التي لا يمكن إجراؤها بهذا الأسلوب كدراسات الجودة ومراقبة المواصفات، كل ذلك أدى إلى التفكير بأسلوب علمي آخر، وتعميم النتائج على المجتمع وفق قواعد علمية.

ويسمى الأسلوب الذي يتم بموجبه الحكم على المجتمع الكلي، عن طريق بحث جزء منه أسلوب العينة Smapling ولضمان دقة النتائج التي نحصل عليها باتباع الأسلوب، يجب أن تمثل العينة المجتمع الأصلي الذي سحبت منه اصدق تمثيل ، علماً أن حجم العينة يعتمد، على مدى تجانس المجتمع، لأنه كلما كانت وحدات المجتمع موضوع البحث متجانسة، كلما أمكن الاكتفاء بعدد قليل من الوحدات، أي الاكتفاء بعينة صغيرة.

٢ ــ ٤ ــ مميزات أسلوب العينة :

١ ــ توفير الجهد، والوقت، والتكاليف، لأن الدراسة تقتصر، على بحث جزء من المجتمع الأصلي وليس كل المجتمع.

٢ ــ زيادة دقة النتائج نظراً لإمكان فرض الرقابة وأحكام العمل أثناء عملية جمع البيانات بالمقارنة بأسلوب الحصر الشامل.

٣ ــ إمكانية جمع بيانات أكثر تفصيلاً، عن الظاهرة المدروسة نظراً لصغر حجم عينة البحث بالمقارنة مع حجم المجتمع.

٤ ــ إظهار نتائج البحث بسرعة، وخلال فترة قصيرة نسبياً.

٥ ــ يستخدم أسلوب العينة في الحالات التي يتعذر فيها استخدام أسلوب الحصر الشامل، أو في الحالات التي يكون فيها إجراء البحث باستخدام الحصر الشامل غير منطقي مثل فحص عينات الدم.

٢ ــ ٥ : عشوائية العينة :

هي طريقة لاختبار (n) وحدة من مجتمع حجمه (N) وبشكل عشوائي وتعني العشوائية أن يكون لكل وحدة من وحدات المجتمع نفس الفرصة أو نفس الاحتمال في أن تؤخذ ضمن العينة الممكن سحبها، أي أن تتساوى فرص الدخول في العينة أمام كل مفردات المجتمع، ويمكن الحصول على عينة عشوائية بسيطة **بعدة طرق أهمها :**

الطريقة الأولى : وتتلخص بترقيم مفردات المجتمع بشكل متسلسل، وتكتب هذه الأرقام على قصاصات من الورق، ثم توضع على كرات متشابهة في إناء ما، وتخلط جيداً وتتم بعد ذلك عملية السحب لمفردات العينة وحدة فوحدة، ونميز هنا بين نوعين من عمليات السحب:

أ ــ السحب مع الإعادة : وفيها تعاد كل وحدة مسحوبة إلى الإناء ، قبل إجراء عملية السحب التالية، وإن احتمال سحب كل وحدة من وحدات هو (1/N) في جميع السحبات.

ب ــ السحب بدون إعادة : نسحب الكرة الأولى من الإناء فتكون هي المفردة الأولى في العينة، واحتمالها $\frac{n}{N}$ ولا تعاد الكرة إلى الإناء مرة ثانية، وإن احتمال سحب المفردة الثانية هو $\frac{n-1}{N-1}$ وهكذا إلى المفردة الأخيرة.

الطريقة الثانية :

طريقة جداول الأرقام العشوائية، ولتقريب هذه الفكرة إلى الأذهان نأخذ المثال التالي:

لنفرض أن حجم المجتمع N = 628 وحدة ونريد أن نسحب منه عينة n=10 باستخدام جدول الأرقام العشوائية .

نختار أي عمود من الأعمدة العشرة الموجودة في الجدول (أنظر جدول الأرقام العشوائية في النهاية) ولنفرض أننا اخترنا العمود رقم (٢) ونختار منه الأرقام الثلاثة الأولى ولنستعرض هذه الأرقام من العمود (٢) ونختار منه الأرقام الثلاثة الأولى ولنستعرض هذه الأرقام ن العمود (٢) ونختار الأعداد المتميزة العشرة الأولى الواقعة بين 628 -001 فتكون : 134, 331, 462, 093, 210, 307, 382, 004, 426, 303 وهي عبارة عن مفردات العينة.

وعموماً عندما نعمد إلى أسلوب العينة يكون الهدف منها الاستدلال على الصفات أو الميزات الرئيسية للمجتمع الإحصائي الذي سحبت منه والوصول إلى قرارات لا يمكن الوصول إليها عادة إلا بطريقة التعداد أو الحصر الشامل، وبصورة عامة يمكن تصنيف طرق اختيار العينات إلى قسمين :

٢ ــ ٦ : الطرق الاحتمالية والطرق التقديرية لاختيار العينات:

في الطرق الاحتمالية لاختيار العينات، يكون هناك لكل مفردة من المفردات في المجتمع، فرصة معلومة ومتساوية مع بقية المفردات في الدخول ضمن العينة وهذا ما يطلق عليه بالعينات العشوائية، أما في الطرق التقديرية لاختيار العينات، فإن الباحث نفسه يقوم بالاختيار بشكل تحكمي في تحديد المفردات التي ستدخل العينة من مفردات المجتمع وهذا ما يسمى بالعينات غير العشوائية التي لا يمكن الاعتماد عليها في تعميم النتائج إلا بتحفظ، بعكس النوع الأول من العينات التي تعتمد أساساً على نظرية الاحتمالات سواءً في أسلوب المعاينة أو طريقة تعميم النتائج وتقدير الأخطاء. وفيما يلي نشرح الطرق الرئيسية للمعاينة الاحتمالية والمعاينة التقديرية :

المعاينة العشوائية البسيطة :

تعد هذه المعاينة، النوع الأساسي للمعاينة الاحتمالية، وهي شائعة الاستعمال وسهلة الاستخدام من وجهة النظر الإحصائية، وتعتبر أيضاً الأساس لعدد من أساليب المعاينة الأكثر تعقيداً .

والصفة المميزة للمعاينة العشوائية البسيطة، هي أن تتساوى فرص الدخول في العينة أمام كل مفردات المجتمع ويتم الحصول على العينة من خلال سحب عدد من المفردات الواحدة بعد الأخرى دون إعادتها، ويعد أفضل طريقة لاختيار مفردات العينة هي استخدام جداول الأرقام العشوائية أو بواسطة الحاسب الآلي عن طريق استخدام برامج العينة.

المعاينة المنتظمة :

تعتمد هذه الطريقة على استعمال قائمة تتضمن مفـردات المجتمـع وتحديد فترة السحب $K = \frac{N}{n}$ فعلى سبيل المثال لو أردنا اختيار عينة مؤلفة من (١٠٠) شخص من قائمة بها (١٠٠٠) اسم نقوم بتحديد فترة السحب $K = \frac{1000}{100} = 10$ واختيار رقم عشوائي يقع بين (0 ,10) ولنفترض أنه (٦) فيكون الاسم ذو الترتيب السادس في القائمة هو الفرد الأول المختار في العينة وإن باقي أفراد العينة هم (... ,6,16,26,36) ومـع ذلـك يجـب الحذر عند استخدام هذه الطريقة لاسيما عند وجود علاقة دورية في ترتيـب العناصر في القائمة وكان طول فترة السحب بين عناصر العينة مساوياً طول الدورة أو من مضاعفاتها.

المعاينة الطبقية :

تستخدم المعاينة، الطبقية عندما يكون المجتمع الإحصائي غير متجانساً ومؤلف من طبقات متجانسة فيما بينها، ويتم تقسيم المجتمع إلى

طبقات بحيث تمثل العينة كل طبقات المجتمع، وعندما يكون هناك تناسب بين عدد المفردات المأخوذة من كل طبقة، مع عدد المفردات في الطبقة تسمى هذه الحالة معاينة طبقية متناسبة مع الحجم، ويفضل عادة تقسيم المجتمع إلى طبقات على أساس عوامل متصلة بموضوع الدراسة، مثال ذلك تقسيم العاملين في شركة ما إلى (مهرة، نصف مهرة، غير مهرة) أو تقسيم مجتمع العمال حسب الدخل (مرتفع، متوسط، منخفض).

المعاينة العنقودية :

وتقوم هذه المعاينة على تقسيم المجتمع إلى مجموعات جزئية والاختيار يكون على شكل (عناقيد) وليس مفردة. ففي سبيل المثال لو كان لدينا في إحدى المدن (١٠٠٠٠) مسكن مسجلة وطلب منا اختيار عينة من (٥٠٠) مسكن ففي هذه الحالة من الأنسب تقسيم المدينة إلى ١٠٠٠ منطقة بكل منطقة (١٠) مساكن ومن ثم نختار عشوائياً ٥٠ منطقة من المناطق الألف ونجمع المعلومات اللازمة عن كل مسكن من المساكن بالمناطق المختارة وهذا يؤدي إلى تخفيض التكلفة والزمن اللازم. بينما لو تم الاختيار للمساكن لـ (٥٠٠) بالطريقة العشوائية البسيطة سيكون جمع المعلومات صعباً وباهظ التكاليف على اعتبار أن هذه المساكن ستكون غالباً مبعثرة على مساحة المدينة كلها يعتبر هذا النوع من المعاينة أقل تكلفة من أنواع المعاينات الأخرى.

المعاينة المتعددة المراحل :

تستخدم هذه المعاينة عندما يكون المجتمع كبيراً ومنتشراً على مساحات واسعة وبالتالي صعوبة الوصول مباشرة إلى الوحدات الإحصائية التي سيتم دراستها، عند دراسة دخل ونفقات الأسر على سبيل المثال نختار عينة عشوائية من المحافظات، ثم نقسم كل محافظة من المحافظات المختارة

إلى عدد من المراكز التي نختار منها العينات عشوائياً، ويمكن بعد ذلك تقسيم المركز إلى وحدات أصغر ونختار العينات منها عشوائياً وتستمر العملية على عدد من المراحل في كل مرحلة نختار عشوائياً المفردات المطلوبة إلى أن نصل إلى عدد من الأسر المتجانسة .

المعاينة المتدرجة :

توجد طريقتان لسحب العينة المتدرجة هما :

ـ **الطريقة الأولى :** تتضمن هذه الطريقة مرحلتين أساسيتين هما :

١ ـ تقسيم مجتمع الدراسة إلى فئات مجتمعية تمثل صفاته الرئيسة والفرعية.

٢ ـ سحب عينة عشوائية أو منتظمة من كل فئة على أن لا يتكرر تمثيل نفسها المفردة في أكثر من فئة .

ـ **الطريقة الثانية :** تقوم هذه الطريقة على سحب العينة المتدرجة حسب نسبة مئوية ثابتة يحددها الباحث بنفسه ويستخدم هذه العينة عندما يكون المجتمع الإحصائي غير متجانس في صفاته ولا في حجومه العددية. ويعمد الباحث إلى تقسيم مجتمع الدراسة إلى صفاته الأساسية، وحسب حجومها العددية ويسحب عينة من كل صفة و يحدد نسبة مئوية ثابتة يأخذها من جميع الفئات التي تكون المجتمع الإحصائي .

المعاينة المساحية :

تعتبر العينة المساحية ذات أهمية كبيرة عند الحصول على عينات تمثل المناطق الجغرافية المختلفة. كما لا يتطلب في هذه الحالة إعداد قوائم كاملة، بجميع الأفراد أو العناصر، داخل منطقة جغرافية معينة. ويتم اختيار المناطق الجغرافية نفسها بطريقة عشوائية، ولكن يجب أن تمثل في كل منطقة إقليمية مختارة، كل الفئات المتمايزة.

ويبدأ الباحث بتقسيم المجتمع الإحصائي إلى وحدات أولية، يختار من بينها عينة بطريقة عشوائية، أو منتظمة، ثم تقسم الوحدات الأولية المختارة إلى وحدات ثانوية، نختار من بينها عينة جديدة، ثم يتم تقسيم الوحدات الثانوية المختارة إلى وحدات ثلاثية، ثم وحدات رباعية، إلى أن يقف الباحث عند مرحلة معينة. واختيار الأشخاص الذين يمكن أن تتم معهم المقابلة ،وتمتاز هذه العينة في أنها : تغطي مساحة جغرافية واسعة من منطقة الدراسة، وتختصر وقت إجراء الدراسة، وتقلل من حركة الباحث أو الباحثين داخل ميدان دراستهم، وتغني الباحث عن متاعب الحصول على قائمة الأسماء وحتى من المشكلات القائمة نفسها .

ونشير هنا إلى أن طرق المعاينة التقديرية تعتمد على اختيار المفردات بطريقة شخصية منها :

معاينة الحصص :

تعتمد هذه الطريقة في المعاينة، على اختيار المفردات، من الطبقات بشكل غير عشوائي واعتماد نسبة مئوية معينة من المفردات من كل طبقة وتستخدم على نطاق واسع في كل من بحوث التسويق واستطلاع الرأي العام لكن على الرغم من أن هذه المعاينة غير مرضية، ونتائجها غير دقيقة فإن كثيراً من باحثي السوق والإداريين يؤيدون هذه الطريقة لما تمتاز به من بساطة وقلة التكاليف والسرعة في الحصول على النتائج .

المعاينة العمدية (المنتفاة) :

إن الباحث في هذه المعاينة هو الذي يحدد المفردات التي ستدخل في العينة وجودة النتائج تتوقف على حكمة الباحث ومهارته في الاختيار مع ذلك تعطي هذه المعاينة نتائج دقيقة عندما الأمر يتعلق بمتوسط ظاهرة ما.

وفي ختام مناقشتنا لبعض أنواع العينات ينبغي القول بأنه إذا كانت طريقة المعاينة ضرورة من ضرورات البحث العلمي، فلا بد من اختيارها اختياراً ممثلاً للمجتمع الإحصائي الأصلي الذي أخذت منه. كما أن كثيراً من عمليات المعاينة تلجأ إلى أكثر من طريقة واحدة من هذه العينات، أو إلى حل وسط مثل العينة المساحية، والعينة الطبقية مثلاً. كما يجب على الباحث التنبه إلى مواقع الخطأ في اختيار العينة وأهمها :

١ ــ **أخطاء التحيز** : وهي الأخطاء الناجمة عن الطريقة التي نختار بها العينة من المجتمع الأصلي للظاهرة .

٢ ــ **أخطاء حجم العينة** : وهي الأخطاء الناجمة عن حجم العينة وتسمى **بأخطاء الصدفة.**

٣ ــ **أخطاء ردود فعل المستجوبين** : وهي الأخطاء الناجمة عن ردود فعل الناس نحو أداة أو وسيلة القياس ذاتها وتسمى أخطاء الأداة .

أنواع أخطاء العينة :

ويعرف خطأ عينة ما، بأنه الفرق بين التقدير الذي نحصل عليه من العينة، وقيمة المعلمة، أو ثابت المجتمع، الذي نحصل عليها بالحصر الشامل أو العد الكامل. وتقسم أخطاء العينة إلى قسمين :

- الأخطاء العمدية أو القصدية أو أخطاء التحيز : وتنشأ هذه الأخطاء من إعطاء معلومات غير صحيحة. فعند سؤال رب الأسرة عن دخله الشهري فإنه يعطي رقماً يختلف عن الرقم الحقيقي خوفاً من ضريبة الدخل. وقد نحصل في بعض الحالات على البيانات من أشخاص غير الأشخاص المستجوبين وقد يهملون استمارة البحث مدّعين أنهم كانوا خارج مكان توزيعها فضلاً عن عدم التمسك التام بأصول العينة العشوائية .

• الأخطاء العشوائية : وهي الأخطاء الناجمة عن العينة نفسها وطريقة اختيارها، ويصعب التحكم فيها وإن كان بالإمكان التقليل منها، بعكس الأخطاء القصدية، التي يمكن التحكم فيها وتقليلها إلى أقل حد ممكن، من خلال الالتزام بقواعد العينة العشوائية والدقة في جمع البيانات. ويعرف الخطأ العشوائي بأنه الفرق بين القيمة التي نحصل عليها من العينة العشوائية والقيمة المتوقعة أو معلمة مجتمع الدراسة .

٢ ــ ٧ : طرق جمع البيانات :

يمكن جمع البيانات والمعلومات بإحدى الطريقتين الرئيسيتين :

١ ــ الطريقة المباشرة :

وتتم عن طريق الاتصال المباشر بمصدر البيان، أو وحدة البحث، وذلك بالاتصال الشخصي بين الباحث أو مندوبه والمصدر وقد يتم الاتصال المباشر بالتليفون، أو الفاكس، كما يشمل أسلوب الملاحظة أو المعاينة لمصدر البيان، وتسجيل البيانات مباشرة، وتمتاز الطريقة المباشرة بإمكانية الحصول على بيانات أكثر دقة.

٢ ــ الطريقة غير المباشرة :

حيث يتم جمع البيانات أو المعلومات من المصدر بطريقة غير مباشرة مثال ذلك، عن طريق صحيفة أو استمارة الاستبيان المرسلة إليه بالبريد Postal Questionnaire. كما تعد عملية جمع البيانات من المصادر التاريخية أو الإحصاءات المنشورة Abstraction from published statistic إحدى الطرق غير المباشرة في جمع البيانات.

ويؤخذ على الطريقة غير المباشرة انخفاض نسبة الاستجابة، علاوة على عدم إمكان مراجعتها مع مصدر البيان للتحقق من دقة البيانات، ومن أهم ضمانات

النجاح لعملية جمع البيانات مستوى جامع البيانات، لذا يجب اختياره من ذوي الكفاءة والاهتمام بتدريبه.

وبشكل عام فإنه يتم تجهيز قائمة بالأسئلة موضع البحث تسمى صحيفة الاستبيان أو الاستمارة الإحصائية .

٢ ــ ٨ــ مكونات الاستمارة الإحصائية (الاستبيان) :

الاستبيان :

عبارة عن كشف يتضمن الأسئلة المطروحة حول المشكلة أو الظاهرة المدروسة.

تتكون الاستمارة من أربعة أجزاء رئيسية هي :

الجزء الأول :

ويعرف باسم صدر الاستمارة وفيه يذكر اسم الجهة المشرفة على البحث، كما يتضمن ملخصاً للأهداف المرجوة من البحث، مع الإشارة في مكان ظاهر من الاستمارة إلى البيانات التي يدلي بها معطي البيان أو يدونها بالسرية التامة وإنها تستخدم لأغراض البحث فقط.

الجزء الثاني :

ويشمل البيانات المميزة لوحدة البحث، فإذا كان البحث يخص الأسرة فيذكر اسم رب الأسرة، والعنوان الدائم، وغير ذلك من البيانات الأخرى المميزة للوحدة.

الجزء الثالث :

وهو الجزء المهم في الاستمارة، ويتضمن مجموعة الأسئلة والاستفسارات المطلوب الإجابة عليها.

الجزء الرابع :

ويشمل ملاحظات جامع البيان إن وجدت، كما قد يتضمن بعض التعليمات والإشارات الخاصة بكيفية استيفاء الاستمارة وغالباً ما يكتفى بالأجزاء الثلاثة الأولى فقط مع وضع التعليمات والإرشادات في كتيب يسلم مع الاستمارة.

هذا ويجب الأخذ بالحسبان الشروط التالية عند وضع الاستمارة الإحصائية:

١ ــ يجب أن تكون الأسئلة محددة وواضحة، ولا تحتمل الإجابات الإنشائية ويستحسن أن تدون الإجابة بنعم أو لا.

٢ ــ يجب تحديد وحدات القياس بدقة فمثلاً في حالة السؤال عن العمر فيجب أن يوضح المطلوب هل هو العمر بالسنوات الكاملة وإهمال الأشهر أم العمر بالسنوات والأشهر والأيام. وكذلك في حالة السؤال عن المرتب يجب أن يوضح هل هو أسبوعياً أم شهرياً أم سنوياً.

٣ ــ التسلسل المنطقي للأسئلة بحيث تكون مفتاحاً للأسئلة التي تليها.

٤ ــ يجب أن يكون عدد الأسئلة قليلاً ما أمكن، بحيث لا توجد أسئلة كثيرة ينبرم المبحوث الإجابة عنها .

٥ ــ تجنب الأسئلة التي تضر بمصالح المستجوب وتثير حرجه.

٦ ــ تجنب الأسئلة الإيحائية التي تؤدي إلى إجابات كيفية.

٧ ــ نظراً للتوسع في استخدام الحاسب الآلي في تبويب البيانات وتحليلها فإنه من المستحسن ترميز الأسئلة والإجابات المحتملة.

الفصل الثالث

أساليب عرض

البيانات الإحصائية وتبويبها

- العرض الجدولي
- أنواع الجداول الإحصائية
- العرض البياني
- الجداول التكرارية
- عرض التوزيعات التكرارية بيانياً

Damascus University


وقل رب زدني علما
جامعة دمشق
Damascus University

١ ــ ٣ : العرض الجدولي :

بعد عملية جمع البيانات سواءً من مصادرها الأصلية، أم استخدام الاستمارات أم من المصادر التاريخية، يكون من الصعب وحتى من المستحيل في معظم الأحيان فهم اتجاهها، أو مدلولها أو الخروج بأية استنتاجات عنها لذلك يقوم الباحث بتبويب تلك البيانات وعرضها في جداول إحصائية هادفة، يمكن إخضاعها للتحليل وتموين فكرة عن طبيعة الظاهرة ومكوناتها واتجاهها وتعتبر هذه الجداول أساساً للتحليل الإحصائي.

وتتلخص طريقة التبويب في تكوين جداول تفريغ، مكونة من ثلاثة أعمدة ، يخصص الأول منها لبيان القيم والصفة والعمود الثاني لوضع الإشارات (|||) نتيجة فرز الاستمارات، واحدة تلو الأخرى والعمود الثالث يشير إلى التكرارات، ولتوضيح ذلك، نأخذ المثال التالي:

مثال : سبحت عينة من (50) ناخب من الناخبين في إحدى الدوائر الانتخابية لمعرفة الحالة التعليمية فكانت النتائج التالية : أمي، أمي، إعدادية، يقرأ ويكتب، إعدادية، أمي ، ابتدائية، يقرأ ويكتب، أمي، إعدادية، ابتدائية و.... هكذا .

المطلوب : تفريغ هذه البيانات في جدول إحصائي .

جدول توزيع الناخبين حسب الحالة التعليمية

الحالة التعليمية	الإشارات	التكرار
أمي	ـــ ـــ ـــ ـــ ـــ	15
يقرأ ويكتب	ـــ ـــ ـــ ـــ ـــ ـــ ـــ ـــ ـــ ـــ	20
ابتدائية	ـــ ـــ ـــ |||	8
إعدادية	ـــ ـــ ـــ ||	7
المجموع	50	50

المصدر : فرضي.

ونظراً لاختلاف طبيعة البيانات المجمعة ونوعيتها، فقد تكون البيانات مرتبطة بالزمن، أو تكون موزعة حسب صفات أخرى غير زمنية، كأن تكون جغرافية، أو صفات تتعلق بخاصية من خواص المجتمع المبحوث، مثل الحالة الزواجية، أو التعليمية أو المهنية ... إلخ. وتسمى العملية التي يتم فيها تجميع البيانات في مجموعات مميزة ومتجانسة عملية التصنيف وبوجه عام تصنيف البيانات الإحصائية وفق مايلي :

١ ــ تصنيف جغرافي :

وعلى أساس هذا التصنيف تجمع الوحدات التي تشترك في صفة مكانية واحدة، في مجموعة مستقلة، والجدول أدناه يبين مثل هذا التصنيف:

مثال : جدول يبين التوزيع النسبي للصادرات السورية حسب الكتل الدولية خلال عام ٢٠٠٤:

الكتل الدولية	البلدان العربية	الاتحاد الأوروبي	بلدان آسيوية مختلفة	بلدان أخرى
الصادرات	29.9	3.2	10.4	0.7

المصدر: الجهاز المركزي للإحصاء ٢٠٠٥ .

٢ ــ تصنيف كمي :

وفي هذا النوع من التصنيفات تجمع الوحدات المشتركة في درجة الصفة المعينة والتي تأخذ شكلاً رقمياً محدداً في مجموعة واحدة، مثال ذلك توزيع الطلبة حسب فئات الدرجات في أحد الاختبارات، في إحدى الكليات :

مثال : جدول يبين توزيع الدرجات وعدد الطلاب في إحدى الكليات :

فئات الدرجات	أقل من 20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
عدد الطلاب	8	20	25	43	18	15	4	1

المصدر : فرضي.

٣ ــ تصنيف تاريخي أو زمني :

وفيه تجمع الوحدات المرتبطة بزمن معين في مجموعة واحـــدة، أي أن تقسيم البيانات يتم حسب زمن الحدوث وتسمى بيانات ذلك النـــوع باســـم السلاسل الزمنية. فمثلاً إذا أردنا دراسة نمو الناتج المحلي الإجمالي خـــلال خمس سنوات في الجمهورية العربية السورية (٢٠٠٠ ــ ٢٠٠٤) فإننا ندرج المعلومات في جدول ابتداء من عام ٢٠٠٠ م وانتهاء بعـــام ٢٠٠٤م وهـــذا الترتيب هو الترتيب السائد في عرض السلاسل الزمنية .

مثال : جدول يبين نمو الناتج المحلي الإجمالي في ســـورية خـــلال الفتـــرة (٢٠٠٠ / ٢٠٠٤).

السنوات	2000	2001	2002	2003	2004
الأسعار الجارية%	10.4	5.3	6.6	4.4	9.5
الأسعار الثابتة%	0.6	3.6	4.1	1.3	1.7

المصدر: الجهاز المركزي للإحصاء ٢٠٠٥ .

أهم القواعد التي يجب مراعاتها عند إعداد الجداول الإحصائية :

١ ــ كتابة عنوان الجدول بشكل واضح يبين محتواه .

٢ ــ كتابة مصدر المعلومات في الأسفل (وزارة النفط، إدارة المرور، جمع شخصي).

٣ ــ ترقيم الأعمدة لسهولة الإشارة إليها .

٤ ــ كتابة رقم للجدول لتسهيل الرجوع إليه عند الحاجة.

٥ ــ تحديد الفترة الزمنية التي تعود إلها البيانات في متن الجدول.

٣ ــ ٢ : أنواع الجداول الإحصائية :

الجدول البسيط : وهو الجدول الذي يشمل بياناً واحداً فقط، فهو يحتوي على عمود واحد للبيان موضوع الدراسة، يقابله عمود آخر للتكرارات.

مثال (١) :

جدول يبين أهم المؤشرات الديموغرافية في الجمهورية العربية السورية في عام ٢٠٠٤

المؤشر	معدل الخصوبة الكلية	وفيات الأطفال الرضع بالألف	نسبة القوة العاملة %	نسبة البطالة %
القيمة	3.58	17.1	29	10.4

المصدر: الجهاز المركزي للإحصاء ٢٠٠٥ .

الجدول المركب : ويتكون من أكثر من عمود واحد للبيان.

مثال (٢) :

جدول يبين قطاع السياحة في الجمهورية العربية السورية لعام (٢٠٠٠/٢٠٠٤).

السنوات	2000	2004
عدد الفنادق	466	523
عدد الأسرة	34209	39985

المصدر: الجهاز المركزي للإحصاء ٢٠٠٥ .

٣ ــ ٣ : العرض البياني :

إن الهدف الأساسي لعلم الإحصاء، هو جمع وعرض البيانات الإحصائية عن ظاهرة ما بأسلوب يسهل معه إدراك اتجاه الظاهرة، وبالتالي فإن العرض

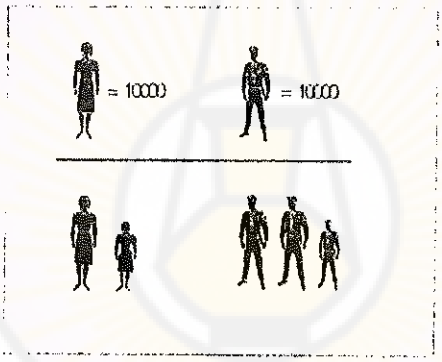
البياني يعتبر أقرب لذهن القارئ وأسهل للتذكر من العرض الجدولي ويعتبر العرض الجدولي أساساً للعرض البياني، وذلك بإعادة عرض محتويات الجداول التي يتم بناؤها، على شكل رسوم بيانية وهندسية، لتكون بمتناول مختلف المستويات والاختصاصات.

وسنتناول بعض من الرسوم والصور والأشكال المعبرة الشائعة الاستخدام .

الأشكال المعبرة: ولتوضيح هذا الأسلوب من العرض البياني لدينا المثالان التاليان :

مثال ١ : بلغ عدد طلاب المرحلة الأساسية فى إحدى السنوات كما يلي ٢٥١٥٠ ذكراً و ١٥٢١١ أنثى. وضح باستخدام الرسم البياني :

يمكن استخدام الرسم البياني التالي لتوضيح هذه البيانات .



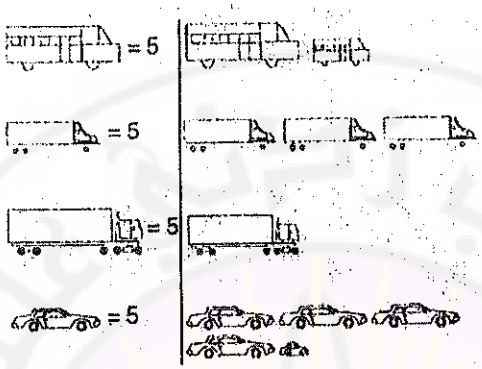
شكل (٣ ــ ١)

في الرسم (ذكر أو أنثى) يمثل ١٠٠٠٠ شخص

مثال ٢ :

أرادات إحدى المؤسسات أن ترصد ما لديها من العربات فوجدت أن لديها سبع حافلات وخمس عشرة ناقلة صغيرة وثلاث ناقلات كبيرة وإحدى وعشرين سيارة ركاب. وضح باستخدام الرسم البياني ما لهذه المؤسسة من عربات :

يمكن استخدام الرسم التالي لتوضيح عدد العربات من كل نوع بحيث يمثل كل شكل عدد خمس عربات .



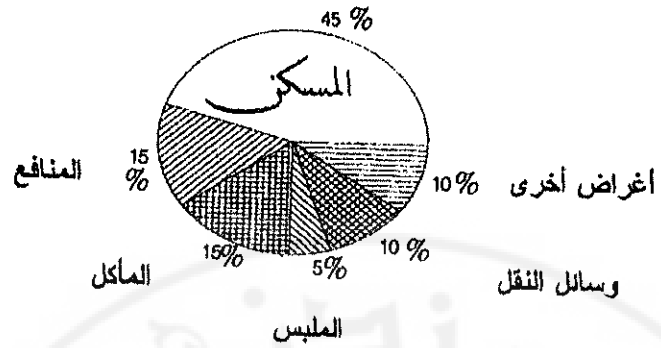
شكل (٣ ــ ٢)

٢ ــ الرسوم الدائرية :

تستخدم الرسوم الدائرية لمقارنة الأعداد المتجانسة كتقسيم ميزانية معينة على عدد من بنودها بنسب معينة كما يوضح ذلك المثال التالي :

مثال ٣ : في دراسة اقتصادية اجتماعية لمجموعة من الأسر المنتجة متوسطة الدخل وجد أن الأسر في المتوسط تنفق حوالي ٤٥% من دخلها على المسكن و ١٥% للمنافع (كالماء والكهرباء والغاز) و ١٥% للمأكل و٥% للملبس و ١٠% لوسائل النقل و ١٠% لأغراض أخرى. وضح تقسيم نفقات هذه الأسر باستخدام الرسوم الدائرية.

عند رسم أقسام هذه النفقات نقوم برسم دائرة بحسب المساحة المتاحة ثم نقسمها بنفس نسب تقسيمات النفقات كما يلي :



شكل (٣ ــ ٣)

ويتم حساب طول القوس (القطاع) من خلال العلاقة التالية :

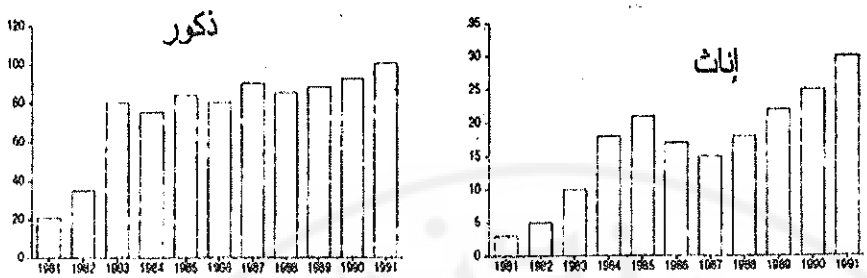
$$\text{القطاع} = \frac{\text{حجم الظاهرة}}{\text{المجموع الكلي}} \times 360^{\circ}$$

٣ ــ الأعمدة البيانية :

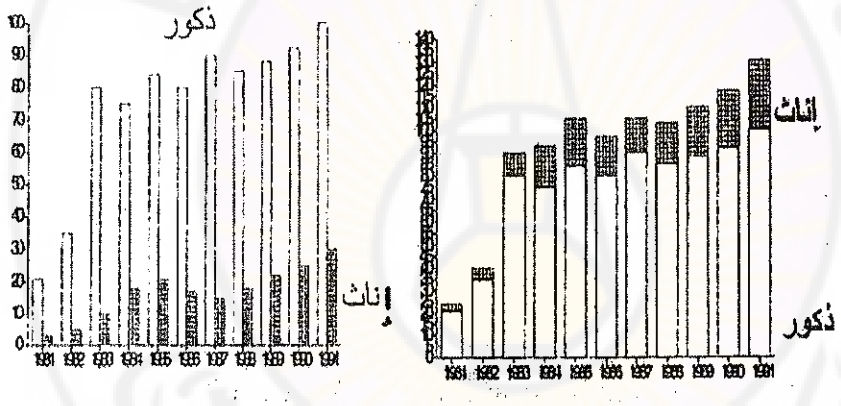
تستخدم الأعمدة البيانية لمقارنة الأعداد المتجانسة أو لتوضيح تطور الأعداد عبر الزمان أو المكان. كما في المثال التالي :

مثال ٤ : الجدول التالي يوضح تطور أعداد الطلاب الملتحقين بإحدى الكليات لجامعة ما من عام ١٩٨١ وحتى عام ١٩٩١، وضح ذلك باستخدام أعمدة بيانية مستقلة ومشتركة ؟.

العام	81	82	83	84	85	86	87	88	89	90	91
ذكور	21	35	80	75	84	80	90	85	88	92	100
اناث	3	5	10	18	21	17	15	18	22	25	30



شكل (٤ ــ ٣)



شكل (٦ ــ ٣)

شكل (٥ ــ ٣)

من خلال الأعمدة البيانية أعلاه نجد أن الشكل (٤ ــ ٣) يعبر عن الأعمدة البيانية البسيطة والتي تستخدم لتمثيل بيانات إحصائية لفترة زمنية واحدة أو عدة فترات أما الشكل (٥ ــ ٣) (٦ ــ ٣) يعبر عن الأعمدة البيانية المزدوجة أو المتلاصقة والتي تستخدم عند المقارنة بين ظاهرتين أو أكثر لعدة سنوات .

٤ ــ الخطوط البيانية :

وهي تستخدم في الغالب لمقارنة الأعداد المتجانسة عبر الزمن كما هو موضح في المثال التالي:

مثال ٥ :

يبين الجدول التالي كمية الواردات من القمح والذرة والشعير في دولة ما خلال الأعوام ١٩٨٥ ــ ١٩٩٥ :

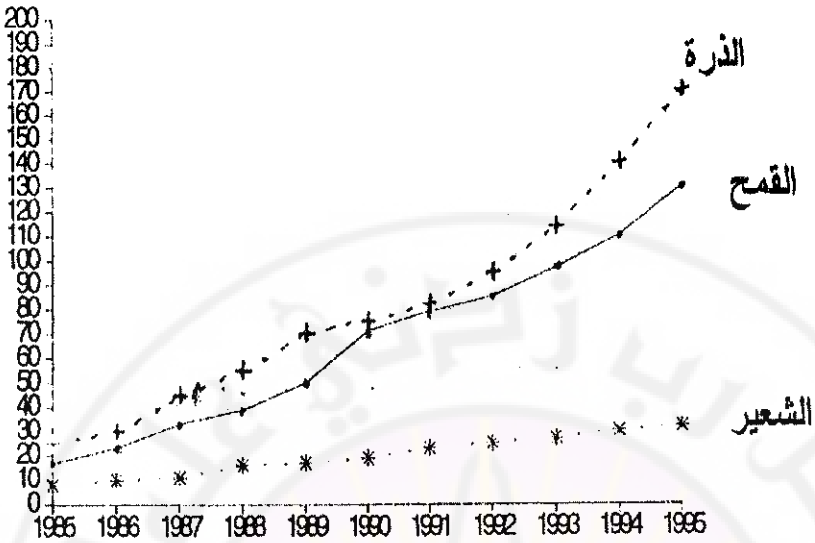
	85	86	87	88	89	90	91	92	93	94	95
القمح	17	23	33	39	50	71	79	85	97	110	130
الذرة	25	30	45	55	70	75	82	95	114	140	170
الشعير	8	10	11	16	17	19	23	25	27	30	32
الإجمالي	50	63	89	110	137	165	165	184	205	280	332

وضح ذلك برسم خطوط بيانية ؟

الحل :

يكون توضيح بيانات هذا الجدول باستخدام الخطوط البيانية كما يلي :

Damascus University



شكل (٦ ــ ٣)

٣ ــ ٤ : الجداول التكرارية :

بعد عملية الجمع للبيانات، لابد من تبويب هذه البيانات في جدول تكراري فإذا كان عدد المشاهدات قليلاً فإن هذه العملية تتم بسهولة ويسر كما يبدو في المثال التالي :

3, 11, 4, 3, 5, 5, 4, 4

فإنه يمكن تبويب هذه البيانات كما يلي :

القيمة	التكرار
3	2
4	3
5	2
11	1

لكن إذا كانت البيانات أو المشاهدات كثيرة ومختلفة فإنه يتم تقسيمها إلى مجموعات جزئية (تسمى الفئات) تضم كل منها عدداً من القيم المتقاربة ويكون عدد القيم داخل كل فئة هو (التكرار) ويسمى الجدول الذي يضم الفئات والتكرارات لهذه القيم بالجدول التكراري، وهو بالتعريف: تبويب للبيانات في فئات متسلسلة بحيث يظهر أمام كل فئة عدد مرات تكرارها.

خطوات إنشاء الجدول التكراري :

إن تركيب الجداول التكرارية لا يعتمد على قواعد محددة معينة ولكن يختلف بحسب نوع البيانات الخام وأهمية الأجزاء التي يراد أن يتركب الجدول منها ومع ذلك هناك أمور مهمة ينبغي مراعاتها عند إنشاء الجداول التكرارية نورد أهمها في ما يلي:

١ ــ نقسم البيانات في الجداول التكرارية إلى مجموعات متجانسة فيما بينها تسمى فئات.

٢ ــ عدد هذه الفئات عادة ما يكون بين ٥ إلى ١٥ فئة.

٣ ــ ليس من الضروري أن تكون هذه الفئات متساوية الطول. ولكن يفضل ذلك إن أمكن.

٤ ــ لكل فئة بداية تسمى حد الفئة الأدنى ونهاية تسمى حد الفئة الأعلى.

٥ ــ الحد الأدنى لأية فئة هو الحد الأعلى للفئة السابقة لها مباشرة.

٦ ــ كل الفئات تشمل أحد حديها فقط الحد الأدنى أو الحد الأعلى .

٧ ــ تحصر وتحصى كل المشاهدات التابعة لكل فئة من الفئات وتكتب أمام الفئة التي تتضمنها وهذا ما يدعى بتكرار الفئة. وقد يستخدم بدلاً عن تكرار الفئات نسبة تلك التكرارات إلى مجموع البيانات الكلي الداخلة في الجدول مضروبة في مئة وتسمى النسبة المئوية لتكرار الفئة.

مثال :

تمثل البيانات التالية عدد الأوراق الانتخابية الموجودة في ثلاثين صندوقاً انتخابياً في إحدى المدن، والمطلوب: إنشاء جدول تكراري لهذه البيانات؟

21	10	18	17	14
15	12	21	25	22
23	19	20	16	26
23	18	14	19	16
19	19	29	18	20
24	27	22	20	24

الحل :

نوزع البيانات في جدول تكراري عدد فئاته (٥) تشمل على حدها الأدنى فقط، ولتحديد طول الفئة الواحدة، نقسم الفرق بين أكبر قيمة وأصغر قيمة بين المشاهدات على عدد الفئات وهو في مثالنا 19=10-29 وبالتالي فإن طول الفئة يساوي تقريباً (٤) وعلى هذا الأساس نقوم بإنشاء الجدول التكراري.

الفئات	التكرارات	نقطة وسط الفئة
14-10	2	12
18-14	6	16
18-22	12	20
22-26	7	24
26-30	3	28

إن طول الفئة عبارة عن المدى العددي الواقع بين حدي الفئة الأدنى والأعلى، ففي الجدول السابق، كان طول الفئة الأولى أي الفئة ١٠ وأقل من ١٤ يساوي (٤) أوراق والأمر نفسه بالنسبة لجميع فئات هذا الجدول نظراً لأن فئاته متساوية فالرقم (١٠) يدعى بالحد الأدنى للفئة الأولى في حين أن الرقم (١٤) يدعى بالحد الأعلى لنفس الفئة وهاتين تدعيان بحدي الفئة وعندما تكون البيانات مبوبة، تفقد المفردات الأصلية هويتها، ومفردات كل فئة من الفئات السابقة تعامل وكأن لها نفس قيمة نقطة وسط فئتها نفسها أو مركز فئتها، ووسط الفئة هو عبارة عن حاصل جمع الحد الأدنى والحد الأعلى للفئة مقسوماً على (٢) أي :

$$\text{وسط الفئة} = \frac{\text{الحد الأدنى} + \text{الحد الأعلى}}{٢}$$

وبالتالي فإن نقطة وسط الفئة تمثل مفردات فئتها.

ويمكن الاستئناس بالعلاقة المقترحة من قبل (Sturges) في تحديد عدد الفئات وشكل العلاقة العام هو :

$$K = 1+3.322\ (\log n)$$

حيث إن :

K تشير إلى عدد الفئات

n عدد المشاهدات

أما بالنسبة لتحديد طول الفئة والذي يسمى أحياناً بمدى الفئة فيتم من خلال العلاقة:

$$\text{طول الفئة (C)} = \frac{\text{المدى العام}}{K} = \frac{R}{K}$$

حيث نحصل على المدى العام: بإيجاد الفرق بين أكبر قيمة وأصغر قيمة في البيانات قبل التبويب، مع محاولة وتقريب النتيجة إلى عدد صحيح في حالة الحصول على كسر.

ومن الضروري التنويه إلى أن هذه العلاقة لا تعتبر قاعدة يمكن الاعتماد عليها دوماً، وذلك لاعتبارات حسابية.

٣ ــ ٥ : عرض التوزيعات التكرارية :

يتم عرض التوزيعات التكرارية بإحدى الطرق الآتية :

١ ــ المدرج التكراري .

٢ ــ المضلع التكراري

٣ ــ المنحني التكراري

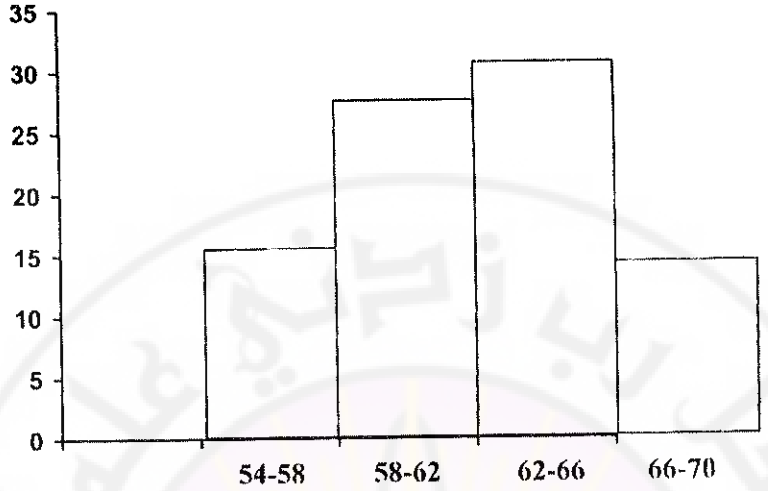
٤ ــ المنحني التكراري المتجمع : (التجميعي) .

وسنناقش فيما يلي كل من هذه الطرق :

١ ــ **المدرج التكراري** : وهو عبارة عن مجموعة من المستطيلات المتساوية القواعد والمختلفة بالارتفاع، ويوضح المثال التالي كيفية رسم المدرج التكراري .

فئات الوزن	54-58	58-62	62-66	66-70
عدد الأشخاص	15	27	30	13

(ملحوظة : الوزن بالكيلوجرام).



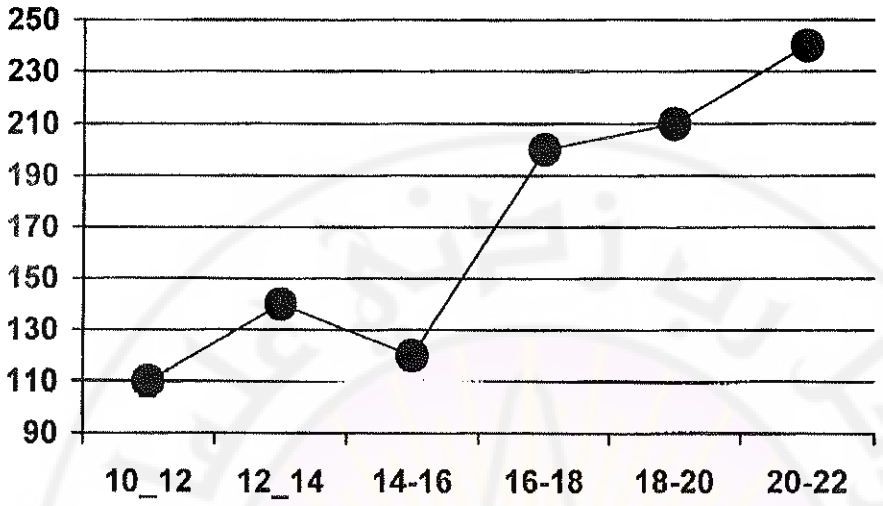
٢ ــ المضلع التكراري :

وهو يتكون من خط بياني يصل بين نقط مراكز الفئات المقابلة لكل تكرار حيث مركز (أو منتصف) الفئة عبارة عن [(بداية الفئة + نهاية الفئة) ÷ ٢] ويمكن أيضاً رسم المضلع التكراري على الشكل المكون للمدرج التكراري وذلك بأخذ منتصفات المستطيلات من الأعلى وتوصيل النقط الناتجة .

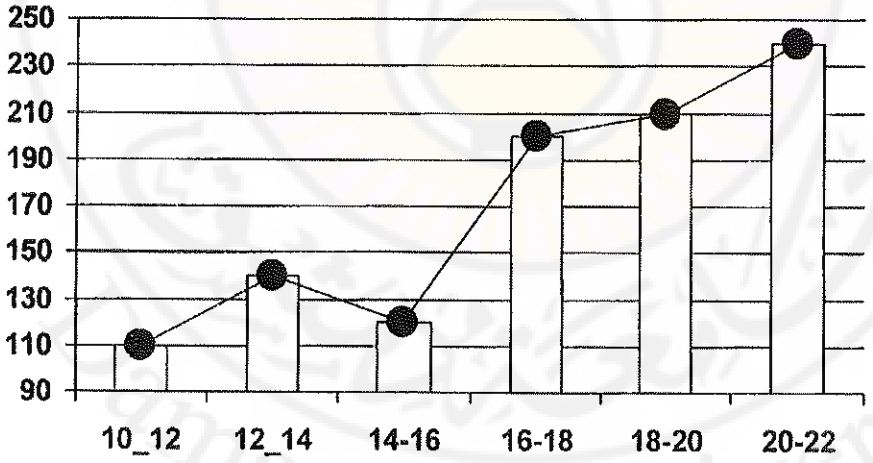
ويوضح المثال التالي طريقة رسم المضلع التكراري .

الدرجات	10-12	12-14	14-16	16-18	18-20	20-22
عدد الطلاب	110	140	120	200	210	240

ويمكن رسمه باستخدام المضلع التكراري كما يلي :



ويمكن رسمه باستخدام مدرج تكراري كما يلي:



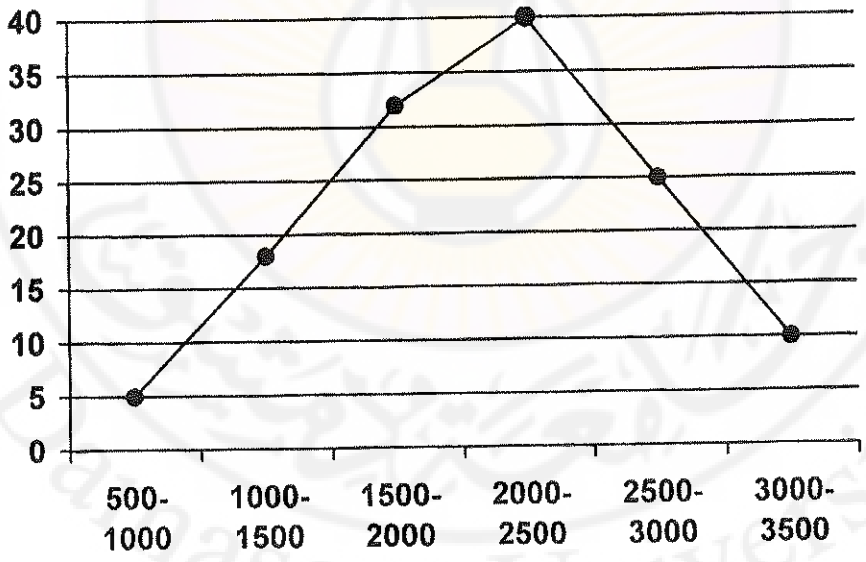
٣ ــ المنحني التكراري :

وهو تمهيد للمضلع التكراري (بمعنى أن الخط البياني لا يكون منكسراً كما في المضلع التكراري).

ويوضح المثال التالي طريقة رسم المنحني التكراري.

فئات الدخل	500-1000	1000-1500	1500-200	2000-2500	2500-3000	3000-3500
عدد الأشخاص	5	18	32	40	25	10

(ملحوظة : الدخل بالوحدات النقدية)



٤ ــ المنحني التكراري التجميعي :

وينقسم المنحني التكراري المتجمع إلى نوعين :

الأول : منحني تكراري متجمع صاعد .

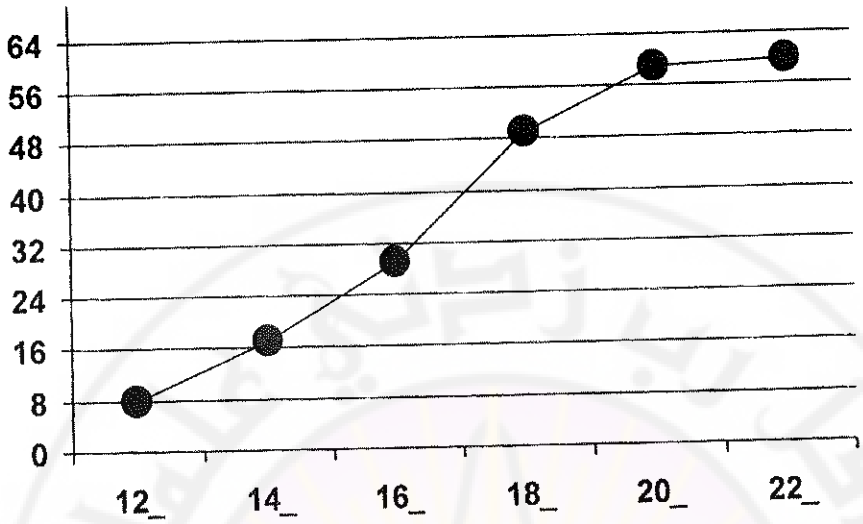
الثاني : منحني تكراري متجمع هابط .

وتستخدم هذه المنحنيات المتجمعة عندما يتطلب البحث تجميع التكرارات من الفئات المتتالية بهدف معين .

وتوضح الأمثلة التالية ذلك .

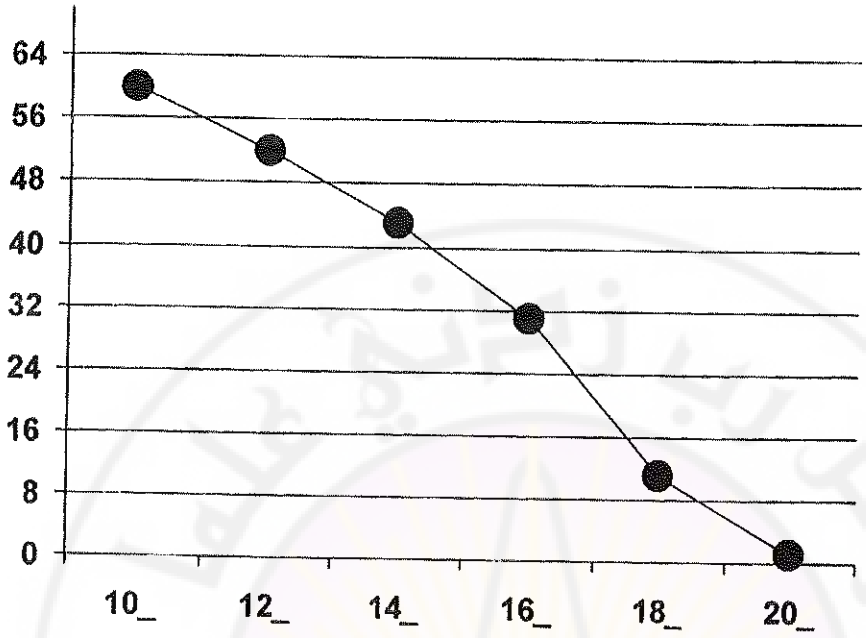
المنحني التكراري المتجمع الصاعد :

فئات العمر	عدد الأشخاص	حدود عليا للفئات	تكرار متجمع صاعد
10-12	8	أقل من 12	8
12-14	9	أقل من 14	17
14-16	12	أقل من 16	29
16-18	20	أقل من 18	49
18-20	10	أقل من 20	59
20-22	1	أقل من 22	60
مجموع	60	—	—



المنحني التكراري المتجمع الهابط :

فئات العمر (سنة)	عدد الأشخاص	حدود عليا للفئات	تكرار متجمع صاعد
10-12	8	أكثر من 10	60
12-14	9	أكثر من 12	52
14-16	12	أكثر من14	43
16-18	20	أكثر من 16	31
18-20	10	أكثر من 18	11
20-22	1	أكثر من 20	1
مجموع	60	_	_



Damascus University

أسئلة

وتمارين غير محلولة

١ ــ **عرف المفاهيم التالية** : الإحصاء، العرض البياني، العينة الطبقية، الجدول الإحصائي، الجدول التكراري، الحصر الشامل، العينة المنتظمة، المضلع التكراري، الخطوط البيانية عشوائية العينة.

٢ ــ ماهي المصادر الأساسية للبيانات الإحصائية مع الشرح .

٣ ــ ما الفرق بين الحصر الشامل، والحصر غير الشامل (العينات) مع مثال على كل منهما.

٤ ــ اشرح مفهوم العينة العشوائية الطبقية؟ وما الفرق بينها وبين عينة الحصص.

٥ ــ أذكر خمسة شروط هامة يجب مراعاتها عند صياغة أسئلة الاستمارة الإحصائية (الاستبيان).

٦ ــ ما هي مراحل الدراسة الإحصائية مع مثال على ذلك .

٧ ــ أذكر أهم الشروط التي يجب مراعاتها عند إعداد الجدول الإحصائي.

٨ ــ ما هي مكونات الاستمارة الإحصائية مع الشرح بالتفصيل ؟

٩ ــ بيّن دور الإحصاء في التخطيط المستقبلي.

١٠ ــ أذكر أهم مراحل إنشاء الجدول التكراري ؟

١١ ــ فيما يلي بيان المعدل السنوي لاستهلاك البنزين (مليون/ ليتر) في إحدى الدول خلال (١٩٩٠ ــ ١٩٩٦)م .

السنة	1990	1991	1992	1993	1994	1995	1996
المعدل	150	170	250	360	220	180	175

المطلوب : عرض تلك البيانات على شكل خطوط بيانية .

١٢ ــ توضح البيانات التالية بعض أهم الإحصاءات الزراعية في الجمهورية العربية السورية عام ٢٠٠٤

اسم المؤشر %	عام 2004
المساحة القابلة للزراعة	31.9
الأراضي المزروعة	80
الأراضي المروية	30
الحراج	3
المراعي	45

المصدر : الجهاز المركزي للإحصاء ٢٠٠٥

المطلوب :

عرض تلك البيانات باستخدام الدوائر مرة ومرة أخرى باستخدام الأعمدة البيانية (المستطيلات).

١٣ ــ توضح البيانات التالية عدد العاملين في الصناعات الاستخراجية وفي صناعة الكيماويات وتكرير النفط للأعوام (١٩٩١ ــ ١٩٩٦) في إحدى الدول.

السنة	العاملين في الصناعة الاستخراجية	العاملين في صناعة الكيماويات وتكرير النفط
1991	1806	6351
1992	1963	7009
1993	2010	7196
1994	1544	7150
1995	2249	7073
1996	2201	6999

المطلوب: اعرض تلك البيانات بالأسلوب الذي تراه مناسباً ؟

١٤ ــ يبين الجدول التالي الحالة التعليمية لمشاهدي نشرة الأخبار المسائية في التلفزيون لعينة عشوائية مؤلفة من (١٠٠) مشاهد من مشاهدي التلفزيون في إحدى المدن .

الحالة التعليمية	ابتدائي	إعدادي	ثانوي	جامعي	دراسات عليا
عدد المشاهدين	15	25	30	15	15

المطلوب: عرض تلك البيانات باستخدام أسلوبين مختلفين للعرض .

١٥ ــ توضح البيانات التالية أعداد الطلاب الملتحقين بالمعهد العالي للبترول والمعهد الوطني للإدارة، في إحدى الجامعات العربية من (١٩٩٥ ــ ١٩٩٧).

المعهد	1995	1996	1997
المعهد العالي للبترول	206	220	280
المعهد الوطني للإدارة	201	415	476
الإجمالي	407	635	856

المطلوب :

اعرض تلك البيانات بالأسلوب الذي تراه مناسباً .

١٦ ــ الجدول التالي يوضح طلبة كلية العلوم السياسية المقبولين في إحدى السنوات حسب النسبة المئوية للدرجات في الثانوية العامة في إحدى الدول العربية :

فئات الدرجات	65-70	70-75	75-80	80-85	85-90	90-95	59-100
عدد المقبولين	38	100	126	312	89	26	10

المطلوب :

ـــ أرسم المدرج التكراري، المضلع التكراري، المنحني التكراري، ماذا تستنتج ؟

١٧ ـــ يبين الجدول التالي الحالة الاجتماعية لعينة عشوائية مؤلفة من (١٠٠) عامل من عمال الصناعة البتروكيميائية في إحدى الدول العربية فكانت كالآتي :

الحالة الاجتماعية	عازب	متزوج	مطلق	أرمل	المجموع
عدد العمال	32	63	3	2	100

المطلوب:

١ ـــ هيئ جدول لتفريغ البيانات السابقة.

٢ ـــ ارسم المدرج التكراري .

١٨ ـــ يبين الجدول التالي أهم مؤشرات قطاع الصحة في الجمهورية العربية السورية لعام ٢٠٠٤ .

متوسط عدد السكان لكل طبيب	متوسط عدد السكان لكل صيدلي	متوسط عدد السكان لكل طبيب أسنان
694	1413	1174

المصدر: الجهاز المركزي للإحصاء ٢٠٠٥.

المطلوب :

١ ــ عرض تلك البيانات على شكل مستطيلات ؟

٢ ــ عرض تلك البيانات على شكل دائرة .

١٩ ــ البيانات التالية تمثل التوزيع النسبي للإنفاق على وسائل الأعلام في إحدى الدول .

وسيلة الأعلام	الجرائد	التلفزيون	الراديو	المجلات	البريد المباشر
النسبة المئوية	38.5	20.5	18.4	17.5	5

المصدر: فرضي

المطلوب :

عرض تلك البيانات على شكل مستطيلات .

٢٠ ــ تم مقابلة ٥٠٠ ناخب فور إدلائهم بأصواتهم في إحدى الدوائر الانتخابية فكان الجدول التالي :

نوع المهنة	عدد الناخبين
ربات منازل	130
موظفون إداريون	250
مدرسين	50
عمال فنيون	20
آخرون	50

المصدر : فرضي.

المطلوب :

عرض البيانات أعلاه بصورتين بيانيتين مختلفتين .

٢١ ــ الجدول التالي يبين قطاع النقل والاتصالات في الجمهورية العربية السورية لعام ٢٠٠٤

قطاع النقل	العدد
نسبة الطرق الإسفلتية من إجمالي الطرق	75
عدد الهواتف الأرضية لكل 1000	148
عدد الهواتف الخليوي لكل 1000	121
عدد المشتركين بالانترنيت لكل 1000	9

المصدر : ج . م . للإحصاء ٢٠٠٥

المطلوب :

١ ــ استخراج النسبة المئوية لكل جزء إلى المجموع الكلي .

٢ ــ عرض البيانات السابقة على شكل دائرة.

٢٢ ــ بفرض الجدول التالي يوضح ميزانية السنة الحالية لإحدى الدول (٢٠٠٥ ــ ٢٠٠٤) (بمليارات الوحدات النقدية)

القسم	الصحة والتعليم	الدفاع	البلدية	العدالة	أخرى
الميزانية	200	150	50	20	80

المطلوب:

عرض تلك البيانات بالأسلوب الذي تراه مناسباً.

الفصل الرابع

مقاييس النزعة المركزية

Measures Of Central Tendency

- مقدمة
- الوسط الحسابي
- الوسيط
- الربيعيات
- المنوال
- العلاقات التقريبية بين المتوسطات
- الوسط الهندسي

Damascus University

وقل رب زدني علما
جامعة دمشق
Damascus University

٤ ــ ١ : مقدمة

يحاول الباحث عند دراسته لمجموعة من البيانات، أن يجد قيمة ممثلة، لها، تحمل صفاتها وتعبر عن خصائص وصفات المجموعة التي تنتمي إليها، لذلك يبحث عن نقطة تجمع أو تمركز لهذه البيانات. لقد بينا فيما سبق أن مركز الفئة هو ممثل لجميع المفردات الواقعة بين حديها، أي أنه يوجد قيمة عددية واحدة تمثل حدي الفئة فعند تخرج الطالب من كلية العلوم السياسية، يختصر الحديث عن الدرجات التي حصل عليها في جميع المواد بقيمة واحدة تعبر عن معدله العام، وكأنه قد حصل في جميع المقررات درجة هذا المعدل. إن نقطة التجمع أو التمركز للبيانات عند قيمة معينة تدعى بالنزعة المركزية وتعرف النزعة المركزية بأنها ميل المفردات للتمركز عند قيمة معينة، فأطول الناس تتجمع أو تتمركز عند قيمة متوسطها مثلاً (١٨٠) سم، وكذلك عدد الكريات الحمراء في جسم الإنسان وأجور العمال وأعمار الناخبين.. إلخ.

وتقاس النزعة المركزية بالمتوسطات التالية: الوسط الحسابي، الوسيط، المنوال، الوسط الهندسي، الوسط التوافقي، ويطلق بعض الإحصائيين على هذه المقاييس تسمية مقاييس الموضع ؟!.

الوسط الحسابي
Arithmetic Mean

٤ ــ ٢ : الوسط الحسابي :

يعتبر الوسط الحسابي من أهم مقاييس النزعة المركزية أو مقاييس الموضع، وذلك لوضوح معناه وسهولة حسابه واستخدامه في معظم التوزيعات الإحصائية، ويعرف الوسط الحسابي لمجموعة من القيم، بأنه

مجموع هذه القيم مقسوماً على عددها، ويرمز له بـ ($\overline{X}$) إذا كان محسوباً من عينة و بـ (μ) إذا كان محسوباً من مجتمع، فإذا كان لدينا مجموعة من المفردات قيمها: $x_1, x_2, x_3, \ldots, x_n$ وعددها n فإن الوسط الحسابي لهذه المجموعة من المفردات يحسب وفق الصيغة التالية :

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}, i = 1,2,3,.....n$$

ويتم حساب الوسط الحسابي لنوعين من البيانات :

بيانات غير مبوبة، بيانات مبوبة .

ـ حساب الوسط الحسابي لبيانات غير مبوبة :

هناك أسلوبان لحسابه :

أ ـ **الأسلوب المباشر** : ويتم حسابه على النحو التالي :

$$\overline{X} = \frac{\sum x}{n}$$

مثال : باع أحد الموزعين من إحدى الصحف : ٩٩ ، ٨٥، ١٠٤، ٨٧ صحيفة وذلك خلال (٥) أيام متتالية، **والمطلوب** : احسب الوسط الحسابي بالأسلوب المباشر لعدد الصحف المباعة في اليوم الواحد .

$$\overline{X} = \frac{\sum x}{n} = \frac{87+104+95+85+99}{5} = \frac{470}{5} = 94 \text{ صحيفة}$$

ويمكن التحقق من خلال العلاقة. $\overline{nx} = \sum x = (5)(94) = 470$

إن أهم خاصة من خواص الوسط الحسابي هي : أن مجموع انحرافات القيم عن وسطها الحسابي يساوي صفراً .

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

أي

$$\sum x - n\bar{x} = 0$$

وبالقسمة على n :

$$\frac{\sum x}{n} - \frac{n\bar{x}}{n} = \bar{x} - \bar{x} = 0$$

وبالتطبيق على مثالنا السابق :

$$(87\text{-}94) + (104\text{-}94) + (95\text{-}94) + (85\text{-}94) + (99\text{-}94) = 0$$

ب ــ أسلوب الانحرافات :

بالاعتماد على الخاصة السابقة يمكن التوصل إلى أن مجموعة انحرافات القيم عن أية قيمة أخرى غير الوسط الحسابي لا تساوي الصفر ويمكن البرهان على ذلك من خلال :

$$\sum (x - A) = \sum d$$

حيث A وسط حسابي مفترض نختاره من بين البيانات المعطاة $A \neq \bar{X}$ ، d انحراف كل مفردة عن قيمة الوسط المفترض A .

$$\sum x - nA = \sum d$$

بالقسمة على n :

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

$$\overline{X} - A = \frac{\sum d}{n}$$

$$\overline{X} = A + \frac{\sum d}{n}$$

وبالتطبيق على مثالنا السابق نجد أن :

نختار قيمة A الوسط الحسابي المفترض، ولنفرض أنها مساوية إلى ٨٥ :

$$d = x - A$$

$$\overline{X} = 85 + \frac{(87-85)+(85-85)+(99-85)}{5}$$

$$\overline{X} = 85 + 9 = 94$$

ـ حساب الوسط الحسابي لبيانات مبوبة :

ويتم حسابه وفق ثلاثة أساليب :

أ ـ الأسلوب المباشر :

ويعتبر الأسلوب الأكثر استخداماً، ويتم حسابه من جداول تكراري ويقوم على حساب وسط الفئة وضربه بالتكرار المقابل لهذه الفئة ثم أخذ المجموع وقسمته على مجموع التكرارات ويعطى بالعلاقة التالية:

$$\overline{X} = \frac{\sum fx}{\sum f}$$

f = التكرارات

x وسط الفئات (مراكز الفئات)

تمرين :

يبين الجدول التكراري التالي الأجور الشهرية، لعينة عشوائية من العمال في إحدى الصناعات البترولية. (الأجور مقدرة بمئات الوحدات النقدية).

فئات الأجور الأسبوعية	80-120	160-120	200-160	240-200	280-240	Σ
(f) عدد العمال	10	25	50	10	5	100

المطلوب : أوجد قيمة الوسط الحسابي لأجور العمال في هذه الصناعة بالأسلوب المباشر.

الحل :

نقوم بإنشاء الجدول المساعد.

فئات الأجور	التكرارات f	مراكز الفئات x	$f.x$
80-120	10	100	1000
160-120	25	140	3500
200-160	15	180	9000
240-200	10	220	2200
280-240	5	260	1300
Σ	100		17000

بتطبيق قانون الوسط الحسابي $\overline{X} = \frac{\sum fx}{\sum f} = \frac{17000}{100} = 170$

ب ــ أسلوب الانحرافات :

تعتمد هذه الطريقة على افتراض أية قيمة على أنها وسط حسابي مفترض وليكن A، ثم نحسب الانحراف عن هذا الوسط المفترض ، وبقسمة مجموع الانحرافات عن الوسط المفترض على مجموع التكرارات وجمعها مع الوسط المفترض نحصل على الوسط الحسابي $\overline{X}$ وذلك من خلال العلاقة التالية:

$$\overline{X} = A + \frac{\sum fd}{\sum f}$$

حيث d = x-A

وبالتطبيق على المثال السابق:نفرض أننا اخترنا وسطاً مفترضاً A=180

فئات الأجور	f	x	d=x-A	f.d
120-80	10	100	80-	800-
160-120	25	140	40-	1000-
200-160	50	180	0	0
240-200	10	220	40+	400+
280-240	5	260	80+	400+
Σ	100			-1000

$$\overline{X} = A + \frac{\sum fd}{f} = 180 + \frac{(-1000)}{100} = 180 - 10 = 170$$

وهي متوسط أجر العامل الواحد في هذه الصناعة وهي نفس القيمة السابقة.

ج ــ أسلوب الانحرافات المختصرة :

تعتبر هذه الطريقة من أبسط الطرق لحساب الوسط الحسابي ويشترط عند تطبيقها الأسلوب أن تكون أطوال فئات التوزيع التكراري متساوية، وأن يتم اختيار الوسط الفرضي حصراً من بين نقاط وسط الفئات، ويعطي بالعلاقة التالية :

$$\overline{X} = A + \left[\frac{\sum fu}{\sum f}\right].C$$

حيث أن :

$$U = \frac{d}{C} = \frac{X - A}{C}$$

C طول الفئة .

وبالعودة إلى التمرين السابق وبفرض أن A=180

فئات الأجور	f	X	$U = \frac{x - A}{C}$	fu
120-180	10	100	2-	20-
160-120	25	140	1-	25-
200-160	50	180	0	0
240-200	10	220	1+	10
280-240	5	260	2+	10
Σ	100			25-

$$\overline{X} = A + \frac{\sum f.u}{f}.C = 180 + \frac{(-25)}{100}.40 = 180 - 10 = 170$$

وهي النتيجة السابقة التي حصلنا عليها بالأسلوبين المباشر والانحرافات .

٣ ــ ٤ : الخصائص الرياضية والجبرية للوسط الحسابي :

١ ــ من الاستعراض السابق لحساب الوسط الحسابي بالطرق الثلاث، نجد أنها تعطي قيمة واحدة للوسط ويفضل استخدام أسلوب الانحرافات عندما تكون أطوال الفئات غير متساوية وأسلوب الانحرافات المختصرة عندما تكون أطوال الفئات متساوية.

٢ ــ وجدنا إن مجموع انحرافات القيم عند وسطها الحسابي يساوي الصفر.

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

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

٣ ــ إن حاصل ضرب الوسط الحسابي بمجموع التكرارات يساوي حاصل مجموع التكرارات بأوساط الفئات.

$$\overline{nx}=\sum_{i=1}^{n}x_i$$

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

وبالتطبيق على المثال نجد أن :

$$100 * 170 = 17000$$

٤ ــ إن مجموع انحرافات القيم عن أية قيمة غير الوسط الحسابي لا تساوي الصفر.

$$\sum(x-A)\neq 0$$

$$\sum f(x-A)\neq 0$$

٥ ــ يعتبر من أكثر المتوسطات تأثراً بالقيم المتطرفة على أحد نهايتي التوزيع، وهذا ما يفقده صفته التمثيلية عندما يكون التوزيع ملتوياً .

٦ ــ صعوبة حسابه عندما تكون إحدى نهايتي التوزيع مفتوحة .

٧ ــ يستخدم كمعيار أي كمقياس تقارن به المفردات الأخرى عندما يكون التوزيع متماثلاً.

الوسيط

Median

٤ ــ ٤ : الوسيط :

يعرف الوسيط بأنه القيمة التي يتساوى عن طرفيها عدد القيم المرتبة تصاعدياً أو تنازلياً، أو أنه القيمة التي نصف المفردات المرتبة دونها والنصف الآخر أكبر منها.

حساب الوسيط لبيانات غير مبوبة:

أ ــ إذا عدد المفردات فردياً، فإن الوسيط هو القيمة المفردة التي تقع في وسط الترتيب، ويتم تحديد موقعه من خلال العلاقة :

$$\frac{n+1}{2}$$

تمرين :

لدينا القيم التالية : 13 , 25 , 19 , 27, 15, 21, 17 نرتب هذه القيم تصاعدياً 13, 15, 17, 19, 21, 25, 27 نطبق العلاقة السابقة :

$$\frac{n+1}{2} = \frac{7+1}{2} = 4$$

فتكون المفردة التي ترتيبها الرابع في الصف التصاعدي هي قيمة الوسيط للبيانات السابقة وهي تساوي (19) .

ب ــ أما إذا كان عدد المفردات زوجياً، فإن موقع الوسيط يتحدد وفق نفس العلاقة السابقة.

13, 15, 17, 19, 21, 25, 28, 31

$$\frac{n+1}{2} = \frac{8+1}{2} = 4.5$$

أي أن الوسيط يقع بين المفردتين اللتين ترتيبهما (٤) و (٥) أي بين القيمتين (١٩) و (٢١) وبأخذ الوسط الحسابي للمفردتين اللتين يقع بينهما الوسيط هي قيمة الوسيط. أي :

$$\frac{19+21}{2} = 20 \quad .$$

الوسيط لبيانات مبوبة :

يتم حساب الوسيط للبيانات المبوبة من خلال العلاقة التالية :

$$Me = L_1 + \left[\frac{\frac{n}{2} - \sum f_1}{f_{me}}\right].C$$

حيث : L_1 تشير إلى الحد الأدنى للفئة الوسيطية.

n: مجموع التكرارات .

$\sum f_1$: مجموع التكرارات التي تسبق تكرار الفئة الوسيطة f_{me} .

f_{me} : تكرار الفئة الوسيطية .

C : طول الفئة الوسيطية .

Me :قيمة الوسيط

ولإيضاح كيفية تطبيق هذه العلاقة ندرس المثال التالي :

تمرين :

قامت مصلحة الضرائب بدراسة حجم المبيعات الحقيقية لمخازن بيع المستهلك في إحدى الدول، فأخذت عينة عشوائية حجمها ١٥٠ مخزناً وحسبت مبيعاتها السنوية.

فأعطت الجدول التكراري التالي (المبيعات بملايين الوحدات النقدية).

فئات المبيعات	12-6	18-12	24-18	30-24	36-30	42-36	Σ
عدد المخازن	12	18	32	43	32	13	150

والمطلوب:

أوجد وسيط المبيعات لهذه المخازن .

الحل :

من أجل حساب الوسيط لهذه البيانات نتبع الخطوات التالية :

١ ــ ننشئ عموداً للتكرار التجميعي (الصاعد أو الهابط)، والتكرار التجميعي الصاعد لفئة ما يساوي إلى تكرار هذه الفئة مضافاً إليها مجموع تكرار جميع الفئات التي تسبقها.

٢ ــ نحدد موقع الفئة الوسيطية وذلك خلال قسمة مجموع التكرارات على العدد /٢/

أي : $\frac{n}{2}$

٣ ــ نبحث عن القيمة $\frac{n}{2}$ أو أكبر منها مباشرة في عمود التكرار التجميعي الصاعد، وبذلك تكون الفئة المقابلة هي الفئة الوسيطية.

٤ ــ نطبق العلاقة السابقة ونحسب قيمة الوسيط .

وبالتطبيق على التمرين السابق نجد :

فئات المبيعات	f	التكرار التجمعي الصاعد
12-6	12	12
18-12	18	30
24-18	32	62
30-24	43	105
36-30	32	137
40-36	13	150
	$\Sigma f = 150$	

١ ــ ترتيب الوسيط : $\frac{n}{2} = \frac{150}{2} = 75$

٢ ــ الفئة الوسيطية هي الفئة الرابعة :

$$Me = L_1 + \left[\frac{\frac{n}{2} - \sum f_1}{f_{me}}\right].C$$

$$Me = 24 + \left[\frac{75 - 62}{43}\right].6 = 25.8$$

التفسير :

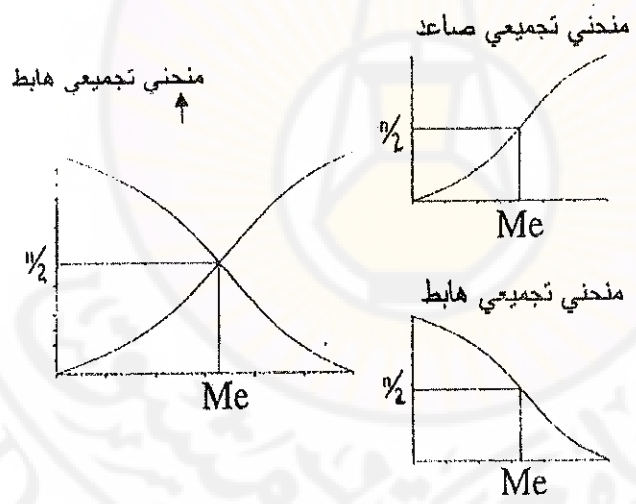
إن نصف مبيعات المخازن كانت مبيعاتها أقل من ٢٥,٨ مليون وحدة نقدية وإن نصف المخازن كانت مبيعاتها أكبر من هذه القيمة .

٤ ــ ٥ : تحديد الوسيط بيانياً :

يمكن تحديد قيمة الوسيط بيانياً وذلك من خلال رسم منحني التكرار التجميعي الصاعد أو الهابط أو الاثنين معاً. سنتناول التمرين السابق لتحديد الوسيط بيانياً.

الخطوات :

نحدد النقطة n/2 على المحور العمودي (محور التكرار التجميعي الصاعد أو الهابط) ثم نرسم من هذه النقطة خطاً أفقياً موازياً لمحور الفئات إلى أن يلاقي المنحني، ثم نسقط من هذه النقطة عموداً على محور الفئات وتكون النقطة هي قيمة الوسيط .



من خلال الشكل نلاحظ أن نقطة تقاطع منحني التكرار التجميعي الصاعد والهابط والتي تُسقط على المحور الأفقي هي قيمة الوسيط.

خصائص الوسيط :

١ ــ إن الوسيط هو مقياس موقع أو ترتيب وليس مقياس قيم كالوسط الحسابي.

٢ ــ يمكن حسابه عندما تكون إحدى نهايتي التوزيع مفتوحة .

٣ ــ بالمقارنة مع الوسط الحسابي يعتبر الوسيط أقل تأثراً بالقيم المتطرفة على جانب التوزيع.

٤ ــ لا يتصف بمزايا رياضية وجبرية كالوسط الحسابي، لذلك تعد طرق حسابه تقريبية، ويمكن ملاحظة ذلك من خلال ضرب الوسيط بمجموع التكرارات، إذ سوف ينتج قيمة تختلف عن مجموع البيانات الأصلية .

٥ ــ تعتد قيمة الوسيط غير ثابتة عندما يكون عدد المفردات قليلاً.

الربيعيات
Quartiles

٤ ــ ٦ : الربيعيات :

عند دراستنا للوسيط، قمنا بتقسيم البيانات الإحصائية إلى قسمين متساوين، لكن قد تدعوا الحاجة أحياناً على تقسيم البيانات إلى أكثر من قسمين إلى أربعة أقسام مثلاً (الربيعيات) أو إلى عشرة أقسام (العشريات) أو إلى مائة جزء (الميئنيات).

وعند دراستنا للربيعيات نميز بين الربيع الأول (Q_1) أو الأدنى وهو عبارة عن المفردة التي يكون ٢٥% من مفردات التوزيع أقل منها و٧٥% أكبر منها، أما الربيع الثالث (Q_3) فيعرف بأنه المفردة التي يكون ٧٥% من مفردات التوزيع أقل من قيمتها و و٢٥% من مفردات التوزيع أكبر منها .

ـــ حساب الربيعيات لبيانات غير مبوبة :

يتم تحديد موقع الربيع الأول لبيانات غير مبوبة من العلاقة $\frac{n+2}{4}$ أما الربيع الثالث يتم تحديده من خلال العلاقة : $\frac{3n+2}{4}$

تمرين :

ليكن لدينا البيانات التالية المرتبة تصاعدياً 6.9,11,15,17.21,27

أوجد قيمة الربيع الأول والثالث لهذه البيانات :

ـــ الربيع الأول (Q1) : يتم تحديد موقعه .

$$\frac{n+2}{4} = \frac{7+2}{4} = 2.25$$

نلاحظ أن الربيع الأول أقرب للقيمة الثانية منه للثالثة، لذلك نأخذ الفرق بين القيمتين (9,11) ونحسب منه قيمة الربيع الأول.

$$\mathbf{Q_1 = 9 + (11\text{-}9) * 0.25 = 9.5}$$

الربيع الثالث Q3 : نحدد موقعه أولاً

$$\frac{3n+2}{4} = \frac{(3).(7)+2}{4} = 5.75$$

وبنفس الطريقة السابقة نحصل على قيمة الربيع الثالث

$$\mathbf{Q_3 = 17 + (21 - 17)\ 0.75 = 20}$$

حساب الربيعيات لبيانات مبوبة :

إن حساب الربيعيات مبوبة مشابهة تماماً لطريقة حساب الوسيط.

نحن نعلم أن موقع الوسيط يحدد بالعلاقة $\frac{n}{2}$ بينما موقع الربيع الأول يحدد من العلاقة $\frac{n}{4}$ والربيع الثالث من العلاقة $\frac{3n}{4}$

وبالعودة إلى التمرين السابق في بحث الوسيط، والمتعلق بتقدير وسيط المبيعات لمخازن بيع المستهلك.

والمطلوب :

أوجد قيمة الربيع الأول والثالث لمبيعات تلك المخازن .

الحل :

المبيعات بملايين الوحدات النقدية	F	التكرار التجميعي الصاعد
12-6	12	12
18-12	18	30
18-24	32	62
30-24	43	105
36-30	32	137
42-36	13	150
Σ	150	

ولإيجاد قيمة الربيع الأول تطبق العلاقة التالية :

$$Q_1 = L_1 + \left[\frac{\frac{n}{4} - \sum f_1}{f_{Q1}}\right].C$$

حيث :

L_1 : تشير إلى الحد الأدنى لفئة الربيع الأول.

$\frac{n}{4}$: ربع مجموع التكرارات . (ترتيب الربيع الأول).

Σf_1 : مجموع التكرارات السابقة لتكرار فئة الربيع الأول .

f_{Q1} : تكرار فئة الربيع الأول .

C_{Q1} : طول الفئة . الربيعية الأول.

ولتطبيق العلاقة السابقة نحدد أولاً فئة الربيع الأول من العلاقة :

$$\frac{n}{4} = \frac{150}{4} = 37.5$$

نبحث عن القيمة 37,5 أو أكبر منها مباشرة في حقل التكرار التجميعي الصاعد فتكون هي فئة الربيع المطلوبة : (٢٤ ــ ١٨).

$$Q_1 = 18 + \left(\frac{37.5 - 30}{32}\right).6 = 19.4$$

والتفسير : أن ٢٥% من مبيعات المحطات كانت أقل من ١٩,٤ مليون وحدة وأن ٧٥% من المبيعات هي أكبر من ١٩,٤ مليون وحدة .

الربيع الثالث يعطى بالعلاقة :

$$Q_3 = L_1 + \left[\frac{\frac{3n}{4} - \sum f_1}{f_{Q3}}\right].C$$

ولتحديد فئة الربيع الثالث نطبق العلاقة :

$$\frac{3n}{4} = \frac{3*150}{4} = 112.5$$

نبحث عن القيمة ١١٢,٥ أو أكبر منها مباشرة في حقل التكرار التجميعي الصاعد فتكون فئة الربيع المطلوب : (٣٦ ــ ٣٠).

وبتطبيق العلاقة :

$$Q_3 = 30 + \left[\frac{112.5 - 105}{32}\right].6 = 31.4$$

والتفسير :

إن ٧٥% من مبيعات المحطات كانت أقل من ٣١,٤ مليون وحدة نقدية وأن ٢٥% من المبيعات كانت أكثر من ٣١,٤ مليون وحدة .

ملاحظة :

في التوزيعات المتماثلة يعطى الوسط الحسابي للربيعين الأول والثالث يعطي قيمة تقريبية لقيمة الوسيط.

$$\frac{Q_1 + Q_3}{2} = Med$$

المنوال
Mode

٤ ــ ٧ : المنوال :

يعرف المنوال بأنه القيمة الأكثر تكراراً أو الأكثر شيوعاً بين المفردات، وقد يكون هناك أكثر من منوال واحد لمجموعة من البيانات .

حساب المنوال لبيانات غير مبوبة :

تمرين :

لدينا أطوال مجموعة من الطلاب في إحدى الكليات في الجامعة. ماهو الطول الأكثر شيوعاً بين هؤلاء الطلاب 160,170,175,180,175,187,175

من خلال ذلك نلاحظ الطول الأكثر تكراراً بين مجموعة الطلاب هو الطول 175 لذلك يعتبر هو الطول المنوالي .

عند حساب المنوال يجب أن نميز بين نوعين من الجداول التكرارية التي يجب من خلالها حساب المنوال.

أ ــ جداول تكرارية ذات أطوال فئات متساوية :

وفي هذا النوع من الجداول فإن الفئة المنوالية هي الفئة التي تقابل أكبر تكرار.

ب ــ جداول تكرارية ذات أطوال فئات مختلفة :

وفي هذا النوع من الجداول يجب توحيد أطوال الفئات، أو حساب التكرار الحقيقي للفئات، ونقصد بالتكرار الحقيقي التكرار المشاهد منسوباً إلى طول الفئة المشاهدة مضروباً بطول فئة معياري.

مثال (١) :

حدد الفئة المنوالية للجدول التكراري التالي :

الفئات	15-0	25-15	35-25	45-35	50-45	60-50	75-60
التكرار	25	18	20	18	11	17	21

كما هو واضح الجدول فإن أطوال الفئات مختلفة ولتحديد الفئة المنوالية لابد من توحيد أطوال الفئات، وكما نلاحظ أيضاً من الجدول فإن طول الفئة الأكثر تكراراً هو (١٠) ونعتبر القيمة ١٠ هو طول ا لفئة المعياري والذي سنوجد على أساسه جميع أطوال الفئات من العلاقة:

$$\text{fc : التكرار المعدل} = \frac{\text{التكرار المشاهد}}{\text{طول الفئة المشاهد}} \times \text{طول الفئة المعياري}$$

فالنسبة للفئة الأولى التكرار المعدل $\frac{25}{15}\times 10$

والثانية $\frac{18}{10}\times 10$ وهكذا بالنسبة لبقية الفئات فنحصل على الجدول التكراري المعدل ا .

الفئات	15-0	25-15	35-25	45-35	50-45	60-50	75-60
التكرار المعدل	16.66	18	20	18	22	17	14

طرق حساب المنوال للبيانات المبوبة :

توجد عدة طرق لحساب المنوال نذكر منها : طريقة الفروق وطريقة العزوم وطريقة الرسم البياني وجميع هذه الطرق تقريبية ولا يعتمد عليها في الدراسات الإحصائية المعمقة.

١ ــ طريقة الفروق :

تعتمد هذه الطريقة على حساب الفرق بين التكرار الحقيقي للفئة المنوالية والتكرار الحقيقي للفئة السابقة واللاحقة. فإذا كان الفرق بين التكرار المعدل والحقيقي للفئة المنوالية والتي تليها فإن قيمة المنوال تكون أقرب إلى الحد الأدنى للفئة المنوالية والعكس صحيح، وتعتبر الفئة المنوالية هي الفئة المقابلة لأكبر تكرار في الجدول المعدل في حال عدم تساوي أطوال الفئات.

ويتم حساب المنوال بأسلوب الفروق من خلال العلاقة التالية .

$$Mod = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) .C$$

حيث أن :

L_1 : الحد الأدنى للفئة المنوالية.

Δ_1 : الفرق بين تكرار الفئة المنوالية وتكرار الفئة التي تسبقها .

Δ_2 : الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة .

C : طول الفئة المنوالية.

مثال (٢) :

يبين الجدول التكراري التالي الأجور اليومية لعينة عشوائية مكونة من ١٠٠ عامل من عمال الصناعات البتروكيميائية في إحدى الدول.

فئات الأجور	120-100	140-120	160-140	180-160	200-180	Σ
عدد العمال	9	22	50	12	7	100

والمطلوب: أوجد الأجر الأكثر تكراراً لهؤلاء العمال.

فئات الأجور	f
120-100	9
140-120	22
160-140	50
180-160	12
200-180	7

وكما هو واضح من الجدول المساعد فإن الفئة المنوالية هي الفئة المقابلة لأكبر تكرار (٥٠) وذلك لأن أطوال فئات الجدول متساوية .

نطبق العلاقة :

$$Mod - L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right).C$$

$$\Delta_1 = 50 - 22 = 28$$
$$\Delta_2 = 50 - 12 = 38$$
$$L_1 = 140, C = 20$$

وبالتعويض نجد :

$$Mod = 140 + \left[\frac{28}{28+38}\right].20 = 148.48$$

التفسير : إن الأجر الأكثر شيوعاً هو ١٤٨,٤٨ وحدة نقدية.

طريقة العزوم :

يمكن استخدام العزوم لحساب المنوال، وذلك باستخدام تكرار الفئتين السابقة واللاحقة لفئة المنوال من الجدول التكراري المعدل في حال عدم تساوي أطوال الفئات، وذلك من خلال العلاقة:

$$Mod = L_1 + \left(\frac{f_2}{f_1 + f_2} \right) .C$$

حيث :

L_1 : الحد الأدنى للفئة المنوالية.

f_1 : تكرار الفئة السابقة لفئة المنوال مباشرة.

f_2 : تكرار الفئة اللاحقة لفئة المنوال مباشرة.

مثال (٢) :

أوجد المنوال بطريقة العزوم ثم بطريقة الفروق وعلق على النتيجة للجدول التكراري التالي:

الفئات	15-0	40-15	50-40	65-50	80-65	90-80	100-90	مجموع
التكرار	8	29	25	34	32	10	6	**144**

الحل :

بما أن الأطوال مختلفة لذلك نعدل أطوال الفئات ونعتبر أن طول الفئة المعياري يساوي ٥.

وبالتالي فإن أطوال الفئات المعدل يصبح على النحو التالي :

$\frac{8}{15}*5=2.66$ ، الفئة الثانية = $\frac{29}{25}*5=5.8$ = الفئة الأولى.

ويصبح الجدول على النحو التالي :

الفئات	15-0	40-15	50-40	65-50	80-65	90-80	100-90
التكرار	2.66	5.8	12.5	11.33	10.55	5	3

١ ــ أسلوب الفروق :

$$Mod = 40 + \frac{6.7}{6.7+1.17}.10 = 48.5$$

٢ ــ أسلوب العزوم

$$Mod = 40 + \frac{11.33}{5.8+11.33}.10 = 46.6$$

ومن خلال المقارنة نلاحظ أن الاختلاف في قيمة المنوال ويرجع إلى أن أغلب أساليب حسابه هي تقريبية.

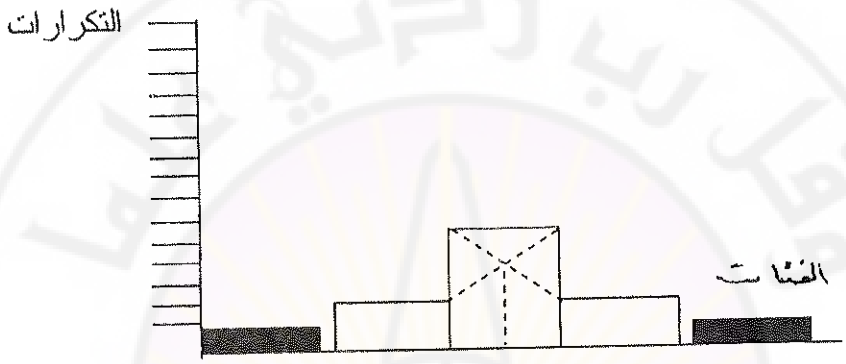
٤ ــ ٨ : تحديد المنوال بيانياً :

عند حساب المنوال من الجدول التكراري يجب ملاحظة أن تحديد فئة المنوال يتوقف على التكرار الحقيقي للفئات بالنسبة للجداول ذات أطوال الفئات غير المنتظمة ويقابل المنوال قمة المنحني دوماً ولتحديده نتبع الخطوات التالية:

١ ــ نرسم المدرج التكراري للفئة المنوالية والفئة السابقة والفئة اللاحقة بعد التعديل إذا كانت أطوال فئات الجدول مختلفة.

٢ ــ نصل بين رأس الفئة المنوالية الذي يمثل الحد الأدنى ورأس الفئة التي تليها الذي يمثل الحد الأدنى للفئة ما بعد المنوالية ثم نصل بين رأس الفئة المنوالية الذي يمثل الحد الأعلى ورأس الفئة التي تسبقها .

٣ ــ نسقط من نقطة التلاقي عموداً على المحور الأفقي والقيمة التي تقابل نقطة التلاقي تمثل قيمة المنوال.



شكل (٦ ــ ٢)

٤ ــ ٩ : خصائص المنوال :

١ ــ لا يتأثر المنوال بالقيم المتطرفة، لأنه يعتمد في حسابه على الفئتين السابقة واللاحقة.

٢ ــ إن قيمة المنوال ثابتة، لا تتغير، وتقابل قمة المنحنى دوماً لذلك يمكن استخدامه كمعيار للحكم على القيمة التمثيلية لبقية المتوسطات من أجل تحديد طبيعة التوزيع.

٣ ــ قليل الاستخدام بالمقارنة بالوسط الحسابي في الدراسات الإحصائية المعمقة نظراً لأن أغلب طرق حسابه تقريبية .

٤ ــ نلجأ إلى اعتماده كمقياس للمقارنة عندما يكون التوزيع ملتوياً باعتباره لا يتأثر بالقيم المتطرفة.

٤ ــ ١٠ : العلاقة التقريبية بين المتوسطات :

كنا قد بينا في السابق أن الوسط الحسابي للتوزيع التكراري عندما يكون عدد المشاهدات كبيراً بأنه النقطة التي يتوازن عندها التوزيع أو نقطة مركز ثقله، أما بالنسبة للوسيط فإنه عبارة عن القيمة التي تقسم مفردات التوزيع إلى قسمين متساويين، والمنوال هو القيمة التي تقابل قمة المنحني للتوزيع التكراري وتتساوى المتوسطات الثلاثة في التوزيعات المتماثلة أي :

$$\overline{X} = Med = Mod$$

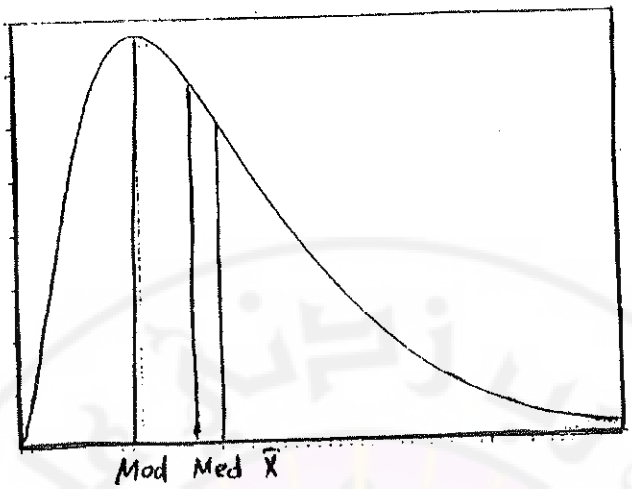
وتختلف قيم المقاييس الثلاثة عن بعضها بعضاً في التوزيعات غير المتماثلة أي الملتوية ويميز بين نوعين من الالتواء، الالتواء الموجب (نحو اليمين) والالتواء السالب (نحو اليسار).

ولفهم طبيعة الالتواء واستنتاج العلاقة التجريبية بين المتوسطات الثلاث، بفرض أضفنا فئة جديدة إلى النهاية العليا للتوزيع. تكرارها أقل من تكرار الفئة المنوالية سينتج لدينا توزيع جديد، يحافظ المنوال فيه على قيمته السابقة ويبقى مقابلاً لقمة المنحني في التوزيع، لأنه لا يتأثر إطلاقاً بالقيم المتطرفة، أما الوسيط سوف يتأثر بالقيمة الجديدة ولكن تأثره بها أقل من تأثر الوسط الحسابي لأن الذي يحدد قيمة الوسيط موقع المفردة وليس قيمتها، وعلى العموم نميز بين حالتين :

أ ــ إذا كان التوزيع موجب الالتواء تكون العلاقة على النحو التالي :

$$\overline{X} > Med > Mod$$

ويقع الوسيط في هذه الحالة على بعد ثلثي المسافة عن المنوال، وثلث المسافة عن الوسط الحسابي ويأخذ الشكل التالي :



ب ــ إذا كان التوزيع سالب الالتواء تكون العلاقة على النحو التالي :

$$\bar{X} < Med < Mod$$

وبما أن الوسيط موقعة دوماً في هذه بين الوسط الحسابي والمنوال، نكتفي بالمقارنة بين الوسط الحسابي والمنوال لمعرفة جهة الالتواء وهنا تأخذ العلاقة التجريبية لبيرسون بين المتوسطات تأخذ الشكل التالي :

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

وتستخدم هذه العلاقة في إيجاد قيمة المنوال أو قيمة الوسيط، بشرط التأكد من أن التوزيع موضوع الدراسة متماثل أو قريب من التماثل، ويمكن الاعتماد على هذه العلاقة في إيجاد قيمة الوسط، بينما الوسيط والمنوال قيمهما تقريبية ومبنية على فرضيات معينة، لذلك تستخدم هذه العلاقة في تقدير قيمة الوسط عندما تكون إحدى نهايتي التوزيع مفتوحة.

الوسط الهندسي
Geometric Mean

٤ ـ ١٣ : الوسط الهندسي :

يعرف الوسط الهندسي لمجموعة من القيم عددها (n) بأنه الجذر النوني لحاصل جداء هذه القيم بعضها ببعض، فلو كان لدينا عدد من المفردات الموجبة $x_1, x_2, x_n, \ldots$ فإن الوسط الهندسي لهذه المفردات يعطى بالعلاقة :

$$G = \sqrt[n]{x_1 . x_2 . x_3 \ldots\ldots x_n}$$

ـ حساب الوسط الهندسي لبيانات غير مبوبة :

لتوضيح ذلك لدينا التمرين التالي :

أوجد الوسط الهندسي لمجموعة القيم التالية :

$$100 , 100 , 300$$

$$G = \sqrt[3]{100 * 100 * 300} = 144,2$$

ويمكن حساب الوسط الهندسي باستخدام اللوغاريتمات لبيانات غير مبوبة :

$$G = (X_1, X_2, X_3, \ldots . X_2)^{1/n}$$

$$LogG = 1/n \, [Log(X_1 . X_2 . X_2 . . . X_2])$$

$$Log G = \frac{1}{n} (Logx_1 + Logx_2 + Logx_3 + \ldots . Logx_n)$$

$$LogG = (\sum Log\, x)/n$$

وبالتطبيق على التمرين السابق :

$$LogG = \frac{2+2+2.47}{3} = 2.16$$

$$LogG = 2.16 \Rightarrow G = 144,2$$

وبالاستناد إلى ذلك يمكن تعريف لوغاريتم الوسط الهندسي بأنه الوسط الحسابي للوغاريتمات المفردات.

تمرين :

لدينا سلسلة زمنية لإنتاج البترول الخام بإحدى الدول مقدراً بألف طن ولعدد من السنوات.

والمطلوب :

احسب الوسط الهندسي لمعدل النمو السنوي بالنسبة للعام السابق باستخدام أسلوب اللوغاريتمات.

السنوات	1991	1992	1993	1994	1995	1996
كميات الإنتاج	147.6	15562	1612	10676	1714	1969

الحل :

نقوم بإنشاء الجدول المساعد :

السنوات	كميات الإنتاج	النمو بالنسبة للعام السابق	Logx
1991	1470.6	--	--
1992	1556.2	1.058	0.0245
1993	1612	1.03538	0.0152
1994	1067.6	0.6622	-0.1792

1995	1714	1.6054	0.2055
1996	1969	1.1487	0.0602
Σ			0,1246

$$LogG = \sum Log\, {}^{x}\!/_{n} = \frac{0,1246}{5} = 0,0252$$

$$G = 10^{0.0252} \times 100 = 1,0597$$

وعليه يكون الوسط الهندسي لمعدل النمو السنوي يساوي ٥,٩٧% .

حساب الوسط الهندسي لبيانات مبوبة :

يُعطى من خلال العلاقة التالية :

$$LogG = \frac{\sum f(Logx)}{\sum f}$$

وبالعودة إلى التمرين السابق، في بحث الوسيط والمتعلق بمبيعات مخازن المستهلك والمطلوب: احسب الوسط الهندسي لمبيعات تلك المخازن؟

بإنشاء الجدول المساعد نجد أن :

فئات المبيعات	f	X	Logx	flogx
12-6	12	9	0.9954	11.448
12-18	١٨	15	10176	21.168
18-24	32	21	1.322	42.304
36-30	43	27	1.413	61.533
36-30	32	33	1.518	48.576
40-36	13	39	1.519	20.683
Σ	150			205.7

نطبق القانون :

$$LogG = \frac{\sum f(Logx)}{\sum f}$$

$$LogG = \frac{205.712}{150} = 1.3714$$

$$G = 23.518$$

٤ ــ ١٢ : خصائص الوسط الهندسي :

١ ــ أن تطبيقات الوسط الهندسي تنحصر في حساب معدلات النمو عموماً والأرقام القياسية ومعدلات الفائدة .

٢ ــ إن الوسط الهندسي، كالوسط الحسابي، محدد رياضياً بدقة وتدخل في حسابه جميع مفردات التوزيع وأن مجموع انحرافات لوغاريتمات المفردات عن لوغارتيم الوسط الهندسي يساوي الصفر أي :

$$\sum f (Log\ x - LogG) = 0$$

وأيضاً نجد أن :

$$LogG \cdot \sum f = \sum (f. Logx)$$

٣ ــ يعطي قيمة أكثر تمثيلاً من الوسط الحسابي عند التعامل مع توزيعات شديدة الالتواء.

٤ ــ يقتصر حسابه على القيم الموجبة فقط، كما يعتبر مفهوماً مجردا صعب الفهم والإدراك.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل الخامس

مقاييس التشتت

of Dispesion Measures

- مقدمة
- أهمية مقاييس التشتت
- المدى
- الانحراف الربيعي
- الانحراف المتوسط
- الانحراف المعياري
- أهمية الانحراف المعياري
- الالتواء والتفلطح
- تمارين غير محلولة

Damascus University


وقل رب زدني علما
جامعة دمشق
Damascus University

٥ ــ ١ : مقدمة :

سبق أن درسنا طرق عرض البيانات الإحصائية، وبناء التوزيعات التكرارية، وعرضها بيانياً وبعض خواصها، ثم حساب بعض مقاييس النزعة المركزية، غير أن هذه المقاييس لا تعطي فكرة كاملة عن حقيقة المجموعة التي يمثلها التوزيع، فعلى الرغم من تجمع المفردات حول قيمة معينة نجد أنها في الوقت نفسه تبتعد عن هذه القيمة، أو أن التكرارات تميل إلى التناقص عن تكرار مقاييس الموضع كلما اقتربنا من نهايتي التوزيع وهذا ما يعرف أو يدعى بالتشتت .

ويُعرف التشتت : بأنه تناثر أو انتشار قيم الظاهرة المدروسة حول أحد المتوسطات، وهو يقيس مقدار تباعد المفردات عن قيمة النزعة المركزية ويقاس التشتت بالمقاييس التالية: المدى الانحراف الربيعي، الانحراف المتوسط، الانحراف المعياري .

٥ ــ ٢ : أهمية مقاييس التشتت : قد تتساوى قيم المتوسطات لمجموعتين من البيانات بينما نلاحظ أن انتشار أو تشتت إحدى المجموعتين هو أقل أو أكثر من الأخرى لذلك نقول عن مفردات المجموعة أو العينة التي تظهر تشتتاً أقل بأن بياناتها أكثر تجانساً، وبالتالي فإن القيمة التمثيلية لمتوسطها تعتبر أكبر وخلاصة القول، إن مقاييس التشتت تعتبر مكملة ومتممة لمقاييس النزعة المركزية، فكلما قل التشتت، كلما ازدادت القيمة التمثيلية لمتوسطها والعكس صحيح.

فعلى سبيل المثال : كانت لدينا البيانات الفرضية التالية :

مجموعة (١) 500, 500, 500, 500, 500

مجموعة (٢) 480, 490, 500, 510, 520

من خلال هذا المثال نلاحظ أن $\overline{X} = 500$ لكل مجموعة إلاَّ أن انتشار أو تشتت المجموعة (١) حول المتوسط هو مساو إلى الصفر، بينما المجموعة (٢) فإن تشتتها بالقيمة المطلقة هو أكبر من الصفر. هذا يعني أن مقياس النزعة المركزية غير كاف للحكم على تماثل أو عدم تماثل المجموعتين، لذلك يجب أن نرفق مقاييس التشتت مع مقاييس النزعة المركزية للحكم على التشابه أو عدمه، وكما نلاحظ من المثال السابق فإن البيانات في المجموعة الأولى أكثر تجانساً من المجموعة الثانية بالإضافة إلى ذلك المجموعة (١) تعكس قيمة تمثيلية أكبر لمتوسطها (مقياس النزعة المركزية) من مقياس المجموعة (٢) من هنا نستنتج إن مقاييس التشتت تبرز بعض الميزات والخصائص التي لا تبرزها مقاييس النزعة المركزية.

إن فكرة مقاييس التشتت تقوم على قياس تباعد أو انتشار المفردات أو انتشارها حول مقياس النزعة المركزية، وهنا نميز بين مقاييس التشتت التي تبني فكرتها على أخذ الفرق بين قيم مفردتين (المدى، الانحراف الربيعي) وبين مقاييس التشتت التي تأخذ جميع المفردات بالحسبان (الانحراف المتوسط والانحراف المعياري) وتحسب مقاييس التشتت بشكل مطلق أو نسبي، ونستخدم التشتت المطلق عندما تتساوى قيم المتوسطات لمجموعتين أو أكثر من البيانات لإجراء المقارنة بينهما، وخلاف ذلك يستخدم التشتت النسبي للمقارنة بين المجموعات، وهو الأكثر استخداماً أثناء عمليات المقارنة، لأنه يقبل التفسير والمقارنة والذي يُعطى بالعلاقة:

$$\text{مقياس التشتت النسبي} = \frac{\text{مقياس التشتت المطلق}}{\text{مقياس النزعة المقابلة له}} \times ١٠٠$$

تمرين :

إذا كان متوسط الأجور اليومية لعمال أحد المصانع $\overline{X} = 500$ وحدة نقدية ومقياس التشتت المقابل ٥٠ وحدة نقدية، ومتوسط أجور العاملات $\overline{X} = 300$ وحدة نقدية ومقياس التشتت المقابل ٤٢ وحدة نقدية .

المطلوب :

أيهما يعرض تشتتاً أو انتشاراً أكبر أجور العمال أم أجور العاملات، ثم بين أي المتوسطات يعكس قيمة تمثيلية أكبر؟

الحل :

$$\text{التشتت النسبي لأجور العمال} = \frac{50}{500} \times 100$$

$$= ١٠٠ \%$$

$$\text{التشتت النسبي لأجور العاملات} = \frac{42}{300} \times 100$$

$$= ١٤ \%$$

للوهلة الأولى نلاحظ أن التشتت في أجور العمال هو أكبر من التشتت في أجور العاملات فيما لو تمت المقارنة بالقيم المطلقة، ونظراً لعدم تساوي المتوسطات فإننا نحصل على نتيجة مخالفة تماماً للنتيجة السابقة إذا تمت المقارنة النسبية، لذلك نلاحظ أن تشتت أجور العاملات (١٤%) أكبر من تشتت أجور العمال (١٠ %).

ويعتبر مقياس النزعة المركزية لأجور العمال أكثر تمثيلاً من مقياس النزعة المركزية لأجور العاملات، على اعتبار أن تشتت أجور العمال أقل من تشتت أجور العاملات.

تمرين :

إذا كان الوسط الحسابي ومقياس التشتت المقابل لأوزان وأطوالها، مجموعة من الأنابيب كالتالي:

متوسط الطول $\overline{X} = 160$ Cm

مقياس التشتت المقابل = 12 Cm

متوسط الوزن $\overline{X} = 60$ kg

مقياس التشتت المقابل = 3 kg

نلاحظ من خلال التمرين، أنه لا يجوز المقارنة بين القيم المطلقة لمقاييس التشتت وذلك لاختلاف الوحدات (Cm/ Kg) لذلك نلجأ لحساب التشتت النسبي :

مقياس التشتت النسبي للطول = $\frac{12}{160} \times 100 = 7.5\%$

مقياس التشتت النسبي للوزن = $\frac{3}{60} \times 100 = 5\%$

واضح ، أن مقياس التشتت النسبي للطول أكبر من مقياس التشتت للوزن، إلا أنه ليس لأي مقياس تشتت أي معنى أو مدلول إلا إذا ما قورن مع غيره من مقاييس التشتت المماثلة.

المدى
Rang

٥ ــ ٣ : المدى :

يعتبر المدى من أبسط مقاييس التشتت من حيث المفهوم، وهو عبارة عن الفرق بين قيمة أكبر مفردة وأصغر مفردة في التوزيع .

ــ حساب المدى لبيانات غير مبوبة :

لتوضيح ذلك لدينا التمرين التالي :

لدينا القيم التالية : 9, 15, 6, 23, 14, 19, 37

والمطلوب : احسب المدى :

$$R - 37 - 6 = 31$$

هذا يعني أن المدى هو الفرق بين أكبر وأصغر قيمة .

ــ حساب المدى لبيانات مبوبة: المدى هو الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى .

ويعطى المدى النسبي (R%) بالعلاقة التالية :

$$R\% = \frac{R}{X} \times 100$$

٥ ــ ٤ : خصائص المدى :

١ ــ يتأثر المدى بالقيم الحدية فقط ولا يأخذ بالحسبان جميع قيم المجموعة لذلك لا يكشف عن انتشار البيانات الواقعة بين القيم الحدية.

٢ ــ لا يمكن حساب المدى بدقة في التوزيعات التكرارية المفتوحة .

٣ ــ أغلب استخداماته في قياس تشتت الأسعار اليومية للسندات والأوراق المالية، ووصف الأحوال الجوية، ومراقبة جودة الإنتاج .

الانحراف الربيعي

Interquartile Deviation

٥ ـ ٥ : الانحراف الربيعي :

يعرف الانحراف الربيعي بأنه نصف المدى الربيعي الذي هو عبارة عن متوسط الفرق بين قيمة الربيع الثالث (الربيع الأعلى) والربيع الأول (الربيع الأدنى) .

المدى الربيعي : $Q=Q_3-Q_1$

الانحراف الربيعي : $$Q.D=\frac{Q_3-Q_1}{2}$$

الانحراف الربيعي النسبي: هو مقياس تشتت نسبي له أهمية في التحليل الاحصائي لمقارنة تشتت توزعين :

$$Q.D\%=\frac{Q.D}{Med}x100$$

أو

$$Q.D\%=\frac{Q_3-Q_1}{2Med}x100$$

حساب الانحراف الربيعي لبيانات غير مبوبة :

لتوضيح ذلك لدينا التمرين التالي :

لدينا القيم التالية والمرتبة تصاعدياً .

41, 45, 47, 49, 50, 52, 55, 60, 61, 63, 65, 67, 70

والمطلوب : حساب الانحراف الربيعي

ترتيب الربيع الأول $\frac{n+1}{4}=\frac{13+1}{4}=3.5$

قيمة الربيع الأول $\frac{47+49}{2}=48$

ترتيب الربيع الثالث : $\frac{3(n+1)}{4}=\frac{3(14)}{4}=10.5$

قيمة الربيع الثالث $\frac{63+65}{2}=64$

المدى الربيعي $Q_3-Q_1=64-48=16$

نصف المدى الربيعي $=\frac{Q_3-Q_1}{2}=\frac{16}{2}=8$

حساب الانحراف الربيعي لبيانات مبوبة :

ويتم من خلال العلاقة التالية : $Q.D=\frac{Q_3-Q_1}{2}$

وبالعودة إلى التمرين السابق (بحث الوسيط) فإن الانحراف الربيعي والانحراف الربيعي النسبي لمبيعات مخازن المستهلك :

$$Q.D=\frac{Q_3-Q_1}{2}=\frac{31.4-19.4}{2}=6$$

$$Q.D\%=\frac{6}{25.81}\times 100=23.2\%$$

أو

$Q.D\%=[\ \frac{31.4-19.4}{2}\ ]\div[\ \frac{19.4+31.4}{2}\ ]*100=23.6\%$ ليس لـــه أي مدلول ما لم يقارن بمقاييس تشتت مماثل من عينة أخرى.

٥ ــ ٦ : خصائص الانحراف الربيعي :

١ ــ إن الانحراف الربيعي هو مقياس التشتت الوحيد الذي يمكن حسابه عندما تكون إحدى نهايتي التوزيع مفتوحة.

٢ ــ إن أسلوب المدى الربيعي يهمل نصف المفردات في التوزيع، حيث أن ربع هذه المفردات أقل من قيمة الربيع الأول وربع المفردات أكبر من الربيع الثالث .

٣ ــ الانحراف الربيعي يعطي فكرة كاملة عن التشتت، لأنه يهتم بنصف المفردات فقط الواقعة بين Q_3, Q_1 .

٤ ــ إن مقياس النزعة المركزية المقابل للانحراف الربيعي هو الوسيط .

الانحراف المتوسط
Mean Deviation

٥ ــ ٧ : الانحراف المتوسط :

هو متوسط الانحرافات المطلقة للمشاهدات عن وسطها الحسابي $\overline{X}$ ويرمز له بـــ M.D ويعطى بالعلاقة التالية :

$$M.D = \frac{\sum |x - \bar{x}|}{n}$$

حسابه لبيانات غير مبوبة :

تمرين : لدينا القيم التالية 40, 45, 50, 55, 60, 65, 70

المطلوب : أوجد قيمة الانحراف المتوسط لهذه البيانات :

$$\overline{X} = \frac{\sum x}{n} = \frac{385}{7} = 55$$

$$M.D = \frac{|40-55| + ... + |70-55|}{7}$$

$$= M.D = \frac{60}{7} = 8.57$$

حسابه لبيانات مبوبة :

يعطى بالعلاقة التالية :

$$M.D = \frac{\sum f|x-\overline{x}|}{\sum f}$$

تمرين :

أوجد الانحراف المتوسط لدرجات عينة عشوائية من الطلاب في أحد المقررات :

فئات الدرجات	50-40	60-50	70-60	80-70	90-80	100-90
عدد الطلاب	2	9	15	11	2	1

الحل :

ننشئ جدول الحسابات المساعد :

فئات الدرجات	f	X	U	fu	$\|x-\overline{x}\|$	$f\|x-\overline{x}\|$
50-40	2	45	-2	-4	21.25	42.5
60-50	9	55	-1	-9	11.25	101.25
70-60	15	65*	0	0	1.25	18.75

80-70	11	75	1+	11+	1.25	96.25
90-80	2	85	2+	4+	18.75	37.5
100-90	1	95	3+	3+	28.75	28.75
Σ	40			5+		325.5

حساب الوسط الحسابي .

$$\overline{X} = A + \frac{\sum f.u}{\sum f} \times C = 65 + \frac{+5}{40} \times 10 = 66.25$$

علماً بأن A = 65

الانحراف المتوسط :

$$M.D = \frac{\sum f|x - \bar{x}|}{\sum f}$$

$$M.D = \frac{325.5}{40} = 8.12$$

وليس لهذه القيمة أي مدلول ما لم تقارن بنفس المقياس لعينة أخرى.

٥-٨ : خصائص الانحراف المتوسط :

١ ــ من الممكن حساب الانحراف المتوسط بدلالة الوسيط بدلاً من الوسط الحسابي.

٢ ــ يعتبر هذا المقياس أكثر شمولاً من سابقيه (المدى، الانحراف الربيعي) لأنه يأخذ جميع قيم التوزيع في حساب التشتت .

٣ ــ عدم اتصافه بالمفهوم الجبري (لإهماله الإشارات الجبرية عند حساب الانحرافات).

٤ ــ لا يمكن حسابه إذا كانت إحدى فئات الجدول التكراري مفتوحة.

الانحراف المعياري
Standard deviation

٥ ــ ٩ : الانحراف المعياري :

هو معلمة أساسية من معالم التوزيعات الاحتمالية، باعتبار أنه يمكن التعرف على طبيعة التوزيعات الاحتمالية من خلال معلمتين أساسيتين هما :

الوسط الحسابي : الذي يسمح لنا بتحديد نقطة ثقل التوزيع .

الانحراف المعياري : يسمح لنا بالتعرف على الكيفية التي تتوزع بها المفردات حول نقطة الوسط الحسابي .

ويعرّف الانحراف المعياري لمجموعة من البيانات: بأنه الجذر التربيعي للتباين، والتباين هو الوسط الحسابي لمربعات انحراف المفردات عن وسطها الحسابي ويعد من أكثر مقاييس التشتت أهمية في التحليل الإحصائي .

حسابه لبيانات غير مبوبة :

أ ــ الأسلوب المباشرة : إذا كان لدينا عدد من المشاهدات :

$$X_1, X_2, X_3, \ldots X_n .$$

ووسطها الحسابي $\bar{X}$ فإن الانحراف المعياري لهذه المشاهدات يعطى بالعلاقة التالية:

$$S = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + (x_n - \bar{x})^2}{n}}$$

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

ويمكن تبسيط العلاقة السابقة بفك الأقواس فنحصل على الصيغ التالية :

$$S = \sqrt{\frac{\sum x^2}{n} - \left[\frac{\sum x}{n}\right]^2}$$

٢ ــ أسلوب الانحرافات :

نحن نعلم أن :

$$\bar{x} = A + \frac{\sum d}{n}$$

نعوض X بالصيغة السابقة للانحراف المعياري :

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

نحصل على العلاقة :

$$S = \sqrt{\frac{\sum d^2}{n} - \left[\frac{\sum d}{n}\right]^2}$$

حيث أن $d = X - A$ وأن A: وسط حسابي مفترض

تمرين :

ليكن لدينا البيانات التالية والمتعلقة بعدد الساعات الأسبوعية لدراسة أحد الطلاب في الأسابيع الأخيرة قبل بدء الامتحان : 15، 9 ، 14، 6، 23، 19، 33 .

المطلوب :

حساب الانحراف المعياري بمختلف الأساليب السابقة :

الحل :

نعد الجدول التالي لحساب الانحراف المعياري لهذه البيانات :

رقم الأسبوع	ساعات الدراسة X	$(x-\bar{x})^2$	X^2	d= x-A	d^2
1	15	4	225	0	0
2	9	64	81	-6	36
3	14	9	196	-1	1
4	6	121	36	-9	81
5	23	36	529	+8	64
6	19	4	361	+4	16
7	33	256	1089	+18	324
Σ	119	496	2517	14	522

الأسلوب المباشر :

$$\bar{x} = \frac{\sum x}{n} = \frac{119}{7} = 17$$

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

ساعة

$$S = \sqrt{\frac{494}{7}} = 8.4$$

$$S = \sqrt{\frac{\sum d^2}{n} - \left[\frac{\sum d}{n}\right]^2} = \sqrt{\frac{2517}{7} - \left[\frac{119}{7}\right]^2} = 8.4 \text{ ساعة}$$

٢ ــ أسلوب الانحرافات :

$$S = \sqrt{\frac{\sum d^2}{n} - \left[\frac{\sum d}{n}\right]^2} = \sqrt{\frac{522}{7} - \left[\frac{14}{7}\right]^2} = 8.4$$

علماً بأن الوسط الفرضي : $A = 15$

الانحراف المعياري لبيانات مبوبة :

١ ــ الأسلوب المباشر :

يعطى بالعلاقة وذلك حسب تعريف الانحراف المعياري مباشرة .

$$S = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}}$$

تمرين :

بالعودة إلى التمرين المتعلق بأجور العمال الشهرية لعينة عشوائية من العمال حجمها $n=100$ عامل في إحدى الصناعات البترولية (الأجور بمئات الوحدات النقدية).

المطلوب :

أوجد الانحراف المعياري بالأسلوب المباشر .

فئات الدرجات	f	X	f.x	$(x-\bar{x})^2$	$F(x-\bar{x})^2$
120-80	10	100	1000	4900	49000
160-120	25	140	3500	900	22500
200-160	50	180	9000	100	5000
240-200	10	220	2200	2500	25000
280-240	5	260	1300	8100	40500
Σ	100		17000		14200

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{17000}{100} = 170$$

$$S = \sqrt{\frac{142000}{100}} = 37.68$$

٢ ــ أسلوب الانحرافات :

ويعطى بالعلاقة التالية :

$$S = \sqrt{\frac{\sum fd^2}{\sum f} - \left[\frac{\sum fd}{\sum f}\right]^2}$$

حيث $d=X-A$ وسط حسابي مفترض .

فئات الدرجات	F	X	$d=X-A$	d^2	$f.d^2$	f.d
120-80	10	100	-80	46000	64000	-800
160-120	25	140	-40	1600	40000	-1000
200-160	50	180	0	0	0	0
240-200	10	220	+40	1600	16000	+400
280-240	5	260	+80	6400	3200	+400
Σ	100				152000	10000

اعتبرنا الوسط الفرضي $A=180$.

$$S = \sqrt{\frac{152000}{100} - \left(\frac{-1000}{100}\right)^2} = 37.68$$

٣ ــ أسلوب الانحرافات المختصرة :

ويعطى بالعلاقة التالية :

$$S = \sqrt{\frac{\sum fu^2}{\sum f} - \left[\frac{\sum fu}{\sum f}\right]^2}.C$$

فئات الأجور	f	X		U^2	$f.u^2$	f.u
120-80	10	100	-2	4	40	-20
160-120	25	140	-1	1	25	-25
200-160	50	180	0	0	0	0
240-200	10	220	+1	1	10	+10
280-240		260	+2	4	20	+10
Σ		100			95	-25

وباعتبار أن A=180

$$S = \sqrt{\frac{95}{100} - \left(\frac{-25}{100}\right)^2}.40 = 37.68$$

الانحراف المعياري النسبي

٥ ــ ١٠ : معامل الاختلاف Coefficient of radiation :

يعرف معامل الاختلاف بأنه حاصل قسمة الانحراف المعياري على الوسط الحسابي مضروباً بـ ١٠٠ ونرمز له بـ C.V % .

$$C.V\% = \frac{S}{\overline{X}} \times 100$$

ويستخدم لمقارنة تشتت توزعين، مختلفين بالاعتماد على وحدات قياس ثابتة في التوزيعين.

تمرين :

بالعودة إلى التمرين السابق والمطلوب أوجد الانحراف المعياري النسبي (معامل الاختلاف، وفسره علماً بأن $\overline{X} = 170$

$$C.V\% = \frac{37.68}{170} \times 100 = 22.16\%$$

والتفسير أن نسبة انتشار المفردات حول وسطها الحسابي يساوي ٢٢,١٦% من قيمة الوسط الحسابي. وليس لهذا المقياس من مدلول ما لم يقارن بالمقياس نفسه لعينة أخرى.

تمرين :

سحبت عينة عشوائية n=100 عامل من الصناعات البترولية، وكان متوسط الأجور اليومية للعامل في هذه العينة ٢٠٠ وحدة نقدية وبانحراف معياري 20 وحدة نقدية، كما سحبت عينة من حجم مماثل من صناعة أخرى

فكان متوسط الأجور اليومية للعمال في هذه الصناعة 250 وحدة نقدية وانحراف معياري قدره 30 وحدة نقدية .

والمطلوب : بيان أي الأجور في الصناعتين تعرض تشتتاً أكبر ثم بيان أيهما يعكس ثقة أكبر بمقياس النزعة المركزية .

الحل :

$$C.V\% = \frac{S}{\overline{X}} \times 100 = \frac{20}{200} \times 100 = 10\% \quad \text{الصناعة الأولى}$$

$$C.V\% = \frac{S}{\overline{X}} \times 100 = \frac{30}{250} \times 100 = 12\% \quad \text{الصناعة الثانية}$$

من خلال المقارنة بين معاملي الاختلاف في الصناعتين : نجد أن 10%<12% وبالتالي فإن أجور العمال في الصناعة البترولية حول المتوسط والبالغ $\overline{X} = 200$ أكثر تجانساً من أجور العمال حول المتوسط والبالغ $\overline{X} = 250$ في الصناعة الأخرى، وبالمقارنة نجد أن تشتت أجور العمال في الصناعة البترولية أقل من مثيله في الصناعة الأخرى وبالتالي فإن الوسط الحسابي في الصناعة البترولية أكثر تمثيلاً لأجور العمال في هذه الصناعة من مثيله في الصناعة الأخرى.

٥ ــ ١١ : خصائص الانحراف المعياري :

١ ــ يعد الوسط الحسابي والانحراف المعياري المعلمتين الأساسيتين في النظرية الإحصائية لاسيما في التوزيعات الاحتمالية وذلك من أجل تحديد طبيعة التوزيع (شكل المنحنيات).

٢ ــ أن قيمة الانحراف المتوسط تساوي ٤/٥ من قيمة الانحراف المعياري، وذلك في التوزيعات المتماثلة والقريبة من التماثل .

$$M.D = \frac{4}{5}.S$$

٣ ــ إن مقياس التشتت المقابل للوسط الحسابي هو الانحراف المعياري

٤ ــ لا يمكن حساب الانحراف المعياري من جدول تكراري إحدى نهايتيه مفتوحة.

٥ ــ يأخذ بالحسبان انتشار أو تباعد جميع المفردات، لذلك فهو أكثر شمولاً من أسلوبي المدى والانحراف الربيعي .

٦ ــ قيمته العددية أكبر من أي مقياس تشتت آخر ما عدا المدى .

٧ ــ له مزايا وخصائص رياضية وجبرية غير موجودة في أي مقياس آخر من مقاييس التشتت.

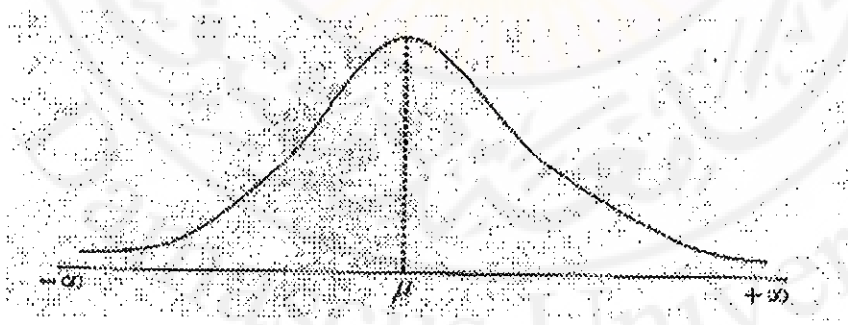
٥ ــ ١٢ : أهمية الانحراف المعياري :

يعتبر التباين الذي هو مربع الانحراف المعياري مهماً للغاية في التحليل الإحصائي، وهو يستعمل في أغلب التطبيقات وبصورة خاصة في مجال الاستدلال الإحصائي، وقد اكتسب أهميته خصوصاً من تمتعه بالميزات الجبرية، فلو كانت قيمته صغيرة دل ذلك على تمركز المعطيات حول الوسط الحسابي، وعلى العكس تماماً عندما تكون قيمته كبيرة فيدل على عدم التجانس وعلى التشتت الأكبر عن قيمة الوسط الحسابي لقيم الظاهرة المدروسة. ولعل من أهم خواص الانحراف المعياري، تلك المتعلقة بنظرية (تشيبيشيف) وتطبيقاتها على التوزيع الطبيعي المعياري .

التوزيع الطبيعي : يُعد التوزيع الطبيعي من أهم التوزيعات العشوائية المستمرة، نظراً لاستخداماته الواسعة والمتعددة في مختلف المجالات

التطبيقية، إذ يتم تمثيل التوزيع الطبيعي بواسطة منحنى يأخذ شكل الجرس ومعرف بمؤشرين هما : الوسط الحسابي $\overline{X}$ والانحراف المعياري S_x (للعينة)، أو الوسط الحسابي μ والانحراف المعياري σ_x للمجتمع، وبالاضافة لذلك نذكر بعضاً من الخصائص الأخرى التي يتصف بها هذا التوزيع، ومنها :

- أنه متماثل ومتناظر، فالمساحات الواقعة على طرفي الوسط الحسابي متساوية، أي لو أطبقنا الطرف الأيمن منه على الطرف أيسر لانطبقت على بعضها الآخر.
- النهايات الطرفية تمتد إلى مالا نهاية، أي أنها تكاد تلامس المحور الأفقي، فهي متقاربة وموازية له؛
- يمثل منحني التوزيع الطبيعي عدداً غير محدود أو عدداً كبيراً من الحالات.
- إن منحني كثافة الاحتمال متناظر حول الوسط الحسابي ويمتد من $[+\infty, -\infty]$ الذي يرمز له غالباً بـــ $N(\mu, \sigma_x)$ أو $N(\overline{X}, S_x)$ ، فالشكل رقم (١٤) يبين لنا ذلك، كالآتي.



الشكل رقم (١٤) : منحني كثافة الاحتمال للتوزيع الطبيعي

إن تابع كثافة الاحتمال يأخذ الشكل الآتي :

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} * e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma x}\right)^2}$$

إن المساحة المحصورة بين المنحني ومحور السينات تساوي الواحد الصحيح، أي:

$$\int_{-\infty}^{+\infty} f(x).d = \int_{-\infty}^{+\infty}\left[\frac{1}{\sigma_x \sqrt{2\pi}} * e^{-\frac{1}{2}\left(\frac{X1-\mu}{\sigma x}\right)^2}\right] * d_x = 1$$

وعندما نعبر عن Xi بدلالة الوحدات المعيارية، نجري تحويلاً للمتحول العشوائي X_i على الشكل الآتي :

ـ بدلالة المجتمع

$$Z = \frac{|x_i - \mu|}{\sigma_x}$$

ـ بدلالة العينة

$$Z = \frac{|X_i - \overline{X}|}{S_x}$$

وبتعويض قيمة Z في كثافة احتمال التوزيع الطبيعي أعلاه، نجد :

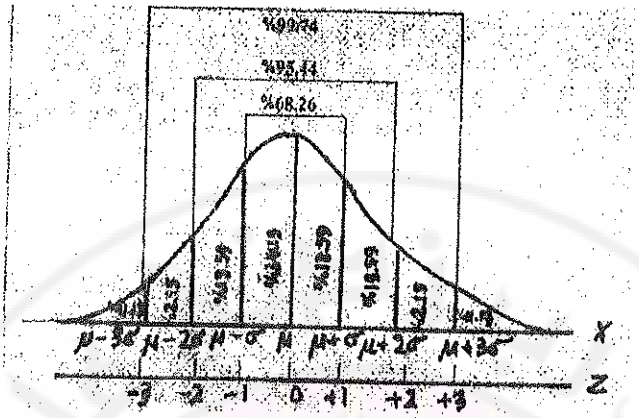
$$\int_{-\infty}^{+\infty}\left[\frac{1}{\sigma_x \sqrt{2\pi}} * e^{-\frac{1}{2}\left(\frac{X1-\mu}{\sigma x}\right)^2}\right] * d_x = \int_{-\infty}^{+\infty}\left[\frac{1}{\sigma_x \sqrt{2\pi}} * e^{\frac{Z^2}{2}}\right] * \frac{d_x}{d_z} = 1$$

وبذلك نحصل على متحول عشوائي جديد يدعى بالمتحول العشوائي المعياري، وسطاؤه: الوسط الحسابي المساوي للصفر، وتباينه يساوي الواحد الصحيح، وكثافة الاحتمال تعطى، كالآتي:

$$\int_{-\infty}^{+\infty} f(z)d_z = \int_{-\infty}^{+\infty}\left[\frac{1}{\sqrt{2\pi}} * e^{\frac{z^2}{2}}\right] * \frac{d_x}{d_z} = 1$$

حيث تدعى Z بالدرجة المعيارية، وهي عبارة عن إمكانية تحويل أي قيمة خاضعة للتوزيع الطبيعي إلى قيمة طبيعية معيارية (خاضعة لتوزيع طبيعي معياري)، والتي تستخدم في مقارنة المشاهدات من مجموعات مختلفة، وبعبارة أخرى هي درجة ابتعاد قيمة X_i عن وسطها الحسابي $\overline{X}$ بالانحرافات المعيارية .

واعتماداً على تابع كثافة الاحتمال للتوزيع الطبيعي المعياري، يمكننا معرفة نسبة أو احتمال أو عدد القيم التي توجد في مجال محدد حول الوسط الحسابي، في حال كان التوزيع التكراري خاضعاً للتوزيع الطبيعي وسطاؤه (N ($\overline{X}$ ، S_x) أو N (μ, σ_x)) ، أي : ($\overline{X} = 0, \sigma_{x=1}$)، فالتمثيل البياني لهذا التوزيع موضح في الشكل رقم ١٥)، كالآتي:



الشكل رقم (١٥) : شكل المنحني الطبيعي مع تحديد المساحات

وبالنظر إلى الشكل رقم (١٥) وإلى جدول المساحات الواقعة تحت المنحني الطبيعي في الجداول الإحصائية الملحقة، نذكر بعض المجالات العددية الشهيرة، ومنها :

$$\Pr\{\bar{X} - 1S_x < X < \bar{X} + 1S_x\} = 68.27\%$$

$$\Pr\{\bar{X} - 1.96S_x < X < \bar{X} + 1.96S_x\} = 95\%$$

$$\Pr\{\bar{X} - 2S_x < X < \bar{X} + 2S_x\} = 95.45\%$$

$$\Pr\{\bar{X} - 2.58S_x < X < \bar{X} + 2.58S_x\} = 99\%$$

$$\Pr\{\bar{X} - 3S_x < X < \bar{X} + 3S_x\} = 99.73\%$$

مثال: أخذت عينة عشوائية من ١٠٠٠ موظف، فتبين أن متوسط مرتبهم الأسبوعي ٣٠٠٠ وحدة نقدية، وبانحراف معياري قدره ٣٠٠ وحدة نقدية، فإذا علمت بأن مرتبات الموظفين في المجتمع الذي أخذت منه هذه العينة تتوزع توزعاً طبيعياً،

والمطلوب:

١ ــ أوجد عدد الموظفين الذين يتراوح أجرهم الشهري بين ٢٧٠٠ وحدة نقدية و ٣٦٠٠ وحدة نقدية .

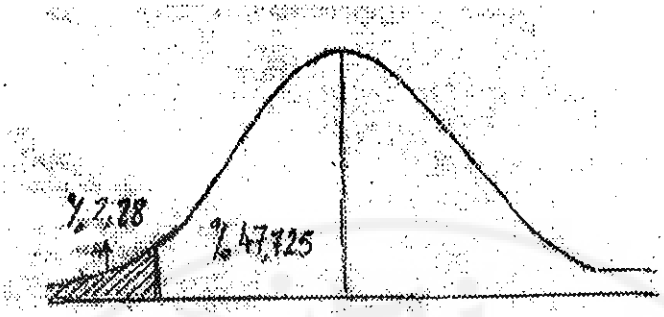
٢ ــ أوجد احتمال أن يكون موظفاً مرتبه الأسبوعي يقل عن ٢٤٠٠ وحدة نقدية ؟.

٣ ــ ما هي نسبة الموظفين الذين يزيد مرتبهم الأسبوعي عن ٢٤٠٠ وحدة نقدية؟

٤ ــ ما هو احتمال أن يكون المرتب الأسبوعي لموظف ما أكبر من ٣٣٠٠ وحدة نقدية.

الحل :

١ ــ لإيجاد عدد الموظفين الذين يتراوح مرتبهم الأسبوعي مابين ٢٧٠٠ و ٣٦٠٠ وحدة نقدية، وجب علينا ن نوجد المساحة المحصورة تحت المنحنى الطبيعي وبين هاتين القيمتين، بعد تحويلهما إلى قيمتين طبيعيتين معياريتين، ثم نقوم بضرب هذه المساحة في عدد الموظفين الكلي (حجم العينة المدروسة)، أي :



$$\Pr\left\{\frac{x_i-\bar{x}}{S_x}\le X\le\frac{x_i-\bar{x}}{S_x}\right\}=\Pr\left\{\frac{2700-3000}{300}\le X\le\frac{3600-3000}{300}\right\}$$

$$\Pr\{-1\le X\le 2\}=\Pr(0\le X\le -1)+\Pr(0\le X\le 2)$$

$$=\frac{0.6827}{2}+\frac{0.9545}{2}=0.34153+0.47725=0.8186$$

عدد الموظفين = المساحة التي أوجدناها مضروبة في حجم العينة ، أي ٠,٨١٨٦ × ١٠٠٠ = ٨١٨,٦ موظفاً ، تقريباً ٨١٩ موظفاً؛

وبطريقة أخرى ، نوجد مدى ابتعاد ٣٣٠٠ عن ٣٠٠٠ بالانحرافات المعيارية من خلال تحويل القيمة ٣٣٠٠ إلى قيمة طبيعية معيارية، أي :

$$Z=\frac{|X_i-\bar{X}|}{S_x}=\left|\frac{2700-3000}{300}\right|=|-1|$$

فهي تحصر ما مساحته تحت المنحنى الطبيعي ما مقداره $\frac{\%68.27}{2} = \%34.135 = 0.34135$ ثم نوجد مدى ابتعاد القيمة الثانية وبالطريقة نفسها، أي :

$$Z = \frac{|X_i - \bar{X}|}{S_x} = \left|\frac{3300 - 3000}{300}\right| = |2|$$

فهي تحصر ما مساحته تحت المنحنى الطبيعي ما قيمته

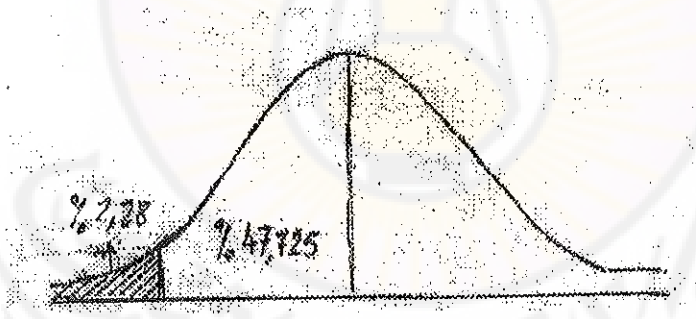
$$\frac{\%95.45}{2} = \%47.725 = 0.47725$$

فالمساحة المطلوبة هي مجموع المساحتين مابين [-1σx, +2σx]، أي :

٠,٤٧٧٢٥ + ٠,٣٤١٣٥ = ٠,٨١٨٦، وعليه يكون :

عدد الموظفين = ٠,٨١٨٦ * ١٠٠٠ = ٨١٨,٦ = ٨١٩ موظفاً .

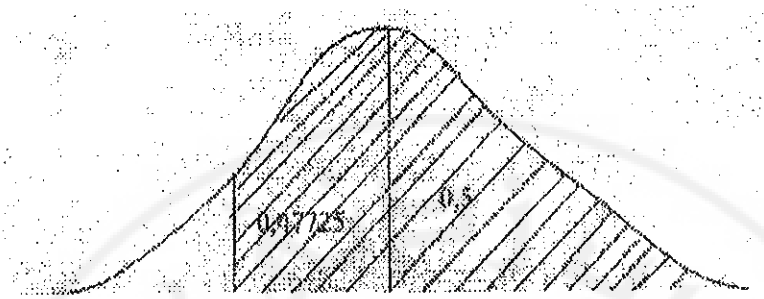
٢ ـــ إن احتمال أن يكون موظفاً مرتبه الأسبوعي يقل عن ٢٤٠٠ وحدة نقدية، هو :



$$\Pr\left\{X \le \frac{x_i - \bar{x}}{S_x}\right\} = \Pr\left\{X \le \frac{2400 - 3000}{300}\right\} = \Pr(-\infty \le X \le -2)$$

$$\Pr\{0 \le X \le -\infty\} = \Pr(0 \le X \le -2) = 0.5 - \frac{0.9545}{2} = 0.5 - 0.47725 = 2.275\%$$

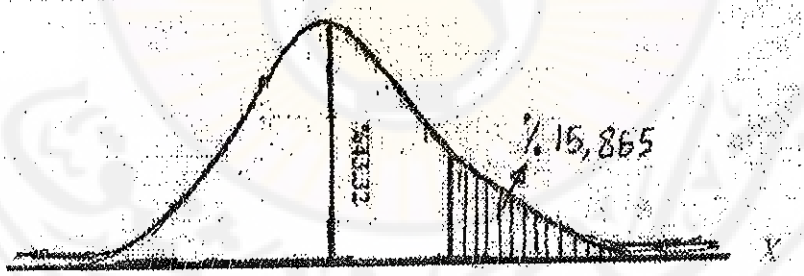
٣ ـــ إن نسبة أن يكون موظفاً مرتبه الأسبوعي يزيد عن ٢٤٠٠ وحدة نقدية، هي :



$$\Pr\left\{X \le \frac{x_i - \bar{x}}{S_x}\right\} = \Pr\left\{X \le \frac{2400-3000}{300}\right\} = \Pr(-2 \le X \le +\infty)$$

$$\Pr\{-2 \le X \le 0\} = \Pr(0 \le X \le +\infty) = 0.5 + \frac{0.9545}{2} = 0.5 + 0.47725 = 0.977725 = 97.77\%$$

٥ ـــ إن احتمال أن يكون المرتب الأسبوعي لموظف ما أكبر من ٣٣٠٠ وحدة نقدية، هو:



$$\Pr\left\{X \le \frac{x_i - \bar{x}}{S_x}\right\} = \Pr\left\{X \le \frac{3300-3000}{300}\right\} = \Pr(0 \le X \le +\infty) - pr(1 \le X \le +\infty)$$

$$= 0.5 - \frac{0.6827}{2} = 0.5 - 0.34135 = 0.15865 = 15.865\%$$

مقياس الإلتواء (Skewness)

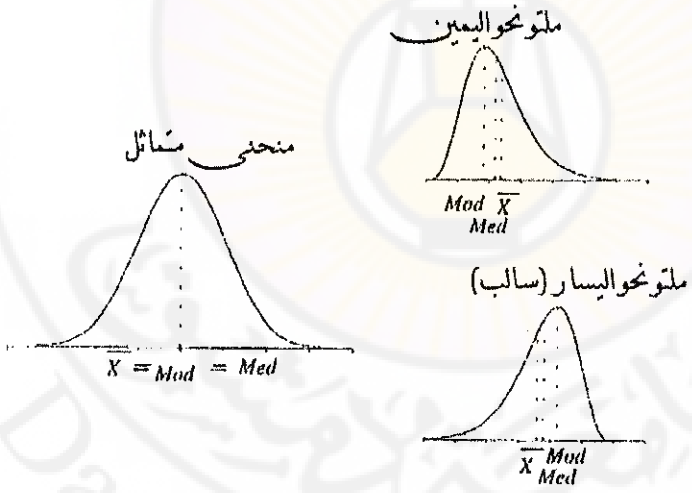
٥ ــ ١٣ : الالتواء : عند دراستنا لمقاييس النزعة المركزية ، تعرضنا لمفهوم الالتواء، وتحديد طبيعة التوزيع وبينا أنه في التوزيع المتماثل تتساوى المتوسطات الثلاثة الوسط والوسيط والمنوال أي : $\overline{X} = Med = Mod$

وفي التوزيعات غير المتماثلة : (الملتوية) يبدو أن هناك نوعين من الإلتواء

(١) الإلتواء نحن اليمين : $\overline{X} > Med > Mod$

(٢) الإلتواء نحو اليسار : $\overline{X} < Med < Mod$

وبيناً أنه لتحديد جهة الإلتواء يكتفى بالمقارنة بين مقياسين فقط وهما الوسط الحسابي والمنوال والشكل (٥ ــ ١) يوضح ذلك.



شكل (٥ ــ ١)

إن الانحراف المعياري يسمح لنا بالتعرف على الكيفية التي تتوزع بها المفردات حول نقطة الوسط الحسابي، وبالتالي التعرف على شكل المنحنى واتساعه، وقد بينا أنه لتحديد جهة الإلتواء نحسب الفرق بين الوسط

والمنوال فإذا كان $\overline{X} - Mod > 0$ كان الالتواء نحو اليمين موجب إما إذا كان $\overline{X} - Mod < 0$ كان الإلتواء نحو اليمين موجب إما إذا كان $\overline{X} - Mod < 0$ الإلتواء نحو اليساء سالب، ولتحديد شدة الإلتواء يتم ذلك من خلال معامل "بيرسون":

$$SK = \frac{(\bar{x} - Mod)}{S}$$

أو

$$SK = \frac{3(\bar{x} - Med)}{S}$$

إذا كان SK=0 فهذا يعني أن التوزيع التكراري متماثل.

SK<0 فهذا يعني أن التوزيع التكراري ملتوٍ نحو اليسار (سالب).

SK>0 فهذا يعني أن التوزيع التكراري ملتوٍ نحو اليمين (موجب).

تمرين :

حصلنا من خلال عينة لدرجات الطلاب في أحد المقررات على البيانات التالية:

$$, S = 10 \ \overline{X} = 65 \text{Med} = 63, \text{Mod} = 73,$$

والمطلوب : حدد جهة الإلتواء بالاعتماد على المنوال :

$$SK = \frac{(\bar{x} - Mod)}{S}$$

$$SK = \frac{65 - 73}{10} = -0.8$$

وبالتالي نجد أن SK < 0 والإلتواء سالب نحو اليسار .

التفلطح Kurtosis

٥ ــ ١٤ : **التفلطح** : هو مقياس يقيس درجة ارتفاع أو انخفاض أي منحنى توزع تكراري بالنسبة للمنحنى الطبيعي، علماً بأن المنحنى الطبيعي هو توزيع متماثل حول وسطه الحسابي، ويتصف بعلاقات نسبية ثابتة بين اتساعه وارتفاعه ويأخذ شكل الجرس المقلوب ويرمز له $N(x,\delta^2)$ ويرمز إلى التفلطح بـــ (K)، ويحسب من العلاقة التالية :

$$K = \frac{M^4}{S^4}$$

M^4 : العزم الرابع حول الوسط الحسابي، ويحسب من العلاقات التالية :

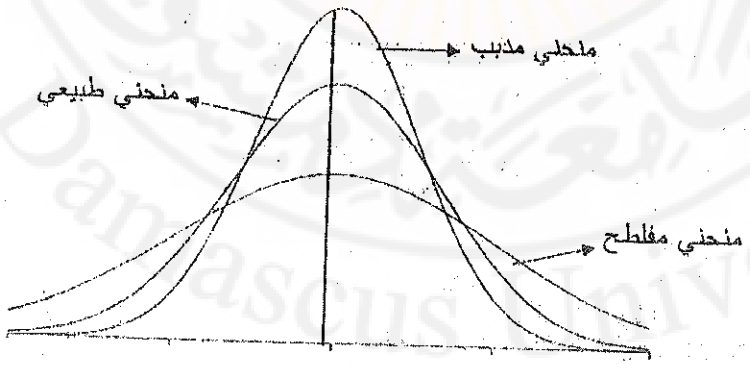
(١) بيانات غير مبوبة : $$M_4 = \frac{\sum (x-\bar{x})^4}{n}$$

(٢) بيانات مبوبة : $$M_4 = \frac{\sum f(x-\bar{x})^4}{n}$$

ــ إذا كان معامل التفلطح K=3 فإن التوزيع التكراري طبيعي .

ــ أما إذا كان K <3 فإنه يعني أن التوزيع هو مفلطح أو منبسط.

ــ أما إذا كان K>3 فإنه يعني أن التوزيع هو مدبب، أي أن قيمة المنحنى المدروس أعلى من حده قمة المنحنى الطبيعي.



شكل (٥ ــ ٢)

تمارين غير محلولة

للفصلين الرابع والخامس

مسألة (١) :

أخذت عينة عشوائية من مخازن بيع المستهلك في مدينة دمشق وكان التوزيع التكراري لمبيعاتها على النحو التالي :

المبيعات بالألف	40-30	50-40	60-50	70-60	80-70	Σ
عدد المخازن	17	77	141	55	10	**300**

والمطلوب:

١ ــ أوجد قيمة مقياسين من مقاييس النزعة المركزية وفسرهما.

٢ ــ أخذت عينة عشوائية مماثلة من مخازن بيع المستهلك في مدينة حلب، فوجد أن وسيط مبيعاتها (٦٠٠٠٠) ليرة ومداها الربيعي (١٢٥٠٠) ليرة

والمطلوب :

أي المخازن مبيعاتها ذات تشتت أكبر، المخازن في مدينة دمشق أم المخازن في مدينة حلب؟ برر إجابتك إحصائياً.

٣ ــ إذا رغبنا في تقدير مبيعات المخازن في المستقبل من خلال النتائج التي توصلنا إليها من الفقرات السابقة، هل نثق بصورة أقوى عند التقدير بمقياس النزعة لمبيعات المخازن في دمشق أم في حلب ولماذا ؟

مسألة (٢) :

أخذت عينة عشوائية من الأسر القاطنة في مدينة ما فكانت دخولها موزعة وفقاً للجدول التكراري الآتي، الدخول مقدره بآلاف الوحدات :

فئات الدخول	50-56	56-62	62-68	68-74	74-80	Σ
عدد الأسر	10	22	45	17	6	**100**

والمطلوب :

١ ــ أوجد الوسط الحسابي للدخول لهذه الأسر بالأسلوب الذي تراه مناسباً.

٢ ــ أوجد مقياس التشتت المقابل لمقياس النزعة المركزية المحسوب في الطلب الأول.

٣ ــ أخذت عينة عشوائية من حجم مماثل من أسر مدينة ثانية. فتبين أن الوسط الحسابي للدخول في هذه المدينة $\overline{X} = 52$ وحدة ومقياس التشتت المقابل (٦,٥) وحدة.

والمطلوب: أي البيانات تعرض تشتتاً أقل ؟.

مسألة (٣) :

أخذت عينة عشوائية من الأسر القاطنة في مدينة ما فكانت دخولها موزعة وفق الجدول التكراري التالي :

فئات الدخل	30-25	35-30	40-35	44-40	50-40	Σ
عدد الأسر	**3**	**17**	**25**	**13**	**2**	**60**

والمطلوب :

١ ــ ما هو أفضل مقياس للنزعة المركزية والتشتت لبيانات هذا الجدول ثم اذكر محاسن هذا المقياس.

٢ ــ حدد طبيعة التوزيع لبيانات الجدول السابق مدعماً إجابتك بالحسابات اللازمة.

٣ ــ احسب معامل الالتواء وماذا تستنتج ؟

٤ ــ احسب قيمة الربيع الأول والثالث وفسرهما في ضوء بيانات الجدول السابق.

مسألة (٤) :

أخذت عينة عشوائية من ١٠٠ مخزن تجاري في المدينة الأولى فكان التوزيع التكراري لمبيعاتها بالاف الوحدات النقدية :

المبيعات	10-14	14-18	18-22	22-26	26-30	$\sum$
التكرارات	8	10	35	30	17	100

وقد أخذت عينة عشوائية أخرى من الحجم نفسه من محلات تجارية مشابهة في مدنية ثانية، فوجد أن متوسط مبيعاتها $\overline{X} = 20000$ وإن نصف مداها الربيعي ٥٠٠٠ وحدة نقدية والمطلوب:

١ ــ أي المحلات مبيعاتها ذات انتشار أكبر، مبيعات محلات المدينة الأولى أم مبيعات محلات المدينة الثانية ؟

٢ ــ احسب معامل الاختلاف لمبيعات محلات المدينة الأولى وفسره .

٣ ــ ما هو مقياس النزعة المركزية الذي تختاره لقياس مبيعات محلات المدينة الأولى ولماذا؟.

مسألة (٥) :

تدعي كثير من المنظمات بأن لكثرة الإنجاب أثر سلبي على صحة الطفل ووزنه عند الولادة، ومن أجل دراسة هذا الغرض سحبت عينة عشوائية من أوزان الأطفال حديثي الولادة فكان الجدول التكراري التالي :

نقاط وسط الفئات	2.1	2.3	2.5	2.7	2.9	3.1	3.3	Σ
التكرارات	6	18	22	24	18	8	4	100

والمطلوب :

١ ــ أثبت أن مجموع الانحرافات عن وسطها الحسابي يساوي الصفر في ضوء بيانات الجدول السابق .

٢ ــ هل تعتقد أن التوزيع السابق هو توزيع متماثل أم توزيع ملتو وإذا كان ملتوياً حدد اتجاه الالتواء مدعماً إجابتك بالحسابات اللازمة .

مسألة (٦) :

لدينا الجدول التكراري التالي والمتضمن الأجور اليومية لعينة عشوائية n=100 من عمال صناعة الزجاج والخزف في إحدى البلدان العربية، (الأجور بالوحدات النقدية).

فئات الأجور	120-80	160-120	200-160	240-200	280-240	Σ
عدد العمال	10	25	50	10	5	100

والمطلوب :

١ ــ ارسم المدرج التكراري لبيانات الجدول السابق ثم حدد المنوال بيانياً على نفس الشكل.

٢ ــ احسب الانحراف الربيعي النسبي لأجور العمال في هذه الصناعة .

٣ ــ حدد طبيعة التوزيع الذي تخضع له تلك البيانات (توزيع متماثل أو ملتو) مدعماً إجابتك بالحسابات اللازمة .

٤ ــ أوجد قيمة معامل التفلطح وماذا تستنتج ؟

مسألة (٧) :

فيما يلي بيانات عن أعمار مجموعة من موظفي إحدى الشركات

فئات الأعمار	25-15	35-25	45-35	55-45	65-55	Σ
عدد الموظفين	12	24	100	40	24	200

والمطلوب :

١ ــ احسب منوال الأعمار لبيانات الجدول السابق وفسره .

٢ ــ حدد طبيعة التوزيع، مدعماً إجابتك بالحسابات اللازمة .

٣ ــ حدد الوسيط بيانياً .

٤ ــ احسب المدى الربيعي لبيانات الجدول السابق .

مسألة (٨)

في بحث لدراسة أثر التدخين على الصحة كان التوزع التكراري كما يلي :

فئات السجائر	17-12	22-17	27-22	32-27	35-32	40-35
عدد المدخنين	12	21	25	19	13	10

والمطلوب :

١ ــ أوجد قيمة الوسط الحسابي .

٢ ــ أوجد قيمة الربيع الثالث فقط وفسره .

٣ ــ ماهو أفضل مقياس للنزعة المركزية لبيانات الجدول السابق .

مسألة (٩) :

فيما يلي توزيعاً تكرارياً لعينة مكونة من (٣٦٠) أسرة من الأسر القاطنة في المدن.

عدد أفراد الأسرة	3	5	7	9	11	Σ
عدد الأسر	20	70	180	60	30	360

كما سحبت عينة أخرى من الأسر القاطنة في الأرياف فحصلنا على التوزيع التكراري التالي:

عدد أفراد الأسرة	1	3	5	7	9	11	13	Σ
عدد الأسر	7	15	40	105	170	220	143	700

والمطلوب :

١ ــ أوجد الوسط الحسابي لكلا التوزعين .

٢ ــ أوجد المنوال لكل من التوزعين السابقين وفسرهما.

٣ ــ أي البيانات أكثر تجانساً حول متوسطها برأيك، أجري كافة الحسابات اللازمة.

مسألة (١٠) :

أجريت دراسة إحصائية على مجموعة من الطيارين في قاعدة جوية فأسفرت الدراسة عن الآتي:

فئات ساعات التدريب	14-10	18-14	22-18	26-22	30-26	Σ
عدد الطيارين	12	28	84	20	16	100

والمطلوب:

١ ــ أثبت صحة العلاقة التالية :

$$\overline{X}\sum f = \sum fx$$

في ضوء بيانات الجدول السابق .

٢ ــ احسب مقياساً للنزعة المركزية وآخر للتشتت لبيانات الجدول السابق .

٣ ــ حدد المنوال بيانياً ثم تحقق من ذلك حسابياً .

٤ ــ أخذت مجموعة من الطيارين من قاعدة أخرى مجاورة فكان متوسط ساعات التدريب للطيارين $\overline{X} = 25$ وانحراف معياري 5 ساعات.

والمطلوب:

ــ أي القاعدتين تعرض تشتتاً أكبر في ساعات التدريب؟.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل السادس

الارتباط والانحدار

Correlation & Regression

٦ ــ ١ : مقدمة

٦ ــ ٢ : تحليل الارتباط

ــ تحديد طبيعة المتغيرات

ــ رسم شكل الانتشار

٦ ــ ٣ : مقاييس الارتباط وأنواعه

٦ ــ ٤ ــ معادلة الانحدار الخطي (خواصها واستعمالاتها)

٦ ــ ٥ : الخطأ المعياري للتقدير

٦ ــ ٦ : تحليل التباينات

٦ ــ ٧ : معامل التحديد

٦ ــ ٨ : اختبار معنوية معامل الارتباط

٦ ــ ٩ : تحليل الارتباط للبيانات النوعية

ــ معامل الاقتران

ــ معامل ارتباط الرتب " سبيرمان"

ــ تمارين غير محلولة


وقل رب زدني علما
جامعة دمشق
Damascus University

٦ ــ ١ : مقدمة في تحليل الارتباط :

من المعروف أن كل ظاهرة من الظواهر ، لا تنشأ من تلقاء نفسها وبدون غيرها ، وإنما تقع ضمن نطاق تأثيرات الظواهر الأخرى ، هذا يعني أن هناك علاقة كائنة بين الظواهر، الأمر الذي يتوجب علينا معرفة هذه العلاقة وطبيعتها ، وفق أسلوب إحصائي يدعى بالارتباط .

إن دراسة الارتباط تقوم على بحث العلاقة بين ظاهرتين أو أكثر ، لمعرفة ما إذا كان تغير، إحدى الظاهرتين مرتبطاً بتغير الأخرى ، ومعرفة طبيعة ، ونوع العلاقة وشدتها ، بين الظاهرتين المدروستين . فمثلاً هناك علاقة بين الدخل والإنفاق ، وبين الدخل والإدخار، وبين الإنتاج والمبيعات، وبين معدلات النمو ومعدلات الاستثمار ، وبين العمر وضغط الدم .. الخ .

إن هذه العلاقات بين المتغيرات السابقة تسمى بالارتباط ، وإن تحليل العلاقات الارتباطية بين المتغيرات يدعى بتحليل الانحدار ، فالانحدار يبحث في تفسير تغير أحد المتحولات بدلالة مجموعة أخرى من المتحولات ، بينما يهدف الارتباط إلى قياس شدة العلاقة بين متحول ومتحولات أخرى كل منها على حدة.

ومن الضروري الإشارة إلى أن وجود علاقة ارتباطية بين ظاهرتين لا يعني بالضرورة وجود علاقة سببية بينهما ، فقد يكون الارتباط نتيجة عوامل خارجة عن نطاق المتغيرين موضوع الدراسة، هذا وعند دراسة العلاقة الارتباطية بين ظاهرتين موضوع الدراسة، يجب أن يتوافر فيها الشروط الآتية:

- أن تكون العلاقة بين المتغيرات قائمة على المنطق .
- أن تكون إحدى الظواهر مسببة والأخرى ناتجة .

- أن تكون الظاهرتان قابلتين للقياس بوحدة قياس معينة لكل منهما أو بوحدة قياس مشتركة لكليهما.

- أن تكون القياسات المأخوذة للظاهرتين متقابلة من حيث الزمان أو المكان أو كليهما.

٦ ـ ٢ ـ تحليل الارتباط :

ـ تحديد طبيعة المتغيرات :

ونقصد بتحديد طبيعة المتغيرات تحديد كل من المتغير المستقل والمتغير التابع : إن المتغير المستقل بالتعريف هو المتغير الذي تكون تغيراته غير مرتبطة بتغيرات الظاهرة الثانية ونرمز له بـ (X) بينما المتغير التابع هو المتغير الذي تكون تغيراته مرتبطة أو مرهونة بتغيرات المتغير المستقل ونرمز له بـ (y)،على سبيل المثال العلاقة بين الدخل والإنفاق ،نجد أن الدخل هو المتغير المستقل (x) والإنفاق هو المتغير التابع (y) وكذلك الأمر العلاقة بين العمر وضغط الدم ، حيث العمر هو المتغير المستقل (x) وضغط الدم هو المتغير التابع (y) وأيضاً المبيعات (x) والإنتاج (y)

فمن أجل تحديد أي المتغيرين هو التابع وأيهما المستقل هناك عدة طرق أو قواعد متبعة في التحديد:

أ) طريقة المحاكمة أو المناقشة العقلانية :

فمثلاً عند دراسة العلاقة بين الدخل والإنفاق ، فمن المنطق أن الدخل هو المتغير المستقل والإنفاق هو المتغير التابع ، على اعتبار أن أي تغير في الدخل يرافقه تغيراً في الإنفاق والعكس غير صحيح ، وأيضاً العلاقة بين الإنتاج والمبيعات ، فالذي يحدد الإنتاج هو المبيعات والعكس غير صحيح .

ب) طريقة الأسبقية الزمنية :

تعتمد هذه الطريقة أن المتغير الأسبق في الحدوث زمنياً هو المتغير المستقل، والثاني هو المتغير التابع .

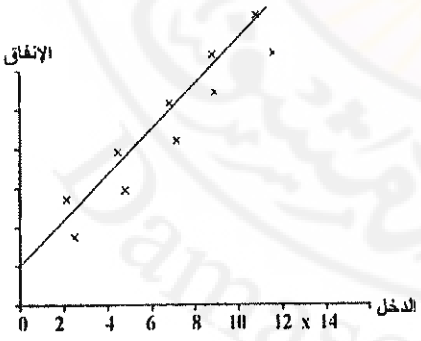
جـ) إذا صادف حدوث المتغيرين معاً فللباحث الخيار في تحديد أي المتغيرات هو المستقل وأيهما هو التابع .

وخلاصة القول : إن الطريقة الأكثر استخداماً في تحديد المتغير المستقل والتابع هي طريقة المحاكمة العقلانية ثم الأسبقية الزمنية في الحدوث .

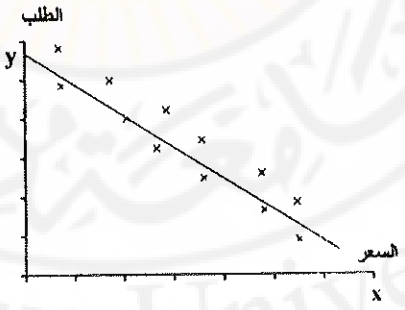
رسم شكل الانتشار :

لتحليل سلوك متغيرين أو ظاهرتين ، نقوم بتمثيل قيم الأزواج المتقابلة للظاهرتين على مستوى ديكارتي واحد ، بحيث تمثل قيم (x) على المحور الأفقي وقيم (y) على المحور العمودي .

إن الهدف من رسم شكل الانتشار هو تحديد نوع العلاقة وطبيعتها بين المتغيرين ، هل العلاقة خطية ، أم من الدرجة الثانية أم أسية وهل الارتباط طردي (موجب) أم عكسي (سالب) أم انعدام ارتباط. لذلك نميز هنا بين أشكال الانتشار التالية للعلاقة بين متغيرين :



شكل (6 - 1)



شكل (6 - 2)

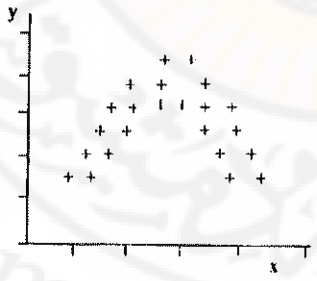
نلاحظ من الشكل (٦-١) أن نقاط الانتشار متوضعة حول خط مستقيم، تتزايد فيه قيم (y) مع تزايد قيم (x) ، وهذا يعني أن العلاقة خطية طرديـــة بين المتغيرين (x, y) وإن شكل معادلة التمثيل من الدرجة الأولـــى $y = f(x)$.

ومن الشكل رقم (٦-٢) نجد أن النقاط أيضاً منتشرة حول خط مستقيم، تتناقص فيه قيم (y) مع تزايد قيم (x) ، أي أن العلاقة خطية عكسية .. مثال السعر وحجم الطلب على السلع .

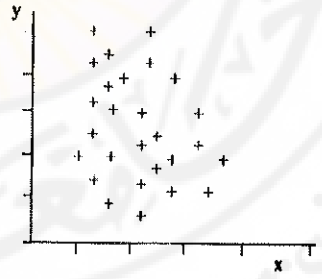
ومن الشكل رقم (٦-٣) نلاحظ أن نقاط الانتشار مبعثـــرة مـــن دون ترابط حول اتجاه معين مما يدل على أنه لا توجد علاقـــة بـــين المتغيـــرين (انعدام الارتباط) .

ومن الشكل رقم (٦-٤) : نجد أن نقاط الانتشار تقع داخل حزمة لهـــا شكل منحني ويستنتج من ذلك أن الارتباط غير خطي، وأن معادلة التمثيـــل تظهر من الشكل التالي :

$$y = a_0 + a_1 x + a_2 x^2$$



شكل (6 - 3)



شكل (6 - 4)

يتضح مما تقدم، إن إجراء تحليل الارتباط بين متغيرين أو ظاهرتين يعتمد بالدرجة الأولى على تصور العلاقة القائمة بينهما.

ونشير إلى أن هناك عدة أنواع للارتباط، منها الارتباط البسيط المستقيم ، الارتباط البسيط غير المستقيم ، والارتباط الخطي المتعدد ، والارتباط الذاتي ، وسوف تقتصر دراستنا على الارتباط البسيط المستقيم فقط.

٦ ــ ٣ : معادلة الانحدار الخطي Linear Regression

إذا كان لدينا متغيران (y,x) أحدهما متغير مستقل (x) والثاني متغير تابع (y) لظاهرتين قيد الدراسة ، مثل العلاقة بين الدخل والإنفاق والعمر وضغط الدم ... الخ ، فإنه يوجد أسلوبان لتحديد ثوابت معادلة التقدير والتي تأخذ الصيغة التالية:

$$\tilde{y} = a + bx$$

حيث (a) : تشير إلى ثابت خط الانحدار ، أي قيمة ($\tilde{y}$) عندما ($x = 0$) . (ثابت خط الانحدار)

b : معامل الانحدار ، (ميل خط الانحدار) ، ويدل على مقدار التغير في y (القيم المقدرة) لكل وحدة تغير واحدة في المتغير المستقل (x) .

١- أسلوب الرسم الحر :

لتوضيح هذا الأسلوب ، نسوق المثال التالي :

أخذت عينة عشوائية مكونة من (٧) أسر في مدينة ما ، بغية دراسة العلاقة بين الدخل والإنفاق السنوي ، لهذه الأسر ، على الثقافة ، فأعطت البيانات التالية : وقد قيس كل من الدخل والإنفاق بآلاف الوحدات النقدية .

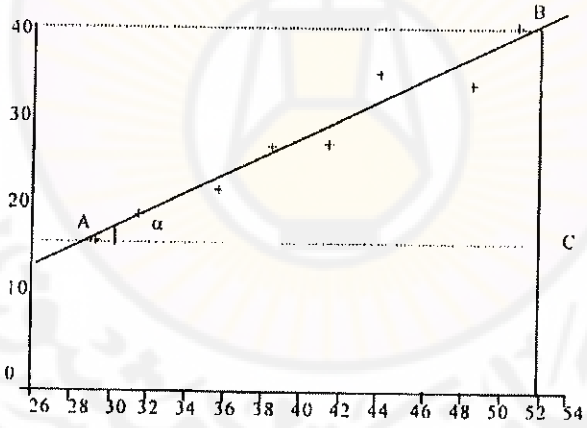
الدخل (x)	38	32	42	48	40	44	50
الإنفاق (y)	24	21	27	30	27	33	36

والمطلوب : إيجاد معادلة خط الانحدار ، بأسلوب الرسم الحر ، ثم بين بالاعتماد على شكل الانتشار الناتج نوع العلاقة وطبيعتها بين المتغيرين .

الحل :

من أجل تحديد نوع العلاقة وطبيعتها بين المتغيرين ، نرسم شكل الانتشار ، ونلاحظ من الشكل أن العلاقة بين المتغيرين خطية ، لأنه يمكن وصف العلاقة المتوسطة بين المتغيرين بمعادلة خط مستقيم من الشكل :

$$\tilde{y} = a + bx$$



شكل (6 – 5)

وواضح أيضاً من الشكل (٦–٥) ، أن نقاط الانتشار متوضعة حول خط مستقيم تتزايد فيه قيم (y) بتزايد قيم (x) ونستنتج من ذلك أن العلاقة بين المتغيرين طردية موجبة .

ولتحديد الثوابت وإيجاد المعادلة (معادلة التقدير) بأسلوب الرسم الحر نقوم بتوفيق أو تمرير مستقيم ، يمر من بين نقاط الانتشار ، على اعتبار أنه أفضل مستقيم يمثل العلاقة المتوسطة بين المتغيرين، ثم نحدد على الخط الممهد الذي رسمناه النقطتين (A, B) ثم نسقط من النقطة (B) خطاً عمودياً على المحور (x) يوازي محور (y) ، ومن النقطة (A) نرسم خطاً أفقياً يوازي المحور (x) ويتقاطع مع المحور العمودي الذي قمنا برسمه في النقطة (c) فيكون ميل المستقيم (b) هو عبارة عن "tan (α)

$$\tan(\alpha) = \frac{Bc}{Ac}$$

أما الثابت (a) فهو عبارة عن نقطة تقاطع خط الانحدار (الخط الممهد) مع المحور (y)، وبالتالي نحصل على ثوابت المعادلة .

ومن عيوب هذا الأسلوب ، أنه يمكن تمرير عدد غير متناه من المستقيمات من بين نقاط الانتشار ، وسوف نحصل لكل مستقيم على ثوابت (a , b) مختلفة من مستقيم لآخر، لهذا السبب فإن هذه الطريقة قليلة الاستخدام في التوفيق وتعتمد بشكل كبير على خبرة الباحث ومهارته.

٢- أسلوب المربعات الصغرى : (المعادلات الطبيعية)

إن أسلوب المربعات الصغرى ، يبني نظريته ، على أساس أن أفضل مستقيم يمثل العلاقة المتوسطة بين (x , y) هو ذلك المستقيم الذي يحقق العلاقة :

$$\sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 = \min$$

نبدل y بقيمتها في المعادلة الأساسية :

$$y = a + b\,x$$

فتصبح العلاقة $F = \sum (y_i - a - b x_i)^2$

ولحساب (a , b) نحسب المشتقات الجزئية للتابع F بالنسبة لكل من (a , b) ونعدمها، فنجد أن:

$$\frac{\partial F}{\partial a} = \frac{\partial}{\partial a} \sum (y_i - a - b x_i)^2$$

$$= \frac{\partial}{\partial a} \sum (y_i - a - b x_i)^2$$

$$= 2 \sum (y_i - a - b x_i) \cdot (-1)$$

وبالإصلاح بعد التقسيم على (٢ـ) نحصل على المعادلة الطبيعية الأولى .

$$\sum y = na + b \sum x \qquad (1)$$

ولإيجاد المعادلة الثانية نشتق بالنسبة لـ b فنجد أن :

$$\frac{\partial F}{\partial a} = 2 \sum (y_i - a - b x_i)(-x_i)$$

وبالإصلاح بعد التقسيم على (٢ـ) نحصل على المعادلة الثانية :

$$\sum x y = a \sum x + b \sum x^2 \qquad (2)$$

نسمي هاتين المعادلتين بالمعادلتين الطبيعيتين لحساب الثوابت (a , b) وبالعودة إلى التمرين السابق ، المتعلق بالدخل والإنفاق السنوي على الثقافة لـ (7) أسر ، أوجد معادلة خط الانحدار بأسلوب المربعات الصغرى؟ ومن أجل ذلك نقوم بإنشاء جدول مساعد لحساب المجاميع:

$$\sum x , \sum y , \sum x^2 , \sum x y$$

i	x	y	x . y	x^2	Y^2	$\tilde{y}$	$(y-\tilde{y})^2$	$(\tilde{y}-\bar{y})^2$
1	38	24	912	1444	576	25.21	1.4641	9.42
2	32	21	672	1024	441	20.59	0.1681	59.13
3	42	27	1134	1764	729	28.29	1.6641	0.001
4	48	30	1440	2304	900	32.91	8.4681	21.43
5	40	27	1080	1600	729	26.75	0.0625	2.34
6	44	33	1452	1936	1089	29.83	10.0489	2.40
7	50	36	1800	2500	1296	34.45	2.4025	38.06
Σ	294	198	8490	12572	5760	203.03	24.2783	132.781

وبالتعويض في المعادلات :

$$\sum y = na + b \sum x \qquad (1)$$

$$\sum x \,.\, y = a \sum x + b \sum x^2 \qquad (2)$$

$$198 = 7a + 294\ b \qquad (1)$$

$$8490 = 294\ a + 12572\ b \qquad (2)$$

نضرب المعادلة الأولى بقيمة (٤٢) (طريقة الحذف بالتعويض) نجد أن :

$$8316 = 294\ a + 12348 \qquad (3)$$

وبطرح (٣) من (٢)

$$174 = 0 + 224\ b$$

$$b = \frac{174}{224} = 0.77$$

وبالتعويض في (١) نجد

$$198 - 294\ (0.77) = 7\ a$$

$$a = \frac{28.38}{7} = -4.05$$

نعوض في معادلة خط الانحدار :

$$\tilde{y} = -4.05 + 0.77$$

التفسير :

تمثل (a) القيمة الوسطية للاتفاق السنوي ($\tilde{y}$) عندما تكون قيمة الدخل السنوي معدوماً ($x = 0$) ، وهي ثابت مستقيم الانحدار .

وتمثل (b) قيمة معامل الانحدار ، وإشارتها موجبة ، وهذا يعني أن العلاقة طردية بين المتغيرين ، وتفسيرها بازدياد قيمة الدخل السنوي (x) وحدة واحدة فإن قيمة الاتفاق السنوي ($\tilde{y}$) تزداد بمقدار (b) وسطياً.

أسلوب المربعات الصغرى : (الطريقة المباشرة) .

يمكن حساب الثوابت بالطريقة المباشرة وفق الصيغ التالية :

$$b = \frac{\sum x \cdot y - n \bar{x} \bar{y}}{\sum x^2 - n(\bar{x})^2}$$

$$a = \bar{y} - b \bar{x}$$

وبالتطبيق ينتج لدينا :

$$b = \frac{8490 - 7(42)(28.28)}{12572 - 7(42)^2} = 0.77$$

$$a = 28.28 - 0.77\ (42) = -4.05$$

$$\tilde{y} = -4.05 + 0.77$$

- **خصائص خط الانحدار :**

١- إن خط الانحدار يمر بنقطة إحداثيها ($\bar{x}$, $\bar{y}$) .

٢- مجموع الفروق بين القيم الفعلية والقيم المقدرة للمتغير التابع أصغر ما يمكن .

- **استعمالات خط الانحدار :**

١- في التقدير : أي تقدير القيمة الوسطية للمتغير التابع من أجل قيمة معينة من قيم (x) في العينة وذلك بالتعويض في معادلة خط الانحدار .

$$\tilde{y} = -4.05 + 0.77$$

من أجل قيم (x) الواردة في التمرين السابق نجد أن :

$$\tilde{y} = -4.05 + 0.77$$

$$\tilde{y} = -4.05 + 0.77\ (38) = 25.21$$

نلاحظ أن القيمة الوسطية (النظرية) من أجل $x = 38$ هي القيمة $= \tilde{y}\ 25.21$.

وهكذا بالنسبة لبقية القيم الواردة في الجدول السابق (يترك للطالب التحقق من ذلك) .

٢- في التنبؤ : أي نتوقع قيمة المتغير التابع (y) بدلالة قيم (x) الواقعة ضمن مجال تغيرات (x) وذلك لضمان بقاء العلاقة خطية بين المتغيرات وهذا ما يدعى بالتمديد الداخلي .

تمرين :

قدر إنفاق أسرة يبلغ دخلها (٣٣) وحدة نقدية .

نعوض في معادلة خط الانحدار :

$$\tilde{y} = -4.05 + 0.77\ (x)$$

$$-4.05 + 0.77\ (33) = 21.36$$

وهي القيمة المتوقعة للإنفاق ، لأسرة دخلها (33) وحدة نقدية ، أما بالنسبة لتقدير قيم ($\tilde{y}$) من أجل قيم (x) الواقعة خارج المجال أو حدود بيانات العينة ، في هذه الحالة لا نستطيع استخدام معادلة خط الانحدار في إجراء التنبؤ أو التقدير ، لأننا لا نستطيع أن نقرر بأن العلاقة ستبقى خطية بين المتغيرين موضوع الدراسة فيما وراء بيانات العينة .

رسم خط الانحدار :

من أجل الرسم البياني لخط الانحدار نوضح ذلك من خلال التمرين السابق :

$$= -4.05 + 0.77\ \tilde{y}$$

$$(x)$$

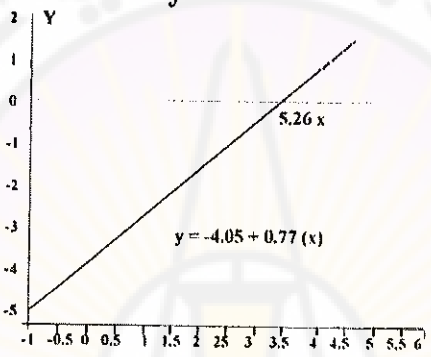
$$= 0\ \tilde{y}$$

أي أن : $-4.05 + 0.77\ x = 0$

$$4.05 = 0.77\ x \Rightarrow x = \frac{4.05}{0.77} = 5.25$$

أي أن :

$$= -4.05\ \tilde{y}$$



شكل (6 – 6)

٦ ــ ٤ : الخطأ المعياري للتقدير لمعادلة خط الانحدار :

هو عبارة عن مقياس للتشتت ، ويقيس انتشار أو تشتت (y) الفعلية حول معادلة خط الانحدار ($\tilde{y}$) ، بمعنى أن هذا المقياس يبين مقدار انحراف (y) عن ($\tilde{y}$) المحسوبة من معادلة خط الانحدار ، والخطأ المعياري للتقدير يعرف رياضياً ، بأنه الجذر التربيعي لمتوسط مجموع مربع انحرافات القيم المقدرة عن القيم الفعلية للمتغير التابع ، ويعطى بالعلاقة التالية :

$$S_{y\tilde{y}} = \sqrt{\frac{\Sigma\,(y - \tilde{y})^2}{n}}$$

وبتعويض y بقيمتها بالعلاقة والإصلاح نحصل على العلاقة التالية :

$$S_{y\tilde{y}} = \sqrt{\frac{\sum y^2 - (a\sum y + b\sum yx)}{n}}$$

إن استخدامات هذا المقياس تشابه تماماً استخدامات الانحراف المعياري بالنسبة لمقاييس النزعة المركزية ، فهو يستخدم كمعيار لمقدار الثقة التي نضعها في القيم المحسوبة $(\tilde{y}_i)$ من معادلة الانحدار بدلالة المتغير (x).

وكلما كبرت قيمة هذا المقياس كلما قلت الثقة بالقيم المقدرة وأصبحت المصداقية أقل ، وبالعكس كلما قلت قيمة هذا المقياس كلما ازدادت الثقة بالقيم المقدرة $\tilde{y}$ والمحسوبة من معادلة الانحدار، أي ازداد التقارب بين القيم الفعلية والقيم النظرية ، وإذا كانت قيمة هذا المقياس $S_{y\tilde{y}}=0$ في هذه الحالة تكون نقاط الانتشار واقعة على خط مستقيم يمر من مركز الإحداثيات دون انتشار أو تشتت وهذا ما ندعوه بالارتباط التام وفيه يكون :

$$\tilde{y} = bx$$

$$\sum y_i = \sum \tilde{y}_i$$

$$y_i = \tilde{y}_i$$

$$a = 0$$

$$\bar{y} = \bar{x}$$

إن مربع الخطأ المعياري للتقدير $s_{y\tilde{y}}$ يساوي التباين غير المفسر .

وبالعودة إلى التمرين السابق، والمتعلق بدخول ونفقات الأسر على الثقافة .

والمطلوب : حساب الخطأ المعياري للتقدير لمعادلة خط الانحدار . بتطبيق الصيغة الأولى :

$$S_{y\tilde{y}} = \sqrt{\frac{\sum (y-\tilde{y})^2}{n}} = \sqrt{\frac{24.2783}{7}} = 1.862$$

- وبتطبيق الصيغة الثانية :

$$S_{y\tilde{y}} = \sqrt{\frac{\sum y^2 - (a\sum y + b\sum yx)}{n}}$$

$$= \sqrt{\frac{5760 - [-4.05(198) + 0.77(8490)]}{7}} = 1.862$$

وبتربيع الناتج نحصل على قيمة التباين غير المفسر أي :

$$= (1.862)^2 = 3.468 \quad \tilde{y} S^2 y$$

يستخدم الخطأ المعياري للتقدير في التقدير المجالي لقيم $(\tilde{y})$.

من خلال العلاقة نحصل على التقدير المجالي لقيمة $(\tilde{y})$.

$$p_r[\tilde{y} - Z_{\alpha/2} \cdot S_{\tilde{y}} y \le y \le \tilde{y} + Z_{\alpha/2} \cdot S_{\tilde{y}} y] = (1-\alpha)$$

من أجل احتمال $(1-\alpha)$ يمكن إيجاد القيمة المجالية لـ $(\tilde{y})$ باستخدام (y) وبالعودة إلى التمرين السابق .

والمطلوب :

قدر قيمة الإنفاق لأسرة دخلها (33) وحدة نقدية عند مستوى دلالة $\alpha = 0.05$

$$\tilde{y} = -4.05 + 0.77(33) = 21.36$$

$$\tilde{y} - Z_{\alpha/2} \cdot S_{\tilde{y}} y \le y \le \tilde{y} + Z_{\alpha/2} \cdot S_{\tilde{y}} y$$

$$21.36 - 1.96 \cdot (1.862) \le y \le 21.36 + 1.96(1.862)$$

$$17.17 \le y \le 25.009$$

والتفسير أنه باحتمال 95 % ، (α = 5 %) إن إنفاق الأسرة التي يبلغ دخلها (33) وحدة نقدية . لن يقل عن 17.17 وحدة نقدية ولن يزيد عن 25.009 وحدة نقدية ومن الضروري الإشارة إلى أن حجم العينة يجب أن يكون $n > 30$ مفردة حتى يكون التقدير صحيحاً ، وفيما لو قل حجم العينة عن ذلك $n \leq 30$ في هذه الحالة نستخدم توزيع (t) ستودنت من أجل (v=n-1) . وتصبح علاقة الخطأ المعياري للتقدير :

$$S_{y\tilde{y}} = \sqrt{\frac{\Sigma (y - \tilde{y})^2}{n-1}}$$

٦ ــ ٥: تحليل التباينات :

لقد رأينا من خلال دراستنا للتمرين السابق ، (المتعلق بالعلاقة بين الدخل والإنفاق على الثقافة) ، أن استخدام معادلة خط الانحدار قد ساعد على التقليل من أخطاء التقدير ، فوجدنا أن الخطأ المعياري للتقدير أقل بكثير من الانحراف المعياري للوسط الحسابي للمتغير التابع (y) ، ويعرف التباين الكلي (S_y^2) بأنه متوسط مجموع مربع انحرافات التغير التابع (y) عن وسطه الحسابي ويعطي بالعلاقة التالية ، وذلك لعينات $n > 30$:

$$S_y^2 = \frac{\Sigma (y - \tilde{y})^2}{n}$$

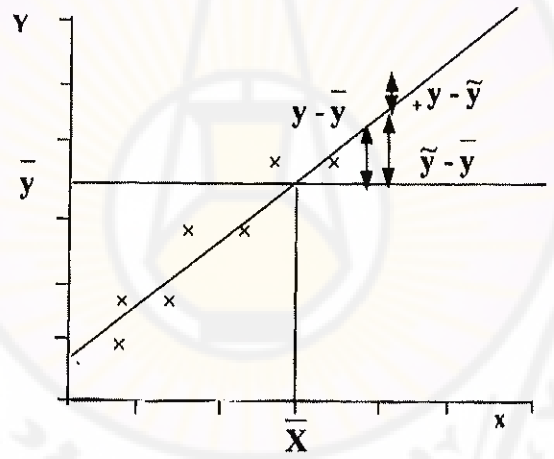
أو

$$S_y^2 = \frac{\Sigma y^2}{n} - \left(\frac{\Sigma y}{n} \right)^2$$

ومن الممكن إظهار التباين الكلي على شكل الانتشار كما هو موضح بالرسم شكل (٦ ــ ٧) ويقسم التباين الكلي إلى قسمين :

١ – **التباين غير المفسر** : هو مربع الخطأ المعياري للتقدير الذي تحدثنا عنه سابقاً وبعبارة أخرى هو عبارة عن التباين الذي لا يمكن لمعادلة الانحدار أن تفسره بدلالة المتغير المستقل (x) .

٢ – **التباين المفسر** : هو عبارة عن التباين الذي تفسره معادلة خط الانحدار ، وكلما زادت قيمته دل ذلك على إمكانية تحديد قيم المتغير (y) بدلالة المتغير (x) بدقة أكبر وتزداد قيمة هذا المعامل كلما ازداد الفرق بين $\tilde{y}$ و y ، أي كلما أزداد ميل معادلة خط الانحدار (b) إلى النقطة التي يمر فيها الخط من مركز الإحداثيات، وبالتالي يكون الارتباط تاماً ، وفي هذه الحالة يكون التباين الكلي مساوياً للتباين المفسر .



شكل (6 - 7)

ويعطي التباين المفسر بالعلاقة التالية :

$$S_{\tilde{y}}^2 = \frac{\sum \left(\tilde{y} - \bar{y} \right)^2}{n}$$

أو :

$$S_{\tilde{y}}^2 = \frac{b\left(\sum x\, y - \bar{x} \sum y\right)}{n}$$

وبالعودة إلى التمرين السابق ، **والمطلوب** : حساب التبـــاين المفـــسر والتباين غير المفسر.

التباين المفسر :

$$S^2 \tilde{y} = \frac{\Sigma \left(\tilde{y} - \bar{y} \right)^2}{n}$$

$$\frac{132.78}{7} = 18.96$$

أما التباين غير المفسر الذي هو عبارة عن مربع الخطأ المعياري للتقدير $S\, y\, \tilde{y} = 1.862$ الذي تم حسابه سابقاً.

وهو التباين غير المفسر $S^2_{y\tilde{y}} = 3.467$

ومنه التباين الكلي = التباين المفسر + التباين غير المفسر .

$$S^2 y = s\, \tilde{y} + S^2_{y\tilde{y}}$$

$$18.969 + 3.467 = 22.436 \cong 22.77$$

ومنه نستنتج أن هذه القيمة هي نفسها قيمة التبـــاين الكلـــي (مربـــع الانحراف المعياري).

$$S^2 y = (4.77)^2$$

$$S^2 y = 22.436$$

معامل الارتباط (بيرسون) :

يستخدم هذا المقياس لقياس درجة أو شدة العلاقة بين متغيـــرين ، وإن قيمته تتراوح بين [1 + , 1-] أي أن $r \in$ [1 + , 1-] وإشارته دوماً من إشارة (b) معامل الانحدار ، وكلما اقتربت قيمة هذا المعامل من (1 ±) دل ذلك على قوة أكبر للعلاقة بين المتغيرين وبالعكس .

نحن نعلم :

$$S^2 S^2 y = \tilde{y} + S^2 \tilde{y}\ y$$

وبقسمة الطرفين على $S^2 y$ نجد :

$$\frac{S^2 y}{S^2 y} = \frac{S^2 y \tilde{y}}{S^2 y} + \frac{S^2 \tilde{y}}{S^2 y}$$

$$1 = \frac{S^2 y \tilde{y}}{S^2 y} + \frac{S^2 \tilde{y}}{S^2 y}$$

ومن العلاقة الأخيرة :

$$r^2 = 1 - \frac{S^2 \tilde{y}}{S^2 y}$$

نرمز للتباين المفسر النسبي بـ r^2 وهو ما يعرف بمعامل التحديد ويعطي بالعلاقة :

$$r^2 = 1 - \frac{S^2 y \tilde{y}}{S^2 y}$$

أو

$$r^2 = \frac{S^2 \tilde{y}}{S^2 y}$$

٦ ـ ٦ : معامل التحديد

إن معامل التحديد هو عبارة عن مربع معامل الارتباط ، وعندما يكون الارتباط تاماً أي $r = 1$ في هذه الحالة يكون ($S^2 \tilde{y} = S^2 y$) وتكون قيمة ($r^2 = 1$) ، وعندما يكون الارتباط معدوماً ($r = 0$) وبالتالي فإن ($r^2 = 0$) أي أن $r^2 \in [0 , 1]$ وهو يدل على النسبة المئوية من تباين المتغير التابع ، التي يفسرها تباين المتغير المستقل ، أو هو النسبة المئوية من التباين الكلي ، التي تفسرها معادلة الانحدار .

وبالعودة إلى التمرين السابق ، والمطلوب : حساب معاملي الارتبـاط والتحديد وفسرهما

$$r^2 = 1 - \frac{S^2 y\tilde{y}}{S^2 y} = 1 - \frac{3.467}{22.436} = 0.84$$

$$r^2 = \frac{S^2 \tilde{y}}{S^2 y} = \frac{18.969}{22.436} = 0.84$$

والتفسير أن 84 % من تباين المتغير التابع ، يفسرها تباين المتغير المستقل ، أو أن 84 % من التباين الكلي ، تفسره معادلة الانحدار ، وإن 16 % من التباين الكلي لا تفسره المعادلة أي 84 % من تباين الإنفاق يفسرها تباين الدخل و 16 % تعود لمتغيرات أخرى لم تدخل في معادلة الانحدار .

معامل الارتباط r :

وهو عبارة عن الجذر التربيعي لمعامل التحديد :

$$r = \sqrt{r^2} = \sqrt{0.84} = 0.91$$

وقيمة معامل الارتباط ، تدل على أن العلاقة بين المتغيرين قوية فـي مثالنا .

كما يمكن حساب معامل الارتباط بالطرق المباشرة الآتية :

$$r = \sqrt{\frac{a\sum y + b\sum xy - n\bar{y}^2}{\sum y^2 - n\bar{y}^2}}$$

أو من العلاقة :

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right]\left[n\sum y^2 - (\sum y)^2\right]}}$$

وبالتطبيق على التمرين السابق نجد أن :

$$r = \sqrt{\frac{-4.05(198) + 0.77(8490) - 7(28.28)^2}{5760 - 7(28.28)^2}} = 0.92$$

$$r = \frac{7(8490) - (294)(198)}{\sqrt{[7(12572) - (86436)][7(5760) - (39204)]}} 0.92$$

معامل ارتباط الرتب ، سبيرمان (Superman) :

إن معامل ارتباط بيرسون ، يستخدم لقياس شدة أو قوة العلاقة بين المتغيرات الكمية، وفي حالة البيانات النوعية نستخدم معامل ارتباط سبيرمان، وتعرف البيانات النوعية: بأنها عبارة عن الأشياء أو الصفات غير القابلة للقياس مثال (اللون ، العرق ، الجنسية ، الحالة الزواجية) .

إن معامل ارتباط الرتب سبيرمان، يعطي قياساً للارتباط للبيانات الكمية والنوعية التي لها صفة الترتيب .

ويعطي بالعلاقة التالية :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

حيث تشير :

r_s : معامل ارتباط الرتب .

n : حجم العينة لأزواج من القيم (x , y) .

d : الفرق بين رتب أزواج القيم (x , y) .

لدينا عينة عشوائية من (7) أسر لدراسة الدخل والإنفاق على الثقافة

والمطلوب :

حساب معامل ارتباط الرتب لسبيرمان .

الدخل X	الإنفاق Y	رتب x	رتب y	الفرق d	مربع الفـرق d^2
38	24	2	2	0	0
32	21	1	1	0	0
42	27	4	3.5	0.5	0.25
48	30	6	5	1	1
40	27	3	3.5	0.5	0.25
44	33	5	6	1	1
50	36	7	7	0	0
					$\sum d^2 = 2.5$

لاحظ أن : 3.5= 2/(3+4) [المفردة (٣) ، (٤)] .

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(2.5)}{7(49-1)} = 1 - \frac{13}{336} = 0.96$$

ملاحظة :

عادة ما يعطي بيرسون ، وسبيرمان ، قيماً متقاربة .

٦ ــ ٧ : اختبار معنوية معامل الارتباط:

إذا سحبنا عدداً كبيراً من العينات العشوائية حجمها (n) وحسبنا معاملات الارتباط ، من هذه العينات ، فإن هذه المعاملات تخضع لمنحنى توزيع طبيعي (التوزيع العيني لمعاملات الارتباط)، إذا كان المجتمع الإحصائي (N) يتوزع طبيعياً ، بصرف النظر عن حجم العينات المسحوبة.

أما إذا كان المجتمع ، لا يتوزع طبيعياً وكانت حجوم العينات المسحوبة (n > 30) فإن توزيع معاملات الارتباط تقترب من شكل التوزيع

الطبيعي ، وإذا كان $n \leq 30$ والمجتمع لا يخضع لتوزيع طبيعي فإن شكل التوزيع لمعاملات الارتباط يخضع لتوزيع (t) ستودنت .

اختبار معامل الارتباط :

يتم اختبار معامل الارتباط المحسوب من العينة بالعلاقة :

$$t = r\frac{\sqrt{n-2}}{\sqrt{1-r^2}}$$

وذلك من أجل عدد من درجات الحرية ودرجات الحرية هي عبارة عن عدد المشاهدات المستقلة في العينة مطروحاً منها العدد(2) .

أي : $V = (n-2)$

ومستوى دلالة محدد $\alpha = 5\%, 10\% \ldots$

نقارن النتيجة التي نحصل عليها من الاختبار السابق (t) بالقيمة النظرية المستخرجة من الجداول الخاصة بالتوزيع (t) .

وبالعودة إلى التمرين السابق ، والمطلوب : اختبار معنوية معامل الارتباط ، أو بيان الأهمية الإحصائية لمعامل الارتباط عند مستوى دلالة

$\alpha = 0.05$

الحل :

نحن نعلم أن : $r = 0.91$

$n = 7$

$V = 7 - 2 = 5$

وبالتطبيق نجد أن :

$$t = r \cdot \frac{\sqrt{n-2}}{\sqrt{1-r^2}}$$

$$= 0.91 \frac{\sqrt{7-2}}{\sqrt{1-(0.91)^2}} = 4.914$$

من خلال المقارنة بين قيمة $t = 4.914$ (المسحوبة) و $t = 2.57$ (الجدولية) عند ومستوى دلالة $\alpha = 0.05$ نجد أن (النظرية) $t > t$ (المحسوبة) وبالتالي نرفض فرضية الاستقلال ، والتي تنص بأنه لا توجد علاقة بين الظاهرتين المدروستين في المجتمع الذي سحبت منه العينة ، ونقرر أن هناك ارتباطاً حقيقياً في العينة ، أو بمعنى آخر الارتباط في العينة لا يساوي الصفر ، والعينة المسحوبة تمثل المجتمع الإحصائي الذي سحبت منه ولبيان الأهمية الإحصائية لمعامل الارتباط نحسب ما يسمى بالخطأ المعياري لمعامل الارتباط الأهمية إحصائية .

الخطأ المعياري لمعامل الارتباط (S.E.C.r) الذي يُعطى بالعلاقة التالية:

وبالعودة إلى التمرين السابق والمتعلق بدخل ونفقات الأسر على الثقافة، تكون قيمة الخطأ المعياري لمعامل الارتباط :

$$S.E.C.r = \sqrt{\frac{1 - 0.84}{7 - 2}} = 0.17$$

ويستخدم الخطأ المعياري لمعامل الارتباط لبيان فيما إذا كان معامل الارتباط ، في المجتمع الإحصائي المسحوب منه العينة مساوياً الصفر أم لا ، وبالتالي الحكم على معنوية معامل الارتباط في العينة ، فإذا كان معامل الارتباط في العينة أكبر بثلاث مرات من خطئه المعياري أو يساويه أي (S.E.C.r) $r \geq 3$ فالعينة المسحوبة عشوائية وتمثل

المجتمع الإحصائي الذي سحبت منه . أفضل تمثيل ، ولمعامل الارتباط في العينة أهمية إحصائية.

وإذا كان (S.E.C.r) 3 > r فالعينة المسحوبة لا تمثل مجتمعها الإحصائي ، الذي سحبت منه ، وليس لمعامل الارتباط أية أهمية إحصائية ، والارتباط الذي تبينه العينة يعود لقوى الحظ والصدف .

وبالنسبة للتمرين المشار إليه سابقاً فإن قيمة معامل الارتباط أكبر من قيمة ثلاثة أمثال خطئه المعياري وبالتالي فإن لمعامل الارتباط في العينة أهمية إحصائية ويعبر عن الارتباط بين هاتين الظاهرتين في المجتمع الإحصائي، وهذا يعني بأن معامل الارتباط للمجتمع الإحصائي الذي أخذت منه العينة العشوائية لا يمكن أن يساوي الصفر، وإنما أية قيمة أخرى موجبة كانت أم سالبة.

تمارين غير محلولة

تمرين (1) :

قامت إحدى الشركات برصد إنفاقها على الإعلان وحجم المبيعات ، وتم التوصل إلى النتائج التالية :

الإنفاق على الإعلان	90	80	80	85	87	92	85	90	93	95
حجم المبيعات	5	2	4	6	4	7	5	6	5	7

وقدرت المبيعات والإنفاق على الإعلان بالألف وحدة نقدية

والمطلوب :

١ - ارسم شكل الانتشار ، لبيانات الجدول السابق ثم استنتج نوع وطبيعتها العلاقة بين المتغيرين .

٢ - أوجد معادلة الانحدار وفسر الثوابت .

٣ - احسب الخطأ المعياري للتقدير وبيّن أهميته .

٤ - احسب معاملي الارتباط والتحديد وفسرهما .

تمرين (2) :

يبين الجدول التالي سنوات الخبرة ، وحجم المبيعات ، بالألف وحدة نقدية ، لعدد من مندوبي المبيعات في إحدى شركات بيع مستلزمات المكتبات.

رقم البائع	1	2	3	4	5	6	7	8	9	10
سنوات الخبرة	1	3	4	4	6	8	10	10	11	13
حجم المبيعات	80	97	92	102	103	111	119	123	117	136

والمطلوب :

١ - احسب معامل الانحدار وثابت مستقيم الانحدار وفسرهما .

٢ - ارسم معادلة الانحدار الناتجة وبيّن على الشكل التباين غير المفسر .

٣ - قدر باحتمال % 95.5 . حجم المبيعات ، لبائع عدد سنوات خبراته (12) سنه .

4 - احسب التباين الكلي .

5 - احسب الخطأ المعياري للتقدير لمعامل الارتباط وبين استخداماته .

تمرين (3) :

أخذت عينة عشوائية من (6) قطع من الأراضي الزراعية المتماثلة في الصفات ، بغية دراسة أثر أحد أنواع الأسمدة ، في متوسط الغلة في وحدة مساحة من القمح فحصلنا على النتائج الآتية :

رقم القطعة	1	2	3	4	5	6
كمية السماد في وحدة المساحة	5	7	9	6	4	5
متوسط الغلة في وحدة المساحة	160	210	231	170	164	175

حيث كانت وحدة الأوزان المستعملة هي الكغ .

والمطلوب :

١ - ارسم شكل الانتشار وبين الهدف منه .

٢ - أوجد معادلة مستقيم الانحدار وفسر ثوابتها .

٣ - ارسم مستقيم الانحدار .

٤ - احسب معامل الارتباط (سبيرمان) و (بيرسون) وقارن بينهما.

تمرين (4) :

تم التوصل من خلال دراسة أجريت ، عن العلاقة بين دخول ونفقات (4 أسرة على الثقافة في مدينة دمشق إلى المعلومات التالية :

$$\sum X Y = 19960 \ , \ \sum X = 560 \ , \ \sum Y = 483$$

$$\sigma_X = 3 \ , \ \sigma_Y = 1.7$$

حيث تمثل (x) الدخل الذي يتراوح بين [59 , 28] وحدة نقدية ، وتمثل (y) الإنفاق الذي يتراوح بين [47 , 27] وحدة نقدية . وقد قيس كل من الدخل والإنفاق بآلاف الوحدات النقدية .

والمطلوب :

١ - ما هي الشروط الواجب توافرها في دراسة العلاقة الارتباطية بين متغيرين .

٢ - أوجد معادلة الانحدار ثم بيّن طبيعة العلاقة بين المتغيرين .

٣ - ارسم معادلة الانحدار ثم اذكر خواص هذه المعادلة .

٤ - قدر النفقات لأسرة دخلها (60) وحدة نقدية ؟ برر إجابتك بما تراه مناسباً .

٥ - ميّز بين المفاهيم التالية ، الارتباط ، الارتباط التام ، الارتباط المعدوم.

٦ - بيّن الأهمية الإحصائية لمعامل الارتباط في العينة ثم اختبره وفق توزيع (t) .

تمرين (5) :

بفرض وجود علاقة خطية بين حجم الطلب وسعر السلعة ، ثم التوصل من خلال دراسة أجريت عن العلاقة بين سعر السلعة ، وحجم الطلب على هذه السلعة إلى المعلومات التالية :

$$\sum X Y = 1596 \quad , \quad n = 14$$

$$\bar{X} = 18 \quad \sum X^2 = 6888$$

$$\bar{y} = 11 \quad \sum Y^2 = 2290$$

علماً أن حجم الطلب على هذه السلعة عندما توزع مجاناً كـــان (20) وحدة .

والمطلوب :

١ – أوجد معادلة الانحدار وبين الغاية أو الهدف من حسابها .

٢ – احسب معاملي التحديد والارتباط وفسرهما .

٣ – ارسم مستقيم الانحدار الناتج .

٤ – ما الفرق بين الارتباط والانحدار .

تمرين (6) :

تم التوصل من خلال دراسة عن العلاقة بين الدخل والإنفاق الشهري على الثقافة لعينة عشوائية من (8) أسر في مدينة حلب إلى البيانات التالية :

$$\sum X Y = 1895 \quad \bar{X} = 25 \quad \sum X^2 = 5864 \ ,$$

$$, \quad \sum Y^2 = 651 \quad \bar{y} = 8.625$$

حيث تمثل (x) الدخل الشهري الذي يتراوح بين [31 , 18] وحدة نقدية وتمثل (y) الإنفاق الذي يتراوح بين [14 , 3] وحدة نقدية ، وقد قيس كل منهما بمئات الوحدات النقدية .

والمطلوب :

١ – احسب التباين غير المفسر واشرح مدلوله .

٢ – ارسم معادلة الانحدار وحدد عليها التباين المفسر .

٣ – قدر باحتمال % 95.5 إنفاق أسرة يبلغ دخلها (28) وحدة نقدية ، ثم قدر النفقات الشهرية على الثقافة لأسرة دخلها الشهري (35) وحدة نقدية .

٤ – احسب معامل التحديد وفسره ؟ ثم بين الأهمية الإحصائية لمعامل الارتباط .

تمرين (7) :

أخذت عينة عشوائية من درجات (10) طلاب من كلية التجارة لمقرري الإحصاء والرياضة فحصلنا على الجدول التالي :

رقم الطالب	1	2	3	4	5	6	7	8	9	10
درجته في الإحصاء	70	50	85	40	90	90	44	70	60	80
درجته في الرياضة	67	50	79	41	85	90	40	60	59	73

والمطلوب :

١ – حدد المتغير المستقل والمتغير التابع ، مبرراً إجابتك بما تراه مناسباً .

٢ – أوجد معامل ارتباط الرتب لسبيرمان ، وفسره .

٣ – أوجد معامل ارتباط بيرسون .

٤ – ارسم شكل الانتشار وبيّن نوع العلاقة وطبيعتها بين المتغيرين .

تمرين (8) :

أخذت عينة عشوائية من (8) أسر حيث جمعت معلومات تتعلق بدخولها وادخاراتها فتبين ما يلي :

$$\sum X\,Y = 7400$$
$$\sum X^2 = 34200$$
$$\sum X = 520$$
$$\sum Y = 112$$
$$\sum Y^2 = 1689$$

حيث (x) تمثل دخل الأسرة وتمثل (y) ادخارها .

والمطلوب :

١ – احسب ثوابت المعادلة وفسرها .

٢ – ارسم مستقيم الانحدار الناتج واذكر خواصه .

٣ – هل من أهمية إحصائية لمعامل الارتباط ، علل ذلك بما تراه مناسباً .

٤ – لقد وجد أن أسرة دخلها (56) . وحدة نقدية وتدخر (18) وحدة نقدية، فهل تعتقد أن ادخارها يقع ضمن المجال وذلك باحتمال 95 % .

الفصل السابع
السلاسل الزمنية

٧ ــ ١ : تعريف السلسلة الزمنية

٧ ــ ٢ : مجالات السلاسل الزمنية

٧ ــ ٣ : العوامل المؤثرة في السلسلة الزمنية

٧ ــ ٤ : تحليل السلسلة الزمنية

٧ ــ ٥ : التذبذب في السلسلة الزمنية

٧ ــ ٦ : التغيرات الموسمية (الفصلية ، الشهرية)

٧ ــ ٧ : التغيرات الدورية والتغيرات العشوائية


وقل رب زدني علما
جامعة دمشق
Damascus University

٧ ــ ١ : تعريف السلسلة الزمنية :

تعرف السلسلة الزمنية لأية ظاهرة بأنها التطور التاريخي لهذه الظاهرة، وبكلام آخر هي القيم أو المقاييس التي تتخذها الظاهرة في فترات زمنية متتابعة، قد تكون أياماً أو أسابيع أو أشهر أو سنين .

٧ ــ ٢ : مجالات السلاسل الزمنية :

- **السلاسل الزمنية الاقتصادية :**

من السلاسل الزمنية الاقتصادية : أسعار الأسهم في الأيام المختلفة، صادرات دولة ما في الأشهر المختلفة، واردات دولة ما في الأشهر المختلفة، أرباح الشركات في السنوات المختلفة، وأسعار السلع المختلفة في الأشهر المختلفة ... إلخ .

- **السلاسل الزمنية الفيزيائية :**

من السلاسل الزمنية الفيزيائية : كميات الأمطار في الأيام المختلفة، سلسلة درجات حرارة الجو في الساعات المختلفة، وسلسلة الهزات الأرضية، .... إلخ.

- **السلاسل الزمنية للتسويق :**

من السلاسل الزمنية للتسويق سلسلة حجم مبيعات سلعة معينة في أشهر مختلفة، والتنبؤ بحجمها في الأشهر القادمة والتعرف على العلاقات بين حجم المبيعات وأثر الإعلانات الدعائية عليها .

- **السلاسل الزمنية السكانية :**

من السلاسل الزمنية السكانية سلاسل النمو السكاني للمدن من السنوات المختلفة .

- **السلاسل الزمنية لمراقبة الإنتاج :**

من السلاسل الزمنية لمراقبة الإنتاج سلاسل إنتاج الشركات في السلع الإنتاجية يومياً أو أسبوعياً أو شهرياً أو فصلياً أو سنوياً .

- **السلاسل الزمنية الثنائية :**

وهي السلاسل الزمنية التي تحدث مع الزمن فتأخذ إحدى قيمتين (0)، أو (١) (فشل أو نجاح) وتحدث مثل هذه السلاسل في نظرية الاتصالات الكهربائية والهندسية والميكانيكية.

- **السلاسل الزمنية النقطية :**

السلاسل الزمنية النقطية تحدث بشكل عشوائي في أوقات غير متوقعة كسلاسل كوارث حوادث سقوط الطائرات، أو سلاسل حوادث القطارات أو سلاسل حوادث السيارات.

٧ ــ ٣ : العوامل المؤثرة على السلسة الزمنية :

تتكون السلسلة الزمنية من أربعة عناصر أساسية هي :

أولاً : تغيرات الاتجاه العام : Trend

يمثل الاتجاه العام للسلسلة الخطوة الأولى من خطوات دراسة السلسلة الزمنية بتمثيلها بيانياً بالرسم، فنحصل على خط متعرج يسمى خط الاتجاه العام للسلسلة الزمنية،أو المنحنى التاريخي للسلسلة. وعلى الرغم من تعرج خط الاتجاه العام، إلا أنه يحدد اتجاهاً خاصاً يمثل الظاهرة مبيناً نموها، أو تناقصها، أو ثباتها مع مرور الزمن، ويسمى اتجاه خط السلسلة الزمنية صعوداً أو هبوطاً "**الاتجاه العام للسلسلة الزمنية**"، لذا، فإن ا لاتجاه

العام للسلسلة الزمنية هو الطريق الذي تسلكه المتغيرات ممهدة أو منتظمة في تغيرها المتسلسل، خلال فترة طويلة من الزمن.

ثانياً : التغيرات الموسمية (الفصلية) :

هي التغيرات الطارئة على الظاهرة على مدار مواسم مختلفة، لفترة زمنية تمثل موضوع القياس، وتختلف مدتها باختلاف طبيعة الظاهرة بشرط أن لا يزيد طول الفترة الزمنية عن سنة واحدة كحد أعلى. ويتم قياس التغيرات الموسمية من حساب متوسط قيم الظاهرة في كل موسم من المواسم التي تتعرض لها الظاهرة، ثم ينسب كل متوسط للمتوسط العام لهذه المتوسطات الموسمية. إذ يعتبر المتوسط العام (١٠٠%)، وتنسب إليه القيم المختلفة "**القيم الشهرية مثلاً**" فنحصل على أرقام تدل على مدى تغيرات للظاهرة، هل هي فوق هذا المتوسط أم دون.

وتتمثل أهمية دراسة التغيرات الموسمية في تخطيط عمليات الإنتاج أو توقيت الإعلان عن السلع أو التوسع بالمشروعات المختلفة، والكشف عن أوقات التغير، مسببات التغير، والاستعداد لمواجهتها .

ثالثاً : التغيرات الدورية :

تختلف التغيرات الدورية عن التغيرات الموسمية في أنها تمتد لفترة زمنية أطول من سنة وينشأ عنها تغيرات عامة. ومن المظاهر الشائعة لدورة كاملة: فترات الرخاء، الارتداد، الكساد، والانتعاش. وهذه المظاهر تعزى إلى العوامل التي تتحكم في الحياة الاقتصادية للبلاد، بينما تعزى التغيرات الموسمية لأسباب عديدة منها: الجوار، العادات والتقاليد الاجتماعية.

وتأخذ الدورة شكل موجة يتحدد طولها بالفترة الزمنية بين قاعي موجتين متتاليتين، أو قمتيين متتاليين. وقد تستغرق الدورة عدداً كبيراً من السنوات، قد تمتد إلى خمسين عاماً. وهذه الدورة طويلة، أما الدورة المتوسطة فتمتد بين (٨ ــ ١٢) سنة، أما الدورة القصيرة فتكون بين (٣ــ٤) سنوات.

رابعاً : التغيرات الفجائية أو العشوائية :

تؤثر هذه التغيرات على السلسلة بشكل عشوائي ومفاجئ وغير منتظم، فهي قد تكون ناتجة عن حدوث ظواهر طبيعية، مثل الزلازل والبراكين، أو ناتجة عن الحروب والحرائق، وهكذا فإن من الصعب التنبؤ بحدوثها. لذا، فإن من الصعب تحديد حجم هذه التغيرات أو تحديد مدى تأثيرها على الظاهرة، أي :

$$\text{التغيرات العشوائية} = \frac{\text{القيمة الحقيقية}}{\text{القيمة الاتجاهية} \times \text{القيمة الدورية} \times \text{القيمة الموسمية}}$$

تمتاز التغيرات العشوائية بميزات عديدة منها: حدوثها لا يخضع لقاعدة أو قانون، فقد تتكرر أو لا تتكرر، تحدث مرة بالزيادة وأخرى بالنقصان، لا تستقر طويلاً. لذا تسمى بالتغيرات قصيرة الأجل، وتعتمد على عوامل عارضة لا تكون في الحسبان ومن الصعب التنبؤ بها.

٧ ــ ٤ : طرائق تحليل السلسلة الزمنية :

١ ــ الاتجاه العام : Trend

يستخدم في قياس الاتجاه العام طرق مختلفة، معظمها ذات قيمة عفوية لا يمكن الاعتماد عليها، إلا لغايات السرعة في الوقوف على طبيعة الظاهرة، ومن هذه الطرق :

طريقة الرسم الحر : Freehand Method

يتم تمهيد خط الاتجاه العام للسلسلة الزمنية "**المنحنى التاريخي للسلسلة الزمنية**" باليد بشكل تقريبي، وقد يكون ذلك خطاً مستقيماً يسمى " **خط الاتجاه العام للسلسلة الزمنية**" وقد يكون منحنياً. ومن خط الاتجاه العام للسلسلة يمكن أن نستنتج، بصفة عامة، ما إذا كانت قيم الظاهرة في زيادة مضطردة، أو انخفاض مستمر أو غير ذلك. ولا يعتمد على طريقة التمهيد باليد بشكل كبير، لأنها تقديرية تختلف من شخص لآخر لاعتمادها على مهارته في رسم خط يمر بأكبر عدد ممكن من النقاط وبقيتها تتوزع على جانبيه بالتساوي ويمثل السلسلة أفضل تمثيل.

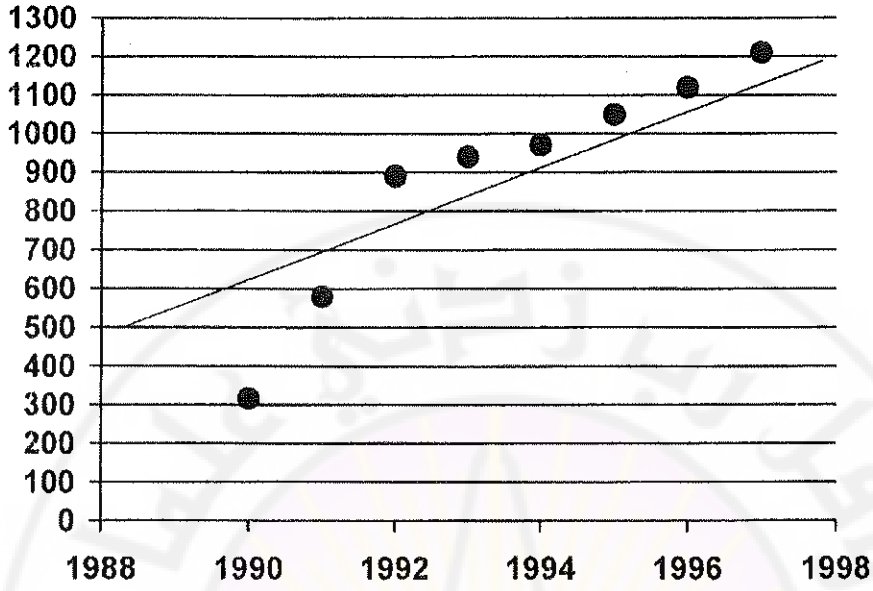
مثال :

يمثل الجدول التالي تطور عدد الطلبة في كلية العلوم السياسية في إحدى الجامعات خلال ثماني سنوات **والمطلوب:** أرسم خط الاتجاه العام للسلسلة الزمنية لعدد الطلبة في هذه الكلية:

السنة	1990	1991	1992	1993	1994	1995	1996	1997
عدد الطلبة	317	580	890	940	970	1050	1120	1210

الحل :

يمكن تمثيل السلسلة الزمنية كما يلي :



٢ ــ طريقة المتوسطات النصفية :

تعتبر هذه الطريقة أدق من طريقة التمهيد باليد. وتتلخص هذه الطريقة بالخطوات الآتية:

١ ــ تقسم السلسلة إلى مجموعتين متكافئتين .

٢ ــ نوجد المتوسط الحسابي لكل من هاتين المجموعتين .

٣ ــ نحدد مكان كل متوسط بنقطة في المستوى البياني .

٤ ــ نصل بين النقطتين اللتين تمثلان هذين المتوسطين بخط مستقيم فيتكون ا لخط الناتج وهو خط الاتجاه العام .

أما إذا أردنا تقييم هذه الطريقة فإن نتائجها يمكن الشك بها، وذلك لأن كل متوسط من المتوسطين يتأثر بالقيم المتطرفة، ولحساب معادلة خط الاتجاه العام نستعمل القانون القائل بأن معادلة الخط المستقيم المار بنقطتين (y_2, x_2)، (y_1, x_1) هي :

$$\frac{y_2 - y_1}{x_2 - x_1}$$

مثال :

يمثل الجدول التالي مبيعات إحدى الشركات بآلاف الوحدات النقدية خلال (١٠) سنوات.

السنوات	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
المبيعات بالاف الوحدات النقدية	150	160	170	180	200	210	220	240	270	280

أرسم خط الاتجاه العام للسلسلة الزمنية باستخدام طريقة المتوسطات النصفية:

الحل :

لرسم خط الاتجاه العام للسلسلة الزمنية نتبع الخطوات الآتية :

- يجب تحديد مفردات السلسلة والسلسلة مكونة من عشر سنوات.
- نأخذ المفردات الخمس الأولى كمجموعة أولى ونستخرج متوسطها الحسابي .

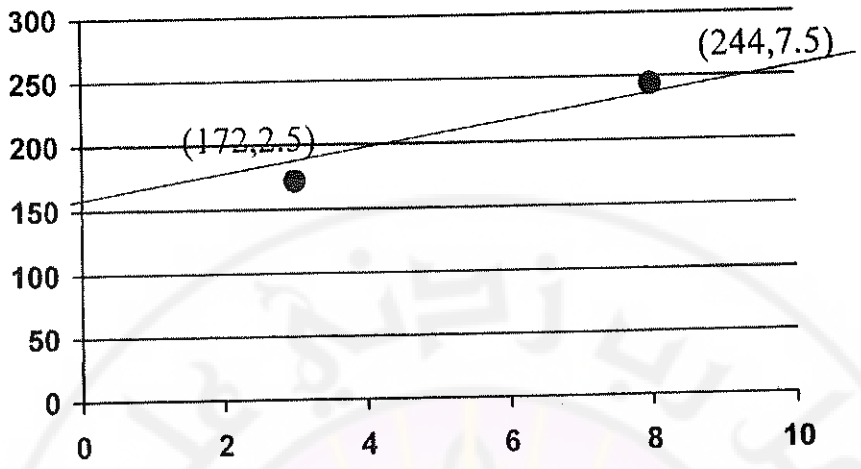
$$\overline{y_1} = \frac{150+160+170+180+200}{5} = 172$$

- نأخذ المفردات الخمس الأخيرة كمجموعة ثانية ونستخرج متوسطها الحسابي .

$$\overline{y_2} = \frac{210+220+240+270+280}{5} = 244$$

فتكون القيمتان ١٧٢، ٢٤٤ هما النقطتان المطلوبتان، وفي الحقيقة فإن ١٧٢ تمثل الإحداثي الرأسي للنقطة الأولى كما أن ٢٤٤ تمثل الاحداثي الرأسي للنقطة الثانية. أما الإحداثي الأفقي للنقطة الأولى فيمثل بمستوى يتلاءم مع نصف الفترة الزمنية التي غطتها الفترة ا لأولى، والإحداثي الأفقي للنقطة الثانية يتمثل بمستوى يتلائم مع نصف الفترة الزمنية التي غطتها الفترة الثانية. وهكذا. فالاحداثي الأفقي للنقطة الأولى يقع على مستوى سنتين ونصف، والفترة الثانية يقع على مستوى سبع سنوات ونصف.

- نعين النقطتين على الرسم بحيث يكون إحداثياً النقطة الأولى هما (٢,٥ ، ١٧٢)، وإحداثياً النقطة الثانية هما (٧,٥ ، ٢٤٤).
- نصل بين النقطتين بخط مستقيم فيكون هو خط الاتجاه العام المطلوب .



٣ ــ طريقة المتوسطات المتحركة :

تعتمد طريقة المتوسطات المتحركة على أخذ متوسطات متتابعة ومتداخلة لمجموعة متتابعة من المشاهدات، وينتج عنها إزالة بعض تعرجات الظاهرة في خط المنحنى التاريخي للسلسلة الزمنية، وبالتالي تمهيده على افتراض أن السلسلة تتألف من (١٠) مفردات رمز لها بالرمز ($X_1, X_2 \ldots X_n$) . ولنفرض أن المتوسط بطول (٣) مفردات فإن الطريقة تكون على النحو التالي:

$$\overline{X_1} = \frac{X_1 + X_2 + X_3}{3}$$

$$\overline{X_2} = \frac{X_2 + X_3 + X4}{3}$$

وهكذا حتى تنتهي جميع المفردات، ثم نأخذ المتوسطات، ونعينها كإحداثيات لنقاط المنحنى التاريخي، ونصل بينها، فيكون الخط الناتج هو الاتجاه العام المطلوب .

مثال :

يمثل الجدول التالي تطور عدد الطلبة في كلية العلوم السياسية بإحدى الجامعات خلال عشر سنوات:

السنوات	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
عدد الطلبة	150	160	170	180	200	210	220	240	270	280

أوجد المتوسطات المتحركة بطول (٣) .

$$X_1 = \frac{150+160+170}{3} = 160$$

$$X_2 = \frac{160+170+180}{3} = 170$$

$$X_3 = \frac{170+180+200}{3} = 183.3$$

وهكذا حتى نهاية جميع المفردات كما هو موضح في الجدول المرفق .

وترتب المتوسطات المتحركة في الجدول على النحو التالي :

السنوات	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
عدد الطلبة	150	160	170	180	200	210	220	240	270	280
المتوسطات المتحركة	-	160	170	183.3	196.7	210	223.3	243.3	263.3	

مثال :

يمثل الجدول التالي سلسلة زمنية لظاهرة معينة عبر إحدى عشرة سنة .

السنوات	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
القيم	50	36.5	43	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8

أوجد المتوسطات المتحركة بطول خمس سنوات .

الحل :

$$X_1 = \frac{50 + 36.5 + 43 + 44.5 + 38.9}{5} = 42.6$$

$$X_2 = \frac{36.5 + 43 + 44.5 + 38.9 + 38.1}{5} = 42.6$$

وهكذا بقية قيم الجدول كما هو موضح بالجدول المرفق .

وترتب المتوسطات المتحركة في الجدول كما يلي :

السنوات	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
القيم	50	36.5	43	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8
المتوسطات المتحركة	-	-	42.6	40.2	39.4	38.6	38.0	38.4	37.6		

مثال :

يمثل الجدول التالي بيانات عن كميات سلسلة زمنية لظاهرة معينة :

السنوات	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
كمية الإنتاج	50	36.5	43	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8

أوجد ما يلي :

١) المتوسطات المتحركة بطول خمس سنوات .

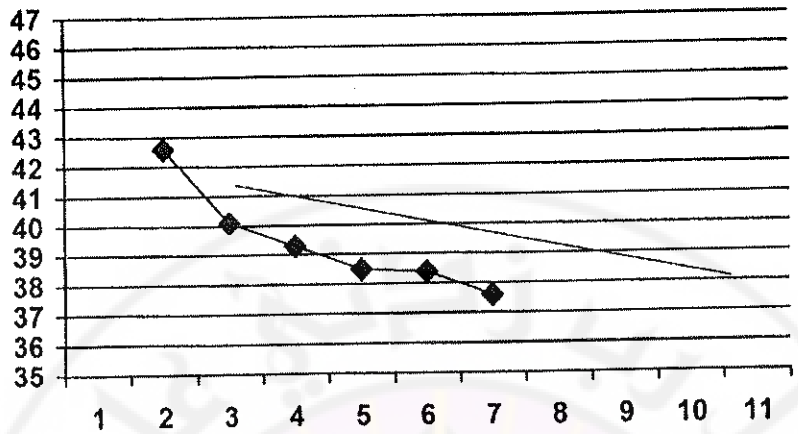
٢) رسم المتوسطات المتحركة .

٣) رسم خط الاتجاه العام .

الحل :

١) الجدول التالي يمثل المتوسطات المتحركة للجدول السابق :

رمز السنوات	0	1	2	3	4	5	6	7	8	9	10
السنوات	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
القيم	50	36.5	43	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8
المتوسطات المتحركة	-	-	42.6	40.2	39.4	38.6	38.0	38.4	37.6	-	-



٤ ــ طريقة المربعات الصغرى :

تتبع نفس خطوات معادلة الخط المستقيم أو المنحنى، ويكون الخط المستقيم أو المنحنى ممثلاً للقيمة الاتجاهية تمثيلاً صادقاً عندما يكون مجموع مربعات الانحراف لهذه القيم الاتجاهية عن القيم الأصلية أصغر ما يمكن، ومعادلة هذه الطريقة هي :

$$\hat{y} = a + bx$$

ويمكن تمثيل شكل الانتشار للسلسلة الزمنية بخط مستقيم على اعتبار أن (a,b) قيمتان ثابتتان، حيث (x) تعبر عن أرقام السنوات ابتداء بالصفر، وهي تعبر عن القيم الأصلية في السلسلة الزمنية، وبعد استخراج قيمة (y)، نكون بذلك قد استخرجنا القيمة الاتجاهية للسلسلة في كل سنة من السنوات، ثم تؤخذ هذه القيمة الاتجاهية وتعين كنقطة على الرسم البياني، ثم يوصل بينهما خط مستقيم، ويكون المستقيم الناتج هو خط الاتجاه العام للسلسلة.

مثال :

يمثل الجدول التالي سلسلة زمنية لمبيعات شركة ما بآلاف الوحدات النقدية .

السنة	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
المبيعات	150	160	170	180	200	210	220	240	270	280

أحسب القيم الاتجاهية التي يمكن منها رسم خط الاتجاه العام لتلك السلسلة الزمنية .

الحل : نرمز للسنوات بترقيمها، حيث نفرض أن رقم السنة الأولى (صفر) تكونها سنة الأساس:

السنة x	المبيعات y	x	X^2	xy
1981	150	0	0	0
1982	160	1	1	160
1983	170	2	4	340
1984	180	3	6	540
1985	200	4	16	800
1986	210	5	25	1050
1987	220	6	36	1320
1988	240	7	49	1680
1989	270	8	64	2160
1990	280	9	81	2520
المجموع	2080	45	285	10750

$$\hat{y} = a + bx$$

$$(1) \quad \sum y = na + b\sum x$$

$$(2) \quad \sum xy = a\sum x + b\sum x^2$$

نستخدم المعادلات أعلاه بفرض أن خط الاتجاه العام مستقيم، ثم نعوض فيها:

$$(1) \qquad (45) / 2080 = 10a + 45b$$

$$(2) \qquad (10)10570 = 45a + 285\ b$$

بعد ضرب المعادلتين بالأرقام المشار إليها أعلاه نحصل على :

$$(3) \qquad 93600 = 450a + 2025b$$

$$(4) \quad 105700 = 450\ a + 2850\ b$$

بطرح المعادلة (٤) من المعادلة (٣) :

$$-12100 = -825b$$

$$b = 14.67$$

بالتعويض بقيمة (b) في معادلة (١) نحصل على :

$$2080 = 10a + 45\ (14.667)$$

$$a = 142$$

وبالتالي تصبح معادلة خط الاتجاه العام للسلسلة الزمنية

$$\hat{y} = 142 + 14.67x$$

وبتعويض قيمة (x) في السنة الأولى في هذه المعادلة تصبح $\hat{y}$ قيمة النقطة الأولى حيث يقطع خط الاتجاه العام المحور العمودي وذلك لأن قيمة (x) في هذه السنة صفر أي أن الإحداثي للنقطة الأولى (١٤٢ ، ٠).

مثال :

يمثل الجدول التالي السلسلة الزمنية لكميات سلسلة زمنية لظاهرة معينة .

السنوات	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
القيم	50	36.5	43	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8

جد ما يلي :

١) الاتجاه العام بطريقة المربعات الصغرى .

٢) معادلة الاتجاه العام للسلسلة باعتبار سنة ١٩٩٢ كسنة أساس .

الحل :

السنوات	كمية الإنتاج	x	X^2	xy
1987	50	0	0	0
1988	36.5	1	1	36.5
1989	43	2	4	86
1990	44.5	3	9	133.5
1991	38.9	4	16	155.6

190.5	25	5	38.1	1992
195.6	36	6	32.6	1993
270.9	49	7	36.7	1994
333.6	64	8	41.7	1995
396.9	81	9	41.1	1996
338	100	10	33.8	1997
2110.1	385	55	438.9	المجموع

بتعويض المجاميع في المعادلتين (١) ، (٢) ، نحصل على :

(1) $438.9 = 11a + 55b$

(2) $2110.1 = 55a + 385b$

نضرب المعادلة (١) بـ (٥) ثم نطرح منها المعادلة (٢) فنحصل على :

(3) $2194.5 = 55a + 275b$

بطرح المعادلة (٣) من المعادلة الأولى (٢) ينتج :

$$84.4 = -110b$$
$$b = -0,767$$

أي أن الاتجاه العام لهذه السلسلة يساوي (0,767-) في العام الواحد، وبالتعويض في المعادلة (١) نجد أن : a=43.73 والمعادلة هي :

$$\hat{y} = 43.73 - 0.767x$$

الطلب (٢) :

السنوات	كمية الإنتاج	x	xy	X^2
1987	50	5-	250-	25
1988	36.5	4-	146-	16
1989	43	3-	129-	9
1990	44.5	2-	89-	4
1991	38.9	1-	38.9-	1
1992	38.1	0	0	0
1993	32.6	1+	32.6	1
1994	36.7	2	77.4	4
1995	41.7	3	125.1	9
1996	41.1	4	164.4	16
1997	33.8	5	169	25
المجموع	438.9	0	84.4-	110

يمكن إيجاد المعادلة بأسلوب المعادلات الطبيعية السابقة أو بالطريقة المباشرة التالية:

$$b = \frac{\sum xy - n\bar{x}\,\bar{y}}{\sum x^2 - n\bar{x}^2} = \frac{(-84.4) - (11)(0)(399)}{110 - 11(0)} = -0.767$$

$$a = \bar{y} - b\bar{x} = 39.9 - (-0.767)(0) = 39.9$$

وبالتالي المعادلة هي : $\hat{y} = 39.9 - 0.767x$

ـــ نلاحظ أن قيم (b) لم تتغير لأن ميل خط الاتجاه العام لا يمكن أن يتغير أما قيمة (a) فتغيرت حيث أنها تمثل الجزء المقطوع من المحور الرأسي بواسطة خط الاتجاه العام أو امتداده، وهذا الجزء لابد أن يتغير إذا غيرنا نقطة الأصل الخاصة بالسلسلة. وعلى أساس معادلة الاتجاه العام يمكن حساب القيم الاتجاهية في قيم الاتجاه العام (x) حسب ترتيب السنوات. فإذا أردنا حساب القيمة الاتجاهية لعام ١٩٨٧ يكون الحل كالآتي :

$$\hat{y} = 39.9 - 0.767(-5) = 43.735$$

مثال :

يمثل الجدول التالي بيانات عن كميات استهلاك السكر بآلاف الأطنان في دولة ما خلال ثمان سنوات .

منتصف السنة x	1990	1991	1992	1993	1994	1995	1996	1997
كميات السكر y	273.7	293.5	315	336.8	364.4	394.4	424.2	458.7

أوجد ما يلي :

١) معادلة خط الاتجاه العام (سنة الأساس منتصف عام ١٩٩٠) .

٢) الاتجاه العام السنوي والربع سنوي والشهري .

الحل :

X^2	xy	y	x
0	0	273.7	0
4	293.5	293.5	1
4	630	315.0	2
9	1010.4	336.8	3
16	1457.6	364.4	4
25	1972	394.4	5
36	2454.2	424.2	6
49	3210.9	458.7	7
140	11119.6	2860.7	28

وبتعويض المجاميع نجد أن :

$$b = \frac{11119.6 - 8(3.5)(357.58)}{140 - (8)(3.5)^2} = 26.4$$

$$a = 357.58 - 26.4(3.5) = 265.19$$

معادلة الاتجاه العام :

$$\hat{y} = 265.19 + 26.4x$$

وبذلك يتضح أن الاتجاه العام= ٢٦,٤ ألف طن سكر في العام الواحد بالمتوسط.

٢) الاتجاه العام الربع سنوي = $\frac{26.41}{4} = 6.6$ ألف طن سكر .

الاتجاه العام الشهري = $\frac{26.41}{12} = 2.20$ ألف طن سكر .

مثال :

يمثل الجدول الآتي بيانات عن كميات استهلاك إحدى الدول من الدقيق (ألف طن).

السنوات	الربع الأول	الربع الثاني	الربع الثالث	الربع الرابع
2000	45	55	62	64
2001	53	49	59	60
2002	62	64	67	69
2003	75	66	69	70
2004	71	68	75	79

أوجد ما يلي :

١) معادلة خط الاتجاه العام. (نقطة الأساس الربع الثاني لعام ٢٠٠٢)

٢) الاتجاه العام السنوي .

٣) القيمة الاتجاهية في الربع الرابع عام ٢٠٠٢ .

الحل :

السنة	الفصل	x	كميات استهلاك الدقيق y	xy	X^2
2000	1	9-	45	405-	81
	2	8-	55	440-	64
	3	7-	62	434-	49
	4	6-	64	384-	36
2001	1	5-	53	265-	25

16	196-	49	4-	2	
9	177-	59	3-	3	
4	120-	60	2-	4	
1	62-	62	1-	1	2002
0	0	64	0	2	
1	67	67	1	3	
4	138	69	2	4	
9	225	75	3	1	2003
16	264	66	4	2	
25	345	69	5	3	
36	420	70	6	4	
49	497	71	7	1	2004
64	544	68	8	2	
81	675	75	9	3	
100	790	79	10	4	
570	1482	1282	10		المجموع

وبتعويض المجاميع نجد أن :

$$b = \frac{1482 - 20(0.5)(64.1)}{570 - 20(0.5)^2} = 1.49$$

$$a = 64.1 - 1.49(0.5) = 56.65$$

وبالتالي معادلة الاتجاه العام : $\hat{y} = 56.65 + 1.49x$ (محسوبة من نقطة الأصل الربع الثاني)

٢) من المعادلة السابقة يتبين أن الاتجاه العام للربع سنوي = ١,٤٩ ألف طن من الدقيق، أي أن استهلاك هذه الدولة من الدقيق يزيد كل ربع عن

الربع السابق له بمقدار ١,٤٩ ألف طن من الدقيق بسبب النمو السكاني العام فيها، وبذلك يكون الاتجاه العام السنوي :

الاتجاه العام السنوي ٥,٩٦ = ٤ × ١,٤٩ ألف طن من الدقيق .

٣) القيمة الاتجاهية في الربع الرابع لعام ٢٠٠٢ : ٥٩,٦ = ٥٦,٦٥ + (٢)١,٤٩ ألف طن من الدقيق استهلاك هذه الدولة .

ـ مزايا طريقة المربعات الصغرى :

تمتاز طريقة المربعات الصغرى بميزات عديدة منها: عدم اعتمادها على التقدير الشخصي، يمكن بمقتضاها صياغة القيم صياغة رياضية واضحة المعالم، يمكن استنباط قيم أخرى عن الحاضر والماضي والمستقبل بموجب هذه المعادلة، كما تساعد في عملية التنبؤ شريطة أن لا تكون الفترة المتنبأ بها بعيدة جداً عن الحاضر .

ـ عيوب طريقة المربعات الصغرى :

تعاب طريقة المربعات الصغرى بأمور عديدة منها: اعتمادها على الصيغ الرياضية التي قد لا يفهمها أو لا يتقنها البعض، افتراضها خضوع الظواهر في سلوكها للمعادلات الرياضية وهو أمر غير واقعي في أحيان كثيرة، وتعقد عملياتها الحسابية خصوصاً إذا كانت الأرقام كبيرة إلا أن هذه العيوب لا تقلل من أهمية وفائدة هذه الطريقة، خاصة إذا ما قورنت بالطرق الأخرى.

٧ ـ ٥ : التذبذب في السلسلة الزمنية :

تتذبذب قيم السلسلة الزمنية من وقت لآخر وهذا يؤدي إلى جعل السلسلة الأصلية غير ملساء، لذا، نلجأ إلى استعمال المعدلات المتحركة

للتقليل من هذه التذبذبات والحصول على سلسلة ملساء يمكن تقدير إحدى مركبات السلسلة المسماة بمركبة التذبذب، من المعادلة الآتية:

التذبذب = المشاهدة الأصلية في السلسلة ــ المعدل المتحرك المقابل لها .

التذبذب = $(y - \bar{y})$

ويمكن استعمال هذا في التعرف فيما إذا كان طول المعدل المتحرك (L) فردياً، أما إذا كان طول المتوسط المتحرك زوجياً، نلجأ إلى أسلوب آخر في إيجاد متوسط متحرك مركزي.

٧ ــ ٦ : التغيرات الموسمية (الفصلية، الشهرية):

لإيجاد المركبة الفصلية نتبع الخطوات الآتية :

١) جعل السلسلة ملساء عن طريق أخذ معدل متحرك بطول مناسب، فإذا كانت البيانات في السنوات المختلفة معطاة للفصول الأربعة نأخذ متوسطاً مركزياً متحركاً طوله (٤)، وإذا كانت البيانات معطاة في السنوات المختلفة للأشهر الإثني عشر فإننا نأخذ متوسطاً مركزياً متحركاً طول (١٢) وهكذا .

٢) نقسم كلاً من البيانات في السلسلة الأصلية على قيم المتوسط المتحرك المركزي المناظر لها ونضرب الناتج في (١٠٠).

٣) نرتب الأعداد التي حصلنا عليها بحيث نجد المتوسط الحسابي لنسبة كل فصل على انفراد.

٤) نوجد مجموع متوسطات الفصول المختلفة.

٥) نوجد التغيرات الفصلية من العلاقة التالية :

$$\text{مركبة الفصل} = \frac{\text{متوسط الفصل}}{\text{مجموع متوسطات الفصول}} \times n \times ١٠٠$$

حيث (n) عدد الفصول في السنة الواحدة .

مثال :

يمثل الجدول التالي السنوات والفصول والمشاهدات لسلسلة زمنية معينة .

السنوات	2000			2001			2002			2003			2004			2005		
الفصول	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
المشاهدات	3	8	5	10	8	12	6	12	4	7	13	11	6	5	4	8	10	6

أوجد ما يلي :

١) مركبات الفصل للسلسلة الزمنية الواردة أعلاه .

الحل :

١) لإيجاد المركبات الفصلية للسلسلة الزمنية نتبع الخطوات الآتية :

أولاً : نجد المتوسطات المتحركة بطول (٣) مساوياً لعدد فصول السنة .

$$\overline{y}_1 = \frac{3+8+5}{3} = 5.3$$

$$\overline{y}_2 = \frac{8+5+10}{3} = 7.7$$

وهكذا حتى نهاية السلسلة بنفس الأسلوب السابق كما هو موضح بالجدول .

ثانياً : نفرغ المشاهدات، المتوسطات المتحركة، $(\frac{y}{\overline{y}} \times 100)$ في الجدول التالي :

2005			2004			2003			2002			2001			2000			السنوات
3	2	1	3	2	1	3	2	1	3	2	1	3	2	1	3	2	1	الفصول
6	10	8	4	5	6	11	13	7	4	12	6	12	8	10	5	8	3	المشاهدات
-	8	7.3	5.7	5	7.3	10	10.3	8	7.7	7.3	10	8.7	10	7.7	7.7	5.3	-	المتوسط المتحرك y
-	125	110	70	100	82	110	126	88	52	164	60	138	80	130	65	151	-	$\frac{y}{\bar{y}} \times 100$

ثالثاً : نحسب متوسط كل فصل على السنوات المختلفة .

متوسط الفصل الأول = $\frac{130+60+88+82+110}{5} = 94$

متوسط الفصل الثاني = $\frac{151+80+164+126+100+125}{6} = 124.3$

متوسط الفصل الثالث = $\frac{65+138+52+110+70}{5} = 87$

رابعاً : نوجد مجموع متوسطات الفصول : ٣٠٥,٣ = ٨٧ + ١٢٤,٣ + ٩٤

خامساً : نوجد مركبات الفصل على النحو التالي :

$$\text{مركبة الفصل} = \frac{\text{متوسط الفصل}}{\text{مجموع متوسطات الفصول}} \times n \times 100$$

الفصل الأول = $\frac{94\times 3}{305.3}\times 100 = 92.37$

الفصل الثاني = $\frac{124.3\times 3}{305.3}\times 100 = 122.14$

الفصل الثالث = $\frac{87 \times 3}{305.3} \times 100 = 85.43$

٧ ــ ٧ : التغيرات الموسمية الشهرية :

يمكن إيجاد التغيرات الموسمية للسلسلة الزمنية باتباع الخطوات الآتية :

١. نوجد المتوسطات الشهرية .

٢ . نوجد متوسط متوسطات الأشهر جميعاً .

٣ . نوجد النسبة الموسمية لكل شهر

$$\text{النسبة الموسمية} = \frac{\text{متوسط الشهر}}{\text{متوسط المتوسطات}} \times ١٠٠\%$$

٤ ــ نجد التغير الموسمي من العلاقة الآتية :

التغير الموسمي = نسبة الموسم ــ ١٠٠%

فإذا زادت النسبة عن ١٠٠% يدل ذلك على وجود تغير موسمي تصاعدي والعكس إذا نقصت النسبة عن ١٠٠ يدل ذلك على وجود تغير موسمي تنازلي.

مثال :

أوجد التغيرات الموسمية للسلسلة الزمنية التالية :

12	11	10	9	8	7	6	5	4	3	2	1	الشهر
347	325	302	269	245	223	216	231	250	278	281	318	1995
364	342	321	288	262	242	236	249	268	299	309	342	1996
394	367	345	309	284	259	251	269	287	320	328	367	1997
417	389	364	328	305	282	273	290	311	342	349	392	1998
452	442	396	356	330	305	296	214	324	370	378	420	1999
483	454	427	392	359	225	322	341	362	398	412	453	2000
516	491	457	415	388	357	347	370	392	429	440	387	2001
560	526	493	448	419	389	380	398	423	463	477	529	2002
3533	333	3105	2809	2592	2282	2321	2362	2617	2899	2974	3208	المجموع

الحل :

لإيجاد التغيرات الموسمية للسلسلة الزمنية نتبع الخطوات الآتية :

١) نوجد المتوسطات الشهرية كما يلي :

$$\text{متوسط شهر (١)} = \frac{٥٢٩ + ٣٨٧ + ٤٥٣ + ٤٢٠ + ٣٩٢ + ٣٦٧ + ٣٤٢ + ٣١٨}{٨} = ٤٠١$$

متوسط شهر (٢) = ٣٧٢

متوسط شهر (٣) = ٣٦٢

متوسط شهر (٤) = ٣٢٧

متوسط شهر (٥) = ٢٩٥

متوسط شهر (٦) = ٢٩٠

متوسط شهر (٧) = ٢٨٥

متوسط شهر (٨) = ٣٢٤

متوسط شهر (٩) = ٣٥١

متوسط شهر (١٠) = ٣٨٨

متوسط شهر (١١) = ٤١٧

متوسط شهر (١٢) = ٤٤٢

٢) نوجد متوسط المتوسطات :

$$\frac{٤٠١ + ٣٧٢ + ٣٦٢ + ٣٢٧ + ٢٩٥ + ٢٩٠ + ٢٨٥ + ٣٢٤ + ٣٥١ + ٣٨٨ + ٤١٧ + ٤٤٢}{١٢} = ٣٥٤$$

٣) نوجد النسبة الموسمية لكل شهر على انفراد :

$$\text{النسبة الموسمية لشهر (١)} = \frac{٤٠١}{٣٥٤} \times ١٠٠\% = ١١٣\%$$

$$\text{النسبة الموسمية لشهر (٢)} = \frac{٣٧٢}{٣٥٤} \times ١٠٠\% = ١٠٥\%$$

$$\text{النسبة الموسمية لشهر (٣)} = \frac{٣٦٢}{٣٥٤} \times ١٠٠\% = ١٠٢\%$$

$$\text{النسبة الموسمية لشهر (٤)} = \frac{٣٢٧}{٣٥٤} \times ١٠٠\% = ٩٢\%$$

$$\text{النسبة الموسمية لشهر (٥)} = \frac{٢٩٥}{٣٥٤} \times ١٠٠\% = ٨٣\%$$

$$\text{النسبة الموسمية لشهر (٦)} = \frac{٢٩٠}{٣٥٤} \times ١٠٠\% = ٨٢\%$$

٢٨٥

النسبة الموسمية لشهر (٧) = ـــــــــــ × ١٠٠% = ٨١%

٣٥٤

٣٢٤

النسبة الموسمية لشهر (٨) = ـــــــــــ × ١٠٠% = ٩٢%

٣٥٤

٣٥١

النسبة الموسمية لشهر (٩) = ـــــــــــ × ١٠٠% = ٩٩%

٣٥٤

٣٨٨

النسبة الموسمية لشهر (١٠) = ـــــــــــ × ١٠٠% = ١١٠%

٣٥٤

٤١٧

النسبة الموسمية لشهر (١١) = ـــــــــــ × ١٠٠% = ١١٨%

٣٥٤

٤٤٢

النسبة الموسمية لشهر (١٢) = ـــــــــــ × ١٠٠% = ١٢٥%

٣٥٤

٤) نوجد التغير الموسمي لكل شهر :

التغير الموسمي لشهر (١) = ١١٣% ـ ١٠٠% = ١٣% .

التغير الموسمي لشهر (٢) = ١٠٥% ـ ١٠٠% = ٥%

التغير الموسمي لشهر (٣) = ١٠٢% ـ ١٠٠% = ٢%

التغير الموسمي لشهر (٤) = ٩٢% ـ ١٠٠% = ـ٨%

التغير الموسمي لشهر (٥) = ٨٣% ـ ١٠٠% = ـ١٧%

التغير الموسمي لشهر (٦) = ٨٢% ـ ١٠٠% = ـ ١٨%

التغير الموسمي لشهر (٧) = ٨١% ـ ١٠٠% = ـ ١٩%

التغير الموسمي لشهر (٨) = ٩٢% ـ ١٠٠% = ـ ٨% .

التغير الموسمي لشهر (٩) = ٩٩% ـ ١٠٠% = ـ١%

التغير الموسمي لشهر (١٠) = ١١٠% ـ ١٠٠% = ١٠%

التغير الموسمي لشهر (١١) = ١١٨% ـ ١٠٠% = ١٨%

التغير الموسمي لشهر (١٢) = ١٢٥% ـ ١٠٠% = ٢٥% .

وإذا كانت قيم الظاهرة متأثرة باتجاه عام معين، فلا بد من استبعاد هذا الاتجاه قبل حساب النسب الموسمية، حيث أننا إذا لم نستبعده فإن التغيرات الموسمية لا تكون دقيقة، فإذا كان الاتجاه العام تصاعدياً (تزايداً)، وكان التغير الموسمي تصاعدياً فإن النسب الموسمية تكون أكثر مما يجب أن تكون، والعكس إذا كان الاتجاه العام تصاعدياً وكان التغير الموسمي تنازلياً (تناقصياً) فإن النسب الموسمية تكون أقل مما يجب أن تكون. ولاستبعاد الاتجاه العام نوجد معادلة الاتجاه العام للقيم الموجودة لدينا ومنها نحسب القيم الاتجاهية للظاهرة، ثم نقسم كل قيمة واقعة على القيم الاتجاهية

المقابلة لها مضروباً بـــ ١٠٠ فنحصل على نسب تدل على الظاهرة مستبعداً منها الاتجاه العام من خلال العلاقة :

$$\frac{y}{\hat{y}} \times 100$$

مثال :

يمثل الجدول الآتي بيانات عن صادرات دولة عربية ما بالمليون وحدة نقدية

السنوات	الربع الأول	الربع الثاني	الربع الثالث	الربع الرابع
1996	45	55	62	64
1997	53	49	59	60
1998	62	64	67	69
1999	65	66	69	70
2000	71	68	75	79

أوجد ما يلي :

١) التغيرات الموسمية بفرض أنه لا يوجد اتجاه عام في صادرات هذه الدولة.

٢) كم تكون قيمة صادرات هذه الدولة دون تأثير الموسم في الربع عام ٢٠٠٠ .

٣) التغيرات الموسمية بفرض أن قيمة الاتجاه العام الربع المرابع السنوي هي : (b=1.114)

الحل :

لإيجاد التغيرات الموسمية للسلسلة الزمنية نتبع الخطوات الآتية :

١) توجد المتوسطات الفصلية (الربعية) كما يلي :

$$\text{متوسط الربع الأول} = \frac{٤٥ + ٥٣ + ٦٢ + ٦٥ + ٧١}{٥} = ٥٩,٢$$

$$\text{متوسط الربع الثاني} = \frac{٥٥ + ٤٩ + ٦٤ + ٦٦ + ٦٨}{٥} = ٦٠,٤$$

$$\text{متوسط الربع الثالث} = \frac{٦٢ + ٥٩ + ٦٧ + ٦٩ + ٧٥}{٥} = ٦٦,٤$$

$$\text{متوسط الربع الرابع} = \frac{٦٤ + ٦٠ + ٦٩ + ٧٠ + ٧٩}{٥} = ٦٨,٤$$

٢) نوجد :

$$\text{المتوسط العام للفصول} = \frac{٥٩,٢ + ٦٠,٤ + ٦٦,٤ + ٦٨,٤}{٤} = ٦٣,٦$$

٣) نوجد النسب الموسمية لكل ربع :

النسبة الموسمية للربع الأول = ٥٩,٢ / ٦٣,٦ × ١٠٠% = ٩٣,١%

النسبة الموسمية للربع الثاني = ٦٠,٤ / ٦٣,٦ × ١٠٠% = ٩٤,٩%

النسبة الموسمية للربع الثالث = ٦٦,٤ / ٦٣,٦ × ١٠٠% = ١٠٤,٤%

النسبة الموسمية للربع الرابع = ٦٨,٤ / ٦٣,٦ × ١٠٠% = ١٠٧,٥%

٤) نوجد التغير الموسمي لكل ربع :

التغير الموسمي للربع الأول = ٩٣,١% ـ ١٠٠% = ـ ٦,٩% .

التغير الموسمي للربع الثاني = ٩٤,٩% ـ ١٠٠% = ـ ٥,١% .

التغير الموسمي للربع الثالث = ١٠٤,٤% ـ ١٠٠% = ٤,٤%

التغير الموسمي للربع الرابع = ١٠٧,٥% ـ ١٠٠% = ٧,٥%

المبلوغات	1996	1997	1998	1999	2000	y	النسب الموسمية	التغير الموسمي
الربع الأول	45	53	62	65	71	59.2	93.1%	-6.9%
الربع الثاني	55	49	64	66	68	60.4	94.9%	-5.1%
الربع الثالث	62	59	67	69	75	66.4	104.4%	4.4%
الربع الرابع	64	60	69	70	89	68.4	107.5%	7.5%

٢) تدل النسبة الموسمية في الربع الرابع من كل عام على أن صادرات هذه الدولة تكون ١٠٧,٥% من قيمتها بدون تأثير الموسم .

أي صادرات الدولة دون تأثير الموسم في الربع الرابع لعام ٢٠٠٠ :

١٠٠

مليون وحدة نقدية ٨٢,٨ = ٨٩ × ـــــــــــــــ

١٠٧,٥

٣) حيث أن الاتجاه العام معروف ويساوي ١,١١٤ في ربع السنة، لذا، يجب استبعاده مثل حساب التغيرات الموسمية .

الربع	المتوسطات	المتوسطات بعد استبعاد الاتجاه العام	النسب الموسمية
الأول	59.2	59.2	95.6
الثاني	60.4	59.286	95.8
الثالث	66.4	64.172	103.7
الرابع	68.4	65.058	105.2

٦٥,٠٥٨ + ٦٤,١٧٢ + ٥٩,٢٨٦ + ٥٩,٢

المتوسط العام = ـــــــــــــــــــــــــــــــ = ٦١,٩

٤

٥٩,٢

النسب المئوية للربع الأول = ـــــــــــــــ × ١٠٠% = ٩٥,٦% .

٦١,٩

٥٩,٣

النسب المئوية للربع الثاني = ـــــــــــــ × ١٠٠% = ٩٥,٨% .

٦١,٩

٦٤,٢

النسب المئوية للربع الثالث = ـــــــــــــ × ١٠٠% = ١٠٣,٧% .

٦١,٩

٦٥,١

النسب المئوية للربع الرابع = ـــــــــــــ × ١٠٠% = ١٠٥,٢% .

٦١,٩

من النسب الموسمية يمكن الاستنتاج بأن نسب التغيرات الموسمية هي الآتي:

التغير الموسمي في الربع الأول ـ ٤,٤% = ١٠٠% ـ ٩٥,٦%

التغير الموسمي في الربع الثاني ـ ٤,٢% = ١٠٠% ـ ٩٥,٨ %

التغير الموسمي للربع الثالث ٣,٧% = ١٠٠% ـ ١٠٣,٧%

التغير الموسمي في الربع الرابع ٥,٢% = ١٠٠% ـ ١٠٥,٢%

يلاحظ أنه قد يكون معلوماً لدينا معادلة الاتجاه العام السنوي يجب قسمته على (٤) واستبعاده عن القيم الربع سنوية ، أما إذا كانت القيم شهرية يجب قسمته على (١٢).

٧ ــ ٨ : التغيرات الدورية والتغيرات العشوائية :

تتغير التغيرات الدورية في الميدان الاقتصادي بحركات متموجة طويلة المدى تتعاقب فيها الارتفاعات والانخفاضات في النشاط الاقتصادي حول منحني الاتجاه العام ومن الناحية الإحصائية تختلف التغيرات الدورية عن التغيرات الموسمية ففي الأول نجد أنها لا تتكرر في فترات زمنية منتظمة، بينما التحركات الدورية تتكرر بعد فترات تزيد عن سنة، لكن ليس من السهل دوماً التمييز بين التغيرات الدورية والتغيرات العشوائية؛ لذلك في الغالب، يجري تقديرهما معاً على النحو التالي :

$$Y = T.S.C.I$$

١ ــ الطريقة الأولى : استبعاد أثر الاتجاه العام من خلال تقسيم قيم السلسلة الأصلية على قيم الاتجاه العام التي تم حسابها، ثم تقسيم قيم السلسلة الفعلية على القيم الاتجاهية :

$$Y / T = S. C. I$$

ثم نقوم باستبعاد استبعاد العامل الموسمي بعد حساب الدليل الموسمي ليبقى لدينا :

$$Y/S = C . I$$

٢ ــ الطريقة الثانية : تقدير (S.T) جداءات الاتجاه العام بالعامل الموسمي، أي جداء قيم الاتجاه العام المحسوبة بالدليل الموسمي ومن ثم تقسيم قيم السلسلة الأصلية على جداء (S.T)

$$Y = S.T.C.I$$

$$Y = S.T.C.I / T.S = C. T$$

من خلال الاستعراض السابق يمكن القول أن تحليل السلاسل الزمنية يعطي مقدمة لإجراء عملية التنبؤ بالسلاسل الزمنية وإن المعالجة الرياضية قد لا تكفي في بعض الأحيان ولابد من الجمع بين خبرة وكلمة الباحث وبين التحليل الرياضي لإعطاء نتائج دقيقة في التنبؤ وهذه النتائج تتطلب توفر ثلاث شروط :

١ ــ إمكانية الحصول على المعلومات الخاصة بالماضي .

٢ ــ افتراض أن بعض النواحي المختلفة المتعلقة بنمط الماضي سوف تستمر في المستقبل.

٣ ــ إمكانية التعبير عن هذه المعلومات رقمياً .

تمارين

١) تمثل البيانات الآتية كميات الحمضيات لإحدى الدول العربية بالطن في السنوات الست التالية :

السنوات	1996	1997	1998	1999	2000	2001
الأول	60	80	95	130	120	150
الثاني	80	90	100	120	130	140
الثالث	90	70	80	105	110	120
الرابع	70	80	85	90	100	110

أوجد ما يلي :

أ) معادلة خط الاتجاه العام للسلسلة الزمنية

ب) التذبذب للسلسلة الزمنية

ج) التغيرات الفصلية للسلسلة الزمنية

د) التغيرات الموسمية للسلسلة الزمنية .

٢) تمثل السلسلة الزمنية التالية صادرات سورية من زيت الزيتون بالمليون دولار:

السنة	1995	1996	1997	1998	1999	2000	2001
الكمية	5	7	10	12	15	17	20

أوجد مايلي :

ـ معادلة خط الاتجاه العام للسلسلة الزمنية باستخدام طريقة المربعات الصغرى.

٣) يمثل الجدول التالي السلسلة الزمنية لواردات دولة عربية ما من النفط بالمليون وحدة نقدية.

السنة	1995	1996	1997	1998	1999	2000	2001	2002	2003
كمية الواردات	15	20	22	27	29	35	42	50	60

أوجد ما يلي :

ـــ خط الاتجاه العام للسلسلة الزمنية حسب طريقة المربعات الصغرى .

٤) يمثل الجدول التالي مشتريات إحدى الشركات خلال إحدى عشرة سنة ؟

السنوات	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
بالاف الوحدات النقدية	50	100	120	150	100	150	200	250	300	350	400

أوجد ما يلي :

ـــ خط الاتجاه العام للسلسلة الزمنية باستخدام منتصف السلسلة .

ـــ معادلة خط الاتجاه العام للسلسلة الزمنية باستخدام طريقة المربعات الصغرى .

٥) يمثل الجدول التالي ثلاثة فصول لأربع سنوات :

السنوات	الفصل الأول	الفصل الثاني	الفصل الثالث
2000	20	25	15
2001	25	10	30
2002	15	5	12
2003	10	25	30

أوجد ما يلي :

١) معادلة خط الاتجاه العام للسلسلة الزمنية بطريقة المربعات الصغرى .

٢) التغيرات الموسمية .

٦) يوضح الجدول التالي استهلاك الطاقة الكهربائية الشهرية في إحدى المدن بملايين الكيلووات ساعة .

الشهر	1	2	3	4	5	6	7	8	9	10	11	12
1995	32	28	28	25	23	22	23	25	27	30	33	35
1996	34	31	30	27	25	24	25	26	29	32	34	37
1997	37	33	32	29	27	25	26	28	31	35	37	40
1998	39	35	34	31	29	27	28	31	33	37	39	42
1999	42	38	37	33	31	30	31	33	36	40	42	45
2000	45	41	40	36	34	32	34	36	39	43	46	49
2001	49	44	43	39	37	35	36	39	42	46	49	52
2002	53	48	46	43	40	38	39	42	45	49	53	59

أوجد ما يلي :

١) خط الاتجاه العام بالرسم الحر

٢) معادلة خط الاتجاه العام للسلسلة الزمنية بطريقة المربعات الصغرى.

٣) القيم الاتجاهية .

٤) التغيرات الموسمية.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل الثامن

الأرقام القياسية

Index Numbers

٨ ــ ١ : تعريف الأرقام القياسية

٨ ــ ٢ : استخدامات الأرقام القياسية

٨ ــ ٣ : الصيغ البسيطة للأرقام القياسية

٨ ــ ٤ : الأرقام القياسية التجميعية البسيطة .

٨ ــ ٥ : الصيغ المرجحة للأرقام القياسية

ــ الأرقام القياسية المرجحة بأوزان فترة الأساس

ــ الأرقام القياسية التجميعية المرجحة بأوزان فترة المقارنة

ــ الرقم القياسي الأمثل

٨ ــ ٦ : اختبار الأرقام القياسية:

ــ اختبار الانعكاس في الزمن

ــ اختبار الانعكاس المعاملي

٨ ــ ٧ : بعض الأرقام القياسية ا لهامة

ــ الرقم القياسي لنفقة المعيشة

ــ الرقم القياسي لكمية الصادرات والواردات

ــ القوة الشرائية للنقود.

ــ تمارين

وقل رب زدني علما
Damascus University

٨ ــ ١ : تعريف :

الرقم القياسي هو مؤشر إحصائي يقيس مقدار التغير النسبي الذي طرأ على قيمة ظاهرة معينة (السعر ــ الكمية ــ القيمة ــ الأجر ــ ..... إلخ) بالنسبة لأساس معين .

والأساس قد يكون فترة زمنية معينة أو مكان معين حيث تؤخذ قيمة الظاهرة حينئذ كأساس لحساب الرقم القياسي .

فعلى سبيل المثال يمكن مقارنة سعر سلعة معينة أو عدة سلع في بلد ما بسعرها في دولة أخرى تكون هي بمثابة الأساس، مثال ذلك مقارنة أسعار العملات في أي بلد بأسعارها في بورصة نيويورك أو بورصة طوكيو.

كذلك فإنه يمكن مقارنة سعر سلعة ما في فترة زمنية معينة (هي فترة المقارنة) بسعرها في فترة زمنية أخرى(هي فترة الأساس)، والتي هي غالباً ما تكون فترة سابقة لفترة المقارنة.

٨ ــ ٢ : استخدام الأرقام القياسية :

١ ــ دراسة التغيرات التي طرأت على الظروف الاقتصادية من فترة إلى أخرى.

٢ ــ التعرف على الاتجاه العام لقيم الظواهر الاقتصادية والاجتماعية وذلك عن طريق تركيب سلسلة تتابعية من الأرقام القياسية تغطي فترة الدراسة المطلوبة .

٣ ــ تفيد بعض الأرقام القياسية في الوقوف على الوضع الاقتصادي للدولة مثال ذلك الأرقام القياسية للصادرات والواردات واستخداماتها في حساب معدلات التبادل الدولي.

والفكرة العامة للرقم القياسي مبنية على أساس إيجاد نسبة مئوية لقيمة متغير ما أو ظاهرة معينة في فترة أو مكان معين هي فترة أو مكان المقارنة إلى قيمة نفس المتغير أو الظاهرة في فترة أخرى أو مكان آخر هي فترة أو مكان الأساس، وفي الحقيقة، فإن الأسعار وأيضاً الكميات تشكل المجال الأكبر والأكثر أهمية بالنسبة لاستخدامات الأرقام القياسية، غير أن هناك ظواهر اقتصادية واجتماعية يحتاج الأمر لدراستها إلى تركيب الأرقام القياسية مثال ذلك الأرقام القياسية المستخدمة في النشاط الصناعي والتجاري والرقم القياسي لنفقة المعيشة.

وتتوقف أهمية الرقم القياسي المحسوب على الأسس التي تم بموجبها اختيار فترة الأساس، فإذا ما تم تحديد فترة الأساس بطريقة عفوية كانت النتيجة الحصول على أرقام قياسية ليست ذات مدلول اقتصادي أو تخطيطي والقواعد التي تجب مراعاتها عند اختيار فترة الأساس هي:

١ ــ أن تتميز فترة الأساس بالاستقرار والبعد عن الاضطرابات والتذبذبات التي قد تتعرض لها الظاهرة كالحروب والأزمات الاقتصادية كالكساد والتضخم .

٢ ــ بالنسبة لمكان الأساس فيجب أن يكون ذلك المكان ذا أهمية تميزه عن باقي الأماكن الأخرى، مثال ذلك عند اختيار أسعار العملات في البورصات العالمية مثل بورصة طوكيو ونيويورك كأساس.

٣ ــ يجب أن يكون احتمال تعرض فترة الأساس للتغير أقل ما يمكن، فمثلاً لا يجوز اتخاذ الشهر كأساس لقصره ولاحتمال تعرض قيمة الظاهرة المقاسة فيه لتغيرات عرضية. وإنما تفضل فترة الأساس سنة في أغلب الأحيان، وقد يفضل أخذ متوسط عدد السنوات كأساس.

٤ ــ فيما عدا حالات خاصة فإن فترة الأساس تكون سابقة لفترة المقارنة.

٥ ــ يجب أن يكون البعد الزمني بين فترتي الأساس والمقارنة محدوداً حتى يكون الرقم القياسي المستنتج منطقياً من وجه النظر الاقتصادية.

المصطلحات المستخدمة في تركيب الأرقام القياسية :

P_n ــ السعر في فترة المقارنة .

P_o ــ السعر في فترة الأساس .

q_n ــ الكمية في فترة المقارنة .

q_o ــ الكمية في فترة الأساس .

يمكن تمييز صيغتين أساسيتين من صيغ الأرقام القياسية هما :

٨ ــ ٣ ــ الصيغ البسيطة للأرقام القياسية :

١ ــ المنسوب :

ويمكن الحصول عليه بنسبة قيمة المتغير في فترة المقارنة إلى قيمة نفس المتغير في فترة الأساس، ويمكن أن يحسب ذلك لعدد من المتغيرات من خلال الصيغة التالية :

منسوب السعر = $\frac{P_n}{P_o} \times 100$

مثال :

الأرقام التالية توضح قيمة الصادرات من النفط في الأعوام من ٩٦ حتى ٢٠٠٠ بالمليون وحدة نقدية :

السنـــــــة	1996	1997	1998	1999	2000
قيمة الصادرات من النفط	2.493	2.659	2.557	2.629	4.781

المصدر : فرضي .

والمطلوب :

حساب الأرقام القياسية (المناسيب) لقيمة الصادرات من النفط بفرض سنة 1996 (سنة أساس) .

الحل :

السنة	قيمة الصادرات من النفط	الرقم القياسي باعتبار ١٩٩٦=١٠٠
1996	2.493	100
1997	2.659	106.6
1998	2.557	102.5
1999	2.629	105.4
2000	4.781	191.8

التفسير :

يتضح من بيانات الجدول السابق أن قيمة الصادرات من النفط ارتفعت إلى ١٠٦,٦% في سنة ١٩٩٧ بالمقارنة بالقيمة في سنة ١٩٩٦ كما أن قيمة الصادرات من النفط قفزت إلى ١٩١,٨% في سنة ٢٠٠٠ بالمقارنة بنفس سنة الأساس وهي ١٩٩٦ أي أن مقدار الزيادة في قيمة الصادرات تعادل ٩١,٨% .

مثال :

البيان التالي يوضح أسعار مجموعة من السلع في عامي ١٩٩٦ ــ ٢٠٠٠ (السعر للكيلوجرام بالليرات) .

السلعة	أسعار سنة 1996	أسعار سنة 2000
أسماك	300	700
دواجن	350	750
لحوم	400	700

والمطلوب :

استنتاج مناسيب السعر في عام ٢٠٠٠ باعتبار سنة ١٩٩٦ سنة أساس.

الحل :

السلعة	منسوب السعر باعتبار ١٩٩٦ أساس
أسماك	$\frac{700}{300}\times 100 = 233\%$
دواجن	$\frac{750}{300}\times 100 = 214\%$
لحوم	$\frac{750}{400}\times 100 = 175\%$

وإذا كانت مجموعة السلع بالقائمة تنتمي إلى مجموعة سلعية فإنه يمكن استنتاج رقم قياسي موحد باستخدام المناسيب المستنتجة وذلك بإيجاد الوسط الحسابي لتلك المناسيب وتكون صيغة الرقم القياسي في تلك الحالة هي :

$$I_{\bar{p}} = \frac{\sum \frac{P_n}{P_o}}{n} \times 100 = \text{متوسط المناسيب}$$

حيث :

P_n ــ سعر سنة المقارنة

P_o ــ سعر سنة الأساس

n ــ عدد سلع القائمة

ففي المثال السابق يمكن الحصول على متوسط المناسيب كالتالي :

الوسط الحسابي لمناسيب أسعار السلع الداخلة في المجموعة :

$$I_{\bar{p}} = \frac{\frac{700}{400} + \frac{750}{350} + \frac{700}{300}}{3} * 100$$

$$\frac{175 + 214 + 233}{3} * 100 = \frac{622}{3} * 100 = 207.3\%$$

٨ ــ ٤ ــ الأرقام القياسية التجميعية البسيطة :

هي عبارة عن النسبة بين مجموع قيم عدة متغيرات في فترة معينة (المقارنة) إلى مجموع قيم نفس المتغيرات في فترة الأساس . وبالتالي يمكن التعبير عن صيغة الرقم التجميعي البسيط كالتالي :

ـ في حالة إذا كان المتغير المقيس يعبر عن السعر .

$$I_{P_{(S)}} = \frac{\sum P_n}{\sum P_O} \times 100$$ الرقم القياسي التجميعي البسيط =

ـ في حالة ما إذا كان المتغير المقيس يعبر عن الكمية .

$$I_{q_{(S)}} \frac{\sum q_n}{\sum qo} \times 100$$ الرقم القياسي التجميعي البسيط =

حيث :

P_o ، P_n أسعار المقارنة والأساس .

q_o ،q_n كميات المقارنة والأساس .

مثال : في المثال الخاص بأسعار السلع الغذائية أحسب الرقم القياسي التجميعي البسيط لأسعار تلك المجموعة .

الحل :

$$I_{P_{(S)}} \frac{\sum P_n}{\sum P_O} \times 100$$ الرقم القياسي التجميعي البسيط للأسعار =

$$\frac{700+750+700}{400+350+300} \times 100 = \frac{2150}{1050} \times 100$$

= ٢٠٤,٨%

مثال :

البيانات التالية توضح الكمية المصدرة من مجموعة السلع في عامي ١٩٩٦، ٢٠٠٠.

السلعة	الوحدة	الكمية المصدرة في 1996	الكمية المصدرة في 2000
أسماك	طن	70	200
اسمنت	طن	20	80
بترول خام	ألف برميل	400	900

والمطلوب :

حساب الرقم القياسي التجميعي لكميات تلك المجموعة من السلع .

الحل :

الرقم القياسي التجميعي البسيط للكميات = $I_{q_{(n)}} \frac{\sum q_n}{\sum qo} \times 100$

$$= \frac{900+80+200}{900+20+70} \times 100 = \frac{1180}{490} = \%240.8$$

إن الأهمية النسبية للسلع الداخلة في كل مجموعة تُظهر بوضوح أثر (السلعة) المتغير ذي القيمة الكبيرة نسبياً على قيمة الرقم القياسي الناتج، فمثلاً المثال الخاص بالكميات المصدرة يظهر جلياً أثر كمية البترول المصدرة حتى ولو تساوت وحدات القياس على الرقم القياسي نظراً لكبر الكمية المصدرة من النفط بالقياس بباقي السلع، ولذا يفضل استخدام صيغة أفضل لتركيب الرقم القياسي لمجموعة من (السلع) المتغيرات ذات أصناف غير مختلفة بحيث يمكن اعتبار المجموعة بمثابة سلع متجانسة، مثال ذلك مجموعة سلع غذائية ذات صفة واحدة مثل السلع البروتينية الحيوانية،

ومجموعة سلع مواد البناء ومجموعة السلع ذات المنشأ الواحد مثل البترول ومنتجاته.

٨ ــ ٥ ــ الصيغ المرجحة للأرقام القياسية :

لمعالجة المشكلة الناجمة عن مساواة الأهمية النسبية للسلع الداخلة في تركيب الرقم القياسي التجميعي البسيط فإنه يمكن التغلب على تلك المشكلة بمحاولة إعطاء أوزان بحيث يستطيع ذلك الوزن المخصص لكل متغير من إظهار أهميته الاقتصادية .

فمثلاً في حالة تركيب رقم قياسي خاص بالأسعار فإنه يمكن استخدام الكميات المقابلة كأوزان مناسبة، كذلك فإنه في حالة تركيب رقم قياسي للكميات فإنه يمكن استخدام الأسعار المقابلة كأوزان مناسبة .

وفي كل تلك الحالات فإنه يمكن اتخاذ قيمة المتغير في فترة الأساس للترجيح (الوزن) كما يمكن استخدام قيمة المتغير في فترة المقارنة للترجيح، وبناءً على ذلك يمكن كتابة الصيغ التالية للتعبير عن الأرقام القياسية التجميعية المرجحة .

الأرقام القياسية التجميعية المرجحة بأوزان فترة الأساس :

ويسمى الرقم القياسي الناتج في هذه الحالة باسم رقم لاسبير وتكون صيغة الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس في هذه الحالة :

$$I_{P_{(0)}} \frac{\sum p_n q_n}{\sum p_o q_o} \times 100$$

واضح من هذه الصيغة أنه يفترض أن أذواق المستهلكين لم تتغير وأنهم سوف يستمرون في استهلاك نفس كميات السلع حتى بعد تغير أسعارها بالزيادة أو النقصان .

وتكون صيغة الرقم القياسي التجميعي للكميات المرجح بأسعار سنة الأساس هي :

$$I_{q_{(a)}} \frac{\sum q_n p_a}{\sum q_a p_a} \times 100$$

وفي هذه الصيغة يفترض ثبات الأسعار في فترتي الأساس والمقارنة حتى بعد تغير كمياتها بين الفترتين سواء بالزيادة أو النقصان .

الأرقام القياسية التجميعية المرجحة بأوزان فترة المقارنة :

ويسمى الرقم القياسي الناتج في هذه الحالة باسم رقم باش، وعليه يكون الرقم القياسي للأسعار المرجح بكميات سنة المقارنة ويُعطى بالصيغة التالية :

$$I_{P_{(n)}} \frac{\sum p_n q_n}{\sum p_a q_n} \times 100$$

واضح أن هذه الصيغة تقيس التغير في النفقات اللازمة للحصول على كميات السلع في فترة المقارنة بأسعار فترة المقارنة وأسعار فترة الأساس .

وفي هذه الصيغة يفترض أنه تم شراء نفس كميات فترة المقارنة في فترة الأساس أو بمعنى آخر فإن الفرض هنا أن المستهلك قد اشترى في فترة الأساس نفس الكمية التي أصبح يستهلكها في فترة المقارنة وبعد تغير

الأسعار، وهذا الفرض قد يكون غير مقبول. كذلك فإن صيغة باش للرقم القياسي للكميات المرجح بأسعار سنة المقارنة هي :

$$I_{q_{(P)}} \frac{\sum p_n q_n}{\sum q_n p_n} \times 100$$

واضح هنا أيضاً أنه يفترض أن المستهلك سيقيّم ما يستهلكه في فترة الأساس بأسعار المقارنة وهو فرض لا يجوز في كل الحالات .

ويتضح مما سبق أن كلٍ من رقم لاسبير ورقم باش وإن اختلفت قيمتهما إلا أنها تشير إلى الاتجاه نحو التغير وأن كلا الرقمين يعتمد على مقارنة القيم مع اختلاف الغرض المستخدم لحساب القيمة، وأن هذا الاختلاف ناتج عن اختلاف الأوزان المستخدمة .

- **الرقم القياسي الأمثل :**

سبق أن أوضحنا أن رقم لاسبير يبنى على الترجيح بأوزان فترة الأساس مما يجعل صيغة الرقم متحيزة إلى الأعلى، بينما يبنى رقم باش على الترجيح بأوزان فترة المقارنة مما يجعل صيغة الرقم متحيزة إلى أسفل. لأجل ذلك اقترحت عدة صيغ للمعالجة، ومن أهم الصيغ التي اقترحت في هذا الشأن هي صيغة فيشر Fisher حيث اقترح صيغة تعتمد على الوسط الهندسي لرقمي لابيز مكاش وسمى الرقم الناتج بالرقم القياسي الأمثل، وبناءً عليه تكون صيغة الرقم القياسي الأمثل على الصورة التالية :

$I_{P(F)}$ = $\sqrt{}$ رقم لاسبير للأسعار × رقم باش للأسعار

$$I_{p_{(F)}} = \sqrt{\frac{\sum p_n q_n}{\sum p_o q_n} \times \frac{\sum p_n q_o}{\sum p_o q_o}} \times 100$$

وبنفس الأسلوب تكون صيغة الرقم القياسي الأمثل للكميات على الصورة التالية :

الرقم القياسي الأمثل للكميات:

$$I_{q_{(I)}} = \sqrt{\frac{\sum q_n p_n}{\sum q_o p_n} \times \frac{\sum q_n p_o}{\sum q_o p_o}} \times 100$$

• **رقم مارشال ايجورث** : يعتمد هذا الرقم على تثقيل الأسعار أو الكميات لمجموع الكميات أو الأسعار في فترة الأساس والمقارنة. أي أخذ متوسط استهلاك الفرد في الكميات التي اشتراها في سنة الأساس والمقارنة والتثقيل بها الأسعار وقيم حسابه على النحو التالي :

الرقم القياسي للأسعار $$I_{P_{(Mc)}} = \frac{\sum p_n (q_o + q_n)}{\sum p_n (q_o + q_n)} \times 100$$

الرقم القياسي للأسعار $$I_{q_{(Mc)}} = \frac{\sum q_n (p_o + p_n)}{\sum q_n (p_o + p_n)} \times 100$$

مثال :

البيانات التالية توضح أسعار وكميات قائمة من السلع في عامي ١٩٩٥، ٢٠٠٠.

السلعة	السعر في 95	السعر في 2000	الكمية المستهلكة في 95	الكمية المستهلكة في 2000
A	2	3	10	8
B	5	6	15	15
C	10	12	3	3
D	12	20	4	5

وباعتبار سنة ١٩٩٥ سنة أساس .

المطلوب : حساب الأرقام القياسية التالية :

١) الرقم القياسي التجميعي للأسعار مرجحاً بكميات الأساس .

٢) الرقم القياسي التجميعي للأسعار مرجحاً بكميات المقارنة.

٣) الرقم القياسي التجميعي للكميات مرجحاً بأسعار الأساس .

٤) الرقم القياسي التجميعي للكميات مرجحاً بأسعار المقارنة.

٥) الرقم القياسي الأمثل للأسعار .

٦) الرقم القياسي الأمثل للكميات .

الحل :

السلعة	Po	Pn	qo	qn	Po qo	Po qn	Pn qo	Pnqn
A	2	3	10	8	20	16	30	24
B	5	6	15	15	75	75	90	90
c	10	12	3	30	30	36	36	36
D	12	20	4	5	48	60	80	100
المجموع					173	181	236	250

١ ــ الرقم القياسي التجميعي للأسعار المرجح بكميات الأساس .

$$I_{P_{(L)}} = \frac{\sum p_n q_o}{\sum q_o p_o} \times 100 = \frac{236}{173} \times 100 = \%136.4$$

٢ ــ الرقم القياسي التجميعي للأسعار المرجح بكميات المقارنة .

$$I_{P_{(P)}} = \frac{\sum p_n q_n}{\sum p_o q_n} \times 100 = \frac{250}{181} \times 100 = \%138.4$$

٣ ــ الرقم القياسي التجميعي للكميات المرجح بأسعار الأساس :

$$I_{q_{(L)}} = \frac{\sum p_o q_n}{\sum p_o q_o} \times 100 = \frac{181}{173} \times 100 = \%104.6$$

٤ ــ الرقم القياسي التجميعي للكميات المرجح بأسعار المقارنة:

$$I_{q_{(P)}} = \frac{\sum q_n p_n}{\sum p_n q_o} \times 100 = \frac{250}{236} \times 100 = \%105.9\%$$

٥ ــ الرقم القياسي الأمثل للأسعار :

$$I_{q_{(F)}} = \sqrt{\frac{\sum p_n q_n}{\sum p_o q_n} \times \frac{\sum p_n q_o}{\sum p_o q_o}} \times 100$$

$$= \sqrt{\frac{250}{181} \times \frac{236}{173}} \times 100 = 137.3\%$$

٦ ــ الرقم القياس الأمثل للكميات :

$$I_{q_{(F)}} = \sqrt{\frac{\sum q_n p_n}{\sum q_o p_n} \times \frac{\sum q_n p_o}{\sum q_o p_o}} \times 100$$

$$= \sqrt{\frac{250}{236} \times \frac{181}{172}} \times 100 = 105.3\%$$

٨ ــ ٦ : الاختبارات المطبقة على الأرقام القياسية :

بعد أن تعرفنا على الصيغ المختلفة للأرقام القياسية فإن هناك بعض الاختبارات التي تجري على تلك الصيغ للتعرف على مدى اتساقها. ويعتبر الرقم القياسي مثالياً إذا اجتاز تلك الاختبارات، وأهم الاختبارات ما يلي :

ــ اختبار الانعكاس في الزمن :

إذا فرضنا أن سعر سلعة ما قد زاد في عام ٢٠٠٠ إلى الضعف أي ٢٠٠% مما كان عليه سنة ١٩٩٠ فهذا يعني أن سعر هذه السلعة في سنة ١٩٩٠ كان نصف سعره في عام ٢٠٠٠ أي أن الرقم القياسي في عام ١٩٩٠ يجب أن يكون ٥٠% باعتبار أن عام ٢٠٠٠ سنة أساس وهذا معناه أن حاصل ضرب الرقمين بدون الضرب في ١٠٠ يساوي الواحد الصحيح أي أن:

القاعدة : الرقم القياسي المطلوب * بديله الزمني = ١

$$\frac{p_o}{p_n} \times \frac{p_n}{p_o} = 1$$

وفي مثالنا السابق لو فرضنا أن سعر السلعة كان في سنة ١٩٩٠ وحدتين نقديتين وفي سنة ٢٠٠٠ أصبح أربع وحدات نقدية فإن :

$$\frac{2}{4}\times\frac{4}{2}=1$$

وتطبيقاً على ذلك نجد أن الرقم التجميعي البسيط لكل من الأسعار والكميات يقبلان الانعكاس في الزمن حيث أن :

$$\frac{\sum p_o}{\sum p_n}\times\frac{\sum p_n}{\sum p_o}=1$$

$$\frac{\sum q_n}{\sum q_o}\times\frac{\sum q_o}{\sum q_n}=1$$

بينما لا يقبل كل من الرقم القياسي التجميعي للأسعار المرجح سواء بكميات الأساس أو المقارنة وأيضاً الرقم القياسي التجميعي الكميات المرجح بأسعار الأساس أو أسعار المقارنة الانعكاس في الزمن حيث أن حاصل جداء الرقم القياسي المطلوب * بديله الزمني لا يساوي الواحد الصحيح .

$$\frac{\sum p_n q_o}{\sum p_o q_o}\times\frac{\sum p_o q_n}{\sum p_n q_n}\neq 1$$

وكذلك فإن :

$$\frac{\sum p_n q_n}{\sum p_o q_n}\times\frac{\sum p_o q_o}{\sum p_n q_o}\neq 1$$

ويمكن الوصول لنفس النتيجة في حالة الأرقام القياسية المرجحة للكميات.

بينما يقبل الرقم القياسي الأمثل للأسعار وللكميات الانعكاس في الزمن حيث أن صيغة الرقم الأمثل للأسعار هي :

$$\sqrt{\frac{\sum p_n q_n}{\sum p_o q_n} \times \frac{\sum p_n q_o}{\sum p_o q_o}}$$

وبديلها الزمني هو :

$$\sqrt{\frac{\sum p_o q_n}{\sum p_n q_n} \times \frac{\sum p_o q_o}{\sum p_n q_o}} \times 100$$

وحاصل ضربهما يساوي الواحد الصحيح .

ويمكن بنفس الطريقة إثبات أن صيغة الرقم الأمثل للكميات تقبل أيضاً انعكاساً في الزمن.

اختبار الانعكاس المعاملي :

إذا فرضنا أن سعر سلعة ما قد زاد من أربع وحدات نقدية في سنة ما إلى ستة وحدات نقدية في سنة المقارنة، وأن المبيعات أيضاً من هذه السلعة قد زادت خلال نفس الفترة من وحدتين إلى أربعة، معنى ذلك أن الرقم القياسي للسعر أصبح $(\frac{6}{4} \times 100) = 150\%$ والرقم القياسي للكمية يكون $(\frac{4}{4} \times 100) = 200\%$.

وهذا يعني أن القيمة قد زادت في سنة المقارنة إلى ثلاثة أضعاف أمثالها في سنة الأساس لأن القيمة = السعر × الكمية حيث أن :

$$\frac{6}{4}\times\frac{4}{2}\times 100 = \%300$$

ويعتبر الرقم القياسي متسقاً (أو مثالياً) إذا كان حاصل ضرب الرقم القياسي للأسعار لسنة ما في الرقم القياسي للكمية لنفس السنة يعطي منسوب القيمة للسنة ذاتها بالنسبة لسنة الأساس، وعليه تكون :

القاعدة : الرقم القياسي المطلوب * بديله المعاملي = الرقم القياسي للقيمة

ففي المثال السابق نجد أن :

$$\frac{q_n}{q_n}\times\frac{p_n}{p_o} = \frac{p_n q_n}{p_o q_o}$$

والنتيجة هي منسوب القيمة .

معنى ذلك أن الرقم القياسي الذي استبدلت فيه رموز الأسعار (p) بالكميات (q) مع الإبقاء على دليل الزمن هو ما يسمى باسم البديل المعاملي، فإذا كان نتيجة حاصل ضرب الرقمين يعطي منسوب القيمة دل ذلك على أن الرقم القياسي المختبر قد اجتاز اختبار الانعكاس في المعامل وحققه .

وتطبيقاً على ذلك فإن الرقم التجميعي البسيط للأسعار $\frac{\sum p_n}{\sum p_o}$

فإن بديله المعاملي $\frac{\sum q_n}{\sum q_o}$

وحاصل ضربهما لا يعطي منسوب القيمة أي أن :

$$\frac{\sum p_n}{\sum p_o} \times \frac{\sum q_n}{\sum q_o} \neq \frac{\sum P_n q_n}{\sum p_o q_o}$$

أي لا تكون النتيجة منسوب القيمة .

وعلى هذا الأساس فإنه يمكن استنتاج أن جميع الأرقام التجميعية سواء البسيطة أو المرجحة لا تقبل الانعكاس في المعامل .

بينما يجتاز الرقم القياسي الأمثل سواء للأسعار أو الكميات هذا الاختبار هو :

وبديله المعاملي هو $\sqrt{\frac{\sum p_n q_o}{\sum p_o q_o} \times \frac{\sum p_n q_n}{\sum p_o q_n}}$

$$\sqrt{\frac{\sum q_n p_o}{\sum q_o p_o} \times \frac{\sum q_n p_n}{\sum q_o p_n}}$$

وحاصل ضربهما هو :

$$\sqrt{\frac{\sum q_n p_o}{\sum q_o p_o} \times \frac{\sum q_n p_n}{\sum q_o p_n}} \times \sqrt{\frac{\sum p_n q_o}{\sum p_o q_o} \times \frac{\sum p_n q_n}{\sum p_o q_n}} =$$

$$\frac{\sum p_n q_n}{\sum p_o q_o}$$ وهو منسوب القيمة

ويمكن إثبات نفس النتيجة في حالة الرقم القياسي الأمثل للكميات حيث أن الرقم الأمثل للكميات هو :

$$\sqrt{\frac{\sum q_n p_n}{\sum q_o p_n} \times \frac{\sum q_n p_o}{\sum q_o p_o}}$$ وبديله المعاملي هو

$$\sqrt{\frac{\sum p_n q_n}{\sum p_o q_n} \times \frac{\sum p_n q_o}{\sum p_o q_o}}$$

وحاصل ضربهما هو

$$\sqrt{\frac{\sum p_n q_n}{\sum p_o q_n} \times \frac{\sum p_n q_o}{\sum p_o q_o}} \times \sqrt{\frac{\sum q_n p_n}{\sum q_o p_n} \times \frac{\sum q_n p_o}{\sum q_o p_o}}$$

$= \frac{\sum p_n q_n}{\sum p_o q_n}$ وهو منسوب القيمة

٨ ــ ٧ : بعض الأرقام القياسية الهامة :

بعد أن درسنا كيفية تركيب الأرقام القياسية نستعرض بعض الأرقام القياسية الهامة في المجال التجاري والاقتصادي مع توضح كيفية تركيبها ومدلولها ومن هذه الأرقام :

- الرقم القياسي لنفقة المعيشة .
- الرقم القياسي لكمية الصادرات والواردات

• الرقم القياسي القوة الشرائية للنقود .

الرقم القياسي لنفقة المعيشة :

يمكن تعريف نفقة المعيشة بأنها قيمة ما تنفقه الوحدة (فرد أو أسرة) من السلع والخدمات على ما تستهلكه من سلع وخدمات خلال فترة زمنية قياسية هي الشهر غالباً . وطالما أن الموضوع يخص استهلاك الوحدة سواء كان فرداً أو أسرة معيشية فإن الأمر هنا يتعلق بصيغة رقم قياسي لأسعار التجزئة التي يتعامل بها فعلاً الفرد أو الأسرة المعيشية مع إجراء عملية الترجيح المناسبة، والتي غالباً ما تبنى على أساس أوزان يجري تحديدها وتغييرها كل فترة زمنية مع تغير العادات والأنماط الاستهلاكية والإنفاقية والتي تعتبر أبحاث ميزانية الأسرة أساساً لتحديد اتجاهها .

كيفية تركيب الرقم القياسي لنفقات المعيشة :

١ ــ تبوب جميع السلع والخدمات الإنسانية في عدة مجموعات رئيسية غالباً ما يكون عددها تسع مجموعات هي :

الطعام والشراب والدخان ــ الملابس وملبوسات القدم ــ الإيجار والوقود والقوى المحركة ــ الأثاث والتأثيث ــ العناية الطبية والخدمات الصحية ــ النقل والمواصلات ــ خدمات الترفيه والتسلية والثقافة ــ السلع والخدمات المتنوعة ــ المدفوعات التحويلية .

٢ ــ يتم تركيب رقم قياسي لأسعار كل مجموعة من المجموعات التسع السابق ذكرها وهو عبارة عن المتوسط الحسابي لمناسيب أسعار السلع والخدمات الداخلة في المجموعة.

٣ ــ ترجح الأرقام القياسية للأسعار الناتجة بأوزان نسبية تحدد على أساس الأهمية النسبية لكل مجموعة وتقدر من أبحاث خاصة أهمها أبحاث ميزانية الأسرة وأبحاث الاستهلاك.

• **الرقم القياسي لكمية الصادرات والواردات :**

يركب الرقم القياسي لكمية الصادرات والواردات باستخدام صيغة لاسبير أي:

$$\frac{\sum p_n q_o}{\sum p_n q_o}$$

ويحسب لكل مجموعة سلعية على حدة ثم يحسب الوسط الحسابي لأرقام المجموعات المرجح بالقيمة ($\sum p_o q_o$) ويفيد حساب الرقم القياسي لكمية الصادرات والرقم القياسي لكمية الواردات في حساب ما يعرف اقتصادياً باسم معدل التبادل التجاري الإجمالي وصيغته :

$$\text{معدل التبادل التجاري الإجمالي} = \frac{\text{الرقم القياسي لكمية الواردات}}{\text{الرقم القياسي لكمية الصادرات}} \times ١٠٠$$

تدل زيادة هذا الرقم عن ١٠٠% على أنه بكمية معينة من الصادرات أمكن الحصول على كمية أكبر من الواردات بسبب انخفاض أسعار الواردات.

• **الرقم القياسي للقوة الشرائية للنقود :**

يدلنا الرقم القياسي للأسعار يدلنا على مقدار التغير النسبي بين فترتين زمنيتين لذا يمكن التعرف على مقدار التغير في قيمة النقود في الداخل حيث أن هذا التغير يعبر عن الاتجاه العكسي للتغير في أسعار جميع

السلع والخدمات، أي أن القوة الشرائية للنقود هي مقلوب الرقم القياسي للأسعار والصيغة العامة هي :

$$\text{القوة الشرائية للنقود} = \frac{1}{\text{الرقم القياسي لتكاليف المعيشة}} \times 100$$

وعند استنتاج رقم القوة الشرائية للنقود فإنه يجب حساب رقم قياسي لأسعار جميع السلع والخدمات، ولا يجب الاقتصار على مجموعة معينة من السلع أو الخدمات فقد يكون التغير في أسعار تلك المجموعة من السلع والخدمات راجعاً إلى ظروف خاصة بطلبها أو عرضها .

إلى جانب الأرقام القياسية السابق إيضاحها هناك عدة أرقام قياسية أخرى تقوم الجهات المعنية بتركيبها وحسابها نذكر منها الرقم القياسي للإنتاج الصناعي ويحسب باستخدام صيغة باش:

$$\frac{\sum q_n p_n}{\sum q_o p_n} \times 100$$

والرقم القياسي لأسعار الصادرات والواردات .. وغيرها من الأرقام الأخرى.

تمارين

تمرين رقم (١) :

البيانات التالية تبين أسعار الجملة وكذلك الكميات المستهلكة من عدة سلع غذائية رئيسية في بلد ما في السنوات ١٩٩٥. ٢٠٠٠ :

السلعة / السنة	الأسعار بالألف		الكميات المستهلكة	
	1995	2000	1995	2000
A	1000	2000	50	60
B	800	2000	100	130
C	1200	3600	10	9

المطلوب :

حساب كل من الأرقام القياسية الآتية باعتبار سنة ١٩٩٥ كأساس :

١ ــ حساب المناسيب البسيطة للكميات المستهلكة بالنسبة لكل من هذه السلع .

٢ ــ حساب الرقم القياسي التجميعي البسيط لأسعار هذه المجموعة من السلع .

٣ ــ حساب الرقم القياسي الأمثل (صيغة فيشر) لأسعار هذه المجموعة من السلع .

٤ ــ اشرح أهم عيوب الرقم التجميعي البسيط بوجه عام في حساب الأرقام القياسية.

تمرين رقم (٢) :

فيما يلي أسعار الملابس الأساسية وكمية المباع منها في سنتي ١٩٨٩ ، ١٩٩٥:

		رباط العنق	القميص	البدلة	الحذاء
سنة	سعر	1	2.5	20	8
1989	كمية	1000	1500	100	400
سنة	سعر	2	2	3	6
1995	كمية	900	200	120	500

أ ــ احسب الرقم القياسي التجميعي البسيط لأسعار هذه المجموعة لسنة ١٩٩٥ مقارناً بسنة ١٩٨٩ .

ب ــ احسب الرقم القياسي لكميات هذه المجموعة سنة ١٩٩٥ مقارناً بسنة ١٩٨٩ .

تمرين رقم (٣) :

يتضمن الجدول التالي افتراضية عن أسعار وكميات ثلاث سلع .

السلعة	1990		2000	
	الكمية	السعر	الكمية	السعر
A	135	10	400	14
B	30	8	35	10
C	100	20	80	30

والمطلوب :

تركيب الأرقام القياسية الآتية باعتبار سنة ١٩٩٠ كأساس .

١ ــ الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير).

٢ ــ الوسط الحسابي البسيط لمناسيب الأسعار .

تمرين رقم (٤) :

سجلت البيانات التالية عن أسعار ثلاث سلع والكميات المباعة منها في سنتين متتاليتين:

السلعة	الأسعار		الكميات	
	1999	2000	1999	2000
A	5	10	12	25
B	8	12	15	18
C	4	9	10	20

احسب الرقم القياسي التجميعي للكميات المرجح بأسعار سنة الأساس ثم اختبر إمكانية انعكاسه في المعامل والزمن (١٩٩٩ = ١٠٠) .

تمرين رقم (٥) :

فيما يلي بيان عن أسعار أربع سلع غذائية والكميات المباعة منها في عامي ١٩٨٢، ١٩٨٩:

السلعة / السنة	الأسعار		الكمية	
	1982	1989	1982	1989
A	7	14	50	60
B	4	11	30	35
C	2	5	70	100
D	5	6	20	25

والمطلوب : باستخدام سنة ١٩٨٢ سنة أساس :

١ ــ حساب الرقم القياسي الأمثل للأسعار (وفق صيغة فيشر) .

٢ ــ حساب الرقم القياسي الأمثل للكميات (وفق صيغة فيشر)

٣ ــ أذكر مع الشرح الشروط الواجب أخذها في الاعتبار عند اختيار سنة الأساس لتركيب الأرقام القياسية .

تمرين رقم (٦) :

ــ فيما يلي بيان بالأسعار والكميات المتداولة بسوق الأوراق المالية لثلاثة أسهم:

السهم	1988		1989	
	السعر	الكمية	السعر	الكمية
A	10	2	200	8
B	8	4	10	8
C	4	7	5	24

المطلوب :

حساب الرقم القياسي لاسبير والرقم القياسي باش (للأسعار).

تمرين رقم (٧) :

السلعة	فترة الأساس		فترة المقارنة	
	السعر	الكمية	السعر	الكيمة
A	7	35	6	40
B	9	30	8	35

المطلوب:

احسب الرقم القياسي الأمثل للأسعار السابقة.


وقل رب زدني علما
جامعة دمشق
Damascus University

المصطلحات الإحصائية

إنكليزي ــ عربي

Statistical Terminology

A

Absolute error	الخطأ المطلق
Absolute value	القيمة المطلقة
Accuracy	دقة
Adjustment data	تعديل البيانات
Aggregate value	القيمة التجميعية
Allowance value	حدود السماح
Analysis of covariance	تحليل تمام التباين
Analysis of variance	تحليل التباين
Applied statistics	الإحصاء التطبيقي
Arbitrary constant	ثابت اختياري
Area sampling	المعاينة المساحية
Arithmetic mean	الوسط الحسابي
Array	ترتيب
Association	اقتران
Assupption	فرض
Asymmetry	عدم التماثل

Asymmetrical	غير متماثل
Asymptotic	تقريبي
Attribute	صفة
Automatic	تلقائي، آلي
Average	متوسط
Axis	محور
Axis of abscissas	المحور السيني أو الأفقي
Axis of ordinates	المحور العمودي

B

Base	أساس أو قاعدة
Base Period	فترة الأساس
Basic statistics	إحصاءات أساسية
Battery of test	الاختبارات
Bell-shaped curve	نى جرسي
Between groups Bias	بين المجموعات
Binominal distribution	ثنائي الحدين
Bivariate table	ذو متغيرين
Breakdown	م

Business Cycles	دورات اقتصادية
Business indicators	مؤشرات اقتصادية

C

Calculation	حساب
Calcutus of probability	حساب الاحتمالات
Census	تعداد
Census enumerator	عداد
Census schedule	استمارة التعداد
Causation	سببية
Central tendency	النزعة المركزية
Chain indexes	الأرقام القياسية المتسلسلة
Chart	خريطة
Chi-Squared test	اختبار كاي مربع
Classification	تصنيف
Cluster sampling	المعاينة العنقودية
Code	رمز
Coefficient	معامل
Coefficient of correlation	معامل الارتباط

Coefficient of rank correlation	معامل الارتباط الجزئي
Column	عمود
Compute	يحسب
Confidence intervals	فترات الثقة
Confidence limits	حدود الثقة
consistency	اتساق
constant	ثابت
Contingency	توافق
Contingency table	جدول التوافق
Continuous	متصل، مستمر
Control	مراقبة
Coordinate	إحداثي
Coordinate axis	محور الإحداثيات
Corrected rate	معدل مصحح
Correction factor	عامل التصحيح
Correlation	ارتباط
Correlation table	جدول ارتباطي
Correlation ratio	نسبة الارتباط

Cost of living	نفقة المعيشة
Covariable	متغاير
Covariance	تمام التباين
Critical value	قيمة حرجة
Curves	منحنيات
Cycle	دورة
Cyclical variations	تغيرات دورية

D

Data	بيانات
Deduction	استنباط
Degrees of freedom	درجات الحرية
Dependent variable	متغير تابع
Derivative function	دالة مشتقة
Design of experiments	تصميم التجارب
Determinant	محددة
Deviation	انحراف
Diagram (figure)	شكل
Differential	تفاضلي

Discrete	متقطع ، منفصل
Dispersion	تشتت
Distribution	توزيع
Distribution function	دالة التوزيع
Double sampling	معاينة مزدوجة
Downward bias	تحيز نحو الأدنى
Downward trend	اتجاه هابط

E

Econometric s	اقتصاد قياسي
Effect	تأثير
Exponential equation	معادلة أسية
Extrapolation	استكمال خارجي
Efficiency of estimates	فعالية التقدير
Elimination	حذف

F

Factor	عامل
Factor analysis	تحليل عوامل
Factorial analysis	تحيلي عاملي

Family budget	ميزانية الأسرة
Fluctuations	تقلبات
Forecasting	تنبؤ
Forms	استمارات
Frame	إطار
Frequency	تكرار
Frequency curve	منحنى التكرار
Frequency density	كثافة التكرار
Frequency distribution	توزيع التكرار
Frequency polygon	مضلع التكرار
Frequency table	جدول التكرار
Function	دالة
Function relationship	علاقة دالة
F table	جدول " ف "

G

Gaussian curve	منحنى غوس
Geometric mean	وسط هندسي
Graduation	تدريج

Graph	رسم بياني
Graphic analysis	تحيلي بياني
Graphic method	طريقة بيانية
Graphic presentation	عرض بياني
Gross national product	الإنتاج القومي الاجمالي
Grouping	تجميع / تصنيف
Grouping error	خطأ التجميع
Growth curve	منحنى النمو

H

Harmonic equation	دالة توافقية
Harmonic mean	وسط توافقي
Head of household	رب الأسرة
Histogram	مدرج تكرار
Homogeneous	متجانس
Household	أسرة
Houseing census	تعداد المساكن
Hypothesis	افتراض / فرضية

I

Ideal number	الرقم الأمثل
Identity	متطابقة
Independent variable	متغير مستقل
Index	دليل أو مقياس
Index number	الرقم القياسي
Infinite	لانهائي
Integration	تكامل
Interclass correlation	ارتباط بين الفترات
Interpolation	استكمال داخلي
Interval	فترة، مجال
Interview	مقابلة / استجواب
Inverse correlation	ارتباط عكسي
Inverse functtion	دالة عكسية

J

Joint distribution	توزيع مشترك
Join regression	انحدار مشترك
J - curve	منحنى راني

K

Kurtosis	تفرطح

L

Labels	علامات مميزة
Labor force	قوة العمل
Large sample	عينات كبيرة
Law of large numbers	قانون الأعداد الكبيرة
Least squares	المربعات الصغرى
Least squares method	طريقة المربعات الصغرى
Level of living	مستوى المعيشة
Living conditions	ظروف معيشية
Logarithmic	لوغاريتم
Logistic curve	منحنى لوجستي
Long cycles	دورات طويلة
Lorenz curve	منحنى لورنز

M

Master sample	عينة رئيسة
Matrix	مصفوفة
Maximum variation	أقصى اختلاف

Mean	وسط
Mean absolute error	وسط الخطأ المطلق
Median	الوسيط
Model	نموذج
Moment	عزم
Moving average	متوسط متحرك
Multi-phase sampling	معاينة متعددة الأوجه
Multiple correlation	ارتباط متعدد
Multiple factor analysis	تحليل عاملي متعدد
Multi-Stage sampling	معاينة متعددة المراحل
Multivariate analysis	تحليل المتغيرات المتعددة

N

National accounts	حسابات قومية
National income	الدخل القومي
Negative correlation	ارتباط سالب
Non-Linear correlation	ارتباط غير خطي
Non-Statistics	إحصاء لامعلمي

O

Obsevation	مشاهدة
Orthogonal	متعامد

P

Parabola	قطع مكافئ
Parameter	معلمة، ثابت
Partial correlation	ارتباط جزئي
Partial regression	انحدار جزئي
Percentage	نسبة مئوية
Periodicity	دوية
Periodic movements	تقلبات دورية
Poulation (statistical)	مجتمع (احصائي)
Positive correlation	ارتباط موجب
Predication	تنبؤ
Price change	تغيرات الأسعار
Price index	الرقم القياسي للأسعار
Probability	احتمال
Probability curve	منحنى الاحتمال

Probability density	كثافة الاحتمال
Proportional allocation of the sample	التوزيع التناسبي للعينة
Purposive sample	عينة عمدية
Purposive selection	اختيار عمدي

Q

Qualitative analysis	تحليل وصفي
Quality control	مراقبة الجودة
Quantitative analysis	تحليل كمي
Quantity index	الرقم القياسي للكميات
Questionnaire	استبيان / استمارة
Quota sampling	عينة الحصص

R

Random distribution	توزيع عشوائي
Random error	خطأ عشوائي
Random numbers	أرقام عشوائية
Random sample	عينة عشوائية
Random selection	اختيار عشوائي
Random variable	متغير عشوائي

Range	مدى
Rank correlation	ارتباط الرتب
Rank correlation coefficient	معامل ارتباط الرتب
Rate	معدل
Ratio	نسبة
Ratio estimate	تقدير نسبي
Ratio test	اختيار النسب
Regression	انحدار
Regression coefficient	معامل الانحدار
Regression equation	معادلة الانحدار
Regression estimate	تقدير الانحدار
Regression line	خط الانحدار
Representative sample	عينة ممثلة
Root-mean square deviation	جذر متوسط مربع الانحراف
Rounding error	خطأ التقريب
Row	صف

S

sample	عينة

Sample census	مسح بالعينة
Sample design	تصميم العينة
Sampling	معاينة
Sampling error	خطأ المعاينة
Sampling method	طريقة المعاينة
Sampling unit	وحدة المعاينة
Scale	مقياس الرسم
Scatter	انتشار
Scatter diagram	شكل الانتشار
Seasonal index	الرقم القياسي الموسمي
Seasonal variation	تغيرات موسمية
Secondary trend	اتجاه ثانوي
Second degree curve	منحنى من الدرجة الثانية
Secular trend	اتجاه عام
Sensitivity	حساسية
Sequential analysis	تحليل تتابعي
Significance	معنوية
Simple correlation	ارتباط بسيط
Size of sample	حجم العينة
Small samples	عينات صغيرة
Source of error	مصدر الخطأ
Supurious correlation	ارتباط وهمي

Square	مربع
Standard normal distribution	توزيع طبيعي معياري
Standard deviation	انحراف معياري
Standard error	خطأ معياري
Standard error of estimate	خطأ معياري للتقدير
Standard error of the difference between means	خطأ معياري للفرق بين الوسطين
Statistical induction	استنتاج إحصائي
Statistical inference	استدلال إحصائي
Statistical method	الطريقة الإحصائية
Statistical probability	احتمال إحصائي
Statistical processing	التجهيز الإحصائي
Statistics	إحصاء
Stratified sample	عينة طبقية
Stratum	طبقة
survey	مسح (استقصاء)
Symmetrical	متماثل
System	نظام
Systematic	نظامي
Systematic sampling	معاينة منتظمة

T

" t ' table	جدول " ت " ستودنت
table	جدول
Tabulation	جدول أو تبويب

Test of goodness of fit	اختبار جودة التوفيق
Test of homogeneity	اختبار التجانس
Test of significance	اختبار المعنوية
Third degree curve	منحنى من الدرجة الثالثة
Time series	سلسلة زمنية
Time seies curve	منحنيات السلاسل الزمنية
Trend analysis	تحليل الاتجاه
Trend line	خط الاتجاه
Trial census	تعداد تجريبي
True mean	وسط حقيقي
Type	نوع ، نمط
Typical value	قيمة نمطية

U

U-curve	منحنى ذو شعبتين
Unbiased	غير متحيز
Unit of measurement	وحدة القياس
Universe (statistical	مجتمع (إحصائي)
Upward trend	اتجاه صاعد

V

Variable	متغير
Variance	تباين
Variance analysis	تحليل التباين
Variance rtio	نسبة التباين

Variation	اختلاف
Variation coefficient	معامل اختلاف
Vital Statistics	إحصاء حيوي
W	
Weights	تثقيلات، أوزان
Weight avarage	متوسط مثقل أو مرجح
X	
X-axis	محور السينات
Y	
Y-axis	محور العينات
Z	
Z-test	اختبار (ن)

Damascus University

الجداول
الإحصائية



وقل رب زدني علما
Damascus University

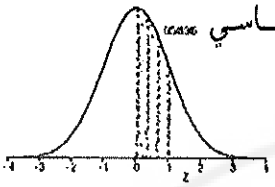
جدول رقم (1)

جدول الأرقام العشوائية

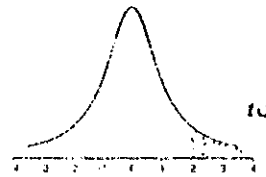
	1	2	3	4	5	6	7	8	9	10
1	000134	134485	679737	337743	214524	595648	560378	263838	146505	730892
2	016764	810654	967096	217831	815516	455968	047221	979812	313182	361437
3	095476	331699	887056	228844	939443	995951	902647	476533	147775	179660
4	934443	462342	882060	605488	430315	493912	830909	566670	471835	457463
5	805361	792746	257553	685939	789415	738990	863570	833689	979355	182857
6	670087	093261	194104	742389	676938	373792	946204	211184	419402	857067
7	760907	657684	262992	798609	617247	724031	275524	397985	425289	133406
8	113400	210456	873954	826175	155879	503846	440498	748132	161108	675694
9	175026	307058	244304	271874	484860	980719	062302	516458	138515	461724
10	878240	382272	537974	984299	607545	589857	787793	557232	314328	715454
11	779986	784040	246772	037428	943163	732084	474118	653965	290960	431722
12	498197	004955	846473	678487	895366	510523	264787	745585	369990	965279
13	274583	619411	809110	810902	920701	815422	098366	198124	248743	659873
14	322839	426424	138714	362812	087630	927775	295736	765548	092923	484176
15	354858	303042	339228	351545	953693	971864	967050	693517	615394	521983
16	357250	880208	403440	943072	211660	483043	881200	689656	924270	247856
17	656227	025944	429949	884011	457441	380407	150041	206947	533770	981969
18	028334	243014	743639	501366	180085	550866	755121	868428	721282	746140
19	541812	376729	954893	670700	510615	858287	390098	553456	160300	267504
20	726485	091091	361595	837529	826911	285934	762241	182030	037440	438023
21	810601	386378	199374	691120	363874	741727	280144	753757	680007	752865
22	325127	297280	921742	390053	484269	715908	018023	219688	000892	108108
23	640916	160035	217761	756608	533650	488496	252832	461034	111535	513549
24	114444	004315	220171	576064	706307	061956	603973	629221	941897	193569
25	305560	539347	521428	008050	288341	744520	496665	652596	737160	196115
26	195009	418433	178521	006231	042671	064997	083073	574505	145000	514404
27	376181	304142	315082	778824	333885	124678	384126	813098	124981	300500
28	022605	017805	385284	353000	735614	584795	015782	637247	622569	562455
29	825662	225607	160532	124971	951791	099351	972809	655876	821156	306904
30	207770	200921	066542	621318	973926	418894	601150	984525	644506	362960
31	971201	115129	317753	664694	740804	361810	143731	065636	563223	370052
32	400118	391167	719085	086734	600439	226241	966328	204474	402837	256477
33	014788	895905	885641	841756	054860	280091	790989	559258	354668	059636
34	848489	988069	705169	219471	857514	011362	873589	907299	333557	454484
35	061103	508647	146095	433854	189195	420229	198572	412411	694621	810477
36	637902	580908	261863	231756	649436	528688	821517	551425	827655	309571
37	737773	613469	732826	969553	179550	086001	689612	928159	456902	696314
38	221622	683631	603194	194080	443694	750158	201449	019876	112761	039281
39	702701	453917	399211	259997	461773	769792	181112	484485	095186	910185
40	837596	739666	901375	499563	721578	224058	638959	560651	898188	773168
41	699525	458282	671898	445350	197269	007266	869872	081419	273546	646059
42	440585	285228	987286	668750	658662	908196	733969	177417	193199	757370
43	073076	653438	410702	593716	332765	524483	746111	177172	149847	671238

جدول رقم (2)

جدول التوزيع الطبيعي القياسي

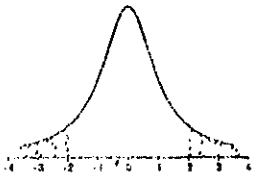


Z	0	1	2	3	4	5	6	7	8	9
0.0	.000000	.003989	.007978	.011967	.015953	.019939	.023922	.027903	.031881	.035856
0.1	.039828	.043795	.047758	.051717	.055670	.059618	.063559	.067495	.071424	.075345
0.2	.079260	.083166	.087064	.090954	.094835	.098706	.102568	.106420	.110261	.114092
0.3	.117911	.121719	.125516	.129300	.133072	.136831	.140576	.144309	.148027	.151732
0.4	.155422	.159097	.162757	.166402	.170031	.173645	.177242	.180822	.184386	.187933
0.5	.191462	.194974	.198468	.201944	.205401	.208840	.212260	.215661	.219043	.222405
0.6	.225747	.229069	.232371	.235653	.238914	.242154	.245373	.248571	.251748	.254903
0.7	.258036	.261148	.264238	.267305	.270350	.273373	.276373	.279350	.282305	.285236
0.8	.288145	.291030	.293892	.296731	.299546	.302338	.305106	.307850	.310570	.313267
0.9	.315940	.318589	.321214	.323815	.326391	.328944	.331472	.333977	.336457	.338913
1.0	.341345	.343752	.346136	.348495	.350830	.353141	.355428	.357690	.359929	.362143
1.1	.364334	.366500	.368643	.370762	.372857	.374928	.376976	.378999	.381000	.382977
1.2	.384930	.386861	.388768	.390651	.392512	.394350	.396165	.397958	.399727	.401475
1.3	.403199	.404902	.406582	.408241	.409877	.411492	.413085	.414657	.416207	.417736
1.4	.419243	.420730	.422196	.423641	.425066	.426471	.427855	.429219	.430563	.431888
1.5	.433193	.434478	.435745	.436992	.438220	.439429	.440620	.441792	.442947	.444083
1.6	.445201	.446301	.447384	.448449	.449497	.450529	.451543	.452540	.453521	.454486
1.7	.455435	.456367	.457284	.458185	.459071	.459941	.460796	.461636	.462462	.463273
1.8	.464070	.464852	.465621	.466375	.467116	.467843	.468557	.469258	.469946	.470621
1.9	.471284	.471933	.472571	.473197	.473810	.474412	.475002	.475581	.476148	.476705
2.0	.477250	.477785	.478308	.478822	.479325	.479818	.480301	.480774	.481237	.481691
2.1	.482136	.482571	.482997	.483414	.483823	.484222	.484614	.484997	.485371	.485738
2.2	.486097	.486447	.486791	.487126	.487455	.487776	.488089	.488396	.488696	.488989
2.3	.489276	.489556	.489830	.490097	.490358	.490613	.490863	.491106	.491344	.491576
2.4	.491802	.492024	.492240	.492451	.492656	.492857	.493053	.493244	.493431	.493613
2.5	.493790	.493963	.494132	.494297	.494457	.494614	.494766	.494915	.495060	.495201
2.6	.495339	.495473	.495603	.495731	.495855	.495975	.496093	.496207	.496319	.496427
2.7	.496533	.496636	.496736	.496833	.496928	.497020	.497110	.497197	.497282	.497365
2.8	.497445	.497523	.497599	.497673	.497744	.497814	.497882	.497948	.498012	.498074
2.9	.498134	.498193	.498250	.498305	.498359	.498411	.498462	.498511	.498559	.498605
3.0	.498650	.498694	.498736	.498777	.498817	.498856	.498893	.498930	.498965	.498999
3.1	.499032	.499065	.499096	.499126	.499155	.499184	.499211	.499238	.499264	.499289
3.2	.499313	.499336	.499359	.499381	.499402	.499423	.499443	.499462	.499481	.499499
3.3	.499516	.499533	.499550	.499566	.499581	.499596	.499610	.499624	.499638	.499650
3.4	.499663	.499675	.499687	.499698	.499709	.499720	.499730	.499740	.499749	.499758
3.5	.499767	.499776	.499784	.499792	.499800	.499807	.499815	.499821	.499828	.499835
3.6	.499841	.499847	.499853	.499858	.499864	.499869	.499874	.499879	.499883	.499888
3.7	.499892	.499896	.499900	.499904	.499908	.499912	.499915	.499918	.499922	.499925
3.8	.499928	.499931	.499933	.499936	.499938	.499941	.499943	.499946	.499948	.499950
3.9	.499952	.499954	.499956	.499958	.499959	.499961	.499963	.499964	.499966	.499967
4.0	.499968	.499970	.499971	.499972	.499973	.499974	.499975	.499976	.4999[illegible]	.9978

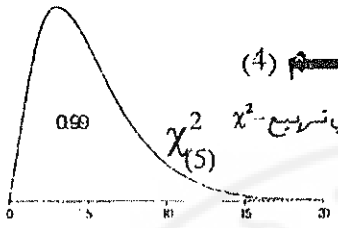


جدول رقم (3)

جدول توزيع ستيودنت t



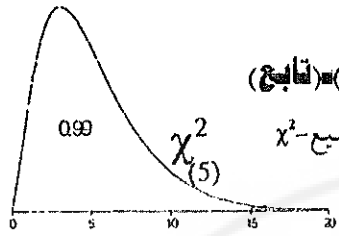
	.45	.40	.35	.30	.25	.20	.15	.10	.05	.025	.01	.005
باتجاه واحد												
باتجاهين	.90	.80	.70	.60	.50	.40	.30	.20	.10	.05	.02	.01
df												
1	0.1584	0.3249	0.5095	0.7265	1.0000	1.3764	1.9626	3.0777	6.3138	12.7062	31.8206	63.6570
2	0.1421	0.2887	0.4447	0.6172	0.8165	1.0607	1.3862	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.1366	0.2767	0.4242	0.5844	0.7649	0.9785	1.2498	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.1338	0.2707	0.4142	0.5686	0.7407	0.9410	1.1896	1.5332	2.1319	2.7764	3.7470	4.6041
5	0.1322	0.2672	0.4082	0.5594	0.7267	0.9195	1.1558	1.4759	2.0151	2.5706	3.3649	4.0322
6	0.1311	0.2648	0.4043	0.5534	0.7176	0.9057	1.1342	1.4398	1.9432	2.4469	3.1427	3.7075
7	0.1303	0.2632	0.4015	0.5491	0.7111	0.8960	1.1192	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.1297	0.2619	0.3995	0.5459	0.7064	0.8889	1.1081	1.3968	1.8596	2.3060	2.8965	3.3554
9	0.1293	0.2610	0.3979	0.5435	0.7027	0.8834	1.0997	1.3830	1.8331	2.2622	2.8214	3.2499
10	0.1289	0.2602	0.3966	0.5415	0.6998	0.8791	1.0931	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.1286	0.2596	0.3956	0.5399	0.6974	0.8755	1.0877	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.1283	0.2590	0.3947	0.5386	0.6955	0.8726	1.0832	1.3562	1.7823	2.1788	2.6810	3.0546
13	0.1281	0.2586	0.3940	0.5375	0.6938	0.8702	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.1280	0.2582	0.3933	0.5366	0.6924	0.8681	1.0763	1.3450	1.7613	2.1448	2.6245	2.9769
15	0.1278	0.2579	0.3928	0.5357	0.6912	0.8662	1.0736	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.1277	0.2576	0.3923	0.5350	0.6901	0.8647	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.1276	0.2573	0.3919	0.5344	0.6892	0.8633	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.1274	0.2571	0.3915	0.5338	0.6884	0.8621	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.1274	0.2569	0.3912	0.5333	0.6876	0.8610	1.0655	1.3277	1.7291	2.0930	2.5395	2.8610
20	0.1273	0.2567	0.3909	0.5329	0.6870	0.8600	1.0640	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.1272	0.2566	0.3906	0.5325	0.6864	0.8591	1.0627	1.3232	1.7208	2.0796	2.5176	2.8314
22	0.1271	0.2564	0.3904	0.5321	0.6858	0.8583	1.0614	1.3212	1.7172	2.0739	2.5083	2.8188
23	0.1271	0.2563	0.3902	0.5317	0.6853	0.8575	1.0603	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.1270	0.2562	0.3900	0.5314	0.6848	0.8569	1.0593	1.3178	1.7109	2.0639	2.4922	2.7969
25	0.1269	0.2561	0.3898	0.5312	0.6844	0.8562	1.0584	1.3164	1.7081	2.0595	2.4851	2.7874
26	0.1269	0.2560	0.3896	0.5309	0.6840	0.8557	1.0575	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.1268	0.2559	0.3894	0.5306	0.6837	0.8551	1.0567	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.1268	0.2558	0.3893	0.5304	0.6834	0.8546	1.0560	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.1268	0.2557	0.3892	0.5302	0.6830	0.8542	1.0553	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.1267	0.2556	0.3890	0.5300	0.6828	0.8538	1.0547	1.3104	1.6973	2.0423	2.4573	2.7500
35	0.1266	0.2553	0.3885	0.5292	0.6816	0.8520	1.0520	1.3062	1.6896	2.0301	2.4377	2.7238
40	0.1265	0.2550	0.3881	0.5286	0.6807	0.8507	1.0500	1.3031	1.6839	2.0211	2.4233	2.7045
45	0.1264	0.2549	0.3878	0.5281	0.6800	0.8497	1.0485	1.3007	1.6794	2.0141	2.4121	2.6896
50	0.1263	0.2547	0.3875	0.5278	0.6794	0.8489	1.0473	1.2987	1.6759	2.0086	2.4033	2.6778
55	0.1262	0.2546	0.3873	0.5275	.6790	0.8482	1.0463	1.2972	1.6730	2.0041	2.3961	2.6682
60	0.1262	0.2545	0.3872	0.5272	0.6786	0.8476	1.0455	1.2958	1.6707	2.0003	2.3902	2.6604
70	0.1261	0.2543	0.3869	0.5268	0.6780	0.8468	1.0442	1.2938	1.6669	1.9944	2.3808	2.6480
80	0.1261	0.2542	0.3867	0.5265	0.6776	0.8461	1.0432	1.2922	1.6641	1.9901	2.3739	2.6387
90	0.1260	0.2541	0.3866	0.5263	0.6772	0.8456	1.0424	1.2910	1.6620	1.9867	2.3685	2.6316
100	0.1260	0.2540	0.3864	0.5261	0.6769	0.8452	1.0418	1.2901	1.6602	1.9840	2.3642	2.6259
110	0.1260	0.2540	0.3863	0.5259	0.6767	0.8449	1.0413	1.2893	1.6588	1.9818	2.3607	2.6213
120	0.1259	0.2539	0.3862	0.5258	0.6765	0.8446	1.0409	1.2886	1.6577	1.9799	2.3578	2.6174
∞	0.1257	0.2533	0.3853	0.5244	0.6745	0.8416	1.0364	1.2816	1.6449	1.9600	2.3263	2.5758



جدول رقم (4)

جدول توزيع كاي تربيع - χ^2

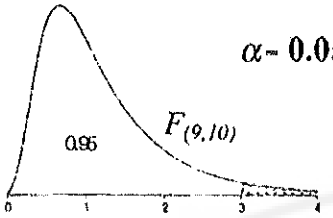
د.ح	.99	.985	.98	.95	.9	.8	.7	.6	.5
1	0.0002	0.0004	0.0006	0.0039	0.0158	0.0642	0.1485	0.2750	0.4549
2	0.0201	0.0302	0.0404	0.1026	0.2107	0.4463	0.7133	1.0217	1.3863
3	0.1148	0.1516	0.1848	0.3518	0.5844	1.0052	1.4237	1.8692	2.3660
4	0.2971	0.3682	0.4294	0.7107	1.0636	1.6488	2.1947	2.7528	3.3567
5	0.5543	0.6618	0.7519	1.1455	1.6103	2.3425	2.9999	3.6555	4.3515
6	0.8721	1.0160	1.1344	1.6354	2.2041	3.0701	3.8276	4.5702	5.3481
7	1.2390	1.4184	1.5643	2.1674	2.8331	3.8223	4.6713	5.4932	6.3458
8	1.6465	1.8603	2.0325	2.7326	3.4895	4.5936	5.5274	6.4226	7.3441
9	2.0879	2.3349	2.5324	3.3251	4.1682	5.3801	6.3933	7.3570	8.3428
10	2.5582	2.8372	3.0591	3.9403	4.8652	6.1791	7.2672	8.2955	9.3418
11	3.0535	3.3634	3.6087	4.5748	5.5778	6.9887	8.1479	9.2373	10.3410
12	3.5706	3.9104	4.1783	5.2260	6.3038	7.8073	9.0343	10.1820	11.3403
13	4.1069	4.4757	4.7654	5.8919	7.0415	8.6339	9.9257	11.1291	12.3398
14	4.6604	5.0572	5.3682	6.5706	7.7895	9.4673	10.8215	12.0785	13.3393
15	5.2293	5.6534	5.9849	7.2609	8.5468	10.3070	11.7212	13.0297	14.3389
16	5.8122	6.2628	6.6142	7.9616	9.3122	11.1521	12.6244	13.9827	15.3385
17	6.4078	6.8842	7.2550	8.6718	10.0852	12.0023	13.5307	14.9373	16.3382
18	7.0149	7.5165	7.9062	9.3905	10.8649	12.8570	14.4399	15.8932	17.3379
19	7.6327	8.1588	8.5670	10.1170	11.6509	13.7158	15.3517	16.8504	18.3377
20	8.2604	8.8105	9.2367	10.8508	12.4426	14.5784	16.2659	17.8088	19.3374
21	8.8972	9.4708	9.9146	11.5913	13.2396	15.4446	17.1823	18.7683	20.3372
22	9.5425	10.1390	10.6000	12.3380	14.0415	16.3140	18.1007	19.7288	21.3370
23	10.1957	10.8147	11.2926	13.0905	14.8480	17.1865	19.0211	20.6902	22.3369
24	10.8564	11.4974	11.9918	13.8484	15.6587	18.0618	19.9432	21.6525	23.3367
25	11.5240	12.1867	12.6973	14.6114	16.4734	18.9398	20.8670	22.6156	24.3366
26	12.1981	12.8821	13.4086	15.3792	17.2919	19.8202	21.7924	23.5794	25.3365
27	12.8785	13.5833	14.1254	16.1514	18.1139	20.7030	22.7192	24.5440	26.3363
28	13.5647	14.2900	14.8475	16.9279	18.9392	21.5880	23.6475	25.5093	27.3362
29	14.2565	15.0019	15.5745	17.7084	19.7677	22.4751	24.5770	26.4751	28.3361
30	14.9535	15.7188	16.3062	18.4927	20.5992	23.3641	25.5078	27.4416	29.3360
31	15.6555	16.4403	17.0423	19.2806	21.4336	24.2551	26.4397	28.4087	30.3359
32	16.3622	17.1663	17.7827	20.0719	22.2706	25.1478	27.3728	29.3763	31.3359
33	17.0735	17.8966	18.5271	20.8665	23.1102	26.0422	28.3069	30.3444	32.3358
34	17.7891	18.6309	19.2754	21.6643	23.9523	26.9383	29.2421	31.3130	33.3357
35	18.5089	19.3691	20.0274	22.4650	24.7967	27.8359	30.1782	32.2821	34.3356
36	19.2327	20.1110	20.7829	23.2686	25.6433	28.7350	31.1152	33.2517	35.3356
37	19.9602	20.8565	21.5419	24.0749	26.4921	29.6355	32.0532	34.2216	36.3355
38	20.6914	21.6054	22.3040	24.8839	27.3430	30.5373	32.9920	35.1920	37.3355
39	21.4262	22.3577	23.0693	25.6954	28.1958	31.4405	33.9316	36.1628	38.3354
40	22.1643	23.1130	23.8376	26.5093	29.0505	32.3450	34.8720	37.1340	39.3354



جدول رقم (4)×(تابع)

جدول توزيع كاي تربيع - χ^2

ح	.4	.3	.2	.1	.05	.02	.015	.01	.005
1	0.7083	1.0742	1.6424	2.7055	3.8415	5.4119	5.9165	6.6349	7.8795
2	1.8326	2.4079	3.2189	4.6052	5.9915	7.8240	8.3994	9.2103	10.5966
3	2.9462	3.6649	4.6416	6.2514	7.8147	9.8374	10.4650	11.3449	12.8382
4	4.0446	4.8784	5.9886	7.7794	9.4877	11.6678	12.3391	13.2767	14.8603
5	5.1319	6.0644	7.2893	9.2364	11.0705	13.3882	14.0978	15.0863	16.7496
6	6.2108	7.2311	8.5581	10.6446	12.5916	15.0332	15.7774	16.8119	18.5476
7	7.2832	8.3834	9.8033	12.0170	14.0671	16.6224	17.3984	18.4753	20.2778
8	8.3505	9.5245	11.0301	13.3616	15.5073	18.1682	18.9739	20.0902	21.9550
9	9.4136	10.6564	12.2421	14.6837	16.9190	19.6790	20.5125	21.6660	23.5893
10	10.4732	11.7807	13.4420	15.9872	18.3070	21.1608	22.0206	23.2093	25.1882
11	11.5298	12.8987	14.6314	17.2750	19.6751	22.6179	23.5029	24.7250	26.7569
12	12.5838	14.0111	15.8120	18.5493	21.0261	24.0540	24.9628	26.2170	28.2996
13	13.6356	15.1187	16.9848	19.8119	22.3620	25.4715	26.4034	27.6882	29.8194
14	14.6853	16.2221	18.1508	21.0641	23.6848	26.8728	27.8268	29.1413	31.3194
15	15.7332	17.3217	19.3107	22.3071	24.9958	28.2595	29.2349	30.5779	32.8013
16	16.7795	18.4179	20.4651	23.5418	26.2963	29.6332	30.6293	32.0001	34.2674
17	17.8244	19.5110	21.6146	24.7690	27.5871	30.9950	32.0111	33.4085	35.7182
18	18.8679	20.6014	22.7595	25.9894	28.8693	32.3461	33.3817	34.8053	37.1564
19	19.9102	21.6891	23.9004	27.2036	30.1435	33.6873	34.7419	36.1907	38.5820
20	20.9514	22.7745	25.0375	28.4120	31.4104	35.0196	36.0926	37.5662	39.9968
21	21.9915	23.8578	26.1711	29.6151	32.6706	36.3434	37.4344	38.9322	41.4011
22	23.0307	24.9390	27.3015	30.8133	33.9245	37.6596	38.7682	40.2895	42.7960
23	24.0689	26.0184	28.4288	32.0069	35.1724	38.9682	40.0939	41.6382	44.1808
24	25.1063	27.0960	29.5533	33.1962	36.4150	40.2704	41.4130	42.9798	45.5586
25	26.1430	28.1719	30.6752	34.3816	37.6525	41.5662	42.7254	44.3144	46.9285
26	27.1789	29.2463	31.7946	35.5632	38.8852	42.8559	44.0313	45.6419	48.2903
27	28.2141	30.3193	32.9117	36.7412	40.1133	44.1401	45.3312	46.9631	49.6452
28	29.2486	31.3909	34.0266	37.9159	41.3372	45.4189	46.6256	48.2783	50.9936
29	30.2825	32.4612	35.1394	39.0875	42.5570	46.6928	47.9148	49.5881	52.3360
30	31.3159	33.5302	36.2502	40.2560	43.7730	47.9618	49.1989	50.8922	53.6720
31	32.3486	34.5981	37.3591	41.4218	44.9854	49.2265	50.4784	52.1915	55.0030
32	33.3809	35.6649	38.4663	42.5847	46.1942	50.4866	51.7530	53.4856	56.3278
33	34.4126	36.7306	39.5718	43.7451	47.3998	51.7428	53.0235	54.7753	57.6479
34	35.4438	37.7954	40.6756	44.9031	48.6023	52.9952	54.2900	56.0608	58.9636
35	36.4746	38.8592	41.7780	46.0588	49.8019	54.2440	55.5529	57.3424	60.2755
36	37.5049	39.9220	42.8788	47.2122	50.9985	55.4889	56.8114	58.6193	61.5813
37	38.5349	40.9840	43.9783	48.3635	52.1925	56.7309	58.0670	59.8932	62.8847
38	39.5644	42.0451	45.0763	49.5127	53.3837	57.9692	59.3187	61.1628	64.1827
39	40.5935	43.1054	46.1731	50.6598	54.5724	59.2043	60.5670	62.4287	65.4767
40	41.6222	44.1649	47.2686	51.8051	55.7586	60.4365	61.8122	63.6914	66.7673



جدول توزیع F بسطوی معنوي $\alpha = 0.05$

$df_2 \backslash df_1$	1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
31	4.160	3.305	2.911	2.679	2.523	2.409	2.323	2.255	2.199	2.153
32	4.149	3.295	2.901	2.668	2.512	2.399	2.313	2.244	2.189	2.142
33	4.139	3.285	2.892	2.659	2.503	2.389	2.303	2.235	2.179	2.133
34	4.130	3.276	2.883	2.650	2.494	2.380	2.294	2.225	2.170	2.123
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217	2.161	2.114
36	4.113	3.259	2.866	2.634	2.477	2.364	2.277	2.209	2.153	2.106
37	4.105	3.252	2.859	2.626	2.470	2.356	2.270	2.201	2.145	2.098
38	4.098	3.245	2.852	2.619	2.463	2.349	2.262	2.194	2.138	2.091
39	4.091	3.238	2.845	2.612	2.456	2.342	2.255	2.187	2.131	2.084
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077

المصادر والمراجع
العربية والأجنبية

أولاً – المراجع العربية:

- أبو صالح، محمد صبحي؛ عوض، عدنان محمد، ١٩٩٠ ــ مقدمة في الإحصاء ــ مركز الكتب الأردني .
- أبو عمة، عبد الرحمن، راضي، الحسني عبد البر، الهندي، محمود إبراهيم، ١٩٩٥، الإحصاء والاحتمالات، الرياض، جامعة الملك سعود .
- البلداوي، عبد الحميد، الإحصاء للعلوم الإدارية والتطبيقية، ١٩٩٧، عمان، جامعة الإسراء.
- البشير، زين العابدين عبد الرحيم، ١٩٩٧، الاستدلال الإحصائي، الرياض، جامعة الملك سعود .
- الجاعوني، فريد؛ إسماعيل، حبيب؛ غانم، عدنان ، ١٩٩٩، مبادئ الإحصاء، جامعة صنعاء.
- الجراد، خلف مطر، مجموعة محاضرات في الرياضيات الاقتصادية والإدارية، ١٩٩٦، جامعة دمشق .
- الحلاق، عمر؛ عبيدو، أميرة، ١٩٩١، المدخل إلى الإحصاء الحيوي، منشورات جامعة حلب.
- العتوم، شفيق، ١٩٨٢ ــ مقدمة في الأساليب الإحصائية، الجامعة الأردنية ــ كلية الاقتصاد والعلوم الإدارية .
- العلي ، إبراهيم محمد، ١٩٧٩، مدخل في نظرية العينات ــ منشورات جامعة حلب .

- العلي ، إبراهيم محمد، ١٩٨٠، الإحصاء الرياضي، جامعة حلب .
- العلي، إبراهيم محمد، ١٩٨٥، نظرية الاحتمالات، جامعة حلب.
- الأفندي، عبد القادر، ١٩٧٥، نظرية الإحصاء ، جامعة حلب .
- اوسيليفان، جورج ، بانكروفت، جوردن، ١٩٨٣، الرياضيات والإحصاء لدراسات المحاسبة والأعمال، دار ماكجرو هيل للنشر، ترجمة مقدسي، جمال سامي، ومراجعة السيد محمد الغزي .
- حميدان، عدنان؛ مخول، مطانيوس؛ الجاعوني، فريد؛ ناصر آغا، عمار، ٢٠٠٥، الإحصاء التطبيقي، منشورات جامعة دمشق.
- حميدان، عدنان؛ الجاعوني، فريد؛ ناصر آغا، عمار؛ العواد، منذر، ٢٠٠٣، مبادئ الإحصاء ، منشورات جامعة دمشق.
- حميدان، عدنان؛ النعيمي، قاسم؛ ناصر آغا، عمار، ٢٠٠٤، مبادئ الإحصاء، منشورات جامعة دمشق، مركز التعليم المفتوح .
- حيدر، ناظم، ١٩٨٢، مبادئ الإحصاء ــ منشورات جامعة دمشق.
- عبد العزيز، عمر عبد الجواد؛ بلعربي ، عبد الحفيظ، ١٩٩٩، مقدمة في الطرق الإحصائية مع تطبيقات، تجارية ــ دار زهرات للنشر والتوزيع ــ الأردن .
- عبد المجيد ، سمير، ١٩٩٤، علم الإحصاء ، الرياض، دار مرامر للطباعة والنشر.
- عربش، شفيق؛ مخول، مطانيوس وآخرون، محاضرات في مبادئ الإحصاء خلال السنوات الماضية في كلية الاقتصاد بجامعة دمشق.
- علي، إبراهيم، ١٩٩٨، مبادئ عامة في العينات وبعض أنواع العينات ــ ورشة العمل حول مسح الهجرة الداخلية في الجمهورية العربية السورية .

- علي ، نبيل ، ١٩٩٤، مقدمة في مبادئ الإحصاء، منشورات جامعة دمشق.
- شجبيل، موراي ر.، ١٩٧٢، سلسلة ملخصات شوم ــ نظريات ومسائل في الإحصاء ــ دار ماكجرو هيل للنشر.
- رمضان ، محمد، ١٩٨٤، الإحصاء الاجتماعي والعينات، جامعة دمشق.
- طيوب، محمود؛ غانم، عدنان، ٢٠٠٢، مبادئ الإحصاء ــ جامعة صنعاء .
- كابوس، أمل، ١٩٨٤، الإحصاء الحيوي ــ منشورات جامعة حلب .
- كاربنتر، ر أ ي ل؛ فاسو، إلين ستوري؛ ترجمة: حسب الله سيد؛ غندور، محمد جلال سيد محمد، ١٩٩٨، الإحصاء للمكتبين ، دار المريخ للنشر، الرياض.
- كنجو ، أنيس، ١٩٩٥، تقنية المعاينة الإحصائية ، جامعة الملك سعود، الرياض.
- مخول، مطانيوس، غانم، عدنان، ٢٠٠٦ ، مبادئ الإحصاء ــ جامعة دمشق.

ثانياً- المراجع الأجنبية:

1- Kolmogorov A.N, Theory Of Probability and Mathematical Statistics, Moscow " Seince", 1967 .

2-Mitropolskyi. A.K., Statistical Calculations Technics. Moscow, "Seicnce", 1978.

3- Druzhynin n.k. Sampling and Experiment, Gernal Logical Principles of Organization, Moscow "Statistica" 1977.

4- Fisher R.A., statistical Methods for an analyst, Moscow 1958.

5- Cohen J.Statistical Power Analysis, Hillsdale, New Jersey, 1988.

6- Collyer C. and Enns J., analysis of Variance: the Basic Designs. Chicogo, 1976.

7- Kirk R., Experimental Dezign: Procedures for the Behavioral Sciences., California, 1982.

8- Ferguson G.A., Statistical Analysis in Psychology and Education. London 1981: Mc graw HiLL.

9- Sprinthau R. C, Basic Statistical Analysis. Readiny U.S.A 1982: Addison – Wesley.

10- Sche -ffer W.C., Statistics for Biological Sciences, 2nd Ed. Reading. MA. U.S.A 1979: Addison
11– Wesley. Bradley, James V.Distribution-Free Statistical Tests, Prentice-Hall, Inc., Englewood Cliffs, N.Y., 1968. De Groot, Morris H., Probability and Statictics, Addison-Wesley, Inc., 1975.
11- Freedman, David, Robert Pisani, and Roger Purves. Statistics. W.W. Norton @ Co., Inc., 1978.
12- Gibbons, Jean Dickinson, Nonparametric Methods of quantitative Analysis. American Sciences Press, Inc., 1985, 2nd edition.
13- Groebner, David F., and Patrick W. Shannon. Business Statistics. Charles Merril Co. 1981.
14- Hogg, R.V., and A.T.Craig, Introduction to Mathematical Statistics, 1978, 4th edition.
15- Mendenhall, W., Introduction to Probability and Statistics., PWS Publishers, 1983, 6th edition.
16- PWS publishers, 1983, 6th edition.
Ott, L.An Introduction to Statistical Methods and Data Anaysis., 17-PWS-Kent Publishing Co. 1988, 3rd edition.


وقل رب زدني علما
جامعة دمشق
Damascus University

اللجنة العلمية:

- الدكتور مطانيوس مخول
- الدكتور هاني عمران
- الدكتور محمود طيوب

المدقق اللغوي: الدكتورة منيرة فاعور

حقوق الطبع والترجمة والنشر محفوظة لمديرية الكتب والمطبوعات الجامعية

وقل رب زدني علما
جامعة دمشق
Damascus University