الإحصاء في المكتبات

الدكتور عدنان غانم

أستاذ مساعد في قسم الإحصاء التطبيقي

كلية الاقتصاد ـ جامعة دمشق

ملاحظة المطلوب هو الفصل الأول والثاني والخامس

مقدمة: لكلمة **الإحصاء** معاني كثيرة ومتباينة فعند سماع هذه الكلمة يتبادر إلى أذهان البعض من الناس مفهوم الجداول والأرقام الخاصة بعد مفردات معينة كعدد السكان أو عدد المساكن والبعض الآخر من الناس يُبني مفهوم الإحصاء لديه بناءً على ما تنشره الصحف من بيانات مختلفة أو بيانات استطلاعات الرأي أو تلك البيانات التي توضح المواد المصدرة والمستوردة في بلد ما.. إلى غير ذلك من الأمثلة المتداولة يومياً.

بدأ استخدام كلمة إحصاء لأول مرة في مجالات متعلقة باعداد جداول ذات طابع تنظيمي في التنظيمات الحكومية مثل جداول جمع الضرائب وغيرها.

أما في الوقت الحاضر فقد تعددت أساليب استخدام الإحصاء في معظم مجالات الحياة وأصبحت طرائق البحث العلمي تتطلب حضور الإحصاء بشكل كبير نظراً لفائدته في تصميم التجارب العلمية وإمكانية الحصول على النتائج التي يسعى الباحثون إلى تحقيقها عن طريق هذا العلم، بل أصبح الإلمام بأساليب استخدام الإحصاء البسيطة وكيفية تطبيقها ضرورة ملحة من مقتضيات الحضارة المعاصرة.

يشكل الإحصاء في **التخطيط** أساساً يعتمد عليه، في رسم أية خطة مستقبلية تحددها السياسة العليا للدولة لتحقيق أهدافاً معينة في مختلف المجالات والأنشطة، ومن المعروف أنه لايمكن أن توضع أية خطة دون توفر البيانات الإحصائية، وكلما كانت البيانات دقيقة ومفصلة كلما كانت الخطة أسلم وأقرب إلى الواقع.

كما أنه يحتل مكانة في **الاقتصاد** وبخاصة في تطوير الاقتصاد ونموه لأن توفر البيانات الاحصائية عن اقتصاد دولة ما يشكل الأساس في وضع السياسات الاقتصادية الملائمة للنظريات الاقتصادية.

أما أهمية الإحصاء في **العلوم الإدارية** فتنبع من خلال إستخدام أدواته في العديد من الموضوعات التي يمكن أن تخضع للقياس والتعبير عنها كمياً (بحوث العمليات، نظرية القرارات الإدارية).

ويعد الإحصاء في **العلوم الإجتماعية** وسيلة مهمة لقياس مدى رفاهية الناس وتقدم المجتمع ورقي مستوى الأفراد ثقافياً واجتماعياً، حيث يمكن جمع البيانات الإحصائية المتعلقة بالحالة الثقافية "مثقفين، أميين" عددهم ونسبتهم في المجتمع وتوزيعهم حسب النوع أو المناطق.

لقد توخى الباحث البساطة والسهولة في العرض الرياضي، وابتعد عن البراهين المعقدة، ليكون هذا الكتاب في متناول جميع الدارسين والمهتمين بعلم الإحصاء.

يتكون هذا الكتاب من سبعة فصول توزعت على النحو التالي:

1 . جمع وعرض البيانات الإحصائية.

2 . التوصيف الكمي للبيانات.

3 . مبادئ الاحتمالات.

4 . المتغيرات العشوائية والتوزيعات الاحتمالية.

5 . الارتباط والانحدار .

6 . السلاسل الزمنية.

وقد ضمّن الباحث كل فصل من الفوصل السابقة عدداً من التمارين المحلولة وأخرى غير محلولة لإتاحة الفرصة للطالب كي يختبر مدى تفهمه وإدراكه للمعلومات المعطاة وألحقت في نهاية الكتاب الجداول الإحصائية اللازمة، بالإضافة إلى قائمة بالمراجع العربية والأجنبية وأخرى بالمصطلحات الإحصائية.

إن الباحث يضع هذا الكتاب بين أيديكم، نأمل منكم جميعاً طلاباً وباحثين، أن يتفضلوا بتقديم أي نقد بناء يساهم في إغنائه وإخراجه مستقبلاً بشكل نجعل منه أكثر نفعاً وفائدة، وبما يسمح لنا في متابعة مسيرة البحث العلمي، وأخيراً أقدم جزيل الشكر والتقدير لكل من ساهم في إخراج هذا الكتاب، والله ولي التوفيق.

الدكتور عدنان غانم

أستاذ مساعد بقسم الإحصاء التطبيقي

كلية الاقتصاد . جامعة دمشق

الفصل الأول

جمع وعرض البيانات الإحصائية

. جمع البيانات الإحصائية

. مصادر البيانات الإحصائية

. أساليب جمع البيانات الإحصائية

. طرق جمع البيانات الإحصائية

. تفريغ وتبويب البيانات الإحصائية

. عرض البيانات الإحصائية (العرض الجدولي، العرض البياني)

. التوزيعات التكرارية

. أسئلة وتمارين غير محلولة

جمع وعرض البيانات الإحصائية

1-1- جمع البيانات الإحصائية :

تعتبر عملية جمع البيانات الإحصائية أهم مراحل العملية الإحصائية حيث تعتبر عملية جمع البيانات بمثابة الأساس الذي تبني عليه باقي مراحل الدراسة الإحصائية للظاهرة ، فإذا ما حدث خلل ما في عملية جمع البيانات ترتب على ذلك التوصل إلى نتائج مضللة عن تلك الظاهرة.

وقبل أن نبدأ بعملية جمع البيانات فإنه يجب أولاً التعرف على العناصر الأساسية التالية:

2-1- مصادر البيانات الإحصائية :

يمكن أن نميز بين نوعين أساسيين من مصادر العمليات الإحصائية:

I . المصادر الأصلية أو المباشرة :

وهي عبارة عن الوحدات الأصلية التي نتلقى منها المعلومات أو البيانات، مثال ذلك جمع البيانات الخاصة بالدخل والانفاق مباشرة من الأسر عن طريق الاستبيان.

وتمتاز عملية جمع البيانات بالطريقة المباشرة بإمكان مراجعتها والتحقيق من دقتها مع مصدر البيان نفسه. ولكن يعاب عليها كثرة الجهد والتكاليف.

II . المصادر التاريخية أو غير المباشرة :

وهي عبارة عن تلك السجلات التي تقوم هيئات رسمية أو شبه رسمية بإصدارها بصفة دورية، حيث يمكن الرجوع إليها للحصول على تلك البيانات أو المعلومات الإحصائية دون الرجوع إلى الوحدات الأصلية ومن أمثلتها: (سجلات السكان، سجلات المواليد والوفيات، السجلات المالية، سجلات الطلبة....)

وتمتاز هذه الطريقة بالسهولة والتوفير في الجهد والتكاليف والوقت بينما يعاب عليها عدم إمكان التحقق من صحة البيانات .

3-1- أساليب جمع البيانات الإحصائية :

إذا ما تقرر جمع البيانات الإحصائية من مصادرها الأصلية أو المباشرة فإن عملية جمع البيانات يمكن أن تتم باتباع أحد أسلوبين أساسيين هما :

I . 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 ولذا فإن هذا الأسلوب يتبع في حالات معينة تقتضيها طبيعة البحث.

II . أسلوب العينات : Sampling Method

هو اتباع أسلوب علمي يقتصر بموجبه على بحث عدد محدد من وحدات المجتمع والخروج بنتائج يمكن تعميمها على المجتمع وفق قواعد علمية، ولضمان دقة النتائج التي نحصل عليها باتباع ذلك الأسلوب فإنه يجب أن تمثل العينة المجتمع الأصلي الذي سحبت منه أصدق تمثيل. ويتميز أسلوب العينة بتوفير الجهد والوقت والتكاليف.

أنواع العينات :

أولاً : العينات الاحتمالية وتشمل :

I . العينة العشوائية البسيطة Simple Random Sample :

يمكن الحصول على عينة عشوائية بسيطة بإحدى طريقتين:

الطريقة الأولى : طريقة القرعة وتتلخص في عمل بطاقة لكل وحدة من وحدات المجتمع ثم نختار بطريقة عشوائية عدداً من تلك الوحدات .

الطريقة الثانية : طريقة جداول الأرقام العشوائية، حيث نعطي مفردات المجتمع المدونة في الإطار أرقاماً مسلسلة ثم نختار أحد الأعمدة أو الصفوف بالجدول العشوائي لاختيار وحدات العينة (انظر جداول الأرقام العشوائية في نهاية الكتاب).

الطريقة الثالثة: دواليب الحظ

الطريقة الرابعة: توليد الأرقام العشوائية بواسطة الحاسوب.

II . العينة الطبقية Staratified sample :

وتستخدم في حالة المجتمعات غير المتجانسة حيث يمكن تمييز وحدات المجتمع على شكل مجموعات متجانسة فيما بينهما وتسمى كل مجموعة حينئذٍ طبقة، حيث نقوم بحسب عدد من وحدات كل طبقة بطريقة عشوائية وتكون العينة عبارة عن إجمالي الوحدات التي تم اختيارها من كل طبقة على حدة، وغالباً ما يتناسب عدد الوحدات التي يتم اختيارها من كل طبقة مع حجم الطبقة.

III . العينة المتعددة المراحل Multistage Sample :

تستخدم في حالة المجتمعات المنتشرة على مساحات كبيرة، وفيها يقسم المجتمع الأصلي إلى عدد الوحدات الابتدائية حيث نختار عدداً منها بطريقة عشوائية كمرحلة أولى، وهكذا تتعاقب مراحل المعاينة ويستخدم ذلك النوع من العينات لغرض التوفير في تكاليف انتقال الباحثين خاصة في حالة المجتمعات الكبيرة .

V . العينة المنتظمة Systematic Sample :

تعتمد على قائمة تتضمن مفردات المجتمع وعلى تحديد فترة السحب $K = \frac{N}{n}$

فإذا كان لدينا مجتمع مكون من 90 وحدة ومطلوب اختيار عينة من 10 وحدات فإن فترة السحب = 90/10 = 9، نختار بعد ذلك رقماً عشوائياً بين 9,1 وليكن الرقم 5 فتكون وحدات العينة هي الوحدات ذات الرقم المسلسل بالإطار 5، 14، 23، 41، 50، 59، 68، 77، 86.

وتمتاز طريقة العينة المنتظمة بالسهولة والبساطة في إجرائها ولكن يعاب عليها عدم صلاحيتها إذا ما وجدت علاقة دورية في ترتيب الوحدات الأولية المكونة للمجتمع مع طول فترة المعاينة.

IV . العينة العنقودية: تقوم على تقسيم المجتمع إلى مجموعات جزئية والاختيار يكون على شكل عناقيد وليس مفردة مثال : في إحدى المدن (10,000) مسكن وطلب منا إختيار عينة من (500) مسكن في هذه الحالة يتم تقسيم المدينة إلى (1000) منطقة وبكل منطقة (10) مساكن ومن ثم نختار عشوائياً (50) منطقة من المناطق الـ (1000) وتجمع المعلومات اللازمة عن كل مسكن من المساكن بالمناطق المختارة.

ثانياً . العينات غير الاحتمالية :

I . عينة الحصص حيث تقسم العينة المطلوبة إلى حصص تشمل كل حصة مجموعة من الوحدات يتم جمع البيانات منها على أن تترك عملية اختيار تلك الوحدات للباحث نفسه واعتماد نسبة معينة من المفردات في كل طبقة.

فلو أننا حددنا العينة بـ 400 طالب من كلية الاقتصاد بجامعة دمشق موزعين على التخصصات الأربعة اقتصاد وإدارة ومحاسبة ومصارف فإذا فرض أن نسبة الطلبة من كل تخصص هي : 1:1:5:3 معنى ذلك أننا سنقوم ببحث 120 طالباً من تخصص اقتصاد 200 من تخصص إدارة، 40 من تخصص المحاسبة، 40 من تخصص المصارف.

ونترك عملية اختيار هؤلاء الطلبة للقائمين بعملية جمع البيانات أنفسهم وبالطبع ستتم عملية المقابلة وجمع البيانات مع الطلبة الذين يمكن مقابلتهم أو من السهل جمع المعلومات منهم. وتمتاز هذه العينة بالبساطة وقلة التكاليف والسرعة في الحصول على النتائج.

II . العينة العمرية (المنتقاة) : في هذه المعاينة الباحث هو الذي يحدد المفردات التي ستدخل في العينة وجودة النتائج تتوقف على حكمة الباحث ومهارته في الاختيار.

- **أخطاء العينة:**

1 . الأخطاء الشخصية (التحيز) : تنتج عن التقصير في خطوات البحث بالعينة، وأسبابها عدم مراعاة الأساليب العلمية في سحب العينات.

2 . الأخطاء الاحتمالية (الخط والصدف): تصاحب عملية الاختيار وناتجة عن دراسة جزء من المجتمع.

1-4- طرق جمع البيانات الإحصائية :

يمكن جمع البيانات والمعلومات بإحدى طريقتين رئيسيتين :

I . الطريقة المباشرة :

وذلك عن طريق الاتصال المباشر بمصدر البيان أو وحدة البحث وذلك بالاتصال الشخصي بين الباحث أو مندوبه والمصدر.

وقد يتم الاتصال المباشر بالهاتف أو بالانترنت كما تشمل أيضاً أسلوب الملاحظة وتسجيل البيانات مباشرة وتمتاز الطريقة المباشرة بإمكان الحصول على بيانات أكثر دقة.

II . الطريقة غير المباشرة :

حيث يتم جمع البيانات أو المعلومات من المصدر بطريقة غير مباشرة مثال ذلك عن طريق إرسال صحيفة الاستبيان إليه بالبريد Postal Questionnaire. كما تعتبر عملية جمع البيانات من المصادر التاريخية أو الإحصاءات المنشورة . إحدى الطرق غير المباشرة في جمع البيانات .

ويعيب الطريقة غير المباشرة انخفاض نسبة الردود وعلى عدم إمكان مراجعتها مع مصدر البيان للتحقق من دقة البيانات.

- **مكونات الاستمارة الإحصائية :**

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.

الجزء الثاني :

ويشمل البيانات المميزة لوحدة البحث (الاسم، الجنس، العمر، مكان الإقامة ....)

الجزء الثالث :

وهو الجزء الهام 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.

- **أهم الشروط الواجب مراعاتها عند وضع أسئلة الاستمارة الإحصائية :**

1. يجب أن تكون الأسئلة محددة وواضحة.
2. يجب أن الأسئلة بعيدة عن التأويل والتفسير
3. يستحسن تجنب الأسئلة التي تثير الحرج أو التي تحتاج للرجوع بالذاكرة.

5-1- تفريغ وتبويب البيانات الإحصائية :

بعد عملية جمع البيانات سواء من مصادرها الأصلية باستخدام الاستمارات أو من المصادر التاريخية، يكون من الصعب فهم مدلولها واتجاهها، لذلك يقوم الباحث بتبويب تلك البيانات وعرضها في جداول إحصائية بحيث يمكن باستعراضها تكوين فكرة عامة عن طبيعة الصفة أو الصفات محل الدراسة وهذه الجداول تعتبر أساساً للتحليل الإحصائي .

- **تبويب البيانات :**

I . التبويب اليدوي :

يعتمد على جدول التفريغ الذي يتكون من ثلاثة أعمدة يخصص الأول منها لبيان القيم والعمود الثاني لوضع الإشارات نتيجة فرز الاستمارات واحدة تلو الأخرى ويخصص العمود الثالث لوضع عدد تلك الإشارات مقابل الإجابة المعنية وهذا ما يطلق عليه اسم التكرار ويرمز له بالحرف (f) .

II . التبويب الآلي :

يعتمد على الحاسب الآلي وفيه تترجم البيانات المدونة في الاستمارات الإحصائية إلى رموز معينة مقسمة إلى عدة أعمدة يحمل كل عمود منها رقماً مسلسلاً ويقسم كل عمود إلى سطور .

مثال : أخذت عينة عشوائية من (93) عاملاً من عمال إحدى الصناعات بهدف دراسة الحالة التعليمية لهؤلاء العمال فكانت النتائج التالية: أمي، أمي، يقرأ ويكتب، اعدادي... وهكذا لبقية العمال.

المطلوب: تفريغ هذه البيانات في جدول تفريغ.

جدول رقم (1) التوزيع التكراري للعمال حسب الحالة التعليمية

الحالة التعليمية	الإشارات	التكرار
أمي	𝍸 𝍸 𝍸	15
يقرأ ويكتب	𝍸 𝍸 𝍸	13
ابتدائي	𝍸 𝍸 | |	12
اعدادي	𝍸 𝍸 | | |	13
ثانوي	𝍸 𝍸 𝍸 |	16
جامعي	𝍸 𝍸 𝍸	14

العمود الأول والثالث يكونان معاً "جدول التوزيع التكراري" للمتغير وهو يعطينا عدد المرات التي ظهرت فيها كل من قيم المتغير في العينة .

وعندما تكون قيم المتغير كثيرة (سواء أكان كمياً أو وصفياً) أو يكون المتغير كمياً متصلاً فمن الصعب أن نخصص لكل قيمة من قيم المتغير خانة خاصة بها ولذلك نجمع القيم المتقاربة في عدد مناسب من الفئات. وبشكل عام ليس هناك قاعدة ثابتة للحصول على العدد الملائم من الفئات ولو أن بعض الإحصائيين ينصحون باستخدام قاعدة سترجس التالية:

$$\text{عدد الفئات} = 1 + 3.322 \log (n)$$

حيث n = عدد القيم.

وعموماً فإننا نأخذ عادة عدداً من الفئات يتراوح بين (5 وَ 12) فئة طبقاً لحجم العينة وعادة نراعي أيضاً أن تكون حدود الفئات ملائمة ويتم ذلك بأن نحدد طول المدى وهو عبارة عن الفرق بين أكبر قيمة وأصغر قيمة فيها ثم نحدد طول الفئة :

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}}$$ (يتم تقريب الناتج إلى أقرب رقم صحيح)

وكما سبق فإن العمودين الأول والثالث يكونان جدول التوزيع التكراري المطلوب.

في المثال السابق اخترنا فئات متساوية الطول ولكن هذه ليست قاعدة ثابتة، فهناك الجداول التكرارية ذات فئات غير متساوية

وإذا لم يكن هناك الحد الأعلى لفئات الجدول قيل إن الجدول مفتوح من أعلى وبالمثل إذا لم يمكن هناك أصغر فئة قيل إن الجدول مفتوح من أسفل.

أشكال الفئات :

a) الشكل العام :

60 وأقل من 50 ويمكن كتابتها 60 – 50

وهذه الصورة العامة للفئات سواء أكان المتغير متصلاً أو منفصلاً.

b) الفئات المحددة البداية والنهاية :

30-39

40-49

لاحظ أنه في هذا الشكل من الفئات يعين كل من الحد الأدنى والحد الأعلى لكل فئة ولذلك فإن هذا الشكل يستخدم في حالة المتغيرات المنفصلة (المتقطعة) فقط .

6-1- عرض البيانات الإحصائية :

تعتبر عملية عرض البيانات الإحصائية امتداداً لعملية التفريغ والتبويب بحيث يصعب الفصل بينهما وكلتا العمليتين تهدفان إلى ترتيب البيانات المجموعة عن الظاهرة محل الدراسة وتنسيقها بطريقة تساعد على فهم مدلولها والاستفادة منها تمهيداً لعرضها وإبراز أهميتها.

- **العرض الجدولي للبيانات الإحصائية :**

العرض الجدولي هو عبارة عن عرض البيانات على شكل جدول، والجدول الإحصائي عبارة عن ترتيب البيانات العددية في صورة صفوف وأعمدة بحيث يمكن قراءة الجدول بالاتجاهين والهدف منه الاختصار وإبراز أهمية البيانات.

وتسمى العملية التي يتم فيها تجميع البيانات في مجموعات مميزة ومتجانسة عملية التصنيف وبوجه عام تصنف البيانات الإحصائية طبقاً لإحدى القواعد التالية :

I . تصنيف جغرافي :

وعلى أساس هذا هذا التصنيف تجمع الوحدات التي تشترك في صفة مكانية واحدة في مجموعة مستقلة مثال ذلك: التوزيع النسبي للصادرات والمستوردات السورية حسب الكتل الدولية لعام 2004.

جدول رقم (4)

الكتل الدولية	الصادرات	المستوردات
البلدان العربية	29.9	16.4
بلدان الاتحاد الأوروبي	53.2	15.2
بلدان أوربية أخرى	2.3	17.1
البلدان الأمريكية	3.5	9.6
بلدان أسيوية مختلفة	10.4	26.5
بلدان أخرى	0.7	15.2

م.م للإحصاء 2004

ومن الأمثلة الأخرى للتنصيف الجغرافي أيضاً توزيع سكان سورية على المحافظات المختلفة.

II . تصنيف تاريخي أو زمني:

وفيه تجمع الوحدات المرتبطة بزمن معين في مجموعة واحدة أي أن تقسيم البيانات يتم حسب زمن حدوثها وتسمى بيانات ذلك النوع من الجداول باسم السلسلة الزمنية.

مثال ذلك : نسبة صادرات النفط والقطن من إجمالي الصادرات السورية خلال الفترة (1999- 1995).

جدول رقم (5)

السنة	النفط الخام	القطن	حصة النفط والقطن من اجمالي الصادرات
1995	62.5	5.6	68.1
1996	69	4.5	73.5
1997	63.5	6.7	70.2
1998	59	10	69
1999	67.1	4.7	71.8

المصدر: المجموعة الاحصائية السورية ، إعداد 1996-2000

III . تصنيف نوعي أو وصفي :

حيث تجمع الوحدات ذات النوع الواحد أو التي تشترك في صفة واحدة في مجموعة واحدة مثال ذلك.

جدول التركيب النسبي للواردات السورية حسب طبيعة المواد خلال الفترة (1996 . 1999)

جدول رقم (6)

السنة	الواردات %		
	خام	نصف مصنعة	مصنعة
1996	8	47	45
1997	10	48	42
1998	10	48	42
1999	11.4	47.8	40.8

المصدر: المجموعة الإحصائية السورية إعداد 1997-2000

V . 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 العاملين في إحدى الشركات حسب فئات العمر

جدول رقم (7)

فئات السن	أقل من 20	25-20	30-25	35-30	40-35	45-40	50-45	55-50
عدد العاملين	10	18	24	36	60	30	14	5

المصدر: فرضي

أشكال الجداول التكرارية :

يمكن التمييز بين ثلاثة أشكال عامة للجداول الإحصائية هي :

I . الجدول البسيط :

وهو الجدول الذي يشمل بياناً واحداً فقط فهو يحتوي على عمود واحد للبيان موضوع الدراسة يقابله عمود آخر بعدد التكرارات المقابلة، مثال ذلك: توزيع موظفي الدولة حسب الفئة.

توزيع 500 موظف بإحدى الوزارات حسب فئات الراتب (بالدولار)

(جدول رقم 9)

فئات الراتب	100-50	150-100	200-150	250-200	300-250	المجموع
عدد الموظفين	42	58	180	140	80	500

II . الجدول المزدوج :

وهو الجدول الذي صنفت فيه الوحدات حسب ظاهرتين مثال ذلك الجدول التالي والخاص بتوزيع 600 رجل حسب العمر وعمر الزوجة.

جدول رقم 10

عمر الرجال / عمر الزوجات	30-20	40-30	50-40	60-50	70-60	80-70	90-80	المجموع
20-15	35	63	22	2	-	-	-	122
25 – 20	25	72	33	13	5	1	-	149
30 – 25	17	51	30	16	8	2	1	125
35-30	2	33	28	19	15	7	3	107
40 – 35	1	18	20	15	7	4	-	65
45 – 40	-	3	12	10	5	1	1	32
المجموع	80	240	145	75	40	15	5	600

- **الشروط الواجب مراعاتها عند إنشاء الجداول الإحصائية :**

1 . كتابة عنوان الجدول بشكل واضح يبين محتواه.

2 . كتابة المصدر الذي أخذت منه البيانات (وزارة التعليم العالي، جمع شخصي، مصدر فرضي).

3 . تسجل الملاحظات الخاصة في أسفل الجدول مع الإشارة إليها بعلامات خاصة.

4 . يجب أن تذكر وحدات القياس ورقم الجدول .

- **العرض البياني :**

كما سبق وأوضحنا فالهدف الأساسي لعلم الإحصاء هو جمع وعرض البيانات المجموعة عن ظاهرة أو ظواهر معين بأسلوب يسهل معه إدراك اتجاه الظاهرة. ولذلك فإنه إلى جانب العرض الجدولي للبيانات الإحصائية يسعى الباحث غالباً إلى عرض النتائج بعد تجميعها وتبويبها متبعاً

أحد أساليب العرض البياني الآتية :

I . الأشكال المعبرة :

وفي هذه الحالة تستخدم رسومات أو صور أو أشكال تعبيرية تمثل القيم المختلفة للظواهر المطلوب عرضها، وهذا الأسلوب يقتصر استخدامه على القيم الإجمالية للظاهرة أو الظواهر محل الدراسة ولغرض المقارنات فقط.

II . الرسوم الدائرية:

في هذه الحالة نفرض أن القيمة الكلية للظاهرة المدروسة تمثل إجمالي مساحة الدائرة ثم نقسم الدائرة إلى قطاعات جزئية بحيث تناسب مساحات تلك القطاعات مع قيم المجموعات الجزئية التي تم توزيع الظاهرة عليها على أن تميز تلك القطاعات عن بعضها بألوان مختلفة لضمان الإيضاح، وتستخدم هذه الطريقة في إظهار نسب التوزيع الهيكلي للظاهرة.

مثال :

الجدول التالي يوضح طلبة الدراسات في كلية الاقتصاد بجامعة دمشق حسب تخصصهم والمطلوب عرض تلك البيانات باستخدام دائرة مجزأة .

جدول رقم (11)

التخصص	عدد الطلبة
اقتصاد	14
إدارة	12
محاسبة	24
تأمين ومصارف	6
إحصاء	4
المجموع	60

الحل :

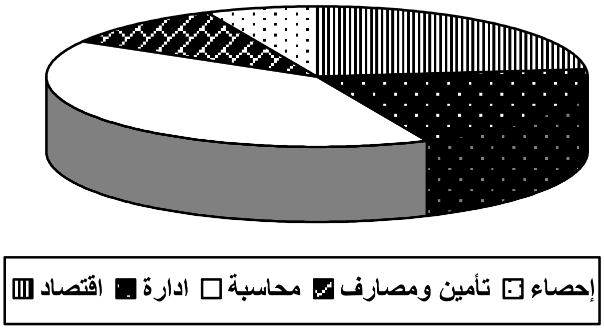
1 . نرسم دائرة باستخدام مقياس رسم مناسب.

2 . نحسب مقدار زوايا القطاعات المختلفة حسب ما هو وارد بالجدول التالي:

جدول رقم (12)

التخصص	عدد الطلبة	النسبة المئوية%	قيمة زاوية القطاع
اقتصاد	**14**	**0.23**	**83**
ادارة	**12**	**0.20**	**72**
محاسبة	**24**	**0.40**	**144**
تأمين ومصارف	**6**	**0.10**	**36**
إحصاء	**4**	**0.07**	**25**

وبذلك نقسم الدائرة إلى قطاعات وتأخذ الشكل التالي :



III . الأعمدة أو المستطيلات :

تستخدم الأعمدة أو المستطيلات ذات القواعد المتساوية لعرض البيانات الإحصائية والفكرة الأساسية في هذه الحالة تبنى على اساس أعمدة أو مستطيلات متساوية القاعدة يتناسب ارتفاعها مع قيم الظاهرة المعروضة مع استخدام مقياس رسم مناسب يمكن من عرض أكبر القيم المعروضة. وعند استخدام طريقة المستطيلات أو الأعمدة في عرض البيانات الإحصائية فإنه يمكن التمييز بين الحالات الآتية :

الحالة الأولى :

إذا كان الرسم يعرض ظاهرة واحدة في عدة فترات زمنية أو أوجه مختلفة لظاهرة واحدة فيخصص لكل فترة زمنية أو كل وجه عمود مستقل مع ترك مسافات متساوية بين كل عمود وآخر وبذلك نحصل على رسم عبارة عن مجموعة من الأعمدة المتجاورة.

مثال :

البيان التالي يوضح عدد الخريجين (دكتوراه) من الجامعات السورية خلال الفترة (1999 - 2004).

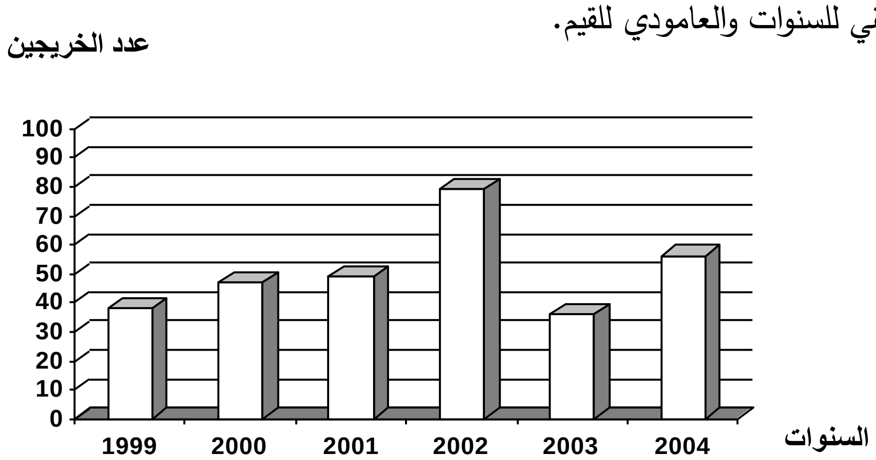
جدول رقم (13)

السنة	عدد الخريجين (دكتوراه)
1999	38
2000	47
2001	49
2002	79
2003	36
2004	56

المصدر: المجموعة الاحصائية السورية لعام 2005

المطلوب: عرض تلك البيانات باستخدام الأعمدة .

الحل : باستخدام ورقة الرسم البياني للسهولة فإننا نقوم برسم المحورين الأفقي ويخصص المحور الأفقي للسنوات والعامودي للقيم.



الحالة الثانية : الأعمدة المتلاصقة :

تستخدم عند عرض ظاهرة لعدة فترات زمنية أو عدة أماكن جغرافية، بمعنى آخر عند المقارنة بين ظاهرتين أوأكثر لعدةسنوات، مثال ذلك الصادرات والواردات لنفس الفترة وعدد المواليد وعدد الوفيات لنفس السنة .. الخ وفي هذه الحالة تمثل كل فترة أو كل منطقة بعدد من الأعمدة المتلاصقة يساوي عدد أوجه الظاهرة مع تمييز كل عمود بلون خاص.

مثـال : البيان التالي يوضح قيمة الصادرات والواردات بملايين الليرات السورية في الفترة من 1995-1999 . والمطلوب عرض تلك البيانات باستخدام الأعمدة .

جدول رقم (14)

السنة	الصادرات	الواردات
1995	44562	52856
1996	44887	60385
1997	43953	45211
1998	32443	43725
1999	38880	43010

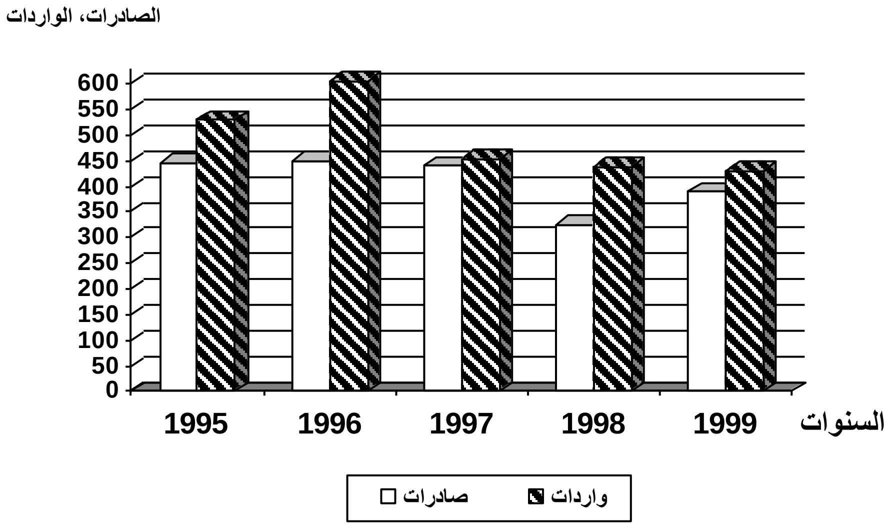
المصدر : المجموعة الإحصائية السورية (1996-2000)

والمطلوب: عرض تلك البيانات باستخدام الأعمدة (المستطيلات).

الحل: يخصص لكل سنة على المحور الأفقي عمودان أحدهما يمثل الصادرات والآخر الملاصق له يمثل الواردات مع ترك مسافات منتظمة بين أعمدة كل سنة.

شكل رقم (14)

تطور قيمة الصادرات والواردات



الحالة الثالثة : الأعمدة المجزأة

إذا كانت القيمة الإجمالية للظاهرة موزعة على مجموعات فرعية مميزة، أي تجزئة القيم الاجمالية وكل جزء يخصص له ارتفاع يتناسب مع حجم الظاهرة بالنسبة إلى الحجم الكلي بمعنى آخر العمود الذي يمثل إجمالي قيمة الظاهرة يقسم بما يتناسب وقيم المجموعات الجزئية مع التمييز بين أقسام كل عمود بألوان مميزة .

مثال :

البيان التالي يوضح أعداد العاملين في إحدى الدوائر الحكومية والوضع الوظيفي لهم في أعوام 2000، 2001، 2002.

والمطلوب : تمثيل تلك البيانات باستخدام أسلوب الأعمدة (المستطيلات) المجزأة.

العاملون حسب الوضع الوظيفي

جدول رقم (15)

البيان	السنوات		
	2000	**2001**	**2002**
العاملون الدائمون	120	175	255
العاملون بعقود سنوية	130	288	340
العاملون بعقود موسمية	105	90	100
إجمالي	355	553	695

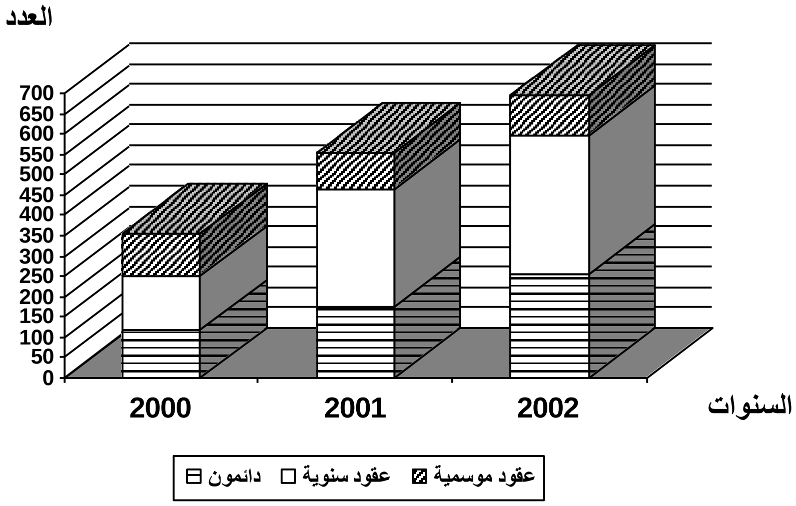
المصدر: فرضي

الحـل : لاحظ هنا أن البيـان يمثـل ظـاهرة واحـدة وهـي عـدد العـاملين فـي كـل سـنة وهـذا العـدد (الظـاهرة) مجزأة إلى مجموعات جزئية .

1 . نرسم أعمدة تناسب ارتفاعاتها مع إجمالي قيمة الظاهرة وهي إجمالي عدد العاملين.

2 . نقوم بتجزئة كل عمود بما يناسب مع عدد العاملين في كل مجموعة جزئية.

والشكل التالي يوضح ذلك:

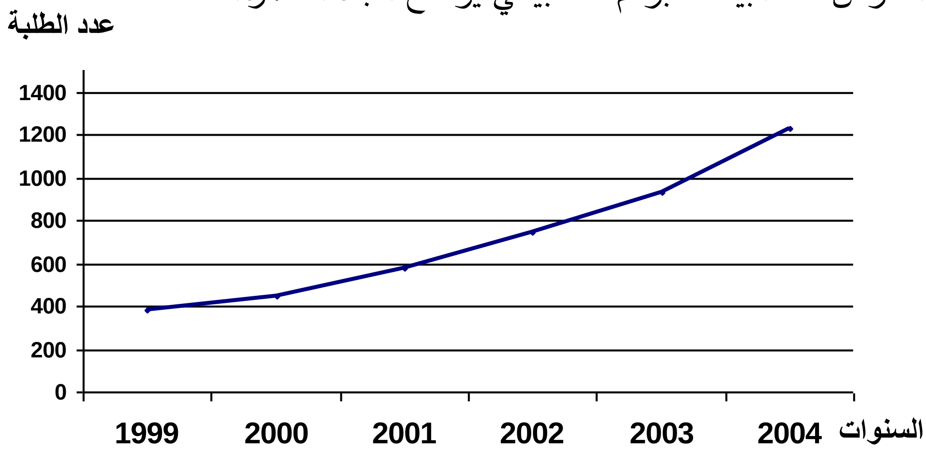


V . الخطوط البيانية: تستخدم في الغالب لمقارنة البيانات المتجانسة عبر الزمن ويمكن تمثيل العلاقة بين ظـاهرتين أو متغيرين x، y بخط بياني من خلال رسم النقاط والوصل بينها لنحصل على خطوط مستقيمة متصلة.

مثال : الجدول التالي يوضح أعداد الطلبة المسجلين في كلية العلوم الإدارية في إحدى الجامعات الخاصة خلال الفترة من 1999 حتى 2004 :

السنة الدراسية	عدد الطلبة
1999	380
2000	445
2001	580
2002	750
2003	930
2004	1230

والمطلوب : عرض تلك البيانات برسم خط بياني يوضح اتجاه الظاهرة.



تطور أعداد الطلبة المسجلين بكلية العلوم الإدارية

كما يمكن أن يشمل الرسم أكثر من خط واحد ، مثال ذلك: عرض البيانات الخاصة بالقمح والشعير على نفس الرسم.

1-7- التوزيعات التكرارية :

تعتبر جداول التوزيع التكراري للبيانات الكمية أساساً لمعظم العمليات الإحصائية الوصفية والتحليلية وجدول التوزيع التكراري بالتعريف هو عبارة عن فئات متسلسلة وما يقابلها من تكرارات. فمثلاً لو استعرضنا جدول التوزيع التكراري التالي :

توزيع المقترضين حسب فئات القرض من أحد المصارف الخاصة (القيمة بالألف وحدة نقدية).

جدول رقم (17)

فئات القرض	45-35	55-45	65-55	75-65	85-75	95-85	105-95	المجموع
عدد المقترضين	3	4	7	14	15	6	1	50

يتضح من الجدول السابق أن القيمة (3) تكررت ثلاث مرات بينما ظهرت القيمة (4) مرات وهكذا ...

وتجدر الإشارة إلى أنه في حالة الظواهر التي تتغير تغيراً متصلاً كأطوال الأشخاص أو أعمارهم فإن قيمتها تنتشر في المدى بين القميتين الصغرى والكبرى في المجموعة بالتدريج بدون أن تتجمع في بعض النقط دون الأخرى لأن المتغير المتصل معناه أن القيم لاتقفز من قيمة لأخرى أكبر منها وإنما بالتدريج بحيث لاتمر بكل القيم المتوسطة

التكرار التجميعي الصاعد والهابط والتكرار النسبي:

يستخدم هذا النوع من التكرارات لمعرفة عدد المفردات في التوزيع الأكبر أو الأقل من قيمة معينة هي الحدود الدنيا أو العليا للفئات التي يتكون منها جداول التوزيع التكراري الأصلي.

. التكرار التجميعي في حالة البيانات المبوبة :

لنعود إلى الجدول السابق والخاص بتوزيع المقترضين حسب فئات القرض فإنه يمكن استنتاج التكرار التجميعي الصاعد والتكرار التجميعي الهابط كالتالي :

جدول التكرار المتجمع الصاعد والهابط

جدول رقم (19)

فئات القروض	التكرار (عدد المقترضين) f	التكرار التجميعي الصاعد		التكرار التجميعي الهابط	
		الحدود	التكرار المتجمع	الحدود	التكرار المتجمع
45-35	3	أقل من 35	صفر	35 فأكثر	50
55-45	4	أقل من 45	3	45 فأكثر	47
65-55	7	أقل من 55	7	55 فأكثر	43
75-65	14	أقل من 65	14	65 فأكثر	36
85-75	15	أقل من 75	28	75 فأكثر	22
95-85	6	أقل من 85	43	85 فأكثر	7
105-95	1	أقل من 95	49	95 فأكثر	1
		أقل من 105	50	105 فأكثر	صفر
المجموع	50				

وبذلك يمكن استخدام الجدول التكراري التجميعي الصاعد لحساب تقدير لعدد المفردات التي تحمل قيمة الظاهرة عند حدود تقل عن قيمة معينة، فمثلاً نلاحظ أن المقترضين الذين تقل قروضهم عن (45) ألف وحدة نقدية هم ثلاثة مقترضين وبذلك تكون نسبتهم إلى إجمالي المقترضين هي 3/50 = 6% من المقترضين .

وكذلك فإن عدد اقترضوا أقل من (60) ألف وحدة نقدية يساوي (14) مقترضاً ونسبتهم = 14/50 = 0.28 أو (28%) من إجمالي المقترضين .

. التكرار النسبي :

يفيد التكرار النسبي في تكوين فكرة استدلالية عن اتجاه الظاهرة ، ويتم تحديد التكرار النسبي الخاص بكل فئة بواسطة قسمة تكرار كل فئة على مجموع التكرارات، $(fc = \frac{f}{\sum f})$

مثال: الجدول التكراري المرفق يوضح فئات الأسهم وعدد المساهمين في إحدى الجمعيات التعاونية.

المطلوب: حساب التكرارات النسبية

جدول رقم (20)

فئات عدد الأسهم	عدد المساهمين	التكرار النسبي
15-10	**5**	**0.125**
20-15	**9**	**0.225**
25-20	**11**	**0.275**
30-25	**5**	**0.125**
35-30	**4**	**0.100**
40-35	**4**	**0.100**
45-40	**2**	**0.05**
المجموع	**40**	**1.000**

من خلال الجدول السابق نجد أن نسبة المساهمين الذين يمتلكون أسهماً مابين 10 إلى 15 سهماً تعادل 0.125 من إجمالي المساهمين، كذلك نجد أن المساهمين الذين يمتلك كل منهم من 15 إلى 20 سهماً تبلغ 0.225 من إجمالي المساهمين أي 22.5% من المساهمين في تلك الجمعية وهكذا.

V . التمثيل الهندسي والبياني للجداول التكرارية :

يمكن تمثيل بيانات الجدول التكراري عن طريق رسم كل من :

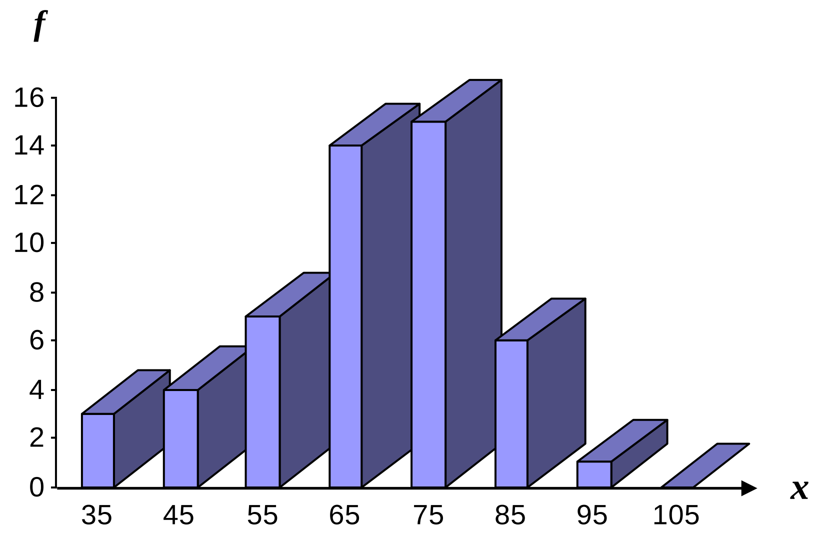
1) المدرج التكراري .

2) المضلع التكراري

3) المنحني التكراري

ولرسم المدرج التكراري يخصص المحور الأفقي لقيم الظاهرة أما المحور العامودي فيخصص للتكرارات المقابلة ولتوضيح ذلك نعود للتمرين السابق والمتعلق بالمقترضين من أحد المصارف الخاصة.

نرسم الشكل التالي ليوضح المدرج التكراري لتوزيع 50 مقترضاً من أحد المصارف الخاصة



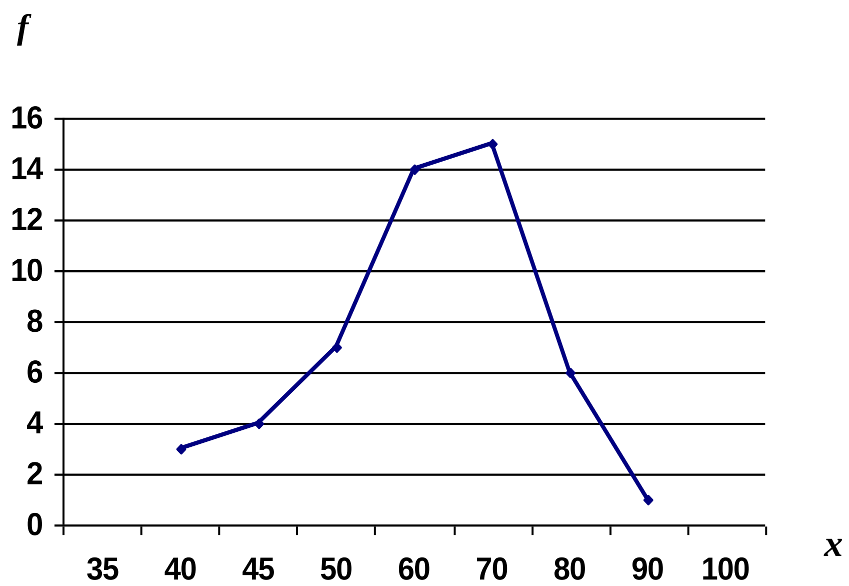
المدرج التكراري لتوزيع المقترضين

ولرسم المضلع التكراري نأخذ مراكز الفئات (مركز الفئة هو نصف مجموع حدي الفئة الأدنى والأعلى)، والتكرارات المناظرة لها، ثم نوصل هذه النقط بخط منكسر فنحصل على المضلع التكراري كما هو موضح في الجدول التالي:

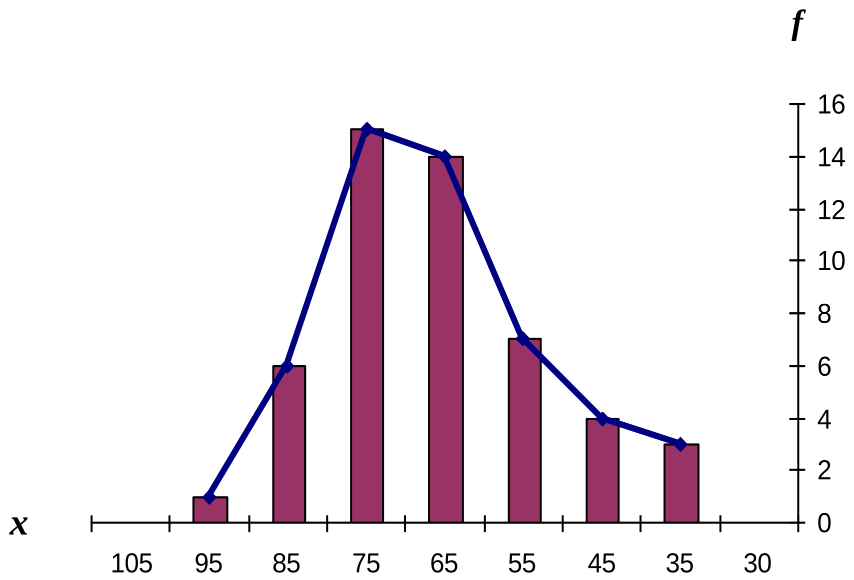
جدول رقم (21)

فئات	45-35	55-45	65-55	75-65	85-75	95-85	105-95
مراكز الفئات	40	50	60	70	80	90	100
التكرارات	3	4	7	14	15	6	1

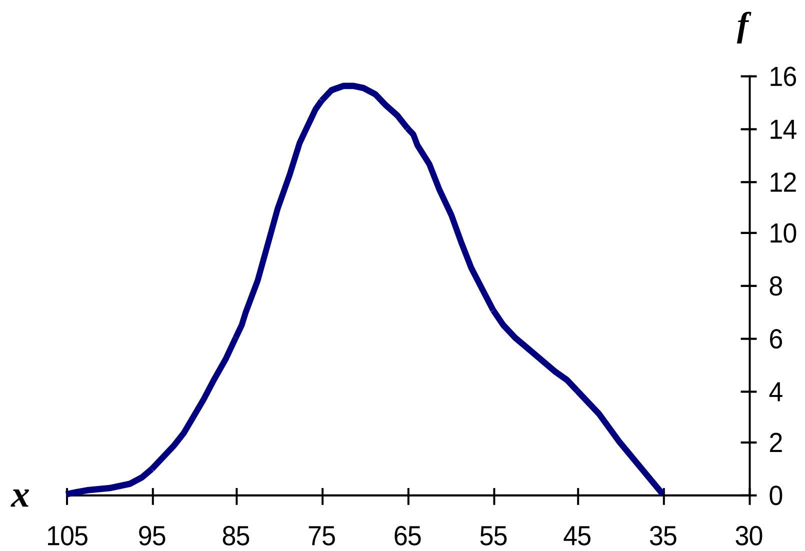
وبذلك نحصل على الشكل التالي للمضلع التكراري:



كما يمكن رسم كل من المدرج التكراري والمضلع التكراري على نفس الشكل :



وبـإجراء عمليـة تهـذيب للمضـلع التكـراري فإنـه يمكـن الحصـول علـى المنحنـي التكـراري حيـث لايشترط هنا أن يمر المنحني بجميع نقط المضلع التكراري .



المنحني التكراري لتوزيع المقترضين

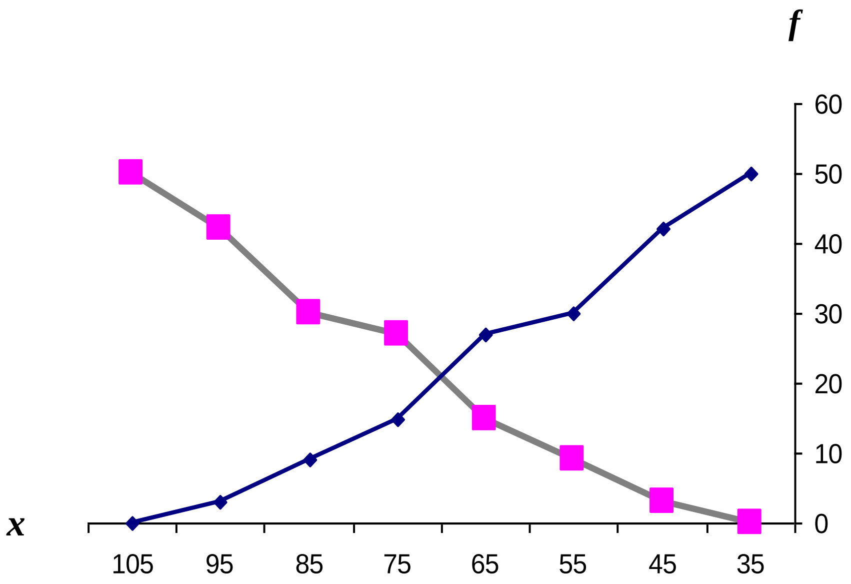
وفي حالة عرض الجدول التكراري في فئات غير منتظمة فإنه لابد من تعديل التكرارات المقابلة لكل فئة حيث أنه كما سبق القول فإن المساحة المخصصة لكل عمود تعبر عن حجم الظاهرة المقابلة لكل فئة من فئات الجدول مع مراعاة تفاوت أطوال القواعد للأعمدة الناتجة.

التمثيل البياني للجدول التكراري التجميعي الصاعد والهابط :

يمكن تمثيل الجدول التكراري التجميعي الصاعد بيانياً بواسطة تحديد نقط المنحنى بنفس الأسلوب السابق شرحه في حالة رسم المنحنى التكراري حيث يخصص المحور الأفقي لحدود قيم الظاهرة والمحور العامودي للتكرارات المتجمعة، ويمكن تمثيل كل من منحنى التكرار التجميعي الصاعد والهابط على نفس الرسم .

مثال :

المطلوب رسم منحنى التكرار التجميعي الصاعد والمنحنى التكراري التجميعي الهابط للمثال الخاص بتوزيع المقترضين من أحد المصارف .



وفي حالة العدد المناسب من التكرارات فإن المنحنيين يلتقيان في نقطة تقابل منتصف التكرارات.

IV . خصائص التوزيعات التكرارية:

1 . تظهر المفردات ميلاً للتمركز حول قيمة معينة والتناقص على جانبي هذه القيمة، تدعى هذه القيمة النزعة المركزية وتقاس بالمقاييس التالية: الوسط الحسابي، الوسيط، المنوال، الوسط الهندسي، الوسط التوافقي.

2 . تبتعد القيم عن نقطة التمركز هذا الاختلاف يدعى التشتت ويقاس بالمقاييس التالية: المدى، الانحراف الربيعي، الانحراف المتوسط، الانحراف المعياري.

3 . تختلف التوزيعات عن بعضها البعض على النحو التالي:

a) . منحني متماثل ينطبق على المنحني الطبيعي

b) . منحني غير متماثل وبالتالي لاينطبق على المنحني الطبيعي ويكون ملتوي نحو اليمين أو ملتوي نحو اليسار .

2) الالتواء Skewness :

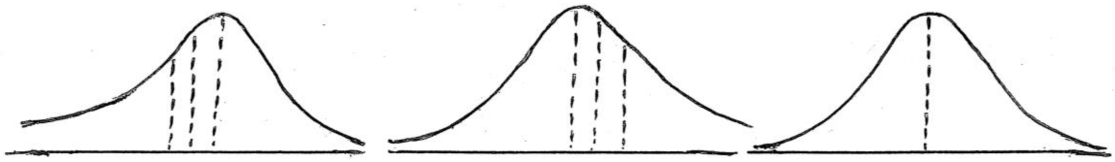
ومعناه مدى تماثل أو درجة التماثل أو البعد عنها فإذا كان الجدول التكراري وزعت التكرارات على عدد مناسب من الفئات مع توافر شرط التجانس فإن تلك القيم تأخذ شكل توزيع متماثل .
وعلى ذلك يمكن التمييز بين :

a) المنحنى التكراري المتماثل Symmetrical Curve

يشبه الجرس المقلوب وله محور تماثل عامودي يمر بنقطة النهاية العظمى له (قمة المنحنى) ويقسمه إلى قسمين متساويين ويتحقق عندما ينطبق على المنحنى الطبيعي.

b) المنحنى التكراري غير المتماثل Skewed Curve :

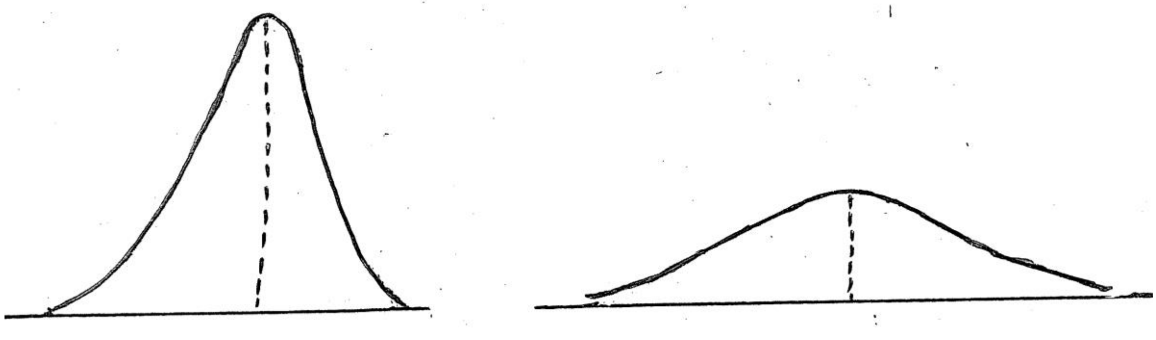
وهو عبارة عن منحنى تكراري لاينطبق على المنحنى الطبيعي فقد يكون ملتوياً نحو اليمين أو ملتوياً نحو اليسار.



منحني تكراري متماثل **منحني تكراري ملتو إلى اليمين** **منحني تكراري ملتو إلى اليسار**

3 . التفرطح Kurtosis :

تختلف المنحنيات التكرارية ذات القمة الواحدة من حيث مدى الاتساع فقد يكون المنحنى التكراري مدبباً حيث تضيق قمته بشكل ملحوظ ويسمى منحنى مدبب Lepto-kurtic .
وقد تتسع قمة المنحنى فيقال له حينئذ منحنى مفرطح Platy Kurtic Curve
ويعرف التفرطح بأنه قياس درجة علو أو انخفاض قمة التوزيع بالنسبة للتوزيع المتماثل.



منحني تكراري مفرطح **منحني تكراري مدبب**

أسئلة وتمارين غير محلولة

1 . ماهي أساليب جمع البيانات الإحصائية وماهي مزايا وعيوب كل أسلوب .

2 . ماهي المصادر الأساسية للبيانات الإحصائية مع الشرح .

3 . أذكر خمسة شروط هامة تجب مراعاتها عند صياغة أسئلة الاستمارة الإحصائية؟

4 . ماهي التصنيفات المختلفة للبيانات الإحصائية؟ أذكر مثالاً لكل تصنيف ؟

5 . ما الفرق بين الإحصاء الوصفي والإحصاء الاستدلالي ؟

6 . ميز بين المفاهيم الآتية : الجدول الإحصائي، المدرج التكراري، التوزيع التكراري.

7 . عرف مع الشرح المصطلحات الإحصائية الآتية :

مجتمع البحث . وحدة البحث . المتغير المتصل والمنقطع . المتغير الكمي، والمتغير الوصفي، الاستبيان، الدراسة التحضيرية، الرقم الاحصائي، النموذج، المتغير التابع.

8 . اذكر مع الشرح نوعين من العينات العشوائية موضحاً ظروف استخدام كل منهما.

9 . اذكر مع الشرح الشروط الهامة التي تجب مراعاتها عند إعداد الجدول الإحصائي.

10 . فيما يلي بيان بتطور المعدل اليومي لأقصى استهلاك من الكهرباء بالمليون واط ساعي في إحدى الدول .

السنة	1996	1997	1998	1999	2000
المعدل	4.5	7	18	24	32

والمطلوب عرض تلك البيانات بالأسلوب الذي تراه مناسباً .

11 . البيان التالي يوضح أعداد الأطباء بإحدى المدن حسب التخصص في عامي (2004, 2000) .

التخصص	عام 2000	عام 2004
بشري	208	310
أسنان	24	86
علاج طبيعي	9	23
المجموع	241	419

والمطلوب عرض تلك البيانات هندسياً باستخدام الدوائر ومرة أخرى باستخدام الأعمدة.

12 . البيان التالي يوضح أعداد المرضى المعالجين المراجعين بإحدى المستشفيات موزعين حسب النوع خلال أربعة أعوام (الأرقام بالألف).

السنة	ذكور	إناث	إجمالي
1998	18	22	40
1999	24	18	42
2000	28	16	44
2001	32	24	56

والمطلوب عرض تلك البيانات بأسلوبين مختلفين .

13 . البيانات التالية توضح أعداد الخريجين من الجامعات السورية خلال الفترة (2002، 2003، 2004).

أعداد الخريجين من الجامعات السورية

السنة / خريج	2002	2003	2004
دكتوراه	79	36	56
ماجستير	548	494	448
الإجمالي	4169	4164	3582

المصدر: المجموعة الإحصائية السورية لعام 2005

والمطلوب عرض تلك البيانات بالأسلوب الذي تراه مناسباً .

14 . البيانات التالية توضح معدل الناتج المحلي الإجمالي في سورية بالأسعار الجارية للفترة (2000 حتى 2004).

السنوات	2000	2001	2002	2003	2004
بالأسعار الجارية	10.4	5.3	6.6	4.4	9.4

والمطلوب : عرض تلك البيانات بأسلوب الأعمدة .

15 . الجدول التالي يوضح توزيع أطوال عينة عشوائية من الطلاب بإحدى الكليات في جامعة دمشق.

فئات الأطوال	165-160	170-165	175-170	180-175	185-180	المجموع
عدد الطلبة	10	20	40	25	5	100

والمطلوب :

1 . رسم المنحنى التكراري الصاعد والهابط في نفس الشكل .

2 . نسبة الطلاب الذين أطوالهم أقل من 170 سم .

3 . نسبة الطلاب الذين أطوالهم أكثر من 180 سم .

4 . عدد الطلاب الذين أطوالهم مابين (165 وَ 170).

الفصل الثاني

التوصيف الكمي للبيانات

. مقدمة

. مقاييس النزعة المركزية (المنوال، الوسيط، الوسط الحسابي، الوسط الهندسي، الوسط التوافقي).

. العلاقة بين المنوال والوسيط والوسط الحسابي .

. مقاييس التشتت (المدى المطلق، نصف المدى الربيعي، الانحراف المتوسط، الانحراف المعياري).

. الالتواء

. التفرطح

. الدرجة المعيارية

. تمارين غير محلولة

التوصيف الكمي للبيانات

1-2 مقدمة :

في الفصل السابق استعرضنا طرق جمع البيانات عن المجتمع موضع الدراسة وكذلك طرق عرض هذه البيانات جدولياً وبيانياً، ورغم القيمة الكبيرة لطرق العرض هذه فإنها عموماً لاتساعدنا إلا في الحصول على انطباع عن الظاهرة المدروسة. ولكن إذا احتاج الأمر إلى صورة أكثر وضوحاً للظاهرة موضع الدراسة نلجأ إلى التوصيف الكمي للبيانات حيث نستخدم مقاييس أو مؤشرات كمية محسوبة من البيانات المتاحة بحيث تظهر أهم الخصائص التي نريد معرفتها عن ظاهرة ما في المجتمع الذي ندرسه. (مقاييس النزعة المركزية، مقاييس التشتت)

2-2- مقاييس النزعة المركزية:

سنكتفي هنا بثلاثة مقاييس للنزعة المركزية تجمع بين الشيوع وبساطة الحساب ووضوح المعنى في كل منها .

- **المنوال Mode :**

يعرف المنوال بأنه القيمة الأكثر شيوعاً (أو الأكثر تكراراً) في البيانات.

ففي مجموعة البيانات التالية :

14 15 23 16 15 19 15 18 19

نجد أن المنوال هو القيمة 15 إذ أنها تظهر ثلاث مرات. ولتحديد قيمة المنوال نضع البيانات في جدول تكراري فيكون المنوال هو القيمة المقابل لأعلى تكرار.

أما إذا كانت البيانات مبوبة في جدول تكراري ذي فئات فإنه لايمكن إيجاد قيمة المنوال بمجرد النظر، إذ لايمكن تحديد أي القيم أكثر شيوعاً في البيانات، إلا أنه يمكن افتراض أن قيمة المنوال هي إحدى قيم المتغير داخل الفئة ذات أكبر تكرار والتي تسمى بالفئة المنوالية كما في المثال التالي :

مثال (2-1) : الجدول التكراري التالي يوضح توزيع عينة عشوائية من 200 عضو هيئة تدريسية تبعاً لفئات الأعمار في إحدى الجامعات التابعة لوزارة التعليم العالي في إحدى الدول.

فئات الأعمار	40-35	45-40	50-45	55-50	60-55	65-60	المجموع
عدد الأعضاء	20	35	80	43	16	6	200

نلاحظ أن أكبر تكرار في هذا الجدول هو 80 وهو المقابل للفئة 45 إلى أقل من 50 وتسمى هذه الفئة بالفئة المنوالية ومن ثم يمكن القول أن المنوال هو أحد قيم هذه الفئة ولأنه من الصعب الوصول إلى القيمة الحقيقية للمنوال فإننا نلجأ إلى تقديرها بأي من الطرق التالية:

I . مركز الفئة المنوالية : أسهل تقريب لقيمة المنوال هو مركز الفئة المنوالية وفي مثالنا هذا يكون المنوال طبقاً لهذه الطريقة هو 47.5 ولكن هذه الطريقة لا تراعي شكل التوزيع وتفضل

عليها طريقة الفروق وطريقة العزوم ونحن سوف نتناول طريقة الفروق باعتبارها الأكثر شيوعاً واستخداماً.

II . طريقة الفروق :

تعتمد طريقة الفروق على التكرارين السابق واللاحق وتكرار الفئة المنوالية كأساس لحساب المنوال وفي هذه الحالة فإن بعد المنوال عن بداية الفئة المنوالية يتناسب مع هذه الفروق 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 .

لو عدنا لبيانات الجدول السابق والمتعلق بأعمار أعضاء هيئة التدريس فإن منوال أعمار هؤلاء الأعضاء هو :

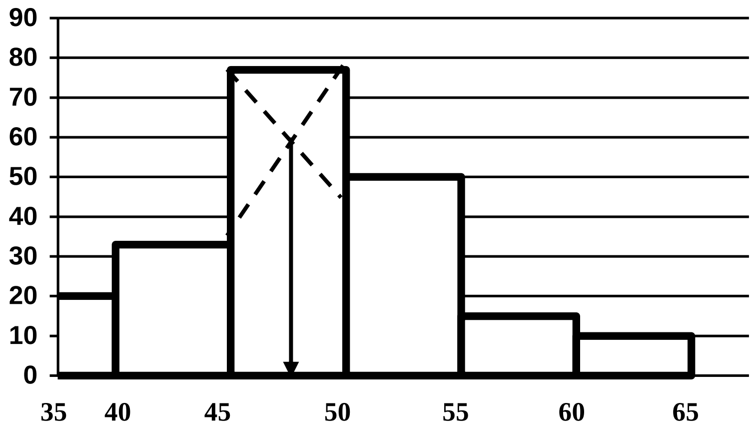
$$\begin{aligned} Mod &= 45 + \left(\frac{80-35}{(80-35)+(80-43)} \right) .5 \\ &= 45 + \left(\frac{45}{45+37} \right) .5 \\ &= 45 + 2.70 = 47.70 \end{aligned}$$

التفسير: إن العمر الأكثر شيوعاً أو تكراراً بين هؤلاء الأعضاء هو 47.70 وبالتقريب (48) سنة تعطينا هذه الطريقة أحسن تقريب لقيمة المنوال ويمكن الوصول إلى نفس القيمة بيانياً كما يلي :

طريقة الرسم :

عند رسم المدرج التكراري نلاحظ أن الفئة المنوالية هي قاعدة أطول عمود وبتوصيل نقطة الحد الأعلى للفئة السابقة بالحد الأعلى للفئة المنوالية والحد الأدنى للفئة المنوالية بالحد الأدنى للفئة اللاحقة لها فإن العمود النازل من نقطة التقاطع يحدد قيمة المنوال.

وباستخدام هذه الطريقة في التمرين السابق نجد أن :



نجد من الرسم أن قيمة المنوال هي 48 سنة تقريباً (وبالدقة تساوي 47.70).

ملاحظة :

عند إيجاد قيمة المنوال من جدول تكراري ذي فئات غير متساوية يجب أولاً التخلص من أثر اختلاف أطوال الفئات أي أنه يجب أولاً الحصول على قيمة التكرارات المعدلة ثم استخدامها لإيجاد المنوال بالطريقة السابقة .

. خصائص المنوال :

1 . لا يتأثر بالقيم المتطرفة .

2 . يمكن حسابه في التوزيعات المفتوحة .

3 . طرق حسابه تقريبية .

• الوسيط Med :

الوسيط هو القيمة التي تتوسط البيانات أي هي القيمة التي يكون عدد المفردات الأصغر منها مساوياً لعدد المفردات الأكبر منها. وعلى ذلك فإنه لإيجاد قيمة الوسيط يجب أولاً ترتيب البيانات ترتيباً تصاعدياً أو تنازلياً ثم نحدد مكان المفردة الوسيطية ويسمى ترتيب الوسيط. ثم بعد ذلك نوجد قيمة الوسيط.

لإيجاد الوسيط من مجموعة البيانات التالية :

14 15 23 16 15 19 15 18 19

نبدأ بترتيبها تصاعدياً أي :

14 15 15 15 16 18 18 19 23

ويلاحظ هنا أن ترتيب الوسيط هو (5) حيث أن هناك (4) قيم أصغر منها و (4) قيم أكبر منها وبالتالي فإن قيمة الوسيط هي قيمة المفردة رقم (5) في البيانات المرتبة أي أن: الوسيط=16.

يلاحظ في هذا المثال أن عدد المفردات = 9 وأن ترتيب الوسيط وهو (5) يمكن الحصول عليه

على الصورة $\frac{1+9}{2}$ وإذا حذفت القيمة الأخيرة وهي (23) من البيانات أصبح عدد المفردات (8) ونجد أنه في هذه الحالة يقع الوسيط بين قيمة المفردتين رقم 5,4 إذ أن ترتيب الوسيط في هذه الحال $\frac{1+8}{2} = (4.5)$ أي بين القيمتين (16,15) فيكون الوسيط $\frac{16+15}{2} = (15.5)$ ويمكن تلخيص خطوات حساب الوسيط في الخطوات الثلاث التالية:

1 . ترتيب القيم تصاعدياً أو تنازلياً.

2 . إيجاد ترتيب الوسيط = $\frac{n+1}{2}$ حيث أن n عدد المفردات .

3 . إيجاد قيمة الوسيط وهو قيمة المفردة رقم = $\frac{n+1}{2}$

ويلاحظ أنه إذا كان عدد المفردات كبيراً فإنه يمكن اعتبار ترتيب الوسيط هو :

$$\frac{\text{عدد المفردات}}{2} \text{ أي } = \frac{n}{2} = \frac{\sum f}{2}$$

. حساب الوسيط لبيانات مبوبة (جدول تكراري)

ويتم على النحو التالي :

1 . تكوين الجدول المتجمع الصاعد (أو الهابط).

2 . إيجاد ترتيب الوسيط وهو = $\frac{\sum f}{2}$ ثم البحث عن القيمة $\sum f/2$ وما يساويها أو أكبر منها مباشرة لتحديد الفئة الوسيطية.

3 . إيجاد قيمة الوسيط وذلك بطريقتي الحساب والرسم.

I . بطريقة الحساب :

بعد تكوين جدول التكرارات المتجمعة الصاعدة (أو الهابطة) ومن ثم ترتيب الوسيط من خلال $f/2$ ،نبحث عن القيمة $\sum f/2$ أو أكبر منها مباشرة في عمود التكرارات المتجمعة الصاعدة والفئة المقابلة لها تكون الفئة الوسيطية وأخيراً نطبق العلاقة التالية :

$$Med = L_1 + \left(\frac{\sum f/2 - \sum f_1}{f_{me}} \right) . C$$

حيث أن :

L_1 : الحد الأدنى للفئة الوسيطية .

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

f_{me} : تكرار الفئة الوسيطية

C : طول الفئة الوسيطية .

وبالعودة إلى التمرين السابق والمتعلق بأعمار أعضاء هيئة التدريس في إحدى الجامعات نجد أن

وسيط أعمار هؤلاء الأعضاء هو :

$$\sum f/_{2} = \frac{200}{2} = 100$$

$$Med = 45 + \left(\frac{100-55}{80}\right).5$$

$$= 45 + 2.81 = 47.81$$

الفئات	f	ت.ت.ص
40-35	20	20
45-40	35	55
50-45	80	135
55-50	43	178
60-55	16	194
65-60	6	200

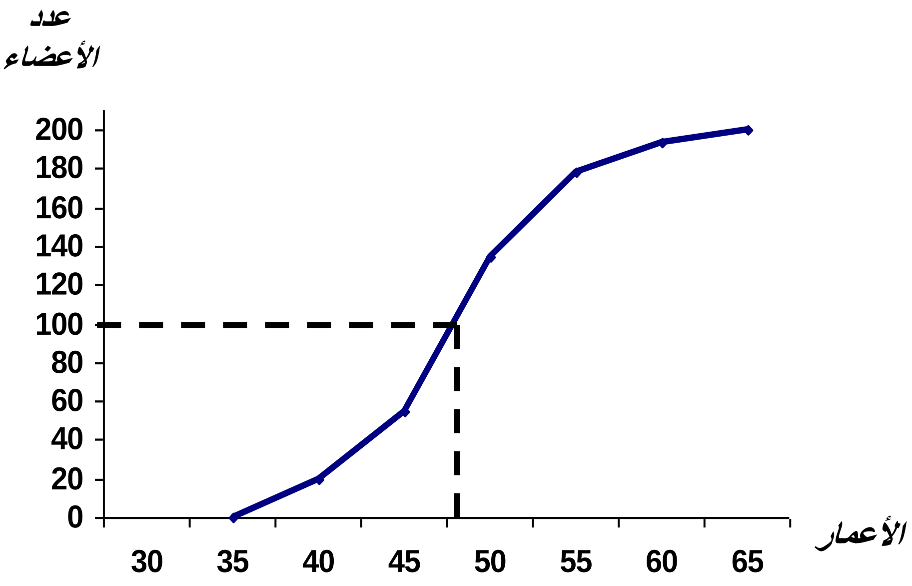
التفسير : إن نصف الأعضاء كانت أعمارهم أقل من (47.81) سنة وإن نصف الأعضاء الآخر كانت أعمارهم أكبر من هذا العمر .

II . طريقة الرسم :

نرسم المنحنى المتجمع الصاعد (أو الهابط) ثم نوجد القيمة المقابلة لترتيب الوسيط من الرسم فتكون هي قيمة الوسيط.

مثال : أوجد الوسيط لبيانات الجدول السابق والمتعلق بأعمار أعضاء هيئة التدريس إحدى الجامعات .

حدودا لفئات	التكرار المتجمع الصاعد
أقل من 35	.
أقل من 40	20
أقل من 45	55
أقل من 50	135
أقل من 55	178
أقل من 60	194
أقل من 65	200



ملاحظة : يمكن تحديد قيمة الوسيط بيانياً أيضاً من خلال رسم منحني التكرار التجميعي الهابط أو الاثنين معاً (الصاعد والهابط) نقطة التقاطع اسقاطها على المحور الأفقي هي قيمة الوسيط .

ثم نوجد ترتيب الوسيط من العلاقة = $\frac{\sum f}{2}$ فيكون ترتيب الوسيط $\frac{200}{2} = 100$

من الرسم نوجد القيمة المقابلة لتكرارات 100 فتكون قيمة الوسيط هي (47.81) سنة.

● **خصائص الوسيط :**

1 . يمكن حسابه من الجداول التكرارية المفتوحة .

2 . لا يتأثر بالقيم المتطرفة.

3 . قيمته غير ثابتة عندما يكون حجم العينة صغيراً.

● **الوسط الحسابي :**

يعرف الوسط الحسابي لمجموعة من القيم بأنه القيمة التي لو أعطيت لكل مفردة فإن مجموعها لايتغير ، أي أن: (عدد المفردات) * (الوسط الحسابي) = مجموع القيم

وبصورة عامة إذا كان المتغير x يأخذ القيم x_1, x_2, x_3, x_n فإن الوسط الحسابي لهذه القيم يرمز له بالرمز $\bar{x}$ ويعبر عنه على أنه :

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots\dots x_n}{n}$$

ويمكن كتابة ذلك في صورة مختصرة :

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

(حيث n عدد قيم المتغير x)

مثال : أوجد الوسط الحسابي لأجور (5) عمال كانت أجورهم الاسبوعية على النحو التالي:

2150 2520 1950 1875 2125 (ليرة سورية)

الأسلوب المباشرة :

ويتم من خلال تطبيق العلاقة التالية : $\bar{x} = \dfrac{\sum x}{n}$

$$= \frac{1}{5}(2150 + 2520 + ...) = \frac{10620}{5} = 2124$$

أي الوسط الحسابي لأجر العامل = 2124 ليرة.

. حساب الوسط الحسابي لبيانات مبوبة (جدول تكراري):

مثال :

أحسب الوسط الحسابي لأجر العامل إذا علمت أن الأجور الأسبوعية لعينة عشوائية من 200 عامل من عمال إحدى الصناعات كانت وفق الجدول التالي (بمئات الوحدات النقدية):

فئات الأجر الأسبوعي	40-20	60-40	80-60	100-80	120-100	140-120	المجموع
عدد العمال	9	22	68	64	27	10	200

الحل :

من تعريف الوسط الحسابي نعلم أن :

الوسط الحسابي (للأجر الأسبوعي للعامل) = $\dfrac{\text{جملة الأجور الأسبوعية}}{\text{عدد العمال}}$ $\bar{X} = \dfrac{\sum fx}{\sum f}$

حيث أن (x) مركز الفئات (f) التكرارات.

من الجدول نلاحظ أن هناك (9) عمال كل منهم يحصل على أجر يتراوح بين (20)، (40) وحدة نقدية أي أن كل منهم يحصل في المتوسط على أجر قدره (30) وحدة $\dfrac{40+20}{2} = 30$ (وهو مركز الفئة الأولى) وبالتالي فإن أجور العمال في هذه الفئة الأولى هي (30)(90)=270 وحدة نقدية كذلك نجد أن هناك (22) عاملاً يحصل كل منهم على أجر يتراوح بين (40),(60) وحدة أي أن أجر العامل منهم هو (50) في المتوسط $\dfrac{60+40}{2} = 50$ أي مركز (الفئة الثانية) فيكون بذلك أجور العمال في الفئة الثانية هو (22)(50)=1100.

وبالمثل يمكن إيجاد أجور العمال في باقي فئات الجدول، وبجمع هذه الأجور نحصل على جملة الأجور. ويمكن ترتيب هذه الحسابات كما في الجدول التالي :

فئات الأجر	عدد العمال (f)	مركز الفئات (x)	fx
40 – 20	9	30	270

1100	**50**	**22**	**60 – 40**
4760	**70**	**68**	**80 – 60**
5760	**90**	**64**	**100-80**
2970	**110**	**27**	**120-100**
1300	**130**	**10**	**140-120**
16160		**200**	**المجموع**

ويكون الوسط الحسابي لأجر العامل $\frac{16160}{200} = 80.8$ وحدة نقدية أي أن أجور هؤلاء العمال بالمتوسط هو (80,8) وحدة نقدية.

ملاحظة : يلاحظ أنه لإيجاد قيمة الوسط الحسابي في حالة جدول تكراري ذي فئات غير متساوية فإننا نتبع نفس الطريقة كما في الأمثلة السابقة (الفئات المتساوية) دون أي اختلاف.

. الوسط الحسابي المرجح :

في بعض الأحيان يكون من المهم ترجيح بعض القيم $x_1, x_2, \ldots. x_n$ بأوزان $w_1, w_2, \ldots\ w_n$ وهذه تعتمد على الأهمية المرتبطة بهذه القيم في الدراسة ويعرف الوسط الحسابي هنا بالوسط الحسابي المرجح ويحسب بالعلاقة التالية:.

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \ldots\ldots\ldots x_n wn}{w_1 + w_2 + \ldots\ldots w_n} = \frac{\sum xw}{\sum w}$$

مثال :

إذا أعطى الامتحان النهائي في مقرر ما وزناً يساوي ثلاثة أمثال الامتحانات الشهرية، وإذا حصل طالب في الامتحان النهائي على 85 درجة وفي الامتحانات الشهرية على 90,70 درجة فإن متوسط درجاته يكون :

$$\bar{x} = \frac{(70)(1) + (90)(1) + (85)(3)}{1 + 1 + 3} = 83$$

● خصائص الوسط الحسابي :

1 . يتأثر بالقيم المتطرفة على جانبي التوزيع.

2 . صعوبة حسابه من جدول مفتوح .

3 . معيار للمقارنة.

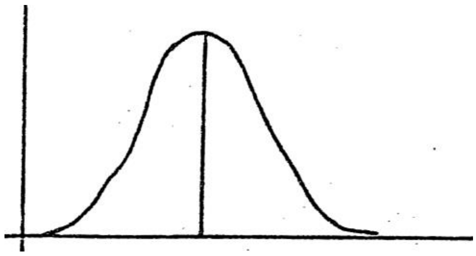
3-2 : العلاقة بين المنوال والوسيط والوسط الحسابي :

يلاحظ أنه في التوزيعات التكرارية المتماثلة (الشكل المرفق) تتطابق قيم كل من المنوال والوسيط والوسط الحسابي، ومن الناحية التطبيقية يندر وجود التوزيعات المتماثلة، فقد نحصل على توزيعات ملتوية ناحية اليمين أو ناحية اليسار، وفي التوزيعات التكرارية وحيدة المنوال المتماثلة والقريبة من التماثل تحقق العلاقة التجريبية التالية:

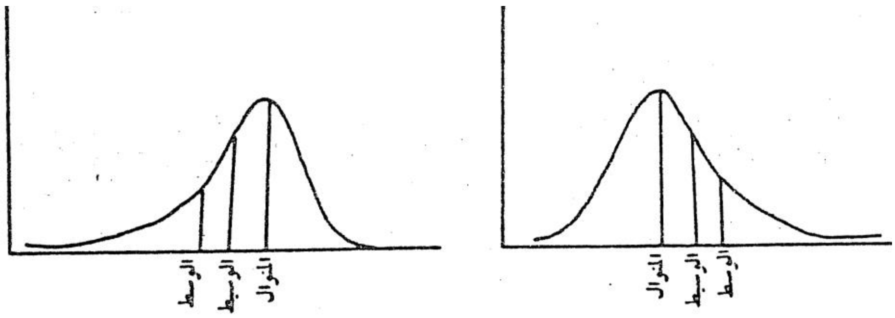
$$\bar{x} - Mod = 3(\bar{x} - Med)$$

أما في التوزيعات التكرارية الملتوية التواء شديداً فلا تتحقق هذه العلاقة.

الأشكال المرفقة توضح الموضع النسبي للوسط والوسيط والمنوال للمنحنيات التكرارية الملتوية إلى اليمين والملتوية إلى اليسار على الترتيب.



المنوال = الوسيط = الوسط الحسابي



في المنحنيات المتماثلة يتطابق الوسط والوسيط والمنوال أي أن تطابق المقاييس الثلاثة دليل على تماثل التوزيع كما أن التباعد بين هذه المقاييس يعتبر دليلاً على عدم التماثل لذا فإن هذه الفروق يمكن أن تستخدم لقياس الالتواء كما سنوضح فيما بعد.

- **الوسط الهندسي :**

يعرف الوسط الهندسي لمجموعة من القيم $x_1, x_2.....x_n$ على أنه الجذر النوني لحاصل ضرب هذه القيم .

$$\text{الوسط الهندسي} = \sqrt[n]{x_1.x_2.x_3.......x_n}$$

أغلب استخداماته في حساب معدلات النمو ومعدلات الفائدة وحساب الأرقام القياسية .

- **الوسط التوافقي :**

الوسط التوافقي لمجموعة من القيم هو مقلوب الوسط الحسابي لمقلوبات هذه القيم أي أن:

$$\text{الوسط التوافقي} = \frac{n}{\sum 1/x}$$

أغلب استخداماته في مجال انتاجية العمل والأسعار والعلاقة الثلاثية (السرعة، الزمن، المسافة)

4-1- مقاييس التشتت :

لقد استعرضنا كيفية تحليل البيانات الإحصائية حسابياً عن طريق المتوسطات وجدير لكن هذه

المتوسطات (مقاييس الموضع) لاتعطي وصفاً كاملاً للبيانات إذ أنها تبين القيمة التي تتركز حولها بقية قيم العينة فقط ولا تعطي أي إيضاح عن مدى تباعد أو تقارب باقي القيم من هذه القيمة أي أنها لاتوضح مدى التجانس في البيانات أو مدى تشتتها لذا فإننا نحتاج إلى مقياس آخر يوضح مدى تقارب أو تباعد البيانات وهذه المقاييس تسمى بمقاييس التشتت وهي: المدى المطلق، نصف المدى الربيعي . الانحراف المتوسط . الانحراف المعياري.

- **المدى المطلق (R) :**

هو الفرق بين أكبر وأصغر قيمة في مجموعة البيانات.

ومن عيوب هذا المقياس أنه يعتمد على قيمتين فقط من البيانات ويتجاهل باقي القيم وأغلب استخدامات هذا المقياس في وصف الأحوال الجوية وقياس تشتت الأسعار اليومية للسندات والأسهم والأوراق الحالية ومراقبة جودة الإنتاج .

- **نصف المدى الربيعي (Q.D) :**

يعرف على أنه نصف الفرق بين أكبر وأصغر قيمة من القيم المتبقية بعد استبعاد الربع الأدنى والربع الأعلى من البابنات (الفرق بين الربيع الثالث والربيع الأول مقسوماً على (2)).

وللحصول على نصف المدى الربيعي نتبع الخطوات التالية :

(1) ترتيب البيانات ترتيباً تصاعدياً (أو تنازلياً).

(2) تحديد ترتيب (موضع) القيمة التي تقع في آخر الربع الأول

لبيانات مبوبة $Q_1 = \frac{\sum f}{4}$. لبيانات غير مبوبة $K_1 = \frac{n+1}{4}$

(3) إيجاد قيمة Q_1 بالرسم أو بالحساب مع ملاحظة أن Q_2 = الوسيط

(4) تحديد ترتيب القيمة التي تقع في آخر الربع الثالث

لبيانات مبوبة $K_3 = \frac{\sum f}{4}$. لبيانات غير مبوبة $K_3 = \frac{3(n+1)}{u}$

(5) نوجد قيمة Q_3 من الرسم أو الحساب (بنفس طريقة إيجاد الوسيط).

(6) $Q.D = \frac{Q_3 - Q_1}{2}$ نصف المدى الربيعي

حساب المدى الربيعي لبيانات مبوبة :

إن حساب الربيعيات لبيانات مبوبة مشابهة تماماً لطريقة حساب الوسيط ويتم ذلك على النحو التالي :

$$Q_1 = L_1 + \left(\frac{\sum f/4 - \sum f_1}{f_{Q1}} \right).C$$ الربيع الأول

حيث أن :

L_1 : تشير إلى الحد الأدنى لفئة الربيع الأول .

$f/4$: ربع مجموعة التكرارات (ترتيب الربيع الأول).

$\sum f_1$: مجموع التكرارات السابقة لتكرار فئة الربيع الأول .

f_{Q1} : تكرار فئة الربيع الأول .

$$Q_3 = L_3 + \left(\frac{3\sum f/4 - \sum f_1}{f_{Q3}} \right).C$$ الربيع الثالث

حيث أن :

L_3 : تشير إلى الحد الأدنى لفئة الربيع الثالث .

$3\sum f/4$: ترتيب الربيع الثالث .

$\sum f_1$: مجموع التكرارات السابقة لتكرار فئة الربيع الثالث.

f_{Q3} : تكرار فئة الربيع الثالث .

وبالتطبيق على المثال السابق المتعلق بأعمار أعضاء هيئة التدريس في إحدى الجامعات (بحث الوسيط) نجد أن:

$$\sum f/4 = \frac{200}{4} = 50$$

$$3\sum f/4 = \frac{3(200)}{4} = 150$$

ت.ت.ص	f	الفئات
20	**20**	**40-35**
55	**30**	**45-40**
135	**80**	**50-45**
178	**43**	**55-50**
194	**16**	**60-55**
200	**6**	**65-60**

الربيع الأول :

$$Q_1 = 40 + \left(\frac{50-20}{30}\right).5$$

$$40 + 5 = 45$$

التفسير : أن 25% من أعضاء هذه الجامعة أعمارهم كانت أقل (45) سنة وأن 75% من هؤلاء العمال كانت أعمارهم أكبر من (45) سنة.

الربيع الثالث :

$$Q_3 = 50 + \left(\frac{150-135}{43}\right).5$$

$$= 50 + 1.70 = 51.70$$

التفسير : أن 75% من أعضاء هذه الجامعة أعمارهم كانت أقل من (51.70) سنة وأن 25% من هؤلاء العمال كانت أعمارهم أكبر من (51.70) سنة .

نصف المدى الربيعي :

$$Q.D = \frac{51.7-45}{2} = 3.35$$

. خصائص الانحراف الربيعي (نصف المدى الربيعي) :

1 . يمكن حسابه من التوزيعات التكرارية المفتوحة

2 . مقياس النزعة المركزية المقابل له هو الوسيط .

. الانحراف الربيعي النسبي (Q.D%) :

يعتبر من مقاييس التشتت النسبية له أهمية في التحليل الاحصائي لمقارنة تشتت توزيعين مختلفين، ويعطى بالعلاقة التالية :

$$Q.D\% = \frac{Q.D}{Med}.100$$

أو

$$Q.D\% = \frac{Q_3.Q_1}{2Med}.100$$

وبالتطبيق على التمرين السابق والمتعلق بأعمار عينة عشوائية من أعضاء هيئة التدريس إحدى الجامعات نجد أن :

$$Med= 47.81 \text{ , } Q_1= 45 \text{ , } Q_3 = 51.70 \text{ , } Q.D = 3.35$$

$$Q.D\% = \frac{3.35}{47.81}.100 = 7\%$$

وبفرض أخذت عينة عشوائية من حجم مماثل من الأعضاء في جامعة أخرى وكان الوسيط للأعمار (40) سنة وبانحراف ربيعي مقداره (4) سنة والمطلوب :

أيهما يعرض تشتتاً أكبر أعمار الأعضاء في الجامعة الأولى أم أعمار الأعضاء في الجامعة الثانية؟ برر إجابتك إحصائياً. $Q.D\% = \frac{4}{40}.100 = 10\%$ (الجامعة الثانية)

من خلال المقارنة نجد أن Q.D% (10%) > Q.D% (7%)

وبالتالي يمكن القول أن عينة أعضاء الجامعة الثانية تعرض تشتتاً أكبر من أعمار عينة أعضاء الجامعة الأولى .

- **الانحراف المتوسط أو متوسط الانحرافات (M.D) :**

يعرف الانحراف المتوسط لمجموعة القيم $x_1, x_2, \ldots. x_n$ كالتالي:

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

أي أنه متوسط القيم المطلقة لانحرافات القيم عن وسطها الحسابي (القيمة المطلقة لرقم هي الرقم بإشارة موجبة ويعبر عنها بخطين حول الرقم فتكون القيمة المطلقة للرقم)

مثال : أوجد الانحراف المتوسط لمجموعة القيم :

2 3 6 8 11

الحل :

$$x = \frac{2+3+6+8+11}{5} = 6$$ الوسط الحسابي

$$\frac{1}{5}|2-6|+|3-6|+\ldots\ldots.. = \frac{4+3+2+5}{5} = 2.8$$ =الانحراف المتوسط

إذا كانت القيم $x_1, x_2, \ldots. x_n$ تحدث بتكرارات $f_1, f_2, \ldots. f_n$ فالانحراف المتوسط يمكن كتابته على صورة

$$M.D = \frac{1}{n}\sum f\,|x - \bar{x}|$$ الانحراف المتوسط

- **الانحراف المعياري (S) :**

هو الجذر التربيعي لخارج قسمة مربعات انحرافات القيم عن وسطها الحسابي على عدد القيم.

ويعتبر الوسط الحسابي والانحراف المعياري في أكثر المقاييس الاحصائية أهمية في التحليل الاحصائي ويعتبران الأساس في النظرية الاحصائية على اعتبار أن الوسط الحسابي يحدد لنا مركز التوزيع الطبيعي والانحراف المعياري يحدد لنا كيفية الانتشار على جانبي الوسط الحسابي.

- **حساب الانحراف المعياري لبيانات غير مبوبة :**

. الأسلوب المباشر :

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

حيث أن : $\bar{x} = \frac{\sum x}{n}$

ويمكن تبسيط العلاقة السابقة فنحصل على الصيغة التالية :

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

. حساب الانحراف المعياري لبيانات مبوبة (جدول تكراري):

1 . الأسلوب المباشر :

ويتم من خلال العلاقة التالية :

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

2 . أسلوب استخدام مراكز الفئات :

$$S = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

مع العلم أن جميع هذه الأساليب لحساب الانحراف المعياري تعطي نفس النتائج وبالتطبيق على المثال في بحث الوسط الحسابي والمتعلق بأجور العمال في إحدى الصناعات نجد أن:

فئات الأجر	عدد العمال	مراكز الفئات (x)	fx	fx^2
40-20	9	30	270	8100
60-40	22	50	1100	55000
80-60	68	70	4760	333200
100-80	64	90	5760	518400
120-100	27	110	2970	326700
140-120	10	130	1300	169000
المجموع	200		16160	1410400

$$S = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$$

$$S=\sqrt{\frac{1410400}{200}-(80.8)^2}=22.87$$

- **خصائص الانحراف المعياري :**

1. لايمكن حسابه من جداول تكرارية مفتوحة .
2. يتأثر بضرب وقسمة المفردات ولا يتأثر بالجمع والطرح.
3. مقياس النزعة المركزية المقابل له هو الوسط الحسابي.

- **معامل الاختلاف (الانحراف المعياري النسبي) (C.D%)**

نحن نعلم أن مقياس التشتت يستخدم للمقارنة بين مدى تقارب البيانات من مقاييس النزعة لها، وجدير بالملاحظة هنا أنه إذا تساوت مقاييس النزعة للمجتمعات المختلفة فإنه يمكننا مقارنة درجات التجانس في هذه المجتمعات عن طريق مقاييس التشتت. أما إذا اختلف مقياس النزعة أو وحدة القياس المستخدمة من مجتمع لآخر فإننا لانستطيع مقارنة التجانس عن طريق مقاييس التشتت مباشرة بل لابد وأن نلجأ إلى تعديلها أو إلى مقياس آخر نتلافى فيه هذا الاختلاف، هذا المقياس يسمى بمقياس (أو معامل) الاختلاف ويعرف كالتالي :

$$C.V\%=\frac{S}{\bar{x}}\times 100$$

يستخدم هذا المقياس بشكل واسع في التحليل الإحصائي لمقارنة تشتت توزيعين مختلفين ولتوضيح ذلك ندرس الأمثلة التالية:

مثال (1) :

إذا كان الانحراف المعياري لدرجات مجموعة من الطلبة في مادة (A) هو (10) ومتوسط الدرجات (50) درجة وفي المادة (B) كان متوسط الدرجات (30) بانحراف معياري (6) درجات فهل نستطيع الحكم على أن مستوى الطلبة في المادة (A) أكثر تشتتاً منه في المادة (B) أم لا؟ ولماذا ؟ .

الحل : من الطبيعي أننا لا نستطيع الحكم مباشرة بذلك نظراً لاختلاف متوسط الدرجات ولذا لا نستطيع المقارنة مباشرة بين الرقمين (10)، (6) ولابد من حساب معامل الاختلاف فيكون:

معامل اختلاف المادة (A) = $\frac{10}{50}\times 100 = 20\%$

معامل اختلاف المادة (B) = $\frac{6}{30}\times 100 = 20\%$

وعلى ذلك يكون التجانس في مستوى الطلبة في المادتين من نفس الدرجة ويمكننا أيضاً استنتاج أن مستوى الطلبة في المادة (A) أفضل منه في المادة (B) نظراً لارتفاع متوسط درجاتهم في المادة (A) عنها في المادة (B).

2-5- الالتواء :

يعرف الالتواء على أنه درجة التماثل أو البعد عن التماثل للتوزيع، ويقاس بالعلاقة التالية:

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

وفي حالة استخدام الوسيط تصبح العلاقة:

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

وإذا كان معامل الالتواء مساوياً صفراً كان التوزيع متماثلاً بينما إذا كان : SK>0 هذا يعني أن التوزيع ملتوياً نحو اليمين وبالعكس إذا كان : SK < 0 فإن التوزيع ملتوي نحو اليسار.

مثال :

أوجد معامل بيرسون للالتواء باستخدام المنوال ومرة أخرى باستخدام الوسيط لأجور مجموعة من العمـال إذا علمـت أن متوسـط الأجـر (79.76) الوسـيط = (79.06) الانحـراف المعيـاري = (15.6) والمنوال = (77.5)

الحل :

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

$$0.14 = \frac{2.26}{15.6} = \frac{77.5 - 79.76}{15.6} = \text{ ، } SK = \frac{3(\bar{x} - Med)}{S}$$

المعامل الثاني للالتواء = $\frac{(79.06 - 79.76)3}{15.6}$

$$0.13 = \frac{2.1}{15.6} = \frac{(0.7)3}{15.6} =$$

ونذكر هنا أن قيمة معامل الالتواء قد تختلف بحسب الأسلوب المستخدم لحسابه وعلى البيانات المتوفرة.

. معامل الالتواء الربيعي :

هناك مقاييس أخرى للالتواء منها معامل الالتواء الربيعي ويعرف كالآتي:

معامل الالتواء الربيعي = $\frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1}$

أوجد معامل الالتواء الربيعي إذا علمت أن $Q_1=44$ ، $Q_2=48$ ، $Q_3=52$

الحل :

معامل الالتواء الربيعي = $\frac{(52 - 48) - (48 - 44)}{(52 - 44)} = 0$

أي أن التوزيع متماثل في هذه الحالة.

6-2- التفرطح :

يعرف بأنه مقياس لقياس درجة علو أو انخفاض قمة التوزيع بالنسبة للتوزيع الطبيعي ويعطى بالعلاقة الآتية : $K = \frac{m^4}{S^4}$

حيث أن : m^4 : العزم الرابع حول الوسط الحسابي .

S^4 : مربع التباين .

وإذا كان $(k=3)$ كان التوزيع متماثلاً بينما إذا كان $(k<3)$ كان التوزيع منبسط أما إذا كان $(k>3)$ فإن التوزيع مدبب.

- **الدرجة المعيارية :**

تعرّف الدرجة المعيارية بأنها النسبة الناتجة في قسمة انحراف القيمة الخام عن المتوسط الحسابي للتوزيع التكراري مقسوماً على الانحراف المعياري وتعطى بالعلاقة التالية:

$$Z = \frac{|x - \bar{x}|}{s}$$

ويمكن القول أن الدرجة المعيارية عبارة عن وحدة لقياس الانحراف المعياري عن الوسط الحسابي وتستخدم الدرجات المعيارية المقدرة بوحدات الانحرافات المعيارية لمقارنة قيمتين مأخوذتين من مجموعتين احصائيتين مختلفتين ولا يجوز مقارنة القيمتين على أساس القيم الخام .

تمرين : كانت درجتا طالبين في أحد المقررات في مجموعتين مختلفتين على النحو التالي:

الطالب الثاني (B)	الطالب الأول (A)	المقاييس
64	86	درجة الطالب
58	75	متوسط درجات الطلاب
4	10	الانحراف المعياري

والمطلوب : أيهما كان تحصيله أفضل بالنسبة لمستوى الطلاب ؟

$$Z = \frac{|x_i - \bar{x}|}{s}$$

$$Z_{(A)} = \frac{|x_i - \bar{x}|}{s} = \frac{|86 - 75|}{10} = 1.1$$

$$Z_{(B)} = \frac{|x - \bar{x}|}{s} = \frac{|64 - 58|}{4} = 1.5$$

$$Z_{(B)} > Z_{(A)}$$

وبالتالي فإن تحصيل الطالب (B) أفضل من تحصيل الطالب (A) على الرغم أنه قبل التحويل إلى درجات معيارية كانت درجات الطالب (A) أفضل من درجات الطالب (B)

تمارين غير محلولة

1 . أخذت عينة عشوائية مكونة من (200) مودع من المودعين لدى أحد المصارف الخاصة في مدينة دمشق فكانت الودائع موزعة في الجدول التكراري المرفق (القيم بالمليون) بتاريخ 2004/12/31 :

فئات الودائع	10-0	20-10	40-20	50-40	80-50	المجموع
عدد المودعين	7	42	75	60	16	200

المطلوب :

1 . حساب الوسط الحسابي والانحراف المعياري ومعامل الاختلاف لودائع هؤلاء المودعين.

2 . حساب المنوال للودائع بطريقة الفروق ثم استخدم النتائج في حساب معامل الالتواء للودائع .

3 . بفرض أن توزيع أرصدة ودائع المصرف خاضعة للتوزيع الطبيعي، وأن متوسط رصيد المودع فيها كان مساوياً (11) مليون وانحراف معياري (1.5) مليون والمطلوب ما نسبة أرصدة المودعين الذين أرصدتهم (12.5 وَ 14).

2 . أخذت عينة عشوائية ملونة من (200) موظف في احدى الدوائر الحكومية في مدينة دمشق، فكانت فئات الأعمار موزعة وفق الجدول التكراري المرفق :

الأعمار بالسنة	25-15	35-25	45-35	55-45	65-55	المجموع
عدد الموظفين	12	14	100	48	26	200

والمطلوب :

1 . حساب منوال العمر.

2 . باستخدام المنوال أحسب معامل الالتواء للتوزيع السابق.

3 . إذا كان الوسط الحسابي لأطوال التلاميذ في المدرسة (A) الابتدائية هو (120)سم والمنوال (117) سم بانحراف معياري (7) سم والوسط الحسابي لأطوال التلاميذ في المدرسة (B) المتوسطة هو (140) سم والمنوال (142) سم بانحراف معياري (12)سم.

1 . قارن بين تشتت الأطوال في المدرستين.

2 . حدد التواء المنحني التكراري في كل مدرسة .

3 . فيما يلي توزيع تكراري لعينة عشوائية مكونة من (100) عامل موزعين حسب عدد وحدات إنتاجهم .

عدد وحدات الانتاج	أقل من 10	20-10	30-20	40-30	50-40	60-50	60 فأكثر	المجموع
عدد العمال	3	5	10	25	30	20	7	100

والمطلوب:

1 . إيجاد منوال الانتاج وتفسير النتيجة ؟

2 . إيجاد وسيط الإنتاج ومقياس مناسب لتشتت التوزيع .

3 . إذا افترضنا أن الفئة الأخيرة كانت 60 وأقل من 70 فأوجد الانحراف المعياري للتوزيع وكذلك معامل الاختلاف.

4 . فيما يلي التوزيع التكراري للدخل الإجمالي لعينة عشوائية من الأسر في أحدى المدن

الدخل بالدولار	أقل من 100	300-100	500-300	800-500	1000-800	1200-1000
عدد الأسر	18	24	40	26	6	2

والمطلوب:

a) أحسب مقياساً للنزعة المركزية وآخر للتشتت لدخل هذه الأسر.

. إذا كان انفاق هذه الأسر على الثقافة كالآتي :

فئات الانفاق	أقل من 10	15-10	20-15	30-20	60-30	100-60
عدد الأسر	18	24	40	26	6	2

b) قارن بين تشتت دخول هذه الأسر وتشتت إنفاقهم على الثقافة .

الفصل الثالث

مبادئ الاحتمالات

. مقدمة

. تعاريف هامة

. خواص الاحتمالات

. الاحتمال الشرطي

. نظرية بايز

. تمارين غير محلولة

3-1- مقدمة :

يُعد علم الاحتمال من الوسائل الهامة جداً التي يستخدمها الاحصاء الرياضي الذي بني أساساً على نظرية الاحتمال، حيث يهتم بالتنبؤ بأحداث مستقبلية غير مؤكدة الحدوث، أي يتضمن عنصر المخاطرة .

تقوم نظرية الاحتمال عموماً على مفهومين هما:

- التجربة الاحتمالية أو العشوائية (Random Experiment) : بالتعريف هي كل عملية تؤدي إلى ملاحظة أو قياس ظاهرة ما.
- الحادثة (Event): هي الناتج من التجربة العشوائية والتي لا يمكن معرفة نتائجها قبل إجراؤها.

تعريف :

مجموعة كل النتائج الممكنة لتجربة عشوائية ما تسمى فراغ العينة ويرمز لها بالرمز S.

الحادث هو مجموعة جزئية من فراغ العينة. ولتوضيح هذا التعريف لدينا المثال التالي :

مثال (1) :

عند قذف قطعة نقد متوازنة مرة واحدة. هذه التجربة تؤدي إلى إحدى نتيجتين إما صورة النقد (Head) ويرمز لها عادة بالرمز H أو إلى الكتابة على النقد (Tail) ويرمز لها عادة بالرمز T. عندئذ فإن فراغ العينة لهذه التجربة هو . S= {H, T} وبالتالي ستكون نتيجة التجربة إحدى الحالات التالية :

(a) هي H

(b) إما H أو T

(c) كلاهما T, H (وهذا غير محتمل حدوثه) .

(d) ليست H.

مثال (2) :

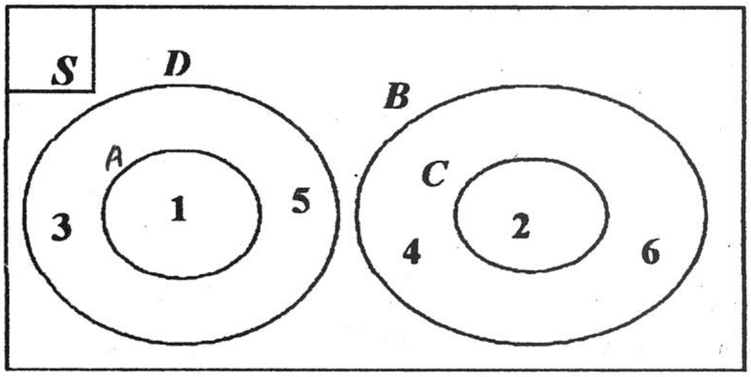
في تجربة قذف حجر النرد. نعلم أن هناك 6 نتائج ممكنة. فراغ العينة لهذه التجربة هو : S = {1, 2, 3, 4, 5, 6}.

يمكن أن نعرف الحوادث الآتية على فراغ هذه التجربة :

A هي حادثة الحصول على العدد 1 و B هي حادثة الحصول على عدد زوجي و C هي حادثة الحصول على عدد زوجي لا تتجاوز 3 وكذلك D هي حادثة عدد فردي، ويمكن تمثيل تجربة قذف حجر النرد في شكل فن المبين في الشكل المرفق وعلى ذلك يكون لدينا:

$A = \{ 1 \}$ (1) $B = \{2, 4, 6 \}$ (2)

$C = \{ 2 \}$ (3) $D = \{ 1, 3, 5\}$ (4)



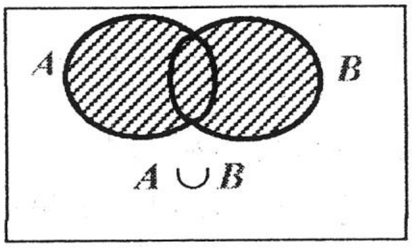
3-2- تعاريف هامة :

. الاجتماع union ورمزه U :

إن اجتماع مجموعتين B, A هو مجموعة العناصر التي تنتمي إلى A أو إلى B أو إليهما معاً. ويرمز لذلك بالرمز $A \cup B$، حيث إن $A \cup B$ هو مجموعة العناصر التي تنتمي إلى واحدة منها على الأقل .

ففي مثالنا السابق نجد أن :

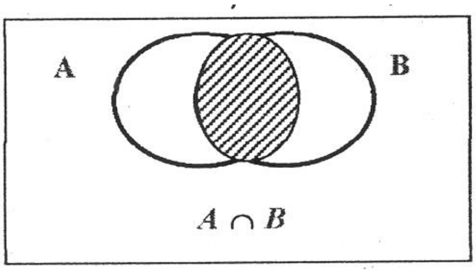
$B \cup C = \{2,4,6\}$ و $A \cup B = \{1,2,4,6\}$



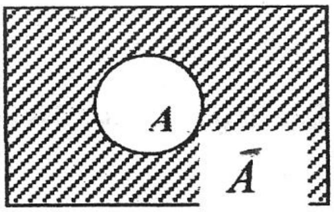
. التقاطع Intersection ورمزه ∩ :

إن تقاطع مجموعتين A و B هو مجموعة العناصر التي تنتمي إلى A وإلى B بنفس الوقت ويرمز له بالرمز $A \cap B$، ولو عدنا للمثال السابق لوجدنا أن :

$B \cap C = \{2\}$ و $A \cap B = \phi$ (أحياناً تكتب $A \cap B$ بالصورة AB).



. **متممة مجموعة (أو تتمة مجموعة) :**



إن متممة مجموعة A ويرمز لها بالرمز A^c (أو $\overline{A}$) .

هي مجموعة العناصر الموجودة في المجموعة الشاملة ولا تنتمي إلى A أي أن $\overline{A}$ هي نفي A (أو ليس A).

من الواضح أنه إذا كانت S هي المجموعة الشاملة فإن $A \cup \overline{A} = S$. ففي مثالنا السابق تلاحظ أن $\overline{A} = \{2,3,4,5,6\}$.

. **المجموعة الخالية :**

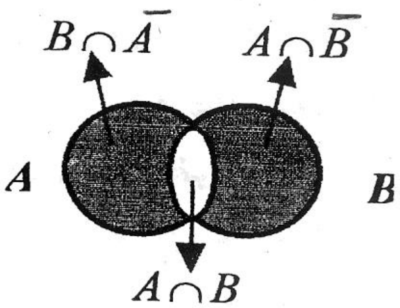
هي تلك المجموعة التي لا تتضمن أي عناصر فيها ويرمز لها بالرمز ϕ أو بالرمز $\{\}$ ونجد في المثال السابق أن $D \cap C = \phi$.

. **المجموعات الجزئية :**

إذا كان كل عنصر ينتمي إلى A ينتمي أيضاً إلى B فيقال أن A محتواه في B وتكتب بالصورة $A \subset B$ (وعند التساوي نكتب) $B \subseteq A$، $B \subset A$ وتقرأ : A محتواة في B أو A مجموعة جزئية في B.

. **الفرق بين مجموعتين :**

إن الفرق بين المجموعتين B و A هو مجموعة العناصر التي تنتمي إلى A ولا تنتمي إلى B.



ويرمز لذلك بالرمز $A \cap \overline{B}$ وهذا موضح في شكل فن المقابل أي :

$A \cup B = A\overline{B} \cup AB \cup B\overline{A}$ حيث إن :

$\overline{A}B = \overline{A} \cap B$

وبالمثل $A\overline{B} = A \cap B$ وكذلك $AB = A \cap B$.

العمليات على المجموعات :

إذا كانت لدينا المجموعات الجزئية A, B, C من S فإن عمليات المجموعات تحقق عدة خواص منها الآتي :

. **قوانين التبديل :**

$$A \cup B = B \cup A$$

$$A \cap B = B \cap A$$

. **قوانين الدمج أو التجميع :**

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

. قوانين التوزيع :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

. قوانين دي مورغان :

$$(\overline{A} \cup \overline{B}) = 1- [(A \cap B)]$$

$$(\overline{A} \cap \overline{B}) = 1 - [(A \cup B)]$$

وهذه القوانين تعطينا العلاقة بين عمليات الاتحاد والتقاطع.

3-3- خواص الاحتمال:

(1) $0 \leq P (A) \leq 1$ (الاحتمال موجب ودائماً أكبر من الصفر)

(2) $P (\phi) = 0$ ، $P(S)=1$

(3) مجموع الحالات الممكنة تساوي الواحد الصحيح .

وما يمكن ملاحظته من هذا التعريف هو أن الاحتمال P عدد ليس له وحدات ولا يمكن أن يكون سالب القيمة ولا يمكن أيضاً أن يتجاوز الواحد الصحيح .

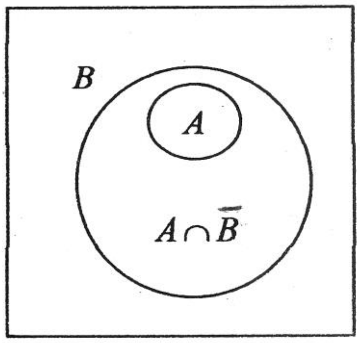
ملاحظات :

- إن احتمال الحادث المتمم للحادث A يعطى :

$$\mathbf{P (\overline{A}) = 1 - P (A)}$$

- إن احتمال الحادث المستحيل يساوي الصفر، أي: $\mathbf{P (\phi) = 0}$:
- إذا كانت لدينا الحوادث A, B وكان $A \subset B$ فإن : $P (A) \leq P (B)$:

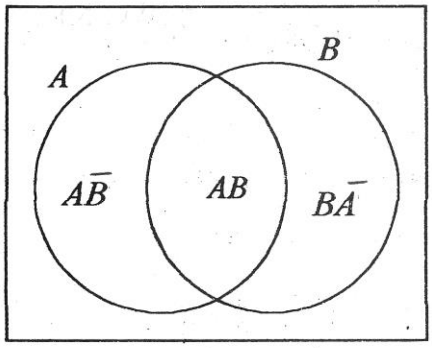
أنظر شكل فن :



- **لتكن A حادثة ، فإن احتمالها $P(A) \le 1$**
- **إن احتمال اجتماع حادثين A, B يعطى :**

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

أنظر شكل فن المرفق :



مثال (3) :

لدينا قاعدتان لإطلاق صواريخ أرض/جو، وفجأة ظهرت طائرة معادية في مجال القاعدتين حيث أطلقت كل قاعدة صاروخاً واحداً على الطائرة، إذا كان:

$P(A) = 0.93$, $P(B) = 0.89$ وكذلك $P(A \cap B) = 0.87$

المطلوب : أوجد احتمال إصابة الطائرة المعادية لصاروخ واحد على الأقل ؟

الحل :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.93 + 0.89 - 0.87 = 0.95$$

مثال (4) :

صندوق به 5 كرات اثنتان بيضاوان وكرة واحدة سوداء وكرتان حمراوان، سحبت كرة من هذا الصندوق بصورة عشوائية، فما احتمال أن تكون الكرة المسحوبة بيضاء أو سوداء ؟ .

الحل :

لتكن A حادثة الحصول على كرة بيضاء و B حادثة الحصول على كرة سوداء .

بما أن هذه الحوادث متنافية الحدوث أي بيضاء أو سوداء، فإننا نطبق قانون الجمع فنحصل على :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= \frac{2}{5} + \frac{1}{5} - 0 = \frac{3}{5}$$

- إذا كانت C و B و A ثلاثة حوادث فإن :

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - p(A \cap B)$$

$$P(A \cap C) - P(B \cap C) - P(A \cap B \cap C)$$

4-3- الاحتمالي الشرطي Condition Probability :

يقوم هذا الاحتمال على إيجاد احتمالات معينة إذا علم تحقق حوادث أخرى ويُعبر عنه كالتالي:

P(A/B) ويقرأ احتمال حدوث (A) علماً بأن الحادث (B) قد وقع، ويعرّف بالعلاقة التالية:

$$P(A/B)=\frac{P(A\cap B)}{P(B)}$$

بحيث : $P(B) > 0$

وبالمثل : إذا كان $P(A) > 0$ فإن $P(B/A)=\frac{P(A\cap B)}{P(A)}$ أحياناً P (A / B) يقرأ " احتمال حدوث A شريطة أن B قد حدث .

مثال 5 :

لدينا 20 كرة متماثلة من حيث الشكل، وضعت في صندوقين، وكانت (13) منها بلون أحمر (R) و (7) بلون أصفر (Y) تبعاً للجدول الآتي:

اللون	الصندوق الأول (A)	الصندوق الثاني (B)	المجموع
R	**5**	**8**	**13**
Y	**3**	**4**	**7**
المجموع	**8**	**12**	**20**

اختير منها كرة عشوائياً :

(1) أوجد احتمال أنها حمراء .

(2) إذا كانت الكرة المختارة من اللون الأحمر فما هو احتمال أنها من الصندوق الأول.

الحل :

(1)

$$P(R)=(1/2)\,(5/8) + (1/2)\,(8/12)$$
$$= (5/16) + (8/24) = 0.64$$

(2)

$$P(A/R)=\frac{P(A\cap R)}{P(R)}=\frac{P(A)\,P(R/A)}{P(R)}$$

$$=\frac{(1/2)(5/8)}{0.64}=0.48$$

مثال (6) :

إذا كان لدينا : $P (A \cap B) = 0.3$ ، $P (B) = 0.5$ ، $P (A) = 0.4$

فإن : $P(A/B)=\frac{P(A\cap B)}{P(B)}=\frac{0.3}{0.5}=0.6$

وبالمثل نجد أن : $P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{0.3}{0.4}=0.75$

خواص الاحتمال الشرطي :

(1) $P(A/B)$ ، $P(B) \geq 0$

(2) $P(A/A) = 1$

(3) إذا كان $A_1, A_2, \dots$ حوادث متنافية:

فإن $A_i \cap A_j = \phi \quad \forall\, i \neq j$:

$$P(A_1 \cup A_2 \cup \dots / B) = P(A_1/B) + P(A_2/B) + \dots$$

ملاحظات هامة :

(1) من خلال تعريف الاحتمال الشرطي يمكن كتابة :

$$P(A \cap B) = P(A)\, P(B/A)$$

أو :

$$P(A \cap B) = P(B)\, P(A/B)$$

3-5- : قاعدة بايز Bayes' Theorem :

هي إحدى الطرائق المناسبة لحساب احتمال حادث (E) إذا كان هذا الحادث يأتي من قبل عدة حوادث متعددة، أي علينا تحديد كل مرحلة من مراحل التجربة الإحصائية ولتقريب فكرة هذه القاعدة إلى الأذهان نأخذ المثال الآتي :

مثال (7) :

ثلاثة صناديق تحتوي على كرات متساوية الحجم، الصندوق B_1 يحتوي على كرتين بلون أحمر (R) و 4 بيضاء (W)، والصندوق B_2 يحتوي على كرة واحدة حمراء (R) ودائرتين بلون أبيض (W)، كما أن الصندوق B_3 يحتوي على 5 كرات حمراء (R) و4 بيضاء (W)، **المطلوب:** سحبت كرة عشوائياً ما احتمال أن تكون حمراء؟

ومن تعريف الاحتمال الشرطي نجد أن:

$$P(R) = P(B_1)\, P(R/B_1) + P(B_2)\, P(R/B_2) + P(B3)\, P(R/B_3)$$

$$= \left(\frac{1}{3}\right)\left(\frac{2}{6}\right) + \left(\frac{1}{6}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{5}{9}\right) = \frac{4}{9}$$

سحبنا كرة عشوائياً فكانت حمراء **المطلوب:** أي صندوق تم اختيارها؟ . وللإجابة على هذا التساؤل فإننا نحسب الاحتمالات الشرطية التالية :

وهي $P(B_3/R), P(B_2/R), P(B_1/R)$:

$$P(B_1 / R) = \frac{P(B_1 \cap R)}{P(R)}$$

$$= \frac{P(B_1)P(R/B_1)}{P(B_1)P(R/B_1) + P(B_2)P(R/B_2) + P(B_3)P(R/B_3)}$$

$$= \frac{(1/3)(2/6)}{(1/3)(1/6) + (1/2)(4/9) + (1/2) + (5/6)}$$

$$= \frac{2}{8} = \frac{1}{4}$$

$$P(B_2 / R) = \frac{P(B_2)P(R/B_2)}{P(B_1)P(R/B_1) + P(B_2)P(R/B_2) + P(B_3)P(R/B_3)}$$

$$= 1/8$$

الاحتمال الأخير :

$$P(B_3 / R) = \frac{P(B_3)P(R/B_3)}{P(B_1)P(R/B_1) + P(B_2)P(R/B_2) + P(B_3)P(R/B_3)}$$

$$= \frac{(1/2)(5/9)}{\frac{4}{9}} = \frac{5}{8}$$

تمارين غير محلولة

تمرين (1) :

يحتوي صندوق على (4) كرات حمراء مرقمة من (1) إلى (4)، و (3) كرات بيضاء مرقمة من (1) إلى (3)، وجميع هذه الكرات من الحجم نفسه .

والمطلوب :

أ . سحبت بصورة عشوائية كرة من الصندوق، فما احتمال أن تكون هذه الكرة حمراء أو تحمل رقم زوجي ؟

ب . أعيدت الكرة المسحوبة سابقاً، ثم سحبت الكرات بالتتالي، ما احتمال تعاقب الكرات الحمراء والبيضاء في عملية السحب.

تمرين (2) :

أطلق رامي ثلاث طلقات متلاحقة على هدف، إذا علمت أن احتمال إصابة الهدف (0.6) والمطلوب:

. ما هو احتمال إصابة الهدف بطلقة واحدة فقط .

. احسب عدد الطلقات الواجب رميها على الهدف ليكون احتمال إصابة الهدف (0.9) على الأقل.

تمرين (3) :

في أحد المختبرات (10) حاسبات، (6) منها تعمل بلغة البيسك فقط و(3) تعمل بلغة الفورتران فقط و (1) تعمل بكلتا اللغتين، فإذا اختيرت إحدى الحاسبات بصورة عشوائية، ما احتمال:

. أن تكون الحاسبة تعمل بلغة البيسك ولاتعمل بلغة الفورتران.

. أن تكون الحاسبة تعمل بلغة البيسك والفورتران معاً.

. أن تكون الحاسبة تعمل بلغة البيسك .

. أن تكون الحاسبة تعمل بلغة الفورتران .

. أن تكون الحاسبة تعمل بلغة البيسك أو الفورتران .

تمرين (4) :

يحتفظ مشفى بسيارتي إسعاف للطوارئ واحتمال أن تكون السيارة جاهزة للتحرك عند الحاجة إليها هو (0.9)، وإذا علمت أن توفر أحد السيارتين مستقل عن الأخرى .

والمطلوب :

1 . أوجد احتمال أن لا تتوفر أيّ منهما .

2 . ما هو احتمال تلبية الطلب لسيارة إسعاف في حالة الطوارئ .

تمرين (5) :

إذا كان لدينا احتمال هطول مطر في يوم معين (0.1)، واحتمال وجود رياح نشطة في ذلك اليوم (0.05)، واحتمال وجود مطر ورياح نشطة (0.03).

والمطلوب :

1 . أوجد احتمال وجود مطر أو رياح نشطة في ذلك اليوم.

2 . أوجد احتمال هطول مطر في ذلك اليوم علماً أن الرياح نشطة .

تمرين (6) :

يتنافس أحمد، محمد، حسين على المرتبة الأولى في امتحان مقرر الإحصاء، فإذا كان احتمال فوز أحمد = (2/3) من احتمال فوز محمد، واحتمال فوز حسين ضعف احتمال فوز محمد، فما هو احتمال فوز كل منهم ؟.

تمرين (9) :

إذا كانت (B, A) حادثتين ، حيث كان $P(A\cup B)=3/4$، $P(B)2/5$، $P(A)=x$. أوجد قيمة x في كل من الحالات التالية :

1 . إذا كان A و B متنافيتين .

2 . إذا كان B, A مستقلتين .

الفصل الرابع

المتغيرات العشوائية والتوزيعات الاحتمالية

Random Varibles and Probablity Distribution's

. مقدمة

. المتغيرات العشوائية

. التوزيعات الاحتمالية المنفصلة.

1 . التوزيع المنتظم

2 . محاولة بيرنولي .

3 . توزيع ثنائي الحدين

4 . توزيع بواسون .

. التوزيعات الاحتمالية المستمرة .

1. التوزيع الطبيعي :

2. التوزيع الطبيعي المعياري

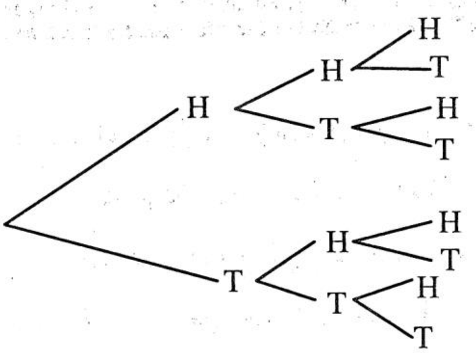
. تمارين غير محلولة

1-4- مقدمة :

عند إجراء التجارب الإحصائية وتعيين فضاء العينة لكل منها، ليس من الضروري دراسة نقاط فضاء العينة بالتفصيل ، وبمعنى آخر لا يكون الاهتمام منصب على النتائج في حد ذاتها، وإنما على القيم العددية المرتبطة بهذه النتائج الممكنة، لذلك فإن هذه القيم العددية هي ما نعبر عنه بقيم المتغير العشوائي، الذي هو عبارة عن نتائج التجربة الإحصائية غير المعروفة النتائج مسبقاً

تمرين (1) :

ألقيت قطعة نقود ثلاث مرات وكان المتغير العشوائي x يمثل عدد مرات ظهور صورة.
والمطلوب : وضح القيم الممكنة للمتغير العشوائي بطريقة الشجرة البيانية .



$HHH \rightarrow X = 3$
$HHT \rightarrow X = 2$
$HTH \rightarrow X = 2$
$HTT \rightarrow X = 1$
$THH \rightarrow X = 2$
$THT \rightarrow X = 1$
$TTH \rightarrow X = 1$
$TTT \rightarrow X = 0$

من خلال ذلك، نلاحظ أنه عند إلقاء قطعة النقود مرة واحدة، فإن فضاء العينة :

$S = (x : xi \in \{ H , T \}]$

واحتمال الحصول على كل من نتائج هذه التجربة :

X_i	0	1	2	3
$f(x_i)$	1/8	3/8	3/8	1/8

وبذلك نقول عن فضاء عينة (S) أنه فضاء عينة متقطع إذا كانت S مجموعة منتهية أو مجموعة غير منتهية ولكنها قابلة للعد .

المتغيرات العشوائية :

نقول عن متحول عشوائي (x) بأنه متحول عشوائي متقطع، إذا كان منطلقه (S) متقطعاً، أي إذا كانت (S) مجموعة منتهية أو قابلة للعدد، لذلك يأخذ المتغير العشوائي المتقطع قيماً صحيحة فقط. مثال { حجر النرد، أوراق اللعب ، عدد أفراد الأسرة، عدد سكان مدينة ما }.

وفي خلاف ذلك نقول عن متحول عشوائي (x) بأنه متحول عشوائي مستمر إذا كان فضاء العينة مستمراً مثل أوزان الطلاب أو أطوالهم، .. الخ .

4 . 2 . التوزيع الاحتمالي للمتحول العشوائي المتقطع :

إذا كان x متحولاً عشوائياً متقطعاً معرفاً على فضاء العينة S حيث صورته مجموعة منتهية:

$X(s) = \{ x_1, x_2, x_3, \ldots . x_n \}$

فإن تابع الاحتمال للمتحول العشوائي (X) يكتب على الشكل التالي :

$f(x_i) = P(X = x_i)$

حيث x_i تقابل P_i :

أو على شكل جدول

X_n	X_3 ……	X_2	X_1	X
$f(x_n)$	$f(x_3)$ …….	$f(x_2)$	$f(x_1)$	$f(x_i)$

إن الجدول يمكن تسميته بجدول الكثافة الاحتمالية للمتحول العشوائي (x) ونسمي التابع =(x_i) $f\ P(x = x_i)$ بتابع الكثافة الإحتمالية أو تابع التوزيع الاحتمالي، إذا تحقق الشرطين التاليين :

1 - $0 \le f(x_i) \le 1$

2- $\sum f(x_i) = 1$

لنعد الآن مثالنا السابق: عند إلقاء قطعة النقود ثلاث مرات، كان المتغير العشوائي (x) يمثل عدد مرات ظهور الصورة، والمطلوب :

1 . تحديد فضاء العينة ؟

2 . كتابة تابع الاحتمال

الحل :

. فضاء العينة $S = \{ HHH, HHT, HTH, HTT, THH, THT, TTT\}$

حيث أن $S = 8$ عناصر .

تابع الإحتمال :

3	2	1	0	X_i
1/8	3/8	3/8	1/8	$f(x_i)$

وأنه يحقق الشرطين السابقين .

$1- \sum f(x_i) = 1/8 + 3/8 + 3/8 + 1/8 = 1$

$2- 0 \le f(x_i) \le 1$

4 . 3 . الأمل الرياضي Mathematical Expecation :

يعد الأمل الرياضي للمتغير العشوائي من أهم المواضيع في الاحتمالات، ولتقريب فكرة الأمل الرياضي إلى الأذهان . نفترض أن (3) قطع نقدية رميت (10) مرات.

. فإذا كان عدم ظهور (H) قد حدث مرتين .

. وظهور (1H) حدث ثلاث مرات .

. وظهور (2H) حدث ثلاث مرات .

. وظهور (3H) حدث مرتين .

فلو أخذنا متوسط ظهور H في الرقعة الواحدة لقطع النقود الثلاث لوجدنا أن:

$$\frac{(0\times 2)+(1\times 3)+(2\times 3)+(3\times 2)}{10}=1.5H$$

وباختصار فإن الأمل الرياضي هو عبارة عن قيمة المتحول مضروباً باحتمال الحصول عليها، ويعطى التوقع الرياضي لمحتول عشوائي متقطع x يأخذ القيم x_i المقابلة للإحتمالات P_i بالعلاقة :

$$E(x)=\sum_{i=1}^{n} x_i \; f(x_i) \qquad i = 1, 2, 3, 4, \ldots n$$

ويطلق أحياناً اسم المتوسط للمتحول العشوائي x على المقدار E (x). ففي تمريننا يكون الأمل الرياضي أو التوقع لعدد مرات ظهور صورة من خلال ثلاث رميات هو :

E (x) = [(0). (1/8)] + (1).(3/8) + (2).(3/8) + (3). (1/8) = (12/8) = 3/2

التباين : Variance

إن تباين المتغير العشوائي المتقطع X الذي يأخذ القيم x_i والمقابلة للاحتمالات ذات $f(x_i)$ هو المقدار :

$$\operatorname{var}(x)=\sum_{i=1}^{n}\left(x_i-\mu\right)^2 f(x_i)$$

حيث إن : $E(x)=\mu$

ولحساب التباين في تمريننا (1) السابق:

x_i	$f(x_i)$	$x_i - \mu$	$(x_i - \mu)^2$	$(x_i-\mu)^2 f(x_i)$
0	1/8	0-3/2 = -3/2	2.25	0.28125
1	3/8	1 – 3/2 = -0.5	0.25	0.09735
2	3/8	2-3/2 = 0.5	0.25	0.09735
3	1/8	3-3/2 = 1.5	2.25	0.28125

- $\sum_{j=1}^{n} f(x_i)=1$

- var (x) = σ^2 = 0.28125 + (0.09735) . 2 + 0.28125 = 0.76

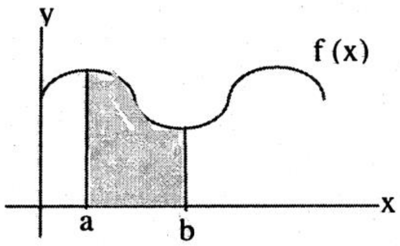
4 . 4 : المتغيرات العشوائية المستمرة :

من المعروف أنه لايمكن تمثيل التوزيع الاحتمالي للمتغير المتصل بجدول ولكن نعبر عنه بمعادلة ويكون الاحتمال عبارة عن مساحة تحت المنحني لتلك المعادلة، أي أن المتغير المستمر يأخذ قيمة في مدى محدد، لذلك توجد لأي متغير عشوائي مستمر دالة f(x) بحيث أن المساحات تحت منحني هذه الدالة (التابع) تعطي الاحتمالات المناظرة للفترات التابعة لها على المحور الأفقي .

● **تابع الكثافة الاحتمالية للمتحول العشوائي المستمر (X) :**

يعبر عن احتمال الحادث $P(a \le x \le b)$ بالنسبة للمتحول العشوائي المستمر بالعلاقة:

$P(a \le x \le b) = \int_a^b f(x).dx$ ، وإنه لإيجاد الاحتمال أو المساحة $P(a \le x \le b)$ والمعرفة بالدالة f(x) لابد من حساب المساحة تحت منحني هذه الدالة شكل (1 . 3) .



الشكل (1 . 4)

وإن احتمال الحادثة :

$$P(x = a) = \int_a^b f(x).dx$$

وإن مجموعة الحوادث :

$$a < x < b \,,\; a < x \le b \,,\; a \le x < b \,,\; a \le x \le b$$

لها جميعها نفس قيمة الاحتمال والمساو $\int_a^b f(x).dx$

ونقول عن تابع ما أنه تابع كثافة احتمالية لمتغير عشوائي مستمر، إذا كانت قيمة التابع f(x) تحقق الشرطين التاليين :

$$1- \quad \int_{-\infty}^{+\infty} f(x).dx = 1$$

$$2- \quad f(x) \ge 0$$

تمرين (2) :

ليكن لدينا تابعاً معيناً معطى بالعلاقة :

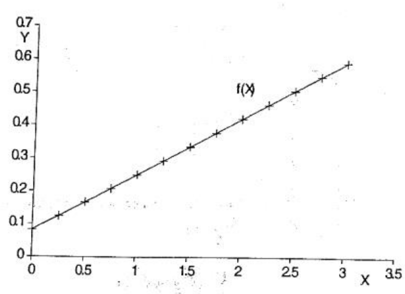
$$f(x)=\begin{cases}1/6x+1/12 & \forall x:0\le x\le 3\\ 0\end{cases}$$

والمطلوب :

1 . ارسم هذا التابع .

2 . أثبت أن هذا التابع هو تابع كثافة احتمالي لمتحول عشوائي مستمر .

الحل :



. لكي يكون هذا التابع تابع كثافة احتمالي لمتحول عشوائي مستمر، لابد أن يحقق الشرطين التاليين :

1 . $f(x)\ge 0$ وذلك من أجل جميع قيم x_i

2 . $\int_0^3 f(x).dx=1$

.كما نلاحظ أن الشرط الأول محقق وذلك من أجل جميع قيم x

$$P(0\le x\le 3)=\int_0^3 f(x).dx$$

$$=\int_0^3(\tfrac{1}{6}x+\tfrac{1}{12}).dx=\tfrac{1}{6}\int_0^3(x+\tfrac{1}{2}).dx$$

$$=\tfrac{1}{6}\left[\tfrac{x^2}{2}+\tfrac{x}{2}\right]_0^3=\tfrac{1}{12}\left[x^2+x\right]_0^3$$

$$=\tfrac{1}{12}\left[(9+3)-(0)\right]=\tfrac{1}{12}[12]=\tfrac{12}{12}=1$$

4. 5 . التوزيعات الاحتمالية المنفصلة:

سبق لنا دراسة المتغيرات العشوائية وتابع كثافته الإحتمالية، والأمل الرياضي، والتباين، لذا فإننا سندرس هنا بعض التوزيعات المتقطعة الخاصة لما لها من أهمية في الحياة العملية، بتفسير كثير من الظواهر .

• **التوزيع المنتظم** : يعرّف التوزيع المنتظم، بأنه التوزيع الذي يكون فيه لجميع قيم المتغير العشوائي نفس الاحتمال .

لنفرض أن المتغير العشوائي X يأخذ القيم $X = \{ x_1, x_2, ... x_n\}$ باحتمال متساوٍ ، فإن تابع الكثافة الاحتمالية يعطى بالعلاقة :

$$f(x) = \begin{cases} 1/n & \forall : x = x_1, x_2 \dots . . . x_n \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

أي أن : $P (X = x_i) = P_i = 1/ n$

حيث تمثل n عدد القيم الممكنة للمتحول العشوائي X.

خواص التوزيع المنتظم :

1 . $P_i > 0$

2 . $\sum P_i = \sum 1/n = n. 1/n = 1$

3 . $F(x) = P(x \le x_i) = \sum_{i=1}^{s} P_i = s/n$

4 . الأمل الرياضي $E(x) = \sum P_i x_i = 1/n \sum x_i = \bar{x}$

5 . التباين $\text{var}(x) = \sum P_i (x-\mu)^2 = 1/n \sum (x-\bar{x})^2$

6 . الانحراف المعياري $\delta_x = \sqrt{\text{var}(x)}$

• **محاولة بيرنولي** :

ثمة تجارب كثيرة يكون لها نتيجتان فقط، تحقق الحالة المرغوبة أو عدم تحققها. فإما أن يكون الإنتاج سليماً أو معطوباً، فنتيجة رمي قطعة النقد في الهواء نحصل على صورة أو كتابة، فلو رمزنا لتحقق الحالة المرغوبة بـ P فيكون عدم تحققها مساوياً إلى $q = 1-p$.

فكل تجربة لها نتيجتان فقط، إما q أو P وتسمى مثل هذه التجارب بمحاولة بيرنولي، وإن فراغ العينة لتجربة بيرنولي لها نتيجتان فقط ، ونرمز لحالة تحقق الحادثة المرغوبة بـ $P(x)=1$ ولحالة عدم تحقق الرغبة بـ $P(x)=0$، أي أن $X = \{ 0, 1\}$.

وعليه يجب أن تحقق محاولة بيرنولي الشروط التالية :

1 . كل تجربة لها نتيجتان فقط .

2 . كل التجارب مستقلة، أي أن نتيجة أي تجربة ليس لها أي تأثير على نتائج التجارب الأخرى .

3 . احتمال النجاح والفشل يظل ثابتاً في جميع المحاولات .

إن تابع الكثافة الاحتمالية يعطى بالعلاقة :

$$f(x)=\begin{cases} P^x(1-P)^{1-x} & x=\{0,1\} \\ 0 & \text{خلاف ذلك} \end{cases}$$

أو يعطي تابع الكثافة على شكل جدول :

Xi	0	1
Pi	(1-P)	P

تمرين (3) :

لنفرض أننا رمينا حجر النرد وكانت الحادثة المرغوبة هي ظهور الوجه {x = 6}، أوجد تابع الكثافة لمحاولة بيرنولي :

$$f(x) = (1/6)^1 .\ (1-1/6)^{1-1} = 1/6$$

أو :

$$f(x) = 0\ (1-1/6) + 1.\ (1/6) = 1/6$$

خواص محاولة بيرنولي :

1 . $P + q = 1$

2 . $0 \le P \le 1$

$$E(x) = \sum x_i P_i = 0\ (1-P) + 1.P = P$$

3 . الأمل الرياضي $E(x) = P$

4 . التباين : $\operatorname{var}(x) = E(x^2) - \mu^2 = \sum P_i x_i^2 - \mu^2$

$$= 0.\ (1-P) + 1.p - P^2 = P - P^2 = P\ (1-P) = P.q$$

5 . الانحراف المعياري :

$$\delta = \sqrt{P(1-P)} = \sqrt{P.q}$$

- **توزيع ثنائي الحدين The Binomial distribution :**

نفرض أن تجربتنا العشوائية تتكون من تكرار محاولة بيرنولي عدداً من المرات، (nمرة) وتعرف التجارب من هذا النوع بأنها تجارب ثنائية. وحينما ندرس تجربة ثنائية ذات (n) مرة، فإن كل خصائص هذه التجربة ترتبط بوسيطي هذه التجربة الثنائية (P,n) وهما احتمال النجاح، وعدد مرات تكرار التجربة، ويجب أن تحقق التجربة الإحصائية الشروط التالية:

1 . تتألف التجربة من عدد (n) من التكرارات المتماثلة .

2 . نتيجة كل محاولة للتجربة أحد ناتجين: " نجاح " أو " فشل " .

3 . إن احتمال النجاح (P) أو الفشل (q) يبقى ثابتاً في جميع المحاولات المكررة للتجربة.

4 . التكرارات مستقلة عن بعضها البعض، أي ليس هناك من تأثير لنتيجة أي محاولة على نتائج المحاولات الأخرى .

5 . نهتم فقط بعدد النجاحات (K) أو الحالات المرغوبة ، خلال جميع المحاولات (n).

إن تابع التوزيع الثنائي يعطى بالعلاقة:

$$f(x)=\begin{cases} C_k^n P^k q^{n-k} & k=0,1,2,........n \\ 0 \end{cases}$$

وإن $0 \le k \le n$

وهو ما يعرف بالتوزيع الثنائي أو بتوزيع ذي الحدين ويمكن عرض تابع الكثافة على الصيغة الجدولية التالية :

n	k	2	1	0	k
P^n	$C_k^n P^k q^{n-k}$	$C_2^n P^2 q^{n-2}$	$C_1^n P^1 q^{n-1}$	q^n	P

. التوقع (الأمل الرياضي) $E(x) = nP$

. التباين

$Var(x) = np$

الانحراف المعياري : $\delta=\sqrt{\operatorname{var}(x)}=\sqrt{nPq}$

تابع التوزيع التراكمي : $F(x)=P(x=k\le m)=\sum_{j=0}^{m} C_k^n P^k q^{n-k}$

تمرين (4) :

إذا كان احتمال إصابة الهدف لشخص ما $P = 1/5$، وأتيحت له فرصة الرماية $n=(5)$ مرات.

والمطلوب :

1 . أثبت أن هذا التوزيع الإحتمالي هو تابع ثنائي الحدين ؟

الحل :

لكي يكون هذا التوزيع ثنائي الحدين يجب أن يحقق العلاقة :

$$\sum_{k=0}^{n} C_k^n\ P^k\ q^{n-k}=1$$

$$\sum_{k=0}^{5} C_k^n \left(\tfrac{1}{5}\right)^k \left(\tfrac{4}{5}\right)^{n-k}$$

$$= \begin{array}{l} C_0^5\left(\frac{1}{5}\right)^0\left(\frac{4}{5}\right)^5 + C_1^5\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^4 + C_2^5\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^3 \\ + C_3^5\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^2 + C_4^5\left(\frac{1}{5}\right)^4\left(\frac{4}{5}\right)^1 + C_5^5\left(\frac{1}{5}\right)^5\left(\frac{4}{5}\right)^0 \end{array}$$

$$= 0.32768 + 0.4096 + 0.2048 + 0.0512 + 0.064 + 0.00032 = 1$$

والشرط الأول محقق حيث :

$$\sum_{k=0}^{n} C_k^n P^x q^{n-1} = 1$$

والشرط الأول محقق حيث : $0 \le C_k^n P^k q^{n-k} \le 1$

2 . احسب احتمال اصابة الهدف مرتين : $P (x = 2)$

$$P(x=2) = C_2^5\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^3 = 0.2048$$

3 . احسب احتمال الإصابة لأكثر أو يساوي مرتين $F(x) = P (x \ge 2)$

$$P (x \ge 2) = P (x = 2) + P (x = 3) + P (x = 4) + P (x = 5)$$

$$= 1- [P (x = 0) + P (x = 1)]$$

$$P(x=0) = C_0^5\left(\frac{1}{5}\right)^0\left(\frac{4}{5}\right)^5 = 0.32768$$

حيث أن :

$$P(x=1) = C_1^5\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^4 = 0.4096$$

وعليه يكون الاحتمال المطلوب :

$$= 1 – [0.32768 + 0.4096] = 0.26272$$

4 . احسب التوقع الرياضي والتباين والانحراف المعياري للمتحول العشوائي x

$$n = 5 \quad , \quad P = \frac{1}{5} \quad , \quad q = \frac{4}{5}$$

$$E(x) = nP = 5\left(\frac{1}{5}\right) = 1$$

$$\operatorname{var}(x) = nPq = 5\left(\frac{1}{5}\right)\left(\frac{4}{5}\right) = \frac{4}{5}$$

$$\delta = \sqrt{\frac{4}{5}} = 0.894$$

- **توزيع بواسون Possion is distribution :**

يصف توزيع بواسون المتغيرات العشوائية المتقطعة النادرة الحدوث ويعطينا عدد النجاحات في فترات زمنية معينة أو في منطقة محددة، والفترة الزمنية يمكن أن تكون ثانية أو دقيقة أو يوماً

والمنطقة المحددة يمكن أن تكون وحدة قياس للمساحة أو الحجم، فعدد المرضى الذين يراجعون غرفة الإسعاف خلال ساعة، وعدد حوادث السيارات خلال أسبوع في منطقة محددة، وعدد المكالمات الهاتفية التي تصل إلى مكتب الاتصال كل عشرة دقائق، وعدد كلها أمثلة على تجارب بواسون .

ويجب أن تحقق التجارب في توزيع بواسون الشروط التالية :

1 . عدد مرات النجاح (x) قليل خلال عدد كبير من محاولات بيرنولي (n).

2 . احتمال النجاح (P) صغير جداً من خلال عدد كبير من محاولات بيرنولي .

3 . في عدد فترات زمنية منفصلة عن بعضها البعض فإن حدوث النجاحات أي فترة مستقل عن حدوث النجاحات في أي فترة أخرى .

ويعطى التوزيع الإحتمالي لبواسون ، بالعلاقة التالية :

$$P(X=K)=\frac{e^{-\lambda}.\lambda^{k}}{K!}$$

حيث λ : هي معدل عدد النجاحات في الفترة الزمنية المعينة، أو المنطقة المحددة .. ،

$e = 2.718$

$x = k = 0, 1, 2, \dots$

خصائص توزيع بواسون :

1- $\lambda > 0$

2- $\sum_{k=0}^{\infty} P_k = 1$

3- $E(x) = var(x)\ \lambda = nP$

تمرين (5) :

إذا كان متوسط عدد المكالمات الهاتفية في أحد المراكز في الدقيقة الواحدة (2) مكالمة،

المطلوب:

1 . ما احتمال أن يتلقى المركز (4) مكالمات هاتفية خلال دقيقيتين ؟

الحل :

$\lambda = 4$ متوسط المكالمات خلال دقيقتين ، $K=4$ عدد المرات أو المكالمات .

$$P(K=4)=e^{-\lambda}.\frac{\lambda^{k}}{k!}$$

$$=e^{-4}.\frac{4^4}{4!}$$

$$=\frac{1}{e^4}.\frac{4^4}{4\times3\times2}=0.194$$

2 . ما احتمال أن يتلقى المركز (4) أو (5) مكالمات خلال (3) دقائق ؟

$\lambda = 6$

- $K = 4$

$$P(K=4) = e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$= e^{-6}.\frac{6^4}{4!} = 0.049$$

- $K=5$

$$P(K=5) = e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$= e^{-6}.\frac{\lambda^4}{K!}$$

$$= e^{-6}.\frac{6^5}{5!} = 0.58$$

- $P(K = 4) \cup P(K = 5) = 0.049 + 0.058 = 0.107$

• . التوزيع الطبيعي : Normal Disbritoution

يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية المستمرة وهو عبارة عن مجموعة المنحنيات البيانية أو التوزيعات التكرارية، التي تعد في غاية الأهمية في مجال البحوث طالما أن العديد من الظواهر تتبع التوزيع الطبيعي في مجتمع الدراسة، ويمكن من خلال هذا التوزيع الذي يشبه الجرس المقلوب إيجاد النسبة المئوية للحالات المدروسة التي تقع بين نقطتين فعندما يكون المتحول العشوائي (x) مستمراً، فإن عدد القيم الممكنة يصبح لانهائياً، لذلك فإن توزيع الاحتمالات على قيم المتحولات يعطي بدلالة تابع الكثافة الاحتمالية لمتغير عشوائي مستمر وهي من الشكل :

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

ونلاحظ أن لهذا التوزيع وسيطين هما الوسط الحسابي (μ) والانحراف المعياري (δ) وإن مجال تحوله $-\infty < x < +\infty$

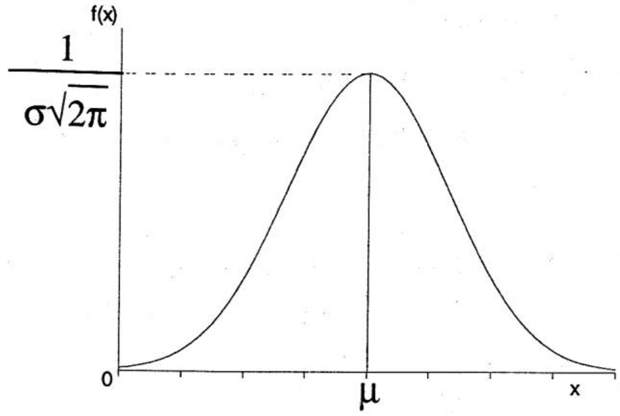
خواص التوزيع الطبيعي :

1 . إن المساحة التي يحددها تابع الكثافة f(x) على المحور (0x) تساوي الواحد الصحيح، أي:

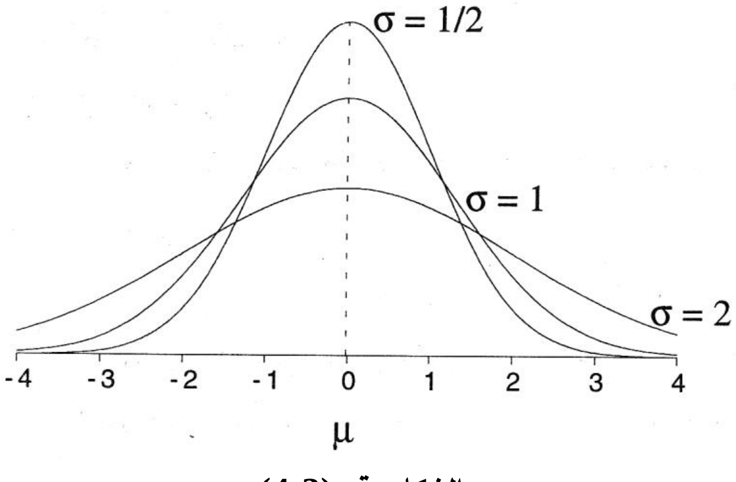
$$\int_{-\infty}^{+\infty} f(x).dx = 1$$

2 . إن منحني هذا التابع متناظر حول النقطة $x= \mu$ وهي النقطة التي يتمركز عندها التوزيع أو يبلغ نهايته العظمى. ثم ينتشر على جانبيها بصورة متناظرة وله نقطتي إنعطاف في الموضعين $x = \mu - \delta$ والنقطة $x = \mu + \delta$

ونلاحظ أنه من أجل قيم صغيرة لـ δ يكون انتشار المنحني على جانبي μ ضئيلاً، فيما يمتد إلى مسافات أبعد ويأخذ شكلاً أكثر انبساطاً كلما ازدات δ . الشكل رقم (3 . 4) وإن المساحة التي يحددها المنحني مساوية الواحد في جميع الأحوال .



الشكل رقم (2-4)



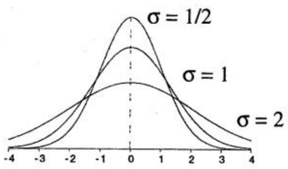
الشكل رقم (3-4)

3 . الأمل الرياضي :

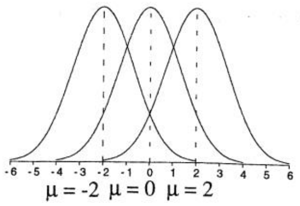
$$E(x) = \int_{-\infty}^{+\infty} f(x).dx = 1$$

4 . إن تابع الكثافة لمنحني التوزيع الطبيعي يحوي كما رأينا وسيطين (μ ، δ) ومن أجل قيم مختلفة لهما، نحصل على منحنيات متعددة لها جيمعاً شكل الجرس المقلوب ويمثل الشكلان أعلاه مجموعتين من المنحنيات الطبيعية .

تختلف منحنيات المجموعة الأولى بعضها عن بعض بمتوسطها كما في الشكل رقم (4 . 4) فقط بينما تختلف المجموعة الثانية بانحرافاتها المعيارية الشكل رقم (4 . 5).

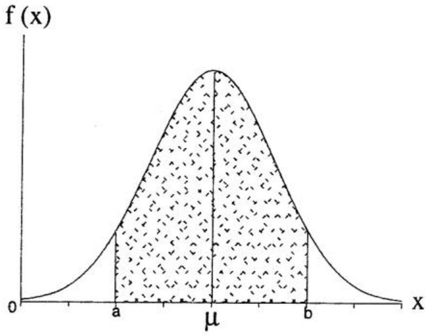


الشكل (4 . 4)



الشكل (4 . 5)

وإن احتمال أن يقع x في المجال [a , b] هو عبارة عن المساحة المحدودة بهذا المنحني والمحور (0x) والمستقيمين x=a و x = b حيث a < b الشكل رقم (4 . 6) .



الشكل رقم (4 . 6)

ولحساب الاحتمال :

$$P(a \le x \le b) = \int_a^b \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2} .dx$$

إذا كانت الظاهرة تخضع لمنحني توزيع طبيعي في المجتمع. فإن المساحات تتوزع وفق النسب التالية :

$$P (\mu . \delta \le x \le \mu + 1. \delta) = 68.27\%$$

$$P (\mu - 1.96\, \delta \le x \le \mu + 1.69\, \delta) = 95\%$$

$$P (\mu = 3\delta \le x \le \mu + 3\delta) = 99.73\%$$

• **التوزيع الطبيعي المعياري : Standard Normal Distribution**

يعرف التوزيع الطبيعي المعياري بأنه التوزيع الطبيعي الذي وسطه الحسابي يساوي الصفر وتباينه يساوي الواحد الصحيح أي أن المتغير العشوائي (Z) يخضع للتوزيع الطبيعي ذا الوسط الحسابي μ=0 و δ^2=1 ونعبر عنه Z: N(0,1). إذا كان (x) يخضع للتوزيع الطبيعي وسطاؤه

(μ) و (δ^2)، فإن المتغير (Z) الذي نحصل عليه من التحويل $Z = |\frac{x-\mu}{\delta}|$ يكون توزيعه توزيعاً طبيعياً معيارياً وكل قيمة من قيم (x) يقابلها قيمة من قيم (Z).

. خصائص التوزيع الطبيعي المعياري :

1 . المساحة التي يحددها التوزيع تحت المنحني تساوي الواحد الصحيح .

2 . منحني متناظر وله وسطه الحسابي (μ=0) و (δ^2=1)

3 . البيانات تتوزع بشكل منتظم حول الوسط الحسابي .

ولإيجاد تابع الكثافة لمنحني التوزيع الطبيعي المعياري (0,1) ~ N نجري عملية التحويل لتابع الكثافة لمنحني التوزيع الطبيعي (μ , δ^2) N

$$Z^2(\frac{x-\mu}{\delta})^2$$

والإنحراف المعياري δ = 1 إذا يصبح تابع الكثافة لمنحني التوزيع الطبيعي من الشكل :

$$\varphi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2}$$

وهو ما يعرف بالتوزيع الطبيعي المعياري .

حيث أن :

$$Z = \frac{x-\mu}{\delta}$$

$$-\infty < Z < +\infty$$

إن هذا التابع متناظر حول وسطه الحسابي حيث (μ = 0) كما أن انحرافه المعياري δ =1 .

نسمي المتغير العشوائي الجديد (Z)، المتغير المعياري أو الدرجة المعيارية، ولحساب قيمة الاحتمال أو المساحة نطبق العلاقة :

$$P(a \le x \le b = \int_a^b \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)}.dx$$

نحول المتغير العشوائي (x) إلى المتغير المعياري (Z) في تابع الكثافة الاحتمالية (μ, δ^2) ~N ليصبح (0,1) ~ N .

ولإيجاد التكامل السابق، لابد أيضاً من تحويل حدود التكامل إلى الدرجات المعيارية أو المتغير المعياري .

عندما x=a يصبح الحد الأدنى للتكامل $Z = \frac{a-\mu}{\delta}$

عندما x=b يصبح الحد الأعلى للتكامل $Z = \frac{b-\mu}{\delta}$

وإن : $dz = \frac{dx}{\delta} \Rightarrow dx = dz.\delta$

وبالتعويض في علاقة منحنى التوزيع الطبيعي $N\ (\mu, \delta^2)$

$$P(a \le x \le b) = \int_{\frac{a-\mu}{\delta}}^{\frac{b-\mu}{\delta}} \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}Z^2}\,\delta.dz$$

وبما أن $\delta = 1$ تصبح العلاقة :

$$P(a \le x \le b) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\delta}}^{\frac{b-\mu}{\delta}} e^{-\frac{1}{2}Z^2}\,\delta.dz$$

إن إجراء التكامل السابق شاق ويتطلب جهداً كبيراً، لذلك تم وضع جدول خاص لحساب هذه التكاملات ويدعى بجدول المساحات تحت منحني التوزيع الطبيعي المعياري $N\sim(0,1)$ وقد حسبت المساحات في هذا الجدول لنصف المنحني على اعتبار أن منحني التوزيع القياسي متناظر حول النقطة $\mu = 0$ وهي عبارة عن الوسط الحسابي وتقسم المنحني إلى قسمين متساويين ومتناظرين تماماً، وقد حسبت المساحات بدءاً من النقطة $Z = 0$ وبتزايد مقداره 0.01 ويقسم جدول المساحات، إلى سطر وعمود وتبدأ قيم (Z) في السطر بالقيمة $Z = 0$ وتزايد مقداره 0,01 حتى القيمة 0,09 وتبدأ قيم العمود من القيمة $Z = 0$ وتزايد مقداره 0.1 حتى القيمة $Z = 3$ حيث المساحة المحصورة $P\ (Z \le 0) = P\ (Z \ge 0) = 0.5$ وأن الهدف من الاستعراض السابق هو معرفة نسبة أو احتمال المفردات الواقعة في مجال محدد أو خارجه تحت المنحني الطبيعي من خلال التوزيع الطبيعي المعياري بطريقة رياضية تعتمد على تحويل القيم الطبيعية (x) قيم طبيعية معيارية (Z) وذلك لمعرفة مدى ابتعادها عن الوسط الحسابي μ أو $\bar{x}$ بالانحرافات المعياري وذلك من خلال جدول المساحات (Z) والموجود في نهاية الكتاب.

تمرين (6) :

إن طول الشخص هو متحول عشوائي مستمر (x) يخضع للتوزيع الطبيعي بمتوسط x=170 وانحراف معياري $\delta = 5$

1 . ماهو احتمال أن يقع x في المجال $P\ (160 \le x \le 185)$

2 . ماهو احتمال أن يقع x في المجال $P\ (x \ge 170)$

3 . ماهو احتمال أن يقع x في المجال $P\ (x \le 160)$

الحل :

$$P(160 \le x \le 185) = P\left(\frac{160-170}{5} \le Z \le \frac{185-170}{5}\right)$$

$$= P(-2 \le Z \le 3)$$

$$= P(0 \le Z \le 2) + P(0 \le Z \le 3)$$

$$= 0.4772 + 0.49865 = 0.97585$$

$$2\text{-}\ P(Z \ge 0) = P(0 \le Z \le 3) = 0.5$$

$$3\text{-}\ P(Z \le 2) = [\, P(Z \ge 0) - P(0 \le Z \le 2)$$

$$= 0.5 - 0.4772 = 0.0228$$

تمرين (7):

في إحدى الكليات تقدم لامتحان الاحصاء (500) طالب، وكان متوسط الدرجات بهذا المقرر (65) درجة وبانحراف معياري (15) درجة، فإذا كانت درجة النجاح لأي طالب في هذا المقرر (50) درجة، وبفرض أن درجات امتحان الطلاب في هذا المقرر تخضع للتوزيع الطبيعي

المطلوب:

1 . أحسب عدد الطلاب الراسبين في الامتحان لهذا المقرر.

2 . أحسب عدد الطلاب الذين حصلوا على أكثر من (65) درجة.

(يترك الحل للطالب)

تمارين غير محلولة

تمرين (1) :

وجد أحد الباحثين الاجتماعيين أن مشاهدي التلفزيون يفضلون القناة A بنسبة (75%)، بينما (25%) يفضلون القناة B ، سئل أربعة مشاهدين .

والمطلوب :

1 . احسب احتمال أن نجد بينهم اثنين يفضلون مشاهدة القناة (A) ؟

2 . احسب احتمال أن نجد بينهم واحد على الأقل يفضل مشاهدة القناة (A) ؟

3 . أحسب التوقع والتباين والانحراف المعياري للمشاهدين الذين يفضلون مشاهدة القناة (A)؟

تمرين (2) :

إذا كان من المعلوم ، أن الأطفال يفضلون برنامجي الأطفال B, A بالتساوي، سئل ثمانية أطفال، فأجاب ستة منهم أنهم يفضلون البرنامج A ،

والمطلوب :

1 . ماهو احتمال هذه النتيجة .

2 . ماهو التوقع والانحراف المعياري .

تمرين (3) :

أوجد الأمل أو التوقع الرياضي للتوزيع الاحتمالي التالي :

3	1	0	-1	-2	X
0.25	0.35	0.1	0.2	0.1	F(x)

تمرين (4) :

من خلال دراسة أجريت على سجلات الطلاب في كلية الاقتصاد تبين أن نسبة النجاح (50%)، قام أحد الطلاب بامتحان (12) مقرر والمطلوب :

1 . ماهو احتمال نجاح الطالب بـ (10) مقررات على الأكثر؟

2 . ماهو احتمال نجاح الطالب بـ (5) مقررات على الأقل ؟

3 . ماهو احتمال نجاح الطالب بجميع المقررات .

تمرين (5) :

إذا كان أطوال طلاب الكلية يتبع توزيعاً طبيعياً (36 ، 168) ~N، وسحبنا عينة عشوائية من الطلاب بحجم 500 = n طالباً .

والمطلوب :

1 . احسب عد د الطلاب الذين تزيد أطوالهم عن 184 سم.

2 . احسب عدد الطلاب الذين يقل أطوالهم عن 156 سم.

3 . احسب عدد الطلاب الذين تزيد أطوالهم عن 165 وتقل عن 174 سم.

تمرين (6) :

كان الإعتقاد السائد لدى المسؤولين عن الصناعات البترولية أن متوسط الأجر الاسبوعي للعامل 3100 وحدة نقدية بانحراف معياري 250 وحدة، وإذا علمت أن الأجور كانت تتبع توزيعاً طبيعياً في هذه الصناعة، وسحبنا بصورة عشوائية عاملاً من عمال هذه الصناعة، فما هو احتمال أن يكون أجره :

1 . $P (3600 \leq x \leq 3350)$

2 . $P (3350 \leq x \leq 3850)$

3 . $P x \leq 2455)$

على اعتبار أن الأجر هو متغير عشوائي مستمر ونرمز له بـ (x)

تمرين (7) :

تبين من الدراسات الاحصائية أن احتمال إصابة الأطفال بالشلل هو (0,016) فإذا افترضنا أن عدد الأطفال حديثي الولادة (250) خلال فترة زمنية محددة، المطلوب :

1 . ما احتمال أن يوجد (5) أطفال من بين هذا العدد مصابين بهذا المرض؟

2 . ما احتمال أن يوجد أكثر من ثلاثة أطفال مصابين بهذا المرض من بين (5) أطفال.

تمرين (8) :

ليكن لدينا x متغيراً عشوائياً متقطعاً له جدول التوزيع الاحتمالي التالي :

X	0	1	2	3
$P(X=x)$	0.1	0.35	0.4	0.15

والمطلوب :

. احسب التوقع الرياضي للمتغير العشوائي x ، ثم احسب التباين والانحراف المعياري.

الفصل الخامس

الارتباط والانحدار

- . مقدمة
- . العلاقة الارتباطية
- . خطوات دراسة العلاقة الارتباطية
- . الارتباط الخطي
- . معامل بيرسون للارتباط
- . معامل ارتباط سبيرمان
- . اختبار معنوية معامل الارتباط
- . معامل التحديد
- . الانحدار الخطي
- . معامل الاقتران ومعامل التوافق
- . تمارين غير محلولة

الارتباط والانحدار

5-1- مقدمة :

كان اهتمامنا حتى الآن مركزاً على دراسة متغير عشوائي واحد بمعنى أن دراستنا قاصرة على قياسات لظاهرة كمية واحدة (الأجر، العمر، الدخل، الإنتاج.. إلخ)

ولكن تقابلنا مواقف كثيرة يكون اهتمامنا بها موجهاً، لا إلى متغير واحد وإنما إلى متغيرين أو أكثر وتكون الأسئلة المطروحة منصبة على العلاقة بينهما فمثلاً: هل لو زاد دخل الأسرة زاد إنفاقها على الغذاء؟ وهل لو زاد دخل الأسرة زاد إنفاقها على الثقافة؟

إن دراسة هذه البيانات تهدف للإجابة على السؤالين التاليين :

1 . هل هناك علاقة بين المتغيرين ؟

2 . إذا كان هناك علاقة بين المتغيرين فكيف نعبر عنها بمعادلة.

5-2- العلاقة الارتباطية :

إن دراسة العلاقة الكائنة بين المتغيرين أو أكثر هدفها معرفة فيما إذا كان هناك إرتباط (علاقة) بين المتغيرين وقوة هذه العلاقة، وعند دراسة العلاقة الارتباطية يجب توافر الشروط التالية:

1 . أن تكون العلاقة بين المتغيرين منطقية.

2 . أن تكون أحدهما مسببة والأخرى ناتجة.

3 . أن تكونا قابلتين للقياس.

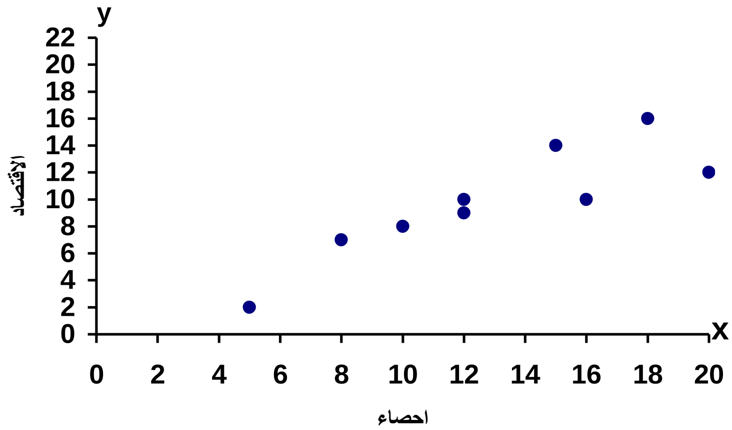
4 . أن تكون القياسات متقابلة من حيث الزمان أو المكان.

5-3- خطوات دراسة العلاقة الارتباطية :

I . شكل الانتشار Scatter Diagram :

ونحصل عليه من خلال التمثيل البياني لقيم (x) و (y) كما هو موضح في المثال التالي الذي يمثل درجات عشرة طلاب في مادتي الإحصاء والإقتصاد.

درجات الإحصاء	درجة الإقتصاد
20	12
18	16
16	10
15	14
14	12
12	10
12	9
10	8
8	7
5	2

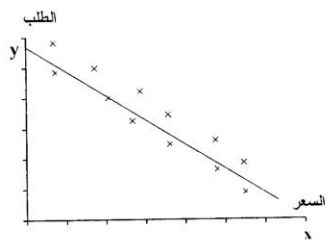


بالنظر إلى الشكل المرفق نلاحظ النقاط تنتشر تقريباً حول خط مستقيم بما يوحي بوجود ارتباط بين المتغيرين فالقيم الكبيرة في كلا المتغيرين تتواجد سوياً وكذلك القيم الصغيرة أي أن هناك علاقة طردية بين المتغيرين.

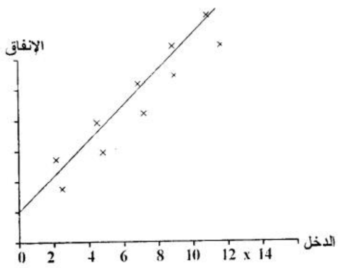
وبشكل عام لو رسمنا المشاهدات بهذه الطريقة فإننا نحصل على مجموعة من النقاط تعطينا انطباعاً بوجود علاقة أو عدم وجودها والآن لننظر إلى الأشكال التالية وننظر ماذا يمكن استنتاجه منها.

واضح من الشكلين في (A) و (B) أنهما يمثلان علاقة ارتباطية والارتباط بينهما، في الشكل (A) يمثل علاقة طردية و (B) يمثل علاقة عكسية . فكلما زاد المتغير (x) قل المتغير (y).

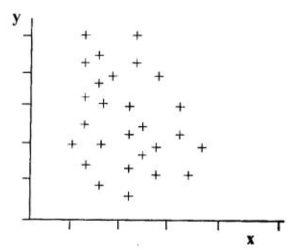
أما في الشكل (C). فإن النقاط مبعثرة بطريقة لاتوحي بوجود أي علاقة بين المتغيرين فالقيم (x) تصاحبها قيم صغيرة وقيم كبيرة للمتغير (y) وكذلك القيم الكبيرة لنفس المتغير (x) تصاحبها أيضاً قيم صغيرة وقيم كبيرة للمتغير (y) .. ومن ثم فإنه من الصعب استنتاج وجود علاقة بين المتغيرين .

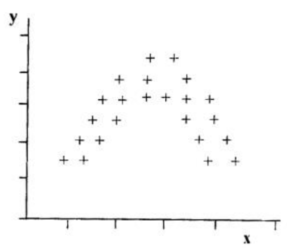


الشكل (B)



الشكل (A)





الشكل (C) **الشكل (D)**

أما الشكل (D) فهو يمثل علاقة غير خطية من نوع خاص لن نهتم بها في دراستنا الحالية وإنما سيكون اهتمامنا بالعلاقات الخطية فقط.

مما سبق يمكن القول: إن شكل الانتشار مهمته تحليل سلوك المتغيرين والهدف منه تحديد نوع وطبيعة العلاقة بين المتغيرين.

2 . **تحديد طبيعة المتغيرات :** نحن نعلم أن المتغير هو عبارة عن القيم للمشاهدات التي تخص الظاهرة المدروسة خلال فترة معينة (كمية أو نوعية)، ويمكن تقسيم المتغير إلى نوعين من المتغيرات هي:

- المتغير المستقل (x) : هو المتغير الذي تكون تغيراته غير مرتبطة بتغيرات المتغير الآخر.
- المتغير التابع (y): هو المتغير الذي تغيراته مرتبطة بتغيرات المتغير المستقل ويتم تحديد المتغير المستقل من خلال:

طريقة الأسبقية الزمنية : وتقوم على فكرة المتغير الأسبق حدوث زمنياً هو المتغير المستقل مثل: الدخل والانفاق على الغذاء، فالدخل هنا هو الأسبق زمنياً لذلك هو المتغير المستقل والانفاق هو المتغير التابع. وتبقى طريقة الأسبقية الزمنية هي الطريقة الأكثر استخداماً وشيوعاً.

4-5- الارتباط الخطي Linear Correlation :

معامل الارتباط (r) :

إذا كانت هناك علاقة إحصائية بين متغيرين فإننا نسمي هذا ارتباطاً، والارتباط قد يكون طردياً أو عكسياً والأهم من ذلك هو "درجة" هذا الارتباط والتي تتراوح مابين (1+) و (1-) أي rε=[-1,+1] وكلما اقتربت (r) من الواحد، دل ذلك على قوة العلاقة بين المتغيرين وبالعكس ومثال ذلك العلاقة بين الدخل والانفاق على الثقافة وغيرها من العلاقات .

5-5- معامل بيرسون للارتباط Pearsons Correlation Coeffcient :

يعرَّف معامل بيرسون للارتباط بالمعادلة التالية :

$$r=\frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2\Sigma(y-\bar{y})^2}}=\frac{\Sigma(x-\bar{x})(y-\bar{y})}{n.s_x.s_y}$$

ولنحسب الآن معامل الارتباط لدرجات الطلاب في الإحصاء والاقتصاد في المثال السابق وحيث أن :

$\bar{y}=10$ ، $\bar{x}=13$ على الترتيب فإن الانحرافات ستكون :

$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$	$y-\bar{y}$	$x-\bar{x}$

4	14	49	2	7
36	30	25	6	5
0	0	9	0	3
16	8	4	4	2
4	2	1	2	+1
0	0	1	0	-1
1	1	1	-1	-1
4	6	9	-2	-2
9	15	25	-3	-5
64	64	64	-8	-8
138	140	188	0	0

$$r = \frac{140}{\sqrt{(188)(138)}} = 0.87$$

وهو رقم يقرب من الواحد ومن ثم فإنه يمكن الاستنتاج بأن ثمة علاقة قوية ولأن هذا العدد موجب فإننا نستنتج أن العلاقة طردية.

ويمكن أن نستخدم صيغة أبسط في حساب هذا المعامل وهي:

$$r = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2}} = 0.87$$

$$r = \frac{\Sigma xy - n\bar{x}\bar{y}}{\sqrt{[\Sigma x^2 - n(\bar{x})^2][\Sigma y^2 - n(\bar{y})^2]}} = 0.87$$

y^2	x^2	xy	y	x
144	400	240	12	20
256	324	288	16	18
100	256	160	10	16
196	225	210	14	15
144	196	168	12	14
100	144	120	10	12
81	144	108	9	12
64	100	80	8	10
49	64	56	7	8
4	25	10	2	5
1138	1878	1440	100	130

$$r = \frac{140}{\sqrt{(188)(138)}} = 0.87$$

ملاحظة : إن معامل الارتباط لايتأثر بإضافة (أو طرح) أو القسمة على مقدار ثابت على جميع قيم أي من المتغيرين ولتوضيح ذلك لدينا المثال التالي :

6-5- معامل ارتباط سبيرمان (ارتباط الرتب) Rank Correlation (r_s) :

يقدم معامل بيرسون مقياساً ممتازاً للارتباط ويفقد ميزته في أحوال أخرى أي لايصلح للاستخدام إذا كان المتغير لايمكن قياسه مثل تقديرات الطلاب أي يستخدم هذا المعامل لقياس المتغيرات النوعية (الوصفية) على شكل رتب بأن ترتب قيم كل متغير بينهما وبين نفسها ثم نطرح الرتب من بعضها لتكون d = فروق الرتب ومن ثم نحسب معامل ارتباط الرتب من المعادلة:

$$r = 1 - \frac{6\Sigma d^2}{\sqrt{n(n^2-1)}} = 0.87$$

مثال :

أخذت عينة عشوائية مكونة من (5) طلاب (خريجيون ومستجدون) من إحدى الكليات لمعرفة كيفية ترتيب أولياتهم في كيفية قضائهم لأوقات فراغهم، فكان الجدول الآتي:

الأنشطة	تلفزيون	أفلام	كتب	مجلات	انترنيت
رتب الخريجين	1	2	3	4	5
رتب المستجدين	2	3	1	4	5

والمطلوب :

قياس شدة العلاقة بين رأيي الطلاب حول الأنشطة وذلك لمعرفة تقارب وجهات نظر الطلاب حول قضاء وقت فراغهم أو (قياس تقارب وجهات النظر بين الطلاب)

	الخريجون	المستجدون		
الأنشطة	رتب (x)	رتب(y)	d	d^2
تلفزيون	1	2	-1	1
أفلام	2	3	-1	1
كتب	3	1	2	4
مجلات	4	4	0	0
انترنت	5	5	0	0

$$r_S = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6(6)}{5(25-1)} = 0.70$$

والعلاقة جيدة، ونرى أن المتخرجين والمستجدين اتفقوا في الرأي على شيئين فقط من حيث الأفضلية في قضاء وقت الفراغ (المجلات والانترنت).

5-7- اختبار معنوية معامل الارتباط :

لإجراء هذا الاختبار نعتبر المجتمع الذي أخذت منه العينة (n)، ومعامل الارتباط (R) هو تقدير غير متحيز لمعامل الارتباط (r) وأن المجتمع الأصلي يخضع للتوزيع الطبيعي ويمكن تحديد فيما إذا كان معامل الارتباط (r) لعينة حجمها (n) دالاً على وجود علاقة خطية بين المتغيرين، ويمكن إجراء اختبار معنوية معامل الارتباط بالاعتماد على توزيع (Z) الطبيعي وحساب الخطأ المعياري لمعامل الارتباط (Sr) وذلك على النحو التالي :

$$S_r = \sqrt{\frac{1-r^2}{n-2}}$$

حيث أن : (r^2) معامل التحديد (مرجع معامل الارتباط).

ونميز حالتين :

I. $r \geq 3Sr$ وبالتالي العينة المسحوبة عشوائية وتمثل المجتمع الاحصائي المحسوبة منه ولمعامل الارتباط أهمية احصائية.

II. $r < 3Sr$ العينة المسحوبة غير عشوائية ولاتمثل المجتمع الإحصائي المسحوبة منه وليس لمعامل الارتباط أهمية إحصائية.

مثال: نعود إلى التمرين السابق والمتعلق بدرجات الإحصاء والاقتصاد للطلاب والمطلوب: بيان الأهمية الإحصائية لمعامل الارتباط باستخدام التوزيع الطبيعي.

الحل :

$$S_r = \sqrt{\frac{1-(0.87)^2}{10-2}} = 0.17$$

$$0.87 > 3\ (\ 0.17)$$

بالتالي العينة المسحوبة عشوائية وتمثل المجتمع الإحصائي المحسوبة منه ولمعامل الارتباط أهمية احصائية.

5-8- معامل التحديد (r^2) :

بالتعريف هو عبارة عن مربع معامل الارتباط وتتراوح قيمته أو درجته بين الصفر والواحد الصحيح أي : $r^2 \in [0,1]$ وهو يقيس التحسن الناتج عن استخدام خط الانجدار بدلاً من استخدام الوسط الحسابي كأساس لتقدير قيم المتغير التابع أو يقيس النسبة المئوية من تباين المتغير التابع التي يفسرها المتغير المستقل أو النسبة المئوية من التباين في المتغير التابع الذي

تفسره معادلة خط الانحدار والباقي لاتستطيع معادلة خط الانحدار تفسيره.

9-5- الانحدار الخطي Regression line :

عند دراسة الارتباط كان الاهتمام بالعلاقة الكائنة بين (x) و (y) وماهي درجة قوة هذه العلاقة، أما الانحدار يهدف إلى تمثيل هذه العلاقة بمعادلة رياضية بحيث يمكن استخدامها للتنبؤ بقيم أحد المتغيرين إذا علمنا قيمة المتغير الآخر. مثال: التنبؤ بإنفاق أسرة

. معادلة خط الانحدار : بالتعريف هي المعادلة التي تصف العلاقة بين متغيرين موضوع والغاية منها تقدير القيم التي يمكن أن يأخذها المتغير التابع بدلالة المتغير المستقل وتعطى هذه المعادلة بالعلاقة التالية :

$$\hat{y} = a + bx + e$$

حيث أن :

$\hat{y}$: القيمة المقدرة للمتغير التابع (القيمة الوسطية).

a : ثابت خط الانحدار

b: معامل الانحدار (ميل خط الانحدار)

e: الخطأ العشوائي

. إيجاد معادلة خط الانحدار : ويتم إيجاد هذه المعادلة بطريقتين :

1) طريقة الرسم الحر (طريقة الرسم البياني):

تعتمد هذه الطريقة على رسم شكل الانتشار للظاهرة المدروسة باعتبار المحور (x) يمثل المتغير المستقل والمحور الأفقي (y) المتغير التابع ويتم رسم خط الانحدار بحيث يمر بين منتصف نقاط الانتشار ويتم تحديد معادلة خط الانحدار على النحو التالي :

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

a: عبارة عن نقطة تقاطع خط الانحدار مع المحور (y)

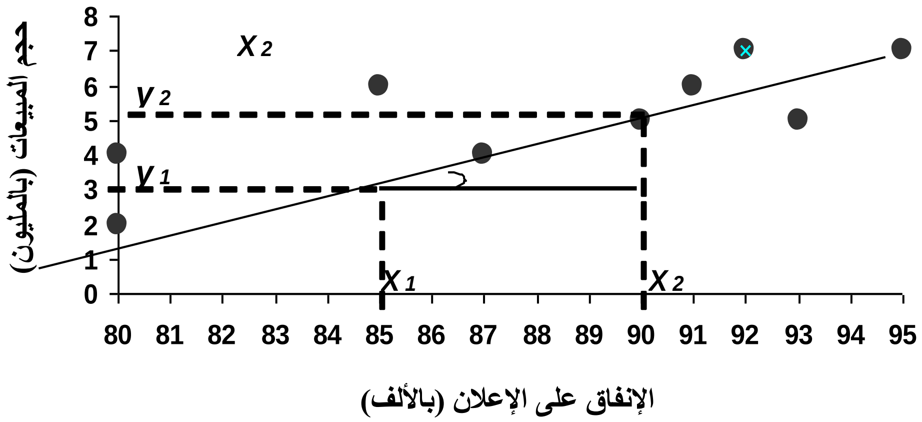
b: عبارة عن ميل خط الانحدار (ظل الزاوية) مع المحور (x)

ويتم حساب ظل الزاوية من خلال أخذ نقطتين ما على خط الانحدار وإجراء عملية الاسقاط على المحور (x) والمحور (y) كما هو موضح من خلال التمرين التالي:

أخذت عينة عشوائية مكونة من (10) شركات فكان انفاقها على الاعلان وحجم مبيعاتها السنوية على النحو التالي (بألوف الوحدات النقدية).

10	9	8	7	6	5	4	3	2	1	الشركات
85	93	95	90	92	87	85	80	80	90	الانفاق على الاعلان

حجم البيانات	5	2	4	6	4	7	6	7	5	5



a= +1 (نقطة التقاطع مع المحور y) .

$$b = \frac{y_2 - y_1}{x_2 - x_1} = \frac{5.5 - 3}{90 - 85} = 0.5$$ (ظل الزاوية = $\frac{\text{المقابل}}{\text{المجاور}}$)

وبالتالي معادلة خط الانحدار :

$$\hat{y} = 1 + 0.5x$$

يعاب على هذه الطريقة بأنها تقريبية وتختلف من شخص لآخر بسبب اختلاف أسلوب الرسم إلا أنها تعطي فكرة عن معادلة خط الانحدار .

2) . **الطريقة الرياضية** : تعتبر أدق من الطريقة السابقة باعتبارها تعتمد الأسلوب الرياضي ويمكن إيجاد ثوابت المعادلة من خلال أسلوبين :

a) أسلوب المربعات الصغرى: وتتلخص هذه الطريقة بإيجاد (a) و (b) التي تجعل مجموع مربعات الانحرافات أصغر ما يمكن أي إيجاد (a) و (b) التي تجعل الكمية $(\Sigma(y-\hat{y}) \rightarrow \min)$ أي أصغر ما يمكن، بمعنى آخر $(\Sigma(y-\hat{y})=0)$ وبالتالي : $(y-a-bx)^2 = 0$ وبالاشتقاق الجزئي الأول بالنسبة لكل من (a) و (b) ونعدم هذه المشتقات سنحصل على جملة المعادلتين التاليتين :

$$(1)\ \sum y = na + b\sum x$$
$$(2)\ \sum xy = a\sum x + b\sum x^2$$

وبحل جملة المعادلتين نحصل على قيم الثوابت (a)، (b) وبالتالي على معادلة خط الانحدار.

b) . الأسلوب المباشر: ويقوم هذا الأسلوب على حل المعادلتين السابقتين لنحصل على الثوابت:

$$b = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \frac{\Sigma xy - n\,\bar{x}\,\bar{y}}{\Sigma x^2 - n\bar{x}^2}$$

$$b = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

ومع ذلك نجد أن الأسلوب المباشر هو الأسلوب الأكثر استخداماً في التطبيقات.

تمرين : نعود إلى المثال المتعلق بنفقات الإعلان وحجم المبيعات والمطلوب : أوجد معادلة خط الانحدار.

y^2	x^2	xy	y	x
25	8100	450	5	90
4	6400	160	2	80
16	6400	320	4	80
36	7225	510	6	85
16	7569	344	4	87
49	8464	644	7	92
36	8100	540	6	90
49	9025	665	7	95
25	8649	465	5	93
25	7225	415	5	85
281	77157	4513	51	877

وبالتعويض بالمعادلتين نجد أن :

(1) $51 = 10a + b877$

(2) $4513 = 877a + b77157$

وبضرب المعادلة الأولى في (877) والثانية في (10) نحصل على :

$44727 = 8770a + 769129b$

$45130 = 8770a + 771570b$

وبطرح المعادلة الأولى من الثانية

$-403 = 0 - 2441b$

وبالتالي :

$$b = \frac{-403}{-2441} = 0.165$$

نعوض في المعادلة (1) نجد أن : $51 = 10a + 877\ (0.165)$

$a = -9.37$

وبذلك تكون المعادلة المطلوبة هي : $\hat{y} = -9.37 + 0.165x$

أي أنه كلما زاد الإنفاق على الإعلان بمقدار وحدة واحدة (ألف وحدة نقدية) زاد حجم المبيعات بمقدار (165) أي (165) ألف وحدة نقدية.

ويمكن استخدام هذه المعادلة لتقدير حجم المبيعات عند أي قيمة من الإنفاق فمثلاً إذا كان هذا الإنفاق هو (80) مثلاً فإن :

$$\hat{y} = -9.37 + 0.165(80) = 3.83$$

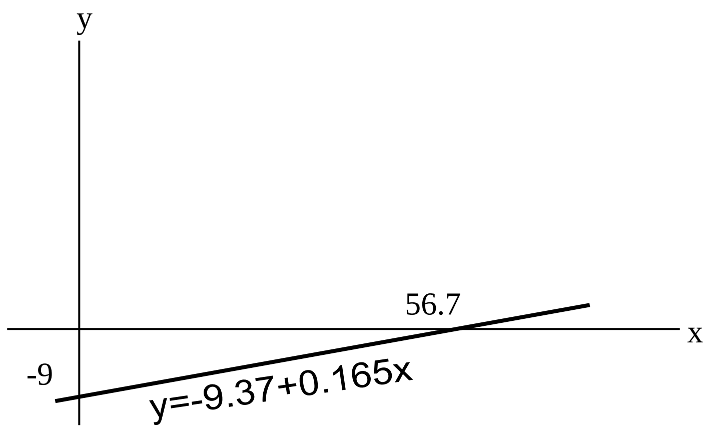
أي أن حجم المبيعات المتوقعة هو : (3830) وحدة .

● **رسم خط الإنحدار:** لتحديد أي خط مستقيم يكفي رسم نقطتين وبالعودة للتمرين المطلوب:

أرسم خط الانحدار بيانياً.

الحل :

$$x = 0 \Rightarrow y = -9.37$$
$$y = 0 \Rightarrow -9.37 + 0.165x = 0$$
$$x = 56.78$$



● **تفسير الثوابت** : إن (b) تمثل معدل التغير في y المناظر لتغير في (x) مقدار وحدة واحدة. ففي المثال السابق يزداد حجم المبيعات بمقدار (0.165) كلما ازداد الإنفاق على الإعلان وحدة واحدة (ألف وحدة). أي أن مردود الوحدة النقدية الواحدة في الإعلان . في المتوسط يبلغ 165 وحدة نقدية مبيعات.

● **العلاقة بين ميل خط الانحدار ومعامل الارتباط :**

توجد علاقة وثيقة بين ميل خط الانحدار (b) ومعامل الارتباط (r) والتي يمكن الحصول عليها بالصيغة التالية :

$$r = b\frac{S_x}{S_y}$$

أي أن المتغير (y) يزداد بمقدار جزء من وحدته المعيارية يعادل معامل الارتباط كلما ازداد المتغير (x) وحدة واحدة من وحداته المعيارية.

● **خط انحدار x على y :**

هو أحسن خط مستقيم يمثل خط الانتشار بحيث تكون مربعات الأخطاء في (y) حول هذا المستقيم أقل ما يمكن ومن ثم فإنه أصلح ما يمكن للتنبؤ بقيم (y) إذا عرفنا قيمة (x).

ولكن للتنبؤ بقيم x بمعلومية y يجب أن نحصل على الخط الذي يجعل :

1 . مجموع أخطاء x صفراً.

2 . مجموع مربعات أخطاء x أصغر ما يمكن.

وسنجد أن الثابتين (b`, a`) في المعادلة :

$$\hat{x} = a' + b'y + e$$

يحققان المعادلتين :

$$(1)\ \Sigma x = na' + b'\Sigma y$$

$$(2)\ \Sigma xy = a'\Sigma y + b'\Sigma y^2$$

ويسمى المستقيم الناتج بخط انحدار x على y والميل b' يرتبط بمعامل الارتباط بالعلاقة:

$$b = r\frac{S_y}{S_x}$$

مثال :

في المثال السابق فإن خط انحدار x على y يمكن الحصول عليه بحل المعادلتين :

$$(1)\quad 877 = 10a' + b'\,51$$

$$(2)\quad 4513 = a'51 + b'281$$

وهي تعطينا $b' = 1.92$ $a' = 77.9$

وبالتالي معادلة خط الانحدار هي :

$$\hat{x} = 77.9 + 1.92y$$

تجدر الإشارة هنا إلى أنه يمكن الحصول على تقديرات (b') ، (a') بالأسلوب المباشر :

$$b' = \frac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma y^2 - (\Sigma y)^2}$$

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

وبالتطبيق على المثال السابق نجد أن :

$$b' = \frac{10(4513) - (877)(51)}{10(281) - (51)^2}$$

$$a' = 87.7 - (1.92)(5.1) = 77.9$$

وهي نفس النتائج السابقة .

ومن خلال الاستعراض السابق نلاحظ أن :

I . أن كلا خطي الانحدار يمران بالنقطة $(\bar{x})$ ، $(\bar{y})$.

II . $r = \sqrt{b.b'}$

ونلاحظ أن : (b, b') يكون دائماً موجباً، لكن نحن نعلم أن إشارة (r) من إشارة (b) و (b').

وبالتطبيق على المثال السابق نجد أن :

$$r = \sqrt{b.b'}$$

$$r = \sqrt{(0.165)(1.92)} = 0.56$$

ويمكن التأكد من ذلك من خلال الصيغة المباشرة لحساب معامل الارتباط:

$$r = \frac{n\Sigma x\Sigma y - \Sigma x\Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2[n\Sigma y^2 - (\Sigma y)^2}}$$

$$= \frac{40.3}{71.5} = 0.56$$

ونلاحظ أن الإشارة موجبة لأن كلاً من (b) و (b') موجب أما لو كان كل من (b) و (b') سالباً فإنه يجب اختيار الإشارة السالبة.

ونشير هنا إلى أنه أيضاً يمكن استخدام معادلة خط الانحدار (x/y) في التنبؤ، فعلى سبيل المثال: أوجد القيمة المتوقعة للانفاق على الإعلان إذا كان حجم المبيعات (800) وحدة، نعوض في المعادلة :

$$\hat{x} = 77.9 + 1.92y$$

$$= 77.9 + 1.92(800) = 1613.9$$

القيمة المتوقعة للانفاق على الإعلان هي : 1613.9 وحدة نقدية.

10-5- الخطأ المعياري للتقدير :

لقد بيّنا في السابق كيفية استخدام معادلة خط الانحدار في التنبؤ لبيانات داخل حدود العينة ولبيانات خارج حدود العينة مع افتراض ثبات العلاقة بين المتغيرين (أي ستبقى العلاقة خطية بين المتغيرين) وعكس ذلك لايمكن إجراء هذا التقدير، ولكن السؤال الذي يطرح نفسه، ما مدى جودة هذا التنبؤ؟ أو بتعبير آخر: ما مدى ثقتنا في القيمة المقدرة طبقاً لهذه المعادلة؟ للجواب على هذا السؤال يجب معرفة الخطأ المعياري للتقدير (الانحرافات) ونسميه بالخطأ العشوائي والذي يعرف بأنه عبارة عن الجذر التربيعي لمتوسط مجموع مربع انحرافات القيم المقدرة عن القيم الفعلية للمتغير التابع وهو يشابه الانحراف المعياري لتوزيع تكراري الذي يقيس مدى التشتت حول الوسط الحسابي بينما الخطأ المعياري يقيس مدى التشتت حول خط الانحدار، ويبين هذا المقياس مقدار انحراف القيم الفعلية عن القيم المقدرة للمتغير التابع بواسطة معادلة خط الانحدار، وإذا كانت قيمته مساوية الصفر، هذا يعني أنه ليس هناك تشتتاً أو انتشاراً وأن جميع النقاط وقعت على خط الانحدار وبالتالي يكون الارتباط بين المتغيرين تام، ولتوضيح ذلك لدينا التمرين التالي .

تمرين : الجدول التالي يمثل سنوات الخبرة وحجم المبيعات بآلاف الوحدات النقدية لعدد من

مندوبي المبيعات في إحدى شركات التأمين .

رقم التابع	1	2	3	4	5	6	7	8	9	10
سنوات الخبرة (x)	1	3	4	4	6	8	10	10	11	13
حجم المبيعات (y)	80	97	92	102	103	111	119	123	117	136

والمطلوب : أحسب الخطأ المعياري للتقدير وبيّن ماذا تقيس هذه القيمة ؟

$(\hat{y}-\bar{y})^2$	$\hat{y}-\bar{y}$	$(y-\hat{y})^2$	$(y-\hat{y})$	$\hat{y}=80+4x$	y	x	رقم البائع
576	-24	16	-4	84	80	1	1
256	-16	25	5	92	97	3	2
144	-12	16	-4	96	92	4	3
144	-12	36	6	96	102	4	4
16	-4	1	-1	104	103	6	5
16	4	1	-1	112	111	8	6
144	12	1	-1	120	119	10	7
144	12	9	+3	120	123	10	8
256	16	49	-7	124	117	11	9
576	24	16	4	132	136	13	10
2272		170			1080	70	

إن معادلة خط الانحدار للتمرين أعلاه هي :

$$\hat{y}=80+4x$$

في هذا الجدول حصلنا على القيم $(\hat{y})$ بواسطة المعادلة، التي تأخذ بعين الاعتبار عدد سنوات الخبرة، ونلاحظ أنها أيضاً تختلف كثيراً عن المتوسط (108) وأن مجموع مربعات الانحرافات عن هذا المتوسط هي (2272) أي أن اختلاف قيم (x) هو المسؤول عن الجزء الأكبر من الفروق في قيم $(\hat{y})$ ونقول أن معادلة خط الانحدار قد خسرت جزءاً كبيراً من التباين في قيم (y) ويبقى جزء غير مفسر هو ما نسميه بالاخطاء العشوائي (الخطأ المعياري للتقدير) ويعطى بالعلاقة التالية :

$$Sy\hat{y}=\sqrt{\frac{\Sigma(y-\hat{y})^2}{n}}$$

بالتطبيق على بيانات التمرين :

$$4.12=\sqrt{\frac{170}{10}}$$

التفسير : هذا المقياس لايقبل التفسير لكن يمكن القول كلما صفرت هذه القيمة كانت الثقة أكثر بمعادلة خط الانحدار.

ونشير هنا، إلى أنه يمكن حساب الخطأ المعياري للتقدير بالعلاقات التالية :

$$Sy\hat{y}=\sqrt{\frac{\Sigma y-(a\Sigma y+b\Sigma xy)}{n}}$$

$$Sy\hat{y}=sy\sqrt{1-r^2}$$ (خط انحدار y/x)

$$Sy\hat{y}=sx\sqrt{1-r^2}$$ (خط انحدار x/y)

وبتطبيق العلاقة الأولى نجد أن :

$$\sum xy = 8128 \ , \ \sum y^2=119082$$

$$Sy\hat{y}=\sqrt{\frac{119082-(80*1080+4(8128)}{10}}=4.12$$

11-5- التباين المفسّر والتباين غير المفسّر :

نحن نعلم أن التباين الكلي = التباين المفسّر + التباين غير المفسّر.

ويعرّف التباين المفسّر $S^2\hat{y}$ بأنه عبارة عن جزء من التباين الكلي وهو التباين الذي يمكن تفسيره بمعادلة خط الانحدار وكلما زادت قيمته دل ذلك على إمكانية تحديد قيم (y) بدلالة (x) بدقة أكثر.

أما التباين غير المفسر فهو أيضاً جزء من التباين الكلي (مربع الخطأ المعياري للتقدير) وهو التباين الذي لايمكن تفسيره بمعادلة خط الانحدار وكلما صغرت قيمته كلما زادت الثقة أكثر بمعادلة خط الانحدار.

ويمكن التعبير عما سبق على النحو التالي :

$$SST = SSR + SSE$$

حيث أن :

SST : مجموع مربعات انحرافات قيم (y) الفعلية عن وسطها الحسابي $(\bar{y})$ ويعطى بالعلاقة التالية: $\Sigma(y-\bar{y})^2$

SSR : مجموع مربعات انحرافات قيم $(\hat{y})$ المقدرة عن الوسط الحسابي $(\bar{y})$ ويعطى بالعلاقة التالية: $\Sigma(\hat{y}-\bar{y})^2$

SSE : مجموع مربعات انحرافات قيم (y) الفعلية عن القيم المقدرة $(\hat{y})$ ويعطى بالعلاقة التالية: $\Sigma(y-\hat{y})^2$

$$\frac{\Sigma(y-\bar{y})^2}{n}=\frac{\Sigma(\hat{y}-\bar{y})^2}{n}+\frac{\Sigma(y-\hat{y})^2}{n}$$

$$\frac{\Sigma y^2}{b}-\left(\frac{\Sigma y}{n}\right)^2=\frac{b(\Sigma xy-\bar{x}\Sigma y)}{n}+\frac{\Sigma y^2-(a\Sigma y+b\Sigma xy)}{n}$$

وبالتطبيق على المثال السابق، نجد أن :

SST= 2272 + 170 = 2442 (التباين الكلي)

وباستخدام صيغ التباين المفسّر والتباين غير المفسر والمتضمنة ثوابت معادلة خط الانحدار نجد

أن :

$$SSR = \frac{4(8128-(7)*(1080)}{10} = 227$$ (التباين المفسّر)

$$SSE = (4.12)^2 = 17$$ (التباين غير المفسّر)

$$SST = \frac{119080}{10} - \left(\frac{1080}{10}\right) = 244$$ (التباين الكلي)

$$SST = 227 + 17 = 244$$

- **تحديد التباين المفسّر والتباين غير المفسر بيانياً :**

نحن نعلم أن معادلة خط الانحدار للتمرين السابق هي :

$$\hat{y} = 80 + 4x$$
$$x = 0 \Rightarrow \hat{y} = 80$$
$$y = 0 \Rightarrow 80 + 4x = 0$$
$$x = -20$$

- **العلاقة بين التباينات (المفسّر وغير المفسّر) ومعامل التحديد (r^2):**

يمكن إيجاد معامل التحديد (r^2) من خلال هذه العلاقة :

$$r^2 = \frac{s^2\hat{y}}{s^2y} = \frac{\text{التباين المفسر}}{\text{التباين الكلي}}$$

$$r^2 = 1 - \frac{s^2y\hat{y}}{s^2y}$$

وبالتطبيق على التمرين السابق نجد أن :

$$r^2 = \frac{2272}{2442} = 0.93$$ (بالاعتماد على الصيغ المباشرة)

أو

$$r^2 = \frac{227}{244} = 0.93$$ (بالاعتماد على الصيغ التي تتضمن ثوابت المعادلة)

- **ملاحظات :**

1) . كلما اقتربت القيم الفعلية (y) من خط الانحدار صَّغر الخطأ المعياري للتقدير.
2) . عندما يمر خط الانحدار بمركز الاحداثيات هذا يعني أن جميع نقاط الانتشار وقعت على هذا الخط وبالتالي الارتباط التام بين المتغيرين .
3) . عندما معامل الانحدار (b=0) فإن خط الانحدار ينطبق على خط ($\bar{y}$) وبالتالي الارتباط معدوم بين المتغيرين وأصبح ($s^2\hat{y} = 0$) التباين المفسّر .

4) . عندما $(y=\hat{y})$ فإن جميع نقاط الانتشار وقعت على خط الانحدار أي أن : $(sy\hat{y}=0)$ والارتباط تام بين المتغيرين.

5) . إذا كان : $(s^2 y\hat{y}=s^2 y)$ لايوجد علاقة ارتباطية بين (y, x)

6) . إذا كان : $(s^2 y\hat{y}=0)$ العلاقة تامة بين (y, x) أي $(r=1)$.

12-5- معامل الاقتران ومعامل التوافق :

. معامل التوافق بالتعريف : هو عبارة عن مقياس يقيس شدة العلاقة الارتباطية بين متغيرين وكل منهما ينقسم إلى صفتين في الجداول (2×2) ويعطى بالعلاقة التالية :

$$r_c = \frac{AD-BC}{AD+BC}$$

كلما اقتربت قيمة (r_c) من الواحد كانت العلاقة بين المتغيرين قوية جداً، وإذا كانت $(r_c>0.5)$ هذا يعني أن هناك ترابط (اقتران) بين المتغيرين (الظاهرتين) وبالعكس إذا كانت $(r_c<0.5)$ العلاقة ضعيفة بين المتغيرين وتأخذ عادة (r_c) بالقيمة المطلقة.

تمرين (1) : يبين الجدول التالي البيانات المستخدمة في دراسة العلاقة بين رضا أو عدم رضا مجموعة من المتعاملين مع أحد المصارف وكونهم موظفين في القطاع العام أو في القطاع الخاص.

موقع الوظيفة / درجة الرضا	في القطاع العام	في القطاع الخاص	المجموع
راضٍ	50	35	85
غير راضٍ	40	45	85

والمطلوب: قياس مقدار الترابط بين رضا أو عدم رضا المتعامل مع هذا المصرف كونه في القطاع العام أو القطاع الخاص.

$$r_c = \frac{(50)(45)-(35)(40)}{(50)(45)+(35)(40)} = 0.23$$

وبالتالي العلاقة ضعيفة بين المتغيرين (عدم وجود اقتران بين موقع الوظيفة، قطاع عام، قطاع خاص ودرجة الرضا).

. معامل التوافق :

يستخدم هذا المعامل عند دراسة العلاقة الارتباطية بين الظواهر التي تحمل أكثر من صفتين، لذلك هذا المعامل أعم من معامل الاقتران ويتم حسابه من خلال العلاقة التالية :

$$rA = \sqrt{\frac{G-1}{G}}$$

$$G=\Sigma\left(\frac{fij^2}{fifj}\right)$$

إذا كانت ($rA>0.5$) كان هناك علاقة ارتباطية (توافق)، وكلما اقتربت من الواحد دل ذلك على متانة العلاقة الارتباطية (التوافق).

تمرين (2) : أعيد تصنيف المتعاملين مع المصرف في التمرين السابق حسب مستواهم التعليمي فكانت النتائج على النحو التالي :

المستوى التعليمي / درجة الرضا	ثانوية وإعدادية	جامعة أو معهد	فوق جامعية	Σ
راض	20	25	50	95
غير راض	30	25	20	75
Σ	50	50	70	170

والمطلوب: حساب معامل التوافق للبيانات السابقة :

$$G=\Sigma\left(\frac{fij^2}{fifj}\right)$$

$$=0.08+0.024+0.13+0.167+0.376+0.08=1.073$$

$$rA=\sqrt{\frac{1.073-1}{1.073}}=0.26$$

وبالتالي العلاقة ضعيفة بين المتغيرين (عدم وجود توافق بين المستوى التعليمي ودرجة الرضا).

تمارين غير محلولة

1 . كانت حركة القروض والودائع (بملايين الوحدات النقدية) في مجموعة من البنوك التجارية كما يلي :

القروض (x)	75	89	84	77	110
الودائع (y)	92	112	116	90	85

1 . أوجد معادلتي خطي انحدار (x/y)، (y/x) .

2 . أوجد معامل الارتباط بين المتغيرين (y,x)

2 . في دراسة للعلاقة بين المتغيرين y, x تبين أن :

المقياس	$\bar{x}$	S
المتغير (x)	100	7
المتغير (y)	60	5

وبفرض أن معامل الارتباط بين المتغيرين = 0.7

والمطلوب :

1 . أوجد معادلة خط انحدار y/x

2 . أحسب قيمة x بفرض أن $y= 110$.

3 . حكمان في مسابقة طلب منهما تقييم 10 متسابقين قدموا الدرجات الآتية :

الحكم الأول	**10**	**8**	**9**	**8**	**7**	**6**	**5**	**6**	**7**	**8**
الحكم الثاني	**5**	**7**	**7**	**6**	**8**	**10**	**9**	**8**	**5**	**4**

والمطلوب :

إيجاد معامل ارتباط الرتب (سبيرمان) بين تقديرات الحكمين.

4 . الأرقام التالية تمثل سعر الدولار واليورو بالليرة السورية خلال عدة أيام مختلفة.

الدولار	**50**	**52**	**51**	**53**	**50**	**52**	**53**	**54**	**54**	**55**
اليورو	**54**	**52**	**55**	**57**	**54**	**58**	**60**	**62**	**63**	**65**

المطلوب :

1 . حدد نوع وطبيعة العلاقة بين المتغيرين

2 . أحسب معامل الارتباط بين قيمة الدولار وقيمة اليورو.

5 . ظاهرتان y,x توفرت عنهما البيانات التالية :

5	6	8	7	7	6	4	8	5	4	x
5	4	9	7	8	5	4	5	4	3	Y

المطلوب:

1 . حساب معادلة انحدار y/x

2 . التنبؤ بقيمة y عندما تكون $x=12$ بفرض أن العلاقة ستبقى خطية.

3 . حساب معامل الارتباط الخطي البسيط بين الظاهرتين.

6 . في دراسة إحصائية عن (20) شركة تتناول العلاقة بين المبيعات (y) ومصاريف الإعلان (x) توفرت لديك البيانات الآتية :

$\sum x^2= 80 \qquad \sum x=10$

$\sum xy=60 \quad \sum y^2=50 \quad \sum y=20$

المطلوب:

1 . أوجد معامل الارتباط الخطي (بيرسون) بين قيم (x)، (y)

2 . ماهي قيمة المبيعات المتوقعة عند صرف (100) وحدة نقدية على الإعلان.

3 . أوجد معادلة خط الانحدار بفرض أن العلاقة خطية.

7 . تم التوصل من خلال دراسة أجريت عن العلاقة بين درجات الاحصاء (x) ودرجات الرياضيات (y) لعينة ملونة من (25) طالب إلى النتائج الآتية :

$\sum x^2 = 106250 \qquad \sum x = 1625$

$\sum xy = 98300$

$\sum y^2=97750 \qquad \sum y=1500$

والمطلوب :

a) أوجد معادلة خط الانحدار وما الغاية من هذه المعادلة.

b) ماهو تقديرك لدرجات امتحان الرياضيات لطالب حصل على (65) درجة في الاحصاء.

8 . إذا كانت درجات ثمانية من الطلبة في مادتي الرياضيات والاقتصاد كالآتي:

75	60	75	80	75	65	60	70	**الرياضيات**
70	75	65	70	60	75	80	50	**الاقتصاد**

والمطلوب:

أوجد معامل ارتباط الرتب (سبيرمان).

الفصل السادس

السلاسل الزمنية

. مقدمة

. تعريف السلسلة الزمنية

. الاتجاه العام

. معادلة خط الاتجاه العام الخطية

. معادلة خط الاتجاه العام غير الخطية

. تمارين غير محلولة

6-1- مقدمة :

يعتبر التنبؤ وسيلة هامة وأساسية في وضع الخطط التنموية لاقتصاد أي بلد، يستخدمها المخطط لرؤية الوضع المستقبلي للظاهرة محل الدراسة.

وبالتالي يعرف التنبؤ بأنه تقدير كلي للقيم المتوقعة للمتغيرات التابعة في المستقبل القريب بناءً على ماهو مقترح من معلومات عن الماضي والحاضر، والتحليل الاحصائي للسلسلة الزمنية هو محاولة اكتشاف النسق الذي سارت عليه هذه الظاهرة في الماضي للحصول على نموذج رياضي يساعدنا في تقدير قيم الظاهرة في المستقبل.

6-2- تعريف السلسلة الزمنية :

عبارة عن مجموعة القيم لظاهرة ما خلال فترة زمنية معينة والهدف منها التعرف على التغيرات التي تطرأ على قيم هذه الظاهرة.

6-3- مكونات السلسلة الزمنية :

إن كل قيمة من قيم السلسلة الزمنية تتكون من أربعة أجزاء :

1. **الاتجاه العام (T):** وهو من أهم الأجزاء في التنبؤ طويل الأمد، ويمكن القول أن الاتجاه العام عبارة عن المسار للخط الممثل لقيم الظاهرة، وهو من التغيرات المنتظمة التي تظهر آثارها بوضوح بعد فترة زمنية طويلة الأمد والتغيرات يمكن أن تكون نحو الزيادة أو النقصان وقد تتجه نحو الزيادة فترة طويلة. مثال : معدلات الناتج القومي وتغيراته خلال (20) سنة.
2. **التغيرات الموسمية (S)** : عبارة عن تغيرات منتظمة خلال فترات زمنية لاتتعدى السنة (يوم، أسبوع، شهر، فصل، سنة) مثال: نشاط المصارف في بداية الشهر، أو العمالة في فصل الصيف أو كمية الأمطار في شهر كانون ... إلخ.
3. **التغيرات الدورية (c)** : عبارة عن التغيرات التي تطرأ على السلسلة الزمنية وطولها أكثر من سنة مثل دورات نمو الرخاء الاقتصادي.
4. **التغيرات العشوائية (I):** عبارة عن التغيرات التي تحدث بصورة فجائية مثل: الأوبئة، الزلازل، البراكين.

وعلى هذا الأساس تصبح دراسة السلاسل الزمنية عبارة عن تقدير أو تحديد وحساب كل من مكونات السلسلة (y) الأربع ومدى تأثيرها خلال فترة السلسلة كلها، ونحن سنتهتم هنا فقط بالاتجاه العام على اعتبار أنه من أهم الأجزاء في التنبؤ طويل المدى .

6-4- الاتجاه العام (T) :

في كثير من الحالات وبعد دراسة البيانات الاحصائية يتم تحديد منحني السلسلة (نوع وطبيعة العلاقة) الذي يكون عادة أحد نوعين: خطي وهو مجال دراستنا الآن أو غير خطي لذلك عندما تكون العلاقة خطية يمكن التعبير عن ذلك بالمعادلة الخطية التالية:

$$\hat{y} = a + bt$$

حيث :

t : الزمن.

b : الزيادة أو النقصان وحدة واحدة عند تغير الزمن.

6-5- معادلة خط الاتجاه العام الخطية :

يتم إيجاد معادلة خط الاتجاه العام الخطية بعدة طرق منها :

a) **الطريقـة البيانيـة :** وتعتمد هذه الطريقة على رسم شكل الانتشار للظاهرة المدروسة باعتبار المحور الأفقي يمثل الزمن (t) والمحور العامودي يمثل قيم الظاهرة (y) وقد تم تناول هذه الطريقة بالتفصيل سابقاً في بحث الارتباط والانحدار.

b) **طريقة المتوسطات المتحركة:** وتعتمد هذه الطريقة على أخذ المتوسط لعدد من السنوات أو الفصول للظاهرة المدروسة ويعتبر الناتج ممثلاً للسنة التي تقع في منتصف مجموعة السنوات والوسط المتحرك يمكن أن يكون من (3) سنوات أو (4) سنوات أو (5) سنوات.

c) **طريقـة المتوسـطات النصـفية:** وتعمتد هذه الطريقة على تقسيم البيانات في السلسلة الزمنية إلى مجموعتين متساويتين بحيث نقوم بحساب متوسط الجزء الأول ومتوسط الجزء الثاني لنصل لحساب القيم الاتجاهية للظاهرة المدروسة.

. إن الطرق السابقة يعاب عليها بأنها تقريبية وتعطي فكرة عن معادلة الاتجاه العام وبالتالي لايمكن تحديد قيم الاتجاه العام بدقة.

d) **الطريقـة الرياضـية:** تعتبر أدق من الطرق السابقة باعتبارها تعتمد الأسلوب الرياضي ويتم إيجاد معادلة خط الاتجاه العام بأسلوبين.

I . أسلوب المربعات الصغرى : نحن نعلم أن معادلة خط الإتجاه العام من الشكل :

$$\hat{y} = a + bt$$

ويتم إيجاد ثوابت هذه المعادلة خلال حل جملة المعادلتين الآتيتين :

$$(1) \quad \sum y = na + b\sum t$$
$$(2) \quad \sum yt = a\sum t + b \sum t^2$$

لنحصل فيما بعد على معادلة خط الاتجاه العام

II . الأسلوب المباشر : ويقوم على أساس حل المعادلتين السابقتين لنحصل على الثوابت.

$$b = \frac{n\Sigma ty - (\Sigma t)(\Sigma y)}{n\Sigma t^2 - (\Sigma t)^2}$$

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

مثال :

الجدول التالي يمثل عدد الأطباء البشريين في إحدى المدن العربية خلال الفترة (2000-2004) .

السنة	2000	2001	2002	2003	2004
عدد الأطباء	1338	1555	1921	2133	2263

والمطلوب :

1 . أوجد معادلة خط الاتجاه العام.

2 . ماهو العدد المتوقع للأطباء في سنة (2007) ثم سنة (2012)

الحل :

بما أن السنوات متتالية (وعددها فردي) فلو أخذنا السنة الوسيطة كنقطة الصفر (سنة الأساس) فسيكون لدينا الجدول التالي :

t^2	ty	Y	T
4	-2676	1338	-2
1	-1555	1555	-1
0	0	1921	0
1	2133	2133	1
4	4526	2263	2
10	2428	9210	0

وبالتعويض نجد أن :

(1) $9210 = 5a + b\ (0)$

(2) $2428 = a(0) + b(10)$

أي أن :

$$b = \frac{2428}{10} = 242.8 \qquad a = \frac{9210}{5} = 1842$$

وبالتالي معادلة خط الاتجاه العام : $\hat{y} = 1842 + 242.8t$

حيث أن سنة الأساس هي (2002) والفترة الزمنية سنة كاملة أي أن متوسط النمو السنوي هو (242.8) طبيباً وعدد الأطباء في سنة الأساس (2002) هو (1842) طبيباً.

2 . عدد الأطباء المتوقع سنة (2007) سيكون :

$$\hat{y} = 1842 + 242.8t$$

$=1842 + 242.8\ (5) = 3056$ طبيباً

وبالتالي عدد الأطباء المتوقع سنة (2007) هو (3056) طبيباً .

$\hat{y} = 1842 + 242.8t(10) = 2470$ طبيباً

وبالتالي عدد الأطباء المتوقع سنة (2012) هو (2470) طبيباً.

أما إذا كان عدد السنوات زوجياً فيتم التعامل على النحو التالي :

مثال (2) : الجدول التالي يتضمن الودائع المصرفية بأحد المصارف الخاصة في إحدى الدول (بملايين الوحدات النقدية) خلال الفترة (1999 ,2004).

السنوات	1999	2000	2001	2002	2003	2004
الودائع	31	46	55	60	77	69

والمطلوب :

1 . أوجد معادلة خط الاتجاه العام.

2 . حساب القيمة المتوقعة للودائع المصرفية في عام (2008)

الحل : باعتبار عدد السنوات زوجي سوف نختار السنة (2001.5) سنة الأساس، لكن هذا سوف يجعل قيم المتغير المستقل فيها كسور لذلك سنضرب الناتج بـ (2) للتخلص من الكسور.

t	Y	Ty	t^2
-5	31	-155	25
-3	46	-138	9
-1	55	-55	1
1	60	60	1
3	77	231	9
5	69	345	25
0	338	288	70

وبالتعويض في المعادلتين نجد أن :

(1) $338 = 6a + b\,(0)$

(2) $288 = a(0) + b(70)$

أي أن :

$$a = \frac{338}{6} = 56.33 \qquad b = \frac{288}{70} = 4.11$$

وبالتالي معادلة خط الاتجاه العام : $\hat{y} = 56.33 + 4.11t$

حيث أن سنة الأساس هي منتصف عام (2001) والفترة الزمنية نصف سنة.

ملاحظة (1) : من خلال المثالين السابقين نلاحظ أننا قد استخدمنا طريقة المربعات الصغرى في حساب ثوابت معادلة خط الاتجاه العام لكن مع افتراض أن ($\sum t=0$) نجد أن: المعادلة (1) أصبحت:

$$\sum y = na$$

وبالتالي الثابت (b) يمكن حسابه من خلال : $a = \bar{y}$

من المعادلة (2) نجد أن : $\sum ty = b\sum t^2$

ومنه نجد أن الثابت (b) يمكن حسابه من خلال :

$$b = \frac{\Sigma ty}{\Sigma t^2}$$

ملاحظة (2) :

في حالة وجود عدد زوجي من السنوات تكون سنة الأساس هي منتصف السنة الوسيطة والوحدة الزمنية نصف سنة وعليه يجب علينا هنا مضاعفة معدل التغير (النصف سنوي) للحصول على معدل التغير السنوي وبالتالي يمكن القول أن :

متوسط النمو السنوي للودائع هو (8.22) مليون وحدة نقدية .

2 . القيمة المتوقعة للودائع المصرفية في عام (2008)

$$\hat{y} = 56.33 + 8.22t$$

$$= 56.33 + 8.22\ (8.5) = 126.2$$

أي أن القيمة المتوقعة للودائع المصرفية عام (2008) هو 126.2 مليون وحدة نقدية.

مثال (3) : الجدول التالي يتضمن المبيعات السنوية (y) لشركة صناعية في إحدى الدول (بملايين الوحدات النقدية) خلال الفترة (2000-2004) .

السنة	2000	2001	2002	2003	2004
المبيعات	2	6	10	13	16

والمطلوب :

1 . أوجد معادلة خط الاتجاه العام . 2 . أوجد حجم المبيعات المتوقعة في عام 2006

الحل : باعتبار عدد السنوات فردي سوف نأخذ السنة الوسيطة (سنة الأساس) فسيكون لدينا الجدول التالي :

t	Y	ty	t^2
-2	2	-4	4
-1	6	-6	1
0	10	0	0
+1	13	13	1
+2	16	32	4
6	47	35	10

وبالتعويض في المعادلتين نجد أن :

$$(1) \quad 47 = 5\,a + b\ (0)$$

$$(2) \quad 35 = a\ (0) + b\ (0)$$

أي أن :

$$a = \frac{\Sigma y}{n} \Rightarrow a = \frac{47}{5} = 9.4$$

$$b = \frac{\Sigma ty}{\Sigma t^2} \Rightarrow b = \frac{35}{10} = 3.5$$

وبالتالي معادلة الاتجاه العام هي :

$$\hat{y} = 9.4 + 3.5t$$

حيث أن سنة الأساس هي (2002) والفترة الزمنية سنة كاملة أي أن متوسط النمو السنوي للمبيعات هو (3.5) مليون وحدة نقدية وحجم المبيعات في سنة الأساس (2002) هو (9.4) مليون.

2 . حجم المبيعات المتوقعة في عام (2006)

$$\hat{y} = 9.4 + 3.5t$$
$$= 9.4 + 3.5(4) = 23.4$$

وبالتالي حجم المبيعات المتوقع سنة (2006) هو)(23.4) مليون وحدة نقدية .

في دراستنا للسلاسل الزمنية نجد في كثير من الحالات ضرورة التحويل من الوحدات السنوية إلى الوحدات الأسبوعية أو الشهرية أو الربع سنوية .. إلخ وفي مثل هذه الحالات يتوجب تقسيم كلاً من الثوابت (a) ، (b) وكذلك (x) على (52) أسبوعياً أو (12) شهراً أو (4) (ربع السنة) حسب الترتيب وبالتطبيق على المثال (3) وكان المطلوب :

1 . حساب المبيعات الربع سنوية (أي كل ثلاثة أشهر)

تصبح المعادلة على شكل ربع السنوية التالية :

$$\hat{y} = \frac{a}{4} + \frac{b}{4}\left(\frac{x}{4}\right)$$

وهذا يعني عندما (x=0) تعني الأول من كانون الثاني (2002)، بينما (x=1) تعني الأول من شهر نيسان (2002) أي بعد ثلاثة أشهر من التاريخ السابق وهكذا ...

ملاحظة :

في بعض الأحيان تعطى قيم الاتجاه العام كل ربع سنة ولكنها محسوبة بالأسلوب السنوي وفي مثل هذه الحالات نجد أن معادلة الاتجاه العام تمثل أربعة أضعاف معادلة الاتجاه العام الربع سنوية وتصبح المعادلة السابقة على النحو التالي :

$$\hat{y} = 4\left(\frac{a}{4} + \frac{b}{4}\left(\frac{x}{4}\right)\right)$$

تمارين غير محلولة

1 . الجدول التالي يتضمن متوسط الإنتاج السنوي لاحدى المؤسسات في احدى الدول خلال الفترة (1999-2003) بملايين الوحدات :

السنة	1999	1997	1998	1999	2000	2001	2002	2003
الانتاج بملايين الوحدات	6	5	7	4	5	5	6	7

والمطلوب:

1 . بفرض وجود علاقة خطية، أوجد معادلة خط الاتجاه العام للإنتاج .

2 . قدر متوسط الانتاج المتوقع في هذه المؤسسة عام (2010).

2 . الجدول التالي يتضمن مبيعات إحدى الشركات بملايين الوحدات النقدية في إحدى الدول خلال الفترة (1998-2005) .

السنة	1998	1999	2000	2001	2002	2003	2004	2005
المبيعات	9	10	13	10	12	14	14	16

والمطلوب:

1 . أوجد معادلة خط الاتجاه العام، بفرض أن العلاقة خطية.

3 . الجدول التالي يتضمن مبيعات إحدى الشركات الكبرى بملايين الوحدات النقدية في إحدى الدول خلال الفترة (1995-2001).

السنة	1995	1996	1997	1998	1999	2000	2001
المبيعات	12	12	15	14	16	16	11

والمطلوب :

1 . أوجد معادلة خط الاتجاه العام بفرض أن المبيعات تأخذ شكل خطي.

2 . أوجد قيمة المبيعات المتوقعة عام (2005) لهذه الشركة.

4 . الجدول التالي يتضمن أرباح إحدى الشركات الصناعية بملايين الوحدات النقدية في إحدى الدول خلال الفترة (1998-2005)

السنة	1998	1999	2000	2001	2002	2003	2004	2005
المبيعات	2	4	5	7	8	6	8	11

والمطلوب: 1 . أوجد معادلة خط الاتجاه العام

2 . قدر الأرباح المتوقعة لهذه الشركة عام (2010)

الفصل السابع
الأرقام القياسية

. تعريف الرقم القياسي
. استخدامات الأرقام القياسية
. أهمية الأرقام القياسية
. الأرقام القياسية البسيطة
. الأرقام القياسية المرجحة
. مناسيب الأسعار
. مناسيب الكميات
. اختبارات الأرقام القياسية
. فوائد الأرقام القياسية
. تمارين غيرمحلولة

الأرقام القياسية

1-7- تعريف الرقم القياسي :

هو مقياس إحصائي لقياس التغيرات الحاصلة في متغير واحد، أو مجموعة من المتغيرات، لظاهرة معينة، خلال فترة زمنية معينة أو من منطقة جغرافية إلى أخرى. ويعرف الرقم القياسي بأنه مؤشر إحصائي يستخدم في قياس التغير الحاصل في ظاهرة من الظواهر الاقتصادية، أو الاجتماعية، ويستخدم لتوضيح التحسن أو عدم التحسن في خطة زمنية معينة أو لتوضيح زيادة أو نقصان الانتاجية في مصنع ما أو لدولة معينة.

فمثلاً الرقم القياسي للناتج القومي لدولة ما يعرف ويحسب من العلاقة التالية :

$$I(n/0) = \frac{\text{كمية الإنتاج القومي في سنة المقارنة}}{\text{كمية الإنتاج القومي في سنة الأساس}} \times 100$$

ومن خلال هذه النسبة المئوية يمكننا معرفة مقدار ارتفاع أو انخفاض أو ثبات الناتج القومي وهو ما يعكس تحسن الأداء والادارة والعمالة أو البطالة ... إلخ.

وهذه العلاقة مفيدة أيضاً في معرفة الرقم القياسي لمنتج سلعي ما، فعلى سبيل المثال لو كان لدينا الرقم القياسي الحالي لمنتج سلعي ماهو (95%) عن الرقم الأساسي هذا يدل على أن هناك نقصاً بمقدار (5%) عن كمية الإنتاج في السنة الأساس ولو افترضنا أن كمية الانتاج في سنة الأساس كانت (950) وحدة، فإن مقدار النقص هو :

$$950 = \frac{5}{100} = 47.5 \approx 48 \text{ وحدة}$$

وبالمثل لو كان الرقم القياسي هو (115%) فهذا يعني أن هناك زيادة في كمية الإنتاج عن سنة الأساس بمقدار (15%) أي ما يعادل (146=145.5) وحدة .

ويمكن القول أن النقصان قد يعكس خصائص سلبية مثل نقص المواد الأولية أو العمالة أو سوء الإدارة ... إلخ وبالعكس نجد أن الزيادة قد تعكس خصائص إيجابية مثل توفر المواد الأولية وحسن التسويق والجودة ... الخ.

. العلاقة بين معدل النمو والرقم القياسي :

ويتضح ذلك من خلال التالي :

. الرقم القياسي يعبر دوماً عن تغير الظاهرة المدروسة بالزيادة أو النقصان

$$I(n/0) = \frac{y_n}{y_0} \times 100$$

. معدل النمو يصف مدى التغير الذي يظهر في ظاهرة ما

$$a(n/0) = \frac{\Delta y}{y_0} = \frac{y_n - y_0}{y_0}$$

حيث أن :

y_0: حجم الظاهرةعند سنة الأساس.

y_n : حجم الظاهرة في سنة المقارنة .

. معدل النمو = (الرقم القياسي . 1) .

. الرقم القياسي : مقياس إحصائي مصمم لإظهار التغيرات في متغير أو مجموعة متغيرات مرتبطة من حيث الزمان أو المكان.

7-2- استخدامات الأرقام القياسية :

يمكن إيجاز استخدامات الأرقام القياسية فيمايلي :

- مقارنة تكاليف المعيشة من عام بالنسبة لعام سابق له.
- قياس التغيرات الحاصلة في الظواهر الاقتصادية (الصادرات، الواردات، الإنتاجية)
- التنبؤ بالأعمال التجارية والنشاطات الاقتصادية المستقبلية للحصول على معلومات تفيد في عملية اتخاذ القرار وترشيده.

7-3- أهمية الأرقام القياسية :

ويبدو ذلك من خلال التالي :

- قياس التغير الحاصل في الأسعار في فترة زمنية معينة لاكتشاف سبب التغير وأثره على النشاط الاقتصادي حتى يمكن التحكم به .
- قياس التغير الحاصل في تكاليف المعيشة في مدينة ما خلال سنة معينة مع تكاليف المعيشة خلال سنوات سابقة.
- قياس التغير الحاصل في مستوى الدخل الحقيقي أو الاسمي بمعنى آخر لمعرفة القوة الشرائية لوحدة النقد مع العلم أن :

$$\text{القوة الشرائية} = \frac{1}{\text{الرقم القياسي للأسعار}} = \frac{1}{\text{الرقم القياسي لتكاليف المعيشة}}$$

$$\text{الدخل الحقيقي} = \frac{\text{الدخل في سنة المقارنة}}{\text{الرقم القياسي لتكاليف المعيشة}} \times 100$$

- قياس التغير الحاصل في حجم الإنتاج الصناعي والزراعي والموجودات في المخازن والتجارة الخارجية والمبيعات.
- قياس التغير الحاصل في عدد المشتغلين عامة في كل نشاط من نواحي النشاط الاقتصادي

كل على حده، والتغير الحاصل في عدد العاطلين عن العمل (مقياس البطالة).

ملاحظة : إن الغاية من إنشاء الرقم القياسي هو تحديد مدى التغير ، **لكن هناك صعوبات تواجه الاحصائي عند تكوين الرقم القياسي،** من هذه الصعوبات :

a) تحديد سنة الأساس التي نقارن بها، بشكل عام نستطيع القول أنه يجب أن تكون سنة طبيعية خالية من أية مشكلات (كوارث، أوبئة، أزمات...)

b) اختيار السلع التي ستدخل في حساب الرقم القياسي ويعتمد هذا الاختيار على نوع الرقم القياسي مثلاً: الرقم القياسي لنفقات المعيشة لطبقة معينة في الناس يختلف عن الرقم القياسي لكافة طبقات المجتمع) .

c) اختيار الأسلوب المناسب لحساب الرقم القياسي، حيث يفضل اختيار الأسلوب الذي يجتاز أكبر عدد ممكن من اختبارات الأرقام القياسية.

4-7- الأرقام القياسية البسيطة :

تلعب الأرقام القياسية دوراً هاماً في الحسابات العملية للدول أو الشركات وحتى الفرد لما لها من انعكاسات اقتصادية هامة فزيادة الرقم القياسي للأسعار البضائع قد يعني التضخم الاقتصادي وبعكسه قد يعني الركود بكافة النواحي الاقتصادية، ويعرف الرقم القياسي للسعر بأنه النسبة المئوية لسعر سلعة معينة خلال سنة معينة إلى سعر نفس السلعة خلال سنة الأساس ويعطى بالعلاقة التالية :

$$PI = \frac{P_n}{P_0} \times 100$$

حيث أن :

P_0 : السعر في سنة الأساس .

P_n : السعر في سنة المقارنة.

وهناك عدة أنواع للأرقام القياسية للأسعار منها

. الرقم القياسي التجميعي البسيط للأسعار :

يمكن إيجاد الرقم القياسي التجميعي البسيط للأسعار من الصيغة التالية :

الرقم القياسي التجميعي البسيط $P_I(S) = \frac{\Sigma P_n}{\Sigma P_0} \times 100$

إن الرقم القياسي البسيط يعكس تطور أسعار بين فترتين زمنيتين، وبالرغم من سهولة التعامل مع هذا الرقم إلا أن نتيجته تتغير عند تغير الوحدة المستعملة في قياس السعر لإحدى السلع المدروسة.

5-7- الأرقام القياسية المرجحة :

لتلافي عيوب الرقم القياسي البسيط التجميعي نعطي الأهمية النسبية للسلع الداخلة في تكوين الرقم القياسي فيكون الناتج رقماً قياسياً مرجحاً، مثلاً: عند تركيب رقم قياسي للأجور يجب ترجيحه بعدد عمال كل فئة من فئات الأجور، وعند تركيب رقم قياسي للكميات يجب ترجيحه بالأسعار وهكذا. وهناك عدة أرقام قياسية مرجحة:

I . الرقم القياسي للأسعار مرجحاً بكميات سنة الأساس (رقم لاسبير للأسعار):

اقترح لاسبير ترجيح الرقم القياسي للأسعار بكميات سنة الأساس، وأصبح يسمى الرقم القياسي الناتج بهذه الطريقة باسم رقم لاسبير للأسعار ويحسب من العلاقة التالية:

$$P_I(L) = \frac{\Sigma p_n q_0}{\Sigma p_0 q_0} \times 100 \quad \text{رقم لاسبير للأسعار}$$

حيث p_n أسعار سنة المقارنة، أسعار سنة الأساس، p_0 كميات سنة المقارنة، q_n كميات سنة الأساس، q_0

إن رقم لاسبير يعبر عن أثر التغير في السعر كما لو بقيت الكميات المشتراة في سنة الأساس هي نفسها في سنة المقارنة، لذلك يعتبر هذا الرقم غير واقعي لأنه قائم على جمود الاستهلاك لاسيما إذا كان هناك تباعد بين فترة الأساس والمقارنة، لكنه يوفر الوقت والمال عند جمع البيانات عن الكميات والأسعار.

مثال : يوضح الجدول التالي أسعار وكميات أربع سلع مباعة في سنتي 2004 ,2005

2005		2004		
كميات المبيعات	أسعار	كميات المبيعات	أسعار	السلعة
250	40	200	28	A
360	20	300	16	B
460	15	400	10	C
660	10	600	4	D

والمطلوب: حساب :

a) الرقم القياسي التجميعي البسيط للأسعار.

b) الرقم القياسي المرجح للأسعار (رقم لاسبير للأسعار).

الحل :

الرقم القياسي التجميعي البسيط للأسعار = $\frac{85}{58} \times 100 = 146.55\%$

b) رقم لاسبير للأسعار:

السلعة	2004		2005			
	أسعار	كميات المبيعات	أسعار	كميات المبيعات	p_0q_0	p_nq_0
A	28	200	40	250	5600	8000
B	16	300	20	360	4800	6000
C	10	400	15	460	4000	6000
D	4	600	10	660	2400	6000
المجموع	58	1500	85	1730	16800	26000

رقم لاسبير للأسعار = $\frac{26000}{16800}\times 100 = 154.76\%$

هذا يدل على أن الأسعار عام (2005) ارتفعت عن مستواها عام (2004) بمقدار (54.76%) أي أن هذا الرقم يقيس تغير الأسعار على فرض أن مستوى الحياة بقي دون تغيير.

II . الرقم القياسي للأسعار مرجحاً بكميات سنة المقارنة (رقم باش للأسعار):

قدم باش رقماً قياسياً آخر استخدم فيه كميات سنة المقارنة لترجيح الأسعار ويعطى بالعلاقة التالية: رقم باش للأسعار = $P_I(P)=\frac{\Sigma p_n q_n}{\Sigma p_0 q_n}\times 100$

إن رقم باش يفيد في قياس التكلفة الاجمالية لمجموعة سلعية في سنة المقارنة بالنسبة لتكلفة هذه المجموعة السلعية في سنة الأساس، لكنه يتطلب تكلفة من حيث الوقت والمال عند جمع البيانات اللازمة لحسابه.

مثال : يبين الجدول التالي أسعار وكميات أربع سلع مباعة في سنتي 2004م 2005 كمايلي:

	2004		2005	
السلعة	أسعار	كميات المبيعات	أسعار	كميات المبيعات
A	28	200	40	250
B	16	300	20	360
C	10	400	15	460
D	4	600	10	660

والمطلوب: أحسب رقم باش للأسعار.

الحل :

السلعة	2004		2005			
	أسعار	كميات المبيعات	أسعار	كميات المبيعات	p_0q_0	p_nq_0
A	28	200	40	250	7000	10000
B	16	300	20	360	5760	7200
C	10	400	15	460	4600	6900
D	4	600	10	660	2640	6600
المجموع	58	1500	85	1730	20000	30700

رقم باش للأسعار $\frac{30700}{20000}\times 100 = 153.5\%$

التفسير الاقتصادي : لو كنا نعيش في عام (2004) بنفس مستوى الحياة عام (2005) فإننا نشعر بأن الأسعار ارتفعت بمقدار (53.5%)

III . الرقم القياسي الأمثل للأسعار (رقم فيشر للأسعار) :

يطلق عليه اسم الرقم القياسي الأمثل، وهو عبارة عن المتوسط الهندسي لكل من رقمي لاسبير وباش أي الجذر التربيعي لحاصل ضرب رقم لاسبير برقم باش. ويتم حساب رقم فيشر للأسعار بالعلاقة التالية:

$$\text{رقم فيشر للأسعار} = \sqrt{\text{رقم لاسبير} \times \text{رقم باش}}$$

مثال : إذا كان رقم لاسبير للأسعار 154.76 ، ورقم باش للأسعار 153.5%

أوجد رقم فيشر للأسعار

الحل : رقم فيشر للأسعار = رقم لاسبير × رقم باش $P_{I(f)}$

$$\sqrt{154.76\% * 153.5\%} = 154.13\%$$

V . رقم مارشال أيجورث القياسي للأسعار:

يستخدم رقم مارشال أيجورث القياسي الصيغة التجميعية المرجحة باستخدام طريقة السنة النموذجية حيث الأوزان هي المتوسط الحسابي لكميات سنة الأساس وكميات سنة المقارنة، ويتم حسابه من خلال العلاقة التالية :

رقم مارشال . أيجورث القياس للأسعار $$P_I(Me) = \frac{\Sigma p_n(q_0 + q_n)}{\Sigma p_0(q_0 + q_n)} \times 100$$

مثال :

	2004		2005	
السلعة	أسعار	كميات المبيعات	أسعار	كميات المبيعات
A	28	200	40	250
B	16	300	20	360
C	10	400	15	460
D	4	600	10	660

والمطلوب : أوجد رقم مارشال . أيجورث القياسي للأسعار؟

الحل :

السلعة	2004		2005				
	أسعار Po	كميات q_o	أسعار P_n	كميات q_n	q_o+q_n	$p_n(q_o+q_n)$	$p_o(p_o+q_n)$
A	28	200	40	250	450	18000	12600
B	16	300	20	360	660	13200	10560
C	10	400	15	460	860	12900	8600
D	4	600	10	660	1260	12600	5040
المجموع						56700	36800

رقم مارشال . أيجورث القياسي للأسعار $\frac{56700}{36800} \times 100\% = 154.08\%$

التفسير : هذا يعني أن مستوى الحياة في سنة (2004) مساوياً للوسط الحسابي لمستوى الحياة في سنتين و (2004 وَ 2005) وأردنا أن نعيش بنفس مستوى الحياة في عام (2005) لشعرنا أن الأسعار ارتفعت بمقدار (54%).

● **الرقم القياسي التجميعي البسيط للكميات :**

في حساب الرقم القياسي التجميعي البسيط للكميات تؤخذ كميات سنة المقارنة وتجمع ثم تقسم على مجموع كميات سنة الأساس كما في العلاقة الآتية :

الرقم القياسي التجميعي البسيط للكميات = $q_I(S)=\frac{\Sigma q_n}{\Sigma q_n}\times 100$

مثال :

يوضح الجدول التالي أسعار وكميات أربع سلع في سنتي 2004 و 2005

	2004		2005	
السلعة	**أسعار**	**كميات المبيعات**	**أسعار**	**كميات المبيعات**
A	**28**	**200**	**40**	**250**
B	**16**	**300**	**20**	**360**
C	**10**	**400**	**15**	**460**
D	**4**	**600**	**10**	**660**

والمطلوب: أوجد الرقم القياسي التجميعي البسيط للكميات.

الحل :

الرقم القياسي التجميعي البسيط للكميات = $\frac{1730}{1500}\times 100=155.3\%$

ويوجد ثلاثة أرقام قياسية مرجحة للكميات هي :

II . الرقم القياسي للكميات مرجحاً بأسعار سنة الأساس (لاسبير للكميات):

يتم حساب الرقم القياسي مرجحاً بأسعار سنة الأساس وفق المعادلة الآتية :

رقم لاسبير للكميات = $q_I(L)=\frac{\Sigma p_0 q_n}{\Sigma p_0 q_0}\times 100$

مثال : يوضح الجدول التالي أسعار وكميات أربع سلع مباعة في سنتي 2000 وَ 2005

	2004		2005	
السلعة	**أسعار**	**كميات المبيعات**	**أسعار**	**كميات المبيعات**
A	**28**	**200**	**40**	**250**
B	**16**	**300**	**20**	**360**
C	**10**	**400**	**15**	**460**
D	**4**	**600**	**10**	**660**

والمطلوب : أوجد رقم لاسبير للكميات ؟

الحل :

p_oq_n	q_oq_o	2005		2004		السلعة
		كميات q_n	أسعار P_n	كميات q_o	أسعار Po	
7000	5600	250	40	200	28	A
9300	4800	360	20	300	16	B
4600	4000	460	15	400	10	C
2640	2400	660	10	600	4	D
23600	16800	1730	85	1500	58	المجموع

رقم لاسبير للكميات = $\frac{23600}{16800}\times 100 = 140.48\%$

III . الرقم القياسي للكميات مرجحاً بأسعار سنة المقارنة (رقم باش للأسعار):

يتم حساب الرقم القياسي للكميات مرجحاً بأسعار سنة المقارنة من المعادلة الآتية:

رقم باش للكميات = $q_I(P) = \frac{\Sigma p_n q_n}{\Sigma p_n q_0}\times 100$

مثال :

2005		2000		
كميات المبيعات	أسعار	كميات المبيعات	أسعار	السلعة
4	25	3	20	A
6	30	5	25	B
3	25	2	15	C
5	40	4	30	D
6	25	4	20	F
5	40	3	35	J

المطلوب :

a) الرقم القياسي التجميعي البسيط للكميات.

b) رقم لاسبير للكميات

c) رقم باش للكميات

				2005		2004		السلعة
p_nq_n	p_nq_o	p_oq_n	p_oq_0	كميات q_n	أسعار P_n	كميات q_o	أسعار Po	
100	75	80	60	4	25	3	20	A
180	150	150	125	6	30	5	25	B
75	50	45	30	3	25	2	15	C
200	160	150	120	5	40	4	30	D
150	100	120	80	6	25	4	20	F
200	120	175	105	5	40	3	35	J
905	655	720	520	29	185	21	145	المجموع

. الرقم القياسي التجميعي البسيط للكميات = $\frac{29}{21}\times 100 = 138.1\%$

. رقم لاسبير للكميات = $\frac{720}{520} \times 100 = 138.46\%$

. رقم باش للكميات = $\frac{905}{655} \times 100 = 138.17\%$

V . الرقم القياسي الأمثل للكميات (رقم فيشر للكميات):

يعبر عن الرقم القياسي الأمثل للكميات بالمتوسط الهندسي لكل من رقم لاسبير للكميات ورقم باش للكميات ويكون وفق الصيغة التالية :

رقم فيشر للكميات = $\sqrt{\text{رقم لاسبير للكميات} \times \text{رقم باش للكميات}}$

مثال :

إذا كان رقم لاسبير للكميات 138.46% ورقم باش للكميات 138.17%

المطلوب : أوجد رقم فيشر للكميات :

الحل :

رقم فيشر للكميات = رقم لاسبير للكميات × رقم باش للكميات

$$\sqrt{138.17\% \times 138.46\%} = 138.11\%$$

6-7- مناسيب الأسعار :

يمكن حساب الرقم القياسي على أساس المتوسط الهندسي لمناسيب أسعار السلع الداخلة في تكوينه. والمتوسط إما أن يكون حسابياً أو هندسياً أو توافقياً، وهذه المتوسطات إما أن تكون بسيطة أو مرجحة ولحساب الرقم القياسي تبعاً لهذه الطريقة نبدأ بحساب منسوب السعر لكل سلعة على انفراد. لنفرض أن السلع (A، B، .... فإن منسوب السعر يكون :

منسوب سعر السلعة (A) 100 × $\frac{\text{سعرها سنة المقارنة}}{\text{سعرها سنة الأساس}}$

$$P_{R_A} = \frac{P_n}{P_0}$$

منسوب سعر السلعة (B) 100 × $\frac{\text{سعرها سنة المقارنة}}{\text{سعرها سنة الأساس}}$

$$P_{R_B} = \frac{P_n}{P_0}$$

وهكذا ...

وبذلك، يكون الرقم القياسي للأسعار لهذه السلع باستخدام المتوسط الحسابي على النحو التالي:

الرقم القياسي للأسعار = $P_{I(\bar{x})} = \frac{\Sigma\left(\frac{p_n}{p_0}\right)}{n} \times 100$

حيث n عدد السلع .

والرقم القياسي للأسعار لهذه السلع باستخدام المتوسط الهندسي البسيط يحسب من المعادلة الآتية :

الرقم القياسي للأسعار :

$$P_{I(G)} = \sqrt[n]{\frac{p_n}{p_0} \times ....\frac{pn`}{po`}} \times 100$$

7-7- مناسيب الكميات :

يمكن حساب الرقم القياسي على أساس المتوسط الهندسي لمناسيب كميات السلع الداخلة في تكوينه. والمتوسط إما أن يكون حسابياً أو هندسياً أو توافقياً، وهذه المتوسطات إما أن تكون بسيطة أو مرجحة. ولحساب الرقم القياسي تبعاً لهذه الطريقة نبدأ بحساب منسوب الكمية لكل سلعة على انفراد وبنفس أسلوب مناسيب الأسعار.

7-8- اختبارات الأرقام القياسية :

تخضع الأرقام ا لقياسية لثلاث اختبارات لاختبار، مدى صحتها وهي:

I . اختبار الانعكاس في الزمن :

يأخذ اختبار الانعكاس في الزمن سنة الأساس كسنة مقارنة وسنة المقارنة كسنة أساس وذلك على النحو التالي :

البديل الزمني لرقم لاسبير $\frac{\Sigma p_0 q_n}{\Sigma p_n q_n} \times 100$

البديل الزمني لرقم باش = $\frac{\Sigma p_0 q_0}{\Sigma p_n q_n} \times 100$

والقاعدة الأساسية لاختبار الانعكاس في الزمن هي :

الرقم القياسي × البديل الزمني له = 1

ويلاحظ أن الرقم القياسي الأمثل ورقم مارشال والرقم القياسي التجميعي البسيط يجتازون هذا الاختبار، بينما لايجتازه كل من رقمي لاسبير وباش.

II . اختبار الانعكاس في المعامل :

في اختبار الانعكاس في المعامل يتم استبدال الأسعار بالكميات أو العكس الكميات بالأسعار مع بقاء الزمن على حالة، والناتج هو البديل المعاملي للرقم القياسي.

البديل المعاملي لرقم لاسبير $= \frac{\Sigma q_n p_0}{\Sigma q_0 p_0} \times 100$

البديل المعاملي لرقم باش $= \frac{\Sigma q_n p_n}{\Sigma q_0 p_n} \times 100$

أي أن البديل المعاملي للرقم القياسي للأسعار المرجح بالكميات هو نفسه الرقم القياسي للكميات المرجح بالأسعار والعكس بالعكس. ولكي يجتاز الرقم القياسي اختبار البديل المعاملي يجب أن يكون:

الرقم القياسي × بديله المعاملي = الرقم القياسي للقيمة

ويلاحظ أن رقمي باش ولاسبير لايجتازان هذا لاختبار بينما يجتازه الرقم القياسي الأمثل.

III . الاختبار الدائري :

إذا تم احتساب الرقم القياسي لسلسلة زمنية متحرك أي أن كل فترة زمنية تعتبر فترة مقارنة بالنسبة للفترة السابقة لها مباشرة، ثم قمنا بضرب هذه السلسلة من الأرقام في بعضها فإننا نحصل على الرقم القياسي للفترة الأخيرة بأساس الفترة الأولى. أي أنه إذا حسب الرقم القياسي للأسعار في 2001 بأساس 2000 و 2002 بأساس 2001 ولعام 2003 بأساس 2002 ثم ضربنا جميع هذه الأرقام في بعضها فإننا نحصل على الرقم القياسي لأسعار 2003 بأساس أسعار 2000 ، إذا تحقق هذا فإننا نقول بأن الرقم القياسي يجتاز الاختبار الدائري، ويلاحظ أن كلاً من أرقام لاسبير وباش وفيشر لاتجتاز هذا الاختبار.

7-9- فوائد الأرقام القياسية :

I . تستعمل الأرقام القياسية في شتى نواحي الحياة لقياس التغير الذي يطرأ عليها كالأرقام القياسية المتعلقة بالعلوم الاجتماعية والإدارية والزراعية والتحليل المالي.

II . تساعد على تحليل العوامل التي تساهم في تغير الظاهرة فتبين مدى مساهمة كل من هذه العوامل في إحداث التغير الكلي.

III . تستخدم الأرقام القياسية في الرقابة على تنفيذ الخطط حيث يستفاد منها في تحديد مدى تنفيذ الخطط الموضوعة.

تمارين غير محلولة

1) يمثل الجدول التالي أسعار وكميات ثلاث سلع

	2000		2005	
السلعة	أسعار	كميات المبيعات	أسعار	كميات المبيعات
A	50	10	60	15
B	75	15	100	20
C	100	20	150	30

والمطلوب :

أوجد مايلي :

. الرقم القياسي التجميعي البسيط للأسعار.

. الرقم القياسي التجميعي البسيط للكميات.

. رقم فيشر للأسعار

. رقم فيشر للكميات .

2) تمثل البيانات التالية كميات وأسعار خمس سلع .

	2000		2005	
السلعة	أسعار	كميات المبيعات	أسعار	كميات المبيعات
A	60	70	20	40
B	50	65	30	50
C	40	60	40	60
D	70	90	60	90
F	100	120	90	110

والمطلوب: أوجد مايلي :

a) الرقم القياسي التجميعي البسيط للأسعار

b) الرقم القياسي التجميعي البسيط للكميات

c) رقم فيشر للأسعار

d) رقم فيشر للكميات

F) مناسيب أسعار السلع F, D, C, B, A

J) متوسط مناسيب الأسعار باستخدام المتوسط الحسابي

H) متوسط مناسيب الأسعار باستخدام المتوسط الهندسي.

K) متوسط مناسيب الكميات باستخدام المتوسط الحسابي.

L) متوسط مناسيب الكميات باستخدام المتوسط الهندسي.

الملاحق

الرموز والأحرف الأبجدية

الرموز والأحرف الأبجدية المستخدمة في هذا الكتاب:

الأحرف الأبجدية: أولا :

الأبجدية العربية	الأحرف الأبجدية اللاتينية			الأحرف الأبجدية اليونانية		
	حرف صغير	حرف كبير	اللفظ	حرف صغير	حرف كبير	اللفظ
أ	a	A	إ	α	A	ألفا
ب	b	B	بي	β	B	بيتا
ج	c	C	سي	γ	Γ	قاما
د	d	D	دي	δ	Δ	ديلتا
هـ	e	E	إي	ε		ابسيلون
و	f	F	إف	ζ	Z	زيتا
ز	g	G	جي	η	H	إيتا
ح	h	H	اتش	θ	θ	تيتا
ط	i	I	آي	i	÷	يوتا
ي	j	J	جي	k	K	كبا
ك	k	K	كي	λ	Λ	لمدا
ل	l	L	ال	μ	M	ميو
م	m	M	ام	v	N	نيو
ن	n	N	ان	ξ	Ξ	كسي
س	o	O	أو	o	O	أوميكرون
ع	p	P	بي	π	Π	بي
ف	q	Q	كيو	ρ	P	رو
ص	r	R	آر	σ	∑	سيكما
ق	s	S	اس	τ	T	تاو
ر	t	T	تي	ν	Y	يويسيلون
ش	u	U	يو	φ	ϑ	فاي
ت	v	V	في	χ	X	كاي
ث	w	W	دبل يو	ψ	Ψ	بسي
خ	x	X	إكس	ω	Ω	أوميغا
ذ	y	Y	واي	ϖ	ς	اس
ض،ظ،غ	z	Z	زد			

ثانياً : الأرقام المستخدمة:

الأرقام العربية	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
الأرقام الهندية	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
الأرقام الرومانية		**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**

ثالثاً : الرموز والإشارات المستخدمة:

العملية	الرمز	العملية	الرمز	العملية	الرمز
ضرب	* أو ×	**أكبر أو يساوي**	≤	**اللوغاريتم العشري**	**Log(x)**
تقسيم	/ أو ÷	**أكبر من**	<	**اللوغاريتم الطبيعي**	**Ln(x)**
الجمع	+	**أصغر أو يساوي**	≥	**التوقع الرياضي**	**E(x)**
الطرح	**-**	**أصغر من**	>	**التباين**	σ^2
يؤدي إلى	⇒	**يساوي**	=	**الانحراف المعياري**	**σ**
واو العطف	**&**	**للتقريب**	#	**الانحراف المعياري للعينة**	**S**
المضاريب	**II**	**يتطابق**	≅	**العاملي**	**!**
المجاميع	∑	**يساوي تقريباً**	≈	**المحدد**	**Δ**
مجموع المربعات	$\sum X^2$	**لانهاية**	∞	**النسبة المئوية**	**%**
الاجتماع	**U**	**التقاطع**	∩	**الوسط الحسابي للمجتمع**	**µ**
التوافيق	C^k_n	**الجذر**	$\sqrt{X}$	**الوسط الحسابي للعينة**	**X**
حجم العينة	**N**	**حجم العينة**	**n**	**الاحتمال التراكمي**	**F(Xs)**

قائمة المراجع بالعربية

- أبو صالح، محمد صبحي؛ عوض، عدنان محمد ، 1990 . مقدمة في الإحصاء . مركز الكتب الأردني.
- أبو عمة، عبد الرحمن، راضي، الحسني عبد البر، الهندي، محمود إبراهيم، 1995، الإحصاء والاحتمالات، الرياض، جامعة الملك سعود.
- الجاعوني، فريد؛ إسماعيل، حبيب؛ غانم، عدنان، 1999، مبادئ الإحصاء، جامعة صنعاء.
- أوسيليفان، جورج، بانكروفت، جوردن، 1983، الرياضيات والإحصاء لدراسات المحاسبة والأعمال، دار ماكجرو هيل للنشر، ترجمة مقدسي، جمال سامي، ومراجعة السيد محمد الغزي.
- حميدان، عدنان؛ الجاعوني، فريد؛ ناصر آغا، عمار؛ العواد، منذر، 2003، مبادئ الإحصاء، منشورات جامعة دمشق.
- طيوب، محمود؛ غانم، عدنان، 2002، مبادئ الإحصاء، جامعة صنعاء.
- مخول، مطانيوس، غانم، عدنان، مبادئ الإحصاء، جامعة دمشق 2006

المراجع باللغة الروسية

1. جرسيموفيتش، أ . ن، الإحصاء الرياضي، الطبعة الثانية، موسكو 1983.
2. ايفتشينكو، ج،1 ، ميدفيديف، يو، ان ، نظرية الاحتمالات والاحصاء الرياضي، موسكو، 1984 .
3. بوروفكوف، أ ، أ ، الاحصاء الرياضي، تقدير المعالم واختيار الفرضيات، موسكو، 1988 .
4. كينينكا، ب، ف، نظرية الاحتمالات، موسكو ، 1994 .

المراجع الأجنبية

1- Kolmogrove A.N, Theory Of Probability and Mathematical Statistics. Moscow “ Seince’, 1967.

2- Mitropolskyi. A.K., Statistical Calculations techiscs. Moscow, “Seince”, 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 gersey, 1988.

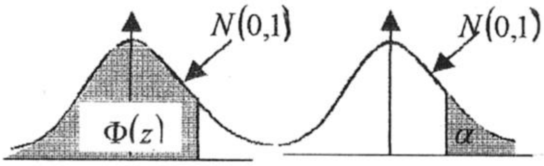
6- Collyer c. and Enns J., analysis of Variance: The Basic Designs. Chicogo, 1976.

7- Kirk R. Experimental Dezign: Procedures for the Behavioral Sciences. California, 1982.

جدول الأرقام العشوائية في الفترة (0,1)

9588	0198	3820	0894	6705	2742	4569
4900	7466	7496	2909	4978	3722	3079
7521	2725	3902	4063	0949	8802	7023
5843	8325	0528	9800	2068	9029	9665
3949	5233	7119	8639	8131	2475	7942
8715	3299	6692	9803	4528	3506	4945
8401	7484	8384	0224	1242	3068	3077
9207	0149	8058	6870	8529	7588	6798
7736	7574	8588	1319	1872	8386	2689
6520	7841	7373	6333	0333	7685	6023
4165	7946	3533	1031	3257	7900	7313
4838	1276	1246	7801	6225	0457	4350
7947	9142	2410	4292	9920	7496	3669
0253	3400	6426	1920	4342	1146	5190
8348	7762	5936	2109	1036	0439	5598
6720	7203	2655	2512	6737	4896	0476
0628	4188	5390	4364	9341	5278	5372
3661	2808	6778	9171	3486	0070	8395
9784	8036	3352	4960	3833	1634	2902
6119	7492	6927	0885	5729	5241	3936
7482	6518	9499	2966	7333	3457	5776
5938	9499	8154	6110	2970	2117	8560
4725	7083	3201	7851	0122	0435	5401
4298	3395	5982	9762	3821	6014	4061
8745	3018	6243	7117	6795	7009	4831
2084	5778	1475	9227	6156	7372	7152
2578	7162	3083	1147	2882	2377	9183
0045	7556	2820	1923	6061	6569	1805
8138	7994	0808	5852	3319	8838	7975
4187	5360	9691	0736	2104	5497	3998
9239	6149	2790	8068	0064	5110	2736
6640	0233	9391	4707	6740	8089	4086
5314	1778	0038	4353	3054	3854	7375
1982	4704	4771	5948	5567	8559	6118
6178	7492	1538	8617	8481	2830	0787
2817	6883	8082	0589	6475	0674	8508
3270	8611	9218	7240	6695	9635	9953
2826	4639	7201	6681	2385	4960	2165
1734	1450	5790	0361	7877	7867	2580
4870	1693	8312	6607	5553	3707	1932
8207	3300	7158	1835	8846	8869	1195
8248	1060	7053	8547	9854	4010	3488
7596	8039	2547	3006	3395	4834	9408
4959	4082	2318	5127	578	4345	2857
3329	7550	9481	2988	4642	3842	4167

Standard Normal Distribution التوزيع الطبيعي المعياري



$$P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw, \quad [\Phi(-z) = 1 - \Phi(z)]$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319
0.1	0.5398	0.5438	0.5478	0.5478	0.5557	0.5596	0.5636	0.5675	0.5714
0.2	0.5793	0.5832	0.5871	0.5871	0.5948	0.5987	0.6026	0.6064	0.6103
0.3	0.6179	0.6217	0.6255	0.6255	0.6331	0.6368	0.6406	0.6443	0.6480
0.4	0.6554	0.6591	0.6628	0.6628	0.6700	0.6736	0.6772	0.6808	0.6844
0.5	0.6915	0.6950	0.6985	0.6985	0.7054	0.7088	0.7123	0.7157	0.7190
0.6	0.7257	0.7291	0.7324	0.7324	0.7389	0.7422	0.7454	0.7486	0.7517
0.7	0.7580	0.7611	0.7642	0.7642	0.7704	0.7734	0.7764	0.7794	0.7823
0.8	0.7881	0.7910	0.7939	0.7939	0.7995	0.8023	0.8051	0.8078	0.8106
0.9	0.8159	0.8186	0.8212	0.8212	0.8264	0.8289	0.8315	0.8340	0.8365
1.0	0.8413	0.8438	0.8461	0.8461	0.8508	0.8531	0.8554	0.8577	0.8599
1.1	0.8643	0.8665	0.8686	0.8686	0.8729	0.8749	0.8770	0.8790	0.8810
1.2	0.8849	0.8869	0.8888	0.8888	0.8925	0.8944	0.8962	0.8980	0.8997
1.3	0.9032	0.9049	0.9066	0.9066	0.9099	0.9115	0.9131	0.9147	0.9162
1.4	0.9192	0.9207	0.9222	0.9222	0.9251	0.9265	0.9279	0.9292	0.9306
1.5	0.9332	0.9345	0.9357	0.9357	0.9382	0.9394	0.9406	0.9418	0.9429
1.6	0.9452	0.9463	0.9474	0.9474	0.9495	0.9505	0.9515	0.9525	0.9535
1.7	0.9554	0.9564	0.9573	0.9573	0.9591	0.9599	0.9608	0.9616	0.9625
1.8	0.9641	0.9649	0.9656	0.9656	0.9671	0.9678	0.9686	0.9693	0.9699
1.9	0.9713	0.9719	0.9726	0.9726	0.9738	0.9744	0.9750	0.9756	0.9761
2.0	0.9772	0.9778	0.9783	0.9783	0.9793	0.9798	0.9803	0.9808	0.9812
2.1	0.9821	0.9826	0.9830	0.9830	0.9838	0.9842	0.9846	0.9850	0.9854
2.2	0.9861	0.9864	0.9868	0.9868	0.9875	0.9878	0.9881	0.9884	0.9887
2.3	0.9893	0.9896	0.9898	0.9898	0.9904	0.9906	0.9909	0.9911	0.9913
2.4	0.9918	0.9920	0.9922	0.9922	0.9927	0.9929	0.9931	0.9932	0.9934
2.5	0.9938	0.9940	0.9941	0.9941	0.9945	0.9946	0.9948	0.9949	0.9951
2.6	0.9953	0.9955	0.9956	0.9956	0.9959	0.9960	0.9961	0.9962	0.9963
2.7	0.9965	0.9966	0.9967	0.9967	0.9969	0.9970	0.9971	0.9972	0.9973
2.8	0.9974	0.9975	0.9976	0.9976	0.9977	0.9978	0.9979	0.9979	0.9980
2.9	0.9981	0.9982	0.9982	0.9982	0.9984	0.9984	0.9985	0.9985	0.9986
3.0	0.9987	0.9987	0.9987	0.9987	0.9988	0.9989	0.9989	0.9989	0.9990
	بعض القيم الشائعة								
α	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.240	2.576	2.807	3.291

المصطلحات الاحصائية

Statistical Terminology

إنكليزي - عربي

A

Absolute	مطلق
Absolute value	القيمة المطلقة
Accuracy	دقة
Adjustment data	تعديل البيانات
Aggregation	تجميع
Allowance Value	حدود السماح
Aggregate	كلي، شامل
Analysis of variance	تحليل التباين
Applied Statistics	الإحصاء التطبيقي
Arbitrary constant	ثابت اختياري
Area	مساحة
Area under the normal curve	مساحة تحت المنحني الطبيعي
Array	ترتيب
Association	اقتران
Alternative hypothesis	الفرض البديل
Asymmetry	عدم التماثل
Asymmetrical	غير متماثل
Asymptotic	تقريبي
Attribute	صفة
Average	متوسط
Axis	محور
Axis of abscissas	المحور السيني أو الأفقي
Axis of ordinates	المحور العمودي

B

Base	أساس أو قاعدة
Bayes theorem	نظرية بايز
Basic statistics	إحصاءات أساسية
Battery of test	مجموعة الاختبارات
Best fit	التوفيق الأفضل

Between Groups	بين المجموعات
Bias in surveys	التحيّز في المسوح
Binominal Distribution	توزيع ثنائي الحدين
Bivariate	ذو متغيرين
Breakdown	تقسيم
Business Cycles	دورات اقتصادية
Business indicators	مؤشرات اقتصادية

C

Calculation	حساب
Calculus of probability	حساب الاحتمالات
Census	تعداد
Census enumerator	عداد
Census schedule	استمارة التعداد
Causation	سببية
Central tendency	النزعة المركزية
Central limit theorem	نظرية الحد المركزي
Change	تغير
Chi-Squared test	اختبار كاي مربع
Classification	تصنيف
Class frequency	تكرار الفئة
Code	رمز
Coefficient	معامل
Coefficient of Correlation	معامل الارتباط
Coefficient of Detemination	معامل التحديد
Column	عمود
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
Classical probability	احتمال كلاسيكي
Conditional probability	احتمال شرطي
Continuous distribution	توزيع متصل (مستمر)
Critical region	منطقة حرجة
Curves	منحنيات

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 بياني
Differential	تفاضلي
Discrete distribution	توزيع متقطع
Dispersion	تشتت
Distributions of probability	توزيعات احتمالية
Distribution function	دالة التوزيع
Double sampling	معاينة مزدوجة
Downward bias	تحيز نحو الأدنى
Downward trend	اتجاه هابط

E

Effect	تأثير
Efficiency of estimates	فعالية التقدير
Elimination	حذف
Enumerator	عداد
Equation	معادلة
Error	خطأ
Error of Estimate	خطأ التقدير
Error of Variance	تباين الأخطاء

Estimate / Estimation	تقدير
Estimation equation	معادلة التقدير
Event	واقعة / حادثة
Expected value	القيمة المتوقعة
Experimental error	الخطأ التجريبي
Experimental units	وحدات تجريبية
Exponential Curve	منحني آسي
Exponential equation	معادلة آسية

F

“ F “ Table	جدول " ف "
Factor	عامل
Factor analysis	تحليل عوامل
Factorial analysis	تحليل 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

G

Gaussian curve	منحنى غوس
Geometric mean	وسط هندسي
Graduation	تدريج
Graph	رسم بياني
Graphic analysis	تحليل بياني
Graphic presentation	عرض بياني
Goodness of fit test	اختبار جودة التوفيق
Grouping	تجميع / تصنيف

Grouping error	خطأ التجميع
Grouped data	بيانات مبوبة

H

Harmonic equation	دالة توافقية
Harmonic mean	وسط توافقي
Heteroscedasticity	اختلاف التباين
Histogram	المدرج التكراري
Homogeneous	متجانس
Hypothesis testing	اختبار الفروض

I

Ideal number	الرقم الأمثل
Identity	متطابقة
Independent variable	متغير مستقل
Index	دليل أو مقياس
Index number	الرقم القياسي
Infinite	لانهائي
Inferential statistics	احصاء استدلالي
Interclass correlation	ارتباط بين الفترات
Inersection of sets	تقاطع المجموعات
Interval estimation	فترة، مجال التقدير
Interview	مقابلة / استجواب
Inverse correlation	ارتباط عكسي
Inverse function	دالة عكسية

J

Joint distribution	توزيع مشترك
Joint regression	انحدار مشترك

K

Kurtosis	تفرطح

L

Large sample	عينات كبيرة
Law of large numbers	قانون الأعداد الكبيرة
Least squares	المربعات الصغرى
Least squares method	طريقة المربعات الصغرى
Level of significane	مستوى المعنوية
Likelihood	إمكان
Limit	نهاية أو حد
Linear	خطي / مستقيم
Linear correlation	ارتباط خطي
Linear equation	معادلة خطية
Linear interpolation	استكمال خطي
Linear regression	انحدار خطي
Linear trend	اتجاه مستقيم

M

Master Sample	عينة رئيسة
Matrix	مصفوفة
Miximum variation	أقصى اختلاف
Mean	وسط
Method of least squanes	طريقة المربعات الصغرى
Median	الوسيط
Model	نموذج
Moment	عزم
Muliple correlation	ارتباط متعدد
Multiple factor analysis	تحليل عاملي متعدد
Multipleication of probabilities	ضرب الاحتمالات
Mutually exclusive events	احداث متنافية

N

Negative correlation	ارتباط سالب
Non-Linear correlation	ارتباط غير خطي
Non-Statistics	إحصاء لا معلمي
Normal curve	منحنى طبيعي
Normal distribution	توزيع طبيعي
Normal equation	معادلة طبيعية
Null hupothesis	فرضية العدم

O

Observation	مشاهدة
One – tailed test	اختبار من طرف واحد

P

Parabola	قطع مكافئ
Parameter	معلمة، ثابت
Partial correlation	ارتباط جزئي
Partial regression	انحدار جزئي
Percentage	نسبة مئوية
Population (statistical)	مجتمع (إحصائي)
Positive correlation	ارتباط موجب
Prediction	تنبؤ
Point estimation	تقدير بنقطة
Poisson distribution	توزيع بواسون
Power of a test	قوة الاختبار
Probability	احتمال
Prior probability	الاحتمال المسبق
Probability density	كثافة الاحتمال
Probability sampling	المعاينة الاحتمالية

Q

Qualitative analysis	تحليل وصفي
Quality control	مراقبة الجودة
Quantitative analysis	تحليل كمي
Quanitity 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 coefficent	معامل ارتباط الرتب
Rate	معدل
Ratio	نسبة
Ratio estimate	تقدير نسبي
Ratio test	اختبار النسب
Regression	انحدار
Regression coefficient	معامل الانحدار
Regression equation	معادلة الانحدار
Regression estimate	تقدير الانحدار
Regression line	خط الانحدار
Rejection region	منطقة الرفض

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
Sample size	حجم العينة
Sample space	فضاء العينة
Scatter diagram	شكل الانتشار
Second degree curve	منحنى من الدرجة الثانية
Set theory	نظرية المجموعات
Sensitivity	حساسية
Sequential analysis	تحليل تتابعي
Significance level	مستوى المعنوية
Simple correlation	ارتباط بسيط
Small samples	عينات صغيرة
Sources of error	مصدر الخطأ
Spurious 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
Survey	مسح (استقصاء)
Student's – distribution	توزيع (t) ستودنت

T

' T ' Table	جدول " ت " ستودنت
Table	جدول
Tabulation	جدول أو تبويب
Test	اختبار
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
Two – tailed test	اختبار من طرفين
Type I error	خطأ من النوع الأول
Type II error	خطأ من النوع الثاني

U

Union of sets	اتحاد المجموعات
Unbiased estimate	تقدير غير متحيز
Unit of measurement	وحدة القياس
Universe (statistical)	مجتمع (إحصائي)
Upward tend	اتجاه صاعد

V

Variable	متغير
Variance	تباين
Variance analysis	تحليل التباين
Variance ratio	نسبة التباين
Variation	اختلاف

Variation coefficent	معامل اختلاف
Venn diagram	شكل فن

W

Weight average	تثقيلات، أوزان
Weight average	متوسط مثقل أو مرجح

X

X – axis	محور السينات

Y

Y-axis	محور العينات
Yates continuity correction	تصحيح ياتس

Z

Z – test	اختبار (ز)

مواقع الانترنت

1- http://www.stutsft.com/textbook/stathome.html
2- http://www.bmj.com/collections/statsbk/11.shtml.
3- http://davidmlance.com/hyperstat/intro.html.
4- http://www.psychstat.missouristate.edu/introbook/sbko?.htm
5- http://www.penhall.com/groebner