

بسم الله الرحمن الرحيم

جامعة السودان المفتوحة

برنامج التربية

الإحصاء الإداري

رمز المقرر ورقمه: إدر

المؤلف: أ.د. أحمد حمد النوري

التحكيم العلمي: د\ زين العابدين عبدالرحيم

التصميم التعليمي :

منشورات جامعة السودان المفتوحة، الطبعة الأولى ٢٠٠٦

جميع الحقوق محفوظة لجامعة السودان المفتوحة، لا يجوز إعادة إنتاج أي
جزء من هذا المقرر، وبأي وجه من الوجوه، إلا بعد الموافقة المكتوبة من
الجامعة.

مقدمة المقرر

عزيزي الدارس

مرحباً بك إلي هذا المقرر "الأحصاء الإداري"

والذي يحتوي علي سبعة وحدات دراسية تشتمل علي دراسة تنظيم البيانات بصورة تمكن من الوصول إلي المعلومات حيث تدرس مقاييس النزعة المركزية ، التشتت ، والالتواء وعوامل الارتباط والانحدار الخطي والأرقام القياسية وكيفية تحليل السلاسل الزمنية .

ستجد هذا المقرر شيق لو كنت من المهتمين بالرياضيات والتفاضل والتكامل حاول الرجوع إليها حينها ستجد في ذلك خير معين لعملية الفهم .

كما نرجو منك أن تكون جاداً في التعامل مع الأنشطة والتدريبات وأسئلة التقويم الذاتي التي ستقابلها من خلال هذه الوحدات التي يحتويها المقرر .

والله خير معين والذي علم بالقلم علم الإنسان ما لم يعلم .

الأهداف العامة للمقرر



ينتظر منك عزيزي الدارس،

بعد دراسة هذا المقرر "الإحصاء الإداري" أن تصبح قادراً على

أن:

١. تفهم مصادر البيانات وأساليب جمعها وكيفية عرضها.
٢. تشرح معاني مقاييس النزعة المركزية، الوسط الحسابي، الوسيط، المنوال.
٣. تتعلم كيفية استخدام مقاييس التشتت، الانحرافات، المدى والوتشتت النسبي أو معامل الاختلاف.
٤. تتعرف على درجة الالتواء وكيفية استخدام الغروم.
٥. تحل مسائل الارتباط ومعامل الاختلاف والتوافق.
٦. توضح أهمية الأرقام القياسية لعمل الإحصاء الإداري.
٧. تعد المنحى التاريخي وتحليل الزلازل الزمنية.

محتويات المقرر

الوحدة	عنوان الوحدة	الصفحة
الأولى	جمع وعرض البيانات	١
الثانية	مقاييس النزعة المركزية	٥٤
الثالثة	مقاييس التشتت	٨٨
الرابعة	الالتواء والتقرطح والعزوم	١١٠
الخامسة	الارتباط والانحدار الخطي	١٢٦
السادسة	الأرقام القياسية	١٦٤
السابعة	تحليل السلاسل الزمنية	١٨٧



محتويات الوحدة

الصفحة	الموضوع
٣	مقدمة
٣	تمهيد
٣	أهداف الوحدة
٤	١. مصادر البيانات
٤	١.١ المصادر التاريخية
٥	٢.١ المصادر الميدانية
٧	٢. أساليب جمع البيانات
٧	١.٢ الحصر الشامل
٧	٢.٢ المعاينة
٢٢	٣. عرض البيانات
٢٢	١.٣ العرض الجدولي للبيانات
٣٥	٢.٣ العرض البياني للبيانات
٥٠	الخلاصة
٥١	لمحة مسبقة عن الوحدة التالية
٥٢	مسرد المصطلحات
٥٣	إجابات التدريبات

مقدمة

تمهيد

عزيزي الدارس،

مرحباً بك إلى هذه الوحدة

تحتوي هذه الوحدة على ثلاثة أقسام رئيسية، تبدأ بمصادر جمع البيانات ثم تنتقل إلى أساليب جمع البيانات ومن هذه الأساليب أسلوب الحصر الشامل وأسلوب المعاينة، تنتهي هذه الوحدة بالقسم الثالث وهو عرض البيانات، حيث نقدم لك في هذا القسم العرض الجدولي للبيانات وكذلك طرق عرض البيانات.

تتخلل هذه الوحدة العديد من التدريبات والأنشطة وأسئلة التقويم الذاتي. أرجو أن تكون خير معين لك في هذه الوحدة وأن تستفيد من هذه الوحدة في فهم بقية وحدات هذا المقرر، مرحباً بك مرة أخرى في هذه الوحدة التي نأمل أن تساعدنا في تطويرها بالنقد الهادف. مع تمنياتنا لك بالتوفيق والنجاح.

أهداف الوحدة

عزيزي الدارس،

بعد فراغك من دراسة هذه الوحدة ينبغي أن تكون قادراً على أن:

- ١- تبين المصادر الأساسية لجمع البيانات الإحصائية.
- ٢- تعدد أساليب جمع البيانات الإحصائية.
- ٣- تقارن بين أسلوب الحصر الشامل وأسلوب المعاينة.
- ٤- تشرح العرض الجدولي للبيانات.
- ٥- تناقش طرق العرض للبيانات.



مقدمة

عزيزي الدارس ماذا تعرف عن متطلبات تصميم البحث العلمي؟

تصميم البحث العلمي يتطلب خطة تحتوي على العناصر الأساسية المتعلقة بمشكلة البحث، أول هذه العناصر هي

تصميم مشكلة البحث وصياغتها بصورة واضحة مع تحديد المنهج الذي يجب اتباعه ، وينقسم منهج البحث العلمي عادة إلى قسمين :

١. المنهج النموذجي وهو يحدد الإطار النظري للبحث في استنباط الأهداف واختبار الفروضات.
٢. المنهج التطبيقي فهو عبارة عن تطبيق المنهج النموذجي باستخدام المفاهيم الإحصائية لجمع البيانات وتحليلها .

وهذا يتطلب تحديد نوع البيانات ومصادرها وما إذا كانت بيانات شاملة أو بيانات معاينة وكيفية اختيار العينات . ويتطلب أيضا وضع خطة للعرض الجدولي والبياني وطرق التحليل واستخلاص النتائج وتعتبر عملية جمع البيانات من أهم العمليات الإحصائية وأكثرها حساسية وتحتاج إلى دقة كبيرة ، وتتوقف عليها جميع النتائج الخاصة بالبحث ، فكلما كانت هذه العملية دقيقة كلما كانت النتائج التي يتم التوصل إليها دقيقة أيضا والعكس صحيح .

١. مصادر البيانات Data Sources

البيانات الإحصائية لها مصدران أساسيان :

١.١ المصادر التاريخية (أو الثانوية)

وهي مصادر غير مباشرة وتشمل الإحصاءات المنشورة والكتب والمطبوعات والوثائق التي تصدرها الحكومات والهيئات الرسمية وشبه الرسمية ، وكذلك المنظمات الدولية كالأمم المتحدة ومنظمة العمل الدولية وغيرها ، وتقوم هذه الجهات بحكم اختصاصها بجمع البيانات وجدولتها وإصدارها في نشرات دورية من الأمثلة على ذلك الكتاب الديمغرافي السنوي الذي تصدره الأمم المتحدة عن التعدادات العامة للسكان في مختلف دول العالم والتقارير التي تصدر عن الجهاز المركزي للإحصاء حول السكان في السودان .

والمصادر التاريخية (الثانوية) يعاب عليها أنها ربما تحتوي على أخطاء مقصودة أو غير مقصودة كالأخطاء الناتجة عن عدم فهم صاحب المصدر الثانوي لما أخذه من المصدر الأصلي أو أخطاء بسبب تحيز صاحب البحث الثانوي مثل ذكر جزء وترك جزء آخر أو إغفال التعاريف أو الشروط الواردة في المصدر الأصلي ، أو أخطاء بسبب تحيز صاحب البحث الثانوي مثل ذكر جزء وترك جزء آخر أو و إغفال التعاريف أو الشروط الواردة في المصدر الأصلي ، ولذلك فإن المصدر الأصلي أفضل كثيرا من المصدر الثانوي.

٢.١ المصادر الميدانية (أو الأولية)

وهي مصادر مباشرة وتشمل الوحدات الأصلية التي تستقضي منها المعلومات بصفة مباشرة ، ويتم جمع البيانات طبقا لهذه المصادر عن طريق:

(١) **المشاهدة** : ففي حالة بحث عدد المدارس والمستشفيات مثلا يمكننا جمع بيانات بمجرد المشاهدة دون مقابلة الأفراد ، ويتم جمع البيانات طبقا لهذه الوسيلة في استمارة بحث عن طريق المشاهدة .

(٢) **المقابلة** : ويتم جمع البيانات طبقا لهذه الوسيلة في استمارة استبيان عن طريق المقابلة حيث يقوم الباحث بمقابلة الأشخاص ويوجه الأسئلة إليهم ثم يدون الإجابة ، ومن مميزات المصادر الميدانية أن الباحث يستطيع التحقق من صحة المعلومة وأن كان يعاب عليها كثرة ما تتطلبه من نفقات.

(٣) **المراسلة** : وطبقا لهذه الوسيلة يعد الباحث استمارة استبيان يرسلها بالبريد إلى الأشخاص المبحوثين على عناوينهم للإجابة على الأسئلة التي تضمنتها الاستمارة ثم إرسالها إلى الباحث بعد ذلك ، وعلى الرغم من أن هذه الطريقة تؤدي إلى انخفاض التكلفة إلا أنها ليست شائعة الاستخدام نسبة عدم التجاوب العالية كما أنها غير مجدية بالنسبة للمجتمعات التي لا تعرف القراءة والكتابة .

(٤) **الهاتف** : وفيها يتم الاتصال بالوحدة عن طريق الهاتف وأخذ المعلومات منها .

وقد أتاحت التقنيات الحديثة طرقاً أخرى لجمع البيانات منها استخدام الإنترنت والفاكس .

أسئلة تقويم ذاتي



- ١- عرف المنهج النموذجي والمنهج التطبيقي؟
- ٢- أذكر مصادر البيانات الإحصائية؟
- ٣- كيف يتم جمع المعلومات طبقاً لوسيلة المقابلة؟

٢. أساليب جمع البيانات

Methods of Data Collection

هنالك طريقتان لجمع البيانات الإحصائية وهما طريقة الحصر الشامل وطريقة المعاينة .

١.٢ الحصر الشامل Full Count

ويعني أخذ المعلومات من جميع مفردات المجتمع المراد دراسته . ومن أهم خصائصه الشمول الكامل بالمستوي الزمني والمكاني لجميع مفردات الإطار . والإطار يقصد به تمثيل بشكل ما لجميع مفردات المجتمع التي تخضع للدراسة ، فمثلاً إذا كان المجتمع المراد دراسته هو طلاب كلية الطب بجامعة الجزيرة فإن الإطار يمكن أن يكون - مثلاً - قائمة بأسماء هؤلاء الطلاب .

أسلوب الحصر الشامل يمكن استخدامه إذا كانت مفردات المجتمع محدودة أو في الحالات التي تتطلب جمع البيانات عن خصائص المجتمع على المستوى القومي ، مثل ذلك التعداد العام للسكان والمنشآت والتي تقوم به مصلحة الإحصاء على فترات زمنية منتظمة .

من مزايا الحصر الشامل الإلمام بجميع البيانات الخاصة بالمجتمع ككل كما يمكن من إجراء جميع الدراسات المطلوبة كما يؤدي إلى دقة النتائج وعدم تحيزها ، ويعاب على أسلوب الحصر الشامل أنه يكون عرضة أخطاء الشمول الناتجة عن تكرار أو نقص الحصر إلى جانب أخطاء المحتوى والتي تصيب بعض خصائص المفردات بالإضافة إلى كثرة

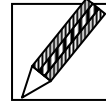
النفقات وطول الوقت المطلوب لإجراء البحث واستخراج النتائج.

٢.٢ المعاينة Sampling

وفي هذه الطريقة يتم جمع البيانات باستخدام عدد محدود من مفردات الإطار ثم نعمم نتائج الدراسة على جميع مفردات المجتمع وتعتبر المعاينة الإحصائية من الوسائل الأكثر استخداماً لأنه وباتباع الطرق والنظريات الإحصائية العلمية يمكن معها تعميم النتائج فيما بعد على المجتمع. ومن مزايا طريقة المعاينة السهولة والبساطة واختصار الوقت والتكاليف كما يؤدي التحكم في عدد محدود من مفردات الإطار إلى نتائج ربما تكون أكثر دقة من الحصر الشامل ، إلا أن من عيوب طريقة المعاينة هي أنه إذا كانت العينة غير ممثلة للمجتمع فإن هذا يؤدي إلى أخطاء التحيز بحيث تصبح نتائجها غير مطابقة تمام كما لو تمت الدراسة باستخدام التعداد الكامل كما سنرى فيما بعد. وربما يتوجب على الباحث أن يستخدم طريقة المعاينة في الحالات التي يؤدي فيها الشمول الكامل إلى فقدان المجتمع بأكمله ، مثل ذلك عندما يفحص الطبيب الدم لمريض فإنه يكون مضطراً لأخذ عينة من الدم لأنه إذا أخذ الدم كله لمات ذلك المريض ، أو إذا أراد أحد الباحثين الزراعيين دراسة نوع معين من أمراض الجذور لنوع من أنواع النباتات فلا بد له أن يأخذ عينة وإلا فقد النبات بأكمله.

تدريب (١)

قارن بين أسلوب الحصر الشامل وأسلوب المعاينة ؟



أولاً المفهوم الإحصائي للمعاينة

كما ذكرنا سابقاً فإن نظرية العينات تقوم على أساس دراسة خصائص مجتمع معين عن طريق دراسة عينة ، تمثل ذلك المجتمع.

وفكرة المعاينة هي فكرة مأخوذة من السلوك البشري العام حيث أن فكرة دراسة المجتمع عن طريق دراسة جزء منه هي فكرة شائعة وتمارس يومياً في مختلف مناحي الحياة ، فالتاجر يقوم بفحص جزء من السلع المراد شراؤها وصاحب المصنع يقوم بضبط الجودة لإنتاجه بأخذ عينة من الإنتاج وفحصها ، لكن مثل هذا السلوك غالباً ما يتم خارج الإطار

الإحصائي بافتراض أن مفردات المجتمع تحت الاختبار متجانسة .

على أن المفهوم الإحصائي يحدد تحديدا قاطعا لكي تحكم على الكل يفحص الجزء لابد من اتباع أسس منظمة وسليمة يتم بها اختبار هذا الجزء ، وهذه الطريقة المنظمة تعرف باسم طريقة المعاينة والجزء المختار يسمى عينة ، وأهم الأسس التي تقوم عليها نظرية المعاينة هي فكرة الصيغة العشوائية والتي تمثل أحد أساليب جمع البيانات والتي تمثل نظرية متكاملة من النظريات الإحصائية التي تحكمها قوانينها الرياضية .

نشاط



بعد أن تعرفت على مصادر جمع البيانات وطرق جمع البيانات أعطي أمثلة لعناوين بعض الدراسات والبحوث ثم حاول توضيح أنسب طريقة لجمع البيانات يمكن أن يستخدمها الباحث

ثانيا تصميم العينة Sample Design

إذا تعين إجراء بحث عن طريق العينة فإنه يجب على الباحث أن يأخذ في الاعتبار المرتكزات الآتية :

(أ) تعريف المشكلة المراد بحثها ، ويعتبر هذا التعريف ضروريا للتعرف على نوع وكمية البيانات المطلوبة .

(ب) تعريف وتحديد المجتمع المراد بحثه ، ويبدأ ذلك بتحديد الإطار الذي تسحب منه مفردات العينة .

(ج) تحديد وحدة المعاينة المطلوبة فإذا كان البحث يتعلق بموضوع اقتصادي مثلا فإن أنسب وحدة للمعاينة هي المنشأة الاقتصادية أو الحقل .

(د) تحديد البيانات المطلوبة من أفراد العينة مع تحديد طريقة جمعها سواء كان ذلك عن طريق الاتصال المباشر أو غير المباشر .

(هـ) تنظيم ومراقبة العمل الميداني وذلك بوضع جدول زمني لمراحل العمل وتوفير العدد اللازم من الباحثين والمشرفين الميدانيين .

(و) ربما احتاج الباحث في بعض الأحيان أن يقوم بمسح تمهيدي للاستفادة منه في تصميم وتحديد حجم العينة والتأكد من سلامة استمارة الاستبيان الموضوعة .

ثالثا الإطار Sample Frame

كما ذكرنا سابقا فإن الإطار يمثل مفردات المجتمع الخاضعة فعلا للمعاينة أي جميع الوحدات الأولية التي يتكون منها المجتمع والتي هي في متناول اليد ولكي يكون الإطار مناسباً يجب أن تتوافر فيه الصفات التالية :

- أن يشتمل بقدر الإمكان على مفردات المجتمع المراد دراسته .

- أن تكون مفرداته مرتبة رقمياً وسهلة التمييز وفي نفس الوقت تعكس خصائص المجتمع .

إن أهم شرط من شروط الإطار هو شموله لجميع مفردات المجتمع حتى تتحقق قاعدة إعطاء فرص متساوية لجميع المفردات للظهور في العينة .

رابعا وحدات المعاينة Sample Units

بداية لابد من معرفة الفرق بين وحدة الدراسة ووحدة المعاينة . وحدة المعاينة هي الوحدة التي تبنى على أساسها المعاينة (ماذا نسحب ؟) ، أما وحدة الدراسة فهي الوحدة التي تؤخذ منها المعلومات . فإذا أراد أحد الباحثين اختيار عدد من الفصول بكلية التربية لبحث ظاهرة عدم الإكمال في الامتحانات فإن وحدة الدراسة هي الطالب أما وحدة المعاينة فهي الفصل لأن الوحدة التي ستسحب هي الفصل بينما التي ستؤخذ منها المعلومات هي الطالب .

و وحدات المعاينة قد تكون وحدات طبيعية بسيطة كأفراد المجتمع أو مجموعة من هذه الوحدات كالأسر؛ كما قد تكون وحدات مصطنعة كتقسيم الطلاب إلى مجموعات واعتبار كل مجموعة وحدة معاينة .

أسئلة تقويم ذاتي



- ١- ما هي طرق جمع البيانات الإحصائية؟
- ٢- ماذا نعني بالحصص الشامل؟
- ٣- ما هي الحالات التي يستخدم فيها أسلوب الحصص الشامل؟
- ٤- أذكر مزايا وعيوب طريقة المعاينة؟
- ٦- متى يتوجب على الباحث استخدام طريقة المعاينة؟
- ٧- ما هي أهم الأسس التي تقوم عليها طريقة المعاينة؟
- ٨- ما هي الصفات التي يجب أن تتوفر في الإطار حتى يكون مناسباً؟
- ٩- ما الفرق بين وحدة الدراسة ووحدة المعاينة؟
- ١٠- عرف ما يلي: المجتمع الإحصائي، العينة، الإطار، وحدة المعاينة؟

خامسا طرق المعاينة Sampling Techniques
هناك طرق كثيرة للمعاينة يمكن تقسيمها إلى مجموعتين متميزتين :

Random المعاينة العشوائية
Purposive المعاينة العمدية
Sampling

والمعاينة العشوائية هي التي يتم أخذها على أساس عشوائي أو احتمالي مكن معه تعميم نتائج العينة على المجتمع ، وكلمة عشوائي هنا تعني إعطاء فرص متكافئة لجميع مفردات المجتمع عند أخذ العينة بحيث يكون لكل مفردة احتمال (موجب) معروف ومحدد سلفا .

والعينة العشوائية الممتازة تحتوي على أنواع المفردات المختلفة للمجتمع بنسب قريبة جدا من النسب الموجودة لهذه الأنواع في المجتمع الأصلي.

أما المعاينة العمدية فهي التي يتم أخذها بطريقة غير عشوائية وتسمى أحيانا بالمعاينة الفرضية فيلجأ إليها إذا لم تكن هنالك أي معلومات متوفرة عن المجتمع المراد دراسته أو إذا أريد اختيار عينة صغيرة الحجم من مجتمع كبير نسبيا ، كذلك نستخدم في الحالات التي يراد فيها تكوين فكرة سريعة عن مشكلة معينة أو لتجربة استثمارية الاستبيان في مسح تمهيدي.

المعاينة العشوائية Random Sampling

المعاينة العشوائية تنقسم إلى عدة أنواع حسب الطريقة التي ينفذ بها مبدأ العشوائية .

(أ) المعاينة العشوائية البسيطة Simple Random Sampling

إذا أردنا سحب عينة حجمها n من مجتمع ما ، فإنها توصف بأنها عشوائية بسيطة إذا كان لكل مجموعة ذات حجم n يمكن سحبها من ذلك المجتمع نفس الفرصة في أن تكون هي المسحوبة . ويتحقق ذلك عندما يكون لكل مفردة من مفردات المجتمع نفس الفرصة للظهور في العينة ، ولتطبيق مبدأ العشوائية نستخدم ما يسمى بالصندوق المثالي أو نلجأ للجدول العشوائية .

مثال (١)

إذا أردنا أن نأخذ عينة عشوائية بسيطة تتكون من ١٠ طلاب في فصل به ٢٠٠ طالب فإننا نبدأ بترقيم جميع الطلاب في الفصل أما إذا استخدمنا جداول الأرقام العشوائية فإننا نرقم أفراد المجتمع من ١ إلى ٢٠٠ كما سبق ثم نستعمل جدول الأرقام العشوائية رقم (٢) الملحق بنهاية هذا الكتاب وتقوم بسحب عينة عشوائية بسيطة ذات حجم ١٠ من هذا المجتمع.

ويمكن عمل ذلك بعدة طرق إحداها أن نقرأ هذا الجدول رأسياً بحيث يكون عدد الخانات أو والأعمدة مساوية لعدد خانات الرقم الذي يمثل مجموع المجتمع محل الدراسة والعدد الذي يكون أصغر من عدد المجتمع نأخذه والذي يكون أكبر من عدد المجتمع نتركه حتى نصل إلى أرقام يكون عددها مساوية لعدد مفردات العينة.

وإذا ما نظرنا إلى الجدول وقرأنا أول ثلاثة أرقام في الصف والعمود الأول نجد ٧٧٢ يليه رأسياً ٠٣٣ وهنا نأخذ الثاني ونترك الأول وهكذا نجد أن

العينة المكونة من ١٠ طلاب هي

٣٣	٩٧	١٣٦	١٨٤	١٧٣
١٣٣	١٥٧	١٥٤	٧٧	٧٨

وهذه الأرقام مأخوذة من العمودين الأول والثاني من الملحق رقم (٢) ومن مزايا المعاينة العشوائية البسيطة إنها سهلة الاختيار إلا أن عيبها الرئيسي أنه لا ينصح باستخدامها إلا إذا كان المجتمع الأصلي متجانساً من حيث التركيب والتكوين.

(ب) المعاينة الطبقية

المعاينة الطبقية يلجأ إليها في حالة معرفة التركيب النسبي للمجتمع الأصلي عندما يكون هذا المجتمع مكوناً من عدة طبقات بينها اختلاف واضح. فقد تعتمد النتيجة الإحصائية في بعض الأحيان على بعض الصفات كالعمر والجنس والمهنة ، وفي هذه الحالة يتم تقسيم المجتمع حسب الصفات ثم بعد ذلك يتم تحديد حجم العينة بتخصيص نصيب كل صفة أو طبقة من حجم العينة بحيث نضمن تمثيل العينة لكل طبقات المجتمع.

ويتم توزيع العينة بين الطبقات بثلاث طرق هي

١. التوزيع المتساوي وفيه يتم توزيع العينة على جميع طبقات الإطار بالتساوي.
٢. التوزيع المتناسب وفيه يتم توزيع العينة على طبقات الإطار بصورة تتناسب مع حجم كل طبقة .
٣. التوزيع الأمثل وفيه يؤخذ عدم التجانس داخل الطبقات وحجم الطبقة وتكلفة مشاهدة الوحدة فيها في الاعتبار بحيث يزيد عدد المفردات المأخوذة من الطبقات غير المتجانسة وذات الحجم الكبير والأقل تكلفة .

ولكن وفي كل الحالات فإن المعاينة الطبقية يراعى فيها أول ما يراعى أن تكون الطبقات متباينة فيما بينها متجانسة في داخلها وفي هذه الحالة وحدها تكون أخطاء المعاينة للعينة الطبقية أصغر من أخطاء المعاينة العشوائية البسيطة .

مثال (٢)

إذا كان الفصل في المثال السابق به ١٢٠ طالب و ٨٠ طالبة فإن أخذ عينة طبقية بتوزيع متناسب من هذا الفصل حسب النوع يتم كما يلي

٢. نقسم الفصل إلى قسمين بالتناسب أي ١٢٠ : ٨٠ أي ١٢٠ طالب و ٨٠ طالبة .
٣. تؤخذ من كل فئة عينة عشوائية بسيطة بحيث يكون إجمالي حجم العينة يساوي ١٠ طالبة وطالبات .

ومن ذلك يتضح أننا لابد أن تأخذ ٤ من الطالبات و ٦ من الطلاب وتتميز المعاينة الطبقية على غيرها لأنها تستهدف دراسة كل طبقة من طبقات المجتمع على حد إلى جانب دراسة المجتمع بأكمله .

وتعطي المعاينة الطبقية أفضل النتائج إذا كانت الصفة التي يتم على أساسها التقسيم الطبقي ترتبط مع الصفة المطلوبة دراستها في العينة .

(ج) المعاينة المنتظمة

وفيها يتم ترتيب مفردات المجتمع بأرقام سلسلة ثم تحدد فترة المعاينة بقسمة حجم المجتمع الكلي على حجم العينة. فإذا كان طول هذه الفترة k مثلاً، نختار وحدة عشوائياً من الـ k وحدة الأولى، ثم نختار كل k وحدة بعد ذلك. فإذا كان ترتيب الوحدة المختارة r مثلاً نختار الوحدات ذات الرتب $r+k, r+2k, r+3k, \dots$ وهكذا .

مثال (٣)

تعود إلى المثال السابق ونأخذ عينة منتظمة مقدارها ١٠ طلاب من عدد ٢٠٠ طالب . فإذا افترضنا أن حجم المجتمع يساوي (N) وحجم العينة يساوي (n) فإن فترة المعاينة هي :

$$1 = N/n = 200/10 = 20$$

فإذا بدأنا بالطالب رقم ١ فإن أرقام الوحدات التي تدخل العينة المنتظمة تكون على النحو التالي

١	٢١	٤١	٦١	٨١
١٠١	١٢١	١٤١	١٦١	١٨١

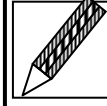
والسؤال هنا هو كيف يتم تحديد رقم المفردة الأولى في المعاينة المنتظمة ؟ فمثلاً إذا بدأنا بالطالب رقم ٥ فإن المعاينة المنتظمة تكون على النحو التالي

٥	٢٥	٤٥	٦٥	٨٥
١٠٥	١٢٥	١٤٥	١٦٥	١٨٥

وهكذا فإن اختيار المعاينة المنتظمة يعتمد على اختيار المفردة الأولى وفترة المعاينة ، ولتحقيق مبدأ العشوائية في المعاينة المنتظمة يتم السحب اعتباراً من رقم مفردة يتحدد عشوائياً ، ويمكن أن يتم ذلك باختيار مفردة واحدة من العشرين مفردة الأولى بطريقة الصندوق المثالي:

ومن مزايا المعاينة المنتظمة السهولة وقلة التكلفة والجهد .

وضح الفرق بين المعاينة الطبقية
والمعاينة المنتظمة



(د) المعاينة متعددة المراحل Multi Stage Sampling

هذا النوع من المعاينة يستخدم في الدراسات الكبيرة ذات الطابع القومي حيث إنه يستفاد من التنظيم الإداري للدولة في تقسيم وحدات المعاينة ، والمعاينة متعددة المراحل يتم فيها عشوائيا اختيار عدد معين من وحدات المعاينة الأولية ثم من داخل كل وحدة يتم اختيار وحدات المعاينة التالية ثم من داخل كل وحدة يتم اختيار وحدات أقل حجما عشوائيا ويمكن تكرار العمل أكثر من مرحلة ، على سبيل المثال لقد كان التعداد الأول في السودان لعام ١٩٥٦ تعداد بالعينة واستخدمت فيه المعاينة متعددة المراحل . كانت المديرية في ذلك الوقت تمثل وحدات المعاينة الأولية وقد تم اختيار عدد من المراكز من داخل كل مديرية بطريقة عشوائية في المرحلة الثانية ، في المرحلة الثالثة تم اختيار عدد من المعموديات من داخل كل مركز بطريقة عشوائية . ومن كل عمودية تم اختيار عدد من المشايخ بطريقة عشوائية في المرحلة الرابعة وفي المرحلة الخامسة والأخيرة تمت معاينة جميع الأسر الواقعة في نطاق هذه المشايخ المختارة . مثل هذا النوع من المعاينة يسمى معاينة من خمس مراحل .

من أهم مزايا المعاينة متعددة المراحل انخفاض تكلفة انتقال الباحثين ففي المثال السابق نجد أن الأسر التي يتم اختيارها بهذه الطريقة تكون متمركزة في أقل عدد من القرى مما يقلل تكاليف دراستها مقارنة بالمعاينة العشوائية البسيطة حيث يتم الاختيار من كل المديرية . ولكن من عيوب المعاينة متعددة المراحل هو تراكم حجم الخطأ العشوائي ذلك لأن خطأ عدم التجانس يزداد حجما من مرحلة إلى أخرى ، لهذا السبب لا بد وأن نتأكد أن المرحلة الأولية تم تقسيمها تقسيما يضمن التمثيل الكافي للمجتمع الأصلي .

(هـ) المعاينة العنقودية Cluster Sampling

في بعض الأحيان يتم تقسيم المجتمع إلى مجموعات ثم يتم التعامل مع كل مجموعة كوحدة معاينة قائمة بذاتها

وتستخدم هذه الطريقة في حالات استبيانات الرأي بصورة منتظمة بحيث أنه وبعد اختيار المرحلة الأولى للمعاينة يمكن إعادة استخدامها بصورة مستمرة لدراسة التغيرات في الرأي العام ، وتختلف المعاينة العنقودية عن المعاينة الطبقية في مبدأ العناقيد والذي يحدد أن تكون العناقيد متباينة في داخلها متجانسة فيما بينها .

وفي المعاينة الطبقية يتم الاختيار من داخل الطبقات أما في المعاينة العنقودية يتم الاختيار بين العناقيد ، والمعاينة بمرحلتين أو أكثر تعتبر أمثلة للمعاينة العنقودية .

والمعاينة العنقودية تمتاز بالسهولة وانخفاض التكلفة إلا أن عيبها الأساسي يتلخص في زيادة حجم الخطأ الناجم عن تقسيم مجمع غير متجانس.

(و) المعاينة المتعددة الأغراض Multi Purpose Sampling

في هذه الطريقة يتم سحب عينة كبيرة من مفردات المجتمع لدراسة ظاهرة محددة ثم يتم سحب عينة أصغر من العينة الأولى لدراسة ظاهرة أخرى ثم عينة أصغر من العينة الثانية لدراسة ظاهرة أخرى وهكذا ، على سبيل المثال من إطار يشمل جميع النساء نختار عينة لدراسة مستوى الدخل ، ثم نختار عينة أصغر من النساء بغرض دراسة ظاهرة العنف ضد المرأة ومن هذا الإطار الجديد يمكن أن نختار عينة أصغر لدراسة النساء المتزوجات فقط بغرض دراسة استعمال وسائل منع الحمل وهكذا .

ويلاحظ أن هذه الطريقة للمعاينة تمتاز بثبات وحدة المعاينة في كل المراحل إلا أنها لا تصلح

في الحالات التي يتوجب أن يقسم فيها المجتمع إلى طبقات أو مراحل حيث أن وحدات المعاينة تختلف في طبيعتها من مرحلة إلى أخرى.

♦ المعاينة العمدية Purposive Sampling

في شرحنا للمعاينة الاحتمالية وجدنا أن مبدأ الاختيار العشوائي لمفردات العينة يمثل المرتكز الأساسي للعملية الإحصائية .

أما في المعاينة العمدية فإن الباحث لا يهتم كثيراً بمبدأ العشوائية بل وفي بعض الأحيان يتحكم في اختيار أو عدم اختيار المفردة في العينة .

والمعاينة العمدية لا علاقة لها بالعملية الإحصائية إلا من قبيل الصدفة ، ذلك لأنه لا يجوز هنا تطبيق قواعد نظرية العينات ويتم إجراء العينة بإحدى طريقتين:

(أ) طريقة الاختيار الوسيط Average Selection Method

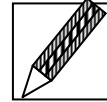
وفيها يختار الباحث مفردات المعاينة من المجتمع تحت الدراسة حسب وجهة نظره على إنها تمثل مفردات وسط ، فمثلا في حالة اختيار عينة لذوي الدخل المحدود ربما يختار الباحث الأشخاص من ذوي الدخل المتوسطة لا هي بالعالية ولا هي بالمنخفضة وبذلك تصبح عينة عميدة بحيث يعتمد عليها في العملية الاختيارية الوسطية .

(ب) طريقة الزمرة Quota Sampling

وفي هذه الطريقة يتم الاختيار بطريقة صدفية بحتة دون الاهتمام بالأسلوب العشوائي على سبيل المثال إذا أراد تاجر فواكه أن يفحص صلاحية كمية من البرتقال فإنه ربما يختار صندوق أو صندوقين من هذا البرتقال ويقوم بفحص مفرداته ويقرر على ضوء ذلك سلامة أو عدم سلامة البرتقال . ونلاحظ في هذه الطريقة أن ذلك التاجر قد قام باستخدام أسلوب التقسيم الطبقي بطريقة عمدية وبأقل تكلفة حيث لا يحتاج الأمر إلى صبر إحصائي لمعرفة أخطاء المعاينة . وقد ينجح ذلك التاجر في اتخاذ قراره إذا كان على علم مسبق ببعض المعلومات عن ذلك البرتقال . من كل ما تقدم يتضح لنا جليا أن أسلوب المعاينة هو أمر شائع بين الناس في الحياة العادية ، ولكن من المهم أن تفرق بين المعاينة العلمية القائمة على الاحتمال والاختيار العشوائي والتي تمكن من قياس الأخطاء والتحكم فيها وبين المعاينة غير العشوائية والتي لا تمكننا من الوصول إلى تقديرات إحصائية بالمعنى العلمي كما لا يترتب عليها إجراءات لاحقة أو دراسات متأنية .

تدريب (٣)

وضح بعض الحالات التي تستخدم فيها المعاينة العنقودية؟



سادسا أخطاء المعاينة Sampling Errors

كما ذكرنا سابقا أن الغرض من المعاينة هو دراسة الجزء بغرض التعميم على الكل أي تحليل مفردات العينة وتعميم نتائجها على المجتمع ، وعند تنفيذ هذه العملية فإن هنالك نوعان من الأخطاء المحتملة وهي:

أ) الخطأ العشوائي Random Error

وهو الخطأ الذي ينتج عن طريقة الصدفة الناتجة عن طبيعة الاختيار العشوائي ويتوقف هذا الخطأ على طريقة المعاينة وحجم العينة وتباين المجتمع ، فكلما زاد حجم العينة أو زادت درجة التجانس بين مفردات المجتمع يقل مقدار الخطأ العشوائي ، ولذلك فإنه يمكن التحكم في الخطأ العشوائي بزيادة حجم العينة ، كما أن مبدأ العشوائية لو طبق تطبيقاً سليماً لا يمكن معه تقدير حجم الخطأ من العينة نفسها .

ب) خطأ التحيز Bias Error

وهذا الخطأ يحدث عادة بسبب الاختيار غير العشوائي و ينتج عن التحيز عند اختيار المفردات وهنالك أمور كثيرة تؤدي إلى التحيز في الاختيار تجمل أهمها فيما يلي

١. اختيار مفردات العينة وفق الخاصية المراد قياسها .
٢. التحيز الإرادي والذي يكون الهدف منه تحقيق غرض ما في نفس الباحث.
٣. عدم وجود إطار سليم لمفردات المجتمع عند إجراء المعاينة .
٤. الاستبدال المقصود لبعض وحدات مختارة من العينة لوحدات بديلة غير مختارة .
٥. الفشل في الحصول على قسم كبير من المفردات المختارة في العينة .

وخطأ التحيز ليس قاصراً على أسلوب المعاينة فقط وإنما يشمل الحصر الشامل أيضاً نتيجة لعدم الدقة في القياس أو عدم الشمول الناتج عن الكم أو المحتوى.

سابعاً أخطاء غير المعاينة

وهي أخطاء شائعة تحدث أثناء مراحل جمع وتجهيز البيانات ويمكن أن نجملها فيما يلي:

١. أخطاء التبليغ وتحدث من الباحث بسبب النسيان أو عدم الفهم للمادة المبحوثة (وهذا يكون في التعداد السكاني).
٢. عدم التجاوب في التزويد بالمعلومات الكافية مما يؤدي إلى خطأ من المبحوث أما لعدم الاهتمام أو لعدم وجود المعلومات الكافية عن المادة المبحوثة .
٣. أخطاء الترميز وتحدث عادة في مرحلة تجهيز البيانات عند استبدال الكلمات برموز رقمية غير مطابقة .

قم بكتابة وتلخيص أهم المعلومات التي ذكرت عن أساليب جمع البيانات مع توضيح أهمية كل أسلوب وأعرض ما قمت به على مشرفك الأكاديمي



أسئلة تقويم ذاتي



- ١- عرف المعاينة العشوائية والمعاينة العمدية؟
- ٢- ما هي الحالات التي تستخدم فيها المعاينة العمدية؟
- ٣- عدد أنواع المعاينة العشوائية؟
- ٤- ما هي مزايا المعاينة العشوائية البسيطة؟
- ٥- متى يتوجب على الباحث استخدام المعاينة الطبقية؟
- ٦- ما هي العينة المنتظمة ومتى تستخدم؟
- ٧- ما هي مزايا المعاينة المنتظمة؟
- ٨- ما هي مزايا وعيوب المعاينة متعددة المراحل؟
- ٩- في المعاينة العمدية يتم إجراء العينية بإحدى طريقتين ما هما؟
- ١٠- تحدث باختصار عن أخطاء المعاينة وأخطاء غير المعاينة؟
- ١١- ميز بين :
أ- الخطأ العمدى والخطأ العشوائى
ب- المجتمع والعينة
ج- المجتمع والإطار

٣. عرض البيانات Data Presentation

١.٣ العرض الجدولي للبيانات

بانتهااء عملية جمع البيانات تبدأ عملية المراجعة . والمراجعة قد تكون مكتبية او ميدانية الغرض منها محاولة اكتشاف الأخطاء قبل تصنيف البيانات وتبويبها . بعد الانتهاء من عملية المراجعة تبدأ عملية تبويب البيانات وذلك بغرض تنظيمها وتلخيصها وإعدادها بصورة يسهل معها استخلاص أهم النتائج المضمنة فيها . يبدأ عرض البيانات الإحصائية بتصنيفها حيث يلزم التفرقة بين البيانات الوصفية والكمية .

البيانات الوصفية

البيانات الوصفية هي عبارة عن مجموعة أوصاف لبعض الخصائص حسب ورودها في استمارة الاستبيان . وتنقسم الجداول التي يمكن بواسطتها تبويب البيانات الوصفية إلى نوعين: جداول عامة وأخرى تلخيصية . يشتمل النوع الأول على كافة المعلومات التي ترد عادة في استمارة الاستبيان وهي جداول تفصيلية في معظم محتوياتها . وتستخدم كمراجع للدراسة ، ونجدها غالبا كملحق في نهاية التقرير للبحث .

ولتوضيح ذلك نعرض المثال التالي

مثال (١)

كانت المتغيرات في نتيجة الإحصاء لـ ٥٠ طالبا في الفصل الدراسي الثالث بقاعدة العلوم الاجتماعية في كلية الاقتصاد والتنمية الريفية للعام الدراسي ١٩٩١/٩٠م كما يلي:

D	B	C	A	C	B	A	B	B	A
D	A	D	B	D	A	B	B	C	B
F	F	B	A	C	B	A	C	A	C
F	B	A	B	F	B	B	D	B	B
D	A	C	B	B	A	A	D	B	B

البيانات الواردة في مثال (١) هي بيانات تفصيلية ولا يمكن الاستفادة منها في دراسة ، وذلك لعدم وضوحها وصعوبة استنتاج أي معالم عن مستوى الطلاب في الفصل فمثلا لا

يمكننا معرفة عدد الطلاب الذين نالوا تقدير A بسهولة من البيانات أعلاه وخاصة إذا كان عدد الطلاب كبيراً. لذلك أصبحت الحاجة إلى إيجاد طريقة لتنظيم وتلخيص مثل هذه البيانات في صورة سهلة ضرورية. ولذلك نلجأ إلى ما يعرف بالجدول التلخيصية.

والجدول التلخيصية تشتق عادة من الجداول التفصيلية العامة وتستخدم بغرض إظهار المؤشرات والمقاييس الإحصائية اللازمة لوصف البيانات الواردة في الدراسة وأهم أنواع الجداول التلخيصية هو الجدول التكراري.

وسنقوم بتعريف وشرح طريقة عمل الجدول التكراري فيما يلي:

٢. الجداول التكرارية الوصفية :

أولى الخطوات في تكوين الجدول التكراري الوصفي هي تكوين جدول تفريغ البيانات ومنه نستنتج جدولاً آخر يسمى جدول التوزيع التكراري على النحو التالي؛ انظر جدول رقم (١)

جدول رقم (١)

تفريغ البيانات لتقديرات الطلاب في مثال (١)

التقدير	العلامات	التكرار (عدد الطلاب)
A	/// ###	12
B	### ### ###	20
C	// ###	7
D	// ###	7
F	////	4
المجموع		50

ونلاحظ إننا في العمود الأول نكتب الصفة وهي في هذه الحالة التقدير. أما في العمود الثاني فيكون العنوان العلامات. ونضع لكل قراءات علامة أمام التقديرات الموجودة في العمود الأول. والعلامة عبارة عن خط رأسي مائل (/) فإذا وصل عدد العلامات أي أربعة فان الخط الخامس يكتب بشكل عرضي لتشكيل المجموعة حزمة خماسية. وبعد تفريغ جميع البيانات تعد الحزم أمام كل صفة ويكتب عددها في العمود الثالث الذي يسمى عمود التكرارات ويقصد بها عدد الطلاب أمام كل تقدير. ولا بد وان يكون مجموعها يساوي مجموع التقديرات التفصيلية وهي في هذه الحالة ٥٠ تقديراً.

وإذا حذفنا العمود الثاني من الجدول السابق (عمود العلامات) فإننا نحصل على جدول مألوف يتكون من عمودين

ويسمى جدول التوزيع التكراري كما هو موضح في الجدول رقم (٢) التالي:

الجدول رقم (٢)
التوزيع التكراري لتقديرات توزيع الطلاب في مثال رقم (١)

التقدير	التكرار (عدد الطلاب)
A	١٢
B	٢٠
C	٧
D	٧
F	٤
المجموع	٥٠

المصدر: استاذ المادة لعام ١٩٩٠-الدفعة ١١

الفصل الدراسي الخامس

لا بد من ملاحظة أن أي جدول إحصائي يحتوي على عنوان يوضح نوعية الجدول وطبيعة البيانات المعروضة فيه، كما يوضح مصدر هذه البيانات.

الجدول التكرارية الرقمية

والجداول التكرارية الرقمية تستخدم في حالة قياس الظواهر بمقياس كمي (رقمي) مثل الأجور ودرجات الطلاب وقيم الإنتاج والاستهلاك وحسابات الدخل وتوزيع السكان و غيرها . ولتوضيح طريقة تكوين الجدول التكراري الرقمي نأخذ المثال التالي:

مثال (٢)

الجدول التالي يوضح عدد جوالات الذرة التي حصل عليها كل مزارع في ٥٠ حواشة بمشروع الجزيرة للموسم الزراعي ٩٠-١٩٩١م .

جدول رقم (٣)

عدد جوالات الذرة التي حصل عليها كل مزارع

25	30	38	42	51	34	42	54	34	42
39	40	50	26	52	38	47	35	53	28
34	41	35	31	41	36	53	41	36	32
37	44	45	37	45	46	29	46	38	48
40	33	44	45	44	40	31	42	43	27

أولى الخطوات التي يجب اتخاذها لتنظيم هذه البيانات هي ترتيبها ترتيبا تصاعديا أو تنازليا وذلك لإعطاء فكرة عامة عن توزيع إنتاج الذرة. فإذا أخذنا بالترتيب التصاعدي للبيانات السابقة نتحصل على الجدول التالي:

جدول رقم (٤)
الترتيب التصاعدي للبيانات في الجدول رقم (٣)

٥٤	٥٠	46	44	42	40	38	35	33	29
٥٣	٤٨	45	44	41	40	37	35	32	28
٥٣	٤٧	45	43	41	39	37	34	31	27
٥٢	٤٧	45	42	41	38	36	34	31	26
٥١	٤٦	44	42	40	38	36	34	30	25

ونلاحظ أن نصف الإنتاج يتراوح بين ٤٤,٣٥ لحواشة الواحدة وأن أعلى إنتاجية كانت ٥٤ جوال وأقل إنتاجية كانت ٢٥ جوال. وعلى الرغم من أن فكرة الترتيب قد أبرزت لنا بعض المؤشرات إلا أنها ما زالت هنالك صعوبة في تفسير هذه البيانات. وتزداد هذه الصعوبة إذا كان المجتمع كبيرا. ولتنظيم هذه البيانات وتلخيصها ووضعها في جدول تكراري نكون أولا جدولا للتفريغ كما فعلنا في حالة البيانات الوصفية مع استبدال الصفة في العمود الأول بمجموعات تسمى فئات. وقبل أن نقوم بتكوين جدول التفريغ نستعرض فيما يلي طريقة اختيار الفئات.

أولا تحديد عدد الفئات

لتحديد عدد الفئات نتبع الخطوات التالية

أولا نحدد مدى البيانات

وهو يمثل الفرق بين أكبر قيمة في المفردات وأصغر قيمة. وفي المثال رقم (٢) يكون المدى كما يلي

المدى (Range- R) = أكبر قيمة - أصغر قيمة

$$R = 54 - 25 = 29 \text{ (جوال)}$$

ولا توجد قاعدة ثابتة لتحديد عدد الفئات ولكن عادة يقسم المدى إلى أعداد مناسبة من الفئات لا يقل عن ٥ ولا يزيد عن ١٥ فئة تقريبا. كما يمكن الاستعانة بالقاعدة التالية لتحديد عدد الفئات

$$NC = 1 + 3.222 \log N$$

$$NC = 1 + 3.222 \log 50 = 6$$

حيث n هو عدد المفردات وفي هذا المثال نجد ان عدد الفئات يساوي ٦ فئات تقريبا

ثانيا طول الفئة

بتحديد المدى وعدد الفئات يمكننا بكل سهولة حساب طول الفئة من القانون التالي

$$C = R/NC$$

حيث C تمثل طول الفئة و R هو المدى الكلي و n عدد الفئات وبالرجوع للمثال رقم (٢) نجد ان طول الفئة هو ٥.

$$C = 29/6 = 5$$

ومن الأفضل دائما أن يكون طول الفئة رقميا دائريا ٥ او ١٠ أو ١٥... الخ

ثالثا حدود الفئات

بعد تحديد عدد وأطوال الفئات يصبح من الضروري ترسيم حدودها بحيث لا تتداخل مع بعضها البعض. وهنا من الأفضل دائما أن نبدأ الحد الأدنى للفئة الأولى بأصغر عدد في المجموعة ما أمكن ذلك. وعليه سيكون الحد الأدنى للفئة الأولى هو ٢٥ ثم نضيف طول الفئة لنحصل على الحد الأعلى للفئة الأولى وهو ٣٠، ثم نبدأ الحد الأدنى للفئة الثانية وهو يساوي الحد الأعلى للفئة الأولى وهو ٣٠ ثم نضيف طول الفئة لنحصل على الحد الأعلى للفئة الثانية... وهكذا.

باستخدام الخطوات السابقة يمكن تحديد فئات المثال رقم (٢) على النحو التالي

$$(٢٥-٣٠) ، (٣٠-٣٥) ، (٣٥-٤٠) ، (٤٠-٤٥) ، (٤٥-٥٠) ، (٥٠-٥٥)$$

إذا كانت أطوال الفئات متساوية فيمكن كتابة الحد الأدنى فقط وترك الحد الأعلى الذي يتحدد تلقائيا بطول الفئة.

رابعا تكوين الجدول

بعد الفراغ من تحديد عدد وأطوال حدود الفئات نقوم بتكوين جدول تفريغ البيانات الكمية بنفس الطريقة التي اتبعناها عند تكوين جدول البيانات الوصفية على النحو التالي

جدول رقم (٥)
جدول تفريغ البيانات لانتاج الذرة في مثال (٢)

التكرار (عدد الحواشات)	العلامات	الفئات (إنتاج الذرة بالجوال)
5	###	-25
8	/// ###	-30
10	### ###	-35
13	///### ###	-40
8	/// ###	-45
6	/###	-50
50		المجموع

لاحظ اننا اكتفينا بكتابة الحدود الدنيا للفئات لان اطوال الفئات متساوية وهي تقرأ كما يلي: من ٢٥ الى اقل من ٣٠ ، من ٣٠ الى اقل من ٣٥ ... وهكذا
ويمكننا حذف عمود العلامات للحصول على جدول التوزيع التكراري من الجدول رقم (٢-٥) ، بحيث يصبح الجدول من عمودين فقط : الاول يمثل فئات إنتاج الذرة بالجوال والثاني يمثل التكرارات لها أي عدد الحواشات المقابلة لكل فئة ، ويكتب كما يلي

جدول رقم (٦)
جدول التوزيع التكراري لانتاج الذرة في مثال رقم (٢)

التكرار f	الفئات
5	25-
8	30-
10	35-
13	40-
8	45-
6	50
50	المجموع

خامسا الحدود الفعلية للفئات
ذكرنا سابقا ان الفئة تتكون من حد ادنى وحد اعلى والذان يسميان بحدود الفئات (Class Limits) . نلاحظ ان حدود الفئات لها توزيع منقطع (Discrete) ولكن الحدود الفعلية

للفئات يجب ان يكون لها توزيع متصل (Continuous) ولذلك نقوم باستبدال حدود الفئات بحدود متصلة تسمى حواجز الفئات (Class Boundaries) وتعرف في بعض الأحيان بالحدود الفعلية للفئات. ويتم تكوين حواجز الفئات بطرح ٠.٥ من الحد الأدنى للفئة لنحصل على الحاجز الأدنى للفئة . كذلك نضيف ٠.٥ للحد الأعلى للفئة لنحصل على الحاجز الأعلى للفئة ... وهكذا . في المثال السابق يصبح التوزيع التكراري لحواجز الفئات كما يلي

جدول رقم (٧)
جدول التوزيع التكراري لانتاج الذرة

التكرار f	حواجز الفئات
5	29.5-24.5
8	34.5-29.5
10	39.5-34.5
13	44.5-39.5
8	49.5-44.5
6	54.5-49.5
50	المجموع

ونلاحظ اننا نرمز للتكرار Frequency f بالحرف

سادسا: مراكز الفئات

تحتسب مراكز الفئات والتي سنرمز لها بالحرف (X) من العلاقة التالية :

حيث L = الحاجز الأدنى للفئة ، U = الحاجز الأعلى للفئة
وإذا قمنا بتكوين مراكز الفئات للجدول رقم (٧-٢) نحصل على الجدول التالي:

جدول التوزيع التكراري لإنتاج الذرة في مثال (٢)
جدول رقم (٨)

مراكز الفئات X	التكرار f	حواجز الفئات
27	5	29.5-24.5
32	8	34.5-29.5
37	10	39.5-34.5
42	13	44.5-39.5
47	8	49.5-44.5
52	6	54.5-49.5
	50	المجموع

أنواع الجداول التكرارية

ويسمى الجدول التكراري جدول منتظم إذا تساوت جميع فئاته. أما إذا اختلفت أطوال الفئات فيكون الجدول غير منتظم. ولا يلجأ الباحثون للجداول غير المنتظمة إلا إذا كانت البيانات تتطلب ذلك كأن يتمركز عدد كبير من المفردات في مدى ضيق بينما ينتشر جزء كبير منها على مدى واسع. على سبيل المثال عند دراسة توزيع الدخل للأفراد نجد على طرفي التوزيع أصحاب الدخل المنخفضة وأصحاب الدخل العالية تنتشر على مدى واسع. والجدول رقم (٩) يعطي مثالا للجدول التكراري غير المنتظم

جدول رقم (٩)
جدول التوزيع التكراري لمستويات الدخل الشهري
لعينة من الاسر (بالجنيحات السودانية)

التكرار f	فئات الدخل
500	600-
300	1000-
400	1500-
600	2000-
1000	2500-
1500	3000-
800	4000-
200	6000-
120	10000-
50	20000-
5450	المجموع

المصدر بيانات افتراضية

من الواضح ان فئات الدخل في جدول رقم (٩) لها اطوال متباينة . على سبيل المثال طول الفئة الاولى يساوي ٤٠٠ بينما طول الفئة الثانية ٥٠٠ والفئة قبل الاخيرة ١٠٠٠٠٠ ، وبالتالي هو جدول تكراري غير منتظم .

الجدول التكرارية المفتوحة

في بعض الدراسات يصعب تحديد الحد الادنى للفئة الاولى وفي هذه الحالة يسمى الجدول التكراري مفتوحا من اسفل . كما يصعب احيانا تحديد الحد الاعلى للفئة الاخيرة فيسمى الجدول التكراري مفتوحا من اعلى . ويمكن ان يكون الجدول مفتوح من الطرفين .

جدول رقم (١٠)
التوزيع العمري لعينة من النساء المتزوجات في المسح
الصحي والديمغرافي
السودان ٨٩-١٩٩٠م

فئات العمر	عدد النساء f
اقل من 15	380
15-	983
20-	1355
25-	970
30-	1047
35-	630
40 فأكثر	540

المصدر وزارة التخطيط - مصلحة الاحصاء - المسح الصحي والديمغرافي ص ١٧-
 الخرطوم - مايو ١٩٩١

ج- الجداول التكرارية المتجمعة

في بعض الأحيان يتطلب البحث معرفة عدد أو نسبة المفردات التي تقل عن قيمة معينة أو تلك التي تزيد عن حد معين. يتم ذلك عن طريق الجداول التكرارية المتجمعة، وهي نوعان:

الجدول التكراري المتجمع الصاعد

التكرار المتجمع الصاعد المقابل للحاجز الأعلى لفئة ما هو عدد القيم التي تقل عن ذلك الحاجز. أي هو مجموع تكرارات هذه الفئة والفئات التي قبلها. الجدول التكراري يسمى الجدول التكراري المتجمع الصاعد. بمعنى آخر فإنه في الجدول التكراري المتجمع الصاعد يخصص العمود الأول للحواجز العليا للفئات والعمود الثاني للتكرار المتجمع الصاعد حيث يجمع تكرار كل فئة إلى مجموعة التكرارات السابقة لها. ولتوضيح هذا نوجد الجدول التكراري المتجمع الصاعد من جدول رقم (١١) الذي يتضمن انتاج الذرة في مشروع الجزيرة.

جدول رقم (١١)
الجدول التكراري للمتجمع الصاعد لإنتاج الذرة بمشروع الجزيرة

المتجمع الصاعد	فئات الانتاج
0	اقل من 24.5
5	اقل من 29.5
13	اقل من 34.5
23	اقل من 39.5
36	اقل من 44.5
44	اقل من 49.5
50	اقل من 54.5

يلاحظ ان التكرار المتجمع الصاعد المقابل لأقل من الحاجز الأدنى للفئة الأولى يساوي صفر والذي لأقل من الحد الأعلى للفئة الأخيرة يساوي المجموع أي ٥٠ طالبا .
الجدول التكراري للمتجمع الهابط:
في بعض الاحيان قد يتطلب الامر الحصول على التوزيع التكراري للمتجمع لجميع القيم الاكبر من او المساوية للحاجز الأدنى للفئة . وهذا يسمى التوزيع التكراري للمتجمع الهابط . وفيه تمثل فئات الانتاج الحواجز الدنيا للفئات والتكرار المتجمع الهابط يمثل تراكم التكرارات من اسفل الى اعلى .

جدول رقم (١٢)
الجدول التكراري للمتجمع الهابط لإنتاج الذرة بمشروع الجزيرة

المتجمع الهابط	فئات الانتاج
50	24.5 فأكثر
45	29.5 فأكثر
37	34.5 فأكثر
27	39.5 فأكثر
14	44.5 فأكثر

49.5 فأكثر	6
54.5 فأكثر	0

الجدول التكراري النسبي

التكرار النسبي لفئة هو تكرار الفئة مقسوماً على التكرار الكلي لجميع الفئات ويعبر عنه كنسبة مئوية .
ويتم اعداد الجداول التكرارية النسبية دائماً من الجداول التكرارية البسيطة ، في الجدول التالي نوضح التكرارات النسبية لانتاج الذرة في مشروع الجزيرة .

جدول رقم (١٣)

الجدول التكراري النسبي لانتاج الذرة في مثال (٢)

فئات الانتاج	التكرار f	التكرار النسبي
29.5-24.5	5	10
34.5-29.5	8	16
39.5-34.5	10	20
44.5-39.5	13	26
49.5-44.5	8	16
54.5-45.5	6	12
المجموع	50	100

٢.٣ العرض البياني للبيانات

يؤدي العرض البياني إلى زيادة إيضاح البيانات ويمكن للقارئ أن يكون فكرة سريعة بمجرد النظرة السريعة الفاحصة للرسوم البيانية ، ولكن يؤخذ عليها أنه لا يمكن الاستفادة منها إلا في إلقاء نظرة أو فكرة عن البيانات حيث أنها توضح قدر بسيط من المعلومات .

طرق عرض البيانات

تنقسم الأشكال والرسوم البيانية إلى قسمين رئيسيين :
عرض البيانات للجداول البسيطة وعرض البيانات للجداول التكرارية :

أولاً: العرض البياني للجداول البسيطة :

(أ) الخط البياني Line Chart

ويستخدم لتوضيح نمط اتجاه ظاهرة ما خلال فترة زمنية محددة ، ولرسم الخط البياني يوضع الزمن على المحور الأفقي والظاهرة تحت الدراسة على المحور الرأسي ، ثم تحدد النقاط حسب الإحداثيات المعطاة ثم توصل بخطوط مستقيمة على سبيل المثال يمكن تمثيل البيانات الخاصة

بتطور التعليم الابتدائي والمتوسط في السودان للفترة 1980/74 والواردة في الجدول (١٤) بخط بياني انظر شكل (14)

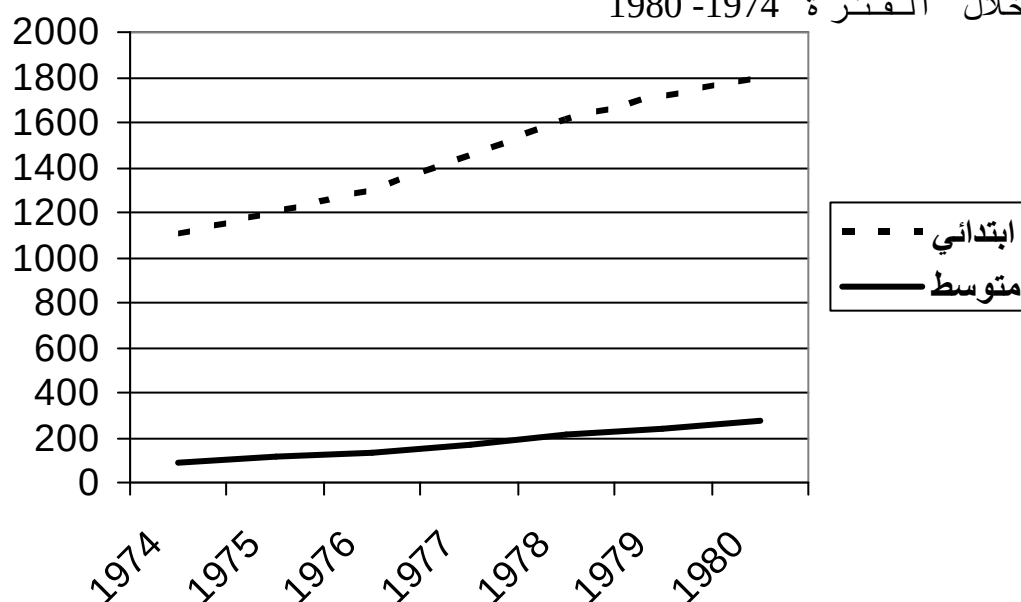
جدول (١٤)
توزيع أعداد تلاميذ المدارس الابتدائية والمتوسطة في السودان خلال الفترة

العام	عدد التلاميذ (بالآلاف)	
	ابتدائي	متوسط
١٩٧٤	١١٠٠	٩٠
١٩٧٥	١٢٠٠	١٢٠
١٩٧٦	١٣٠٠	١٣٠
١٩٧٧	١٥٠	١٧٠
١٩٧٨	١٦١٠	٢١٠
١٩٧٩	١٧٢٠	٢٤٠
١٩٨٠	١٨٠٠	٢٧٩

المصدر الإحصاء التربوي - وكالة التخطيط التربوي ١٩٨٥ م .

شكل (١)

توزيع أعداد تلاميذ المدارس الابتدائية والمتوسطة في السودان خلال الفترة 1980-1974



في بعض الأحيان تكون الظاهرتان تحت الدراسة مقاسة بنفس الوحدات ولها علاقة ببعضهما البعض ، كما يكون للفرق بينهما معني معين ، على سبيل المثال الدخل و الإنفاق

يعبر عنه بالادخار وهو الفرق بينهما ، الصادرات والواردات يعبر عنه بالميزان التجاري ، والمواليد والوفيات يعبر عنه بمعدل النمو السكاني وهكذا . في مثل هذه الحالات تمثل كل ظاهرة بخط بياني كالمعتاد ويمثل الفرق بينهما المساحة المحصورة بين الخطين . تتضح هذه المسألة عند مقارنة المواليد و الوفيات بمحافظة الجزيرة في الفترة من ١٩٨٠ - ١٩٩٠ م .

جدول (15)

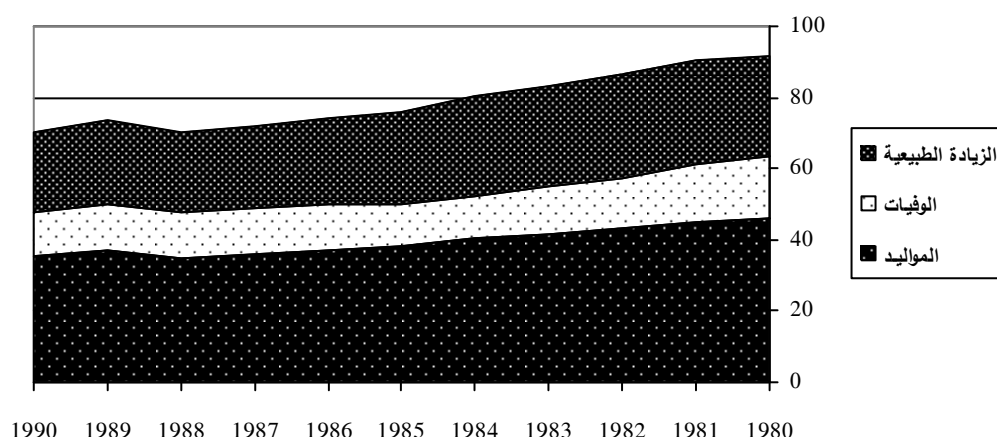
المواليد والوفيات والزيادة الكلية محافظة الجزيرة ١٩٨٠ - ١٩٩٠ م (بالآلاف) .

السنة	المواليد	الوفيات	الزيادة الطبيعية
١٩٨٠	٤٥.٩	١٧.٤	٢٨.٥
١٩٨١	٤٥.٩	١٦.١	٢٩.٠
١٩٨٢	٤٣.٣	١٤.٢	٢٩.١
١٩٨٣	٤١.٥	١٣.٦	٢٧.٩
١٩٨٤	٤٠.٢	١٢.٠	٢٨.٢
١٩٨٥	٣٨.٠	١٢.١	٢٥.٩
١٩٨٦	٣٧.٢	١٢.٦	٢٤.٦
١٩٨٧	٦.٠	١٣.١	٢٢.٩
١٩٨٨	٣٥.١	١٢.٨	٢٢.٣
١٩٨٩	٣٦.٨	١٣.١	٢٣.٧
١٩٩٠	٣٥.٢	١٢.٦	٢٢.٦

المصدر التقرير الختامي للمسح الصحي والديمغرافي -
مصلحة الإحصاء ١٩٩١ م .

الشكل (٢)

المواليد والوفيات والزيادة الطبيعية بمحافظة الجزيرة
١٩٨٠ - ١٩٩٠ م



ملحوظة

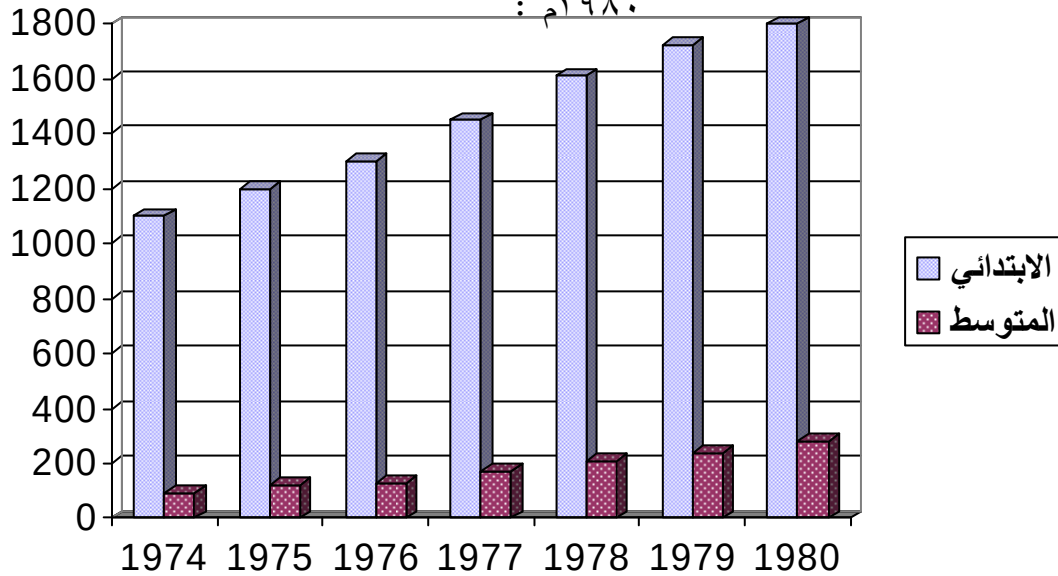
عند رسم الخط البياني لابد من تعيين مقياس رسم مناسب على كل من المحورين ، فإذا كان مقياس الرسم على المحور الأفقي كبيراً مقارنة مع مقياس الرسم للمحور الرأسي فإنه يمكن كسر المحاور وذلك سيؤدي إلى الزيادة في قيمة التغيرات الظاهرة على الرسم والعكس صحيحاً ، وتستخدم طريقة كسر المحاور كعلاج لمثل هذه الحالات.

(ب) الأعمدة البيانية Bar chart

الأعمدة البيانية تمثل أعمدة رأسية قواعدا متساوية وتتناسب ارتفاعاتها مع قيم الظاهرة تحت الدراسة . المحور الأفقي في الأعمدة البيانية يمثل الصفة المميزة للبيانات والمحور الرأسي يمثل القيم المختلفة للظاهرة المعنية ، ففي مثال تلاميذ المدارس المتوسطة فإن تطور أعداد التلاميذ خلال الفترة ١٩٧٤ - ١٩٧٦ يعبر عنه بالتغير في ارتفاعات الأعمدة كما هو موضح.

شكل (3)

تلاميذ المدارس المتوسطة في السودان خلال الفترة ١٩٧٤ - ١٩٨٠م :



ويلاحظ في الأعمدة البيانية عدم كسر المحور الأفقي ، ولكن يجوز كسر الأعمدة إذا كان الرسم يحتوي على عمود أو أكثر أطول بكثير من باقي الأعمدة .

ويلاحظ في الأعمدة البيانية عدم كسر المحور الأفقي ، ولكن يجوز كسر الأعمدة إذا كان الرسم يحتوي على عمود أو أكثر أطول بكثير من باقي الأعمدة .
وكسر العمود يمكن أن يكون جزئيا مع تكملته أفقيا أو أن يكسر الأعمدة إذا كان الرسم يحتوي على عمود أو أكثر أطول بكثير من باقي الأعمدة .
وكسر العمود يمكن أن يكون جزئيا مع تكملة أفقيا أو أن يكسر الجزء الزائد من العمود من أعلى مع كتابة القيمة العددية داخله الجزء المكسور على سبيل المثال إذا أخذنا الثلاث سنوات الأولى في جدول (3) وعبرنا عنها بأعمدة بيانية نجد أن أعداد طلاب المدارس الابتدائية أكبر بكثير من طلاب المدارس المتوسطة وفي هذه الحالة يمكننا كسر الأعمدة الخاصة بطلاب المدارس الابتدائية .
ملاحظات على الأعمدة البيانية

١. لا يجوز كسر المحاور .
٢. لابد وأن تكون قواعد الأعمدة متساوية وأن تكون المسافات بينهما أيضا متساوية .
٣. لا يجوز استخدام الأعمدة البيانية في حالة البيانات المتصلة .
٤. يفضل استخدام النسب المئوية بدلا عن القيم الرقمية .

(ج) الرسوم الدائرية

وهي عبارة عن دائرة ذات نصف قطر مناسب بحيث يتم تقسيمها إلى قطاعات تتلاقى في المركز وتتناسب مساحة هذه القطاعات مع القيم الجزئية لاظهار الأهمية النسبية لكل مشاهدة ، ويمكن حساب الزاوية المركزية للقطاع باستخدام القانون التالي:

$$\text{الزاوية المركزية للقطاع} = \frac{\text{عدد المشاهدات}}{360} \times$$

ونوضح ذلك بالمثل التالي

جدول (16)

يمثل وسائل منع الحمل المستخدمة لعينة من النساء بمركز ابن حيان بجزيرة الفيل لعام ١٩٩١

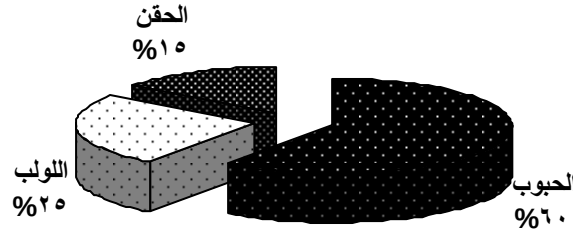
الوسيلة	عدد النساء	النسبة المئوية	الزاوية المركزية
الحبوب	٦٠	٦٠	٢١٦
اللؤلؤ	٢٥	٢٥	٩٠
الحقن	١٥	١٥	٥٤
المجموع	١٠٠	١٠٠	٣٦٠

المصدر: مكتب الإحصاء - مركز ابن حيان - جزيرة الفيل ١٩٩١م

يلاحظ أن مجموع النساء المراد تمثيلهم يساوي (1000) امرأة قمنا بتوزيعهن بالنسبة المئوية كأول خطوة ، وفي الخطوة الثانية حسبنا الزاوية المركزية لكل وسيلة كما يلي:
الزاوية المركزية للحبوب = $216 = 360 \times 60 / 100$
الزاوية المركزية للؤلؤ = $90 = 360 \times 25 / 100$
الزاوية المركزية للحقن = $54 = 360 \times 15 / 100$
الآن يمكن رسم الدائرة بعد تحديد نصف قطر مناسب وليكن في هذه الحالة ٤ سم كما يلي:

شكل (4)

استخدام وسائل منع الحمل بمركز ابن حيان جزيرة الفيل ١٩٩١م



والرسوم الدائرية تستخدم استخدما فعالا في مقارنة البيانات ، وهنا نرسم كل دائرة بحيث يتناسب نصف قطرها مع الجذر التربيعي لحجم الظاهرتين ويحسب ذلك من العلاقة التالية:

$$\frac{R_1}{R_2} = \sqrt{\frac{X_1}{X_2}}$$

حيث أن X_1, X_2 تمثلان أجزاء المجاميع و R_1, R_2 تمثلان أنصاف أقطار كل من الدائرتين . وفيما يلي نقوم بعملية مقارنة

بالدوائر على بيانات صادرات دولة الكويت عامي ١٩٧٣ و ١٩٨٣ حسب ما هو موضح في الجدول (17).

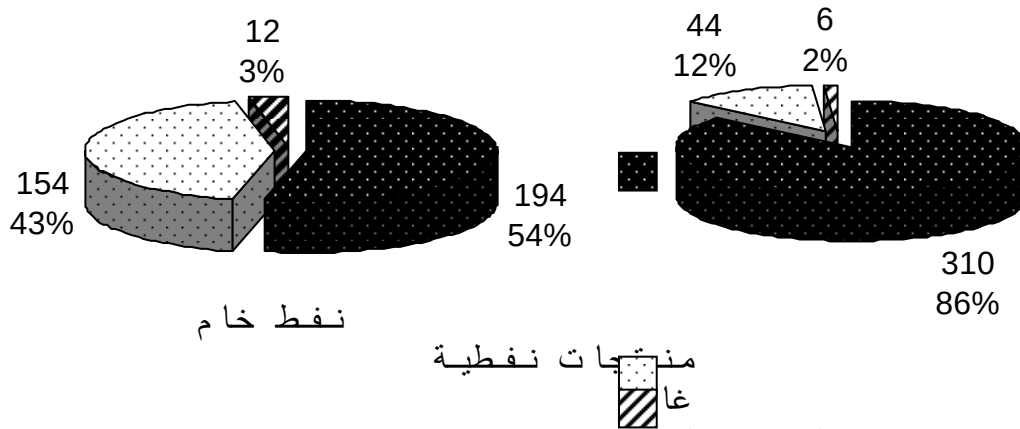
جدول (17)

قيم صادرات النفط لدولة الكويت 1983-1973

بيان ١٩٨٣			بيان ١٩٧٣			
الزاوية	%	القيمة	الزاوية	%	القيمة	
١٩٤	٥٣.٧	١٥٧٨١٧ ١	٣١.٠	٨٦.٢	٩١٣٠٢٢	نفط خام
١٥٤	٤٢.٩	١٢٥٩٢٦ ٩	٤٤	١٢	١٣٠٨٧٠	منتجات نفطية
١٢	٣.٤	١٠٠٧٥٢	٦	١.٥	١٦٠٢٤	غاز
٣٦.٠	١٠٠	٢٩٣٨١٩ ٢	٣٦.٠	١٠٠	١٠٥٩٩١ ٦	المجموع
		١٧١٤.١				الجزر التربيعي
					١٠٢٩.٥	للمجموع
		٤.٣			٢.٦	نصف القطر/سم

شكل (5)

قيم صادرات النفط الكويتي ١٩٧٣ - ١٩٨٣ +



مقياس الرسم : ١ : ٤٠٠
المصدر : مصطفى الشلقاني : الإحصاء للعلوم الاجتماعية والتجارية ١٩٨٩م

ثانيا : العرض البياني للتوزيعات التكرارية

(أ) المدرج التكراري Histogram

يستخدم المدرج التكراري لعرض بيانات الجدول التكراري وهو عبارة عن مستطيلات تمثل التكرارات المختلفة بحيث تتناسب مساحتها مع التكرارات وتتناسب أطوال قواعدها مع أطوال الفئات ، وهناك نوعان من المدرجات التكرارية

المدرج التكراري المنتظم

في حالة الجدول التكراري ذو الفئات المتساوية تكون المستطيلات متساوية في العرض وتصبح النسبة بين مساحاتها كالنسبة بين تكراراتها .

مثال

فيما يلي جدول تكراري يبين أوزان ١٠٠ طالب في إحدى الجامعات:

جدول (١٨)

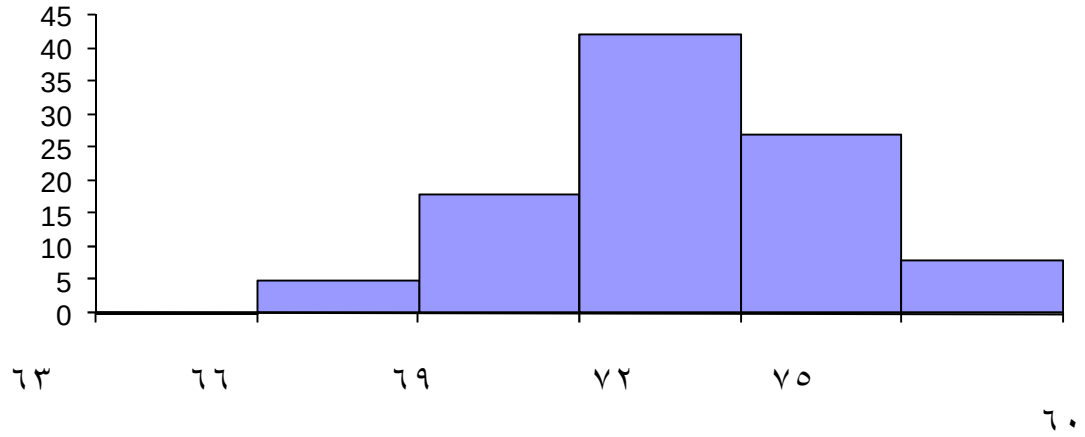
أوزان الطلاب في إحدى الجامعات

عدد الطلاب f	فئات الأوزان
٥	٦٠-
١٨	٦٣-
٤٢	٦٦-
٢٧	٦٩-
٨	٧٢-٧٥
١٠٠	المجموع

المصدر: مواردي دي سبيجل - نظريات وسائل في الإحصاء - ماكروهل ١٩٧٢.

ولرسم المدرجة التكراري المنتظم يتم تقسيم المحور الأفقي إلى (٥) أقسام متساوية تمثل بمقياس رسم مناسب بحيث يسمح بتمثيل أكبر تكرار في الجدول وهو ٤٢ بعد ذلك نرسم خمس مستطيلات قواعدها متساوية وارتفاع كل منها يساوي تكرار الفئة المعنية ويمثل مركز الفئة منتصف قاعدة المستطيل.

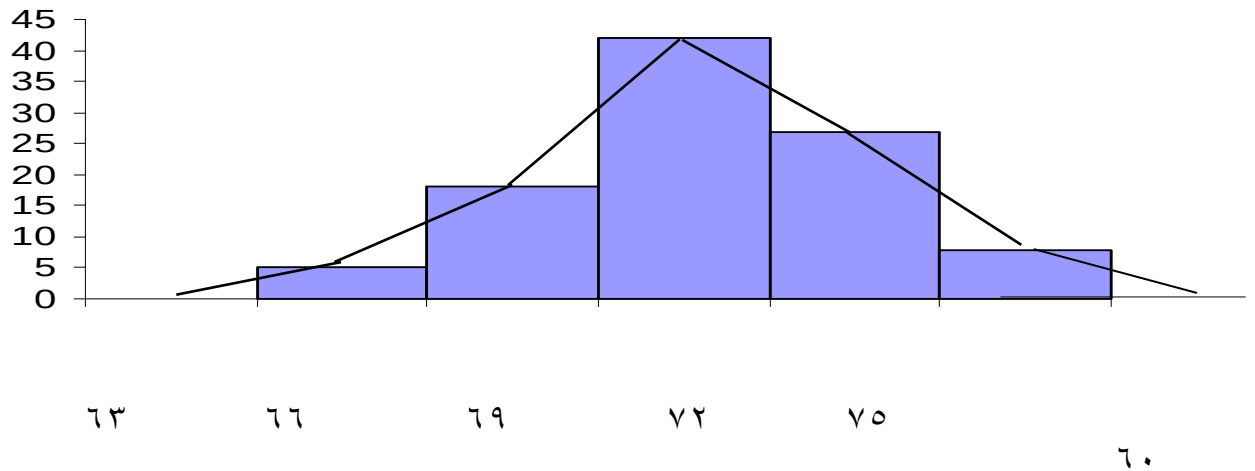
شكل (6) المدرج التكراري المنتظم لأوزان الطلاب في المثال أعلاه.



(ب) المضلع التكراري Frequency Polygon

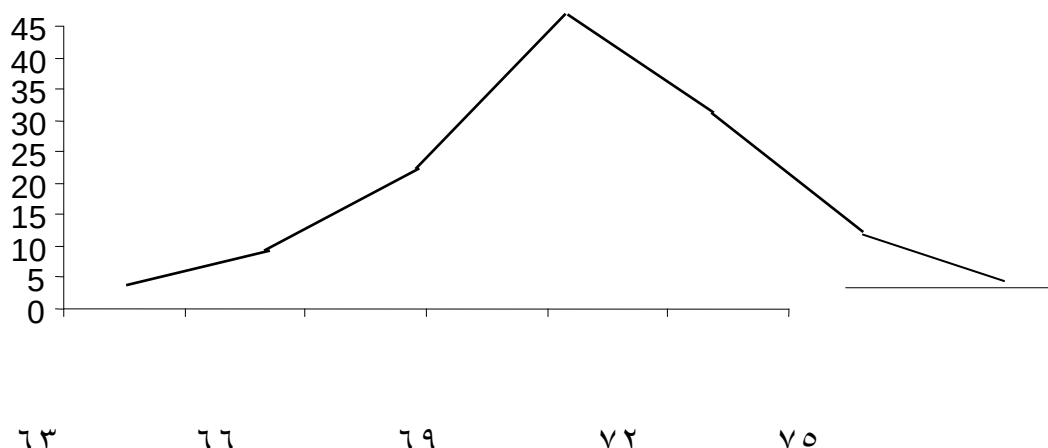
وهو عبارة عن خط متقطع يمر بمنصف المستطيلات (مراكز الفئات) من أعلى بعد إضافة وصلتين إلى ما بعد مركز الفئة الدنيا ومركز الفئة العليا كما في شكل (٧)

شكل رقم (٧) المضلع التكراري لأوزان طلاب إحدى الجامعات



وفي المضلع التكراري فإن مساحة المثلثات الخارجية تساوي مساحة المثلثات الداخلية وعليه فإن مجموع مساحة المستطيلات تحت المدرج التكراري تساوي المساحة المحصورة بين المضلع التكراري والمحور الأفقي ويمكن رسم المضلع التكراري مباشرة دون رسم المدرج التكراري وذلك بتحديد

نقاط ارتفاعات المستطيلات مقابل مراكز الفئات ثم نوصل
النقط بخطوط مستقيمة لنحصل على المضلع التكراري .
شكل (٨) المضلع التكراري لأوزان طلاب إحدى الجامعات



٦٠

(ج) المدرج التكراري غير المنتظم :

تظهر أهمية العلاقة بين مساحات المستطيلات التي تمثل
الفئات وبين ارتفاعات الأعمدة التي تمثل تكرارات الفئات
إذا كانت أطوال الفئات غير متساوية .

ولتحقيق التناسب بين الأعمدة يتوجب تعديل التكرارات
لنحصل على ما يسمى بالتكرار المعدل عن طريق قسمة
التكرار الأصلي للفئة على طول الفئة .

سوف نستخدم بيانات مستويات الدخل الشهري لعينة من الأسر
الواردة في جدول (٩) وقبل رسم المدرج التكراري غير
المنتظم نقوم أولاً بحساب التكرارات المعدلة في جدول (١٩)
التكرارات المعدلة لمستويات الدخل الشهري لعينة من
الأسر (بالجنيهاً السودانية) :

جدول (١٩)

التكرارات المعدلة لمستويات الدخل الشهري لعينة من
الأسر (بالجنيهاً السودانية) :

فئات الدخل	التكرارات f	طول الفئة f	التكرار المعدل fc	× ١٠٠ fc
-600	500	400	1.25	125
-1000	300	200	1.5	150
-1200	400	300	1.33	133
-2000	600	500	1.2	12
-2500	1000	500	2.0	200

180	1.8	1000	1800	-3000
40	0.4	2000	800	-4000

ملحوظة

لقد قمنا بحذف فئات الدخل الثلاثة الأخيرة لتسهيل عملية الرسم كما تم ضرب التكرارات المعدلة ($fc \times 100$) في رقم ثابت (١٠٠) للتخلص من الكسور شكل (٩) يبين المدرج التكراري غير المنتظم للبيانات في جدول (١٩) ويتم الرسم بنفس الطريقة التي تم بها الرسم في حالة المدرج التكراري المنتظم وذلك بعد تعديل التكرارات. والمضلع التكراري في التوزيعات غير المنتظمة يتم رسمه مباشرة عن طريق مراكز الفئات أو من المدرج التكراري ، ولكن في الحالتين يجب أولاً تعديل التكرارات قبل الرسم .

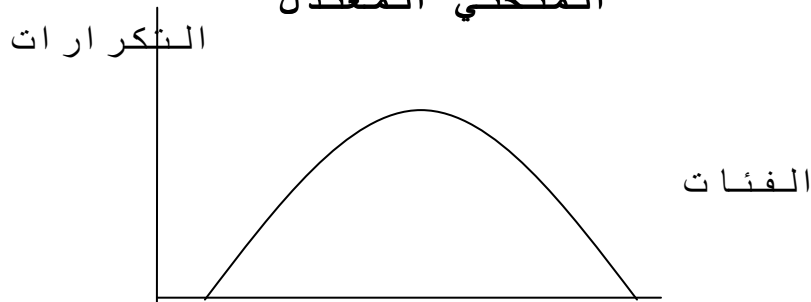
(ب) المنحني التكراري frequency curve

في علم الإحصاء يكون الغرض دائماً من الرسومات والأشكال البيانية هو الوقوف على الاتجاه العام لها ومن ثم إيجاد المعادلة الرياضية التي خضع لها التوزيع التكراري ، ولتكوين المنحني التكراري فإنه يتم تحديد نقاط المضلع التكراري ثم يرسم منحني يمر بهذه النقاط أو معظمها على أن يكون توصيل النقاط في شكل تمهيد باليد أو دقيقاً باستخدام أحد الأساليب الرياضية .

والمنحنيات التكرارية نوعان

المنحنيات المتماثلة في هذا النوع يتماثل المنحني حول محور رأسي يمر بنقطة النهاية العظمى ويقسم المساحة أسفل المنحني إلى قسمين متطابقين كما في شكل (١٠) وأكثر المنحنيات المتماثلة استخداماً هو المنحني المعتدل كما سنرى لاحقاً وهو ناقوس الشكل وله قمة واحدة ويمتد طرفاه إلى ما لا نهاية .

شكل رقم (١٠) المنحني المعتدل

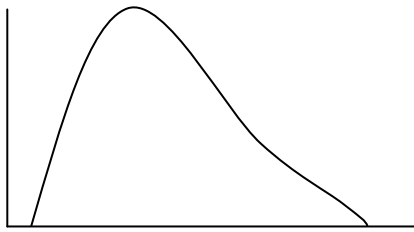


وفي الحياة العملية المنحنيات التكرارية المتماثلة قليلة الحدوث وجميع المشاهدات الإحصائية تؤول إلى شبه التماثل.

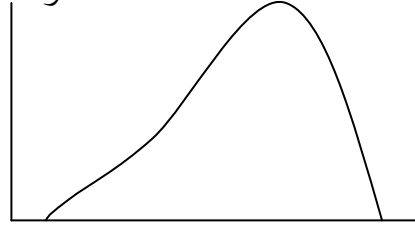
المنحنيات غير المتماثلة

أما المنحنيات التكرارية غير المتماثلة فهي التي كون عدم تماثلها واضحا وهي تسمى منحنيات ملتوية ، والمنحني الملتوي يكون أحد طرفيه أطول من الطرف الآخر ، فإذا زيل المنحني إلى اليسار يكون الالتواء سالبا وإذا كان زيل المنحني إلى اليمين يكون الالتواء موجب . انظر الشكل (١١) و (١٢)

شكل (١٢) ملتو لليمين)



شكل (١١) ملتو لليسار (التواء سالب)
التواء موجب)

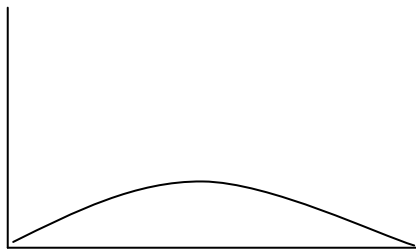


المنحنيات المفرطة Kurtic curves

وهنا يكون الاختلاف من المنحني المعتدل في شكل قمة المنحني ، فالمنحني قد يكون مفرطاً إذا كان للتوزيع التكراري قمة حذباء بحيث أنها أعلى من قمة التوزيع المعتدل . انظر شكل (١٣) وشكل (١٤) .

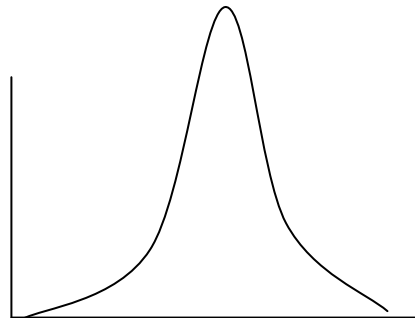
شكل (١٤) منحني

منحني



شكل (١٣) منحني مدبب

منحني مدبب

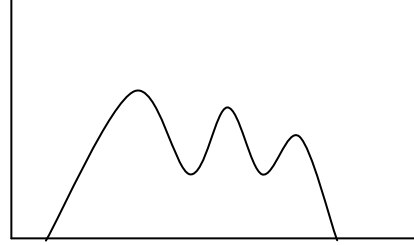


مفرطح

المنحنيات متعددة القمم Multi-modal curve

عندما تكون مفردات المجتمع غير متجانسة ينتج عن ذلك منحنى تكراري متعدد القيم ، وفي الغالب تظهر مثل هذه الحالة إذا كان المجتمع يتكون من عدة مجموعات مختلفة متداخلة انظر الشكل (١٥)

شكل (١٥)
منحنى متعدد القمم



منحنيات أخرى

وهناك أمثلة أخرى للمنحنيات التكرارية مثل المنحنى النوني والمنحنى الأسّي والمنحنى الأسّي المعكوس انظر أشكال (١٦) و (١٧) ، (١٨) .

شكل (١٨)

المنحنى

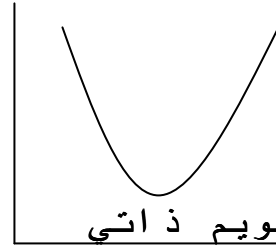
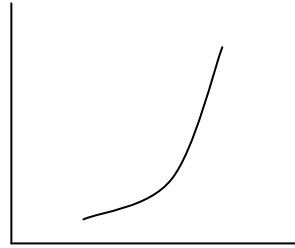
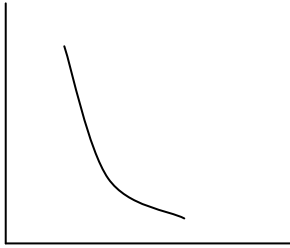
شكل (١٧)

المنحنى الأسّي

الأسّي المعكوس

شكل (١٦)

المنحنى النوني



أسئلة تقويم ذاتي

?

- ١- عرف البيانات الوصفية ؟
- ٢- ما هي أو خطوة في تكوين الجدول التكراري الوصفي ؟
- ٣- ما هي الخطوات التي تتبع في تحديد عدد الفئات ؟
- ٤- أذكر أنواع الجداول التكرارية ؟
- ٥- تنقسم الأشكال والرسومات البيانية إلى قسمين ما هي ؟
- ٦- تحدث باختصار عن العرض البياني للجدول البسيطة ؟
- ٧- تحدث باختصار عن العرض البياني للتوزيعات التكرارية ؟

الخلاصة

عزيزي الدارس،

ماذا تناولنا في هذه الوحدة؟
قبل الإجابة على هذا السؤال تأكد من أنك قد حققت الأهداف التعليمية الواردة بداية هذه الوحدة.

نبدأ بتلخيص هذه الوحدة بتقسيم منهج البحث العلمي، وينقسم عادة إلى قسمين هما المنهج النموذجي والمنهج التطبيقي، ثم ننتقل إلى مصادر جمع البيانات حيث هناك مصدران أساسيان هما المصادر التاريخية والمصادر الميدانية، وضحنا بعد ذلك أسلوبين لجمع البيانات حيث تم جمع البيانات عن طريق أسلوب الحصر الشامل ويعني ذلك التعداد الكامل لجميع مفردات الدراسة، والأسلوب الثاني هو أسلوب المعاينة حيث يتم استخدام عدد محدود من مفردات الإطار ثم تعميم على جميع مفردات المجتمع. وأوضحنا أيضا المرتكزات التي يجب أن يأخذها الباحث في الاعتبار عند إجراء بحث عن طريق المعاينة. كما قسمنا المعاينة إلى مجموعتين متماثلتين هما المعاينة العشوائية والمعاينة العمدية. وتحدثنا عن أخطاء المعاينة الخطاء غير المعاينة.

في هذه الوحدة تناولنا العرض الجدولي والعرض البياني للبيانات وذكرنا أن عرض البيانات يبدأ بتصنيفها حيث يلزم ذلك التفرقة بين البيانات الكمية والبيانات الوصفية. وأخيرا تحدثنا عن طريق العرض البياني للبيانات وقسمنا الأشكال والرسوم البيانية إلى قسمين أولا عرض البيانات للجدول البسيطة وثانيا عرض البيانات للجدول التكرارية

مرحبا بك في نهاية هذه الوحدة أرجو أن تكون استفدت من المعلومات التي وردت فيها وأن تكون خير معين لك في دراسة بقية وحدات هذا المقرر نتمنى لك دواما التوفيق والنجاح.

لمحة مسبقة عن الوحدة التالية

عزيزي الدارس

سنتناول في الوحدة التالية مقاييس النزعة المركزية وهي الوسط الحسابي، الوسيط، المنوال ، الوسط الهندسي والوسط التوافقي.

مسرد المصطلحات

- **المعاينة Sampling** تعني جمع البيانات باستخدام عدد محدود من مفردات الإطار ثم تعميم نتائج الدراسة على جميع مفردات المجتمع.
- **الحصر الشامل Full Count** يعني التعداد الكامل لجميع مفردات المجتمع المراد دراسته.
- **الخطأ العشوائي Random Error** وهو الخطأ الذي ينتج عن طريق الصدفة الناتجة عن طبيعة الاختيار العشوائي ويتوقف هذا الخطأ على طريقة المعاينة وحجم العينة وتباين المجتمع.
- **خطأ التحيز Base Error** وهو الخطأ الذي لا يتوقف على أسلوب الاختيار العشوائي وإنما ينتج للتحيز عن اختيار المفردات.
- **الخط البياني Line Chart** يستخدم لتوضيح نمط اتجاه ظاهرة خلال فترة زمنية محددة ويرسم الخط البياني بوضع نقطة على المحور الأفقي والظاهرة تحت الدراسة على المحور الرأسي.
- **الرسوم الدائرية** هي عبارة عن دائرة نصف قطرها يتم تقسيمها إلى قطاعات تتلاقى في المركز وتتناسب مساحة هذه القطاعات مع القيم الجزئية لإظهار الأهمية النسبية لكل مشاهدة.
- **المدرج التكراري Histogram** وهو عبارة عن مستطيلات تمثل التكرارات المختلفة بحيث تتناسب مساحتها مع التكرارات وتتناسب أطوال قواعدها مع أطوال الفئات.
- **المضلع التكراري Frequency Polygon** هو عبارة عن خط متقطع يمر بمنصف المستطيلات (مراكز الفئات) من أعلى بعد إضافة وصلتين إلى ما بعد مركز الفئة الدنيا ومركز الفئة العليا.

إجابات التدريبات

تدريب (١)

- الحصر الشامل يعني التعداد الكامل لجميع مفردات المجتمع المراد دراسته. ويمكن

استخدامه إذا كانت مفردات المجتمع محدودة أو في الحالات التي تتطلب جمع البيانات عن خصائص المجتمع عن المستوى القومي مثل التعداد العام للسكان.

- أما المعاينة فتعني جمع البيانات باستخدام عدد محدود من مفردات الدراسة ثم نعمم نتائج الدراسة على جميع مفردات المجتمع. تستخدم طريقة المعاينة في الحالات التي يؤدي فيها الشمول الكامل إلى فقدان المجتمع بأكمله. تمتاز طريقة المعاينة عن طريقة الحصر الشامل بالسهولة والبساطة واختصار الوقت.

تدريب (٢)

- في المعاينة الطبقية يتم تقسيم المجتمع حسب الفئات ثم بعد ذلك يتم تحديد حجم العينة بتخصيص نصيب كل صفة وطبقة عن حجم العينة حيث نضمن تمثيل العينة لكل طبقات المجتمع، ويراعى في توزيع العينة بين الطبقات أن تكون الطبقات متباينة فيما بينها متجانسة في داخلها.

- في المعاينة المنتظمة يتم ترتيب مفردات المجتمع بأرقام متسلسلة ثم تحدد فترة المعاينة بقسمة حجم المجتمع الكلي على حجم العينة ثم بعد ذلك تسحب المفردات التي تقع بانتظام على أبعاد فترة المعاينة وتكرارها.

تدريب (٣)

تستخدم المعاينة العنقودية في حالات استطلاع الرأي العام بحيث أنه في المرحلة الأولى للمعاينة يمكن إعادة استخدامها بصورة مستمرة لدراسة التغيرات في الرأي العام.



محتويات الوحدة

الصفحة	الموضوع
٥٧	مقدمة
٥٧	تمهيد
٥٨	أهداف الوحدة
٥٩	١. الوسط الحسابي
٥٩	١.١ تعريف الوسط الحسابي
٥٩	٢.١ طريقة حسابه
٦٥	٣.١ مزايا وعيوب الوسط الحسابي
٦٦	٢. الوسيط
٦٦	١.٢ تعريف الوسيط
٦٦	٢.٢ طريقة حسابه
٧٣	٣.٢ مزايا وعيوب الوسيط
٧٤	٣. المنوال
٧٤	١.٣ تعريف المنوال
٧٤	٢.٣ طريقة حساب المنوال
٧٨	٣.٣ مزايا وعيوب المنوال
٨٠	٤. الوسط الهندسي
٨٢	٥. الوسط التوافقي
٨٣	٦. الأهمية النسبية للمتوسطات المختلفة

الصفحة	الموضوع
٨٥	الخلاصة
٨٥	لمحة مسبقة عن المقرر
٨٦	إجابات التدريبات
٨٧	مسرد المصطلحات

المقدمة

تمهيد

عزيزي الدارس،

مرحباً بك إلي هذه الوحدة وهي الوحدة الثانية في هذا المقرر، وقد انتهينا في الوحدة السابقة من طرق جمع البيانات وعرضها جدولياً. الآن نريد أن نبحث عن مقياس يكون معبراً عن الظاهرة موضوع الدراسة لها. أي نريد الحصول على قيمة واحدة تعبر عن جميع القيم. فمتوسط الدخل - مثلاً - في بلد ما يعبر عن جميع الدخول في هذا البلد، أي يعبر عن المستوى العام للدخل. كذلك فإن متوسط طول الطلاب في الفصل يعبر عن جميع الأطوال، أي عن المستوى العام لطول الطلاب السائد في الفصل. ومن خصائص البيانات - أي بيانات - أن لها (نزعة) أو ميلاً لأن (تتركز) حول قيمة معينة متوسطة. وهذه القيم التي تتركز حولها البيانات تسمى بالمتوسطات أو مقاييس النزعة المركزية. ففي مثال الطول نجد أن معظم أطوال الطلاب في الفصل تدور أو تتركز حول قيمة، هذه القيمة هي متوسط الطول.

وتتحدث هذه الوحدة عن مقاييس النزعة المركزية وهي الوسط الحسابي والوسيط والمنوال والوسط الهندسي والوسط التوافقي. حيث قمنا بتعريف لكل من هذه المقاييس ووضحنا كيفية حساب كل مقياس وأخيراً أوضحنا الأهمية النسبية لهذه المتوسطات أرجو أن تكون وحدة مفيدة وإن تستفيد من المعلومات الواردة فيها مع خالص تمنياتنا بالتوفيق والنجاح.

أهداف الوحدة



عزيزي الدارس بعد فراغك من دراسة هذه الوحدة ينبغي أن تكون قادراً على

ن

١. تعرف الوسط الحسابي والوسيط والمنوال والوسط الهندسي والوسط التوافقي.
٢. تحسب قيمة المتوسطات من البيانات المبوبة وغير المبوبة.
٣. تعدد مزايا وعيوب كل من الوسيط والوسط الحسابي والمنوال.
٤. تذكر الأهمية النسبية للمتوسطات المختلفة.

١ . الوسط الحسابي Arithmetic Mean

١ . ١ تعريف الوسط الحسابي

يعرف الوسط الحسابي لمجموعة من القيم بأنه القيمة التي لو حلت محل قيمة كل مفردة في المجموعة، لكان مجموع هذه القيم الجديدة مساوياً لمجموع القيم الأصلية. وبالتالي فإن الوسط الحسابي يساوي مجموع القيم على عددها.

٢ . ١ طريقة حسابه

أولاً في البيانات غير المبوبة

كما ذكر في التعريف فإن الوسط الحسابي لمجموعة من القيم يتم حسابه بجمع كل القيم ثم القسمة على عددها. أي أن:

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عددها}}$$

مثال (١)

إذا كانت أعمار أربعة أشخاص هي: 8 سنوات، ١٢ سنة، ٩ سنوات، ١١ سنة
أحسب الوسط الحسابي لأعمار هؤلاء الأشخاص.

الحل

بما أن

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عددها}}$$

$$\frac{8 + 12 + 9 + 11}{4} = \frac{40}{4} = 10 \text{ years} = \text{الوسط الحسابي}$$

إذاً

وللانتقال من مجرد مثال محدد إلى الحالة أو القاعدة العامة لحساب الوسط الحسابي، فإنه يمكن التعبير عن ذلك باستخدام الرمز. ففي المثال السابق إذا فرض أن القيم الأربع هي:

$$X_1, X_2, X_3, X_4$$

وإذا رمزنا للوسط الحسابي بالرمز $\bar{X}$ (وتقرأ X -bar) فإن قيمته حسب التعريف هي:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{4} = \text{الوسط الحسابي}$$

كذلك يمكن الانتقال من حالة أربع قيم إلى الحالة العامة، حيث يكون لدينا أي عدد من القيم. فإذا فرض أن عدد القيم يساوي N وأن القيم هي:

$$X_1, X_2, X_3, \dots, X_N$$

فإن الوسط الحسابي هو:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$$

ويلاحظ أن كتابة الوسط الحسابي بالشكل السابق يأخذ وقتاً وحيزاً أكبر، بينما الاتجاه دائماً إلى الاختصار في التعبير. لذا فإننا سوف نستخدم رمزاً ليعبر عن مجموع القيم، هذا الرمز هو $\sum X$ (ويقراً $\sum X$). وبذلك فإن الوسط الحسابي يأخذ الصورة المبسطة التالية:

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$$

حيث $\bar{X}$ تعبر عن الوسط الحسابي، $\sum_{i=1}^n X_i$ تعبر عن مجموع القيم، N هي عدد القيم.

نشاط

إذا كان الوسط الحسابي للدخل الشهري لخمسـة أشخاص ٤٠٠٠ دينار
فما هو مجموع دخولهم؟



مثال (٢)

البيانات التالية تمثل درجات بعض الطلاب في إحدى المواد ٨١، ٧٤، ٩٢، ٦٧، ٨٦، أحسب الوسط الحسابي لدرجات الطلاب في هذه المادة.

الحل

في هذه الحالة فإن (عدد الطلاب) هو: $N = 5$

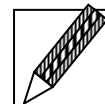
والقيم هي: $X_1 = 81, X_2 = 74, X_3 = 92, X_4 = 67, X_5 = 86$

وبالتالي فإن الوسط الحسابي هو:

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} = \frac{81 + 74 + 92 + 67 + 86}{5} = \frac{400}{5} = 80$$

تدريب (١)

إذا كانت أوزان خمسة أشخاص هي: ٦٠، ٧٠، ٣٠، ٤٠، أحسب الوسط الحسابي لأوزان هؤلاء الأشخاص.



ثانياً في البيانات المبوبة

في المثال رقم (١) كان لدينا شخص واحد عمره ٨ سنوات، وشخص واحد عمره ١٢ سنة، وآخر ٩ سنوات، ورابع ١١ سنة. أي أن كل عمر من هذه الأعمار تكرر مرة واحدة. ولكن ماذا يكون الوضع لو كان هناك أكثر من شخص واحد أعمارهم ٨ سنوات، وأكثر من شخص واحد أعمارهم ١٢ سنة... وهكذا؟

في هذه الحالة نأخذ المثال التالي:

مثال (٣)

افترض أن هناك ٣ أشخاص عمر كل منهم ٨ سنوات، ٥ أشخاص عمر كل منهم ٩ سنوات، ٤ عمر كل منهم ١١ سنة، ٢ عمر كل منهما ١٢ سنة، كما في الجدول التالي:

جدول رقم (١)

العمر	عدد الأشخاص
٨	٣
٩	٥
١١	٤
١٢	٢

أحسب الوسط الحسابي للعمر .

الحل

لحساب الوسط الحسابي للعمر فإنه يجب أولاً حساب مجموع الأعمار، ثم نقسم هذا المجموع على عدد الأشخاص (والذين يبلغ عددهم في هذا المثال $3 + 5 + 4 + 2 = 14$ شخصاً).
ويلاحظ أنه للحصول على مجموع الأعمار يجب ضرب كل عمر من الأعمار المختلفة في عدد الأشخاص المقابل له في الجدول. فالأشخاص الثلاثة الذين يبلغ عمر كل منهم ٨ سنوات، يكون مجموع أعمارهم يساوي $3 \times 8 = 24$ سنة، والخمسة أشخاص التاليين لهم أعمارهم يساوي $5 \times 9 = 45$ سنة،... وهكذا فإننا نحصل على مجموع الأعمار بضرب القيم في خانة العمر في الأرقام المقابلة لها في خانة التكرار (عدد الأشخاص) ثم نجمع حواصل الضرب.
فإذا رمزنا للعمر بالرمز X ، ولعدد الأشخاص أو التكرار بالرمز f ، فإن خطوات الحل يمكن تبسيطها كما هو مبين في الجدول التالي:

جدول رقم (٢)

العمر (X)	التكرار (f)	العمر × التكرار (f X)
٨	٣	$3 \times 8 = 24$
٩	٥	$5 \times 9 = 45$
١١	٤	$4 \times 11 = 44$
١٢	٢	$2 \times 12 = 24$
المجموع	$\sum f = 14$	$\sum fX = 137$

ومن الجدول السابق يتضح أن:

عدد الأشخاص = $\sum f = 14$ شخصاً ، ومجموع الأعمار = $\sum fX = 137$ سنة.
وبالتالي فإن الوسط الحسابي هو:

$$\bar{X} = \frac{137}{14} = 9.786 \text{ years}$$

أي أن معادلة الوسط الحسابي تصبح في هذه الحالة:

$$\bar{X} = \frac{\sum fX}{\sum f}$$

حيث $\sum fX$ هي مجموع القيم. ونحصل عليهما بضرب القيم في التكرارات ثم أخذ المجموع.
؛ $\sum f$ هي عدد القيم (أو مجموع التكرارات).

وبالتالي فإن المعنى لم يتغير، وهو أن الوسط الحسابي يساوي مجموع القيم على عددها. ولنتقدم خطوة أخرى، ونأخذ حالة الجداول التكرارية ذات الفئات.

مثال (٤)

الجدول التالي يمثل توزيع مائة شخص حسب فئات الوزن (بالكيلوجرام).

جدول رقم (٣)

فئات الوزن	عدد الأشخاص
٣٠ -	9
٤٠ -	15
٥٠ -	22
٦٠ -	25
٧٠ -	18

80-90	11
المجموع	١٠٠

والمطلوب حساب الوسط الحسابي للوزن.

الحل

يختلف هذا المثال عن المثال السابق رقم [٣] في أن هذا المثال يحتوي على فئات للوزن بدلاً عن قيم محددة. وكما نعلم فإن القاعدة في هذه الحالة هي افتراض أن كل المفردات التي تقع في فئة واحدة متساوية في القيمة، وأن قيمتها تساوي مركز الفئة التي تقع بها.

فمثلاً في الفئة الأولى: يفترض أن وزن كل شخص من الأشخاص التسعة يساوي ٣٥ كجم (مركز الفئة : $\frac{30 + 40}{2} = 35$).

وفي الفئة الثانية : نفترض أن وزن كل شخص من الأشخاص الـ ١٥ يساوي ٤٥ كجم (مركز الفئة : $\frac{40 + 50}{2} = 45$.. وهكذا..

وبالتالي فإن خطوات الحل تكون كما يلي:

١. حساب مراكز الفئات. وسوف نرسم لها بالرمز X .

٢. نضرب مركز كل فئة X في تكرار هذه الفئة f ، أي نكون عموداً لحاصل الضرب $f X$.

كما في الجدول التالي

جدول رقم (٤)

$f X$	مراكز الفئات (X)	التكرار (f)	فئات الوزن
315	٣٥	٩	٣٠ -
675	٤٥	١٥	٤٠ -
1210	٥٥	٢٢	٥٠ -
1625	٦٥	٢٥	٦٠ -
1350	٧٥	١٨	٧٠ -
935	٨٥	١١	٨٠ - ٩٠

المجموع	١٠٠		٦١٤٠
---------	-----	--	------

٣. نوجد قيمة الوسط الحسابي من العلاقة:

$$\bar{X} = \frac{\sum fX}{\sum f} = \frac{6410}{100} = 64.1 \text{ kgs}$$

٣. ١ مزايا وعيوب الوسط الحسابي

المزايا

١. يمتاز بالسهولة والبساطة في العمليات الحسابية.
٢. تدخل جميع القيم في حسابه.

العيوب

١. يتأثر بالقيم الشاذة أو المتطرفة. والقيمة الشاذة هي التي تكون إما كبيرة جداً أو صغيرة جداً بالنسبة لمعظم القيم وبالتالي فهي ترفع قيمة الوسط أو تخفض قيمته عن معظم القيم.
٢. لا يمكن إيجاده بيانياً (عن طريق الرسم).
٣. لا يمكن إيجاده من القيم الوصفية.

أسئلة تقويم ذاتي

١. عرف الوسط الحسابي؟
٢. كيف يتم حساب الوسط لمجموعة القيم؟
٣. عدد مزايا وعيوب الوسط الحسابي؟



٢ . الوسيط The Median

٢ . ١ تعريف الوسيط

يعرف الوسيط لمجموعة من القيم بأنه القيمة التي تتوسط المجموعة بعد ترتيبها تصاعدياً أو تنازلياً.

يرمز للوسيط بالرمز $\tilde{X}$ ، وتقرأ (X- wave) .

٢ . ٢ طريقة حساب الوسيط

أولاً من البيانات غير المبوبة

استنتاجاً من التعريف السابق للوسيط فإن أول ما يجب عمله لحساب قيمة الوسيط هو ترتيب القيم ترتيباً تصاعدياً أو تنازلياً، ثم أخذ القيمة التي تقع في المنتصف لتكون هي الوسيط. هذا بفرض أن عدد القيم فردي. أما إذا كان عدد القيم زوجياً فنأخذ القيمتين اللتين في المنتصف ويكون الوسيط هو مجموع القيمتين مقسوماً على ٢.

مثال (٥)

أحسب الوسيط للأوزان التالية لبعض المرضى: ٥٢ ٥٨ ٥٠ ٦٣ ٥٥

الحل

نرتب أولاً القيم تصاعدياً أو تنازلياً.

تصاعدياً 50, 52 , 55, 58, 63

نلاحظ أن القيمة التي في المنتصف هي الثالثة في الترتيب وأن قيمتها تساوي ٥٥ كيلوجرام.

وبهذا فإن $\tilde{X} = 55 \text{ kg}$.

وسوف نصل لنفس النتيجة إذا كان الترتيب تنازلياً.

ومن المثال السابق يلاحظ أن هناك فرقاً بين ترتيب الوسيط وقيمة الوسيط. ففي المثال نجد أن: ترتيب

الوسيط = ٣ ، بينما قيمة الوسيط = ٥٥ كيلوجرام.

وعموماً فإن ترتيب الوسيط هو الذي يحدد لنا قيمته. ويمكن حساب ترتيب الوسيط من العلاقة:

$$\text{ترتيب الوسيط} = \frac{\text{عدد القيم} + 1}{2}$$

أو:

$$\text{ترتيب الوسيط} = \frac{n+1}{2}$$

ففي المثال السابق كان $n=5$ ∴ ترتيب الوسيط $= \frac{5+1}{2} = \frac{6}{2} = 3$
ثم نبحت في القيم بعد ترتيبها تصاعدياً (أو تنازلياً) عن القيمة التي ترتيبها ٣، فتكون هي قيمة الوسيط (تساوي ٥٥ كجم في المثال).

مثال (٦)

أحسب الوسيط للأوزان التالية لبعض المرضى: ٥٨ ٥٢ ٥٠ ٦٣ ٥٧ ٥٥

الحل

١. نرتب القيم أولاً ترتيباً تصاعدياً كما يلي: 50, 52, 55, 57, 58, 63
٢- نحسب ترتيب الوسيط. وهنا نلاحظ أن هناك قيمتين في المنتصف. ويكون ترتيب الأولى هو $\frac{n}{2}$ ،
والثانية $\frac{n}{2} + 1$. أي أن:

$$\text{ترتيب القيمة الأولى} = \frac{n+1}{2} , \text{ ترتيب القيمة الثانية} = \frac{n}{2} + 1$$

ومن المثال:

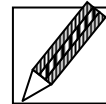
ترتيب القيمة الأولى $= \frac{6}{2} = 3$ ، وترتيب القيمة الثانية $= 3 + 1 = 4 = \frac{6}{2} + 1$
أي أن قيمة الوسيط بين القيمتين الثالثة والرابعة. وبالتالي فإن الوسيطان هما ٥٥ ، ٥٧ . ويكون الوسيط هو:

$$\tilde{X} = \frac{\tilde{X}_1 + \tilde{X}_2}{2} = \frac{55 + 57}{2} = 56 \text{ kg}$$

تدريب (٢)

أحسب الوسيط للأطوال التالية لبعض الطلاب:

١٨٠، ١٦٩، ١٥٤، ١٦٥، ١٨٧



ثانياً من البيانات المبوبة ذات الفئات

يمكن حساب الوسيط في حالة البيانات المبوبة ذات الفئات بطريقتين إما حسابياً أو بيانياً (عن طريق الرسم) كما يلي:

أ (حسابياً)

وتكون خطوات الحل بالطريقة الحسابية كما يلي:

١. تكوين الجدول المتجمع الصاعد من الجدول التكراري البسيط (هذه الخطوة تناظر ترتيب القيم تصاعدياً).

٢. حساب ترتيب الوسيط. وهو في هذه الحالة:

$$\text{ترتيب الوسيط} = \frac{\text{مجموع التكرارات}}{2}$$

$$\text{أي أن ترتيب الوسيط} = \frac{N}{2} = \frac{\sum f}{2}$$

٣. تحديد الفئة التي تحتوي على الوسيط من الجدول التكراري المتجمع الصاعد والتي تسمى الفئة الوسيطة، ثم حساب قيمة الوسيط من العلاقة التالية:

$$\tilde{X} = L_1 + \left(\frac{N/2 - \sum f_1}{f_m} \right) C$$

حيث:

L_1 : لحد الأدنى للفئة الوسيطة.

$\sum f_1$: هي التكرار المتجمع الصاعد المقابل للحد الأعلى للفئة قبل الفئة الوسيطة .

C : طول الفئة الوسيطة.

f_m : التكرار المناظر للفئة الوسيطة.

أي أن موقع الوسيط في الفئة الوسيطة يبعد عن بداية الفئة بمقدار يتناسب مع نسبة التكرارات المتبقية لإكمال رتبة الوسيط و تكرار الفئة الوسيطة.

ومن المثال التالي تتضح خطوات الحل السابقة عملياً.

مثال (٧)

الجدول التالي يوضح أوزان ١٠٠ شخص بالكيلوجرام. أحسب الوسيط للأوزان.

جدول رقم (٥)

f	فئات الوزن C
٩	٣٠ -
١٥	٤٠ -
٢٢	٥٠ -
٢٥	٦٠ -
١٨	٧٠ -
١١	٨٠ - ٩٠
١٠٠	المجموع

الحل

تكوين الجدول المتجمع الصاعد كما يلي

جدول رقم (٦)

الحدود العليا للفئات	متجمع صاعد
أقل من 30	٠
أقل من ٤٠	٩

أقل من ٥٠	٢٤
أقل من ٦٠	٤٦
أقل من ٧٠	٧١
أقل من ٨٠	٨٩
أقل من ٩٠	١٠٠

ترتيب الوسيط = ٥٠
 { الفئة الوسيطة

$$\text{ترتيب الوسيط: } \frac{\sum f}{2} = \frac{N}{2} = \frac{100}{2} = 50$$

ثم نبحث أين يقع ترتيب الوسيط بين التكرارات المتجمعة الصاعدة. وفي هذا المثال نجد أن ترتيب الوسيط يساوي ٥٠ وبالتالي فهو يقع بين ٤٦ وهي تمثل هنا $\sum f_1$ ، و ٧١ والتي تسمى التكرار المتجمع الصاعد اللاحق لترتيب الوسيط. ومنهما الاثنان فإن الفئة الوسيطة هي من ٦٠ وأقل من ٧٠، وتكرارها هو ٢٥ (لاحظ أن تكرار الفئة الوسيطة يساوي أيضاً الفرق بين التكرارين المتجمعين السابق واللاحق). وبهذا فإن الحد الأدنى للفئة الوسيطة يساوي ٦٠ وطولها يساوي ١٠. والآن يمكن التعويض مباشرة في المعادلة كما يلي:

$$\begin{aligned}\tilde{X} &= L_1 + \left(\frac{N/2 - \sum f_1}{f_m} \right) C \\ \tilde{X} &= 60 + \left(\frac{50 - 46}{25} \right) \times 10 \\ &= 60 + \frac{4}{25} \times 10 \Rightarrow 60 + 1.6 = 61.6 \text{ kg}\end{aligned}$$

نشاط



أحسب الوسيط من بيانات الجدول التالي:

الحدود العليا للفئات	التكرار المتجمع الصاعد
أقل من ١٠	٦
أقل من ٢٠	١٦
أقل من ٣٠	٣١
أقل من ٤٠	٤٨
أقل من ٥٠	٦٠
أقل من ٦٠	٦٨

وإذا علمت أن أقل قيمة للبيانات هي صفر، فكون الجدول التكراري الأصلي لهذه البيانات.

ب) بيانياً

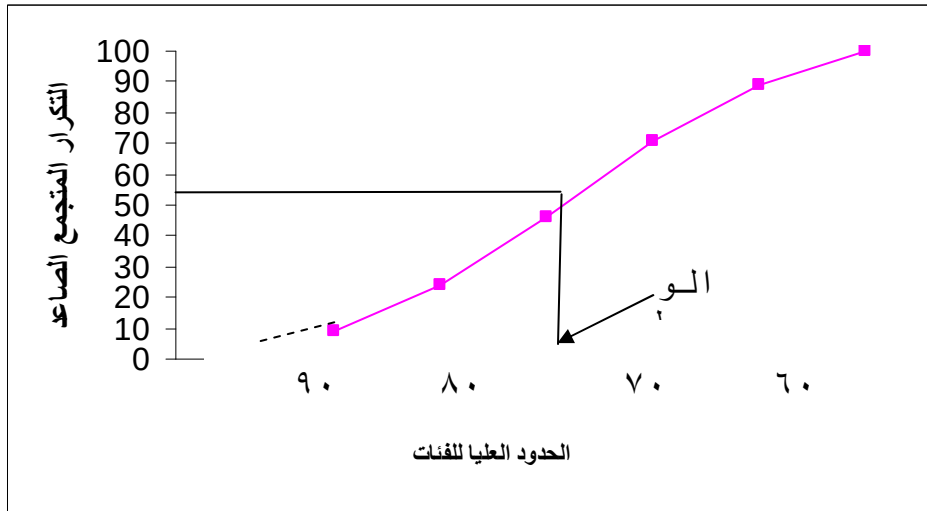
وتكون خطوات استخراج الوسيط بيانياً كما يلي:

- ١- تكوين الجدول المتجمع الصاعد (كما سبق).
- ٢- رسم المنحنى المتجمع الصاعد.
- ٣- حساب ترتيب الوسيط (يساوي مجموع التكرارات مقسوماً على ٢).
- ٤- نحدد ترتيب الوسيط على المحور الرأسي. ومن هذه النقطة نرسم خطاً موازياً للمحور الأفقي حتى يلتقي مع المنحنى المتجمع الصاعد. ومن نقطة الالتقاء هذه ننزل عموداً رأسياً على المحور الأفقي. وتكون نقطة تقاطع العمود مع المحور الأفقي هي قيمة الوسيط.

ونلاحظ هنا أن دقة النتيجة تتوقف إلى حد كبير على دقة الرسم.

والشكل التالي يبين خطوات العمل السابقة لإيجاد الوسيط.

شكل (١) إيجاد الوسيط بيانياً من بيانات مثال (٧)



ومن الرسم يتضح أن قيمة الوسيط تساوي ٦١.٥ كجم (أو بالضبط ٦١.٦ كجم).

٣.٢ مزايا وعيوب الوسيط

المزايا:

١. لا يتأثر بالقيم الشاذة أو المتطرفة.
٢. يمكن إيجاده بيانياً (برسم المنحنى المتجمع الصاعد).

العيوب:

١. لا تدخل جميع القيم في حسابه، إذ أنه يعتمد على قيمة واحدة فقط (أو قيمتين على الأكثر).
٢. في حالة البيانات المبوبة ذات الفئات تستخدم طرق تقريبية في حسابه.
٣. لا يمكن إيجاده من القيم الوصفية.

نشاط

أحسب الوسيط لبيانات الجدول التالي (حسابياً وبيانياً):

فئات العمر	-٥	-١٠	-١٥	-٢٠	-٢٥	-٣٠	٣٥-٤٠	المجموع
التكرار	١٨	٢٠	١٢	١٥	١٤	١١	١٠	١٠٠

(تنبيه: لاحظ أن ترتيب الوسيط هو واحد من التكرارات المتجمعة الصاعدة. وفي هذه الحالة لا يستخدم القانون، بل نحصل على قيمة الوسيط مباشرة من الجدول المتجمع الصاعد).



أسئلة تقويم ذاتي



- ١- عرف الوسيط؟
- ٢- كيف يتم حساب الوسيط في حالة البيانات المبوبة ذات الفئات؟
- ٣- عدد خطوات استخراج الوسيط بيانياً؟
- ٤- عدد مزايا وعيوب الوسيط؟

٣ . المنوال *The Mode*

١.٣ تعريف المنوال

بأنه القيمة الأكثر شيوعاً أو n يعرف المنوال لمجموعة من القيم عددها تكراراً. وهو ربما لا يوجد، وإذا وجد ربما لا يكون فريداً

يرمز للمنوال بالرمز $\hat{X}$ ، وتقرأ ($\hat{X}$) .

مثال (٨)

إذا كانت القيم التالية تمثل أوزان مجموعة من المرضى بأحد المستشفيات (بالكيلوجرام):

٥٠ ، ٥٥ ، ٥٣ ، ٥٥ ، ٦٢ ، ٥٧ ، ٥٥ ، ٥٩

فما هو المنوال لهذه الأوزان ؟.

الحل

من التعريف - مباشرة- نجد أن المنوال = ٥٥ كجم.

وذلك لأن القيمة ٥٥ تكررت أكثر من غيرها.

٣ . ٢ طريقة حساب المنوال

طريقة حساب المنوال في البيانات المبوبة ذات الفئات:

لاحظنا في المثال رقم [٨] أن حساب المنوال في حالة البيانات غير المبوبة يتم مباشرة وحسب التعريف بمجرد النظر إلى البيانات. وكذلك بنفس الطريقة في حالة البيانات المبوبة بدون فئات يكون المنوال هو القيمة المناظرة لأكبر تكرار. أما في حالة البيانات المبوبة ذات الفئات فإن المنوال يقع في الفئة التي تتناظر أعلى تكرار وتسمى الفئة المنوالية. وفيما يلي نتناول طريقتين لحساب المنوال، إحداها حسابية والأخرى بيانية.

١/ إيجاد المنوال حسابياً :

يمكن الحصول علي المنوال حسابياً بأكثر من طريقة. وسوف نحسب المنوال وفقاً للقانون التالي:

$$\hat{X} = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times C$$

حيث :

L_1 هو الحد الأدنى للفئة المنوالية، وهي الفئة التي يقابلها أكبر تكرار في الجدول.

Δ_1 هو الفرق بين التكرار المنوالي والتكرار السابق.

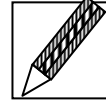
Δ_2 ، هو الفرق بين التكرار المنوالي والتكرار اللاحق.

C طول الفئة المنوالية.

والتكرار المنوالي هو أكبر تكرار في الجدول.

ولذلك فإنه وقبل استخدام القانون السابق لاستخراج المنوال فإنه يجب أولاً تحديد التكرار المنوالي (أكبر تكرار في الجدول) والتكرار السابق له واللاحق أيضاً. وتكون الفئة التي تقابل أكبر تكرار هي الفئة المنوالية.

تدريب (٣)



إذا كانت درجات الطلاب في مادة الإحصاء هي:

٩٤،٨٥،٨٩،٧٥،٨٥،٨٥،٨٠

فما هو المنوال لهذه الدرجات؟

مثال (٩)

في المثال رقم (٧) الخاص بالأوزان أحسب منوال الوزن.

الحل

جدول (٧)

عدد الأشخاص	فئات الوزن
9	٣٠ - ٤٠
15	٤٠ - ٥٠
22	٥٠ - ٦٠
25	٦٠ - ٧٠
18	٧٠ - ٨٠
11	٨٠ - ٩٠
100	المجموع

التكرار

التكرار

التكرار

الفئة المنوالية

من الجدول فإن: $\Delta_1 = 25 - 22 = 3$ ، $\Delta_2 = 25 - 18 = 7$

الفئة المنوالية هي: ٦٠-٧٠ ويكون حدها الأدنى هو ٦٠.

$$C = 70 - 60 = 10$$

$$\hat{X} = L_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times C = 60 + \frac{3}{3+7} \times 10 = 60 + \frac{3}{10} \times 10 = 60 + 3 = 63 \text{ kg}$$

إذاً:

نشاط



الجدول التالي يمثل دخول مجموعة من العمال في أحد المصانع في اليوم بالدينارات (مقربة إلى أقرب ٥ دينارات) = .

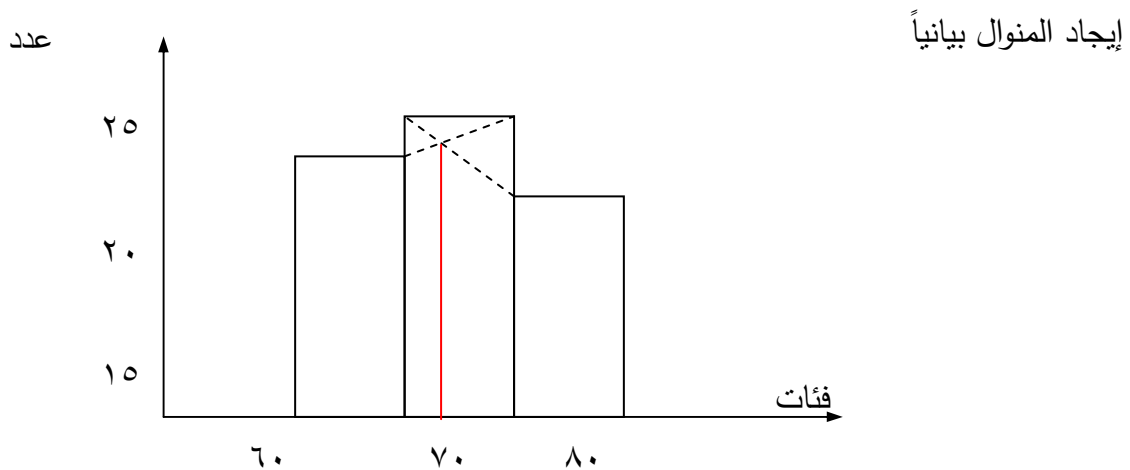
الدخل	٥٠	٥٥	٦٠	٦٥	٧٠	٧٥	٨٠	٨٥	٩٠	٩٥	١٠٠
التكرار	٤	٧	١٠	١٢	٩	٨	٦	٥	٣	٢	١

والمطلوب:

٢ / إيجاد المنوال بيانياً

يتم إيجاد المنوال بيانياً عن طريق رسم المدرج التكراري، وغالباً ما يكتفى برسم ثلاثة مستطيلات فقط والتي تمثل الفئة المنوالية والفئة السابقة وكذلك اللاحقة، ثم يتم استخراج المنوال من داخل الفئة المنوالية بالطريقة الآتية:

شكل رقم (٢)



أي أنه يتم توصيل الرأس الأيمن العلوي لمستطيل الفئة المنوالية بالرأس الأيمن العلوي لمستطيل الفئة السابقة، وتوصيل الرأس الأيسر العلوي لمستطيل الفئة المنوالية بالرأس الأيسر العلوي لمستطيل الفئة اللاحقة. ثم من نقطة التقاطع نسقط عموداً على المحور الأفقي فيقابلة في نقطة تكون هي المنوال. ويتضح من الرسم أن قيمة المنوال تساوي ٦٣ كجم تقريباً.

نشاط



الجدول التالي يبين توزيع مجموعة من الطلاب حسب أطوالهم
(بالسنتيمتر)

فئات الطول	-١٥٠	-١٥٤	-١٥٨	-١٦٢	-١٦٦	-١٧٠	-١٧٤
عدد الطلاب	١٢	١٩	٣٥	٤٢	٣٧	٢٢	١٥

والمطلوب:

٣.٣ مزايا وعيوب المنوال

المزايا

١. بسيط من حيث الفكرة أو طريقة إيجاده.
٢. لا يتأثر بالقيم الشاذة أو المتطرفة.
٣. يمكن إيجاده من القيم الوصفية.

العيوب

١. رغم تعدد طرق حسابه إلا أنها كلها طرق تقريبية لاسيما في حالة البيانات المبوبة ذات الفئات.
٢. في بعض الحالات - وحسب التعريف - لا يمكن إيجاده؛ أي لا يكون هناك منوال للقيم (إذا لم تكن هناك قيمة متكررة أكثر من غيرها). وفي بعض الحالات الأخرى يوجد أكثر من منوال (كما في حالة تكرار بعض القيم بنفس الدرجة وأكثر من باقي القيم).

نشاط



إذا كان أحد المصانع يتكون من ثلاثة أقسام، وكان توزيع العمال حسب أعمارهم في الأقسام الثلاثة كما يلي:

القسم الأول		القسم الثاني		القسم الثالث	
العمر	عدد العمال	العمر	عدد العمال	العمر	عدد العمال
٢٥	٧	٣٠	٨	٢٣	٤
٣٨	١٠	٣٥	١٢	٣٢	١١
٤٠	٥	٤٢	٦	٤٥	٥

المطلوب

١ / حساب متوسط العمر في كل قسم.

٢ / حساب منوال العمر في كل قسم .

أسئلة تقويم ذاتي



١ / عرف المنوال؟

٢ / أشرح باختصار طريقة حساب المنوال في البيانات المبوبة ذات

الفئات؟

٣ / عدد مزايا وعيوب المنوال؟

٤. الوسط الهندسي Geometric Mean

الوسط الهندسي لمجموعة من القيم عددها N هو الجذر النوني لحاصل ضرب هذه القيم. فإذا كان لدينا القيم $X_1, X_2, X_3, \dots, X_N$ فإن الوسط الهندسي لها- ويرمز له بالرمز G هو:

$$G = \sqrt[N]{(X_1)(X_2)(X_3) \dots (X_N)}$$

مثال (١٠)

الوسط الهندسي للأرقام ٢، ٤، ٨ هو:

$$G = \sqrt[3]{(2)(4)(8)} = \sqrt[3]{64} = 4$$

في حالة المجتمعات الكبيرة حيث تكثر المفردات يكون من الأفضل استعمال اللوغاريتمات لحساب الوسط الهندسي، وذلك بالعلاقة التالية:

$$\text{Log } G = \frac{1}{N} \{ \text{Log } X_1 + \text{Log } X_2 + \text{Log } X_3 + \dots + \text{Log } X_N \}$$

وعن طريق جدول اللوغاريتمات يمكن إيجاد العدد المقابل للوسط الحسابي للوغاريتمات الأعداد ويكون هذا العدد هو الوسط الهندسي.

مثال (١١)

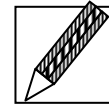
الوسط الهندسي للأرقام ٣، ٥، ٦، ٧، ١٠، ١٢ هو:

$$\begin{aligned} \text{Log } G &= \frac{1}{6} \{ \text{Log } 3 + \text{Log } 5 + \text{Log } 6 + \text{Log } 7 + \text{Log } 10 + \text{Log } 12 \} \\ &= \frac{1}{6} [0.4771 + 0.6990 + 0.7782 + 0.8451 + 1.0000 + 1.0792] = 0.0808 \end{aligned}$$

والعدد المقابل لهذا اللوغاريتم يساوي ٦.٤٣ وهو قيمة الوسط الهندسي للقيم السابقة.

تدريب (٤)

أوجد الوسط الهندسي للأرقام ١، ٣، ٩



الوسط الهندسي للبيانات المبوبة

إذا كانت لدينا التكرارات $f_1, f_2, f_3, \dots, f_N$ هي مراكز الفئات المقابلة لها فإن الوسط الهندسي من الجدول التكراري يحسب بالقانون التالي:

$$\text{Log } G = \frac{1}{N} \sum_{i=1}^N f_i \text{Log } X_i$$

أي أننا نحتاج لتكوين عمود للوغريتمات مراكز الفئات $\text{Log } X_i$ ، ثم نقوم بتكوين عمود لحواصل الضرب $f_i \text{Log } X_i$ ثم نطبق القانون.

مثال (١٢)

من الجدول التالي جد الوسط الهندسي.

C	4-	8-	12-	16-	20-	sum
f	3	5	9	5	3	25

الحل:

نكون عمود لكل من مراكز الفئات، و لوغريثماتها، و حواصل الضرب.

جدول رقم (٨)

C	f_i	X_i	Log X_i	f_i Log X_i
4-	3	6	Log 6 = 0.7782	2.3346
8-	5	10	Log 10 = 1.0000	5.0000
12-	9	14	Log 14 = 1.1461	10.3149
16-	5	18	Log 18 = 1.2553	6.2765
20-	3	22	Log 22 = 1.3424	4.0272
sum	25			27.9532

وبهذا فإن:

$$\text{Log } G = \frac{1}{N} \sum_{i=1}^N f_i \text{Log } X_i$$

$$\text{Log } G = \frac{1}{25} (27.9532) = 1.1183$$

والعدد المقابل لهذا اللوغريتم يساوي 13.13 وهو قيمة الوسط الهندسي للقيم السابقة.

نشاط

من الجدول التكراري التالي جد الوسط الهندسي

١٧٤-	١٧٠	١٦٦	١٦٢	١٥٨	١٥٤	١٥٠	فئات الطول
١٧٨	-	-	-	-	-	-	



٥ . الوسط التوافقي Harmonic Mean

ويستخدم في قياس النزعة المركزية لمجتمع تتزايد مفرداته تزايداً غير مضطرب كالأسعار والسرعة والمعدلات بصفة عامة. الوسط التوافقي يرمز له بالرمز H .

فإذا كان لدينا القيم $X_1, X_2, X_3, \dots, X_N$ فإن الوسط التوافقي لها هو:

$$H = \frac{N}{\sum \frac{1}{X}}$$

مثال (١٣)

الوسط التوافقي للمفردات ٢، ٣، ٦ هو:

$$H = \frac{N}{\sum_{i=1}^N \frac{1}{X_i}} = \frac{3}{\left[\frac{1}{2} + \frac{1}{3} + \frac{1}{6}\right]} = \frac{3}{\left[\frac{3+2+1}{6}\right]} = \frac{3(6)}{6} = 3$$

الوسط التوافقي للبيانات المبوبة:

إذا كانت لدينا التكرارات $f_1, f_2, f_3, \dots, f_N$ هي مراكز الفئات المقابلة لها فإن الوسط التوافقي من الجدول التكراري يحسب بالقانون التالي:

$$H = \frac{N}{\sum_{i=1}^N \frac{f_i}{X_i}}$$

أي أننا نحتاج لتكوين عمود لحاصل قسمة كل تكرار f_i على مركز الفئة المقابل له X_i ثم نجمع

ونطبق القانون.

مثال (١٤)

من الجدول التالي جد الوسط التوافقي.

C	4-	8-	12-	16-	20-	sum
f	3	5	9	5	3	25

الحل

نكون عمود لكل من مراكز الفئات، و حاصل قسمة كل تكرار f_i على مركز الفئة المقابل له X_i

جدول رقم (٩)

C	f_i	X_i	f_i / X_i
4-	3	6	0.5000
8-	5	10	0.5000
12-	9	14	0.6429
16-	5	18	0.2778
20-	3	22	0.1364
sum	25		2.0571

وبهذا فإن:

$$H = \frac{N}{\sum_{i=1}^N \frac{f_i}{X_i}} = \frac{25}{2.0571} = 12.15$$

٦. الأهمية النسبية للمتوسطات المختلفة

➡ الوسط الحسابي أفضل مقاييس النزعة المركزية لأنه يأخذ في الاعتبار جميع مفردات المجتمع، كما يسهل حسابه الجبري، وله استخدامات مهمة في الإحصاء بكافة أنواعه، أذ هو - إضافة إلى ذلك هو الأفضل إذا كان التوزيع منتظماً.

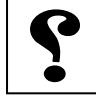
➡ الوسيط هو أسهل مقاييس النزعة المركزية تحديداً، لأنه يعتمد على ترتيب المفردات فقط كما يمكن حسابه في حالة الجداول التكرارية المفتوحة ولا يتأثر بالقيم الشاذة كالوسط الحسابي.

➡ المنوال من مميزاته أنه يعبر عن القيمة الأكثر شيوعاً أو تكراراً، فهو يجذب الانتباه كما أنه لا يتأثر بالقيم المتطرفة إلا أنه ربما لا يوجد وإذا وجد ربما لا يكون فريداً، وهو شديد الاختلاف بين العينات من نفس المجتمع.

👉 الوسط الهندسي والوسط التوافقي هما متوسطان لهما استخداماتهما المعينة؛ فهما يعتمدان على نوع البيانات تحت الدراسة كما ذكرنا سابقاً، ولا يلجأ إليهما كثيراً لأنهما يتأثران تأثيراً مباشراً وكبيراً بالقيم المتطرفة.

👉 لاختيار المتوسط لوصف مفردات المجتمع يتوجب على الباحث أن يختار المتوسطات التي تمثل البيانات أفضل تمثيل، والتي تتلاءم والتعبير الأقرب إلى الصواب، ولكن دائماً يكون الوسط الحسابي هو الأفضل ما لم تكن البيانات موزعة توزيعاً شاذاً أو متطرفة.

أسئلة تقويم ذاتي



- ١ / عرف الوسط الهندسي؟
- ٢ / أشرح باختصار طريقة حساب الوسط الهندسي للبيانات المبوبة؟
- ٣ / كيف يتم حساب الوسط التوافقي للبيانات المبوبة؟
- ٤ / تحدث باختصار عن الأهمية النسبية للمتوسطات المختلفة؟

الخلاصة

عزيزي الدارس،

ماذا تعلمت من هذه الوحدة؟ معلومات كثيرة ومفيدة وردت في هذه الوحدة نلخصها فيما يلي:
بدأنا بتعريف الوسط الحسابي وشرحنا طريقة حسابه من البيانات المبوبة وغير المبوبة ثم قمنا بتعريف الوسيط ووضحنا طرق حسابه وعرفنا المنوال وكذلك وضحنا طرق حسابه. كما وضحنا مزايا وعيوب كل من المتوسطات الثلاثة السابقة (الوسط الحسابي - الوسيط - المنوال) ثم انتقلنا إلى الوسط الهندسي والوسط التوافقي وعرفناهما ووضحنا طريقة حساب كل منهما وأخيراً قمنا بتوضيح الأهمية النسبية للمتوسطات المختلفة.

قدمنا لك العديد من الرسوم والاشكال والجداول التي تعينك في فهم المادة العلمية كما ذودناك بالتدريبات وأسئلة التقويم الذاتي والأنشطة. فارجو أن تكون وحدة مفيدة وممتعة وان تساعدنا في تطوير المادة باسهماك ومقترحاتك مع خالص تمنياتنا لك بالتقدم والنجاح.

لمحة مسبقة عن الوحدة التالية

عزيزي الدارس الوحدة التالية تتحدث عن مقاييس التشتت - المدى،.....، الانحراف المتوسط، نصف المدى الربيعي، التباين، الانحراف المعياري ومعامل الاختلاف.

مسرد المصطلحات

- **الوسط الحسابي: Arithmetic Mean**
الوسط الحسابي لمجموعة من القيم هو القيمة التي لو حلت محل قيمة كل مفردة في المجموعة لكان مجموع هذه القيم الجديدة مساوياً لمجموع القيم الأصلية. يساوي مجموع القيم على عددها.
- **الوسيط: The Median**
هو القيمة التي تتوسط المجموعة بعد ترتيبها تصاعدياً أو تنازلياً.
- **المنوال: The Mode**
هو القيمة الأكثر شيوعاً أو تكراراً.
- **الوسط الهندسي: Geometric Mean**
الوسط الهندسي لمجموعة من القيم عددها N هو الجذر النوني لحاصل ضرب هذه القيم.
- **الوسط التوافقي: Harmonic Mean**
يستخدم في مقياس النزعة المركزية، لمجتمع تتزايد مفرداته تزايد غير مضطرد كالأسعار والسرعة.

إجابة التدريبات

تدريب (١)

الوسط الحسابي = $\frac{\text{مجموع القيم}}{\text{عددها}}$

$$٥٠ = \frac{٢٠٠}{٤} = \frac{٤٠+٣٠+٧٠+٦٠}{٤}$$

تدريب (٢)

الوسيط = $\frac{n+1}{2}$

٢

أولاً نرتب القيم تصاعدياً:

١٥٤، ١٦٥، ١٦٩، ١٨٠، ١٨٧

$$\text{ترتيب الوسيط} = \frac{١+٥}{2} = \frac{6}{2} = 3$$

الوسيط = ١٦٩

تدريب (٣)

المنوال للدرجات: ٩٤، ٨٥، ٧٩، ٧٥، ٨٥، ٨٥، ٨٠، ٧٩

المنوال هو ٨٥ وذلك لأن ٨٥ تكررت أكثر من غيرها.

تدريب (٤)

الوسط الهندسي للأرقام ٩، ٣، ١ هو:

$$3 \sqrt[3]{1 \times 3 \times 9} = 27 = \sqrt[3]{27} = 3$$



محتويات الوحدة

الصفحة	الموضوع
٩٠	مقدمة
٩٠	تمهيد
٩٠	أهداف الوحدة
٩١	١. المدى المطلق
٩١	٢. نصف المدى الربيعي
٩٥	٣. الانحراف المتوسط
٩٨	٤. الانحراف المعياري
١٠٣	٥. التشتت النسبي (معامل الاختلاف)
١٠٧	الخلاصة
١٠٧	لمحة مسبقة عن الوحدة التالية
١٠٨	إجابة التدريبات
١٠٩	مسرد المصطلحات

مقدمة

تمهيد

عزيزي الدارس،

مرحباً بك إلى هذه الوحدة تحتل هذه الوحدة المرتبة الثالثة من وحدات هذا المقرر، لقد تناولنا في الوحدة السابقة مقاييس النزعة المركزية (المتوسطات). ولكن هذه المقاييس لا تعطي فكرة كاملة عن خصائص التوزيعات التكرارية، بحيث لا يمكن إجراء مقارنة متكاملة بين ظاهرتين من معرفة متوسطيهما فقط. فقد يكون المتوسطان متساويين في القيمة بينما تكون مفردات إحدى الظاهرتين متقاربة أي أكثر تجانساً، ومفردات الأخرى متباعدة عن بعضها البعض - أي أكثر تشتتاً .

لذلك تأتي هذه الوحدة ببعض المقاييس التي تحدد مدى تقارب وتباعد مفردات الظاهرة وسوف نتناول المدى المطلق، نصف المدى الربيعي، الانحراف المتوسط، التباين، الانحراف المعياري ومعامل الاختلاف. أرجو أن تكون قد استوعبت الوحدات السابقة حتى يتيسر لك فهم هذه الوحدة نتمنى لك دوماً النجاح والتوفيق.

أهداف الوحدة

عزيزي الدارس بعد فراعك من دراسة هذه الوحدة ينبغي قادراً على أن:



- ١ / تعرف مقاييس التشتت التي وردت في هذه الوحدة.
- ٢ / تشرح كيفية حساب كل مقياس من مقاييس التشتت.
- ٣ / تعدد خطوات حساب الانحراف المعياري.
- ٤ / تذكر الحالات التي يستخدم فيها معامل الاختلاف.

١. المدى المطلق Rang

يعرف المدى المطلق لمجموعة من المفردات عددها بأنه الفرق بين أكبر مفردة وأصغر مفردة من بين تلك المفردات ويرمز له بالرمز R .

$$R = X_1 - X_2$$

حيث: X_1 أكبر مفردة، X_2 أصغر مفردة

مثال (١)

لمجموعة المفردات التالية جد المدى المطلق.

١٩،٦٠،١١،٦،١٢،٨،٤،٣١

الحل

$\begin{aligned} \therefore R &= X_1 - X_2 \\ \therefore R &= 60 - 4 = 56 \end{aligned}$
--

٢. نصف المدى الربيعي Inter –quartile Range

لا حظنا مما سبق إن أهم خصائص المدى غير المرغوب فيها ثائرة بالقيم الشاذة لذا فمن الواجب إيجاد مقياس أو مقاييس أخرى تستبعد هذه القيم الشاذة من الطرفين ومن أهم هذه المقاييس نصف المدى الربيعي، ويمكن حسابه بترتيب البيانات وتقسيمها إلى أربعة أقسام يستبعد منها ربع القيم الصغرى من ناحية، وكذلك ربع القيم الكبرى من ناحية أخرى.

بعد ذلك فإننا نسمي القيمة (النقطة) التي تكون دونها ربع القراءات الربيع الأدنى ويرمز لها بالرمز R_1 أما القيمة (النقطة) التي تحدد ثلاثة أرباع القراءات فتسمى الربيع الأعلى ويرمز لها

بالرمز R_3 والفرق بينهما هو ما يسمى المدى الربيعي اما نصف المدى بين الربع الثالث والربع الاول فيسمى نصف المدى الربيعي.

ويرمز له بالرمز R أي ان:

$$R = \frac{(R_3 - R_1)}{2}$$

ويعتبر نصف المدى الربيعي مقياس للتشتت ويستبعد القيم المتطرفة من الجانبين الاعلى والادنى . ويلاحظ ان القيمة (النقطة) التي تكون دونها نصف القراءات (وتسمى بالربع الثاني) وهي القراءات التي تقسم البيانات الى نصفين ويرمز لها بالرمز R_2 وقد سبقت الاشارة اليها في الفصل السابق على انها الوسيط عند دراسة مقاييس النزعة المركزية وسوف نتناول طريقة حساب نصف المدى الربيعي في حالة البيانات المبوبة كالآتي:

نصف المدى الربيعي للبيانات المبوبة:

لحساب نصف المدى الربيعي للبيانات المبوبة نقوم بحساب الربع الاول والثالث بطريقة متشابه لحساب الوسيط مع الأخذ في الاعتبار ترتيب الربع . ويتم حساب الربع من العلاقة العامة

$$R = L_R + \left[\frac{\frac{\sum f}{k} - \sum f_R}{f_R} \right] C$$

التالية:

$$L_R = \text{الحاجز الادنى للفئة الربيعية} \quad \sum f = \text{مجموع التكرارات}$$

$k = \text{ترتيب الربع}$ $\sum f_k = \text{التكرار المتجمع الصاعد السابق}$

$f_R = \text{تكرار الفئة الربعية}$ $C = \text{طول الفئة}$

وبهذه الطريقة يمكن حساب الربعات الثلاثة كما يلي:

$$R_1 = L_1 + \left[\frac{\frac{\sum f}{4} - \sum f_1}{f_1} \right] C$$

$$R_2 = L_2 + \left[\frac{\frac{2\sum f}{4} - \sum f_2}{f_2} \right] C$$

$$R_3 = L_3 + \left[\frac{\frac{3\sum f}{4} - \sum f_3}{f_3} \right] C$$

ونلاحظ ان الربع الثاني هو الوسيط.

سوف نوضح طريقة حساب نصف المدى الربيعي من بيانات انتاج الذرة بمشروع الجزيرة.

مثال (٢)

من بيانات انتاج الذرة بمشروع الجزيرة جد نصف المدى الربيعي؟

فئات الإنتاج	التكرار f
24.5-	5

29.5-	8
34.5-	10
39.5-	13
44.5-	8
49.5-	6
المجموع	50

الحل

١. ترتيب الربع الاول: $\frac{\sum f}{n} = \frac{50}{4} = 12.5$ ، أي ان R_1 يقع في الفئة (29.5-) وتكرارها 8
 ٢. ترتيب الربع الثاني: $\frac{2(50)}{4} = \frac{100}{4} = 25$ ، أي ان R_2 يقع في الفئة (39.5-) وتكرارها 13
 ٣. ترتيب الربع الثالث $\frac{3(50)}{4} = \frac{150}{4} = 37.5$ ، أي ان R_3 يقع في الفئة (44.5-) وتكرارها ٨
- اولاً تقوم بحساب الربعين الأول والثالث كما يلي:

$$R_1 = L_1 + \left[\frac{\frac{\sum f}{4} - \sum f_1}{f_1} \right] C = 29.5 + \left[\frac{12.5 - 5}{8} \right] \times 5 = 34.19$$

$$R_2 = L_2 + \left[\frac{\frac{2\sum f}{4} - \sum f_2}{f_2} \right] C = 39.5 + \left[\frac{25 - 23}{13} \right] \times 5 = 40.27$$

$$R_3 = L_3 + \left[\frac{\frac{3\sum f}{4} - \sum f_3}{f_3} \right] C = 44.5 + \left[\frac{37.5 - 36}{8} \right] \times 5 = 45.44$$

وعليه فان نصف المدى الربيعي هو :

$$QD = \frac{1}{2}(R_3 - R_1)$$

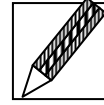
$$\therefore QD = \frac{1}{2}(45.44 - 34.19) = 5.6$$

ومن مزايا نصف المدى الربيعي عدم تأثره، بالقيم الشاذة المتطرفة كما يمكن حسابه في حالة الجداول التكرارية المفتوحة، من عيوبه انه لا يأخذ جميع القيم في الاعتبار عند حسابه كما لا يصلح في التحليل الإحصائي بنفس طريقة الربيعات يمكن حساب التشتت باستخدام العشريان وذلك بتقسيم المجموعة إلى عشرة اقسام او المئينات بتقسيم المجموعة الى مائة جزء... وهكذا.

تدريب (1)

أحسب المدى المطلق لمجموعة المفردات التالية:

١٥،٢٩،٣٥،٨،٨٩



أسئلة تقويم ذاتي



- ١- ما معنى التشتت؟
- ٢- عرف المدى المطلق
- ٣- ما هي أهم خصائص المدى غير المرغوب فيها؟
- ٤- أشرح باختصار كيفية حساب نصف المدى الربيعي للبيانات المبوبة؟
- ٥- عدد مزايا وعيوب نصف المدى الربيعي؟

٣. الانحراف المتوسط MD – Middle Deviation

يعرف الانحراف المتوسط لمجموعة من المفردات عددها N بأنه متوسط مجموع القيم المطلقة

لأنحرافات المفردات عن وسطها الحسابي ويرمز له بالرمز MD.

$$MD = \frac{\sum_{i=1}^N |X_i - \bar{X}|}{N}$$

وخطوات حسابه هي:

١/ نحسب أولاً الوسط الحسابي $\bar{X}$ للمفردات؛ حيث $\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$.

٢/ نحسب انحراف كل مفردة X_i عن وسطها الحسابي $\bar{X}$ لنكون عمود $X_i - \bar{X}$.

٣/ نأخذ القيم المطلقة لهذه الانحرافات لنكون عمود $|X_i - \bar{X}|$.

٤/ نطبق القانون.

مثال (٢)

لمجموعة المفردات التالية جد: (أ) الوسط الحسابي (ب) الانحراف المتوسط.

8, 4, 6, 10, 2

الحل

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} = \frac{6+8+2+10+4}{5} = \frac{30}{5} = 6 \quad \text{أ}$$

ب/ لحساب الانحراف المتوسط نكون أولاً الأعمدة الداخلة في حسابه.

X_i	$X_i - \bar{X}$	$ X_i - \bar{X} $
8	8 - 6 = 2	٢
6	6 - 6 = 0	٠
10	10 - 6 = 4	٤

2	2 - 6 = - 4	٤
4	4 - 6 = - 2	٢
sum	$\sum (X_i - \bar{X}) = 0$	١٢

ويكون:

$$MD = \frac{\sum_{i=1}^N |X_i - \bar{X}|}{N} = \frac{12}{5} = 2.4$$

الانحراف المتوسط من البيانات المبوبة:

يحسب الانحراف المتوسط من البيانات المبوبة وفقاً للعلاقة التالية:

$$MD = \frac{\sum_{i=1}^N f_i |X_i - \bar{X}|}{\sum f}$$

حيث: f_i التكرار، X_i مراكز الفئات.

وخطوات حسابه هي:

$$\bar{X} = \frac{\sum_{i=1}^N f_i X_i}{N}$$

١ / نحسب أولاً الوسط الحسابي .

٢ / نكون عمود لانحراف كل مركز فئة X_i عن الوسط الحسابي $\bar{X}$.

٣ / نأخذ القيم المطلقة لكل انحراف لنحصل على عمود $X_i - \bar{X}$.

٤ / نضرب كل تكرار f_i في الانحراف المطلق المقابل له $|X_i - \bar{X}|$ لنكون عمود $f_i \times |X_i - \bar{X}|$ ونجمع لنحصل على $\sum f_i |X_i - \bar{X}|$.

٥ / نطبق القانون .

مثال (٣)

من الجدول التالي جد:

أ/ الوسط الحسابي .

ب/ الانحراف المتوسط .

C	4-	6-	8-	10-	12-	sum
f	2	3	5	3	2	15

الحل

أ/ نحسب الوسط الحسابي بنفس الطريقة السابقة.

C	f	X	fX	$X_i - \bar{X}$	$ X_i - \bar{X} $	$f_i X_i - \bar{X} $
٤-	٢	٥	10	-٤	٤	8
٦-	٣	٧	21	-٢	٢	6
٨-	٥	٩	45	٠	٠	0
١٠-	٣	١١	33	٢	٢	6
١٢-	٢	١٣	26	٤	٤	8
المجموع	١٥		135	٠		٢٨

$$\therefore \bar{X} = \frac{\sum_{i=1}^N f_i X_i}{N} = \frac{135}{15} = 9$$

ب/ لحساب الانحراف المتوسط فإن:

$$MD = \frac{\sum_{i=1}^N f_i |X_i - \bar{X}|}{\sum f} = \frac{28}{9} = 3.11$$

أسئلة تقويم ذاتي



- ١ / عرف الانحراف المتوسط؟
- ٢ / ماهي خطوات حساب الانحراف المتوسط للبيانات غير المبوبة؟
- ٣ / أشرح كيفية حساب الانحراف المتوسط من البيانات المبوبة؟

٤ . الانحراف المعياري Standard Deviation

يعتبر الانحراف من أكثر مقاييس التشتت استخداماً. فإذا كان لدينا N من المفردات $X_1, X_2, X_3, \dots, X_N$ ووسطهما الحسابي هو $\bar{X}$ ، فإن هذه المفردات تكون متقاربة من بعضها البعض إذا كانت قريبة من وسطها الحسابي $\bar{X}$ ، أي إذا كانت انحرافاتهما عن $\bar{X}$ صغيرة. وبالتالي فإن انحرافات المفردات عن وسطها الحسابي يمكن استخدامها لقياس التشتت. ويمكن أن يتم ذلك بأخذ متوسط هذه الانحرافات. وحيث أن مجموع انحرافات القيم عن وسطها الحسابي دائماً يساوي صفراً، لأن بعض هذه الانحرافات موجب والآخر سالب، والموجب يتلاشى السالب، فيمكن التغلب على ذلك بأخذ متوسط مربعات هذه الانحرافات والذي يسمى التباين **Variance**.

أي أن التباين، ويرمز له عادة بالرمز σ^2 ويقراً (sigma square)، ويكتب كما يلي:

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}$$

ويلاحظ أن وحدات التباين هي مربع الوحدات الأصلية، ونظراً لأنه يفضل أن يكون مقياس التشتت بنفس الوحدات الأصلية، فيؤخذ الجذر التربيعي للتباين كمقياس للتشتت ويسمى الانحراف المعياري **Standard Deviation**، ويرمز له بالرمز σ .

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}}$$

أي أن:

تعريف:

الانحراف المعياري هو الجذر التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابي.

وخطوات حسابه هي:

$$١/ \text{ نحسب أولاً الوسط الحسابي } \bar{X} \text{ للمفردات؛ حيث } \bar{X} = \frac{\sum_{i=1}^N X_i}{N}.$$

$$٢/ \text{ نحسب انحراف كل مفردة } X_i \text{ عن وسطها الحسابي } \bar{X} \text{ لنكون عمود } X_i - \bar{X}.$$

٣/ نربع كل انحراف لنكون عمود $(X_i - \bar{X})^2$.

٤/ نطبق القانون.

مثال (٤)

أحسب الانحراف المعياري للقيم الآتية: ٢٣، ٢٨، ٢٤، ٢٩، ٣٢، ٢١، ٢٥، ٣٤

الحل

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} = \frac{23+28+24+29+32+21+25+34}{8} = \frac{216}{8} = 27$$

نحسب أولاً الوسط الحسابي.

لحساب الانحراف المعياري نكون الجدول التالي:

X	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
23	-٤	١٦
28	١	١
24	-٣	٩
29	٢	٤
32	٥	٢٥
21	-٦	٣٦
25	-٢	٤
34	٧	٤٩
216	٠	١٤٤

$$\sigma^2 = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}} = \sqrt{\frac{144}{8}} = \sqrt{18} = 4.242$$

صيغة مختصرة للانحراف المعياري:

يمكن كتابة صيغة الانحراف المعياري:

$$\sigma^2 = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}}$$

على الصورة المختصرة التالية:

$$\sigma^2 = \sqrt{\frac{\sum_{i=1}^N X_i^2}{N} - \bar{X}^2}$$

وتتطلب هذه الطريقة حساب:

١- الوسط الحسابي $\bar{X}$.

٢- مجموع مربعات القيم X^2 .

مثال (٥)

أحسب الانحراف المعياري للبيانات التالية:

٩ ، ٧ ، ٥ ، ٣ ، ١

الحل

لحساب الانحراف المعياري نكون الجدول التالي:

X	1	3	5	7	9
X ²	1	9	25	49	81

ويكون: $\bar{X} = \frac{\sum X}{N} = \frac{25}{5} = 5$, $\sum X^2 = 165$

$$\therefore \sigma^2 = \sqrt{\frac{\sum_{i=1}^N X_i^2}{N} - \bar{X}^2} = \sqrt{\frac{165}{5} - 25} = \sqrt{33 - 25} = \sqrt{8} = 2.83$$

مثال (٦)

أطرح ٢٠ من كل قيمة من القيم الموجودة في المثال رقم [٤] ثم أحسب الانحراف المعياري للقيم الجديدة وقارن النتائج.

الحل

بعد طرح ٢٠ من كل قيمة تصبح القيم: ١٤ ، ٥ ، ١ ، ١٢ ، ٩ ، ٤ ، ٨ ، ٣

ولحساب الانحراف المعياري نكون الجدول التالي:

X	٣	٨	٤	٩	١٢	1	٥	١٤
X ²	9	64	16	81	144	١	٢٥	١٩٦

$$\bar{X} = \frac{\sum X}{N} = \frac{56}{8} = 7, \quad \sum X^2 = 536$$

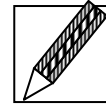
ويكون:

$$\therefore \sigma^2 = \sqrt{\frac{\sum_{i=1}^N X_i^2}{N} - \bar{X}^2} = \sqrt{\frac{536}{8} - 49} = \sqrt{67 - 49} = \sqrt{18} = 4.242$$

نلاحظ في المثالين (٤)، (٦) أن قيمة الانحراف فيهما متساوية. ومن هذا نستنتج أن طرح كمية من جميع القيم لا تؤثر على قيمة الانحراف المعياري.

تدريب (٢)

وضح أهم وأول خطوة في حساب الانحراف المعياري من البيانات غير المبوبة؟



الانحراف المعياري للبيانات المبوبة:

إذا كانت البيانات مصنفة في جدول تكراري، فإنه يمكن حساب الانحراف المعياري باستخدام الصيغة الآتية:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N f_i (X_i - \bar{X})^2}{\sum f}}$$

ولحساب الانحراف المعياري في هذه الحالة نتبع الخطوات الآتية:

$$١/ \text{ نحسب أولاً الوسط الحسابي } \bar{X} = \frac{\sum_{i=1}^N f_i X_i}{N}$$

٢/ نكون عمود لانحراف كل مركز فئة X_i عن الوسط الحسابي $\bar{X}$.

٣/ نربع كل انحراف لنحصل على عمود $(X_i - \bar{X})^2$.

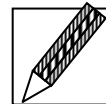
٤/ نضرب كل تكرار f_i في الانحراف المربع المقابل له $(X_i - \bar{X})^2$ لنكون عمود $f_i (X_i - \bar{X})^2$

ونجمع لنحصل على $\sum f_i (X_i - \bar{X})^2$.

٥/ نطبق القانون.

تدريب (٣)

وضح السبب في عدم استخدام الانحراف المعياري لمقارنة تشتت أطوال واوزان عدد من الطلاب؟



مثال (٧)

البيانات التالية تمثل توزيع مائة عامل في أحد المصانع حسب الأجر اليومي (بالدينار):

فئات الأجر	١٥	٢٥	٣٥	٤٥	٥٥	٦٥	٧٥-٨٥	المجموع
عدد العمال	٦	١٢	١٨	٢٤	٢٠	١٢	٨	١٠٠

أحسب الانحراف المعياري للأجر اليومي للعمال.

الحل

نكون الجدول التالي:

Class	f	X	fX	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^2$
١٥-	6	٢٠	120	-٣٠.٨	٩٤٨.٦٤	٥٦٩١.٨٤
٢٥-	١٢	٣٠	٣٦٠	-٢٠.٨	٤٣٢.٦٤	٥١٩١.٦٨
٣٥-	١٨	٤٠	٧٢٠	-١٠.٨	١١٦.٦٤	٢٠٩٩.٥٢
٤٥-	24	٥٠	١٢٠٠	-٠.٨	٠.٦٤	٠٠١٥.٣٦
٥٥-	20	٦٠	١٢٠٠	٩.٢	٨٤.٦٤	١٦٩٢.٨٠
٦٥-	12	٧٠	٨٤٠	١٩.٢	٣٦٨.٦٤	٤٤٢٣.٦٨
75-85	8	٨٠	٦٤٠	٢٩.٢	٨٥٢.٦٤	٦٨٢١.١٢
المجموع	١٠٠		٥٠٨٠			٢٥٩٣٦.٠٠

ومن هذا نجد أن:

$$\therefore \bar{X} = \frac{\sum_{i=1}^N f_i X_i}{N} = \frac{5080}{100} = 50,8 \text{ Dinars}$$

و لحساب الانحراف المعياري فإن:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N f_i (X_i - \bar{X})^2}{\sum f}} = \sqrt{\frac{25936}{100}} = \sqrt{259.36} = 16.1 \text{ Dinars}$$

من الناحية العملية يمكن تقريب القيمة ٥٠.٨ لتكون ٥١.



١/ البيانات التالية تمثل درجات عشرة طلاب في مادة التاريخ:

٥٩ ، ٥٩ ، ٨٦ ، ٧١ ، ٤٢ ، ٦٥ ، ٧٨ ، ٨٩ ، ٩٣ ، ٥٩

احسب الانحراف المعياري للدرجات.

٢/ فيما يلي توزيع ١١٠ طلاب في إحدى المدارس الثانوية حسب درجاتهم في

مادة الإحصاء:

الدرجات	-٣٠	-٤٠	-٥٠	-٦٠	-٧٠	-٨٠	-٩٠	المجموع
عدد الطلاب	٩	١١	٢٠	٣٠	٢٠	١١	٩	١١٠

5. التشتت النسبي (معامل الاختلاف)

Coefficient of Variation

إذا أردنا على سبيل المثال مقارنة تشتت أطوال و أوزان عدد من الطلاب، فإنه لا يمكن استخدام الانحراف المعياري لهاتين الظاهرتين، وذلك لاعتماده على وحدات القياس (حيث تقاس الأطوال بالسنتيمتر والأوزان بالكيلوجرام). ولذا نلجأ إلى مقياس آخر للتشتت يعتمد على الانحراف المعياري ولكنه لا يتأثر بوحدات القياس، ويسمى بمعامل الاختلاف. وهو مقياس نسبي للتشتت، ويستخدم في الحالات التالية:

١/ مقارنة تشتت مجموعات من القيم تختلف في وحدات القياس.

٢/ مقارنة تشتت مجموعات من القيم تقاس بنفس الوحدات ولكن تختلف في أوساطها الحسابية.

٣ / مقارنة تشتت مجموعات من القيم تختلف في وحدات قياسها وأوساطها الحسابية معاً.

تعريف:

يعرف معامل الاختلاف على أنه النسبة المئوية للانحراف المعياري بالنسبة للوسط الحسابي.

يرمز لمعامل الاختلاف بالرمز CV. ويحسب بالعلاقة:

$$CV = \frac{c}{\bar{X}} \times 100$$

مثال (٨)

أحسب معامل الاختلاف للبيانات التالية:

٩ ، ٧ ، ٩ ، ٥ ، ١٠ ، ٦ ، ٤ ، ٨ ، ٥ ، ٧

الحل

$$\sum X = 70, \sum X^2 = 526$$

$$\bar{X} = \frac{70}{10} = 7$$

ويكون:

كما أن:

$$\sigma = \sqrt{\frac{526}{10} - (7)^2} = \sqrt{52.6 - 49} = \sqrt{3.6} = 1.90$$

$$CV = \frac{c}{\bar{X}} \times 100 = \frac{1.90}{7} \times 100 = 27.14\%$$

وبهذا فإن:

مثال (٩)

من البيانات الآتية وضح أي المجموعتين أكثر تشتتاً:

المجموعة الأولى: $\bar{X} = 10, \sigma = 5$

المجموعة الثانية: $\sum X = 100, N = 5, \sigma = 7$

الحل

بالنسبة للمجموعة الأولى فإن:

$$CV_1 = \frac{c}{\bar{X}} \times 100 = \frac{5}{10} \times 100 = 50\%$$

بالنسبة للمجموعة الثانية:

$$\bar{X} = \frac{\sum X}{N} = \frac{100}{5} = 20$$

$$CV_2 = \frac{c}{\bar{X}} \times 100 = \frac{7}{20} \times 100 = 35\%$$

إذاً:

أي أن المجموعة الأولى أكثر تشتتاً.

نشاط



١/ البيانات التالية تمثل توزيع طلاب إحدى المدارس الثانوية حسب العمر:

فئات العمر بالسنوات	-١٤	-١٦	-١٨	-٢٠	-٢٢	المجموع
عدد الطلاب (التكرار)	٢٠	٤٠	٨٠	٤٠	٢٠	٢٠٠

والمطلوب : حساب الانحراف المتوسط والمعياري لأعمار الطلاب.

٢/ عند دراسة أعمار مجموعتين من الطلاب حصلنا على المعلومات الآتية:
المجموعة الأولى:

الوسط الحسابي للعمر = ١٩ سنة، الانحراف المعياري للأعمار = ٢ سنة
المجموعة الثانية:

الوسط الحسابي للعمر = ٣١ سنة ، الانحراف المعياري للأعمار = ٣ سنوات
فأي المجموعتين أكثر تشتتاً.

٣/ احسب معامل الاختلاف للبيانات التالية :

١٥ ، ٥ ، ١٠ ، ٢٠ ، ١٥ ، ١٠ ، ٨ ، ٢ ، ٥ ، ٢٠

أسئلة تقويم ذاتي



- ١- عرف الإنحراف المعياري؟
- ٢- عدد خطوات حساب الانحراف المعياري.
- ٣- أكتب صيغة مختصرة للانحراف المعياري من البيانات المبوبة؟

الخلاصة

عزيزي الدارس،

مرحباً بك في نهاية هذه الوحدة. تناولت في هذه الوحدة مقاييس التثنت التي تحدد تقارب أو تباعد مفردات الظاهرة عن المدى المطلق والانحراف المتوسط والتباين والانحراف المعياري ومعامل الاختلاف. حيث قمنا بتعريف كل واحد من هذه المقاييس مع شرح لطريقة حسابه. أرجو أن تكون وحدة ممتعة ومفيدة وان تستفيد من الأمثلة الواردة في حل التدريبات والنشاطات التي وردت في هذه الوحدة حتى تستوعب هذه الوحدة والوحدات القادمة . أرجو أن تكون قد حققت الأهداف التعليمية الواردة في بداية هذه الوحدة وأن تتأكد دائماً في نهاية كل وحدة من أنك قد حققت تلك الأهداف. ختاماً نسأل الله لك التوفيق والنجاح.

لمحة مسبقة عن الوحدة التالية

عزيزي الدارس،

الوحدة التالية تتحدث عن الالتواء والنفرطح والعزوم

إجابات التدريبات

تدريب (١)

المدى المطلق **Rang** = الفرق بين أعلى مفردة وأدنى مفردة = $89 - 8 = 81$

تدريب (٢)

أهم وأول خطوة هي حساب الوسط الحسابي

تدريب (٣)

لا يمكن استخدام الانحراف المعياري لهاتين الظاهرتين وذلك لاعتماده على وحدات القياس (حيث تقاس الأوزان بالكيلوجرام والأطوال بالسنتيمتر).

مسرد المصطلحات

- المدى المطلق Rang
هو الفرق بين أكبر مفردة واصغر مفردة من بين تلك المفردات ويرمز له بالرمز **R**
- الانحراف المتوسط Mean Deviation
هو متوسط مجموعة القيم المطلقة لانحراف المفردات عن وسطها الحسابي ويرمز له بالرمز **MD**
- الانحراف المعياري Standard Deviation
هو الجذر التربيعي لمتوسط مربعات انحراف القيم عن وسطها الحسابي
- التشتت النسبي (معامل الاختلاف) Coefficient of Variation
هو النسبة المئوية للانحراف المعياري بالنسبة للوسط الحسابي.



محتويات الوحدة

الصفحة	الموضوع
112	مقدمة
112	تمهيد
113	أهداف الوحدة
114	١. مقاييس الالتواء
117	٢. مقاييس التفريط
117	٣. العزوم
121	٤. حساب معامل الالتواء باستخدام العزوم
121	٥. حساب معامل التفريط باستخدام العزوم
124	الخلاصة
124	لمحة مسبقة عن الوحدة التالية
125	إجابات التدريبات

مقدمة

تمهيد

عزيزي الدارس،

مرحباً بك في رابع وحدات هذا المقرر. يعتبر المنحنى التكراري المعتدل مرجعاً هاماً في الدراسات والتحليلات الإحصائية وعليه تعتمد الكثير من النظريات الإحصائية مثل نظرية التقدير ونظرية القرار وغيرها. إلا أنه في الحياة العملية يندر وجود توزيع معتدل اعتدالاً تاماً بحيث يتمثل حول وسطه الحسابي. والذي نحصل عليه في الواقع عادة هو التوزيع التكراري الذي يكون قريباً من المعتدل أو شبه المعتدل. والمنحني شبه المعتدل إما أن يكون ملتوياً نحو اليمين أو نحو اليسار أو تكون قمته أكثر أو أقل تفرطحاً من قمة التوزيع المعتدل.

لهذه الأسباب فإنه يهمننا في القياسات الإحصائية أن نتعرف على درجة الالتواء والتفرطح للتوزيع التكراري. تحتوي هذه الوحدة على خمس أقسام في القسم الأول تناولنا مقاييس الالتواء ومعامل الالتواء، وفي القسم الثاني تناولنا مقاييس التفرطح وفي القسم الثالث تناولنا العزوم، وفي القسم الرابع وضحنا كيفية حساب معامل الالتواء باستخدام العزوم، أما في القسم الأخير فتناولنا حساب معامل التفرطح باستخدام العزوم. تتخلل هذه الوحدة العديد من الأنشطة والتدريبات والأمثلة التي تعينك على الفهم والاستيعاب. مع خالص تمنياتنا بالتوفيق والنجاح.

أهداف الوحدة



عزيزي الدارس

بعد فراغك من دراسة هذه الوحدة ينبغي أن تكون قادراً على أن :

- ١- تشرح مقاييس الالتواء.
- ٢- تذكر مزايا وعيوب معامل الالتواء.
- ٣- تناقش مقاييس التفرطح.
- ٤- تحسب قيمة العزوم من البيانات المبوبة وغير المبوبة.
- ٥- تحسب معامل الالتواء ومعامل التفرطح عن طريق العزوم.

١. مقاييس الالتواء

عزيزي الدارس سبق أن ذكرنا في دراستنا لمقاييس النزعة المركزية أنه إذا كانت المنحنيات معتدلة (متماثلة) فإن مقاييس الوسط الحسابي والوسيط والمنوال تكون متطابقة، أي لها نفس القيم. أما في حالة عدم التماثل فإنها تختلف .

لنرمز للوسط الحسابي و الوسيط و المنوال بـ $\bar{X}$, $\tilde{X}$, $\hat{X}$ بالترتيب

١/ إذا كان المنحنى ملتوياً نحو اليسار فيكون الوسط الحسابي أصغرها وتطبق العلاقة:

$$\bar{X} < \tilde{X} < \hat{X}$$

٢/ إذا كان المنحنى ملتوياً نحو اليمين، فيكون الوسط الحسابي أكبرها وتطبق العلاقة:

$$\bar{X} > \tilde{X} > \hat{X}$$

في الحالة الأولى يكون الالتواء سالباً بحيث يكون ذيل التوزيع الأطول نحو اليسار. أما في الحالة الثانية فيكون الالتواء موجباً بحيث يكون ذيل التوزيع الأطول نحو اليمين . ويتم تحديد نوعية الالتواء باستخدام الفرق بين الوسط والمنوال فإذا كان:

$$(أ) \quad (\bar{X} - \hat{X}) \quad \text{قيمة موجبة يكون الالتواء موجبا.}$$

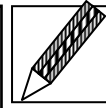
(ب) قيمة سالبة يكون الالتواء سالبا. $(\bar{X} - \hat{X})$

(ج) تساوي صفر يكون التوزيع معتدلاً. $(\bar{X} - \hat{X})$

وتبين الحالات الثلاثة في شكل (١):

تدريب (١)

وضح متى تكون مقاييس الوسط الحسابي والوسيط والمنوال متطابقة؟



معامل الالتواء

لقد تعرفنا فيما سبق على العلاقة بين الوسط الحسابي والوسيط والمنوال عندما يكون التوزيع التكراري شبه متماثل وأوضحنا أن العلاقة التالية تتحقق:

$$(\bar{X} - \hat{X}) = 3(\bar{X} - \tilde{X})$$

وقد استخدم كارل بيرسون هذه العلاقة في اشتقاق مقياسين للالتواء B_1 , B_2 على النحو التالي:

$$\begin{aligned} B_1 &= (\bar{X} - \hat{X}) / \sigma_x \\ B_2 &= 3(\bar{X} - \tilde{X}) / \sigma_x \end{aligned}$$

ويسمى المعامل B_1 معامل بيرسون الأول B_2 معامل بيرسون الثاني للالتواء.

مثال:

في توزيع أجور ٦٥ عاملا في إحدى الشركات (الوحدة الثانية) كان:
الوسط الحسابي $\bar{X} = 79.76$ و الوسيط $\tilde{X} = 79.06$ والمنوال $\hat{X} = 77.5$ وكان الانحراف المعياري $c = 15.6$. أوجد معاملي بيرسون الأول والثاني للالتواء.
لإيجاد معاملي بيرسون نقوم بالتعويض المباشر في المعادلتين السابقتين:

* معاملا بيرسون الأول والثاني للالتواء:

$$B_1 = \frac{79.76 - 77.50}{15.60} = 0.14$$

$$B_2 = \frac{3(79.76 - 79.06)}{15.60} = 0.13$$

بما أن B_1 , B_2 موجبة فإن التوزيع ملتوي التواء موجبا أي أنه ملتوي ناحية اليمين ولكن من الواضح أن درجة الالتواء خفيفة للغاية وذلك لان :

$$B_1 = 0.14, \quad B_2 = 0.13$$

مما يعني أن لهم قيم متقاربة وعليه فإن العلاقة التقريبية بين الوسط والوسيط والمنوال قد تحققت الأمر الذي يدل على قرب التوزيع من التماثل . ويمكن قياس الالتواء من الانحراف الربيعي عن طريق معامل الالتواء الربيعي B .

$$B = (R_3 - 2R_2 + R_1) / (R_3 - R_1)$$

يمكن حساب معامل المدى الربيعي من بيانات إنتاج الذرة في مثال رقم ٢ الفصل الثاني حيث وجدنا أن:

$$R_1 = 34.19$$

$$R_2 = 40.27$$

$$R_3 = 45.44$$

وعليه فإن معامل الالتواء الربيعي :

$$B = (45.44 - 2 \times 40.27 + 34.19) / (45.44 - 34.19) = 0.08$$

والأساس في هذا المعامل هو أنه في حالة التوزيعات المتماثلة تقع الربيعات الأول والثالث على بعدين متساويين من الوسيط، في حين يختلف بعداهما عنه في حالة التوزيعات الملتوية أي أنه عندما يكون قيم B أقرب إلى الواحد الصحيح.

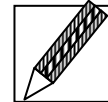
مزايا وعيوب معامل الالتواء

من مزايا هذا المعامل انه يمكن إيجاده بالحساب أو بالرسم باستخدام المنحنى التكراري المتجمع الصاعد . وكذلك يمكن حسابه في حالة الجداول المفتوحة.

ولكن عيبه الأساسي أنه لا يأخذ جميع القيم في الاعتبار، بل يعتمد على القيم الواقعة بين الربيعين الأول والثالث فقط. كما أنه لا يمكن استخدامه في التحليل الإحصائي الرياضي.

تدريب (٢)

- قارن بين مزايا وعيوب معامل الالتواء.



أسئلة تقويم ذاتي

- ١- كيف يتم تحديد نوعية الالتواء؟
- ٢- عدد مزايا وعيوب معامل الالتواء؟
- ٣- تحدث باختصار عن مقاييس الالتواء؟



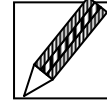
٢. مقاييس التفرطح

عزيزي الدارس ماذا تعرف عن كلمة التفرطح ومتى تستخدم؟

ونستعمل كلمة تفرطح لوصف تدبب المنحنى التكراري بالنسبة للمنحنى المعتدل فإذا كان المنحنى أكثر تفرطحاً من المنحنى المعتدل يوصف بأنه مفرطح وهو الذي يتسع في الوسط وتتحني قمته عن قمة

المنحني المعتدل. إذا كان المنحني أقل تفرطحاً من المنحني المعتدل يوصف بأنه منحني مدبب وهو الذي يأخذ مدى ضيقاً في وسطه وترتفع قمته عن قمة المنحني المعتدل. أما إذا كان المنحني التكراري له قمة مطابقة للمنحني المعتدل فاه يوصف بأن تفرطحه متوسط (Mesokartik) أنظر شكل (٢).
وتستخدم العزوم لقياس معاملي الالتواء والتفرطح . لذلك فأنا سنقوم بشرح مفهوم العزوم وطريقة حسابها:
تدريب (٣)

- وضح متى يوصف المنحني بأنه مفرطح.



٣. العزوم Moments

إذا كان لدينا مجموعة من المفردات $X_1, X_2, \dots, X_N$ فإن العزم الرائي حول الصفر ويرمز له بالرمز (X_r) يحسب من العلاقة التالية:

$$\bar{X}_r = \frac{\sum X^r}{n}$$

وعلى وجه الخصوص عندما تكون $r = 1$ فإن العزم الأول حول الصفر هو الوسط الحسابي $\bar{X}$ كذلك العزم الرائي حول الوسط الحسابي ويرمز إليه بالرمز (m_r) يحسب من العلاقة التالية:

$$m_r = \sum (X - \bar{X})^r / n$$

فإذا كانت $r = 1$ فإن العزم الأول حول الوسط الحسابي $= 0$ أما إذا كانت $r = 2$ فإن العزم الثاني حول الوسط الحسابي يساوي التباين c^2 .

وهناك أيضاً العزم الرائي حول أي نقطة A ويرمز له بالرمز (m'_r) ويحسب من العلاقة التالية:

$$m'_r = \sum (X - A)^r / n$$

ونلاحظ أنه إذا كانت $A = \bar{X}$ فإن العزم الرائي حول النقطة A يساوي العزم الرائي حول الوسط الحسابي . وكذلك إذا كانت $A = 0$ فإن العزم الرائي حول النقطة A يساوي العزم الرائي حول الصفر. وعليه فإن m'_r أكثر عمومية وتعتمد قيمته على نقطة الأصل A .

ونلاحظ أيضاً أن $d_1 = x_1 - A$ هي انحرافات القيم عن أي وسط فرضي A والتي استخدمناها في حساب الوسط الحسابي والانحراف المعياري في الأقسام السابقة.

وستستخدم العزم الرائي حول النقطة A في اشتقاق علاقات مختلفة بين العزوم وهي :

$$\begin{aligned}
m_1 &= 0 \\
m_2 &= m'_2 - m_1' \\
m_3 &= m'_3 - 3m'_1m'_2 + 2m_1'^3 \\
m_4 &= m'_4 - 4m'_1m'_3 + 6m_1'^2m'_2 - 3m_1'^4
\end{aligned}$$

ونقوم بإثبات m_2 فيما يلي:

$$\begin{aligned}
m_2 &= \sum (X - \bar{X})^2 / n \\
m_2 &= \left[\sum X^2 / n - (\sum X / n)^2 \right] \\
d &= X - A \\
X &= A + d \\
\bar{X} &= A + \bar{d} \\
(X - \bar{X}) &= (d - \bar{d}) \\
\sum (X - \bar{X})^2 / n &= \sum (d - \bar{d})^2 / n \\
&= \left\{ \sum d^2 / n - (\sum d / n)^2 \right\} \\
&= \left\{ \sum (X - A)^2 / n - (\sum X - A / n)^2 \right\} \\
\therefore m_2 &= m'_2 - m_1'^2
\end{aligned}$$

مثال :

أوجد العزوم الثلاثة الأولى حول الرقم ٤ لمجموعة المفردات التالية ١٠ ، ٨ ، ٧ ، ٣ ، ٢ ومن ثم أوجد العزم الثاني والثالث حول الوسط الحسابي:

$$\begin{aligned}
m'_1 &= \sum (x - 4) / n = \frac{(2 - 4) + (3 - 4) + (7 - 4) + (8 - 4) + (10 - 4)}{5} = 2 \\
m'_2 &= \sum (x - 4)^2 / n = \frac{(2 - 4)^2 + (3 - 4)^2 + (7 - 4)^2 + (8 - 4)^2 + (10 - 4)^2}{5} = 13.2 \\
m'_3 &= \sum (x - 4)^3 / n = \frac{(2 - 4)^3 + (3 - 4)^3 + (7 - 4)^3 + (8 - 4)^3 + (10 - 4)^3}{5} = 59.6
\end{aligned}$$

ثم تستخدم العلاقة بين العزوم لحساب العزم الثاني والثالث حول الوسط الحسابي كما يلي:

$$\begin{aligned}
m_2 &= m'_2 - m_1'^2 = 13.2 - (2)^2 = 9.2 \\
m_3 &= m'_3 - 3m'_1m'_2 + 2m_1'^3 \\
&= 59.6 + 3(2)(13.2) + 2(2)^3 = 3.6
\end{aligned}$$

حساب العزوم من البيانات المبوبة:

إذا كانت $x_1, x_2, \dots, x_n$ هي مراكز فئات لها تكرارات $f_1, f_2, \dots, f_n$ فإن العزوم السابقة يمكن أن تكتب بتكرارات كما يلي:

$$\bar{X}_r = \sum f x^r / n$$

$$m_r = \sum f (X - \bar{X})^r / n$$

$$m'_r = \sum f (X - A)^r / n$$

وبما أن هذه الصيغ غير ملائمة للحساب نسبة لطول العملية الحسابية ولتسهيل الأمر سنلجأ إلى طرق الترميز السابقة وذلك لاستخدامها بصورة مختصرة لإيجاد العزوم . ولتوضيح هذه الطرق أكثر دعنا نتوصل إليها بالخطوات التالية:

$$m'_r = \sum f d^r / \sum f \quad \text{ضع } d = x - A \text{ عليه فإن.}$$

والآن ضع $u = d/c$ فتصبح :

$$m'_r = c^r \left(\sum f u^r / \sum f \right)$$

ويمكن استخدام هذه الصيغ للحصول على m_r بتطبيق معادلة اشتقاق العزوم.

مثال (٤)

أوجد العزوم الأربعة حول الوسط الحسابي لتوزيع إنتاج الذرة بمشروع الجزيرة .
يتم أولاً تنظيم العمل وفق الجدول التالي:

جدول رقم (١)

حساب العزوم من بيانات إنتاج الذرة بمشروع الجزيرة

فئات الإنتاج	f	x	d=x-A	u=d/c	fu	fu ²	fu ³	fu ⁴
-27.5	٥	٢٧	-١٥	-٣	-١٥	٤٥	-١٣٢	٤٠٥
-29.5	٨	٣٢	-١٠	-٢	-١٦	٣٢	-٦٤	١٢٨
-34.5	١٠	٣٧	-٥	-١	-١٠	١٠	-١٠	١٠
-39.5	١٣	٤٢	٠	٠	٠	٠	٠	٠
-44.5	٨	٤٧	٥	١	٨	٨	٨	٨
-49.5	٦	٥٢	١٠	٢	١٢	٢٤	٤٨	٩٦

المجموع	٥٠		-٢١	١١٩	-١٥٣	٦٤٧
---------	----	--	-----	-----	------	-----

$$m'_1 = c \left(\sum fu / \sum f \right) = (5)(-21/50) = -2.1$$

$$m'_2 = c^2 \left(\sum fu^2 / \sum f \right) = (5)^2 (119/50) = 59.5$$

$$m'_3 = c^3 \left(\sum fu^3 / \sum f \right) = (5)^3 (153/50) = -382.5$$

$$m'_4 = c^4 \left(\sum fu^4 / \sum f \right) = (5)^4 (647/50) = 8087.5$$

الان نحسب العزوم الرابعة الاولى حول الوسط الحسابي بالاستفادة من هذه النتائج

$$m_1 = 0 \quad (\text{حسب التعريف})$$

$$m_2 = m'_2 - m_1'^2 = 59.5 - (2.1)^2$$

$$m_2 = 54.1$$

$$m_3 = m'_3 - 3m'_1m'_2 + 2m_1'^3 = 382.5 + (3)(-2.1)(59.5) + 2(2.1)^3$$

$$m_3 = -26.1$$

$$m_4 = m'_4 - 4m'_1m'_3 + 6m_1'^2m'_2 - 3m_1'^4$$

$$= 8087.5 - (4)(2.1)(-382.5) + (6)(-2.1)^2(59.5) - 3(2.1)^4 = 12902.65$$

$$m_4 = 12902.65$$

نشاط



١- أوجد العزوم حول الصفر:

(أ) الأول (ب) الثاني (ج) الثالث (د) الرابع .

لمجموعة الأرقام ١٠، ٨، ٧، ٣، ٢ .

٢- أوجد العزوم حول الوسط الحسابي:

(أ) الأول (ب) الثاني (ج) الثالث (د) الرابع .

لمجموعة الأرقام بالمسألة ١ .

٣- أوجد العزوم حول النقطة (٤) :

(أ) الأول (ب) الثاني (ج) الثالث (د) الرابع .

لمجموعة الأرقام بالمسألة ١ .

٤- باستخدام نتائج المسائل ٣ و ٢ حقق العلاقة بين العزوم :

$$(أ) m_2 = m'_2 - m_1'^2$$

$$(ب) m_3 = m'_3 - 3m'_1m'_2 + 2(m_1')^3$$

$$(ج) m_4 = m'_4 - 4m'_1m'_3 + 6(m_1')^2m'_2 - 3(m_1')^4$$

٤ . حساب معامل الالتواء باستخدام العزوم

حساب معامل الالتواء باستخدام العزم الثالث حول الوسط الحسابي والعزم الثاني حول الوسط الحسابي ويرمز إليه بالرمز α_3 ، ويعرف كما يلي :

$$\alpha_3 = \frac{m_3}{\sqrt{(m_2)^3}}$$

مثال

أوجد معامل الالتواء باستخدام العزوم لبيانات إنتاج الذرة بمشروع الجزيرة.

الحل

كنا قد أوجدنا العزوم الأربعة حول الوسط الحسابي لهذا التوزيع وما علينا الآن إلا التعويض في القانون السابق لإيجاد (α_3) معامل الالتواء.

$$\alpha_3 = \frac{m_3}{(\sqrt{54.1})^3} = \frac{-26.1}{(7.4)^3} = -0.064$$

ونلاحظ أن التوزيع له التواء سالبا أي أنه ملتوي ناحية اليسار كما أن الالتواء خفيفة. عندما يكون التوزيع معتدلا تكون قيمة $\alpha_3 = 0$

٥ . حساب معامل التفرطح باستخدام العزوم

لحساب معامل التفرطح بدلالة العزوم فأنا نستخدم العزم الرابع حول الوسط الحسابي . ويرمز لمعامل التفرطح بدلالة العزم بالرمز α_4 ويحسب من العلاقة:

$$\alpha_4 = \frac{m_4}{m_2^2}$$

مثال

أوجد معامل التفرطح باستخدام العزوم لبيانات إنتاج الذرة بمشروع الجزيرة ؟

الحل

لقد أوضحنا من قبل إن العزم الرابع حول الوسط الحسابي لهذه البيانات يساوي . وعليه فإن معامل التفرطح يساوي :

$$\alpha_4 = \frac{m_4}{m_2^2} = 12902.65 / (54.1)^2 = 4.41$$

وبما أن تفرطح المنحني المعتدل يساوي (٣) فإن المنحني لبيانات إنتاج الذرة بمشروع الجزيرة يعتبر منحني مدبب تدبباً كبيراً إذ أنه يزيد كثيراً عن تفرطح التوزيع المعتدل.

نشاط



١. العزمان الثانيان حول الوسط الحسابي لتوزيعين هما ٩ ، ١٦ بينما العزمان الثالثان حول الوسط الحسابي هما ١٢.٨ ، -٨.١ على الترتيب. أي التوزيعين أكثر التواء إلى اليسار؟
٢. إذا كان العزم الأول حول الرقم (٢) هو ٥ فما هو الوسط الحسابي ؟
٣. متوالية عددية حدها الأول (a) وفرقها الثابت (d) أوجد عزمها الأول والثاني حول الوسط الحسابي.
٤. الانحراف المعياري لتوزيع متماثل هو
٥. كم يجب أن يكون العزم الرابع لكي يكون هذا التوزيع:
أ/ مدبب ب/ مفرطح ج/ متماثل

أسئلة تقويم ذاتي



- ١- ناقش باختصار مقاييس التفرطح؟
- ٢- متى يوصف المنحني بأنه مفرطح / مدبب، متوسط التفرطح؟
- ٣- تحدث باختصار عن العزوم؟
- ٤- أشرح كيفية حساب العزوم من البيانات المبوية؟
- ٥- أكتب معادلة حساب معامل الالتواء باستخدام العزوم؟
- ٦- كيف يتم حساب معامل التفرطح باستخدام العزوم؟

الخلاصة

عزيزي الدارس مرحباً بك في نهاية الوحدة وهي رابع وحدات هذا المقرر في القسم الأول تحدثنا عن مقاييس الإلتواء وذكرنا أنه إذا كانت المنحنيات معتدلة فإن مقاييس الوسط الحسابي والوسيط والمنوال تكون متطابقة، أما في حالة عدم التطابق فإنها تكون ملتوية نحو اليسار ويكون الوسط الحسابي أكبرها أو ملتوية نحو اليمين ويكون الوسط الحسابي أصغرهما. ويتم تحديد نوع الإلتواء باستخدام الفرق بين الوسط والمنوال.

تحدثنا أيضاً عن معامل الإلتواء ووضحنا مزاياه وعيوبه. ثم انتقلنا إلى مقاييس التفرطح وذكرنا أن كلمة تفرطح تستخدم لوصف تدبب المنحنى التكراري بالنسبة للمنحنى المعتدل، حيث يوصف المنحنى بأنه مفرطح أو مدبب أو متوسط التفرطح. ثم تحدثنا عن العزوم وطريقة حسابه من البيانات المبوبة وغير المبوبة، وتحدثنا أيضاً عن كيفية حساب معامل التفرطح باستخدام العزوم. مرحباً بك مرة أخرى في نهاية هذه الوحدة أرجو أن تكون موفق في استيعابها. مع خالص تمنياتنا بالنجاح والتوفيق.

لمحة مسبقة عن الوحدة التالية

عزيزي الدارس، الوحدة التالية بعنوان الارتباط والانحدار الخطي. سنقوم فيها بدراسة أساليب قياس مقدار واتجاه العلاقة بين متغيرين وقوة الارتباط بينهما ومن ثم تحديد علاقة رياضية تربط المتغيرين وتمكن من التنبؤ.

إجابات التدريبات

تدريب (١)

تكون مقاييس الوسط الحسابي والوسيط والمنوال متطابقة إذا كانت المنحنيات معتدلة (متماثلة)

تدريب (٢)

- المزايا
- يمكن إيجاده بالحساب العادي أو بالرسم البياني باستخدام المنحنى التكراري للمجتمع الصاعد. وكذلك يمكن حسابه في حالة الجداول التكرارية المفتوحة.
- العيوب
- لا يأخذ جميع القيم في الاعتبار
- لا يمكن استخدامه في التحليل الإحصائي الرياضي

تدريب (٣)

يوصف المنحنى بأنه مفرطح عندما يكون المنحنى أكثر استواءً من المنحنى المدبب أو المنحنى متوسط التفرطح.



محتويات الوحدة

الصفحة	الموضوع
١٢٨	المقدمة
١٢٨	تمهيد
١٢٩	اهداف الوحدة
١٣٠	١. الارتباط الخطي البسيط
١٣٧	٢. معامل الارتباط الخطي البسيط
١٤٤	٣. معامل ارتباط الرتب
١٤٧	٤. معامل الاقتران
١٤٩	٥. معامل التوافق
١٥٢	٦. الإنحدار الخطي البسيط
١٦٣	الخلاصة
١٦٣	لمحة مسبقة عن الوحدة التالية

المقدمة

تمهيد

عزيزي الدارس،

مرحباً بك في الوحدة الخامسة من مقرر الإحصاء الإداري، تتألف هذه الوحدة (الارتباط والانحدار الخطي) من ستة أقسام رئيسة. شرحنا في القسم الأول المفهوم الإحصائي للارتباط والعلاقة الدالية والإحصائية بين متغيرين (ظاهرتين) وشكل الانتشار بينما تناولنا في القسم الثاني معامل الارتباط الخطي البسيط من بيانات مبوبة وغير مبوبة ، تحدثنا في القسم الثالث عن معامل ارتباط الرتب وفي القسمين الرابع والخامس تناولنا معامل الاقتران ومعامل التوافق على التوالي.

وختاماً تحدثنا في القسم الأخير عن الانحدار الخطي البسيط من خلال معادلة خط الانحدار وكيفية رسم خط الانحدار والعلاقة بين معامل الارتباط ومعامل الانحدار.

عزيزي الدارس ، تحتاج في حياتك العملية إلى معرفة العلاقة بين متغيرين أو أكثر كي تتنبأ بقيمة واتجاه أحدهما إذا علمت قيمة المتغير الآخر ، كأن تعرف التحصيل الدراسي أو العلاقة بين زيادة الدخل الشهري وزيادة ساعات العمل، ولا شك أن دراسة المقاييس الإحصائية السابقة تعطيك تنبأً علمياً دقيقاً في هذا الجانب.

ستجد بين ثنايا هذه الوحدة تعاريف وأسئلة وتمارين تعينك على الفهم الجيد لهذا الموضوع. أرجو أن تستفيد من دراستك لها كما آمل منك مساهمتك في تقييم هذه الوحدة علمياً بما يدفع في زيادة قيمتها العلمية ولك من الله التوفيق.

أهداف الوحدة



عزيزي الدارس

بعد فراغك من دراسة هذه الوحدة ينبغي أن تكون قادراً على أن:

١. تكتب تعريفاً شاملاً وموجزاً لكل من المقاييس الإحصائية الواردة في محتويات هذه الوحدة.
٢. تحل (مستخدماً أي من هذه المقاييس) مسألتين من ثلاثة مسائل إحصائية حلاً صحيحاً في حدود ستة دقائق.
٣. تفرق بين استخدامات المقاييس الإحصائية الواردة في هذه الوحدة.
٤. تستخدم ثلاثة من المقاييس الإحصائية المشار إليها في هذه الوحدة في التنبؤ بعلاقات مستقبلية تخصك لهذا الشهر.
٥. تقارن العلاقة بين معامل الارتباط ومعامل الإنحدار.

١ . الارتباط الخطي البسيط

Simple Linear Correlation

عزيزي الدارس،

عند دراستنا للمقاييس الإحصائية المختلفة في الوحدة السابقة كان إهتمامنا ينصب على دراسة متغير واحد وهو المتغير تحت الدراسة . ولكن في الحياة العملية قد يصادف الباحث أوضاعاً متعددة . فقد تكون مفردات العينة محل الدراسة عبارة عن أزواج من القيم لظاهرتين مختلفتين أو أكثر ويود الباحث دراسة العلاقة بينهما .

على سبيل المثال العلاقة بين الراتب الشهري ومدة الخدمة والعلاقة بين الطول والوزن لمجموعة من الطلاب والعلاقة بين كمية السماد المستخدم والإنتاج . كل هذه الظواهر لها علاقات مألوفة وهناك الكثير غيرها .

ومن أهم أهداف الحصول على علاقة بين ظاهرتين هي الوصول إلى معادلة تمكن من التنبؤ بقيمة أحد المتغيرين إذا علمنا قيمة المتغير الآخر . وفي هذا الوحدة نقوم بدراسة أساليب قياس مقدار واتجاه العلاقة بين متغيرين وقوة الارتباط بينهما ومن ثم تجديد علاقة رياضية تربط المتغيرين وتمكن من التنبؤ .

أ. المفهوم الإحصائي للارتباط

لكي نبدأ دراسة الارتباط لابد من التفرقة بين العلاقة الدالية والعلاقة الإحصائية.

تعريف العلاقة الدالية

العلاقة الدالية بين متغيرين هي علاقة متناظرة تمثلها علاقة رياضية محددة بحيث تحدد قيم أحد المتغيرين تماماً حسب قيم المتغير الآخر .

على سبيل المثال نأخذ العلاقة الدالية بين مساحة الدائرة ونصف قطرها ويعبر عنها رياضياً بالمعادلة :

$$A = \pi r^2$$

حيث تشير A إلى مساحة الدائرة في حين تشير r إلى نصف القطر . فإذا عرف نصف قطر الدائرة أمكن الحصول على مساحتها .

العلاقة الدالية بين متغيرين

خطية تتبع خطاً مستقيماً
غير خطية تتبع خطاً منحنياً

فإذا كانت العلاقة خطية يعبر عنها بمعادلة الخط المستقيم:

$$y = a + bx$$

وتعبر تلك المعادلة عن علاقة خطية دالية بين متغيرين x و y يسمى الأول المتغير المستقل والثاني المتغير التابع . ويمكن إيجاد الخط المستقيم بمعلومية الثوابت a و b . وفي مثل هذه الحالة تقع جميع النقاط (x,y) على الخط المستقيم ويكون الارتباط بين المتغيرين ارتباطاً خطياً . ويعني ذلك أن أي زيادة أو نقصان في المتغير المستقل (x) تتبعه زيادة أو نقصان في المتغير التابع (y) بمعدل ثابت هو ميل الخط b في معادلة الخط المستقيم .

العلاقة بين متغيرين

عكسية ، ميل
الخط سالباً
(الزيادة في

طرديّة ميل الخط
المستقيم موجباً
(الزيادة في

أسئلة تقويم ذاتي



١. عرف العلاقة الدالية بين متغيرين؟

٢. أكمل الجمل التالية:

أ. العلاقة الدالية بين متغيرين يمكن أن تكون..... وتتبع خطأً مستقيماً كما يمكن أن تكون وتتبع خطأً منحنياً.

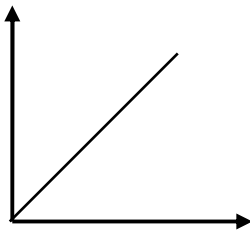
ب. العلاقة بين متغيرين يمكن أن تكون..... حيث يكون ميل الخط موجباً، والزيادة في قيمة X يبتعها في قيمة Y . كما يمكن أن تكون..... حيث يكون ميل الخط سالباً، والزيادة في قيمة X يبتعها في قيمة Y .

ويمكن التعبير عن معادلة الخط المستقيم بالرسم شكل رقم (١) وفيه تخصص محور الأحداثيات الأفقية لقيم x والإحداثيات الرأسية لقيم y المقابلة لها . فإذا كان الخط المستقيم ينطلق من أصل الأحداثيات الرأسية والأفقية يكون ميله يساوي واحد ويصبح الخط المستقيم $y = x$ شكل رقم (١-أ) وتسمى المعادلة معادلة الخط المستقيم المار بنقطة الأصل.

أما إذا كان الخط المستقيم لا يمر بأصل الأحداثيات بل يقطع محور الأحداثيات الرأسية في النقطة (a) عندما تساوي قيمة (x) صفراً فيصبح الخط المستقيم $y = a + bx$ شكل رقم (١-ب)

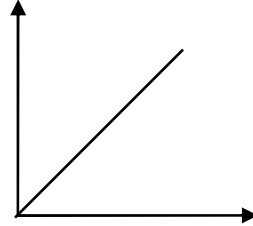
$$y = a + bx$$

[١-ب]



$$y = x$$

[١-أ]



شكل رقم (١)

إن العلاقة الخطية الدالية هي علاقة مثالية بحيث أن العلاقة بين متغيرين قد لا تكون بالمعنى التام ولكن توجد علاقة من درجة ما بين المتغيرين . وفي هذه الحالة تسمى العلاقة علاقة إحصائية . والعلاقة الإحصائية بين المتغيرين يعبر عنها بالإرتباط بينهما . على سبيل المثال نفترض أننا بصدد إيجاد العلاقة بين متوسطات الدرجات في مادتي الرياضيات X والإحصاء Y في أحد الفصول الدراسية والتي كان توزيعها كما يلي :

جدول (١)

الرياضيات	٦٥	٦٣	٦٧	٦٤	٦٨	٦٢	٧٠	٦٦	٦٨	٦٧	٦٩	٧١
الإحصاء	٦٨	٦٦	٦٨	٦٥	٦٩	٦٦	٦٨	٦٥	٧١	٦٧	٦٨	٧٠

وبنظرة سريعة إلى هذه الدرجات يظهر أن هناك علاقة بين درجات الطلاب في هاتين المادتين بإعتبار أن الدرجة إذا كانت مرتفعة في الرياضيات كانت أيضاً مرتفعة في الإحصاء . ولكن هذه العلاقة ليست علاقة سببية بمعنى أنه إذا كانت درجة الطالب في الرياضيات مرتفعة يصبح من الضروري إرتفاع درجته في الإحصاء ، بل هذه العلاقة إحصائية بحيث أن هناك إتجاه عام لإرتفاع درجات الإحصاء مع إرتفاع درجات الرياضيات .

نشاط

البيانات التالية تمثل حركة القروض والودائع بملايين الدينارات في أحد البنوك التجارية:

القروض (X)	٧٥	٨٩	٩٥	٨٤	٧٢
الودائع (Y)	٩٢	١١٢	١١٦	٩٠	٨٥

جد معادلتني خطي إنحدار (Y/X) و (X/Y) ثم استخدم النتائج لحساب معامل الارتباط بين المتغيرين X و Y .

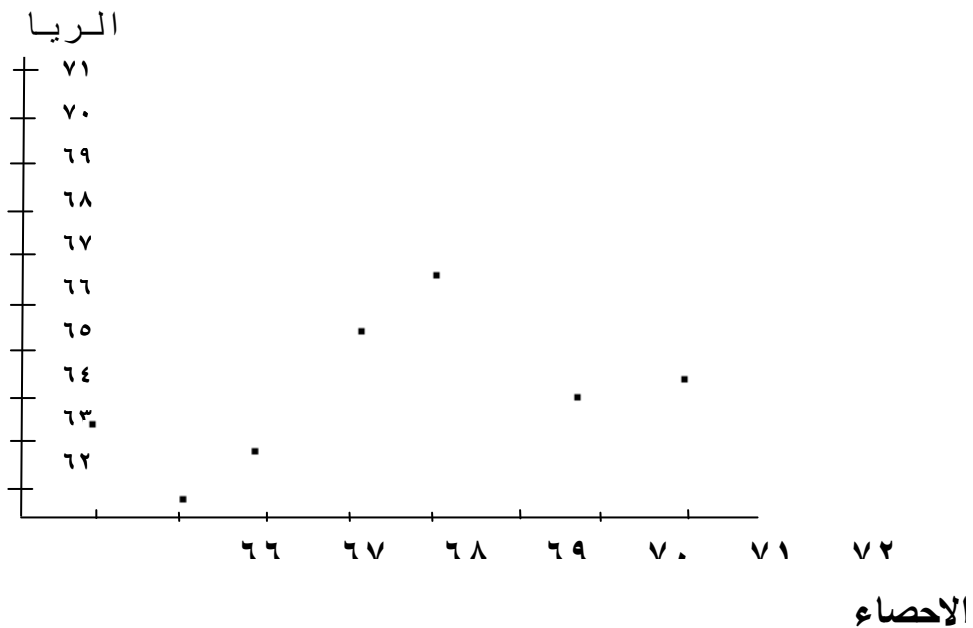


ب - شكل الإنتشار : Scatter Diagram

يمكن بيانياً رصد جميع أزواج الدرجات في المثال السابق بحيث يخصص لكل زوج منهما نقطة بالنسبة إلى محورين متعامدين يمثل أحدهما درجات الرياضيات بينما يمثل المحور الثاني درجات الإحصاء فنحصل على ما يسمى بشكل الإنتشار شكل رقم (٢) .

شكل رقم (٢) ' ١

شكل الإنتشار للعلاقة بين درجات الطلاب في الرياضيات والإحصاء

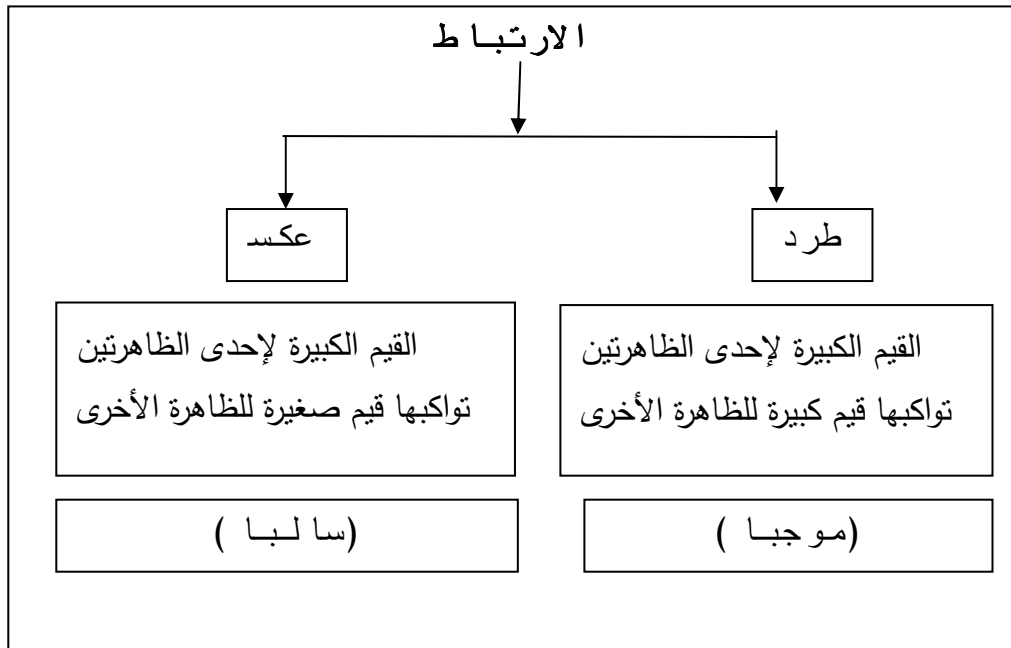


ونلاحظ من الشكل أن النقاط تنتشر حول محور ولذلك يبدو من المعقول وصف العلاقة بين درجات الرياضيات والإحصاء بالارتباط ومن ثم التعبير عنها بخط مستقيم يمر ببعض النقاط ويوسط معظمها وتمثله معادله جبرية من النوع :

$$y = a + bx$$

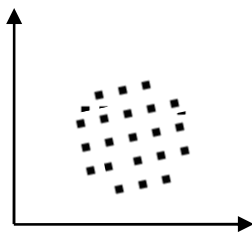
وسنتعرض لهذه المعادلة بالتفصيل في الأقسام القادمة . يهمننا الآن أن نبين أن شكل الإنتشار يمكن من تكوين فكرة مبدئية عن العلاقة بين متغيرين . فمن ناحية قد تتجمع النقاط حول إتجاه معين كما رأينا في هذا المثال . بحيث يقال أن المتغيرين بينهما إرتباط من درجة ما . وفي بعض الحالات قد تكون هذه النقاط مبعثرة بلا إتجاه محدد وبالتالي ليس هناك إرتباط بين المتغيرين .

والارتباط بين متغيرين قد يكون طردياً (موجباً) أو عكسياً (سالباً) . الشكل يبين أنواع الإرتباط .

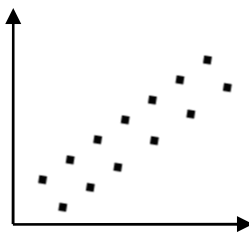


شكل (٣)

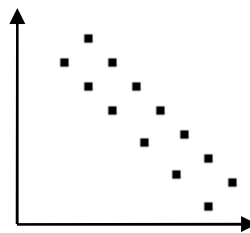
يبين أنواع الإرتباط



[ج] لا يوجد إرتباط



[ب] إرتباط طردي



[أ] إرتباط عكسي

والارتباط إما أن يكون قوياً أو ضعيفاً . وإحصائياً يقاس الارتباط بمقياس رقمي يتراوح بين [-1] و [+1] بحيث تحدد الإشارة الجبرية اتجاه الارتباط موجباً أو سالباً بينما تعكس القيمة العددية درجة الارتباط وقوته . والمقياس الذي يوضح درجة واتجاه الارتباط يسمى بمعامل الارتباط . وسنكتفي هنا بدراسة معامل الارتباط الخطي . وهناك أكثر من طريقة لحساب معامل الارتباط الخطي والتي تختلف باختلاف البيانات المتعلقة بالظاهرتين تحت الدراسة وسنقوم الآن بشرح أهم هذه المعاملات .

أسئلة تقويم ذاتي



١. أرسم معادلة الخط المستقيم في الحالات التالية:
أ. عندما يمر بنقطة الأصل
ب. عندما لا يمر بنقطة الأصل
٢. أشرح كيف يتم تكوين شكل الانتشار؟
٣. بين أنواع الارتباط وكيف يتم قياسه؟

٢. معامل الارتباط الخطي البسيط

Simple Linear Correlation

ويطلق عليه معامل ارتباط بيرسون *Pearson* ، بالنسبة للمتغيرين x ، y يحسب ارتباط بيرسون على أنه متوسط مجموع حواصل ضرب القيم المعيارية للمتغيرين ويرمز له بالرمز r ، ويعبر عنه رياضياً كما يلي :

$$r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{n\sigma_x\sigma_y}$$

وهذا المقياس يعبر عن كل خصائص الارتباط الخطي البسيط من ناحية القوة والاتجاه فهو عبارة عن مجموع ضرب الانحرافات للمتغيرين x ، y عن وسطهما الحسابي مقسوماً على انحرافهما المعياري . فعندما يكون انحراف x سالباً وانحراف y سالب يكون الارتباط موجب أيضاً . أما إذا كان أحد الانحرافات موجب والثاني سالب فيكون الارتباط سالب .

للأغراض الحسابية يمكن تحويل المعادلة السابقة إلى عدة أشكال :

$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

والمعادلة السابقة هي أكثر المعادلات إستخداماً ذلك أنها تستخدم القيم الأصلية للمتغيرين x ، y دون اللجوء إلى حساب الانحرافات والوسط الحسابي والانحراف المعياري للمتغيرات . نوضح في الجدول التالي خطوات حساب معامل الارتباط البسيط بين درجات الرياضيات والإحصاء الواردة في الجدول رقم (١) .

جدول رقم (٢)

خطوات حساب معامل الارتباط البسيط بين درجات الرياضيات والإحصاء

الرياضيا ت X	الإحصاء ت Y	X	Y	XY	X^2	Y^2
٦٥	٦٨	٥	٨	٤٠	٢٥	٦٤
٦٣	٦٦	٣	٦	١٨	٩	٣٦
٦٧	٦٨	٧	٨	٥٦	٤٩	٦٤
٦٤	٦٥	٤	٥	٢٠	١٦	٢٥
٦٨	٦٩	٨	٩	٧٢	٦٤	٨١

٣٦	٤	١٢	٦	٢	٦٦	٦٢
٦٤	١٠٠	٨٠	٨	١٠	٦٨	٧٠
٢٥	٣٦	٣٠	٥	٩	٦٥	٦٦
١٢١	٦٤	٨٨	١١	٨	٧١	٦٨
٤٩	٤٩	٤٩	٧	٧	٦٧	٦٧
٦٤	٨١	٧٢	٨	٩	٦٨	٦٩
١٠٠	١٢١	١١٠	١٠	١١	٧٠	٧١
$\sum Y^2 = 729$	$\sum X^2 = 618$	$\sum XY = 647$	$\sum Y = 91$	$\sum X = 80$		

$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

بالتعويض في المعادلة أعلاه نجد أن معامل الارتباط الخطي البسيط :

$$r = \frac{(12)(657) - (80)(91)}{\sqrt{\{(12)(618) - (80)^2\} \{(12)(729) - (91)^2\}}}$$

نلاحظ أننا قمنا بطرح مقدار ثابت من كل درجة من الإحصاء والرياضيات لتسهيل العمليات الحسابية ، وذلك لا يؤثر على معامل الارتباط حيث أنه لا يعتمد على إختيار نقطة الأصل .

ونلاحظ أن قيمة r أعلى من ٠.٥ مما يدل على أن هنالك علاقة طردية موجبة ذات إرتباط متوسط . وتجدر الإشارة إلى أنه لا توجد حدود فاصلة تحدد قوة وضعف الارتباط ولكن من الخبرة يمكن تحديد ذلك .

ومن أهم خصائص معامل الارتباط الخطي لبيرسون أنه لا يتأثر بإضافة أو طرح أي مقدار ثابت بالنسبة لجميع قيم الظاهرتين . كما يتمتع معامل الارتباط بهذه الخاصية في حالة الضرب والقسمة . ويمكن استخدام هذه الخواص في حساب معامل الارتباط كما بينا في المثال السابق . ومن الإحصائيات

المتوفرة هناك حدود تقريبية لقيم r يتم العمل بها ، انظر جدول رقم (٣) ، وتتغير الإشارة عندما تكون r سالبة .

جدول رقم (٣)

درجة قوة الارتباط *

قيمة r	درجة قوة الارتباط
<0.3	لا يوجد ارتباط يذكر
$0.3 - 0.49$	ارتباط ضعيف
$0.5 - 0.69$	ارتباط متوسط
$0.7 - 0.89$	ارتباط قوي
$0.9 - 1.00$	ارتباط قوي جداً

*ويمكن أن تكون هذه القيم سالبة في حالة الارتباط العكسي .

معامل الارتباط الخطي لبيرسون من البيانات المبوبة :

في هذه الحالة ندخل التكرارات f على معادلة معامل الارتباط الخطي البسيط لتأخذ الشكل :

$$r = \frac{N \sum f_{xy} u_x u_y - (\sum f_x u_x)(\sum f_y u_y)}{\sqrt{\{N \sum f_x u_x^2 - (\sum f_x u_x)^2\} \{n \sum f_y u_y^2 - (\sum f_y u_y)^2\}}}$$

حيث :

u_x = انحرافات مراكز فئات X عن وسط فرضي .

u_y = انحرافات مراكز فئات Y عن وسط فرضي .

N = مجموع التكرارات .

f_x = التكرارات (الهامشية) المقابلة للفئة x .

f_y = التكرارات (الهامشية) المقابلة للفئة y .

$$\begin{aligned}
f_x u_x &= \text{تكرار الفئة } x \text{ مضروباً في الانحراف المقابل للمتغير } x . \\
f_y u_y &= \text{تكرار الفئة } y \text{ مضروباً في الانحراف المقابل للمتغير } y . \\
f_{xy} u_x u_y &= \text{حاصل ضرب إنحرافي } x \text{ و } y \text{ في التكرار المشترك لفئتيهما.}
\end{aligned}$$

مثال (١)

أوجد معامل الارتباط الخطي بين درجات الرياضيات y ودرجات الإحصاء x لمجموعة مكونة من ٣٠ طالب في الجدول التالي :

جدول رقم (٤)

حساب معامل الارتباط الخطي

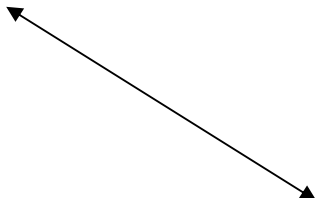
المجموع	٩٩-٩٠	-٨٠	-٧٠	-٦٠	-٥٠	$y \backslash x$
٣					٣	-٥٠
٥				٤	١	-٦٠
١٠				٨	٢	-٧٠
٨	١	٦	١			-٨٠
٤	٤					٩٩-٩٠
١٠٠	٥	٦	١	١٢	٦	المجموع

وفي الجدول يمثل العمود الأول فئات المتغير x والسطر الأول فئات المتغير y . أما الأرقام في الخلايا داخل الجدول فتتمثل التكرارات المشتركة للفئتين المقابلتين . فمثلاً الرقم (٣) في الخلية التي تقع في تقاطع الصف الثاني والعمود الثاني يمثل عدد الوحدات التي تتراوح قيمة x فيها بين ٥٠ إلى أقل من ٦٠ وقيمة y بين ٥٠ إلى أقل من ٦٠ ... وهكذا .

ولحساب معامل الارتباط الخطي لبيرسون نقوم بتنظيم خطوات الحل في جدول (٥-٦) على النحو التالي :

جدول رقم (5)

حساب معامل الارتباط الخطي من بيانات مبوبة

$f_{xy}u_xu_y$	$f_xu_x^2$	f_xu_x	u_x	f_x	-٩٤.٥	-٨٤.٥	-٧٤.٥	٦٤. -٥	-٥٤.٥	Y X
١٢	١٢	-٦	-٢	٣					123	-٥٤.٥
٦	٥	-٥	-١	٥				4ε	21	-٦٤.٥
٠	٠	٠	٠	١٠				08	0٢	-٧٤.٥
٨	٨	٨	١	٨	21	66	01			-٨٤.٥
١٦	١٦	٨	٢	٤	16ε					-٩٤.٥
٤٢	٤١	٥		٣٠	٥	٦	١	١٢	٦	f_y
					٢	١	٠	-١	-٢	u_y
				-٨	١٠	٦	٠	-١٢	-١٢	f_yu_y
				٦٢	٢٠	٦	٠	١٢	٢٤	$f_yu_y^2$
				٤٢	١٨	٦	٠	٤	١٤	$f_{xy}u_xu_y$

ونلاحظ في الجدول أن القيم المحصورة بين المربعات داخل الخلايا تمثل التكرار المشترك للفئتين المقابلتين مضروباً في حاصل ضرب إنحرافيهما عن الوسط الفرضي. فمثلاً العدد ١٢ في العمود الأول تم حسابه بضرب التكرار المشترك للفئتين الأولتين X و Y وهو (٣) في إنحرافيهما وهما -٢ و -٢ .

$$f_{xy}u_xu_y = -2(-2)(3) = 12$$

ويحتوي الجدول على الأعمدة التالية :

$u_x =$ الانحرافات بطريقة الترميز لمراكز الفئات بالنسبة للمتغير X ، حيث نضع (٠) مقابل أي فئة (يفضل الوسطي) ثم -١ ، -٢ ، ... للفئات قبلها بالترتيب ، و +١ ، +٢ ، ... للفئات بعدها .
فمثلاً في حالة X وضعنا الصفر مقابل الفئة ٧٤.٥ - ٨٤.٥ مما نتج عنه أن الانحراف المقابل للفئة ٥٤.٥ - فمثلاً ، -٢ .

$u_y =$ الانحرافات بطريقة الترميز لمراكز الفئات بالنسبة للمتغير y .

باقي الأعمدة تم حسابها مباشرة عن طريق ضرب التكرارات المشتركة لكل فئتين متقابلتين في انحرافاتها .
وقد تم حساب $\sum f_x u_x u_y$ أو $\sum f_y u_x u_y$ بواسطة جمع الأرقام الموجودة في المربعات الجانبية بكل خلية بالنسبة لكل صف ولكل عمود . ونلاحظ أن المجاميع الكلية في الصف الأخير والعمود الأخير متساوية وهذا أمر مهم بالنسبة لمراجعة حساب الارتباط .
ومن الجدول نحصل على :

$$r = \frac{N \sum f_{xy} u_x u_y - (\sum f_x u_x)(\sum f_y u_y)}{\sqrt{\{N \sum f_x u_x^2 - (\sum f_x u_x)^2\} \{n \sum f_y u_y^2 - (\sum f_y u_y)^2\}}}$$

$$r = \frac{(30)(42) - (-8)(5)}{\sqrt{\{(30)(62) - (-8)^2\} \{(30)(41) - (5)^2\}}}$$

$$r = \frac{1300}{1471.1} = 0.884$$

أي أن الارتباط طردي قوي بين درجات الرياضيات و الإحصاء .



الجدول التالي يمثل تكلفة السماد وعائد الإنتاج بالآف الدينارات
السودانية:

تكلفة X(السماد)	٩	١٠	٩	١٠	٤	٥	٧
المبيعات Y()	١٦٠	١٨٠	١٥٠	١٤٠	١٢٠	١٢٠	١٤٠

أ- أرسم شكل الانتشار للمتغيرين X و Y.

٣. معامل ارتباط الرتب Rank Correlation

في الحالات التي يكون فيها أحد المتغيرين أو كليهما ترتيبياً ، فإننا لا نستطيع استخدام معامل الارتباط الخطي البسيط . لذلك نشأت الحاجة إلى إيجاد مقياس يستخدم في مثل هذه الحالة ، ومن أمثلة البيانات الترتيبية ترتيب عدد من العاملين حسب درجة رضائهم الوظيفي أو حماسهم للعمل ويعتمد معامل ارتباط الرتب لسيبرمان على فروق الرتب المتناظرة بين قيم المتغيرين ، والذي يحسب من المعادلة :

$$r = 1 - \left[\frac{6 \sum d^2}{n(n^2 - 1)} \right]$$

حيث:

d = فروق الرتب بين قيم المتغيرين ، n = عدد أزواج القيم.

والبيانات قد تكون في شكل رتب أصلاً ، أو قد تكون قيماً تحتاج لترتيب . في الحالة الأخيرة تعطى رتب لقيم كل من المتغيرين على حدة حيث نأخذ أصغر القيم للمتغير الرتبة (١) والتي تليها في الكبر الرتبة (٢) وهكذا . وكذلك ترتب قيم المتغير الآخر . وإذا تساوت قيمتان أو أكثر عند ترتيب قيم متغير نعطي كل منهما متوسط رتبها إذا لم تكن متساوية .

فمثلاً إذا تساوت قيم كل المفردات التي كان سيكون ترتيبها لو أنها مختلفة ٦ ، ٥ ، ٤ فإن كل منها تعطى الرتبة ٥ . وبالمثل إذا تساوت قيمتا المفردتين اللتان ترتيبهما ٨ ، ٧ فتعطى كل منهما الرتبة ٧.٥ وهكذا . ثم نكوّن جدولاً نضع فيه قيم المتغيرين ورتب كل منهما ثم نحسب الفروق بين الرتب ثم مربعاتها ونعوض في المعادلة الأخيرة لنحصل على معامل الارتباط .

مثال (٢)

حصل ١٠ طلاب على التقديرات الآتية في مادتي الإحصاء والرياضيات .

جدول رقم (٦)

رقم الطالب	الإحصاء	الرياضيات
١	C	B
٢	F	C
٣	+C	C
٤	A	B
٥	C	C
٦	D	C+
٧	B	C
٨	+C	A
٩	D	F
١٠	C	D

أحسب معامل الارتباط .

ونبين خطوات الحل في الجدول التالي .

جدول رقم (٧)

حساب معامل ارتباط الرتب

رقم الطالب	الإحصاء Y	الرياضيات X	رتب X	رتب Y	فروق الرتب d	d^2
١	C	B	٦	٢.٥	٣.٥	١٢.٢٥
٢	F	C	١٠	٦.٥	٣.٥	١٢.٢٥
٣	C+	C	٣.٥	٦.٥	٣.٠	٩
٤	A	B	١	٢.٥	١.٥	٢.٢٥
٥	C	C	٦	٦.٥	٠.٥	٠.٢٥
٦	D	C+	٨.٥	٤	٤.٥	٢٠.٢٥
٧	B	C	٢	٦.٥	٤.٥	٢٠.٢٥
٨	C+	A	٣.٥	١	٢.٥	٦.٢٥
٩	D	F	٨.٥	١٠	١.٥	٢.٢٥
١٠	C	D	٦	٩	٣.٠	٩
المجموع						٩٤

من هذا الجدول نجد أن معامل الارتباط :

$$r = 1 - \left[\frac{6 \sum d^2}{n(n^2 - 1)} \right]$$

$$r = 1 - \left[\frac{6(94)}{10(100 - 1)} \right]$$

$$r = 1 - \frac{564}{999} = 0.43$$

ويلاحظ أنه وفي هذا المثال الارتباط طردي ضعيف بين درجات الطلاب في الإحصاء والرياضيات .
ومن أهم ما يمتاز به هذا القياس للارتباط أنه لا يصلح فقط للبيانات الترتيبية وإنما يمكن استخدامه أيضاً لقياس الارتباط في حالة البيانات الرقمية . إلا أنه يؤخذ عليه أنه أقل دقة بالمقارنة بمعامل ارتباط بيرسون في هذه الحالة.

أسئلة تقويم ذاتي



١. عرف معامل ارتباط بيرسون وأكتب معادلته؟
٢. ما هي أهم خصائص معامل ارتباط بيرسون؟
٣. على ماذا يعتمد معامل ارتباط الرتب لسبيرمان وكيف يتم حسابه؟

نشاط



أحسب معامل ارتباط سبيرمان للرتب من واقع البيانات الواردة في الجدول التالي والذي يمثل تقديرات مجموعة من الطلاب في مادتين:

D	F	B	F	D	A	D	تقديرات المادة الأولى (X)
C	D	A	D	B	C	D	تقديرات المادة

٤. معامل الإقتران Contingency Coefficient

في بعض الأحيان قد تصادفنا بيانات وصفية ليس في طبيعتها صفة للترتيب أي لا يمكن وضع رتب لها أو لا معنى للرتب . فإذا أخذنا مثلاً فصائل الدم نجد الفصائل A ، B وغيرها أو المدخنين فنجد أن التقسيم يمكن أن يكون يدخن أم لا . وهناك الكثير من الأمثلة التي يكون للمتغيرين فيها صفتان أولي وثانية .

ويمكن التعبير عن جدول الإقتران بالجدول التالي .

جدول (٦)

إقتران بين الصفات

X_2	X_1	X Y
B	A	Y_1
D	C	Y_2

ونلاحظ أن A هي عبارة عن التكرارات المشتركة في الصفة الأولى X والصفة الثانية Y وهكذا بالنسبة لبقية خلايا الجدول . ويستخدم الإحصائيون في قياس معامل الإقتران ما يعرف بمعامل ييل (Yule) للإقتران والذي يأخذ الشكل التالي :

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

مثال (٣)

أوجد معامل إقتران بين لون العيون ولون الشعر لمجموعة من الأفراد من البيانات في جدول (٩-٦)

جدول (٧)

الإقتران بين لون الشعر ولون العيون

لون العيون x	أزرق	عسلي
لون الشعر y	أشقر	أسود
٨	٦	٥

من هذا الجدول يمكن حساب معامل الإقتران مباشرة بتطبيق القانون السابق :

$$R_C = \frac{AD - BC}{BC + AD} = \frac{40 - 42}{42 + 40} = \frac{-2}{82} = -0.024$$

ونلاحظ أنه لا يوجد إرتباط بين لون الشعر ولون العيون إذ أن معامل الإرتباط قريب جداً من الصفر .

نشاط



في تجربة لدراسة أثر التدخين في الإصابة بمرض السرطان كانت النتائج كما يلي:

التكرارات المشتركة للتدخين والإصابة بالسرطان

الإصابة التدخين	أصيب بالمرض	لم يصب بالمرض
يدخن	١٠	٩٠
لا يدخن	٤	٩٤

جد مدي علاقة التدخين بالإصابة بالسرطان .

٥. معامل التوافق

في حالة تعدد صفات المتغيرات لا يمكن تطبيق قانون بيل . على سبيل المثال إذا كانت البيانات الخاصة بالعيون تحتوي على أكثر من صفتين (أسود ، أزرق ، عسلي ،.....) أو لون الشعر يحتوي أكثر من صفتين أيضاً (أشقر ، بني ، أسود) في مثل هذه الحالة لابد وأن يشتمل جدول الإقتران على جميع الصفات الواردة بالنسبة للمتغيرين في جدول مبسط كما يلي :

جدول إقتران بين عدد من الصفات

المتغير x \ المتغير y	X_1	X_2	X_3	المجموع
Y_1	C_{11}	C_{12}	C_{13}	C_1
Y_2	C_{21}	C_{22}	C_{23}	C_2
Y_3	C_{31}	C_{32}	C_{33}	C_3
المجموع	$\sum C_1$	$\sum C_2$	$\sum C_3$	$\sum C$

جدول رقم (٨)

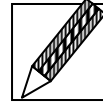
حيث $j=1,2,3,\dots$ $i=1,2,3,\dots$

وفي هذه الحالة يعرف معامل الإقتران بمعامل التوافق ويرمز له بالرمز R ويمكن حساب قيمته باستخدام العلاقة التالية :

$$R_c = \sqrt{\frac{C-1}{C}}$$

حيث أن $C = \sum C_1$

تدريب (١)



أوجد معامل التوافق بين الطول والوزن لعينة مكونة من ٤٥ طالباً بإحدى الجامعات باستخدام البيانات التالية .

X الطول	متوسط	نحيف	المجموع	Y الوزن
طويل	٧	٣	١٥	
متوسط	٦	٦	١٥	
قصير	٦	٤	١٥	
المجموع	١٧	١٥	٤٥	

لإيجاد معامل الإقتران نحسب أولاً قيمة C كالآتي :

$$\sum C_1 = \frac{(7)^2}{15 \times 17} + \frac{(5)^2}{15 \times 13} + \frac{(3)^2}{15 \times 15} = 0.360$$

$$\sum C_2 = \frac{(6)^2}{15 \times 17} + \frac{(3)^2}{15 \times 13} + \frac{(6)^2}{15 \times 15} = 0.347$$

$$\sum C_3 = \frac{(4)^2}{15 \times 17} + \frac{(5)^2}{15 \times 13} + \frac{(6)^2}{15 \times 15} = 0.347$$

$$C = \sum C_1 + \sum C_2 + \sum C_3 = 1.058$$

$$R_C = \sqrt{\frac{C-1}{C}} = \sqrt{\frac{1.058-1}{1.058}} = 0.23$$

لا يوجد ارتباط يذكر بين أطوال وأوزان الطلاب .

٦. الإنحدار الخطي البسيط

Simple Linear Regression

كما ذكرنا فيما سبق أنه من أهم أهداف البحث العلمي الحصول على علاقة بين المتغيرات تمكن من التنبؤ في المستقبل . وذلك باستنباط قيم أحد المتغيرين بدلالة المتغير الآخر . وأوضحنا من توزيع النقاط بشكل الانتشار أن العلاقة الإحصائية بين ظاهرتين . ربما تأخذ شكل الخط المستقيم ويطلق عليها الإنحدار الخطي البسيط .

وتشير معادلة خط الإنحدار إلى أنحدار أحد المتغيرين وليكن Y ويسمى متغير تابع على المتغير الآخر وليكن X ويسمى متغير مستقل ، وتسمى العلاقة عندئذ خط إنحدار Y على X ويرمز له بـ $E(Y/X)$ وكذلك معادلة خط إنحدار المتغير X على المتغير Y ويرمز له بـ $E(X/Y)$.

• معادلة خط إنحدار Y على X :

تكتب هذه المعادلة على الصورة الآتية :

$$E(Y / x) = a + \beta x$$

حيث a, β ثابتان ، ويتحدد الإنحدار متى أمكن تحديدهما و في نموذج الانحدار البسيط لـ Y على X نفترض أن أي قيمة لـ Y ، مثلاً y ترتبط بقيمة معينة لـ X ، مثلاً x ، بالمعادلة :

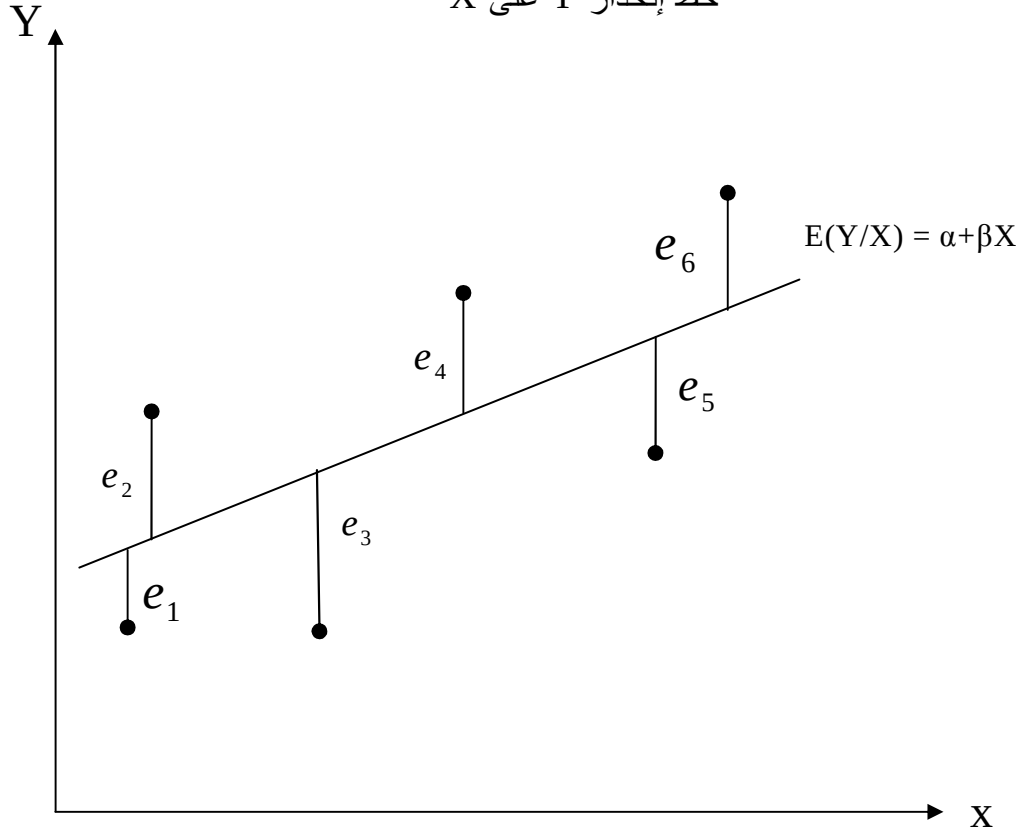
$$y = a + \beta x + e$$

حيث e متغير عشوائي .

ويمكن توضيح هذه الصورة بالشكل التالي :

شكل رقم (٥)

خط إنحدار Y على X



في الشكل النقاط تمثل ٦ أزواج مرتبة (x_1, y_1) , (x_2, y_2) , ..., (x_6, y_6) من المشاهدات . أما الخط المستقيم فيمثل خط انحدار Y على X . أما الخطوط الرأسية فتتمثل الأخطاء العشوائية $e_1, e_2, \dots, e_6$ أي انحرافات القيم الفعلية لـ y عن القيم المقابلة بالخط .
نلاحظ أن الخط المستقيم يمر ببعض النقاط ويتوسط معظمها والمقصود منه تمثيل العلاقة الخطية بين المتغيرين أصدق تمثيل. والسؤال هو كيف يتم رسم مثل هذا الخط ؟

أ- التمهيد باليد :

وهنا يتم تمهيد الخط المستقيم المناسب لنقاط شكل الانتشار للبيانات الخاصة بالمتغيرين X ، Y .
ويعاب على هذه الطريقة أنها تقريبية جداً وتخضع للتقدير الشخصي . ولهذا كان لابد من إيجاد طريقة لتوفيق خط الانحدار بحيث لا تعتمد على إنطباع الأشخاص وإنما على البيانات الخاصة بالمتغيرين Y و X .

ب - طريقة المجموعات :

وفي هذه الطريقة يتم قسمة البيانات الخاصة بالمتغيرين X ، Y إلى مجموعتين ثم يؤخذ المتوسط لمفردات التنغير Y في المجموعة الأولى ويسمى $\bar{Y}_1$ والمتوسط لمفردات Y في المجموعة الثانية ويسمى $\bar{Y}_2$ وتطبق نفس هذه الطريقة على المتغير X لنحصل على $\bar{X}_1, \bar{X}_2$ وبهذه الطريقة يكون لدينا مجموعتان إحداثياتهما $(\bar{X}_1, \bar{Y}_1), (\bar{X}_2, \bar{Y}_2)$ ، وعليه يمكن حساب الثابت β الذي يمثل ميل الخط المستقيم من العلاقة :

$$\beta = \frac{\bar{Y}_2 - \bar{Y}_1}{\bar{X}_2 - \bar{X}_1}$$

وبالتعويض في إحدى المعادلتين يمكن إيجاد الثابت a .

مثال (٥)

جد خطأ مستقيماً يمثل البيانات التالية .

١٠	٩	٧	٥	٣	٢	X
١٧	١٥	١١	٧	٣	١	Y

الحل

أولاً: نقسم قيم X إلى مجموعتين هما $(٥, ٣, ٢)$ و $(١٠, ٩, ٧)$.

ثم نقسم قيم Y إلى مجموعتين أيضاً هما $(٧, ٣, ١)$ و $(١٥, ١١, ١٧)$.

ثانياً : نحسب الوسط الحسابي لكل مجموعة من المجموعات الأربعة كما يلي :

$$\begin{aligned}\bar{X}_1 &= \frac{2+3+5}{3} = 3.33 \\ \bar{X}_2 &= \frac{7+9+10}{3} = 8.67 \\ \bar{Y}_1 &= \frac{1+3+7}{3} = 3.67 \\ \bar{Y}_2 &= \frac{17+15+11}{3} = 14.33\end{aligned}$$

ثالثاً : الإحداثيات هي : $(\bar{X}_1, \bar{Y}_1) = (3.33, 3.67), (\bar{X}_2, \bar{Y}_2) = (8.67, 14.33)$ وعليه فإن ميل الخط الخط β يساوي :

$$\beta = \frac{\bar{Y}_2 - \bar{Y}_1}{\bar{X}_2 - \bar{X}_1} = \frac{14.33 - 3.67}{8.67 - 3.33} = 2.00$$

وبتعويض ميل المعادلة مع أي نقطتين X, Y نجد أن $a = -3$ والمعادلة المطلوبة هي

$$\boxed{Y = 2X - 3}$$

وهذه الطريقة وإن كانت تعتمد على بيانات المتغيرات الأصلية إلا أنها تعتمد على الأوساط الحسابية للنقاط على طرفي التوزيع وهي تعطي نتائج غير ملائمة في حالة التوزيعات المتطرفة أو إذا كانت المفردات لا تقبل القسمة على ٢ . لتلافي مثل هذه الأمور فإننا نلجأ إلى صيغة رياضية لخط الإنحدار تسمى بطريقة المربعات الصغرى .

ج - طريقة المربعات الصغرى :

إذا نظرنا لشكل الإنتشار في شكل رقم (٥-٦) نجد أن المسافة بين أي نقطة والخط المستقيم هي $e_i = D_i$ حيث $(i = 1, 2, \dots, n)$ وتسمى هذه القيم بالانحرافات ونسبة لأن هذه الانحرافات تقع على الخط مباشرة أو على جانبيه فهي إما أن تكون سالبة أو موجبة أو قيمتها صفر، وعليه يمكن الإنطلاق من الافتراض أن مجموع هذه الانحرافات صغير للغاية .

وأساس طريقة المربعات الصغرى هو اعتبار الخط الثاني الذي يطابق النقاط أحسن مطابقة وهو الخط الذي يكون فيه مجموع مربعات انحرافات النقط عنه أصغر ما يمكن . فإذا كان الخط أو الانحراف في تمثيل النقطة (X_i, Y_i) عن خط الانحدار هو e_i فإن :

$$e_i = Y_i - \alpha - \beta X_i$$

وعليه فإن مجموع مربعات الانحرافات هو :

$$\sum e_i^2 = \sum (Y_i - \alpha - \beta X_i)^2$$

ولكي يكون $(\sum e_i^2)$ نهاية صغرى بمعنى أن الأبعاد الرأسية للنقطة عن الخط المقترح أصغر ما يمكن فإننا نفاضل $(\sum e_i^2)$ بالنسبة لكل من α و β ونجعل نتيجة كل منهما تساوي الصفر فنحصل على المعادلتين التاليتين :

$$\sum Y_i = n\alpha - \beta \sum X_i$$

$$\sum X_i Y_i = \alpha \sum X_i + \beta \sum X_i^2$$

حيث n عدد أزواج القيم المعلومة .

وتسمى هذه المعادلات بالمعادلات المعتادة ، ويمكن حل المعادلات المعتادة جبرياً لنحصل على مقدرين لقيمتي الثابتين α و β كما يلي :

$$\hat{\alpha} = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n \sum X^2 - (\sum X)^2} = \frac{\sum Y}{n} - \hat{\beta} \frac{\sum X}{n}$$

$$\hat{\beta} = \frac{n \sum XY - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2}$$

ويسمى الثابت $\hat{\beta}$ عادة معامل إنحدار Y على X والثابت $\hat{\alpha}$ هو الجزء الذي يقطعه خط إنحدار Y على X من المحور Y.

مثال (٦)

أوجد معادلة إنحدار Y الإنفاق على X الدخل للبيانات التالية ، ومن ثم جد تقدير الإنفاق عندما يكون الدخل ٢٥٠٠ دينار سوداني .

الحل

نقوم بتفصيل الحل في الجدول التالي :

جدول رقم (٩)

معادلة الإنحدار (Y) الإنفاق على (X) الدخل

الدخل X	الإنفاق Y	XY	X^2	Y^2
٨	٨	٦٤	٦٤	٦٤
١٠	٩	٩٠	١٠٠	٨١
١٢	١٢	١٤٤	١٤٤	١٤٤
١٢	١٠	١٢٠	١٤٤	١٠٠
١٣	١٠	١٣٠	١٦٩	١٠٠
١٥	١٣	١٩٥	٢٢٥	١٦٩
٢٠	١٩	٣٨٠	٤٠٠	٣٦١
$\sum X = 90$	$\sum Y = 81$	$\sum XY = 1123$	$\sum X^2 = 1246$	$\sum Y^2 = 1019$

أولاً نحسب معامل الإنحدار :

$$\hat{\beta} = \frac{(7)(1123) - (90)(81)}{(7)(1246) - (90)^2} = 0.92$$

$$\hat{\alpha} = \frac{(81)(1246) - (90)(1123)}{(7)(1246) - (90)^2} = 0.23$$

وعليه فإن معادلة خط إنحدار (X/Y) تصبح :

$$Y = -0.23 + 0.92 X$$

وتكون قيمة الإنفاق Y عندما يكون الدخل ٢٥٠٠ دينار هو :

$$Y = -0.23 + 0.92 (2500) = 2299.77$$

أي أن الإنفاق يساوي : ٢٣٠٠ دينار تقريباً .

• معادلة خط إنحدار (X/Y)

بصورة مماثلة لخط إنحدار (Y/X) يمكن توفير خط إنحدار (X/Y) على الصورة التالية :

$$X_i = \pi + \rho Y_i + e_i$$

حيث Y ، X متغيران ، يمثل X المتغير التابع ويمثل Y المتغير المستقل ، أما π, ρ فهما ثابتا المعادلة تماماً مثل α و β في معادلة انحدار (Y/X) ، أما e_i فهو المتغير العشوائي .
نلاحظ في هذه المعادلة أن التغير في X ينشأ من مصدرين هما التغير في المتغير المستقل Y والذي يمكن قياسه ثم المتغير العشوائي e والذي لا يمكن قياسه . وكما فعلنا في حالة إنحدار (Y/X) يمكن استخدام طريقة المربعات الصغرى لنحصل على المعادلات المعتادة التالية :

$$\sum X_i = n\pi + \rho \sum Y_i$$

$$\sum X_i Y_i = \rho \sum Y_i + \rho \sum Y_i^2$$

حيث n عدد أزواج القيم المعلومة .

وتسمى هذه المعادلات بالمعادلات المعتادة ، ويمكن حل المعادلات المعتادة جبرياً لنحصل على تقدير لقيمتي الثابتين ρ, π كما يلي :

$$\hat{\pi} = \frac{(\sum X)(\sum Y^2) - (\sum Y)(\sum XY)}{n \sum Y^2 - (\sum Y)^2}$$

$$\hat{\rho} = \frac{n \sum XY - (\sum X)(\sum Y)}{n \sum Y^2 - (\sum Y)^2}$$

مثال (٧)

جد معادلة خط إنحدار (X/Y) ، من بيانات الدخل والإنفاق في المثال السابق ثم جد تقدير الدخل إذا كان الإنفاق ٢٠٠٠ دينار ؟ .

الحل

لحل هذه المسألة لا نحتاج لجدول آخر حيث أن جميع البيانات المطلوبة لحساب $\hat{\pi}, \hat{\rho}$ موجودة في جدول رقم (٩-٦) ، وعليه فإن :

$$\hat{\pi} = \frac{90(1019) - 90(1123)}{7(1019) - (81)^2} = -16.36$$

$$\hat{\rho} = \frac{7(1123) - 90(81)}{7(1019) - (81)^2} = 1.00$$

وعليه فإن معادلة خط إنحدار (X/Y) هي :

$$X = -16.36 + Y$$

فإذا كان الإنفاق ٢٠٠٠ دينار فإن الدخل يساوي :

$$X = 16.36 + (2000) = 1984$$

ملاحظة :

قد يلاحظ الدارس أن المتغير المستقل في هذا المثال لا يصلح أن يكون كمتغير تابع ، ولكن هنالك أكثر من العلاقات الخطية يمكن تبديل المتغيرات داخلها . وقد أوردنا نفس المثال في معادلة (X/Y) أشرح بعض المفاهيم الإحصائية المتعلقة بين الارتباط والإنحدار .

ج - العلاقة بين معامل الارتباط ومعامل الإنحدار :

عندما يكون الإنحدار بين ظاهرتين مستقيماً ، فإن الارتباط بينهما أيضاً يكون كذلك ، وتكون هناك علاقة رياضية مباشرة بين العاملين نبينها فيما يلي :

أولاً : معادلة إنحدار (Y/X) :

$$\beta = \frac{\sum XY - (\sum X)(\sum Y)}{n\sigma_x}$$

فإذا قمنا بضرب طرفي المعادلة في $\frac{\sigma_x}{\sigma_y}$ نجد أن :

$$\beta \frac{\sigma_x}{\sigma_y} = \frac{\sum XY - (\sum X)(\sum Y)}{n\sigma_y} \Rightarrow r = \beta \frac{\sigma_x}{\sigma_y} \rightarrow *$$

وعليه فإن معامل الارتباط يساوي معامل الانحدار مضروباً في الانحراف المعياري للمتغير X مقسوماً على الانحراف المعياري للمتغير Y .

ثانياً : معادلة إنحدار (X/Y)

بالمثل إذا قمنا بضرب معامل إنحدار (X/Y) في $\frac{\sigma_x}{\sigma_y}$ نحصل على :

$$r = \rho \frac{\sigma_x}{\sigma_y} \rightarrow **$$

$$r^2 = \beta \rho$$

وإذا ضربنا المعادلتين $[*]$ و $[**]$ ببعضهما البعض نحصل على :

أي أن معامل الارتباط يساوي الجذر التربيعي لمعامل الانحدار $r = \sqrt{\beta \rho}$ ويستخدم هذه العلاقة لحساب معامل الارتباط بمعلومية معاملي خطي الانحدار .

مثال (٨)

جد معامل الارتباط من بيانات الدخل و الإنفاق في المثال السابق .

الحل

لقد أوجدنا معاملي خطي الانحدار (Y/X) و (X/Y) حيث كان $\beta = 0.92, \rho = 1.00$ وعليه فإن معامل الارتباط r هو :

$$r = \sqrt{\beta \rho} = \sqrt{0.92(1.00)} = \sqrt{0.92} = 0.96$$

وهي نفس النتيجة التي توصلنا إليها عند حساب معامل الارتباط الخطي لبيرسون من البيانات مباشرة .



- في دراسة عن الخصوبة لعينة من الأسر بمحافظة الجزيرة [5]
كانت العلاقة بين عمر الأم وعدد الأطفال كما يلي :

عدد الأطفال	العمر	٢-١	٤-٢	٦-٤	٨-٦
٢٠-	١٢				
٢٥-	٧	٥			
٣٠-	٥	٨			
٣٥-		٧	٤		
٤٠-		٦	٣		
٤٥-		٤	٩		
٥٠-		٢	٧	٥	
٥٥-٦٠				١٠	١٥

أسئلة تقويم ذاتي



١. متي نستخدم المقاييس الإحصائية التالية:

- معامل الاقتران.

- معامل التوافق

٢. عرف الإنحدار الخطي البسيط

٣. أكتب معادلة خط الإنحدار (X,Y) .

٤. أذكر الطرق التي يتم بها رسم خط الإنحدار.

الخلاصة

عزيزي الدارس،

ما الذي ناقشناه في الوحدة السادسة

- تعرضنا في الوحدة السادسة إلى الارتباط والانحدار الخطي. من خلال المقاييس الإحصائية الستة الواردة فيها.
 - تهدف هذه المقاييس إلى دراسة أساليب قياس مقدار واتجاه العلاقة بين متغيرين وقوة الارتباط بينهما ومن ثم تحديد علاقة رياضية تربط بين المتغيرين وتمكن من التنبؤ بقيمة أحد المتغيرين إذ علم قيمة المتغير الآخر.
 - يوجد اختلاف بين العلاقة الدالية والعلاقة الإحصائية.
 - توجد أنواع مختلفة من أشكال الارتباط وقوته وقياسه رقمياً. كما توجد عدة طرق رسم خط الانحدار.
 - هنالك علاقة بين معامل الارتباط ومعامل الانحدار.
- أرجو أن تراجع هذه الخلاصة ما ورد في هذه الوحدة. حتى توفق في فهم هذه الوحدة والاستفادة منها.

لمحة مسبقة عن الوحدة التالية:

تتحدث الوحدة التالية عن الأرقام القياسية من خلال تعريف الرقم القياسي ومجالات تطبيقها. تتعرض في ذلك إلى مناسيب الأسعار وخصائصها وأواعها كما تبين سلاسل المناسيب والمشكلات التطبيقية التي تواجه استعمال هذه السلاسل. كما تشمل طرق حساب الأرقام القياسية مبينة مزايا وعيوب كل طريقة. أتمني لك التوفيق في دراستك لها.



محتويات الوحدة

الصفحة	الموضوع
١٦٦	المقدمة
١٦٦	تمهيد
١٦٧	اهداف الوحدة
١٦٨	١. الرقم القياسي
١٦٩	٢. مناسيب الأسعار
١٧٣	٣. سلاسل المناسيب
١٧٦	٤. استخدام مقاييس النزعة المركزية
١٧٧	٥. الطريقة التجميعية البسيطة
١٧٩	٦. الوسط البسيط للمناسيب
١٧٩	٧. الطريقة التجميعية المرجحة
١٨٣	الخلاصة
١٨٤	لمحة مسبقة عن الوحدة التالية
١٨٥	إجابات التدريبات
١٨٦	مسرد المصطلحات

المقدمة

تمهيد

عزيزي الدارس،

مرحباً بك في رحاب الوحدة السادسة من المقرر الإحصاء الإداري والذي يشتمل على تسع وحدات دراسية.

تشمل هذه الوحدة (الأرقام القياسية) سبعة أقسام رئيسية يتناول الأول تعريف الرقم القياسي وتطبيقاته بينما يتحدث القسم الثاني عن مناسيب الأسعار وخصائصها معرباً منسوب السعر. يشتمل القسم الثالث توضيحاً لمفهوم سلاسل المناسيب. بينما تشمل الأقسام من الرابع وحتى السابع طرقاً مختلفة لحساب الأرقام القياسية.

تأتي أهمية هذه الوحدة (الأرقام القياسية) في استعمالها لدراسة التغيرات لمجموعة مرتبطة من المتغيرات بالنسبة للزمان أو المكان أو أي خاصية أخرى.

تجد بين ثنايا هذه الوحدة شرحاً وافياً للأرقام القياسية وطرق حسابها كما تتخلل الوحدة أسئلة تقويم ذاتي وتدرّيات مختلفة وأمثلة توضيحية للشرح.

أرجو أن تستفيد من هذه الوحدة في حياتك العملية كما لا يفوتني تنبيهك إلى نقدها وتقويمها بما يسهم في تطويرها علمياً – مع تمنياتنا لك بالنجاح والتوفيق.

أهداف الوحدة



عزيزي الدارس،

بعد فراغك من دراسة هذه الوحدة ينبغي أن تكون قادراً على أن:

- تعرف الرقم القياسي تعريفاً علمياً.
- تستخدم الرقم القياسي في مقارنة تكاليفك لهذا العام مع ثلاث أعوام سابقة.
- تقارن بين المناسيب الكمية والمناسيب القيمة.
- تعدد الطرق المختلفة لحساب الأرقام القياسية.
- تشرح الصيغ المختلفة لقياس الطريقة التجميعية المرجحة.

١ . الرقم القياسي Index Number

تعريف

الرقم القياسي هو دليل إحصائي الغرض منه دراسة التغيرات لمجموعة مرتبطة من المتغيرات بالنسبة للزمن ، المكان الجغرافي، أو أي خاصية أخرى مثل الدخل ، الوظيفة وغير ذلك . وتسمى أحيانا المجموعة من الأرقام القياسية لسنوات أو أماكن مختلفة وما إلى ذلك بالسلسلة القياسية.

تطبيقات الأرقام القياسية

عزيزي الدارس

ماذا أنت فاعل إذا طلب منك مقارنة تكاليف المعيشة بمدينتك خلال سنة معينة مع سنوات سابقة؟ أو مقارنة إنتاج الزيوت خلال سنة معينة في مدينة الأبيض بالإنتاج في مدينة الخرطوم. لا شك أن استخدام قياس إحصائي يبنى على (الأرقام القياسية) يفيدك في تقديم تقرير علمي يستند عليه.

وعلى الرغم من أن الأرقام القياسية تستخدم أساسا في الأعمال والاقتصاد، فإنه يمكن تطبيقها في مجالات كثيرة مختلفة. ففي مجال التعليم على سبيل المثال، نستخدم الأرقام لمقارنة ذكاء الطلبة في مناطق مختلفة أو سنوات مختلفة.

كثير من الوكالات الحكومية والخاصة تقوم بحساب أرقام قياسية أو أدلة كما تسمى في أغلب الأحيان، وذلك بهدف التنبؤ بأحوال الأعمال والاقتصاد، وكذلك الحصول على معلومات عامة وما إلى ذلك . فمثلا هناك الأرقام القياسية للأجور، الأرقام القياسية للإنتاج ، الأرقام القياسية للبطالة وغير ذلك. ومن أكثر الأرقام المعروفة الرقم القياسي لتكاليف المعيشة أو الرقم القياسي للمستهلك والذي يعده مكتب إحصاءات العمل. وفي كثير من عقود العمل يظهر شرط معين للتدرج والذي بمقتضاه تغطي زيادة تلقائية في الأجور مقابلة في الرقم القياسي لتكاليف المعيشة.

في هذا الوحدة سنهتم أساسا بالأرقام القياسية التي تظهر التغيرات بالنسبة للزمن، على الرغم من أن الطرق التي ستشرح يمكن تطبيقها في الحالات الأخرى .



- ١- عرف الرقم القياسي؟
٢- أذكر أهداف تطبيق الأرقام القياسية؟

٢. مناسب الأسعار Price Relative

من أبسط الأمثلة للرقم القياسي هو منسوب السعر ، وهو نسبة السعر لسلعة واحدة في فترة المقارنة إلى سعرها في فترة أخرى تسمى بفترة الأساس أو فترة الإسناد. وللتسهيل سوف نفترض أن الأسعار ثابتة لأي فترة. فإذا لم يكن هذا صحيحا فإنه يمكن استخدام متوسط ملائم للفترة حتى نجعل هذا الفرض صحيحا.

إذا كانت P_0 تمثل سعر السلعة خلال فترة الأساس و P_n سعرها خلال فترة المقارنة، فإنه بالتعريف:

$$(١) \quad \text{منسوب السعر} = P_n / P_0$$

ويعبر عنه بشكل عام في صورة نسبة مئوية بضربه في ١٠٠.
وبشكل أكثر عمومية إذا كانت P_a ، P_b هي أسعار سلعة خلال الفترات a ، b علي الترتيب فإن منسوب السعر في الفترة b بالنسبة للفترة a يعرف بأنه $\frac{P_b}{P_a}$ ويرمز له بالرمز $\frac{P_b}{P_a}$ ، وسنجد أن هذا الرمز مفيد فيما بعد. بهذا الرمز فإن منسوب السعر بالمعادلة (١) يمكن أن يرمز له بالرمز $(P_{a/n})$.

مثال (١)

أفترض أن أسعار المستهلكين لعنصر معين في السنوات ١٩٦٠ ، ١٩٥٥ هي ٣٠.٥٢ ديناراً على الترتيب . فإذا أخذنا ١٩٥٥ كسنة أساس و ١٩٦٠ سنة المقارنة فإن :

$$\text{منسوب السعر} = p_{1955/1960} = \frac{\text{السعر في ١٩٦٠}}{\text{السعر في ١٩٥٥}} = \frac{30}{25} = 1.2 = 120\%$$

أو باختصار ١٢٠ بحذف علامة % كما هو متبع غالباً في المؤلفات الإحصائية. هذه النتيجة تعني ببساطة أن سعر العنصر سنة ١٩٦٠ أصبح ١٢٠% من سعره في سنة ١٩٥٥ أي زاد بنسبة ٢٠% .

مثال (٢)

بأخذ ١٩٦٠ كسنة أساس و ١٩٥٥ هي سنة المقارنة في المثال (١) فإن :

$$\text{منسوب السعر} = p_{1960/1955} = \frac{\text{السعر في ١٩٥٥}}{\text{السعر في ١٩٦٠}} = \frac{٢٥}{٣٠} = \frac{٥}{٦} = ٨٣.٣٣\%$$

أو باختصار $\frac{٨٣}{٣}$ وهذه يعنى أنه في ١٩٥٥ كان سعر العنصر هو ٨٣.٣٣ % من سعره في ١٩٦٠ ، أي أنه كان ينقص بنسبة ١٦.٦٦ % .

لاحظ أن منسوب السعر لفترة معينة بالنسبة لنفس الفترة سيكون دائما ١٠٠ % أو ١٠٠ وعلى وجه الخصوص فإن منسوب السعر المقابل لفترة الأساس يصبح دائما ١٠٠ . وهذا يوضح الرمز الذي يستخدم غالبا في المؤلفات الإحصائية بكتابة ، على سبيل المثال ، $١٩٥٥ = ١٠٠$ للإشارة إلى أن سنة ١٩٥٥ أخذت كسنة أساس :

خصائص مناسيب الأسعار:

إذا كانت ... P_a ، P_b ، P_c تعبر عن أسعار الفترات a ، b ، c ... على الترتيب فإن الخصائص التالية تتحقق لمناسيب الأسعار المرتبطة بها . الإثباتات تحصل عليها مباشرة من التعاريف .

أ- خاصية التطابق Identity Property

$$P_a/a = 1.0$$

وهذه تقرر ببساطة أن منسوب السعر لفترة معينة بالنسبة لنفس الفترة تساوى ١ أو ١٠٠ % .

ب- خاصية الانعكاس في الزمن Time Reversal Property

وهذه تقرر أنه إذا حللنا فترتين كلا محل الأخرى ، فإن مناسيب الأسعار المقابلة تكون كلا منها

$$\text{معكوس الأخرى} . (P_{a/b})(P_{b/a}) = 1.00 \text{ or } (P_{a/b}) = \frac{1}{(P_{b/a})}$$

ج- خاصية الدورية أو الدائرية Cyclical Property

$$(P_{a/b})(P_{b/c})(P_{c/a}) = 1.00$$

$$(P_{a/b})(P_{b/a})(P_{c/a}) = 1.00$$

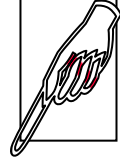
د- خاصية الدورية أو الدائرية المعدلة Modified Cyclical Property

$$(P_{a/b}) (P_{b/c}) = (P_{a/c})$$

$$(P_{a/b}) (P_{b/c}) (P_{c/d}) = (P_{a/d})$$

وهذه نحصل عليها مباشرة من الخاصيتين (٣)، (٢) أو ب وج.

نشاط



مناسيب الأسعار :

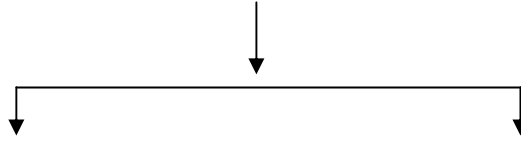
١ - متوسط أسعار التجزئة بالدولار للطن من الفحم المباع في بلد معين خلال السنوات ١٩٥٨ - ١٩٥٣ موضح بالجدول التالي :
أ. باستخدام ١٩٥٣ كأساس ، أوجد مناسيب الأسعار المقابلة للسنوات ١٩٥٦ و ١٩٥٨ .

ب. باستخدام ١٩٥٦ كأساس . أوجد منسوب السعر المقابل لجميع السنوات المعطاة

ج. باستخدام ١٩٥٥-١٩٥٣ كأساس، أوجد منسوب السعر لجميع السنوات المعطاة

السنة	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦	١٩٥٧	١٩٥٨
-------	------	------	------	------	------	------

مناسيب الأسعار



مناسيب القيمة
Value Relative

مناسيب الكمية
Quantity Relatives

مناسيب الكمية Quantity Relatives

بدلاً من مقارنة أسعار السلعة ، قد نهتم بمقارنة كميات أو أحجام السلع، مثل كمية أو حجم الإنتاج ، الاستهلاك ، التصدير ، وغيرها . في مثل هذه الحالات نتكلم عن مناسيب الكمية أو مناسيب الحجم للتسهيل، كما في حالة الأسعار، نفترض أن الكميات ثابتة في أي فترة إذا لم يكن هذا صحيحاً فإنه يمكن استخدام متوسط ملائم لجعل هذا الفرض ممكن.

إذا كانت q_0 تعبر عن كمية أو حجم السلعة المنتجة، المستهلكة ، المصدرة وغير ذلك خلال فترة الأساس ، بينما q_n تعبر عن كمية الإنتاج، الاستهلاك وغير ذلك المقابلة، خلال فترة المقارنة، فأنا نعرف :

$$(٢) \text{ منسوب الكمية} = q_n / q_o$$

ويعبر عنها بصفة عامة في شكل نسب مئوية.

كما في حالة مناسيب السعر، فأنا نستخدم الرمز $q_{a/b} = q_b / q_a$ للتعبير عن منسوب السعر في الفترة b بالنسبة للفترة a . نفس الملاحظات التي تتعلق بمناسيب السعر تنطبق على مناسيب الكمية.

مناسيب القيمة Value Relatives

إذا كان p هو سعر السلعة خلال فترة ما و q هي الكمية أو الحجم المنتج ، المباع وغير ذلك خلال الفترة ، إذن pq تسمى القيمة الإجمالية . بهذا فإذا بيعت ١٠٠٠ وحدة بسعر ٣٠ جنيها لكل وحدة فإن القيمة الإجمالية هي $30000 = (1000) \times (30)$.

إذا كانت P_o تعبر عن السعر و q_o عن الكمية لسلعة خلال فترة الأساس بينما p_n تعبر عن السعر المقابل و q_n عن الكمية المقابلة خلال الفترة المعطاة، كذلك فإن القيمة الإجمالية خلال هذه الفترات هي V_o لفترة الأساس و V_n للفترة المعطاة فأنا نعرف:

$$(٣) \text{ منسوب القيمة} = V_n / V_o = P_n q_n / P_o q_o = P_n / P_o \times q_n / q_o$$

نفس التعليقات، والرموز والخصائص التي تتعلق بمناسيب السعر والكمية يمكن ان تنطبق على مناسيب القيمة وعلى وجه الخصوص إذا كانت $P_{a/b}$ تعبر عن منسوب السعر $q_{a/b}$ عن منسوب الكمية و $V_{a/b}$ عن منسوب القيمة للفترة b بالمقارنة بالفترة a، بهذا كما في المعادلة (٣) .

$$V_{a/b} = P_{a/b} q_{a/b}$$
والذي يسمى خاصية الانعكاس في المعامل.

أسئلة تقويم ذاتي



- ١- عرف منسوب السعر؟
- ٢- أذكر خصائص مناسيب الأسعار؟
- ٣- أشرح مفهوم مناسيب أ- الكمية ب- مناسيب القيمة

٣. سلاسل المناسيب Series Relatives

إذا كانت $P_1, P_2, P_3, \dots$ تمثل الأسعار خلال الفترات المتتالية $1, 2, 3, \dots$ إذن $P_{1/2}, P_{2/3}, P_{3/4}, \dots$ تمثل مناسيب الأسعار لكل فترة زمنية بالمقارنة بالفترة الزمنية السابقة لها وتسمى بمناسيب الوصلات.

مثال:

إذا كانت أسعار سلعة خلال السنوات ١٩٥٤، ١٩٥٥، ١٩٥٦، ١٩٥٣ هي ٨، ١٢، ١٥، ١٨ دينار على الترتيب ، فإن الوصلات النسبية هي:

$$P_{1953/1954} = 12/8 = 150\%$$

$$P_{1954/1955} = 15/12 = 125\%$$

$$P_{1955/1956} = 18/15 = 120\%$$

يمكن التعبير دائما عن مناسيب الأسعار لفترة معينة بالمقارنة بفترة الأساس بدلالة وصلة المناسيب. هذا

$$P_{5/2} = P_{5/4} P_{4/3} P_{3/2}$$

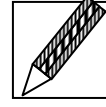
مثال :

من المثال السابق ، منسوب السعر لسنة ١٩٥٦ بالمقارنة بسنة الأساس ١٩٥٣ هو :

$$P_{1953/1956} = P_{1953/1954} P_{1954/1955} P_{1955/1956} \\ 12/8 \times 15/12 \times 18/15 = 18/8 = 225\%$$

مناسيب السعر بالنسبة لفترة أساس ثانية، والتي سبق أن أوضحنا يمكن أن نحصل عليها باستخدام وصلة المناسيب تسمى أحيانا بسلسلة المناسيب بالنسبة لهذا الأساس ، أو المناسيب مسلسلة إلى أساس ثابت.

تدريب (١)



في الأمثلة ١ ، ٢ مجموعة سلسلة المناسيب للسنوات ١٩٥٥ ، ١٩٥٦ ،
١٩٥٤ بالمقارنة بسنة الأساس ١٩٥٣ اعط الاجابة لما يلي :

- (1) $P_{1953/1954}$
- (2) $P_{1953/1955}$
- (3) $P_{1953/1956}$

الأفكار السابقة قابلة للتطبيق أيضا في حالة مناسيب الكميات ومناسيب القيمة.

مشاكل تطبيقية:

في نواحي التطبيق الفعلي لا نهتم بدرجة كبيرة بالمقارنة بين أسعار ، كميات أو قيم سلع بمفردها بقدر اهتمامنا بالمقارنة بين مجموعات كبيرة من هذه السلع. على سبيل المثال ، عند حساب الرقم القياسي لنفقات المعيشية لا نهتم فقط بأسعار اللبن في فترة واحدة بالمقارنة بفترة أخرى ولكن نرغب أيضا في مقارنة أسعار البيض ، اللحم ، الخبز الإيجار والملابس وغيرها . بحيث يمكن أن نحصل على صورة عامة. وبالطبع يمكن وضع قائمة بمناسيب أسعار كل السلع . ولكن هذا لا يعد مرضيا . فما نرغب فيه هو رقم قياسي واحد والذي يمكن أن يقارن الأسعار في الفترتين في المتوسط.

وليس من الصعب التنبؤ بأن حسابات الأرقام القياسية المتضمنة مجاميع من السلع تتضمن كثيرا من المشاكل التي يجب حلها عند حساب الرقم القياسي لتكاليف المعيشية ، على سبيل المثال فيجب أن نقرر ما هي السلع التي يجب أن تدخل ضمن الرقم وكذلك كيفية ترجيحها بما يتناسب مع أهميتها النسبية. فيجب أن نجمع بيانات تتعلق بأسعار وكميات هذه السلع كذلك فإننا نواجه بمشكلة التعرف في حالة وجود درجات مختلفة لنفس النوع من السلع، أو ماذا نفعل في حالة ما إذا كانت بعض الأنواع من المواد أو الآلات متاحة في احد السنوات ولكنها لم تكن موجودة في سنة الأساس وفي النهاية يجب أن نقرر كيف نضع هذه المعلومات معا بحيث تنتهي بالحصول على رقم قياسي واحد لتكلفة المعيشة له دلالة عملية.



- ١- أشرح مفهوم سلاسل المناسيب؟
- ٢- ماهي المشكلة العملية التي تجابه تطبيق الرقم القياسي كمفهوم إحصائي

٤. استخدام مقاييس النزعة المركزية

بما انه من الأفضل دائما أن نصل إلى رقم قياسي واحد يلخص كمية كبيرة من المعلومات، فانه من السهل التحقق من أن مقاييس النزعة المركزية، مثل تلك التي درست في الفصل الثالث، تلعب دورا مهما في حساب الأرقام القياسية وكما أن هناك طرقا عديدة موجودة لحساب مقاييس النزعة المركزية، فإن هناك طرقا كثيرة لحساب الأرقام القياسية لكل منها مزاياه وعيوبه.

فيما يلي سوف نقوم باختيار عدد قليل من الطرق الشائعة الاستخدام في النواحي العملية مستخدمين أنماطا عديدة من طرق مقاييس النزعة المركزية . وعلى الرغم من أننا سنقتصر على الأرقام القياسية للأسعار أولا ، فإننا سوف نوضح كيف يمكن بسهولة تعديلها لتطبق في حالة الكمية أو القيمة.

اعتبارات نظرية Theoretical Considerations

من المستحب من الناحية النظرية أن تحقق الأرقام القياسية لمجموعة من السلع الخواص التي تحققها المناسيب (أي الأرقام القياسية لسلعة واحدة). وأي رقم قياسي له خاصية معينة يذكر عنه انه يحقق الاختبار المرتبط بهذه الخاصية. بهذا، فعلى سبيل المثال، الأرقام القياسية التي لها خاصية الانعكاس في الزمن يقال عنها أنها تحقق اختبار الانعكاس في الزمن ، وهكذا.

ولم يكتشف رقم قياسي للان تحقق كل الاختبارات ، على الرغم من انه في كثير من الحالات تتحقق هذه الاختبارات تقريبا. يحقق رقم فيشر المثالي ، وبهذا يقترب من أي رقم قياسي نافع آخر من تحقيق الخصائص التي تعتبر مهمة، ومنها جاءت تسمية "المثالي" ومن وجهة النظر العملية ، يمكن استخدام أرقام قياسية أخرى كذلك وسوف نقوم باختبار بعضها.

نشاط

- أثبت أن (أ) $(P_{a/b}) (P_b/c) = p_a/c$ (ب) $(P_{a/b})(P_{b/a}) = 1$



رموز :

من المعتاد استخدام الرمز $p_n^{(1)}, p_n^{(2)}, p_n^{(3)}, \dots$ للتعبير عن أسعار السلعة الأولى ، الثانية والثالثة ، ... خلال الفترة المعينة n . الأسعار خلال فترة الأساس يرمز لها بالرمز $p_o^{(1)}, p_o^{(2)}, p_o^{(3)}, \dots$ الخ ... الأرقام $1, 2, 3, \dots$ هي أدلة ويجب ألا تثير اللبس مع الأسس. وبهذه الرموز فان سعر السلعة في خلال الفترة n يمكن الإشارة إليها بالصورة $p_n^{(t)}$.

وكما في الفصول السابقة ، نستخدم رمز التجميع على الدليل j . على سبيل المثال ، بافتراض أن هناك مجموعة N من السلع ، فان مجموع أسعاره خلال الفترة n يمكن كتابته على الشكل :

$$\sum_{j=1}^N P_n^{(j)} \quad \text{أو} \quad \sum P_n^{(j)}$$

ومن السهل حذف الأدلة معا وكتابة $\sum P_n$ ، وهو ما سوف نتبعه إذا لم يؤد هذا إلى أي التباس . ويجب أن نحفظ نصب أعيننا حقيقة أننا نستخدم مجموعة متكاملة من الرموز ، في هذه الرموز ، فأن $\sum P_o$ تعبر عن الأسعار لجميع السلع خلال فترة الأساس . ونستخدم رموزا مماثلة للكميات والقيم.

٥. الطريقة التجميعية البسيطة

Simple Additive Method

في هذه الطريقة لحساب الرقم القياسي للأسعار ، فأننا نعبر عن مجموع أسعار السلع في سنة المقارنة كنسبة مئوية من مجموع أسعارها في سنة الأساس.

$$(٤) \quad \text{الرقم القياسي التجميعي البسيط} = \sum P_n / \sum P_o$$

حيث :

$$\sum P_o = \text{مجموع أسعار السلع في سنة الأساس.}$$

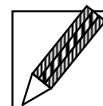
$$\sum P_n = \text{المجموع المقابل لأسعار السلع في سنة المقارنة.}$$

حيث يعبر عن النتيجة كنسبة مئوية كما هو الحال بالنسبة للأرقام القياسية بشكل عام. وعلى الرغم من أن هذه الطريقة لها الميزة بأنها سهلة التطبيق، إلا أن لها عيبين كبيرين يجعلان استخدامها غير مستحب.

١ / لا يؤخذ في الحسبان الأهمية النسبية للسلع المختلفة . فمثلا طبقا لهذه الطريقة ، فإن أوزانا متساوية تعني أن نفس الأهمية سوف تعطي للألبان ولمعجون الحلاقة عند حساب الرقم القياسي لتكلفة المعيشة.

٢ / الوحدات المستخدمة في تمييز السعر ، مثل ، الجرام وغيرها تؤثر على قيمة الرقم القياسي.

تدريب (٢)



– الجدول التالي يوضح متوسط أسعار الجملة في بلد والإنتاج من الألبان ، والزبد والجبن للسنوات ١٩٥٨ ، ١٩٥٠ ، ١٩٤٩ أحسب رقم قياسي بسيط الأسعار الجملة لمنتجات هذه الألبان لسنة ١٩٥٨ مستخدما كأساس

(أ) ١٩٤٩ (ب) ١٩٥٠ – ١٩٤٩

	الكميات المنتجة (ملايين الكيلوجرامات)			الأسعار لكل (كيلوجرام)		
	١٩٥٨	١٩٥٠	١٩٤٩	١٩٥٨	١٩٥٠	١٩٤٩
لبن	١٠.٤٣٦	٩٧١٧	٩٦٥٧	٤.١٣	٣.٨٩	٣.٩٥
زبد	١٥.٥	١١٥.٥	١١٧.٧	٥٩.٧	٦٢.٢	٦١.٥
جبن	٨٢.٧٠	٧٤.٣٩	٧٧.٩٣	٣٨.٩	٣٥.٤	٣٤.٨

٦. الوسط البسيط للمناسيب

Simple Average Method

في هذه الطريقة هناك عديد من الصيغ تعتمد على الطريقة المستخدمة في الحصول على أوساط مناسيب الأسعار، مثل الوسط الحسابي، الوسط الهندسي، الوسط التوافقي، الوسيط، وما إلى ذلك فإذا استخدمنا الوسط الحسابي، على سبيل المثال فإننا نحصل على :

$$\text{الوسط الحسابي البسيط للرقم القياسي لمناسيب الأسعار} = \sum p_n / p_o / N \quad (٥)$$

حيث :

$$\sum p_n / p_o = \text{مجموع مناسيب أسعار جميع السلع.}$$

$$N = \text{عدد مناسيب أسعار السلع المستخدمة.}$$

ويمكن حساب الأرقام القياسية باستخدام أنواع أخرى من مقاييس النزعة المركزية وعلى الرغم من أن هذه الطريقة تتخلص من العيب الثاني الموجود في الطريقة التجميعية البسيطة ولكن يظل العيب الأول موجودا بها.

٧. الطريقة التجميعية المرجحة

Weighted Average Method

للتغلب على عيوب الطريقة التجميعية البسيطة، فأنا نرجح أسعار كل سلعة باستخدام معامل ملائم ويستخدم غالبا كمية أو حجم السلعة المباعة خلال فترة الأساس، أو سنة المقارنة أو سنة نموذجية (والتي تتضمن متوسط عدد من السنوات) هذه الأوزان تشير إلى أهمية السلعة المعينة وهناك ثلاث صيغ ممكنة تعتمد على ما إذا كنا سنستخدم كميات سنة الأساس أو سنة المقارنة أو سنة نموذجية وتعبر عنها بالرموز :

$$q_t, q_o, q_n \text{ على الترتيب .}$$

أ - رقم لاسبيرز القياسي أو طريقة سنة الأساس Laspeyer's Price Index

$$(٦) \text{ الرقم القياسي التجميعي المرجح (سنة الأساس) .} = \frac{\sum p_n q_o}{\sum p_o q_o}$$

ب - رقم باش القياسي أو طريقة سنة المقارنة Paache Price Index

$$(٧) \text{ الرقم القياسي التجميعي المرجح باستخدام كميات سنة المقارنة.} = \frac{\sum p_n q_n}{\sum p_o q_n}$$

ج - طريقة السنة النموذجية :

إذا اعتبرنا أن q_t تعبر عن وزن الكمية خلال فترة نموذجية ، فأنا نعرف الرقم القياسي التجميعي المرجح باستخدام كميات السنة النموذجية =

$$(٨) \frac{\sum p_a q_t}{\sum p_o q_t}$$

عندما تكون $t, n = t = 0$ فإن هذه الصيغة تؤول إلي الصيغة (٦) و (٧) على الترتيب.

رقم فيشر المثالي Fisher Ideal Index

$$(٩) \text{ نعرف الرقم المثالي لفيشر :} F1 = \sqrt{\left[\frac{\sum p_n q_o}{\sum p_o q_o} \right] \left[\frac{\sum p_n q_n}{\sum p_o q_n} \right]}$$

وهذا الرقم القياسي هو الوسط الهندسي للرقمين القياسين لكل من لاسبيرز وباش الموضحين بالمعادلتين (٦) و (٧) وكما سبق أن أوضحنا فإن رقم فيشر المثالي يحقق كلا من اختباري الانعكاس في الزمن والانعكاس في المعامل ، وهذا مما يعطيه بعض المزايا النظرية عن الأرقام القياسية الأخرى.

رقم مارشال - أدجورث القياسي : Marshal – Edgworth Index

يستخدم رقم مارشال أدجورث القياسي الصيغة التجميعية المرجحة باستخدام طريقة السنة النموذجية حيث الأوزان هي الوسط الحسابي لكميات سنة الأساس وكميات سنة المقارنة .أي :

$$\text{رقم مارشال - أدجورث القياسي للأسعار} = \frac{q_t = \frac{1}{2} (q_o + q_n)}{\text{وبالتعويض بهذه القيمة } [q_t] \text{ في المعادلة (٨) نحصل على :}}$$

$$\frac{\sum p_n (q_o + q_n)}{\sum p_o (q_o + q_n)}$$

مثال :

الجدول التالي يبين السعار والكميات للطماطم واللحوم في سوق ودمدني لشهري يناير وفبراير عام ١٩٩٤ أوجد أرقام باش ، لاسبيرز ، فيشر ، ومارشال - أدجورث القياسي؟

جدول رقم (١)

الأسعار والكميات للطماطم واللحوم في سوق ود مدني

الكميات		الأسعار		
يناير	فبراير	يناير	فبراير	
٥٠٠٠	٥٥٦٦	٨٠٠	٢٥٠	الطماطم بالصفحة
٢٠٠٠	١٠٠٠	٨٠٠	١٠٠٠	اللحمة بالكيلو

الحل:

١ - رقم لاسبيرز القياسي- وهو يعتمد على الطريقة التجميعية بكميات سنة الأساس (يناير) أي :

$$G_L = \frac{\sum P_n q_o}{\sum p_o q_o} = \frac{250 \times 5000 + 1000 \times 2000}{800 \times 5000 + 800 \times 2000}$$

$$\frac{\sum p_n q_n}{\sum p_o q_n} = \frac{1800 \times 1000 + 5566 \times 250}{1800 \times 800 + 5566 \times 800} = \text{رقم باش القياسي}$$

$$\left(\frac{\sum p_n q_n}{\sum p_o q_o} \right) \times \left(\frac{\sum p_n q_o}{\sum p_o q_o} \right) = \text{رقم فيشر المثالي} = 0.56 = 0.542 \times 0.58$$

$$\frac{\sum p_n (q_o + q_n)}{\sum p_o (q_o + q_n)} = \text{رقم مارشال المثالي - ادجورث القياسي}$$

$$\frac{(2000 + 1800)1000 + (5566 + 5000)250}{(2000 + 1800)800 + (5566 + 5000)800} = 0.56$$

كل هذه الأرقام تشير إلى انخفاض كبير في أسعار الطماطم واللحوم بودمدني خلال شهري يناير وفبراير . والسبب في ذلك هو الكميات الكبيرة المعروضة من سلعة الطماطم خلال الفترة .

أسئلة تقويم ذاتي



١ - أذكر الطرق المستخدمة لحساب الأرقام القياسية؟

٢ - كيف نستخدم أي من هذه الطرق في حساب الرقم القياسي؟

٣ - متى نستخدم كل من الآتي لحساب الرقم القياسي؟
أ - رقم فيشر المثالي.

ب - رقم مارشال - أدجورث القياسي.

الخلاصة

عزيزي الدارس،

هل أنت مستوعب لما تناولناه في هذه الوحدة؟ فلنعد إليه سوياً

* في الأقسام الثلاثة الأولى تناولنا الرقم القياسي Index Number وتطبيقاته ومناسيب الأسعار وخصائصها والمناسيب الكلمية ومناسيب القيمة وسلاسل المناسيب والمشكلات التطبيقية لهذه المفاهيم.

* وفي القسم الرابع ذكرنا أن مقاييس النزعة المركزية تلعب دوراً مهماً في جانب الأرقام القياسية كما أن هنالك عدة طرق لحساب الأرقام القياسية منها:

أ- الطريقة التجميعية البسيطة.

ب- الوسط البسيط للمناسيب

ج- الطرق التجميعية المرجحة

* في حالة استخدام الطريقة التجميعية المرجحة هنالك ثلاث صيغ ممكنة تعتمد على ما إذا كنا سنستخدم سنة الأساس أو سنة المقارنة أو سنة نموذجية يعبر عنها بالرموز q_t, q_o, q_n على الترتيب

طريقة سنة الأساس رقم لاسبيرز القياسي Laspeyers Index Price

طريقة سنة المقارنة رقم باش القياسي Paache Piece Index

طريقة السنة النموذجية

* ثم تعرضنا في بقية الأقسام إلى رقم فيشر المثالي Fisher's Ideal Index

والذي يمثل الوسط الهندسي للرقمين القياسيين لكل من لاسبيرز وباش.

• رقم مارشال - ادجورث القياسي Marshal- Adgworth الذي يستخدم الصيغة

التجميعية المرجحة باستخدام طريقة السنة النموذجية .

أتمنى أن تكون مستوعباً لهذه الوحدة وأسأل الله لك التوفيق.

لمحة مسبقة عن الوحدة التالية

عزيزي الدارس،

تتناول الوحدة التالية تعريف السلاسل الزمنية والعوامل التي تؤثر فيها كما تشرح مفهوم المنحنى التاريخي. والعوامل المؤثرة في السلسلة الزمنية وتحليل السلاسل الزمنية والمتوسطات المتحركة كتمهيد للسلاسل الزمنية في الأقسام من الثاني والخامس نشرح في الوحدة تقدير الاتجاه العام والمتغيرات الموسمية والدورية والعشوائية . وأخيراً قابلية البيانات للمقارنة والتنبؤ. ثم تختم الوحدة بخلاصة لما ورد فيها ومسرد لمصطلحاتها ولمحة مسبقة للوحدة التي تليها. أتمنى لك التوفيق في دراستك لها.

إجابات التدريبات

تدريب (١)

المدى المطلق **Rang** = الفرق بين أعلى مفردة وأدنى مفردة = $89 - 8 = 81$

تدريب (٢)

أهم وأول خطوة هي حساب الوسط الحسابي

تدريب (٣)

لا يمكن استخدام الانحراف المعياري لهاتين الظاهرتين وذلك لاعتماده على وحدات القياس (حيث تقاس الأوزان بالكيلوجرام والأطوال بالسنتيمتر).

مسرد المصطلحات

- الرقم القياسي: Index Number

هو دليل إحصائي الغرض منه دراسة التغيرات لمجموعة مرتبطة من المتغيرات بالنسبة للزمن أو المكان الجغرافي أو أي خاصية أخرى.

- منسوب السعر: Price Relative

هو نسبة السعر لسلعة واحدة في فترة المقارنة إلى سعرها في فترة أخرى تسمى بفترة الأساس أو فترة الإسناد.

- سلاسل المناسيب: Series Relative

إذا كانت p_1, p_2, p_3 تمثل الأسعار خلال الفترات المتتالية ١.٢.٣ إذن تمثل مناسيب الأسعار لكل فترة زمنية بالمقارنة بالفترة الزمنية السابقة لها وتسمى بمناسيب الوصلات.



محتويات الوحدة

الموضوع	الصفحة
المقدمة	١٨٩
تمهيد	١٨٩
أهداف الوحدة	١٩٠
١. السلاسل الزمنية	١٩١
١.١ المنحنى التاريخي	١٩٢
٢.١ العوامل المؤثرة في السلاسل الزمنية	١٩٣
٣.١ تحليل السلاسل الزمنية	١٩٦
٤.١ المتوسطات المتحركة تمهيد السلاسل الزمنية	١٩٧
٢. تقدير الاتجاه العام	٢٠٠
٣. تقدير المتغيرات الموسمية	٢٠٥
٤. تقدير المتغيرات الدورية	٢٠٧
٥. تقدير المثيرات العشوائية	٢٠٨
٦. قابلية البيانات للمقارنة	٢٠٨
٧. التنبؤ	209
الخلاصة	210
لمحة مسبقة عن الوحدة التالية	211
مسرد المصطلحات	٢١٢

المقدمة

تمهيد

عزيزي الدارس

مرحباً بك إلي دراسة الوحدة السابعة من مقرر الإحصاء الإداري. تتناول هذه الوحدة مفهوم السلاسل الزمنية وتنقسم إلي سبعة أقسام:

- القسم الأول تعريف السلاسل الزمنية والعوامل المؤثرة فيها وتحليلها والمتوسطات المتحركة كتمهيد للسلاسل الزمنية.
 - القسم الثاني يتحدث عن تقدير الاتجاه العام والذي يمكن أن يتم عن طريق أربعة طرق نتعرض لها بالتفصيل.
 - يتناول القسم الثالث تقدير المتغيرات الموسمية (الدليل الموسمي) والذي يمكن حسابه عن طريق أربعة طرق .
 - في القسم الرابع نتحدث عن تقدير التغيرات الدورية .
 - القسم الخامس يتناول تقدير التغيرات العشوائية ويتناول القسمين السادس والسابع على التوالي موضوعي قابلية البيانات للمقارنة والتنبؤ.
- تتخلل الوحدة عدد من التدريبات والأنشطة وأسئلة التقويم الذاتي بهدف تعزيز الفهم الموفي لك. أمل أن تجد عندك هذه الوحدة - على صغر حجمها- العناية والاهتمام اللازمين ولا يفوتني أن الفت انتباهك إلي ضرورة إجراء الأنشطة الواردة فيها وعرض إجابتك على مشرفك الأكاديمي. أما أسئلة التقويم الذاتي فما استعصى عليك فإجابته داخل النص الذي يسبق السؤال ختاماً نرجو لك دراسة موفقة في هذه الوحدة.

أهداف الوحدة



عزيزي الدارس

بعد فراغك من دراسة هذه الوحدة ينبغي أن تكون قادراً على أن:

- ١- تعرف مفهوم السلاسل الزمنية لأي ظاهرة تعريفاً علمياً.
- ٢- ترسم المنحنى التاريخي لظاهرة ما.
- ٣- تعدد العوامل المؤثرة في السلسلة الزمنية.
- ٤- تشرح الطرق المختلفة التي يمكن بها تقدير الاتجاه العام على السلاسل الزمنية.
- ٥- تميز بين تقديرات المتغيرات الموسمية والدورية والعشوائية.

١ . السلاسل الزمنية

أسألك عزيزي الدارس هل سمعت بهذا المصطلح من قبل؟؟ أحسب أنها المرة الأولى التي يرد فيها لسمعك، ولكن ما هو التعريف العلمي للسلاسل الزمنية؟

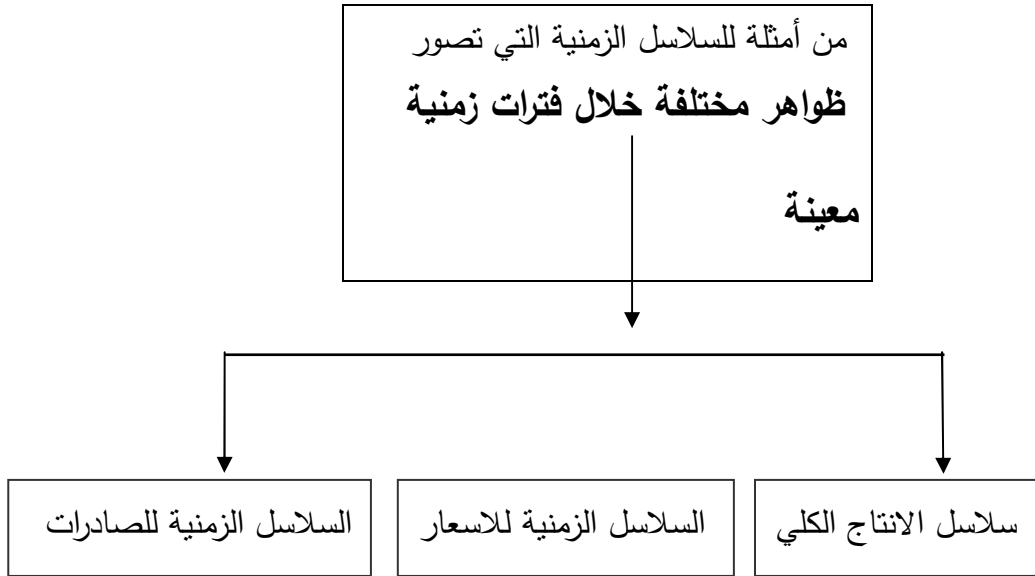
تعريف

السلسلة الزمنية لأي ظاهرة هي عبارة عن مجموعة من المشاهدات أخذت على فترات زمنية نتيجة تعقب هذه الظاهرة لفترة زمنية طويلة نسبياً وبصورة متتابعة وفي أغلب الأحيان تكون هذه الفترة الزمنية منتظمة .

والسلسلة الزمنية تتكون من متغيرين أحدهما مستقل وهو الزمن (t) والآخر تابع وهو قيمة الظاهرة محل الدراسة (Y) .

ويمكن كتابة العلاقة بين الزمن وقيمة الظاهرة في صورة دالة كما يلي :

$$Y = f(t)$$



وبالنسبة للفترة الزمنية فإن المشاهدات تؤخذ على فترات زمنية متساوية وفي أوقات محددة مثل أول كل أسبوع أو أول كل شهر أو أول كل سنة ... الخ .

وعند أخذ المشاهدات على فترات زمنية متساوية فإننا نجد أن فترات السلسلة الزمنية تكون قد غطت المشاهدات لفترة زمنية كاملة مثل الصادرات الشهرية أو الإنتاج الشهري وفي هذه الحالة نعتبر أن المشاهدات قد أخذت في منتصف الفترة الزمنية للسلسلة إما قيمة الظاهرة فقد تعتبر إما بالزيادة أو بالنقصان وهي لا تتزايد أو تتناقص باستمرار ولكنها تتأرجح بين الزيادة والنقصان من وقت لآخر حسب الفترات الزمنية المأخوذة .

نشاط

بعد أن صار لديك تعريفاً علمياً للسلاسل الزمنية هل يمتلك إعطاء ثلاثة أمثلة من بيئتك المحلية يمكنك دراستها علمياً باستخدام هذا المفهوم

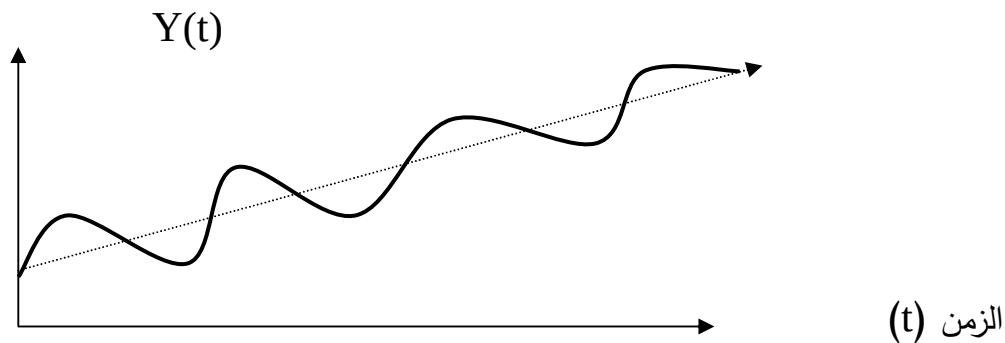


1.1 المنحنى التاريخي Historical Curve

الآن بعد معرفتك لمصطلح السلاسل الزمنية إليك السؤال التالي:

ما الغرض من تحليل السلسلة الزمنية لظاهرة ما ؟ والإجابة هي: التعرف على التغيرات التي تطرأ عليها لمعرفة أسبابها ونتائجها ونوع العلاقة التي تربطها بغيرها من الظواهر وإمكانية التنبؤ بقيمتها الموجودة في السلسلة لفترة سابقة أو لاحقة .

ويمكن تمثيل السلسلة الزمنية المتضمنة للمتغير y بتكوين الشكل البياني المقابل للزمن t ثم نعين نقطة لكل فترة زمنية . وعند توصيل هذه النقاط نحصل على خط متعرج يطلق عليه اسم المنحنى التاريخي . أنظر شكل (٨-١)



شكل رقم (8-1)

شكل يوضح خط الاتجاه العام

ومن المفيد التفكير في المنحنى التاريخي كما أوضحنا في شكل (٨:١) كنقطة تتحرك مع مرور الزمن وذلك فيما يشبه التحرك المادي للذرة تحت تأثير قوى مادية . وعلى أي حال ، فبدلاً من القوى المادية فإن الحركة قد تكون ناتجة عن قوى اقتصادية أو اجتماعية أو ديمغرافية أو أي قوة أخرى .

أسئلة تقويم ذاتي



* من خلال دراستك لما تقدم أجب عن الأسئلة التالية:

١/ عرف السلاسل الزمنية؟

٢/ ما الغرض من تحليل السلسلة الزمنية لظاهرة ما؟

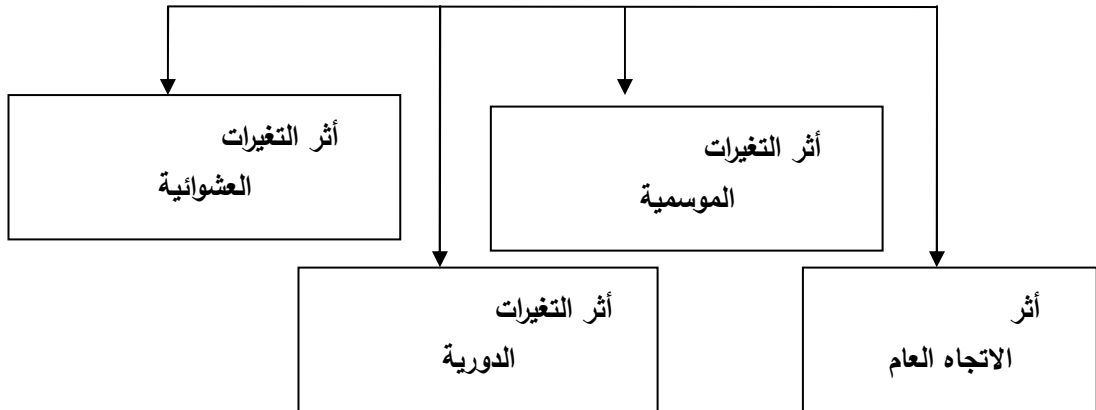
٣/ أكمل الجملة التالية : تتكون السلسلة الزمنية من متغيرين أحدهما وهو والآخر وهو

٤/ أشرح كيف يمكننا رسم المنحنى التاريخي لظاهرة ما بيانياً؟

2.1 العوامل التي تؤثر في السلسلة الزمنية :

عزيزي الدارس من معلوماتك السابقة تدرك أن دراسة أي ظاهرة من الظواهر وتحليلها لا بد وأن يتأثر بعدة عوامل وكذا عند دراسة أو تحليل أي سلسلة زمنية نلاحظ أنها تتأثر بعامل أو أكثر وبدرجات متفاوتة بالعوامل أو

الآثار التالية :



أ/ الاتجاه العام Trend

تعريف

يمثل التحركات طويلة المدى ويمثل الاتجاه العام إذا كان خطياً بخط انحدار المشاهدات $Y(t)$ على الزمن t . ويعبر عنه بخطوط متقطعة كما في الشكل (٨ : ١) وقد يكون الاتجاه العام موجبا بالزيادة أو سالبا بالنقصان .

ب/ التغيرات الموسمية Variations Seasonal

الموسم هو وحدة زمنية أقل من سنة ، قد تكون أسبوعية أو ربع سنوية . والمقصود بالتغيرات الموسمية.

تعريف

هي تلك التغيرات التي تتكرر في شكل زيادة أو نقصان في فترات (مواسم) معينة وبفترة تكرار منتظمة .

. أي أنه تشير إلى النمط المتماثل لحركة السلسلة الزمنية في الأشهر المتقابلة خلال السنوات المتتالية . مثال لذلك الزيادة في مبيعات المحلات التجارية السابقة لأعياد الفطر والاضحي أو بداية العام الدراسي وغيرها. وفي الشكل (٨ : ١) لا تظهر أي تغيرات موسمية نظرا لان الشكل يوضح الأرقام السنوية فقط.

ج/ التغيرات الدورية Cyclical Varations

تعريف

وهي تلك التغيرات التي تطرأ على الظاهرة بصفة دورية وعلى فترات متباعدة تزيد طول كل منها عن السنة وفترة التكرار غير منتظمة .

وهذه التغيرات تتكرر وتستعيد سيرتها بعد فترة من الزمن . وبمعنى آخر فهي تشير إلى الذبذبات طويلة المدى حول خط الاتجاه العام . وهي قد تتبع أو لا تتبع نفس النمط بعد كل فترة زمنية متساوية . ومن الأمثلة الهامة للتحركات الدورية ما يسمى بدورات الأعمال والتي تمثل فترات الرخاء ، والركود ، والكساد ، ثم انتهاء الأزمة.

د/ التغيرات العشوائية Random Variations

تعريف

وهي التغيرات التي تحدث نتيجة لسبب طارئ (عشوائي) غير متوقع مثال الحروب ،
الزلازل والأوبئة أو الفيضانات والاضرابات وغيرها

وهي لا يمكن التنبؤ بها أو تحديد حجمها أو اتجاهه حيث أنها لا يحكمها قانون أو قاعدة معينة.

نشاط

* إلى أي من التحركات المميزة في السلاسل الزمنية تنتمي أساساً
الأحداث التالية



أ/ اشتعال النار في مصنع أدى إلى
تأخير الإنتاج ثلاثة أسابيع.

ب/ عهد من الرهبانية.

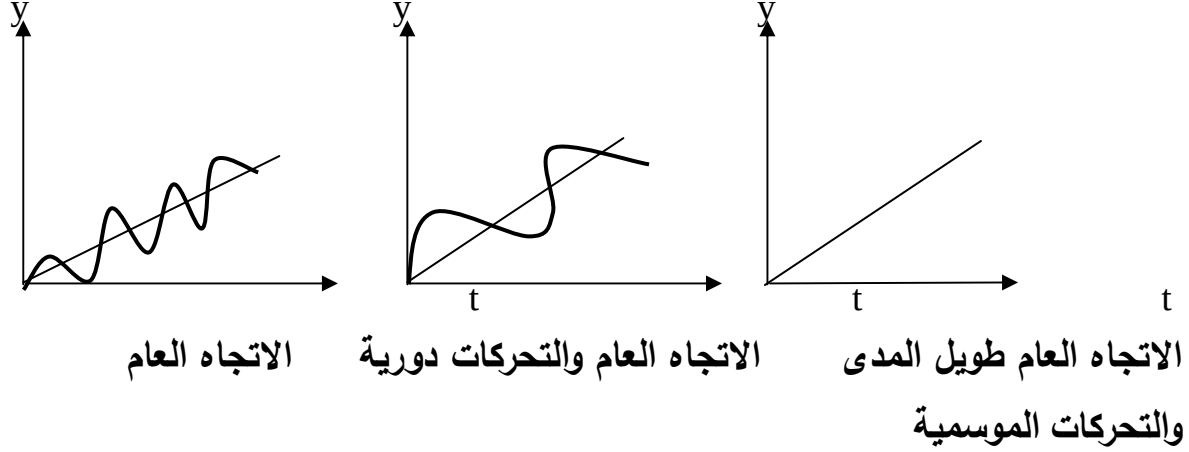
ج/ مبيعات فترة عيد الفطر في

٣.١ تحليل السلاسل الزمنية Analysis of Time S

سبق وان أشرنا إلى الغرض من تحليل السلاسل الزمنية ونضيف إلى ذلك أيضاً يتم تحليل
السلسلة الزمنية بغرض دراسة اثر التغيرات الخاصة بالاتجاهات العامة والموسمية والدورية والعشوائية
عن طريق عزل هذه التغيرات بأنواعها المختلفة التي تطرأ على الظاهرة في الزمن . ولتوضيح الطرق
التي تستخدم في عزل التغيرات انظر الشكل (٨.٢) ، حيث يوضح

خط الاتجاه العام طويل المدى

الشكل (أ) خط الاتجاه العام طويل المدى ، والشكل (ب) خط الاتجاه العام طويل المدى ومعه تحركات دورية والشكل (ج) خط الاتجاه العام طويل المدى ومعه تغيرات دورية وموسمية .



شكل رقم (8.2)

فإذا افترضنا إن النموذج الإحصائي الذي يعبر عن السلسلة الزمنية y هو حاصل ضرب المتغيرات الأربعة السابقة ، تحركات الاتجاه العام (T) ، تحركات موسمية (S) ، تحركات دورية (C) ، وتحركات عشوائية (I) فيمكن التعبير عن النموذج كما يلي :

$$Y_{(t)} = T_{(t)} \times S_{(t)} \times C_{(t)} \times I_{(t)}$$

وتحليل السلسلة الزمنية يتضمن فحص وتقدير العوامل من T , S , C , I والتي يشار إليها بتفكيك السلسلة الزمنية إلى المكونات الأساسية للتحركات .
ويجب أن نشير إلى أن بعض الإحصائيين يفضلون اعتبار Y كمجموع للتغيرات وليست مضروباً . أي أن :

$$Y_{(t)} = T_{(t)} + S_{(t)} + C_{(t)} + I_{(t)}$$

وهنا سنلتزم بالمفكوك الضربي إلا إن الأسلوب مشابه تماماً في حالة المفكوك الجمعي . ومن الناحية العملية ، فإن قرار اتخاذ طريقة التفكيك تعتمد على درجة النجاح بالنسبة لتحقيق الغرض .

أسئلة تقويم ذاتي



- ١/ أذكر العوامل التي تؤثر في السلسلة الزمنية ؟
- ٢/ عدد الأغراض من تحليل السلسلة الزمنية؟
- ٣/ سم الطرق المختلفة التي يتم بها عزل المتغيرات التي تطرأ على ظاهرة ما في زمن معين؟

٤.١ المتوسطات المتحركة - تمهيد السلاسل الزمنية

إذا كانت لدينا مجموعة من الأرقام :

$$Y_1, Y_2, Y_3, \dots, Y_n \quad \text{-----} (2)$$

فإننا نعرف الوسط المتحرك من الدرجة N بأنه يعطى بمتتابعة من الأوساط الحسابية .

$$(3) \quad \frac{Y_1 + Y_2 + \dots + Y_n}{N}, \frac{Y_2 + Y_3 + \dots + Y_{N+1}}{N}, \frac{Y_3 + Y_4 + \dots + Y_{N+2}}{N}$$

المجاميع في البسط بالمعادلة (٣) تسمى المجاميع المتحركة من الدرجة (N) .

مثال

إذا كان لدينا الأرقام 2, 6, 1, 5, 4 .

$$\frac{2 + 6 + 1}{3}, \frac{6 + 1 + 5}{3}, \frac{1 + 5 + 6}{3}, \frac{5 + 3 + 7}{3}, \frac{3 + 7 + 2}{3}$$

ومن المعتاد أن نضع كل رقم في الوسط المتحرك في مكانه الملائم بالنسبة للبيانات الأصلية . في هذا المثال يجب أن نكتب :

البيانات الأصلية 2, 6, 1, 5, 3, 7, 2

الوسط المتحرك من الدرجة ٣ هو 3, 4, 3, 5, 4

كل رقم في الوسط المتحرك عبارة عن متوسط الأرقام الثلاثة الواقعة فوقه . إذا كانت البيانات

معطاة سنوياً أو شهرياً ، فإن الوسط المتحرك من الدرجة N يسمى على الترتيب N . سنة متوسط

متحرك أو N شهر متوسط متحرك .بهذا نتحدث عن ٥ سنوات متوسطات متحركة ، ١٢ شهرا متوسطات متحركة ،.. وغيرها.ومن الواضح أنه يمكن استخدام وحدات أخرى للزمن .
المتوسطات المتحركة لها خاصية أنها تتجه إلى التقليل من كمية الاختلاف الموجودة في مجموعة من البيانات في حالة السلاسل الزمنية تستخدم هذه الخاصية لاستبعاد التقلبات غير المرغوب فيها وتسمى العملية بتمهيد السلاسل الزمنية .

إذا استخدمنا في (٣) ، الوسط الحسابي المرجح ، وكانت الترجيحات محددة مقدما، فإن المتتابة الناتجة تسمى الأوساط المتحركة المرجحة من الدرجة N .

تدريب (١)

٢/١ / الجدول التالي يوضح متوسط الإنتاج الشهري ، في بلد معين ، من فحم البيتومينس بمليون الكيلوجرامات للسنوات من ١٩٤٨ - ١٩٥٨ . أحسب :

(أ) ٥ سنوات متوسط متحرك .

(ب) ٤ سنوات متوسط متحرك.

(ج) ٤ سنوات متوسط متحرك مركزي .

السنة ١٩٤٨ ١٩٤٩ ١٩٥٠ ١٩٥١ ١٩٥٢ ١٩٥٣ ١٩٥٤ ١٩٥٥ ١٩٥٦ ١٩٥٧ ١٩٥٨

أسئلة تقويم ذاتي



١. أشرح مفهوم الوسط المتحرك من الدرجة (N).
٢. أكمل العبارة التالية:
 - المتوسطات المتحركة لها خاصية.....
 - تستخدم هذه الخاصية في.....
 - وتسمى هذه العملية ب.....

٢ . تقدير الاتجاه العام

يمكن تقدير الاتجاه العام بعدة طرق :

طريقة المربعات الصغرى	←	تقدير الاتجاه العام ←
طريقة التمهيد باليد	←	
طريقة المتوسط المتحرك	←	
طريقة أشباه المتوسطات	←	

- أ/ **طريقة المربعات الصغرى** : المعطاة بالفصل السادس من الممكن استخدامها للحصول على معادلة لخط الاتجاه العام الملائم أو لمنحنى الاتجاه العام . من المعادلة يمكن أن نحسب القيمة الاتجاهية T .
- ب/ **طريقة التمهيد باليد** : والتي تتكون من توفيق خط الاتجاه العام أو منحنى الاتجاه العام الذي يمكن استخدامه لتقدير T بالنظر إلى الشكل البياني . وعلى أية حال ، فهذه لها مضار حيث أنها تعتمد كثيرا على التقدير الشخصي .
- ج/ **طريقة المتوسط المتحرك** : باستخدام المتوسطات المتحركة من درجات ملائمة ، فإن الأنماط الدائرية ، الموسمية وغير المنتظمة يمكن حذفها ، تاركة فقط حركة الاتجاه العام .

مساوئ هذه الطريقة

- إنّ البيانات في بداية ونهاية السلسلة تفقد . ففي المثال أعلاه نبدأ بسبعة أرقام وباستخدام متوسط متحرك من الدرجة ٣ فستنتهي بخمسة أرقام .
 - إنّ المتوسطات المتحركة قد تولد تحركات دائرية أو غيرها ليست موجودة في البيانات الأصلية .
 - إنّ المتوسطات المتحركة تتأثر بشدة بالقيم المتطرفة.
- وللتغلب على هذه الصعوبة ، فإننا نستخدم أحيانا متوسطا متحركا مرجحا بأوزان ملائمة في هذه الحالة فان القيم المتطرفة أوزانا أقل .

د/ طريقة أشباه المتوسطات : تتكون من تقسيم البيانات إلى مجموعتين (يفضل أن تكون متساويتين) ثم نحصل على متوسط كل جزء ، وهذا يعطينا نقطتين على خط السلسلة الزمنية . وبرسم خط الاتجاه العام بين هذين النقطتين يمكن تحديد القيم الاتجاهية . ويمكن تحديد القيم الاتجاهية بدون الرسم . وعلى الرغم من أن هذه الطريقة بسيطة في تطبيقها إلا أن لها مساوئ منها:

- إنّها قد تؤدي إلى نتائج غير جيدة إذا استخدمت بدون تمييز .
- قابلة للتطبيق فقط في حالة ما إذا كان الاتجاه العام خطأ أو يقرب إلى خطين على الرغم من أنه يمكن مد صلاحيتها في الحالات التي يمكن تجزئة البيانات فيها إلى عدد من الأجزاء في كل جزء يكون الاتجاه العام فيه خطيا .

أسئلة تقويم ذاتي



١. أذكر طرق تقدير الاتجاه العام.
٢. أذكر مساوئ استخدام طريقتي المتوسطات المتحركة أشباه المتوسطات.

مثال : طريقة المربعات الصغرى :

فيما يلي بيان بالكميات المباعة من إنتاج احد المصانع في الفترة ١٩٧٠ إلى ١٩٧٦ بآلاف الوحدات :

السنوات	٧٠	٧١	٧٢	٧٣	٧٤	٧٥	٧٦
المبيعات	٧	١٠	١٢	١٥	١٨	٢٠	٢٤

المطلوب : إيجاد معادلة خط الاتجاه العام لهذه السلسلة؟

الحل:

بما أن السلسلة تتكون من عدد فردى من السنين هو ٧ لذلك نعتبر نقطة الأصل في منتصف السلسلة وهى سنة ١٩٧٣ وإذا أخذت انحرافات السنين عن نقطة الأصل نحصل على أرقام سهلة للمتغير (x) وهو الزمن كما يوضح الجدول رقم (٨:١) . وبتحديد قيمة (x) واعتبار قيم المبيعات (y) يمكن تطبيق قوانين حساب معاملي الانحدار الخطى البسيط كما يلي :

$$\beta = \frac{\sum x_i Y_i}{\sum x_i} , \quad \alpha = \frac{\sum Y_i}{N}$$

ملحوظة:

نلاحظ الاختلاف في معاملي الانحدار هنا عنهما في الباب السادس عند دراستنا للانحدار الخط البسيط وهو أن مجموع انحرافات القيم عند منتصف السلسلة يساوى صفرا ، أى أن :

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

حساب معادلة خط الاتجاه العام :

الزمن	x	Y	xY	X ₂
١٩٧٠	-3	7	-21	9
١٩٧١	-2	10	-20	4
١٩٧٢	-1	12	-12	1
١٩٧٣	0	15	0	0
١٩٧٤	1	18	18	1
١٩٧٥	2	20	40	4
١٩٧٦	3	24	72	9
المجموع	$\Sigma x = 0$	$\Sigma Y = 106$	$\Sigma xy = 77$	$\Sigma x^2 = 28$

$$\therefore \alpha = \frac{106}{7} = 15.143$$

$$\beta = \frac{77}{28} = 2.75$$

وعليه فإن معادلة خط الاتجاه العام هي :

$$Y = 15.143 + 2.75x$$

حيث نقطة الأصل هي سنة ١٩٧٣ . ونلاحظ أن قيمة Y عند سنة ١٩٧٣ هي القيمة الاتجاهية لهذا العام . ولإيجاد القيمة الاتجاهية لأي عام نعوض في المعادلة خط الاتجاه العام فمثلا القيمة الاتجاهية لعام ٧٦ هي :

$$Y_{1976} = 15.147 + 2.75(3) = 23.4$$

ويمكن استخدام معادلة الاتجاه العام كذلك للتنبؤ بالقيمة الاتجاهية لأي سنة خارج السلسلة بمسافة معقولة (أي ليست بعيدة). فمثلا القيمة الاتجاهية للمبيعات في سنة ١٩٨٠ هي :

$$15.147 + 2.75(7) = 34.4$$

ملحوظة :

إذا كانت السلسلة الزمنية تحتوى على عدد زوجي من السنين فلا توجد عندئذ سنة في منتصف السلسلة تجعل $\sum x = 0$ ولذلك يجب أن تؤخذ وحدة الزمن مساوية لنصف السنة أي ٧٣.٥

نشاط

1 / أوجد القيم الاتجاهية لبيانات الجدول التالي باستخدام طريقة أشباه المتوسطات ، حيث نأخذ كمتوسط :

أ - الوسط الحسابي ب - الوسيط =

السنة	البيانات	٥ سنوات مجموع متحرك	٥ سنوات متوسط متحرك
١٩٤٨	٥٠.٠		
١٩٤٩	٣٦.٥		
١٩٥٠	٤٣.٠	٢١٢.٩	٤٢.٦
١٩٥١	٤٤.٥	٢٠١.٠	٤٠.٢
١٩٥٢	٣٨.٩	١٩٧.١	٣٩.٤
١٩٥٣	٣٨.١	١٩٢.٨	٣٩.٦
١٩٥٤	٣٢.٦	١٩٠.٠	٣٨.٠
١٩٥٥	٣٨.٧	١٩٢.٢	٣٨.٤
١٩٥٦	٤١.٧	١٨٧.٩	٣٧.٦
١٩٥٧	٤١.١		
١٩٥٨	٣٣.٨		

٣ . تقديرات المتغيرات الموسمية

الدليل الموسمي

لتحديد المعامل الموسمي S في السلسلة الزمنية يجب أن نقدر كيف تتغير البيانات في السلسلة الزمنية من شهر إلى شهر خلال سنة نموذجية . مجموعة الأرقام التي توضح القيم النسبية لمتغير خلال اشهر السنة تسمى الدليل الموسمي للمتغير . فإذا كنا نعلم على سبيل المثال أن أرقام المبيعات خلال يناير ، فبراير ، مارس ، ... هي ٩٠ ، ١٢٠ ، ٥٠ ، ... في المائة من متوسط المبيعات الشهرية خلال العام كله ، فإن الأرقام ٩٠ ، ١٢٠ ، ٥٠ ، ... تعطي الدليل الموسمي ويشار إليها أحيانا بالأرقام القياسية الموسمية وسط (المتوسط) للدليل الموسمي للسنة كلها يجب أن يكون ١٠٠% ، أي أن مجموع الأرقام القياسية يجب أن يكون ١٢٠٠% .

وهناك عدة طرق متاحة لحساب الدليل الموسمي :

أ/ طريقة متوسط النسب المئوية : في هذه الطريقة يعبر عن بيانات كل شهر كنسبة مئوية من المتوسط في السنة . ثم نحصل على وسط النسبة المئوية للأشهر المتقابلة في مختلف السنوات وذلك باستخدام الوسط الحسابي أو الوسيط . فإذا استخدمنا الوسط الحسابي فمن الأفضل تجنب القيم المتطرفة والتي يمكن أن تحدث . وال ١٢ نسبة مئوية تعطي الدليل الموسمي . فإذا كان متوسطها ليس ١٠٠% (إذا كان المجموع لا يساوي ١٢٠٠%) فيجب تعديله بالضرب في معامل ملائم .

ب/ طريقة النسبة المئوية للاتجاه العام أو النسبة للاتجاه العام : في هذه الطريقة فإن بيانات كل شهر يعبر عنها كنسبة مئوية من القيم الاتجاهية في الشهر . وباستخدام وسط ملائم لهذه النسب للأشهر المتقابلة نحصل على الدليل المطلوب . وكما في الطريقة الأولى نعدل هذه القيم إذا لم يكن متوسطها ١٠٠% .

لا حظ أن قسمة كل من القيم الشهرية Y على القيم الاتجاهية T ينتج :

$$Y/T = CSI$$

من السلسلة الزمنية . وينتج عن عمليات الحصول على متوسط Y/T الأدلة الموسمية والتي تحتوى على التغيرات الدورية وغير المنتظمة وعلى وجه الخصوص إذا كانت كبيرة . وهذه قد تكون من المساوئ المهمة لهذه الطريقة ، في هذه الطريقة نحسب ١٢ شهرا متوسطا متحركا .

ج/ طريقة النسب المئوية للمتوسط المتحرك أو النسبة للمتوسط المتحرك: وبما أن النتائج التي حصلنا عليها تقع بين الأشهر المتتالية بدلا من وقوعها في منتصف الشهر كما هو الحال في البيانات الأصلية ، فإننا نحسب ٢ شهر متوسط متحرك لهذه ال ١٢ شهرا متوسطا متحركا . وتسمى النتيجة ١٢ شهرا متوسطا متحركا مركزيا . بعد ذلك ، نعبر عن البيانات الأصلية لكل شهر كنسبة مئوية من ال ١٢ شهر متوسطا متحركا مركزيا المقابل له . ويحسب الدليل المطلوب بأخذ متوسط النسب للأشهر المتقابلة . وكما سبق فإننا نعدل هذه النسب إذا لم يكن متوسطها ١٠٠% .

لاحظ أن السبب المنطقي وراء استخدام هذه الطريقة يجئ من طبيعة السلسلة الزمنية . ال ١٢ شهرا متوسطا متحركا مركزيا Y_i يعمل على استبعاد التحركات الموسمية وغير المنتظمة . S, I وهذه مكافئ للقيم المعطاة ب TC فإن قسمة البيانات على TC تنتج SI . والعملية التالية في الحصول على أوساط الأشهر المتقابلة تعمل على حذف المتغيرات العرضية I وهذا ينتج دليلا ملائما S .

د/ طريقة الوصلات النسبية : في هذه الطريقة يعبر عن بيانات كل شهر كنسبة مئوية من بيانات الشهر السابق له . ثم نحصل على متوسط ملائم للنسب الموصلة للأشهر المتقابلة.

ومن هذه الاثنى عشر متوسط النسب الموصلة يمكن أن نحصل على النسبة المئوية لكل شهر بالنسبة لشهر يناير والذي يعتبر مثل ١٠٠% . وبعد أن نفعل ذلك فإنه من المعتاد أن نجد أن شهر يناير التالي تقابله نسبة مئوية قد تكون إما أعلى أو أقل من ١٠٠% وهذا يعتمد على ما إذا كان الاتجاه العام في زيادة أو نقصان . باستخدام ذلك نقوم بتعديل النسب المئوية المختلفة التي حصلنا عليها بالأخذ في الاعتبار هذا الاتجاه العام . وهذه النسب المئوية النهائية ، والمعدلة بحيث يكون متوسطها ١٠٠% تعطي الدليل الموسمي المطلوب.

تحرير البيانات من أثر الموسمية :

إذا قسمنا البيانات الموسمية الشهرية الأصلية على الأرقام القياسية الموسمية المقابلة ، فإن

البيانات التي نحصل عليها تسمى بيانات لا موسمية أو بيانات معدلة لاستبعاد التغيرات الموسمية

. مثل هذه البيانات تتضمن الاتجاه العام ، التغيرات الدورية والتغيرات غير المنتظمة أسئلة تقويم

ذاتي



١. ؟ عرف الدليل الموسمي.
٢. أشرح طرق حساب الدليل الموسمي.
٣. كيف يمكن حساب البيانات الموسمية؟

٤ . تقدير التغيرات الدورية

بعد تخلص البيانات من اثر الموسم ، فإنه يمكن تعديلها أيضا للتخلص من اثر الاتجاه العام وذلك بقسمة البيانات ببساطة على القيم الاتجاهية المقابلة . وطبقا لطبيعة السلسلة الزمنية فإن عملية التعديل للتخلص من التغيرات الموسمية والقيم الاتجاهية تقابل قسمة Y على ST ، بحيث ينتج C, I أي التغيرات الدورية وغير المنتظمة . وباستخدام متوسط متحرك لعدد بسيط من الأشهر (٥ ، ٣ أو ٧ أشهر على سبيل المثال ، بحيث لا نحتاج إلى الحصول على قيم مركزية بعد ذلك) نستطيع استبعاد التغيرات غير المنتظمة I حيث يبقى فقط التغيرات الدورية . وطالما أمكن عزل هذه التغيرات فإنه يمكن دراستها بالتفصيل . فإذا كانت الدورات متكررة فإنه يمكن تكوين دليل الدورية بطريقة مشابهة لتكوين الدليل الموسمي .

٥ . تقدير التغيرات العشوائية أو غير المنتظمة

يمكن تقدير التقديرات العشوائية أو غير المنتظمة وذلك باستبعاد أثر الاتجاه العام والتغيرات الموسمية والتغيرات الدورية.

ويمكن تحقيق ذلك بقسمة البيانات الأصلية Y على S, C, T وينتج عن ذلك I من السلسلة الزمنية . ومن الناحية العملية وجد أن التحركات غير المنتظمة تتجه إلى أن تكون ذات حجم صغير وأنها غالبا ما تتجه إلى أن تتبع نمط التوزيع الطبيعي ، أي أن انحرافات صغيرة تحدث بتكرارات كبيرة ، إما الانحرافات الكبيرة فتحدث بتكرارات صغيرة .



أذكر كيف يمكن تقدير الآتي:

أ- التغيرات الدورية.

ب- التغيرات العشوائية أو غير المنتظمة.

٦ . قابلية البيانات للمقارنة

يجب التزام الحذر عند مقارنة البيانات حيث يجب أن تكون مثل هذه المقارنة ممكنة . على سبيل المثال ، فعند مقارنة بيانات شهر فبراير ، يجب أن نلاحظ أن شهر مارس ٣١ يوما بينما شهر فبراير قد يكون إما ٢٨ أو ٢٩ يوما . كذلك عند مقارنة أشهر فبراير لسنوات مختلفة يجب أن نتذكر أنه خلال السنوات الكبيسة يكون شهر فبراير ٢٩ . كذلك فإن عدد أيام العمل خلال الأشهر المختلفة لنفس السنة أو لسنوات مختلفة قد تختلف نتيجة لأيام الأجازات ، الاضرابات أو الأعطال وغيرها . ومن الناحية العملية ، لا توجد قاعدة محددة لإجراء التعديلات الأربعة لهذه التغيرات . ويترك تقدير الحاجة لهذه التعديلات لتوجهات الباحث .

٧ . التنبؤ Prediction

الدراسة السابقة يمكن استخدامها في المشكلة الهامة الخاصة بالتنبؤ بالسلاسل الزمنية . وعلى أية حالة ، يجب أن نتأكد من أن المعالجة الرياضية للبيانات لا تحل في حد ذاتها المشكلة . ولكن بالجمع بين الإحساس العام ، والخبرة والقدرة على الحكم السليم للباحث وبين التحليل الرياضي يمكن أن يكون له قيمة في كل من التنبؤ طويل المدى والتنبؤ قصير المدى.

الخلاصة

عزيزي الدارس،

ما الذي تناولناه في هذه الوحدة من المقرر؟؟

هل بمقدورك أن تلخص أهم خطواته؟؟

هنا نعمل معاً على تلخيص الخطوات الأساسية في تحليل السلاسل الزمنية فيما يلي:

تلخيص الخطوات الأساسية في تحليل السلاسل الزمنية :

- ١/ أجمع البيانات الخاصة بالسلاسل الزمنية ، وابذل كل مجهود للتأكد من أن البيانات يمكن الاعتماد عليها . في جمع البيانات يجب أن نضع نصب أعيننا الهدف من تحليل السلسلة الزمنية. على سبيل المثال ، فإذا أراد شخص التنبؤ بسلسلة زمنية معينة قد يساعد ذلك الحصول على سلسلة زمنية على علاقة بها وكذلك معلومات أخرى . وقد يكون من الضروري تعديل البيانات لجعلها قابلة للمقارنة مثل التعديل للسنوات الكبيسة ، وغيرها.
- ٢/ أرسم السلسلة الزمنية ، لاحظ من الناحية الوصفية وجود الاتجاه العام طويل المدى ، التغيرات الدورية والتغيرات الموسمية.
- ٣/ أرسم منحني الاتجاه العام أو خط الاتجاه العام واحصل على القيم الاتجاهية باستخدام إما طريقة المربعات الصغرى ، طريقة التمهيد باليد ، طريقة المتوسطات المتحركة أو طريقة شبه المتوسطات.
- ٤/ إذا كانت هناك تغيرات موسمية ، احصل على الدليل الموسمي ثم عدل البيانات وذلك للتخلص من أثر الموسم أي ، جعل البيانات لا موسمية.
- ٥/ خلص البيانات اللاموسمية من أثر الاتجاه العام . بهذا تحتوى البيانات الناتجة (نظريا) على التغيرات الدورية أو غير المنتظمة . متوسط متحرك نستخدم فيه ٥ ، ٣ أو ٧ أشهر يفيد في حذف التغيرات غير المنتظمة وإظهار التغيرات الدورية .

٦/ أرسم التغيرات الدورية التي حصلت عليها في الخطوة الخامسة ، لاحظ أي تكرارية (أو شبه تكرارية) التي يمكن أن تحدث .
٧/ بتجميع نتائج الخطوات من ١ - ٦ مع أية معلومات أخرى متاحة ، أجرى التنبؤ (إذا كان مرغوباً فيه) وإذا كان ممكناً ناقش مصادر الخطأ وحجمه.

ختاماً ٠٠

نتمنى أن تكون مدركاً لما ورد في هذه الوحدة حتى نتمكن من الانتقال معاً لما يليها.

لمحة مسبقة عن الوحدة التالية

تتناول الوحدة القادمة من الجزء الثاني من المقرر موضوعاً لا غنى لأي دارس علمي عنه . حيث تتناول مبادئ نظرية الاحتمالات شارحة التجارب العشوائية ومبينة أنواع الاحتمالات والأحداث والنظريات العلمية في هذا الجانب.
تختتم الوحدة القادمة كما بدأنا بخلاصة لها ثم إجابات للتمرينات الواردة فيها وتعريفاً بالمصطلحات العلمية.
نرجو لك التوفيق فيما أنت مقبل عليه من دراسة هذه الوحدة.

مسرد المصطلحات

- السلاسل الزمنية Time Series :
عبارة عن مجموعة من المشاهدات أخذت على فترات زمنية نتيجة تعقب ظاهرة معينة لفترة زمنية طويلة نسبياً وبصورة متتابعة وأغلب الأحيان تكون هذه الفترة الزمنية منتظمة.
- المنحنى التاريخي Historical Curve :
هو الخط المتعرج الناتج عن رسم بياني يمثل السلسلة الزمنية المتضمنة للمتغير (y) تقابل الزمن (t) .
- الدليل الموسمي للمتغير:
هو مجموعة الأرقام التي توضح القيم النسبية لمتغير خلال أشهر السنة.