الفهرس

الوحدة الأولى

الاحصاء واساة

ك٦

2كغ الاحصناء Statistics and Life

1.2 ^ introductionوالخياة

يقتصر معنى الاحصاء لدى الشخص العادي على الأرقام والاستبانات وا لجداول العددية التي تصف ظاهرة معينة، أو على الرسوم البيانية والأشكال التصويرية التي تعرض التغير في ظاهرة ما خلال فتره زمنية أو في مناطق جغرافية متعددة، وما شابه ذلك.

فكثيراً ما نطالع في الصحف اليومية جداولاً تبين معدل سقوط المطر على الأماكن المختلفه في البلد أو سقوطه على منطقه معينة خلال فترة زمنية محددة.

وفي المدارس ترى رسوماًبيانية تظهر أعداد الطلبة في الصفوف المختلفة. أما في مكتب مدير التربية والتعليم لمنطقة معينة فقد ترى جداول تبين اعداد المعلمين واعداد الطلبة في مدارس تلك المنطقة. ولو أمعنت النظر لوجدت أن الاحصاء يقوم بوصف طرق متعددة لجمع البيانات والمشاهدات، ومن ثم يتم تنظيمها وعرضها كما تشاهد في الكراسات الاحصائية التي تصدرها الوزارات والمؤسسات.

أما المجالات التي يوفرها الإحصاء للباحثين فهي كثيرة ومتعددة ومنها على سبيل المثال لا الحصر الطرق العلمية والوسائل الكمية التي يزودهم بها . وتستعمل هذه الطرق والوسائل في الأنشطة الحياتية المختلفة.ففي العلوم تمتد التطبيقات الإحصائية من مجال تحليل البيانات والتفسير والتنبوء إلى مجال تصميم التجارب مرورا بتقدير المجاهيل المطلوب معرفتها إما بنقطة أو بفترة ثقة ومن ثم إلى اختبار الفرضيات.

أمافي الصناعة فهناك التطبيقات القصيرة المدى مثل القرارات العملية اليومية، والتطبيقات طويلة المدى في التخطيط واتخاذ القرارات.

وكثيرا ما تستعمل المؤسسات والشركات الطرق الإحصائية لتحليل نماذج التغير والتتبؤ المستقبلية لتلك المؤسسات أو للأقتصاد بشكل عام لترسي قواعد التخطيط والتحكم.

وبالإضافة إلى التنبؤ فإن حقولا مثل التحكم في الانتاج، والتحكم في النوعية، تستعمل الطرق الإحصائية كقاعدة أساسية.

كذلك تستعمل المفاهيم الإحصائية في دراسات القوى العاملة، واختيار الموظفين، وأبحاث التسويق، والإعلان، والتحكم في المخزون، والتجارب الصناعية، والتحليل المالي وتدقيق الحسابات وتوظيف رؤوس الأمه ١١ ' والتنمية. فهل هناك أوسع من هذه المجالات !. أما في علم الادارة العامة والعلوم اي ة فإن الطرق الإحصائية تستعمل بكثرة وخاصة في تحليل القضايا السياسية والاجتماعية. ويستعمل الاحصاء في انجاز الدراسات

الكمية عن المجتمعات وظاهرة الفقر وحوادث المركبات وفي النماذج الانتخابية والأمور التربوية وقضايا الصحة العامة؛ كالعلاقة بين التدخين وأنواع متعددة من المرض، والعلاقة بين ضغط. لدم، والعمر، والوزن.

ويعتمد الباحثون إلى درجه كبيرة على مسح العينات حين يرغبون في الحصول على البيانات الإحصائية التي تتعلق بأنشطة الانسان والأمور المتعلقه به، إن كثيراً من الطرق الإحصائية قد تم استعمالها في السابق كما تستعمل في الوقت الحاضر لجمع البيانات وتحليلها وعرضها لغرض التخطيط واتخاذ القرارات.

ويلعب التحليل الإحصائي دوراً هاماً في كثير من حقول النشاط الانساني، وهو مفيد جداً في تبادل المعلومات والوصول إلى الاستنتاجات والاستدلالات من البيانات المتوفرة ومن ثم فهو مفيد في الإرشاد إلى التخطيط المنطقي واتخاذ القرارات .

وبالإمكان عزيزي الدارس، بعد ما تقدم من سرد أن نضع تتريفاً بسيطاً لعلم الإحصاء فهو العلم الذي يهتم بجمع البيانات وتصنيفها وتبويبها وعرضها. واستخراج المقاييس الإحصائية المناسبة للظواهر المدروسة لغرض التفسير والتليل والتنبؤ.

2.2 تطبيقات الإحصاء Applications of Statistics

ان الطرق الاحصائية متنوعة، وذات مدى واسع في مجالات تطبيقها، حتى يعجز المرء أو يكاد عن حصرها ولكن ساعرضها عرضاً سريعاً، ولكي أعطيك فكرة عن قوة الطرق الإحصائية أسرد لك بعض الأمثلة:

1.2.2تحليل الانحدار والارتباط

Regression Analysis and Correlation

يعد تحليل الانحدار إحدى الطرق الإحصائية التي تستعمل لدراسة العلاقة بين الظواهر المختلفة وإيجاد نماذج رياضية تصف اعتماد بعض هذه الظواهر على ظواهر أخرى، وقد تم استعماله في تحليل (اقتران الاستهلاك)، وهو العلاقة بين نفقات المستهلك ودخله.كما يستعمل تحليل الانحدار في كثير من البيانات الاقتصادية العددية لتقدير الاتجاه السابق والحالة الحاضرة ومن ثم التنبؤ حول النشاط الاقتصادي المستقبلي.

ويستعمل موضوع الارتباط في تقدير وجود علاقة بين متغيرين أو ظاهرتين، وفي وصف هذه العلاقة من حيث كونها خطية أو غير خطية، طردية أم عكسية، ومن ثم قياس قوة هذه العلاقة.

ومن الأمثلة على الإنحدار والإرتباط دراسة العلاقة بين نتائج امتحان شهادة

الدراسة الثانوية العامة وتحصيل الطالب في الجامعة، فباستعمال الانحدار والارتباط يستطيع المرء تقدير العلاقه بين الظاهرتين المذكورتين. (المعدل في امتحان شهادة الدراسة الثانوية والمعدل في الجامعة)

وعندما توضح العلاقة يمكن الإجابة عن الاسئلة التالية :

هل العلاقة طردية، بمعنى أنه كلما ارتفع المعدل في امتحان شهادة الدراسة الثانوية ارتفع المعدل في الجامعة، والعكس بالعكه؟ وهل العلاقه خطية، أي أن هناك ارتفاعا ثابتا في المعدل في الجامعة لكل ارتفاع وحدة واحدة في معدل امتحبان شهادة الدراسه الثانويه؟.

هل تستطيع التعبير عن المعدل في الجامعة، رياضياً بدلالة المعدل في امتحان شهادة الدراسة الثانوية؟ وكمثال آخر على استعمال موضوع الإنحدار والإرتباط تأتي دراسة العلاقه بين تكاليف الدعاية والاعلان عن سلعة معينة وبين المبيعات من تلك السلعة.

2.2.2 التحكم في النوعية Quality Control

تستعمل الطرق الاحصائية في تنفيذ التحكم في نوعية المنتوجات المصنعة، وفي إبقاء معدل مستوى العمليات الإنتاجية ضمن حدود مقبولة، هذا إضافة إلى قياس التغير وضبطه في تلك العمليات.فعلى سبيل المثال، في أحد مصانع مساحيق صابون الغسيل تعتبر العبوة صالحة إذا كان وزنها 200 غم مثلا، وغير صالحة إذا زاد الوزن أو نقص عن ذلك كثيراً .ويستطيع المسؤول عن المصنع باستخدام الطرق الإحصائية أن يقرر نسبة الصالح من العبوات، وأن يحدد الكمية التي يمكن التغاضي عنهاضمن حدود الخطأ المقبول لديه.وبالتالي فإنه يستطيع في أي وقت يشاء أن يقرر الاستمرار في تعبئة العبوات أو التوقف عن ذلك حتى يتم اصلاح الآلات الضرورية.وهو بذلك يقرر ما إذا كان الاختلاف في الوزن قد حدث نتيجة عطل في الآلات أو كان بسبب الصدفة.

وفي مثل هذه التطبيقات تستعمل الطرق الاحصائية للتفريق بين التغيرات التي يمكن عزو ها إلى الصدفة وتلك التي تكون كبيرة جداً فلاتعزى للصدفة. إذ أن التغيرات من النوع الثاني يمكن تحليلها ومعالجتها، ويؤدي التحكم في النوعية في كثير من الأحيان إلى تحسينات ملموسة في نوعية الإنتاج وفي تخفيض التكلفة عن طريق تقليل إعادة العمل والإتلاف.

3.2.2 طرق Sampling MethodsA^y

لا يتمكن الباحث في كثير من الدراسات من القيام بإجراء مسح شامل لجميع عناصر الموضوع قيد الدراسة، أي جميع عناصر المجتمع. وقد يكون ذلك إما لأن دراسة العنصر تؤدي إلى إهلاكه، مثل معرفة صلاحية الذخيرة، فهي لا تتم إلا باطلاق تلك الذخيرة، وإما لأن المسح الشامل للمجتمع كله يتطلب وقتاً وجهداً كبيرين علاوة عن التكاليف المالية الباهظة والحاجة إلى الإستعانة بفنيين كثيرين.

في مثل هذه الحالات يكتفى بدراسة عينه تؤخذ من المجتمع، إذ تستعمل الطرق الإحصائية لاختيار تلك العينة. ومن ثم تجصع البيانات المتعلقة بها وتحلل للخروج باستنتاجات يمكن تعميمها على المجتمع، ويتم كل ذلك باستعمال الطرق الإحصائية.

ومن الأمثلة على ذلك، دراسة عينة من العائلات في بلد معين لتحديد أنماط الدخل والإنفاق، أو المستوى المعيشي لجميع سكان ذلك البلد بناء على النتائج التي يتم التوصل إليها من دراسة العينة.

ولقد أصبح من المألوف والمرغوب فيه استعمال الإحصاء في تحديد حجم العينة، وكيفيه اختيارها، وتحليل نتائجها، وخاصة في الدراسات السكانية والدراسات التربوية والاجتماعية والاقتصادية.

والواقع أن استعمال الإحصاء لا يقتصر على ما سبق بل يتعدى ذلك إلى حقول المعرفة الأخرى كالطب والزارعة والصناعة والعلوم، حيث يحتاج الباحث في هذه المجالات إلى استخدام الطرق الاحصائية للتوصل إلى إجابات علمية عن تساؤلات مثل : ما العلاقة بين التحصيل العلمي للطلبة والمؤشرات الاجتماعية والاقتصادية لعائلاتهم؟ - ما أنماط البطالة في بلد معين، وما علاقة هذه الأنماط بالمستوى العلمي أو العمر؟ - ما العلاقات الوراثية بين الآباء والأبناء؟

- ما العلاقه بين البدانة والجلطة القلبية؟
- ما أثر أنماط معينة من الري و أنفاع معينة من السماد في ناتج المحاصيل؟ وغير ذلك كثير.

4.2.2 تصميم التجارب Design of Experiments

يعتبر موضوع تصميم التجارب أحد التطبيقات الهامة في الإحصاء ويحتاجه الباحث أو الدارس عند إجراءكثيرمن البحوث العلمية التي لابدأن يكون من متطلباتها: أ- تحديد الهدف من البحث.

ب- إجراء تجارب عملية، ومعرفة ما إذا كان من الكافي دراسه عينة؟ وفي هذه الحال يجب تحديد مجتمع الدراسة، وتعيين المتغير المراد قياسه وهو ما يسمى بالمتغير التابع.

ج- تحديد المتغيرات الأخرى التي تؤثر بالمتغير التابع، وتسمى هذه المتغيرات بالمتغيرات المستقلة وذلك لتحديد العلاقات بينها وقياسها.

وبعد تحديد هذه النقاط يتم تصميم التجربة باختبار أحد النماذج الاحصائية المناسبة لها، ثم تجمع البيانات ويقام بتخليلها بغية التوصل إلى استنتاجات وقرارات محددة. وتستعمل نماذج تصميم التجارب في الدراسات الزراعية والطبية والتربوية والصناعية وغيرها.فيمكن على سبيل المثال دراسة العلاقات والتأثيرات على الناتج الزراعي لمحصول معين باعتباره المتغير التابع واعتبار أنواع الأسمدة وطرق الري وتوقيتها وأنواع البذور المستعملة متخيرات مستقلة.

5.2.2 الإحصاءات الحيوية Vital Statistics

أظهرت الحاجة في الكثير من الدول والمنظمات الإقليمية والدولية إلى أهمية توفير بيانات تخص الحالات والمؤشرات الاجتماعية ومن أهمها تعداد السكان Census والتي توفر بيأنات عن الحالات الحيوية ذات الأهمية البالغة مثل عدد الولادات والوفيات والعمر والجنس والنساء في سن الحمل ومعدات الخصوبة وغيرها لما لهذه البيانات من أهمية كبيرة في التخطيط لمراقبة الأمراض والحد منها وتحديد معدلات النمو السكاني وبناء جداول الحياة LifeTables.

أسئلة التقويم الذاتي (1)

اعط أمثلة عن بعض الطرق الاحصائية التي تستعمل لدراسة العلاقة بين الظواهر.

هجدشأ تتطلب البحوث في كثير من العلوم جمع المعلومات عن العالم.من حولنا وتنظيمها.تسمى مثل هذه المعلومات بيانات.ويلعب علم الإحصاء دوراً هاماً في تحديد طرق جمع البيانات المطلوبة، والحكم على مدى صلاحية هذه الطرق، ومدى تمثيل البيانات المجموعة للمجتمع الذي جمعت منه.

وبعد أن تتم عملية جمع البيانات يأتي دور علم الإحصاء في تنظيم هذه البيانات، وعرضها وتحليلها و استقراء النتائج منها.ومن هنا نشأت صلة وطيدة بين البحث العلمي والإحصاء .وتظهر هذه الصلة عند دراستنا لخطوات البحث العلمي، التي يمكن تلخيصها ليمالي: ١١

3ل1 تحديد مشكلة البحث وأبعاده وأهدافه Defining The Research Problem, its Dimensions and Goals

عزيزي الدارس، قبل إجراء أي بحث يجب أن تكون هناك مشكلة ما تتطلب حلا، بعبارة أخرى أن هناك مشكلة قابلة للبحث وهذا من أهم الاعتبارات التي يجب أن تراعى عند تحديد مشكلة البحث، كما ويفضل أن تكون المشكلة قيد البحث أصلية ما أمكن، وغير مكررة وذات قيمة وكنلك يجب مراعاة الكلفة والوقت لإنجاز البحث.

إن تحديد مشكلة البحث من قبل الباحث قد تأتي من خبرته الشخصية أو إحساسه بالواقع الذي يعيشه وتفاعله معه.وعند ذاك على الباحث أن يقرر الإطار الذي يقع ضمنه البحث والخطة التي سيسير عليها والعوامل التي ستتم دراستها والطريقة التي سيتبعها.

•ئل(1) ■ قع

شعر باحث من أعضاء هيئة التدريس في جامعة ما أن هناك تذمراً لدى الطلبة فيما يتعلق بتسجيل المقررات المطلوبة منهم. فقام بتحديد أهمية مشكلة تسجيل المقررات وقرر أن يدرس أسبابها ويقدم اقتراحات عملية لازالة هذه الأسباب، ترى هل يعمد الباحث إلى دراسة المشكلة مع كل طالب في الجامعة أم يختار عينة يقوم بدراسة أفرادها وآرائهم ثم يحللها ويعمم نتائجه على أنها تتطبق على مجتمع الطلبة. بـــن ـ أ---------------------------------ئ-------------- —— ——

أين يستعمل الإحصاء في هذا المثال؟ إنه يستعمل في تحديد مجتمع الدراسة، أي المجتمع الذي سيدرسه فعلا .كما يستعمل أيضا في تحديد المجتمع الهدف، أي المجتمع الذي يهدف الباحث إلى تعميم استنتاجاته عليه. كذلك يحتاج الباحث إلى تعريف السمة أو الخاصية التي يريد دراستها، وإلى تحديد وحدة المعاينة أو المشاهدة التي يسجل قياساته عنها، وبعد ذلك يكون عليه تقرير حجم العينة التي يأخذها وكيفية اختيارها.كل هذه المفاهيم تقع ضمن تطبيقات الإحصاء، وهكذا يتضح لك، عزيزي الدارس، حاجة البإحث إلى استعمال الطرق الإحصائية في البحث العلمي.

2.3 صياغة الغروض5ح5^0^ Stating

الفرض عبارة محددة حول سمات المجتمع أو صفاته تكون صالحة للإختبار، بحيث يستطيع الباحث قبولها أو رفضها أو الحكم بان المعلومات المتوفرة لديه غير كاقية لاتخاذ أحد القرارين.

وتتخذ صياغة الفرضية شكلين أساسيين وهما:

1- صيغة الإثبات حيث تصاغ الفرضية بشكل يثبت وجود علاقة بين المتغيرات Variables أو المشاهداتObservations.

2- صيغة النفي حيث تصاغ الفرضية بشكل ينفي وجود هذه العلاقة.

ففي مثالنا السابق (مثال 1) ربما يضع الباحث الفرضية التالية: إن سبب تذمر الطلبة حول تسجيل المقررات هو النقص في عدد شعب المقرر المطروح.

يقوم الباحث بجمع البيانات من الطلبة حول رأيهم في سبب التذمر، ومن ثم يحلل هذه البيانات ويختبر الفرضية السابقة.

وأنت ترى، عزيزي الدارس، أن استعمال الطرق الإحصائية هو حجر الأساس فيكلما سبق.

م (2)ج

يدعي. مدير مصنع للمصابيح الكهربائية أن معدل عمر المصباح أكثر من 1000 ساعة. إن هذا الإدعاء مجرد فرضية احصائية، فهي خاضعه للإختبار .ويمكنناذلك إحصائياًبأخذ عينة من هذه المصابيح وتسجيل أعمارهاثم تحليل البيانات التي نحصل عليها، والحكم بصحة الفرضية أو عدم صحتها.

3.3 تصميم البحث Research Design

قبل أن يبدأ الباحث في تنفيذ بحثه، عليه أن يختار المنهج الذي سيسير عليه.

ومن مناهج البحث المعروفة ما يلي:

أ- الدراسة الميدانية Field Study

وفي هذا النوع من البحوث يقوم الباحث بجمع البيانات عن الأفراد قيد الدراسة، وذلك إمابتعبئة الاستبانات أو اجراء الاختبارات أو المقابلة الشخصية، وغيرها.

وبعد جمع البيانات يتم تنظيمها وتحليلها واجراء اختبار الفرضيات التي يكون الدارس قد قام بصياغتها.وكما هو واضح فإن الإحصاء يلعب الدور الأكبر في تحليل نتائج الدراسة الميدانية واتخاذ القرارات بناء على ذلك،

ب- المنهج التجريبي Experimental Study Method

يستعمل لمنهج التجريبي لتحديد أسباب ظاهرة معينة أو تأثير مؤثر معين، إضافة إلى تحديد العلاقات بين المتغيرات المختلفة. وفي كثير من الأحيان تقسم العينة المدروسة إلى مجموعتين، تسمى الأولى المجموعة الضابطة وهي التي لا تجري عليها أي تجارب. وتسمى الثانية المجموعة التجريبية أي تلك التي تخضع للتجربة. يقارن الباحث النتائج بين المجموعتين، الضابطة والتجريبية ليقرر ما إذا كان هناك أي تأثير للتجربة التي قام بها. ولا حاجة إلى القول أن أي استنتاج علمي في مثل هذه الدراسات يفرض استعمال الطرق الإحصائية ويتم تبريره بناء عليها.

ج- المنهج التاريخي Historical Method .

يقوم الباحث عند استعماله المنهج التاريخي في البحث بدراسة ظاهرة معينة عبر حقبة زمنية محددة قد تبدأ بالماضي وتستمر في الحاضر. وهو يرجع في مثل هذه الدراسة إلى المصادر الرئيسة التي كانت معاصرة للظاهرة وتتضمن الآثار والوثائق والمخطوطات والرسائل وكتب الرحالة والتراجم والسير الذاتية والمذكرات الشخصية والمشاهدة الشخصية وروايات شهود العيان وغيرها. كما أنه قد يرجع إلى مصادر ثانوية كالدراسات والمعلومات التي تنقل عن المصادر الرئيسة، وبمقدور الباحث اجراء بحوث التاريخ دون استعمال الطرق الإحصائية، لكنه إذا أراد أن يعقد مقارنات معززة بالمقاييس العددية والتبريرات العلمية فإن عليه استعمال الإحصاءبشكل ملحوظ.

د- تحليل المحتوى Content Analysis

من أبرز الأمثلة على تحليل المحتوى الذي يستعمل فيه الإحصاء تحليل النصوص الأدبية لمترفة كونها من إنتاج كاتب واحد أو عدة كتاب.وكذلك تحليل

النصوص من لغات مختلفة للمقارنة بين هذه اللغات، وهناك دراسات احصائية كثيرة في هذا المجال.

4.3 تحديد مصادر جمع البيانات
Defining Data Collection Sources

بالإمكان تحديد مصادر جمع البيانات والمعلومات بنوعين رئيسيين وهما:

1- المصادر الأولية Primary Sources:

حيث أن البيانات في هذه المصادر متوفرة لدى أفراد معينين يتوجب على الباحث استحصالها منهم ويتم ذلك عن طريق:

أ- المقابلة الشخصية: وفيها يقوم أشخاص مدربون ومؤهلون بإجراء مقابلات شخصية مع الأفراد الذين يراد جمع البيانات عنهم ويقوم الشخص الذي يجري المقابلة بتدوين الإجابات المعطاة له عن الأسئلة التي يطرحها على من يريد مقابلته.

ب- المكالمة الهاتفية: وهي مثل المقابلة الشخصية، لكنهاتتم باستعمال الهاتف.

ج- الاستبانة: وهي عبارة عن مجموعة من الأسئلة تمت صياغتها بإحكام بحيث يكون لكل منها إجابة محددة. وتوزع أوراق الاستبانات على مجموعة من الأفراد يقومون بالإجابة عنها، ثم يتم جمعها إذا كان الأفراد في مكان واحد أو يسهل الاتصال بهم لذلك الغرض. أما إذا كانت العينة التي يريد الباحث منها تتبئة الاستبانة أشخاصاً في أماكن متفرقة أو بعيدة فإن عليه استعمال الاستبانة البريدية، حيث يقوم بإرسال الاستبانات بالبريد. ولما كانت صياغة الاستبانة مهمة جداً فإنها يجب أن تكون على درجة متناهية من الإحكام، ولذلك يجب مراعاة عدة متطلبات في تصميمها.

د- المشاهدة المباشرة: حيث تجمع البيانات في هذه الحالة عن طريق المشاهدة المباشرة، مثل تسجيل عدد المكالمات الهاتفية التي تصل إلى مكتب الهاتف في الساعة.

2- المصادر الجاهزة (أو الثانوية) Secondary Sources

وهنا البيانات المتوفرة في الدوريات والنشرات و السجلات سواء كانت حكومية أو لمنظمات محلية أو إقليمية أو دولية. وعلى سبيل المثال سجل الصادرات والواردات لبلد ما، أو سجلات المرضى في مستشفى ما.

53 تنظيم البيانات وعرضها
Presentation & Organization of Data

بعد جمع البيانات يقوم الباحث بتنظيمها وعرضها بالطرق الجدولية، أو الرسومات البيانية، مثل المستطيلات أو الدوائروغيرها. وفي كثير من الأحيان يحتاج الباحث إلى تلخيص البيانات التي قام بجمعها في توزيعات تكرارية وعرض هذه التوزيعات بيانياً كالمضلع التكراري والمنحنى التكراري.

وسيأتي تفصيل ذلك في الوحدة الثانية.

6.3 تحليل البيانات واختبار صحة الفرضيات
Data Analysis & Hypotheses Testing

يبدأ التحليل الاحصائي بحساب بعض مقاييس النزعة المركزية، ومقاييس التشتت، والعزوم، والمقاييس التي تبرز الحاجة لحسابها عند تطبيق النظريات الاحصائية المتقدمة والتي يحتاجها الباحث لاختبار فرضياته، أو تحديدتقديراته وا لوصول إلى اتخاذ القرارات بصدد المشكلة التي يقوم بدراستها.

7.3 اتخاذ القرار والاستنتاج ووضع التوصيات
Decision Making, Conclusion, Recommendation

في ضوء نتيجة اختبار الفرضيات يمكن للباحث أن يتخذ قراره بقبول أو رفض أي من هذه الفرضيات ويبني استنتاجه حول البيانات التي تمت دراستها والمؤشرات الإحصائية التي تم احتسابها لغرض الوصول إلى التوصيات المتعددة حول مشكلة البحث.

وإذا ما نظرنا إلى الدور الذي يلعبه الإحصاء في تنفيذ البحث العلمي وجدنا أثره الهام في ذلك وخاصة في البنود الأربعة الأخيرة من مراحل البحث المذكورة.ويتجلى ذلك الأثر في تصميم وتنفيذ الدراسات البحثية، حيث يتم اختيار منهج البحث، وتحديد مصادر البيانات وطرق جمعها، ومن ثم تنفيذ جمع البيانات وتصنيفها وتحليلها واستقراء النتائج باتخاذ القرارات أو التنبؤات عن الظواهر التي مثلتها البيانات.

أسئلة التقويم الخاتي (2)

عدد المصادر التي يتم منها جمع البيانات.

4. الإحصاء الوصفي والإجصاء الاستنتاجي

Descriptive and inferentiai sgtfstics

بعد دراستك القسم الرئيسي الثاني (الاحصاء والحياة) يتضح لك أن هناك نوعين من الاحصاء، الأول هو الاحصاء الوصفي، وهو ذلك الجزء من الإحصاء الذي يعنى بجمع البيانات وتنظيمها وتصنيفها وعرضها عن طريق الجداول أو الرسوم البيانية وغيرها. فعندما يقوم مدير المدرسة بتسجيل عدد الطلبة في كل صف في مدرسته ومن ثم عرض هذه الاعداد على شكل مستطيلات فإنه يستعمل الإحصاء الوصفي في هذه الحالة. وعندما يزور طبيب الاسنان إحدى المدارم ويفحص أسنان كل طالب ويضع جدولاً يبين فيه عدد الطلبة الذين يعانون من تسوس في أسنانهم ثم يلخص هذا الجدول حسب عدد الأسنان التي أصابها التسوس فيذكر عدد الطلبة الذين يعانون تسوساً:في سن واحد، وعدد الذين يعانون تسوساً في سنين فإنه يستعمل الإحصاء الوصفي في هذه الحالة أيضاً.

أما النوع الثاني من الإحصاء فهو الاحصاء الاستتتاجي، وهو ذلك الجزء من الاحصاء الذي يعنى بتحليل البيانات ليتوصل إلى التنبؤ أو الاستقراء أو اتخاذ القرارات. ويلعب هذا الجزء من الاحصاء دوراً هاماً في تخطيط وتصميم التجارب التي تجمع من أجلها البيانات.

إذن فالإحصاء يهتم بطرق جمع البيانات وتمثيلها وعرضها (الاحصاء الوصفي) ومن ثم تحليلها وتفسيرها والتوصل إلى الاستنتاجات بناء عليها (الاحصاء الاستنتاجي).

والآن ما هو جمع البيانات؟

إنه عملية الحصول على القياسات أو التدادات أو قيم المشاهدات للتجارب التي يجريها الاحصائي، مستعملا في ذلك مختلف الطرق المتاحة له من إجراء مقابلة شخصية، أو الحديث على الهاتف، أو تعبئة استبانة، أو اجراء القياسات على التجارب الميدانية.

وكلما كان جمع البيانات دقيقاً زادت ثقة الدارس في الاعتماد عليها، ولا يكون تحليل البياناتصحيحاً ومفيداً إذا كان هناك أخطاء في جمع تلك البيانات.

وبعد أن يتم جمع البيانات يحتاج الباحث إلى (تنظيم وعرض تلك البيانات) أى القيام بعملية وضع البيانات في جداول منسقة وعرضها بطرق مناسبة : أشكال هندسية ورسوم بيانية، وتوزيعات تكرارية مما سيأتي ذكره في الوحدة الثانية.

كل ما سبق يقع ضمن الإحصاء الوصفي، ثم يأتي الإحصاء الاستنتاجي الذي يشمل:

أ-تحليل البيانات:

أي إيجاد قيم لمقاييس واقترانات متينة تتحدد قيمها من البيانات قيد الدراسة، فيحسب الباحث مثلاً الوسط الحسابي للبيانات، أو مدى تلك البيانات، أو بعض المقاييس التي تظهر له تباعد البيانات أو تقارب بعضها من بعض.

ب- استقراء النتائجواتخاذ القرارات:

وهو أهم أهداف علم الإحصاء وأكثرها فائدة حيث يشمل معظم الدراسات الاحصائية و النظريات القائمة عليها والتطبيقات العملية لها.وهو باختصار يتألف من الطرق الاحصائية التي تؤدي إلى الاستنتاجات التي يتوصل إليها الباحث من تحليل البيانات و غالباً ما تكون على شكل تقديرات أو تنبؤات أو قرارات رفض أو قبول للفرضيات الاحصائية.ويحتمد الإحصاء على نظرية الاحتمال وتطبيقاتها والنظريات الاحصائية التي بنيت عليها.

وهكذا، تصنف الطرق الاحصائية إلى طرق الإحصاء الوصفي، وطرق الإحصاء (الاستنتاجي) فالطرق التي تهتم بالبيانات المتوفرة فقط ولا تحاول التعميم من العينة المدروسة إلى مجتمع أكبر هي طرق الإحصاء الوصفي أما المعالجات التي تؤدي إلى تنبؤ أو استنتاج أو تعميم إلى مجموعة كبيرة كان قد تم مشاهدة بعض عناصرها فهي طرق الإحصاء الاستنتاجي.

5 ٠ تدزيج للقياسةح٦3ح5 Measurement

أنت في حياتك اليومية، عزيزي الدارس، كثيراً ما تقوم بإعطاء قيم عددية للمشاهدات أو الظواهر أو الحالات الاجتماعية، فمثلاً تقول : لدى عائلة خمسة أطفال ذكور وثلاث إناث، أو تقول: كانت علامة سعيد 87 من مائة في أحد الامتحانات، أو كانت درجة حرارة المريض 38.5 درجة مئوية، أو طول قطعة مستقيمة 0.2 1سم. هل هناك أية فروق في المقاييس التي استعملت في كل من هذه الحالات؟ هل التدريج على مقياس الحرارة يعطي نفس المعنى على المسطرة التي نقيس.بها أطوال المستقيمات؟ ما معنى: طول مستقيم صفر سم؟ وما معنى : درجة حرارة قطعة من الجليد صفر درجه مئوية؟ هل تحمل كلمة صفر في الحالتين المعنى ذاته؟

دعنا نحاول الإجابة عن هذه التساؤلات، إن ذلك يتطلب أن نقوم بدراسة تدريج القياس. وسننظر في ذلك إلى أربعة مستويات هي تلك المستويات التي يختلف فيها تدريج قياس عن تدريج آخر من حيث قابلية المقارنة، و إمكانية المقارنة بين المشاهدات. وهي تتدرج من أدنى مستوى يمكننا مقارنة المشاهدات فيه إلى أعلى مستوى وهذه هي:

1.5 التدريج الاسمي Nomina؛ Scaie

يستعمل التدريج الاسمي كمقياس لتحديد هوية الأفراد أو العناصر، وبالتالي فهو يصنف عناصر الظاهرة التي تختلف في النوعية لا في الكمية. ومن الممكن أن تكون التصنيفلت عبارة عن الأثواع المختلفة لظاهرة ما.وكثيراً ما تستعمل الأعدل د لتحديد هوية المفردات، وفي هذه الحالة لا يكون للعدد ذلك المدلول الكمي الذي يفهم منه عادة. فمثلا، يمكن استعمال العددين 1,0 ليدلا على التصنيف حسب الجنس فيجط الصفر ليدل على الذكر والعدد 1 ليدل على الأنثى. لاحظ أن العددين 1,0هنا ايدلان على القيم العددية، ولذلك لا تجري عليهما عمليات الجمع أو الطرح أو أخذ المعدل، هذا من جهة، ومن جهة أخرى، فانت تستطيع تعيين أي عددين آخرين ليدلا على الذكر والأنثى.

ومن الممكن أن تكون التصنيفات حسب الفئات أو المسميات مثل فئات الدخل، فنقول : الفئة الخاصة، الفئة العليا، الفئة الوسطى، الفئة الدنيا، أو حسب لون العين فنقول : زرقاء، سوداء، خضراء، عسلية.

أو حسب مكان الاقامة حيث تغطي أسماء المدن التي يقيم بها الأفراد. ومن الممكن أن نعطي هذه المدن أرقاماً لتدل عليها فيعطى سكان مدينة القدس الرقم 1، وسكان مدينة نابلس الرقم 2، وسكان مدينة إربد الرقم 3، وسكان مدينة عمان الرقم 4، وسكان

مدينه الزرقاء الرقم 5 . وكما ذكرنا سابقاً لا يكون للأرقام في هذا الحال ذلك المدلول الكمي للأعداد، وبالتالي فلا يقارن الرقم 5 مع الرقم 1 من حيث الكمية، أي اننا لا نستطيع أن نقول : الرقم 5 يحمل قيمه أكبر من الرقم 1، لأن المقصود هنا الأسم والهوية التي عبرنا عنها بالرقم، لا القيمة العددية لذلك الرقم.

2.5 التدريج الترتيبي Ordinal Scale

يقع هذا التدريج في مستوى أعلى من مستوى التدريج الاسمي، فبالإضافة إلى خواص التدريج الاسمي فإن التدريج الترتيبي يسمح بالمفاضلة، أي ترتيب العناصر حسب سلم معين.

وعندما تسجل قياسات العناصر أو المشاهدات حسب التدريج الترتيبي تكون هذه العناصر مرتبة من أعلى إلى أسفل أو بالعكس، مثل : الرتب العسكرية والتقدير الذي يحصل عليه خريج الجامعة وهو : مقبول، وجيد وجيد جداً، و ممتاز. وواضح أن التدريج الترتيبي يسمح بالمفاضلة أي بمعرفة أي شخص أو عنصر أطول أو أقصر من الآخر، أكبر أو أصغر، أفضل تقديراً أو أقل تقديراً، وهكذا.

3.5 تدريج الفترة Interval Scale

يعطي هذا التدريج معنى للفروق بين المشاهدات، وبذلك يقع في مستوى أعلى من مستوى التدريج الترتيبي وعليك أن تلاحظ أن وحدات القياس في تدريج الفترة متساوية، وأن الصفر في هذا التدريج رمز اصطلاحي وليس هو الصفر المطلق الذي يعني العدم.مثل هذا هو التدريج على ميزان الحرارة المئوي.وهو نوع من تدريج الفترة، ووحدة القياس فيه واحدة أي أن المسافة على الميزان بين أي درجتين هي نفسها المسافة بين أي درجتين أخريين. أما الصفر على ميزان الحرارة المئوي فهو اصطلاحي، أي أننا اصطلحنا أن نعين الدرجة صفر مئوية لتدل على درجه تجمد الماء النقي، ولذلك فإن الصفر على ميزان الحرارة المئوي لا يعني العدم وبالتالي فهو ليس الصفر المطلق. ويتضح هذا الأمر عندما تأخذ مثالاً آخر على تدريج الفترة، وليكن تدريج ميزان الحرارة الفهرنهايتي. لاحظ أن وحدة القياس في هذا التدريج هي واحدة، أي أنها هي نفسها بين كل درجتين ولاحظ أيضا أن الصفر في هذا التدريج صفر اصطلاحي وهو هنا يقابل (0-32 X 5/9) أي (17.78-) درجة مئوية، بينما يقابل الصفر على ميزان الحرارة المئوي 32 درجة فهرنهايت، وبالتالي فإن الصفر على كل من هذين التدريجين اصطلاحي. ويظهر

تدريج الفترة عدد الوحدات التي يكون فيها شخص أو عنصر أكبر أو أصغر، أكثر أو أقل، من شخص أو عنصر آخر . فإذا كانت درجة حرارة الماء في وعاء هي 90 درجة مئوية، وفي وعاء آخر 30 درجة مئوية فإن الفرق بين درجتي الحرارة هو 60 درجة مئوية، وبالتالي فإنك تستطيع إيجاد الفرق بين القيم على تدريج الفترة، ويكون لهذا الفرق معنى واضح، فنقول في مثالنا السابق: درجة حرارة الماء في الوعاء الأول أعلى من درجة حرارته في الوعاء الثاني بمقدار 60 درجة مئوية. ولكن لا معنى للنسب في تدريج الفترة، ففي مثالنا السأبق ليس هناك معنى لقولك: إن سخونة الماء في الوعاء الأول ثلاثة أضعاف سخونة الماء في الوعاء الثاني، حيث أنك لا تطيق وضع يدك في الوعاء الأول ولكنك تستطيع ذلك في الوعاء الثاني. إذا لا معنى هنا لعبارة: ثلاثة أضعاف. ويعود السبب في ذلك إلى أن التدريج على ميزان الحرارة المئوي تدريج فترة، ولا يوجد فيه صفر مطلق.

ومن الأمثلة على تدريج الفترة التدريج على مقياس الضغط الجوي (البارومیتر)، وقياس معامل الذكاء، والعلامات التي يحصل عليها الطلبة في الامتحانات.

4.5 التدريج النسبي Ratio Scale

يقع هذا التدريج في أعلى مستوى من مستويات القياس، وأهم فرق بينه وبين تدريج الفترة أن التدريج النسبي يعطي معنى للصفر المطلق، أي أن الصفر على هذا التدريج يعني العدم، وبالتالي فهو ليس صفرا اصطلاحيا بل إن له معنى محددا.ومن ثم فإنك تستطيع إعطاء معنى للنسب بين القيم المقاسة بالتدريج النسبي.وأهم الأمثلة على التدريج النسبي هي قياسات الطول والمساحة والحجم والوزن والقوة وغيرها.

فعلى سبيل المثال : إذا قلنا أن طول النقطة صفر فإن هذا يعني أن طولها معدوم. وبالتالي فهناك معنى للصفر المطلق. واذا قلنا : طول قطعة مستقيمة 20سم وطول قطعة مستقيمة أخرى 0 اسم فإن هناك معنى واضحاً لقولنا :طول القطعة الأولى ضعف طول القطعة الثانية.

ا:رم (؟]

شجثي ببعريج سري و ي سبي؟

6المتتيرات وؤابغهقا5ا- ر ذق '

إن من أهداف وغليات البحوث في العلوم الاجتماعية ان نفهم الظواه الاجتماعية ونفسرها ونصوغ التنبوءات . ومن أجل ذلك يجب تحويل أفكارنا عن هذه الظواهر إلى بيانات فعلية، عن طريق القياس.

وتسمى السمات أو الصفات التي نقيسها على أفراد عينة معينة متغيرات، وهكذا فإن المتغير هو تلك السمة، أو الصفة، أو الكمية، التي تتغير قيمتها من عنصر إلى آخر، أومن مشاهدة إلى أخرى. فلو أردت قياس أطوال طلاب أحد الصفوف لحصلت على عدد من القياسات يمثل كل منها طول أحد الطلبة، أي أن الطول متغير. وكمثال آخر، لو سجلت درجات الحرارة في مدينة معينة، كل يوم، لمدة شهر فإنك ستحصل على عدد من القيم التي تمثل درجات الحرارة، وبالتالي فإن درجة الحرارة في تلك المدينة تعتبر متغيرا. إليك أمثلة أخرى على المتغيرات:

1- ضغط الدم عند الشخص.

2- مستوى السكر في الدم عند الشخص.

3- عدد الأطفال لدى العائلة.

4- المصروف الشهري لطالب الجامعة.

والأمثلة على المتغيرات كثيرة جداً، ولا حصر لها، ومنها: كمية نزول المطر اليومي في فصل الشتاء، والصادرات، والواردات في بلد ما.

أما الثوابث فهي السمات والخواص التي لا تتغير؛ وهي تصف ماهية المواد في ظروف معينة مثل: الكثافة النوعية لعنصر ما في ظرف محدد. فمثلا: الكثافة النوعية للماء النقي في درجة الحرارة العادية هي 1 غم لكل سم مكعب. واليك أمثلة أخرى على الثوابت؛ فمعامل التمدد لعنصر الحديد النقي ثابت، ومعامل الاحتكاك بين مادتين محددتين ثابت. وهناك ثوابت كثيرة في القوانين الفيزيائية مثل تسارع جذب الأرض، ومثل ثابت نفاذية الفراغ للتأثير الكهربائي، العدد الذري لعنصر معين. فالعدد الذري للهيدروجين 1، والحدد الذري للذهب 79.

وهناك ثوابت من نوع آخر تقع ضمن اهتمامات الإحصائي أو الباحث في الدراسات الاجتماعية أو الاقتصادية أو الطبية أو الصناعية إو الزراعية وغيرها. وهي تلك الثوابت التي تصف المجتمعات مثل: معدل عمر المصابيح الكهربائية التي ينتجها مصنع معين على مدى عام واحد، أو معدل الدخل السنوي للفرد في بلد معين في سنة معينة. لاحظ هنا أن معدل عمر المصابيح في المصنع المذكور على مدى عام واحد

ثابت، لكنه غير معلوم. ولاحظ أيضاً انك لو أخذت عينة من مصابيح ذلك المصنع وأردت معرفة عمر الحياة لكل منها لوجدت أن عمر المصباح متغير.

in

المتغيرات الوصفية والمتغيرات الكمية
Qualitative Variables & Quantitative Variables

تصنف الظواهر من حيث طبيعة بياناتها

الحياة على عمر محدد لكل مصباح ويمكن أن يتغير هذا العمر من مصباح لآخر، وكنلك متدل أعمار المصابيح التي أخنتها في التينة، فهو متغير أيضاً وسيأخذ قيماً متعددة تبعاً للعينات المختلفة التي تأخذها من المصابيح لقياس أعمارها.

S

قهوتبب (1)

صنف ما يلي إلى متغيرات وثوابت:
أ- درجة تجمد الماء النقي تحت الضغط الجوي العادي على سطح البحر.
ب- طول الطالب في الصف الثالث الاعدادي في مدرسة ما.
ج- دخل الفرد السنوي في الأردن.
د- "لكثافة النوعية للذهب الصافي.
ه- العدد الذري للهيدروجين.
و- درجة الحرارة اليومية لبلد ما في الساعة الثامنة صباحاً.

أتسكلة!الققويم1لفأقي(4) ما هو المتغير ؟

أ- متغيرات وصفية (نوعية)

وهي الظواهر التي لا يمكن قياسها بالأرقام العددية مباشرة مثل الجنس، وتخصص الطالب، الجنسية ء والمستوى المعيشي (غني، متوسط، فقير، معدوم) إلى آخره من هذه الصفات.

ب- متغيرات كمية

وهي الظواهر التي يمكن قياسها بالأرقام العددية والتعبير عنها بها فمثلا عند قياس وزن الفرد نستعمل وحدات القياس كالكيلو غرام مثلا ويكون لكل فرد رقم عددي يمثل وزنه.ومثال آخر عدد المقاعد الدراسية لكل قاعة تدريس، إذ يعبر. عن ذلك برقم عددي كأن نقول بأن القاعة أ فيها 40 مقتداً والقاعة ب فيها 25 مقعداً.

1,7 أنواع المتغيرات الكمية

Types of Quantitative Variables

تقسم المتغيرات الكمية إلى نوعين:

أ- المتغيرات المنفصلة (المتقطعة) Discrete Variables

المتغير المنفصل هو المتغير الذي يأخذقيماًعددية صحيحة لاتضم كسراً، وتمثل عدد الوحدات أو المشاهدات الإحصائية، بعبارة أخرى هناك قفزات بين القيم التي يأخذها المتغير.

مثال (3)

عدد الأفراد الذين يسددون فواتير الكهرباء هو متغير منفصل، لأن عدد الأفراد يمكن أن يكون ......، 0، 1، 2، 3 ولو كان لدينا 20 فرداً يقومون بتسديد الفواتير في يوم ما، فان قيم المتغير هنا ستكون 0، 1، 2، 3، ......، 20. وهذه قيم بالإمكان عدّها.

مثال (4)

عدد المرضى الذين يتم إدخالهم إلى مستشفى ما في اليوم الواحد هو متغير منفصل، لأن عددهم يمكن أن يكون 0 أو 1 أو 2 أو 3 أو ...... وهكذا.

لاحظ عزيزي الدارس، أن عدد عناصر المجال يمكن عدها.

ب- المتغيرات المتصلة (المستمرة) Continuous Variable

هذا النوع من المتغيرات يأخذ قيماً ضمن مدى معين أو قيمة ما من القيم، وهذا

في الفقرة السابقة تعرفنا على المتغيرات وأنواعها من حيث طبيعتها وصفية أو

يعني عدم وجود قفزات بين قيمة كما هو الحال في المتغير المنفصل، أي أننا سنلاحظ وجود كسورا في قيم المتغير المتصل.

:--ن:

مثال (5)

الأجور المدفوعة لقاء استهلاك الكهرباء هي متغير متصل، فلو كانت قيمة إحدى الفواتير المسددة (25.350) دينار فهذا يعني وجود كسر قيمته (0.350) من الدينار وهي قيمة من بين عد غير محدود من القيم في مجال قيم الفواتير من 25 دينار إلى 26 دينار مثلا.

! ■٠٦

طول الطالب كذلك هو متغير متصل بسبب عدم وجود قفزات عند قياس الطول . فمهما كان طول أحد الطلبة قريباً من طول طالب آخر فإن هناك طالب ثالث سيقع طوله

بين طولي الطالبين السابقين.ومن ناحية أخرى لو أن وحدة قياس الطول المستخدمة هي السنتمتر وأن أجزاء السنتمتر هي المليمتر فهذا يعني وجود كسور للسنتمتر هي المليمتر.

تهوجبب (2)

صنف ما يلي إلى متغيرات وصفية وكمية ثم ت منفصلة ومتصلة :	سنف الكمية منها إلى متغيرات
أ- عدد المكالمات الهاتفية التي ترد إلى مقسم ما. ب- المعدات التراكمية لمجموعة من الطلبة.	و- الدخل الشهري للأسرة.
ج- مسقط رأس الفرد (مكان ولادته).	ز- درجة حرارة المريض. ح- عدد حولدث السير.
د- عمر الفرد.	ط- أنواع اللحوم.
ه- مستوى التحصيل العلمي.	

كمية، أو من حيث قيمها منفصلة أو. متصلة، وفي هذا البحث سنتناول العلاقة بين المتغيرات.

في دراستنا لبعض الظواهر الطبيعية أو الاجتماعية أو التربوية نجد أن هناك علاقات بين المتغيرات التي نقوم بمشاهدتها، وهي إما أن تكون خطية أو غير خطية وتكون بين متغيرين اثنين أو بين عدة متغيرات.فمثلاً يمكن دراسة العلاقة بين تحصيل الطالب في الجامعة وبين علامته في شهادة الدراسة الثانوية. ولو رمزنا إلى معدل الطالب في الجامعة بالحرف ٧ ولعلامته في شهادة الدراسة الثانوية بالحرف *، وافترضنا أن ٧ اقتران خطي في لا ٠٠ يسمى لامتغيراًمستقلاً، ويسمى ٧ متغيراًتابعاً. وفي هذه الحالة فإننا نستطيع تنبؤ قيمة ٧ المقابلة لأي قيمة معطاه للمتغير X . وهكذا فإن المتغير الذي نريد تقدير قيمته أو التنبؤ بها يسمى (متغيراً تابعاً) . أما المتغير الذي نعتمد عليه في الوصول إلى التنبؤيسمى (المتغير المستقل).

ولا يعني مصطلحا"مستقل" و "تابع" وجود علاقة سببية بين المتغيرات؛ أي لا يعني أن أحد المتغيرات سبب في حدوث المتخير الآخر، وإتما تعني أنه بمعرفة قيمة المتغير المستقل X يمكننا معرفة قيمة المتغير التابع ٧ بواسطة اقتران رياضي بين X و ٧ . وبهذا المعنى تكون قيم ٧ تابعة ومعتمدة على قيم لا، اما المتغير لافقد يكون وقد لا يكون سببا في التغير الذي يحدث في ٧.

فمثلاً، عند تقدير حجم المبيعات Y لأحد المتاجر من معرفة تكاليف الدعاية Xالتي ينفقها صاحب المتجر فإننا نستعمل المعالة الرياضية التي تربط المتغير Y بالمتغير *ا. وفي هذه الحالة يكون X هو المتغير المستقل و٧ هو المتغير التابع، لكن هذا لا يعني أن التغير في تكاليف الدعاية هو الذي يسبب التغير (الزيادة أو النقصان) في حجم المبيعات.

تهوتبب (3)

أراد طبيب دراسة تأثير العمر على ضغط الدم لدى الرجال البالغين، فدرس 30 حالة سجل فيها عمر الرجل وضغط الدم لديه. فما هو المتغير المستقل في هذه الدراسة؟ وما المتغير للتابع؟

تدريب (4)

قام باحث زراعي باجراء تجارب على قطع أرض متماثلة استعمل فيها كميات نتافة من الأسمدة وسجل كمية المحصول من كل قطعة، ثم وجد علاقة رياضية تربط كمية المحصول بدلالة كمية السماد.فما المتغير المستقل؟ وما المتغير التابع؟

أكفة التقيجبم اللفاًقي (5)

ماذا يعني مصطلح المتغير المستقل ومصطلح المتغير التابع؟

ؤ.:الم٠|٠ئ p٠iat|-a|-t٠ :ة, ا

درست في هذه الوحدة مراحل البحث العلمي ومنها تحديدمشكلة البحث وأبعاده وأهدافه وتصميم منهج البحث.وتتطلب هذه المرحلة من البلحث تعريف المجتمع الإحصائي تعريفا دقيقا.

المجتمع الإحصائي مم يتكون؟ إنه يتكون من مجموعة العناصر أو جميع الأفراد الذين ينضوون تحت لواء الدراسة الاحصائية.ويهدف تعريف المجتمع الاحصائي إلى تعيين الحدود الصريحة لعملية جمع البيانات ولعملية الاستقراء والاستنتاجات التي يمكن الحصول عليها من الدراسة.

ويمكن أن تكون عناصر المجتمع أفراداً أو عائلات أو موظفين أو مدارس أو مؤسسات أو غير ذلك . لكنه يجب أن يكون ذلك المجتمع معرفاً تماما ومحدداً بحدود الزمن والفراغ (الطبيعي والجغرافي) بحيث يستطيع الباحث معرفة انتماء أي عنصر لذلك المجتمع أو عدم انتمائه. وهذا ما يتم العمل عليه في أعمال المسح، وتسمى كشوف أعمال المسح التي تسجل فيها جميع عناصر المجتمع (الإطار) أو (إطار المعاينة) . أما (التعداد) فهو المسح الذي يحاول أن يضم كل عنصر في المجتمع. ونحن نقول (يحاول) لأن الباحث في المجتمعات الكبيرة جداً لن يستطيع الوصول إلى جميع الأفراد.

تعريف (1)

المجتمع هو مجموعة العناصر أو الأفراد الذين ينصب عليهم الاهتمام في دراسة معينة.

تعريف (2)

وحدة المعاينة Unit of observation : هي أي عنصر أو فرد في المجتمع الذي ندرسه.

تعريف (3)

العّينة : هي مجموعه جزئية من المجتمع.

1.9 طرق جمع البيانات الإحصائية

يتم جمع البيانات الاحصالية بإحدى الطرق الأية:

1.1.9 طريقة المسح الشامل

فيها تجمع البيانات من جميع أفراد المجتمع الإحصائي فمثلا:

إذا أردنا القيام بتعداد عام للسكان في بلد قليل السكان نسبياً، كالأردن مثلاً، فإننا نقوم بعد أفراد جميع العائلات في المجتمع الاردني، وإذا أردنا التعرف على مستوى طلاب جامعة اليرموك في الرياضيات فإننا نقوم برصد علامات جميع هؤلاء الطلاب في الرياضيات لدى المسجل العام، أو نقوم بتصميم فحص عام في الرياضيات نطبقه على جميع طلبة الجامعة ومن خلال نتائجهم في هذا الفحص، نحاول التعرف على مستوى طلبة الجامعة في الرياضيات.

2.1,9 طريقة العينة

هناك عدة حالات يتعذر فيها المسح الشامل وعندها نلجأ إلى دراسة جزء من المجتمع الإحصائي يسمى (العينة) .وحجم العينة هو عدد عناصرها.

ومن أسباب اللجوء إلى العينة الممثلة لملمجتمع الاحصائي ما يلي:

أ- فساد عناصر المجتمع نتيجة أخذ المشاهدات من تلك العناصر، فمثلا: إذا أردنا معرفة مدى صلاحية البيض الذي تنتجه مزرعة ما للأكل فمن غير المعقول أن نقوم بتكسير كل إنتاج تلك المزرعة من البيض للتأكد من صلاحيته، لذا نجد أن أخذ جزء من هذا البيض لفحصه أمر ضروري ومن ثم نعمم نتيجة الفحص على المجتمع الإحصائي، وهو هنا، إنتاج تلك المزرعة من البيض.

ب- غالباً ما يتعذر الوصول إلى جميع عناصر المجتمع الاحصائي وإجراء الدراسة عليها فمثلاً: إذا أردنادراسة كميات الأمطار التي سقطت في الأردن منذ عام 1930 حتى الآن من أجل التعرف على أثرها على كميات إنتاج القمح في الأردن.فقد لا يكون لدينا سجلات كافية تحتوي على المعلومات المطلوبة منذ عام 1930، لذا نكتفي بدراسة عينة من السجلات المتوفرة.

ج- إن دراسة العينة أدنى كلفة من المسح الشامل وتستغرق وقتاً وجهداً أقل.

د- قد يطلب إجراء المسح الشامل عدداً كبيراً من الأشخاص المدربين لإتمام جمع البيانات، أما العينة فتحتاج إلى عد أقل من هولاء وبذلك يتم التوفير في عدد العاملين. هذا من جهة، ومن جهة ثانية، إذا كان عدد المطلوبين قليلا فإنه يمكن تدريبهم على جمع البيانات بصورة أفضل، وبالتالي يمكن التقليل من الأخطاء والفروق الفردية في جمع البيانات.

ه- قد يحتاج الباحث إلى نتائج (سريعة) تعطيه مؤشرا عن ظاهرة معينة، وبالتالي فهو لا يستطيع دراسة المجتمع كله في ذلك الوقت القصير. لذا يلجأ إلى دراسة عينة فقط من ذلك المجتمع. ومن أهم الأمثلة على ذلك استطلاع الرأي العام بعد حوادث معينة،

أو قبيل الانتخابات، أو قبيل اجراء استفتاء عام حول موضوع مهم.

و- قد يكون المجتمع متصلاً، كأن تكون مجموعة عناصره غير قابلة للعد، وبالتالي يتعذر اجراء مسح شامل لها، مثل :دراسة مخزون الأردن من الفوسفات، أو دراسة مخزون دولة الكويت من البترول، حيث يتعذر إجراء مسح شامل، ويكتفى بدراسة عينة.

وفي بعض الأحيان، وحتى لو كان المجتمع منفصلا، أي أن مجموعة عناصره قابلة للعد، فإنه قد لا يكون من السهل إجراء مسح شامل للمجتمع، إذا كان ذلك المجتمع كبيرا جدا مثل سكان الصين لذلك يكتفى بدراسة عينة من ذلك المجتمع. والسؤال الذي ينشأ الآن هو:كيف نختار عينة من مجتمع احصائي؟

2.9طرق اختيار Sampling TechniquesM

درسنا معاً في الفقرة 9. 1 أن جمع البيانات يتم بإجراء المسح الشامل أو بطريقة العينة، وفي هذه الفقرة سنتعرف معاً على الطرق التي نختار بها العينة، ونبدأ ببعض

التعاريف لكي نعطي العبارات معانيها الدقيقة:

تحويبف (4)

المجتمع !^-Target Populations: هو كل المجتمع الذي نطلب المعلومات عنه.

فمثلاً: جميع طلبة الأردن من عمر 9 إلى 18 سنة.

أو: جميع الأسر في بلد معين.

أو :جميع المصابيح الكهربائية التي يصنعها مصنع في عام معين.

فعوتبف (5)

مجتمع الدر!Study Populations :هو مجموعة الأفراد التي يتاح لنا إجراء الدراسة عليها.

فمثلا: إذا كان المجتمع الهدف هو جميع طلبة الأردن من عمر 9 إلى 18 سنة، فمن الممكن أن يكون مجتمع الدراسة جميع الطلبة في مدينة عمان من عمر 9 إلى 18

ثد») ' ض

سمة !^^Population Characteristic : هي الخاصية أو الوجه الذي نريد قياسه لدى المجتمع.

فمثلاً: إذا أراد طبيب الوقوف على حالة سلامة الأسنان لدى الأطفال من عمر 6إلى 12 عاماً في المملكة الأردنية الهاشمية، فإن سمة المجتمع هنا هي : (الحالة الصحية لأسنان الطفل). ومن الواضح أن سمة المجتمع هذه يمكن قياسها بأحد تدريجات القياس التي مر ذكرها في الفقرة 5 من هذه الوحدة. إن قياس السمة يتغير من فرد إلى آخر، ولذلك فهو متغير. ونحن نحصل على سمة المجتمع من دراستنا لقياسات السمة لدى الأفراد جميعهم أو عينة منهم.

والهدف الرئيس هنا اختيار عينة تمثل المجتمع وتؤدي إلى إحراز معطومات عن سمة المجتمع بشكل دقيق يتناسب مع التكلفة والجهد المستعملين.

إن هدف دراسة التينة التوصل إلى استقراءات واستنتاجات عن المجتمع الهدف مبنية على معلومات في العينة. وهناك كميتان من المعلومات تحتويهما العينة، وبالتالي فهما يؤثران في دقة عملية الاستقراء التي نتبناها. الأولى منهما حجم العينة المختارة من المجتمع، والثانية مقدار التغير في البيانات، ويمكن التحكم في هذا عادة بطريقة اختيار العينة.

وهناك نوعان من الطرق لاختيار العينات.

1.2,9 العيناتغير لاحتمالية Non ٠ Probabilistic Sampling

وهي العينات التي لايخضع اختيارها إلى أي قوانين احتمالية، وإليك، فيمايلي، أهم أنواع العينات غير الاحتمالية:

1- المعاينة الوصولية Accessibility Sampling

باعتبار الأولوية لسهولة الوصول والحصول على المشاهدات، وفي هذه الحالة نختار العينة التي نستطيع أن نسجل القياسات والمشاهدات عن أفرادها بسهولة. فعلى سبيل المثال: إذا أردنا دراسة عادات التدخين بين طلبة الجامعة فإننا نطلب متطوعين للإجابة عن الأسئلة التي ستطرح.ومن سلبيات هذه الطريقة عدم كون العينة ممثلة للمجتمع بصورة واضحة جلية.

2- المعاينة الهادفة أو الحكمية أو الغرضية

Purposive or Judgemental Sampling

والنظرة هنا مختلفة عن سابقتها، إذ أن الباحث في هذه الحالة يختار عينته بناء على حكم شخصي ورأي ذاتي، إذ يعتبر العينة المختارة ممثلة للمجتمع، مع علمه أن المجتمع يحتوي على أنواع مختلفة من الأفراد ذوي مقاييس مختلفة وتباين في سهولة الوصول إليهم.ومن الممكن أن تعطي هذه الطريقة نتائج جيدة إذا كان رأي الباحث

وحكمه صائبين . وتحمل المعاينة الهادفة، في طياتها إلغاء مصادر التحريف المتوقعة، لكنه لا مناص من وجود التحريف، وذلك إما بسبب التحيز الشخصي، أو بسبب الجهل في جوانب بعض صفات المجتمع، وخاصة عند وجود ارتباط غير مكتشف بين طريقة المعاينة والمقياس الذي نريد دراسته.

2.2.9 العينات الاشالية Probability Sampling

هي تلك العينات التي يخضع اختيارها إلى قوانين احتمالية، ونورد لك فيما يلي أهم أنواع العينات الاحتمالية:

يالعينة العشوائية لبسيطة Simple Random Sample

قحويف (7)

إذا اخترنا عينة ذات حجم n من مجتمع حجمه N بحيث يكون لكل عينة ذات حجم •نفس امكانية الاختيار مثل أي عينة اخرى من الحجم نفسه فعندئذ، تسمى هذه الطريقة (المعاينة العشوائية البسيطة)، وتسمى العينة الحاصلة (عينة عشوائية بسيطة).

وتستعمل المعاينة العشوائية البسيطة للحصول على تقديرات لمعدلات المجتمع والمجموع الكلي للمجتمع، وللنسب الخاصة به.

وهناك جداول تسمى جداول الأرقام التشوائية يمكن الاعتماد عليها لاختيار مثل هذه العينات.

ي؟ كيفية اختيار العينة العشوائية البسيطة:

(أ) ليكن حجم المجتمع N، أعط كل عنصر من عناصرمجتمع الدراسة رقماً متسلسلاً من صفر إلى (1- N) وليكن عدد خانات الرقم المتسلسل هو عدد خانات الرقم الذي أعطي لآخر عنصر في المجتمع. فمثلا إذا كان حجم المجتمع 1500 وأردنا اختيار عينة عشوائية بسيطة منه، فإننا نعطي أفراد ذلك المجتمع الأرقام التالية:

, 1499, 1498, 1497,0002,0001,000

(ب) استعمل جداول الأرقام العشوائية واقرأ منه عمودياً بحيث يكون عدد منازل كل عدد مساوياًلعددمنازل الأرقام المتسلسلة في (أ). فإذاكان الرقم الذي قرأته من الجدول أحد الأرقام المتسلسلة فاعتبره عنصرامن عناصر العينة، وإلافاتركه وانتقل إلى عدد آخر، ولا تكرر الأعداد التي أخذتها عناصر في العينة ومعنى ذلك انك ترفض أي عدد أخذته في قراءة سابقة.

مثال (7)

عدد الطلبة المسجلين في مقرر الإحصاء (500)، والمطلوب اختيار عينة عشوائية بسيطة حجمها (15)من هؤلاء الطلبة.

الحل:

نعطي الطلبة أرقاماً متسلسة، تبدأ بالصفر وتنتهي بالعدد 499، وبذلك تكون أرقام الطلبة 000، 001، 002، ...، 497، 498، 499 .نذهب إلى جدول الأرقام العشوائية ونبدأ بالعمود الأول من اليسار ونقرأ الأعداد واحداً تلو الآخر بحيث ننظر إلى ثلاث منازل من جهة اليسار فإذا كان العدد الذي نقرؤه ضمن أرقام الطلبة نأخذه في العينة وإلا فاقرأ العدد الذي يليه، وهكذا.

نضع أمامك جزءاً من جدول الأعداد العشوائية لنشرح طريقة القراءة:

لاحظأن العدد الأول 702 لا يقع 70218

ضمن أرقام الطلبة فنرفضه ونقرأ (023)49

العدد الثاني 023 وهو من ضمن أرقام 93606

الطلبةفنأخذ الطالب رقم 023، (242)74

العدد الذي يليه 936 مرفوض والذي 67778

يليه 242 نأخذه والذي يليه 677 مرفوض

نستمر في هذه العملية حتى نحصل على العينة المطلوبة والتي حجمها 15، وفي مثالنا تكون العينة هي:

و17، 366، 199، 358، 066، 033، 242، 023

3و4، 368، 109، 480، 412، 028، 206

سقلكقؤيم1لااتي(6)____لكا

كيف يمكن اختيار عينة عشوائية بسيطة؟ ٤يالعينة الطبقية العشوائية Stratified Random Sampling

هدف تصميم إجراء العينة في هذه الحالة هو الحصول على أعظم كمية من المعلومات بتكلفة معطاة، والواقع ان طريقة العينة العشوائية البسيطة، وهي تصميم أساس في المعاينة، تؤدي في كثير من الأحيان إلى تقديرات جيدة لسمات المجتمع بتكاليف منخفضة.

ونتعرف الآن على طريقة ثانية تؤدي إلى زيادة في المعلومات عند إعط معينة.

فحرتبفه (8)

عند تجزئة المجتمع إلى مجموعات غير متداخلة ومن ثم اختيار عينة. عشوائية بسيطة من كل مجموعة فإننا نسمي هذه الطريقة : العينة الطبقية العشوائية، ومثال ذلك أخذ عينة من محافظة أو إمارة فيها ثلاث مدن كبيرة ومنطقة ريفية، والمجتمع الهدف المطلوب هو جميع الرجال والنساء من عمر 21 سنة فما فوق.

وتقضي طريقة العينة الطبقية العشوائية بأن تؤخذ عينة عشوائية من كل مدينة وعينة عشوائية من المنطقة الريفية، وهذا يعني أن المدن الثلاث والمنطقة الريفية تمثل أربع طبقات منفصلة عن بعضها البعض، ولذلك نأخذ عينة عشوائية من كل واحدة. وهناك ثلاثة أسباب تجعل المعاينة الطبقية العشوائية تؤدي إلى زيادة في المتلومات عند تحديد تكلفة معينة، هي:

أ- أن البيانات في كل طبقة بمفردها تكون أكثر تجانساً منها في المجتمع ككل.

ب- أن تكلفة تتفيذ المعاينة الفعلية تميل إلى الانخفاض عنهافي حالة المعاينة العشوائية البسيطة وذلك بسبب الملاءمة الادارية.

ج- عند استعمال المعاينة فإنه يمكن الحصول على تقديرات لمؤشرات المجتمع لكل طبقة دون الحاجة إلى أخذ عينات إضافية.

د- أن التجانس في كل طبقة (تقليل التغير ضمن كل طبقة) يؤدي إلى تقديرات تبايناتها أصغر من تباينات التقديرات التي نحصل عليها من التينة العشوائية البسيطة ذات الحجم نفسه.

وستدرس، عزيزي الدارس، موضوع التباين في الوحدة الثانية، ويكفي الآن فهم التباين بمعنى أنه يقيس تباعد أو اختلاف العناصر أو القيم عن بعضها البعض، فالتباين الصغير يعني أن التباعد أو الإختلاف بين القيم صغير، والتباين الكبير يعني أن التباعد أو الاختلاف بين القيم كبير.

فكيفية اختيار عينة طبقية عثوائية:

أول خطوة في المعاينة الطبقية العشوائية تعيين الطبقات بوضوح، وبعد ذلك يأتي وضع كل وحدة معاينة من المجتمع في الطبقة الملائمة.

قد لا تكون هذه العملية بسيطة كما تبدو، فمثلا: إذا أردت أن تحدد الطبقات التي

تنتمي إليها الأسر في وحدات حضرية وأخرى ريفية، فماذا تقرر عن قرية فيها 1000(ألف) نسمة؟ هل هذه القرية ريفية أم حضرية؟ يمكن أن نعتبرهأ ريفية إذا كانت بعيدة ؤمنفصلة عن مدينة كبيرة، أما إذا كانت بجوار مدينة كبيرة وقريبة منها فيمكن اعتبار ها حضرية.

من ثم أصبح ضرورياً أن تحدد ماذا تعني بكلمة :حضر، وكلمة ريف، لكي تستطيع وضع كل وحدة معاينة ضمن طبقة واحدة فقط.

بعد تقسيم وحدات المعاينة في طبقات، تختار عينة عشوائية بسيطة من كل طبقة باستعمال الطرق التي تم شرحها سابقا.

ويجب أن تتأكد أن العينات المختارة من الطبقات عشوائية ومستقلة عن بعضها البعض، بمعنى أنه يجب استعمال طرق معاينة عشوائية في كل طبقة على حدة، لكي نضمن أن المشاهدات التي نحصل عليها من طبقة معينة لا تعتمد على المشاهدات المختارة من طبقة أخرى.

*تقسيم العينة على الطبقات:

أنت تذكر بالتأكيد أن الغرض من تصميم العينة الحصول على تقديرات ذات تباينات صغيرة بأقل كلفة ممكنة و لقد سبق أن شرحنا ذلك. فإذا كان لديك k من الطبقات وأردت اختيار عينة حجمها n من المجتمع الكلي، فهناك عدة طرق لتقسيم الحجم P على الطبقات بحيث يكون:

٢١٦ :حجم العينة من الطبقة الأولى

2٢١ : حجم العينة من الطبقة الثانية

n_K : حجم العينة من الطبقة ذات الرقم k

بحيث يكون n : n٦ ب ٢١2 ... +*•0 ، أي أن حجم العينة الكلي يساوي مجموع العينات من الطبقات جميعها. ويمكن تحديد حجم العينة من أي طبقة بحيث يتناسب مع حجم الطبقة.فإذا كان حجم المجتمع N وكانت حجوم الطبقات كما يلي:

N1 حجم الطبقة الأولى

N حجم الطبقة الثانية

Nk حجم الطبقة التي رقمها k

وكان حجم العينة المطلوب P ، فإننا نحدد حجم العينة من كل طبقة كما يلي:

n٢N1xr

مثال (8)

إذا كان طلبة السنة الأولى في إحدى الجامعات موزعين حسب الكليات كما

يلي:

الهندسة	الاقتصاد	العلوم	الآداب
200	150	180	120

فكيف يمكن اختيار عينة حجمها (130) بطريقة المعاينة الطبقية العشوائية وعلى أساس

الكلية؟ الحل:

حجم المجتمع:

N = 200 + 150 + 180 + 120 = 650

من الواضح ان الطبقات هنا هي الكليات.

حجم العينة من كلية الاداب:

ا"اا =200x^ = 40
1 650

حجم العينة من كلية العلوم:

"2=150x^=30
2 650

حجم العينة من كلية الاقتصاد:

n, SOx^Se
3 650

حجم العينة من كلية الهندسة:

٢١ه=130-(40ه30 36) = 130-106 = 24

ثم نستخدم جدول الأرقام العشوائية، وبالطريقة التي شرحناها في مثال (7)، لاختيار عينات

عشوائية بسيطة من طلاب الكليات المختلفة:

40من200لطلاب الآداب،30من150 لطلاب لاقتصاد ...وهكذا.

يو ههركل .... ..

اذاكانت طبقات أحد المجتمعات تحوي العناصركمافي الجدول الآتي: الطبقة الاولى الثانية الثالثة الرابعة الخامسة

الحجم500 400 280 200 220

واراد باحث اختيار عينة حجمها 160 من هذا المجتمع، فما حجم العينة من كل طبقة؟

ت يداو - - "_رغ:م

المحاينة المنتظمة Sampling عأ٤ة٢٢مع٤*5٧5: إذا اخترناعنصراً بطريقة عشوائية من أول n من العناصر في إطار المعاينة، وكان رقم هذا العنصر >ا ومن ثم اخترنا عنصراً رقمه k من ثاني n من العناصر وعنصراً آخر رقمه k من ثالث n من العناصر. ..وهكذا فإن هذه الطريقة تسمى طريقة العينة المنتظمة.

تعتبر المعاينة المنتظمة بديلا مفيدا للمعاينة التشوائية البسيطة للأسباب التالية:

ا~ المعاينة المنتظمة أكثر سهولة في التنفيذ، ولنلك فهي أقل تترضاً لحصول أخطاء في المقابلات من المعاينة العشوائية البسيطة.

2- في كثير من الأحيان يكون من الصعب عليك أن تحدد مجتمع الدراسة، وبالتالي يصعب إجراء معاينة عشوائية بسيطة.

هئل (9) R

إذا أردت أخذ عينة حجمها 30 من زبائن محل تجاري كبير فإنه يصعب عليك أن تحدد مجتمع الدراسة، وبالتالي يصعب اجراء معاينة عشوائية بسيطة لكنه من السهل بمكان أن تستعمل المعاينة المنتظمة، وذلك بأن تختار رقما بطريقة عشوائية من 1 إلى 20 مثلا.وليكن العدد الذي اخترته 9 فيعتبر هذا هو العنصر الأول في العينة، وتقابل الشخص التاسع الذي يخرج من المحل التجاري مثلاً، ثم تقابل الشخص 29/69,49 ...، وهكذا إلى أن تحصل على عينة حجمها 30.

من الواضح أن هذه طريقة سهلة حتى على شخص غير مدرب وغالباً ما تحصل على معلومات من المعاينة المنتظمة اكثر مما تحصل عليه في حالة المعاينة العشوائية البسيطة مقابل الكلفة نفسها٠

وغالبأما تكون ا لعينة المختارة بالطريقة المنتظمة موزعة بشكل أكثر تجانساً على جميع أفراد المجتمع، وبالتالي يمكن أن تعطي معلومات عن المجتمع أكثر مما تعطيه عينة مماثلة تم اختيارها بالطريقة العشوائية البسيطة.

هئل (10)

إذا أردت اختيار عينة حجمها n=2OO من مجموعة من بطاقات التسجيل في الجامعة عددها NOOO ، لندرس معاً عدد البطاقات التي فيها أخطاء.

إن طريقة المعاينة المنتظمة تقضي بأن تختار بطاقة. عشوائياً من الخمس بطاقات الأولى (ولتكن البطاقه الرابعة)، ومن ثم تضيف 5 وتسحب البطاقة التاسعة وبعدها البطاقة 14،ثم 19، وهكذا. وآخر بطاقة تسحبها هي رقم 999، كما في الجدول الآتي:

رقم البطاقة	رقم البطاقة المختارة
1 2 3 4 5	4
6 8 9 10	9
* * * *	ز* ٤ ز٠ ذ٠

وهكذا إلى أن تصل إلى آخر خمس بطاقات فتختار البطاقه التي رقمها 999,

لدا

قهوببب (6)

ناقش طرق جمع البيانات الاحصائية في كل مما يأتي:

أ- رغبة الإذاعة في التعرف على رأي المستمعين حول برنامج متين.

ب- رغبة إدارة الجامعة في التعرف على رأي الطلبة حول أسعار وجبة الغداء في مطعم الجامعة.

ج- معرفة الصحف اليومية التي يفضلها طلبة الجامعات الأردنية مع نسبة تفضيل كل صحيفة منها.

قهوجببه (7)

أراد رئيس الدوريات الخارجية معرفة معدل سرعة السيارات الذاهبة من إربد إلى عمان، وللقيام بهذه المهمة قامت دورية بتسجيل سرعة كل سابع سيارة تمر من نقطة معينة:

أ- ما مجتمع الدراسة؟

ب- ما العينة في هذه الدراسة؟

ج-. كيف يمكن للدورية أن تعمل مسحاً شاملا؟

د- ما طريقة المعاينة التي استعملتها الدورية؟

ه- إذا كان عدد السيارات التي قامت الدورية بتسجيل سرعتها مائة سيارة، فكم كان حجم المجتمع؟

قدويبب (8)

هل يستطيع صانعو السيارات استعمال المسح الشامل لتقدير عمر طول الحياة للإطارات المصنوعة في مصنعهم. اشرح بالتفصيل إذا كانت الإجأبه بنعم أو لا.

قهوببي (9)

قامت إحدى المجلات باستطلاع الرأي العام في عمان حول استعمال حزام الأمان في السيارات حيث قابل مندوبها 2000 شخصاً، فوجد أن 42% يحبذون استتمال الحزام و58% لا يحبذون ذلك.

أ- ما المجتمع لهذه الدراسة ؟

ب- ما العينة في هذه الدراسة؟

ج- ما الطريقة التي تقترحهالسحب العينة التي استطلعت المجلة رأيها؟

الم لمة والا مص 'ه Illllics:

تنكر عزيزي الدارم، علد دراستنا للمجتمع والعينة، أن العينة هي جزء من المجتمع الذي نقوم بدراسته وكلما كان اختيارنا للعينة أكثر دقة وتستطيع فيه بيانات التينة أن تمثل المجتمع بشكل أفضل كانت نتائج المقاييس الاحصائية التي تحسب من بيانات العينة أقرب للدقة لنتائج بيانات المجتمع.

ومن الأهمية بمكان التمييز بين المقاييس صائية المحسوبة لكل من المجتمع والعينة، ومن هنا يمكن اعتماد ما سيتم تتاوله في .م٠٠_ .ات التالية بأن :

1.10 الإحصاءة Statistic

إن المقاييس الإحصائية التي تحسب من بيانات العينة يطلق عليها الإحصاءة وعادة ما يرمز لها بالاحرف الانجليزية وعلى سبيل المثال: الوسط الحسابي للعينة: x

الانحراف المعياري للعينة: S*

الارتباطين المتغيرين في كر)(العينة: ٢x,y

2.10 ,^Parameter

إن المقاييس الإحصائية التي تحسب من بيانات المجتمع يطلق عليها المعلمة وعادة ما يرمز لها بالاحرف الاغريقية وعلى سبيل المثال: الوسط الحسابي للمجتمع: ٧إ

الانحراف المعياري للمجتمع: o_x

الارتباطبين المتغيرين في ٢,٧ للمجتمع: Px_ry

إن الرموز السابقة ستتكرر علينا في الوحدات اللاحقة مع تبيان ماهية وكيفية حساباتها. نرجو عدم نفاد الصبر.

0 عصيستي

ا٠؛؛رلا'ز;٦غرحزذ.?٦

لقد تم في هذه الوحدة (الاحصاء والحياة) بحث عدد من المفاهيم ألخصها لك كما يلي:

1- الاحصاء والحياة

لقد اصبح الإحصاء في العصر الحاضر عنصراً هاما في حياة الانسان المتحضر، فهو يحتاجه ويستعمله لاتخاذ القرارات حيال الظواهر العشوائية وفي مواجهة المستقبل المجهول.

2- البحث العلمي والاحصاء

يستحوذ الاحصاء على قدر كبير من اهتمام الباحثين وخاصة في البحوث الاجتماعية والانسانية والتربوية والزراعية وغيرها.وأهم النواحي التي يستتمل فيها الإحصاء في البحث العلمي هي جمع البيانات وتنظيمها وعرضها وتحليلها، وتصميم التجارب، واختبار الفرضيات، واتخاذ القرارات.

3- الإحصاء الوصفي والإحصاء الاستنتاجي

يعرف الإحصاء بأنه العلم الذي يبحث في جمع البيانات وتنظيمها وعرضها وتحليلها واستقراء النتائج منها. إن (الجمع والتنظيم والعرض) هو موضوع الإحصاء الوصفي، وأما التحليل والاستقراء فهما أهم مواضيع الإحصاء الاستنتاجي.

4- تدريج القياس

تسجل البيانات وتقاس المشاهدات بأحد أنواع التدريج التالية:

أ- لالتدريج الإسمي :وتستعمل كمقياس لتحديد هوية الأفراد أو العناصر

ب- التدريج الترتيبي : و هو يسمح بترتيب العناصر حسب سلم معين.

ج- تدريج الفترة :وهو يعطي معنى للفروق بين القياسات.

د- التدريج النسبي : يقع في أعلى مستوى من مستويات التدريج، ويعطي معنى للصفر المطلق.

5- المتغيرات والثوابت

المتغير هو الصفة أو السمة التي تأخذ قيماً متعددة لأفراد عدد من المشاهدات مثل الطول لطلبة أحد الصفوف؛ أما الثابت فهو القيمة الثابتة التي تكون من صفات المادة أو العنصر مثل الكثافة النوعية لمادة معينة، أو معامل التمدد للحديد النقي، وما شابه ذلك .وتقسم المتغيرات إلى منفصلة، وهي التي تأخذ عددا معدوداً (محدوداً أو غير محدود) من القيم، مثل عدد أفراد الأسرة؛ والمتغيرات المتصلة، وهي

التي تأخذ جميع القيم على فترة أو فترات مثل درجة الحرارة والحجم وغير ذلك.

6- المتغيرات الوصفية (النوعية) والكمية

المتغيرات الوصفية هي المتغيرات التي لا يمكن قياسها بالأرقام العددية بينما المتغيرات الكمية يمكن قياسها بالأرقام العلدية، وتنقسم الى قسمين منفصلة والتي تأخذ قيماً عددية صحيحة ومتصلة تقع ضمن مدى معين ٠

7" المتغيرات المستقلة والتابعة

مصطلح متغير مستقل، ومتغير تابع يعني أنه بمعرفة قيمة المتغير المستقل يمكننا معرفة قيمة المتغير التابع بواسطة إقتران رياضي بين لاو ٧، ولهذاتكون قيم ٧ تابعة ومعتمدة على قيم لا.

8- المجتمع والعينة

المجتمع :مجموعة من جميع الأفراد أو العناصر التي تكون ضمن اهتمام الباحث في دراسته، مثل جميع طلبة الصف الثالث في مدارس البنين في المملكة الأردنية الهاشمية.

والعينة: مجموعة جزئية من المجتمع يقوم الباحث بدراستها لعدم قدرته دراسة كافة أفراد المجتمع.

يتم جمع البيانات الإحصائية باحدى الطرق التالية:

1- طريقة المسح الشامل.

2- طريقة العينة.

9٠ المعلمة والاحصاءة

المعلمة تطلق على كلب مقياس إحصائي محسوب من جميع بيانات المجتمع ويرمز له بحرف إغريقي، أما الإحصاءة فتطلق على كل مقياس إحصائي محسوب من بيانات العينة.

S &

تهويبي (1)
أ- ثابت.

لبا منير
ج متحير. د-
ثابت.

ه- ثابت

و- ملتير

تجدديئس (2)

أ- المكالمات الهاتفية : كمي - منفصل

ب- المعدلات التراكمية: كمي -متصل

ج- مسقط رأس الفرد :وصفي.

د- عمر الفرد :كمي - متصل.

ه- التحصيل العلمي:وصفي.

و- الدخل الشهري :كمي - متصل

ز- حرارة المريض : كمي - متصل.

ح- حوادث السير : كمي - منفصل.

ط- أنواع اللحوم :وصفي.

قشوييب (3)

المتغير المستقل في هذا التدريب هو عمر الرجل والمتغير التابع هو ضغط الدم.

وذلك لأن الطبيب يسجل عمر الرجل ويسجل ضغط الدم لديه ويحاول أن يجد علاقة رياضية تربط ضغط الدم بدلالة العمر لكي يتمكن من تنبؤ ضغط الدم لدى رجل عرف عمره.

قكويي (4)

من الواضح أن كمية الأسمدة هي المتغير المستقل حيث أن تغيرها كان تحت تحكم الباحث الزراعي حيث قرر أن يستعمل كمية معينة من الأسمدة لكل قطعة من الأرض تحت التجربة.

والمتغير التابع هو كمية المحصول حيث يحاول الباحث إيجاد علاقة تربط كمية المحصول بكمية السماد المستعمل وبالتالي يريد أن يستنتج أنه إذا استعمل كمية معينة من السماد فهل يستطيع أن يتنبأ بكمية المحصول التي سيحصل عليها.

قفوتيب (5)

حجم المجتمع:

N = 500 ب 400 ب 280 ب 200 ب 220 = 1600

حجم العينة من الطبقة الأولى:

n٦=5OOx—50 = —
1 Ifion

حجم العينة من الطبقة الثانية:

١٦2=400يح=40
2 1600

حجم العينة من الطبقة الثالثة:

3=280 لا=28
3 1600

حجم العينة من الطبقة الرابعة:

0ع=200بر=20
4 1600

حجم العينة من الطبقة الخامسة:

١٦5=22Ox لا=22
5 1600

تهويبي (6)

أ- للوقوف على رأي المستمعين حول برنامج معين، يمكن للأذاعه أن تطرح عدداً من الأسئلة في الصحف المحلية وتطلب أن يجيب عنها الراغبون.وبالطبع هذه الطريقه كأنها محاولة إرسال استبانة لجميع المستمعين وا لإستجابة تكون ممن يرغب في إعطاء رأيه، وهناك طرق أخرى لجمع البيانات في هذه الحالة وهي على سبيل المثال : إرسال استبانة لعينة تختارها الإذاعة من دليل الهاتف.

ب- رغبة إدارة الجامعة في التعرف على رأي الطلبة حول أسعار وجبة الغداء في مطعم الجامعة، من الممكن الحصول على البيانات في هذه الحالة بتصميم استبانة توزع على الطلبة الذين يحضرون إلى المطعم على مدى عدة أيام، ويطلب من الطالب تعبئة الاستبانة مرة واحدة.

ويمكن الإكتفاء بتوزيع الاستبانة على عينة منتظمة من الطلبة الذين يردون إلى المطعم فتعطى الاستبانة لكل خامس طالب مثلا ويطلب منه تعبئة الاستبانة.

ج- معرفة الصحف المفضلة ونسبة التفضيل يمكن الوقوف على ذلك بتوزيع استبانة وجمعها من الطلبة أو من عينة منهم. ومن الممكن أن يكون ذلك بأخذ رأي الطلبة مباشرة وذلك عند التسجيل مثلا، فيسأل الطالب عن الصحف المفضلة لديه ونسبة التفضيل لديه ويسجل ذلك مباشرة.

قدويب (7)

أ- مجتمع الدراسة: جميع السيارات التي استعملت الطريق من إربد إلى عمان في ذلك

اليوم.

ب- العينة منتظمة وتكونت من السيارات التي كان رقمها المتسلسل عند مرورها من النقطة المعينه:7/14/21 ٠٠٠٠٠,٠٠٠٠٠

ج- يمكن للدورية أن تتمل مسحاً شاملاً وذلك بتسجيل سرعة كل سيارة لدى مرورها عن النقطة المعينة.

د- الطريقة المنتظمة.

ه- حجم العينه 100 وهذيساوي حجم المجتمع. إذاًحجم المجتمع يساوي 700 سيارة.

تشوببتب (8)

لايستطيع صانعو اطارات السيارات. اجراء المسح الشامل لتقدير عمر الإطارات وذلك لأن تسجيل عمر الإطار يعني استهلاك ذلك الإطار، أي استعماله حتى يتلف وبالتالي قياس عمره، وهذا غير ممكن حيث أنه لا معنى لأن يقوم المصنع بإتلاف جميع الإطارات التي يصنعها لكي يعرف طول الحياة أو عمر الإطار.

قشوببب (9)

أ- مجتمع الدراسة: جميع سائقي السيارات في عمان.

ب- العينة: 2000 سائق قام مندوب المجلة بمقابلتهم.

ج- الطريقة المقترحة لسحب العينة تقسيم المنطقة إلى عدد من المحطات، يقف مندوبو المجلة عند هذه المحطات ويقومون بمقابلة عدد محين من السائقين بطريق المقابلة والعينة تكون منتظمة، فعلى سبيل المثال : المحطة التي يمر منها سيارات كثيرة يمكن للمندوب أن يقابل كل خامس سائق أما المحطات غير المزدحمة فيقابل كل ثالث سائق وهكذا.

14 .مسرد المصطلحات ;٠ - 'ر ;٦■

- الإحصاء الاستنتاجي Inferential Statistics : هو فرع الاحصاء الذي يعنى بتحليل البيانات و استقراء النتائج منها.
- الإحصاء الوصفي Descriptive Statistics : هو فرع الإحصاء الذي يعني بجمع البيانات وتبويبها وعرضها.
- الإحصاء والحياة Statistics and Life : هو الموضوع الذي يبحث علاقة الإحصاء في الحياة ومدى الحاجة لاستعمال الإحصاء في حياتنا الحاضرة.
- التحكم بالنوعية ٣٥١*م0ع٤٧اا2لما0: إحدى التطبيقات الإحصائية وتستعمل في الصناعة لغرض الحكم على استمرار خط الانتاج أو إيقافه مثلاً.
- تحليل الانحدار وبالإرتباط Regression Analysis and Correlation : الارتباط هو قياس قوة العلاقة الخطية بين متغيرين، أما الانحدار فهو تحديد العلاقة الخطية بين المتغيرين بمعادلة رياضية.

-تحليل المحتوى 5أ75اا3٤٨٣٥٤٦٥٢: إحدى الطرق في مناهج لبحث لعلمي.

- التدريج الاسمي Nominal Scale : أقل التدريجات مستوى ويعني بتسمية العناصر المقاسة.
- التدريج الترتيبي Scale اة0ال٣0:هو التدريجذو المستوى الثاني ويعنى بترتيب الحناصر المقاسة مثل الرتب العسكرية.
- تدريج القياس Measurement Scale : هي التدريجات التي تستعمل في القياس وتصنف حسب مستواها في المقارنة أو بوجود معنى للصفر المطلق أو عدمه.
- تصميم لبحث Design ٣٤١•ع5ع3: إحدخطوات البحث لعلمي ن تصممطريقة البحث.
- تصميم التجاربب Design of Experiments : إحدى الطرق الإحصائية المستعملة في الحياة العملية وخاصة في الزراعة والدراسات الاجتماعية.
- تطبيقات الإحصاء Applications of Statistics : هي تلك التطبيقات التي تستعمل فيها الطرق الاحصائية وتشمل فروعاً هامة في العلوم والزراعة والتربية والصناعة والطب.
- الثوابت Constants : القيم الثابتة في الطبيعة مثل ثابت الجذب العام وسرعة الضوء وغيرها.
- الدراسة الميدانية Field Study : هي الدراسة العملية التي تستعمل فيها طرق المقابلة

أو الاجابة عن الاستبيانات،

- طرق المعاينة Sampling Methods : هي الطرق التي تستعمل لاختيار العينة.
- العينات الاحتمالية Probability Sampling : هي العينات التي نختارها باستعمال النظرية الاحتمالية ومن هذه العينات العينة العشوائية البسيطة.
- العينات غير الإحتمالية Non-Probability Sampling : هي العينات التي لا تختار بناء على نظرية الاحتمال.
- العينة العشوائية البسيطة Simple Random Sampling: إحدى طرق اختيار العينة والتي تضمن أن لكل عنصر في المجتمع نفس فرصة الاختيار مثل أي عنصر آخر.
- المتغيرات Variables :كل ظاهرة تأخذ قيماً متعددة تسمى متغيراً.
- المتغيرات المتصلة Continuous Variables : هي المتغيرات التي تأخذ جميع القيم على فترة معينة أو عدد من الفترات.
- المتغيرات المستقلة Independent Variables : عند دراسة متغيرين لعناصر معينة بغرض التنبؤ بأحدهما بناء على الاخر ويسمى الأخير المتغير المستقل.
- المتغيرات المنفصلة Discrete Variables : هي المتغيرات المحدودة أو التي يمكن وضع قيمها في مقابلة واحد إلى واحد مع مجموعة الأعداد الصحيحة الموجبة.
- المجتمع والعينة Population and Sample : المجتمع هو مجموعة جميع العناصر قيد الدراسة والعينة هي جزء (مجموعة جزئية) من المجتمع.
- المعاينة الهادفة أو الحكمية Purposive or Judgmental Sampling : هي طريقة أخذ عينة لغرض محدد أو بحكم محدد يقرره الباحث.
- المنهج التاريخي Historical Method : أحد مناهج البحث العلمي ويعتمد على دراسة الوثائق وا لمقارنة.
- المنهج التجريبي Experimental Method : هو المنهج الذي يحتاج إلى اجراء تجارب لاستعمالها في البحوث.
- وحدة المعاينة أو المشاهدة Unit of Observation : هي الوحدة في المجتمع التي نقوم بمشاهدة إحدى خاصياتها.

15, المراجع -

أ- المراجع العربية:

1- أبو زينه، فريد؛ لطفية، لطفي؛ الخليلي، خليل. الطرق الاحصائية في التربية والعلوم الانسانية: الجزء الأول. عمان :دار الفرقان للنشر والتوزيع، 984 ٦ .

2- أبو صالح، محمد، الطرق الإحصائية : لإصدار الثاني.عمان :دار اليازوري العلمية للنشر والتوزيع، 2001.

3- أبو صالح، محمد، الموجز في الطرق الإحصائية : الطبعة العربية الأولى. عمان : دار اليازوري العلمية للنشر والتوزيع، 2002.

4- غرايبة، فوزي وآخرون، أساليب البحث العلمي في العلوم الاجتماعية والنفسية: الطبعة الثالثة، عمان: دار وائل، 2002.

ثآ١إ

الوحدة الثانية

الإحصاء الوصفيلمتغيرواحد

Univariate Descriptive Statistics

محتويات الوحدة

m.-.

ء' ■ض وع

اه 1 تمهيد

عزيزي الدارس، مرحبا بك في هذه الوحدة، تتألف هذه الوحدة (الإحصاء الوصفي لمتغير واحد) من سبعة أقسام رئيسية، نشرح في القسم الرئيسي الأول منها عرض البيانات الأولية بواسطة الجداول أو الطرق البيانية ومنها طريقة المستطيلات والخطوط المنكسرة والمنحنى والدائرة إضافة إلى الطرق التصويرية. وبعد انتهاءك من طرق عرض البيانات الأولية ننقلك إلى أسلوب آخر لعرض البيانات بتلخيصها ووضعها على شكل جدول يتضمن عمودين وعدد من الصفوف يطلق على العمود الأول الفئات والعمود الثاني التكرارات وهي عدد الحالات التي تقابل هذه الفئات ويسمى هذا جدول التوزيع التكراري البسيط، وسنشرح لك أيضاً كيفية بناء هذا الجدول يدوياً.

أما القسم الثاني (تمثيل التوزيعات التكرارية بيانياً) فستتعلم فيه طريقة رسم المدرج التكراري والمضلع التكراري والمنحنى التكراري، وهذه الطرق البيانية ستعرفك وتدربك على تمثيل الجداول للبيانات التجميعية في الجدول التكراري البسيط بمنحنيات تسهم في تعزيز فهمك لوصف البيانات من خلال التعرف على صفاتها وخواصها.

وتوخيا للتسلسل المنطقي في استكمال وصف البيانات، نستعرض في القسم الثالث (التماثل والالتواء والتفرطح في التوزيعات التكرارية) بغية تعرفك على وصف التوزيعات التكرارية من خلال الرسوم البيانية التي تمثلها، بحيث تستطيع التمييز بين التوزيع التكراري المتماثل والتوزيع التكراري غير المتماثل، أو الملتوي نحو اليمين أو الملتوي نحو اليسار. مضاف إلى كل ذلك معرفة طبيعة قمة منحنى التوزيع التكراري فيما إذا كان مدبباً أو غير مدبب، بما في ذلك متوسط التفرطح أو كبير التفرطح.

أما في القسم الرابع (مقاييس النزعة المركزية) فسيتبين لك أنواع هذه المقاييس وهي الوسط الحسابي والوسيط والمنوال، وبذلك تنتقل إلى وصف البيانات الإحصائية بدقة أكثر، ويعود السبب لكون هذه المقاييس تعطي قيماً عددية لمقاييس الموقع في التوزيع.

رغم أهمية مقاييس النزعة المركزية، إلا أنها غير كافية لوصف البيانات الإحصائية، إذ أنها لا تعطي فكرة عن تغيرالبيانات والتباعد فيما بينها، وهذا دفعنا إلى شرح (مقاييس التشتت)كالمدى والمدى الربيعي والتباين والانحراف المعياري في القسم الخامس بحيث تستطيع اكمال وصف البيانات.

أما في القسم السادس (أثر التحويلات الخطية على مقاييس النزعة المركزية ومقاييس التشتت) فتضمن كيفية حساب قيم هذه المقاييس لو طلب تحويل البيانات من

قيمها الأصلية إلى قيم جديدة دون تحويل كل مفردة من البيانات.

وفي القسم السابع (مقاييس التشتت النسبية) بينا أن مقاييس التشتت عند استخدامها لملمقارنة بين مجموعتين أو أكثر تختلف في وحدات قياسها أو تتباعد مقاييس النزعة المركزية من النوع الواحد بينها حتى وأن قيست بنفس وحدات القياس، لذا لا بد من مقاييس خاصة تحسب التشتت في هذه الحالة تسمى مقاييس التشتت النسبية ومنها معامل الاختلاف بنوعية المعياري والربيعي وكذلك الدرجة المعيارية. وشرح مفاهيم العشيريات والمئينات وطرق حساب المئينات.

2.1 أهداف الوحدة

بعد فراغك من دارسة هذه الوحدة يجب أن تكون قادراً على أن:

1- تعرض البيانات النوعية في جداول مناسبة وحسب التصنيفات الملائمة أو المطلوبة وتمثلها بيانيا.

2- تلخص البيانات الإحصائية وتعرضها على شكل جدول توزيع تكراري بسيط يتضمن عمودين.

3- تمثل التوزيعات التكرارية على شكل مدرج تكراري أو مضلع أو منحنى.

4- تصف التماثل أو الالتواء أو التفرطحللتوزيعات التكرارية بعد تمثيلها بيانيا.

5- تحسب مقاييس النزعة المركزية (الومط الحسابي، الوسيط والمنوال).

6- تحسب مقاييس التشتت بنوعيها المطلق والنسبي.

7" تفسر دلالات كل من مقاييس النزعة المركزية بصفتها مقاييسموقع، ومقاييس التشتت بصفتها مقاييس تباعد البيانات وتغيرها.

8- تناقش اثر التحويلات الخطية للبيانات على مقاييس النزعة المركزية والتشتت.

9- تحسب المئينات.

3.1 أقسام الوحدة

القسم الرئيسي الأول من الوحدة يرتبط ارتباطاً وثيقاً بالهدفين الأول والثاني حيث نعرض لك البيانات الأولية بالطرق المناسبة ثم نلخصها بوضعها في جدول توزيع تكراري بسيط.

وفي القسم الثاني تمثل التوزيع التكراري بيانياً من خلال المدرج التكراري

وا لمضلع التكراري، وهذا ما يحقق لك الهدف الثالث لهذه الوحدة.

أما الهدف الرابع فيتحقق بدراستك القسم الثالث، فيما المادة التي يتطلبها الهدف الخامس فقد تم شرحها في القسم الرابع، وتناولنا في القسم الخامس كل ما تحتاجه لتحقيق الهدف السادس.

والقسمان الرابع والخامس يحققان الهدف السابع، حيث اعطينا فيهما نبذة عن دلالات مقاييس النزعة المركزية، وشرحنا المقارنة بين صفاتها، وتناولنا متنى التشتت والحاجة إلى دراسته.

و تم في القسم السادس دراسة أثر التحويلات الخطية على كل من مقاييس النزعة المركزية والتشتت وبهذا تحقق لك الهدف الثامن.

أما في القسم السابع فقد تناولنا فيه كيفية حساب العشيريات والمئينات كمقاييس موضعية للبيانات وهذا مايحقق الهدف التاسع.

4,1 القراءات المساعدة

إن مادة الوحدة التي بين يديك كافية لاستيعاب ما تتضمنه، ولاثراء معلوماتك حول موضوع الوحدة ننصح بالرجوع إلى القراءات التالية:

1- أبو صالح، محمد صبحي؛ عوض، عدنان محمد، مقدمة في الإحصاء :نيويورك : جون وايلي، 1982.

2- أبو صالح، محمد صبحي؛ عوض، عدنان محمد، مبادىء الإحصاء . الجزء الاول، ص 70 1-51. عمان : دار الفرقان للنشر والتوزيع، 986 1 .

5ه1 ما تحتاج إليه لدراسة الوحدة

وفر الجو الهاديء والمكان المناسب للدراسة وابدأ بقراءة الوحدة بعد أن تكون قد حضرت الأشياء التالية:

ورقاً للكتابة، وقلم رصاص، وآلة حاسبة، ومسطرة، وفرجاراً لرسم الدوائر.

أثناءدراستك لمادة الوحدة ستجد في ثناياها أسئلة التقويم الذاتي التي تساعد ك في مراجعة وتلخيص الوحدة كما ستجد تدريبات عليك القيام بها، فهي سترسخ الأفكار المطروحة في الوحدة وتمنحك فرصة لاختبار تعلمك ومدى استيعابك وفهمك لهذه الوحدة، وإذا واجهتك أي صعوبات أو كان لديك أي ملاحظات، اتصل بمرشدك الذي سيرحب بك، ويساعدك على ايجاد الحلول المناسبة لأسئلتك.

2.عرضالبيانات الإحصائية

1. 2 عرض لبيانات الأولية

تواجهنا في الحياة العملية كميات كبيرة من البيانات. منها ما هو خاص بالوزارات والمؤسسات ومنها ما يتعلق بنتائج التجارب في العلوم السلوكية والعلوم الطبيتية والزراعية وغيرها.فإذا ما عرضنا هذه البيانات بطريقة المقال ضمن التقارير أو الصحف اليومية فإنها بلا شك ستكون مملة ويصعب استيعابها والمقارنة بين مفرداتها، ولذا كان من الضروري عرض هذه البيانات بطرق شيقة سهلة، ومن هذه الطرق:

1.1.2 طريقة الجداول

هي عبارة عن وضع البيانات في جدول، وكثيراً ما تستعمل هذه الطريقة في عرض تغير ظاهرة ما مع الزمن، أو مع مسميات كالبلدان والمدارس وغيرها، أو مع الزمن والمسميات معاً . وعند استعمال هذه الطريقة يجب مراعاة نكر ما يأتي: أ- عنوان الجدول.

ب- الوحدات المستعملة.

ج- مذكرات المصادر التي أخذت منها البيانات.

د- مذكرات تفسيرية تفسر سبب شذوذ بعض البيانات، إن وجدت.

?

كان عدد رواد مكتبة الجامعة في أحد الأسابيع كما في الجدول رقم (1) الآتي: جدول رقم (1)

اليوم	عدد الرواد
السبت	**760**
الأحد	618
الاثنين	81.5
الثلاثاء	780
الأربعاء	450
الخميس	327

2.1.2 طريقة المستطيلات أو الأعمدة

قتلخص هذه الطريقة في وضع المسميات على محور أفقي او عمودي ورسم مستطيل على كل مسمى يكون ارتفاعه ممثلا للقيمة المقابلة لذلك المسمى وذلك باستعمال مقياس رسم مناسب.

ونستعمل هذه الطريقة لعرض تغير ظاهرة مع الزمن، أو مع مسميات، أو كليهما معا، حيث يمكن استتمالها للمقارنة بين قيم الظواهر حسب المسميات على مدى عدة سنوات، كأن نقارن بين اعداد الطلبة حسب تخصصاتهم في الجامعة على مدى ست سنوات.

(2)٠ ج

لائل الجذول رقم

(2) الآتي : أعداد خريجي إحدى الجامعات. جدول رقم (2)

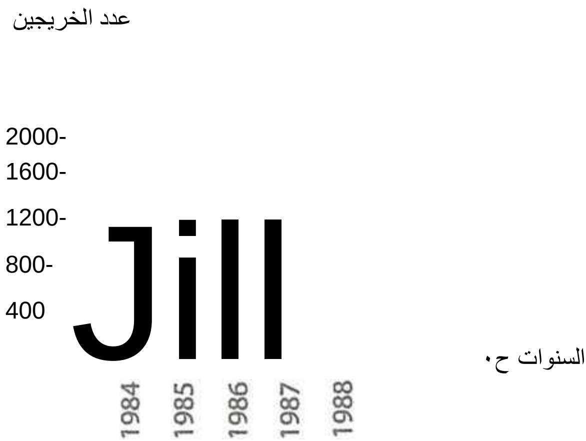
السنة	عدد الخريجين
1984	800
1985	100
1986	1400
1987	1200
1988	1600

اعرض هذه البيانات بطريقة المستطيلات .

الحل:

لتعرض هذه البيانات بطريقة المستطيلات كما يظهر في الشكل رقم (1)، عليك أن تمثل السنوات على الخط الأفقي ثم ترسب مقابل كل سنة مستطيلاً يمثل عدد الخريجين في تلك السنة، على أن يكون ارتفاع المستطيل متناسباً مع عدد الخريجين الذي يمثله حسبمقياس رسم مناسب.

لاحظ أن السنة على الخط الأفقي تمثل بقطعة مستقيمة وتكتب السنة في مركز تلك القطعة، وتكون السنوات على الخط الأفقي غير متلاصقة لأن المقصود بالسنوات في هذه الحال مثلاً هو التتبير عن مسميات، ولو وضعت أسماء بلدان مثلاً على الخط الأفقي فإنك تمثل كل بلد بقطعة مستقيمة منفصلة عن القطعة الأخرى وتضع اسم البلد في مركز تلك القطعة.


عدد الخريجين
2000-
1600-
1200-
800-
400
1984
1985
1986
1987
1988
السنوات ح٠

شكل رقم (1)

3.1.2 طريقة الخط المنكسر

تستعمل هذه الطريقة لعرض البيانات الناتجة من تغير ظاهرة أو عدة ظواهر مع مسميات، أو مع الزمن، أو مع كليهما، مثل تغير درجة حرارة مريض مع الساعات، أو تغير أعداد الطلاب مع السنوات، أو تغير أعداد الطلاب مع الكليات.

مثال (3)

اعرض البيانات في المثال (2) بطريقة الخط المنكسر.

الحل:

لعرض هذه البيانات بطريقة الخط المنكسر . ارسم محورين متعامدين يمثل المحور الأفقي منهما السنوات والمحور العمودي يمثل أعداد الخريجين بمقياس رسم مناسب ٠

ارصد النقاط التي احداثيها السيني هو السنة واحداثيها الصادي هو عدد الطلبة، ثم صل بين كل نقطتين متتاليتين بقطعة مستقيمة ليظهر لك عرض البيانات على طريقة الخط المنكسر.

لاحظ في مثال (3) أن احداثي النقطة الاولى هي (800، 1984) حيث تعين 1984 على الخط الأفقي الذي يمثل السنوات وتتين 800 على الخط العمودي الذي يمثل عدد الخريجين.

عدد الخريجين
٨
2000
1600.
1200-
800 -
400
اكنوات < ١١'١١١
1984 1985 1986 1987 1988

الشكلرقم (2)

2. 4.1 طريقة الخط المنحني

وهذه الطريقة تماثتل طريقة الخط المنكسر وتحصل عليها بتمهيد الخط المنكسر ليصبح منحنى، وتستعمل هذه الطريقة عندما تتغير الظاهرة على فترات زمنية قصيرة وكثيرة.

2. 5,1طريقة لدائرة

وأهم استعمالات هذه الطريقة استعمالها لتقسيم الكل إلى أجزائه، فيمتل المجموع الكلي بدائرة كاملة، ويمثل كل جزء بقطاع دائرة يكون قياس زاويته مساوياً 360مضروباً في نسبة الجزء إلى المجموع الكلي، وذلك لأن مجموع قياس الزوايا حول نقطة هو360.

نقال (4)

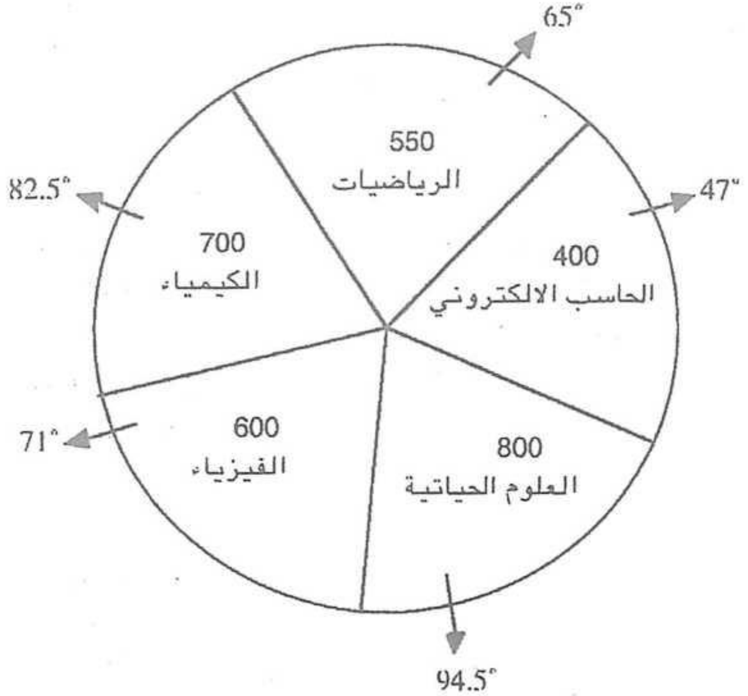
إذا كان عدد الطلبة في إحدى كليات العلوم هو3050 طالباً وطالبة موزعين على الأقسام الأكاديمية كما في الجدول الآتي:

الحاسب الالكتروني الرياضيات الكيمياء الفيزياء العلوم الحياتية

400	550	700	600	800

فاعرض هذه البيانات بطريقة الدائرة:

قياس زاوية قطاع العلوم الحياتية = $\frac{360^\circ \times 800}{3050} = 94.43^\circ$



6.1.2 الطريقة التصويرية

وتفتعمل هذه الطريقة لعرض البيانات بصورة مبسطة مشوقة كما هو الحال في التقارير الحكومية، وكتب علم النفس، وللدعاية، وفي كتب الأطفال.

فإذا عرضت البيانات المتعلقة بقيمة الودائع السنوية في عدد من المصارف (البنوك) فإنك ترسم صورة كيس نقود واحد ليمثل كل 20000 دينار أردني على سبيل المثال، فإذا بلغت الودائع في البنك : (أ) قيمة 100000 دينار فإنك ترسم خمسة اكياس لتمثل هذا المبلغ، وإذا كانت الودائع في البنك (ب) ما قيمته 90000 دينار ترسم صورة أربعة أكيلس ونصف الكيس مقابل هذا المبلغ، وكماتلاحظ فلن هذه الطريقة ليست دقيقة جدا.

قئويبيه (1)

بلغت أعداد طلبات العمل لدى 5 مكاتب عمل كما في الجدول أدناه:

المكتب	عدد الطلبات
1	1700
2	1200
3	2100
4	1090
5	300

اعرض هذا الجدول بطريقة المستطيلات وطريقة الخط المنكسر.

2.2 التوزيعات التكرارية

إذا كان لديك عدد كبير من البيانات مثل علامات شهادة الدراسة الثانوية في الأردن، فكيف تستطيع عرض هذه البيانات بطريقة مفيدة تمكنك من فهمها وتكوين فكرة عامة عنها؟

إذا فكرت في عرض هذه البيانات بإحدى طرق عرض البيانات الأولية مثل طريقة الجداول - فلا بد أن تلاحظ أنك تحتاج إلى كتابة جداول طويلة لتتمكن من عرض البيانات المطلوبة، كما ستلاحظ أيضاً صعوبة تكوين فكرة عامة عن البيانات التي لديك.

إذن، ماذا تقترح؟ إن إحدى الطرق المفيدة في مثل هذا الحال، حال وجود بيانات كثيرة، هي عرض البيانات بواسطة التوزيع التكراري؛ التي هي إحدى الطرق التي تتمكن بواسطتها من تنظيم البيانات الكثيرة بحيث لا تفقد هذه البيانات من أهميتها إلا الشيء اليسير أو ربما لا تفقد شيئاً.

أما الطريقة الأساسية لبناء التوزيع التكراري فهي اجراء تقسيم مدى قيم البيانات إلى فئات، وحصر عدد البيانات الواقعة ضمن كل فئة . والمدى كما تعلم هو الفرق بين أكبر قيمة في البيانات و أقل قيمة فيها.

هفأل (5)

ا] ٠ ئخا ك٦اخت ألبفاك_ ات 1 ر~ككا خب عدد Uكا'ت ١صخحة ١ش بك ب عدد ح 1 اسةفيذصل تخرجهم

ح ئحة:

١; 12 15 10 10 14

9 13 12

فإن الجدول رقم (3) يمثل التوزيع التكراري لعدد الساعات التي سجل فيها الهلابة.

تلاحظ في بناء هذا التوزيع أننا بدأنا من أدنى قيمة وهي 9 ورتبنا القيم ترتيباً تصاعدياً إلى أن وصلنا أعلى قيمة 16 كما يظهر في العمود الأول.

جدول رقم (3)

عد الساعات	التكرار
9	3
10	2
11	0
12	4
13	2
14	1
15	2
16	1

أما عناصر العمود الثاني فتمثل عدد المرات التي تكررت فيها كل قيمة، فالقيمة التي لم تظهر في البيانات يكون تكرارها صفرا.

من هذا المثال، يتبين لديك أن التوزيع التكراري هو عبارة عن جدول يتألف من:

1- فئات قيم المشاهدات أو القياسات.

2- التكرارات المقابلة لكل فئة أو قياس.

1.2.2 بناء التوزيع التكراري

من السهل أن ترى، عزيزي الدارس، أنه إذا كان مدى البيانات صغيراً أمكنك بناء التوزيع التكراري مباشرة كما في المثال (5). أما إذا كان المدى كبيراً أو كان عدد البيانات كذلك. فإنه يجدر بك في هذه الحالة أن تقسم قيم البيانات إلى فئات يتراوح عددها ما بين 5 و15 فئة، حسب عدد البيانات وبعد هذا التقسيم عليك افراغ البيانات على الفئات وجمع التكرارات المقابلة لكل فئة، عندها تحصل على توزيع تكراري.

مال (6)

الجدول رقم (4) يمثل التوزيع التكراري للرواتب الشهرية بالدينار الأردني لسبعين موظفا في إحدى الوزارات.

جدول رقم (4)

حدود الفئة	مركز الفئة	التكرار
99-90	94.5	15
109100	104.5	08
19-110	114.5	19
129-120	124.5	10
139-130	134.5	18

من هذا المثال تلاحظ أن التوزيع التكراري جدول يحتوي على فئات غير متداخلة يقابل كلا منها عدد العناصر الموجودة فيها ونعبر عن الفئة. إما ب (حدود الفئة) أو (بمركزها) . أما مجموع التناصر في كل فئة فتتبر عنه ب (التكرار).

وأنت تلاحظ ومن هذا الجدول أن الحد الأدنى للفئة الاولى هو 90 والحد الأعلى هو99، أمامركزتلك الفئة الأولى فهويساوي أي 94.5، واماتكرارالفئة الأولى فإنه يساوي 5 1 .

ولشرح الخطوات المتبعة في بناء التوزيع التكراري دعنا ندرس المثال الآتي:

r

هال (7)

تمثل البيانات التالية المصروف الشهري (لأقرب دينار أردني) لستين طالباً.

41	33	25	35	45
48	46	36	39	22
48	32	27	48	44
46	33	29	42	24
49	23	23	46	25
36	41	48	37	35
41	43	47	39	36
43	48	33	41	43
38	46	36	26	48
32	34	22	28	47
24	24	39	33	38
45	46	44	23	43

والمطلوب وضع البيانات أعلاه في جدول توزيع تكراري.

إذا أردت أن تعرض هذه البيانات في توزيع تكراري ذي فئات فاتبع الخطوات الآتية:

٦, خذ فئات متساوية وافرض أن عددها 7.

(لاحظ أنك أنت الذي قررت عدد الفئات ليكون 7، وبامكانك أخذ أي عدد آخر مثل 5 أو 8 مثلا).

2- حدد المدى للبيانات، وهو عبارة عن أكبر قيمة ناقصاً أصغر قيمة في تلك البيانات. وفي مثالنا هذا يكون المدى يساوي 27

27 = 22 - 49، لأن أكبر قيمة هي 49 وأصغر قيمة فيها هي 22 .

3- أجد طول الفئة C وذلك بقسمة المدى على عدد الفئات، ثم قرب الجواب دائماً إلى أعلى بحيث لا يحتوي طول الفئة على خانات عشرية أكثر من تلك المستعملة في البيانات. فإذا كانت البيانات أعداداً صحيحة فإن طول الفئة يجب أن يكون عدداً صحيحاً. أما إذا كانت البيانات تحتوي على خانة عشرية واحدة فإن طول الفئة يجب أن يكون عدداً صحيحاً أو عدداً يحتوي على خانة عشرية واحدة.

وفي مثلنا يكبن —=3.9 . ويما لن البيانات معطاة لاقرب عد محيح (لا يوجد كسور عشرية) فأنت تقرب العدد 3.9 إلى أعلى فيصبح طول الفئة c=4.

4- عين الحد الأدنى لأول فئة، ويجب أن يكون هذا الحد مساوياً أو أصغر من أقل قيمة في البيانات.ويجب أن تكون هذه القيمة من حيث الدقة كالبيانات الأصلية.فإذا كانت البيانات الأصلية أعداداً صحيحة كان الحد الأدنى عدداً صحيحاً. وإذا كانت تلك البيانات تحتوي على خانة عشرية واحدة فإن الحد الأدنى يمكن أن يحتوي على خانة عشرية واحدة. لذلك يمكنك أخذ الحد الأدنى للفئة الأولى 22.

5- بعد تعيينك الحدالأدنى للفئة الأولى عين الحد الأدنى الفعلي لتلك الفئة، وهو عبارة عن الحد الأدنى ناقصاً 0.5 (لأن البيانات أعداد صحيحة، وقد طرحنا 0.5 لأن البيانات أعداد صحيحة، فمعنى ذلك أن درجة الدقة 1 لذلك نطرح نصف وحدة دقة أي 0.5). أما اذا كانت البيانات مقربة لأقرب رقم عشري واحد فمعنى ذلك أن درجة الدقة 0.1 وبالتالي نصف وحدة الدقة في هذه الحالة يساوي 0.05 ، إذن فالحد الأدنى الفعليللفئة الأولى هو:

22-0.5=21.5

لأن البيانات الأصلية كانت اعداداً صحيحة.

6- عين الحد الأعلى الفعلي للفئة الأولى، وذلك بإضافة طول الفئة إلى الحد الأدنى الفعلي لتلك الفئة.

فيكون الحد الأعلى الفعليللفئة لأولى هو: 25.5 =4+21.5

عين الحد الأعلى للفئة الأولى وهو يساوي الحد الأعلى الفعلي لتلك الفئة ناقصاً 0.5 ، لأن البيانات أعداداً صحيحة، ولذا فالحد الأعلى للفئة الإولى ـــ

25.5 - 0.5- 25.

بهذا تكونقد حصلت على حدود الفئة الأولى وهي 22 - 25 وكذلك على الحدود الفعلية للفئة الأولى وهي 21.5 - 25.5.

7- عين الحدود الدنيا والعليا لجميع الفئات الباقية، وذلك بإضافة طول الفئة لكل حد، ثم عين الحدود الدنيا الفعلية والعليا الفعلية وذلك باضافة طول الفئة لكل حد فعلي.

أنظر العمودين (1)، (2) في جدول رقم (5) حيث تلاحظ، على سبيل المثال، أنك تحصل على الحد الأدنى للفئة الثانية باضافة 4 إلى 22 فتحصل على 22)4=26 وتحصل على الحد الأعلى للفئة الثانية بقولك.

25(4=29

أما الحد الأدنى الفعلي للفئة الثانية فهو

21.5+4-25.5

وأما الحد الأعلى الفعلي للفئة الثاتية فهو

25.5+ 4= 29.5

وهكذا تبنى جميع الفئات وتعين حدودها الدنيا والعليا، وحدودها الدنيا الفعلية والعليا الفعلية.

8- عين مراكز الفئات (ونعبر عنها بالرمز |X ليدل على مركز الفئة أ، أي أن ,X هو مركز الفئة الأولى، فيما X2 هو مركز الفئة الثانية، وهكذا).

ويتم تعيين مركز أي فئة بأن نقسم مجموع حديها الأدنى والأعلى على 2، بذا يكون ركز ,لفئةالاولى x١ = 22٠25

23.5 =

عين مراكز الفئات الأخرى بنفس الطريقة السابقة، أو باضافة طول الفئة إلى مركز الفئة التي قبلها، إذا كانت الفئات متساوية.بذا يكون مركز الفئة الثانية:

23.5+4-27.5

وهكذا.

9- افرغ البيانات المعطاة لديك على الفئات التي انشأتها، وذلك باستعمال خط عمودي لكل قراءة وخط مائل للقراءة الخامسة في كل فئة، لتسهيل الجمع.

10- عد الخطوط التي حصلت عليها في كل فئة وسجل ذلك في عمود التكرارات. وسنعبر عن التكرار للفئة i بالرمز fj ؛ أي أن تكرار الفئة الأولى هو fl وتكرار الفئة الثانية f 2 وهكذا.

اجمع التكرارات لجميع الفئات وقارنه بعدد البيانات، وتأكد أن مجموع التكرارات يساوي عدد البيانات الأصلية.

ويوضح جدول رقم (5)جميع هذه الخطوات.

جدول رقم (5): التوزيع التكراري للمصروف الشهري لستين طالباً.

(1) حدود الفئة	(2) الحدود الفعلية للفئة	(3) مركز الفئة Xi	(4) افراغ البيانات	(5) التكرار fj
25-22 29-	25.5-21.5	23.5	11 ''''	10•
26	29.5-25.5	27.5	1	4
33-30 37-	33.5-29.5	31.5	1 ///	6
34 41-38	37.5-33.5	35.5	//// لككئ	8
45-42	41.5-37.5	39.5	1 ////1 ////	9
49-46	45.5-41.5	43.5	1	9
المجموع	49.5-45.5	47.5		17
				60

عند عرض التوزيع التكراري لا تكتب عمود افراغ البيانات، وفي بعض الأحيان لاتكتب عمود الحدود الفعلية للفئات بل تكتفي بعرض العمودين (1) و (5)، أو العمودين (2) و(5)، أو (3)و (5) .

ويمكن تلخيص خطوات بناء التوزيع التكراري فيما يأتي: - عين عدد الفئات المتساوية.

- عين المدى.
- عين طول الفئة بقسمة المدى على عدد الفئات ثم قم بالتقريب لأقرب وحدة.
- عين لحد الأدنى للفئة الأولى ثم اطرح منه 0.5 لتتين الحد الأدنى الفعلي للفئة الأولى، إذا كانت البيانات أعداداً صحيحة.
- عين لحد الأعلى الفعلي للفئة الأولى بإضافة طول الفئة إلى الحد الأدنى الفعلي لتلك الفئة، ثم اطرح منه 0.5 لتعيين الحد الأعلى للفئة الأولى.

- عين الحدود الدنيا والعليا والحدود الدنيا الفعلية والعليا الفعلية للفئات الباقية، وذلك بإضافة طول الفئة لكل حد على التوالي.
- عين مراكز الفئات وذلك بتعيين مركز الفئة الأولى ثم إضافة طول الفئة له لتعيين مركز الفئة الثانية وهكذا.
- افرغ البيانات على الفئات.
- سجل مجموع تكرارات كل فئة أمامها في عمود التكرارات.
- اجمع التكرارات لجميع الفئات وقارن حاصل الجمع بعدد البيانات للتأكد من أنهما متساويان ا

عند بناء التوزيع التكراري يجب مراعاة النقاط التالية:

1- أن تكون الفئات منفصلة عن بعضها البعض (غير متداخلة فيما بينها).

2- أن تكون الفئات متساوية في الطول ما أمكن.

3- أن تكون الفئات كافية لاحتواء جميع البيانات.

وهذا يعني أن أي قيمة في البيانات إنما يمكن وضعها في فئة واحدة فقط.

وبهذا نتمكن من افراغ جميع البيانات في فئات التوزيع التكراري وبدون التباس، وسيكون مجموع التكرارات مساويا لعدد البيانات.

لقد قمت ببناء التوزيع التكراري في المثال (7)؛ فهل تستطيع أن تبنى توزيعاً تكرارياً لبيانات مقربة لرقم عشري واحد؟ (أي إذا احتوت البيانات على خانة عشرية واحدة مثل 23.4)

كيف تقوم ببناء هذا التوزيع؟

الأمر بسيط للغاية، اتبع الخطوات العشر السابقة وانتبه إلى ما يلي:

1- في الخطوتين (5)، (6) اطرح 0.05 من الحد الأدنى لكي تحصل على الحد الأدنى الفعلي، لأن البيانات الأصلية فيها خانة عشرية واحدة، أي أن وحدة دقة القياس 0.1 وبالتالي فإن نصف وحدة دقة يساوي 0.05). وإذا كانت البيانات أعداداً صحيحة فاطرح 0.5 كما في الخطوتين (5)، (6).

2- أما إذا كانت البيانات تحتوي على خانة عشرية واحدة فإننا نطرح 0.05 في الخطوتين (5)، (6) عند تحديد الحد الأدنى الفعلي... وإذا كانت البيانات تحتوي على خانتين عشريتين فإننا نطرح 0.005 في الخطوتين (5)، (6) عند تحديد الحد الأدنى الفعلي.

تدويص (2) لحا

تمثل البيانات التالية عدد انامات المعتمدة التي نجح فيها ثلاثون طالبا في نهاية التة الثالئآ الجاسين.

60	87	70
72	78	67
81	86	88
67	62	92
73	77	56
83	75	61
87	55	64
59	67	68
91	81	90
82	57	91

ضع هذه البيانات في توزيع تكراري ذي 6 فئات متساوية.

2.2.2 التكرارات النسبية

بعد بناء التوزيع التكراري ومعرفة تكرار كل فئة فيه، قد يتبادر إلى ذهنك أن الاهتمام ربما ينصب على نسبة الأفراد أو العناصر في كل فئة لا على العدد بحد ذاته. فعلى سبيل المثال :لوقرأت توزيعاًتكرارياًيمثل المعدات المئوية لجميع الطلبة الناجحين في امتحان الدراسة الثانوية في الأردن في عام معين، فقد تسأل :ما نسبة الحاصلين على معدل ما بين 90،85 أو على متدل أعلى من 90 ؟ في هذه الحالة لا يكون اهتمامك منصباً على عدد الطلبة الحاصلين على معدل ما بين 85،90 بل على نسبة عدد الطلبة في تلك الفئة إلى المجموع الكلي لعدد الطلبة الناجحين. وهذا يقودك إلى تعريف التكرار النسبي.

نهوبن (1) ______________________________ ئا

لكل التت لى فئة هو تتة تكل تلك افة لى مجموع التكرارات.

فإذا كان مجموع التكرارات n وكان تكرار فئة معينة f ، فإن التكرار النسبي لتلك الفئة هو ٤

p =

n

ويسمى الجدول الذي يعطينا الفئات (أو مراكزها) مع تكراراتها النسبية (التوزيع التكراري النسبي) وعليه فإن الجدول رقم (6) يمثل التوزيع التكراري النسبي للمثال (7).

جدول رقم (6)

النكراد السبي	حدود الفئة
0.167	25-22
0.067	29-26
0.100	33-30
0.133	37-34
0.150	41-38
0.150	45-42
0.233	49-46
1.000	المجموع

لاحظ أن مجموع التكرارات النسبية يجب أن يساوي 1 .

قشوبيب (3)

جد التوزيع التكراري النسبي للبيانات في تدريب (2).

3.2.2 التوزيع التكراري المئوي

إذا ضربت كل تكرار نسبي في 1 00% فإنك تحصل على التكرار المئوي. ويسمى الجدول الذي يعطي الفئات (أو مراكزها) مع تكراراتها المئوية "التوزيع التكراري المئوي".

ومما سبق ترى أن الجدول رقم (7) الآتي يمثل التوزيع التكراري المئوي للمثال

جدول رقم (7)

حدود الفئة	التكرار المئوي %
25-22	16.7
29-26	٦٢■
33-30	10.0
37-34	13.3
41-38	15.0
45-42	15.0
49-46	23.3

قهوتبي (4)

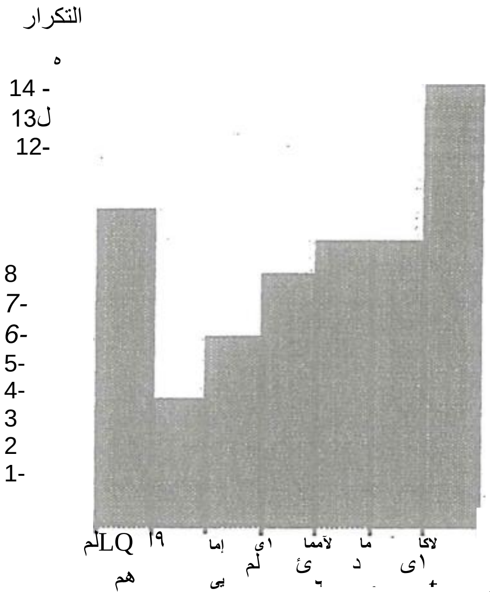
اكتب التوزيع التكراري المئوي للبيانات في تدريب (2).

3.يل ر (-

سندرس سويا ثلاث طرق لتمثيل التوزيعات التكرارية بيانيا، وهذه هي:

1.3 المدرج التكراري

هو عبارة عن تمثيل تكرار كل فئة من فئات التوزيع التكراري ذي الفئات المتساوية بمستطيل حدود قاعدته هي الحدود الفعلية لتلك الفئة، يتناسب ارتفاعه مع تكرارها. ومعنى هذا أنك ستأخذ محورين متعامدين، على المحور الأفقي منها ترصد الحدود الفعلية لكل فئة من فئات التوزيع التكراري وتقيم على كل فئة مستطيلا يتناسب ارتفاعه مع تكرار تلك الفئة، كما هو موضح في الشكل رقم (4) والذي يمثل المدرج التكراري للتوزيع التكراري المعطى في جدول رقم (5).



دينار

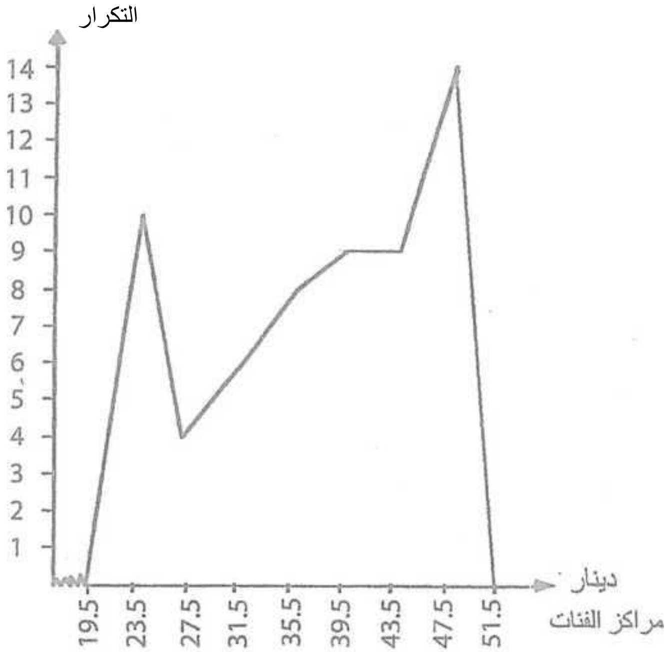
الحدود الفعلية للفئات

الشكلرقم (4)

2.3 المضلع التكراري

المضلع التكراري مضلع مغلق تحصل عليه بتنصيف الأضلاع العلوية للمستطيلات في المدرج التكراري ثم وصل هذه النقاط بعضها مع بعض . ولكي تغلق الخط المنكسر الذي تحصل عليه اعتبر أن هناك فئتين متطرفتين، واحدة إلى أقصى

اليسار والثانية إلى أقصى اليمين، وأن تكرار كل منهما صفر. (أي أن ارتفاع كل من المستطيلين المقامين على هاتين الفئتين صفر).وبعد اقامة المستطيلات المذكورة خذ مركز كل من هاتين الفئتين واغلق المضلع كما في الشكل رقم (5) الذي يمثل المضلع التكراري للتوزيع التكراري في جدول رقم (5).



الشكل رقم (5)

اعتبر مركز كل فئة احداثياً أفقياً واعتبر تكرار هذه الفئة هو الاحداثي العمودي لتلك النقطة . ارصد جميع هذه النقط ثم صل بينها بخطوط مستقيمة . عين مركز فئة قبل الفئة الأولى مباشرة واعتبر تكرارها صفراً، ثم عين مركز فئة بعد الفئة الأخيرة مباشرة واعتبر تكرارها صفراً أيضاً . ارصد هاتين النقطتين واغلق بواسطتهما المضلع التكراري كما هو موضح بالشكل رقم (5) أما النقاط فهي:

(43.5,9)، (39.5,9)، (35,5,8)، (3.56)، (27,5,4)، (23.5,10)،

(19.5,0)، (51.5,0)، (47.5,14)

3.3 المنحنى التكراري

إذا مهدت المضلع التكراري، عزيزي الدارس، وجعلته منحنى بدلا من خطوط منكسرة فإنك تحصل على المنحنى التكراري. لاحظ أنه ينبغي عدم رسم المنحنى التكراري إلا إذا كانت الفئات كثيرة العدد، ذات طول صغير، وكان عدد البيانات كبيراً، وهي من النوع المتصل مثل الزمن والوزن.

إن الطريقة التي مثلت بها التوزيع التكراري بيالياً تستطيع استعمالها لتمثيل كل من التوزيع التكراري النسبي، والتوزيع التكراري المئوي، وذلك باستعمال التكرار النسبي أو التكرار المئوي على المحور العمودي بدلاً من التكرار.

هناك عدة أمور إذا فقدت كلها أو بعضها لا يكون المضلع مضلعاً تكرارياً، وهي: 1- يجب أن يكون مجموع التكرارات مساوياً لمجموع البيانات الأصلية.

2- يجب أن لا يكون هناك أي تكرار سالب؛ أي ألا يقع أي جزء من المضلع التكراري أسفل المحور الأفقي.

3- يجب أن يكون هناك قيمة واحدة فقط لتكرار أي فئة من الفئات؛ أي أن تقابل كل فئة قيمة واحدة للتكرار، ولا يجوز أن تقابلها قيمتان أو اكثر.

تدريب (5)

ارسم المدرج التكراري والمضلع التكراري للتوزيع التكراري الذي قمت ببنائه في تدريب (2).

أسئلة التقويم الذاتي (1)

عدد طرق تمثيل التوزيعات التكرارية بيانياً.

4.التماثلوالالتواءوالتفرطحفيالتوزيعاتالتكرارية ق من أهم الصفات التي تميز التوزيعات صفات التماثل والالتواء والتفرطح، وتعتبر هذه الصفات من أهم خواص أشكال التوزيعات . ونقوم هنا بوصفها وصفاً انشائياً:

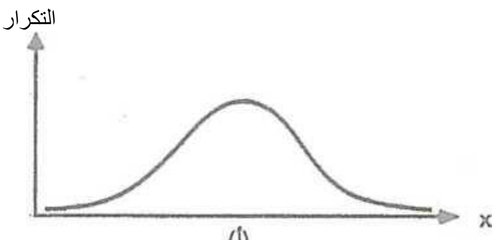
1٠٠ إن أول تمييز في الشكل هو التمييز بين التوزيعات المتماثلة والتوزيعات غير المتماثلة.

فما هو التوزيع المتماثل؟

يعتبر التوزيع متماثلا إذا أمكن إقامة عمود على المحور الأفقي يقسم التوزيع إلى قسمين ينطبقان على بعضهما تمام الانطباق، بحيث لو وضعت مرآة عمودية على المحور الأفقي فإنها تقسم التوزيع إلى قسمين أحدهما صورة طبق الأصل للآخر.وبعبارة أخرى يكون التوزيع متماثلاً إذا أمكن طيه بحيث ينطبق أحد نصفيه على الآخر تمام الانطباق. وتسمى النقطة التي تقع على المحور الأفقي والتي يتم الطي عندها (نقطة تماثل)، كما يسمى الاحداثي العمودي المار بها (محور تماثل) المنحنى.

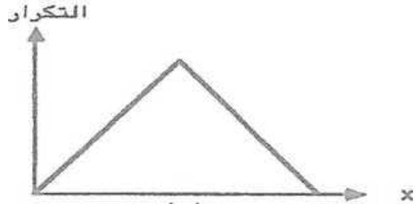
ولا يوجد في الحياة العملية إلا عدد قليل من التوزيعات المتماثلة، ولكن هناك كثير من التوزيعات التي تكون تقريباًمتماثلة.

ويظهر في الشكل رقم (6) أ.ب، ج، د بعض التوزيعات المتماثلة.



(أ)

الشكل رقم (6-أ)



(ب)

الشكل رقم (6-ب)

التكرار

٦ حبا–هش—————————تب

(ج)

الشكل رقم (6-ج)

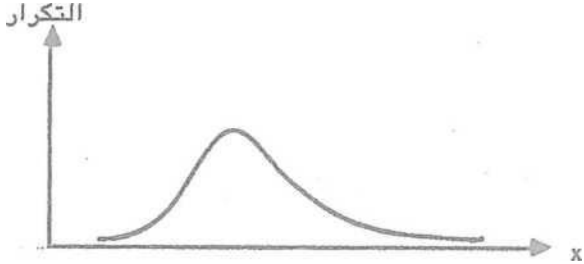
التكرار

خ

الشكل رقم (6-د)

أما التوزيعات التي يكون عدم تماثلها واضحا فتسمى توزيعات ملتوية. ويكون التوزيع ملتويا إذا امتد أحد طرفيه إلى اليمين كثيراً أو إلى اليسار كثيراً، و يكون ملتوياً أيضاً إذا كان عالياً من جهة ومنخفضاً من الأخرى.

و إذا كان طرف التوزيع ممتداً إلى اليمين (أي في الاتجاه الموجب) قلنا هذا التوزيع ملتو نحو اليمين أو موجب الالتواء كما في الشكل رقم (7). أما إذا كان ممتداً نحو اليسار (أي في الاتجاه السالب) فنقول أن هذا التوزيع ملتو نحو اليسار أو سالب الالتواء كما في الشكل رقم (8).



ا لتوز يع ملتو نحو ا ليمين (موجب الالتواء)

الشكل رقم (7)

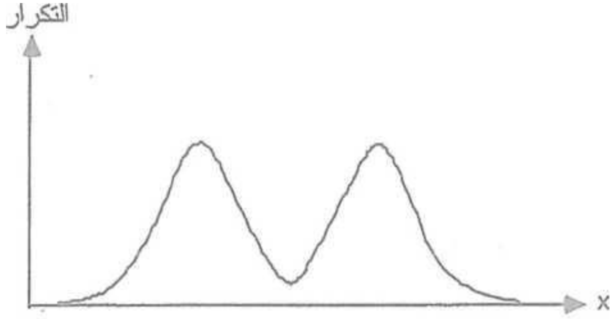
التكرار

L-A,

التوزيع ملتو نحى اليسار
(سالبا لا لتواء)
الشكل رقم (8)

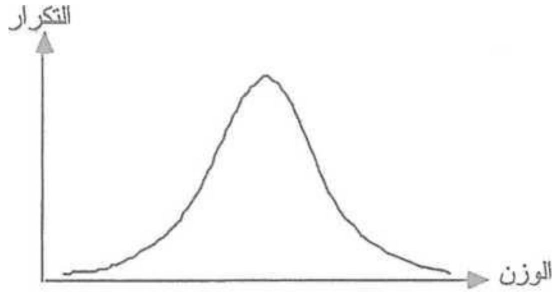
ولترسيخ فكرة التوزيعات المتماثلة إليك بعض الأمثلة على ذلك، افرض أنك رميت زهرة نزد منتظمه 1200 مرة وبنيت توزيعاً تكرارياً للأعداد التي ظهرت إلى أعلى، فالغالب أنك ستجد التكرارات متقاربة، و إذا رسمت المضلع التكراري، فإنك ستجده في الغالب على شكل مستطيل، وهذا يعني ان التوزيع متماثل.

وهذا مثال آخر إذا سجلت المسافات التي يسجلها المنتخب الوطني في رمي القرص في بلد معين، ثم سجلت المسافات التي يسجلها منتخب رمي القرص في إحدى المدارس الاعدادية، وبعد ذلك بنيت التوزيع التكراري لجميع المسافات التي سجلتها، ومن ثم رسمت المضلع التكراري والمنحنى التكراري لذلك التوزيع فإنك في الغالب ستحصل على شكل يشبه الشكل رقم (9) الذي يظهر أنه متماالل.وفيه يكون الجزء الأيمن خاصاً بالمنتخب الوطني الذي سجل مسافات طويلة، والجزء الأيسر خاصاً بمنتخب المدرسة الذي سجل مسافات أقصر.



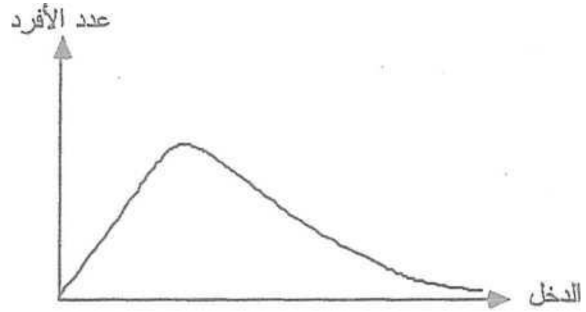
الشكل رقم (9)

أما إذا سجلت أوزان 100 رجل اخترتهم عشوائياً ثم وضعت القيم التي حصلت عليها في توزيع تكراري ورسمت المضلع التكراري لهذا التوزيع ثم قمت بتمهيده فإنك في الغالب ستجد المنحنى التكراري الذي حصلت عليه بتمهيد المضلع التكراري يشبه شكل الجرس ويكون هذا التوزيع متماثلا كما هو واضح في الشكل رقم (1 0).



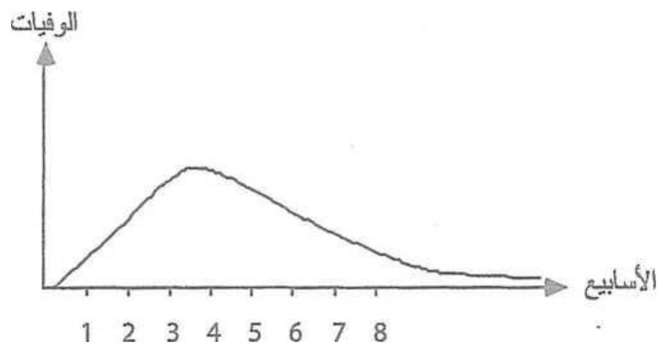
الشكل رقم (10)

أما الأمثلة على الالتواء فمنها توزيع الدخل في قطر معين، حيث يكون في الغالب عدد كبير من الأفراد ذوي دخل قليل؛ أما الافراد الذين دخلهم كبير فيكون عددهم صغيراً وعادة يكون المنحنى التكراري للدخل ملتوياً نحو اليمين كما يظهر في الشكل رقم



الشكل رقم (11)

وقد أظهرت بعض الاحصائيات أن المنحنى التكراري لوفيات حديثي الولادة حسب العمر (من لحظة الولادة حتى الأسبوع الخامس) ملتوياً نحو اليمين كما يظهر في الشكل (12).

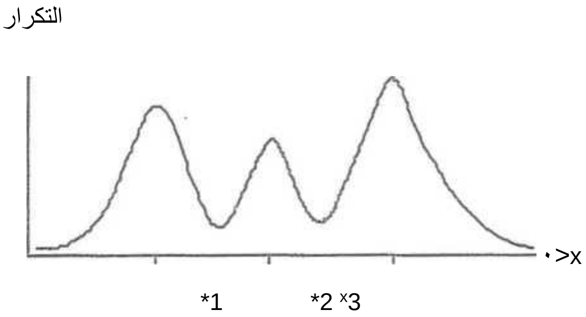


الشكلرقم (12)

2- ها قد عرفت التمييز الأول الذي نصف به التوزيعات التكرارية، أما التمييز الثاني فهو كونها ذات منوال واحد أو عدة منوالات.والمنوال هو القيمة التي يكون تكرارها أكبر من تكرار القيم التي في جوارها، أي النقطة التي يقابلها قمة. وفي الحالة التي يكون فيها للتوزيع منوال واحد نقول أن التوزيع أحادي المنوال، وحين يكون هناك منوالان نقول أن التوزيع ثنائي المنوال، وهكذا.فالشكل رقم (0 1) يظهر توزيعاً أحادي المنوال، لأن هناك نقطة واحدة يقابلها قمة.

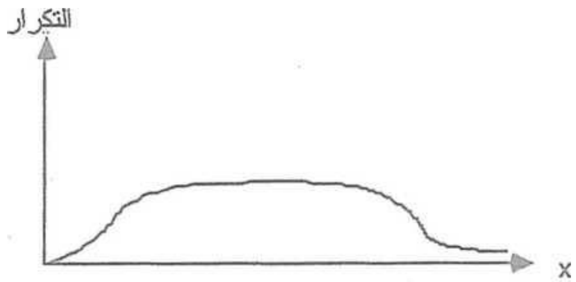
أما الشكل رقم (9) فهو ثنائي المنوال لأن فيه نقطتين يقابل كلا منهما قمة.

لاحظ الشكل رقم (13) تجد أن Xl ,Xj ,3* كلها منوالات لأنه يقابل كلا منها قمة، أي أن التكرار على كل منها أكبر من التكرار على النقاط اللاتي في جوارها.



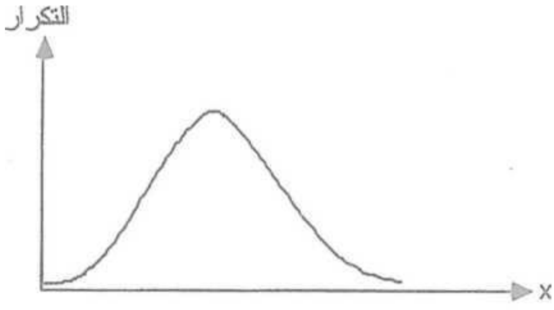
الشكل رقم (13)

3- والتمييز الثلث الذي نصف به التوزيع التكراري هو تلرطح نتله التوزيع، واللفرطح هو الاستواءوعدم كون التوزيع مدبباً، فإذا قلت أن التوزيع كبير التفرطح فذلك يتني أنه كبير الاستواء أو غير مدبب، كما يظهر في الشكل رقم (14).



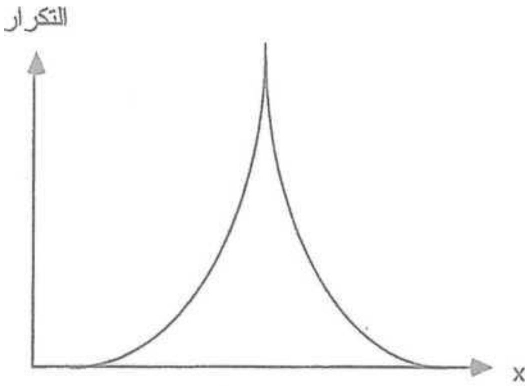
الشكل رقم (14)

أما التوزيع الذي يكون متوسط التفرطح فانه لا يكون مستويا بالكامل ولا يكون مدبباً بالكامل، كما يظهر في الشكل رقم (15)



الشكل رقم (15)

وبعض التوزيعات التكرارية يكون قليل التفرطح، أو بعبلرة أخرى يكون مدببا كما هو واضح لك في الشكل رقم (1 6).



الشكل رقم (16)

هذا ولا حاجة إلى تذكيرك أن الوصف السابق للتوزيعات التكرارية كان مبنيا على الأشكال التوضيحية التي رسمتها لتمثل المضلع أو المنحنى التكراري.

مثال (8)

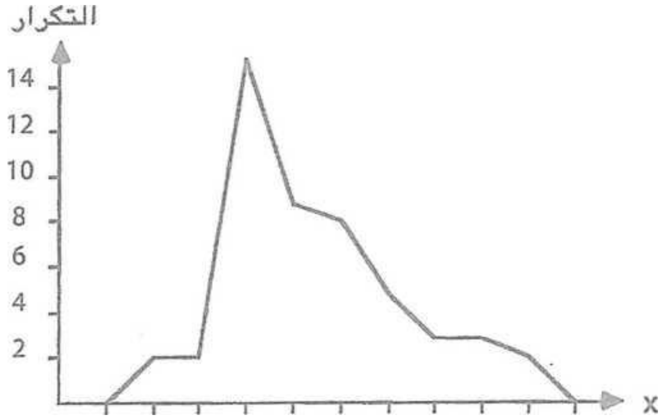
الجدول رقم (8) يمثل التوزيع التكراري لعلامات 50 طالبافي أحد الامتحانات حاول أن تصف هذا التوزيع.

الجدول رقم (8)

X	f
9	2
10	2
لالا	16
12	9
1٢	8
14	5
15	3
16	3
17	2

الحل:

ارسم المضلع التكراري لهذا التوزيع فتجده كما في الشكل رقم (17).



89101112131415161718

الشكل رقم (17)

إن مجرد النظر إلى هذا التوزيع يبين أن تكرار العلامة 11 هو 16، ولايوجد علامة أخرى يقابلها تكرار أكبر من 16 أو.يساوي 16 .لذلك فإن هذا التوزيع أحادي المنوال حيث قيمة المنوال 11، لاحظ أن التوزيع موجب الالتواء حيث يمتد طرفه نحو اليمين وبالنظر إلى المدرج التكراري أو المضلع التكراري للتوزيع فإن الصفتين السابقتين تتضحان من الشكل رقم (17).

4- بالاضافة إلى الصفات الثلاث السابقة للتوزيعات التكرارية تجدفي كثيرمن الأحيان أن هناك تسميات معينة لبعض التوزيعات التكرارية. وتصف لك التسميات التوزيع وصفاً دقيقاً بدون ذكر صفات التماثل والالتواء والمنوال.وأكثر ما تتضح هذه التسميات في الأمثلة الآتية:

أ- توزيع متجانس (مستطيل) وشكله كما في الشكل رقم (18).

التكرار

الشكل رقم (18)

فعندما نقول عن توزيع ما أنه متجانس فإننا نعلم أنه متماثل وأنه ليس له منوال

واحد.

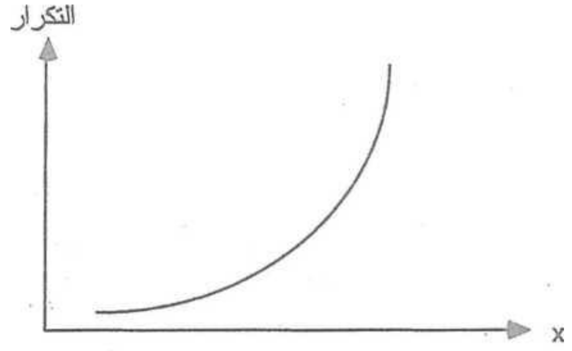
ب- توزيع لا وهو كما في الشكل رقم (19)

التكرار

الشكل رقم (19)

قم

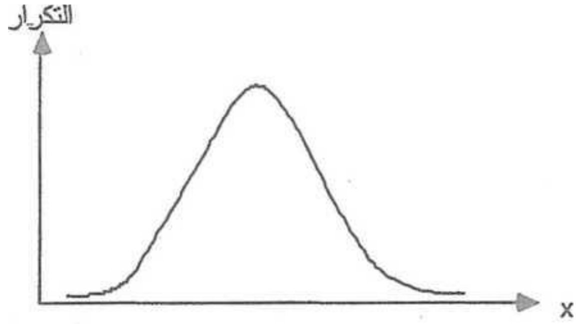
ج- توزيع ل وهو كما في الشكل رقم (20)



الشكل رقم (20)

فإذا كان التوزيع على شكل لوالتواؤه إلى اليسار فهو توزيع أحادي المنوال تهفي أقصى ليمين.

د- توزيع شكل الجرس وهو كما في الشكل رقم (21).



الشكل رقم (21)

وهو توزيع متماثل وأحادي المنوال.

قهويبب (6)

صف التوزيع التكراري الآتي:

العلامة	التكرار
8	2
9	3
10	9
11	7
12	5
13	4
14	3
15	2
16	2
17	1
18	1

أصلة ائجيهأنص.'(2)

1- اذكر التمييزات الأربعة التي نستطيع بها وصف المضلع التكراري لتوزيع معين.
2- عرق منوال التوزيع التكراري.
3- عرف الالتواء إلى اليمين.
4- ماذا نعني بقولنا : هذا التوزيع التكراري كبير التفرطح.

قرص If I ||ى|طذاةدات لظابر٠ة معية، هتكن ذ٦ائج تجره ط أَر علامات الطلبة الناجحين في فحص شهادة الدراسة الثانوية في بلد ما في إحدى السنوات، والمطلوب منك أن تعطي تحليلا لهذه النتائج.

لا شك أن أول عمل تقوم به هو وضع هذه النتائج في جدول، لكنك سرعان ما تدرك كثرة مفردات ذلك الجدول وصعوبة استنتاج أي فائدة منه. لذا فإنك تعمد إلى وضع هذه البيانات في توزيع تكراري غير مجمع، أي أنك ترتب البيانات ترتيباً تصاعدياً وترصد تكرار كل علامة، من أدنى علامة إلى أعلاها. لا تكون قد خسرت أية معلومة من البيانات، لكن اعطاء تحليل دقيق وسريع لها لايزال صعباً. لذلك عليك أن تضع هذه البيانات في توزيع تكراري ذي فئات متساوية وعددها معين ومعقول إنك بهذه الخطوة تكون قد كثفت البيانات وجعلت طريقة عرضها أسهل، وإن فقدت بعض المجلومات عنيا. فأنت الآن لا تعلم التكرارات التي تقابل كل قيمة داخل الفئة، لكنك تعلم عدد المفردات (التكرار) للفئة بكاملها . إنك تضحي بشيء من المعلومات في سبيل تسهيل عرض البيانات وسرعة الاستفادة منها.

حالما تضع البيانات في توزيع تكراري ذي فئات يصبح بالإمكان دراسة شكل التوزيع وطبيعته كالتماثل والالتواء وعدد المنوالات. لكن هذا لا يكفي، فأنت تود ايجاد مقاييس عددية تصف لك البيانات التي بين يديك . وأول المقاييس التي تفكر فيها مقاييس النزعة المركزية هو التوسط أو المعدات. وهي عدة مقاييس عددية تعين موقع التوزيع. فقد يكون هناك توزيعات تكرارية متشابهة في طبيعتها وشكلها، لكنها تختلف في مواقعها. وفي مثل هذه الحال تكون معرفة المعدات أو مقاييس النزعة المركزية ذات فائدة في دراسة الفرق بين هذه التوزيعات التكرارية.

على هذا الأساس يمكن القول أن المعدلات تعطيك قيمة تمثل العينة أو المجتمع الذي تدرسه، لكن بصورة غير دقيقة. ويتبلور هذا التمثيل إلى حد ما في ايجاد القيمة التي ر تتمركز حولها معظم المشاهدات . وهذه القيمة المتوسطة تمثل المجتمع أكثر من أية وحدة من مفرداته.وتتلخص الاستعمالات الرئيسة للمعدات فيمايلي:

1- أنها تعطيك قيمة حالة أو ظاهرة معينة تمثل مجموعة معينة من الأفراد، وعادة ما تكون هذه القيمة مفيدة في الحياة العملية. فأنت مثلاً تستطيع تحديد عدد السمال المطلوبين لانجاز عمل معين، إذا علمت معدل ما يستطيع العامل الواحد القيام به في اليوم.

2- أنها تعطيك قاعدة لمقارنة حالة معينة لمجموعة تحصيل مع نفس تلك الحالة لمجموعة ثانية. فمثلا لمقارنة تحصيل صف معين مع صف آخر، يمكنك استعمال معدل كل منهما كقاعدة للمقارنة.

3- أنها تسمح بتقدير حالة أو صفة لأفراد كثيرين إذا علمنا قياسات عن جزء من المجموع الكلي لهؤلاء الافراد، فمثلا إذا أردت تقديرحياة نوع من المصابيح الكهربائية التي ينتجها مصنع ما، فإنك تأخذ عدداًمن هذه المصابيح وتسجل حياة كل منها ثم تنتعمل معدل الحياة لملعينة المدروسة كتقدير لمعدل حياة ذلك النوع من المصابيح.

ويجب أن لا يغيب عن ذهنك أن المعدلات تعطي صورة (جزئية) عن المعلومات المحتواة في البيانات.

وإذا ما حقق المعدل (المتوسط) كل الصفات الآتية أو معظمها فإنه يعد مقبولا، وهذ ء الصفات:

1- يجب أن يكون المتوسط معرفا تتريفاً دقيقاً.

2- يجب أن يبنى على جميع أو معظم المشاهدات.

3- يجب أن يكون من السهل فهمه وتفسيره.

4- يمكن حسابه بسهولة وسرعة معقولتين.

5- يخضع للعمليات الجبرية بسهولة.

6- لا يتأثر كثيراً بلالقيم المتطرفة أو الشاذة.

7- لا يتأثر كثيراً باختلاف العينات من مجتمع واحد.

ومن مقاييس النزعة المركزية التي سندرسها معك في هذه الوحدة ما يلي:

1 .5 الوسط الحسابي قحوبف (2)

إذا كان لدينا n من الأعداد (قيم المشاهدات)، فإن الوسط الحسابي لها هو حاصل قسمة مجموعها على عددها. أي أن: *كثتقثتك=ج

n

حيث يمثل الوسط الحسابي.

ويمكن استعمال الرمز S (سيغما) الذي يعني جمع الحدود التي في داخله. وبذلك يصبح تعريف الوسط الحسابي.

كره،

او ئ-,

n

n

هثل(9)____لح

جد الوسط الحسابي للمشاهدات التالية: 22,17,0,5,15

الحل.

الوسط الحسابي يساوي:

8 و ,بيييد

وبما أنك ستحتاج إلى الرمز فى في دراستك هذه الوحدة وغيرها فعليك أن تتعرف على بعض خواصها، وتفهم معناها.

تعني فى جمع الحدود التي في داخلها، وإذا كان المؤشر إلى أسفل يمين المتغير داخل I يأخذ القيم من عدد صحيح إلى عدد صحيح آخر، فهذا يعني أن عليك أن تضع بدلاً منه القيمة الأولى التي يأخذها المؤشر ثم تزيد واحداً تلو الآخر إلى أن تصل إلى العدد الأخير الموضوع فوق I .

لمثلأ، إبا قنا Xj ٤ فهذا يعنه ان الموشر هو . والقيم النى ياخذها ه 1= اثم

2=! ثم 3=1 ثم 4=أ وثم 1=5 ويلله بكون مع Xj ٤ هو٠

$X_1+X_2+X_3+X_4+X5$

هل تستطيع أن تعرف معنى Xj فى ؟ بالرجوع إلى الشرح أعلاه، ستجد أن

؛X آ تعني XjtXj-X . فى تمفر من الاميلن لا بون بشرء الجمع 2عى

متلفرات لوهاموشرمل ؛X ولكن بكون علية الجمع ط الرلن فه ا مل (5بأ) Iاو (5ب^2؛) ت آو مرها.

بالطبع فإن معنى I يبقى (جمع الحدود في داخلها) وأنت تجد قيمة ذلك بأن

تضع بدلامن أ القيمة الدنيا لها، وهي القيمة الموجودة تحت 2 ، ثم تضيف واحدا واحدا وتعوض بدلامن

ألإلى ن تصل إلى القيمة الموجودة في أعلى I .

فستثلأ يكن (3 + 1) ٤ وسادي

(3+2) + (3+ 1) ويساوي 5+3-9

اما (5*»ي فيكنمناها:

(5- 2 3ه (ي 2 2}+■ (5+■ 2 1): (5ب2ا) ٤

= (1 + 5) + (4 + 5) + (9 -+ 5)

= 6+9+14=29

اما (1-أ5) ٤ مغه ليون

(٦ 5x3؛ (ب 5x2)+ (1 1x5)= (1-51) ؤ

-(5-1)(10-1)+(15-1)

= 4+9+14=27

أما عن خواص S فهي:

1-I(xi+yi){xi£yi

2- إذا كان a ثابتاً فإن ,X لة= |aX 1

أي أنك تستطيع استخراج الثابت خارج إشارة I :

مثال ذلك ,X فى3= ,I 3x

أو مل ا ؟،= اه ٢

= 4(1+ 2)= 12

3- إذا كلن a ثلبتأ فلن a = na ٢

وهذا يعني أنك إذا جمعت العدد الثابت a مرات عددها n فان النتيجة تكون na

غى 5=4x5=20 ٤

بيعظة: لاحظ أنه إذا كان داخل الرمز S حاصل ضرب متغيرين لهما نفس المؤشر فيجب لجراء عملية الضرب وبعدها نجمع حواصل الضرب، فمثلا لإيجاد ,۴ ,X ٤ خذ 1=أوأوجدثم خذ2=1 وأوجد ثم خذ3=1 وأوجد، ثم اجمع هذه القيم كلها. بقله يون 3 ۴3 IX 2؛2 x؛ ,x١f= ٤ ,X نج وسسكفيد منهذا فه ترين الوسط الحسابي في حالة التوزيع التكراري، أي البيانات المجمعة؛ حيث يكون معلوماً لدينا حدود الفئات والتكرارات المقابلة لها وحيث تعتبر أن التكرار يتراكم على مركز الفئة. لنلك فإن أول عمل تقوم به هو ايجاد مراكز الفئات إن لم تكن معطاة لك.

ويعرف الوسط الحسابي في حالة التوزيع التكراري كما يلي: تحوبيفه (3)
اذا كان لديك توزيع تكراري عدد فئاته h وكانت مراكز الفئات (أو القيم) ,$x_h,...،x_2x$
والتكرارات المقابلة لها أب,...,۴o فإن الوسط الحسابي يكون:

h ب f + -2■ +ff

وباستعمال الرمز Z يكون:

دء

حيث ——— = x تمثل مجموع التكرارات

n

مثاق (10)

احسب الوسط الحسابي للتوزيع التكراري الآتي:

التكرار	مركز الفئة X
3	'(0
4	1S
2	20
/	25
2	30

الحل: بما أن مراكز الفئات معطاة، فإننا نحسب الوسط الحسابي مباشرة من التتريف:

-103+15٧4+20٧2+25ا71302365٧

3ب4ب2ب7 +2 18

= 20.28

م(أ1)   - - لتتكاا

جد الوسط الحسابي للتوزيع التكراري الآتي:	

حدود الفئات	التكرار
26-22	Q
31-27	3
36-32	10
41-37	8
46-42	ذا
51-47	8

الحل:

نجد مراز الفئات اولا، ثم نضرب مركز كل فئة في التكرار المقابل لها كما يلي:

مركز الفئة X|	التكرار f,	Xi fi
2٢	9	216
29	3	87
34	“TO	340
٠ وق ٠-'		312
44	12	528
49	ئ	392
المجموع	50	18775

الوسط الحسابي = ٦٢=37.5

قهوتيب (7)   ١١

أوجد الوسط الحسابي للتوزيع التكراري الآتي:

مركز الفئة X,	التكرار . fi
55	10
65	15
75	22
85	13
95	05

أعئلة]1تقويبمالثاتتي (3)

1- عرف الوسط الحسابي.
2“ ما خواص الوسط الحسابي.

2.5 الوسيط

1.2.5 الوسيط للبيانات الأولية (الخام)

في كثير من الأحيان ينصب اهتمام الدارس على معرفة القيمة التي يكون عدد الأفراد الحاصلين على قيمة أقل منها مساوياً لعدد الحاصلين على قيمة أعلى منها، فمثلاً : عندما تحصل على علامة معينة في امتحان ما، هل تقع علامتك في النصف الأول من العلامات أم في النصف الثاني؟ في مثل هذه الحالة، عليك أولا أن تعرف تلك العلامة التي حصل نصف الطلبة على أقل منها ونصفنهم الآخر على علامة أعلى منها. إذا عرفت هذه العلامة يكون بإمكانك الإجابة عن سؤالك السابق.

فمثلا إذا كانت العلامات التي حصل عليها طلبة الصف هي:

20،18،7،10،19,1712،13،8

فما هي العلامة التي حصل نصف الطلبة على أقل منها والنصف الآخر على أعلى منها؟ للإجابة عن هذا السؤال عليك أن ترتب العلامات ترتيباً تصاعديا؛ أي من الأصغر إلى الأكبر، فيكون لديك:

20،19،18،1713،12،10،87

من الواضح أن العلامة المطلوبة هي 13، فهناك أربع علامات أصغر منها وأربع علامات أعلى منها.وتسمى العلامة 13 هذه الوسيط للبيانات المعطاة. وعند تعريف هذا الوسيط انتبه إلى التفريق بين كون عدد البيانات فردياً أو زوجياً.

والأن ما هو الوسيط؟

(jjg د (4)

الوسيط لمجموعة من الأعداد المرتبة ترتيباً تصاعدياً أو ترتيباً تنازلياً هو العدد الأوسط منها إذا كان عددها فردياً، والومنط الحسابي للعددين الأوسطين في المجموعة إذا كانعددها زوجيا.

هثل (12)

ما الوسيط للبيانات التالية: 12،5،17،8،9،2

الحل:

بما ان عدد البيانات زوجي، فإنه يوجد عددان لوسطان بعد ترتيبهم ترتيباً تصاعديا:

17،12،9،8،5،2

فإنه من الواضح أن العددين الأوسطين هما 8، 9

إن يكون لوسيط ح "=8.5

لاحظ أن هناك ثلاثة اعداد أقل من 8.5 وهي (2، 5، 8)، وثلاثة أعداد أكبر منها وهي (17،12،9).

قعوبف (5)

إذا كان $X_1, X_2, ... X$ مجموعة من الأعداد المرتبة تصاعدياً (أو تنازلياً) فإن الوسيط لهذه المجموعة هو العدد x_{n+1} إذا كان n فردياً، وهو العدد □م ت٢ل إذا وا 2 لم2 '2

آ

كان0زوجيا.

لاحظ أن x_n هو العدد الذي رقمه بعد الترتيب اب— وأن X هو العدد 2 2 2 الذي رقمه ق وهكذا.

مثال (13)

ما الوسيط للبيانات التالية:
18,11,17,21,25,27,33,40,5,15,7

الحل:

بعد ترتيب البيانات حسب قيمتها تصاعدياً تصبح:

5،7،1،15،17،18،21،25،27،33،40

لاحظ أن عددها [1 أي فردي، إذن —ـ —=- =6 ، وبذلك يكون العدد 2 2 السادس في الترتيب هو الوسيط ويساوي 18 .

هئل(ا4)

جدالوسيط للأعداد 8,10,11,12,19,23

الحل:

البيانات مرتبة ترتيباً تنازلياً وعددها 6 (أي أنه زوجي)، إذن فالوسيط هو :

الوسط الحسابي للعددين الثالث والرابع، وبالتالي فإنه يساوي: $\frac{11+12}{2} = 11\frac{1}{2}$

مثال (15)

جد الوسيط للأعداد 17,5,14,3,6

الحل:

رتب الأعداد ترتيباً تصاعدياً فتصبح: 17,14,6,5,3

والآن، لاحظ أن عدد البيانات 5 (أي انه فردي). إذن الوسيط هو العدد الذي تبيبه اي انه العد الذو وتييه الثالث، وبلنالى فلوسيط هو العلد 6 .

سديع٠ (8) ل١]

جد الوسيط للبيانات:

1- 30.2 ،19.5 ،11.4 ،9.1 ،8.8 ،26.5 ،25.0 ،12.4 ،17.8

2-جد الوسيط لبيانات 1,100,19,27,12,8

أسلةالققويم الاقي 4) 0

عرف الوسيط لبيانات عددها فردي.

2.2.5 الوسيط للتوزيع التكراري ذي الفئات

عزيزي الدارس، بعد أن تعرفت على طريقة ايجاد الوسيط من البيانات الأولية، فأنت في حاجة إلى معرفة طريقة ايجاد الوسيط للتوزيع التكراري ذي الفئات، حيث يتم عرض كثير من البيانات بواسطة التوزيع التكراري.

ولما كان الوسيط هو القيمة التي يكون عدد التكرارات الأقل منها مساوياً لعدد

التكرارات الأعلى منها، فإنك في هذه الحالة تحتاج إلى معرفة التكرار المتجمع المقابل لأي فئة. وأنت تحصل على هذا التكرار المطلوب بايجاد مجموع التكرارات التي تساوي قيمها الحد الأعلى الفعلي لتلك الفئة أو تقل عنه. ونبدأ دائما بالحد الأدنى الفعلي للفئة الأولى، ونعتبر تكراره المتجمع صفراً، حيث لاتوجد بيانات تقل قيمتها عن ذلك الحد أو تساويه.

مثال (16) ____ ج

جد التكرار المتجمع في التوزيع التكراري:

حدود الفئات		التكرار
28-25	28.5-24.5	5
32-29	32.5-28.5	8
36-33	36.5-32.5	6
40-37	40.5-36.5	9

الحل.

من الواضح أن عدد المشاهدات التي قيمها أقل من 24.5 هو صفر. أما عدد البيانات التي قيمها أقل من 28.5 فيو 5.

استمر في جمع تكرار كل فئة لمجموع تكرارات الفئات التي تسبقها لتحصل على التكرار المتجمع كما في الجدول رقم (9-أ).

جدول رقم (9-أ)

الحدود الفحلية	التكرار المتجمع
28.5-24.5	5
32.5-28.5	13
36.5-32.5	19
40.5-36.5	28

ليكن واضحا لديك أن ما كتبته في الجدول رقم (9-أ) إنما هو لتوضيح مفهوم التكرار المتجمع المقابل لكل فئة، ولأغراض حساب الوسيط. لذا يجب فهمه بالمعنى الآتي الموضح في الجدول رقم (9-ب).

الجدول رقم (9-ب)

الحدود الفعلية للفئات	التكرار المتجمع
أقلمن 24.5	0
أقلمن 28.5	5
أقلمن 32.5	13
أقلمن 36.5	19
أقلمن 40.5	28

اكتب عمود (الحدود الفعلية) وعمود (التكرار) وعمود (التكرار المتجمع) للمثال السابق في جدول واحد .والآن، نعود بك إلى ايجاد الوسيط للتوزيع التكراري.

قحوببف (6)

إذا كان لديك توزيع تكراري مجموع تكراراته n، فإن الفئة الوسيطية هي أول فئة يزيد تكرارها المتجمع عن: او يساويه.

خل_(17)________________________________ج

ما الفئة الوسيطية في المثال السابق ؟

الحل:

28 = 14،0=تل
2

أنظر الجدول (9-أ). ما أول فئة يزيد تكرارها المتجمع عن 14 أو يساويه؟ لاحظ أن العدد 19 هو الذي-يزيد عن 14، لذلك فالفئة الوسيطية هي 32.5-36.5 ولايجاد الوسيط للتوزيع التكراري. افرض ما يلي:

مجموع التكرارات = n

طول الفئة الوسيطية = C

الحد الأدنى الفعلي للفئة الوسيطية = a

تكرار الفئة الوسيطية = f_m

التكرار المتجمع للفئة التي تسبق الفئة الوسيطية مباشرة أ٢١ وبعبارة أخرى 1٢١ هو مجموع التكرارات التي تقل قيمها عن a

نج

n
—-n.
اوسيط لتزيع اتكرلري هو: M =a غب—xC

حيث أشرنا سابقا إلى دلالة كل رمز من الرموز المستخدمة في هذا التعريف.

هئل (18)

جد الوسيط للتوزيع التكراري للمصروف الشهري لستين طالبا والمعطى في الجدول رقم (5).

الحل:

خذ عمود الحدود الفعلية وعمود التكرار ثم أوجد التكرار المتجمع واكتب النتيجة كمافي الجدول رقم (10.

جدول رقم (10)

الحدود الفعلية للفئة	التكرار	التكرار المتجمع
25.5-215	10	10
29.5-25.5	4	14
33.5-29.5	6	20
37.5-33.5	8	28
41.5-37.5	9	37 إ
45.5-41.5	و	46
49.5-45.5	’ه :	60 ■1

لاحظ أن عدد المشاهدات n=6O وطول الفئة 4=ء

جد الفئة الوسيطية وهي أول فئة يكون تكرارها المتجمع 6 أو أكثر. 2
ومن الجدول تجد ان الفئة الوسيطية هي 41.5-37.5
إذن 37.5=a 28= و عليه 9= ط

--28
إذن، الوسيط يساوي: X4~~-2 9~~ ب 37.5= M

= 37.5+ؤ =38.39
9

نويع• (و؛ اغا

إذا كانت المبيعات بالدينار الأردني يوميا وعلى مدى للساعات كمالبوأ . عزز

يلي:

الحدود الفعلية	التكرار
16.5-10,5	5
22.5-16.5	6
28.5-22.5	9
34.5-28.5	6
40.5-34.5	5
46.5-40.5	3
52.5-46.5	2

جد الوسيط.

أسئلة 1لتقويمالةاقي (5)

1- عرف الوسيط في التوزيع التكراري ذاكراً معنى كل رمز تستعمله.

2- عند ايجاد الوسيط في التوزيع التكراري ذي الفئات هل تستعمل حدود الفئة أم حدودها الفعلية؟

3.5 المنوال

تعريف (8)

المنوال لمجموعة من القيم هو القيمة الأكثر تكراراً، أي أن المنوال يكون تلك القيمة التي يقابلها أكبر تكرار في جوارها.

مثال (19)

إذا كانت الأجور اليومية التي يتقاضاها عدد من العمال بالدينار الأردني كما في الجدول الآتي، جد قيمة المنوال.

جدولرقم (11)

عدد الحمال (f)	الأجر اليومي (m)
7	4
8	5
10	6
18	7
11	8
7	9
3	10

الحل.

الجدول التكراري أعلاه يتضمن عمودين الأول هو مركز الفئة (m) والعمود الثاني التكرار (f). يلاحظ أن أكبر تكرار هو18 وا يوجد قيمة أكبر منه أو مساوية له في عمود التكرار، وأنه يقابل مركز الفئة 7 وعليه فإن قيمة منوال الأجر اليومي هو 7.

أما إذا لم يكن هناك قيم يقابلها تكرار أكبر من غيرها فلا يكون هناك منوال.
وفي بعض الأحيان يكون هناك أكثر من منوال، وتكراركل منها أكبر من تكرار
القيم في جوارها. ومن الأمثلة التي يحدث فيها وجود منوالان في توزيع واحد، لو اخذت مجموعة من معلمي المدارس الاعدادية وأخرى من الأساتذة الجامعيين واعتبرت الرواتب
للمجموعتين هي البيانات التي لديك. الواضح في تلك الحال أنك ستجد منوالين.

مثاق (20)

ولعع٠دا،4

ا كاح امه عدد ض مغمي المدارس الثانوية رلصة احدى تمسة ولأقرب

ا 77,421,421/421,77,77,75,77,85,421,450,82,421_جد_المنوال. ا

الحل:

إذا ما وضعنا هذه الرواتب في جدول تكراري كما يلي:

الراتب	75	77	82	85	421	450
التكرار	1	4	1	1	5	1

نجد أن الراتب 77 يقابله أكبر تكرار نسبة إلى مجموعة الرواتب في جواره، فهو منوال للتوزيع حسب التعريف. وكذلك فإن الراتب 421 له أكبر تكرار في جواره أيضاً،

فهو منوال آخر للتوزيع.

وكما سبق وذكرنا في وصف أشكال التوزيعات التكرارية فإنه يمكن اعتبار المنوال هو القيمة التي يقابلها تكرار أكبر من تكرارات القيم في جوارها. فمثلا : إذا كان التوزيع الآتي يمثل علامات 44 طالباً في أحد الامتحانات (العلامة العظمى 12) فإن العلامة 7 هي منوال، لأن تكرارها اكبر من تكرار القيم التي في جوارها، وكذلك العلامة 10 منوال لنفس السبب. كما هو واضح في الجدول الآتي رقم (12).

جدولرقم (12)

العلية، ية	التكرار f 2
ة	3
7	و
8	4
9	5
10	6
12	3
المجموع	44

أما في التوزيعات التكرارية ذات الفئات فإننا نعطي التعريفات الآتية:

1- تسمى الفئة (أو الفئات) التي يقابلها أكبر تكرار الفئة المنوالية (اكفئات المنوالية).

2- مركز الفئة المنوالية هو (المنوال التقريبي)، وإذا كان هناك عدة فئاتمنوالية فإنه يكون لدينا عدة منوالات تقريبية.

؟كتتتتذينإم

هم (21)

أح

جد المنوال أو المنوالات التقريبية للتوزيع الآتي:

حدود الفئة	التكرار
29-25	2
34-30	5
39-35	10
44-40	4
49-45	6
54-50	9
59-55	3

الحل:

واضح أن الفئة 35-39 فئة منوالية، لأن تكرارها 10 وهو أكبر من تكرارات الفنان المجاورة لها، ولنلك فلن مركزها 2 :37 هو موال بقريبي.

واضح أيضاً أن الفئة 54"50 فئة منوالية، لأن تكرارها 9 أكبر من تكرارات الفئات المجاورة لها،

ولذلك فمركزها :52 منوال تقريبي أيضاً.

تدريب (10)

1- جد المنوال للبيانات: = 8,2,6,5,2,3,2
2- جد المنوال التقريبي للتوزيع المعطى في تدريب (9)

أسئلة التقويم الذاتي (6)

1- عرّف المنوال لمجموعة من البيانات.
2- عرّف المنوال التقريبي في التوزيع التكراري ذي الفئات.
3- هل يمكن ان يكون لمجموعة من البيانات ثلاثة منوالات؟

4,5 خصائص مقاييس النزعة المركزية الثلاثة والعلاقة بينهما

4.5. ٦ مقارنة بين صفات الوسط الحسابي والوسيط والمنوال

إن أكثر مقاييس النزعة المركزية استعمالاً في الاحصاء هو الوسط الحسابي؛ فهو سهل الحساب، سهل التعريف، كما يأخذ جميع قيم المتغير قيد البحث بعين الاعتبار.

وبما أن الوسط الحسابي يعرف بمجموع القيم مقسوما على عددها فإنك تستطيع معرفة مجموع القيم إذا علمت الوسط الحسابي وعدد التكرارات . وهناك خاصية مهمة للوسط الحسابي، أن مجموع انحرافات البيانات عن وسطها الحسابي يساوي صفراً، أي ان: (ه اأ ,X) ؛؛ 0 7- ؛X) ت

وبالنظر إلى المعادلة السابقة والمدرج التكراري للبيانات فإن الوسط الحسابي هو نقطة اتزان المدرج التكراري. وبما أن الوسط الحسابي هو نقطة اتزان التوزيع، فإنك لو أضفت أي عدد من القيم المساوية للوسط الحسابي إلى البيانات فإن هذه الاضافة لا تؤثر عليه، أما إذا أضفت مفردات تختلف قيمها عن قيمة الوسط الحسابي فإن قيمته تتغير.

فمثلاً: إذا كانت علامات 12 طالباً في أحد الامتحانات كما يلي:

30,28,26,19,30,-17,29,20,30,21,23,27

فإن الوسط الحسابي للعلامات يساوي 25، فإذا أضفت العلامات 25، 25,25، فإن الوسط الحسابي للطلبة جميعهم (أصبحوا 15 طالباً) يبقى 25، ولكن إذا أضفت طالباً واحداً إلى المجموعة الأصلية من الطلبة (الإثنى عشر طالبا) وكانت علامته20 مثلاًفإن الوسط الحسابي للمجموعه الكلية (13 طالباً) يتغير ويصبح في هذه الحالة:

27-1-23_+_2_______30+20+20_+17+30+19(20

13

=24-3=24.62

13

أما أهم عيوب الوسط الحسابي فهي تأثره الشديد بالقيم المتطرفة، فلو قلت أن مجموعة من الأفراد تبرعوا لعمل خيري بمبالغ بسيطة (أقل من 10 دنانير من كل منهم) ثم تبرع اثنان بمبلغين كبيرين (00 1 من كل منهما) وحسبت الوسط الحسابي للتبرعات لوجدت أن معدلها كبير، وهذا لا يمثل ما دفعه معظم الأفراد.

أما الوسيط فهو سهل التعريف، سهل الحساب، لا يتأثر بالقيم الشاذة ولا يعتمد على جميع القيم دائماً . إن تغير قيمة من القيم قد يؤثر فيه وقد لا يؤثر. وهو يستتمل خاصة في البيانات التي يعرف ترتيبها ولا تعرف قيمها، وكذلك في البيانات الناقصة، أي إذا كان في البيانات مفردات سقطت قيمها. وأهم الأمثلة على ذلك، تلك البيانات التي تقيس الزمن الذي ينجز فيه الطلاب عملا معيناً . فقد ينتهي الوقت المقرر دون أن يتمكن عدد من الطلبة من انجاز ذلك العمل، ولذلك لا تكون هناك قيم لزمن انجاز العمل عند هؤلاء الطلابب، لكنك تترف أنهم يستغرقون وقتا أطول من باقي الطلبة أي أنك تترف ترتيبهم، وبالتالي فأنت تستطيع ايجاد الوسيط لمثل هذه التجربة وا تستطيع ايجاد الوسط الحسابي.

وإذا اخذت عينة من مجتمع ما، ثم عينة من نفس المجتمع فإنك تجد تقارباً بين الوسطين الحسابيين لهاتين العينتين أكثر من التقارب بين وسيطيهما، أي أن الوسط الحسابي بصورة عامة أكثر استقراراً من الوسيط. فالوسط الحسابي لا يتأثر كثيراً من عينة لآخرى، أما الوسيط فيتأثر . فإذا أردت تقدير معدل المجتمع باستعمال أحد مقاييس النزعة المركزية فإن الوسط الحسابي أكثر استقراراًوقبولاللاعتماد عليه.

وهكذا فمن الممكن أن يكون بعد وسط العينة الحسابي عن الوسط الحسابي للمجتمع أصغر من بعد وسيط العينة عن وسيط المجتمع.

أما المنوال فهو أقل الثلاثة استعمالاً، وهو في حال البيانات القليلة العدد عديم الفائدة تقريبا، هذا إن وجد أصلاً.

أما في البيانات الكبيرة العدد معنى معقول، وميزته أنه لا يحتاج إلى حسابات إلا في حالة التوزيعات التكرارية ذات الفئات، كما أنه يتار كثيراً بطريقة اختيار الفئات، فمن الممكن أن توجد فئتان منواليتان غير متقاربتين، بصدفة اختيار الفئات فقط.

ومن جهة ثانية فإن المنوال لا يتأثر كثيراً بالقيم الشاذة، بل أنه ا يتأثر في بعض الأحيان، حتى لو تغيرت قيم بعض مفردات البيانات. هذا ويفضل استعمال الوسط الحسابيفي الحالات لتالية:

ا- إذا كان التوزيع متماثلاً تقريباً.

ب- إذا كان اهتمامك منصباً على القيمة العددية لجميع البيانات أي المجموع الكلي، بدلاً من الاهتمام بقيمة نموذجية.

أما الوسيط فيفضل استعماله:

أ- إذا اردت ليجاد قيمة مملة، أي قيمة نمونجية بدلاً من اهتمامك بالمجموع الكلي، وإذا كان التوزيع ملتويا.

ب- إذا فقدت لديك بعض القيم (وعرف ترتيبها) حيث لا يمكن حساب الوسط الحسابي مباشرة.

أما المنوال فلا يفضل استعماله، إلا إذا كان التوزيع متعدد المنوالات.

3شؤهص (11)

لديك اسات

40,42,48,50,5457,62,6776,83,91,95,100

أ- إذا تغيرت العلامة 67 وأصبحت 89 فهل يتغير الوسيط؟ وما قيمته؟

ب- اذا عدلت العلامة 40 واصبحت 61 فهل يتغير الوسيط؟ وما قيمته؟

ج- هل يتغير الوسط الحسابي في كل من الحالتين السابقتين؟ كم يصبح في كل حالة؟

2.4.5 العلاقة بين المتوسطات الثلاثة

قد يتبادر إلى ذهنك، عزيزي الدارس، السؤال الآتي: هل يمكن لهذه المقاييس الثلاثة أن تكون قريبة من بعضها؛ بل وهل يمكن أن تكون متساوية.وللإجابة عن هذا السؤال تصور أن بيانات ما قد وضعت في الجدول التالي:

الأجر اليومي (m)	عدد العمال (f)
4	
ى	10
6	5
7	10
8	

وطلب منا حساب كل من الوسط الحسابي والوسيط والمنوال لأجرة العامل اليومية.

حيث يتضح أن مجموع التكرار (عدد العمال) -49 . وأن كلا من الوسط الحسابي والوسيط والمنوال مساوي 6، وهذا يعني أن المتوسطات قد تساوت في هذه الحالة نقول أن التوزيع التكراري متماثل Symmetric.

وفي بعض التوزيعات التكرارية الأخرى والقريبة من التماثل، وجد أن العلاقة التالية صحيحة

Mean - Mode = 3*(Mean - Meadian)

إن أهمية العلاقة أعلاه تتجلى عند حساب أي من المتوسطين السابقين يمكن حساب الثالث منها، وفي هذا أهمية كبرى في الجداول المفتوحة إذ لا يمكن حساب مركز الفئة للفئة المفتوحة وبالتالي لا يمكن حساب الوسط الحسابي بينما يمكن ايجاد الوسيط والمنوال من الجداول المفتوحة. وبذلك تكون العلاقة أعلاه واحدة من طرق حساب الوسط الحسابي للبيانات الموضوعة في جداول تكرارية مفتوحة، إذ يمكن إعادة صياغتها كالآتي:

3Median-Mode

Mean:- - ——■

6,مقاييس التشتت

عزيزي الدارس، لقد بسطنا في هذه الوحدة طرق عرض البيانات الإحصائية، وبناء التوزيعات التكرارية وعرضها بيالياً، كما أوردنا وصف أشكالها وبعض خواصها. ثم درست معنا مقاييس النزعة المركزية التي تصف هذه التوزيعات عدديا وهي تعتبر مقاييس موقع، أي أن قيمهاتكون على المحور الأفقي الذي يمثل قيم البيانات.

إن هذه المقاييس غير كافية لتحديد صفات التوزيعات التكرارية والبيانات الإحصائية. فقد يكون لدينا ظاهرتان متساويتان في مقاييس الموقع كالوسط الحسابي والوسيط، إلا أن الظاهرتين مختلفتين. فمثلاً، إذا كانت درجات الحرارة في بلد ما، خلال الليل والنهار، هي؟ 346,2237 923936924 فيكون معدل درجات الحرارة29.3، وإذا كانت درجات الحرارة فيبلد آخر هي 17هم 40؟ ٦9؟42,°23,°35 فإن معدل درجات الحرارة فيهام29.3 .وهذا يعني أن معدل درجات الحرارة في البلدين متساو.

و لو نظرت إلى مفردات البيانات في كل من البلدين لوجدت إختلافا بينهما، وهذا يعني أن الوسط الحسابي لا يكفي لوصف البيانات أو للحكم على تشابهها.

وكمثال آخر، صف دراسي فيه 40 طالبا معظم علاماتهم تقع ما بين 40,80 وعند حساب المعدل وجد أنه يساوي 65، وصف آخر معظم العلامات فيه ما بين 67,64 ومعدله 65.

أعتقد أنه من الواضح جداً لك وجود فروق بين هذين الصفين بالرغم من تساوي الوسطين الحسابيين فيهما. إذن لا بد من استعمال مقاييس آخرى تبين مدى اختلاف البيانات فيما بينها، ومدى التفاوت والتغير بين مفرداتها. فهل هي متقاربة من بعضها البعض أم متباعدة؟. إن مقاييس التشتت تجيب عن هذه التساؤلات.

ومن هذه المقاييس:

1.6 المدى والمدى الربعي قحوبيفد (9) .

يعرف المدى في البيانات بأنه (الفرق بين أعلى قيمة وأصغر قيمة في البيانات). ;.,ة..

فإذا كان المدى صغيرا كانت البيانات محصورة في مسافة قصيرة، وإذا كان المدى كبيرا "١<

كانت_البيانات_تقع_ضمن مسافة_طويلة.

تتجذييكه (0 أ؛

يعرف المدى في البيانات المجمعة (التوزيع التكراري)، بأنه الفرق بين الحد : الأعلى_للفئة العليا و الحد الأدنى_للفئة_الدنيا.

ومن هذا التعريف يظهر أن المدى لا يعتمد على جميع البيانات بل على أكبر قيمة و أصغر قيمة فقط. وهذا يقلل من أهميته، إذ قد يحدث أن تكون القيمتان المتطرفتان (أكبر قيمة وأصغر قيمة) قيمتين شاذتين، عندئذ يكون المدى كبيراً بينما مفردات البيانات ليقت متباعدة عن بعضها البعض، فمثلاً إذا كانت علامات الصف الثاني الابتدائي في مدرسة ما هي: 71,67,65,70,100/30,69,70,65,69,65؛ فإن المدى هو 70-30- 100 .بينما معظم العلامات واقعة بين 71,65 أي أنها متقاربة من بتضها البعض.وأنت تلاحظ من هذا المثال أن معظم علامات الصف كانت متقاربة إلى حد ما إلا أن المدى كان كبيراً. على ماذا تلك هذه الظاهرة؟ إنها توضح أن المدى عنصر لا يعتمد عليه كثيرا كمقياس لتشتت البيانات,

فكيف يمكنك التخلص من هذا النقص؟ إن إحدى الطرق للتخلص من ذلك هي حذف العلامات المتطرفة باعتبار أنها شاذة. احذف العلامتين 100,30، ماذا يصبح المدى للبيانات الجديدة؟ الجواب هو: 71 - 65 = 6

ولكن كم علامة يجب أن تحذف؟

يمكنك أن تحذف نسبة من العلامات العلياوالعلامات الدنيا.

إذا حذفت أعلى 25% من البيانات وأدنى 25% منها ثم حسبت المدى للبيانات الجديدة فإنك تحصل على المدى الربيعي.

الآن، كيف تحسب القيمة التي يزيد عنها 25% من البيانات؟ وكيف تحسب القيمة التي يقل عنها 25% من البيانات؟ نقسم البيانات المرتبة تصاعدياً إلى أربعة أقسام متساوية. إن عدد هذه النقاط 3 وهي من اليسار إلى اليمين، الربيع الأول (الأدنى) Qi، الربيع الثاني 02، الربيع الثالث (الأعلى) ٠3.

الربيعات في البيانات المرتبة تصاعديا هي:

الربيع الأول ,-Q هو القيمة التي يسبقها ربع البيانات و يليها ثلاثة أرباع البيانات. الربيع الثاني Q? وهو الوسيط، أي القيمة التي يسبقها نصف البيانات ويليها نصفها. الربيع الثالث ٠3 وهو القيمة التي يسبقها ثلاثة أرباع البيانات ويليها ربعها.

ويتم حساب الربيع الأول Qi والربيع الثالث ق0 بنفس الطريقة التي تم فيها حساب الوسيط (الربيع الثاني ٠2) وذلك باستعمال الرموز والتعاريف التالية:

ففة الربيع الأول هي أول فة فزيد تكرلرها المتجمع عن —— من البياتاك أو يساويه حيث n هو مجموع التكرارات.

طولفئة الربيع الأول-ع
الحد الأدنى الفعلي لفئة الربيع الاول = ,a
تكرار فئة الربيع الاول = fl
التكرار المتجمع للفئة التي تسبق فئة الربيع الاول = ٢١٦
وبعبارة أخرى أ٢١ هو مجموع التكرارات قبل a٦

قحوتيف (12)

n
n,
الربيع الاول هو xc■■■ ■■ب, Q, =a,
fl ى
وبنف الطريقة: فئة الربيع القالث هي أول فنة فزيد بكرارها المتجمع عن —— . أو يساويه .

ا.7م

3n
الربيع الثالث هو: xcˢ-■■4"3۴ ٢ Q₃=a₃
حيث الرموز كمثيلاتها السابقة، أي:
الحد الأدنى الفعلي لفئة الربيع الثالث : 3ه
تكرار فئة الربيع الثالث = 3۴
التكرار المتجمع للفئة التي تسبق فئة الربيع الثالث = 3ل٢

الجدول الآتي يمثل التوزيع التكراري للزمن (لأقرب دقيقة) الذي استغرقه 50 !
ا طالباً للجاجابة عن لسفلة امتحان شيري.

جدول رقم (13)

الزمن	التكرار
32-36	4
37-41	7
42-46	10
47-51	7
52-56	18
57-61	4

جد: الوسيط والربيع الأول والربيع الثالث.

الحل:

ـــ الفئة الوسيطية هي أول فئة يزيد بكرارها المتجمع عن ـــ أي 25 وبلتلي في الفئة

47 -51 وبتطبيق قانون الوسيط تجد

$$M=Q_2=46.5+\frac{25-21}{7}(5)-49.36$$

- لحساب الربيع الأول، لاحظ أن فئة الربيع الأول هي أول فئة يزيد تكرارها المتجمع عن50 أي 12.5
الحل:

وبالتالي فهي الفئة 42-46 وحدها الادنى الفعلي هو 41.5

'4 -

مجموع التكرارات قبل الحد الأدنى الفعلي هو nil

طول الفئة 5=c

وبتطبيقتعريف لربيع ل أول

42.25=5حرتلتيتعب41.5=اب 10 [1]

- ولحساب الربيع الثاكث، لاحظ أن فئة الربيع الثالث هي أول فئة يزيد تكرارها المتجمع عن 50ج =

3٠7.5 وبلتلى فهى الفئة 52-56 (,حدها الأمنى الفعلى 51.5).

مجموع التكرارات قبل لحد الأدنى الفتلي n_3=28

طول الفئة 5ح0

تكرار فئة الربيع الثالث هو 18= f_3

بتطبيق تعريف الربيع الثالث:

54.14=_(5)_٦٢٠ق" + 5.ل5:ـة

د((13 Q

المدى الربيعي هو الفرق بين الربيع الثالث_والربيع الأول، أي_Q_3-QI

مل 623) ليبج

إذا كانت أكبر قيمة في البيانات 100 وأقل قيمة 20، وكان الربيع الأول إ 45=_Qi_و_الربيع الثالث_37=، فما_هو_المدى_وما_هو_المدى_الربعي.

المدى = أكبر قيمة - أصغر قيمة

أي 80 = 20 - 100

المدى الربعي يساوي أي

87 -45 = 42

تدريب (12)

إذا كان الربيع الأول لمجموعة من العلامات 37 والربيع الثالث 92 وأعلى علامة 99 وأقل علامة 22. جد المدى والمدى الربعي.

أسئلة التقويم الذاتي (7)

1- عرف الربيع الأول.

2- عرف المدى الربيعى.

2.6 التباين والانحراف المعياري

إن أحد مقاييس التشتت التي تخطر على البال هو مجموع لنحرافات البيانات عن وسطها الحسابي، أي (X; — x) 1 > ، لكن هذا المجموع يساوي صفراً دائماً، ولذلك لا بد من حذف الاشارة السالبة لنحصل على مقياس ذي معنى.

وإحدى الطرق التي تزيل بها .الاشارة السالبة هي بتربيع الانحرافات، ونحن نستعمل مربعات الانحرافات هذه في حساب التباين.

تعريف (14)

للبيانات $x_1, x_2, \ldots\ldots, x_n$

يعرف التباين على أنه
$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$$

حيث $\overline{X}$ الوسط الحسابي للبيانات.

لاحظ أنك تقسم على (n - 1) على أساس أن البيانات التي لديك هي مجرد عينة أخذت لتدرسها ومن ثم تعمم النتائج على المجتمع الذي اخذتها منه. وقد تمت القسمة على

(n -1) لكي يكون للتباين 2ة بعض الصفات الاحصائية المرغوب فييها، والتي سوف لن نتطرق لهافي هذا المقرر.

.!عك

احسب التباين لبيانات 7,5,8,4,9,7,2

الوسط لصابى 6عيمياً:إ

انحرافات البيانات عن الوسط الحسابي هي:

(2-6)،(7-6)،(9-6)،(4-6)،(8-6)،(5-6)،(7-6)

ومربعات هذه الانحرافات هي:

1,1,4,4,9,1,16

ولوبلنهو 6ثكثةثث=ق

=—6
6

تدريب (13)

جد التباين للبيانات التالية:

6,15,12,9,7,11,10

أسئلة التقويم الذاتي (8)

كيف تزيل الاشارة السالبة من الانحرافات؟1

ما مجموع انحرافات القيم عن وسطها الحسابي؟2

لقد عرفت التباين لمجموعة من البيانات، فما هو التباين في حالة التوزيعات التكرارية؟ هذا ما سيعطيك إياه التعريف الآتي:

صبه,15) R
٦خججججججججججتتحكأ

إذا كانت مراكز فئات توزيع تكراري هي: ,Xj,x2 ,.....,X

وكانت التكرارات لمقابلة ها f),fj

فالتباين S^2 يكون $S^2=\dfrac{\sum_{i=1}^{h}(x_i-\overline{x})^2 f_i}{n-1}$

حيث $\overline{x}=$ الوسط الحسابي للتوزيع التكراري

$n=$ مجموع التكرارات.

تعريف (16)

الانحراف المعياري هو الجذر التربيعي للتباين أي:

للبيانات الأولية: $S=\sqrt{s^2}=\sqrt{\dfrac{\sum(x_i-\overline{x})^2}{n-1}}$

أو $S=\sqrt{\dfrac{\sum(x_i-\overline{x})^2 f_i}{n-1}}$

للبيانات المجمعة (التوزيع التكراري ذي الفئات).

إن طريقة حساب التباين طويلة وتحتاج إلى حساب أرقام كثيرة، وبخاصة إذا احتوى الوسط الحسابي على كسور، لأن ذلك يستوجب أن تحتوي الانحرافات على كسور، وبالتالي فأنت تحتاج إلى تربيع تلك الكسور. هناك طريقة مختصرة لحساب التباين صالحة للاستعمال بالآلة الحاسبة مباشرة وتعطى بالنظرية: قظوييبة (1)

للبيانات الأولية

وللبيانات المجمعة في توزيع تكراري ذي فئات

[^{2}تل1بتتل]ل:8^2

حيث X - الوسط الحسابي n = مجموع التكرارات

احسب التباين للتوزيع التكراري في الجدول رقم (14). جدول رقم (14)	

التكرار	حدود الفئة
3	24-20
7	29-25
12	34-30
8	39-35
10	44-40

الحل:

جد مراكز ألفئات ثم اضرب كل مركز بالتكرار المقابل له لكي تحسب الوسط الحسابي، وتجد انحرافات مراكز الفئات عن الوسط الحسابي. بعد ذلك ربع هذه الانحرافات واضرب مربعات الاتحرافات في التكرارات المقابلة لها كما يظهر في الجدول رقم (15).

جدول رقم (15)

x)2fj-ٍx)	2(x-ٍx)	x-ٍx	Xi fj	التكرار fi	مركز الفئة Xi
432.39	141.13	-11.88	66	3	22
331.31	47.33	-6.88	189	7	27
42.36	3.53	-1 88	384	12	32
7..ئ4	9 ؤو	3.12	1%		37
	6 ؤ.9ؤ	8.12	420		42
1534.2			1355	40	المجموع

لاطأن 33.88:أ=ت
40

بثلك يكون القبلن 39.34:آ8^2

وكما هو واضح فإن وحدة التباين هي مربع وحدة البيانات، فإذا ما أردت وحدة البيانات خذ الجذر التربيعي واحصل على الانحراف المتياري كمقياس للتشتت.

طبر_(26) R

احسب التباين والانحراف المعياري للتوزيع التكراري التالي:

الحل:

رتب خطوات الحل كما في الجدول رقم (16)

جدول (16)

x_i	f_i	$x_i f_i$	x_i^2	$x_i^2 f_i$
3	2	6	9	18
7	5	35	49	245
11	9	99	121	1089
15	7	105	225	1575
19	2	38	361	722
المجموع	25	283		3649

وبالتعويض بقانون التباين تجد: $S^2 = \frac{1}{24}\left[3649 - 25x\left(\frac{283}{25}\right)^2\right]$

]ل =3649-3203.56]

24=!؛ = 18.56

أما الانحراف المعياري فهو s=٧18.56

-4.31

تهوييب (14)

يمثل الجدول أدناه التوزيع التكراري للأجور الاسبوعية ل 550 عاملاً في مصنع. جد التباين والانحراف المعياري.

الأجرو الأسبوعية	عدد العمال
20-16	22
25-21	17
30-26	101
35-31	180
40-36	120
45-41	50
50-46	35
55-51	10
60-56	15

أسئلة التقويم الذاتي (9)

ما هو الانحراف المعياري ؟

7. أثر التحويلات الخطية على كل من مقاييس النزعة المركزية

أي أن التحويل الخطي يؤثر على الوسط الحسابي بنفس الطريقة التي يؤثر فيها على البيانات

وتوضح لنا النظرية التالية أن التحويل الخطي يؤثر على الوسيط والمنوال بنفس الطريقة التي يؤثر بها على الوسط الحسابي البيانات أيضاً

1.7 أثر التحويلات الخطية على مقاييس ألنزعة المركزية

نظرية (2): التحويل الخطي هو كل اقتران من النوع:

$y=f(x)=ax+b$

إذا حولت مجموعة من البيانات (الأولية أو التي في توزيع تكراري) حسب التحويل الخطي (المعادلة) $y=ax+b$ حيث a, b ثوابت. وتبرز الحاجة لدراسة أثر التحويلات الخطية على مقاييس النزعة المركزية والتشتت في حالة كون البيانات معطاة حسب نظام معين من المقاييس ثم حولت إلى مقياس آخر، وحساب مقاييس النزعة المركزية، والتشتت للبيانات الجديدة مباشرة بدلامن الرجوع إلى البيانات القديمة، وتحويلها، ثم حساب مقاييس النزعة المركزية والتشتت بعد التحويل. فمثلا، ربماتعطى درجات الحرارة بالمقياس المئوي وتكون مقاييس النزعة المركزية والتباين قد حسبت فإنك ستتمكن من حساب مقاييس النزعة المركزية والتشتت بالدرجات الفيرنهايتية باستعمال النظريات التي تعطيك أثر التحويلات الخطية على هذه المقاييس.

مثل (27)

إذا كان لدينا العلامات 4,10,8,6,7

فكم تصبح هذه العلامات تحت تأثير التحويل الخطي $y = 2x - 3$ الحل: العلامة 7 يقابلها

العلامة $y = 2x7-3 = 11$

سيعطيك جدول رقم (17) جميع قيم y المقابلة للعلامات.

جدول رقم (17)

x	7	6	8	10	4
y	11	9	13	17	5

والآن ما هو الوسط الحسابي للعلامات الأصلية (*)، وما هو الوسط الحسابي للعلامات بعد التحويل

منتريف لوسطت بد: 7 خ ن ———— x عنم

أما الوسط الحسبي بعد التحويل فهد 1 ا: ~~17~~ 3 9 ~~11~~ Y ٠

والآن لوعوضت قيمة في التحويل المعطى لوجدت 2x7-3 ه 11

حيث b,a ثابتان، فأن الوسط الحسابي والوسيط والمنوال تتأثر جميعها بنفس صيغة التحويل؛ أي أنه إذا كان الوسط الحسابي X والوسيط M_x والمنوال Dy فإن المقاييس الجديدة تصبح:

1- y=ax + b   2- My=aM*+b   3- $Dy=aD_x tb$

مثات(28)

إذا كان الوسط الحسابي لعلامات صف مكون من 60 طالبا هو 70، والوسيط 68 والمنوال 65 وحولت جميع العلامات وفق المعادلة:

y=O.§x+-2O

حيث x العلامة قبل التحويل، و y العلامة بعد التحويل.

اعتماداً على ذلك أجب عما يلي:

أ- ما قيمة الوسط الحسابي بعد التحويل؟

ب- ما قيمة الوسيط بعد التحويل؟

ج- ما قيمة المنوال بعد التحويل؟

د- ما قيمة مجموع علامات الطلاب بعد التحويل؟ الحل:

أ- y=O.8x+2O

اذن الوسط الحسابي بعد التحويل هو: y=0.8x70 + 20:76

ب- والوسيط بعد التحويل يساوي: 0.8 X 68 + 20 = 74.4

ج- والمنوال بعد التحويل يساوي: 0.8 X 65 +20 = 72

د- مجموع العلامات بعد التحويل : عدد الطلبة x الوسط الحسابي بعد التحويل:

60X76 = 4560

يمكنك كتابة نظرية (1) بالرموز كما يلي:

إذا رمزنا للبيانات الأصلية بالرمز X ثم عدلت هذه البيانات حسب المعادلة y = ax+b حيث b.a عددان ثابتان، فإنك تحصل على أ، ب،ج، د ي منطوق النظرية (1). لاحظ أنه لم يكن هناك أي أثر على التباين أو الانحراف المعياري أو المدى الربعي من جراء جمعك العدد b لكل مفردة من البيانات،

إذا كان تباين مجموعة من البيانات 16 وكان المدى.يساوي 12، وإذا تم تحويل هذه البيانات حسب المعادلة y=3x+l ؛ فما هو تباين البيانات الجديدة؟ فما انحرافها المعياري،_وما_المدى لها؟ الحل:

إن التباين والانحراف المعياري والمدى لا تتأثر من جراء جمع العدد 1 إلى البيانات الأصلية، ولكنها تتأثر بالضرب بالعدد 3 حيث التحويل حسب 3x+l٢٠

التباين للبيانات الجديدة = تباين البياتات الأصلية، أي 16x9 = 144

والانحراف المعياري للبيانات الجديدة هو 2ا: 144ر١ أو هو 12 =|3| 4x

والمدى للبيانات الجديدة هو x,3| المدى للبيانات الأصلية ويساوي 12x3 ه 36

هع8ل (31) g

إذا كان المدى الربعي لمجموعة من البيانات 60، وكان تباين البيانات 16، وقمت بضرب كل مفردة من البيانات في العدد 3؛ فما المدى الربعي والتباين والانحراف المعياري للبيانات الجديدة؟ الحل:

المدى الربعي للبيانات الجديدة , المدى الربعي للبيانات الاصلية مضروباً في |3|

= 3x60 = 180

تباين البيانات الجديدة = 16 X ت3

-144

الانحراف المعياري للبيانات الجديدة = 12 = 144ر٦

مثال (32)

عدلت علامات مجموعة من الطلبة (X) حسب المعادلة $y=0.7x + 30$

فإذا كان الانحراف المعياري للعلامات 8، فما هو الانحراف المعياري بعد التعديل ؟ الحل:

الانحراف المعياري بعد التعديل =

الانحراف المتياري للبيانات الاصلية مضروباً في |0.7| ويساوي

$8x|0.7|=5.6$

صد (16) • ددا

1- أضاف أحد المدرسين 7 علامات لكل طالب، فإذا كان الانحراف المقياري لعلامات الطلبة قبل الزيادة 9، فما الاتحراف المعياري بعد الزيادة ؟

2- إذا ضربت كل علامة في 0.8 وجمعت لها 20 وكان تباين العلامات 25 ومداها 90، فما مدى العلامات وما هو تباينها بعد التعديل ؟

أسئةآلتقيتيم اتخاقي (10)

اكمل العبارة: إذا ضربت كل مفردة من البيانات في عدد ثابت a وجمعت لها عدداً ثابتاً b فإن تباين البيانات الجديدة يساوي.... والانحراف المعياري للبيانات الجديدة يساوي.....

8.مقاييسالتشتتا لنسبية

عزيزي الدارس، إن جميع مقاييس التشتت التي تم التطرق إليها (المدى والانحراف الربيعي والتباين وجذره) هي مقاييس تشتت يطلق عليها بمقاييس التشتت المطلقة. حيث تستخرج جميعها لقياس درجة تباعد القيم فيما بينها أو بينها وبين الوسيط في حالة الانحراف الربيعي، وبينها وبين الوسط الحسابي في حالة الانحراف المعياري وجميع هذه النتائج تقاس بوحدات القياس الأصلية.

ولو أردنا مقارنة تشتت مجموعتين أو أكثر تختلفان في وحدات القياس وكانت الظاهرة الأولى تمثل الطول المقاس بالسنتمتر، فإن جميع المفردات ستكون مقاسة بالسنتمتر وبالتالي فإن جميع مقاييس التشتت المطلقة المستخرجة ستكون مقاسة بالسنتمتر، والظاهرة الثانية تمثل الدخل الشيري المقاس بالدينار فإن جميع المفردات ستكون مقاسة بالدينار وبالتالي فإن جميع مقاييس التشتت المطلقة المستخرجة ستكون مقاسة بالدينار. وهذ لايمكن المقارنة بين مقاييس التشتت ذات النوع الواحد كأن نقول بأن الدينار أصغر من السنتمتر. لذا يتطلب أولاً أن نتخلص من وحدات القيام، ويتم ذلك باستخدام مقايس التشتت النسبية.

1.8 معامل الاختلاف

يتم التخلص من وحدات القياس عند حساب مقاييس التشتت النسبية بقسمة مقاييس التشتت المطلقة على المتوسطات المحسوب بعد القيم عنها، وبهذه الحالة يكون لدينا: ه معامل الإختلاف المعياري

Coefficient of standardized variation = ؤ*%100

- معامل الإختلاف الربيعي

0و+0إCoefficient of quartile variation ـ‍تدون‍ـ‍ـ+ ب%100

أن المقياسين أعلاه هما نسبة مئوية خالية من وحدات القياس، وأن أكثرهما شيوعاً في الاستخدام هو معامل الإختلاف المعياري.

وبالإمكان استخدام معامل الاختلاف للمقارنة بين ظاهرتين مقاستين بنفس وحدات القياس، غير أن هنالك فارقاً كبيراً بين قيمتي متوسطيهما.

)33وثعل (

افترض أن الوسط الحسابي لأوزان عينة من طلبة كلية العلوم الادارية 70 كغم وبانحراف معياري 5 كغم، بينما متوسط أطوالهم 175 سم بانحراف معياري 10 سم. أي من المتخيرين أكثر نشتتاً لهؤلاء الطلب.

الحل:

بانسبة إلى متغير الوزن وليكن X

Coefficient of standardized variation (x) = هذ%100=7٠14

وبالنسبة إلى متغير الطول وليكن Y

Coefficient of standardized variation (Y) = ييل%100=5.71
175

يلاحظ من خلال النتيجتين أن متغير الوزن للطلبة هو أكثر تشتتا من متغير الطول، بينما لو تمت المقارنة على لساس الاتحراف المعياري وبأهمال وحدات القياس لكان الجواب متغير الطول.

عزيزي الدارس، يمكن لنا حساب قيمة معامل الاختلاف الربيعي بمعرفة قيمتي الربيعين الأول (Q؛) والثالث (Q3) حيث:

معمل الإهتلف لرسيعي Coefficient of quartile variation = ا—|م%100 فمثلاً لو كانت قيمة الربيع الأول 25 ديناراً وقيمة الربيع الثالث 100 ديناراً لبيانات متغير ما، فإن معامل الاختلاف الربيعي لهذه البيانات سيكون:

Coefficient of quartile variation = ~:ب%100=60%

2.8 العلامة (الدرجة) المعيارية

هذا مقياس آخر من مقاييس النشتت النسبية، بل ويعتبر من أهمها واستخداماته عديدة، يتلخص هذا المقياس بحساب الوسط الحسابي للقيم والاتحراف المعياري لها أولا ثم طرح قيمة المتغير من وسطه الحسابي وقسمه الناتج على الانحراف المعياري؛ وبهذا تكون العلامة المعيارية وسنرمز لها ب Z. وبالتعبير عن ذلك رياضياً.

يلل-٤7
تاً

فاذا ما أردنا مقارنة قيمتين تنتميان إلى متغيرين مختلفين أو مجموعتين مختلفتين ولمعرفة الترتيب النسبي لكل منها في مجموعته فإن العلامة المعيارية هي أفضل مقياس لذلك.

•ئني(34) ج

حصل أحد الطلاب على علامة 85 في مساق الرياضيات حيث كان الوبسط الحسابي لغلامة المساق 76 وبانحراف معياري بلغ 9 علامات، فيما حصل الطالب نفسه على علامة 75 في مساق الإحصاء حيث كان الوسط الحسابي 60 وبانحرلف. معياري 12، في أي من الامتحانين كان مستوي هذا الطالب أفضل؟   85-76 .   z = —ا:———   9   75-60. ع5   12

ويتضح من النتيجتين أن مستوى الطالب في امتحان الإحصاء أفضل منه في امتحان الرياضيات، على الرغم من أن علامته أعلى في الرياضيات وذلك بسبب أن علامته في الإحصاء تبعد بوحدة وربع معيارية عن متوسط علامة الإحصاء، بينما تبعد علامته في الرياضيات بوحدة معيارية عن متوسط علامة الرياضيات.

■ئ؛ئيك ز17ر ن ٢.٦٠٠.

لديك البيانات التالية 5، 8، 10، 7، 3، 12

احسب كلا من:

1- معامل الاختلاف المعياري.

2- العلامة المعيارية للقيمتين 12،5

أسسئلة لتقوبيم 1 الفاتـي

مقاييس التشتت النسبية أفضل

(11) إذا كانت لديك بيانات فيها قيم شاذة، فأي من لمعرفة تشتت البيانات.

3.8 العشيريات والمئينات

عند استخراجنا للوسيط قمنا بترتيب القيم تصاعدياً ثم قسمنا القيم إلى أربعة أقسام، فكان لديناالربيع الأول وقيمته تعني أن ربع القيم المرتبة تصاعدياتسبقه، والربيع الثاني الذي يمثل الوسيط ويشير إلى أن نصف هذه القيم المرتبة تسبقه، بينما الربيع الثالث يدل على ثلاثة أرباع القيم المرتبة تسبقه، وهي Di ;D2;D3 ;...;DIO وكذلك الحال للمئينيات P حيث تقسم قيم البيانات بعد ترتيبها إلى مئة جزء وهي ?10;...; 111 ;P3، وفي هذه الحالة تكون

Dj : ?10 ; ٥2 : ?20 ; •••; D؟ : P90

وكذلك قيمة المئين الخامس والعشرون هي قيمة الربيع الأول أي أن $Q1=P_{2S}$وأن قيمة المئين الخمسون هي قيمة الوسيط أو الربيع الثاني أي أن ?50=2ن وبهذا تكون قيمة المئين الخامس والسبعون هي قيمة الربيع الثالث أي Q_3 :P75 •

وطريقة حساب المئين j والذي نرمز له ب Pj هو بنفس فكرة حساب؛ الواسيط والربيعين الأول والثالث سوى أتنا نقسم مجموع البيانات على 100 ونضرب الناتج في **j**فنحصل على ترتيب المئين ن فتكون فئة المئين j هي أول فئة التي يزيد تكرارها أو يساوي ترتيبه ثم نستخرج قيمته وفق العلاقة التالية:

$$P_j = a + \frac{\frac{j}{100} * n - n_1}{f_j} * C$$

انسب التين اديين

للجدول التكراري الآتي: جدول رقم (18)

حدود الفئات	التكرار
19.5-9.5	25
29.5-19.5	30
39.5-29.5	42
49.5-39.5	68
59.5-49.5	85
69.5-59.5	35
٦9.5-69.5	15
المجموع	300

الحل:

حدود الفئات	التكرار	التكرار المتجمع
19.5-9.5	25	25
29.5-19.5	30	55
39.5-29.5	42	97
49.5-39.5	68	165
59.5-49.5	85	250
69.5-59.5	35	285
79.5-69.5	15	300
المجموع	300	

ولحساب المئين السبعين، لاحظ أن فئة المئين المطلوب هي أول فئة يزيد تكرارها المتجمع عن

210=300ئق ، وبالتالي فهي الفئة (49.5 - 59.5)، إذن المئين السبعين هو

?70=49.5+ 2 ل١٥ (10)^54.794
§5

قهويب (18)

احسب المئين الخامس والسبعون والمئين الخامس والعشرون لبيانات المثال (22).

أسقلة التقوتيم1لذاتي(12)

ما الفرق بين الربيعات والعشيريات والمئينات.

ق الخلاصة

تعرفت على قسم كبير من الإحصاء الوصفي الذي اشتمكت عليه هذه الوحدة: طرق عرض البيانات ووصفها، وتعريف عدد من مقاييس الموقع، ومقاييس التشتت، وسألخص لك أهم الموضوعات التي وزدت في مادة الوحدة، أرجو أن تجدها مفيدة لك.

(1) طرق عرض البيانات الأولية.

1.1 الجداول.

2.1 طريقة المستطيلات أو الأعمدة.

3.1 طريقة الخط المنكسر.

4.1 طريقة المنحنى.

5.1 طريقة الدائرة.

6.1 الطريقة التصويرية.

(2) التوزيعات التكرارية

أ- التوزيع التكراري جدول يعطى فيه حدود الفئات والتكرارات المقابلة لها، وفي كثيرمن الأحيان تعطى الحدود الفعلية للفئات أومراكز الفئات بدلاًمن حدودها.

ب- التكرار النسبي لأي فئة هو تكرار تلك الفئة مقسوماً على مجموع التكرارات.

ج- التكرار المئوي هو التكرار النسبي موضوعا على شكل نسبة مئفية.

(3) عرض التوزيعات التكرارية بيانيا

أ- المدرج التكراري:

ونحصل عليه برسم مستطيل حدود قاعدتة هي الحدود الفعلية للفئة ويتناسب ارتفاعه مع تكرار تلك الفئة عندما تكون الفئات متساوية.

بب- المضلع التكراري:

ونحصل عليه برصد النقاط التي احداثياتها مراكز الفئات والتكرارات المقابلة لها، ثم نغلق المضلع فناخذ مركز فئة قبل أول فئة ونضع تكرارها صفراً ونلخذ مركز فئة بعد آخر فئة مباشرة ونضع تكرارها صفرا ثم نغلق المضلع.

(4) التماثل والالتواء والتفرطح

ا- يعتبر التوزيع متماثلاً إذا أمكن إقامة عمود على المحور الأفقي له بحيث يقسم التوزيع إلى قسمين ينطبقان على بعضهما تمام الانطباق.

ب- يعقبر القوزيع ملتوياً نحو اليمين إذا كان له طرف ممتد إلى اليمين، ويسمى توزيعاً موجب الالتواء، ويعتبر التوزيع ملتويانحو اليسار إذاكان له طرف ممتد إلى اليسار

ويسمى توزيعا سالب الالتواء.

ج- التفرطح هو الاستواء أو كون. لتوزيع مدبباً أم لا.

(5) مقاييس النزعة المركزية

أ- الوسط الحسابي هو مجموع قيم البيانات مقسوماً على عددها.

ب- الوسيط هو تلك القيمة التي يكون مجموع التكرارات قبلها مساوياً لمجموع التكرارات بعدها.

ج- المنوال: هو القيمة الاكثر تكراراً أو شيوعاً في جوارها، أي القيمة التي يقابلها قمة في المضلع التكراري.

(6) مقاييس التشتت

أ- المدى، ويساوي أعلى قيمة ناقصاً أصغر قيمة.

ب- المدى الربيعي وهو الربيع الثالث ناقصاً الربيع الاول.

ج- التباين ويساوي مجموع مربعات انحرافات البيانات عن وسطها الحسابي مقسوماً على (عدد البيانات -1).

د- الانحراف ألمعياري وهو الجذر التربيعي للتباين.

(7) أثر التحويلات الخطية

أ- أثر التحويلات الخطية على مقاييس النزعة المركزية:

تتأثر مقاييس النزعة المركزية بالتحويلات الخطية بنفس الطريقة التي تتأثر بعا البيانات. فإذا ضربت كل قيمة من البيانات في عدد ثابت وجمعت إليها عدداً ثابتاً توجب عليك عمل الشيء نفسه لمقاييس النزعة المركزية، فتضرب الوسط الحسابي بنفس العدد الثابت وتجمع له العدد الثابت الذي أضفته لكل قيمة في البيانات. وينطبق هذا الأمر على كل من الوسيط والمنوال.

ب- أثر التحويلات الخطية على مقاييت التشتت:

لا تتأثر مقاييس التشتت إذا أضفت قيمة ثابتة لكل قيمة في البيانات، ولكن هذه المقاييس تتأثر بالضرب، فإذا ضربت كل قيمة من البياتات في عدد ثابت a فإن ذلك يؤثر على مقاييس التشتت كما يلي:

المدى بعد التحويل =|جا المدى قبل التحويل

المدى الربعي بعد التحويل =|a| المدى الربعي قبل التحويل

التباين بعد التحويل التباين قبل للتحويل

الانحراف المعياري بعد التحويل =|a| الانحراف المعياري قبل التحويل

حيث |a| القيمة المطلقة للعدد 2، وهي عبارة عن العدد a بدون اشارة، فمثلا

(8) مقاييس التشتت النسبية

أ- معامل الاختلاف المعياري ويساوي الانحراف المعياري مقسوماً على الوسط الحسابي.

ب- معامل الاختلاف الربيعي ويساوي الانحراف الربيعي مقسوماً على الوسيط.

ج- العلامة المعيارية وتساوي حاصل الفرق بين القيمة ووسطها الحسابي مقسوما على الانحراف المعياري.

(9) العشيريات والمئينات

أ- العشير يعني تقسيم البيانات إلى عشرة أقسام بعد ترتيبها وكل عشير هو القيمة التي يكون عشيرته من البيانات أقل منه أو تساويه.

ب- المئين يعني تقسيم البيانات إلى مائة قسم بعد ترتيبها وكل مئين هو القيمة التي يكون ٠ 0.01 مئينته من البيانات أقل منه أو تساويه.

10، لمحة مسبقة ع الوحدة الدراية لشالته ' محة بقةعيالوح $_5$ة لدرا ية ال تالتة . . قبتتةي

بعد دراستك الوحدة الثانية (الإحصاء الوصفي لمتغير واحد) أصبح بإمكانك أن تلخص البيانات وتعرضها على شكل توزيع تكراري، وأن تمثلها تمثيلاً بيانياً ثم تقوم بحساب عدد من مقاييس الموقع (مقاييس النزعة المركزية) وعدد من مقاييس التشتت وتباعد البيانات بعضها عن بعض أو عن مقياس النزعة المركزية المحسوبة لها، وأن تعرف فيما إذا كان منحنى التوزيع التكراري ملتو أو لا ونوعية الالتواء، وطبيعة قمته مفرطحة أو مسطحة أو متوسطة التفرطح.

بعد هذه الدراسة سنوسع دائرة المعرفة لتكون لمتغيرين وليس متغيرا واحدا، وهذا ما ستجده في الوحدة الثالثة (الإحصاء الوصفي لمتغيرين) وستجد في بداية الوحدة الجديدة كيفية إنشاء جدول تقاطع لمتغيرين نوعيين كل منهما مقسم إلى عدة أصناف.

وستتعلم بأن دراسة متغير لوحده لا تكفي وإنما علاقة المتغير بمتغير آخر وهو ما يسمى بالارتباط وكيفية حسابه وما هو اتجاه الارتباط وما هي شدته كل ذلك عبر مجموعة من المقاييس الاحصائية تعتمد في أماسها على الوحدة الثانية. وتزداد معرفتك عندما تجد أن هذا الارتباط يمكن أن يمثل باقتران خطي أو لا خطي وفي مثل هذه الحالة إ يمكن استخدام هذا الاقتران للتبق بقيم معينة.

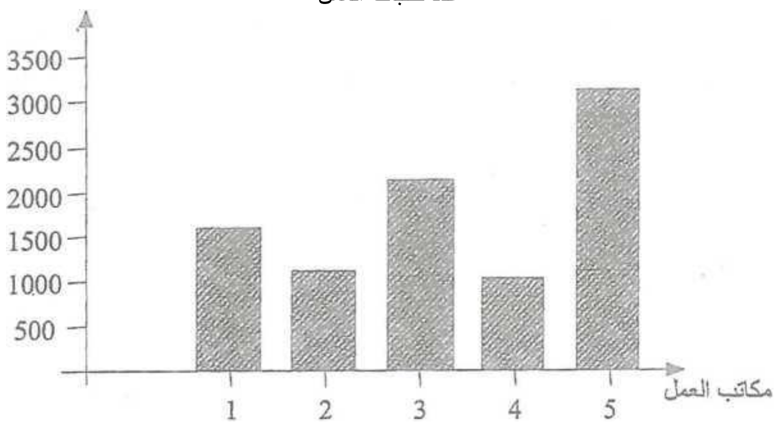
١١. إجابات التدريبات

تثويج (1)

1- العرض بطريقة المستطيلات

أرسم خطين متعامدين، وضع أعداد مكاتب العمل على الخط الأفقي. بحيث يقع كل عدد في منتصف قطعة مستقيمة بمقياس رسم معين وتكون هذه القطع منفصلتين بعضها عن بعض وارسم على كل قطعة مستقيمة مستطيلا ارتفاعه يتناسب مع عدد طلبات العمل التي وردت إلى المكتب كما يظير في الشكل. لاحظ أن الخط العمودي يمثل أعداد طلبة العمل، ولذلك فهو يبين تلك الأعداد بمقياس رسم مناسب.

غدد طلبات العمل



2- عرض البيانات بطريقة الخطر المنكسر

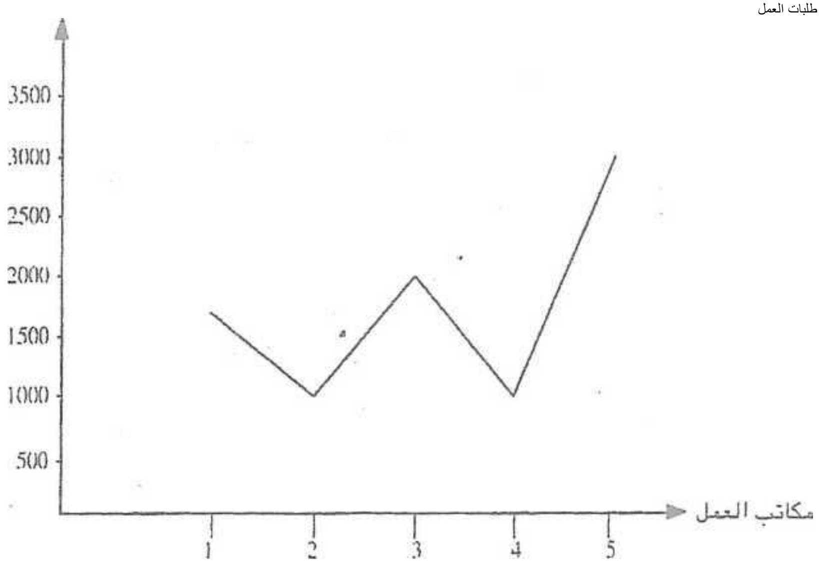
أرسم خطين متعامدين يملل الخط الأفقي منهما مكاتب العمل (المحور السيني) ويمثل الخط العمودي اعداد طلبات العمل (المحور الصادي).

استعمل مقياس رسم مناسب وارصد النقاط التي احداثياتها:

(5,3100) ، (4,1090) ، (3,2100) ، (2,1200) ، (1,1700)

هل هذه الناط، كل نقطة بالتي تليها، بخطوط مستقيمة، تحصل على العرض بطريقة الخط المنكسر كما في الشكل التالي:

عدد طلبات العمل



فقهوييب (2)

1- عدد الفئات المتساوية 6.

2- أصغر قيمة في البيانات 55.

وأعلى قيمة 92 وبذلك فإن المدى يساوي: 37 = 92-55

و- طول الفئة من قسمة -ت—=6.17 وبالتقريب ٠ 7
6
حيث تقرب الجواب إلى أقرب عدد صحيح أعلى.

4- خذ الحد الأدنى للفئة ليكون 55

5- الحد الأدنى الفعلي للفئة الأولى يساوي ٠ 55-0.5 ، أي تساوي 54.5 لأن البيانات أعدادصحيحة، فتطرح 0.5

6- الحد الفعلي للفئة الأولى يساوي 61.5 = 54.5؛7

الحد الإعلى للفئة الأولى يساوي 61.5 = 61.5-0.5

وبذلك فإن حدود الفئة الأولى هي: 55-61

والحدودالفعلية للفئة الأولى 54.5- 61.5

7- أضف 7 (وهو طول الفئة) لكل عدد لبناء جميع فئات التوزيع التكراري كما يظهر في الجدول الآتي:

الفئات	التكرارات
61-55	7
68-62	5
7ذ-69	4
82-76	5
89-83	5
96-90	4
المجموع	30

تهويب (3)

اقسم تكرار كل فئة على مجموع التكرارات (30) تحصل على التكرار النسبي وبذلك تحصل على التوزيع التكراري النسبي كما في الجدول الآتي:

الفئة	التكرار النسبي	التكرار المئوي %
(1)	(2)	(3)
61-55	0.233	23.3
68-62	0.167	16.7
75-69	0.133	13.3
76-%2	0.167	16.7
89-83	0.167	16.7
96-90	0.133	13.3

لاحظ أن مجموع التكرارات النسبية 1-

تدويبب (4)

حول كل تكرار نسبي إلى نسبة مئوية، وذلك بالضرب في 100% وضع الإجابة في عمود (3) في الجدول (2) فيعطيك العمودان (1)، (3) التوزيع التكراري المئوي، ويكون مجموع الاعداد في العمود (3) يساوي 100%.

تهويب (5)

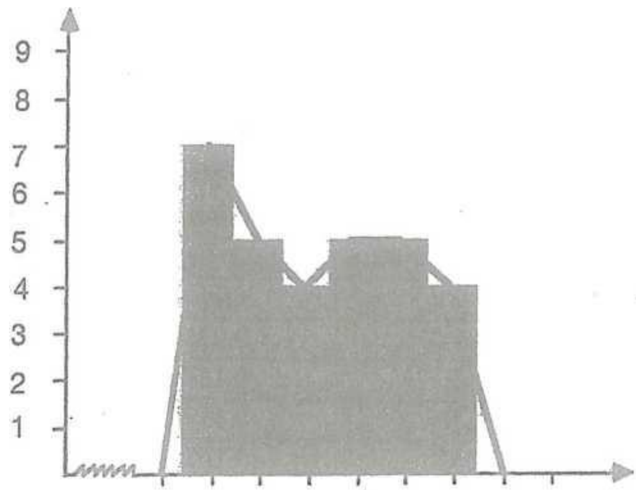
1- نعين مراكز الفئات الواردة في جدول تدريب (2).

2- نمثل عدد الساعات المعتمدة على المحور الأفقي، وعدد الطلاب على المحور العمودي.

3- نرسم المستطيلات التي تمثل كل فئة وتكرارهافنحصل على المضلع التكراري المبين في الشكل رقم (1).

4- نعين على المحور الأفقي مراكز الفئات ونضيف مركزي فئتين أحدهما على يمين مركز الفئة العليا مباشرة، والآخر على يسار مركز الفئة الدنيا.مباشرة.

5- نعين النقط المكون كل منها من الزوج المرتب (مركز الفئة،تكرار الفئة). ونصل بين النقط بخطوط مستقيمة فنحصل على المضلع التكراري المبين في الشكل الآتي:

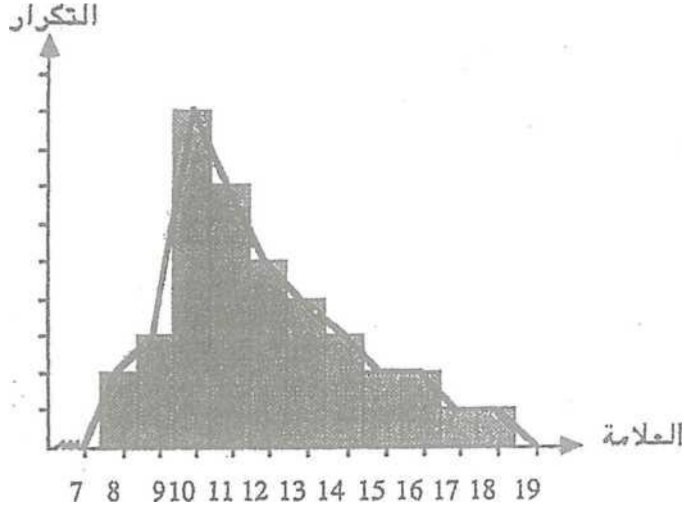


51 58 6572 7986 93 100

تهوبيبه (6)

يتضح لك من الشكل الآتي أن هناك قيمة واحدة تقابل العلامة 10، وبذلك فالتوزيع أحادي المنوال، والمنوال قيمته 10، وهذا واضح من التوزيع التكراري كذلك.

ويتضح أيضاً أن التوزيع ملتو نحو اليمين، حيث أنه ممتد نحو اليمين ولذلك فهو توزيع موجب الالتواء، وهو غير متماثل.



التكرار
العلامة
7 8 910 11 12 13 14 15 16 17 18 19

قهديبب (7)

أضف على التوزيع التكراري عموداً جديداً فيه: (مركز الفئة x التكرار) أي كما يظير في

Xi fi	التكرار fi	مركز الفئة x؛
550	10	55
975	15	65
1650	22	*75*
105	13	85
475	05	95
4755	65	المجموع

الجدول الآتي:

احسب الوسط الحسابي من المعادلة:

٧_Z^x^jfj_4755

n 65

= 73.15

قهوببب ا8٤

-1 نرتب البيانات تصاعدياً فتصبح

8.8، 9.1، 11.4، 12.4، 17.8، 19.5، 25.0، 26.5، 30.2

وبما أن عدد البيانات (9) هو عدد فردي فإن الوسيط

MᵒX(^)X 17.8

-2 البيانات بعد ترتيبها تصاعديا هي

100 ،27 ،19 ،12 ،81

وبما أن عدد البيانات (6) هو عدد زوجي فإن الوسيط

7 2

2 2

تدريي(و

أكتب التوزيغ التكراري في التدريب وأضف إليه عمود التكرار المتجمع فيصبح الجدول

كما يلي:

الحدود الفعلية	التكرار	التكرار المتجمع
16.5-10.5	5	5
22.5-16.5	6	11
28.5-22.5	9	20
34.5-28.5	6	26
40.5-34.5	5	31
46.5-40.5	3	34
52.5-46.5	2	36

n = 36 وبذلك ئ=18
2

عين افئة الوسيطية، وهي أول فئة يزيد تكرارها المتجمع عن 18 أو يساويه وبذلك فإن

الفئة المطلوبة هي 22.5 - 28.5

أرجع إلى تعريف الوسيط و أكتب قيم الرموز التي تحتاجها:

إ ع

M=a, xC

fei

22.5 ·ai

11=111

تلة = 9

إذا (6) M=22.5+ $\frac{18-11}{9}$

٦ 6أب لآ،حلآ: 9

= 22.5+4.7

= ٦٦ ح

تهوييب (10)

(1) من الواضح أن العدد 2 قد تكرر ثلاث مرات، وهو أكبر تكرار في البيانات، حيث أن البيانات 3، 5، 6، § ظهر كل منها مرة واحدة فقط.

إذا المنوال هو 2

(2) بالرجوع إلى الجدول في التدريب (9) ما هو أكبر تكرار؟

من الواضح أن أكبر تكرار هو 9

ولذلك فإن الفئة 22.5 - 28.5 هي فئة منوالية.

وبالتالي فإن المنوال التقريبي هو مركز هذه الفئة

لأميعدس

قيوس(11)

أ- العلامات مرتبة ترتيباً تنازلياً وعددها 13 أي نه فردي وبالتالي الوسيط هو العلامة المقابلة للترتيب، أي أن الوسيط هو العلامة 62 .

إذا تنيرت العلامة 67 ولصبحت 89 فكما هو واضح أن ترتيب العلامة 62 بقي كما كان سابقا (ترتيبها السابع)، وبذلك لا يتغير الوسيط ويبقى 62 .

ب- عند تعديل العلامة 40 لتصبح 61 فإن ترتيب العلامة 62 لا يتغير وبذلك يبقى الوسيط 62 .

لاحظ أن العلامات المرتبة تنازلياً بعد التعديل هي:

100، 95 ،91،83،76 ، 67 ،62 ، 57 ،54 ،50 ، 48، 42 ، 40.

ج- الوسط الحسابي للقيم الأصلية هو 66.538 ، وعندما تعدل العلامة 67 لتصبح 89فإن الوسط الحسابي يتغير لأن قيمة إحدى العلامات تغيرت، وفي هذه الحالة يصبح الوسط الحسابي 68.231.

د- وعندما تعدل العلامة 40 لتصبح 61 فإن الوسط الحسابي يتغير لأن قيمة إحدى العلامات تغيرت، وفي هذه الحالة يصبح الوسط الحسابي 68.154 .

تئوبيب (12)

المدى = أعلى قيمة - أدنى قيمة

ولذلك فإن المدى يساوي: 77 = 22-99

المدى الربعي ٠ الربيع الثالث - الربيع الأول

ويساوي 55 ه 37-92

تهويب (13)

أوجد أولا الوسط الحسابي

ع 7 ٦٧٦Zˣi - 10 + 11 + 7 + 9 + 12+15 + 6:70؛ 10

نستعمل التعريف لايجاد التباين:

عثتئيي: s2

6 (10-10)2_+_(11-10)2+(7-10)^2_+_(9-10)2+(12-10)2_+_(15-10)^2+(6-10)2

: Q^2 +ا2 + (—3)2 + (—1)2 + ت2 + ج5 + (ب)2 ؛ت ? 33

تهويبب (14)

X	أ2	X	f.	؛X
7128	324	396	22	18
8993	529	391	17	23
79184	784	2828	101	28
196020	1089	5940	180	33
173280	1444	4560	120	38
92450	1849	.2150	50	43
80640	2304	1650	35	48
28090	2809	530	10	53
50460	3364	870	15	58
716245		19345	550	المجموع

وضعنا مراكز الفئات ؛X في عمود (1) والتكرارات في عمود (2) ثم ضربنا Xi في fi ووضعنا الناتج X: fj في العمود (3) وأخيراً ضربنا ^x في fi ووضعنا الناتج 1؛ xf في العمود الخامس.

x!:35.،7

550نحسب الوسط الحسابي

[ك-,؛لآ]يإ*8

لايجاد التباين:

= ل716245-550٢(35.17)2]

لت(6 7ح549

= 35934.11 ;65 45

549

S=٧65.45=8.09

تووتييج (15)

باستعمال نظرية (1.2) تجد أن:

الوسط الحسابي بعد التعديل =82- 0.9x80+10

المنوال بعد التعديل ' 0.9x66+10 = 69.4

الوسيط بعد التعديل = 0.9x74+10 = 76.6

تشوجبي (16)

1- الانحراف المعياري بعد التعديل =

الاتحراف المعياري قبل التعديل، لأن التعديل كان بإضافة ثابت لكل علامة وهو العلامه 7، إذن الانحراف المعياري بعد التعديل تع 9

2- مدي العلامات بعد التعديل -272 0.8x90

وتباين العلامات بعد التليل -16- x25 (0.8)

تووجبب (17)

لإيجاد كلاً من معامل الاختلاف المعياري و العلامة المعيارية لابد أولاً من حساب قيمة كل من الوسط الحسابي والانحراف المعياري لقيم التدريب.

الوسط لصسبي 7.5=أ:ت

الشرلس سيره كتوتقة-

ولحساب معامل الاختلاف المعياري نعوض بالصيغة:

Coefficient of standardized variation = 3*100%

= 1!x100%=43.615%
7.5

وبالتتويض بصيغة العلامة المعيارية سنجد بأن القيمة المعيارية للقيمة 5 هي: 0.7642يي٠شب-=ة

S 3.2711

والعلامة المعيارية للقيمة 12 هي:

$_2$=ختتقعئكث:ا,3756

S 3.2711

تهويي (18)

المئين 25 والمئين 75 محسوبة في حل مثال 22 حيث المئين 25 هو والمئين 75 هو Q_3.

12 مسرد المصطلحات

- الالتواء Skewness: هو عدم التماثل ويكون الالتواء موجبا إذا كان المنحنى مائلا إلى اليمين ويكون سالبا اذا كان المنحنى مائلا إلى اليسار.

- الانحراف المعياري Standard Deviation: هو الجذر التربيعي للتباين.
- أثر التحييلات ^^Effect of Linear Transformations: هو أثر التحويل الخطي على الوسط الحسابي أو الوسيط أو المنوال أو التباين والتحويل الخطي هو عبارة عن ضرب البياتات بعدد ثابت وجمع عدد ثابت لها.
- التباين Variance: هو أحد مقاييس التشتت ويعرف بمجموع مربعات انحرافات البيانات عن وسطها الحسابي مقسوماً على (عدد البيانات ناقصاً واحد).
- التفرطح Kurtosis: هو تدبب منحنى التوزيع أو انبساطه.
- التكرار النسبي Relative Frequency: هو نسبة تكرار أي فئة إلى مجموع التكرارات.
- التمثيل بيانياً Graphical Presentation: هو تمثيل البيانات بالطرق البيانية مثل الأعمدة أو الدائرة أو الخط المنكسر أو المنحنى.
- التماثل Symmetry: هو إمكانية تطابق المضلع أو المنحنى التكراري حول عمود مقام على نقطة على المحور الأفقي للمنحنى.

- التوزيع التكراري Frequency Distribution: هو جدول يحتوي على فئات غير متداخلة وازاء كل منها عدد البيانات (التكرار).
- الطريقة التصويرية Pictorial Method: إحدى طرق عرض البيانات وذلك برسم صورة تمثل قيمة معينة.
- طريقة الدائرة Pie Chart: إحدى الطرق البيانية لعرض البيانات الأولية حيث تمثل الدائرة مجموع البيانات كلها وكل ظاهرة تمثل جزءاً من هذه الدائرة.
- طريقة الجداول Tables Method: إحدى طرق عرض البيانات الأولية وتستعمل في عرض ظاهرة أو عدة ظواهر مع الزمن أو المسميات أو كليهما.
- طريقة الخط المنحني Curve Method: هو خط منكسر ممهد.
- طريقة الخط المنكسر Broken Line Method: تستعمل مثل طريقة المستطيلات ولكن بدلاً من رسم مستطيل نقوم برصد النقط التي احداثياتها (الزمن، قيم الظاهرة).
- طريقة المستطيلات أو الاعمدة Bar Method: هي طريقة عرض البيانات بواسطة مستطيلات متباعدة قواعدها على المحور الأفقي وارتفاعاتها تمثل قيم الظاهرة، وتستعمل لعرض ظاهرة أو عدة ظواهر مع المسميات أو الزمن أو كليهما.
- العلامة المعيارية Z-score: وهي أحد مقاييس التشتت النسبي ومن إحدى استحمااتها مقارنة بعد القيم عن وسطها الحسابي بعد إزالة أثر وحدات القياس.
- المدى Range: الفرق بين أعلى قيمة وأصغر قيمة في البيانات.
- المدى الربعي Quartile Range: هو الفرق بين الربيع الثالث والربيع الأول.
- المدرج التكراري Frequency Histogram: هو تمثيل التوزيع التكراري بيانيا على شكل مستطيلات قواعدها على المحور الأفقي وارتفاعاتها تتناسب مع تكرار الفئات في حالة التوزيع التكراري ذي الفئات المتساوية.
- المضلع التكراري Frequency Polygon: وهو تمثيل التوزيع التكراري بمضلع مغلق، ويمكن عمله بتنصيف الأضلاع العليا للمستطيلات في المدرج التكراري ثم وصل النقاط بخطوط مستقيمة وإغلاق المضلع.

" معامل الاختلاف Coefficient of Variation: هو أحد مقاييس التشتت النسبية ويستعمل للمقارنة بين مجموعتين من البيانات تختلفان في وحدات قياس قيم كل منها أو بين مجموعتين من البيانات مقاستين بنفس وحدات القياس ولكن هناك فرق كبير في متوسطيهما.

- مقاييس النزعة المركزية Measures of Central Tendency: هي الوسط الحسابي

والوسيط والمنوال وهي مقاييس موقع.

- المنوال Mode: هي القيمة التي يقابلها أكبر تكرار في جوارها.
- الوسط الحسابي Ari hmatic Mean: هو مجموع القيم مقسوما على غددها.
- الوسيط Median: هو القيمة التي يكون أعلى منها نصف البيانات وأصغر منها النصف الآخر.

13- المراجع

1- أبوزينة، فريد؛ لطفية، لطفي؛ الخليلي، خليل. الطرق الإحصائية في التربية والعلوم الإنسانية. عمان: دار الفرقان للنشر والتوزيع، 1989.

2- أبو صالح، محمد، الطرق الإحصائية: الإصدار الثاني. عمان: دار اليازوري العلمية للنشر والتوزيع، 2001.

3- أبو صالح، محمد، الموجز في الطرق الإحصائية: الطبعة العربية الأولى. عمان: دار اليازوري العلمية للنشر والتوزيع، 2002.

ثاً

الوحدة الثالثة

الإحصاء الوصفي لمتغيرين

Bivariate Descriptive Statistics

n : لويودج:||ا!ااااا

س.1 •

1.1 تمهيد

عزيزي الدارس، مرحباً بك إلى هذه الوحدة، التي تنقسم إلى ثلاث أقسام رئيسة.

في الوحدتين السابقتين استعرضنا ظاهرة واحدة أو متغيرا واحداً وكنا نجمع عنه البيانات ونصفها ونحسب له بعض المقاييس التي تم دراستها. وفي هذه الوحدة سنقوم بدراسة ظاهرتين معاً أو متغيرين أثنين بعد أن نقيس كل عنصر من الظاهرتين، والأساس في هذا القياس أن تكون الظاهرتان هي لوحدة أو مفردة إحصائية واحدة؛ أي أن هناك زوج واحد من القيم المتناظرة لكل وحدة إحصائية، وعلى سبيل المثال قياس كل من وزن وطول الطالب في مدرسة إبتدائية، أو الدخل الشهري والأجور المدفوعة عن استهلاك الكهرباء للعائلة. وقبل معرفة كيفية حساب المقاييس الإحصائية فيما يخص هذا النوع من البيانات سنقوم معاً بتكوين ما يطلق عليه بالجداول المتقاطعة Cross Tables، ثم حساب مقياس معامل الارتباط الذي يقيس العلاقة بين المتغيرين، أي بين ازواج القيم المتناظرة وتفسيره وفهم معناه.

وستعرف عزيزي الدارس، أن معرفة معامل الارتباط Correlation Coefficient ليس كافياً من دون أن نحدد كيفية صياغة علاقة رياضية بينها، وهو ما يطلق عليه بمعادلة خط الانحدار Regression Line Equation.

2.1 أهداف الوحدة

بحد فراغك من دراسة هذه الوحدة يجب أن تكون قادراً على أن:

1-تتشئ جدول التقاطع الثنائي لمتغيرين نوعيين.

2- تعرف مفهوم الارتباط.

3- تستخدم شكل الانتشار كوسيلة تمهيدية لمعرفة العلاقة بين متغيرين كميين من حيث كونها خطية أو غير خطية، أو من حيث كونها طردية الاتجاه أو عكسية.

4- تحسب معامل الارتباط الخطي بين متغيرين وتفسر نتيجته.

5- تحسب معامل إرتباط الرتب وتفسر نتيجته.

6- تحسب معاملي خط انحدار متغير تابع على متغير مستقل.

7- تقدر قيمة متغير تابع عند معرفة قيمة المتغير المستقل.

3.1 أقسام الوحدة

قسمت هذه الوحدة (الإحصاء الوصفي لمتغيرين) إلى ثلاثة أقسام رئيسية: القسم الأول يتناول كيفية إنشاء جداول التقاطع الثنائي لمتغيرين نوعيين، حيث بدأ القسم بإعادة تناول المتغير النوعي من حيث أن لكل وحدة إحصائية قراءتين تمثلان أزواج من القيم المتناظرة (أو يطلق عليها كذلك بالقيم المرتبة) من خلال العصل اليدوي، وهذا ما يحقق الهدف الأول.

وفي القسم الثاني تم تناول موضوع الارتباط من خلال دراسة متغيرين كميين للمفردة الإحصائية الواحدة وكيفية توضيح العلاقة بينهما مستخدمين شكل الانتشار كوسيلة تمهيدية لمعرفة طبيعة العلاقة بين المتغيرين، فيما إذا كانت خطية أو غير خطية واتجاهها إذا كانت طردية أو عكسية، وهذا ما يحقق الهدفين الثاني والثالث. وبعد تعريف المقياس الذي تقاس به العلاقة بين المتغيرين ويسمى معامل الارتباط، سنتعلم كيفية حسابه، ثم ماذا تعني نتيجته، بمعنى آخر تفسير دلالة قيمته وهو الهدف الرابع. كماسيتناول حساب معامل ارتباط بين متغيرين أحدهما على الأقل وصفي من خلال رتبهما وليس قراءاتهما وهذا ما يحقق الهدف الخامس.

أما في القسم الثالث فستتعلم كيفية حساب معالم خط المستقيم، وهما الميل والحد الثابت عندما تكون العلاقة بين المتغيرين خطية وهذا ما يحقق الهدف السادس، وفي ضوء خط المستقيم هذا (الذي يسمى بخط الانحدار) ستستطيع أن تقدر وتتنبئ بقيم المتغير التابع بمعلومية قيمة للمتغير المستقل وهومايحقق الهدف للسابع.

اب4القراءات لمساعدة ك •

لكي ترسخ في ذهنك المفاهيم التي سترد في هذه الوحدة ننصحك بالرجوع إلى القراءات المساعدة التالية:

1- أبو صالح، محمد، الطرق الإحصائية: الاصدار الثاني. عمان: دار اليازوري العلمية للنشر والتوزيع، 2001.

’2 أبو صالح، محمد، الموجز في الطرق الإحصائية: الطبعة العربية الأولى، عمان: دار اليازوري العلمية للنشر والتوزيع، 2002.

5.1 ما تحتاج إليه لدراسة الوحدة

بعد أن تكون قد وفرت المكان المناسب للدراسة، تحتاج إلى توفير بعض المستلزمات الآتية: قلم، ومسطرة، وآلة حاسبة، وبعض الأوراق للكتابة.

أثناء دراستك لمادة الوحدة ستجد في ثناياها أسئلة التقويم الذاتي التي تساعدك في مراجعة وتلخيص الوحدة كما ستجد تدريبات عليك القيام بها، فهي سترسخ الأفكار المطروحة في الوحدة وتمنحك الفرصة لاختبار تعلمك ومدى استيعابك وفهمك لهذه الوحدة.

وفي حالة حاجتك إلى مزيد من التوضيح حول بعض الموضوعات أثناء دراستك لهذه الوحدة، لا تتردد في مناقشتها مع مرشدك، فهو سيكون سعيد بذلك وسيرحب بك كثيرا.

2. جدول التقاطع الثنائيلمتغيرين نوعيين

EWiate Qualitative Coss.Table . .٠.. ٠ '

1 .2 مفهوم جدول التقاطع

في الوحدة الأولى من المقرر عرفت بأن البيانات النوعية أو الوصفية هي الظواهر التي لا يمكن قياسهابالأرقام التددية مباشرة كالجنس وتخصص ألطالب وتقدير العلامات (امتياز، جيدجداً، جيد، مقبول، ضعيف) مثلا.

إن كل واحدة من هذه الظواهر مصنفة إلى أصناف، فظاهرة الجنس مصنفة إلى جزئين هما النكور والاناث، وظاهرة تقدير علامات للطلبة مصنفة إلى أصناف هي امتياز، جيد جداً، جيد، مقبول وضعيف. ومفهوم جدول التقاطع الثنائي أن لدينا ظاهرتين أو متخيرين فقط كلا منهما مصنف إلى عدة أصناف بحيث أن الجدول المطلوب إنشائه يتضمن الظاهرتين وحالة التقاطع أي الاشتراك بينهما. في هذة الحالة سيكون عرض إحدى الظاهرتين على شكل صفوف كل صف يتضمن صنفاً من أصناف الظاهرة الأولى، بينما اسيكون عرض الظاهرة الثانية على شكل أعمدة بحيث كل عمود يتضمن صنفاً من أصناف الظاهرة الثانية. وفي هذه الحالة سيكون لدينا عدد من الخلايا عددها هو حاصل ضرب عدد الصفوف. في عدد الأعمدة، وتمثل كل خلية حالة التقاطع بين إحدى اصناف الظاهرة الأولى مع واحدة من أصناف الظاهرة الثانية، وبالنسبة للظاهرتين الجنس وتقدير العلامات سيكون عدد الصفوف 2 (على أساس أن متغير الجنس سنضعه في الصف وله صنفان) ولتقدير العلامات هناك 5 أعمدة بعدد أصناف تقدير التلامات وعند ذلك ستتكون لدينا 102522 خلايا هي حاصل جميع الحالات المتقاطعة بين الجنس وتقدير العلامات، وشكل هذا الجدول سيكون على النحو الآتي:

الجنس	تقدير العلامات				
	امتياز	جيد جدا	جيد	مقبول	ضعيف
ذكور					
إناث					

إن جدول التقاطع أعلاه يشير إلى وجود صفين أحدهما للذكور والآخر للإناث، وإلى وجود5 أعمدة كل عمود يشير إلى صنف تقدير العلامات، وأن عدد الخلايا (أو المربعات) التي تكونت هي 10. الخلية في الصف الأول (من الذكور) والخلية في العمود الأول من تقدير ألعلامات (إمتياز) تشير إلى عدد الذكور الذين تقدير علاماتهم هي امتياز،

أي عدد الحالات المتقاطعة بين الذكور وعلامات الامتياز، بينما الخلية تحت عمود ضعيف وصف الإناث فتشير إلى الحالات المتقاطعة بين الإناث اللواتي علاماتهن ضعيفة.

ويسمى الجدول أعلاه بجدول التقاطع الثنائي لمتغيرين نوعيين، وفي أغلب الأحيان بجدول الاقتران Contingency Table أو جدول 2x5 ، أي أن هناك صفين وخمسة أعمدة. والصيغة العامة لجداول الاقتران تكتب rxc وتقرأ تمثل عدد الصفوف فيما ء تمثل عدد الأعمدة. (r by c حيث r
والمثال الآتي يبين كيفية احتساب عدد الحالات في كل خلية.

I'

مثثل (1)

البيانات الآتية تمثل أزواج من القيم المتناظرة ل 30 طالباً حسب التخصص (X) وهذه التخصصات هي المحاسبة، الحاسوب، الهندسة والآداب وتقديرات معدلاتهم التراكمية (٧) وتضم امقيار، جيد جدائ جيد مقبول وضعين:

محاسبة	حاسوب	هندسة	آداب	حأسوب	التخصص X
جيد جدا	جيد	أمتياز	مقبول	مقبول	التقدير Y

حاسوب	حاسوب	هندسة	هندسة	محا سدة	التخصص X
جئب	جيد	جيد	جيد جدا	جيد	التقدير Y

A ادا ۱ ۱	حاسوب	هندسة	اداب	حاسوب	التخصص X
جيد	ضعيف	جيد	جيد	امتياز	التقدير Y

محاسدة	اداب	حاسوب	اداب	اداب	التخصص X
جيد جدا	جيد	ضعيف	مقبول	جيب	التقدير Y

حاسوب	حاسوب	هندسة	آداب	حاسوب	التخصص X
جيدذ	جيد جدا	مقبول	جيد جدا	ضعيف	التقدير Y

هذدسة	حاسوب	محاسبة	اداب	حاسوب	التخصص X
جيد جدا	مقبول	مقبول	ضعيف	جيد جدا	٠ التقدير تآ

والمطلوب وضع البيانات في جدول اقتران.

الحل:

لدينا أربعة تخصصات ستشكل صفوف المتغير x وخمسة تقديرات للعلامات ستشكل أعمدة المتغير٦ وعدد الخلايا (20) خلية، لذا يكون الجدول المطلوب 4x5 وعلى النحو الآتي:

Y / X	امتياز	جيد جدا	جيد	مقبول	ضعيف
حاسوب					
هندسة					
آداب					
محاسبة					

لتفريغ بيانات المثال في الجدول 54x، يلاحظ أن الطالب الأول هو من تخصص الحلسوب وتقدير علامته مقبول وعليه نضع مستقيم عمودي على شكل ا في الخلية التي تقع تحت عمود مقبول و أمام تخصص الحاسوب ونسمي هذه الخطوة بتفريغ بيانات المفردة الأولى لإزواج القيم المتناظرة ثم ننتقل إلى المفردة الثانية وهي لطالب تخصصه آداب وتقدير علامته كذلك مقبول، لذا نقوم بوضع ا تحت عمود مقبول وأمام صف آداب ونكرر هذه العملية لجميع المفردات الإحصائية المتبقية (وهنا 28 من الطلبة).

إن الجدول الأخير بعد الانتهاء سيظهر كما يلي:

Y / X	امتياز	جيد جدا	جيد	مقبول	ح لم ولج -t
حاسوب	1	2	4	2	3
هندسة	1	2	2	1	
آداب		1	3	2	1
محاسبة		2	2	1	

وبجمع قيم الخلايا لكل صف ولكل عمود يصبح الجدول

Y / X	امتياز	جيد جدا	جيد	مقبول	ضعيف	المجموع
حاسوب	1	2	4	2	3	12
هندسة	1	2	2	1		6
آداب		1	3	2	1	7
محاسبة		2	2	1		5
المجموع	2	7	11	6	4	30

وهي عدد 30ويلاحظ من الجدول أعلاه بأن مجموع الصفوف أو مجموع الأعمدة يساوي المفردات (الطلبة) المعطى في نص المثال، وأن عدد طلبة الحاسوب 12، والهندسة 6 ، وجيد جداً 2والآداب 7، والمجاسبة 5، بينما عدد الطلبة الذين تقدير علامتهم امتياز كان ، وجيد 11، ومقبول 6 و أخيراً كان عدد الطلبة الذين حصلوا على تقدير ضعيف 4 طلاب، 7 وبالنسبة للقيم داخل الخلايا فيلاحظ أن 4 من طلبة الحاسوب قد حصلوا لى تقدير جيد بينما 3 منهم حصل على تقدير ضعيف وهكذا لبقية القيم.

دريب (1)

تمثل البيانات التالية، الجنس x والبرنامج y الذي يلتحق به الدارسون في إحـــدى جموعات مقرر مبادئ الإحصاء:

الجنس X	ذكر	ذكر	أنثى	ذكر	أنثى	أنثى	أنثى	ذكر
البرنامج Y	حاسوب	حاسوب	تربية	إدارة	إدارة	تربية	تربية	إدارة

الجنس X	ذكر	ذكر	ذكر	أنثى	أنثى	أنثى	ذكر	ذكر
البرنامج Y	تربية	إدارة	تربية	تربية	إدارة	تربية	حاسوب	حاسوب

الجنس X	ذكر	أنثى	ذكر	أنثى	ذكر	أنثى	أنثى	ذكر
البرنامج Y	إدارة	تربية	تربية	إدارة	إدارة	حاسوب	تربية	حاسوب

اعمل جدول تقاطع ثنائي، تمثل فيه الصفوف الجنس x، وتمثـــل فيـــه الأعمـــدة برنامج y.

سئلة التقويم الذاتي (1)

ماذا تمثل خلية التقاطع في جدول التقاطع الثنائي؟

٠3 الارتباط Correlation

1.3 مفهوم الارتباط

درست في الوحدة السابقة المسائل والتدريبات المتتلقة بقياس مشاهدات متغير واحد، وبعبارة أخرى أنحصرت هذة المسائل والتدريبات في ظاهرة واحدة، ففي كل الأمثلة والتدريبات للوحدة الثانية كانت القياسات المحسوبة هي لبيانات تخص ظاهرة واحدة؛ أي أن البيانات التي تم التعامل معها كانت لظاهرة واحدة تم تسجيلها لمجموعة من المفردات الإحصائية كالأفراد مثل: علامات الطلبة في أحد الامتحانات، والأجور اليومية لعدد من العمال على سبيل المثال. وكان الاهتمام منصبا على وصف بيانات الظاهرة كعرضها في رسوم، أو عمل جدول توزيع تكراري، أو حساب مقاييس النزعة المركزية لها؛ أو مقاييس التشتت، وكذلك حساب معاملي الالتواء والتفرطح لمنحناها.

وفي الفقرة السابقة تم التطرق لظاهرتين وصفيتين تخص مفردة إحصائية واحدة، وقلنا أن البيانات فيها أزواج من القيم المتناظرة. وتركز الاهتمام على عرض الظاهرتين في جدول تقاطع ثنائي بعدد من الصفوف والأعمدة يعتمد عدد كل منها على عدد أصناف كل ظاهرة لوحدها، وغرفت بأن الجدول يمكن تسميته بجدول الاقتران.

أما في هذه الوحدة تنقوم بدراسة ظاهرتين القيم فيها أيضاً أزواج متناظرة غير مصنفة ولكنها كمية وليست وصفية،والهدف هنا دراسة العلاقة بين هذه الأزواج المتناظرة لمعرفة ماهية تأثر قيم الظاهرتين ببعضها، بعبارة أخرى هل ثمة علاقة ما موجودة بين الظاهرتين، وعلى سبيل المثال لو كان الاهتمام يتركز فيما إذا كانت هناك علاقة بين وزن وطول الأشخاص، عند ذلك نقوم بأخذ وزن وطول الشخص الأول ثم وزن وطول الشخص الثاتي وهكذا حتى الشخص الأخير ضمن مجموعة ما قد تم تحديدها مسبقا، ثم نقوم بالعمل الإحصائي المناسب والذي سنتطرق إليه لاحقاً لمعرفة إذا كانت هنالك علاقة ما وما قوتها وكيفية اتجاهها.

ومثال آخر على هذا الموضوع لو كنا نبحث عن العلاقة بين للدخل الأسري الشهري وبين ايجار الدار المدفوع، هنا سنقوم بسؤال كل من يستأجر داراً عن قيمة دخله الشهري وكم منه يدفع شهرياً لأجرة الدار، وهذا يعنى أن الاجابتين لكل مستأجر ستكون زوجاً من القيم المتناظرة كلا منها تمثل ظاهرة، وفي كثير من الأحيان نعبر عن المتغيرين (الظاهرتين) بعبارة "متغير ذو بعدين". والأمثلة على نلك كثيرة، فقد يحتاج باحث ما دراسة العلاقة بين عدد الأطفال في الأسرة ومعدل ذكاء الطفل الواحد في الأسرة، في هذه الحال لا بد من جمع البيانات على أساس أن هناك أسر لديها عدد من

الأطفال وهذا متغير نسميه X ، وأن هناك امتحان للذكاء يشارك فية كافة أطفال الأسرة الواحدة وحساب معدل نتائج كل أطفال الأسرة الواحدة ونسميه بالمتغيرية . وهذا يعني أن الباحث ستتوفر لدية أزواج من القيم المتناظرة وعلى النحو الآتي:

XpYj يمثل عدد أطفال الأسرة الأولى ومعدل ذكاء الطفل الواحد في الأسرة الأولى.

2مذل2ت يمثل عدد أطفال الأسرة الثانية ومعدل ذكاء الطفل الواحد في الأسرة الثانية.

3لاو3ت يمثل عدد أطفال الأسرة الثالثة ومعدل ذكاء الطفل الواحد في الأسرة الثالثة.

وهكذاحتى نصل إلى الزوج x_n,y_n الذي يمثل عدد أطفال الأسرة 1 ومعدل ذكاء الطفل الواحد في الأسرة ع.

وبعد جمع البيانات سيكون لدى الباحث سؤال مهم هو: هل العلاقة بين هاتين الظاهرتين خطية أو غير خطية؟ بمعنى آخر عند تعيين الأزواج المتناظرة على المحورين السيتي والصادي هل ستقع على استقامة واحدة أو يمكن أن يمر مستقيم بأكبر قدر ممكن من هذه الأزواج؛ فإذا أمكن ذلك كانت العلاقة خطية وعكس ذلك تكون العلاقة غير شطية •

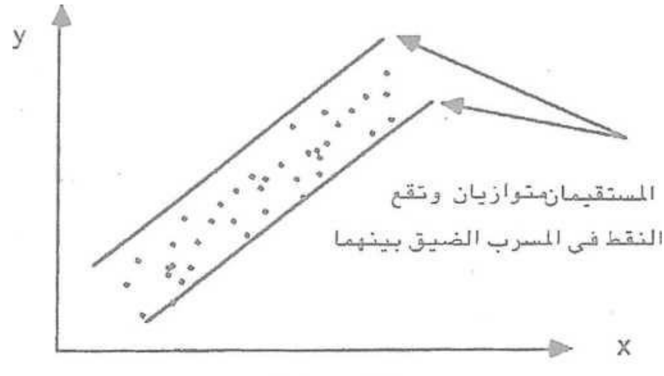
وما سنركز عليه في هذه الفقرة هي الحالة التي تكون فيها العلاقة بين الظاهرتين خطية، كي نستطيع فيما بعد تحديد قوتها واتجاهها. بمعنى هل أن المتغيرين يرتبطان بعلاقة قوية بحيث يؤثر أحدهما على الآخر أم أن العلاقة بينهما ضعيفة أو معدومة. وإذا كان ثمة علاقة فهل هي في نفس الاتجاه أم أنها عكس الاتجاه، وهذا يعني هل أن زيادة قيمة أحدهما ستؤدي إلى زيادة في قيم الآخر (أو أن نقص أحدهما سيؤدي إلى نقص في الآخر). وهذا ما نسميه بنفس الاتجاه، أم أن زيادة أحدهما ستؤدي إلى تقصان الآخر، وهذا ما نسميه بعكس الاتجاه.

ولكي تضع الصررء اف سيجيب لباصئ عطى م الارل اك من غلل:

2.3 شكل الانشار Scatter. Diagram

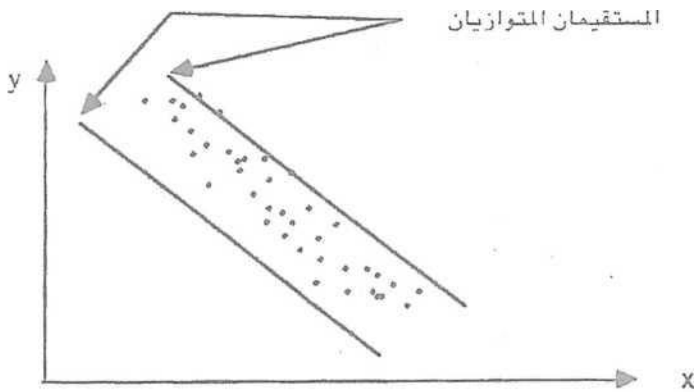
إن التصور الأولي عن العلاقة بين المتغيرين يمكن الاستدلال عليه من تعيين أزواج القيم المتناظرة في المستوى السيني والصادي (المحورين السيني والصادي)، بحيث تمثل قيم أحد المتغيرين على المحور السيني وقيم المتغير الآخر على المحور الصادي، وهذا ما يسمي شكل الانتشار (Scatter Diagram)• حيث أن رسم شكل الانتشار للأزواج المتناظرة يظهر بشكل جيد وجود علاقة بين المتغيرين XjY أو عدم وجودها. فإذا كانت الأزواج المتناظرة واقعة على خط مستقيم أو في مسرب ضيق بين

مستقيمين متوازيين فإن هذا مؤشر على وجود علاقة قوية بين المتغيرين كما في الشكل رقم (1) لآتي:



شكل رقم (1)

وكذلك الحال في الشكل رقم (2) الآتي:

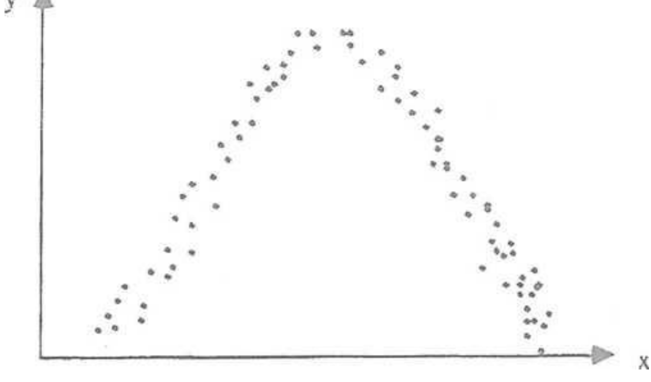


شكل رقم (2)

أما إذا ظهرت الأزواج المتناظرة بصورة مبعثرة في شكل الانتشار، فإن ذلك يدل على عدم وجود علاقة بين المتغيرين كما في الشكل رقم (3) الآتي:

شكل رقم (3)

وقد تظهر الأزواج المتناظرة انتشاراً يختلف عما ظهر في الأشكال السابقة ولكنه يأخذ نمطاً معيناً مما يشير إلى أن العلاقة بين المتغيرين غير خطية والشكل رقم (4) الآتي يبين شكل انتشار لأزواج من القيم العلاقة بينهما تربيعية.



شكل رقم (4)

من ناحية أخرى يمكن الاستدلال من شكل الانتشار (1) السابق أن العلاقة بين المتغيرين طردية، إذيلاحظ كلماتزدادقيمة أحدالمتغيرين تزدادمعهاقيمة المتغير الآخر، بينما في شكل الانتشار شكل (2) تكون العلاقة بين المتغيرين عكسية، حيث يلاحظ كلما ازدادت قيمة أحدهما تقل قيمة الآخر.

مثقفه (2)

البيانات التالية في جدول رقم (1) تبين عدد الأطفال في الأسرة الواحدة ومعدل ذكاء الطفل الواحدآ ٠ ح5

جدول رقم (1)

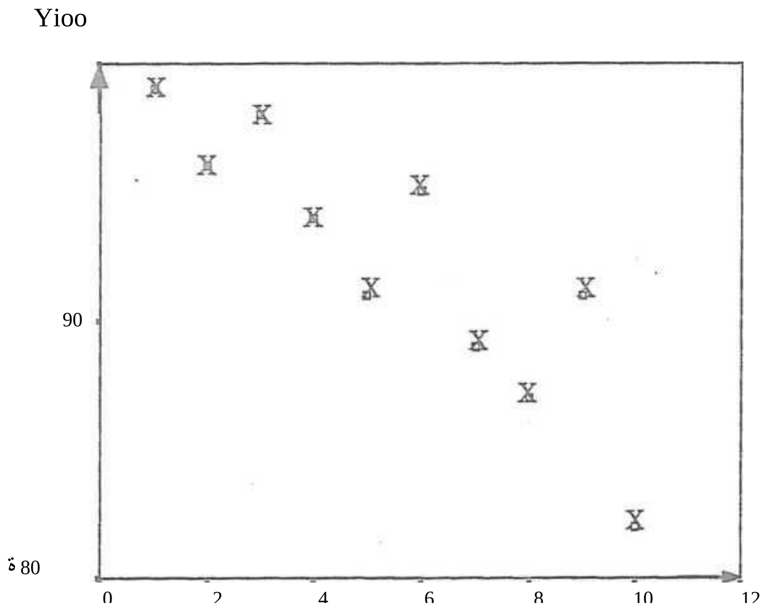
X	1	2	3	4	5	6	7	8	9	10
Y	99	96	98	94	91	95	89	87	91	82

ارسم لوحة الانتشار، واستدل منها على نوع العلاقة بين عدد الأطفال في الأسرة ومحدل ذكاء الطفل فيها.

الحل:

تمثل عدد الأطفال في الأسرة الواحدة على محور السينات ومعدل ذكاء الطفل الواحد على محور الصادرات، ونزصد النقاط العشرة (X, y) فنحصل على لوحة الانتشار كمافي الشكل (5).

سيظهر التقرير النهائي والذي يتضمن شكل الانتشار للمثال على النحو الآتي:



شكل رقم (5)

إن شكل الانتشاز (5) أعلاه هوأقرب مايكون لشكل الانتشار (2) السابق، حيث أن قيم الأزواج المتناظرة قريبة من بعضها، مما يشير إلى وجود علاقة قوية بين عدد الأطفال في

، وكلما زاد عدد الأطفال فيY ومعدل ذكاء الطفل الواحد فيها Xالأسرة

الأسرة انخفض معدل ذكاء الطفل الواحد فيها مما يعني أن العلاقة بينهما عكسية.

تدريب (2)

يبين الجدول أدناه معدلات ثمانية طلاب في الصف الثالث الثانوي (X) ومعدلاتهم التي سجلت في شهادة الدراسة الثانوية (Y).

جدول

X	82.5	66.5	90.3	75.2	60.5	55.8	91.2
Y	85.5	70.1	88.8	76.5	65.0	54.1	90.5

يطلب رسم شكل الانتشار ؟

أسئلة التقويم الذاتي (2)

ماذا نقصد بشكل الانتشار.

3.3 معامل الارتباط الخطي

Linear Correlation Coefficient

في الفقرة السابقة 2.3 من هذا القسم، عرفت بأن أزواج القيم المتناظرة لمتغيرين كميين يمكن أن تكون بينهما علاقة أو قد لا تكون، وأن رسم شكل الاننشار لأزواج القيم المتناظرة سيعطي تصوراً أولياً عن نلك من حيث إذا كانت العلاقة خطية أم غير خطية، وفي هذة الفقرة سنفترض أن العلاقة خطية، حيث أن شكلي (1) و (2) يبينان أن بالإمكان رسم خط مستقيم يمر بأكبر قدر من نقاط الأزواج المتناظرة، مما يعني أن العلاقة بين المتغيرين يمكن تمثيلها بخط مستقيم،وهذا ما سنتطرق إليه في فقرة لاحقة تحت عنوان خط الانحدار. إن الشكلين يبينان بأن العلاقة بين المتغيرين قوية وأن هناك اتجاهاًإما طردياً أو عكسياً.

عزيزي الدارس، قد يتبادر إلى ذهنك سؤال هو: هل بالامكان أن نجد رقماً يساعدنا في تفسير العلاقة بين المتغيرين دون الحاجة إلى رسم شكل الانتشار، والاجابة على ذلك نقول نعم لأن ما بدأنا فيه عند شرح شكل الانتشار هو أنه تصور أولي، لذلك

هنالك مقياس إحصائي يسمى بمعامل الارتباط الخطي البسيط Linear Correlation Coefficient Simple وهو مقياس يقيس قوة واتجاه العلاقة بين متغيرين كميين فقط بينهما علاقة خطية ويرمز لهب* في حالة كون قيم المتغيرين مأخوذة من بيانات العينة.

وإحدى صيغه الحسابية هي:

تنئعئث

حيث E X؛ Yj : حاصل جمع ضرب أزواج القيم المتناظرة.

: حاصل جمع قيم المتغير X.

y. >٦ : حاصل جمع قيم المتغير -

مر٦ز : حاصل جمع مربعات قيم المتغير X.

E y? : حاصل جمع مربعات قيم المتغير Y .

وتسمى هذه المعادلة بمعامل الارتباط بصيغة بيرسون Pearson؛ والقيمة المطلقة لمعامل الارتباط محصورة بين صفر وواحد صحيح؛ فإذا كانت القيمة مساوية للصفر فإن هذا يعني عدم وجود علاقة بين المتغيرين ويستخدم عند ذلك تعبير أن المتغيرين مستقلان، بينما إذا كانت قيمة r مساوية للواحددل ذلك على أن العلاقة بين المتغيرين هي علاقة تامة، وهذا يعني أن الزيادة أو النقصان في قيمة أحد المتغيرين تصاحبها زيادة أو نقصان في قيمة المتغير الآخر بنفس النسبة. وكلما اقتربت قيمة r من الصفر دل ذلك على أن العلاقة بين المتغيرين ضعيفة، وكلما اقتربت من الواحد دل على أن العلاقة بينهما قوية. أما من حيث إشارة r ، فإذا كانت موجبة فإن ذلك يعني أن العلاقة بينهما طردية، أي أن زيادة قيم إحدهماتصاحبه زيادة قيم المتغير الآخر (أو نقصان قيمة أحدهمايصاحبه نقصان في قيمة الآخر). بعبارة أخرى أن قيم المتغيرين تسيران في نفس الاتجاه. وفي حالة كون أشارة r سالبة فهذا معناه أن العلاقة بينهما عكسية، مما نستنتج منه أن قيم المتغيرين تسيران باتجاه متعاكس، أي أن الزيادة في قيم أحدهما يصاحبها نقصان في قيم الآخر.

مثل (3)

احسب معامل الارتباط الخطي البسيط لبيانات المثال (2) الموضوعة في جدول (1) الخاص بعدد الأطفال في الأسرة الواحدة X ومعدل ذكاء الطفل الواحد Y.

جدول (2)

X	y	X*	y2	xy
1	99	1	9801	99
2	96	4	9216	192
3	98	9	9604	294
4	94	16	8836	*376*
5	91	25	8281	455
6	95	36	9025	570
7	89	49	7921	623
8	87	64	7569	696
9	91	81	8281	819
10	82	100	6724	820
55	922	385	85258	4944

حيث يشير جدول (2) بأن

= دلآ;385=لآ;4944=؛^£;922=لآ;55=لآ85258

نقوم بتعويض هذه القيم في صيغة r على النحو الآتي:

- 10*4944-55*922

٧ l • • 385-(55)2 ٧ l • • 85258-(922)2

سهة٧ي•بقني

1= J-2ZL=-O.885

تشير النتيجة الأخيرة إلى أن الارتباط بين المتغيرين اتجاهه عكسي وقوي، أي كلما ازداد عدد الأطفال في الأسرة الواحدة قل معدل نكاء الطفل الواحد، وهي نفس ما تم استنتاجه في رسم شكل الانتشار.

مثةق (4)

) والرياضيات (ل٢).Xيبين الجدول التالي العلامات النيائية لعشرة طلاب في مقرري الفيزياء (

X	90	80	60	70	85	75	40	55	60	80
Y	85	80	55	70	90	70	50	60	65	75

"Y و Xاحسب معامل ارتباط بيرسون الخطي بين المتغيرين

الحل:

نرتب الحل كما في الجدول (3) التالي:

جول03

xv		x^2	y	x
7650	722'5	HOO	85	90
6400	6400	6400	٦٢٠	80
3300	3025	3-00	آ	60
4900	4900	4900	70	7٢
7650	8100	15	90	85
5250	49,00	5625	7o	75
2000	2500	1600	50	40
3300	100	3025	60	55
3900	4225	3600	65	60
6000	5625	6400	75	80
503 0	50500	60475“	700	695

حيث يشير جدول (3) أعلاه إلى أن:

50500 :علال;50475=تآل;50350=؛لاأت£;700=،لال;695 = ،لل

نقوم بتعويض هذه القيم في صيغة r على التحو الآتي:

1050350-695:٤700

٧10*50475-(695)2١ار0ب50500-(700)2

$$\sqrt{n\sum x_i^2-\left(\sum x_i\right)^2}\ \sqrt{n\sum y_i^2-\left(\sum y_i\right)^2}$$

17000 :ل942٠0

i5)fi

تشير النتيجة الأخيرة إلى أن الارتباط بين المتغيرين اتجاهه طردي وقوي جدا،
أي كلما ازدادت قيمة أحد المتغيرين ازدادت قيمة المتغير الآخر والعكس صحيح.

تدريب (3)

> استخدم البيانات الواردة في تدريب (2) من هذه الوحدة، واحسب معامل الارتباط الخطي بين المتغيرين؟

أسئلة التقويم الذاتي (3)

> إذا كان معامل الارتباط بين متغيرين يساوي (0.92)، فما قوة واتجـــاه العلاقـــة بينهما؟

4.3 معامل ارتباط الرتب

Coefficient of Rank Correlation

عرفت في الفقرة السابقة أن معامل الارتباط الخطي (بيرسون) يقيس قوة واتجاه العلاقة بين متغيرين كميين، إلا أنه في الكثير من الأحيان قد يكون المتغيران المدروسان وصفيين، أي لايمكن قياسهما كمياً أو أحدهما لايمكن قياسه كمياً. لذلك نلجأ إلى استخدام الرتب لقيم المتغير الوصفي، بحيث تعطى للقراءات الوصفية رتب بعد أن ترتب القراءات الوصفية تصاعدياً أو تنازلياً وبعدها تعطى للمتخير الآخر رتب بعد ترتيبه أيضاً تصاعدياً أو تنازلياً حسب نوع الترتيب للمتغير الأول، وبشرط أن تكون رتبة القيمة المناظرة للمتغير الثاني هي المقابلة لقيمة رتبة القراءة للمتغير الأول كما هي في الأصل، ويكون ترتيب القيم للمتغيرين بنفس نوع التزتيب؛ وفي حالة وجود قراءات تحمل نفت الصفة فيؤخذ معدل الرتب لها. ويسمى المقياس الاحصائي الذي يقيس العلاقة بين رتب القراءات الوصفية المتتاظرة بمعامل ارتباط الرتب. وهناك عدد من المقاييس الإحصائية التي تقيس الارتباط بين الرتب، من أشهرها معامل ارتباط الرتب لسبيرمان وصيغتة الحسابية:

ي-

حيث dj تمثل الفرق بين رتب المتغيرة ورتب Y المناظرة لها.

و n عدد القراءات

و r_s معامل ارتباط الرتب لسبيرمان.

وقبل عرض مثال لكيفية حساب هذا المعامل، لا بد من الإشارة إلى امكان إستخدامه حتى ولو كانت قراءات المتغيرين كمية، إلا أن من الضرورة التأكيد هنا أن النتيجة التي نحصل عليها في هذه الحالة هي معامل الارتباط بين رتب المتغيرين وليس بين قيمهما الأصلية، كما درست في معامل الارتباط الخطي (بيرسون).

مثال (5)

احسب_معامل_ارتباط_الرتب_لبيانات_المثال_السابق_(4).

الحل:

1- نرتب قيم المتغيرينلة وY اللذين يمثلان العلامات النهائية للفيزياءو العلامات التهائية للرياضيات على التوالي تصاعدياً، كما هو مبين في العمودين الثالث والرابع من الجدول (4):

2- نجدالفرق بين رتب علامتي كل طالب في مقرري الفيزياء(X) والرياضيات (y) كما هو مبين في العمود الخامس من الجدول (4).

3- نجدمربعات الفروق التي تم حسابهافي الخطوة السابقة ونضعهافي العمودالسادس من الجدول (4).

جدول (4)

X	y	رتب x	رتب ۷	الفرق بين الرتب (di)	(d?)
90	85	10	9	1	1
80	80	7.5	8	-0.5	0.25
ا 60	55	3.5	2	1.5	2.25
70	70	5	5.5	-0.5	0.25
85 ؛	90	9	10	-1	1
75)	70	*6*	5.5	0.5	0.25
40	50	1	1	0	0
55 ا	60	2	3	-1	1
60	65	3.5	4	-0.5	0.25
80	*75*	*7.5*	*7*	*0.5*	، aجـ5 ا
؛				—..LJ	6.5

إن القيم تحت عمود رتب X تمثل الرتب التصاعدية للمتغيرلة ، وعلى سبيل المثال القيمة 40 في الصف 7 هي أصغر قيمة للمتغير X وعليه سيكون رتبتها 1 وهذا سنجده عند نفس الصف ولكن تحت عمود رتب X، والقيمة المناظرة ل 40 عند المتغيرآ هي 50 وهي أيضاً أصغر قيمة للمتغير Y وبالتالي تكون رتبتها 1 تحت عمود رتب y.

بينما القيمة الثأنية الأعلى من 40 للمتغير X هي 55 في الصف 8، لذا تكون رتبتها 2 وسنجد هذه القيمة تحت عمود رتب X في نفس الصف بينما القيمة المناظرة ل55 عند المتغير Y هي 60 وهذه القيمة تأتي عند ترتيب قيم Y تصاعدياً بعد القيمة 55 (الموجودة تحت عمود Y في الصف 3) في المرتبة 3 في عمود رتب y وهذا ما نلاحطه في الصف 8.

والقيمة التي تلي 55 للمتغيرة هي 60، وهناك قيمتان ل 60 إحداهافي الصف 3 وا لآخرى في الصف 9 لذا نقوم بحساب المعدل بين الرتبة الثالثة والرابعة وهذا يعني أن نجمع 3 زائداً 4 ونقسم على 2 والنتيجة 3.5 هي رتبة القيمة 60 للمتغير X وسنلاحظ ذلك تحت عمودرتبج عندكل صف X=6O^ ، أمايالنسبة لقيم7 المناظرة X=60،Jفهنالك قيمتين الأولى 55 وهي تأتي بالمرتبة الثانية تصاعدياًلقيم لآلذاسنجد2 تحت عمود رتب (رتب y الصف 3) والثانية 65 وهي تأتي بالمرتبة الرابعة تصاعدياً لقيم Yلذا سنجد 4 تحت عمود (رتب y الصف9) وهكذا لبقية القيم.

4- يشير الجدول (4) أعلاه إلى أن عدد القيم(o 1 دع) ، ومجموع مربعات الفروق Sqdهو 6.5 :دلآفى ، ومجموع قيم الفروق d هو0تهل، ودائماً يكون مجموع الفروق صفر.

5- نعوض بصيغة سبيرمان لحساب معامل إرتباط الرتب، حيث:

»"-ة

مما يعني أن العلاقة بين رتب العلامات النهائية لمقرري الفيرياء والرياضيات طردية وقوية جداً وهو نفس الاستتنتاج الذي توصلنا إليه عند حساب معامل الارتباط الخطي في ألمثال (4.3) والذي كان 0.942 وهي قيمة قريبة، ولكن عليك الانتباه إلى أن معامل سبيرمان هو للارتباط بين الرتب، بينما معامل بيرسون هو للارتباط بين القيم و لاشك في أن الاخير هو الأفضل في حالة كون قيم المتغيرين كمية كما في مثالنا هذا.

كخؤك.ب- (4}

احسب معامل ارتباط الرتب لسبيرمان لمعدات الشهادة الثانوية Y وتقديرات العلامات بعد التخرج من الجامعة X، والخاصة بثمانية طلبة كمامبين أدناه:

مقيول	جيد	متوشط	جيد	امتياز	جيد جدا	متوسط	جيد	X
59	66	70	85	93	85	72	67	Y

آتثلةلااققوببمقفااقيإ4

في أي نوع من المتغيرات يفضل استخدام معامل ارتباط الرتب لسبيرمان

4.الاثحدارالخطي اكبق Regression٠٠٠(|

1.4 مفهوم الانحدار Concept of Regression

عزيزي الدارس، في الفقرات السابقة من هذه الوحدة درست مفيوم الارتباط والارتباط الخطي، وعرفت كيف تحتب قيمة معامل الارتباط الخطي بين متغيرين، وتلمت كيفية قياس قوة العلاقة الخطية بينهما، واتجاه هذه العلاقة سواءاً كانت طردية أم عكسية، إضافة إلى دراستك لشكل الانتشار والذي من خلاله يكون التشخيص الأولي للعلاقة بين المتغيرين.

وفي هذا البند نحاول الإجابة على مجموعة من الأسئلة في مقدمتها:

إذا كانت العلاقة بين المتغيرين هي خطية فكيف يمكن أن نعبر عنها بمعادلة خط مستقيم؟ وما مقدار تأثير إحدى المتغيرين على الآخر؟ وهل بالامكان التتبوء بقيمة أحدهما إذا عرفنا قيمة الآخر.

تتطلب الإجابه عن الأسئلة أعلاه، أن نحدد في البداية المتغير المستقل Independent Variable، المتغير التابع (المعتمد) Dependent Variable، وهذا يعني أنناهنانبحث عن تأثرالمتغيرالمعتمدبقيم المتغيرالمستقل، ولتوضيح ذلك نذكرعلى سبيل المثال، الاعتقاد بأن معدل الطالب في السنة الجامعية الأولى يتأثر بمعدله في امتحان الثانوية العامة؛ أي قبل التحاقه بالجامعة، هنايقال لمتغير معدل الطالب في السنة الجامعية الأولى بأنه متغير تابع (أو معتمد) على معدله في امتحان الثانوية العامة الذي سيكون هنا هو المتغير المستقل. كما يمكن توضيح ذلك لحالة عرضت سابقاً في هذه الوحدة، وهي أن معدل ذكاء الطفل الواحد في الأسرة يتأثر بعدد الأطفال في هذه الأسرة، أي أننا هنا قد اعتبرنا أن معدل ذكاء الطفل الواحد هو المتغير التابع وأن عدد الأطفال في الأسرة هو المتغير المستقل.

إن من شأن هذا التحديد المسبق في معرفة المتغيرين أيهما مستقل وأيهما تابع أن يسهل علينا الاجابة على مجموعة الأسئلة من خلال ما يلي:

ب2.4 معادلة أنحدإر متغير على متغير آخر

Simple Regression Equation

سنقتصر هنا في تناول الموضوع على دراسة متغيرين أحدهما مستقل والآخر تابع، فإذا كان لديك عينة من الأزواج المتناظرة:

(X ه) ....وح3لاو3ت) ’(2مد و 2x) ’ز 1٢ؤ و Xl)

حيث X....... X XjX هي قيم المتغير المستقل X المشاهدة (الفعلية)، وكذلك yi,y2,y3i・・・’・・・yn هي قيم المتغير التابع Y المشاهدة (الفعلية).

إن رسم شكل الانتشار لأزواج القيم المتناظرة سيساعد في الحكم الأولي فيما إذا كان من المعقول تمثيل خط مستقيم على شكل الانتشار أم لا، خاصة وأن هذه الأزواج من القيم لاتقع على استقامة واحدة بشكل عام. فإذا أظهرشكل الانتشار إمكانية تطبيق خط مستقيم يمر بأكبر قدر ممكن من نقاط الأزواج فإننا نفترض بأن معادلة هذا الخط المستقيم ?Y = c(.-rj3x، حيثY تمثل القيمة التقديرية للمتغير التابع Y التي تقابل قيمة المتغير المستقل X والتي تقع على خط الانحدار (المستقيم).

وتمثل a قيمة الحد الثابت في معادلة خط الانحدار، وهي المسافة بين تقاطع الخط أو امتداده مع محور الصادات، وبين نقطة الأصل.

وتمثل 3) ميل الخط المستقيم.

وتمثل X قيمة المشاهدة (الفعلية) للمتغير المستقل.

وهذا يعني أن لكل قيمة من قيم المتغير المستقل هناك قيمتان للمتغير التابع إحداها هي قيمة Y والتي تقع على خط الانحدار، والثانية القيمة الفعلية للمتخير التابع Yالتي قد تقع على خط الانحدار أو خارج الخظ (أعلى أو أسفل). وهذا يعني أن هناك فرق بين القيمة التقديرية والقيمة المشاهدة (أو الفعلية) Y ، فإذا ما رمزنا لهذا الفرق بالحرف ٥،فإنقيمةهذا لفرق هو: e=Y-Y

ويسمى هذا المقدار بالخطأ، مما يعني أن كل قيمة من قيم المتغير المستقل سيقابلها خطأ في تقدير إحدى قيم المتغير التابع. ومن المنطقي أن يكون خط الانحدار (المستقيم) المطلوب هو المستقيم للذي يمر بأكبر عدد ممكن من نقاط شكل الاتتشار (القيم الفعلية للمتغيرين المستقل والتابع). وثمة عددمن الطرق الإحصائية لعمل ذلك، من أهمها طريقة المربعات الصغرى التي تتلخص فكرتهافي حساب قيم الحد الثابت a ، والميل لخط الانحدار P بحيث تجعل مجموع مربعات قيم الأخطاء أصغر ما يمكن

جيؤ2=> minimize

صدن(!) ________________________________ ج

طريقة المربعات الصغرى هي طريقة تطبيق خط انحدار (مستقيم) على مجموعة من النقاط

(x!, yt)١ (x2, y2)j (x3,y3),................

بحيث يكون مجموع مربعات الأخطاء أصغر ما يمكن، أي أن:

$$Ee^2=E(Y-Y)^{2}\text{^ min imize}$$

والمسالة الآن هي كينية حاب تتديرات معالم خط الانحدار، بعبارة أخرى حساب كلاً من الميل P والحد الثابت ة بحيث تكون minimize جت $X(Y-Y)^2$ = 2حق . ثم بعد استخراجهما يتم تعويضهما في المعادلة للحصول على خط انحدار المتغير التابع Y على المتغير المستقل X بطريقة المربعات الصغرى.

إن إحدى الطرق الرياضية لحساب الميل والحد الثابت تتضمن أن تكون معادلة الخط المستقيم هي:

$$Y=a+px$$

وبأخذ المجموع للطرفين نحصل على:

(1) .... IY=na + pZX

وبضرب معادلة الخط المستقيم بالمتغير المستقل X وأخذ المجموع لملطرفين نحصل على:

(2) .... £xY=a£x-t-^x

وبحل هاتين المعادلتين نحصل على:

2(٠لاقع

وبعد حصولنا على قيمة الميل نستخرج قيمة الحد الثابت وفق العلاقة الآتية:

$$Y\text{-}X$$

حيث X تمثل الوسط الحسابي لمشاهدات المتغير المستقل X؛

Y وتمثل الوسط الحسابي لمشاهدات المتغير التابع Y.

وبعد حساب قيمة الميل والحد الثابت نعوض قيمهما المستخرجة في المعادلة X-؟=؛ فنحصل على معادلة انحدار المتغير التابع Y على المتغير المستقلة.

هفةقيأ1)

لبيانات المثال (2) الموضوعة في جدول (1) التي تبين عدد الأطفال في الأسرة الواحدة X ومعدل ذكاء الطفل الواحد Y. يطلب:

1- حساب معاملات خط انحدار معدل ذكاء الطفل الواحد على عدد الأطفال في الأسرة الواحدة.

2- كتابة معادلة خط الانحدار بشكلها النهائي.

الحل:

نلاحظ من مراجعة شكل الانتشار لهذه البيانات في المثال (2) بأن العلاقة بين متغيري معدل الذكاء وعدد الأطفال في الأسرة قوية وخطية. وأن خط الاتحدار المراد تقدير معاملاته والذي سيمر بأكبر قدر من النقاط سيكون ميله سالب وسيقطع المحور الصادي في الاتجاه الموجب، مما يعني أن اشارة الحد الثابت ستكون موجبة.

ومن حل المثال السابق لدينا النتائج التالية:

5258§=2لآ;385=دل;4944م7لآ;922مرل;55بتل

بالتعويض بمعادلة الميل 3ا :

539 ا:1270- 922*55-4«4ر0ا ¥ك?-¥تلعح

p- n£X^2-(Zx)2 385٠0-(55)2 - 825ا

لاحظ أن إشارة الميل سالبة، وهذا يطابق ما تم استنتاجه من شكل الاتتشار. وبتعويض بقيمة فيمعالة الحدلثابت ينتج:

100.667 = (ج٠(1.539-)-أ = عت6-؟ = ى

10 10

لاحظ أن اشارة الحد الثابت موجبة، وهذا يطابق ما تم استنتاجه من شكل الانتشار.

وعليه تكون معادلة خط الاتحدا ر بشكلها النهائي كالآتي:

Y=100.667-1.539X

مثال (12)

في ضوء نتائج المثال السابق جد:

1- القيم التقديرية للمتغير التابع.

2- قيم الأخطاء.

الحل:

لحساب القيم التقديرية للمتغير التابع، بعبارة أخرى لحساب قيم ي، لابد من تعويض قيم X في كل مرة بمعادلة خط الانحدار المستخرجة وهي: ¥=100.667-1.539X، ثم نحسب الفرق بين القيم المشاهدة أو الفعلية Y والقيم التقديرية (التنبؤية) Y والنتائج هي قيم الأخطاء ج.

وستكون النتائج مع بيانات المثال على النحو الآتي:

X	Y	Y 100.667-1.539X	e=Y-Y
1	99	99.128	-0.128
2	96	97.589	-1.589
3	98	96؛050	1.950
4	94	94.511	-0.511
5	91	92.972	-1.972
6	95	[illegible]	3.567
7	89	89.894	-0.894
8	87	88.355	-1.355
9	91	86.816	4.184
10	82	85.277	-3.277

ظل(ا)3 8 Fl

يبين الجدول الآتي أوزان § طلاب بالكغم (Y) والأطوال المقابلة لهم بالسنتمتر

جدول رقم (13)

Y	67	56	75	68	64	70	78	72
X	168	156	176	170	165	172	180	167

يطلب ما يلي:

1) ايجاد معادلة انحدار الوزن (Y- على الطول (X")

2- تقدير قيمة وزن الطالب عندما يكون طوله 172.

3- تقدير قيمة الخطألوزن الطالب عندمايكون طوله 172,

4- حساب معامل ارتباط بيرسون.

الحل:

1- لايجاد معادلة خط انحدار y على x نحسب المقادير المطلوبة كما في الجدول (14) • جدول (14)

XY	V²	X²	Y	X
12024	5184	27889	72	167
14040	6084	32400	78	180
12040	4900	29584	70	172
10560	4096	27225	64	165
1560	4624	28900	68	17(-)
13200	5625	30976	75	176
8736	3136	24336	56	156
11256	4489	28224	67	168
93416	38138	229534	550	1354

حيث يشير الجدول (14) إلى أن:

£x؛ =1354;ألال =550؛Exj y؛ =93416;تل' = 229534;للائ? =38138

نعوض بمعادلة الميل 5) :

889 0حح§262 1354ت550ك93416م8 ¥ختد-¥تلع م

889، 2956’, 2(354)-8*229534 2(تل)-2تلع -P

وبتعويض بقيمة ٠3ا في معادلة الحد الثابت ة نحصل على:

&=¥-3خق=آ-(0.889)٠(آ)=81.71

وعليه تكون معالة خط الانحدار هي:

Y=-81.71 + 0.889X

2- لتقدير قيمة الوزن عندما يكون الطول 172 نعوض بقيمة 172 =7 في معادلة خط الانحدار فتحصل على القيمة:

؛=-81.71+0.889*172=71.198

3- ولحساب قيمة الخطأ لوزن الطالب عندما يكون طوله 172، نعلم من بيانات المثال أن الوزن الفعلي للطالب الذي طوله 172 سنتمتر هو 70 كغم وهذا يعني أن Y=7O، وبذلك يكون الخطأ:

e=Y-Y=70-71.198=-1.198

4- نقوم بالتتويض في صيغة] لحساب معامل ارتباط بيرسون على النحو الآتي:

8*93416-550*1354

تئخ١ت٠أ٧ سة)حبجتب٠لم

r: . 26.28 Q.947

n

وتشير النتيجة الأخيرة إلى أن الارتباط بين المتغيرين اتجاهه طردي وقوي، أي كما اردل طول الطالب اردلد وزنه

3.4 تفسير معامل الإنحدار

Interpretation of Regression Coefficient

نكرنا سابقاً أنه إذا كان المطلوب التنبؤ بقيمة المتغير التابع لآ من معرفة القيمة المناظرة لها والتي تخص المتغير المستقل X ، وكانت العلاقة بين المتغيرين خطية، فكل ما علينا عمله هو إيجاد خط انحدار المتغير التابع Y على المتغير المستقل X بواسطة طريقة المربعات الصغرى؛ حيث أن معادلة خط الانحدار هي:

Y X

حيث نجدفي البداية الميل P ثم نجدالحدالثابت ة ويطلق عليهمامعاملات أو (معالم) الانحدار. وعند دراسة هذه المعادلة والتي هي كما عرفنا معادلة خط مستقيم، فإن قيمة ستساوي قيمة آ عندماتكون x=o٠

وبالتالي فإن قيمة ة هي المسافة بين تقاطع خط المستقيم مع المحور الصادي (حيث هنا قيمة x=o) ونقطة الأصل.

أما P فيمثل كما تعلمت ميل خط الانحدار، وميل المستقيم هو مقدار التغير في الارتفاع Y عندما نتحرك إلى اليمين بمقدار وحدة واحدة باتجاه X . وهذا يعني بالنسبة لخط الاتحدار أن قيمة P هي مقدار التغير في Y الذي يرافق تغيراً مقداره وحدة واحدة في X,

ومن هنا يمكن أن نفسر الميل في مثال (12) حيث 1.539--م، بأن زيادة طفل في الأسرة سيؤدي إلى نقصان في معدل نكا ء الطفل في الأسرة الواحدة بمقدار 1.539.

ومن هنا، إذاكانت 3، سالبة فذلك يعني أن الزيادة في قيمة المتغير المستقل X

يرافقها نقصان في قيمة المتغير التابع Y، وبالتالي تكون العلاقة بين المتغيرين عكسية. أما إذا كانت P موجبة فذلك يعني أن الزيادة في قيمة المتغير المستقل X يرافقها زيادة في قيمة المتغير التابع Y، وبالتالي تكون العلاقة بين المتغيرين طردية. ونحن في هذه الحالة نتكلم عن الارتباط بين المتغيرين ممايشير إلى امكانية حساب معامل الارتباط بمعرفة قيمة الميل وفق العلاقة الآتية:

حيث أن s Sy هما الانحراف المعياري للمتغيرين Y و X على التوالي، و لأن قيمة الانحراف المعياري تكون دائما موجبة لكونها تمثل بعد القيم عن وسطها الحسابي كما مر معك سابقا، عند ذلك يمكن القول أن اشارة P هي نفسها أشارة r ؛ فإذا كانت اشارة الميل ؤإ موجبة فهذا يعني أن معامل الارتباط بين المتغيرين سيكون موجباً وأن العلاقة بين المتغيرين طردية، و إذا كانت اشارة الميل P سالبة فهذا يعني أن معامل الارتباط بين المتخيرين سيكون سالباً، وأن العلاقة بين المتغيرين عكسية. والعكس

صحي.

تهويب (5)

يبين الجدول الآتي العلامات النهائية لتسعة طلاب في مقرر الإحصاء (X) ومقرر الاقتصاد (Y)، يطلب:

1- ايجاد معادلة انحدار العلامات النهائية لمقرر الاقتصاد على العلامات النهائية لمقرر **الإحصاء.**

2- إذا كانت علامة أحد الطلبة في الإحصاء 82 فكم تقدر علامته في الاقتصاد ؟

3- ما تقدير الخطأ الذي حصلت عليه في البند (2) ؟

X	91	87	55	60	77	90	65	82	75
Y	85	90	ه	50	80	90	5ؤ	هة	6 ؤ

جححت:

ؤئتك (10) ______________________________ققي

1- ما هي معادلة الانحدار الخطي؟

2- اشرح طريقة المربعات الصغرى.

3- بين العلاقة بين ميل خط الانحدار لمتغيرين واتجاة العلاقة بينهما.

تتناول هذه الوحدة كما درستها الحلاقة بين الظواهر المختلفة سواء كانت كمية أو وصفية أو كلا منها مصنف إلى أصناف. وكذلك معرفة معالم خط المستقيم وكيفية تقديره عندما تكون للعلاقة بين الظاهرتين قوية وفيما يلي نسرد الصيغ الخاصة والتي تم اتتخدامها للوحدة.

- معامل الارتباط الخطي البسيط لبيرسون: ويقيس قوة واتجاه العلاقة بين متغيرين كميين.

يجهل يكيتجهلى:

- معامل ارتباط الرتب لسبيرمان: يقيس قوة واتجاه العلاقة بين رتب متغيرين على الأقل ا أحدها وصفية.

وُس■': ؛

- ميل خط الانحدار: هو ظل الزاوية التي يقطعها المستقيم مع المحور السيني أو ما يوازيه.

حفئتكعدج

إ(٠ ^{2}تةه

- الحد الثابت لخط الانحدار: بعد نقطة تقاطع المستقيم مع المحور الصادي عن نقطة : ٥٠ ـخط

الاتحت١٠ال_إل٠قدر:_خظ_إسهيم_م_يبس_ثدقة_إنخطينبين_سنفيد ت١بع ;سنق. ا

6. لمحة عن الوحدة الدراسية ائر ابحة

عزيزي الدارس، ستتعرف في الوحدة التالية على تعريف الاحتمال الذي يعتمد على دراسة المجموعات وفضاء العينة وفرضيات الاحتمال وكذلك ستتم دراسة الاحتمال الشرطي والاستقلال بالاضافة إلى المتغيرات العشوائية من حيث تعريفها وتوقعها وتباينها. وسوف تتعرف أيضاً إلى التوزيع الاحتمالي المنفصل وستدرس أحد أهم ! التوزيعات المنفصلة وهو التوزيع ذات الحدين.

7. إجابات التدريبات قوويج (1)

نقوم بتفريغ البيانات لنحصل على الجدول التالي:

y x	حاسوب	إدارة	تزبية
ذكر	HU	UH	اا
أنثى	ا	III	HH ؛؛

وبايجاد التكرار في كل خلية وجمع قيم الخلايا لكل صف وكل عمود يصبح الجدول كما يلي:

y x	حاسوب	إدارة	تربية	المجموع
ذكر	5	5	3	13
أنثى	1	3	7	11
المجموع	6	8	10	24

قهويب (2)

نمثل معدلات الطلاب في الصف الثالث الثانوي على محور السينات ومعداتهم في شهادة الدراسة الثانوية على محور الصادات ونرصد النقاط الثمانية فنحجل على لوحة الانتشار كما في الشكل التالي:

100 £---
Y

go ٥٥

SO

70 0

SO

د

50 60 70 so so 100

إن شكل الانتشار أعلاه هو أقرب ما يكون لشكل الانتشار رقم (1) السابق، حيث أن قيم الأزواج المتناظرة قريبة من بعضها مما يشير إلى وجود علاقة قوية بين معدل الطالب في الصف الثالث الثانوي X ومعدله في شهادة الدراسة الثانوية بآ، وكلما كان معدله في الصف الثالث أعلى ارتفع معدله في شهادة الدراسة الثانوية، وهذا يعني أن العلاقة بينهما طردية.

ثهويب **(3)**
نرتب الحل كما فى الجدول التالى:

xy	y^2	X^2	y	X
7053.75	7310.25	6806.25	85.5	82.5
4661.65	4914.01	4422.25	70.1	66.5
8018.64	7885.44	8154.09	88.8	90.3
5752.8	5852.25	5655.04	76.5	75.2
3932.5	4225.0	3660.25	65.0	60.5
3018.78	2926.81	3113.64	54.1	55.8
8253.6	8190.25	8317.44	90.5	91.2
40691.72	41304.01	40128.96	530.5	522

نجد من الجدول أعلاه أن

= اتإلتة;41304.01 اتل;40128.96=متة;530.5=للاي522=؛تتق40691.72

؛xy-JxXy

ت(530.5)-41304.01*7يت(522)-40128.96*7٧

7*40691.72-(522)(530.5)

= 0.98

وهناتجد قيمة معامل الارتباط مساوية ل 0.984، مما يدل على أن العلاقة طردية وقوية جداً بين معدل الطالب في الصف الثالث الثانوي X ومعدله في شهادة الدراسة الثانوية Y، وهي نفس المؤشرات التي حصلت عليها من شكل الانتشار للتدريب السابق.

تدريب إ(4)

نرتبالحلكما في لجدول لتالي:

d	(di)	رتب x	رتبل7	y	x
4	+2	5	3	67	جيد
6.25	-2.5	2.5	5	72	متوسط
0.25	+0.5	7	6.5	85	جيد* حدا
0	0	8	8	93	امتياز
2.25	-1.5	5	6.5	85	جيد
2.25	-1.5	2.5	4	70	متوسط
9	+3	.5	2	66	جيد*
0	0	1	1	59	مقبول
24					

وبالتعويض بصيغة سبيرمان نحصل على:

؛لاق6 0.714عل-ا:

(٩ة)8ا'' لاها1٠ذل 504

والذي يشير إلى أن العلاقة بين رتب معدلات الشهادة الثانوية X وتقديرات العلامات بعد التخرج من الجامعةل٦ والخاصة بهؤلاء الطلبة طردية وقوية إلى حد ما.

قهوتبي (5)

نكون الجدول التالي:x على yلايجاد معادلة خط انحدار

xv	٢٠	x^	y	x
7735	7225	8281	85	91
7830	8100	H	90	87
2750	2500	3025	50	55
3000	o	3600	50	60
6160	6400	5929	80	77
§100	8100	8100	90	90
3575	3025	4225	55	65
6970	7225	6724	85	82
5175	4761	5625	69	75
51295	49836	53078	654	682

حيث يشير الجدول إلى أن:

z X؛ = 682; z Yi: 654; Zˣi Yi ت51295; z دد :53078; z ت 49836

نعوض بمعادلة الميل 3] :

ك(682)ع ,12578 £ HZXY~ZXZY 9*51295-682*654 15627

p- nZX²-(Zx)² I 9٠53078

وبتعويض قيمة P في معادلة الحد الثابت a ينتج: a=Y-pX=f-(1.24)*m=-21.30

وعليه تكون معادلة خط الانحدار بشكلها النهائي كالآتي:

Y 21.30 + 1.24X

2- لتقدير قيمة علامة الاقتصاد عندما تكون علامة الإجصاء 82، نعوض بقيمة 82= xفي معادلة الانحدار فينتج:

؟=--21.30+1.24*82=80.38

3- ولحساب قيمة الخطأ في تقدير علامة الطالب في الاقتصاد عندما تكون علامته في الإحصاء 82، نعلم من بيانات المثال أن علامة الطالب في الاقتصاد 85 عندما تكون علامتة في الإحصاء هي 82، وهذا يعني أن Y=85، وعليه يكون الخطأ:

e=YY 85-80.38=4.62

28مفبرد المصطلحات

- الانحدار الخطي Linear Regression : معادلة خط مستقيم تمثل العلاقة بين المتغيرين المستقل والتابع (للثابت) يتم فيها تقدير ميل المستقيم والحد المطلق.
- جدول التقاطع الثنائي لمتغيرين نوعيين Bivairate Qualitative Cross Table: جداول تتضمن أصناف متغيرين، بحيث تمثل الصفوف أصناف المتغير الأول والأعمدة أصناف المتغيرالثاني، وكل خلية تمثل التكرارالمشاهدلعددالحالات المشتركة بين صنفين من أصناف المتغيرين.
- شكل الانتشار Scatter Diagram: لوحة يتم فيها رصد مجموعة الأزواج المتناظرة من قيم المتغيرين تحت الدراسة وتعطي تصوراً أولياً فيما إذا كانت العلاقة بين المتغيرين خطية أم غير خطية وقوة هذه العلاقة واتجاهها.
- معامل الارتباط الخطي Linear Coefficient of Correlation: مقياس إحصائي يقيس قوة و اتجاه العلاقة بين متغيرين كميين فقط بينهما علاقة خطية، ويسمى بمعامل ارتباط العزوم ل بيرسون.
- معامل ارتباط الرتي Coefficient of Rank Correlation: مقياس إحصائي يقيس قوة واتجاه العلاقة بين رتب متغيرين كميين أو على الأقل أحدهما وصفي، وقد استخدم في هذه الوحدة معامل ارتباط سبيرمان.

9. المراجع

أ- المراجع العربية:

1. العتوم، شفيق؛ العاروري، فتحي، الأساليب الإحصائية: الجزء الأول. عمان: دار المناهج: (1982).
2. القاضي، دلال؛ عبداله، سهيلة؛ البياتي، محمود. الإحصاء للاداريين والاقتصاديين. عمان: دار الحامد للنشر والتوزيع، (2004).
3. أبو صالح، محمد، الطرق الإحصائية: الإصدار الثاني. عمان: دار اليازوري العلمية للنشر والتوزيع، (2001).
4. أبو صالح، محمد، الموجز في الطرق الإحصائية: الطبعة العربية الأولى. عمان: دار اليازوري العلمية للنشر والتوزيع، (2002).
5. شبيجل، موراي، ملخصات شوم: نظريات ومسائل في الإحصاء، دار ماكجرو هيل للنشر، الطبعة العربية، 1978.

ب- المراجع الأجنبية:

1- Daniel, W.W., Biostatistics: A Foundation for Analysis in the Health Sciences, 5th ed., John Wiley & Sons, 1991،
2- Walpole, R.E., Introduction to Statistics, 3rd ed. MacMillan Publishing Co. ,Inc., 1990،

الو ذك'.حب. ي'٦..د.ر٠

الإحصاء الاستنتاج

IHHg التوزيع الطبيعي

يوصف التوزيع الطبيعي بأنه اقتران رياضي احتمالي لمتغير عشوائي مستمر، والصيغة الرياضية له هي:

f(x)= 7لمe-رجتم ;-٠0هبكحك0

حيث: Xدج الوسط الحسابي للمتغير العشوائي المستمر X.

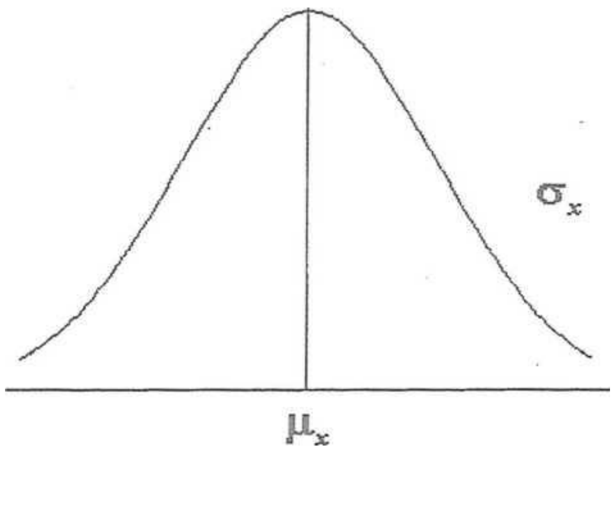
تباين المتغير العشوائي المستمر X،

*٠ الانحراف المعياري للمتغير العشوائي المستمر X (وهو الجذر التربيعي الموجب للتباين).

2.71828 -٥: العدد لنيبيري.

3.14285 = TC النسبة التقريبية.

يسمى منحنى الاقتران أعلاه بمنحنى التوزيع الطبيعي، ويشبه شكله ناقوس الجرس (Bell Shaped)، وعند فحص هذا الاقتران يلاحظ وجود عدد من الثوابت ومتغير واحد هو X ، وبعض هذه الثوابت لها قيم معروفة مثل العدد النيبيري و النسبة التقريبية، أما المتوسط ,رلإ والتباين ئ٠ فهما من الثوابت ولكن قيمتيهما تتغيران بتغير البيأنات التي يحسبان منها، كما تعلمنا في الوحدة الأولى من هذا المقرر، لذا يطلق عليهما بمعلمتي التوزيع الطبيعى. ويبين الشكل رقم (اك الآتى منحنى التوزيع الطبيعي



شكل رقم (1)

بمهرا اسور "سسيي ي "س /٦٨ يم "ير "ى "ي دد "يا يبع اضئري الطبيعي، بينما يمثل المحور الصادي التكرار النسبي لكل قيمة من قيم X، والمنحنى هو منحنى التوزيع الطبيعي، ولو انزلنا عموداً وهمياً من قمة هذا المنحنى على المحور السيني فسيقطعه في نقطة، ولا شك أن القيمة التي تقابل أعلى تكرار نسبي هي المنوال، فتكون نقطة التقاطع هي المنوال، وبما أن بيانات المتغير X مرتبة تصاعدياً لأنها واقعة على المحور السيني، فإن نقطة تقاطع العمود الوهمي مع المحور السيني ستقسم هذه البيانات إلى قسمين متساويين بحيث يكون عدد القيم على يسار نقطة التقاطع يساوي عدد القيم على يمينها. ومن تعريف الوسيط فإن نقطة التقاطع هذه هي الوسيط؛ وهذه النقطة هي كذلك المتوسط الحسابي ج.1إ. ومن دراستنا لمقاييس النزعة المركزية والتشتت والالتواء، نستنتج أن قيمة معامل الالتواء لمنحنى التوزيع الطبيعي تساوي صفر. وبعبارة أخرى أن المنحنى متماثل حول وسطه الحسابي، والانحرلف المعياري ٩ هو أحد مقاييس التشتت ويقيس بعد القيم عن وسطها الحسابي ويمثل هنا شكا، المتحنى. أما ذيلي المنحنى (Tails) فيقتربان من الصفر على الجهتين عندما تقترب قيمة X إلى سالب ما لانهاية؛ وكذلك عندما تقترب قيمة X إلى موجب ما لانهاية.

خصائص التوزيع الطبيعي

Properties of Normal Distribution

وفي ضوء ما تقدم يمكن تلخيص خواص التوزيع الطبيعي بما يلي:

1- متحنى التوزيع الطبيعي متماثل حول العمود المقام على الوسط الحسابي جلل.

2- شكل منحنى التوزيع الطبيعي يشبه شكل الناقوس أو الجرس.

3- في التوزيع الطبيعي مقاييس النزعة المركزية الثلاث، الوسط الحسابي والوسيط والمنوال، متساوية؛ وهذا يعني أن لمنحناه قمة واحدة فقط.

4- المساحة تحت المنحنى وفوق المحور السيني تساوي (1)، لأنها تمثل مجموع التكرارات النسبي.

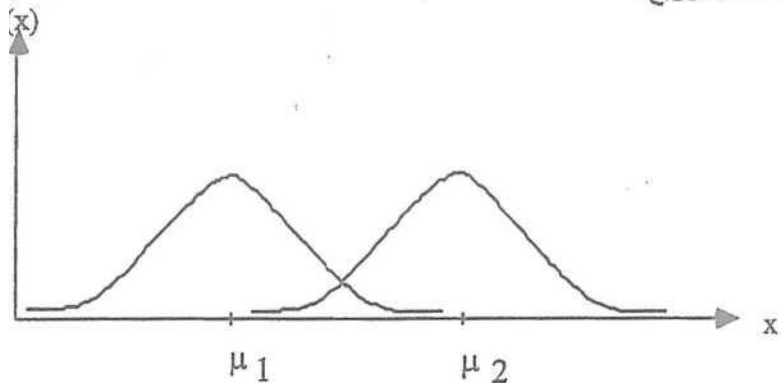
5- يتقارب ذيلي منحنى التوزيع الطبيعي من الصفر، ولكن لاينطبقا على المحور السيني عندما 00— <—X و . X4 + CO

6- هناك معلمتان للتوزيع هما *لمز: المتوسط الحسابي و ٠؛ : التباين لا بد من معرفتهما لنتمكن من رسم المنحنى بشكل تقريبي، فإذا انتقل المتوسط يل! الذي يمثل مركز التوزيع نحو اليمين أو نحو اليسار وبقي الانحراف المعياري ٩ ثابت فسينتقل مركز التوزيع تبعاً لانتقال المتوسط ويبقى شكل المنحنى كما هو دون تغير، أما إذا تغيرب

قيمة الانحراف المعياري ٠x وبقيت قيمة المتوسط يلر ثابتة دون تغير فإن مركز التوزيع لن يتغير وإنما يتغير شكل المنحنى، فينضغط كلما صغرت قيمة ٥x ويتوسع كلما زادت قيمة ي. وإذا تغير الاثنان فإن مركز التوزيع وشكل المنحنى يتغير كذلك.

7- التوزيع الطبيعي مجموعة من العوائل وليس عائلة واحدة، ولتوضيح ذلك لنفترض أن لدينا متغير عشوائي طبيعي (,*)له وسط حسابي رل! وتباينه 2٥ ، وكذلك لدينا متغير عشوائي طبيعي آخر (و*) له وسط حسابي 2لإ وتباينه ؛٠، لنناقش الحالات التالية: أ- عندما يكون $_2$لمإيالإ ولكن 2٠=ل٠٠

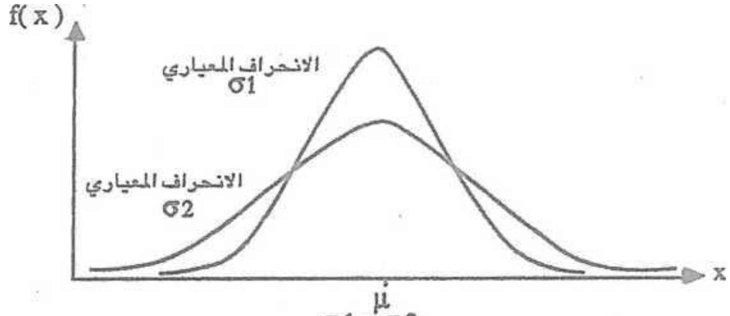
وكمثال لتوضيح ذلك، انظر الشكل رقم (2) عندما 2لإ > ،لإ، إذ تلاحظ أن تباعد كل منحنى عن مركز توزيعة (شكل التوزيع) بقي دون تغير على الرغم من اختلاف متوسط كل توزيع.



شكل رقم (2)

ب- عندما يكون ول!-1ل1 ولكن 2٠*اه.

وكمثال لتوضيح ذلك، تمعن وتدبر الشكل رقم (3) عندما 2أَ>١٠، حيث تلاحظ أن تباعد كل منجنى عن مركز توزيعة (شكل التوزيع) تغير على الرغم من تساوي متوسطي التوزيعين.



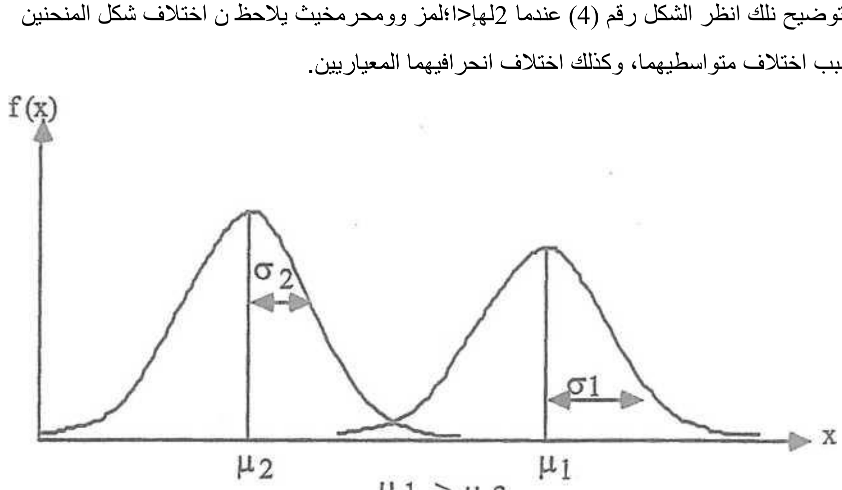
$\sigma 1 < \sigma 2$

)3شكل رقم (

ج-عندمايكون ولإمرل! ولكن 2تءج٩ ٠

ولتوضيح نلك انظر الشكل رقم (4) عندما 2لهإدا؛لمز وومحرمخيث يلاحظ ن اختلاف شكل المنحنين

بسبب اختلاف متواسطيهما، وكذلك اختلاف انحرافيهما المعياريين.



$\mu_1 > \mu_2$

$\sigma_1 > \sigma_2$

شكل رقم (4)

-8 هناك نسب معينة من المساحة تقع ضمن عدد من الانحرافات المعيارية عن الوسط

الحسابي ومنها ما يلي:

المساحة ضمن انحراف معياري واحد عن الوسط ٠

المساحة الواقعة على الفترة (٩اب Mx؛x٠ا-؛ىدا)=68.26% من المساحة الكلية.

المساحة ضمن انحراف معياري واحد ونصف عن الوسط ه

المساجة الواقعة على الفترة (٥٩٠ب p_x;o_x5٠ا-للنت86.64% من المساحة الكلية.

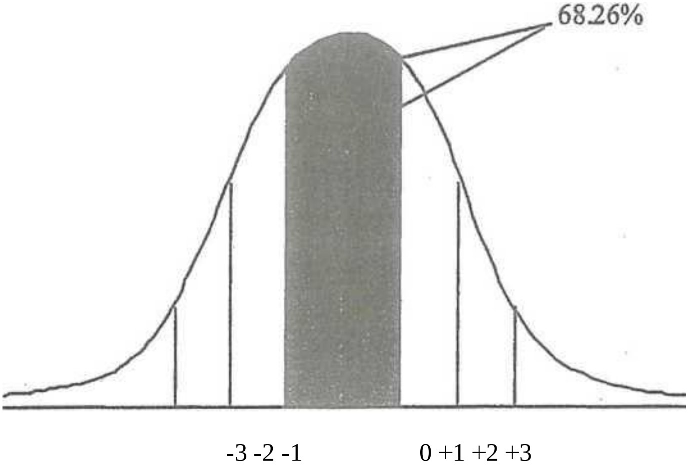
المساحة ضمن انحرافين معياريين عن الوسط ه

المساحة الواقعة على الفترة ($2a_x$;g_x +$2a_x$— لالإ)=95.44% من المساحة الكلية.

المساحةضمن ثلاثة انحرافات معيارية عن الوسط -

المساحة الواقعة على الفترة (٩3 + ,.د);٩3— تلخ)=99.74% من المساحة الكلية.

والشكل رقم (5) الآتي يبين هذه المساحات



ذاد 6826% 95.4% I

99.7%

شكل رقم (5)

تدريب (1)

قارن برسم أشكال تقريبية بين أزواج التوزيعات الطبيعية التالية:

أ- توزيع طبيعي وسطه 50 وتباينه 100، وتوزيع طبيعي آخر وسطه 60 وتباينه 100.

ب- توزيع طبيعي وسطه 40 وتباينه 64، وتوزيع طبيعي آخر وسطه 40 وانحرافه المعياري 4.

ج- توزيع طبيعي وسطه 15 وانحرافه المعياري 5، وتوزيع طبيعي آخر وسطه 25 وانحرافه المعياري 10.

المنحنى الطبيعي المعياري

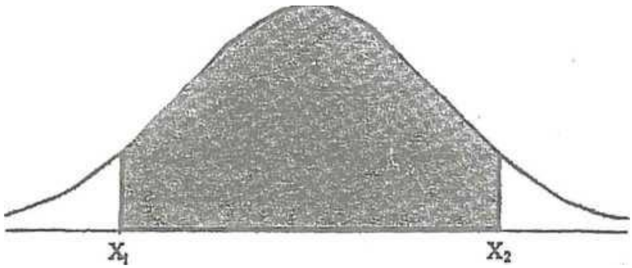
The Normalized Distribution

علمت في الفقرة السابقة أن المنحنى الطبيعي يعتمد على معلمتين هما: متوسط المجتمع \الإ وتباين قيم المجتمع ، كما تعلمت أن من خصائص هذا المنحنى أن المساحة تحته وفوق المحور السيني (الأفقي) مساوية للواحد الصحيح.

وقد نحتاج في العمل التطبيقي إلى حساب جزء من هذه المساحة (التي هي في ا لواقع تمثل احتمالاً)، وتنحصر عادة المساحات المطلوبة، إما بأصغر من قيمة ما، أو أكبر منها. وأخيرا المساحة المحصورة بينها وبين قيمة أخرى قد تكون أكبر أو أصغر منها.

وعلى سبيل المثال لو كان لدينا قيمتين للمتغير العشوائي الطبيعي X هما$X[$ و x_2 بحيث كانت قيمة الة أصغر من قيمة x_2 ، وطلب منا حساب المساحة المحصورة بينهما كما هو مبين في الشكل رقم (6).

يتطلب حساب هذه المساحة مترفة كلاً من متوسط المجتمع ٦لإ وتباينه ؛٠ وتعويضهما في اقتران التوزيع الطبيعي، ثم القيام بعد ذلك باجراء التكامل المحدود لهذا الاقتران، ولا شك أن ذلك ليس بالسهولة، لذا يمكن الاستعانة بجداول خاصة لحساب المساحات للتوزيع الطبيعي، إلا أن مثل هذه الجداول سيكون عددها لانهائي طالما أن متوسط المجتمع يل! وتباينه درك يأخذان قيماً كثيرة جداً وأن قيم المتغير العشوائي كنلك لانهائية.



شكل رقم (6)

ولحل مثل هذه المشكلة، استعان علماء الاحضا ء والرياضيات بالجزء الخاص من اقتران التوزيع الطبيعي الذي يمثل قوة العدد الأسي، ولعلك لاحظت عزيزي الدارس أن هذا المقدارعل ماهو لا العلامة المعيارية (Z-Score) التي تم دراستها في الوحدة الثانية من هذا المقرر، ومن دراستك للعلامة المعيارية فإنك تعلم أن متوسط العلامة المعيارية هو صفر وتباينها واحد صحيح، وفي هذه الحالة وجب تحويل قيم المتغير العشوائي الطبيعي إلى قيم معيارية، وهنا يسمى الاقتران الاحتمالي بالتوزيع الطبيعي المعياري أو القياسي واختصاراً بالتوزيع القياسي، ويسمى منحناه بالمنحنى القياسي الذي سيكون فيه مركز التوزيع أي متوسطه صفراً وتباينه واحد.

التوزيع الطبيعي المعياري هو التوزيع الطبيعي الذي وسطه صفر وتباينه 1، فإذا كان المتغير العشوائي Z يتبع التوزيع الطبيعي المعياري، فإن ذلك يعني أن توزيع Z هو التوزيع_الطبيعي_الذي_معدله (متوسطه)_0٢عإ وتباينه_1=|٠_.

صيق(ا)

لج

؛

باستخدام التعابير والرموز الرياضية، فإن الرمز التالي (؛ ٠ ,X~N(Mx وتقرأ: المتغير العشوائي الطبيعي X يتوزع طبيعياً (استخدم الحرف الأول من كلمة Normal) بمتوسطيلر وتباين ؛ ٠٠

فإذاكان المتغيرالعشوائية معرف بالعلاقة التالية:

Z—X —لدر

aX

عند نلك

Z~N(MZ, ٠ D

وتقرأ: المتغير العشوائي الطبيعي Z يتوزع طبيعياً بمتوسط JLtz وتباين z٠ ، ولأن0--لر وأن 1=ؤ٠.

فإن ذلك يتني أن: Z~N(O,1)

وتقرأ: المتغير العشوائي الطبيعي Z يتوزع طبيعياًبمتوسط 0 وتباين 1٠

إن ذلك يعني أن كل قيمة من قيم المتغير العشوائي الطبيعي يقابلها قيمة من قيم Z بموجب التحويل أعلاه تسمى بالقيمة المعيارية المقابلة لقيم X.

نل>1ل لتقل

ليكن لديك المتغير العشوائي X يتبع التوزيع الطبيعي بمتوسط 50 وتباين 36.

يطلب:

-1 تحويل قيم*2 لى متغير عشوائي يكون توزيعه التوزيع الطبيعي المعياري.

-2 إيجاد_القيم_المعيارية_المقابلة لقيم X) =56;x2 =32;x3_=50 .

الحل:

1) لدينا X • N(5O,36)، وعليه يكون المتغير العشوائي الذي توزيعه التوزيع الطبيعي

المعياري هو: 1 ⵣⵣ = Z

-2

X	Z
56	1
32	−3
50	0

قهويب (2)

إذا كان عمر المصباح الكهربائي الفنتج من قبل أحد المصانع يخضع للتوزيع الطبيعي بمتوسط 830 ساعة وتباين 64 ساعة، جد القيمة المعيارية لمصباح سجل عمره ب 1)815 ساعة. 2) 840 ساعة. 3) 830 ساعة.

أسفلة التقييبم الذاقي (2)

قارن بين خصائص منحنى التوزيع الطبيعي ومنحنى التوزيع الطبيعي المعياري.

حساب المساحات تحت المنحنى الطبيعي المعياري

How to Find The Area Under The Normal Curve

من بين أهم التطبيقات في نظرية الاحتمالات حساب احتمالات أن تكون قيمة المتغير العشوائي أصغر من قيمة معينة أو أكبر منها أو محصور بين قيمتين وهذا يعني حساب المساحة لهذه الحالات، ونظراً لكون التعامل مع حساب المساحة للتوزيع الطبيعي يحتاج إلى اجراء التكامل لاقتران التوزيع الطبيعي، بعد التعويض بقيمة المعالم وهي المتوسط *لمإ والتباين ؛٠ ، وكما سبق وأن بينا ان هذه خطوة ليست بسهلة لجميع الدارسين والباحثين، كما أن عمل جدول للمساحة يتطلب التعويض في كل مرة بمتوسط ما وبتباين ما، ممايعني عمل عدد لانهائي من الجداول، خاصة وأن قيم المتغير العشوائي الطبيعي هي جميع القيم على المحورالسيني. لهذا جاءت فكرة تحويل قيم المتغير العشوائي إلى قيم معيارية، حيث من خصائص هذه القيم أن متوسطها يساوي صفراً وتباينها يساوي واحد، مما يعني أن بالامكان حساب المساحات تحت منحنى الطبيعي القياسي، لأننا نتعامل هنا مع قيمتين فقط هما صفر للمتوسط وواحد للتباين وليس مع قيم لانهائية لها. لذلك تم عمل جدول لحساب هذه المساحات (أو الاحتمالات)، وثمة اختلاف في عرض جداول التوزيع الطبيعي المعياري، منها ما يعطي المساحة إلى يسار أي قيمة معيارية موجبة كانت أو سالبة، ومنها ما يعطي المساحة بين $Z=O$ وأي قيمة موجبة للمتغير 2، ولكن تبقى كل هذه الجداول تعطي نفس المساحات المطلوب حسابها، وسيستعمل هنا الجدول الذي يعطي المساحات التراكمية على يسار القيمة المعيارية المحسوبة بالأساس من تحويل قيمة المتغير العشوائي الطبيعي إلى القيمة المعيارية التي تقابله. ويعبر رياضياً عن نلك: $P(Z<z)$، وتقرأ احتمال بأن تكون قيمة المتغير العشوائي الطبيعي المعياري Z أقل من قيمة رقمية معينة هي. Z (انظر الجدول في الملحق). لذلك عند حساب احتمال أي قيمة لا بد من وضع تلك القيمة بالصيغة السابقة من رسم المنحنى المطلوب ومن تماثل منحنى التوزيع الطبيعي، ومن كون المساحة تحت المنحنى تساوي 1• .حسد,

هئل (2)______________________________ ج

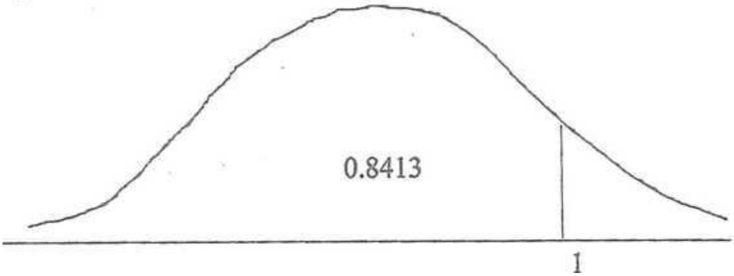
إذاكانلديك(0,1)آ1-2،جدمالي:

P(Z<1) -1	P(z<-2) -2
?(Z<1.25) -3	p(z>1.75) -4
P(-2<Zcl.75) -5	

الحل:

لحساب المساحات (أو الاحتمالات) في هذا المثال للقيم المعيارية Z وليس للقيم الأصليةم.

– انظر الشكل رقم (7)

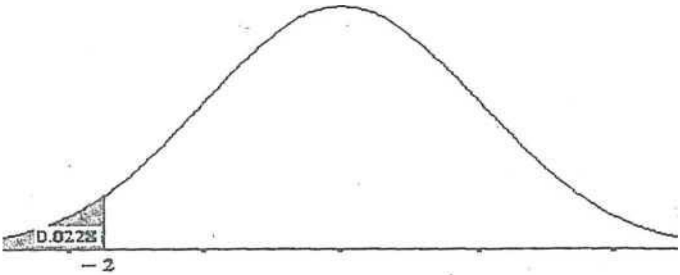


شكل رقم (7)

إن شكل المنحنى المعياري والقيمة المعيارية والمساحة المظللة هي ل P(Z<1) .

فمن جدول التوزيع الطبيعي المعياري نبحث عن قيمة 1 تحت عمود Z ولعدم وجود كسر عشري بعد 1 تأخذ المساحة التي تقابل 1 تحت العمود 00، فيكون الجواب 0.8413. أين: P(Z<l)=0.§413

2-انظر الشكل رقم (8) الآتي:

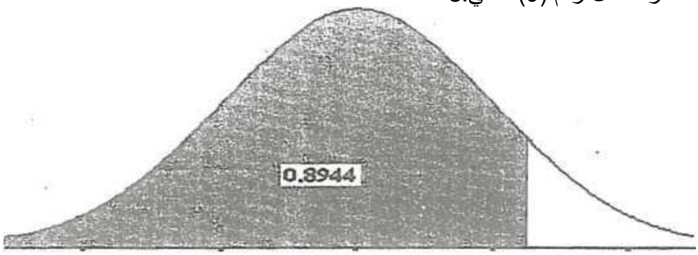


شكل رقم (8)

إن شكل المنحنى المعياري والقيمة المعيارية والمساحة المظللة هي ل P(Z<-2) ،

ومن جدول التوزيع الطبيعي المعياري نبحث عن قيمة 2- تحت عمود 2، ولعدم وجود كسر عشري بعد 2- يأخذ المساحة التي تقابل 2- تحت العمود 00، فيكون الجواب 0.0228؛ أي ن: P(Z<-2)=0.0228

3:-انظر الشكل رقم (9) الآتي



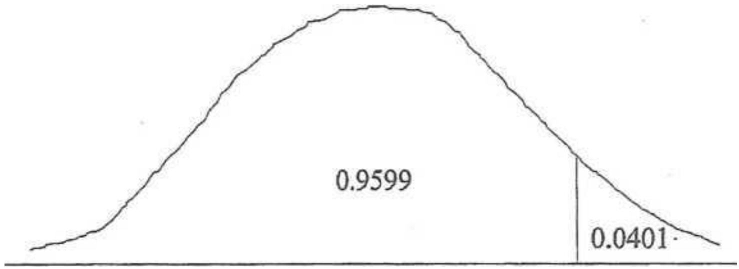
1.25

شكل رقم (9)

إن شكل المنحنى المعياري والقيمة المعيارية والمساحة بالخطوط المظللة هي ل P(Zcl.25) ·

ومن جدول التوزيع الطبيعي المعياري نبحث عن قيمة 1.2 تحت عمود Z ول 0.05، لذلك تأخذ المساحة التي تقابل 1.2 تحت العمود 05، فيكون الجواب 0.8944؛ أي أن: P(Z<1.25):0.§944

4-انظر الشكل رقم (10) لآتي:



1.75

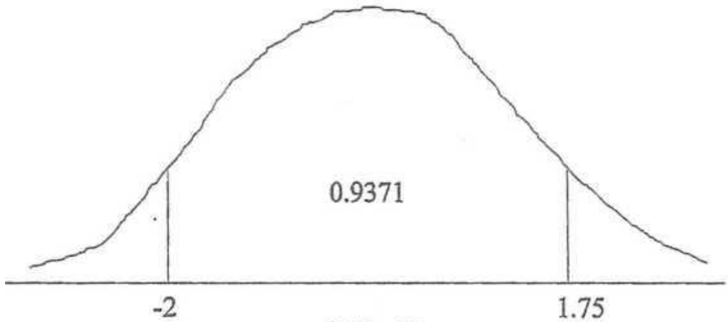
شكل (10)

إن P(Z>l,75) تعني احتمال أن تكون قيمة العلامة المعيارية أكبر من 1.75، بعبارة أخرى ما هي المساحة تحت المنحنى بأكبر من 1.75، إن طبيعة المساحة في جدول التوزيع المعياري المستعمل هو من عند القيمة المعياريه فأقل وليس أكثر كما في المساحة المطلوبة هنا، وبما أن المساحة الكلية تحت المنحنى هي 1 فيمكن من الجدول أن نستخرج المساحة من عند القيمة المعيارية فأقل، فيكون الفرق بينهما المساحة المطلوبة،

ي ان:

$P(Z>1.75)=I-P(Z<1.75)=I-0.9599=0.0401$

5- انظر الشكل رقم (11) الآتي:



شكل (11)

المساحة المطلوبة هنا محصورة بين القيمتين المعيارتين 2- و 1.75، ومن جدول التوزيع المعياري نجد المساحة الأقل من 1.75 ونطرح منها المساحة الأقل من 2-، فتكون النتيجة هي المساحة المطلوبة، أي أن:

$$P(-2Z1.75)=P(Z<1.75)-P(Z<-2)=0.9599-0٠0228=0.9371$$

عزيزي الدارس، تعلمت، في المثال السابق، كيفية حساب المساحة تحت المنحنى لمتغير عشوائي طبيعي معياري، والسؤال الآن ماذا لوكان المتغير هو X~N(pxJ٠D، وسحبت منه مفردة بشكل عشوائي فكيف نحسب الاحتمال لهذه المفردة. كل ما مطلوب عمله هنا هو تحويل قيمة X إلى Z واستخدام نفس الاسلوب الذي اتبعناه في المثال (2).

مثال (3)

لديك متغير عشوائي X يتوزع توزيعا طبيعيا متوسطه 50 وتباينه 16؛ سحبت مفردة منه، فما هو احتمال أن:

$p(x>55)$-I	$p(X<58)$ -2
$P(4O<X<6O)$ -3	$p(X<47)$-4

الحل:

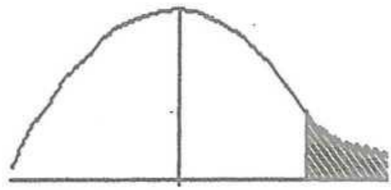
لدينا X~N(5O,16)، ولا يتوفر لدينا جدول لتوزيع طبيعي بمتوسط 50 وتباين 16. لذا يجب أن نحول X إلى Z حسب النظرية (1.4) أي أن:

هيق:عل7
6ني ٦
وهذا يؤدي إلى أن تكون Z~N(O,1)، وبالتالي سيمكننا هذا التحويل من حساب الاحتمالات (أو المساحات) المطلوبة.

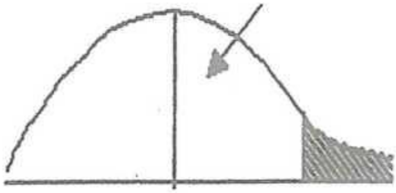
1-القيمة المعيارية المقابلة ل 55 هي:

انظر الشكلين (12) و (13)
وإن هذا يعني أن:

p(Z - 1 = (Z > 1.25) = 1 - p(X > 55) = p(Z > 1.25) = 1 - p(Z • 0.1056 = 0.8944



شكل (12)



شكل (13)

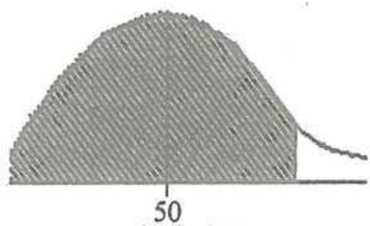
2- القيمة المعيارية المقابلة ل58 هي:

Z=S0 2

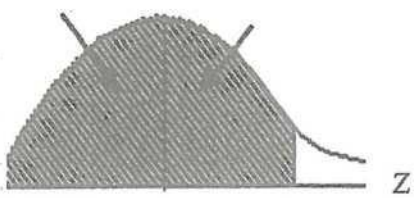
4

انظر الشكلين (14)و (15)
وهذا يعني أن:

p(X < 58) = p(z < 2) = 0.9772



شكل (14)



شكل (15)

- القيمة المعيارية المقابلة ل 40 هي:3

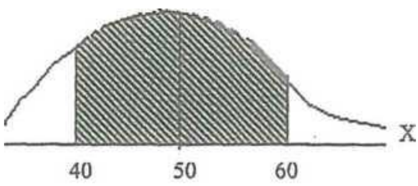
2ـد-2.5

وا لقيمة المعيارية المقابلة ل 60 هي:

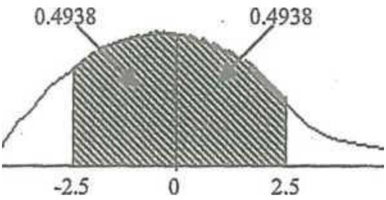
2,ئ:.:
4

$$P(4O < X < 60) = p(-2.5 < z < 2.5) = p(Z < 2.5) - p(Z < -2.5)$$
$$p(4O < X < 60) = 0.9938 - 0.0062 = 0.9876$$

انظر الشكلين (16) و (17)



شكل (16)



شكل (17)

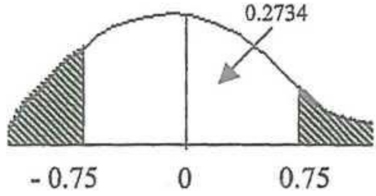
4- القيمة المعيارية المقابلة ل 47 هي:

عغ:-0.75
4

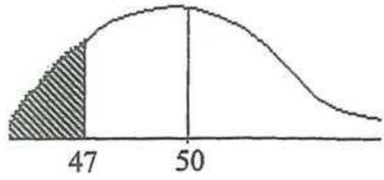
مما يعني أن:

$$p(X < 47) = p(Z < -0.75) = 0.2266$$

انظرالشكلين (18) و (19)



شكل (19)



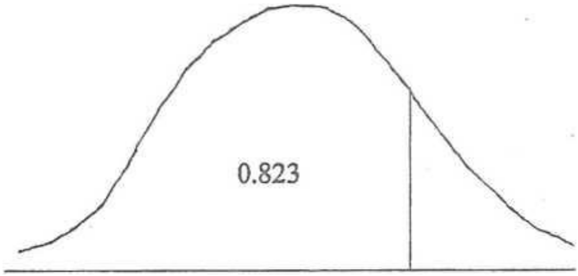
شكل (18)

مثال (4)

إذا كان $Z \sim N(0,1)$ ، فجد قيمة الثابت b بحيث $p(Z < b) = 0.8238$.

الحل:

في هذا المثال لدينا المساحة (الاحتمال) والحالة المطلوبة هي ايجاد القيمة المعيارية 2، يوضح الشكل رقم (20) الآتي المساحة المعطاة لذلك:



شكل رقم (20)

إن سبب رسم المنحنى أعلاه بهذا الشكل يعود إلى أن المساحة المطلوبة هي المساحة الأقل من القيمة المعيارية لأ، ولأن المساحة هي أكبر من 0.5 فهذا سيجعلها في النصف الموجب من المحور السيني، وبالبحث داخل المساحات في الجدول المعياري سنجد بأن المساحة 0.8238 تقابل العلامة المعيارية 0.92 ، أي أن:

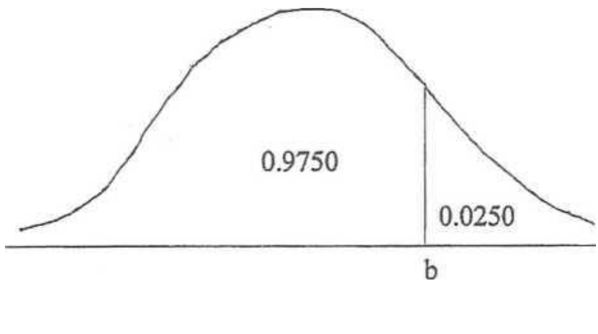
لأ = 0.93

مثال (5)

.p(Z>b)إذا كان (0,1)آ+-ة،فجد قيمة الثابت 1 بحيث 0.0250 = (

الحل:

يوضح الشكل رقم (21) الآتي المساحة المعطاة لذلك:



شكل رقم (21)

$$p(z>b) = 0.0250$$
$$l-p(z<b) = 0.0250$$
$$p(Z<b) = 0.9750$$

وبالبحث داخل المساحات في الجدول المعياري سنجد بأن المساحة 0.9750 تقابل العلامة المعيارية 1.96، أي أن:

$$b=1.96$$

مثاله (6)

لديك (X* N(p ,100 وأن p(X < 40) = 0.0080، جدقيمة المتوسط فر.

الحل:

هذا المثال يختلف عما سبقه في كون المطلوب هو إحدى معالم التوزيع الطبيعي، لذلك نجد العلامة المعيارية وفق الآتي:

العلامة المعيارية للقيمة 40 هي:

٠, 10

ومن الجدول المعياري نجد بأن المساحة 0.0080 تقابل العلامة المعيارية 2.41-، أي أن:

p(Z < - 2.41) ه 0.0080

وهذا يعني أن:

ئي=2.41-
10

ومنها:

^=40 + 2.41*10=64.1

طحلر7) K?

لديك (X-N(25'O؛ وأن0.9970-(36>*)]،جدقيمة لتباين ي.

الحل:

المطلوب هنا إحدى معالم التوزيع الطبيعي ت٠ ، لذلك نجد العلامة المعيارية للقيمة 36 وفق الآتي:

العلامة المعيارية للقيمة 36 وهي:

لل=قة‌إ:ختخ=2

ومن لجدول المعياري نجد ن لسساحة 0.9970 بتبل العلامة المعيرية 2.75، أيأن: 0.9970 = p(Z<2.75)

وهذا يعني أن:

لل 2.75

ومنها:

775 Va =لل=4د0؛ =16

تهويبر (3)

قرر مدرس الإحصاء أن يعطي أعلى 15 هره من طلبته في تلك المادة تقدير (ممتاز)، فإذا كانت علامات الطلبة تخضع لتوزيع طبيعي وسطه 67 وتباينه 169.

أ- فما هي أقل علامة تحصل على تقدير "ممتاز"؟

ب- إذا كانت علامة الرسوب في تلك المادة دون الخمسين، فما نسبة الطلبة الراسبين.

تهويج (4)

إذا كانت قوة نوع خيوط الصوف تخضع لتوزيع طبيعي وسطه 130 وانحرافه المعياري 25، واخترت خيطاً بطريقة عشوائية.

أ- فما احتمال أن تكون قوته أكبر من 160؟

ب- وما احتمال أن تكون قوته أقل من 120؟

أتسئلة التقويم 1لفاتي (3)

اذكر الشروط الواجب توفرهالحساب المساحة تحت المنحنى المعياري.

لآ٠!اي٠اع٦٠اا٠$ ٦

لعلك تذكر مما سبق أن علم الإحصاء يقسم إلى قسمين، الأول هو الإحصاء الوصفي Descriptive Statistics وهو ما تم دراسته في الوحدتين الثانية والثالثة من هذا المقرر، والقسم الثاني ويطلق عليه الإحصاء الاستنتاجي Statistical Inference. والقسم الأخير هو الذي يمكن الدارس والباحث من التوصل إلى معلمات Parameters المجتمع المدروس عن طريق سحب عينة عشوائية، وتسمى طريقة السحب بالمعاينة، وتتلخص فكرة العشوائية أن كل مفزدة أو وحدة أو قراءة في المجتمع لها نفس فرصة الظهور في العينة بصرف التظر عن نوع العينة المختارة عشوائية بسيطة كانت أم عشوائية طبقية أو غيرها، كما ذكرنا في الوحدة الأولى.

والعينة هي جزء من مفردات المجتمع وبالامكان حساب المقاييس الإحصائية الوصفية كالوسط الحسابي والتباين والارتباط وغيرها كما هو الحال لجميع المفردات في المجتمع وتسمى المقاييس الاحصائية المستخرجة من بيانات العينة بالإحصاءة Statisticويرمز لها بالأحرف الانجليزية؛ وعلى سبيل المثال إذا كان الرمزـلإ يمثل الوسط الحسابي لمجتمع المتخير العشوائي X وهي معلمة فإن الرمز X يمتل الوسط الحسابي للعينة المسحوبة التي تضم جزءا من ففردات المتغير العشوائي X وهو إحصاءة، وإذا كان الرمز ؤن يمثل تباين مجتمع المتغير العشوائي Y وهو معلمة، فإن الرمز ؛S يمثل تباين العينة المسحوبة من مجتمع المتغير العشوائي Y وهو إحصاءة.

فإذا كان لمتغير المجتمع اقتران احتمالي فإن لكل إحصاءة محسوبة من بيانات العينة اقتران احتمالي يسمى بتوزيع المعاينة. وعلى هذا الأساس هناك توزيع معاينة للوسط الحسابي للعينة، وتوزيع معاينة لتباين العينة وهكذا لبقية الإحصاءات.

توزيع المعاينة للوسط الحسابي للعينة

Sampling Distribution of the Sample Mean

إذا كان المتغير العشوائي X يتوزع توزيعا طبيعيا بمتوسط يل! وتباين ؛ه وعدد مفرداته N وسحبت عينة عددها (حجمها) n من هذا المجتمع، فإن لهذه العينة وسط حسابي X وتباين ؛S ، وإذاقمناباعادة مفردات العينة إلى المجتمع وسحبنا عينة عشوائية جديدة بنفس الحجم n فإن للعينة الجديدة وسط حسابي X وتباين ؛S ، وإذا أعدنا مفردات العينة إلى المجتمع مرة أخرى وسحبنا عينة عشوائية جديدة وبنفس الحجم فإن للعينة الثالثة وسط حسابية وتباين ي5 ، وهكذانستمربالاعادة والسحب من جديدولكن

بنفس الحجم، وفي كل مرة نحصل على وسط حسابي X وتباين ؛S ، والسؤال الأول هنا هو: هل هذه الأوساط الحسابية للعينات متساوية جميعها وكذلك التباينات، والسؤال الثاني هو: ما علاقة هذه الأوساط بمتوسط المجتمع \إ. وقبل الاجابة عن هذين السؤالين، ننكر أن قيمة الوسط الحسابي لأي مجموعة من البيانات يعتمد على قيمها، أي إذا كانت القيم مختلفة فإن الوسط الحسابي لها ثابت. وعلى الرغم من أنه سيكون ثمة قيمة و لحدة إلا أنها تختلف عن الوسط الحسابي لجميع القيم قبل السحب، لذا يمكن القول أن الأوساط الحسابية للتينات تختلف من عينة إلى أخرى، ولأن السحب في الأساس عشوائيا، فإن الوسط الحسابي للعينة X هو متغير عشوائي، ولكونه متغير عشوائي فإن له وسط حسابي يرمز له ب -لما ويعني المتوسط العام لمتوسطات العينة أو الوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة وكذلك له تباين يرمز له ب ؤى ويعني تباين توزيع المعاينة للوسط الحابي للعينة وأن قيمة الجخر الترهيعي له يسمى بالانحراف المعياري للوسط الصلببي.

قظوية (2)

إذا كان (يع و N(j١c تمم X وكان حجم المجتمع N وسحبت عينة عشوائية حجمها 11مع الارجاع فإن:

1-عدد العينات الممكن تكوينها N^n .

2- (إه’ N X

3-وأن ٦لإ٢لإ

4- ك٢0
n

5-أي أن الاش؛ X
' واع

إن النتيجة 3 السابقة: تشير إلى أن المتوسط العام لمتوسطات العينة هو نفسه متوسط المجتمع للمتغير العشوائي X.

بينماتشير النتيجة 4: إلى أن تبلين الوسط الحسابي للعينات هوتبالن المجتمع للمتغير العشوائي X مقسوما على حجم العينة المسحوبة.

أما النتيجة 5 فهي تلخص النتيجتين 3 و4.

والمثال التالي يوضح ذلك.

مثال (8)

افترض أن مجتمعاً طبيعياً يضم القيم التالية: 1،3، 5، 7، 9، وسحب من هذا المجتمع عينة بحجم 2 مع الارجاع، يطلب:

1- عدد العينات الممكن تكوينها.

2- الوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة.

3- الانحراف المعياري لتوزيع المعاينة للوسط الحسابي للعينة.

الحل:

لللة لا بد م مترفة لسط صع «لفلة لفم سع رلكك X:

N ا 5

X~N(5,8)

1- لن نعوض الآن بمعادلة عدد العينات، ولكن سنستعرض العينات الممكن تكوينها هنا ثم نعدها ونقارن بالنتيجة التي تظهر لو استخدمنا علاقة عدد العينات الممكن تكوينها.

والجدول رقم (1) الآتي يظهر العينات المسحوبة في المرة الأولى والمرة الثانية:

جدول رقم (1)

		السحبة الثانية				
		1	3	5		9
السحبة الأولى	1	(1و1)	(1,3)	(1و5)	(1و7)	(1,9)
	3	(3,1)	(3,3)	(3,5)	(3و7)	(3,9)
	5	(5و1)	(5,3)	(5,5)	(5و7)	(5,9)
	7	(7و1)	(7,3)	(7و5		(7,9)
	9	(9,1)	(9و3)	(9,5)	(9,7)	(9’9)

يلاحظ من الجدول أعلاه تكوين 25 عينة في حالة الارجاع، وكل عينة تضم حجم يساوي 2، و لو قمنا بحساب معاللة عدد العينات الممكن تكوينها، حيث لدينا 5 == Nوهي عدد القيم للمجتمع و 2 = n حجم العينة مع الارجاع، لوجدنا 25==5^2=2(N) وهي نفس ما توصلنا إليه في عدد العينات الكلية الممكن تكوينها.

2" لحتاب قيمة _لإ سنقوم أولاً بحساب الوسط الحسابي لكل عينة ثم حساب متوسط هذه المتوسطات، والجدول رقم (2) الآتي يبين تسلسل العينة وعناصرها ووسطها الحسابي،

حيث أت٦

جدول رقم (2)

	العناصر	الوسط الحسابي X		2(هإ-ت)
1	1,1	1	-4	16
2	1و3	2	-3	9
3	5,1	3	-2	4
4	7,1	4	-1	1
5	91	5	0	0
6	ذاً	2	-3	9
7	3و3	3	-2	4
8	5,3	4	-1	1
9	73			
10	93	6	i	1
11	1,5	3	-2	4
12	3,5	4	-1	1
13	55	5	0	0
14	٦٠٥	6	1	1
15	9.5	7	2	4
16	1,7	4	-1	1
17	3.7	5	0	0
18	[illegible]	6	1	1
- 19	77	7	2	4
20	9.7	8	3	9
21	1.9	5	0	0
22	3.9	6	1	1
23	59	7	2	4
24	٦٥	8	3	9
25	9,9	9	4	16
المجموع		125	0	100

ولحساب متوسط المتو بسطات

والنتيجة الأخيرة تساوي متوسط مجتمع X المحسوب في بداية حل المثال، وهذا

يعني أن:

5=*لإ٢لإ

3- ولحساب تباين المتوسطات، نجد أن:

1:4:ثالألاجة:ر

X N“ 25

وكان بالامكان حساب هذه النتيجة باستخدام النتيجة 3 من النظرية (2) حيث لدينا وحجم العينة المسحوبة 2- 10، لذايكون التباين:

؛ § -ص

X n 2

أما الانحراف المعياري (للمتوسطات) ورمزه _٠ فيحسب وفق الآتي: جمة4٢ت٢

مال (9)

افترض أن العينة المسحوبة لبيانات المثال (8) كانت بحجم 4 مع الارجاع،

فاحسب:

1- عدد العينات الممكن تكوينها.

2- الوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة.

3- الانحراف المعياري لتوزيع المعاينة للوسط الحسابي للعينة.

الحل:

منحل المثال (8.4) لدينا X~N(5,8)

1- عدد العينات الممكن تكوينها:

N^n:5^4 : 625

2- الوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة:

5= *صلل

3- الانحراف المعياري لتوزيع المعاينة للوسط الحسابي للعينة:

تحويج(5)

القيم التالية لمجتمع طبيعي وتمثل درجة حرارقأجسام مرضى.

38، 39، 38، 37، 38، 40، 37، 38، 39، 36 . ١

إذا اختير ستة مرضى بطريقة عشوائية مع الإرجاع، فاحسب:

ا 1- عدد العينات الممكن تكوينها.

2- الوسط الحسابي لدرجة حرارة الجسم لتوزيع المعاينة للوسط الحسابي للعينة.

3- الانحراف المعياري لدرجة حرارة الجسم لتوزيع المعاينة للوسط الحسابي للعينة.

أفلة ل التقويم الةاقي (4)

لـج

ما الفرق بين الانحراف المعياري للمتغير والانحراف المعياري لتوزيع المعاينة للوسط الحسابي للعينة.

تقدير وسط المجتمع ونسبته.

Eiiiiiofi • • ohiidit • p(ii

ر .'ة.. ■•دن:،؛•.’ ٦ ب ٦ح-٢.٢٦'ج / ..

-'.-.-لم:؛.'.'.

Point ةئ.برك٢;. لا٠دأ٠ج

من المنطقي أن نقدر وسط المجتمع بوسط العينة التي نسحبها منه، فإذا كان وسط المجتمع m فمن الطبيعي والمنطقي أن نقدرلا بواسطة وسط العينة X .

وبما أن تونيع المعاينة للإحصاء X يتمركز حول لم حيث $E(X)=7$ وأن تباين X هو — حيث 2٠ هو بملن المبكي، n حبم العينة، واه فى مطم الاميان نجد ان هذا التباين أصغر من تباين أي مقدر آخر.

لذا فإن القيمة X ستستعمل كتقدير نقطي لمعدل المجتمع لم ولما كان تباين Xهر — فيذ يعن لن مان قنبه السعينة للبصاء X يهة مغاا، ولاصة عدما تكبرة،ولنلكمن المتوقع أن تكون لقيمة تتقديرأدقيقاً لمعدل لمجتمع ا.

•:لا -.لآض.

:٩دى٠٦١٠■■■' ٢, .لم - رب •.ب٠.؟ ٦٠٠ ج:٠ذذ

التقدير بنقطة (التقدير النقطي) لمعلمة مجتمع هو قيمة عدديةمفردة تقدر بهات لك ا :المطفة.

؛د.. بم

آ -

لتقدير عدد الكلمات في أحد الكتب، أخذت عينة حجمها 10 صفحات عشوائياً من : ذلله الكتاب وتم عد الكلمات فيهافكانكمالي:

309,310,295,305,300, 297,291,303,317,283 . : كم تقدر وسط عدد الكلمات في الصفحة الواحدة لم؟ إذا كان الكتاب يحتوي على :

400صفحة_فكم_يقدر_عيد_لكلماتف_يه؟ ا

:بيئح: .

حجم ا لعينة 10 ووسط تلك العينة هو:

- 297+ 291-+ 303+317+238-+ 309 ؛310 + 295+305 + 300

10

= 301

إذن تقدر وسط عدد الكلمات في الصفحة الواحدة لز بالوسط الحسابي للعينة وهو 301 = X وتقدير

عدد كلمات الكتاب ه عدد الصفحات X تقدير عدد الكلمات في الصفحة وهو: كلمة120400- 301]400

لتقدير وسط المصروف الشهري بالدينار لدى طلاب إحدى الجامعات قام باحث :

: بمقابلة عينة عشوائية من الطلبة حجمها 20 فوجد أن مصروفهم الشهري كما يلي:

, 50.3 و 49.4 و 48.1 , 50.2 ’ 35.5 , 37.2,42.0,43.6 و : 49.2,38.3 , 36.5,37.8

30.2,38.7,32.6,41.8,33.9 ’ 35.1,44.0’39.0 :

ما هو تتديرك لوسط المصروف الشهري لجميع طلبة الجامعة؟

' دد بكر وسد المسصسيرن الفي لصية التي تم ابيلما بيحمسي 20 طالباً. وتسي وسط العينة وهو:

(35.5 + 37.2 + 42.0 + 43.6 + 44.0 ؛ 39.0 + 49.2 + 38.3 + 36.5 + 37.8 + 48.1
+ 50.2
؛32.6 + 41.8 + 33.9 + 35.1
+ 49.4 + 50.3 + 30.2 + 38.7)/20
= 40.67

طه شر رسا "سسلت لفسه لصلة لمسعة مقلر 40.67 فللا

لمعرفة وسط عدد الأيام التي يقضيها العرضى في س ستنس بيست كسج ا عينة عشوائية حجمها 30 من سجلات المرضى الذين غادروا المستشفى بعد الشفاء فوجد أن : . وسط عدد أيام الإقامة من العينة المذكورة 6.7 يوماً، فما هو وسط عدد الأيام التي يقضيها :

المرضي في ذلك المستشفي؟

اسعسل وسد العبة بمحر لقله لرسد اسبعي، ولنلك بر وسد عد الله لك يقضيها المرضى في ذلك المستشفى بالمقدار 6.7 يوماً، الذي هو وسط العينة التي تمت دراستها.

إن تقدير وسط المجتمع بفترة هو عبارة عن إيجاد تقدير نقطي لوسط المجتمع أولا ومن ثم استعمال هذا المقدار لايجاد قيمتين تعتمدان على التوزيع الإحتمالي لهذا المقدر وعلى معامل الثقة، وبالتالي تمثل هاتان القيمتان الحد الأدنى والحد الأعلى لفترة الثقة المطلوبة.

فإذا كان وسط المجتمع لمإغير معلوم واردت ايجاد فترة ثقة للمعلمة لمإ فإنك تستعمل وسط العينة x كتقدير نقطي لما للمعلمة وتستعمل توزيع المعاينة للأحصاء x لتحدد الحد الأدنى والحد الأعلى لفترة الثقة المطلوبة.

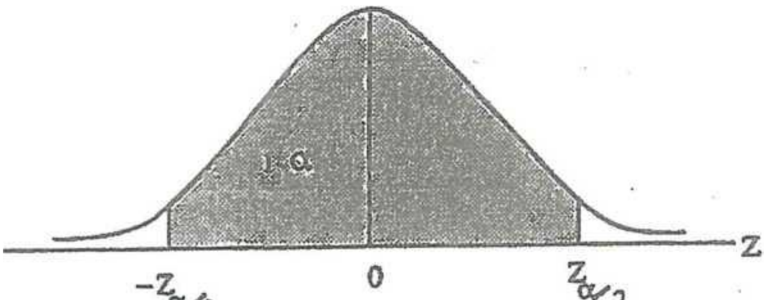
فإذا كانت العينة المدروسة من مجتمع طبيعي تباينه معلوم أو في حالة عدم تحقق ذلك، لذا كان حجم العينة n كبيراً فان تونيع المعاينة للإحصاء x هو التونيع الطبيعي أو تقريبا القنيه الليعى ذه الوسط!) واملن — وفي هذه الحالات تطيه ليجد فقة القة للوسط m من العبارة الاحتمالية:

P(-Za<Z<z^)=l-a

ج هتكخ =7

٠ / ٠

وحيث يي2 هي النقطة على المحور الأفقي لمنحني التوزيع الطبيعي المعياري الذي يقع إلى يمينه من المساحة أنظر الشكل (1).



الشكل (1)

ومن المعادلة السابقة تجد فترة الثقة ذات معامل ثقة (1- a) للوسط ئإه ونسميها فترة ثقة(1-0)00 1% كما يلي:

إذا كان X الوسط الحسابي لعينة عشوائية حجمها n من مجتمع طبيعي تباينه 2ه معلوما فإن فترة ثقة (ى-1) 00ا% للمعلمة لز تكون:

X-Za..؛<1،<X+Z٠،i

حيث Z هي النقطة على المحور الأفقي لمنحنى التوزيع الطبيعي المعياري والتي يقع إلى يمينها مساحة 8 .

أوجد فترة ثقة 95% للوسط اإ في مجتمع طبيعي تباينه 64 إذا اختيرت عينة عشوائية حجمها 9 وكان وسطها الحسابي 52 = X

بما أن 1-a = 0.95

إذن 0.025 = ٠ وبالتالي 1.96 Z من جدول التونيع الطبيعي.

إذن، بالتعويض في فترة الثقة تجد أن فترة 95% ثقة للوسط لإ هي:

52 - 1.96 X في > U. < 52 + 1.96 X في

52-5.227 <1د<52+5.227

أي أن: 46.773 < لإ< 57.227

هي فترة ثقة 95% للوسط لما

عنة عوالية حجمبا 25 = n أخنت من مبتمع طبيعي انصرافه المعيارعي 5 = 0 أعطت المعدل X80

أوجد فترة ثقة 98% لوسط المجتمع لمز.

بما أن $1-a = 0.98$

إذا 0.01=؛ وبالتالي $z_{٠٥} = 2.33$

بالتعويض في فترة الثقة:

أنسسلآ>مل.1س ت-آ

تجد أن فترة ثقة 98% للوسط لا هي:

§0-2.33 X في> U, < 80 ب 2.33 X ئ

ج١٢ جي

وبا لإختصار تجد:

77.67 <دإ<82.33

هي فترة الثقة المطلوبة.

فيس سصل،(العحات السعيالي للعيلح بدلأ من 7، شبيلة ان n ي (n >30) وفي هذه الحالة تكون فترة الثقة ـ

للوسط 1) هي: X - Za/2 . ت■ < P■ < X ب ZQ/2 . T

عينة عشوائية من مجتمع طبيعي حجمبا 81 = n أعطت 6 = 63,8 = X أوجد فترة ثقة 98% لوسط المجتمع لإ٠

من الواضح أن بياين المبتمع به نير معوم لكن حجم العينة أكير من 30 ولفلل تستعمل وبدلاً من 7، كما في (2).

ومن الواضح أن 0.98 = a -1 ولنلك 0.01 =ح ومن جدول الطبيعي المعياري

تجد 2.33 Z٠.O1 ٠

بالتعويض في فترة الثقة المعطاة في (2) تجد:

63 - 2.33 X ه < LI < 63 ب 2.33 X ئ

أ/هبى

وبالإختصار تحصل على:

61.447 > لم >64.553

فترة ثقة 98% لوسط المجتمع 1|٠

إذا كان الوسط الحسابي والإنحراف المعياري لمعدات عينة عشوائية مؤلفة من 36 طالباً في إحدى الجامعات هي، S = 12, X = 66.3 .

اوجد (أ) فترة ثقة 95% (ب) فترة ثقة 99% لمعدل جميع طلبة الجامعة إذا افترضنا أن علامات الطلبة تخضع لتوزيع طبيعي.

تباين المجتمع غير معلوم لكن n كبيرة، وبالتالي تستطيع استعمال الشرح في (2).

()1-42095 ذا 0.025= ؤ

Zo.oz5⁼1-96

نعوض في (2)فنجد: ؛I.96x ب 66.3>لمز> 66.3-1.96 X

لعوض في (2) ٠ :ةقر٦ لإ 36^٧

وبالاختصار تكون فترة 95% ثقة هي:

66.3-3.92 <1<66.3+3.92

62.38 >لمإ> 70.22

(ب) 1-0.99 إذا 0.005= ؤ

١٧

2

z ٥.٠٥؛

وبكون فتري التقة ٩٩٠/ه هى ي 2.57X ب 66.3>لإ> |-66.3-2.57x

دسس بر“ / -٠

وبالإختصار 5.14 ب 66.3 > دج > 66.3-5.14

أي 71.44>لإ>61.16

ويمقارنة (أ)، (ب) تجد أنك تحتاج الى فترة أطول إذا أردت أن تعطي تقديراً أكثر دقة للمعدل لإ، ففي (أ) كان طول الفترة التي ثقتها 95% هو:

70.22-62.38-7.84

أما في (ب) فكان طول الفترة التي ثقتها ٩٩٠/ه هو: 10.28 = 71.44-61.16

لاحظ: تعطي فترة الثقة (1-a) 100 % تقديراً لدقة التقدير النقطي، الذي تستعمله.

فالشكل الذي أمامك يبين فترة ثقة للمعلمة لما معامل ثقتها (a— 1) .

ث X+Z س ة. X

فإذا كانت 1] في مركز فترة الثقة فإن ذلك يعني أن استعمالك X كمقدر للوسط gكان بدون خطأ. ولكن إذا كانت g غير واقعة في مركز الفترة فهناك خطأً في التقدير حصل بسب استعمالك للقيمة X كتقدير للوسط لما.

وقيمة ذلك الخطا هي الفرق بين لمز و X ومن الواضح أن لديك ثتقة (l-a)lOO% وأن الخطأ الحاصل من استعمالك X كمقدر للمعدل g لا يزيد علىً/٦ سأ٠

أج ع٢'ح لآةلاحئكن مبمي غير طبيتي فزن ييعله اسضسلينطيد النهاية المركزية لمعرفة توزيع الوسط إذا كانت n كبيرة وبالتالي تكون فترة الثقة olOO/٠(l-a) للمعدل لا هي:

X-2ص. *T* <g<x + Z٠٥٠T

وفي حالة عدم معرفة a تستعمل S (الانحراف المعياري للتينة) بدلاً منها، شريطة أن تكون n كبيرة.

مثال (8)

درست عينة عشوائية حجمها 1§ سيجارة لتقدير وسط النكوتين في السجائر فوجد إ
_ملغم._أوجد_فترة ثقة_99%_لوسط_النكوتين_في_السجائر.S ملغم،_0.9 =_X: أن_3.6_=

!ثخعزيي:

1-0.99 إذن 0.005 = aاستعمل الفترة في (3)، إذن:

3.6 - 2.57 X - ٦ث > ·LI < 3.6 + 2.57 X ئل

-i

أي 3.6- 0.257 >لما> 3.6.+ 0.257

3.343 > عر >3.857

هي فترة الثقة المطلوبة.

لانذلاهم

طق (و وج

قيست أطوال عينة عشوائية مؤلفة من 50 طالباً في إحدى الجامعات فكان معدل : الأطوال 170.8=*سموالانحراف المعياري لها 7.0 =8 سم.

أوجد فترة ثقة 98 % لمعدل أطوال جميع طلبة الجامعة. . :

50 — n كبيرة، وبذلك تستطيع استعمال (3).

!-« = 0.9

إذن 0.01=^

2

ومن جدول التوزيع الطبيعي المعياري z_0 • oi=2.33 وبالتالي تكون فترة الثقة

98ه/ه هي: جة + 170.8 > لها > 170.8.2.33x^

9*ي: *50ل* لم 50ا ١

أي أن الفترة المطلوبة هي: 173.11 >لما>168.49

يصدف أنك تحتاج لإيجاد فترة ثقة لمعدل توزيع طبيعي تباينه غير معلوم وحجم العينة التي يمكنك دراستها صغير (أي 30 > n).

في مثل هذه الحالة يمكنك استعمال النظرية القائلة بأن تونيع المعاينة للإحصاء

للتج = T هو تونيع t بدرجات حرية عددها (n -1)•

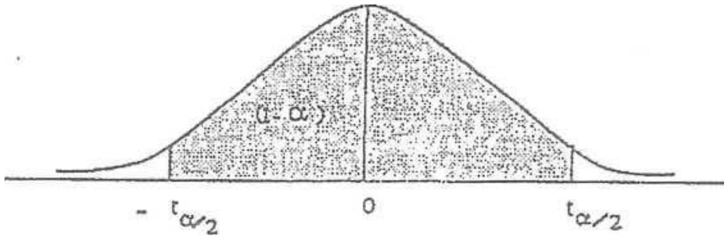
S/١/n

وفي هذه الحالة تكون فترة الثقة (a- 1) 100 % للمعدل لل معطاة بالمتراجحة:

X - ط - *T* >بت>لإ؛ a/2. *T*

حيث ta/2 هي النقطة على المحور الأفقي لتوزيع t ذي درجات الحرية (1 - n) والتي

يقع إلى يمينها ع من المساحة، كمايظهر في الشكل (2).



الشكل (2)

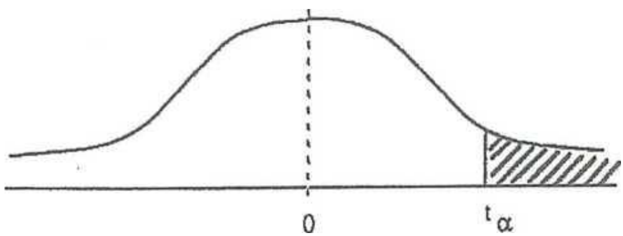
ولايجاد النقاط الحرجة لتوزيع ت (t) فقد صممت جداول خاصة لهذا الغرض.
ولتوضيح استخدام مثل هذه الجداول نقتبس جزءاً منها:

درجات الحرية	**0.10؛**	**0.05؛**	**0.025؛**	**0.01؛**
1	3.078	6.314	12.706	31.821
2	1.886	2.920	4.303	6.965
3	1.638	2.353 2.132	3.182	4.541
4	1.533	2.015	2.77	3.747
5	1.476		6	3.365
		1.753	2.571	
15	1.341	1.746		2.602
16	1.337	1.740	2.131	2.583
17	1.333	1.734	2.120	2.567
18	1.330	1.729	2.110	2.552
19	1.328		2.101	2.539

حيث يمثل الرمز ؛to قيمة النقطة من توقيع؛ بحيث تكون المساحة تحت منحنى التوزيع وعلى يمين

t_a تساوي 0.

أي أن $P(t > t_a) = cc$ كما في الشكل (3).



الشكل (3)

فمثلاً إذا كانت درجة الحرية 4 فإن

،0.05 2.132

وبما أن توزيع t متماثل حول الصفر فإن

،0.95 ه - ،0.05 = -2.132

أخذتعينةعشوائيةحجمها 7 من مجتمعطبيعي فأعطت 1.2 1-٢ة ,0.4= 5. أوجد فترة ثقة 95% لمعدل المجتمع fl,

شروط (4) متوفرة ولذلك نستعمل فترة الثقة.

X - ٤ص• *T* <لإ> بخ أ0ا٠2 ٢

بماأن 0.95 = »-ا إذا 0.025 = ع ومن جدول تونيع ؟ بدرجات حرية 6

الوارد في الملحق في الجدول رقم (4) نجد أن

،0.025 ه 2.447

إذن فترة الثقة 95% للوسط m هي:

11.2-~~2~~ت447تم~~0.4~~>ء •0ي

*fi. * ' r*

أي أن الفترة المطلوبة هي:

1 1.20-0.37 <ل<11.2+0.37

10.83 <لإ<11.57

كالت محتويات 9 عبوات من أحد أنواع المنظفات كالآتي:

10.1 , 10.3,9.9,9.8,10.2,9.7,10.0 , 10.3,9.7 لتراً. :

أوجد فترة ثقة 99% لوسط محتويات العبوات لذلك النوع من المنظفات على افتراض أن محتويات العبوات يخضع للتوزيع الطبيعي.

نحتاج لحساب وسط محتويات العبوات X والانحراف المعياري لتلك المحتويات Sللعينة المعطاة.

٧_ 10.1+10.3 + 9.9 + 9.8 + 10.2 + 9.7+10.0 + 9.7 + 10.3

9

لترا ح10.0

(ه-ا)_ج2 : Z(x-x)2

ع (0-1)2 +(0-3)2 +(-0.1)2 +(-0.2)2 +(0.2)^2 +(-0.3)2 +0+(0,3)2

8

0.01 + 0.09 + 0.01 + 0.04 + 0.04 + 0.09 + 0 + 0.09 + 0.09

= !:0.0575

§

s = ٧0.0575 = 0.24

بما أن 0.99 = 1-a إذا 0.005

2

إذن 3.355 = 005 0؛ من جدول توزيع t ذي درجات الحرية 8

إذن تكون فترة الثقة المطلوبة:

io.o-3.355x ٢ >لز>10.0 + 3.355آل

وبالتالي فإن فترة الثقة هي: 10.27 >لإ> 9.73

ينتج أحد المصانع مصابيح كهربائية تخضع أعمارها تقريبا لتوزيع طبيعي انحرافه المعياري 45 ساعة.

أخذت عينة عشوائية حجمها 25 مصباحاً، فكان الوسط الحسابي لأعمار هذه المصابيح 790 ساعة، أوجد فترة ثقة 98% لمعدل أعمار المصابيح التي ينتجها المصنع؟

أخذت عينة من 50 خيطاً من مصنع للخيوط فوجد أن وسط قوة هذه الخيوط 80.2كغم بانحراف معياري مقداره 6.5 كغم، أوجد فترة تقة 95% لمعدل قوة جميع الخيوط التي ينتجهاذلك المصنع.

أخذت عدة أفلام من إنتاج شركة ما بطريقة عشوائية فوجد أن الفترة الزمنية لها هي: 102,99,96,105,103,98,101 دقيقة. اوجد:

أ- تقديرا نقطيا لمعدل زمن جميع الأفلام التي أنتجتها تلك الشركة.

ب- على فرض أن أزمان الأفلام تخضع لتوزيع طبيعي، أوجد فترة ثقة 95% لمعدل زمن جميع الأفلام التي أنتجتها تبك الشركة؟

من المعلوم أنه إذا كانت Xj,...,X„ عينة من توزيع طبيعي وسطه لمإ وتباينه ٠2 فلن تتدير بنقطة بالإحصاءه p. = X وإن مربيع الخطا المعياري في هذا المقدار يساي — Var (X ومن ناحية أخرع فإن فقة 100 % (1 - a) ثقة للوسط : هي:

إذا كالن 7) مطومة، ويسمى المقار T z حد 100%(1 - a) للخطا فى $\widetilde{7\delta}$ ا 7 تقدير ب•

لهذا إذا تم تحديد المقدار الأكبر المسموح به لهذا الخطأ في التقدير أمكن حساب حجم العينة اللازم لتحقيق ذلك الحد عن طريق حل المتباينة.

z d

في7

حيث d حد المقدار الأكبر المسموح به لهذا الخطأ، وينتج من ذلك أن:

fz٥o١ 2

n> ―――― d

ومن الجدير بالملاحظة هنا أنه إذا كانت ٢ح مجهولة فبالإمكان استخدام مقدار سابق لها مثل الانحراف المعياري لعينة سابقة S وعندها تكون n معطاة بالمتباينة.

|Z ة2٦١

n>

d

لاحظ مدرت مادة طرق الإحصاء من خلال خبراته السابقة في تدريس هذه المادة : أن الوسط الحسابي لعلامات الطلبة في هذه المادة خلال السنوات السابقة كان 75 ويانحراف : معياري 9 علامات. إذا رغب هذا المدرس في تطوير أسلوب تدريس هذه المادة ومن ثم تقدير الوسط الحسابي للعلامات وفق الأسلوب الجديد وبحيث يكون متأكدا بنسبة 95% إن • الخطأ في المقدر الناتج لا يزيد عن 3 علامات فكم طالبا يحتاج أن يخضعهم لهذه التجربة؟

لاحظ أن d = 3, ء٢ = 9, 1 - a = 0.95

إذن z٠=zi.96

2

ولهذا فإن n هي أصغر عدد طبيعي بحيث أن

أي أن

(1.96x9

34.5744 الكا

إذن حجم العينة المطلوب n = 35 طالبا

أراد أحد الباحثين تقدير الوسط الحسابي لكمية الفوسفات في وحدة الحجم في مياه : إحدى البحيرات. ولقد كان معلوما من دراسات سابقة أن الانحراف المعياري لكميات الفوسفات في وحدات الحجم المأخوذة من هذه البحيرة هو 4.

كم عينة كل منها وحدة حجم يجب أن تؤخذ من مياه هذه البحيرة لنكون متأكدين : بنسبة 90% أن الخطا فييقدير ابسط لا ينيد عن 0.8؟

؛طية ؛;ذقمئج٠ الأ ـ٠؛دب٩ ؛٢) لى

: 1. عرف الثقدير بنقطة.

\ ة٠ ماذا نقصد بتقدير وسط المجتمع بفترة؟

: تمظير. ؛لتفدلليية بتقطية Propoithm ق® Estiaaatioa لا٠س p

إنه من المنطقي أن نسبة وجود ظاهرة في مجتمع ما يمكن أن تقدر بنسبة وجود تلك الظاهرة في عينة عشوائية تؤخذ من ذلك المجتمع.

فمثلا إذا أردت أن تقدر نسبة العائلات التي تستعمل الغسالة الأوتوماتيكية، فبإمكانك أن تختار عينة عشوائية وتحسب نسبة عدد العائلات التي تستعمل ذلك النوع من الغسالات. ثم تستعمل النسبة في العينة كتقدير نقطي لملنسبة في المجتمع.

أي أنه إذا كانت نسبة النجاح في تجربة ذات الحدين P يكون بإمكانك تقدير P كما يلي:

خذ عينة عشوائية حجمها n وافرض أن عدد النجاحات في هذه العينة X، يمكن استقمال ح = P

كمقرلنسبة النجاح P ويكون اللقليرالنقطى للمطمة P هونسبة النجاح فه العينة ٠=p

لاحظ أن * هو الإحصاء: عدد النجاحات في العينة، أما X فهو قيمة X التي تحصل عليها من دراسة العينة أيأن هو قيمة X التي تحصل عليها من عينة معينة.

والأمثلة على الحاجة لتقدير نسبة النجاح P كثيرة، فمثلا، تحتاج لمتقدير نسبة الطلبة من عمر 17 وحتى 21 سنة الذين يستعملون النظارات الطبية.

وتحتاج الى تقدير نسبة الطلبة من عمر8 إلى 12 سنة الذين يكتبون باليد اليسرى.

لتقدير نسبة المدخنين بين طلبة إحدى الجامعات قام باحث بمقابلة عينة عشوائية حجمها 200 طالب فوجد أن 70 طالباً يدخنون، فما نسبة الطلبة المدخنين في الجامعة؟

نسبة المذخنين في العينة 35 0 = حك تقدر نسبة المدخنين في الجامعة بنسبة تق" ٠200 ' *٠
الشلن لى العلة وهي 35%٠

٠ فرفة كبة !ابيوت التي يوجد فيه١ كفئة محفية في تفة ط; كذدى عنه؛ عشوائية حجمها 500 بيت ووجد أن 200 منها لديها تدفئة مركزية. ما هو التقدير النقطي : لنسبة البيوت ذات التدفئة المركزية في تلك المدينة؟

النسبة فى العينة 040 0 - تنك = Pي ٠ ٠ 500
التقدير النقطي للنسبة P في المجتمع هي P أي 0.40.

٠٠*®-١/؟.’٩٠.؛٠

(٠٣ ■ رح:؟

أردت معرفة نسبة المواطنين الذين يحبذون معالجة ماء الشربب بالكلور، فأخذت عينة حجمها 1000 مواطن فوجد أن 420 منهم يحبذون ذلك.

ما هو اللقلير القطى لفسبة المواطفن_الحقيقية النن يحيذون قلكة

بدون..

النسبة في العينة 42 0 = p= التقدير النقطي للنسبة الحقيقية في المجتمع - ي' ٠ 1000 - ت
- تتة ٠*
هو2 أي 0.42 .

يتيين أإيتتبيية بفةؤة !ص٣٠أ؟ يةثة JMmaffoB

إن تقدير النسبة بفترة هو عبارة عن إيجاد تقدير نقطي لنسبة النجاح في المجتمع Pثم إيجاد توزيع المعاينة لذلك المقدر واستعمال هذه المعلومات لايجاد فترة ذات معامل ثقة معين تحصر نسبة النجاح P داخلها.

إذا كان من المتوقع أن لا تكون نسبة النجاح غير المعلومة P قريبة جداً من الصفر أو الواحد، وكان حجم العينة n كبيراً فإن بإمكانك استحال النظرية التى تتيد أن تقنبع P = z يقترب هن النوع الطبيعي المعياري إذا كانت n كبيرة.

$\frac{\sqrt{p(l-p)}}{٢٢}$

إذا توفرت هذه الشروط أمكنك وضع العبارة الاحتمالية ...

$\underline{P(-Z_{„}n<z<z_{„}n}$ =l-a

z٠

تفضع لقونبه الطيعى المعيالفي ميا ث = X, P عندالسجاحات في الميلة التي حجمها n.

ومن الصعب استعمال العبارة ا لاحتمالية السابقة بصيغتها المعطاة لايجاد فترة ثقة لسسبة « وتله لن Pفى المقامفبقء١٢ محر مطومة، وﷲ ستس ع = P بدمج من P في المقام فنحصل على فترة الثقة (»- I) 100% التقريبية للنسبة P وهي:

فترة الثقة للنسبة P

إذا كانت = P نسبة النجاح في عينة عشوائية حجمها H، وكان n كبيراً، فإن فترة الثقة (1O.O% a -1) التقريبية لنسبة النجاح P. (p معلمة ذات الحدين، نسبة النجاح في المجتمع) هي:

حيث $z_{\alpha/2}$ هي النقطة على محور الطبيعي المعياري التي يقع إلى يمينها ي من المساحة.

لإيجاد فترة ثقة 95% لنسبة عدد الطلبة في المدارس الإعدادية الذين يستعملون النظارات الطبية، أخذت عينة عشوائية حجمها 900 طالب فوجد أن عدد مستعملي النظارات الطبية 100، أوجد فترة الثقة المطلوبة.

نسبة النجاح في العينة تساوي خ =.ـتة = p
•ح - ثم ي 9 900

بما أن حجم العينة كبير يمكنك استعمال فترة الثقة التقريبية التالية، حيث
0.05 -ه لأن 0.95 = ى-1

P-Z..O25 J!<p<p+Zi جم

بالتعويض نجد:

$$\frac{1}{9}-1.96\sqrt{\frac{\frac{1}{9}x\frac{8}{9}}{900}}<p<\frac{1}{9}+1.96\sqrt{\frac{\frac{1}{9}x\frac{8}{9}}{900}}$$

أي: 0.02 <0.02<p
9• 9

0.091<p< 0.131

لتقدير نسبة عدد الطلبة في إحدى الجامعات الذين رسبوا في ثلاثة مقررات أو أكثر خلال دراستهم، قام أحد الباحثين بدراسة عينة عشوائية حجمها 900 طالب فوجد أن 243 منهم قد رسبوا في ثلاثة مقررات أو أكثر.

أ- قدر نسبة الطلبة في الجامعة الذين رسبوا في ثلاثة مقررات أو أكثر.

ب- أوجد فترة ثقة 95% لهذه النسبة.

أخذت عينة عشوائية حجمها 400 من معلمي المرحلة الإعدادية فوجد أن 80 منهم ; حاصلون على شهادة البكالوريوس.

: أ- قدر نسبة المعلمين في المرحلة الإعدادية الحاصلين على شهادة البكالوريوس.

: ب- أوجد فترة تقة 99% للنسبة الحقيقية للمعلمين في هذه المرحلة الحاصلين على شهادة

: البكالوريوس.

من المعلوم أنه إذا كانت .د...2٠ت$_{ن}$ت$_{1}$ت عينة من توزيع ذات الحدين (بيرنولي) الذي معالمه (1,
P) فإن P تقدر بنقطة بالإحصاءه.

ث = P . ون الخطاالمعياري في هذاالمقدر لـ‍ـك-ـ■=(Var (p ومن ناحية لخرف فإن

فترة 100% (a -1) ثقة للنسبة P هي:

ل ٦ دي

شريطة أن يكون حجم العينة n كبيرا، ويسمى المقدار

■٠١.٥٥

حد 100% (a -1) للخطأ في تقديرو.

ومن الجدير بالملاحظة أن] غالبا ما تكون مجهولة، وإذا كانت هنالك معلومات سابقة من دراسات مماثلة فبالإمكان استخدام قيمة P المعروفة من تلك الدراسات السابقة أما إذا لم تكن هنالك أية فكرة عن قيمة P فإننا نأخذ بمبدأ أسوأ الأوضاع وهو أن تكون قيمة ل = P لأن نلك يودي إلى أكبر خطا معياري للمقر ونلك لأن

الاقتران g(p) = p(l-p),o<p<l

ياخذ أكبر تيمة له عندما خ= P ٠

ولهذا إذا تم تحديد المقدار الأكبر المسموح به للخطأ في تقدير]، أمكن حساب حم لعبة لللمة لتحقة

قه الحد عن لمة حم لهيبة d > سي

حيث d هو المقدار الأكبر المسموح به لهذا الخطأ وينتج من ذلك أن

دم لم-　　　　　　　2١

n>　　لم ى. P(l-P)

إذا كانت P معلومة من دراسات سابقة. أما إذا كانت P مجهولة بالكامل فإن

1 n> z_a 2
4 d

ر
لأ

مثالي إ7٤

في إحدى تجارب علم النفس، يسمح للأشخاص الخاضعين لإحدى التجارب بالاستجابة لأحد مؤشرين A أو B، ويريد الباحث أن يقدر نسبة الأشخاص الذين يختارون المؤشر A. ولنرمز لهذه النسبة بالرمز P.

كم شخصاً يجب أن نخضع لهذه الدراسة كي نكون وإثقين بنسبة 90% أن الخطأ في تقدير1 لا يزيد عن 0.04 فيكلمن الحالتين التاليتين.

أ- إذا كنا نعلم أن P حوالي 0.2.

ب- إذا لم يكن لدينا أية فكرة عن قيمة P.

اأيحلو

= 0.90 أ- 64 . o5=l . d = O.O4, Z_a=z_o

(ه- ومنها 2 (1

n> . P(l-P)

d

p = o.2

لأ. ر

n>. ١ام64٠ (0.2)(0.8)

0.04

n 268.96ج

ومنها حجم العينة المطلوب n = 269

ب- بما ان] خر

معلومة نضع مكانها [z

n> نكون

2

أسوا تيمة p = 1

ة

ر
لأ.

يى (٠٠

n> 420.25

ومنها حجم الغينة المطلوب n = 421

يراد تصميم دراسة طبية لتقدير نسبة المواطنين الذين يعانون من مشاكل في النظر، كم شخصا يجب فحصيم كي نكون واثقين بنسبة 98% أن الخطأ في تقدير هذه النسبة لا يزيد عن 0.05 في الحالتين التاليتين.

أ- إذا لم يكن لدينا أية معلومات سابقة عن هذه النسبة.

ب- إذا كنا نعلم من دراسات سابقة أن هذه النسبة قد تكون حوالي 0.3.

عرف تقدير النسبة بنقطة وتقدير النسبة بفترة.

مغاهيم.ز اا.-الغرمئبغس٠س0ابغا١ 1

الفرضية الإحصائية هي كل عبارة تكون صحتها أو عدم صحتها بحاجة إلى قرار.

إذا كان وسط طول الجندي في أحد الجيوش حوالي 170 سم، فالجمل التالية تعتبر : قرضيات حول هذا الوسط:

العبارة كلاماً:

وسط الطول 170 سم .

وسط الطول أقل من 170 سم

وسط الطول أكثر من 170 سم

وسط الطول لايساوي 170 سم

العبارة رياضياً:

H 170

H 170

H 170

H*17O

في معظم الأحيان هناك نوعان من الفرضيات في المسألة الواحدة. النوع الأول هو الفرضية الصفرية وهي الفرضية التي تبنى على أمل أن يتخذ قرار بعدم صحتها، ونصطلح الآن على أعتبارأي فرضية نود اختبارها فرضية صفرية، ونعبر عن ذلك بالرمز H٥ . وهكذا، فكل فرضية إحصائية تريد اختبارها تسمى فرضية صفرية H_o ٠

إن رفض الفرضية الصفرية H٥ يؤدي الى قبول فرضية أخرى تسمى الفرضية البديلة، ونعبر عنها بالرمز Hj’

وتصاغ الفرضية الصفرية المتعلقة بمعلمة مجتمع معين بشكل يعين قيمة محددة لتلك المعلمة. ففي المثال أعلاه تعتبر الفزضية: معدل الطول 170 سم أي H_o : *[1 =* 170 فرضية صفرية، وهي كما تلاحظ تعين قيمة محددة للمعلمة j] هي 170 سم.

أما الفرضية البديلة فيمكن أن تعين قيمة محددة للمعلمة تحت الدراسة كما تسمح

بأن تأخذ المعلمة قيماً متعددة. ارجع إلى المثال نجد أن الفرضية:

وسط الطول أكبر من 170سم أي170 <لإ ;H) تعتبر فرضية بديلة، وهي كما تلاحظ تعطي إمكانية أن يأخذ الوسط أي قيمة أكبر من 170 مثل 172.5 أوأي قيمة أخرى أكبر من 170 سم.

كيف تختبر الفرضيات الإحصائية؟ أي ما هي القواعد لعملية اتخاذ القرار برفض أو عدم رفض الفرضية الصفرية H٠ لصالح الفرضية البديلة H؛ ؟

سنبدأ بشرح اختبار الفرضية الصفرية 170 =,لر ; H_0

مقابل الفرضية البديلة 70 ايو دز : ٠لل حيث لإهي معدل طول الجندي في أحد الجيوش كما في المثال.

إن الفرضية في هذا المثال تتعلق بمعدل المجتمع ولذلك كان من المنطقي ان يبنى اختبار هذه الفرضية باستعمال معدل العينة (الوسط الحسابي لها).

إحصاء الإختبارA test Statistic.

هو احصاء (اقتران تعين قيمته من العينة) يبنى عليه قرار إختبار الفرضيات. وهكذا فإن هو إحصاء إختبار.X

المنطقة إلحرجة للإختبار هي مجموعة قيم إحصاء الاختبار التي تؤدي الى رفض : الفرضية الصفرية، كل حد من حدود المنطقة الحرجة يسمى قيمة حرجة لإحصاء الاختبار.

: مقابل الفرضية البديلة 10٠19O=H₀:pتريد إختبار الفرضية

أوجد القيم الحرجة : O. باستعمال الوسط الحسابي لعينة حجمها 100 = ه اذا كان 24 =

إذا كانت القاعدة:z, Xوالمنطقة الحرجة لكل من

أكثر من انحرافين معياريين للوسط الحسابي عن القيمةX إذا بعدت H_0ارفض

.190

بما أن 190 # H_t:g فهذا يعني أن القيم الحرجة هي:

-20-190,ج20 + 190

الآن2'4:د:ل،

إذن القيم الحرجة لإحصاء الاختبار X هي:

190-2x2.4= 185.2

190 + 2x2.4 = 194.8

والمنطقة الحرجة هي 185.2 ى* و 194.8 ح* أي ارفض،]] إذا كان الوسط الحسابي للعينة X يساوي أو أصغر من 185.2 أو أكبر من 194.8 أو يساويه.

بما أنه يتم الرفض إذا كان الوسط الحسابي على بعد أكثر من انحرافيين معياريين للوسط عن 190 فهذا يعني أن القيم الحرجة لإحصاء الاختبار Z هي:

2-- ك,2ح ك

ويمكن الحصول على هاتين القيمتين من المعادلة

X■؛ 194.8-190 -2

185.2-190 i 2

والمنطقة الحرجة: ارفضة[إذا كان 2-ك2 أو 2ح 2

كل قرار يبنى على نتائج عينة يكون معرضاً للخطأ، حيث أن مثل هذه القرارات تبنى على متغير عشوائي أو إحصاء مثل X . وبالطبع فمن الممكن أن تحدث نتائج ولو كان احتمال حدوثها صغيراً جداً. وفي حال حدوث نتائج مع أن احتمال حدوثها صغير جداً فهذا يعني أننا وقعنا في الخطأ. وفي اختبار الفرضيات، هناك حالتان بالنسبة للفرضية الصفرية، هما: إما أن تكون هذه الفرضية صحيحة وإما أن تكون غير صحيحة. هناك أيضاً نوعان من القرارات التي يتخذها الإحصائي بصدد الفرضية الصفرية، هما: رفض الفرضية الصفرية أو عدم رفضها.

فإذا كانت صحيحة وكان قرارك رفضها، فمعنى ذلك أنك وقعت في الخطأ. وهو يسمى في هذه الحالة: الخطأ من النوع الأول.

إما إذا كانت الفرضية الصفرية غير صحيحة وكان قرارك عدم رفضها، فمعنى ذلك أنك وقعت في خطأ أيضاً، لكنه في هذه الحالة يسمى: الخطأ من النوع الثاني.

يحدث الخطا من النوع الأول إذا رفضت الفرضية الصفرية وهي في الحقيقة : صحيحة. إ

٠٠٠غ.

ور يحتتفنامن العتنتي بجالرتنت لقتتية اصتنلة. وحيف يالحققة. غير صحيحة بالنظر إلى

هذا التعريف، يظهر لك أنه ربما تكون الفرضية البديلة صحيحة : (الفرضية الصفرية غير صحيحة) ولكن النتائج من العينة لا تؤدي الى رفض الفرضية : الصفرية، في هذه الحالة يكون قد حدث خطأ من النوع الثاني. وبمعنى آخر، يحدث الخطأ من النوع الثاني عندما لا يتمكن الاختبار من التوصل إلى أن الفرضية البديلة هي : الصحيحة. :

أجريت دراسة منذ عدة سنوات حول الساعات في الأسبوع التي يقضيها طلبة الصف الثالث الثانوي في التحضير لامتحان شهادة الدراسة الثانوية فوجد أن معدل عددها 40 ساعة والانحراف المعياري 12 ساعة.

وبسبب عدم معرفتك اتجاه التغير الذي ربما حدث في سلوك الطلبة في هذا الموضوع، قررت إجراء اختبار الفرضية الصفرية 40 : غز H_0: .

مقابل الفرضية البديلة 40محلإ.٠H1

ولإجراء الاختبار أخذت عينة حجمها 100 طالب وعبرت عن وسطها الحسابي بالرمزx إذاكان القرار: رفض x إذاكان على بعدأكثرمن انحرافين معياريين للوسط عن القيمة 40 = لمز.

أوجد القيم الحرجة والمنطقة الحرجة واشرح حالات وقوع الخطأ من النوع الأول : والخطأً من النوع الثاني.

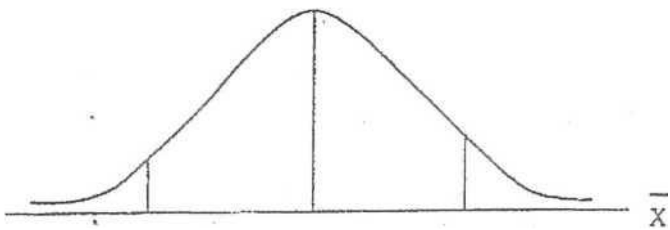
بما أن الفرضية البديلة هي 40 عو دا٠:للأ فهذا يعني وجود قيمتين حرجتين هما بدلالة

x

40-2٩

4O-f-2٠j

لبن، الفمة الحرهة السرع، 7.6و: 00ـر٦ X 2 - 40



وافمة الحهجة المدى 42.4 : ت' ـعـ X 2 ب 40 وامسقة العجة الى سار 37.6 أو إلى يمين 42.4

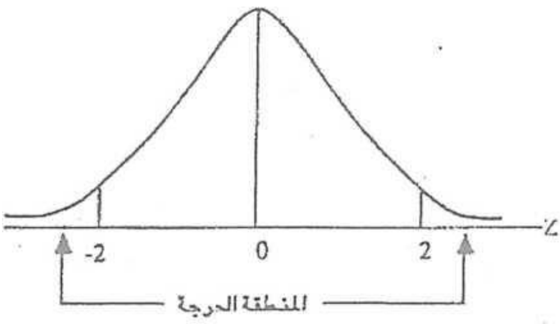
أي: أرفض H٠ إذا كان 37.6ح* و 42.4ح5 أما بدلالة إحصاء الإختبار Z فإن المنطقة الحرجة هي:

ار H إذا كانZC-2 أو 2>z. أنظر الشكل (4) والشكل (5)

؛37.6 ٠ ٠٠40 42.4 أ

40

ل الذطت٦الحرجة الشكل (4)



الشكل (5)

وهي صحيحة) عند تحقيق الحالتين: H_o يحدث خطأ من النوع الأول (إذا رفضت

• p. معدل عدد ساعات الدراسة لدى طلبة الصف الثالث الثانوي 40 = .1

X. معدل العينة 5 يقع في المنطقة الحرجة أي 37.6 ك* أو 42.4 < 2

ويمكن التحقق من (2) إذا وقعت Z في المنطقة الحرجة أي 2- > z أو 2 < z حيث 12/1 2 x-4'o

أما الخطأ من النوع الثاني فيحدث عند تحقق الشرطين:

1 . معدل عدد ساعات الدراسة لدى طلبة الصف الثالث الثانوي لا يساوي 40 أي 40 م: دا

2 . معدل العينة 5 لا يقع في المنقطة الحرجة أي أن 42.4 ج X كا 37.6 أو بتبارة أخرى 2ك ك ك 2-.

والآن، تعرفت على الخطأ من النوع الأول والخطأ من النوع الثاني فكيف يتم حساب احتمالات هذه الأخطاه؟

من تعاريف هذه الأخطاء تستطيع حساب احتمالات حدوثها كما يلي:

احتمال الخطأ من النوع الأول، وتعبر عنه بالرز هع هو:

a = p (H_O رفض ا صحيحة[*)

أي: a تساوي احتمال رفض الفرضية H_o إذا علم أن H٠ صحيحة.

احتمال الخطأ من النوع الثاني، وتعبر عنه بالرمز P هو:

p = p (H_o عدم رفض ا صحيحة Hl)

أي: P تساوي احتمال عدم رفض الفرضية H٠ إذا علم أن Hi صحيحة.

عند حساب الاحتمالات السابقة لاحظ أن الحادث "رفض الفرضية H_o " يعني وقوع إحصاء الإختبار في منقطة الرفض، وأن العبارة " HQ صحيحة" تعني استعمال قيمة المعلمة المحددة لك في الفرضية الصفرية.

إن الحادث "عدم رفض H_o " يعني وقوع إحصاء الإختبار في متممة منطقة الرفض، أي متممة المنطقة الحرجة.

والعبارة ",H صحيحة" تعني إعطاء قيمة معينة للمعلمة تحت الإختبار على أن تكون هذه القيمة من القيم المسموح بها تحت الفرضية البديلة H_l .

ويمكن توضيح الخطأ من النوع الأول والخطأ من النوع الثاني والاحتمالين p , a بالرسم كما يظهر في الشكل (6) الذي يمثل إختبارا حول وسط المجتمع لؤ ٠

0دإ ير:ح[مقابلة الفرضية لبديلة 0دإ<عإ:الل وأخذت قيمة محددة

0دل <ادز لتحسب P المقابلة لها.

V كنم٠ددعكرتبج ندئ

1ط سرج' ةت٠ نئوم

النرع ال ول!

افك (6

م؛ ل ص فت (20) ض امىة تا تر جقغ- وسق س ;ه

الثاني P عندما يكون41 = ع ,38 = غع ,_43_=_عز_,45= ح ٠

a = P جت الفرضية رفض ا صحيحة H_o)

ه p(x< 37.60 أو X > 42.4ددا = 40)

وبتحويل ٤5 إلى2 وايجاد القيم المعيارية المقابلة للقيمتين 42.4,37.6 تجد

a=P(Z<-2j١Z>2)

= 2x (0.5000-0.4772) = 0.045

والآن احسب P لكل قيمة معطاة:

p = p (H٥ عدم رفض I صحيحة H))

P1=P (37.6 -

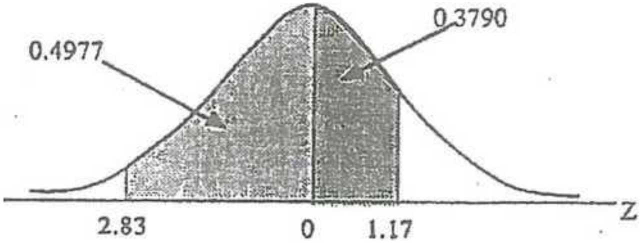
- بم3أَ-6- 41 <z< 42.4-41}

-1(-2.83ك Z اكا.17)

= 0.3790-+0.4977

-0.8767

أنظر الشكل (7)



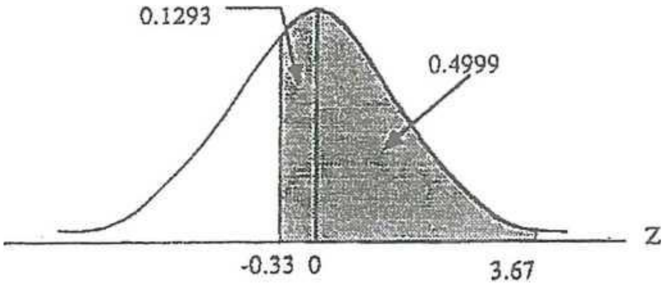
)7الشكل (

(3₂=p(37.6 كا• XS 42.4 ا (38

-·S“,S

-1(-0.33كككك3.67)
= 0.1293 10.4999 = 0,6292

أنظر الشكل (8)



الشكل (8)

03 =p(37.6< X كا 42.4داا=43)

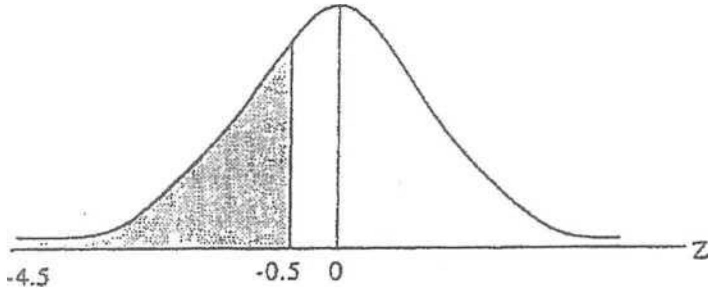
.'S··S

= p(- 4.5 كا كاة-0.5)

-0.5-0.1915

-0.3085

أنظر الشكل (9)



الشكل (9)

(45 = لإا42.4 x> ك37.6)? = م3

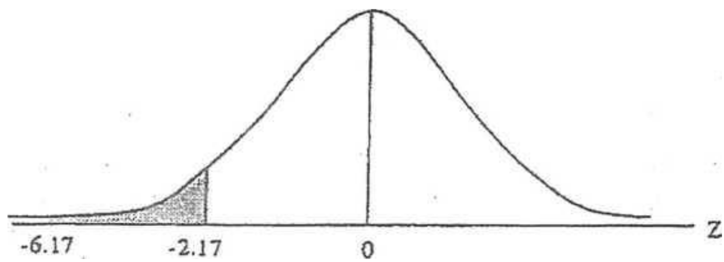
١٥٠

(7ا.2-كةك6.17-)? =

= 0.5-0.4850

-0.015

أنظر لشكل (10)



الشكل (10)

تلاحظ في هذا المثال أن احتمال الخطأ من النوع الثاني يكون كبيرا إذا كانت قيمة لمز في الفرضية البديلة، قريبة من قيمة لما في الفرضية الصفرية. 41= 4! في الفرضية البديلة قريبة من 40 = ^:H٥٠ وبالتالي كان احتمال الخطأ من النوع الثاني كبيرا وذلك لأنه يصعب على الإختبار

التمييز بين القيمتين القريبتين من بعضهما البعض، بالتالي، فإن احتمال عدم رفض H_oبينما H_x صحيحة يكون كبيرا (وهو قيمة و).

أما في حالة الفرضية البديلة التي تكون فيها قيمة لما بعيدة عن قيمة عإ في الفرضية الصفرية فإن الإختبار يستطيع التمييز بين القيمتين، فلا يقبل H_o عندما تكون ,H صحيحة (في مثالنا 45 = 1) بعيدة عن40ملإ وبالتالي كانت P_4 صغيرة جدا).

لاحظ أن احتمال الخطأ من النوع الأول a يتم حسابه من المنطقة الحرجة عندما تكون H_o صحيح، ولاحظ في هذا المثال أن هي مساحة الطرفين تحت توزيع 2 أو ة وإلى يمين القيمة الحرجة اليمنى وإلى يسار القيمة الحرجة اليسرى.

وفي هذا المثال أيضا نفسر احتمال الخطأ من النوع الأول 0.0456 — a بأنه يعني أننا بتكرار تطبيق هذا الاختبار واعتبار الفرضية الصفرية صحيحة، فإن حوالي 4.56% من الإختبارات ستعطي قرارا خاطئا برفض H٥ بينما هي صحيحة.

مستوى الد لالة لاختبار ما، هو احتمال الخطأ من النوع ا لأول a ، ويعبر عنه كنسبة مئوية وهكذا، إذا كان 0.05- *a* تقول: مستوى الدلالة للاختبار 5%.

وعادة أن يكون مستوى الدلالة ا% ل5% هما الأكثر استعمالاً في اختبار الفرضيات. ويكون بناء الاختبار ذي مستوى الدلالة a أولاً: بتحديد وجود طرف واحد للاختبار أو طرفين، أي وجود قيمة حرجة واحدة أو قيمتين ويعتمد هذا على الفرضية البديلة. فإذا كانت الفرضية البديلة من النوع H_1:(X<P،O أو 0دإ > دج :؛H كان للاختبار قيمة حرجة واحدة، أي كان الاختبار ذا طرف واحد؛ وإذا كانت الفرضية البديلة من النوع 0ا) م H,:(i كان هناك قيمتان حرجتان وكان للاختبار طرفان.

ثانياً: عليك أن تجد القيمة الحرجة (القيم الحرجة) بحيث يكون احتمال الخطأ من النوع للأول يساوي Ct ، أي أن المساحة في أقصى يمين أو يسار القيمة الحرجة تساوي Ctفي حالة الاختبار ذي الطرف الواحد. ويكون مجموع المساحتين في أقصى يمين القيمة الحرجة اليمنى وفي يسار القيمة الحرجة اليسرى يساوي Ct في حالة الاختبار ذي الطرفين، علما بأن هذه المساحات تحسب على أساس معرفة التوزيع الاحتمالي لإحصاء الاختبار.

إذا كان معدل نزول المطر على مدى السنوات السابقة في إحدى المحافظات هو 480 ملم وبانحراف معياري 40 ملم. أردت إختبار فيما إذا زاد معدل نزول المطر هذه السنة عن المعدل العام ولنلك أخذت عينة من 36 موقعا ووجدت أن المعدل 490 ملم وإذا قررت : أن ترفض الفرضية إذا كان معدل التينة يبتد عن المعدل العام بأكثر من انحرافيين معياربين : للوسط الحسابي فأوجد المنطقة الحرجة وإختبر الفرضية المذكورة.

بناء على دراسات سابقة، يعتقد أن 0.15 من ظلبة المدارس الثانوية مدخنون. قامت : وزارة الصحة بحملة تتقيفية ضد التدخين لمدة عام، على أمل أن تؤدي الى تقليل النسبة المنكورة، ولإختبار هذه المقولة، أخنت عينة حجمها 400 طالب من المدارس الثانوية فوجد أن عدد المدخنين 48 أوجد المنطقة الحرجة بدلالة P (نسبة النجاح في العينة) ويدلالة 2، وقم بإختبار الفرضية إذا كانت P تبعد عن 0.15 بأكثر من انحرافيين معياريين لنسبة النجاح. :

أجريت دراسة حول عدد ساعات تشغيل الطلبة في احدى الجامعات فوجد أن معدل عدد الساعات 16 ساعة والاتحراف المعياري 7ساعات.

ولمعرفة التغير في عدد ساعات عمل الطلبة قرر باحث اختبار الفرضية الصفرية:
Ho: (16

مقابل الفرضية البديلة 16 ب د) :H٦

وقرر أن يرفض H_o إذا وقع X (معدل العينة) بعيدا عن 16 بأكثر من انحرافين معياريين للوسط الحسابي.

أحسب a ثم أحسب P للقيم 15= Ji ,18 = 1) إذا كان حجم العينة التي أخذها الباحث 50.

سسسسرساترمست

حسب الدراسات السابقة، يبلغ معدل إنتاج الدونم من القمح 250 كغم بانحراف معياري 35 كغم. بعد استتمال نوع معين من السماد الذي يعبقد أنه يساعد في زيادة إ المحصول، أخذت وزارة الزراعة عينة عشوائية من 50 قطعة مساحة كل منها دونم واحد . ' فوجدت أن معدل الإنتاج 260 كغم للدونم. هل توافق أن لهذا السماد تأثير إيجابي في زيادة إ المحصول؟ أذكر الفرضية الصفرية واذكر الفرضية البديلة.

إذا قررت رفض الفرضية الصفرية إذا زاد متدل العينة عن 250 بأكثر من انحرافين

• معياريين للوسط الحسابي، أوجد منطقة الرفض.

احدب احتمال الخطأ من اكوع الأول •   إ٦مج٠آح

اب: 1ذا كفت نسبة نخين فن يتون قرت آنق فذ ندرس

بلد ما هي 0.11. لإختبار فيما إذا زادت هذه النسبة بعد القيام بمسح شامل لتلاميذ المدارس . وتشجيعهم على استعمال النظارة الطبية لمن يحتاج إلى ذلك، أخذت عينة عشوائية من 500 طالبا فوجد أن 65 منهم يستعملون النظارات الطبية. عين الفرضية الصفرية، والفرضية البديلة إذا قررت رفض الفرضية الصفرية، إذا زادت نسبة مستعملي النظارات الطبية P عن : 0.11 بمقدار انحرافين معيارين للنسبة P ، أوجد منطقة الرفض، واختبر الفرضية الصفرية.

احسب a•

٠ئر. ١١سم٩ة٨:؛ج.٦|إ!| ئ . .ذ٠٠ئ||

.ر-;٦٦حزبرزغذة|دحد:همةجقم٢ذخ8تتة 5■'■:

؛ج لآ ؛طزئنم :لرضيغئ: تجيئة يالثنض أهيتفع طبيوي بيايذد معلثث

Testing Hypotheses CGKcernbg the lean ,٠zul DistHs with KaiK ١v٦rbic6

إذا كانت عإ الوسط الحسابي لتوزيع طبيعي تباينه ٠2 معلوم واختيرت عينة عشوائية

حجمها n وكان وسطها الحسابي X فإن توزيع X هو الطبيعي ذو الوسط 1؛ والتباين o^2/n

المطلوب اختبار 0لإ = د! :H_o

مقابل إحدى الفرضيات البديلة:

أ (i) H_1:p.>p٠0

ف (ii)H هلمإ>

أو (iii) 0لإ مج لإ:٩

على مستوى دلالة a

بما أن التوزيع طبيعي، استعمل Z لتعيين القيمة (أو القيم) الحرجة بدلالة Z ثم استعمل

المعادلة ئ:ة لتعيين القيمة (و القيم) الحرجة بدلالة X

<13 ٦٢لآ

بالنسبة للفرضية البديلة (؛)، القيمة الحرجة Z ويقابلها X1 من المعادلة:

Z - Xi ~ نلخم

تمذ “■

بالنسبة للفرضية البديلة (لأ)، القيمة الحرجة Z-ويقابلها *Xz* من المعادلة:

٠ Z- دئ *I*

بالنسبة الفرضية البديلة (إلا)، فإن الاختبار المناسب ذو الطرفين، والقيم الحرجة

هي: , , *z٥(٦ حاعأ٠*

ويقابلها: X1 , *X1* من المعادلتين:

Z — X1 - هداد

دك ٠٠

لأح Z;دن *"٠٠*

ولأجراء الاختبار أحسب x من العينة التي اخترتها، وقارن هذه القيمة بالقيمة الحرجة حسب الفرضية البديلة المعطاة

الحالة (i) : 0لمز حلما:بلآ ، ارفض H_o إذا كان X>X1٠

الحالة (لآ): 0لإ>لإ:ا11 ، ارفض Ho إذا كان $x<x_2$٠

الحالة (iii): 0لإ * عإ:الآ ، ارفض Ho إذا كان $x>x_1$ أو $x<x_2$٠

وهناك طريقة أخرى لاختبار الفرضية HQ السابقة وذلك بتعيين القيم الحرجة بدلالة z ثم حساب قيمة z المقابلة للوسط الحسابي x الذي وجدته من العينة، وبعد ذلك قارن قيمة z التي نحصل عليها با لقيمة الحرجة حسب الفرضية البديلة:

الحالة (1): H > Ho : H٠ ، ارفض H_o إذا كان z >Za .

الحالة (ii) ,هم- ارفض H_o إذا كان Z<-Za.

الالة (ئ) H, :H * Ho ارفض Ho إذا كان z Z أو z < -z٠٥ ٠

فست الر ءبرإت أحد سأحيق فسن تتوتيع طبيعي ,نعزفه ست ج ومعدله لل ، بالإعتماد على عينة بحجم n = 25 .

على مستوى دلالة 5% اختبر الفرضية؟

H_o :H = 2OO مقابل الفرضية البديلة

200ددا: Hj، إذا كت الوسط لص لعيبة حجمها X = 208 غم.

(1) H_o:^ = 2OO

(2) H_t: 200 ب عج

(3)0-0.05

(4) بما أن الفرضية البديلة لا تحدد اتجاهاً واحداً، فإن الاختبار المناسب ذو طرفين، والقيم الحرجة تكون يد-

والمنطقة الحرجة هي: أرفض HQ إذا كان z>1.96

أو1.96-<ة أنظر لشكل (11).

(5) أحسب z من المعادلة عئ =z

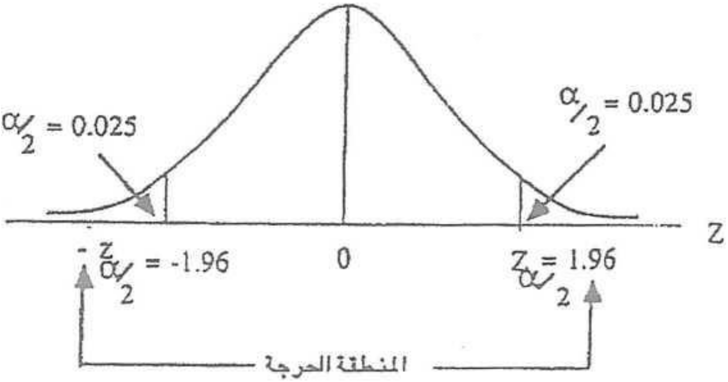
<3؛ ة

أى: 5.71:يخأ:"٢٠

٦٠٠

(6) قارن قيم Z بالقيمة الحرجة:

من الواضح أن 1.96 >5,71 أي أن قيمة Z تقع في المنطقة الحرجة، لذلك أرفض H_o لصالح 200 < P ولاحظ أن النتيجة كانت لصالح 200 < دع لأن Z وقعت في منطقة الرفض اليمنى، أي على يمين Zj



الشكل (11)

يخضع الزمن الذي يحتاجه الطالب للتسجيل في إحدى الجامعات الى توزيع طبيعي انحرافه المعياري 0.7 ساعة ومعدله ئا ساعة.

اختبر الفرضية H_o: p. = 5.2 ساعة

مقابل الفرضية البديلة H؛:p> 5.2

باستعمال مستوى دلالة 5% إذا أعطيت عينة حجمها 16 ووسطاً حسابياً 5.3 = جذ ساعة.

(1) H =5.2

(2) H >5.2

(3) $a = 0.05$

(4) بما أن الفرضية البديلة ذات اتجاه واحد هو" أكبر من" فإن الاختبار المناسب ذو طرف

واحد والقيمة الحرجة 1.645 — 05 ﻩz

(5) إن القيمة الحرجة بدلالة X هي:

1.645:11
ؤ/٦ .د٢

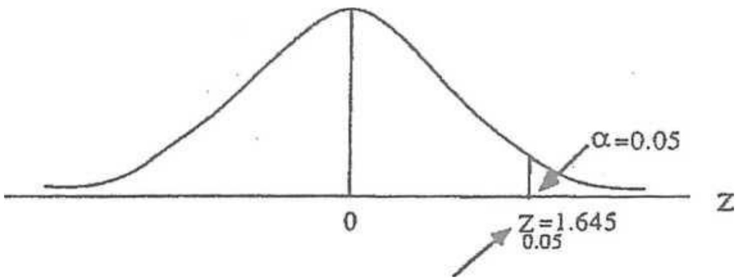
0.05ك ت 1-645

إذن تنلكخيبكل ب 5.2 = Xi

4

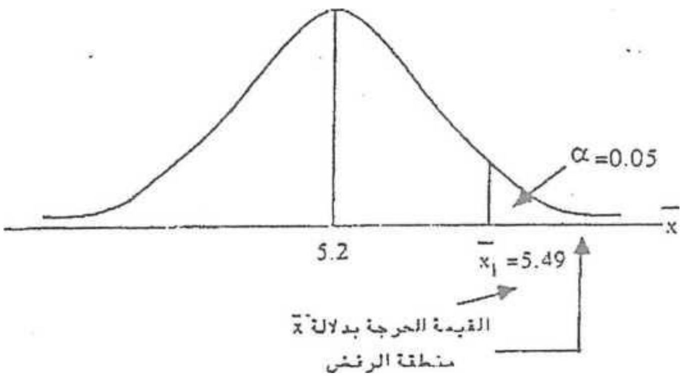
= 5.49

(6) قارن قيمة X التي حصلت عليها من العينة مع القيمة الحرجة X1، تلاحظ أن 5.49> 5.3 إذن لاترفض H أنظر الشكل (12) أ، ب



القيمة!لحدجة بدلالة 2

الشكل (12-أ)



الشكل (12-ب)

أخذت عينة عشوائية حجمها 10 من مجتمع طبيعي تباينه 150.
اختبر الفرضية
مقابل الفرضية 800<.در: ,H باستعمال مستوى معنوية1% إذا كان الوسط :
الصابي لكنة x = 812 ؛

m

قلقية لم

عندما تقارن قيمة Z بالقيمة الحرجة حسبب الفرضية البديلة ينتج 3 حالات:
الحالة800:1ح: وت+ أرفض ه11 إذا كان -
الحالة 2: م•ير:ج]أرفض مت1 إذا كان -
الحالة 3: 0لإ ب ج:Hi أرفض H_o إذا كان ——

إن عملية اختبار الفرضيات في هذه الحالة هي نفسها التي في البند (1.4) مع اختلاف واحد، وهو أنك تستعمل الانحراف المعياري للعينة 5بدلا من الانحراف المعياري للتوزيع •، وذلك لأن ه غير معلومة. أما تبرير إمكانية استعمال إحصاء الاختبار Z في هذه الحال فهو أن حجم العينة كبير أي 30 ة n .

تخضع أعداد حبات التفاح على شجرة التفاح في بستان كبير لتوزيع طبيعي وسطه 150 حبة، بدأ مالك البستان استعمال نوع جديد من السماد وأراد أن يختبر ما إذا زاد الإنتاج تبعا لذلك لذا أخذ عينة من 64 شجرة، فوجد أن الوسط الحسابي لأعداد الحبات في العينة 156 بانحراف معياري 12 حبة. هل تشير هذه البيانات إلى الزيادة في الإنتاج على مستوى دلالة ى = 0.05.

إذا لم يكن هناك زيادة في إنتاج الأشجار التي تم تسميدها فهذا يعني أن معدل عدد الحبات يكون 150= در، ما إذا كان هناك زيادة في الإنتاج فهذا يعني أن المعدل سيكون أكثرمن 150.

فالمطلوب اختبار:

(1) H°:ji = 150

(2) Hj: Ji > 150 مقابل الفرضية

(3) a = 0.05 مستوى الدلالة

(4) الفرضية البديلة ذات طرف واحد (أكبر من)

إذن، القيمة الحرجة $Z_{Q.05} = 1.645$

(5) **أحسب** 2 **من المعادلة:**

0؛Z-X-i

١ 156-150

أفي لآ

(6) قارن قيمة Z التي حصلت عليها مع القيمة الحرجة:

Z=x8=4
12

من الواضح أن 1.645 ح4 **، أي أن قيمة** Z **تقع في المنطقة الحرجة. لذلك ارفض لصالح** H,: Ji >150. **انظر الشكل** (13).

نس اسناً الحرجة

الشكل (13)

وجد في دراسة سابقة أن معدل قيم الفواتير في أحد المستشفيات 70.2 دينار وأن توزيعها يقرب من التوزيع الطبيعي.

أردت اختبار فرضية أن قيم الفواتير قد تغيرت، فدرست 64 فاتورة أخذت عشوائية فوجدأن 73.7-

* أن 11.2 = 8.

باستعمال مستوى دلالة 5% اختبر الفرضية 70.2 = أ:H_o مقابل الفرضية البديلة

: المناسيبة. ;

/7/أ

. خئدل ____ ____ ئؤ

يبلغ معدل طول الجندي في أحد الجيوش 169 سم وفي السنوات الأخيرة بدأ الإقبال على الجندية يزيد وبالتالي صارت القيادة تضع شروطاً أشد بخصوص الطول. اختبر الفرضية التي تقول أن معدل طول الجندي قد ازداد، علما بأن عينة عشوائية حجمها 50 من . أفراد ذلك الجيف أعطت معلل الطول 71,5 اسم والانحراف المعياري 5 سم استعمل = 0.05 a٤.

رئدت• كذئ؛٠. ’> ٢.٠٢-.حخاً٦ذل ت وئ?■ فاق حمهئميب٦وء:ةقب ب،؛تدنمج٦ضكت'،;و

و٠بلمز٢أ٢غ٠ ■ بمبيجب، *وعيدبية ،فيجية،* حبيه.*

تثوثي Tiiig

)■ئ دلآ لآف لز؛لاجئ٢تلأكلاً (.sow؛! خ٠٦ i؛ p٠٠tان عي الايور في همذ الحاًلة هي ط٠ئ

شمجكةية هذه الحالة يفترض أنه طبيعي، ولكن حجم العينة كبير 30 < ه. وبالتالي، فإن المبرر

لاستعمال إحصاءالإختبارZ هونظرية النياية المركزية والتى تجمن لك أن ذ- =2 عر٠/٦
له توزيع قريب من التوزيع الطبيعي المعياري إذا كانت ع كبيرة.

لاحظ أنه في حالة ٠ غير معلومة بإمكانك الاستعاضة عنها باستعمال S، الانحراف المعياري للعينة وذلك لأن حجم العينة n كبير.

على مستوى 5% اختبر الفرضية H_0 :fl5

مقابل الفرضية البديلة 15 تيو.ا-} :Hj

\ إذا اعطيت عينة حجمها n 81 وسطا حسابيا 13.5 =* ، وانحرافا معياريا !

ن 3.3 = 8. ا

(1) H =15

(2) H * 15

(3) a = 0.05 :مستوى الدلالة

(4) الفرضية البديلة ذات طرفين. إذن، فالقيم الحرجة هي:

■ZQ.025 :-1-96, ZO.O25 = 1.96

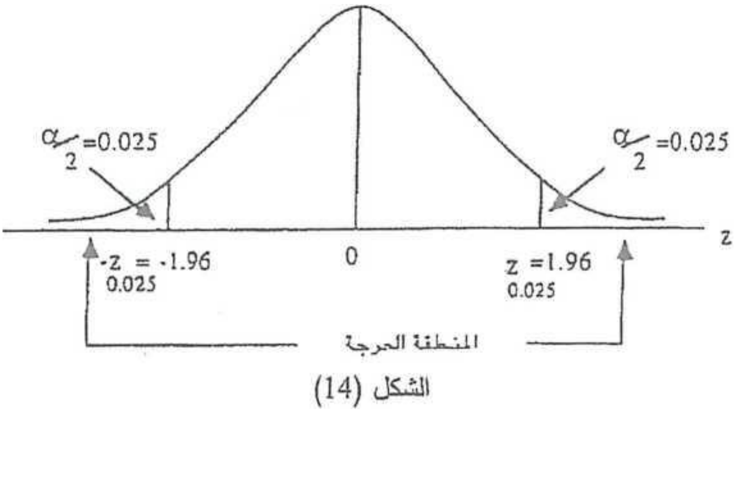
(5) أحسب 2 من المعادلة:

z= X- لأج

= قيتئيل=ي=-4.09

3.3 م 3.3

6. قارن قيمة Z التي حصلت عليها مع القيم الحرجة: من الواضح أن 1.96- > 4.09- إذن، أرفض H_0 لصالح 15يل!. وقدقلت لصالح 15>لإ لأن Z وقعت في المنطقة الحرجة إلى يسار القيمة الحرجة اليسرى.



الشكل (14)

أظهرت دراسة سابقة أن التباين في أعمار المصابيح الكهربائية التي ينتجها أحد المصانع يساوي (170) ساعة تربيع. اختبر الفرضية HQ : Ji = 850 مقابل الفرضية البديلة ؛ 850 ءاً لا:H٠ ■

إذا أعطيت عينة عشوإئية حجمها 100 مصباح وسطا حسابيا 839 = x٠ استعمل : ا مسقي دلالة_5%.

ذ٩1ررح|

؛ ه تييي سن عدد الناًس. ستن في سجت لما تلصةبكة

لمسسكرك الفصول، أخذت عينة عشوائية حجمها 150 طالبا، فأظهرت العينة أن ■5.13 = X وأن : s=3.3 على مستوى دلالة 1% اختبر الفرضية

4ا:ع: H_o مقابل الفرضية اليديلة؛!؛لمب: ب؟

عزيزي الدارس، عليك الإجابة عن جميع التدريبات المطروحة في ثنايا الوحدة، وأنصحك بإعادة دراسة الجزء المتعلق بكل تدريب إذا وجدت الإجابة عسيرة ولم تكن متأكدا منكلنقطة فيها.

تأكد من صحة إجابتك بالعودة الى مفتاح الإجابات المرفق في نهاية الوحدة.

إذا كان لديك أي تعليق أو توضيح أو استفسار فاكتب إلى مرشدك، فإنه يسره أن يساعدك في ذلك.

؛هقلج.. :نفيبةسيأبي بعوط. ؛٠ئ عأ.ش؛ك ترر دم ي,جم ابضيتية ■ئي٠د

ة 'Ute دجTestiag Hypotheses CdSEibg the lean of دكلاًلاً SI؟;؛:؛Normal hpdatfos? wi

لاحظ أن هذه الحالة تختلف عن تلك التي في (1.4) من حيث أن التباين a^2 غير معلوم، وأنها تختلف عن الحالة في (2.4) من حيث حجم العينة صغير، وبالتالي لا تتحقق الشروط التي تبرر لك استعمال إحصاء الاختبار z. بالرجوع الى القسم السابق تجد أن

توزيع الإحصاء:

T=x
ج1 إ/ة

هو توزيع t بدرجات حرية (n -1)، وحيث x هو الوسط الحسابي للعينة العشوائية من توزيع طبيعي، s هو الانحراف المعياري لتلك العينة. إذن والحالة هذه فإن احصاء الاختبار المناسب لاختبار الفرضية 0لإ = ا-ا:H_o

مقابل الفرضية البديلة 0لإ>لإ:٩ أو 0لإ >H_1:jx أو 0دإ ؛؟ لإ:أتلآ هوع[7]

ويتم إجراء الاختبار كما في الجزء (1.4) مع تعديل واحد هو استعمال توزيع t بدرجات حرية (1 -3) بدلا من استعمال التوزيع الطبيعي المعياري.

ويتلخص الإختبار فيما يلي:

لإختبار الفرضية الصفرية 0لإ=لإ:0لل مقابل الفرضية البديلة:

0لإ>لإ:لالآ(ل)

أو 0لإ>لإ:ا٠)

أو 0لإ ؛7 غج: Hj (iii)

علىمستوى دلالة a

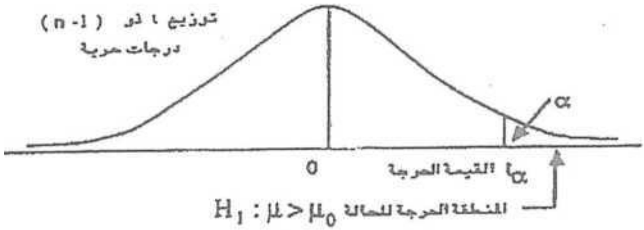
اختر الاختبار المناسب ذا الطرف الواحد أو الطرفين حسب الحالات (i), (ii), (iii)

استعمل إحصنا ء الاختبار1 واستعمل جداول توزيع t بدرجات الحرية (n - 1)، وتكون القيم الحرجة t_a للحالة (1)، $-t_B$ للحالة (ii)، م ط- للحالة (iii)، حيث٤ هي لقيمة على المحور الأفقي لتوزيع t ذي درجات الحرية (n -1) ويقع إلى يمينها *a* من المساحة.

أحس ئ=؛
قلي 1>؟.

قارن t المحسوبة بالقيم الحرجة، t٥ واتخذ القرار:

1. . الحالة 0لإ>لإ :Hj أرفض HQ إذا كان t >t٥
2. الحالة 0لمإ> H أرفض H_o إذا كان t<-t٠
3. الحالة 0لإ ؛7 ح :$H_,$ أرفض Ho إذا كان t>t٥٥ أو ييع-2٤ كما في الأشكال (15) أ، ب، ج.

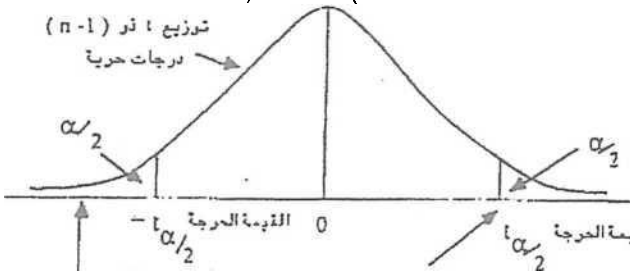


الشكل (15-أ)

L ابغييُ’سجةىر٢ 0 :

المنطتةلحرجة لحالة Hj:p<g0

(الشكل15 ب-)



الشكل (15-ج) Hjالمنطت الحرجة لحال 0لر نج لمز: :

استنتج أحد الباحثين أن معدل عدد الساعات التي يقضيها طلبة إحدى الجامعات في الدراسة أثناء أسبوع الإمتحانات 50 ساعة. اختبر هذه الفرضية مقابل فرضية أن معدل عدد الساعات يختلف عن 50 ساعة، إذا كان الوسط الحسابي لعدد الساعات التي قضاها 10 طلاب أثناء ذلك الأسبوع هو51.7 ساعة بانحراف معياري 6.3ماعة. استعمل مستوى دلالة 5% وافترض أن توزيع عدد الساعات الدراسية تقريباً طبيعي.

1. الفرضية الصفرية H_o: p. — 50

2. الفرضية لبديلة 50 د دي: H؛

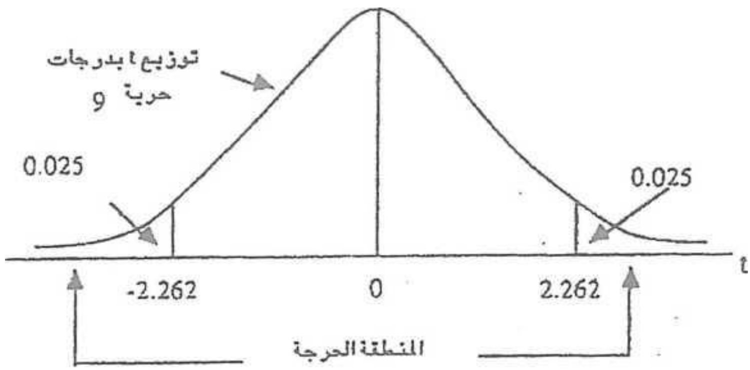
وقد تم استعمال الفرضية الثنائية لأننا لانعرف اتجاه الاختلاف أوالخطأ إذا كان استنتاج الباحث غير صحيح، ولذلك نأخذ الاتجاهين "أكبر من" أصغر من "أي أن الفرضية الصفرية 50 ب H, : P

3. مستوى الدلالة a = 0.05

4. بما أن الفرضية البديلة ذات اتجاهين فإن الاختبار المناسب ذو طرفين، ويما أن توزيع المجتمعطبيعي، تباينه غير معلوم، حجم العينة صغير، نستعمل توزيع ٤بدرجات حرية - (n - 1) ولذلك فالقيم الحرجة هي 2.262 = **0,025**؛ و 2.262-- **0.025**،-

حيث درجات الحرية 9 =1 -10

أنظر الشكل (16)



الشكل (16)

5. احسب٤من لمعادلة ■وتح
؟ه

6.3٠/ 6.3/3.16

6. قارن قيمة t المحسوبة بالقيم الحرجة:

من الواضح أن 2.262 > 0.§5 > 2.262-

إذن فإن t لا تقع في المنطقة الحرجة، وبالتالي فلا ترفض H_o .

كان معدل تحصيل طلبة إحدى المداري الخاصة في امتحان اللغة الانجليزية الذي يتقدمون له عند طلب الالتحاق بالجامعات الأمريكية 400.

اختبر فرضية أن هذا المعدل قد تحسن إذا أعطت نتائج 14 طالباً وسطاً حسابياً ؛

X = 408

بانحراف معياري S - 23 .

اعتبر أن نتائج طلبة المدرسة_تخضع لتوزيع طبيعي وخذ مستوى دلالة

% 1

1. $H_O:M = 4OO$
2. H,:p.>4OO
3. مستوى لدلالة a = 0.01
4. بما أن الفرضية البديلة لها اتجاه واحد فإن القيمة الحرجة **0 01**؛ تحت درجات حرية 13.

وقد تم استعمال توزيع t لأن المجتمع طبيعي، تباينه غير معلوم، وحجم العينة 14

أي أقل من 30.

إذن القيمة الحرجة هي: t٠٠oi =2.65

أنظر الشكل (17)

توذيعابرجات حرية 13

0.01 دی

2.65 الاب الدرجة 0

المبيا س

الشكل (17)

5. أحسب قيمة t من

$$t = \frac{X - gp}{s/\sqrt{n}}$$

408-40_0   8_x_3.742

ذ٦ ةآل ا لا٦

-1.3

6.قارن بين قيمة٤والقيمة لحرجة t_{0}٠o١

من الواضح أن 2.65 > 1.3 أي أن ٤ لا تقع في المنطقة الحرجة، ويالتالي لا ترفض

H_{0}

وبالتالي لا تستطيع استنتاج أن معدل تحصيل الطالبة قد تحسن.

فهلم

. ؛ذلا؛ '!اا'ع درسة سبنة اس٠ئجا بأن كنى الزمه الذي تعت١جه شي

إ خياطة لإنجاز عدد من القطع الجاهزة هو 7 ساعات.

اخقبر هذا الاستتاج مقابل فرضية أن المعدل أقل من 7 ساعات لعينة عشوائية :

: حجمها 23 عاملة أعطت X = 6.5 ساعة وانحراف معياري S =1.4 ساعة.

إفرض أن الزمن موزع حب_التونبع الطبيعي واستتمل_مستوى 5%.

اختبر الفرضية 10=لإ :H_{0}

مقابل الفرضية البديلة 10 >لر

على منتووى دلالة 1 %

إذا أعطت عينة عشوائية حجمها 8 من مجتمع طبيعي النتائج التالية:

. ع ,, 5د7. X^s٠2 11.5 — ؤ

'! ٦اكل -م٦ص ٦ تكوت 2ه غير مطوة وحجم اتيتي-تر تحة_اختبار عإ لتوزيع

طبيعي.

اختبار الفرضيات المتعلقة بالنسبة

TestingHyp heses Coiicerning a 'Proportion

إن اختبار الفرضيات المتعلقة بالنسبة، أي نسبة المجتمع ذي خاصية معينة، يشبه اختبار الفرضيات المتعلقة بالوسط الحسابي، ويتلخص الأمر في تفحص P، النسبة في العينة التي تمتلك الخاصية المطلوبة، فإذا كانت هذه النسبة بعيدة جداً عن النسبة في المجتمع فإنكترفض الفرضية الصفرية-•: ٥ وإلافإنك اترفض مآي] إما إحصاء الاختبار فهو Z حيث:

$$Z = \frac{\bar{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$$

حيث توزيع Z قريب من التوزيع الطبيعي المعياري، وعليه فالمنطقة الحرجة على مستوى دلالة ه لاختبار p = p٥ : H٥ تكون كما يلي: الحالة (؛): H,:P>P٥ أرفض H₀ إذا كان Z>Zaالحالة (ii): H, :P<P٥ أرفض H٠ إذا كان Z <٠Za الحالة (iii): •م •: ,1ج رفض *1 إذا كان ZZأو Z<-Z.٥

ويمكن تحويل القيم الحرجة والمناطق الحرجة بدلالة2 من المعادلة - -P?~م = Z

٧ PQG-PQ)

والتي تعطيك:

Hi:P>P٠ أرفض H₀ إذا كان

P>VZa٠!

؛أ. اذاً •٠٦٠" ■■■ H.:P<P٥ أرفض Ho إذا كان

p<p٥-zf

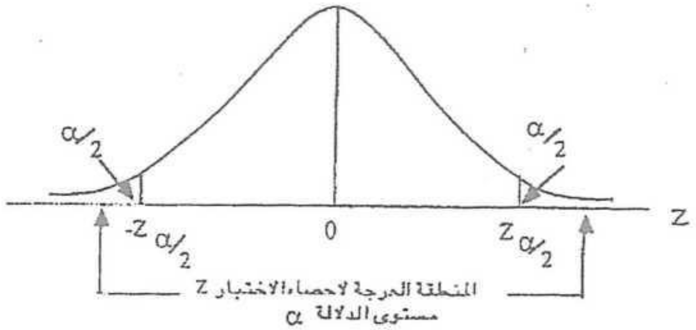
٦ ١٠?، T H,:P*P٥ أرفض Ho إذا كان

P<P٠-Zص٠كقل٢ ق اد P>P٠+Z،2-١RF

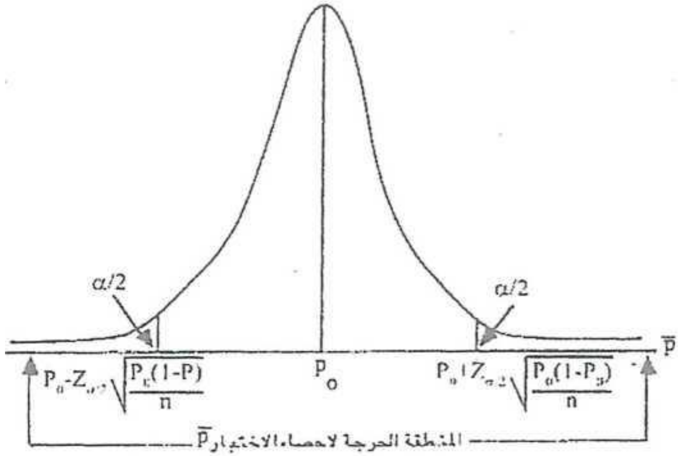
أنظر الشكل (18)، الشكل (19)

تفضل على ذلك التي Z لاحظ أن القيم الحرجة وبالتالي المناطق الحرجة بدلالة

بدلالة $\bar{P}$



الشكل (18)



الشكل (19)

مثال (28)

إذا كانت نسبة مستعملي حزام الأمان في السيارات (قبل تشريع إلزام الاستعمال) هي 0.8 درست عينة عشوائية حجمها 200 سائق بعد صدور تشريع الإلزام فوجد أن 170 منهم يستعملون الحزام، اختبر على مستوى دلالة 5% ما إذا كان التشريع قد زاد نسبة المستعملين له.

من الواضح أن 0.85 إل = p
200
ضع خطوات الحل كما يلي:

1. $H_0: P = 0.80$
2. $H_0: P > 0.80$
3. $a = 0.05$ مستوى الدلالة
4. بما أن الفرضية البديلة ذات اتجاه واحد هو "أكبر من"، فإن الاختبار المناسب ذو طرف واحد، وا لقيمة الحرجة

$$Z_{0.05} = 1.645$$

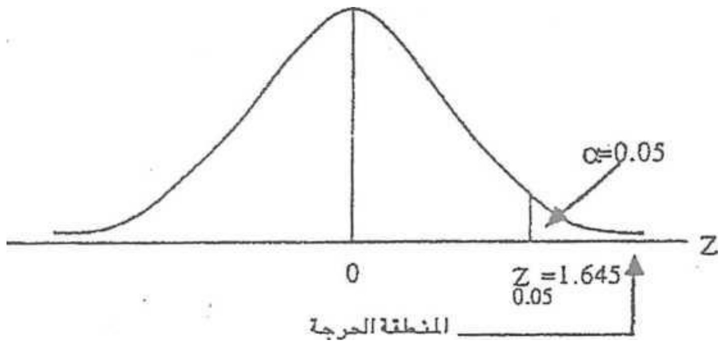
5. أحسب Z من المعادلة

$$z = \frac{p - p_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$$

$$= \frac{0.85-0.80}{\sqrt{\frac{0.80 \times 0.20}{200}}} = \frac{0.05}{0.028} = 1.8$$

6. قارن قيمة Z التي حصلت عليها مع القيمة الحرجة:
من الواضح أن 1.645 < 1.8

إذن أرفض HQ لصالح Hj ، وبالتالي فإن صدور التشريع بإلزامية استعمال الحزام قد زاد من نسبة المستعلمين له. شكل (20)



الشكل (20)

لاحظ أن بإمكانك حل المثال بالطريقة الثانية، أي بدلالة 5، وهي اتباع نفس الخطوات السابقة حتى الخطوة (4)، أما الخطوة (5) فتصبح:

5. احسب القيمة الحرجة على أساس أن2هو إحصناء الاختبار فتجد:

عب٩

= 0.8 + 1.645 ٧٦٢

-0.8+0.046-0.846

6. قارن قيمة P بالقيمة الحرجة

من الواضح أن P = 0.85 أكبر من 0.846

اذن ارفغن H٥ لصالح HP >0.80

إذا كانت نسبة العائلات التي تملك البيوت التي تسكن فيها في مدينة معينة هي 62% اجريت احصائية عن 2000 موظف فوجد أن 1280 شخص من بينهم يملكون البيوت : التي يسكنونها.

اختبر الفرضية H_0:P = O.62 ا

مقابل الفرضية البديلة H,: P > 0.62

استعمل مستوى 1% :

نسبة الطلبة الذين يسجلون لأكثر من 15 ساعة معتمدة في الفصل هي 0.8، تم رفع : قيمة الأقساط فاعتقد المسؤولون أن هذه النسبة ستقل، ولاختبار هذه الفرضية أخذت عينة : عشوائية من 1600 طالب فوجد أن عدد المسجلين لأكثر من 15 ساعة هو 1120 طالبا.

اذكر الفرضية الصفرية والفرضية البديلة واختبر نلك على مستوى دلالة a = 0.01

كيف تستطيع اختبار الفرضيات المتعلقة بنسبة مجتمع ذي خاصية معينة

الوحدة الخامسة

تحليل التباين و

اختبار الاستقلالية

فتحليل التباين

قل٠عهـ of Variance (DVA)

مقدمة:

درسناالمقارنة بين متوسطي مجتمعين لكلمنهما توزيع طبيعي ، والان سنقوم بتعميم ذلك ونتناول المقارنة بين اكثر من متوسطين، فلواراد مثلا احدالباحثين في المجال الطبي دراسة تأثير ثلاثة انواع مختلفة من الادوية A,B,C وقام هذا الباحث باختيار30 مريضا بطريقة عشوائية وقسمهم الى ثلاثة مجموعات ، واعطى لدواء هالى المجموعة الاولى والدواء •الى المجموعة الثانية والدواء C الى المجموعة الثالثة، ثم مجل الباحث طول الوقت اللازم لشفاء كلمريض ، فان هذه التجربة تتطلب المقارنة ين ثلاثة متوسطات هي تعميم للمقارنة بين متوسطي مجتمعين

و الاسلوب الذي يتم فيه اختيار مفردات العينة عشوائيا وتوزيع هذه المفردات عشوائيا على المعالجات المختلفة يسمى بالتصميم التام العشوائي

completely randomized design

واذاطلب منا مثلا اننقارن بين متوسط الانتاج لاربعة انواع مختلفة من القمح ولتكن A,B,C,D ونفرض ان لديناقطعة ارض لزراعة هذه الانواع من القمح فان الشئ الطبيعي ان نقسم قطعة الارض الى قطع صغيرة ولتكن مثلا 20 قطعة متجاورة كماهومبين في الشكل التالي، ثم نوزع المعالجات (انواع القمح) على القطع عشوائيا بحيث تكرركل معالجة عددمساو من المرات ٠

c	D	B	D	A
A	B	A	A	c
A	c	D	c	B
B	D	D	c	B

فان .هذا الاسلوب. الذي يتم به توزيع المعالجات (انواع القمح) على المفردات (قطع الارض) عشوائيا يسمى بالتصميم التام العشوائي.

واهم نقد يوجه الى هذا التصميم هوان التوزيع العشوائي لايضمن ان تكون القطع (الوحدات) التي تقع تحت تأثير احدى المعالجات مشابهة لتلك التي تقع تحت تأثير معالجة اخرى وذلك لانه من الممكن ان تختلف هذه القطع من حيث الظروف المتعلقة بالزراعة (مثل درجة الخصوبة او الرطوبة ، الخ) ممايؤدي الى اختلاف الانتاجية من قطعة الى اخرى . وللتحكم في مثل هذا المصدر المسبب للاختلاف فاننا نقسم الارض الخمسة قطاعات متجانسة ثم نقسم كل قطاع الى اربع قطع متساوية المساحة ثم نخصص هذه القطع عشوائيا على انواع القمح المختلفة كمايتضح من الشكل التالي :

1	2	3	4	5.
B	C	B	A	A
A	B	C	B	D
c	D	A	D	B
D	A	D	C	(3

وتتم ' هذه المقارنة هنا بين متوسطات الانتاجية لكل نوع من انواع القمح وكذلك الانتاجية لكل قطاع وهذا الاسلوب يسمى بتصميم القطاعات الكاملة العشوائية randomized block design وطريقة تحليلالبيانات التي نحصل عليها من التصميم التام العشوائي وتصميم القطاعات الكاملة العشوائية تسمى تحليل التباين (Analysis of Variance (ANOVA

هومقياس لاختلاف مفردات ايةظاهرة. وهذا الاختلاف ناتج عن اسباب كثيرة بعضها يمكن معرفته والبعض الاخر لايمكنمعرفته . وعند تحليل النتائج نقوم بحساب مقدار الاختلاف بين النتائج التي حصلناعليها وهومايسمى بالتباين الكلي ثم نقسمه الى اجزاء ، ونرجع كل جزءالى مسببه وهذامايسمى بتحليل التباين . فمثلا في التصميم التام العشوائي فان الاختلاف في النتائج له سببان:

أ ـــ الاختلاف بين المعالجات.

ب ـــ الخطا التجريبي (النتج بسبب عوامل غير معروفة).

ويسمى تحليل التباين في هذه الحالة بتحليل التباين في اتجاه واحد وكذلك في تصميم القطاعات الكاملة العشوائي فان الاختلاف في النتائج لهثلاثة اسباب: أ ـــ الاختلاف بين المعالجات

ب - الاختلاف بين القطاعات

ج ـــ الخطأ التجريبي (الناتج بسبب عوامل ير معروفة).

ويسمى تحليل التباين في هذه الحالة بتحليل التباين في اتجاهين تحليل التباين في تجاه واحد One - way Analysis of Variance

از .

ج يي ل:..,;

(بم~/بم)ر/ ___ 2

ا-بم

ره* اخر

زت±=7

جدول تحليل التباين في اتجاه واحدANOVA table

مصدر الاختلاف	مجفوع المربعات	درجه الحرية	متوسطات المربعات	النسبة
المعالجات	ح-هلآ=٠ر٢كك	k-l		
الخطأ		k-n	٠ع	!ج MSE
	ك-اهةة I	n-1		

وعندمايكون فرض العدم Ho صحيحا فان القيمة المتوقعة للإحصاءHo تساوي الواحد الصحيح وإلاافان لقيمة المتوقعة ستزيذ عن الواحدلصحيح ولهذا نجري التجربة ونحسب إحصاء الاختبار ونرفض فرض العدم وهو: ؟ ٥٩: ؟،:ق

عندماتكون المحسوبةكبيرةجدا بحيث يمكن استبعاد ان يكون ذلك قد حدث عن طريق الصدفة أي عندماتكون :

$F>F_a(i-$

ولذلك فان الاختبار في هذه الحالة يكون دائما اختبارطرف أيمن مثال:

أراداحدالباحثين في المجال الطبي دراسة تأثير ثلاثة أنواع مختلفة من الأدوية ASC وقام هذاالباحث باختيار30 مريضا بطريقة عشوائية وقسمهم الثلاثة مجموعات ، وأعطى الدواء ٨الى المجموعة الأولى والدواء • إلى المجموعة الثانية والدواء C الى المجموعة الثالثة، ثم سجل الباحث طول الوقت

اللازم لشفاء كلمريض باليوم، فان هذه التجربة تتطلب المقارنة ين ثلاثة متوسطات هي

تعميم للمقارنة بين متوسطي مجتمعين عندمستوى دلالة 0.05؟ الحل:

$ه_3$: دحه = $ح_1$ = ه#

؟،ب =

نوغ الدواء الوقت	A	B	C	[illegible]
	3	6	1	
	4	5	2	
	3	6	3	
	3	4	2	
	2	5	2	
	3	8	3	
	4	4	1	
	3	5	2	
	3	7	2	
	2	10	2	
Ti	30	60	20	T=11O
ni	10	10	10	n=3O
-بم	3	6	2	

110ت2 + 2ب........ + 3 + 4ب3=ه٤ ئ:أز؟=7

4٢10 + 10 + 10 = 30

403.33 : نوهأع : عي

30

ك٣ (3ه)2, (؟؟),(20)2 ■■190

و 10 10 10 لج ؟

530 = 2(2) ب........-؛- 2(4) ب 2(3) :

490-403.33 = 86.66 : ئلا٢ية :٠مت٢كد ئ دخ

530-403.33 = 126.66ت ٠ئ- ■زرك11 = هل،

د 126.66- 86.66 = 40.00

!:!:!:43.33

3-1 ا-ك/

!:!:!:1.481

30-3 رحم

/:!:!:29.250

f 1.481

70.05,2,27) = 3.35

جدول تحليل التباين في اتجاه واحدANOVA table

مصدر الاحتلاف	مجموع المربعات	درجة الحرية	متوسطانت المربعات	النسبة
المعالجات	86.66	2	43.33	29.250
دحقاً	40	*27*	1.481	
	126.66	29		

حيث F المحسوبة اكبرمن F الجدولية ترفض Ho

هناك فروق ذودلالة إحصائية بين المتوسطات عندمستوى دلالة 0.05 أي يوجداختلاف في طول الوقت اللازم لشفاء كل مريض باليوم بين أنواع الأدوية الثلاثة

. انتبارمربعكاي X^2 - Test

يستخدم هذا الاختبار في الحالات الآتية:

> جودة التوفيق

> الاستقلال

> التجانس

وسنكتفي بدراسة هذا الاختبار في حالة الاستقلال

ففي حالات كثيرة نحتاج إلى التعرف عما إذا كان هناك علاقة بين صفتين من صفات مجتمع ما أم لا.

فمثلاً: قد نحتاج إلى معرفة هل هناك علاقة بين مستوى الدخل ومستوى التعليم ؟ أو هل توجد علاقة بين لون العينين ولون الشعر في مجتمع ما ؟ وهكذا وللإجابة على ذلك نتبع الخطوات الآتية:

أ- نضع فرض العدم Ho : لا توجد علاقة بين الصفتين

الفرض البديل HI : توجد علاقة بين الصفتين.

ب- نختار عينة من مجتمع الدراسة ، ثم نصنف مشاهدات هذه العينة حسب مستويات كل صفة من الصفتين ونضعها في جدول يسمى "جدول التوافق" ، وهو يحتوي على التكرارات المشاهدة (Oj) لكل خلية

ج- نحسب التكرار المتوقع (E؛) المناظر لكل تكرار مشاهد (لكل خلية) من العلاقة الآتية:

= (التكرار المتوقع) Ei

مجموع الصف الذي به الخلية X مجموع العمود الذي به الخليةة مجموع التكرارات (حجم العينة)

د- نوجد ?X المحسوبة (الفعلية) من العلاقة الآتية : لتئك) ٢٦ : 2/

ه نوجد 2^ النظرية (الجدولية) بدرجات حرية

2

حيث: Xa] اح)(اح)]٢د

r: عدد الصفوف

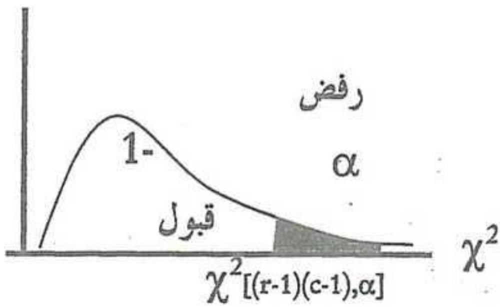
٠: عدد الأعمدة

a: مستوى المعنوية

و- إذا وقعت X^2 المحسبوبة في منطقة القبول ٠٠, نقبل Ho ونرفض H؛

أي لا توجد علاقة بين الصفتين والعكس إذا وقعت المحسوبة خارج منطقة القبول (في منطقة الرفض)

نرفض Ho ونقبل HI أي توجد علاقة ين الصفتين



مثال: لدراسة العلاقة بين التعليم والتدخين. سحبت عينة عشوائية من (400) شخص فأعطت النتائج الآتية:

التعليم	التدخين		Z
	يدخنون	لا يدخنون	
غير متعلم	**170**	**50**	**220**
تعليم متوسط	**50**	**70**	**120**
تعليم عالي	**20**	**40**	60
Z	**240**	**160**	**400**

هل توجد علاقة بين التدخين والتعليم ؟ استخدم مستوى معنوية a = ٠O5الحل:

أ- Ho: لا توجد علاقة بين التدخين والتعليم.

HI : توجد علاقة بين التدخين والتعليم.

ب- نوجد التكرارات المتوقعة Ei كما يلي :

٠٠٠٠٠٠ 120^40:7 220^40:132

E 400 Z 400

ونضعها في جدول كما يلي :

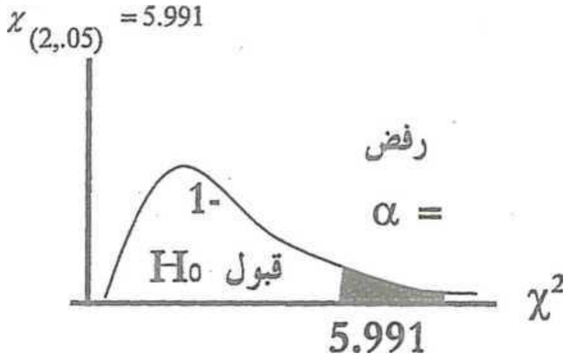
التعليم	التدخين		Z
	يدخنون	لا يدخنون	
غير	132	88	220
تعليم متوسط	*72*	48	120
تعليم عالي	36	24	60
ة	240	160	400

نوجد 2جلاً المحسوبة كما يلي :

في مثل هذه الحالات لا داعي لاستكمال الحل . وإنما نوجد 2أ النظرية من الجدول

هي ٠،

نمهد 132 88



٠٠- المحسوبة تقع خارج منطقة القبول (داخل منطقة الرفض) .٠. نرفض Ho ونقبل HI أي توجد علاقة بين التدخين والتعليم

مثال: لدراسة العلاقة بين لون الشعر ولون العينين في أحد المناطق، أخذت عينة من (200) شخص وتم تصنيفهم في جدول التوافق الآتي :

لون الشعر	لون العينين		المجموع
	بي	أزرق	
اسود	60	20	80
بني	40	30	70
أشقر	30	20	50
المجموع	130	70	200

هل توجد علاقة بين لون العينين ولون الشعر ؟ وذلك عند مستوى معنوية . a=.OI

الحل

أ- م[]: لا توجد علاقة بين لون الشعر ولون العينين.

HI : توجد علاقة بين لون الشعر ولون العينين.

ب- نوجد التكرارات المتوقعة؛E ، كما يلي :

52=هيكبلآج

لون الشعر	لون العينين		المجموع
	بني	أزرق	
اسود	52	28	80
بني	45.5	24.5	70
أشقر	32.5	17.5	50
المجموع	130	70	200

ج نوجد2% المحسوبة كمايلي:

(20-17.5)1 : ?30-32.5) I (30-24,5)1 I (40-45.5)2 I (20-28)1 ı (60-52? ι او-رم٢ 1

17.5 * 32.5 ب 24.5 ب 45.5 ؛ 28 ا 52 رتج z د

= 1.23؛ 2.29 ٥ .66 +1.23 + .19 + .36 = 5.96

د-نوجدةالنظرية921 (2,01)* [بل(ا-مم).(ا-ن]د

رفض

قبول 0شغ

X^2

921

2..المحسوبة تقع في منطقة القبول.. نقبل Ho ونرفض Hl أي لا توجد علاقة بين لون الشعر ولون العينين.

الوحدة السادسة
الاحصاءات الحيوية

الاحصاءات الحيوية

VITAL S**TATISTICS**

النسبة (Ratio) و المعدل (Rate) ة

هنالك العديد من الدوائر الحكومية و الخاصة التي تحفظ ببعض السجلات عن الأمور الصحية و الاجتماعية للمجتمع الذي تتواجد فيه. ففي الاردن مثلاً هنالك دائرة الاحوال المدنية، و دائرة الاحصاءات العامة و وزاة الصحة و غيرها تحتفظ بسجلات عن حجم السكان و أعداد حالات الزواج و الطلاق و الوفيات ضمن فئات العمر المختلفة و اعداد المواليد و انتشاء الامراض المختلفة ضمن فئات العمر المختلقة و غيرها من البيانات الاحصائية الحيوية.

ليس المهم معرفة اعاد حالات الوفيات ضمن فئة معينة بل الاكثر اهمية هو معرفة نسبة هذه الوفيات لهذا سنبدأ بالتعرف على تعرف كل من النسبة و المعدل.

وليكن ى= تكرار وقوع حادث ما خلال فترة زمنية محددة ، وليكن وب7،= عدد الاشخاص الذين كان من الممكن تعرضهم لهذا الحادث خلال نفس الفترة الزمنية وليكن*/حد/لاعداد

ە; ,٥٥; ,٥٥٥; ,٥٥٥٥; ,٥٥٥٥٥; يسمى |شغك|دت٠+٠)

معدل وقوع هذا الحادث. و يسمى العدد ء اساسا. وفي العادة نأخذ *٥٥٥*/طكوتخعر سنلتزم به في *هذا الفصل اما النسبة فهي مقدار على النحو ج/ە/ءكا حيث انه ليس ضروريا از تكون cجزءا من ى .*

مثال (1): اذا كان عدد سكان مدينة 30000 نسمة في يوم منتصف السنة، و بلغ عدد الوفيات في تلك المدينة 600 شخصاً في السنة فإن :

a = ٥٥٥

ههههو = ةبه

واذا اخذنا k=1000 في معدل الوفيات في هذه المدينة يساوي : 2٠ههههها*ج

أي ان الوفيات تحصل بمعدل 20 وفاة لكل 1000 شخص :

مثال (2) : تحتوي مزرعة على 2000 رأس من الغنم، 500 رأس من البقر. لهذا فان نسبة الابقار الى الاغنام في المزرعة تساوي

أي انه توجد 250 رأساً من البقر لكل 1000 رأس من الغنم.

إحصاليات الوفيات :

سنعالج في هذا البند بعض النسب و بعض المعدلات ذات العلاقة بالوفيات تعبر معدلات الوفيات التكرارات النسبية لحدوث الوفيات ضمن مجتمع معين خلال فترة زمنية محددة . يشير مقام معدل الوفاة الى عدد افراد المجتمع الواقع تحت خطر الوفاة . بينما يشير البسط الى عدد تلك الوفيات التي حدثت في المجتمع المشار اليه في المقام . و هنالك عدة انواع من معدلات الوفيات نذكر منها :

1. . معدل الوفاة الخام السنوي Annual crude death rate و يساوي :

الوفات خصهة٠ن. 1/1

حدجمبعافرادالمجعفي 7/1

حيث ارتنم عادة. ومن الجديربالذكر انه من الخطأمقارنة معدلات لوفاة الخام لمجتمعات مختلفة مالم تكن ظروف المجتمعات الصحية و الاجتماعية و غيرها من الظروف مثل فئات الاعمار متماثلة و المثال (I) اعلاه يوضح هذا التعريف.

2. معدل الوفاة المحدد السنوي Annual specific death rate

في كثير من الحالات يكون من الافضل و الاكثر فائدة ايجاد معدل الوفاة ضمن مجموعات جزئية محددة من المجتمع الاحصائي مثل مجموعة الاناث او مجموعة الاطفال او مجموعة الشباب و ما ذلك و تسمى مثل هذه المعدلات بالمعدلات المحددة

مثال (3) : اذا كان عدد الاطفال دون سن العاشرة في مدينة ما يساوي 5000 طفلاً و بلغ عدد الوفيات من الاطفال دون هذا السن 250 فان معدل الوفيات دون سن العاشرة يساوي:

ع* 1000 = 50

5030

أي ان المعدل هو 50وفاة لكل 1000 طفل في هذا السن.

3. معدل الوفاة المعياري Standardized death rate

لقد اشرنا سابقاً انه من الخطأ مقارنة معدلات الوفيات في اكثر من مجتمع بناء على معدلات الوفيات الخام في حالة وجود متغير يؤثر على الوفيات مثل العمر او الجنس ولكن المقارنة تصبح ممكنة عن طريق مقارنة ما يسمى بالمعدلات المعيارية.

فيما يلي طريقة لحساب معدلات الوفاة المعيارية :

تقوم فلسفة هذه الطريقة على ايجاد عدد الوفيات المتوقع في كل مجتمع فيما اذا كان عدد افراد هذا المجتمع مساوياً مجموع عدد افراد المجتمعين المراد مقارنتها معا. و المثال التالي يوضح خطوات ايجاد هذا المعدل المعياري

مثال (4) : يمثل الجدول 1 اعداد السكان و اعداد الوفيات في مدينتين أ ، ب مصنفة حسب فئات الاعمار.

جدول 1

العمر	المدينة ا		المدينة ب	
	عدد السكان	عدد الوفيات	عدد السكان	عدد الوفيات
20'29	**45000**	**90**	**30000**	**45**
30-39	**40000**	**120**	**30000**	**105**
40-49	**35000**	**140**	**40000**	**180 .**
50-59	**30000**	**150**	**50000٠**	**225**
المجموع	**150000**	**500**	**15000**	**555**

لاحظ ان معدل الوفيات الخام في | ح**500٧1000/150000- 3.33** لكل **1000**

وان معدل الوفيات الخام فيب = 0000 555 0)10= 3.70 **لكل** 1000 **لحساب معدل الوفاة المعياري لكل من أ، ب نقوم**

بالخطوات التالية على فرض أن

k=1000

ii **نجد العدد المعياري للسكان في كل من أ ، ب بحيث يكون مجموع عدد السكان في كل منهما مصنفاً حسب الاعمار . وبهذا نحصل**

على العمود الثاني في الجدول 2

2. **نجد معدل الوفاة المحدد حسب فئات العمر في كل من أ، ب وبذبك نحصل على العفودين الثالث و الخامس في الجدول** 2 .

3. **نجد عدد الوفيات في كل من أ، ب وذلك حسب القانون التالي عدد الوفيات المتوقع = (عدد السكان المعياري** X **معدل الوفاة**

المحدد) ب 1000

وبهذا نحصل على العمودين الرابع و السادس في الجدول 2 .

4. **و اخيراً نجد معدل الوفاة المعياري لكل من أ، ب باستعمال القانون التالي**

و نلخص نتائج هذه الخطوات في الجدول 2.

الجدول 2

العمر	العددا لمعيا ري للسكان	المدينة ا		المدينة ب	
		المعدل المعدد	عان الوفيات المتوقع	المعدل المعدد	علد الوفيات المتوقع
20-29	75000	2	150	1.5	112.5
30-39	70000	3٠	210	3.5	245
4049	75000	4	300	4.5	337.5
50-59	80000	5	400	ث	360
المجموع	300000		1060		1055

لهذا فان معدل الوفاة المعياري للمدينة

أ = لأ* 1000 = 3.53لمكق 2000 300000

و معدل الوفاة المعياري للمدينة

ب =!x 1000 = 3.52لكق1000
300000

ومن هنا نلاحظ ان المعدلين المعياريين متساويان تقريباً في حين ان المعدلات الخام كانت مختلفة لهذا لا نستطيع ان نستنتج وجود فروق في معدلات الوفيات بين المدينتين.

5. معدل وفيات الاطفال الرضع

و يساوي

حاخثيقتاحيتثبثفتد 100GX
للحلم تتتنةقى لحياءانعةلوهث لاطقث عد

مثال 5 : اذا كان عدد الاطفال الذين توفاهم الله قبل اتمام عامهم الاول يساوي 50 طفلا بينما يبلغ عدد الاطفال المولودين احياء في نفس العام هو 5000 طفل فان معدل وفيات الاطفال الرضع يساوي

IGOO x!

5000

أي انهيساوي10لكل 1000

6. معدل وفيات الاطفال حديثي الولادة .

ويساوي

د<ض¥،ضعبم!،5حه<صمج28ط X GO□ 1 □
عد 4 لأطثته للمرلىحيت احياء في تقعن للحلم

7. معدل وفيات الامومة.

وهو المعدل الذي يعبر عن وفاة الامهات لاسباب تتعلق باطفالهن كتلك الوفيات التي تحدث عند الولادة او نتيجة مضاعفات قد تنسأ في وقت لاحق او وقت سابق لها بسبب الحمل او غيره ويعرف هذا المعدل على النحو التالي : معدل وفيات الأمومة يساوي.

خحقكلهتاللاثقتظفةن جمم

8. معدل وفيات الاسقاط.

ويساوي

ثجج*ه0٠لم

ومن الجدير بالذكر ان هناك فروقاً بين بعض الاقطار في تعريف حالة اسقاط الجنين تعتمد على فترة الحمل ولكن الاكثر شيوعاً هو اية حالة اسقاط بغض النظر عن فترة الحمل.

احصائيات الامراض :

من المواضيع التي تهم العاملين في المجال الصحي و تحليل الوضع الصحي في المجتمع هو موضوع احصائيات الامراض وفيما يلي عدة مقاييس في هذا المجال:

١. معدل الاصابات:

حتت إلاصديلث ١١١حر-نمق برض عتين خلتن م ويساوي 1000 X

خدلظيمطككل

2. معدل الانتشار :

يقيس هذا معدل مقدار انتشار مرض معين في بلد معين وهو يساوي :

عد_إلاصيثح، آثعييودت(_قيعة_اوجيدذ؛قي_لمحتة_ميتة 1G00X

عددللنسكتفيككلثحظة

3. نسبة حالات الهلاك : وهذه النسبة مفيدة في التعرف على مدى نجاح برنامج مكافحة مرض معين و تعرف هذه النسبة على النحو التالي :

حتد_حاثاك_3تتكيتيي_عئ_صن_هين 1000X

حددحلات٤لاصيكتبيذاللعنص

نسبة حالات الهلاك تساوي

والفترة التي تقاس فيها هذه الحالات اختيارية يمكن ان تكون بأي طول. احصاءات الخصوبة :

تعتبر الخصوبة و نسبة المواليد في بلد ما من اهم المواضيع بالنسبة للعاملين في الحقل الصحي اذ ان معرفة هذه النسب تساعدهم في التخطيط لانشاء مراكز الرعاية الصحية و تنظيم الاسرة و توفير المرافق الضرورية لكل من الام و الطفل ومن اهم مقاييس الخصوبة :

1. معدل الولادة الخام :

ويسوي ج؛ل|

2. معدل الخصوبةا لعام

وسيه ح:ا

وعادة ما يعرف سن الحمل بين 15 الى 44 عاماً او 15 الى 49 عاما

3. معدلالخصوبة المحدد بالعمر

من المعلوم ان توزيع عدد النساء في سن الحمل ليس توزيعاً منتظماً ، بمعنى ان النساء في سن فوق الخامسة و الثلاثين اقل حملاً من النساء في سن الخامسة والعشرين مثلا. لهذا نحتاج الى مقاييس لمعرفة معدل الخصوبة معدل الخصوبة المحددبالعمر يساوي:

حتكلتاجاحقححخيجلة 1000

حدلككسلهعنة1 الحعتعتدعتتصلثحتي

جدول الحياة :

يحتاج العاملون في الحقل الصحي لمعرفة احتمال البقاء على قيد الحياة خلال فترة محددة من الزمن ، و غيره من المقاييس ذات العلاقة بمعدلات الوفاة .هناك جداول يمكن انشاؤها لهذه الأغرض وتسمى بجداول الحياة ومن مكونات هذه الجداول مايلي :

لتكن [تمرب x ع] فترةزمنيةطولهار و بدايتها ي .

لنستعمل الرمز ي للدلالة على عدد الاشخاص الاحياء عند بداية الفترةر، و *الرمز تدك للدلالة على عدد الوفيات خلال الفترة ا »ه د د lx ء و الرمز سرم(لتعببير عن احتماله وفاة شخص ضمن الفترة '١٠ اي ان*

ـــج ت $n^a{\sim}x$ والرمزالدلالة طى عد قرات الحية الحالثة خلل الفقرة

من قبل الاشخاص الذين بقوا على قيد الحياة حتى بداية الفترة أي ان : آ+ط:ه

ويستعمل الرمز ئت للدلالة على عدد سنوات الحياة الحادثة خلال الفترة x و جميع الفترات اللاحقة بها أي ان: y مد ة = T_x

و يستعمل الرمز ربي للدلالة على توقع متوسط الحياة للاشخاص الاحياء في كل فنة عمر محددة أي اني= ه «ج.سنفرض فى هذا البند ان 1ء »و لتوضيح هذه الرموز ناخذ الامثلة التالية:

مثال 6 : تحتوي مزرعة على 100 ارنب وقد تمت مراقبة المزرعة حتى ماتت جميع هذه الارانب وقد لوحظ ان 20منها قد ماتت خلال العام الاول قبل وصولها عاماًكاملا وان 40نهاقدماتت خلال العام الثاني. وان 30 قدماتت خلال العام الثالث. وان العشرة الاخيرة قد ماتت خلال العام الرابع.

جدول 3 : الحياة للارانب في هذه المزرعة

			.حت		احدا	
$\frac{180}{100} = 1.8$	180	§و=آبن0	ههعه	ي	20	0
90 دان	90	40؟ = 60	0.50	80	40	1
!§ = 0.750	30	10؟ = 2S	0.75	40	.30	2
ؤ=0.5 10	5	5تي	1.00	10	10	3 -

مثال 7 : هناك طريقة اخرى لانشاء جداول الحياة فبدلاً من ان نمضي 4 اعوام لانشاء الجدول 3 نستطيع انشاء الجدول خلال عام واحد ويتم ذلك على النحو التالي لنفترض ان المزرعة بها 100 ارنب ولدت خلال العام وان بها 78 عمرها عام وان §0 عمرها عامين وان 45 عمرها ثلاثة اعوام افترض اننا قد راقبنا المزرعة مدة عام و سجلنا عدد الوفيات من كل فئة عمر خلال ذلك العام وكانتت الوفيات كما هو موضح في الجدول 4 .

العمر بالسنوات	العدد عند بداية العام	الوفيات خلال العام
0	100	20
1	78	39
2	80	60
3	45	45

اوجد جدول الحياة لمزرعة مماثلة على فرض ان بها 1000 ارنب عمرها اقل من عام.

الحل: نجد اولاً نسبة الوفيات $n^{d}x$ خلال العام من كل فئة عمر ثم نجد .ح/ي $n^{d}z$ ،

جوه ئ7 مج

فنحصل على الجدول 5 تأكد من صحة الحل

جدول 5

ءلتضفة:	ت٦جشحً- كدمم	لو٦ثذ٦أوة-ة٨٦بيي.	[redacted]	يم١حئ٦.ي:ي- جئذذق١.جحي ى٠ير؛ةغ|٠٤ح؛غ شة ى ة		
1 2 3	0.20 0.50 075 1.00	**1000** 800 400 100	200 400 300 100	900 600 250 50	1800 900 300 50	1.800 1125 0.750 0.500

من الجدير بالملاحظة انه يمكن حساب كينح من معدل الوفاة المحدد بالعمر ئجد

لكل] بدلاًمنلكل 1000 أي انه:

صد اثعقيتة صن الاتحلصن لتيقن احنحما حتتث قطل
حدألاشقاص ض دعجءع ؤكعنج كأتلاً

تحش ارقصن فلالإخ٦عب امن كوم٧ خأبره كم q
حد الاحلص في المجتمح تدبدليةل لحك

وللحصول على قيمة تقريبية للعدد $n^{q}x$ من العدد يرنفترض ان الوفيات خلال

العام تتوزع بشكل منتظم على العام لهذا فان حجم المجتمع عند منتصف العام يساوي

ممه-عد

اذن:

$$= \frac{n^{d_x}}{l_{x-\frac{1}{2}} \quad n^{d_x}}$$

Ixjznh 2mx

ج?7ه2

مثال 8: يمثل الجدول 6 معدل الوفاة المحدد بالعمري جدول 6

0.40	0
0.50	1
0.30	2
0.20	3

احسب احتمال وفاة شخص ضمن الفترة [ممه مم رمم] أي احسبي

الحل: نستعمل القانون

جلب 2

للقل0ب
هعلب2

- **2 X0.4Q**
ل؛علك0ب2

0.30 80
- 2.40 - 240
ت_

و بالمثل يمكن ايجاد بقية الاحتمالات. احسبها بنفسك

جدول الحياة من الدراسات المتابعة :

في بعض الدراسات الصحية نقوم بمتابعة المرضى بعد علاجهم و خروجهم من المستشفيات ففي دراسات الصحة النفسية مثلاً بعد اتمام علاج المريض و خروجه الى المجتمع يقوم الطبيب النفسي بمتابعة حالته للتاكد من مدى تكيفه مع المجتمع الذي يعيش فيه ، ومن التعرف على اسباب امكانية عودته للعلاج مرة اخرى في هذه الدراسات و امثالها يمكن تكوين جداول الحياة وفق الاصطلاحات التالية :

مثلاً اذا اجريت عملية جراحية لقلب المريض فيمكن اعتبار العملية الجراحية ولادة قد تكون ادت الى حياة او موت . واذا توفرت لدينا جداول الحياة لمثل هذه العملية نستطيع ان نقدر توقعاتنا لحياته و احتمال ان يعيشيمن الاشهر او السنين بتد اجراء العملية الجراحية . وكمثال اخر ربما اعتبرت الولادة هي خول المستشفى الوفاة هي الخروج منه . ولهذا عند انشاء هذه الجداول علينا ان نحدد تاريخ الولادة وتاريخ الوفاة (وقد يكون تاريخ الوفاة غير معلوم و عندهانعتبر آخر مرة شوهد فيها الشخص على قيد الحياة تاريخ الوفاة او الانسحاب) وبعد الحصول على البيانات الخام من الضروري تحديد فترة مناسبة لتستعمل في جدول الحياة بطول محدد ، لنعرف الان الرموز التالية التي سنستعملها في هذه الجداول :

X هي الوحدة وقدتكون أي عددطبيعي اوصفر

متم عدد الاشخاص الذين هم تحت المراقبة (احياء) في الزمنء

يح عدد الوفيات التي تحدث بين الزمن-، $x+7$

همحد عدد الانسحابات بيني، $x+7$

لآؤ تقدير احتمال وفاة شخص كان حياً عند الزمنت و توفي قبل الزمن $x+7$ ي2 تقدير احتمال حياة شخص بعد/ير علماً بانه كان حياً عندر

يى تقدير احتمال استمرار حياة شخص-ر من الفترات بعد الولادة. لاحظ ان :

ن)   ئاً-ت=عم

2)   بج-ره-ر-ع»=تاًع

3)   بز مؤمداً

اذا حدث الانسحاب عند بداية الفترة

(ح•=مماً

اذا حدث الانسحاب في نهاية الفترة

ومع هذا سنغض النظر عن وقت الانسحاب و نقدر

5لم ا مهياة،م?

وهذا سنستعمله في هذا البند

من الواضح ان

6) !مق

وكذلك

وذلك لانه حتى تستمر حياة شخص x من الفترات بعد الولادة عليه ان يكون قد استمر على قيد الحياة في الفترات 1,^ السابقة ومن ثم تستمر حياته في الفترة التي ترتيبها تد، وعليه فاحتمال ان تستمر حياته من الفترات تساوي حاصل ضرب هذين الاحتمالين وهو كما ذكرنا سابقاً :

وبعد ذلك نرتب هذه المقادير في جدول مكون من سبعة اعمدة تحمل العناوين التالية على الترتيب:

ولتوضيح ذلك نأخذ المثال التالي :

مثال (9) : اوجد جدول الحياة للبيانات المعطاة في الجدول 7 .

جدول 7

، ، ،'	H		
0	200	14	38
1		7	24
2		6	20
3		3	19
4		3	66
.5			

الحل :

1) نقوم بتعبئة عمود -٦م وفق العلاقة: ر-ئره٦-ر-ة،-ت-ت»=«

فمثلا

هوا مع» – د »

= 200-14-38

-148

وبنفس الاسلوب نجد ان :

0 zzf , 69 =جر, 91=ذم 117’ $n\ 2=2$

2) نقوم بتعبئة عمود خؤ وفق العلاقة

(% ابدائه

-آ1

غكح0٠8=

LOJS 0x٢8W0:

٢٠١

La 6'0=47=

لأءهذه

٣؟١ة

,'م,ء٩ر',٠

ح) : جق٢٦ال٢ةم ٢؛ ٩٢٦ ي٤م

-»٥د

ع 0 0-1 =

/>ع١جه

أ١أ٢

لا-1=٢لآ

ع) : ع٢آ٢١ م مكلآ ١ب٩ ي ٢٤

g مم٣مام قجمس مأا7 32 لتي جيم ٢٥٢ لئم برلم؛ صح م

ع 0’0 =

I8I/H:

(61 "003)أوا =

ؤءلمم١لآه-/٠

والجدول 8 يوضح جدول الحياة المتابع لهذا السؤال

		.•■.أ■" 'حلالا'٦٠ذ٠ث٦'ذ٦نثر	- .	1	ي	
1.0000	0.9227	0.0773	38	14	200	0
0.9227	0.9485	0.0515	24	*7*	148	1
0.8752	0.9439	0.0561	20	6	**117**	2
0.8261	0.9632	0.0368	19	3	91	3
0.7957	0و167	0.0833	66	3	69	4
0.7294					0	5

if