

مبادئ
الإحصاء والقياس الاجتماعي

د. مهدي محمد القصاص
أستاذ علم الاجتماع المساعد
كلية الآداب - جامعة المنصورة
2007

بسم الله الرحمن الرحيم

(وَكُلُّ شَيْءٍ أَحْصَيْنَاهُ كِتَابًا)

صدق الله العظيم
الآية (29) سورة النبا

مقدمة

الإحصاء علم يهتم بالمعلومات والبيانات - ويهدف إلى تجميعها وتبويبها وتنظيمها وتحليلها واستخلاص النتائج منها بل وتعميم نتائجها - واستخدامها في اتخاذ القرارات ، وأدى التقدم المذهل في تكنولوجيا المعلومات واستخدام الحاسبات الآلية إلى مساعدة الدارسين والباحثين ومتخذي القرارات في الوصول إلى درجات عالية ومستويات متقدمة من التحليل ووصف الواقع ومتابعته ثم إلى التنبؤ بالمستقبل .

ولم تعد البحوث الاقتصادية والاجتماعية والإدارية و غيرها في وقتنا المعاصر ، و في ظل التقدم التكنولوجي الهائل في كافة ميادين حياتنا اليومية ، تكتفي بمجرد عرض المشاكل و دراسة الظواهر و تحديد الأسباب و استخلاص النتائج و اتخاذ القرارات بطريقة سطحية مجردة عن أسلوب الإقناع و التقدير والقياس .

ولقد أصبح الاتجاه العام في مثل هذه البحوث و الدراسات هو استخدام طرق القياس الكمية ووسائل الإقناع الإحصائية و ذلك لتحديد الخصائص وإبراز الاتجاهات العامة في الظواهر الاجتماعية والإدارية ، و تحليل العلاقات المتشابكة و المتبادلة بين الظواهر علي أساس موضوع غير متميز .

وعلم الإحصاء يعطي للباحثين في مجال العلوم الاقتصادية والاجتماعية والإدارية ، العديد من الطرق والأساليب اللازمة لضرورة القيام بالدراسات والبحوث الاقتصادية والاجتماعية والإدارية والجغرافية علي أساس من القياس لحركة العديد من المتغيرات المحددة للظواهر موضوع الدراسة .

وتستخدم كلمة الإحصاء لتشير إلي عملية جمع البيانات الكمية و الأساليب المستعملة في معالجة تلك البيانات ، و قد نعني بهذه الكلمة أيضا عملية استخلاص بعض الاستنتاجات من دراسة عينة صغيرة لصياغة تعميمات يمكن تطبيقها علي مجتمعات اكبر حجما .

فبحوث الرأي العام علي سبيل المثال تقوم علي مقابلة و دراسة عينة صغيرة من أفراد المجتمع و لكن نتائجها تستخدم في الاستدلال علي اتجاهات الرأي العام في المجتمع ككل . و بذلك يمكن القول بان الإحصاء يشير إلي طرق تنظيم و تلخيص البيانات والي الأساليب التي تستخدم في تحليل و تفسير النتائج واستخلاصاتها يمكن تعميمها علي مجتمع الدراسة.

فالإحصاء هو علم يبحث في طريق جمع الحقائق الخاصة بالظواهر العلمية الاجتماعية التي تتمثل في حالات أو مشاهدات متعددة ، وفي كيفية تسجيل هذه الحقائق في صورة قياسية رقمية، وتلخيصها بطريقة يسهل بها معرفة اتجاهات الظواهر

وعلاقات بعضها ببعض ، ويبحث أيضاً فى دراسة هذه العلاقات والاتجاهات واستخدامها فى تفهم حقيقة الظواهر ومعرفة القوانين التى تسير تبعاً لها .

ومن هنا يتضح أن الإحصاء لا غنى عنه لأى باحث فى شتى المجالات المختلفة وإذ اعتمد فى بحثه على الأسلوب العلمى. أي أن الإحصاء هو عصا الباحث التى تقوده إلى الطريق الصحيح، وهى الأداة التى تساعد على تفسير الظواهر التى يدرسها وتوضيح النتائج التى يحصل عليها ودلالات البيانات والأرقام التى يحصل عليها .

د. مهدى محمد القصاص

كوم حمادة فى فبراير 2007

وصف المقرر وهدفه

يهدف هذا المقرر إلى تعريف طلاب قسم الاجتماع بعلم الإحصاء وأهميتها ودورها في تسهيل عمل الباحث الاجتماعي في التعامل مع مجتمع البحث بدءاً من أخذ العينات وكيفية جدولة البيانات وتفريغها وتبويبها ووصفها (مقاييس النزعة المركزية والتشتت وأشكال توزيع البيانات) ودرجة ونوع العلاقات بين المتغيرات ومستوى قياسها ودلالاتها واختباراتها كاختبار (ت ، ف ، كا²) الخ ، وذلك بهدف إكساب الطالب مجموعة من الخبرات في مجال الإحصاء الاجتماعي كي تساعده في عرض نتائج البحوث الاجتماعية الكيفية بصورة كمية محدده وواضحة ومختصره ودقيقة .

وفيما يلي وصف المحتوى وهدف كل فصل حيث يهدف إلى تعريف وإفهام واستخدام الطالب لـ :

- 1- التعرف بمعنى كلمة الإحصاء وتطور علم الإحصاء وأهمية الإحصاء للباحث الاجتماعي .
- 2- أنواع المتغيرات المختلفة وكيفية التفرقة بين كل نوع منها وتصنيفها بشكل صحيح .

- 3- العينات والمقصود بها وأنواعها المختلفة وطرق سحب العينات والطرق المختلفة لحساب حجم العينة من المجتمع المفتوح والمغلق .
- 4- القدرة على تبويب البيانات الإحصائية التي يحصل عليها فى بحثه فى جداول تكرارية وأيضاً عرض هذه البيانات بالرسم البيانى بطرقه المختلفة.
- 5- القدرة على وصف وتحليل البيانات من خلال مقاييس النزعة المركزية المختلفة مثل الوسط الحسابى والوسيط والنوال وتعريف الطالب بطرق حساب كل من تلك المقاييس السابقة من البيانات المبوبة والغير مبوبة وتدريب الطالب على تحديد نوع التواء التوزيع .
- 6- القدرة على وصف البيانات من خلال مقاييس التشتت المختلفة مثل المدى والتباين والانحراف المعياري والانحراف المتوسط وتعريف الطالب بطرق حساب كل من تلك المقاييس السابقة من البيانات المبوبة والغير مبوبة .
- 7- القدرة على تحليل التباين بين متغيرين أو أكثر عن طريق حساب قيمة نسبة "ف" ومقارنتها بقيمة "ف" الجدولية لتحديد مدى دلالتها إحصائياً .
- 8- تمكين الطالب من القدرة على استخدام اختبار "ت" لتحديد ودراسة العلاقة بين متغيرين فقط متجانسين وغير

متجانسين عن طريق حساب قيمة "ت" ومقارنتها بقيمة "ت"
الجدولية لتحديد مدى دلالتها إحصائيا .

9- القدرة على استخدام اختبار "كا²" لتحديد ودراسة العلاقة
بين متغيرين عن طريق حساب قيمة "كا²" ومقارنتها بقيمة
"كا²" الجدولية لتحديد مدى دلالتها إحصائيا .

10- تمكين الطالب من القدرة على تقدير قوة العلاقات
بين المتغيرات من خلال استخدام معاملات الارتباط
المختلفة والتنبؤ بقيمة متغير عن طريق معرفة قيمة متغير
آخر من خلال حساب معادلة خط الانحدار بين المتغيرين .

11- تمكين الطالب من القدرة على تقدير ثبات وصدق
الاختبار من خلال حساب قيمة معامل الثبات ومعامل
الصدق .

ولتحقيق الهدف من ذلك المحتوى يستلزم استخدام بعض الوسائل
منها :

- 1- جهاز عرض الشفافيات .
- 2- جهاز كمبيوتر .
- 3- داتا شو .
- 4- سبورة بيضاء وأقلام بألوان مختلفة .

ويتم قياس ذلك من خلال التقويم وفق الأسباب الآتية :

- 1- مناقشات .
- 2- أوراق عمل .
- 3- مجموعات عمل لحل التمارين .
- 4- الاختبار التحريري .

- الفصل الأول
- علم الإحصاء تعريفه و أهميته
- أولا : تعريف علم الإحصاء .
 - ثانيا : أهمية علم الإحصاء .
 - ثالثا : تطور علم الإحصاء .
 - رابعا: علاقة علم الإحصاء بالعلوم الاجتماعية .

أولا : تعريف علم الإحصاء

هو فرع من فروع الرياضيات يشمل النظريات و الطرق الموجهة نحو جمع البيانات ووصف البيانات و الاستقراء و صنع القرارات .⁽¹⁾

و عندما نتكلم عن علم الإحصاء لا نعنى بذلك البيانات الإحصائية وإنما نقصد حينئذ الطريقة الإحصائية . وهى الطريقة التى تمكننا من جميع الحقائق عن الظواهر المختلفة فى صورة قياسية رقمية و عرضها ببيانيا و وضعها فى جداول تلخيصية بطريقة تسهل تحليلها بهدف معرفة اتجاهات هذه الظواهر و علاقات بعضها ببعض⁽²⁾ .

ولقد كان الهدف الرئيسى من علم الإحصاء قديما هو عد أو حصر الأشياء المراد توفير بيانات إحصائية عنها ، وكانت الجهة التى تقوم بإعداد الإحصاءات على مستوى الدولة تعرف بمصلحة التعداد ولذلك كان التعريف القديم لعلم الإحصاء أنه علم العد ، أى العلم الذى يشتمل على أساليب جمع البيانات الكمية عن المتغيرات و الظواهر موضوع الدراسة .

ولكن مع تطور المجتمعات وتشابه جوانب الحياة الاقتصادية والاجتماعية الحديثة بها ، لم يعد مجرد توفير البيانات الكمية عن المتغيرات و الظواهر موضوع الدراسة يفى بحاجات متخذى القرارات و صانعى السياسة العامة إلى تكوين صورة

متكاملة الجوانب عن مجتمعهم والمجتمعات المحيطة به . فقام العلماء بتحديث نظريات علم الإحصاء وأساليبه وأدواته لكي يعين الباحثين وغيرهم على استخلاص استنتاجات معينة من البيانات الكمية التي أمكن لهم جمعها عن طريق العد .

من ذلك على سبيل المثال ، أن نظرية العينات ساعدت الباحثين على استخلاص استنتاجات عديدة من دراسة عدد صغير من الأفراد أو الأشياء - العينة - وتعميم تلك الاستنتاجات على المجتمع الذي سحبت منه العينة بأسره ولذلك يعرف علم الإحصاء حديثاً بأنه : (علم متكامل يتضمن الأسلوب العلمى الضرورى لتقصى حقائق الظواهر واستخلاص النتائج عنها ، كما يتضمن أيضاً أيضاً النظرية اللازمة للقياس واتخاذ القرار فى كافة الميادين الاقتصادية والاجتماعية والسياسية والعسكرية)⁽³⁾

ثانياً: أهميه علم الإحصاء

لقد أصبح لعلم الإحصاء أهميه بالغه فى حياتنا الحديثة فصارت الإحصاءات مألوفة لدينا وتمثل جانباً مهماً من المعلومات التى نطالعها كل يوم مثل جداول النقاط التى تحرزها أندية كره القدم وتنشر فى الصحف والمجلات والتقديرات الخاصة بالتنبؤات الجوية ومؤشرات البورصة وانجازات الحكومة فى مجال الإسكان والتعمير والتغيرات التى تطرأ على أسعار العملات وأثمان السلع . وربما يتساءل المرء عن أهميه الإحصاء بالنسبة لدارس علم

الاجتماع أو علم النفس معتقداً أن الإحصاء موضوع يدخل فى صميم تخصص التجاريين والاقتصاديين والواقع أن الباحث الاجتماعى والمتخصص فى العلوم الاجتماعية بوجه عام يحتاج فى كثير من الأحيان إلى استخدام الأرقام لكي يلخص ويعرض بها مجموعه من المشاهدات التى تتعلق بظاهرة يهتم بدراستها ، فقد يطلب منه أن يقدم تقريراً عن مدي التطور الذى حققه برنامج معين لمحو الأمية بين نزلاء المؤسسة التى يعمل بها ، وقد يكلف بدراسة الأسباب التى تجعل الذكور أكثر تقدماً وحرصاً على التعليم من الإناث فى المدرسة التى يشتغل فيها .

ففى كل مناسبة من هذه المناسبات سيحتاج الباحث أو الدارس إلى أداة من الأدوات الإحصائية لكي يستخدمها فى تلخيص أفكاره والتعبير عنها بصورة محددة ومؤثرة ، فالعبارة التى مؤداها " لقد نجحنا فى محو أمية 90% من العاملين الأميين بالمصنع " أقوى وأشد من العبارة التى مفادها : " لقد نجحنا فى محو أمية عدد كبير من العاملين الأميين بالمصنع " : (4) يحتل الإحصاء (أو الأساليب الإحصائية) أهميه خاصة فى الأبحاث العلمية الحديثة ، إذ لا تخلو أى دراسة أو بحث من دراسة تحليليه إحصائية تتعرض لأصل الظاهرة أو الظواهر المدروسة فتصور واقعها فى قالب رقمى ، وتنتهى إلى ابرز اتجاهاتها وعلاقتها بالظواهر الأخرى . (5)

إن دراسة الإحصاء أمر له فوائد كثيرة بالنسبة لدارسي العلوم الاجتماعية وخاصة بعد أن تفتحت أمامهم مجالات عمل كثيرة فى تنظيمات الشرطة والعلاقات العامة بالشركات ومراكز البحوث وغير ذلك من مجالات العمل المختلفة . بل إن المعرفة بالإحصاء قد تفيد الإنسان على المستوى الشخصي فتكسبه مهارة التخطيط لحياته الاقتصادية الخاصة .

ولكن ينبغي أن نشير إلى أن النتائج التى تسفر عن تطبيق أداة إحصائية أو أكثر ليست نتائج قطعية أو غير قابله للتمحيص والمراجعة . فإذا كانت الأدوات الإحصائية تستطيع أن تعين المرء على وصف البيانات وتصميم التجارب وعلى اختبار العلاقات بين الأشياء والوقائع التى يهتم بها إلا أن ذلك لا يلغى بصيرته السوسولوجية وخبرته المهنية .

وبعبارة أخرى ، يقتصر دور الأدوات الإحصائية على توفير المؤشرات المبدئية التى تساعد الباحث على رفض أو قبول الفروض التى يقوم بدراستها فى حدود درجه معينه من الثقة . والإحصاء أيضا أداة لا تستخدم إلا فى العثور على إجابات عن أسئلة تتصل ببيانات يمكن التعبير عنها بصيغ كمييه . وهناك فى مجال العلوم الاجتماعية موضوعات لا حصر لها لا يمكن صياغة البيانات الخاصة بها فى صورة كمييه على نحو دقيق ، ومن ثم لا يستطيع الباحث استخدام التحليل الإحصائي فى دراستها .

من ذلك على سبيل المثال ، دراسة التجربة الدينية بين جماعه المؤمنين بدين معين ، إذ أن عدد مرات تردد المرء على المسجد أو على الكنيسة في الشهر ليس دليلا فى حد ذاته على انه من الصالحين ، ولكنه مؤشر مبدئي على الصلاح .

ومما يعكس أهميه علم الإحصاء أنها يستخدم فى توجيه عمليه جمع البيانات وفى تفسير العلاقات التى تعكسها تلك البيانات . ومن ابرز المجالات التى تستخدم فيها المعالجات الإحصائية إجراء المقارنة بين عديد من الأشياء فى كثير من المناسبات . ويمكننا القول أن الحياة الإنسانية سلسله من المواقف التى يتخذ فيها الفرد قراره بناء على ما تسفر عنه المقارنة التى يجريها بين عديد من الاحتمالات وهذه المقارنة فى جوهرها عمليه إحصائية تقترن بالقياس والتقييم والتقدير . فنجاح الإنسان فى حياته يتحدد وفق مقياس معين فى ذهنه يقدر به هذا النجاح ، وحرية الفرد فى مجتمعه تقاس أيضا وفق معايير يتعارف عليها الأفراد فى مجتمعهم .

وبعبارة أخرى ، إن حياتنا تذخر بعمليات من القياس والتقدير الإحصائي فنحن علي سبيل المثال ،عندما ننزل إلى السوق لشراء سلعه معينه ، فى موسم التنزيلات ، نهتم وبطريقه لا شعورية بحساب ثمن هذه السلعة بالنسبة إلى إجمالي النقود التى فى حوزتنا ونقدر ما إذا كان الباقي من هذه النقود وسوف

يكفيها حتى نهاية الشهر أم لا وما إذ كانت نسبة التنزيلات على السلعة حقيقية أو مزيفة 00000 الخ فى كل هذه العمليات الفكرية نحن نستعين بعمليات إحصائية ومقارنات مستمرة بين المواقف المختلفة . فضلا عن ذلك ، إن ما نطلق عليه ظاهرة اجتماعيه أو طبيعیه ما هو فى الواقع إلا سلسله متكررة من الواقع التى يمكن رصد حدوثها المستمر عبر فترة من الزمن وبنفس الوتيرة بطريقه إحصائية . (6)

ثالثا : تطور علم الإحصاء

تطور علم الإحصاء وتطبيقاته عبر سنوات طويلة ، وتم ذلك بجهود كثيرة من العلماء من دول مختلفة وكان . التطور بطيئا إلى أن جاء القرن العشرين ليشهد معدلا هائلا للتطور فى النظريات الإحصائية فى مجالات كثيرة .

ويرجع الاهتمام بالإحصاء إلى عصور قديمه ، وان تعداد السكان عند القدماء المصريين وفي الصين أمثلة توضح اهتمام الحكومات منذ القدم بالمعلومات الاجتماعية وذلك لأغراض التنظيم والتخطيط فى أحوال السلم والحرب .

ويبدو أن كلمه إحصاء (statistics) قد ظهرت لأول مره عام 1749 وهى مشتقه من الكلمة اللاتينية (status) أو الايطالية (statista) وتعني كلاهما الدولة السياسية . ومن الطبيعي أن تكون الدولة أول من اهتم بجمع البيانات وذلك لإدارة شؤون البلاد

خاصة عن السكان لأغراض حربية وضريبية ، وامتدت بعد ذلك لتشمل إحصاءات حجم السكان والمواليد والوفيات والإنتاج والاستهلاك والثروة 00000 الخ . وهكذا بدء العلم وتطوره باعتباره علم الدولة أو علم الملوك . (7)

ولقد تطور علم الإحصاء من مجرد فكره الحصر والعد إلى أن أصبح الآن علما له قواعده ونظرياته ويرجع الفضل في ذلك إلى كثير من العلماء من أمثال عائلة برونلي Bernoulli وفرديريك جـاوس F.gauss وكييتليه Quetlet وجولتون F.galton وأخيرا كارل بيرسون Karl.pearson وبولي A.bowley وبول U.yule فيشر L.fisher و..... الخ . (8)

وجاء التطور في علم الإحصاء بصفه عامه ملازما وموازيا للتطور في نظرية الاحتمالات . فقد نشأت نظريه الاحتمالات على أساس رياضى فى (1494) بواسطة باسيولي Lucapacidi . ومن الدراسات الفلكية لكل من كبلر (1630-1517) Keplr وجاليليو (1642-1564) Galilio قاما بتطوير نماذج الاحتمالات . غير أن التاريخ الحقيقي لنظريه الاحتمالات بدء فى القرن السابع عشر حيث وضعت أسسها فى عام 1654 بواسطة كلا من العالمين : باسكان (1623 1662) Pascal,B. عالم الرياضيات والفيزياء

والفيلسوف الفرنسي - وكذا العالم فرمات (Fermat 1608 - 1665) .

وبعد ذلك بثلاث سنوات قام هينجينز Huygens (1629 - 1695) بنشر كتيب صغير في موضوع المعالجة الرياضية لفرص الفوز في مباريات ورق اللعب وزهرة النرد .

وفى نفس الوقت تقريبا قام جروننت grunt (1620 - 1674) بنشر ملاحظاته عن معالجة البيانات المتعلقة بالحكومة خاصة في النواحي الطبيعية والسياسية والتجارية والنمو والوفيات والأمراض.

وقد كان العمل الذى قام به هيجيتتر دافعا للكثيرين لدراسة النظريات والمشاكل المتعلقة بمباريات الصدفة ومنهم برنوللى (1654 - 1705) ودي موافر (De Moivre 1667 - 1754) واربوتنوت Arbuthnott ولابلاس (laplace 1749 - 1827) وجاوس Gauss (1777 - 1855).⁽⁹⁾ ويعد العالم البلجيكى كتيليه (1796 - 1874) أول من وضع قواعد محددة لعلم الإحصاء ، وكلمة إحصاء في الوقت الحاضر ذات معان متعددة فمنها يفهم جمع المعلومات التى تبين الحالة في الدولة مثل عدد المواليد والوفيات وبيانات عن المحاصيل والتجارة الخارجية الخ ويسمى نشر الأجهزة الحكومية لمثل هذه المعلومات في شكل كتب وتقارير " بالإحصاء الرسمي " .

وأخيرا يفهم بالإحصاء فرع من العلم له نظريته الخاصة .
وعلم الإحصاء ، شأنه في ذلك شأن أى فرع آخر من فروع العلم
له أسلوبه وموضوعات البحث الخاص به (10)

وكلمة إحصاء (Statistics) لها ثلاث معاني :

(1) الإحصاءات أو البيانات : مثال ذلك إحصاءات السكان

والمواليد والوفيات والإنتاج - الصادرات - الاستهلاك .

(2) المؤشرات المحسوبة من عينة (العينة هى مجموعة

جزئية من الوحدات محل الدراسة)

(3) علم الإحصاء : وهو فرع من فروع الرياضيات يشمل

النظريات والطرق الموجهة نحو جميع البيانات ووصف

البيانات والاستقراء وصنع القرارات . (11)

ولقد تطور علم الإحصاء وتنوعت طرائقه ، وأصبح له من

القواعد ما يمكنه من القيام كعلم مستقل يمكن الاستعانة به في

رسم وتحديد السياسات الاجتماعية التى ينتهجها المجتمع . كما

برز دور الإحصاء - بما يقدمه من بيانات وإحصاءات - فى

عمليات التخطيط والتنمية التى تمر بها مجتمعاتنا اليوم (12)

ويمكن القول أن الإحصاء تخدم الباحثين فى جميع

الميادين العلمية وصانعى القرارات فى شتى المجالات العملية ،

ولا يكاد يخلو ميدان من ميادين البحث العلمى إلا وطرقته

الإحصاء وساهمت فيه مساهمة فعالة . وقد أثار روبرت

بارسوز فى مستهل كتابه " التحليل الإحصائي " أن كلمة إحصاء لها أكثر من استخدام إلا أن أكثر الاستخدامات شيوعاً هو ذلك الذى يرى أن كلمة إحصاء تشير إلى تلك الأساليب والإجراءات التحليلية المستخدمة فى معالجة البيانات الرقمية.

بمعنى أنه للحصول على معلومات ذات قيمة من تلك البيانات الرقمية فإنها يجب أن تخضع للتحليل الإحصائي **Statistical Analysis** بمساعدة تلك الأساليب والإجراءات والأدوات التي توفرها لنا الإحصاء.

ويذهب كل من **Whittaker, Startup** إلى وجود ثلاثة استخدامات لكلمة إحصاء .

أ- للإشارة إلى الحقائق الرقمية التي جمعت بطريقة منتظمة من الواقع الاجتماعي.

ب- تشير إلى الأساليب المستخدمة فى جمع ، وتصنيف وتحليل البيانات الرقمية.

ج- للإشارة إلى صفة أو خاصية للعينة تحت الدراسة. والقاموس الحديث لعلم الاجتماع الذى وضعه كل من **George and Achilles Theocorson** يقدم رؤية لا تختلف عما سبق فيما يتعلق بكلمة إحصاء سواء من حيث المعنى أو الاستخدام فهي تعني مجموعة من الأساليب التي

تستخدم في جمع ، وتصنيف ، وتبويب وعرض وتحليل البيانات الكمية ، والإحصاء بهذا المعنى لا تقف عند حد الوصف Description بل تتعداه إلى مرحلة الاستنباط Induction والاستدلال Inference كما تستخدم كلمة إحصاء للإشارة إلى البيانات الرقمية والتي عادة ما تسمى " إحصاءات " حيث تأخذ صيغة الجمع .

ومن هنا فان كلمة إحصاء تعني تلك الأساليب والأدوات والإجراءات الإحصائية التي يلجأ إليها الباحث وهو بصدد القيام بدراسة ما في عملية الجمع ، وتصنيف ، وتلخيص وعرض ، و تحليل البيانات الرقمية (13).

رابعاً : علاقة علم الإحصاء بالعلوم الاجتماعية

تأثرت العلوم الاجتماعية وخاصة علم الاجتماع وعلم النفس وعلم السياسة بالتطورات . التي حققها علم الإحصاء ، واستعان العلماء الاجتماعيون بمنهج جديد في دراساتهم . وهو المنهج الاحصائي الذي ينطوي علي نفس خطوات المنهج العلمي في البحث ، حيث يقدم علي عمليتين منطقيتين هما القياس و الاستنتاج ، وإذن يقوم العالم بملاحظة الحقائق في البداية ثم يجري تجاربه ويرصد عددا من النتائج التي يستخلصها من تلك التجارب بنمط أو إطار عام للظاهرة. وبعد

أن يقوم بصياغة نظريته علي ذلك النحو ، ينتقل إلي عملية الاستنتاج التي تعينه علي التنبؤ بسلسلة من النتائج الأخرى .

ومن أشهر الدراسات السوسولوجية التي اعتمدت علي المصادر الإحصائية ، دراسة دور كايم عن الانتحار. وفيها يذهب إلي (انه إذا كان المرء يريد أن يعرف كل ما يتفرع عن الانتحار كظاهرة جمعيه فانه ينبغي أن ينظر إليها في شكلها الجمعي من خلال البيانات الإحصائية) وقد اعتبر دور كايم أن المؤشرات الإحصائية عن الأسباب التي دفعت الأفراد إلي الانتحار بمثابة مصدر لمعرفة الدوافع المفترضة وراء الإقدام عليه . وهكذا نجد أنه قد وضع فروضه علي أساس من الأرقام والإحصاءات التي رأي أنها تعين لنا اقرب نقطة لبدء بحثنا السوسولوجي.

وقد حقق المنهج الإحصائي في السنوات الأخيرة تقدما هائلا ، وخاصة بعد استخدام الحاسبات الالكترونية ، وذلك في ميادين العلوم الاجتماعية المختلفة ، وقد انعكس هذا التقدم بدوره علي التطورات والأدوات الإحصائية ذاتها.

وقد استفاد علماء الاجتماع من المنهج الإحصائي في تطوير أدوات بحثهم وخاصة الاستبيان مما أمكنهم من دراسة آلاف المبحوثين في فتره زمنية وجيزة ، وتوافرت لدي الباحثين إمكانية اختبار العلاقة بين ما يرصدونه من ظواهر علي أرض

الواقع وما يفترضونه من افتراضات يحاولون بها تفسير ذلك
الواقع .

وقد ساعد علم الإحصاء علماء السياسة على اقتحام
مجالات عديدة من البحث السياسي مثل دراسة أنماط المشاركة
السياسية وتكوين الرأي العام والحركات والتنظيمات السياسية .
فلو أن عالم السياسة افترض أن هناك ثمة ارتباط بين مستوي
تعليم الأفراد وتعليم من أدلوا بأصواتهم في الانتخابات فان البيانات
التي يتسنى له الحصول عليها من الواقع عن مشاركة الأفراد فى
التصويت الانتخابي وعن مستوياتهم التعليمية لا تنعقد المقارنة
بينها إلا باستخدام المقاييس الإحصائية التي تكشف عن قوة
الارتباط بين الميل للتصويت فى الانتخابات والمستوي التعليمي
للأفراد . وبدون هذه المقاييس الإحصائية تظل البيانات
والمعلومات الميدانية المتوافرة لدى الباحث بلا قيمة حقيقية.

ويستخدم علماء النفس الأدوات والأساليب الإحصائية أكثر
من غيرهم فى القياس النفسى . ويعد علم النفس التجريبي وعلم
النفس الاكلينيكي وعلم نفس الفروق الفردية من المجالات التي
تعتمد اعتمادا جوهريا على المنهج الاحصائى فى تناولها
لموضوعات الدراسة .

ومن يقرأ مرجعا فى القياس النفسى يجد أن علماء النفس
يذهبون إلي أن كل شيء فى مجال علمهم قابل للقياس تقريبا

فوجد لديهم مقاييس للذكاء والشخصية وللعواطف والميول وللاضطرابات النفسية والأمراض العقلية وكل مقياس من هذه المقاييس يخضع ، فى واقع الأمر لأساليب إحصائية صارمة تحدد مدى ثباته وصدقته فى قياس ما صمم لقياسه ويستخدم فى المقارنة بين النتائج التى يتم التوصل إليها من دراسة عينه محددة من الأفراد وتلك التى يتم التوصل إليها من دراسة عينه أخرى (14)

وقد ظهر اهتمام كبير بتطبيق النظريات والطرق الإحصائية فى العلوم الاجتماعية ، فقد أوضح كيتيليه (1796-1874) عالم الفلك الاجتماعي البلجيكي إمكان استخدام الاحتمالات والإحصاء لوصف وتفسير الظواهر الاجتماعية والاقتصادية وقدم مساهمات هامة فى الطرق الإحصائية فى تنظيم وإدارة الإحصاءات الرسمية - وقدم كذلك طريقه عامه للقياس فى الانثروبولوجيا - وقد ساهم عالم النفس الإنجليزي جالتون Galton (1822-1911) فى تطبيق الطرق الإحصائية فى علم النفس ، ووضع أساس علم القياس النفسى (psychometrics) وبدأ دراسة موضوع الارتباط والاتحاد الذي اهتم به وطوره بعد ذلك عالم الإحصاء الإنجليزي كارل بيرسون Pearson (1857-1936). بالإضافة إلى مساهمات أخرى هامة .

كما قدم سبيرمان Spearman (1863-1945) عالم النفس الإنجليزي مساهمات فعالة في دراسة الارتباط ويعد من الرواد في دراسة وتطوير التحليل العاملي .

وقدم عالم الإحصاء الإنجليزي جولست (1876-1937) Gosset مساهمات هامة في مجال التحليل الإحصائي وخاصة في تفسير البيانات المتعلقة بالعينات كما يعد من الرواد المهتمين بتحليل نتائج العينات الصغيرة . وخلال الفترة السابقة كان الاهتمام كله مركزا علي المفهوم الكلاسيكي للاحتمال .

إن مفهوم التكرار النسبي لم يظهر بصورة ملموسة إلا في بداية القرن العشرين حيث تم صياغتها وظهورها في إطار منطقي بمعرفة فون مايسيس vonmises .

وعلي الرغم من أن الرواد من علماء الإحصاء كان إهتمامهم بوظيفة الاستقراء فان الجانب الأعظم من النظرية الإحصائية تم اكتشافه بعد عام 1920 تقريبا فمنذ مطلع القرن العشرين كان الاهتمام منصبا علي تطبيق الإحصاء علي مشاكل علوم الحياة وعلي التجارب الزراعية والصناعية .

كما أن العمل في هذه المرحلة كان مكثفا ومركزا علي التحليل الإحصائي وأساسه المنطقي ، وتمخض عن ذلك مساهمات قدمها عالم الإحصاء الإنجليزي فيشر Fisher (1890-1962) ومن أعماله البارزة نظرية التقديرات ، وتوزيعات المعاينة للعينات

الصغيرة ، وتحليل التباين وتصميم وتحليل التجارب . ومن العلماء الذين ساهموا كثيرا في نظرية التقديرات واختبارات الفروض كلاً من بيرسون *Pearsson, E.s* وكذلك نيمان *Neyman* - ويعتد الثلاثي فيشر - بيرسون - نيمان مؤسس منهج الاستقراء الإحصائي والذي يعرف حالياً بالاتجاه الكلاسيكي . وهو يعتمد على المعلومات المتاحة من العينة فقط .

وقد ظهر في هذه الفترة اتجاه جديد يعرف بالاستقراء البيزياني *Bayesian inference* وذلك بجهود كل من جفريز *jeffreys* ورافري *Ramsey* وديفتي *Definetti* وجود *Good* وسافج *Savage* وندلي *lindley* وآخرون 0 ويعتمد الاستقراء هنا على بيانات العينة بالإضافة إلى المعلومات المسبقة *Prior. Information* وشهدت هذه الفترة أيضاً عملاً مكثفاً كان فيها الاهتمام منصبا على صنع القرارات ، مما أدى إلى نشوء وظيفة حديثة للإحصاء تحت اسم نظرية القرارات الإحصائية *Statistical Decision theory* ويرجع ذلك إلى أعمال والد *Wald (1939)* ونيومان *Neuman, j* ومورجنسترن *Morgenstern* .

وقد صاحب هذا التطور الكبير في النظريات الإحصائية بداية ظهور مجموعة من التخصصات المختلفة تهتم بمجالات وأهداف خاصة - وقد بلغ هذا التطور قدراً هائلاً يكاد يظهرها

وكأنها علوماً مستقلة . ومن هذه التخصصات بحوث العمليات
Operations Research والإحصاء السكاني
Demography ومراقبة الجودة Quality control
والاقتصاد القياسي Econometrics ونظراً لاعتماد العلوم
المختلفة على الرياضيات في فهم ظواهرها وقياسها وتفسيرها ،
فقد أفردت لها فروعاً خاصة تهتم بدراسة ظواهرها باستخدام
الأساليب الإحصائية والرياضية ومنها على سبيل المثال الإحصاء
الحيوي والاجتماع الرياضي والقياس الاجتماعي وعلم النفس
الرياضي والقياس النفسي والقياس التربوي والاقتصاد الرياضي
والتاريخ الاقتصادي الجديد أو القياس التاريخي (15)

إن الأساليب الرياضية والإحصائية المستخدمة في مناهج
البحث بصفة عامة تستخدم الآن في مجال العلوم الاجتماعية
بنجاح . وقد أمكن عن طريقها التوصل إلى بعض الحقائق العلمية
والنظريات ، ولكنها لم ترق في هذا المضمار إلى ما وصلت إليه
العلوم الطبيعية من نظريات علمية وقوانين .

وتصادف العلوم الاجتماعية صعوبات منهجية تحول دون
تحقيق أهدافها في الوصول إلى ما وصلت إليه الأبحاث الطبيعية ،
ومن بين هذه الصعوبات .

- لا تخضع التفاعلات الاجتماعية لنظام آلى مرتب ، ولا تسير وفق مبدأ الاطراد في تتابع الأحداث مما يسهل عملية الحصول على القوانين التى تحكم نظمها .
 - صعوبة التوصل إلى قوانين التنبؤ الاجتماعى . وقد كان الاعتقاد السائد أن السلوك الاجتماعى والعلاقات الإنسانية التى تربط بين الأفراد فى المجتمع إنما تخضع لنظم وقوابل يصب فيها الأفراد أعمالهم وأفكارهم ولا يكون الخروج عما ترسمه الطبيعة لهم من حدود وما تفرضه من التزامات .
 - ليس لدى بعض العلوم الاجتماعية وحدات معينة تستخدم لقياس الظواهر موضوع الدراسة كما هو فى العلوم الطبيعية التى تستخدم وحدات كمية لوصف ظواهرها والتعبير عنها بمعادلات رياضية والتنبؤ بها بتوافر شروط معينة .
 - عدم استجابة البيئة الاجتماعية موضوع الدراسة للغايات التى يقصدها الباحث وعدم تمكن الباحث من السيطرة على كثير من العوامل التى تلعب دورا كبيرا فى سير الحوادث وارتباط بعضها ببعض الآخر .
- والمزايا التى يجنيها الباحث من الطرق الإحصائية يمكن تلخيصها فيما يلى:-
- تساعد الباحث على إعطاء أوصاف على جانب كبير من الدقة العملية .

- فهدف العلم الوصول إلى أوصاف الظواهر و مميزاتها الطبيعية ، وكلما توصل العلم إلى زيادة في دقة الوصف كلما كان هذا دليلا على التقدم العلمى ونجاح الأساليب العلمية . ودقة الوصف تحتاج دائما إلى اختبار مدى ثبات النتائج التى حصل عليها الباحث. فمجرد الوصول إلى نتائج دون التحقق من ثباتها لا يكفى عادة كأساس يعتمد عليه في تفسير الحقائق وتحقيق الفروض.
- تساعد الإحصاء على تلخيص النتائج في شكل ملائم مفهوم فمجرد ذكر الدرجات لا يكفى للمقارنة بين الجنسين بل إن حساب متوسطى الدرجات قد سهل مهمة المقارنة كثيرا فالبيانات التى يجمعها الباحث لا تعطى صورة واضحة إلا إذا تم تلخيصها في معامل أو رقم أو شكل توضيحي كالرسوم البيانية.
- تساعد الباحث على استخلاص النتائج العامة من النتائج الجزئية .فمثل هذه النتائج لا يمكن استخلاصها إلا تبعا لقواعد إحصائية ، كما يستطيع الباحث أن يحدد درجة احتمال صحة التعميم الذى يصل إليه .
- تمكن الباحث من التنبؤ بالنتائج التى يحتمل أن يحصل عليها في ظروف خاصة . فيما عدا الإحصاء يمكن للباحث أن يتنبأ بنتائج ما يجريه من اختبارات في وقت ما لقدرة

أو قدرات خاصة لما ينتظر للأفراد الذين يختبرهم من نجاح في مهنة معينة أو نوع معين من التعليم.

• في كثير من البحوث يهدف الباحث إلى تحديد أثر عامل خاص دون غيره من العوامل مما لا يتسنى تحقيقه عمليا . وهنا يستطيع أن يلجأ إلى الإحصاء فتعاونه على فصل عامل خاص من العوامل المحتملة وتحديد أثره على حده ، كما تعينه على التخلص من أثر العوامل الأخرى التي لا يستطيع تفاديها في بحوثه والتي تؤثر دائما في نتائج كل بحث ، كعامل الصدفة واختيار العينات .

• وقبل هذا كله تهدي الإحصاء الباحث عند تنظيم خطوات بحثه فهو يحتاج إليها في مرحلة تصميم البحث وتخطيطه ، حتى يمكنه في النهاية أن يخرج من بحثه بالنتائج التي يسعى إلى تحقيقها ، فهي تهديه إلى أضبط الوسائل التي تؤدي إلى التفكير الصحيح من حيث الإعداد أو الاستدلال والقياس أثناء خطوات البحث .

وإذا كان هو حال الإحصاء بالنسبة للبحوث العلمية بوجه عام فإن حاجة البحوث الإنسانية أشد ما تكون إلى تطبيق هذه الوسائل . لذلك كانت البحوث النفسية والتربوية والاجتماعية من أصعب البحوث ، وتحتاج إلى حرص زائد ومهارة فائقة من الباحث .

ويمكن تلخيص أسباب ذلك فيما يلي :-

أ (السلوك البشرى فى تغير دائم، ومدى تغيره من فترة لأخرى أوسع مما نزن ، لدرجة تجعل من الصعوبة بمكان إعطاء تنبؤات علمية دقيقة عنه.

ب (السلوك البشرى كثيرا ما يخدع دراسة ، ذلك لان حقيقته قد تختلف كثيرا عما يبدوا عليه ، فهو يحتاج إلى ضبط فى البحث ودرجة كبيرة من الدقة الإحصائية .

ج) السلوك البشرى معقد تعقيدا كبيرا وتتدخل فيه عوامل قد تزيد أو تختلف عما يتوقعه الباحث .

د) البحوث الإنسانية يقوم بها إنسان . ذلك مما يسمح بتدخل العوامل الشخصية كثيرا فى نواحي القياس والوصف بدرجة قد تكون كبيرة أو صغيرة حسب الطرق التى يستخدمها الباحث . وطرق الضبط الاحصائى خير وسيلة تعين الباحث على استبعاد هذه العوامل الشخصية .

إلا انه ينبغي أن يفهم من ذلك أن الإحصاء هو كل شيء فى البحوث العلمية. فالإحصاء فى يد من لا يجيد تطبيقها واستخدامها استخدام الخبير الفنى ، لا تفيد كثيرا . فهى مرحلة تالية لاكتشاف المشكلة وتحديدها ، وهى تتطلب عادة فروض علمية يتوقعها الباحث بناءً على دراساته السابقة وملاحظاته العديدة ، وهى تتطلب كذلك فى آخر الأمر تفسيراً مبنياً على خبرة

علمية وقدر وافى من المعلومات في الميدان الذى يجرى فيه البحث . وكلما كان الباحث مدركاً للأسس التى بنيت عليها الطرق الإحصائية التى يستخدمها ، كلما سهل ذلك عليه تطبيقها تطبيقاً صحيحاً ، وتفسير النتائج تفسيراً مناسباً (16)

ويتضح لنا من مفهوم الإحصاء أنه يمدنا بمجموعة من الأساليب والأدوات الفنية التى يستخدمها الباحث في كل خطوة من خطوات البحث ابتداءً من المرحلة التمهيدية للبحث وما يتضمنه من عملية اختيار لعينة الدراسة وأسلوب جمع البيانات من الميدان ماراً بمرحلة تصنيف ، وتلخيص ، وعرض وتحليل تلك البيانات حتى مرحلة استخلاص نتائج الدراسة ، ويرى البعض أن وظيفة الإحصاء يمكن أن تتلخص في نقطتين

الأولى :- تتمثل في تلخيص البيانات المتاحة وتقديمها في أبسط وأنسب صورة ممكنة . فالباحث عادة ما يجد نفسه أمام مجموعة كبيرة من البيانات الخام التى لا تفصح عن شئ على حين أنه مطالب باستخلاص حقائق علمية واضحة ومحددة من تلك البيانات سواء كانت بيانات مسوح اجتماعية شاملة . أو بالعينة أو بيانات تعدادات سكانية عندئذ يستطيع الباحث من خلال الإحصاء أن يغير من شكل البيانات بعد تصنيفها وتنظيمها وتلخيصها مستخدماً في ذلك الجانب الوصفى من الإحصاء حيث يمكنه أن يطبق هنا مجموعة من المقاييس الإحصائية التى لا تتعدى حد

الوصف مثل مقاييس النزعة المركزية ومقاييس التشتت ومقاييس الارتباط والاتحادار ... الخ ومن ثم يتبين لدينا أن الوظيفة الإحصائية الأولى للإحصاء هي توصيف البيانات المتاحة والخروج منها بمجموعة من المؤشرات والمعدلات الإحصائية .

الثانية : تتلخص في الاستدلال ، ففي مجال البحوث الاجتماعية ، عادة ما تستخدم العينة sample لتمثل المجتمع الذي سحبت منه ويرجع استخدام العينات في البحوث الاجتماعية إلى عدة أسباب لعل أهمها توفير الوقت ، والجهد ، والإمكانات التي تجعل من المتعذر أحيانا وربما من المستحيل أحيانا أخرى دراسة المجتمع ككل . والعينة ببساطة هي جزء أو قطاع من المجتمع تم اختيارها على أساس إحصائي لكي تمثل المجتمع الذي هي جزء منه وهنا يكون دور الإحصاء هو الوصول إلى تقديرات واستدلالات عن المجتمع ككل من خلال المعلومات المتوفرة عن العينة التي تم سحبها من هذا المجتمع ، إذ إن جُل اهتمام الباحث ليس مجرد العينة المستخدمة في الدراسة بل المجتمع ككل ، باختصار فإن الجانب الاستدلالي من الإحصاء يهتم بتقدير معالم المجتمع Population Parameters فيما يتعلق بالظاهرة موضوع الدراسة مستخدما البيانات والمعلومات المتوفرة لديه عن العينة أو ما يسمى بـ Sample Statistics حول نفس الظاهرة في محاولة الوصول إلى تصميمات Generalizations عن مجتمع الدراسة.

هذا بالإضافة إلى اهتمام الإحصاء الاستدلالي باختبار الفروض العلمية . والإحصائية Hypotheses Teting للدراسة. وإذا كانت تلك هي وظائف الإحصاء فى مجال العلوم الاجتماعية والتي يتضح منها بجلاء مدى ما تقدمه الإحصاء للباحث فهناك كلمة تحذير لابد أن يعيها كل من يفكر فى استخدام الأساليب الإحصائية ألا وهى أن التطبيق غير الصحيح للأسلوب الإحصائي ربما يؤدي إلى نتائج غير صحيحة ومضللة كما أن استخدام الأساليب الإحصائية يجب ألا يكون غاية فى حد ذاته بل انه وسيلة الهدف منها هو تبصير الباحث بما هو بصدد القيام به وتبسيط وتوضيح خطوات البحث العلمى . (17)

وهكذا يتبين لنا مما سبق أن دراسة علم الإحصاء وان ثقلت على نفس بعض الأفراد ، تعد ذات أهمية بالغة لأنها تزود الدارسين بالمهارات البحثية التى لم يعد أى فرض فى غنى عنها ، ونحن نعيش عصر الثورة التكنولوجية وتهيمن على حياتنا لغة الأرقام. (18)

المراجع

1. مصطفى زايد ، الإحصاء ووصف البيانات ، 1989 ، ص 23.
2. فاروق عبد العظيم ، مختار الهانسي ، محمد على محمد ، مبادئ الإحصاء ، دار المعرفة الجامعية ، ص 3 .
3. حسن محمد حسن ، مبادئ الإحصاء الإجتماعي ، دار المعرفة الجامعية ، 2000 ، ص ص 15 - 16 .
4. حسن محمد حسن ، مرجع سابق ، ص 16 .
5. فتحي عبد العزيز أبو راضي ، مبادئ الإحصاء الإجتماعي ، دار المعرفة الجامعية ، ص 4 .
6. حسن محمد حسن ، مرجع سابق ، ص ص 17 - 18 .
7. مصطفى زايد ، مرجع سابق ، ص 19 .
8. فاروق عبد العظيم وآخرون ، مرجع سابق ، ص 3 .
9. مصطفى زايد ، مرجع سابق ، ص ص 19 - 20 .
10. غريب محمد سيد أحمد ، الإحصاء والقياس في البحث الإجتماعي ، دار المعرفة الجامعية ، 1989 ، ص 12 .
11. مصطفى زايد ، مرجع سابق ، ص 23 .
12. غريب محمد سيد أحمد ، مرجع سابق ، ص 13 .
13. اعتماد علام ، يسرى رسلان ، أساسيات الإحصاء الإجتماعي ، دار الثقافة للنشر والتوزيع ، ص ص 7 - 8 .
14. حسن محمد حسن ، مرجع سابق ، ص ص 18 - 20 .

15. مصطفى زايد ، مرجع سابق ، ص 20 .
16. غريب محمد سيد أحمد ، مرجع سابق ، ص ص 14 - 18 .
17. اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص ص 8 - 9 .
18. حسن محمد حسن ، مرجع سابق ، ص 20 .

الفصل الثاني المفاهيم الإحصائية

مقدمة :

أولا : الإحصاء الوصفي والإحصاء الاستدلالي

ثانيا : البيانات

ثالثا : المتغيرات

رابعا : المقاييس الإحصائية

مقدمه :

يزخر كل علم من العلوم بالعديد من المصطلحات والمفردات اللغوية الخاصة به والتي يعد الإمام بها خطوة هامة على طريق الدراسة والفهم المتعمق لموضوعات ذلك العلم وعلم الإحصاء لا يختلف فى هذا الشأن عن غيره من العلوم فهو يتضمن عدد قليل من المصطلحات الأساسية التى نرى أن على الدارس أن يلم بتعريفاتها لكي يعي المقصود منها ويتسنى له معرفة كيفية التعامل معها عندما تعرض له فى دراساته وبحوثه ومن ثم يتفادى الخلط بين المصطلحات المختلفة عندما يحاول اختيار الأداة الإحصائية المناسبة لمعالجة البيانات التى قام بجمعها وتختلف الأساليب الإحصائية فيما بينها من حيث الهدف والتدرج من البساطة إلى التعقيد واختيار الأسلوب الملائم يتحدد وفقا لأهداف الباحث ونوعية البيانات المتاحة .

أولا : الإحصاء الوصفى والإحصاء الاستدلالي

(أ) الإحصاء الوصفى Descriptive statistics

ويهدف إلى إدماج وتلخيص البيانات الرقمية بغية تحويلها من مجرد كم من الأرقام إلى شكل أو صورة أخرى يمكن فهمها واستيعابها بمجرد النظر ومن أغلب الأساليب المستخدمة مقاييس النزعة المركزية ، مقاييس التشتت ومقاييس الارتباط والانحدار

ويتوقف استخدام أي منها على نوعيه البيانات ومستوى القياس سواء أكان اسميا أو وصفيا ، أو ترتيبيا ، أو فنويا ، أو نسبة (1).

ويعتقد بعض الدارسين أن وظيفة الإحصاء تقتصر على معالجة مجموعة البيانات الوفيرة التي جمعها الباحث بقصد استخلاص عدد من الجداول الإحصائية وعرضها في عدد من الأشكال والرسوم البيانية وذلك على نحو ما نشاهده في إحصاءات السكان والاستهلاك والإنتاج وغيرها وقد يحسب المرء أن العمليات الإحصائية تدور في جملتها حول إيجاد المتوسطات ودرجات التشتت في البيانات التي يجمعها الباحثون ولكن في الحقيقة أن ما ذكرناه لا يمثل سوى جانب واحد من جوانب الإحصاء وهو الجانب الوصفي ولهذا يطلق على العمليات الإحصائية التي تقوم بهذه الوظيفة مصطلح الإحصاء الوصفي وعلى هذا يستخدم الإحصاء الوصفي في تنظيم وتلخيص ووصف معلومات خاصة بعينة من العينات فمن عينة محددة من العمال يمكن حساب متوسط الإنتاج الذي ينتجونه وحساب نسبة العمل بين أولئك العمال ومعدل الزيادة في أجورهم وهذه المقاييس كلها وصفية بحتة لا تفيد في حد ذاتها ، في الاستنتاج أو التنبؤ وإنما تصف الكيفية التي تتوزع بها البيانات التي تم الحصول عليها من العمال موضوع البحث (2)

وتعتبر وظيفة الوصف من الوظائف الأولية لعلم الإحصاء التي تستخدم في تلمس حقائق الظواهر المختلفة (اجتماعية ، اقتصادية ، جغرافية .. الخ) وباستخدام أسلوب التحليل الإحصائي للبيانات أصبح من السهولة إمكان تحديد خصائص الظاهرة المدروسة حتى عن طريق الأشكال البيانية التي تمثل بيانات الظاهرية عملية تسهل وتبسط تحديد خصائص الظاهرة واتجاهاتها العامة .

والى جانب ذلك يعتمد الوصف فى الإحصاء على استخدام المقاييس والمؤشرات الإحصائية فى تقصى الحقائق وتحديد الخصائص العامة لتوزيع بيانات الظاهرة دون الوصول إلى نتائج أو استدلاله خاصة بالمجموعات الأساسية التى تنتمى إليها الظاهرة (3) .

وعملية جمع البيانات تعد أقدم وظائف الإحصاء ، وهى تتضمن عدد من الأنشطة يختلف مداها من مجرد بحث يقوم به فرد إلى فريق بحث من عدة مئات أو آلاف . وجمع البيانات يكون بعدد من الأساليب وحسب طبيعة البحث أو العمل ، فقد يكون ذلك باستخدام المجموعات المكتبية أو عن طريق تصميم تجربة أو الملاحظة المنتظمة أو المعيشة أو عن طريق الاستبيان أو الاستبصار أو الأخبار بين الاختبارات ومهما يكن الأمر فإن جمع

البيانات قد يتم إما بفحص كل وحدات المجتمع محل الدراسة أو بفحص جزئي (عينه) .

إن عملية جمع البيانات ليست عملية منفصلة عن وظائف الإحصاء الأخرى فهناك صلة وثيقة - فالهدف واحد وهو الحصول على معلومات أو نتائج - وذلك يكون باستخدام مقاييس وأساليب وصف البيانات - وذلك بعد جمعها - وإذا كانت هذه البيانات خاصة بعينة أى بجزء من المجتمع فإن وصف المجتمع يتطلب استخدام أساليب الاستقراء .. وهذه المقاييس والأساليب لها شروط ومتطلبات يجب مراعاتها وتوفيرها عند جمع البيانات وذلك باستخدام التصميم التجريبي المناسب أو تصميم استمارة استبيان مناسبة واختيار طريقة المعاينة المناسبة وحجم العينة المناسب ومراعاة توفير مستوى القياس المناسب للمتغيرات .. الخ كما أن البيانات التى يتم جمعها يجب أن تكون محل ثقة حتى تكون النتائج المستخلصة منها محل ثقة . أى يجب أن يتوافر فيها الصدق والثبات **Validity and reliability** أن تحديد ذلك واختياره يكون غالبا باستخدام الأساليب الإحصائية . (4)

(ب) الإحصاء الاستدلالي Inferential Statistics

يستند هذا القسم من الأساليب الإحصائية إلى مجموعة من النظريات الإحصائية لعل أهمها نظرية الاحتمالات ونظرية العينات اللتان تمثلان حلقة الوصل بين الإحصاء الوصفي والاستدلالي .

ويسعى هذا النوع من الأساليب الإحصائية إلى الوصول إلى تقديرات لمعالم وخصائص مجتمعات الدراسة من خلال ما هو متوفر من معلومات عن العينات المختارة . من تلك المجتمعات ، فضلا عن اختبار الفروض الإحصائية عن مجتمع البحث على أساس البيانات المتاحة عن عينات الدراسة . ويطلق على هذا النوع من الأساليب أكثر من تسمية تؤدي جميعها إلى نفس المعنى فأحيانا يسمى بالإحصاء الاستدلالي ، أو الاستنباطي Inductive أو التعميمي Generalizing حيث يهدف إلى الوصول إلى تعميمات عن مجمع الدراسة من خلال العينة المسحوبة من هذا المجتمع . ويشمل هذا النوع من الأساليب الإحصائية ، الاحتمالات ، العينات ، اختبار الفروض ، الاستدلال من خلال عينة واحدة أو أكثر وما يتضمنه ذلك من اختبارات مختلفة مثل χ^2 اختبار جاما γ ، فاي ϕ ... الخ (5)

ويقصد بوظيفة الاستدلال اشتقاق النتائج من دراسة وفحص المقدمات والبيانات المتوافرة عن ظاهرة معينة. ولهذا يطلق على عملية الإحصائية التي تستخدم والاستدلالي على أساس المنطق الاستدلالي المبني على نظرية الاحتمالات الرياضية فمن عينة محددة من أعمال أحد المصانع وباستخدام الأسلوب الإحصاء الاستدلالي يكون من الممكن التنبؤ بمعدلات الزيادة في الإنتاج ومقدار التغير في نسبة الغياب وفي هذه الحالة نجد أن

الدقة في التنبؤ تعتمد علي عوامل كثيرة من أهمها ملائمة الأدوات الإحصائية المستخدمة وحجم العينة محل الدراسة والإجراءات الإحصائية اتخذت عند اختيارها .(6)

وتعتبر وظيفة الاستدلال أو الاستقراء من الأهمية بمكان في البحث العلمي فمثلا :

إذا كانت الظاهرة موضوع الدراسة والتحليل ممثلة للمجتمع الذي تنتمي إليه فانه يمكن الحصول على نتائج معنوية عن المجتمع بتحليل بيانات هذه الظاهرة وهو ما يعرف بالاستدلال ويعتمد هذا الأسلوب في البحث على الشروط التي يجب توافرها حتى يكون هذا الاستدلال سليما - وبما أن الاستدلال لا يمكن أن يكون مؤكداً فان لغة الاحتمال تستخدم عند عرض النتائج . (7)

وتعتبر وظيفة الاستقراء لها أهمية كبيرة - فهي تمكن الباحث من الوصول إلى تعميمات عن المجتمع على أساس المعلومات المتاحة من عينة منه . وفي هذه الحالة فان أساليب ومقاييس الوصف يقتصر وصفها على ذلك الجزء (العينة) فقط من المجتمع - ومن هنا تأتي أهمية وظيفة الاستقراء - فهي تمكننا من وصف المجتمع (التعميم) باستخدام بيانات العينة .

إن القوانين في العلوم الطبيعية والاجتماعية تجد برهانها عند الوقائع والحقائق الإحصائية ولذا يعد الاستقراء الاحصائي (Statistical Inference) أساسا لتطور المعرفة العلمية

باعتباره البرهان لهذه القوانين . ووظيفة الاستقراء تحقق مطلبين أساسيين فى البحث : الأول تقدير خواص المجتمع والثانى اختبارات الفروض حول هذه الخواص . ولا تقتصر هذه الوظيفة على مجرد الاستقراء بل تقدم لنا تقييما عن مدى دقة هذا الاستقراء وأكثر من ذلك فهى تمكننا من التحكم فى مستوى الدقة وذلك بعدة طرق منها استخدام الأسلوب المناسب للمعاينة والحجم المناسب للعينة . وباختصار فان هذه الوظيفة للإحصاء تمدنا بالاستقراء المنطقى وتختلف الأساليب المتبعة فى الاستقراء حسب طبيعة محل الاستقراء (8)

ثانيا : البيانات Data :

من الشائع فى مجال البحوث الاجتماعية توافر مجموعة من البيانات الإحصائية التى يحصل عليها الباحث باستخدام أدوات جمع بيانات مناسبة وعادة تتمثل تلك البيانات فى شكل أرقام تعتبر قياسا للمتغيرات تحت الدراسة ولما كانت تلك الأرقام تفتقر إلى الترتيب والتصنيف يطلق عليها البيانات الأولية أو البيانات الخام .Raw Data

وتعرف البيانات الإحصائية أنها كمية من المعلومات على هيئة أرقام وان تلك الأرقام إما أن تكون صحيحة Integers مثل 10 ، 20 ، 30 وهكذا أو تكون أرقاما عشرية أو حقيقية Real Numbers مثل 8.5 ، 10.25 ، 1505 وهكذا : ويتوقف حجم البيانات الخام على حجم المجتمع الأصلي فكلما ازداد حجم هذا

المجتمع يتوقع مزيدا من الأرقام غير المرئية والتي يصعب مع كثرتها وعدم تصنيفها تفهم أو قياس متغير أو أكثر تحت الدراسة ومن ثم كان من الضروري أن يقوم الباحث بتصنيف وتبويب تلك البيانات بالشكل أو بالأسلوب الذى يخدم جيدا هدف الباحث من دراسة المتغيرات أو استنباط نوعية العلاقات أو المعلومات الهامة التى تتعلق بتلك المتغيرات . (9)

ويقصد بتعبير البيانات " أى كمية من المعلومات فى صورة رقمية والصورة الرقمية للبيانات تبدو إما على شكل أرقام صحيحة مثل 10 ، 112 ، 464 . أو على شكل أرقام حقيقية مثل 20.4 ، 61.8 ، 182.1 أى أنها الأرقام التى تحتوى على علامة عشرية . وتعتبر المعلومات الرقمية (البيانات) المادة الخام لأسلوب العمل الإحصائى كما أنها تلعب دورا كبيرا فى تطبيق الأساليب الإحصائية . (10)

وتسمى البيانات المتاحة - المنشورة أو التى تم جمعها - تسمى بيانات خام أو أولية - ذلك أنها تكون غير مجهزة فهى لا تفصح إلا عن القليل من المعلومات . كما أنه يستحيل استخلاص المعلومات منها . وفى سبيل ذلك نستعين بأساليب ومقاييس وصف البيانات . وهذه الأساليب كثيرة ومتنوعة فهى تختلف حسب عوامل أهمها عدد المتغيرات ومستوى قياسها (11)

ولعل أبسط الطرق الإحصائية لتنظيم وتلخيص البيانات طريقة التوزيع التكرارى **Frequency Distribution** ، أو بمعنى ضمنى من التوزيع التكرارى يمكن استخدام وسيلة أو أكثر من الوسائل الثلاث التالية والتي يمكن أن يتحول التوزيع إليها أو إلى أى منها .

(أ) استخدام الجداول الإحصائية **Statistical Tables** فى عملية تصنيف وتبويب البيانات الخام .

(ب) استخدام التمثيل البياني والخرائط فى عرض البيانات الإحصائية (تحويل التوزيع التكرارى إلى منحنيات تكرارية).

(ج) استخدام مقياس أو أكثر من المقاييس الإحصائية مثل المتوسط الخام **Mean** الانحراف المعياري **Standard Deviation** ومعامل الارتباط **Correlation Coefficient** فى تلخيص البيانات الإحصائية فى صورة رقم أو نسبة مئوية ونرى أهمية الوقوف على نوعية البيانات الإحصائية من منظور مستويات القياس الاحصائى نظرا لأهمية تلك البيانات الإحصائية وفقا لمستويات القياس الاحصائى يرجع إلى أن المتغيرات التى تقاس كميًا تنقسم من قيمتها العددية إلى المتغير المتصل والمتغير المتقطع . (12)

ثالثا : المتغيرات Variables :

تشير كلمة المتغيرات إلى الخصائص التي تشترك فيها أفراد المجتمع الإحصائي ولكنها تختلف من فرد إلى فرد آخر فالعمر ، درجة الذكاء ، وطول القامة ، واللياقة البدنية والقدرة على القراءة ، والدخول التي يحصل عليها الأفراد أمثلة للمتغيرات وتتميز هذه المتغيرات بأنها قابلة للقياس الكمي وبإمكانية تحديد قيمة معينة لها .

ويمكن القول بان المتغيرات مفهوم له معنى امبريقي ويعبر عنه بقيم مختلفة وتعبير النوع ، سنوات التعليم والعمر ، والدخل السنوي من المتغيرات الشائعة التي تستخدم فى البحوث الاجتماعية لارتباطها بالخصائص الأساسية للمبحوثين ، ولأهميتها فى تحديد مكانتهم الاجتماعية والاقتصادية وانتماءاتهم الطبقية (13) .

والمتغيرات عبارة عن ظاهرات أو صفات تختلف قيمها باختلاف الحالات . ومن أمثلتها : درجة الحرارة فى مناطق مختلفة أو فى فترات مختلفة لمكان واحد ، كميات الإنتاج الزراعى أو الصناعى . (14)

ويمكن القول بان المتغير هو أى ظاهرة أو حدث أو خاصية تأخذ فيها قيمة تتغير من ظرف لآخر . والمتغير هو الوحدة الأساسية للتحليل الإحصائي ويمكن تعريفه بأنه مجموعة من العناصر أو التقسيمات غير المتداخلة . وهذه المجموعة من

التقسيمات تكون مقياس Scale . وتنقسم المتغيرات إلى مستمرة وغير مستمرة (متقطعة) . المتغير المستمر هو ذلك الذى يأخذ قيما لأى درجة من الدقة - مثل الطول - الوزن - درجة الحرارة أما المتغير غير المستمر فهو الذى يأخذ قيما معينة فقط - مثل عدد الأولاد فى الأسرة عدد الطلاب فى الفصل . وهناك تقسيم آخر للمتغيرات ، حيث تنقسم إلى متغيرات مستقلة ومتغيرات تابعة . فعندما نبحث فى الأثر الذى يحدثه متغير (س) فى آخر (ص) كأثر التدريب على الإنتاجية نقول أن (س) متغير مستقل و (ص) متغير تابع (15)

وتنقسم المتغيرات من قيمها العددية إلى قسمين هما المتغيرات المتصلة Continuous Variables وهى المتغيرات التى يمكن أن تأخذ أى قيمة على المقياس المستخدم فمثلا إذا ارتفعت درجة الحرارة من 20° درجة مئوية إلى 30° درجة مئوية خلال الترمومتر الزئبقي فمعنى ذلك أن الزئبق يكون قد مر بكل القيم الواقعة بين هاتين الدرجتين ، كذلك الحال فى مقياس سرعة السيارة . فإذا زادت السرعة من 30 كيلوا متر / ساعة إلى 60 كيلوا متر / ساعة فإن المؤشر فى المقياس يكون قد مر على كل القيم المحصورة بين هذين الرقمين وبالمثل أيضا الأطوال . وذلك لان طول الشخص قد يكون 168 سم أو 168.1 أو أى قيمة مهما كانت كسرية ، واصغر من المليمتر إذا كان المقياس يسمح بذلك .

والنوع الآخر من المتغيرات يطلق عليه المتغيرات الغير متصلة أو الوثابة **Discrete Variables** وهى التى تختلف قيمها من مرحلة إلى أخرى بدون أن تكون منتظمة كما أن قيمها لا تأخذ إلا أعداد صحيحة **Integers** فعدد الرحلات التى يقوم بها الأشخاص وكمية مياه الفيضان فى الأودية الصحراوية وعدد السيارات المارة فى احد الشوارع وعدد الفصول بالمدارس وعدد الحجرات بالمنازل وحجم الأسرة 000 الخ كلها متغيرات وثابة (غير متصلة) يحصل عليها فى الغالب بالعد⁽¹⁶⁾

والمتغيرات التى تقاس كميًا تنقسم من حيث قيمتها العددية إلى نوعين هامين لا ثالث لهما :

1 - المتغير المتصل Continuous Variable .

لما كان التعريف العام للمتغير **Variable** هو ظاهرة أو صفات تختلف قيمها باختلاف الحالات فإن المتغير يكون متصلًا عندما يأخذ أى قيمة متدرجة على المقياس المستخدم . مثال ذلك قياس درجات الحرارة باستخدام الترمومتر فالمتغير يأخذ أى قيمة بين رقمين صحيحين , بمعنى أن المتغير يمكن أن يأخذ أى قيمة بين 36 درجة ، 37 درجة (36.1 ، 36.2 000 الخ) .

2 - المتغير المتقطع Discrete Variable

عندما يأخذ المتغير قيما محددة يطلق عليه متغيرًا متقطعًا أو بمعنى آخر ، المتغير المتقطع هو الذى يحتوى مداه على عدد

محدود من القيم أو يحتوى عدد لانهاى من القيم ولكن لكل منها قيمة محددة يمكن عدّها أو ترتيبها فى نهاية الأمر تعدد الأولاد أو الأفراد فى الأسرة لابد أن يكون أعدادا صحيحة غير حقيقة مثل 1 ، 2 ، 3 ، 4 ، 00 وهكذا ومن أمثال المتغيرات المتقطعة ، النوع ، الحالة الزوجية **Martial Status** ، عدد أيام الإنتاج فى احد المصانع ، عدد حوادث السيارات وهكذا . (17)

كما يمكن تصنيف المتغيرات إلى عدد من التصنيفات بحسب الغاية من كل تصنيف

وذلك على النحو التالى : -

1 - المتغيرات الكمية والمتغيرات الكيفية :

يمكن تصنيف المتغيرات من حيث طريقة التعبير عنها إلى فئتين هما : المتغيرات الكمية **Quantitative Variables** وهى التى يمكن أن نصفها عددياً بأنها اكبر من أو أقل من قيمة معينة ويعتبر العمر وعدد سنوات التعليم أمثلة لهذه المتغيرات . والفئة الثانية من المتغيرات هى المتغيرات الكيفية **Qualitative Variables** وهى التى تصف الأشياء بصفاتهما مثل متغير النوع الذى ينقسم إلى قسمين : ذكور وإناث . والحالة العملية للفرد حيث تكون إما مزارع أو عامل غير ماهر ، أو عامل ماهر أو موظف أو تاجر وما إلى ذلك من صفات ، وهذه المتغيرات الكيفية يتعذر معالجتها إحصائياً ما لم يميزها عن بعضها بعضاً باستخدام

الأرقام فنرمز لمتغير الإناث برقم 1 و لمتغير الذكور برقم 2 أو العكس ، والرقم فى هذه الحالة لا يعنى أكثر من أنه أداة للتمييز بين المتغيرات الكيفية لتسهيل تفريغ البيانات التى جمعت عنها من ميدان الدراسة تمهيداً لمعالجتها إحصائياً ولا تكون لها قيمة عددية فى حد ذاته .

2- المتغيرات التابعة والمستقلة والضابطة :

ويمكن تصنيف المتغيرات تصنيفاً آخر بحسب دورها فى حدوث الظاهرة محل الدراسة وذلك إلى :

(أ) متغيرات تابعة Dependent Variables

وهى تلك المتغيرات التى نحاول تفسيرها ومعرفة أسباب حدوثها وتحديد مدى إمكان التنبؤ بها .

(ب) متغيرات مستقلة Independent Variables

وهى التى لعبت دوراً مباشراً فى حدوث المتغيرات التابعة ونستخدمها فى تأييد تفسيرنا وفهمنا لما طرأ على هذه المتغيرات من تغيير ، وفى التنبؤ بالحالة التى ستؤول إليها بعد ذلك .

(ج) متغيرات وسيطة Intermediate Variables

وهى تلك المتغيرات التى يمر من خلالها تأثير المتغيرات المستقلة إلى المتغيرات التابعة والمتغيرات الوسيطة بالغة الأهمية فى تفسير حدوث الظواهر الاجتماعية إذ قد يغفل عنها الباحثون أو قد

ينظرون إليها على أنها متغيرات مستقلة لارتباطها المباشر بالمتغيرات التابعة فإذا نظرنا إلى تفسير ظاهرة الانتحار اللامعيارى التى درسها دوركايم ، على سبيل المثال سنجد أن بعض الأفراد ينظرون إلى حالة فقدان المعايير التى تؤدى إلى الانتحار على أنها المتغير المستقل والانتحار هو المتغير التابع ولكن فريقاً آخر من الباحثين الذين ينظرون إلى الظاهرة بطريقة أكثر تفصيلاً ، ويرون أن المجتمع يمر بتغيرات اقتصادية و اجتماعية عاصفة وقوية وهى التى تمثل المتغير المستقل وتكون النتيجة المترتبة على تلك التغيرات انهيار الثقة فى القيم الراسخة والمبجلة لدى الأفراد فتنتشر حالة اللامعيارية وهى تمثل هنا المتغير الوسيط ثم ينتهى الأمر بالانتحار الذى يمثل المتغير التابع . وإذا قارنا بين الطريقتين السابقتين فى تفسير ظاهرة الانتحار نجد أن حالة اللامعيارية كانت متغيراً مستقلاً فى التفسير الأول ثم اعتبرت متغيراً وسيطاً ضابطاً فى التفسير الثانى .

3- المتغيرات غير المستمرة (الوثابة) ، والمستمرة (المتصلة)

Discrete and continuous variables

ذكرنا أن مهمة الباحث هى جمع البيانات عن متغيرات معينة مثل متغير النوع بأن يعرف كم عدد المبحوثين من الذكور وكم عددهم من الإناث ، وعن متغير سعة الوحدة السكنية بأن يحدد عدد الغرف التى يسكن بها كل مبحوث .

وبالنظر إلى المتغيرات السابقة نجد أنها تضم عدداً من المتغيرات غير المستمرة والتي يمكن التعبير عنها بقيم عددية غير قابلة للتجزئة حيث يرمز الباحث للذكور برقم (1) وللإناث برقم (2) ، ولا توجد قيمة وسط بينهما وكذلك الحال بالنسبة لسعة الوحدة السكنية ، فالشقة إما أن تكون غرفة واحدة أو غرفتين أو ثلاث أو أكثر وليس هناك جزء من غرفة . والبيانات التي يتم جمعها عن المتغيرات غير المستمرة تكون بيانات غير مستمرة أيضاً أي أنها غير قابلة للتجزئة ولا نجد لها كسور . فلا يستطيع الباحث أن يدعى أن العينة تتكون من عشرة ذكور ونصف أو أن الشقة تتكون من ثلاث غرف وربع . ويطلق على البيانات الكمية التي يتم جمعها عن المتغيرات غير المستمرة القيم المفردة حيث لا يمكن تبويبها أو تقسيمها إلى فئات متصلة.

وقد يهتم الباحث أيضاً بجمع بيانات عن دخل كل مبحوث في فترة معينة . والدخل يعد من المتغيرات المستمرة التي يمكن أن تأخذ أي قيمة ما بين نقطتين ثابتتين على مقياس معين . وإلى جانب الدخل هناك متغيرات أخرى مثل العمر والطول والوزن تعد أيضاً من المتغيرات المستمرة ، إذ يمكن تقسيم متغير كالدخل إلى أي عدد نشأ من الفئات وكذلك متغير العمر فيمكن القول أن هناك شخصاً يحصل على دخل أسبوعي قدره خمسون جنيهاً وآخر يحصل على تسعة وأربعون جنيهاً ونصف ... وهكذا والبيانات

التي يتم جمعها عن المتغيرات المستمرة تكون بيانات مستمرة أيضاً أى أنها قابلة للتجزئة وبها كسور أو قيم غير صحيحة .

ولذلك فإن هذا النوع من البيانات الكمية يكون ضخماً للغاية عندما يجمعه الباحث من ميدان البحث . فإذا سأل مائة فرد عن دخلهم الأسبوعي فإنه من المتوقع أن يحصل على مائة إجابة تمثل مائة قيمة مختلفة عن بعضها البعض . ولذلك عادة ما يتم تفريغ هذه البيانات فى صورة فئات لكل منها طول معين بحيث تحتوى كل فئة على عدد من القيم المتقاربة لتسهيل عرض البيانات ومعالجتها إحصائياً ، وهذا النوع من البيانات نطلق عليه البيانات أو القيم المبوبة .

والواقع أن التمييز بين المتغيرات غير المستمرة والمستمرة رغم أهميته إلا أنه فى بعض الأحيان نظراً لعدم وجود أداة قياس مضبوطة نجد أن متغيرات كثيرة مستمرة يكون من الضرورى تحديد قيم عددية إجمالية لها ، ومن ذلك مثلاً مقياس الذكاء فهو من الناحية النظرية يعد متغيراً مستمراً ولكن من الناحية العملية نجد أن الاختبارات التى تستخدم فى قياسه تعطى نتيجة إجمالية وقيمة غير مستمرة (18) .

رابعاً : المقاييس الإحصائية

يقصد بالقياس - كمفهوم واسع - انه عملية تعبير عن الخصائص والملاحظات بشكل كمى ووفقاً لقاعدة محدودة .

وعندما نستخدم المقياس والملاحظات بشكل كمي ووفقا لقاعدة محددة . أو بمفهومه وفق الأبعاد الخاصة الملائمة لكل فرع من فروع المعرفة ، فإننا لا نجد غضاضة فى اختيار نسق من المعادلات الرياضية التى تتفق مع تلك الخاصية أو الخصائص قيد البحث - وعامة يمكن القول أن ما تحظى به فروع العلم المختلفة من رياضيات واقتصاد وغيرها من فروع العلوم الاجتماعية من نماذج متعددة ومتباينة تعتمد فى بنيتها الأساسية على المقاييس .

وإن كان هناك اختلاف كبير فى درجة الصعوبة عند التطبيق إذا قورنت النماذج المستخدمة فى العلوم الاجتماعية بغيرها من فروع العلوم الأخرى ففى علم الاجتماع وعلم النفس الاجتماعي كمثال تتصف المتغيرات بالتباين والتعدد بشكل يصعب معه أن نختار رياضيا مناسبا يخدم أهداف البحث الامبريقي لان النفس البشرية (والفرد عامة) - يتصف بالتعقيد واختلاف مستويات العلاقة بينة وبين المحيطين به من أفراد أو بيئات

ولعل ابسط أمثلة القياس نجدها فى الاختبارات التى يتقدم بها الطالب فى مختلف مراحل حياته الدراسية . حيث ترتبط الدرجة التى يحصل عليها فى اختبار على مدى معرفته بالمادة التى يدرسها خلال فترة دراسية معينة وكلما كانت درجة الطالب التى حصل عليها مثلا فى مادة الكيمياء عالية دل ذلك على معرفة أكثر أو تحصيل اكبر لدى الطالب من هذه المادة . ومن هذا المثال

البسيط نجد أن خاصية التحصيل تعبر عنها الدرجة Score التي حصل عليها الطالب من الاختبار .

وتعتبر المقاييس التي تقيس المتغير التابع Dependent Variable واحدة من أكثر المقاييس أهمية عند إيجاد الطرق الإحصائية الملائمة التي تستخدم في تحليل بيانات دراسة أمبريقية معينة . أيضا توجد بعض المقاييس التي يمكن استخدامها في قياس ظاهرة معينة بدقة عالية أو متناهية مثال ذلك المقاييس التي تستخدم في قياس الأطوال والأوزان من جهة أخرى توجد بعض المقاييس التي تفتقر إلى الدقة العالية وإن كانت تحقق قدرا من الدالة فيها علي سبيل المثال مقاييس مستويات القلق النفسي عند الأفراد⁽¹⁹⁾ ويعتمد القياس في التحليل الإحصائي علي القيم العددية التي تستخدم بطرق مختلفة لتحقيق عدة أهداف :-

- أ- تستخدم القيم العددية لترقيم المتغيرات (إجابات الأسئلة) التي يختار من بينها المبحوث في الاستبيان المكتوب.
- ب- وتستخدم القيم العددية في ترتيب مجموعة من المتغيرات فيكون المتغير رقم (1) أعلى من المتغير رقم (2) عندما يكون الترتيب تنازلي للقيم ويكون المتغير رقم (1) أدني من المتغير رقم (2) عندما يكون الترتيب تصاعدي للقيم بعبارة أخرى , تفاوت أهمية القيم بحسب ما إذا كان الترتيب تصاعديا أو تنازليا .

ج- تستخدم القيم العددية أيضا في تحديد المسافة بين الفئات المختلفة من المتغيرات لذلك يجب علي الباحث أن يفهم الكيفية التي تستخدم بها الإعداد في وضع المقاييس الإحصائية⁽²⁰⁾.

ولغرض استخدام المقاييس والأساليب الإحصائية فإنه يجب تحديد مستوى القياس للبيانات أو المتغيرات ولذلك يتم تقسيم مستويات القياس إلى أربعة أنواع هي مستوى القياس الاسمي والترتيبي والفتري والنسبي وهذه المقاييس تختلف من حيث كمية المعلومات التي تحتويها وبالتالي تختلف العمليات الحسابية والإحصائية التي يمكن إجراؤها⁽²¹⁾.

1- المقاييس الاسمية والوصفية *nominal measures* هذا النوع من المقاييس يستخدم المتغيرات التي تستخدم في تصنيف مفردات عينة البحث وذلك بإعطائها قيما عددية والقيمة العددية في هذه الحالة ليس لها دلالة سوي تعريف المتغيرات وتمييزها ويستعين بعض الباحثين بالرموز بدلا من الأرقام في عملية استخدام المتغيرات في تصنيف بعض مفردات عينة البحث ولكن استخدام الرمز لن يفيد كثيرا في حالة تفريغ البيانات بواسطة الحاسب الآلي ومن أمثلة المتغيرات التي تشكل منها المقاييس الوصفية التي تستخدم في تصنيف المبحوثين متغير النوع إذا يعطي الباحث رقم (1) للإناث ورقم (2) للذكور أو يصف

المبحوثين حسب متغير الدين إلي (1) مسلم (2) مسيحي (3) يهودي - والأرقام هنا لا تعني أولوية أو أفضلية متغير علي آخر كما أنها لا تحتل أي قيمة. والواقع أن أرقام السيارات وأرقام المنازل هي أبرز مثال لاستخدام القيم العددية في تصنيف الأشياء فالمنزل رقم (1) ليس يعني أنه أفضل من المنزل (100) أو العكس وإنما الرقم يكون استخدامه بغرض التعرف علي المنزل وتميزه عن المنازل الأخرى (22) ويعد أقل مستوي للقياس , وهو مجرد تقسيم أو تصنيف الأشياء بالاسم فقط ودون تداخل مثال ذلك تقسيم الأشخاص حسب الجنس (ذكور - إناث) وحسب الجنسية (مصري - سعودي - عراقي.....) وتقسيم الجرائم إلي (قتل - خطف - سرقة) وتقسيم الكتب والمراجع بالمكتبة حسب الموضوع (المعارف العامة - الفلسفة - الديانات - العلوم الاجتماعية) وتشمل قياسات خصائص الظاهرة موضوع الدراسة في هذا النوع علي قياسات (23) ثنائية أو ثلاثية ولنضرب مثالا علي ذلك فعند تسجيل حالة التعليم لدي الأشخاص : تعليم متوسط أن تعليم عالي يعطي الشخص من النوع الثاني الرقم (2) وإذا كانت الحالة التعليمية يعطي الرقم (صفر) , وإذا كانت الدراسة تتعلق بانتماء الأشخاص إلي مناطق ريفية أو حضرية فإننا في هذه الحالة نعطي للشخص الريفي الرقم (1) وللشخص الحضري الرقم (2) ويطلق علي المتغيرات التي تقاس بها البيانات الاسمية المتغيرات دمي dummy variables كما أنها في أحيان أخرى

تسمى بالبيانات التصنيفية لأنها تصنف المتغيرات علي أساس خصائصها (24)

ويعتبر التصنيف أبسط العمليات الأساسية في أي فرع من فروع العلم فالتصنيف هو تجميع للمفردات أو العناصر أو المعلومات المتشابهة إلي حد كبير المتماثلة في خصائصها مع بعضها في مجموعة أو مصنف category وذلك بهدف المقارنة بين المجموعات المختلفة علي أساس الخواص مثال ذلك إذا قمنا بتصنيف عدد من الأفراد إلي مجموعات وفق خاصية العقيدة religion (مسلم - مسيحي - يهودي) وقد تقوم أيضا بعمل تصنيف آخر للنزعات السياسية للفئات الدينية الثلاث وهكذا ولا بد من استخدام التصنيف كعملية أساسية تعتمد عليها المقاييس الأعلى كأساس لها أيضا في العلوم الاجتماعية من ذلك لا نبالغ بالقول إن التصنيف يعتبر المستوي الأول في القياس وفي المثال السابق نجد أننا لم نهتم بالتمييز بين الفئات الدينية الثلاث علي أساس الأهمية مثلا فلم نقل أن المسلم أهم من المسيحي أو أن المسيحي أهم من اليهودي فقط ينصب المقياس علي تصنيف وفق الديانة وتمثل الخاصية الأولى للمقياس التصنيفي والتي يمكن أن نحددها في عدم اتصاف المقياس بالترتيب المنطقي من ذلك نلاحظ عدم وجود أي تدخل علي أساس الديانة فالمجموعة كاملة تضم أفراد متماثلين في نوع الديانة ومن ثم لا تتكرر الظاهرة أو المفردة في أكثر من مجموعة وهذه ميزة ثانية وهامة يتصف بها

المقياس التصنيفي والخاصية الثالثة التي تتصف بها المقاييس التصنيفية نجدها في مجال العلاقات بين المفردات أو المقادير في العلوم الرياضية علي سبيل المثال يتصف المقياس بخاصية الانتقالية **transitivity** ويقصد بها أنه إذا كانت هناك علاقة معينة بين متغيرين من أ،ب بحيث أنها تتحقق من (أ) (ب) فإن من الضروري أن تتحقق أيضا من المتغير (ب) نحو المتغير (أ)⁽²⁵⁾.

2- المقاييس الترتيبية **ordinal measures** وهذه المقاييس لا تستخدم فقط لتصنيف المتغيرات وإنما لتعكس أيضا ترتيب تلك المتغيرات بعبارة أخرى يستخدم هذا المقياس في ترتيب الأفراد أو الأشياء من الأعلى أو العكس وذلك وفقا لخصائص معينة يتميز بها المراد ترتيبه فالمكانة الاجتماعية - الاقتصادية والتي تقاس بمتغيرات الدخل والمهنة والتعليم يتم ترتيبها حسب فئات معينة تبدأ تنازليا من الطبقة العليا الطبقة عليا الوسطي - الطبقة الوسطي الطبقة وسطي الدنيا - والطبقة الدنيا - ما دون الطبقة **under class** فإذا أعطينا أرقاما لهذا الترتيب الطبقي فإن رقم (1) يكون له معنى يفيد الرقمي إذا ما قورن برقم (4) وهكذا ويستخدم هذا المقياس أيضا في وصف المتصلات **continuums** مثل المتصل الريفى - الحضري الذي يكون بدايته رقم 1- الريف 2- الأطراف الحضرية 3- الحضر 4- الضواحي فرقم (1) هنا

يشير إلى بداية المتصل ورقم (2) يشير إلى مرحلة أخرى منه وهكذا الحال بالنسبة لباقي المتصل (26) .

وهذا القياس أعلى مستوي من المقياس الاسمي حيث يتم التقسيم علي أساس الرتبة أو الأهمية النسبية مثال ذلك درجات الطلاب علي أساس ممتاز - جيد جدا - جيد - مقبول - ضعيف أو توزيع السكان حسب الحالة التعليمية : أمي - ابتدائي - ثانوي - جامعي - ماجستير - دكتوراه وفي هذا القياس يمكن ترتيب القيم وإجراء المقارنات حيث يمكن القول أن الحاصل علي تقدير جيد مستوي تحصيله أفضل من الحاصل علي تقدير مقبول مثل هذا الترتيب والمقارنة لا نستطيع القيام بها في المقياس الاسمي حيث أن هذا المقياس لا يمكنه تحديد مقدار الفروق بين القيم⁽²⁷⁾ وتعرف القياسات الترتيبية بالبيانات المرتبة في فئات أو حسب خصائصها عن طرق إعطاء القيم الأصلية للمتغيرات رتبا أو أرقام تدرجية أو تنازلية (28) .

وفضلا عن تصنيف الأفراد إلى ثلاث مذاهب دينية يمكن أن ترتب تلك المجموعات الثلاثة وفقا لأهميتها أو لما تمتلكه كل منها من خاصية أو سمات معينة مشتركة وغير مشتركة وقد نجد مثلا أقرب للفهم في الرياضيات عندما نميز بين المقدارين (أ) ، (ب) فنقول أن (أ) < (ب) ونأخذ الشكل الرياضي التالي أ < ب وقد يكون أ > ب ولكن مقدار الفرق في القيمة الدالة علي التمييز

بين أ , ب ليس من خصائص المقياس الترتيبي ومن ثم فإن المقياس الترتيبي هو مستوي أعلى من المقياس التصنيفي في قياس الظواهر أو الخواص وتعتبر خاصية التمييز باستخدام علامات (<) أو (>) الخاصية الثانية إذا أخذنا في الاعتبار الخاصية التصنيف وفق الترتيب وفي العلوم الاجتماعية نجد مثالا لخاصية الترتيب دون الالتزام بالفروق عندما نصنف الأسر وفقا للمكانة الاجتماعية الاقتصادية socio economic status طبقة عليا , متوسط عليا upper middle , متوسط دنيا lower middle وأيضا إلى طبقة دنيا lower class وحقيقة الترتيب هنا هما الرتبة العليا والرتبة الدنيا فقط والخاصية الثالثة لو تخيلنا ترتيبا للأفراد علي متصل continue شريطة ألا يحتل فردان منهما مكانا واحدا أو يتواجدان في نقطة واحدة علي هذا المتصل وذلك مع فرض وجود علاقة أو روابط بين هؤلاء الأفراد علي المتصل ومن ثم يتم جميعهم عشوائيا دون دراية كافية في مجموعة وتكرار ذلك وفق ترتيب لخاصية معينة بحيث يمكن لنا فقط أن نقول أن المجموعة كذا من الأفراد تمثل أعلى التكرارات قياسا بباقي المجموعات أو نقول أن المجموعة كذا تمثل أعلى النقاط نسبيا هذا ويجدد الإشارة أن جميع المفردات دون تكرار ظهور المفردة في أكثر من مجموعة تمثل خاصية يتشابه فيها المقياس الترتيبي مع المقياس التصنيفي والخاصية الرابعة فهي الانتقالية فلو فرضنا قريبا أن أ < ب وأن ب < ج وهذه خاصية

أخري يتشابه فيها هذا المقياس مع المقياس التصنيفي ولكن من المنظور الترتيبي ويجب التنويه إلي ضرورة ملاحظة أن المستوي الترتيبي للمقياس لا يهتم بالفروق - كما قلنا - بين العناصر أو الخواص ومن ثم لا نستطيع أن نستخدم مع هذا المقياس التصنيفي وتوضيح ذلك فالعمليات الحسابية كالتحريك والقسم والضرب والجمع لا يمكن استخدامها أيضا مع المقياس التصنيفي وبافتراضنا أن هناك أربع نقاط متصلة ويرمز لها بالأحرف (أ،ب،ج،د) وبفارق مسافات معينة تقع النقطتان ب،ج بين النقطتين (أ) ،(د) في الشكل التالي متصل

أ ————— ب ————— ج ————— د

فباستخدام المقياس الترتيبي يمكن كتابة العلاقة التالية (اتجاهيا) .

أد = أب + ب ج + ج د ولكن لا يمكن إطلاقا معرفة أطوال المسافات الأربعة المبينة في العلاقة السابقة مثال ذلك الترتيب المستخدم في مقاييس الاتجاهات الذي يبدأ بالموافقة بشدة وينتهي بعدم الموافقة بالمرّة (29)

3- مقاييس الفئات Interval measures

يشير مقياس الفئات إلي تبويب البيانات وتقسيمها إلي رتب معينة تبدأ من أدنى الفئات إلي أعلى الفئات ، وبالإضافة إلي ذلك فهو

يحدد المسافة بين تلك الرتب وتستخدم مقاييس الفئات في تلخيص القيم المتقاربة لتكون فئة واحدة ، ويعتبر الدخل ، والتعليم ودرجات الحرارة والعمر أمثلة علي المتغيرات التي تستخدم في تبويب بياناتها مقاييس الفئات وتتميز الفئات بإمكانية إجراء عمليات الجمع والطرح عليها بمعنى أنه يمكن أن تضيف فئة أخرى كنوع ومدى الفئة أو نقسم الفئة إلي جزأين ليكون كل قسم منها فئة صغيرة علي سبيل المثال ، الفئة العمرية من 16-18 سنة يمكن أن تجمع علي فئة العمر 18-20 سنة وتصبح فئة واحدة هي 16-20 فضلا عن ذلك فإنه يمكن معالجة الفئات معالجات إحصائية متعددة (30)

4- مقاييس الفترة الزمنية والنسبة

Interval and Ratio scale

المقياس الفترى Interval scale وهذا المقياس يعد أقوى من السابق حيث هنا يمكن تحديد الفروق بين القيم مثال ذلك درجات الحرارة المئوية (فهرنهايت) ودرجات الاختبار الرقمية: 40، 80، 65 ، وكذلك عدد ساعات الوقت الإضافي للعمال باعتبارها مقياسا لمستوى التوظيف ويؤخذ علي هذا القياس عدم وجود نقطة الصفر المطلق بمعنى أن الصفر هنا لا يقيس حالة الانعدام الخاصة وبالتالي لا نستطيع إجراء النسبة بين القيم وأن الطالب الحاصل علي (10) درجات مستواه في التحصيل يساوي

خمسة أضعاف آخر حاصل علي (2) درجة ⁽³¹⁾ وتعتبر بيانات الفترة أكثر أنواع البيانات الإحصائية شيوعا واستخداما في أبحاث العلوم الاجتماعية وهي تعكس القيم الأصلية للظواهر كأعمار السكان ، وكميات الإنتاج الزراعي والصناعي ، أعداد السيارات ، مساحات المزارع ومساحات البيئات الحضرية درجات الحرارة ، وكميات الأمطار ⁽³²⁾

- المقياس النسبي **Ratio** . ويعد أقوى مستويات القياس بما يسمح بإجراء النسب بين قيم المتغيرات مثال ذلك الأوزان والأطوال ودرجات الحرارة والسرعة . ⁽³³⁾

- وعلى خلاف ما ذهب إليه بعض الكتابات في الفصل بين مقياس النسبة . من أمثال هنكل **Hinkle** وآخرين ، فإننا نتفق مع ما ذهب إليه بلالوك **Blalock** من عدم الفصل بين نوعي المقياس حيث يعزل ذلك تعليلا منطقيا حين يرى أنه من الصعوبة بمكان أن نجد مقياسا للفترة لا يكون في نفس الوقت مقياس نسبة لان الواقع الامبريقي يشير إلى ضرورة وجود الوحدات القياسية أو المعيارية للمقياس فلا يعقل أن نجد مادة بلا طول أو كتلة أو نجد درجة حرارة بلا وحدة قياس للحرارة وهي إما درجة مئوية يطلق عليها **Centigrade** م° أو درجة فهرنهايت **F⁵ Fahrenheit** وتستخدم تلك المقاييس في حالات تتطلب قياس الفروق أو المسافات الحقيقية بين قيم معينة وهذه خاصية

تجعل مقياس الفترة والنسبة أرقى فى المستوى المقياسى من المقاييس السابقة لى تؤدى تلك المقاييس وظيفتها . فلو كان المطلوب قياس الفروق والمسافات يستخدم مقياس الفترة (الفنوي) (34)

ويتميز مقياس النسب أو المعدلات Ratio بكل الخصائص التى يتصف بها مقياس الفئات من قدره على وضع البيانات فى ترتيب معين فضلا على ذلك فهو يشتمل على الصفر المطلق ، وهذه الخاصية تجعل من الممكن استخدامها فى إجراء كل العمليات الحسابية من جمع وطرح وضرب وقسمة بسهولة تامة . وعلى سبيل المثال ، يمكن القول بسهولة ويسر أن الـ 1000 جرام تزيد على 600 جرام بمقدار 400جرام وأنها ضعف الـ 500 جرام فهذه الأرقام الصفرية لا تحتاج منها إلى استخدام آلات قاسية حسابية لتحديد العلاقة فيما بينها . كما انه من الممكن استخدام هذا المقياس فى حساب النسبة المئوية الخاصة بكل قيمة من القيم الواقعة عليه والواقع أن مقاييس المعدلات قليلا ما تستخدم فى مجال العلوم الاجتماعية ولكنها تستخدم فى ميدان العلوم الطبيعية فى قياس الأوزان والأطوال والوقت .

ولكى نوضح هذه النقطة نقول أن متغيرات كثيرة تستخدم فى مجال العلوم الاجتماعية مثل النوع والعمر والحالة التعليمية لا تتضمن بالضرورة صفرا فى قياسها بينما متغيرات قياس الأوزان

والأطوال تتضمن ذلك الصفر فالكيلو 1000 جرام والمتر 100 سم وهكذا . وفى مجال المعالجات الإحصائية للبحوث الاجتماعية غالبا ما نميل إلى استخدام الفئات الصفرية مثل 10 - 20 ، 20 - 30 لكى نيسر العمليات الحسابية بدلا من استخدام الفئات غير الصفرية مثل 3 - 6 ، 6 - 9 وهكذا (35)

ومن خصائص مقاييس الفترة والنسبة بالإضافة للخصائص التى ذكرناها فى المقاييس السابقين ، توحد نوع وحدة القياس فلا يمكن أن نقيس الفرق بين درجتين من الحرارة إحداهما بالفهرنهايت والأخرى بالدرجة المئوية بل يكون الفرق بين درجتين حراريتين مثل 38 درجة مئوية ، 30 درجة مئوية أى من نفس جنس وحدة القياس . ومن جهة أخرى ، إذا قلنا أنه توجد وحدات قياسية لمقاييس الفترة ، فى العلوم الاجتماعية قد يتعذر تحقيق ذلك ، فمثلا توجد وحدات قياسية أو معيارية لقياس الذكاء ، السلطة ، الهيئة الاجتماعية والتى نجدها متكررة دائما فى الموضوعات الاجتماعية والتنفسية المختلفة الفترة والخاصية الثانية لمقاييس الفترات والنسبة إمكانية استخدام العمليات الحسابية المختلفة من جمع وطرح وضرب وقسمة للدرجات فى عمليات تحليل البيانات فمثلا يمكن إضافة دخل الزوجة إلى الزوج أو إلى دخل باقى أفراد الأسرة . والخاصية الثالثة لمقاييس الفترة إذ يهتم بخاصية تساوى الفروق بين المستويات المختلفة مثال ذلك تقسيم الدرجة الواحدة على مقياس الحرارة (الترمومتر) إلى

تدرّج مقسمة إلى خمسة أقسام يمثل كل جزء منها (2). ومن الدرجة مثلا . ويطلق على هذا النوع من مقاييس الفترة مقياس الفترات المتساوية **Equal intervals Scale**.

ولكى يتم تدرّج فترات متساوية كما قلنا فى مثال مقياس الحرارة يلزم نحدد موضع نقطة مطلقة أو ما نسميه بالاختيار التعسفي لنقطة على المقياس ينسب إليها ترتيب تدرّج القيم تصاعديا وبفروق ثابتة على أساس وحدة القياس النوعية المستخدمة . ويطلق على تلك النقطة نقطة الصفر ومن ثم يطلق على المقياس فى هذه الحالة مقياس النسبة **Ratio Scale** حيث يمكن باستخدام النسب تدرّج القيم والقول بان القيمة كذا اكبر مرتين أو ثلاث مرات عن القيمة الأخرى المعلومة . (36)

ويتبين لنا أنه كلما زاد مستوى القياس للمتغيرات ، أى زادت الدقة فى القياس كلما أمكن استخدام مقاييس وأساليب إحصائية على درجة أفضل ، والثانية هى أن المتغيرات بمستوى قياس معين يكون التعامل معها بالأساليب الإحصائية المخصصة لهذا المستوى من القياس ، كما أنه يمكن أيضا استخدام الأساليب الإحصائية المخصصة لمستويات القياس الأقل . (37)

المراجع

- 1- اعتماد علام ، يسرى رسلان ، أساسيات الإحصاء الإجماعى ، دار الثقافة للنشر والتوزيع ، ص 10 .
- 2- حسن محمد حسن ، أساليب الإحصاء وتطبيقاته ، دار المعرفة الجامعية ، 1992 ، ص ص 19 - 20 .
- 3- فتحى عبد العزيز أبو راضى ، مبادئ الإحصاء الإجماعى ، دار المعرفة الجامعية ، ص ص 2 - 3 .
- 4- مصطفى زايد ، الإحصاء ووصف البيانات ، 1989 ، ص ص 26 - 27 .
- 5- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 10 .
- 6- حسن محمد حسن ، مرجع سابق ، ص 36 .
- 7- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص 3 .
- 8- مصطفى زايد ، مرجع سابق ، ص 32 .
- 9- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 34 .
- 10- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص 6 .
- 11- مصطفى زايد ، مرجع سابق ، ص 30 .
- 12- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص ص 34 - 35 .
- 13- حسن محمد حسن ، مبادئ الإحصاء الاجتماعى ، دار المعرفة الجامعية ، ص ص 36 - 37 .
- 14- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص 7 .

- 15- مصطفى زايد ، مرجع سابق ، ص ص 23 - 24
- 16- فتحي عبد العزيز أبو راضى ، مرجع سابق ، ص ص 7
- 8 -
- 17- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 35
- 18- حسن محمد حسن ، مرجع سابق ، ص ص 37 - 40
- 19- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص ص 35
- 36 -
- 20- حسن محمد حسن ، مبادئ الإحصاء الاجتماعي ، مرجع سابق ، ص 40
- 21- مصطفى زايد ، مرجع سابق ، ص 24
- 22- حسن محمد حسن ، مبادئ الإحصاء الاجتماعي ، مرجع سابق ، ص 41
- 23- مصطفى زايد ، مرجع سابق ، ص 25
- 24- فتحي عبد العزيز أبو راضى ، مرجع سابق ، ص 8
- 25- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 37
- 26- حسن محمد حسن ، مبادئ الإحصاء الاجتماعي ، مرجع سابق ، ص ص 41 - 42 .
- 27- مصطفى زايد ، مرجع سابق ، ص ص 24 - 25 .
- 28- فتحي عبد العزيز أبو راضى ، مرجع سابق ، ص 9 .
- 29- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص ص 38
- 39 - .

- 30- حسن محمد حسن ، مبادئ الإحصاء الإجتماعى ، مرجع سابق ، ص 42 .
- 31- مصطفى زايد ، مرجع سابق ، ص 25 .
- 32- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص 10 .
- 33- مصطفى زايد ، مرجع سابق ، ص 25 .
- 34- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص ص 39 - 40 .
- 35- حسن محمد حسن ، مبادئ الإحصاء الإجتماعى ، مرجع سابق ، ص ص 42 - 43 .
- 36- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 40 ،
- 37- مصطفى زايد ، مرجع سابق ، ص 26 .

الفصل الثالث العينات

مقدمة .

أولاً : تعريف العينة .

ثانياً : أسلوب اختيار العينة (أنواع العينات) .

ثالثاً : شروط اختيار العينة .

رابعاً : الاعتبارات التي تدعو إلى استخدام العينات .

خامساً : إطار المعاينة .

سادساً : مصادر الخطأ في العينات .

سابعاً : العوامل التي تحدد حجم العينة

ثامناً : الأساليب الإحصائية لتحديد حجم العينة .

تاسعاً : التحليل الإحصائي باستخدام العينات .

مقدمه :

إن الإجابة على التساؤلات التي يضعها الباحث أو تحقيق الفروض التي يطرحها في بحثه يتطلب قيامه بجمع بيانات يحصل عليها من ميدان الدراسة ، ثم يقوم بعد ذلك بتحليل هذه البيانات واستخلاص النتائج التي قد تؤكد صحة تلك الفروض أو تدحضها والواقع أن البيانات التي يحتاجها الباحث ما هي في الغالب الأعم إلا ردود وإجابات الناس على أسئلة توجه إليهم ليكشف الباحث بواسطتها عن قيمهم واتجاهاتهم إزاء قضايا ومواقف معينة .

ودراسة المجتمعات الإحصائية تعتمد أساسا على أخذ كل مفردات المجتمع للتعرف على خصائص ومعالم هذا المجتمع وبصفة عامة فإن معالم أي مجتمع (وهي مقادير ثابتة للمجتمع الواحد ولكنها تتغير من مجتمع إلى آخر) هي التي تعطي لهذا المجتمع صفاته دون غيره ونظرا لوجود صعوبات كثيرة تحول دون دراسة جميع مفردات المجتمع بواسطة أسلوب الحصر الشامل ، فإننا نجرى دراستنا على جزء صغير من هذا المجتمع أو ما يسمى بالعينة **Sample** حيث أنه من غير العملي أن يقوم الباحث بالحصول على بيانات من جميع أفراد المجتمع ولكنه يقوم بالحصول على تلك البيانات من قطاع صغير منه وهو ما تعارف عليه علماء الإحصاء بأنه " العينة " .

أولاً : تعريف العينة

هى جزء أو شريحة من المجتمع تتضمن خصائص المجتمع الأصلي الذي نرغب فى التعرف على خصائصه ويجب أن تكون تلك العينة ممثلة لجميع مفردات هذا المجتمع تمثيلاً صحيحاً⁽¹⁾ .

والعينة هى جزء من المجتمع ونقوم بدراستها للتعرف على خصائص المجتمع التى سحبت منه هذه العينة - ولكى تصلح النتائج التى نحصل عليها للتعبير عن المجتمع لا بد وان تكون العينة ممثلة للمجتمع (أى جميع المفردات المراد بحثها) تمثيلاً صحيحاً .⁽²⁾

واستخدام العينات معروف منذ القدم ونشاهد له أمثلة عديدة فى الحياة العملية فالكيميائي فى معمله يقوم بدراسة خواص المادة من واقع عينة من هذه المادة والطبيب يقوم بتحليل دم المريض من واقع عينة صغيرة تتكون من بضعة نقاط من دمه الخ⁽³⁾.

ويتم إتباع دراسة العينات وأسلوب المعاينة وذلك اختصاراً للوقت وتوفيراً للجهد والنفقات ولرفع مستوى العمل البحثى وجعله أكثر دقة وذلك لأن دراسة عدد قليل من المفردات أو الحالات يتيح للباحث فرصة جمع معلومات دقيقة وكثيرة عن كل مفردة أو حالة⁽⁴⁾ .

ثانياً : أسلوب اختيار العينة

هناك أساليب مختلفة لاختيار العينات ولكن نوع العينة وإجراءات سحبها من المجتمع الإحصائي تختلف من موقف لآخر والاعتبار الجوهرى الذى يراعيه الباحث هو الحصول على عينة مناسبة . والواقع أن المعيار الأساسى لكون العينة مناسبة هو أن تحظى العينة برضاء الباحث . بعض الباحثون يلجأون إلى أصدقائهم وجيرانهم وأقاربهم وزملائهم ويعتبرونهم كأفراد ضمن العينة . ويوجد عدة أساليب يعتمد عليها الباحث لاختيار العينات منها (5) :-

(1) العينات الاحتمالية : Non probability sampling

فى تلك الحالات لا تعتمد طريقة اختيار العينة على الأسلوب العشوائى نظرا لأن مجال تطبيقاتها امبريقياً يعتمد على اختيار شريحة أو قطاع معين بطريقة مقصودة . ومن أنواع العينات الاحتمالية العينة المقصودة والعينة بالحصاة .

أ- العينة المقصودة :

إن مجال استخدام هذا النوع من العينات فى الدراسات الاستطلاعية سواء من خلال المقابلات أو الاستبيان بهدف التعرف على اتجاهات فئة معينة من فئات المجتمع حول انتشار وباء معين أو نحو برنامج تليفزيونى أو إذاعى معين وما إلى ذلك وفى هذه الحالة يقتصر الباحث فى اختياره على حي معين من أحياء القاهرة مثلا ثم يقوم الباحث بعد ذلك باختيار عدد من الأسر بهذا

الحي دون أي اختيار عشوائي وهنا تبرز أول عيوب العينة الاحتمالية وتتمثل في صعوبة تعميم النتائج سواء على مستوى القاهرة كمدينة أو حتى التعميم على مستوى حي معين آخر . أما العيب الثاني فيتمثل في صعوبة حصول الباحث على تقدير صحيح للخطأ المتوقع بسبب المجازفة (6) .

ب- اختيار العينة بالحصة : Quota sampling

وفيها يتم اختيار المبحوثين بنسبة توزيعهم في المجتمع الاحصائي مثال اختيار 20% من الإناث 40 % من الذكور وهكذا . ولكن الاختيار الاعتبائي والاختيار بالحصة يعد اختيارا غير اهتمامي ، بمعنى أنه لا يوفر فرصة متكافئة لكل مفردات المجتمع الاحصائي لتظهر في العينة مما يؤدي إلى إخفاق العينة في أن تمثل المجتمع ككل وتستخدم أحيانا في المسوح الاحتمالية للرأي العام وتكون في هذه الحالة أشبه بالعينة الطبقية . ففي هذه الحالة يعطي القائم بالمقابلة حصة معينة يجب استيفاء بياناتها كأن يلتزم بعدد كبير من الإناث فمن يزيد أعمارهن عن أربعين عاما وأيضا يلزم بعدد كبير من الأشخاص تقل دخولهم السنوية عن (300) جنيه . أو أن يخصص له نسبة معينة من الأطباء في مجتمع ما وهكذا بحيث يكون الباحث قادرا على أن يتم الحصة المطلوبة منه (7) .

(2) العينات الاحتمالية : Probability Samples

لقد طور العلماء أساليب المعاينة الاحتمالية لتجنب المخاطر التي تترتب على اختيار عينة غير ممثلة لمجتمع الدراسة وهذه المخاطر لا يمكن تجنبها تماما ولكن هذه الأساليب تمكننا على الأقل من تحديد نسبة الخطأ المحتمل وتعرف العينة الاحتمالية بأنها العينة التي يتم سحبها بحيث يكون لكل مفردة من مفردات المجتمع فرصة معلومة ومتكافئة في أن يكون جزءا من العينة .

يتسم هذا النوع من العينات بالخصائص التالية :-

- أ - لكل مفردة في العينة درجة احتمالات معروفة يفترض وجودها بين باقى مفردات تلك العينة .
- ب - لجميع مفردات المجتمع الأصلي فرص متساوية للظهور في العينة .

يلزم أن تكون الاحتمالات معروفة لدى الباحث حتى يمكن التوصل إلى الثقل الصحيح للعينة أما إذا لم يعرف الباحث تلك الاحتمالات فإنه قد يستحيل عليه أن يستخدم بنجاح الاستنتاج الإحصائي المعتمد على دلالات بحثية . (8)

(3) العينة العشوائية البسيطة: Simple Random sample

العينة العشوائية هي العينة التي تختار بحيث تعطي جميع مفردات المجتمع المراد بحثه نفس الفرصة في الاختيار وهذا يعنى عدم الاهتمام ببعض المفردات أكثر من البعض الآخر وإتاحة الفرصة

المتكافئة أمام كل مفردة للظهور فى العينة ويمكن أن نحقق ذلك بأن نحضر عدا من البطاقات المتشابهة (فى اللون والحجم والوزن وكل شئ) ونكتب على كل بطاقة رقماً يمثل مفردة من مفردات المجتمع وتسحب عددا من هذه البطاقات (بعد خلطها) فنجد أن الأرقام المدونة عليها تعطي لنا المفردات التى تم اختيارها بطريقة عشوائية (9) . وتعرف العينة العشوائية البسيطة بأنها اختياراً بسيطاً بطريقة تتصف بخاصيتين أساسيتين هما :-

أ - أن يتحقق لكل عضو أو مفردة من المجتمع الأصلي درجة احتمال متساوية فى الاختيار

ب - أن يكون اختيار كل مفردة من مفردات العينة بصورة مستقلة عن الأخرى (10)

لو تصورنا أن أحد الأساتذة بقسم الاجتماع يود إجراء دراسة عن اتجاهات طلاب القسم نحو إدمان المخدرات ثم وضع أسماء هؤلاء الطلاب وعددهم 4000 فى حقيبة كبيرة ثم سحب منها 400 اسم أو أنه أعطى رقماً مسلسلاً لكل من هؤلاء الأربعة آلاف طالب تم اختيار 400 رقماً من جدول الأرقام العشوائية وقام بعد ذلك باختيار الطلاب الذين يتطابق رقمهم المسلسل مع الأرقام العشوائية المختارة له فإنه يكون بذلك قد أعطى لكل طالب من الطلاب فرصة متكافئة لى يكون من أحد أفراد العينة .

(4) العينة المنتظمة : Systematic sample :

العينة المنتظمة هي نوع من المعاينة العشوائية بمقتضاها يمكن أن يختار الباحث لو أخذنا في الاعتبار المثال السابق نسبة 10% من عدد الطلاب (400 طالب) ويستطيع الباحث أن يختار هؤلاء الطلاب بطريقة عشوائية فيبدأ بالطالب رقم 8 ثم بعد كل عشر طلاب يقوم باختيار طالب آخر وهكذا أي أنه في هذه الحالة سيختار الطالب رقم 8 ، 18 ، 28 ، 38 وهكذا . وهذه الطريقة في الاختيار مقبولة ما لم يكن اختيار الأرقام من البداية يخفض وراءه تحيز الباحث نحو اختيار طلاب بعينهم . والواقع أن الطريقتين السابقتين من طرق اختيار العينات تلائم الباحثين المبتدئين وغيرهم ممن يريدون تجنب التعقيدات الإحصائية وهناك بالإضافة إلى تلك الطرق أساليب أخرى أكثر تطوراً لسحب العينات توفر للعينة صفات أساسية كأن تكون ممثلة ومقبولة ومناسبة من حيث التكاليف (11)

وتعتبر العينة المنتظمة أكثر أفضلية من العينة العشوائية البسيطة وذلك في حالة توفر قوائم تضم جميع مفردات المجتمع الأصلي غير أن السهولة في العينة المنتظمة يناظر بعض العيوب من أهمها .

أ- توقع نتائج خاطئة إذا تم استخدام هذا النوع من العينات في مجتمعات تتسم بتكرار ظواهر دورية .

ب - اقتصار العشوائية فقط في تحديد الرقم الأول في بداية اختيار العينة . (12)

(5) العينات الطبقيّة : Stratified Samples :

تتميز العينات الطبقيّة على غيرها من العينات بأنّها بالإضافة إلى كونها دراسة للمجتمع ككل فإنّها تتيح لنا دراسة كل طبقة من الطبقات على حده وهذا قد يكون مرغوباً فيه في كثير من الأحيان ففي دراسة لبحث ميزانية الأسرة نحصل على نتائج البحث لكل من الريف والحضر على حده وهما الطبقتان اللتان يتكون منهما المجتمع ، وبذلك تمكنا العينة الطبقيّة من دراسة كل من الريف والحضر إلى جانب دراسة المجتمع المصري ككل . (13)

تعتمد هذه الطريقة على تقسيم المجتمع الإحصائي إلى فئات أو طبقات ثم اختيار عينة من كل طبقة ففي المثال السابق يمكن لباحث أن يقسم الأربعة آلاف طالب بحسب أصولهم الحضرية إلى طلاب من الدلتا ، وطلاب من صعيد مصر ، ثم يقوم باختيار عدد من الطلاب الذين ينتمون إلى كل من هذه التقسيمات بطريقة عشوائية ويتحدد عدد الطلاب الذين سيتم اختيارهم من كل طبقة بحسب نسبة تلك الطبقة إلى المجموع الكلي للمجتمع الأصلي فلو فرضنا على سبيل المثال أن 50% من جملة عدد الطلاب وهم 4000 طالب ، من المدن فإن معنى هذا أن 50% من

العينة التي حجمها 400 طالب يتم اختيارهم من المدن وهكذا .
وعموماً يمكن صياغة تلك العلاقة في القانون التالي :

عدد الأفراد المراد اختيارهم من طبقة معينة =

عدد أفراد الطبقة

$$= \text{حجم العينة المراد سحبها} \times \frac{\text{جملة عدد أفراد المجتمع الإحصائي}}{\text{عدد أفراد الطبقة}}$$

في هذه الحالة من المعتقد أن خطأ المعاينة من المحتمل أن يتناقص ليصل إلى الصفر . فتوزيع الطلاب بحسب موطنهم الأصلي فضلاً عما يعكسه من تباين ثقافي بين الطلاب فإنه يقترب كثيراً من الواقع (14)

وتقوم العينة الطبقية على تقسيم المجتمع الأصلي إلى مجموعات يطلق عليها طبقات فرعية أو شرائح Strata ثم نأخذ عينة من كل شريحة على حده بحيث يتكون لدينا عينة ذات حجم كلي (ن) ومن الأهمية بمكان أن يتحدد تعريف الشريحة الطبقية بضرورة ظهور كل فرد من شريحة واحدة فقط ولا يتكرر في غيرها . وفي الطريقة البسيطة والشائعة من حيث الاستخدام للعينة الطبقية أن تستخدم في الاختيار وعند بداية تصميم نموذج العينة الطبقية على الباحث اتخاذ الخطوات التالية :

- حساب تقديري للمتوسطات الحسابية لكل شريحة على حده .
- حساب تقديري للانحراف المعياري لكل شريحة على حده .

- بعد تقدير قيمة (ع) لكل شريحة نبدأ فى وضع أوزان تبعاً لحجم الشريحة ونسبة هذا الحجم للمجتمع الأصلي (15) .

(6) العينة غير المتناسبة : Disproportionate Sample

يلجأ الباحث عادة إلى مثل هذا النوع من العينات إذا كان يريد أن يرفع نسبة عينة جماعة فرعية معينة . فلو أراد الباحث فى مثلنا السابق أن يعرف رأى الطلاب الذين من أصل قروي فى قضية الإدمان لما يتميزون به من وازع ديني وأخلاقي فإنه فى هذه الحالة يزيد من نسبة تمثيل الطلاب القرويين لأن طبيعة مشكلة البحث تقتضى ذلك فيختار الباحث 200 طالب من المناطق الريفية وباقي الطلاب من المدن ومن الصعيد . ولكن فى هذه الحالة ينبغي على الباحث أن يظهر فى تحليله العوامل التى دفعته لمثل هذا النوع من الاختيار .

(7) العينات العنقودية ذات المرحلة الواحدة ومتعددة المراحل

Single , stage and Multi . stage cluster Samples

فى حالة العينات كبيرة الحجم يلجأ الباحث إلى هذا الأسلوب من أساليب المعاينة لتخفيض نفقات اختيار العينة والعينة العنقودية ذات المرحلة الواحدة تتمثل فيما يقرره احد الباحثين من اختيار حى سكنى معين من إحدى المدن كعينة للدراسة ثم يختار مجموعة من الأسر التى تقطن ذلك الحى لإجراء مقابلة معهم . معنى هذا أن المقابلات التى سيقوم بها الباحث سوف تتجمع فى

حي معين الأمر الذي ساعد على تخفيض الوقت والنفقات ونلاحظ هنا أن اختيار العينة تم على مرحلة واحدة .

- أما العينة العنقودية متعددة المراحل فيلجأ إليها الباحث عند اختيار عينة أكبر حجماً . فلو أردنا أن ندرس اتجاهات الشباب نحو الإدمان فإنه يمكن أن تحصل على خريطة بأحياء المدينة ثم تختار من بينها عدداً من الأحياء الشعبية وعدداً آخر من الأحياء الراقية ثم تختار عدداً من القطاعات داخل الأحياء وبعد ذلك يتم اختيار من تتم مقابلتهم كأفراد داخل العينة . من ذلك يتضح لنا أن أسلوب العينة العنقودية متعددة المراحل وإن كان يحقق الدقة ويرفع درجة تمثيل العينة للمجتمع الأصلي إلا أنه أسلوب يكتنفه التعقيد ولا يستطيع كثير من الباحثين ذوي الإمكانيات المحدودة الاستعانة به (16)

نظراً لضيق الوقت وكثرة التكاليف والجهود اللازمة لاختيار عينة عشوائية بسيطة في معظم الأحيان فإننا قد نجرى الاختيار على مراحل متعددة . فإذا كان المجتمع يتكون من أقسام متجانسة نبدأ باختيار بعض هذه الأقسام عشوائياً (كمرحلة أولى) ثم نختار عينة عشوائية بسيطة من كل قسم من الأقسام التي تم اختيارها (كمرحلة ثانية) وقد يحتاج الأمر إلى اختيار عينة عشوائية بسيطة من كل قسم من الأقسام التي تم اختيارها في

المرحلة الثانية و وهكذا والعينة التي يتم اختيارها بهذا الشكل تعرف بالعينة متعددة المراحل (17) .

ثالثاً : شروط اختيار العينة

1- يجب أن لا تتسم العينة التي تم اختيارها بالتحيز أو المحاباة بمعنى أن تأخذها من بين مفردات المجتمع الأصلي عشوائياً .

2- أن تكون الظاهرة المراد عمل معاينة لها سائدة ومنتشرة في المجتمع الأصلي ولا تكون نادرة الحدوث .

3- يجب أن تكون العينة ممثلة لجميع فئات المجتمع الأصلي .

4- ضرورة افتراض تجانس مفردات المجتمع الأصلي وفي حالة تعذر ذلك في بعض المجتمعات غير المتجانسة يلجأ الباحث إلى تقسيمها إلى مجتمعات صغيرة متجانسة .

5- ضرورة إجراء حصر مسبق لجميع مفردات المجتمع الأصلي المراد بحثه مع تقسيم هذا المجتمع الى وحدات معاينة كل منها داخل قوائم أو ما نسميه إحصائياً بالأطر فعلى سبيل المثال عند دراسة سكان مجتمع ما فإن وحدة المعاينة أما أن تكون الأسرة كوحدة تحليل أو الفرد أو الجماعة وقد يكون المجتمع بالنسبة للمجتمعات الكبيرة .

6- يجب أن يتناسب اختيار حجم ونوع العينة مع الهدف الأساسي للباحث من العينات مع طبيعة المجتمع أو نوع المشكلة موضوع الدراسة وهكذا (18) .

أى أنه يجب أن تتوفر فى العينة الممثلة **Representative sample** مجموعة من الشروط يمكن تلخيصها فى شرطين أساسيين هما :

أ- تكون مفردات العينة ممثلة للمجتمع الذي يجرى عليه البحث تمثيلاً صحيحاً وليست ممثلة لمجتمع آخر . بمعنى أنه إذا تكررت نفس النتائج على عينات أخرى من نفس المجتمع ، كانت العينة التى يجرى عليها البحث عينة ممثلة للمجتمع الأصلي أصدق تمثيل ، وبذلك يمكن أن تكون خصائص مفردات العينة (إحصائيات العينة) متقاربة أو متشابهة مع خصائص المجتمع (معالم المجتمع) الذي تنتمي إليه .

ب - ألا تكون المفردات المختارة ممثلة لجزء (قطاع) من أجزاء المجتمع الأصلي بل يجب أن تمثل جميع أجزاء المجتمع (19) .

رابعاً : الاعتبارات التى تدعوا إلى استخدام العينات

يعتبر السبب الرئيسى لاستخدام العينات هو توفير الوقت والجهد والنفقات فإذا كان المال المخصص لإجراء بحث معين أو

نوع الباحثين وعددهم أو الوقت اللازم لانجاز هذا البحث لا يسمح بإجراء الحصر الشامل فإننا نضطر لاستخدام العينات لدراسة خصائص المجتمع الذي نجرى البحث لدراسته . وقد تكون هذه العوامل الثلاثة متوفرة لدينا ، ومع ذلك نلجأ لاستخدام العينات رغبة في توفير المال أو اختصاراً للوقت أو ادخاراً للجهد أي بهدف حسن توجيهه واستغلال الإمكانيات المادية والفنية . المتاحة في بعض الأحيان يكون المجتمع الذي ندرسه غير محدد ، فإذا أردنا مثلاً فحص إنتاج آلة معينة فالمجتمع هنا يكون ما أنتجته الآلة وما تنتجه الآن وما سوف تنتجه في المستقبل ، لذلك يستحيل في مثل هذه الحالة إجراء حصر شامل ويكتفى بدراسة عينة من إنتاج الآلة .

قد يؤدي أحيانا فحص المفردات إلى تدميرها فإذا أردنا تحليل الدم لشخص مريض فان الحصر الشامل هنا يعنى سحب كل دم المريض بغرض تحليله ، وهذا يعنى قتله ، ولذلك لا بد في مثل هذه الحالة من استخدام العينات . أي تجرى التحليل على عينة من بضعة نقاط من دم المريض ، وسنجد عموماً أنه لا بد من استخدام العينات في الحالات التي يؤدي فيها فحص المفردات إلى إتلافها . (20)

اختيار مفردات العينة :-

إن عملية اختيار مفردات العينة من بين مفردات المجتمع الأصلي أو ما يعرف بأسلوب سحب العينة من المجتمع كواحدة من المشكلات الخاصة بأسلوب المعاينة ، تتوقف أساسا على حجم المجتمع الأصلي . فإذا كان حجم المجتمع صغيرا أي مشتملا على عدد محدد (finite) من المفردات ، فإن المشكلة لا تكون مشكلة اختيار العينة من بين مفردات المجتمع ، بل تكون مشكلة الحصول على عدد كاف من المفردات لغرض البحث . فمثلا إذا أراد الباحث أن يجرى دراسة على كبار الزراعيين بإحدى القرى ، كنموذج لنفس الفئة فى القطر ، فقد يحدد هذه الفئة بأنها تشتمل على كل من يمتلك "100 فدانا أو أكثر من الأراضي الزراعية فى القرية " وفى هذه الحالة يكون عدد هؤلاء الملاك قليلا لدرجة أن العينة تستنفذهم جميعا . كما تكون عملية الاختيار من المجتمع الأصلي عملية مشروطة بتحديد المفردات (عدد الملاك) التى تتكون منها العينة المطلوبة وبالطبع كلما كثرت الشروط اللازمة للعينة كلما صعب الحصول عليها وكلما قل عدد المفردات الذين يتم الاختيار من بينهم . أما إذا كان حجم المجتمع الأصلي كبيرا جداً أي مشتملاً على تحدد عدد غير محدد من المفردات المستوفية لجميع الشروط اللازمة فى العينة فإنه من اللازم إجراء عملية اختيار مفردات العينة إما بواسطة الاختيار غير العشوائى (المعاينة العمدية) أو بواسطة الاختيار العشوائى (21)

يستطيع الباحث أن يسلك شتى السبل ويستخدم كافة الأساليب للحصول على عينة للدراسة ولكنه فى كل الأحوال يجب أن يتوخى الحذر من التحيز فى اختيار العينة كما ينبغي عليه أن يتأكد من أن العينة ممثلة لمجتمع الدراسة حتى تكون التعميمات التى يتوصل إليها من تحليلاته مستمرة وقيمة وإلا انعدمت الفائدة من الدراسة (22)

خامساً : إطار المعاينة : Sampling Frame

الإطار هو حصر شامل لجميع مفردات المجتمع المراد بحثه فقد يكون الإطار عبارة عن قائمة بالمفردات أو مجموعة من البطاقات أو الخرائط أو الخ فعند اختيار العينة يقسم المجتمع إلى أجزاء تسمى وحدات المعاينة (Sampling units) ويكون الإطار عندئذ هو مجموعة القوائم التى تحتوى على هذه الوحدات التى يتكون منها المجتمع . ولما كانت العينات تختار من هذا الإطار وجب أن يكون شاملاً لجميع مفردات المجتمع مع ملاحظة عدم تكرار أي من هذه المفردات لأن عملية التكرار سوف تعطي هذه المفردات فرصة أكبر للاختيار فى العينة وبذلك تتحيز النتائج التى تحصل عليها المفردات التى تكررت فى الإطار ويجب أن يكون الإطار أيضاً متجدداً حتى تعطى المفردات التى تستجد على الإطار القديم نفس الفرصة فى الظهور فى العينة (23)

ويعتبر إطار المعاينة هو المصدر الذي تؤخذ منه العينة أو بعبارة أخرى هو حصر شامل (القائمة أو الدليل) لجميع مفردات وحدات المجتمع الأصلي المراد دراسته .

مثال ذلك قائمة بأسماء العمال فى أحد المصانع ، أو مختلف أنواع الرواسب التى توجد على الشاطئ ، أو موقع المحلات العمرانية الريفية على خريطة إحدى الدول . وعند اختيار العينة من المجتمعات المحدودة يقسم المجتمع الأصلي للظاهرة قيد البحث إلى عدة أقسام تسمى وحدات المعاينة (شخص ، أسرة ، قرية) ويكون إطار المعاينة حينئذ هو عبارة عن القائمة أو مجموعة القوائم التى تتضمن الوحدات التى يتألف منها المجتمع . ويشترط فى إطار المعاينة أن يكون شاملاً لجميع مفردات المجتمع التى يمكن الوصول إليها بسهولة ، وذلك حتى يكون اختيار العينة سليماً . كما يشترط أن يكون إطار المعاينة متجدداً حتى تعطي المفردات أو الوحدات التى تستجد على الإطار القديم نفس الفرصة فى الظهور .

ونظراً لأنه فى المجتمعات غير المحددة *infinite* يستحيل إجراء حصر شامل لكل مفردات المجتمع فى الوقت المتاح للدراسة ، ويكتفى فى هذه الحالة بدراسة عينة بدون تكوين إطار للمعاينة . ويلاحظ على إطار المعاينة وفى مجال الدراسة الجغرافية أنه إما أن يكون إطاراً مكانياً *Spatial* أو غير مكانياً *Non - Spatial* .

1- إطار المعاينة المكاني . هو الإطار الذي يكون فيه المكان Location هو الوحدة الرئيسية ، كما أنه الأساس فى اختيار العينات التى تمثل التغيرات (الاختلافات) المكانية التى يتميز بها مجتمع الأماكن لمنطقة ما تمثيلاً صحيحاً .

فمثلاً إذا كنا بصدد معاينة خريطة بهدف تحديد مساحة الأراضي التى يشغلها نوعا معينا من النشاط الزراعي على هذه الخريطة ، فإننا يجب أن نتأكد من تمثيل كل أجزاء الخريطة تمثيلاً صحيحاً . ويتم ذلك باختيار أحد المعاينات الآتية :-

(أ) المعاينة النقطية : Point - sampling أي معاينة نقط تقاطع شبكة مربعات على خريطة المنطقة .

(ب) المعاينة الخطية : Line – sampling أي نأخذ عينة من قطاعات عرضية مختلفة على الخريطة .

(ج) المعاينة المساحية : Area – sampling أي بأخذ عينة تمثل مساحة مجموعة من المربعات التى تغطي مساحة خريطة المنطقة قيد البحث .

وعلى ذلك يكون إطار المعاينة عبارة عن جميع مفردات المجتمع شكل من أشكال المعاينة الثلاثة .

2- إطار المعاينة غير المكاني - على الرغم من أن طبيعة عمل الجغرافي عند جمعه للبيانات ترتبط بإطار المعاينة المكاني ،

إلا أنه في بعض الأحيان وظروف خاصة نجده يهتم بتحديد إطار معاينة غير مكاني ليلتم دراسته فمثلاً إذا كان يصدد اختيار عينة من أسر أحد الأقسام الإدارية في مدينة ما وذلك لتقدير متوسط الدخل ، فإنه يتحتم عليه اختيار عينة من إطار (أو قائمة) تحتوى على جميع أسر هذا القسم الإداري بالمدينة . ولا يجوز له في هذه الحالة أن يختار العينة من دليل التليفون مثلاً إذ أنه من المعروف أن مثل هذا الدليل لا يتضمن جميع أسر القسم الإداري قيد البحث (24) .

سادساً : مصادر الخطأ في العينات

يتضح لنا مما سبق أن خطأ التحيز أمر متوقع لا محالة في المعاينة الاحتمالية ولا يقتصر هذا التحيز على العينة فقط بل قد نجده أيضاً في عمليات الحصر الشامل حيث تتوافر فرص عديدة للوقوع في مثل تلك الأخطاء . وقولنا بضرورة وقوع أخطاء يبرره عدم التدريب الكامل للقائمين بالبحث أو المساعدين حول كيفية التغلب على العقبات التي قد تواجههم . هذا فضلاً عن عدم الاستخدام الأمثل للأطر المناسبة والممثلة لاختيار العينة بالطرق الإحصائية السليمة (25) .

ويلاحظ أن النتائج التي نحصل عليها من العينة قد لا تماثل تماماً النتائج التي نحصل عليها من الحصر الشامل وذلك لأن العينات عرضه لنوعين من الخطأ .

- 1- خطأ الصدفة (الخطأ العشوائي) أو ما يسميه البعض بخطأ العينة .
- 2- خطأ التحيز .

(1) خطأ الصدفة Random Error

يرجع هذا الخطأ إلى طبيعة الاختيار العشوائي حيث قد تختلف نتائج العينة عن نتائج المجتمع . ويتوقف خطأ الصدفة على كل من حجم العينة وتباين المجتمع وطريقة اختيار العينة وكلما كبرت العينة كلما قل خطأ الصدفة وزادت ثقتنا في النتيجة ، وعلى العكس من ذلك لو زاد تباين مفردات المجتمع لزداد احتمال حدوث الأخطاء العشوائية وعموما لو اختيرت العينة بطريقة عشوائية سليمة لأمكن تقدير هذا النوع من الخطأ من العينة نفسها (26) .

ويتوقف هذا النوع من الخطأ على درجة تباين المجتمع الأصلي وطريقة اختيار العينة وحجمها فكلما كبر حجم العينة قل خطأ الصدفة وبالتالي زادت درجة الثقة في النتائج .

هذا ويمكن التحكم في قيمة هذا الخطأ وتقديره بالطرق الإحصائية وأن كان يصعب تجنب وقوعه إلى حد بعيد . كذلك يجدر الملاحظة أن هذا النوع من الأخطاء يؤثر على العينة وحدها ولا يتأثر به الحصر الشامل بوصفه أحد المصادر الهامة لجمع البيانات .

مثال : فإذا كان لدينا ست أطفال وكانت أعمارهم بالسنة على التوالي 2، 3 ، 4 ، 6 ، 9 ، 12 . أي أن متوسط العمر في هذه المجموعة

$$6 \text{ سنوات} = \frac{36}{6} = \frac{12+9+6+4+3+2}{6} =$$

فإذا سحبنا عينة عشوائية مكونة من حالتين فقط من

هؤلاء الأطفال ولتكن 2 ، 4 فإن متوسط العمر يكون

$$3 \text{ سنوات} = \frac{6}{2} =$$

وهنا نجد فرقا كبيراً بين متوسط العينة ومتوسط المجتمع

الأصلي . وإذا سحبنا عينة أخرى مكونة من حالتين وثالثة ، ورابعة لا يكون هذا الاختيار دقيقاً إلا في حالة سحب الحالتين رقم 3 ، 9 ففي هذه الحالة الأخيرة يمكن القول بأن القيمة المقدرة لأعمار الأطفال تنطبق تماماً على القيمة الحقيقية للأعمار . حيث

$$\text{أن متوسط العينة} \\ \frac{9+3}{2} = 6 \text{ سنوات} =$$

وهو نفس المتوسط الحقيقي للمجموعة . أي أن خطأ الصدفة يرجع إلى الفرق بين القيمة المقدرة من العينة والقيمة الحقيقية في المجتمع الأصلي الذي سحبت منه العينة . ومن هنا لا يستطيع الجزم بأن متوسط القيم في أية عينة هو نفس المتوسط العام للقيم الحقيقية في المجتمع الأصلي ، فقد يكون عمر أحد أفراد العينة

صغيراً فينخفض متوسط العينة وقد يكون كبيراً فيرتفع المتوسط في العينة عن المتوسط الحقيقي ولا يحدث خطأ الصدفة في حالة حدوث التعادل . كذلك لا يمكننا الجزم بحدوث هذا التعادل في أي حالة معينة إذا تركت للصدفة وحدها وكل ما يمكن أن نقوله هنا هو أنه يحتمل حدوث هذا التعادل. (27)

(2) خطأ التحيز Bias Error

هذا الخطأ لا يتوقف على عنصر العشوائية أو الصدفة . ويحدث عادة في اتجاه واحد أي بالزيادة فقط أو بالنقص فقط وتكون خطورته في أنه لا يمكن حصره أو وضع حدود له .

مثل خطأ الصدفة . وهذا النوع من الخطأ ليس قاصراً فقط على العينات بل قد يتعرض له الحصر الشامل نتيجة لعدم الدقة في القياس أو عدم كفاءة الباحثين أو غموض كشوف الأسئلة أو إعطاء بيانات غير صحيحة من قبل المبحوثين أو عدم جمع البيانات عن بعض مفردات المجتمع أو جمع البيانات عن بعض مفردات المجتمع أكثر من مرة أو... الخ

وتتعرض العينات لخطأ التحيز لنفس الأسباب التي يتعرض لها الحصر الشامل بالإضافة إلى الأسباب الآتية :

أ- عدم وجود إطار سليم عند سحب العينة ، فاستخدم إطار قديم أو إطار غير شامل لجميع مفردات المجتمع يؤدي إلى تحيز العينة للمفردات الموجودة في الإطار فقط ، ولو

تكررت بعض المفردات فى الإطار ، فإن ذلك يؤدي إلى
تحيز العينة للمفردات المتكررة .

ب- حالة عدم إمكانية الوصول لبعض مفردات العينة يستعاض
عن هذه الوحدات بوحدة أخرى وذلك قد يؤدي إلى
التحيز ، ففي حالة عدم تمكن الباحث من الحصول على
بيانات بعض الأسر نتيجة لتغيبها خارج المسكن نجد أن
الاستعاضة قد تؤثر على مدى تمثيل العينة للأسر
الصغيرة أو للأسر التي تشتمل على زوجات عاملات .

ج - قد ينشأ التحيز نتيجة لعدم إتباع الطرق السليمة فى
حساب التقديرات (28) ويتسم هذا النوع من الخطأ
بالتحيز غالباً نحو جانب واحد إما بالزيادة أو النقصان
وتزداد أهمية هذا النوع من الخطأ كلما كبر حجم العينة
حيث تقل فرص الخطأ العشوائي .

ويرجع حدوث أخطاء التحيز لعدد من العوامل نذكر من بينها .

- سوء التقدير وعدم توفر الدقة من جانب الباحث وذلك عند
قيامه بعمليات الحصر حيث قد تفوته الدقة الكافية فى حساب
المتغيرات وكذلك عدم توفيق الباحث فى صياغة الفروض
الصحيحة .

- صياغة أسئلة غامضة وغير واضحة للمبحوثين .

- عدم استجابة بعض مفردات العينة لأسئلة المقياس .
- الاختيار المقصود غير العشوائي لمفردات العينة .
- سوء اختيار العينة وقد يحدث نتيجة لسحب العينة من إطار غير كامل .
- عدم دقة القياس (29) .

ويتعرض العمل الإحصائي إلى أنواع كثيرة من الأخطاء أثناء تنفيذه ومنها نوعين رئيسيين من أنواع الأخطاء التي يتعرض لها قياس البيانات والتي من شأنها التأثير على النتائج التي نحصل عليها من العينة وهما أخطاء التحيز والأخطاء الاحتمالية .

وأخطاء التحيز هي الأخطاء الناجمة عن تدخل الباحث في طريقة اختيار العينة فالمعروف مثلاً أن العينة العشوائية تمثل بشكل كبير خصائص المجتمع الذي سحبت منه فإذا اختيرت العينة بطريقة شخصية (أي غير عشوائية) فإن ذلك يؤدي إلى زيادة الأخطاء المتوقعة . كذلك تنشأ هذه الأخطاء نتيجة لتحيز الباحث لوجهة نظر خاصة تجاه القرارات المتخذة ، ويحدث عادة خطأ التحيز في اتجاه واحد أما بالزيادة أو بالنقص

ويمكن أن تعزى أخطاء التحيز لعدة عوامل أهمها :

أ- الاختيار المتعمد (غير العشوائي) للعينة .

ب- استبدال أفراد العينة بمفردات أخرى لعدم تمكن الباحث من الوصول لبعض المفردات الأساسية فى العينة .

ج - سوء التقدير وعد توافر الدقة . فقد لا يوفق الباحث فى التفرقة بين ما هو سبب أو نتيجة أو عدم توفر الدقة فى حصر وحساب المتغيرات المحددة لطبيعة الظاهرة ووضع فروض غير سليمة أما الأخطاء الاحتمالية فهي الأخطاء الناجمة عن احتمالات عدم تماثل النتائج التى نحصل عليها مع خصائص المجتمع . فحتى عندما تؤخذ العينة بالأسلوب العشوائى ، فإنه تظل هناك احتمالات أخطاء فى مدى تمثيل العينة لخصائص المجتمع الذى أخذت منه . ومنهم أهم هذه الأخطاء ما يطلق عليه إحصائياً خطأ الصدفة أو الخطأ العشوائى (30) .

سابعاً : العوامل التى تحدد حجم العينة

عندما يبدأ الباحث فى التفكير فى إجراء دراسته الميدانية يكون من أهم الأسئلة التى ينبغى أن يجيب عنها ذلك السؤال المتعلق بحجم العينة وهل هو مناسب ، كبير ، أم صغير والإجابة عن ذلك السؤال تتوقف على عدة عوامل هي :

1- حجم المجتمع الإحصائى الذى ستسحب منه العينة . حيث يشير إلى مجموع الأفراد الذين سيقوم الباحث بسحب العينة من بينهم ، وهؤلاء الأفراد يشكلون جزءاً من مجتمع أكبر يعرف بالمجتمع الأصلي . فإذا كان الباحث ، على سبيل المثال ، يريد أن

يجرى دراسة على عينة من طلبة كلية الآداب ، فإن عدد هؤلاء الطلبة يمثل المجتمع الإحصائي ، فى حين أن عدد طلبة جامعة المنصورة بجميع كلياتها يكون بمثابة المجتمع الأصلي . وبطبيعة الحال من المعقول أن نقرر أنه كلما كان حجم المجتمع الإحصائي كبيراً كلما تطلب ذلك أن يكون حجم العينة كبيراً . ويقدر ما يشكل حجم العينة نسبة كبيرة من المجتمع الإحصائي بقدر ما تكون العينة ممثلة لذلك المجتمع فالعينة التى عدد مفرداتها 40 طالبا من فصل مدرسي عدد طلابه 50 طالبا تعد عينة ممثلة تمثيلاً صادقاً لذلك الفصل ولكن هذا العدد لا يعتبر عينة ممثلة لمدرسة عدد طلابها 1000 طالب . وبعبارة أخرى ، يعتبر كبر حجم العينة ضماناً لأن تكون العينة ممثلة للمجتمع الإحصائي . وليس معنى هذا أن يزيد الباحث من حجم العينة إلى أن تصبح دراسته الميدانية حصراً شاملاً لكل مفردات المجتمع الأصلي الذي يقوم بدراسته ولهذا يلجأ الباحثون إلى استخدام الأساليب الإحصائية لتحديد الحجم المناسب للعينة التى يقومون بدراستها . فزيادة العينة بعد ذلك الحجم لن يضيف إضافة جوهريّة إلى درجة الضبط التى ينبغى أن تتميز بها النتائج بقدر ما يضيف من أعباء وتكاليف وما يستغرق من وقت .

2- درجة الاختلاف بين مفردات المجتمع الإحصائي .
فإذا كانت درجة الاختلاف كبيرة بين أفراد ذلك المجتمع استدعى الأمر زيادة حجم العينة والعكس صحيح . فعندما يكون هناك تماثل

تام بين أفراد المجتمع . كأن يكونوا متفقيين على قضية عامة ، فإن عينة صغيرة جداً منهم تكفي لكي تمثل المجتمع كله . فلو أننا سألنا . 100 فرد هذا السؤال : هل توافق على عودة الشعب الفلسطيني إلى فلسطين ؟ لكان ردهم كافياً للتعبير عن اتجاهات ملايين العرب نحو القضية الفلسطينية ، بينما لا يكفي هذا العدد كعينة إذا كان السؤال يقصد منه دراسة اتجاهات الأفراد أو نحو السياسية التعليمية .

3- نسبة الخطأ المسموح به أو المقبول ودرجة الثقة التي يرغب الباحث في توافرها في النتائج التي يصل إليها من دراسته للعينة . حيث تعد درجة الضبط المطلوبة في التنبؤ الذي يبني على نتائج دراسة هذه العينة ودرجة الثقة في هذا التنبؤ من العوامل المحددة لحجم العينة . فإذا كان الباحث يسعى إلى التوصل إلى نتائج موثوق بها ويمكن الاعتماد عليها واستخدامها في التنبؤ، فإن حجم العينة التي سيقوم بدراستها ينبغي أن يكون كبيراً ، ولكن كما قلنا سلفاً ، كبر حجم العينة يتطلب وقتاً طويلاً وتكلفة ضخمة ، لهذا السبب اعتاد الباحثون أن يقبلوا حجم العينة الذي يستطيعون بنسبة ثقة 95% أن يعتمدوا على البيانات التي يوفرها لبحثهم وتساعدتهم في استخلاص نتائج يمكن تعميمها على مجتمع الدراسة (31) .

وتتفق آراء كثير من الإحصائيين على أن حجم العينة عينة البحث تتوقف على مجموعة من العوامل تنحصر في : الغرض من البحث ، حجم المجتمع الأصلي ، مدي تباين الظواهر المختلفة في قطاعات المجتمع ، ودرجة الدقة المطلوبة في البحث ، البيانات المتاحة التي يمكن استخدامها في تعميم النتائج ، والإمكانيات المادية .

ونظرا لعدم وجود اتفاق بين الباحثين على وضع حد معين على أساس علمي . أو إحصائي . يحدد الحجم المناسب أو الأمثل للعينة لكي تمثل المجتمع الذي تسحب منه تمثيلاً جيداً ، فإن تقدير حجم العينة على مستوى معظم الدراسات والبحوث - تعتبر واحدة من المشكلات الخاصة بأسلوب المعاينة وتطبيق الأساليب الإحصائية ، وفي مجال العمل الإحصائي يوجد اتجاهان عند تقدير حجم العينة .

الاتجاه الأول: يعتمد على الخبرة السابقة للباحث في هذا المجال ، حيث أظهرت خلاصة الخبرات والتجارب أن حجم عينة في حدود 10% إلى 15% من حجم المجتمع الأصلي يبدو ملائماً في معظم الدراسات والبحوث . ويتميز هذا الاتجاه في تقدير حجم العينة بسهولة ، كما أنه يفيد بعض الباحثين قليلي الخبرة في مجال العمل الإحصائي .

الاتجاه الثاني : يرتبط أساساً بنظرية الاحتمال Theory of probability مما يتطلب من الباحث الإلمام بقدر وافر من المعلومات الإحصائية والرياضية حتى يستطيع استخدام الأساليب الإحصائية في تقدير الحجم الأمثل للعينة .

ويعتمد هذا الاتجاه على تحديد العوامل (المتغيرات) التي يتوقف عليها حجم العينة واعتبارها دلائل رئيسية أو مؤشرات أساسية لهذا الغرض وهو أمر يغفله الاتجاه الأول تماما كما يعتمد هذا الاتجاه على توفير بعض المعلومات عن حجم ومعالم المجتمع الأصلي عن طريق العينات التجريبية أو الاسترشادية .

وتتمثل أهم العوامل والمتغيرات الرئيسية المحددة لحجم العينة في نسبة الخطأ المسموح به (أو درجة الدقة أو الثقة) ، ومعامل التشتت (أو الانحراف المعياري) بين مفردات العينة أو المجتمع أن أمكن ، والاختلاف النسبي بين المتوسط الحسابي للعينة ومتوسط المجتمع (32)

ثامناً : الأساليب الإحصائية لتحديد حجم العينة

يلجأ الباحثون إلى تحديد حجم العينة باستخدام الأساليب الإحصائية تفادياً لتحديده بطريقة تعسفية تثير الانتقادات وتقلل من أهمية العمل العلمي والجهد الذي يبذله الباحث . ويواجه الباحث احتمالين أساسيين عندما يسعى إلى تحديد حجم العينة إحصائياً .

أ - هو ألا يكون على علم بعدد مفردات المجتمع الإحصائي .

ب - هو أن يكون على علم بعدد مفردات المجتمع الإحصائي .

وأخيراً قد تقترح جهة ما على الباحث أن يجرى دراسته على عدد معين من المبحوثين وفي هذه الحالة يميل الباحث إلى تحديد نسبة الخطأ في هذه العينة ليتأكد من أهمية البيانات التي سيحصل عليها ومدى تمثيل تلك العينة للمجتمع الذي سحبت منه. تحديد حجم العينة من مجتمع إحصائي غير معلوم

في كثير من الأحيان لا يجد الباحث بيانات وافية عن عدد أفراد المجتمع الإحصائي الذين سيسحب من بينهم عينة البحث ، وذلك لكبر حجم هذا المجتمع ، أو لعدم توافر إحصاءات رسمية عن أفرادها وفي هذه الحالة يمكن تحديد حجم العينة المطلوب سحبها من مجتمع إحصائي كبير باستخدام المعادلة الآتية :-

$$\text{حجم العينة (ن)} = \frac{Z^2}{\chi^2_m} \times \text{ف (1 - ف)}$$

حيث :

Z : القيمة المعيارية عند مستوى ثقة معين وهى فى جميع أحوال الأبحاث تأخذ أحد رقمين هما :

$Z = 1.96$ عند مستوى دلالة 0.05 أو مستوى ثقة 95%

$Z = 2.58$ عند مستوى دلالة 0.01 أو مستوى ثقة 95%

خ م : الخطأ المعيارى المسموح به وهو أيضاً فى جميع أحوال الأبحاث يأخذ أحد قيمتين هما :

خ م = 0.05 عند مستوى ثقة 95%

خ م = 0.01 عند مستوى ثقة 95%

ف : هى درجة الاختلاف بين مفردات المجتمع الإحصائى وقد اصطلح العلماء على وضعها بقيمة ثابتة أى أن قيم ف = 0.5 دائماً .

تحديد حجم العينة من مجتمع إحصائى معلوم .

عند حساب حجم العينة من مجتمع إحصائى معلوم ، بمعنى إننا نعرف عدد الأفراد الذين يتكون منهم ذلك المجتمع ، فإننا نتبع الخطوات التالية :

- نحسب حجم العينة على أساس أن حجم المجتمع الإحصائى غير معلوم وذلك بالعملية الحسابية السابقة .

- نقوم بعد ذلك بتصحيح حجم العينة ، وذلك باستخدام معادلة تصحيح العينة كالاتى :-

معادلة تصحيح حجم العينة :

$$\text{حجم العينة} = \frac{n_1}{\frac{n_1 - 1}{n} + 1}$$

حيث :

n_1 : حجم العينة من مجتمع غير معلوم .

n : حجم المجتمع الإحصائي .

ومن الملاحظ أن حجم العينة من مجتمع إحصائي معلوم العدد أقل من حجم العينة من مجتمع إحصائي غير معلوم العدد ، ولذلك فإن استخدام معادلة تصحيح معامل حجم العينة قد أسهم في ترشيد حجم العينة المناسب للبحث وإن كان الفرق بين حجمي العينتين ليس كبيراً على ما يبدو .

وفي نهاية الأمر يمكن القول بأن اختيار حجم عينة البحث لم يعد يمثل في الوقت الحالي مشكلة عويصة . فالحاسب الآلي يمكن أن يقدم لنا مقترحات عديدة بهذا الخصوص ، كما أن بعض العلماء قد بذلوا جهداً طيباً في إعداد جداول جاهزة للتغلب على المشكلات المتعلقة بتلك المسألة من ذلك على سبيل المثال جدول حجوم العينات الذي أعده Hush وزميله Backstorm والذي طوره وأضاف إليه Cole (33)

34) التحليل الإحصائي باستعمال العينات

البيانات الإحصائية هي الأساس للتخطيط الاقتصادي والاجتماعي ولكل البرامج الإنمائية ولتتخذى القرار . وبدخول عصر العولمة ومع الوضع الراهن للدول النامية أصبحت هناك ضرورة ملحة ومتزايدة للإحصاءات بوجه عام وللبيانات الاقتصادية والاجتماعية بوجه خاص . واستجابة لهذه الحاجة تسعى، كثيراً من دول العالم النامي إلى النهوض بالعمل الإحصائي إلى المستوى اللازم للوفاء باحتياجات المسئولين عن التخطيط للتنمية الاقتصادية والاجتماعية. كما تبذل جهوداً كبرى في تدريب الكوادر الوطنية القادرة على القيام بإجراء التعدادات والمسوحات وغيرها من نشاطات جمع البيانات وإجراء التحليل بشكل فعال .

"فالإحصاء (سواء تعداداً أو مسحاً بالعينه) من حيث اللغة هو الإلمام بكل المفردات التي يشملها المجتمع الذي نريد دراسته ومعرفة أو صاف كل مفردة في هذا المجتمع معرفة دقيقة ومحددة بالأعداد . أما علمياً هو عبارة عن تصوير رقمي للواقع في المجتمعات المطلوبة دراستها) المجتمعات البشرية أو غير البشرية)" مثال ذلك تعداد السكان ومسح ميزانية الأسرة فهو تصوير رقمي لأحوال السكان ومستوى معيشتهم على الترتيب.

وننوه بداية بأنه يمكن تقسيم الدراسات والبحوث من حيث المجال أي من حيث درجة الشمول لمفردات المجتمع الأصلي إلى بحوث شاملة وبحوث بطريقة العينات. فالبحث الشامل هو الذي ندرس فيه حاله جميع أفراد المجتمع موضوع البحث بهذه الطريقة إذا كان الغرض منه هو الحصر وذلك مثل تعداد السكان التعداد الزراعي.. الخ. وهذا يتطلب تكلفة كبيرة من الوقت والمال والجهد. إن البحث بطريقة العينة فهو الذي نبحث فيه حاله جزء معين (أو نسبة معينة) من أفراد المجتمع الأصلي ثم نقوم بعد ذلك بتعميم نتائج الدراسة على المجتمع كله بتكلفة أقل كثيراً من البحث الشامل .

ومن أمثلة أهم البحوث بالعينة التي تجري على أرض الواقع تلك البحوث التي تستخدم مسوح ميزانية الأسرة وبحوث القوى العاملة والتي عادة ما تجريها الحكومات أو المؤسسات الدولية أو الإقليمية. كما تشمل مسوحات التجارة والصناعة والمساكن وأبحاث استطلاع الرأي .

مميزات البحث بالمعينة وأهميته

واضح أنه من فوائد البحث عن طريق العينة هو اختصار الوقت والجهد اللازمين لإتمام البحث وبالتالي اقتصاد التكاليف. كما يمكن الحصول بسهولة على الردود الكاملة الدقيقة إذا ما استخدمنا جزء من المجتمع الكلي. كما أنه سهل تتبع غير المستجيبين في حالة

البحث بالعينة بينما يكون ذلك صعباً في حالة الحصر الشامل. ويمكن الحصول على بيانات أكثر من أفراد العينة، وحجمها وتلخيصها وتحليلها على وجه السرعة .

كما تُساعدنا بحوث العينات لمعرفة الدقة التي نتجت عن إجراء حصر شامل والطريقة المثلى هي أن نختار عينة وندرسها دراسة دقيقة وبمقارنة نتائجها مع نتائج التعداد يُمكننا معرفة مدى دقة نتائج الحصر الشامل .

مما سبق يتضح مدى أهمية استخدام العينات والدور الذي تلعبه في الدراسات الكثيرة في مُختلف الميادين، وفي الحقيقة أن استخدام الحصر الشامل أصبح لا يُعني عن استخدام العينة في نفس الوقت فإن تحليل النتائج التي نحصل عليها من تعداد شامل تحتاج إلى وقت طويل وقد تضيع الحكمة من التعداد أو تقل الاستفادة منه إذا ما انتظرنا حتى يتم تحليل النتائج. وفي هذه الحالة يتحتم علينا أن نأخذ عينة ونقوم بتحليل نتائجها لتعطي فكرة عن النتائج النهائية .

أهداف المعاينة

يعد تحديد الهدف الرئيسي للمعاينة أو المشكلة المراد دراستها تحديداً واضحاً، وتحديد أهدافه التفصيلية ربما تكون ذا أهمية كبيرة وذلك لتحديد البيانات المطلوب جمعها واستخدامها من قبل

الباحث لكسب ثقة المدى بالبيانات. وبعد ذلك نضع التصميمات المختلفة والممكنة عن طريق الأسئلة المراد الحصول على إجابات عليها. مثلاً يُمكن صياغة أهداف البحث بالسؤال التالي، هل هناك صلة بين التعليم والوعي المصرفي .

إن الغرض الأول من إجراء بحث أو تجربة هو إيجاد إجابات لأسئلة مُعينة حتى نحصل على أساس سليم للتنبؤ عادةً ومنه نستطيع اتخاذ إجراء على نتائج العينة فلا بد أن نترجمها ونفسرها بطريقة تُعطي أقصى الفوائد فنوجد التقديرات الإحصائية المختلفة لمعالم المجتمع كما أنه لا بد من قياس دقة هذه التقديرات. وإن من أهم المسائل في تصميم العينات هو الانتهاء إلى معادلة أو معادلات لحساب التقديرات من بيانات العينة وهذه المعادلة أو المعادلات المختارة لا بد أن تحتفظ بكل المعلومات الخاصة بالمجتمع التي حصلنا عليها من العينة ولا بد من استخدام البيانات لأقصى حد مُمكن .

والتقديرات التي نحصل عليها هي قيم تقريبية لمعالم المجتمع الحقيقية التي نبحث عنها والسؤال المهم هو هل الفرق بين التقدير المحسوب من العينة والقيم الحقيقية للمجتمع صغيراً صغراً كافياً يجعلنا نعتمد على التقدير في دراستنا للمجتمع؟ هنا إذا تم اختيار العينة وحصلنا على التقدير بطرق تعتمد على نظرية الاحتمالات فإنه يُمكننا أن نُقدر دقة هذا التقدير . وإذا كان التقدير

يختلف عن القيمة الحقيقية فإن الباحث يُعاني بعض الخسائر إذا ما استخلص نتائجه على أساس هذا التقدير .

وتقديرات معالم المجتمع التي يُمكن الحصول عليها من العينة كثيرة وأبسطها الوسط الحسابي لعينة عشوائية فمن المعروف بأن هذا المتوسط يُعطى تقديراً لمتوسط المجتمع الذي سحبت منه العينة غير أنه لن يكون مُساوياً تماماً لمتوسط المجتمع وذلك يرجع إلى أخطاء المعاينة. ومن التقديرات الأخرى لمعالم المجتمع التي نحصل عليها من المعاينة هي التباين والتفرع والالتواء.

العوامل التي تحدد حجم العينة

عند اختيار عينة من مجتمع الدراسة تثار قضيتان : الأولى تتعلق بحجم العينة والثانية تتصل بالطريقة التي يتم بها سحب العينة وفى هذا الفصل سنهتم فقط بالأساليب الإحصائية لتحديد حجم العينة :

أولاً : العوامل التي تحدد حجم العينة :

- 1- حجم المجتمع الإحصائي الذي ستسحب منه العينة .
- 2- درجة الاختلاف بين مفردات المجتمع الإحصائي .
- 3- نسبة الخطأ المسموح به أو المقبول ودرجة الثقة التي يرغب الباحث فى توافرها فى النتائج التي يصل إليها من دراسته للعينة.

ثانياً : الأساليب الإحصائية لتحديد حجم العينة :

يلجأ الباحثون إلى تحديد حجم العينة باستخدام الأساليب الإحصائية تفادياً لتحديده بطريقة تعسفية تثير الانتقادات وتقلل من أهمية العمل العلمي والجهد الذي يبذله الباحث ، ويواجه الباحث احتماليين أساسيين عندما يسعى إلى تحديد حجم العينة إحصائياً :

الأول : هو ألا يكون على علم بعدد مفردات المجتمع الإحصائي .

الثاني : هو أن يكون على علم بعدد مفردات المجتمع الإحصائي .

وأخيراً قد تقترح جهة معينة على الباحث أن يجرى دراسته على عدد معين من المبحوثين وفي هذه الحالة يميل الباحث إلى تحديد نسبة الخطأ في هذه العينة ليتأكد من أهمية البيانات التي سيحصل عليها ومن مدى تمثيل تلك العينة للمجتمع الذي سحبت منه .

وفيما يلي نتناول أساليب تحديد حجم العينة في ظل كل احتمال من الاحتمالات السابقة :

1- تحديد حجم العينة من مجتمع إحصائي غير معلوم

في كثير من الأحيان لا يجد الباحث بيانات وافية عن عدد أفراد المجتمع الإحصائي الذي سيسحب من بينهم عينة البحث وذلك لكبر حجم هذا المجتمع أو لعدم توافر إحصاءات رسمية عن أفراده وفي هذه الحالة يمكن تحديد حجم العينة المطلوب سحبها من مجتمع إحصائي كبير أو غير معلوم باستخدام المعادلة التالية :

$$\text{حجم العينة (ن)} = \frac{Z^2}{\chi^2_m} \times \text{ف (1 - ف)}$$

حيث :

Z : القيمة المعيارية عند مستوى ثقة معين وهي فى جميع أحوال الأبحاث تأخذ أحد رقمين هما :

$$Z = 1.96 \text{ عند مستوى دلالة } 0.05 \text{ أو مستوى ثقة } 95\%$$

$$Z = 2.58 \text{ عند مستوى دلالة } 0.01 \text{ أو مستوى ثقة } 95\%$$

خ م : الخطأ المعيارى المسموح به وهو أيضاً فى جميع أحوال الأبحاث يأخذ أحد قيمتين هما :

$$خ م = 0.05 \text{ عند مستوى ثقة } 95\%$$

$$خ م = 0.01 \text{ عند مستوى ثقة } 95\%$$

ف : هى درجة الاختلاف بين مفردات المجتمع الإحصائى وقد اصطلح العلماء على وضعها بقيمة ثابتة أى أن قيم ف = 0.5 دائماً .

مثال :

أوجد حجم عينة من مجتمع احصائى غير معلوم إذا علمت أن مستوى الثقة المطلوب توافره فى البيانات هو 95% ؟

الحل :

$$\text{حجم العينة (ن)} = \frac{Z^2}{خ م^2} \times ف (1 - ف)$$

$$^2(1.96)$$

$$\text{حجم العينة (ن)} = \frac{0.5 \times (0.5 - 1)}{(0.05)^2}$$

$$\text{حجم العينة (ن)} = 0.25 \times 1536.64 = 384.16 \text{ مفردة .}$$

نقرب الكسر لأقرب رقم صحيح فيصبح :

$$\text{حجم العينة} = 385 \text{ مفردة .}$$

2- تحديد حجم العينة من مجتمع إحصائي معلوم

عند حساب حجم العينة من مجتمع إحصائي معلوم بمعنى أننا نعرف عدد الأفراد الذين يتكون منهم ذلك المجتمع فإننا نتبع الخطوات التالية :

(أ) نحسب حجم العينة على أساس أن حجم المجتمع الإحصائي غير معلوم من المعادلة التالية :

$$\text{حجم العينة (ن)} = \frac{Z^2}{\chi^2_m} \times (f - 1)$$

حيث :

Z : القيمة المعيارية عند مستوى ثقة معين وهي في جميع أحوال الأبحاث تأخذ أحد رقمين هما :

$$Z = 1.96 \text{ عند مستوى دلالة } 0.05 \text{ أو مستوى ثقة } 95\%$$

$$Z = 2.58 \text{ عند مستوى دلالة } 0.01 \text{ أو مستوى ثقة } 95\%$$

خ م : الخطأ المعياري المسموح به وهو أيضاً في جميع أحوال الأبحاث يأخذ أحد قيمتين هما :

خ م = 0.05 عند مستوى ثقة 95%

خ م = 0.01 عند مستوى ثقة 95%

ف : هي درجة الاختلاف بين مفردات المجتمع الإحصائي وقد اصطلح العلماء على وضعها بقيمة ثابتة أي أن قيم ف = 0.5 دائماً .

(ب) نقوم بعد ذلك بتصحيح حجم العينة وذلك باستخدام معادلة تصحيح حجم العينة كالتالي :

$$\text{حجم العينة} = \frac{n_1}{\frac{n_1 - 1}{n} + 1}$$

حيث :

n_1 : حجم العينة من مجتمع غير معلوم كما سيتم حسابها في الخطوة (أ) .

حيث n : حجم المجتمع الإحصائي .

مثال :

أوجد حجم عينة من مجتمع إحصائي حجمه 15000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره في البيانات هو 95% ؟

الحل :

الخطوة (أ) حساب حجم العينة من مجتمع غير معلوم :

$$\text{حجم العينة (ن}_1) = \frac{Z^2}{\chi^2} \times \text{ف (ف - 1)}$$

$$\text{حجم العينة (ن}_1) = \frac{(1.96)^2}{(0.05)^2} \times 0.5 (0.5 - 1)$$

$$\text{حجم العينة (ن}_1) = 1536.64 \times 0.25 = 384.16 \text{ مفردة .}$$

نقرب الكسر لأقرب رقم صحيح فيصبح :

$$\text{حجم العينة (ن}_1) = 385 \text{ مفردة .}$$

الخطوة (ب) تصحيح حجم العينة :

$$\text{حجم العينة} = \frac{\text{ن}_1}{\frac{1 - \text{ن}_1}{\text{ن}} + 1}$$

$$\text{حجم العينة} = \frac{385}{\frac{1 - 385}{15000} + 1}$$

$$\text{حجم العينة} = 375.24 \text{ مفردة}$$

نقرب الكسر لأقرب رقم صحيح فيصبح :

$$\text{حجم العينة} = 376 \text{ مفردة .}$$

تحديد نسبة الخطأ في حجم العينة

قد يقرر الباحث إجراء دراسته على عدد معين من الأفراد وفى هذه الحالة التى يحدد فيها الباحث حجم العينة بطريقة تخمينية أو يفرض عليه من الجهة المستفيدة بالدراسة نجده يميل إلى محاولة تحديد نسبة الخطأ فى حجم العينة حتى يطمئن إلى أن البيانات سيحصل عليها والى أن النتائج التى سيتوصل إليها تتمتع بمستوى عالى من الثقة .

وتتحدد نسبة الخطأ فى العينة وفق المعادلة التالية :

$$\text{خطأ العينة} = Z \times \sqrt{\frac{f(1-f)}{n}}$$

حيث :

Z : القيمة المعيارية عند مستوى ثقة معين وهى فى جميع أحوال الأبحاث تأخذ أحد رقمين هما :

$$Z = 1.96 \text{ عند مستوى دلالة } 0.05 \text{ أو مستوى ثقة } 95\%$$

$$Z = 2.58 \text{ عند مستوى دلالة } 0.01 \text{ أو مستوى ثقة } 95\%$$

f : هى درجة الاختلاف بين مفردات المجتمع الإحصائى وقد اصطلح العلماء على وضعها بقيمة ثابتة أى أن قيم $f = 0.5$ دائماً .

ن : عدد مفردات العينة .

مثال :

إذا كان لدينا عينة حجمها 600 مفردة سحبت من مجتمع احصائي كبير العدد فما هي نسبة الخطأ المتوقعة في هذه العينة عند مستوى ثقة بنسبة 95% في البيانات .

الحل :

$$\text{خطأ العينة} = Z \times \sqrt{\frac{f(1-f)}{n}}$$

$$\text{خطأ العينة} = 1.96 \times \sqrt{\frac{0.5(1-0.5)}{600}}$$

$$\text{خطأ العينة} = 1.96 \times 0.0204 = 0.04$$

$$\text{نسبة الخطأ المعياري المتوقعة} = 100 \times 0.04 = 4\%$$

تارين

1- أوجد حجم عينة من مجتمع إحصائي حجمه 20000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره في البيانات هو 95% ؟

2- أوجد حجم عينة من مجتمع إحصائي حجمه 30000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره في البيانات هو 95% ؟

3- أوجد حجم عينة من مجتمع إحصائي حجمه 50000 مفردة إذا علمت أن مستوى الثقة المطلوب توافره في البيانات هو 95% ؟

4- إذا كان لدينا عينة حجمها 800 مفردة سحبت من مجتمع إحصائي كبير العدد فما هي نسبة الخطأ المتوقعة في هذه العينة عند مستوى ثقة بنسبة 95% في البيانات .

5- إذا كان لدينا عينة حجمها 400 مفردة سحبت من مجتمع إحصائي كبير العدد فما هي نسبة الخطأ المتوقعة في هذه العينة عند مستوى ثقة بنسبة 95% في البيانات .

المراجع

- 1- اعتماد علام ، يسرى رسلان ، أساسيات الإحصاء الإجماعى ، دار الثقافة للنشر والتوزيع ، ص 287 .
- 2- فاروق عبد العظيم ، مختار الهانسى ، محمد على محمد ، مبادئ الإحصاء ، دار المعرفة الجامعية ، ص 9 .
- 3- فاروق عبد العظيم ، وآخرون ، مرجع سابق ، ص 9 .
- 4- فتحى عبد العزيز أبو راضى ، مبادئ الإحصاء الإجماعى، دار المعرفة الجامعية ، ص 17.
- 5- حسن محمد حسن ، أساسيات الإحصاء وتطبيقاته ، دار المعرفة الجامعية ، 1992، ص 29 .
- 6- اعتماد علام ، مرجع سابق ، ص 306 .
- 7- المرجع السابق ، ص 307 .
- 8- المرجع السابق ، ص 291 .
- 9- فاروق عبد العظيم وآخرون ، مرجع سابق ، ص 14 .
- 10- اعتماد علام وآخرون ، مرجع سابق ، ص 292 .
- 11- حسن محمد حسن ، مرجع السابق ، ص 30 .
- 12- اعتماد علام ، مرجع سابق ، ص 296 .
- 13- فاروق عبد العظيم وآخرون ، مرجع سابق ، ص 17 .
- 14- حسن محمد حسن ، أساسيات الإحصاء وتطبيقاته ، مرجع سابق ، ص 30 .
- 15- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 297 .

- 16- حسن محمد حسن ، أساسيات الإحصاء وتطبيقاته ، مرجع سابق ، ص ص 29-33 .
- 17- فاروق عبد العظيم وآخرون ، مرجع سابق ، ص 17 .
- 18- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 388 .
- 19- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص 40 .
- 20- فاروق عبد العظيم وآخرون ، مرجع سابق ، ص ص 9-10 .
- 21- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص ص 39-40 .
- 22- حسن محمد حسن ، أساسيات الإحصاء وتطبيقاته ، مرجع سابق ، ص 33 .
- 23- فاروق عبد العظيم وآخرون ، مرجع سابق ، ص ، 10
- 24- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص 44 .
- 25- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص 388 .
- 26- فاروق عبد العظيم وآخرون ، مرجع سابق ، ص ص 11-12 .
- 27- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص ص 289-290 .
- 28- فاروق عبد العظيم وآخرون ، مرجع سابق ، ص ص 12-13 .

- 29- اعتماد علام ، يسرى رسلان ، مرجع سابق ، ص ص
290 - 291 .
- 30- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص ص
10 - 11 .
- 31- حسن محمد حسن ، مبادئ الإحصاء الإجماعى ، دار
المعرفة الجامعية ، ص ص 47 - 50 .
- 32- فتحى عبد العزيز أبو راضى ، مرجع سابق ، ص ص 19
- 20 .
- 33- حسن محمد حسن ، مبادئ الإحصاء الإجماعى ، مرجع
سابق ، ص 69 .

34 - http://www.arab-api.org/course13/c13_1.htm

الفصل الرابع تبويب وعرض البيانات

أولاً : العرض الجدولى للبيانات الإحصائية .

- تبويب البيانات الخام فى جدول تكرارى بسيط .
- تبويب البيانات فى جدول تكرارى ذو فئات .
- تبويب البيانات فى الجدول التكرارى المتجمع الصاعد .
- تبويب البيانات فى الجدول التكرارى المتجمع الهابط .
- الجدول المزدوج .

ثانياً : العرض البيانى للبيانات الإحصائية .

- العرض البيانى للبيانات الغير مبوبة .
 1. طريقة الأعمدة البيانية البسيطة .
 2. طريقة المنحنى البيانى البسيط .
 3. طريقة الخط البيانى المنكسر .
 4. طريقة الدائرة البيانية .
 5. طريقة الأعمدة البيانية المتلاصقة .
 6. طريقة الأعمدة البيانية المجزأة .
- العرض البيانى للبيانات الغير مبوبة .
 1. المدرج التكرارى .
 2. المضلع التكرارى .
 3. المنحنى التكرارى .

تبويب البيانات :

يقصد بتبويب البيانات عرض هذه البيانات (البيانات الخام) فى جداول مناسبة وذلك حتى يمكن تلخيصها وفهمها واستيعابها واستنتاج النتائج منها ومقارنتها بغيرها من البيانات ، كما يسهل الرجوع إليها فى صورة جداول دون الاطلاع على الاستثمارات الأصلية التى قد تحمل أسماء أصحابها مما يخل بمبدأ سرية البيانات الإحصائية .

كما يعتبر عرض وتبويب البيانات الإحصائية الخطوة الثانية (بعد تجميع هذه البيانات الخام) فى مفهوم التحليل الإحصائي، ويلجأ الباحث إلى حصر وتصنيف هذه البيانات وعرضها بطريقة مختصرة تساعد على فهمها وتحليلها إحصائياً للتعرف عليها ووصفها ومقارنتها بغيرها من الظواهر ، والخروج ببعض المدلولات الإحصائية عن مجتمع الدراسة .

عرض البيانات :

تتوقف طريقة عرض البيانات على نوع هذه البيانات وعلى الحقائق المطلوب إبرازها. وهناك طريقتان أساسيتان لعرض وتبويب البيانات الإحصائية وهما :

أولاً : العرض الجدولى للبيانات الإحصائية :

بعد عملية تبويب وتعيين الصفات التى تميز المفردات ، ترصد النتائج فى جداول مناسبة توضح الشكل النهائى للمجموعات المميزة وتسمى هذه العملية التى يتم تجميع البيانات فى مجموعات مميزة ومتجانسة بعملية التصنيف وتصنف البيانات الإحصائية بوجه عام وفقاً لإحدى القواعد التالية :

1- تصنيف جغرافى

2- تصنيف تاريخى أو زمنى .

3- تصنيف نوعى أو وصفى .

4- تصنيف كمى .

ويمكن التمييز بين مجموعة أشكال من الجداول الإحصائية نذكرها فيما يلى :

تبويب البيانات الخام فى جدول تكرارى بسيط :

والمقصود بالجدول البسيط هو ذلك الجدول الذى يتم وضع قيم الدرجات فيه مرتبة ترتيباً تصاعدياً فى عموده الأول أما العمود

الثانى فيسمى بعمود التكرار ويرصد فيه عدد مرات تكرار كل درجة أو حدث .

مثال :

البيانات التالية هى درجات حصل عليها عشرون طالبا فى مادة الإحصاء الاجتماعى بالفرقة الأولى قسم الاجتماع فى امتحان

نهاية العام :

12 11 15 14 12 10 15 13 12 10
14 10 13 12 15 13 12 10 12 15

والمطلوب تبويب هذه البيانات فى جدول توزيع تكرارى بسيط ؟

الحل :

يتم ترتيب البيانات دون تكرار تصاعديا ثم وضع هذه البيانات فى العمود الأول من الجدول وتسمى (س) ثم وضع عدد مرات التكرار باستخدام العلامات فى العمود الثانى أما العمود الثالث فيمثل التكرار ويرمز له بالرمز (ك) .

ك	العلامات	س
4	////	10
1	/	11
6	/ ////	12
3	///	13
2	//	14
4	////	15
20	مج	

مثال :

البيانات التالية هي تقديرات 20 طالباً في مادة الإحصاء بالفرقة الأولى لقسم الاجتماع في العام الجامعي 2006/2005 والمطلوب هو وضع هذه البيانات في جدول بسيط ؟

جيد جداً	جيد	مقبول	جيد جداً	جيد	مقبول	جيد	جيد	مقبول	جيد
مقبول	جيد	جيد	ممتاز	جيد	مقبول	جيد جداً	ممتاز	جيد	ممتاز

الحل :

التكرار	التقدير
5	مقبول
9	جيد
3	جيد جداً
3	ممتاز
20	المجموع

تبويب البيانات في جدول تكراري ذو فئات :

قبل التعرض إلى إعداد هذا الجدول سنقوم أولاً بالتعرف على معنى الفئات وطرق كتابتها .

المقصود بالفئات :

الفئة هي مجموعة من البيانات متشابهة إلى حد كبير جداً في الصفات ، وفي حالة زيادة عدد البيانات الخام التي يتم الحصول عليها من الاستبيان لا يمكن استخدام الجداول البسيطة في التعبير عن هذه الحالات وإلا سنحتاج إلى مئات الصفحات ، وإنما يتم تقسيم البيانات إلى مجموعات متقاربة ومتشابهة في الصفات تسمى فئات .

طرق كتابة الفئات :

يوجد عدة طرق لكتابة الفئات هي :

الطريقة الأولى :

نذكر كلا من الحد الأدنى والحد الأعلى للفئة كما بالجدول التالي :

ك	ف
5	20-10
20	30-20
50	40-30
25	50-40

وتنطق الفئة الأولى مثلاً (من 20 إلى 30) وليس (20 شرطة 30) وهذه الطريقة معيبة لأن نهاية الفئة الأولى هي نفسها بداية

الفئة الثانية وهكذا وفى هذه الحالة لا نعرف إلى أى فئة ينتمى هذا الرقم .

الطريقة الثانية :

نذكر كلا من الحد الأدنى والحد الأعلى للفئة ولكن نقوم بترك فاصل مقدراه الواحد الصحيح بين نهاية الفئة الأولى وبداية الفئة الثانية وهكذا كما بالجدول التالى .

ك	ف
5	19-10
20	29-20
50	39-30
25	49-40

ويعاب على هذه الطريقة أنها لا تصلح فى حالة البيانات التى تحتوى على كسور .

الطريقة الثالثة :

نذكر الحد الأدنى فقط للفئة ونضع بعده شرطة وتنطق الفئة الأولى مثلاً (10 إلى أقل من 20) وهذه الطريقة تصلح لكافة الظواهر.

ك	ف
5	-10
20	-20
50	-30
25	-40

الطريقة الرابعة :

نذكر الحد الأعلى فقط للفئة ونضع قبله شرطة وتنطق الفئة الأولى مثلاً (أكثر من صفر الى 20) وهذه الطريقة تصلح لكافة الظواهر أيضاً ولكنها أقل شيوعاً .

ك	ف
5	20-
20	30-
50	40-
25	50-

خطوات بناء جدول التوزيع التكراري ذو الفئات :

1- حساب المدى = أكبر قيمة - أصغر قيمة

2- حساب عدد الفئات = 3.3 لو (ن)

3- حساب طول الفئة = المدى / عدد الفئات

4- اختيار بداية الفئة الأولى أى الحد الأدنى لها مساوى لأقل قيمة موجودة بالبيانات أو أقل بقليل منها فمثلاً تكون من الأرقام الصفرية لتسهيل الحسابات بعد ذلك .

5- بناء الجدول ووضع العلامات التى تمثل التكرار .

مثال :

قام باحث بجمع بيانات تمثل درجات اختبار مادة الحاسب الآلى لخمسين طالباً من طلاب المرحلة الثانية من الثانوية العامة فى

الجدول التالى :

57	42	51	55	70
53	63	47	60	45
55	82	39	65	33
42	65	61	58	64
55	45	53	52	50
39	63	59	36	25
64	54	49	45	65
78	52	41	42	75
26	48	25	35	30
88	46	55	40	20

والمطلوب هو إعداد جدول توزيع تكرارى ذو فئات للجدول

السابق؟

الحل :

• المدى = أكبر قيمة - أصغر قيمة = 88 - 20 = 68

• عدد الفئات = $3.3 \times \text{نو} = 3.3 \times \text{نو} = (50)$

$5.6 = 1.699 \times 3.3 =$

• نقرب عدد الفئات لأقرب رقم صحيح فتكون

عدد الفئات = 7

• طول الفئة = المدى / عدد الفئات = $9.7 = 7 / 68$

• تقرب طول الفئة لأقرب رقم صحيح فتصبح

طول الفئة = 10

• نختار بداية الفئة الأولى وهو أصغر رقم = 20

• نبدأ في بناء الجدول كالتالي :

التكرار	العلامات	الفئات
4	////	-20
6	/ ////	-30
12	// //// ////	-40
14	//// //// ////	-50
9	//// ////	-60
3	///	-70
2	//	90-80
50	المجموع	

تبويب البيانات في الجدول التكراري
المتجمع الصاعد :

ويقصد بالتكرار المتجمع الصاعد هو تجميع تكرار كل فئة على

جميع التكرارات السابقة لها بحيث يكون مجموع التكرار

التصاعدي للفئة الأخيرة مساوي لمجموع التكرارات .

مثال :

من نفس بيانات المثال السابق كون جدول التكرار المتجمع

الصاعد.

الحل :

بنفس الخطوات السابقة نكون جدول التوزيع التكرارى ذو الفئات
ومنه نكون جدول التوزيع التكرارى المتجمع الصاعد كالتالى :

حدود الفئات	التكرار المتجمع الصاعد (ك.م.ص)
أقل من 20	صفر
أقل من 30	4
أقل من 40	10
أقل من 50	22
أقل من 60	36
أقل من 70	45
أقل من 80	48
أقل من 90	50

تبويب البيانات في الجدول التكرارى المتجمع الهابط :

ويقصد بالتكرار المتجمع الهابط هو تجميع تكرار كل فئة على
جميع التكرارات التالية لها بحيث يكون مجموع التكرار التنازلى
للفئة الأولى مساوى لمجموع التكرارات .

مثال :

من نفس بيانات المثال السابق كون جدول التكرار المتجمع الهابط

الحل :

بنفس الخطوات السابقة نكون جدول التوزيع التكرارى ذو الفئات
ومنه نكون جدول التوزيع التكرارى المتجمع الصاعد كالتالى :

حدود الفئات	التكرار المتجمع الهابط (ك.م.هـ)
20 فأكثر	50
30 فأكثر	46
40 فأكثر	40
50 فأكثر	28
60 فأكثر	14
70 فأكثر	5
80 فأكثر	2
90 فأكثر	صفر

الجدول المزدوج

وهو الجدول الذى يربط بين متغيرين فى نفس الوقت وكل متغير
منهم له فئاته فيتم بناؤه بإتباع عدة خطوات هى :

- 1- تحديد المتغيرين
- 2- تحديد المتغير المستقل والمتغير التابع
- 3- تحديد فئات كل من المتغيرين

4- تكوين الجدول بحيث يحتل المتغير المستقل أعلى الجدول أى يكون أفقياً أما المتغير التابع فيحتل الجزء الأسفل أى يكون عمودياً.

5- وضع العلامات التى تمثل التكرار.

6- إعادة كتابة الجدول بالأرقام .

مثال :

الجدول التالى يوضح البيانات التى حصل باحث فى دراسة بين النوع و مشاهدة البرامج التعليمية لمجموعة من طلاب الصف الثالث الثانوي على النحو التالى :

النوع	مشاهدة البرامج	النوع	مشاهدة البرامج
ذكر	يشاهد	ذكر	لا يشاهد
ذكر	يشاهد	أنثى	لا يشاهد
أنثى	يشاهد	أنثى	لا يشاهد
ذكر	لا يشاهد	أنثى	يشاهد
أنثى	يشاهد	ذكر	يشاهد
أنثى	لا يشاهد	ذكر	يشاهد
أنثى	لا يشاهد	ذكر	لا يشاهد
ذكر	لا يشاهد	ذكر	لا يشاهد
ذكر	يشاهد	أنثى	يشاهد
أنثى	لا يشاهد	أنثى	لا يشاهد

والمطلوب تكوين الجدول المزدوج للعلاقة بين المتغيرين (النوع و مشاهدة البرامج التعليمية) ؟

الحل :

- 1- المتغيرين (النوع - مشاهدة البرامج التعليمية)
- 2- المتغير المستقل هو النوع والمتغير التابع هو مشاهدة البرامج التعليمية .
- 3- فئات المتغير النوع هي (ذكور - إناث)
- فئات المتغير مشاهدة البرامج التعليمية (يشاهد - لا يشاهد)
- 4- تكوين الجدول بحيث يحتل المتغير المستقل أعلى الجدول أى يكون أفقياً أما المتغير التابع فيحتل الجزء الأسفل أى يكون عمودياً .

كالتالى :

إناث	ذكور	النوع
		مشاهدة البرامج التعليمية
		يشاهد
		لا يشاهد

- 5- وضع العلامات .

إناث	ذكور	النوع
		مشاهدة البرامج التعليمية
////	////	يشاهد
/ ///	////	لا يشاهد

- 6- إعادة كتابة الجدول بالأرقام .

مج	إناث	ذكور	النوع
----	------	------	-------

			مشاهدة البرامج التعليمية
9	4	5	يشاهد
11	6	5	لا يشاهد
20	10	10	مج

ثانياً : العرض البياني للبيانات الإحصائية

يعتبر العرض البياني للبيانات الإحصائية بمثابة تلخيص للبيانات الإحصائية في شكل يسهل منه استيعاب خصائص موضوع بحث الدراسة ، وتختلف طرق عرض البيانات المبوبة عن البيانات الغير مبوبة ، وسنتعرض لكل منها بالتفصيل فيما يلي :-

أولاً : العرض البياني للبيانات الغير مبوبة :

والمقصود بالبيانات الغير مبوبة تلك البيانات المفردة أي لا يوجد بها فئات وهناك عدة طرق لعرض البيانات الغير مبوبة .

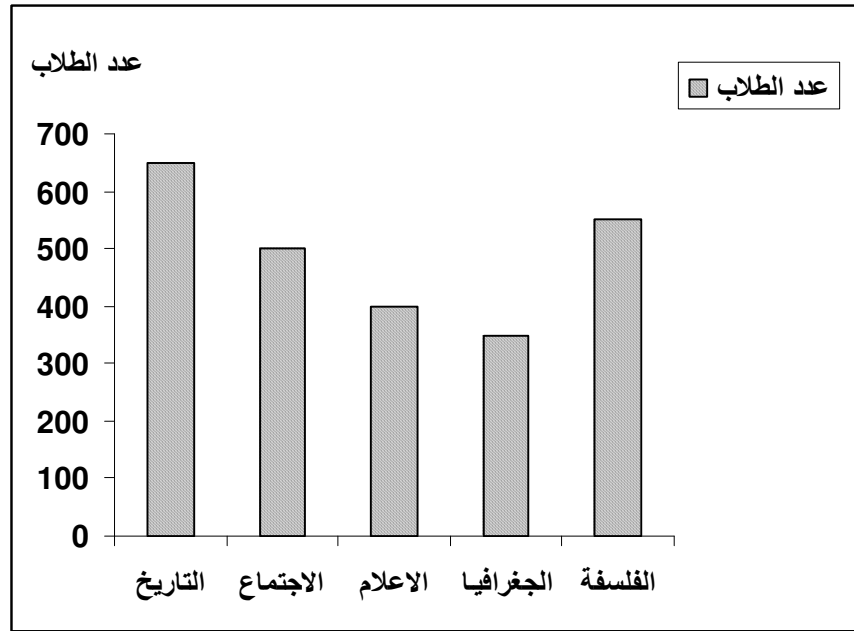
(1) طريقة الأعمدة البيانية البسيطة :

وفي هذه الطريقة يمثل محور السينات قيم المتغير أما محور الصادات يمثل القيمة المقابلة لقيمة المتغير ويتم رسم عمود حول المتغير وارتفاعه يمثل قيمة المتغير .

مثال :

الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية الآداب
جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة
الأعمدة البيانية البسيطة ؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550



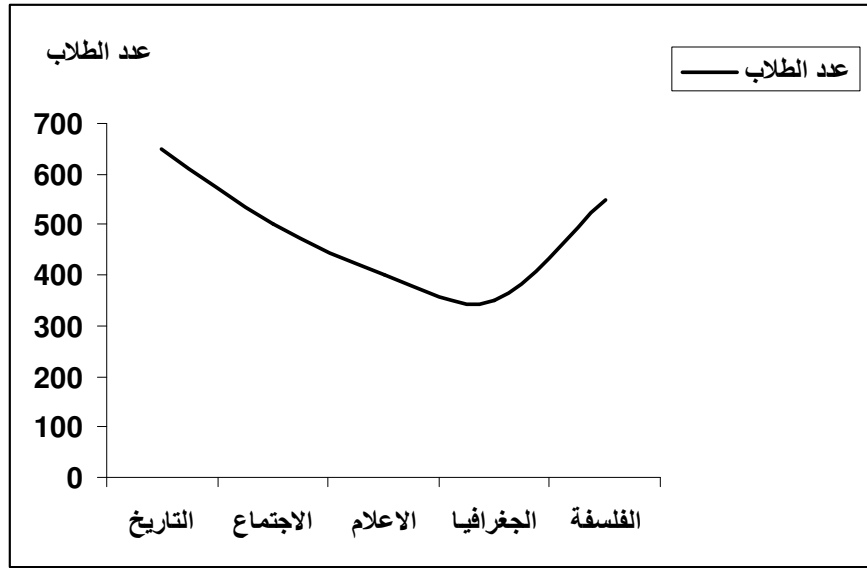
(2) طريقة المنحنى البياني البسيط :

وفي هذه الطريقة يمثل محور السينات المتغير أما محور الصادات
يمثل قيمة المتغير ويتم توقيع نقاط بين كل قيمة من قيم المتغير
على محور السينات والقيمة المقابلة على محور الصادات ثم يتم
توصيل تلك النقاط بخط منحنى باليد .

مثال :

الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة المنحنى البياني البسيطة؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550



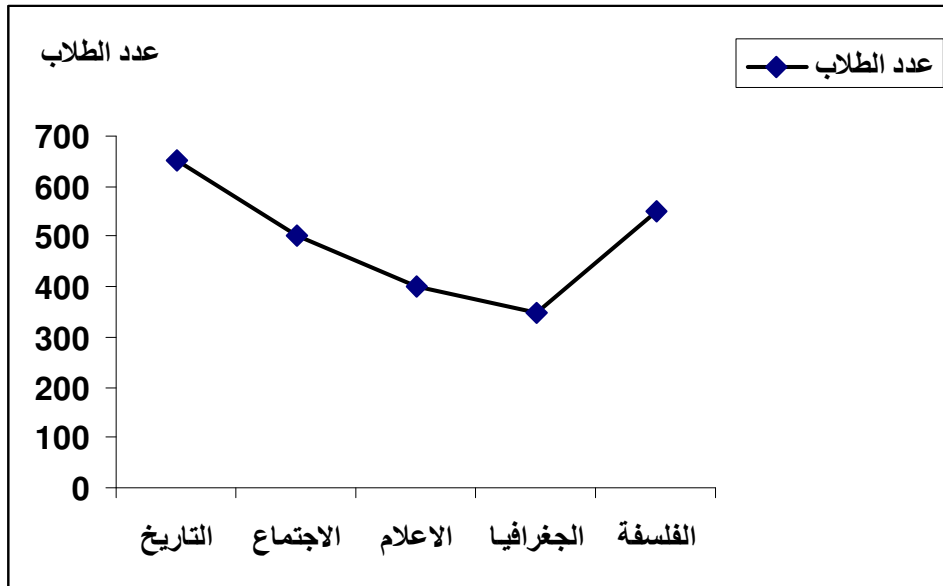
(3) طريقة الخط البياني المنكسر :

وفى هذه الطريقة يمثل محور السينات المتغير أما محور الصادات يمثل قيمة المتغير ويتم توقيع نقاط بين كل قيمة من قيم المتغير على محور السينات والقيمة المقابلة على محور الصادات ثم يتم توصيل تلك النقاط بخط منكسر باستخدام المسطرة .

مثال :

الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية الآداب
جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة
الخط البياني المنكسر؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550



(4) طريقة الدائرة البيانية :

وفى هذه الطريقة يتم رسم دائرة ثم نحسب زاوية قطاع كل قيمة
على حدة ونقوم برسم تلك الزاوية داخل الدائرة حتى تنتهى
الدائرة.

ونحسب زاوية قطاع الجزء من العلاقة :

$$\text{زاوية قطاع الجزء} = \frac{\text{التكرار الفعلى للجزء}}{\text{مجموع التكرارات}} \times 360$$

مثال :

الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية الآداب
جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة
الدائرة البيانية ؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	650	500	400	350	550

الحل :

نحسب مجموع التكرارات = $550+350+400+500+650$

مجموع التكرارات = 2450

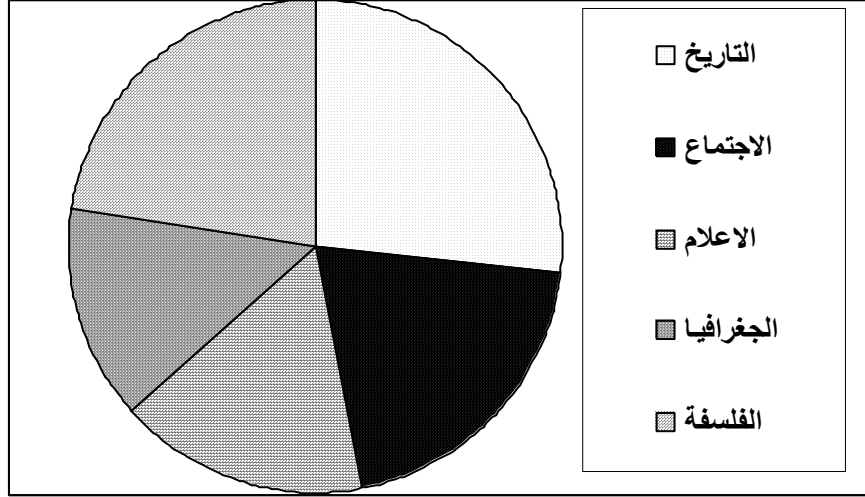
$$^5 95.5 = 360 \times \frac{650}{2450} = \text{زاوية قطاع التاريخ}$$

$$^5 73.5 = 360 \times \frac{500}{2450} = \text{زاوية قطاع الاجتماع}$$

$$^5 58.7 = 360 \times \frac{400}{2450} = \text{زاوية قطاع الإعلام}$$

$$^5 51.4 = 360 \times \frac{350}{2450} = \text{زاوية قطاع الجغرافيا}$$

$$^5 80.8 = 360 \times \frac{550}{2450} = \text{زاوية قطاع الفلسفة}$$



(5) طريقة الأعمدة البيانية المتلاصقة :

تسمى هذه الطريقة أيضا بطريقة الأعمدة البيانية المتجاورة وهي تشبه طريقة العمدة البيانية البسيطة ولكن يتم رسم عدد من الأعمدة متلاصقة يمثل كل منهم احد قيم المتغير .

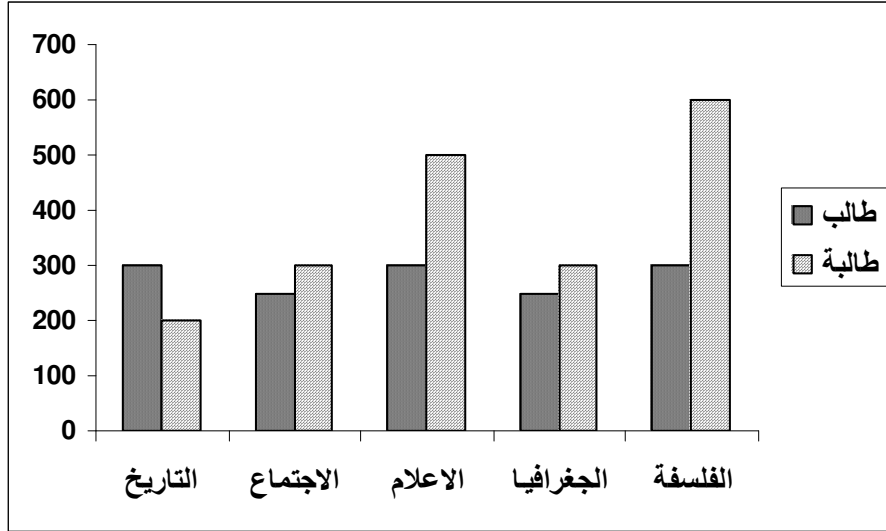
مثال :

الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية المتلاصقة ؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
طالب	300	250	300	250	300

600	300	500	300	200	طالبة
-----	-----	-----	-----	-----	-------

الحل :



(6) طريقة الأعمدة البيانية المجزأة :

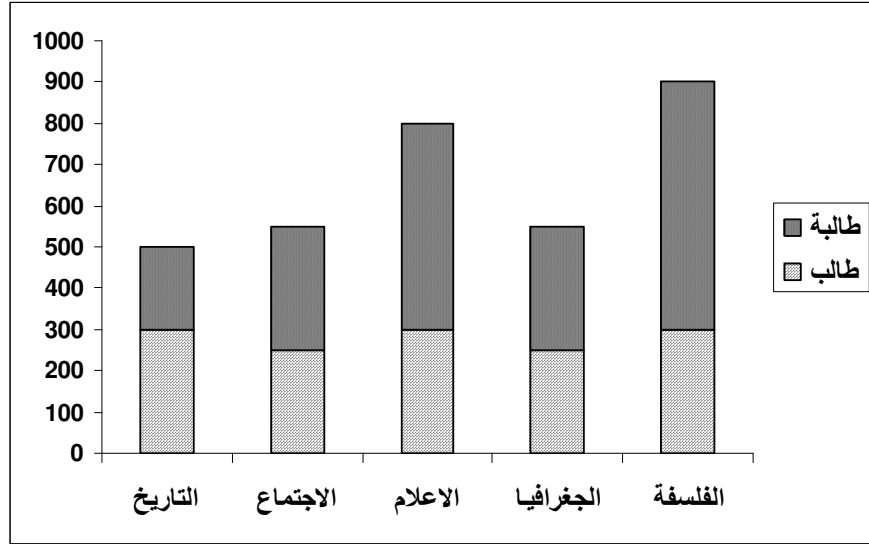
هذه الطريقة تشبه طريقة الأعمدة البيانية البسيطة ولكن يتم رسم عمود يمثل القيمة الأولى للمتغير ثم يليه أو يرتفعه عمود بباقي قيمة المتغير وتكون بادية العمود الثانى هى نهاية العمود الأول .

مثال :

الجدول التالى يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة المنصورة والمطلوب عرض هذه البيانات باستخدام طريقة الأعمدة البيانية المجزأة ؟

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
طالب	300	250	300	250	300
طالبة	200	300	500	300	600

الحل :



ثانياً : العرض البياني للبيانات المبوبة :

والمقصود بالبيانات المبوبة تلك البيانات المقسمة إلى فئات وهناك عدة طرق لعرض البيانات المبوبة .

(1) المدرج التكراري :

أحد طرق عرض البيانات المبوبة حيث يتم تخصيص عمود لكل فئة وتكرارها ، بحيث يكون طول الفئة هي قاعدة العمود والتكرار هو ارتفاع العمود ، ويفضل ترك فراغ كاف قبل الفئة الأولى وفراغ آخر بعد الفئة الأخيرة ، أما بالنسبة لمنتصف العمود فيكون هو مركز الفئة .

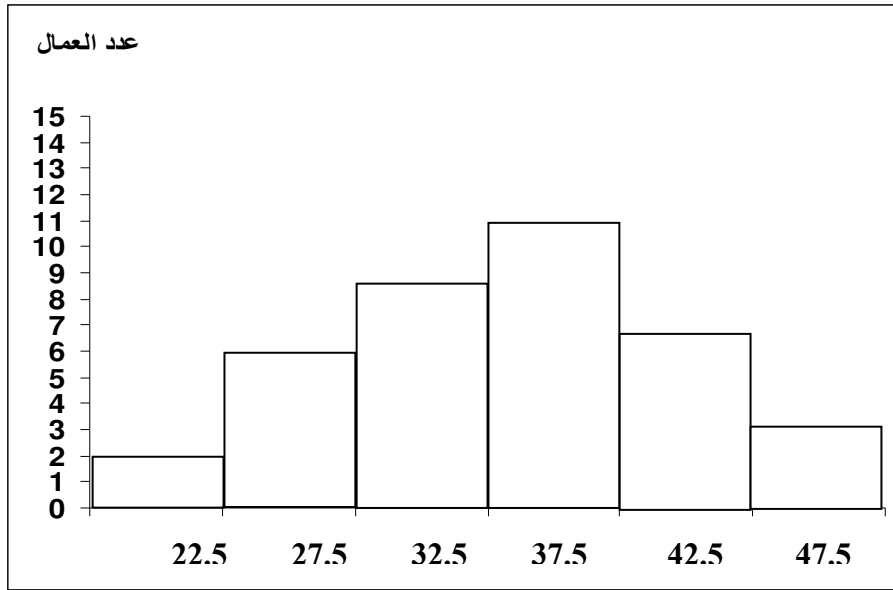
مثال :

اعرض لهذا الجدول بيانياً باستخدام المدرج التكراري ؟

فئات العمر	-20	-25	-30	-35	-40	-45
عدد العمال	2	6	9	11	7	3

الحل :

مركز الفئة	ك	ف
22.5	2	-20
27.5	6	-25
32.5	9	-30
37.5	11	-35
42.5	7	-40
47.5	3	-45



(2) المضلع التكرارى :

تخصص لكل فئة وتكرارها نقطة ، بحيث يكون الاحداثى السينى لها هو مركز الفئة بينما الاحداثى الصادى لها هو التكرار ، نفترض فئة سابقة للفئة الاولى وفئة لاحقة للفئة الأخيرة وتكرار

كل منهما صفر ، ثم نوصل كل نقطتين متتاليتين بخط مستقيم
بالمسطرة .

ملحوظة :

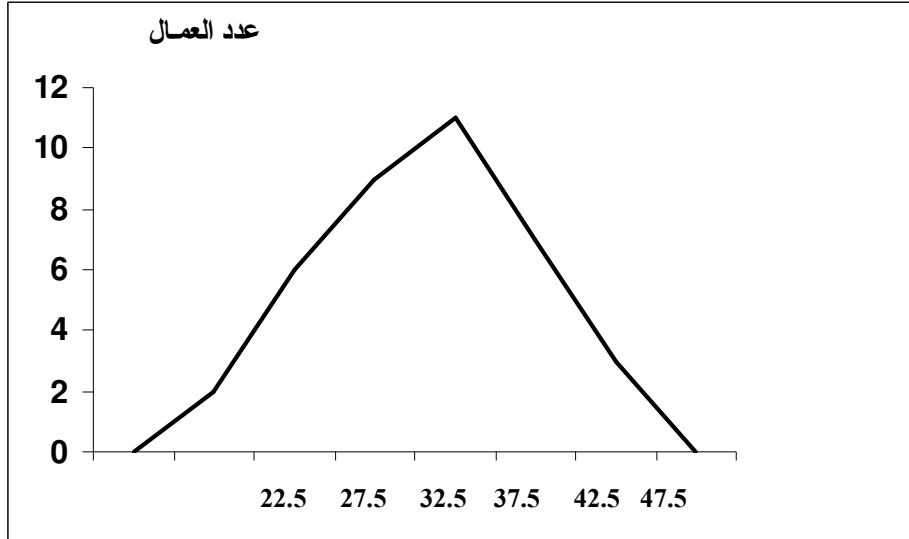
مساحة الشكل تحت المدرج التكرارى = مساحة الشكل تحت
المضلع التكرارى .

مثال :

اعرض لهذا الجدول بيانياً باستخدام المضلع التكرارى ؟

فئات العمر	-20	-25	-30	-35	-40	-45
عدد العمال	2	6	9	11	7	3

الحل :



(3) المنحنى التكرارى :

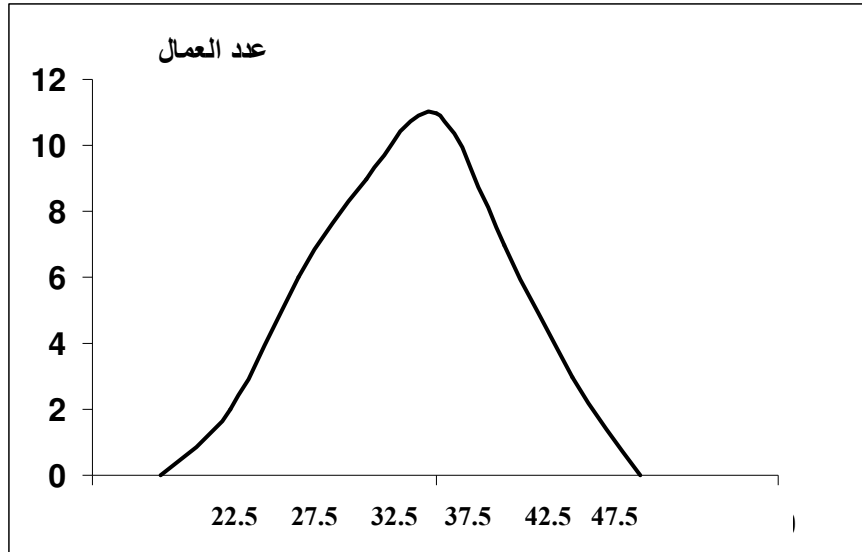
بعد رصد النقاط كما فى الطريقة السابقة نوصل كل نقطتين متتاليتين بمنحنى باليد .

مثال :

اعرض لهذا الجدول بيانياً باستخدام المنحنى التكرارى ؟

فئات العمر	-20	-25	-30	-35	-40	-45
عدد العمال	2	6	9	11	7	3

الحل :



تارين

1- حصل عدد من الطلاب في مادة الإحصاء على الدرجات التالية :

5	4	4	5	3	4	2	3	1	2
3	7	4	1	6	3	2	5	3	4
7	3	2	6	5	3	4	2	4	1

المطلوب : تكوين جدول تكرارى بسيط لهذه الدرجات.

2- تمثل البيانات التالية تقديرات عشرون طالبا في مادة علم النفس والمطلوب وضعها في جدول تكرارى بسيط لتلك التقديرات .

ممتاز	مقبول	جيد جدا	مقبول	جيد
جيد جدا	جيد	ضعيف	جيد	مقبول
جيد	ممتاز	مقبول	ضعيف	جيد
جيد جدا	جيد	مقبول	جيد	مقبول

3- هذه درجات 50 طالبا في اختبار ذكاء ، والمطلوب وضع هذه الدرجات في جدول تكرارى للفئات .

28	39	33	40	27	55	37	35	37	25
29	28	51	29	51	22	36	44	29	34
32	47	38	25	20	41	36	15	42	33
14	18	34	16	10	46	33	27	27	15
16	27	21	24	17	19	36	19	21	46

4- الدرجات التالية تمثل درجات 50 طالبا في أحد الاختبارات:

5	6	5	7	5	6	6	4	5	4
6	6	5	6	6	7	9	8	7	5
5	3	3	5	4	9	7	8	6	7
5	8	8	6	7	7	6	7	7	6
4	6	6	7	6	4	7	7	8	5

والمطلوب : وضع هذه الدرجات في جدول تكرارى للفئات .

5- حصل 80 طالبا في اختبار ذكاء على الدرجات التالية:

46	38	30	20	11	46	23	46	45	18
47	39	33	25	29	49	28	13	36	25
50	43	32	21	19	51	25	15	48	16
49	41	35	27	13	37	29	27	55	37
51	45	21	23	18	50	27	17	12	48
52	42	37	26	14	38	26	14	28	50
53	44	34	22	28	47	30	16	26	36
48	40	31	29	12	35	24	22	20	19

والمطلوب :

- وضع هذه الدرجات في جدول تكرارى للفئات بحيث يكون عدد الفئات .
- تكوين جدول التكرار المتجمع الصاعد .
- تكوين جدول التكرار المتجمع الهابط .

6- الجدول التالي يمثل أعداد الكتب بمكتبة الكلية فى مجموعة من التخصصات :

التخصص	علم الاجتماع	علم النفس	التاريخ	اللغة العربية	الجغرافيا
عدد الكتب	550	350	400	600	300

والمطلوب عرض هذه الجدول بيانياً باستخدام الطرق التالية :

- الأعمدة البيانية البسيطة .
- الخط البياني .
- الخط المنكسر .
- الدائرة البيانية .

7- الجدول التالي يمثل أعداد الذكور والإناث ببعض إدارات أحد الهيئات الحكومية .

الإدارة	الشنون الإدارية	الصيانة	الإحصاء	المعاشات
عدد الذكور	10	20	30	10
عدد الإناث	20	5	60	50

والمطلوب عرض هذه الجدول بيانياً باستخدام الطرق التالية :

- الأعمدة البيانية المتلاصقة .
- الأعمدة البيانية المجزأة .

8- الجدول التالي يمثل فئات درجات مجموعة من الطلاب في اختبار
للتحصيل وتكراراتهم :

-40	-35	-30	-25	-20	-15	-10	-5	الفئات
7	6	5	12	9	8	13	10	التكرار

والمطلوب هو عرض هذا الجدول بيانياً باستخدام الطرق التالية :

- المدرج التكرارى .
- المضلع التكرارى .
- المنحنى التكرارى .

الفصل الخامس مقاييس النزعة المركزية

أولاً : الوسط الحسابي .

ثانياً : الوسيط .

ثالثاً : المنوال .

رابعاً : العلاقة بين الوسط والوسيط والمنوال .

خامساً : تحديد التواء التوزيع من مقاييس النزعة المركزية.

مقاييس النزعة المركزية

إن الأسلوب البياني في تحليل ودراسة الظواهر لتحديد الخصائص والاتجاهات والعلاقات ، يعتمد في دقته على دقة التمثيل البياني نفسه وبذلك ربما تختلف الخصائص من رسم إلى آخر لنفس الظاهرة، وعليه فإنه من الأفضل اللجوء إلى طرق القياس الكمي، حيث يستخدم الباحث الطريقة الرياضية في القياس.

فالهدف الأساسي من استخدام مقاييس النزعة المركزية ومقاييس التشتت هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها ومقدار تشتت البيانات حول هذا المركز (درجة تجانس البيانات) ومن خلال هذين المؤشرين يتمكن الباحث من فهم أبعاد الظاهرة قيد الدراسة.

ومن أهم مقاييس النزعة المركزية التي سنتعرض إليها بالدراسة الوسط الحسابي والوسيط والمنوال ، كما سنتعرض بالدراسة لحساب كل منهم من البيانات المفردة (الغير مبوبة) ومن البيانات المبوبة .

أولاً : الوسط الحسابي (المتوسط)

الوسط الحسابي لمجموعة من القيم هو القيمة التي لو أعطيت لكل مفردة في المجموعة لكان مجموع قيم المفردات الجديدة مساوٍ لمجموع قيم المتغيرات الأصلية .

ويعرف أيضا بأنه مجموع قيم المشاهدات مقسوماً على عددها

ويرمز له بالرمز (س /) أو بالرمز (م)

حساب الوسط الحسابي من البيانات الغير مبوبة
(المفردة)

يحسب المتوسط الحسابي من البيانات الغير مبوبة من العلاقة التالية:

$$\frac{\text{مج س}}{\text{ن}} = \text{م} = \text{س} /$$

حيث :-

س / = م = الوسط الحسابي

مج = مجموع

س = القيمة

ن = عدد الأفراد

مثال :-

احسب الوسط الحسابي لدرجات 8 طلاب في مادة الإحصاء والتي

كان بياناتهم كالتالي :

9 - 8 - 8 - 7 - 6 - 5 - 3 - 2

الحل :

$$6 \text{ درجات} = \frac{48}{8} = \frac{9+8+8+7+6+5+3+2}{8} = \text{س} /$$

حساب الوسط الحسابي من البيانات المبوبة
توجد ثلاث طرق لحساب المتوسط الحسابي من البيانات المبوبة
هي :

1- الوسط الحسابي بطريقة مراكز الفئات

$$\text{س} / = \frac{\text{مج} (\text{س} \times \text{ك})}{\text{مج} \text{ك}}$$

حيث :-

$\text{س} / = \text{الوسط الحسابي}$

$\text{مج} = \text{مجموع}$

$\text{س} = \text{مركز الفئة} = (\text{بداية الفئة} + \text{بداية الفئة التالية}) / 2$

$\text{ك} = \text{التكرار}$

مثال :

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسط الحسابي بطريقة مراكز الفئات .

فئات الدخل	-100	-200	-300	-400	-500	-600	700-800
عدد العمال	10	12	20	28	16	8	6

الحل :

نكون الجدول التالي :

س × ك	س	ك	ف
1500	150	10	-100
3000	250	12	-200
7000	350	20	-300
12600	450	28	-400
8800	550	16	-500
5200	650	8	-600
4500	750	6	800-700
42600	مج	100	مج

$$42600 = \frac{42600}{100} = \text{س} / 426 \text{ جنيه}$$

2- الوسط الحسابي بطريقة الانحرافات

$$\frac{\text{مج} (ح \times ك)}{\text{مج} ك} + أ = \text{س} /$$

حيث :-

س / = الوسط الحسابي

مج = مجموع

ح = الانحراف = س - أ

ك = التكرار

أ = مركز الفئة التي يقابلها أكبر تكرار

مثال :

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسط الحسابي بطريقة الانحرافات .

فئات الدخل	-100	-200	-300	-400	-500	-600	800-700
عدد العمال	10	12	20	28	16	8	6

الحل :

نكون الجدول التالي :

ف	ك	س	ح	ح × ك
-100	10	150	300-	3000-
-200	12	250	200-	2400-
-300	20	350	100-	2000-
-400	28	450	صفر	صفر
-500	16	550	100	1600
-600	8	650	200	1600

1800	300	750	6	800-700
2400-		مج	100	مج

$$426 \text{ جنیہ} = 24 - 450 = \frac{2400-}{100} + 450 = \text{س}$$

3- الوسط الحسابى بطريقة الانحرافات
المختصرة

$$س / أ = \frac{مج - (ح / ك \times ل)}{مج ك}$$

حيث :-

س / أ = الوسط الحسابى

مج = مجموع

ح / ك = الانحراف المختصر = (س - أ) / ل

ك = التكرار

أ = مركز الفئة التى يقابلها أكبر تكرار

ل = طول الفئة

مثال :

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسط الحسابى بطريقة الانحرافات المختصرة .

فئات الدخل	-100	-200	-300	-400	-500	-600	800-700
عدد العمال	10	12	20	28	16	8	6

الحل :

نكون الجدول التالي :

ف	ك	س	ح / ك	ح × ك
-100	10	150	3-	30-
-200	12	250	2-	24-
-300	20	350	1-	20-
-400	28	450	صفر	صفر
-500	16	550	1	16
-600	8	650	2	16
800-700	6	750	3	18
مج	100	مج		24-

$$426 = 24 - 450 = 100 \times \frac{24-}{100} + 450 = \text{س} /$$

$$\text{س} / = 426 \text{ جنيهه .}$$

ثانياً : الوسيط

يعرف الوسيط على أنه القيمة التي تتوسط مجموعة من القيم إذا رتبت ترتيباً تصاعدياً أو تنازلياً .

حساب الوسيط من البيانات الغير مبوبة (المفردة)

يعتمد حساب الوسيط من البيانات الغير مبوبة على عدد تلك البيانات فهناك حالتان هما :

(1) إذا كان عدد المفردات فردى (ن فردية)

يوجد رقم واحد يمثل الوسيط ويحسب ترتيبه من العلاقة:

$$2 / (1+n)$$

مثال :

احسب الوسيط من البيانات التالية

$$61 - 80 - 40 - 10 - 15 - 12 - 20$$

الحل :

نرتب تصاعدي أولاً :

$$80 \quad 61 \quad 40 \quad \boxed{20} \quad 15 \quad 12 \quad 10$$

نحسب ترتيب الوسيط = $(1 + 7) / 2 = 4$ ، ترتيب الوسيط هو الرابع .

الوسيط = 20 .

(2) إذا كان عدد المفردات زوجي (ن زوجيه)

يوجد رقمين يمثلان الوسيط ويحسب عن طريق إيجاد الوسط الحسابي لهما ويحسب ترتيبه من العلاقة :

$$\{ 1 + 2 / ن ، 2 / ن \}$$

مثال :

احسب الوسيط من البيانات التالية :

$$40 - 33 - 20 - 18 - 14 - 15 - 12 - 15$$

الحل :

نرتب تصاعدي أولاً :

$$40 \quad 33 \quad 20 \quad \boxed{18} \quad \boxed{15} \quad 15 \quad 14 \quad 12$$

نحسب ترتيب الوسيط = $(2/8 ، 1 + 2/8) = (4 ، 5) ،$
ترتيب الوسيط الرابع والخامس وقيمة الوسيط متوسط القيمتين اللتان ترتيبهما الرابع والخامس .

$$\text{الوسيط} = 2 / (18 + 15) = 16.5 .$$

حساب الوسيط من البيانات المبوبة
يوجد خمس طرق لحساب الوسيط من البيانات المبوبة هي :
1- الوسيط باستخدام الجدول التكراري
المتجمع الصاعد

$$\text{الوسيط} = \text{الحد الأدنى للفئة الوسيطة} + \frac{\text{ترتيب الوسيط} - \text{ك م ص السابق}}{\text{ك م ص اللاحق} - \text{ك م ص السابق}} \times \text{ل}$$

حيث :-

ترتيب الوسيط = مج ك / 2

ك م ص السابق = التكرار المتجمع الصاعد السابق للفئة الوسيطة

ك م ص اللاحق = التكرار المتجمع الصاعد اللاحق للفئة الوسيطة

ل = طول الفئة .

مثال :

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط باستخدام جدول التكرار المتجمع الصاعد .

70-60	-50	-40	-30	-20	فئات الدخل
10	30	100	40	20	عدد العمال

الحل :

نكون الجدول التالي :

ك م ص	الحدود الدنيا للفئات	ك	ف	
صفر	أقل من 20	20	-20	
20	أقل من 30	40	-30	
ك م ص السابق	أقل من 40	100	-40	الحد الأدنى
ك م ص اللاحق	أقل من 50	30	-50	الحد الأعلى
190	أقل من 60	10	70-60	
200	أقل من 70	200	مج	

ثم نحسب ترتيب الوسيط = $100 = 2/200$
ثم نبحت داخل عمود (ك م ص) عن القيمتين التي ينحصر
بينهما ترتيب الوسيط فنجد أن قيمة ترتيب الوسيط = $100 =$
محصورة بين ($160 - 60$) .

$$44 = 4 + 40 = \frac{400}{100} + 40 = 10 \times \frac{60 - 100}{60 - 160} + 40 = \text{الوسيط}$$

2- الوسيط باستخدام الجدول التكرارى المتجمع الهابط

ترتيب الوسيط - ك م هـ اللاحق
الوسيط = الحد الأعلى للفئة الوسيطة - $\frac{\text{ك م هـ اللاحق}}{\text{ك م هـ السابق - ك م هـ اللاحق}} \times \text{ل}$

حيث :-

ترتيب الوسيط = مج ك / 2

ك م هـ السابق = التكرار المتجمع الهابط السابق للفئة الوسيطة
ك م هـ اللاحق = التكرار المتجمع الهابط اللاحق للفئة الوسيطة
ل = طول الفئة .

مثال :

الجدول التالى يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط باستخدام جدول التكرار المتجمع الهابط .

فئات الدخل	-20	-30	-40	-50	70-60
عدد العمال	20	40	100	30	10

الحل :

نكون الجدول التالى :

ك م ص	الحدود العليا للفئات	ك	ف		
200	20 فأكثر	20	-20		
180	30 فأكثر	40	-30		
ك م هـ السابق	140	40 فأكثر	100	-40	الحد الأدنى
ك م هـ اللاحق	40	50 فأكثر	30	-50	الحد الأعلى
10	60 فأكثر	10	70-60		
صفر	70 فأكثر	200	مج		

ثم نحسب ترتيب الوسيط = $2/200 = 100$
ثم نبحث داخل عمود (ك م هـ) عن القيمتين التي ينحصر
بينهما ترتيب الوسيط فنجد أن 100 محصورة بين (140 - 40)

$$44 = 6 - 50 = \frac{600}{100} - 50 = 10 \times \frac{40 - 100}{40 - 140} - 50 = \text{الوسيط}$$

3- الوسيط بالرسم من الجدول التكراري
المتجمع الماعد
مثال :

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط بالرسم من جدول التكرار المتجمع الصاعد .

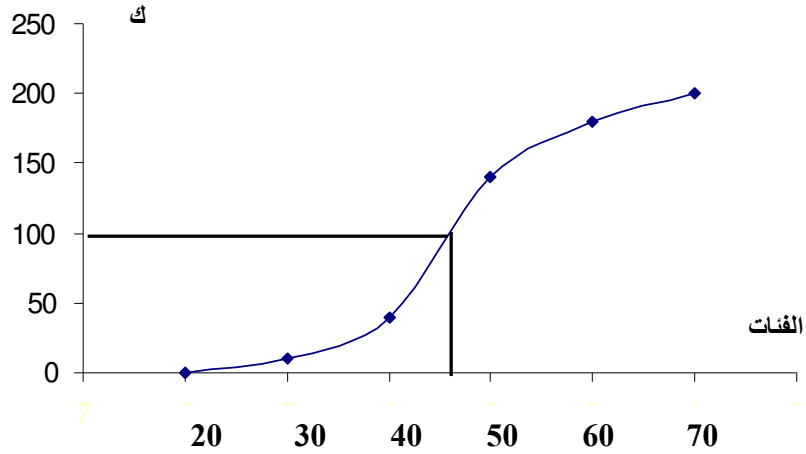
70-60	-50	-40	-30	-20	فئات الدخل
10	30	100	40	20	عدد العمال

الحل :

نكون الجدول التالي :

ك م ص	الحدود الدنيا للفئات
صفر	أقل من 20
20	أقل من 30
60	أقل من 40
160	أقل من 50
190	أقل من 60
200	أقل من 70

ثم نرسم حدود الفئات على محور السينات والتكرار المتجمع الصاعد على محور الصادات ونقوم بتوقيع جميع النقاط ونوصل بينها بخط منحنى باليد كما بالشكل .



ثم نحسب ترتيب الوسيط = $\frac{2}{200} = 2 / ك$ / $100 = 2/200 = 2 / ك$
 ونوقع هذه النقطة على محور الصادات ونرسم منها خط مستقيم
 ليقطع المنحنى في نقطة نقوم بإسقاط عمود من نقطة التقاطع
 ليصل إلى محور السينات لنحصل على قيمة الوسيط عندها .
 الوسيط = 44 .

4- الوسيط بالرسم من الجدول التكراري
المتجمع الهابط

مثال :

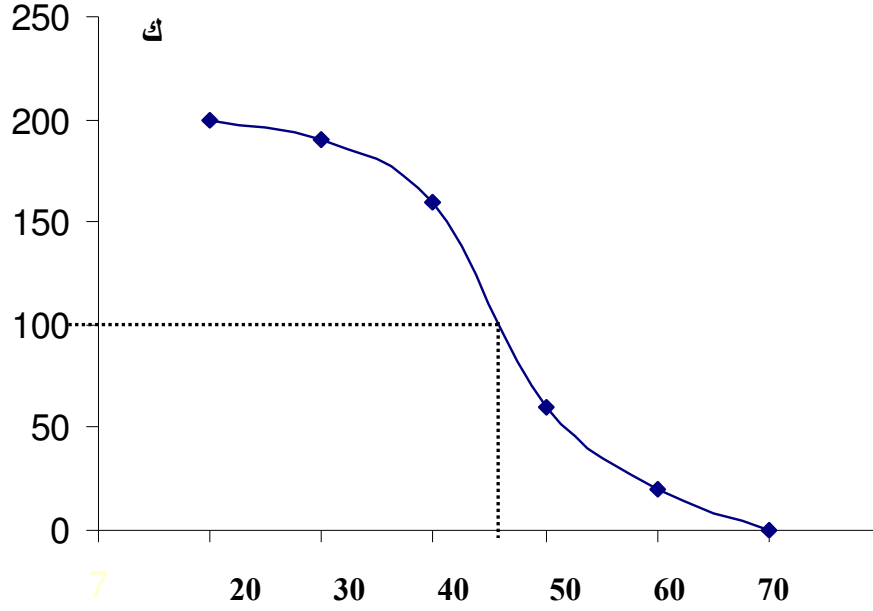
الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط بالرسم من جدول التكرار المتجمع الهابط .

70-60	-50	-40	-30	-20	فئات الدخل
10	30	100	40	20	عدد العمال

الحل :

نكون الجدول التالي :

ك م هـ	الحدود العليا للفئات
200	20 فأكثر
180	30 فأكثر
140	40 فأكثر
40	50 فأكثر
10	60 فأكثر
صفر	70 فأكثر



ثم نحسب ترتيب الوسيط = $\frac{2}{200} = 2 / ك$ = 100
ونوقع هذه النقطة على محور الصادات ونرسم منها خط مستقيم
ليقطع المنحنى في نقطة نقوم بإسقاط عمود من نقطة التقاطع
ليصل إلى محور السينات لنحصل على قيمة الوسيط عندها .
الوسيط = 44 .

5- الوسيط بالرسم من الجدول التكراري
المتجمع الصاعد والهابط معاً
مثال :

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال والمطلوب من واقع بيانات الجدول حساب الوسيط بالرسم من جدول التكرار المتجمع الصاعد والهابط معاً .

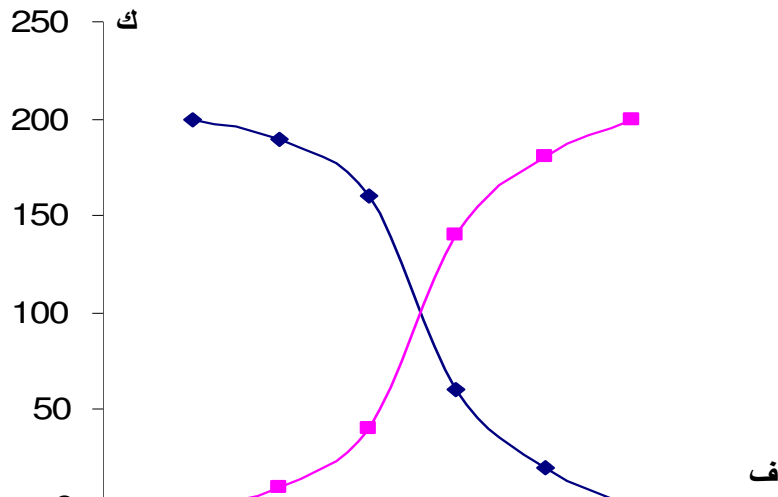
70-60	-50	-40	-30	-20	فئات الدخل
10	30	100	40	20	عدد العمال

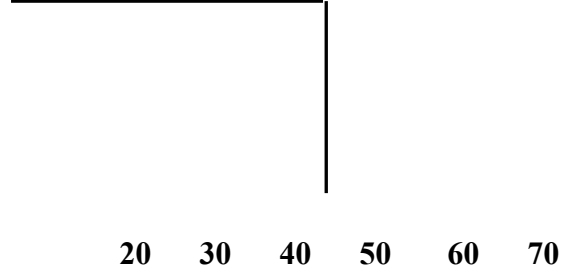
الحل :

نكون الجدولين الصاعد والهابط معاً :

الحدود العليا للفئات	ك م هـ
20 فأكثر	200
30 فأكثر	180
40 فأكثر	140
50 فأكثر	40
60 فأكثر	10
70 فأكثر	صفر

الحدود الدنيا للفئات	ك م ص
أقل من 20	صفر
أقل من 30	20
أقل من 40	60
أقل من 50	160
أقل من 60	190
أقل من 70	200





بعد رسم المنحنيين الصاعد والهابط يتقاطعا في نقطة هذه النقطة
لو قمنا بإسقاط عمود منها رأسياً على محور السينات نحصل على
قيمة الوسيط = 44 .

ولو قمنا برسم خط مستقيم أفقي من نقطة التقاطع ليقطع محور
الصادات نحصل على قيمة ترتيب الوسيط = 100 .

ثالثاً : المنوال :

المنوال هو القيمة الأكثر شيوعاً أو تكراراً .

حساب المنوال من البيانات الغير مبوبة
في حالة تكرار رقم واحد يتم اختياره كمنوال أما في حالة تكرار رقمين بنفس عدد مرات التكرار يتم اختيارهما معاً كمنوال أما إذا زاد أحدهما عن الآخر يتم اختيار ذو التكرار الأكبر وفي حالة عدم تكرار أي رقم يكون المنوال قيمته لاشيء أو لا يوجد منوال .

مثال : احسب المنوال في كل من الحالات

التالية :-

$$8 = \text{المنوال} \quad 12 - 8 - 10 - 8 - 9 - 8 - 7$$

$$10 = \text{المنوال} \quad 10 - 12 - 15 - 10 - 12 - 10$$

$$16 ، 15 = \text{المنوال} \quad 30 - 16 - 20 - 15 - 16 - 15$$

$$\text{المنوال} = \text{لا يوجد} \quad 60 - 50 - 140 - 40 - 30 - 20$$

حساب المنوال من البيانات المبوبة

يوجد أربعة طرق لحساب المنوال من البيانات المبوبة طريقتان

جبريتان وطريقتان بيانيتان وسنتناولهما بالشرح فيما يلي .

أولاً - المنوال بطريقة الفروق لبرسون .

$$\text{المنوال} = أ + \frac{\text{ف}_1}{\text{ف}_1 + \text{ف}_2} \times ل$$

حيث:

أ = الحد الأدنى للفئة المنوالية والمقصود بدايتها .

$$ف_1 = ك - ك_1$$

$$ف_2 = ك - ك_2$$

ك = تكرار الفئة المنوالية

ك₁ = تكرار الفئة التي تسبق الفئة المنوالية

ك₂ = تكرار الفئة التي تلي الفئة المنوالية

ل = طول الفئة

مثال :

أوجد المنوال بطريقة بيرسون من الجدول التالي :

فئات الدخل	-10	-20	-30	-40	-50	-60	80-70
عدد العمال	5	12	22	38	22	12	5

الحل :

	ك	ف	
	5	-10	
	12	-20	
ك ₁	22	-30	
ك	38	-40	أ
ك ₂	22	-50	
	12	-60	
	5	80-70	

ثم نحدد الفئة المنوالية من خلال أكبر رقم في عمود التكرار ثم نحدد الحد الأدنى لهذه الفئة وهو بدايتها وهو $أ = 40$ ، ثم نحدد (ك ، ك₁ ، ك₂) .

$$\text{نحسب ف}_1 = \text{ك} - \text{ك} = 1 \text{ ك} = 38 - 22 = 16$$

$$\text{نحسب ف}_2 = \text{ك} - \text{ك} = 2 \text{ ك} = 38 - 22 = 16$$

$$\text{نحسب ل} = 10$$

ثم نعوض في القانون :

$$\text{المنوال} = 40 + 10 \times \frac{16}{16 + 16}$$

$$\text{المنوال} = 5 + 40 = 45$$

ثانياً - المنوال بيانياً باستخدام

طريقة الفروق لبرسون .

مثال :

أوجد المنوال بيانياً باستخدام طريقة الفروق لبرسون من الجدول

التالى :

فئات الدخل	-10	-20	-30	-40	-50	-60	80-70
عدد العمال	5	12	22	38	22	12	5

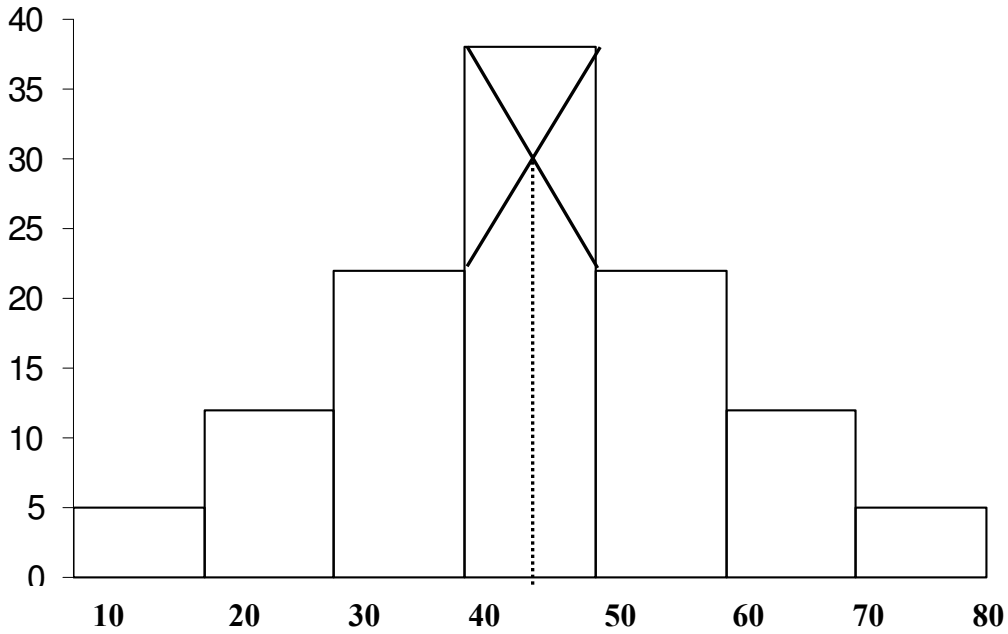
الحل :

نرسم الجدول السابق بالشكل التالى ثم نبحت عن أطول عمود

ونوصل حافتيه بحافتي العمود السابق والتالى فنحصل على تقاطع

هو المنوال .

$$\text{المنوال} = 45$$



ثالثاً : المنوال باستخدام طريقة الرافعة كينج .

$$\text{المنوال} = أ + \frac{1ك}{2ك + 1ك} \times ل$$

حيث:

أ = الحد الأدنى للفئة المنوالية والمقصود بدايتها .

1ك = تكرار الفئة التي تسبق الفئة المنوالية

2ك = تكرار الفئة التي تلي الفئة المنوالية

ل = طول الفئة

مثال :

أوجد المنوال بطريقة الرافعة كينج من الجدول التالي :

80-70	-60	-50	-40	-30	-20	-10	فئات الدخل
5	12	22	38	22	12	5	عدد العمال

الحل :

	ك	ف	
	5	-10	
	12	-20	
1ك	22	-30	
	38	-40	أ
2ك	22	-50	
	12	-60	
	5	80-70	

ثم نحدد الفئة المنوالية من خلال أكبر رقم في عمود التكرار ثم نحدد الحد الأدنى لهذه الفئة وهو بدايتها وهو $40 = أ$ ، ثم نحدد (ك1 ، ك2) .

$$22 = 1ك$$

$$22 = 2ك$$

$$10 = ل$$

ثم نعوض في القانون :

$$\text{المنوال} = 40 + \frac{22}{10} \times$$

$$22 + 22$$

$$45 = 5 + 40 = \text{المنوال}$$

رابعاً - المنوال بيانياً باستخدام
طريقة الرافعة كينج .

مثال :

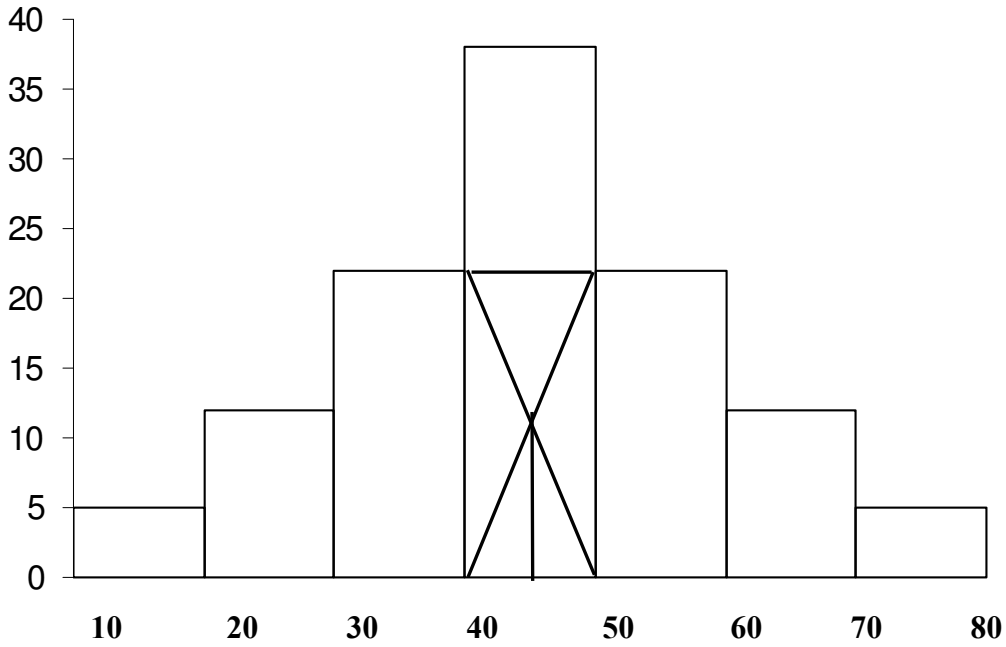
أوجد المنوال بيانياً باستخدام طريقة الرافعة كينج من الجدول
التالي:

فئات الدخل	-10	-20	-30	-40	-50	-60	80-70
عدد العمال	5	12	22	38	22	12	5

الحل :

نرسم الجدول السابق بالشكل التالي ثم نبحت عن أطول عمود
ونصل حافتيه بحافتي العمود السابق والتالي فنحصل على تقاطع
هو المنوال .

$$45 = \text{المنوال}$$



العلاقة بين الوسط والوسيط والمنوال

$$\text{المنوال} = 3 \times \text{الوسيط} - 2 \times \text{الوسط}$$

مثال :

إذا علمت أن قيمة الوسط = 5 وقيمة الوسيط = 10 احسب قيمة المنوال .

الحل :

$$\text{المنوال} = 3 \times \text{الوسيط} - 2 \times \text{الوسط}$$

$$\text{المنوال} = 3 \times 10 - 2 \times 5$$

$$\text{المنوال} = 30 - 10 = 20$$

تحديد التواء التوزيع مباشرة من مقاييس النزعة المركزية :

1- المنحنى معتدل التوزيع :

عندما يكون :

$$\text{المنوال} = \text{الوسيط} = \text{الوسط}$$

2- المنحنى ملتوى التواء موجب :

عندما يكون :

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

3- المنحنى ملتوى التواء سالب :

عندما يكون :

$$\text{المنوال} > \text{الوسيط} > \text{الوسط}$$

مثال

إذا علمت أن قيمة الوسط = 5 وقيمة الوسيط = 10 احسب قيمة المنوال ، ثم حدد نوع التواء التوزيع .

الحل :

$$\text{المنوال} = 3 \times \text{الوسيط} - 2 \times \text{الوسط}$$

$$\text{المنوال} = 3 \times 5 - 2 \times 10$$

$$\text{المنوال} = 15 - 20 = -5$$

نلاحظ أن

$$\text{المنوال} > \text{الوسيط} > \text{الوسط}$$

التوزيع ملتوي التواء سالب .

تمارين

1- احسب الوسط الحسابى والوسيط للدرجات الخام التالية :

$$10 - 4 - 17 - 8 - 2 - 3 - 5$$

من قيمة الوسط والوسيط احسب قيمة المنوال ثم حدد التواء التوزيع .

2- أوجد الوسط الحسابى والوسيط فى كل حالة من الحالات التالية ومنها أوجد قيمة المنوال ثم حدد التواء التوزيع .

$$8 - 11 - 9 - 12 - 7 \bullet$$

$$111 - 102 - 103 - 104 - 107 - 105 \bullet$$

$$24 - 20 - 9 - 18 - 35 - 3 - 39 - 3623 - 22 \bullet$$

3- احسب الوسيط والمنوال لكل حالة من الحالات التالية:

$$6 - 10 - 9 - 4 - 2 - 8 - 5 \bullet$$

$$8 - 5 - 4 - 10 - 7 - 9 - 6 \bullet$$

$$20 - 15 - 15 - 12 - 15 - 10 - 12 - 10 \bullet$$

$$70 - 60 - 40 - 20 - 30 - 25 - 20 \bullet$$

$$18 - 15 - 18 - 10 - 12 - 18 - 15 - 13 \bullet$$

4- الجدول التالى يمثل فئات الأجر الأسبوعى لعمال مصنع .

-10 12	-8	-6	-4	-2	الأجر الأسبوعى
30	50	70	40	10	عدد العمال

والمطلوب :

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد والهابط
- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون
- المنوال بطريقة الرافعة كينج
- المنوال بيانياً بطريقة الرافعة كينج

5- من واقع بيانات الجدول التالي :-

-60 70	-50	-40	-30	-20	ف
10	30	100	40	20	ك

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون
- المنوال بطريقة كينج .

6- من واقع بيانات الجدول التالي :-

ف	-100	-200	-300	-400	-500	-600	-700
ك	10	12	20	28	16	8	6

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد والهابط معاً
- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون
- المنوال باستخدام طريقة الرافعة كينج
- المنوال بيانياً باستخدام طريقة الرافعة كينج .

7- من واقع بيانات الجدول التالي :-

ك	ف
5	-10
12	-20
22	-30
38	-40
22	-50
12	-60
5	80-70
116	المجموع

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الاحرافات
- احسب الوسط الحسابي بطريقة الاحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحني التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحني التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحني التكرار المتجمع الصاعد والهابط معاً
- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون

8- من واقع بيانات الجدول التالي :-

ك	ف
2	-5
4	-10
6	-15
8	-20
10	-25
16	-30
40	-35
24	-40
14	-45
11	-50
5	60-55

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- المنوال بطريقة بيرسون

9- من واقع بيانات الجدول التالي :-

ك	ف
11	-40
20	-50
16	-60
28	-70
13	-80
12	100-90
100	المجموع

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الانحرافات
- احسب الوسط الحسابي بطريقة الانحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- المنوال بطريقة بيرسون
- المنوال بيانياً بطريقة بيرسون

10- من واقع بيانات الجدول التالي :-

ك	ف
10	-100
25	-200
13	-300
28	-400
15	-500
9	700-600
100	المجموع

- احسب الوسط الحسابي بطريقة مراكز الفئات
- احسب الوسط الحسابي بطريقة الاحرافات
- احسب الوسط الحسابي بطريقة الاحرافات المختصرة
- احسب الوسيط باستخدام جدول التكرار المتجمع الصاعد
- احسب الوسيط باستخدام جدول التكرار المتجمع الهابط
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد
- احسب الوسيط بيانياً باستخدام منحنى التكرار المتجمع الهابط
- المنوال بطريقة بيرسون
- المنوال بيانياً

الفصل السادس مقاييس التشتت

- أولاً : المدى .
- ثانياً : التباين والانحراف المعياري .
- ثالثاً : الانحراف المتوسط .
- رابعاً : الالتواء وتحديد اعتدالية التوزيع .

مقاييس التشتت

لا تعتبر مقاييس التمرکز كافية لوصف مجموعة من البيانات وصفاً كاملاً فقد تتساوى بعض العينات في الوسط الحسابي بالرغم من اختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات). فالعينات التالية ذات وسط حسابي واحد (8) ولكنها بلا شك تختلف عن بعضها.

8	8	8	8	8	عينة 1
11	16	6	3	4	عينة 2

فالوسط الحسابي يمثل مركز البيانات لكنه لا يبين مدى التفاف أو بعثرة البيانات حول هذا الوسط ، ولهذا لا بد من وجود مقياس آخر مع المقاييس المركزية لقياس درجة التجانس أو التشتت في داخل هذه البيانات.

إن الدرجة التي تتجه بها البيانات الرقمية للانتشار حول قيمة وسطى تسمى تشتت أو توزيع البيانات .
ومن أهم مقاييس التشتت المدى والتباين والانحراف المعياري والانحراف المتوسط .

أولاً : المدى

المدى هو الفرق بين أكبر قيمة وأصغر قيمة .
حساب المدى من البيانات الغر ميوية

المدى = أكبر قيمة - أصغر قيمة

مثال :

احسب المدى للبيانات التالية :

95 - 200 - 250 - 300 - 110 - 90 - 150 -

100 - 350 - 80

الحل :

نرتب القيم أولاً : (80-90-95-100-110-150-200-

250-300-350)

المدى = 350 - 80 = 270

حساب المدى من البيانات الميوية

المدى = الحد الأعلى للفئة الأخيرة - الحد الأدنى للفئة الأولى

مثال:

احسب المدى للجدول التالى :

الفئات	16-	20-	24-	28-	32-36
عدد المبحوثين	10	15	40	20	15

الحل :

المدى = الحد الأعلى للفئة الأخيرة - الحد الأدنى للفئة الأولى

المدى = 36 - 16 = 20

ثانياً : التباين والانحراف المعياري

يرمز للتباين بالرمز σ^2

بينما يرمز للانحراف المعياري بالرمز ع
أي أنه إذا تم حساب أحدهما فيمكن حساب الآخر لأن الانحراف
المعياري هو جذر التباين .

التباين من البيانات الغير مبوبة
هناك طريقتان لحساب التباين من البيانات الغير مبوبة:
الأولى : باستخدام القانون العام من
الدرجات الخام كالتالي

$$E^2 = \frac{\sum \text{مـجـس}^2}{\text{ن}} - \left(\frac{\sum \text{مـجـس}}{\text{ن}} \right)^2$$

مثال :

احسب التباين والانحراف المعياري للقيم التالية ومنه احسب
الانحراف المعياري لكل من المتغيرين س ، ص على حده .

18	19	19	21	23	س
15	14	18	19	19	ص

الحل :

نكون الجدول التالي :

ص ²	ص	س ²	س
361	19	529	23

361	19	441	21
324	18	361	19
196	14	361	19
225	15	314	18
1467	85	2016	100

ثم نعوض فى القانون العام لحساب التباين :

بالنسبة للمتغير (س)

$$E^2 = \frac{\sum (س \cdot س)}{ن} - \left(\frac{\sum س}{ن} \right)^2$$

$$3.2 = \frac{2016}{5} - \left(\frac{100}{5} \right)^2$$

وبالتالى فان قيمة تباين المتغير س = $E^2 = 3.2$
ومنها فان قيمة الانحراف المعياري = جذر التباين

$$E = \sqrt{3.2} = 1.78$$

بالنسبة للمتغير (ص)

$$E^2 = \frac{\sum (ص \cdot ص)}{ن} - \left(\frac{\sum ص}{ن} \right)^2$$

$$4.4 = \left[\frac{85}{5} \right] - \frac{1467}{5} = ع^2$$

وبالتالي فان قيمة تباين المتغير ص = $ع^2 = 4.4$
ومنها فان قيمة الانحراف المعياري = جذر التباين

$$2.1 = \sqrt{4.4} = ع$$

الثانية : باستخدام الطريقة المختصرة
"طريقة الانحرافات"

$$\frac{\text{مج ح}^2}{\text{ن}} = ع^2$$

حيث ح هو الانحراف = س - م س

مثال :

احسب الانحراف المعياري للقيم التالية :

20	19	13	48	19	32	22	17	35
----	----	----	----	----	----	----	----	----

الحل :

نكون الجدول التالي :

س	حس	ح ² س
35	10	100
17	8-	64
22	3-	9
32	7	49
19	6-	36
48	23	529
13	12-	144
19	6-	36
20	5-	25
225	-	992

حساب المتوسط :

$$م\text{س} = \frac{225}{9} = \frac{\text{مج س}}{1\text{ن}}$$

بعد حساب م_س نحسب عمود ح ومنه نحسب ح² ثم نعوض فى

القانون :

$$ع^2 = \frac{\text{مج ح}^2}{\text{ن}}$$

$$110.22 = \frac{992}{9} = \text{ع}^2$$

$$\sqrt{110.22} = \text{ع} = 10.5$$

التباين والانحراف المعياري من البيانات المبوبة :

يحسب التباين من البيانات المبوبة من العلاقة التالية :

$$\left\{ \left\{ \frac{\text{مج} (\text{ح} \times \text{ك})}{\text{مج ك}} \right\}^2 - \frac{\text{مج} (\text{ح}^2 \times \text{ك})}{\text{مج ك}} \right\} \times \text{ل}^2 = \text{ع}^2$$

حيث :

$$\text{ع}^2 = \text{التباين}$$

$$\text{ل} = \text{طول الفئة}$$

ح = الانحراف ويحسب عن طريق وضع صفر في عموده أمام الفئة التي يقابلها أكبر تكرار ثم نضع من أسفل (1 ، 2 ، 3 ، وهكذا) ومن أعلى نضع (-1 ، -2 ، -3 ، وهكذا)

مثال :

احسب الانحراف المعياري من الجدول التالي :

80-70	-60	-50	-40	-30	-20	-10	فئات الدخل
5	12	22	38	22	12	5	عدد العمال

الحل :

نكون الجدول التالي :

ف	ك	ح	ح × ك	ح × 2 ك
-10	5	3-	15-	225
-20	12	2-	24-	576
-30	22	1-	22-	484
-40	38	0	0	0
-50	22	1	22	484
-60	12	2	24	576
80-70	5	3	15	225
مج	116	-	صفر	2570

ثم نعوض فى القانون :

$$ع^2 = ل^2 \times \left\{ \frac{\text{مج} - (ح \times ك)}{\text{مج} - ك} \right\} - \frac{\text{مج} - (ح \times 2 ك)}{\text{مج} - ك}$$

$$\left\{ \left\{ \frac{0}{116} \right\} - \frac{2570}{116} \right\} \times 2(10) = ع^2$$

$$ع^2 = 2215.5$$

$$ع = \sqrt{2215.5} = 47.1$$

ثالثاً : الانحراف المتوسط

الانحراف المتوسط من البيانات الغير
مبوبة (المفردة)

$$\frac{\text{مجا |س - س'|}}{\text{ن}} = \text{الانحراف المتوسط}$$

حيث :

س = القيمة

س' = متوسط القيم

ن = عدد القيم

مثال :

لمجموعة البيانات التالية احسب الانحراف المتوسط:-

2 - 3 - 5 - 6 - 7 - 8 - 8 - 9

الحل :

$$6 = 8/48 = 8 / (9+8+8+7+6+5+3+2) = \text{نحسب س'}$$

نكون الجدول التالي :

س	س - س'
2	4
3	3
5	1
6	0
7	1
8	2
8	2
9	3

16	مج
----	----

$$2 = \frac{16}{8} = \text{الانحراف المتوسط}$$

الانحراف المتوسط من البيانات المبوبة

$$\text{الانحراف المتوسط} = \frac{\text{مج} (|س - س'| \times ك)}{\text{مج} ك}$$

مثال :

من بيانات الجدول التالي احسب الانحراف المتوسط :-

36-32	-28	-24	-20	-16	الفئات
15	20	40	15	10	عدد المبحوثين

الحل :

نكون الجدول التالي :

ف	ك	س	ح / ح	ح / ك	س - س'	س - س' × ك
-16	10	18	2-	20-	8.6	86
-20	15	22	1-	15-	4.6	69
-24	40	26	0	0	0.6	24
-28	20	30	1	20	3.4	68
36-32	15	34	2	30	7.4	111

358	مج	15	مج	100	مج
-----	----	----	----	-----	----

$$26.6 = 0.6 + 26 = 4 \times \frac{15}{100} + 26 = \text{س} /$$

$$3.58 = \frac{358}{100} = \text{الانحراف المتوسط}$$

الالتواء وتحديد اعتدالية التوزيع

$$\frac{3(m - w)}{ع} = \text{الالتواء}$$

حيث :

م : المتوسط

و : الوسيط

ع : الانحراف المعياري .

مثال :

حدد نوع التوزيع التالي :

$$10 - 40 - 60 - 50 - 20$$

الحل :

حساب المتوسط :

$$10+40+60+50+20 \quad \text{مج س}$$

$$36 = \frac{\quad}{5} = \frac{\quad}{\text{ن}} = \text{م}$$

حساب الوسيط :

نرتب القيم تصاعدياً :

60 50 40 20 10

الوسيط = 40

حساب الانحراف المعياري :

$$\sqrt{\frac{\text{مج ح}^2}{\text{ن}}} = \text{ع}$$

نكون الجدول التالي :

س	ح	ح ²
20	16-	256
50	14	196
60	24	576
40	4	16
10	26-	676
مج	0	1720

$$18.54 = \frac{\sqrt{1720}}{5} = \frac{\sqrt{\text{مج ح}^2}}{\text{ن}} = \text{ع}$$

$$0.64 = \frac{(40 - 36) 3}{18.54} = \frac{(م - و) 3}{\text{ع}} = \text{الالتواء}$$

الالتواء قيمته سالبة فيكون التواء التوزيع سالب .

تارين

1- فيما يلى مجموعة بيانات هي :

**95 - 200 - 250 - 300 - 110 - 90 - 150 -
100 - 350 - 80**

المطلوب حساب :

- المدى
- التباين
- الانحراف المعياري
- المتوسط
- الوسيط
- المنوال
- الانحراف المتوسط
- حدد نوع الالتواء

2- لمجموعة البيانات التالية احسب الانحراف المتوسط :-

2 - 3 - 5 - 6 - 7 - 8 - 8 - 9

3- فيما يلي الدرجات التي حصل عليها عشرة طلاب في اختبار

مادة الإحصاء وهي :

12 - 10 - 9 - 3 - 10 - 4 - 5 - 5 - 3 - 12

المطلوب حساب :

- المدى
- التباين
- الانحراف المعياري
- المتوسط
- الوسيط
- المنوال
- الانحراف المتوسط
- حدد نوع الالتواء

4- فيما يلي أعمار 10 طلاب بالفرقة الأولى قسم الاجتماع

17 - 18 - 19 - 17 - 18 - 20 - 22 - 17 - 18 - 18

19

المطلوب حساب :

- الانحراف المتوسط
- تحديد نوع الالتواء

5- من بيانات الجدول التالي احسب :-

الفئات	-16	-20	-24	-28	36-32
عدد المبحوثين	10	15	40	20	15

- المدى
- التباين
- الانحراف المعياري
- المتوسط
- الوسيط
- المنوال
- الانحراف المتوسط
- حدد نوع الالتواء

الفصل السابع تحليل التباين

مقدمه :

أولاً : طريقة حساب نسبة ف

ثانياً : تحديد مدى دلالة نسبة ف من عدمه .

مقدمه :

دلت الأبحاث الإحصائية التي قام بها فيشر على أهمية تحليل التباين في الميادين المختلفة لعلوم الحياة وخاصة في الكشف عن مدى تجانس العينات ومدى انتسابها إلى أصل واحد أو أصول متعددة .

وبالطبع هناك تساؤل لماذا نستخدم تحليل التباين "النسبة الفائية" للحكم على دلالة الإحصائية للعلاقة بين متغيرين وقد استخدمنا من قبل اختبار "ت" لنفس الغرض .

الإجابة بمنتهى السهولة هو أن اختبار "ت" يستخدم لدراسة العلاقة بين متغيرين فقط لا غير أما إذا زاد عدد المتغيرات عن اثنين فلا يمكن استخدام اختبار "ت" بل نستخدم "نسبة ف" .

وبالتالي فإن "نسبة ف" تصلح في حالة متغيرين أو أكثر .

ويعتمد تحليل التباين في صورته النهائية على قياس مدى اقتراب التباين الداخلي من التباين الخارجي أو مدى ابتعاده عنه وتقاس هذه الناحية بالنسبة التباينية أو النسبة الفائية من خلال العلاقة :

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

حيث أن التباين الكبير هو الأكبر في القيمة والتباين الصغير هو الأصغر في القيمة .

طريقة حساب نسبة ف

• حساب التباين بين المجموعات

مجموع المربعات بين المجموعات

$$\frac{\text{التباين بين المجموعات}}{\text{درجة حرية التباين بين المجموعات}} =$$

$$\text{مجموع المربعات بين المجموعات} = n_1 q_1^2 + n_2 q_2^2 + n_3 q_3^2 + \dots$$

حيث :

n_1, n_2, n_3, \dots هي عدد أفراد المجموعات
 $q_1^2, q_2^2, q_3^2, \dots$ هي مربع انحراف متوسط كل مجموعة عن المتوسط الكلي للمجموعات ويحسب من العلاقة :

$$q_1^2 = (m - 1)^2$$

حيث "م" هو المتوسط الوزني أو الكلي لكافة المجموعات .

درجة حرية التباين بين المجموعات = عدد المجموعات - 1

• حساب التباين داخل المجموعات

مجموع المربعات داخل المجموعات

$$\frac{\text{التباين داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}} =$$

مجموع المربعات داخل المجموعات = $1^2 ع_1 + 2^2 ع_2 + 3^2 ع_3 + \dots$

00000000000

حيث :

$1, 2, 3, \dots$ هي عدد أفراد المجموعات
 $ع_1^2, ع_2^2, ع_3^2, \dots$ هو التباين لكل مجموعة ويحسب من
العلاقة :

$$\left[\frac{\text{مجم س}^2}{ن} \right] - \frac{\text{مجم س}^2}{ن} = ع_1^2$$

درجة حرية التباين داخل المجموعات = مجموع أفراد جميع
المجموعات - عدد المجموعات

تحديد مدى دلالة "نسبة ف" من عدمه

على أى حال نحصل من قانون "نسبة ف" على "ف" المحسوبة
نقوم بمقارنتها بـ "ف" الجدوليه ونتبع الآتى :

إذا كانت قيمة "ف" المحسوبة < قيمة "ف" الجدوليه فان "نسبة
ف" تكون دالة إحصائية .

أما إذا كانت قيمة "ف" المحسوبة > قيمة "ف" الجدوليه فان
"نسبة ف" ليست دالة إحصائية .

مثال :

18	19	19	21	23	درجات الذكور
15	14	18	19	19	درجات الإناث

الجدول السابق يوضح درجات خمس ذكور وخمس إناث فى اختبار ما والمطلوب حساب النسبة الفئوية وبيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 وكذلك عند مستوى دلالة 0.01 ؟

الحل :

نفترض أن درجات الذكور هي "س" ودرجات الإناث هي "ص" ثم نكون الجدول التالى :

ص ²	ص	س ²	س
361	19	529	23
361	19	441	21
324	18	361	19
196	14	361	19
225	15	314	18
1467	85	2016	100

حساب المتوسطات :

$$20 = \frac{100}{5} = \frac{\text{مجموع س}}{ن} = \text{م س}$$

$$17 = \frac{85}{5} = \frac{\text{مجـ ص}}{2\text{ن}} = \text{م ص}$$

حساب المتوسط الكلي :

$$18.5 = \frac{17 + 20}{2} = \frac{\text{م ص} + \text{م ص}}{2} = \text{م}$$

حساب مربع انحراف كل متوسط عن المتوسط الكلي :

$$2.25 = {}^2(1.5-) = {}^2(20 - 18.5) = {}^2(\text{م} - \text{م ص}) = \text{ق ص}^2$$

$$2.25 = {}^2(1.5) = {}^2(17 - 18.5) = {}^2(\text{م ص} - \text{م}) = \text{ق ص}^2$$

حساب التباين :

$$\left[\frac{\text{مجـ ص}}{\text{ن}} \right]^2 - \frac{\text{مجـ ص}^2}{\text{ن}} = \text{ع ص}^2$$

$$3.2 = \left[\frac{100}{5} \right]^2 - \frac{2016}{5} = \text{ع ص}^2$$

$$\left[\frac{\text{مجـ ص}}{\text{ن}} \right]^2 - \frac{\text{مجـ ص}^2}{\text{ن}} = -219$$

$$\frac{\quad}{n} - \frac{\quad}{n} = \text{ع}^2 \text{ص}$$

$$4.4 = \left[\frac{85}{5} \right]^2 - \frac{1467}{5} = \text{ع}^2 \text{ص}$$

حساب مجموع المربعات بين المجموعات :

$$\begin{aligned} \text{مجموع المربعات بين المجموعات} &= \text{ن}^2 \text{ق}^2 \text{س} + \text{ن}^2 \text{ق}^2 \text{ص} \\ \text{مجموع المربعات بين المجموعات} &= 2.25 \times 5 + 2.25 \times 5 \\ \text{مجموع المربعات بين المجموعات} &= 22.5 \end{aligned}$$

حساب مجموع المربعات داخل المجموعات :

$$\begin{aligned} \text{مجموع المربعات داخل المجموعات} &= \text{ن}^2 \text{ع}^2 \text{س} + \text{ن}^2 \text{ع}^2 \text{ص} \\ \text{مجموع المربعات داخل المجموعات} &= 3.2 \times 5 + 4.4 \times 5 \\ \text{مجموع المربعات داخل المجموعات} &= 38 \end{aligned}$$

حساب درجات الحرية :

$$\begin{aligned} \text{درجة حرية التباين بين المجموعات} &= \text{عدد المجموعات} - 1 \\ \text{درجة حرية التباين بين المجموعات} &= 2 - 1 = 1 \\ \text{درجة حرية التباين داخل المجموعات} &= \text{عدد أفراد جميع} \\ &\quad \text{المجموعات} - \text{عدد المجموعات} \end{aligned}$$

$$8 = 2 - 5 + 5 = \text{درجة حرية التباين داخل المجموعات}$$

حساب التباين بين المجموعات :

$$\frac{\text{مجموع المربعات بين المجموعات}}{\text{درجة حرية التباين بين المجموعات}} = \text{التباين بين المجموعات}$$

$$\text{التباين بين المجموعات} = \frac{22.5}{1} = 22.5 \text{ (الأكبر)}$$

حساب التباين داخل المجموعات :

$$\frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}} = \text{التباين داخل المجموعات}$$

$$\text{التباين داخل المجموعات} = \frac{38}{8} = 4.75 \text{ (الأصغر)}$$

حساب نسبة ف :

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

$$4.7368 = \frac{22.5}{4.75} = \text{نسبة ف}$$

حساب "ف" الجدولية :

لحساب "ف" الجدولية نستخدم درجة حرية التباين الكبير = 1 ودرجة حرية التباين الصغير = 8 ونبحث فى جداول النسبة الفائية بدرجتى الحرية السابقتين فنحصل على القيمتين :

"ف" الجدولية = 5.32 عند مستوى دلالة 0.05

"ف" الجدولية = 11.26 عند مستوى دلالة 0.01

تحديد مدى دلالة "نسبة ف"

• "نسبة ف" المحسوبة = 4.7 > "ف" الجدولية عند مستوى دلالة 0.05 = 5.32 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.05 .

• "نسبة ف" المحسوبة = 4.7 > "ف" الجدولية عند مستوى دلالة 0.01 = 11.26 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.01 .

التعليق :

يمكن القول بأن جميع الفروق التي حصل عليها الباحث ليس لها دلالة إحصائية ولا توجد فروق معنوية بين المجموعتين وهذه الفروق ليست إلا مجرد صدفة .

مثال :

-	11	9	7	5	4	س
22	13	11	8	6	3	ص
-	-	16	13	9	7	هـ

الجدول السابق يوضح ثلاث مجموعات من الطلاب فى اختبار ما والمطلوب حساب النسبة الفئوية وبيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 وكذلك عند مستوى دلالة 0.01 ؟

الحل :
نكون الجدول التالى :

س	ص	هـ	س ²	ص ²	هـ ²
4	3	7	16	9	49
5	6	9	25	36	81
7	8	13	49	64	169
9	11	16	81	121	256
11	13	-	121	169	-
-	22	-	-	484	-
36	63	45	292	883	555

حساب المتوسطات :
مـ س

36

$$7.2 = \frac{\quad}{5} = \frac{\quad}{1\text{ن}} = \text{م س}$$

$$10.5 = \frac{63}{6} = \frac{\text{مج ص}}{2\text{ن}} = \text{م ص}$$

$$11.25 = \frac{45}{4} = \frac{\text{مج هـ}}{3\text{ن}} = \text{م هـ}$$

حساب المتوسط الكلى :

$$9.65 = \frac{11.25 + 10.5 + 7.2}{3} = \frac{\text{م هـ} + \text{م ص} + \text{م س}}{3} = \text{م}$$

حساب مربع انحراف كل متوسط عن المتوسط الكلى :

$$6 = {}^2(2.45) = {}^2(7.2 - 9.65) = {}^2(\text{م س} - \text{م}) = \text{ق س}^2$$

$$0.7225 = {}^2(0.85-) = {}^2(10.5 - 9.65) = {}^2(\text{م ص} - \text{م}) = \text{ق ص}^2$$

$$2.56 = {}^2(1.6-) = {}^2(11.25 - 9.65) = {}^2(\text{م هـ} - \text{م}) = \text{ق هـ}^2$$

حساب التباين :

$${}^2 \left[\frac{\text{مج س}}{1\text{ن}} \right] - \frac{{}^2 \text{مج س}}{1\text{ن}} = \text{ع س}^2$$

$$6.56 = \left[\frac{36}{5} \right]^2 - \frac{292}{5} = \text{ع}^2 \text{س}$$

$$\left[\frac{\text{مجـ ص}}{2\text{ن}} \right]^2 - \frac{\text{مجـ ص}^2}{2\text{ن}} = \text{ع}^2 \text{ص}$$

$$36.92 = \left[\frac{63}{6} \right]^2 - \frac{883}{6} = \text{ع}^2 \text{ص}$$

$$\left[\frac{\text{مجـ هـ}}{3\text{ن}} \right]^2 - \frac{\text{مجـ هـ}^2}{3\text{ن}} = \text{ع}^2 \text{هـ}$$

$$12.18 = \left[\frac{45}{4} \right]^2 - \frac{555}{4} = \text{ع}^2 \text{هـ}$$

حساب مجموع المربعات بين المجموعات :

$$\text{مجموع المربعات بين المجموعات} = \text{ن}^2 \text{ق}^2 \text{س} + \text{ن}^2 \text{ق}^2 \text{ص} + \text{ن}^2 \text{ق}^2 \text{هـ}.$$

$$\text{مجموع المربعات بين المجموعات} = 6 \times 5 + 6 \times 0.7225 + 2.56 \times 4$$

$$\text{مجموع المربعات بين المجموعات} = 44.57$$

حساب مجموع المربعات داخل المجموعات :

$$\text{مجموع المربعات داخل المجموعات} = \text{ن}_1 \text{ع}_1^2 + \text{ن}_2 \text{ع}_2^2 + \text{ن}_3 \text{ع}_3^2$$

$$\text{مجموع المربعات داخل المجموعات} = 6 \times 6.56 + 6 \times 36.92 + 4 \times 12.18$$

$$\text{مجموع المربعات داخل المجموعات} = 303.04$$

حساب درجات الحرية :

$$\text{درجة حرية التباين بين المجموعات} = \text{عدد المجموعات} - 1$$

$$\text{درجة حرية التباين بين المجموعات} = 3 - 1 = 2$$

$$\text{درجة حرية التباين داخل المجموعات} = \text{عدد أفراد جميع المجموعات} - \text{عدد المجموعات}$$

$$\text{درجة حرية التباين داخل المجموعات} = 5 + 6 + 4 - 3 = 12$$

حساب التباين بين المجموعات :

مجموع المربعات بين المجموعات

التباين بين المجموعات = $\frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين بين المجموعات}}$

$$\text{التباين بين المجموعات} = \frac{44.57}{2} = 22.27 \text{ (الأصغر)}$$

حساب التباين داخل المجموعات :

التباين داخل المجموعات = $\frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجة حرية التباين داخل المجموعات}}$

$$\text{التباين داخل المجموعات} = \frac{303.04}{12} = 25.2 \text{ (الأكبر)}$$

حساب نسبة ف :

$$\text{نسبة ف} = \frac{\text{التباين الكبير}}{\text{التباين الصغير}}$$

$$\text{نسبة ف} = \frac{25.2}{22.27} = 1.13$$

حساب "ف" الجدولية :

لحساب "ف" الجدولية نستخدم درجة حرية التباين الكبير = 12 ودرجة حرية التباين الصغير = 2 ونبحث فى جداول النسبة الفائية بدرجتى الحرية السابقتين فنحصل على القيمتين :

"ف" الجدولية = 19.41 عند مستوى دلالة 0.05

"ف" الجدولية = 99.42 عند مستوى دلالة 0.01

تحديد مدى دلالة "نسبة ف"

- "نسبة ف" المحسوبة = 1.13 > "ف" الجدولية عند مستوى دلالة 0.05 = 19.41 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.05
- "نسبة ف" المحسوبة = 1.13 > "ف" الجدولية عند مستوى دلالة 0.01 = 99.42 ، لذا فان "نسبة ف" ليست دالة عند مستوى 0.01

تارين

1- الجدول التالي يوضح درجات مجموعتين من الطلاب فى اختبار فى مادة الإحصاء الاجتماعى :

س	19	20	14	17	25	مج = 95
ص	14	11	12	13	20	مج = 70

والمطلوب حساب نسبة "ف" مع بيان عما إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

2- الجدول التالي يوضح درجات مجموعتين من الطلاب فى اختبار فى مادة الحاسب الآلى :

س	18	19	13	16	24	مج = 90
ص	15	12	13	14	21	مج = 75

والمطلوب حساب نسبة "ف" مع بيان عما إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

3- الجدول التالي يوضح درجات مجموعتين من الطلاب فى اختبار فى مادة اللغة الفرنسية :

س	7	15	15	11	12
ص	7	2	8	7	6

احسب الدلالة للفروق القائمة بين تلك الدرجات بطريقة تحليل التباين عند مستوى دلالة 0.05

4- الجدول التالي يوضح درجات 4 مجموعات من الطلاب فى اختبار فى مادة اللغة العربية :

س	49	59	61	60	61
ص	68	55	60	67	60
ع	64	63	54	52	62
هـ	67	55	65	64	59

احسب الدلالة الإحصائية للفروق القائمة بين تلك الدرجات بطريقة تحليل التباين عند مستوى دلالة 0.05 وبين مدى تجانس هذه المجموعات بالنسبة لأصل واحد أو لأصول متعددة .

الفصل الثامن اختبار "ت"

مقدمه :

أولاً : شروط استخدام اختبار "ت"

ثانياً : الحالات المختلفة لحساب "ت"

مقدمة :

يعد اختبار "ت" من أكثر اختبارات الدلالة شيوعاً في الأبحاث النفسية والاجتماعية والتربوية ، وترجع نشأته الأولى إلى أبحاث العالم "ستودنت" ولهذا سمي الاختبار بأكثر الحروف تكراراً في اسمه وهو حرف التاء .

ومن أهم المجالات التي يستخدم فيها هذا الاختبار الكشف عن الفروق بين تحصيل الذكور والإناث في مادة دراسية ما وذلك عن طريق حساب دلالة فرق متوسط تحصيل الذكور عن متوسط تحصيل الإناث .

ويمكن القول أن اختبار "ت" يستخدم لقياس دلالة فروق المتوسطات غير المرتبطة والمرتبطة للعينات المتساوية والغير متساوية .

شروط استخدام اختبار "ت" لدلالة فروق المتوسطات

لا يحق للباحث أن يستخدم اختبار "ت" قبل أن يدرس خصائص متغيرات البحث من النواحي التالية :-

- 1- حجم كل عينة .
- 2- الفرق بين حجم عيني البحث .
- 3- مدى تجانس العينة .
- 4- مدى اعتدالية التوزيع التكراري لكل من عيني البحث .

1- حجم كل عينة

يجب أن يزيد حجم كل من العينتين عن "5" ويفضل أن يزيد عن "30" أما إذا قل حجم أى من العينتين عن "5" فلا يمكن استخدام اختبار "ت" .

2- الفرق بين حجم عينتى البحث : شرط التقارب

يجب أن يكون حجم عينتى البحث متقارباً فلا يكون مثلاً حجم أحد العينتين "500" وحجم الأخرى "30" لأن للحجم أثره على مستوى دلالة "ت" .

3- مدى تجانس العينتين

يقصد بتجانس العينات مدى انتسابها إلى أصل واحد أو أصول متعددة . فإذا انتسبت العينات إلى أصل واحد فهي متجانسة وإذا لم تنتسب العينات إلى أصل واحد فهي غير متجانسة . وبالطبع يصعب بالنسبة للباحث تحديد أصول العينات لتحديد تجانسها لذا يمكنه استخدام النسبة الفئوية لتحديد التجانس . يحدد تجانس العينتين من خلال حساب قيمة النسبة الفئوية حيث تحسب من العلاقة :

التباين الأكبر

$$\frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \text{ف}$$

التباين الأصغر

حيث أن التباين الأكبر هو التباين الأكبر فى القيمة دون التحيز لأحد العينتين ، والتباين الأصغر هو الأصغر فى القيمة دون التحيز لأحد العينتين .

بالطبع نحصل من القانون السابق على قيمة لـ "ف" تسمى بقيمة ف المحسوبة ولتحديد التجانس نحسب قيمة أخرى تسمى ف الجدولية ونحصل عليها من جداول "ف" الإحصائية عند درجة حرية التباين الأكبر ودرجة حرية التباين الأصغر ومستوى الدلالة الذى قيمته إما "0.05" أو "0.01" حيث نحسب درجات الحرية من القانون التالى :

درجة حرية التباين الأصغر = ن - 1

حيث "ن" هى عدد أفراد العينة التى تبيانها هو الأكبر .

درجة حرية التباين الأصغر = ن - 1

حيث "ن" هى عدد أفراد العينة التى تبيانها هو الأصغر .

تحديد التجانس

• إذا كانت قيمة "ف" المحسوبة < قيمة "ف" الجدولية فلا يوجد هناك تجانس .

• أما إذا كانت قيمة "ف" المحسوبة > قيمة "ف" الجدولية فيوجد هناك تجانس .

4- مدى اعتدالية التوزيع التكرارى لكل من العينتين

يكون التوزيع التكرارى معتدلاً عندما تكون قيمة الالتواء الخاص به محصورة بين القيمتين [3- ، 3+] أى واقعة فى الفترة المغلقة 3- و 3+ .

ويحسب الالتواء من القانون التالى :-

$$\frac{(م - و) \times 3}{ع} = \text{الالتواء}$$

حيث:

• "م" هو المتوسط الحسابي ويحسب من العلاقة

$$م = \frac{\text{مج س}}{ن}$$

حيث : "مج س" هي مجموع القيم ، س هي القيم ، ن هي عدد القيم .

• "و" هو الوسيط ، ويحسب عن طريق ترتيب القيم تصاعدياً أو تنازلياً ثم اختيار قيمة الوسيط في حالة أن يكون عدد الأفراد فردياً تكون قيمة الوسيط التي ترتيبها $(ن+1)/2$ أما إذا كان عدد الأفراد زوجياً فتكون قيمة الوسيط هي متوسط القيمتين اللتان ترتيبهما $ن/2$ ، $ن/2 + 1$.

• "ع" هو الانحراف المعياري ويحسب من العلاقة :

$$ع^2 = \frac{\text{مج ح}^2}{ن}$$

من الواضح أن القانون السابق يحسب قيمة التباين فنأخذ للقيمة الناتجة الجذر التربيعي لنحصل على الانحراف المعياري كالتالي .

$$ع = \sqrt{\frac{\text{مج ح}^2}{ن}}$$

حيث :

ع = الانحراف المعياري

ح = الانحراف = س - م

ن = عدد القيم

تحديد مدى دلالة "ت" من عدمه

سنحصل في جميع حالات "ت" على قيمة لـ "ت" نسميها "ت المحسوبة" ثم نقارنها بقيمة لـ "ت" نحصل عليها من الجداول تسمى "ت الجدولية"

- إذا كانت قيمة "ت المحسوبة" < قيمة "ت الجدولية" تكون قيمة "ت" دالة إحصائية .
- أما إذا كانت قيمة "ت المحسوبة" > قيمة "ت الجدولية" تكون قيمة "ت" ليست دالة إحصائية .

الحالات المختلفة لحساب "ت"

1- الحالة الأولى : حساب "ت" لدلالة فرق عينتين متجانستين غير متساويتين في أعداد أفرادهما .

في هذه الحالة تكون ن₁ لا تساوي ن₂ حيث ن₁ ، ن₂ هما عدد أفراد العينة الأولى والثانية على الترتيب .
تحسب دلالة "ت" لفرق عينتين متجانستين ومختلفتين في عدد الأفراد بالمعادلة التالية

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}$$

$$= \sqrt{\left[\frac{1}{2N} + \frac{1}{1N} \right] \left[\frac{N_1 E_1^2 + N_2 E_2^2}{2 - 2N + 1N} \right]}$$

حيث :

- $m_1 =$ المتوسط الحسابي للمجموعة الأولى .
- $m_2 =$ المتوسط الحسابي للمجموعة الثانية .
- $E_1^2 =$ تباين المجموعة الأولى .
- $E_2^2 =$ تباين المجموعة الثانية .
- $N_1 =$ عدد أفراد المجموعة الأولى .
- $N_2 =$ عدد أفراد المجموعة الثانية .

مثال :

2	6	8	3	5	4	7	العينة الأولى
-	13	10	2	15	5	3	العينة الثانية

الجدول السابق يوضح درجات مجموعة من الذكور والإناث في اختبار للذكاء والمطلوب حساب قيمة "ت" من خلال التحقق من شروط اختبار "ت" ومن ثم تحديد هل "ت" دالة إحصائية أم لا ؟ عند مستوى دلالة إحصائية 0.01 ؟

الحل :

قبل أن نبدأ الحل نلاحظ أن

$$6 = 2n \neq 7 = 1n$$

نعتبر أن العينة الأولى هي "س" والعينة الثانية هي "ص" ونقوم ببناء الجدول التالي .

ح ² _ص	ح _ص	ص	ح ² _س	ح _س	س
25	5-	3	4	2	7
9	3-	5	1	1-	4
49	7	15	0	0	5
36	6-	2	4	2-	3
16	4	10	9	3	8
25	5	13	1	1	6
-	-	-	9	3-	2
148	-	48	28	-	35

العينة الأولى :

نحسب لها المتوسط والوسيط والتباين والانحراف المعياري كالتالي:

حساب المتوسط :

$$م\ س = \frac{35}{7} = \frac{مج\ س}{1n}$$

حساب الوسيط :

نرتب قيم المتغير (س) ترتيباً تصاعدياً كالتالي :

8 7 6 5 4 3 2
حيث أن عدد أفراد العينة الأولى فردية لذا فإن قيمة الوسيط هي

القيمة التي ترتيبها (ن+1/2) أي التي ترتيبها (4)

الوسيط = و س = 5

حساب التباين :

$$4 = \frac{28}{7} = \frac{\text{مج ح}^2_{س}}{ن} = ع^2_{س}$$

حساب الانحراف المعياري :

$$2 = \sqrt{4} = \sqrt{ع^2_{س}} = ع_{س}$$

حساب الالتواء :

$$\text{الالتواء} = \frac{(5 - 5) \times 3}{2} = \frac{(م - و) \times 3}{ع} = \text{صفر}$$

العينة الثانية :

نحسب لها المتوسط والوسيط والتباين والانحراف المعياري

كالتالي:

حساب المتوسط :

$$8 = \frac{48}{6} = \frac{\text{مج ص}}{ن} = م_{ص}$$

حساب الوسيط :

نرتب قيم المتغير (ص) ترتيباً تصاعدياً كالتالى :

15 13

10	5
----	---

 3 2

حيث أن عدد أفراد العينة الثانية زوجية لذا فإن قيمة الوسيط هي متوسط القيمتين اللتان ترتيبهما (ن/2 ، ن/2 + 1) أى التى ترتيبها (3 ، 4)

$$7.5 = 2/(10 + 5) = \text{وص}$$

حساب التباين :

$$24.66 = \frac{148}{6} = \frac{\text{مج ح}^2 \text{ص}}{2\text{ن}} = \text{ع}^2 \text{ص}$$

حساب الانحراف المعياري :

$$5 = \sqrt{24.66} = \sqrt{\text{ع}^2 \text{ص}} = \text{ع ص}$$

حساب الالتواء :

$$0.3 = \frac{(7.5 - 8) \times 3}{5} = \frac{(م - و) \times 3}{\text{ع}} = \text{الالتواء}$$

التحقق من شروط اختبار "ت"

1- حجم العينتين :

$$n_1 = 7 < 5$$

$$n_2 = 6 < 5$$

حيث أن حجم كل من العينتين على حده لا بد وأن يكون أكبر من 5
لذا فهذا الشرط متحقق .

2- تقارب العينتين :

$$n_1 = 7 \text{ تتقارب جداً من } n_2 = 6$$

3- تجانس العينتين :

نحسب قيمة "ف" المحسوبة من العلاقة :

$$F_{\text{المحسوبة}} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \frac{24.66}{4} = 6.116$$

لإيجاد قيمة "ف" الجدولية يلزم حساب قيمة كل من درجة حرية
التباين الأكبر ودرجة حرية التباين الأصغر .

$$\text{درجة حرية التباين الأكبر} = n_2 - 1 = 6 - 1 = 5$$

ونلاحظ أننا اخترنا درجة حرية التباين الأكبر من عدد أفراد المجموعة الثانية لأن تباين العينة الثانية هو الأكبر .

درجة حرية التباين الأصغر = $n_1 - 1 = 7 - 1 = 6$
من جداول "ف" عند درجة حرية تباين كبير (5) ودرجة حرية تباين صغير (6) ومستوى دلالة 0.01 نجد أن قيمة "ف" الجدولية = 8.75 .

بمقارنة قيمة "ف" المحسوبة بقيمة "ف" الجدولية نجد أن :
"ف" المحسوبة > "ف" الجدولية (لذا فإنه يوجد تجانس بين العينتين) .

4- اعتدالية التوزيع للعينتين :

$3^- > \text{التواء س} = \text{صفر} > 3^+$

نلاحظ أن قيمة التواء س محصور في الفئة $[3^-, 3^+]$ لذا فإن توزيع العينة س معتدل .

$3^- > \text{التواء ص} = 0.3 > 3^+$

نلاحظ أن قيمة التواء ص محصور في الفئة $[3^-, 3^+]$ لذا فإن توزيع العينة ص معتدل .

حساب قيمة "ت" المحسوبة :

$$t = \sqrt{\left(\frac{1}{2n} + \frac{1}{1n}\right) \left[\frac{2^2 \text{ع} 2n + 2^2 \text{ع} 1n}{2 - 2n + 1n}\right]}$$

بالتعويض فى المعادلة السابقة :

$$t = \sqrt{\left(\frac{1}{6} + \frac{1}{7}\right) \left[\frac{24.66 \times 6 + 4 \times 7}{2 - 6 + 7}\right]}$$

ت المحسوبة = 1.36-

تهمل الإشارة السالبة لقيمة "ت" دائماً فتصبح :

قيمة "ت" المحسوبة = 1.36 .

حساب قيمة "ت" الجدولية :

لإيجاد قيمة "ت" الجدولية يلزم حساب درجة الحرية :

$$\text{درجة الحرية} = n_1 + n_2 - 2 = 7 - 6 + 2 = 11$$

بالبحث فى جداول "ت" عند درجة حرية 11 ومستوى دلالة 0.01

مع الأخذ فى الاعتبار أن البحث يكون فى دلالة الطرفين ، نجد أن

قيمة "ت" الجدولية = 3.11 .

تحديد دلالة "ت"

بمقارنة قيمة "ت" المحسوبة بقيمة "ت" الجدولية :

نجد أن "ت" المحسوبة = 1.36 > "ت" الجدولية = 3.11
وبالتالي فإن "ت" ليست دالة إحصائية .

2- الحالة الثانية : حساب "ت" لدلالة فرق عينتين غير متجانستين و غير متساويتين في أعداد أفرادهما

في هذه الحالة تكون n_1 لا تساوي n_2 أيضاً مثل الحالة السابقة
حيث n_1 ، n_2 هما عدد أفراد العينة الأولى والثانية على الترتيب.
تحسب دلالة "ت" لعينتين غير متجانستين ومختلفين في عدد
الأفراد بالمعادلة التالية :

$$ت = \frac{m_1 - m_2}{\sqrt{\frac{E_1^2}{n_1} + \frac{E_2^2}{n_2}}}$$

حيث :

- m_1 = المتوسط الحسابي للمجموعة الأولى .
- m_2 = المتوسط الحسابي للمجموعة الثانية .
- E_1^2 = تباين المجموعة الأولى .
- E_2^2 = تباين المجموعة الثانية .
- n_1 = عدد أفراد المجموعة الأولى .

ن₂ = عدد أفراد المجموعة الثانية .

مثال :

20	19	13	48	19	32	22	17	35	العينة الأولى
-	-	7	2	14	10	9	3	11	العينة الثانية

الجدول السابق يوضح درجات مجموعة من الذكور والإناث في اختبار للذكاء والمطلوب حساب قيمة "ت" من خلال التحقق من شروط اختبار "ت" ومن ثم تحديد هل "ت" دالة إحصائية أم لا ؟ عند مستوى دلالة إحصائية 0.05 ؟

الحل :

قبل أن نبدأ الحل نلاحظ أن

$$ن_1 = 9 \neq 7 = ن_2$$

نعتبر أن العينة الأولى هي "س" والعينة الثانية هي "ص" ونقوم ببناء الجدول التالي .

س	ح _س	ح _ص ²	ص	ح _ص	ح _ص ²
35	10	100	11	3	9
17	8-	64	3	5-	25
22	3-	9	9	1	1
19	6-	36	14	6	36

36	6-	2	569	23	48
1	1-	7	144	12-	13
-	-	-	36	6-	19
-	-	-	25	5-	20
112	-	56	992	-	225

العينة الأولى :

نحسب لها المتوسط والوسيط والتباين والانحراف المعياري كالتالي:

حساب المتوسط :

$$م س = \frac{225}{9} = \frac{مج س}{1ن} = م س$$

حساب الوسيط :

نرتب قيم المتغير (س) ترتيباً تصاعدياً كالتالي :

48 35 32 22 20 19 19 17 13
حيث أن عدد أفراد العينة الأولى فردية لذا فإن قيمة الوسيط هي
القيمة التي ترتيبها (ن+1/2) أي التي ترتيبها (5)
الوسيط = وس = 20

حساب التباين :

$$110.2 = \frac{992}{9} = \frac{\text{مج ح}^2 \text{س}}{1 \text{ن}} = \text{ع}^2 \text{س}$$

حساب الانحراف المعياري :

$$10.5 = \sqrt{110.2} = \sqrt{\text{ع}^2 \text{س}} = \text{ع س}$$

حساب الالتواء :

$$1.4 = \frac{(20 - 25) \times 3}{10.5} = \frac{(م - و) \times 3}{\text{ع}} = \text{الالتواء}$$

العينة الثانية :

نحسب لها المتوسط والوسيط والتباين والانحراف المعياري كالتالي:

حساب المتوسط :

$$8 = \frac{56}{7} = \frac{\text{مج ص}}{2 \text{ن}} = \text{م ص}$$

حساب الوسيط :

نرتب قيم المتغير (ص) ترتيباً تصاعدياً كالتالى :

14 11 10 9 7 3 2

حيث أن عدد أفراد العينة الثانية فردية لذا فإن قيمة الوسيط هى
القيمة التى ترتيبها (ن+1/2) أى التى ترتيبها (4)
الوسيط = وص = 9

حساب التباين :

$$16 = \frac{112}{7} = \frac{\text{مج ح}^2}{2\text{ن}} = \text{ع}^2$$

حساب الانحراف المعياري :

$$4 = \sqrt{16} = \sqrt{\text{ع}^2} = \text{ع}$$

حساب الالتواء :

$$0.75- = \frac{(9 - 8) \times 3}{4} = \frac{(و - م) \times 3}{\text{ع}} = \text{الالتواء}$$

التحقق من شروط اختبار "ت"

1- حجم العينتين :

$$n_1 = 9 < 5$$

$$n_2 = 7 < 5$$

حيث أن حجم كل من العينتين على حده لا بد وأن يكون أكبر من 5 لذا فهذا الشرط متحقق .

2- تقارب العينتين :

$$n_1 = 9 \text{ تتقارب جداً من } n_2 = 7$$

3- تجانس العينتين :

نحسب قيمة "ف" المحسوبة من العلاقة :

$$F_{\text{المحسوبة}} = \frac{\text{التباين الأكبر}}{\text{التباين الأصغر}} = \frac{110.2}{16} = 6.88$$

لإيجاد قيمة "ف" الجدولية يلزم حساب قيمة كل من درجة حرية التباين الأكبر ودرجة حرية التباين الأصغر .

$$\text{درجة حرية التباين الأكبر} = n_1 - 1 = 9 - 1 = 8$$

ونلاحظ أننا اخترنا درجة حرية التباين الأكبر من عدد أفراد المجموعة الأولى لأن تباين العينة الأولى هو الأكبر .

$$\text{درجة حرية التباين الأصغر} = n_2 - 1 = 7 - 1 = 6$$

من جداول "ف" عند درجة حرية تباين كبير (8) ودرجة حرية تباين صغير (6) ومستوى دلالة 0.05 نجد أن قيمة "ف" الجدولية = 4.15 .

بمقارنة قيمة "ف" المحسوبة بقيمة "ف" الجدولية نجد أن :
 "ف" المحسوبة < "ف" الجدولية (لذا فانه لا يوجد تجانس بين العينتين) .

4- اعتدالية التوزيع للعينتين :

$$3- > 3+ \text{ س } = 1.4$$

نلاحظ أن قيمة التواء س محصور في الفئة [3-, 3+] لذا فان توزيع العينة س معتدل .

$$3- > 3+ \text{ ص } = 0.75$$

نلاحظ أن قيمة التواء ص محصور في الفئة [3-, 3+] لذا فان توزيع العينة ص معتدل .

حساب قيمة "ت" المحسوبة :

$$ت = \frac{م_1 - م_2}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$$

بالتعويض فى المعادلة السابقة :

$$t = \frac{8 - 25}{\sqrt{\frac{16}{7} + \frac{110.2}{9}}}$$

ت المحسوبة = 4.46

حساب قيمة "ت" الجدولية :

لإيجاد قيمة "ت" الجدولية

تحسب من العلاقة التالية :

$$t \text{ الجدولية} = \frac{t \times 1 \left(\frac{1}{n_1} \right) + t \times 2 \left(\frac{2}{n_2} \right)}{\left(\frac{1}{n_1} \right) + \left(\frac{2}{n_2} \right)}$$

حيث :

ت1 : هى "ت" الجدولية للعينة الأولى وتحسب عن طريق حساب

درجة حرية العينة الأولى على حده من العلاقة :

$$8 = 1 - 9 = 1 - n_1 = \text{درجة حرية العينة الأولى}$$

وبالبحث فى جداول "ت" عن درجة حرية 8 ومستوى دلالة 0.05

فى دلالة الطرفين نجد أن قيمة ت = 2.31

ت2 : هي "ت" الجدولية للعينه الثانيه وتحسب عن طريق حساب

درجة حرية العينه الثانيه على حده من العلاقه :

$$\text{درجة حرية العينه الثانيه} = 2n - 1 = 7 - 1 = 6$$

وبالبحث في جداول "ت" عن درجة حرية 6 ومستوى دلالة 0.05

في دلالة الطرفين نجد أن قيمة ت2 = 2.45

ثم نعوض في المعادله التاليه لحساب قيمة "ت" الجدولية :

$$\text{ت الجدولية} = \frac{\text{ت1} \times \left(\frac{1}{n_1}\right) + \text{ت2} \times \left(\frac{2}{n_2}\right)}{\left(\frac{1}{n_1}\right) + \left(\frac{2}{n_2}\right)}$$

$$2.31 \times \left(\frac{9}{110.2}\right) + 2.45 \times \left(\frac{7}{16}\right)$$

$$\text{ت الجدولية} = \frac{2.31 \times \left(\frac{9}{110.2}\right) + 2.45 \times \left(\frac{7}{16}\right)}{\left(\frac{9}{110.2}\right) + \left(\frac{7}{16}\right)}$$

$$\text{ت الجدولية} = 2.33$$

تحديد دلالة "ت"

بمقارنه قيمة "ت" المحسوبه بقيمه "ت" الجدولية

$$\text{نجد أن "ت" المحسوبه} = 4.46 < \text{"ت" الجدولية} = 2.33$$

وبالتالي فإن "ت" دالة إحصائية .

3- الحالة الثالثة : حساب "ت" لدلالة
فرق عينتين غير مرتبطتين ومتساويتين في
أعداد أفرادهما

في هذه الحالة لا نتحقق من شروط اختبار "ت" .
في هذه الحالة تكون $n_1 = n_2$ حيث n_1 ، n_2 هما عدد أفراد
العينة الأولى والثانية على الترتيب .
تحسب دلالة "ت" لفرق عينتين متساويتين في عدد الأفراد
بالمعادلة التالية :

$$t = \frac{\bar{m}_1 - \bar{m}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n-1}}}$$

حيث :

- \bar{m}_1 = المتوسط الحسابي للمجموعة الأولى .
- \bar{m}_2 = المتوسط الحسابي للمجموعة الثانية .
- s_1^2 = تباين المجموعة الأولى .
- s_2^2 = تباين المجموعة الثانية .
- n = عدد أفراد العينة الأولى أو الثانية حيث أنهما متساويتان .

مثال :

2	6	8	3	5	4	7	العينة الأولى
1	13	10	2	15	5	3	العينة الثانية

الجدول السابق يوضح درجات مجموعة من الذكور والإناث في اختبار للذكاء والمطلوب حساب قيمة "ت" ومن ثم تحديد هل "ت" دالة إحصائية أم لا ؟ عند مستوى دلالة إحصائية 0.05 ؟

الحل :

قبل أن نبدأ الحل نلاحظ أن

$$n_1 = n_2 = 7$$

نعتبر أن العينة الأولى هي "س" والعينة الثانية هي "ص" ونقوم ببناء الجدول التالي .

س	ح _س	ح _ص ²	ص	ح _ص	ح _ص ²
7	2	4	3	4-	16
4	1-	1	5	2-	4
5	0	0	15	8	64
3	2-	4	2	5-	25
8	3	9	10	3	9
6	1	1	13	6	36
2	3-	9	1	6-	36
35	-	28	49	-	190

العينة الأولى :

نحسب لها المتوسط والتباين .

حساب المتوسط :

$$5 = \frac{35}{7} = \frac{\text{مج س}}{1\text{ن}} = \text{م س}$$

حساب التباين :

$$4 = \frac{28}{7} = \frac{\text{مج ح}^2\text{س}}{1\text{ن}} = \text{ع}^2\text{س}$$

العينة الثانية :

نحسب لها المتوسط والتباين كالتالي:

حساب المتوسط :

$$7 = \frac{49}{7} = \frac{\text{مج ص}}{2\text{ن}} = \text{م ص}$$

حساب التباين :

$$27.14 = \frac{190}{7} = \frac{\text{مج ح}^2\text{ص}}{2\text{ن}} = \text{ع}^2\text{ص}$$

حساب قيمة "ت" المحسوبة :

$$ت = \frac{2\text{م} - 1\text{م}}{\sqrt{\frac{2\text{ع}^2 + 2\text{ع}^2}{1 - \text{ن}}}}$$

بالتعويض فى المعادلة السابقة :

$$t = \frac{7 - 5}{\sqrt{\frac{27.14 + 4}{1 - 7}}}$$

ت المحسوبة = -0.88

تهمل الإشارة السالبة لقيمة "ت" دائماً فتصبح :

قيمة "ت" المحسوبة = 0.88 .

حساب قيمة "ت" الجدولية :

لإيجاد قيمة "ت" الجدولية يلزم حساب درجة الحرية :

$$\text{درجة الحرية} = n_1 + n_2 - 2 = 7 + 7 - 2 = 12$$

بالبحث فى جداول "ت" عند درجة حرية 12 ومستوى دلالة 0.05

مع الأخذ فى الاعتبار أن البحث يكون فى دلالة الطرفين ، نجد أن

قيمة "ت" الجدولية = 2.18 .

تحديد دلالة "ت"

بمقارنة قيمة "ت" المحسوبة بقيمة "ت" الجدولية

نجد أن "ت" المحسوبة = 0.88 > "ت" الجدولية = 2.18

وبالتالى فإن "ت" ليست دالة إحصائية .

4- الحالة الرابعة : حساب "ت" لدلالة
فرق عينتين مرتبطتين ومتساويتين في أعداد
أفرادهما

يرتبط المتوسطان عندما نجرى اختباراً على مجموعة من الأفراد
ثم نعيد نفس الاختبار على نفس المجموعة في وقت آخر أى أن
العينة التى يجرى عليها الاختبار الأول هى نفسها العينة التى
يجرى عليها الاختبار الثانى وفى هذه الحالة لا تكون $n_1 = n_2$ بل
تصبح هى نفسها .

فى هذه الحالة أيضاً لا نتحقق من شروط اختبار "ت" .
تحسب دلالة "ت" لفرق عينتين متساويتين فى عدد الأفراد
بالمعادلة التالية :

$$t = \frac{m_f}{\sqrt{\frac{مج ح ف^2}{n(n-1)}}}$$

حيث :

• $m_f =$ متوسط الفروق ويحسب من العلاقة :

$$m_f = \frac{مج ف}{n}$$

• $f =$ الفروق = $s_1 - s_2$

• s_1 هى درجات الاختبار الأول

• s_2 هى درجات الاختبار الثانى

• $n =$ عدد الأفراد فى أى من الاختبارين .

• ح ف = ف - م ف
مثال :

11	22	16	23	14	22	24	20	18	26	درجات الاختبار الأول
9	23	11	24	12	18	21	19	16	23	درجات الاختبار الثانى

الجدول السابق يوضح درجات مجموعة من الأطفال في اختبار للذكاء حيث تم إجراء الاختبار مرة ثم بعد إجراء برنامج تدريبي لهم تم إجراء الاختبار مرة أخرى والمطلوب حساب قيمة "ت" للفرق بين درجات الاختبارين ومن ثم تحديد هل "ت" دالة إحصائية أم لا ؟ عند مستوى دلالة إحصائية 0.05 ؟

الحل :

قبل أن نبدأ الحل نلاحظ أن

ن₁ هي نفسها ن₂

نعتبر أن درجات الاختبار الأول هي "س₁" ودرجات الاختبار الثانى هي "س₂" ثم نقوم ببناء الجدول التالى :

س ₁	س ₂	ف	ح ف	ح ² ف
26	23	3	1	1
18	16	2	0	0
20	19	1	1-	1
24	21	3	1	1
22	18	4	2	4
14	12	2	0	0

9	3-	1-	24	23
9	3	5	11	16
9	3-	1-	23	22
0	0	2	9	11
34	-	20	-	-

حساب متوسط الفروق م ف :

$$2 = \frac{20}{10} = \frac{\text{م ف}}{\text{ن}} = \text{م ف}$$

حساب ح ف :

يحسب من العلاقة :

$$\text{ح ف} = \text{م ف} - \text{م ف}$$

حساب قيمة "ت" المحسوبة :

$$ت = \frac{\text{م ف}}{\sqrt{\frac{\text{م ف}^2}{\text{ن}(\text{ن}-1)}}}$$

بالتعويض فى المعادلة السابقة :

2

$$t = \frac{34}{(1-10)10}$$

ت المحسوبة = 3.25

حساب قيمة "ت" الجدولية :

لإيجاد قيمة "ت" الجدولية يلزم حساب درجة الحرية :

$$\text{درجة الحرية} = n - 1 = 10 - 1 = 9$$

بالبحث في جداول "ت" عند درجة حرية 9 ومستوى دلالة 0.05

مع الأخذ في الاعتبار أن البحث يكون في دلالة الطرف الواحد ،

نجد أن قيمة "ت" الجدولية = 1.83 .

تحديد دلالة "ت"

بمقارنة قيمة "ت" المحسوبة بقيمة "ت" الجدولية

نجد أن "ت" المحسوبة = 3.25 > "ت" الجدولية = 1.83

وبالتالي فإن "ت" دالة إحصائية .

تارين

1- القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد فى مقياس توهم المرض :

س	10	6	9	7	8	5	18
ص	3	5	11	6	8	9	-

حيث س هي مجموعة الذكور ، ص مجموعة الإناث .
والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط "ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

2- القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد فى اختبار للذكاء .

س	27	31	38	25	36	21	39
ص	15	23	27	21	30	19	-

حيث س هي مجموعة الذكور ، ص مجموعة الإناث .

والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط
"ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

3- القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين
من الأفراد في اختبار يقيس القدرة على التركيز .

س	9	5	8	6	7	4	17
ص	2	4	10	5	7	8	-

حيث س هي مجموعة الذكور ، ص مجموعة الإناث .

والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط
"ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

4- القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين
من الأفراد في اختبار مادة الإحصاء .

10	9	7	2	3	5	4	8	س
-	-	17	9	4	1	5	6	ص

حيث س هي مجموعة الذكور ، ص مجموعة الإناث .

والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة من خلال التحقق من شروط
"ت" مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

5- القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين
من الأفراد في اختبار مادة الحاسب الآلي .

17	12	14	13	9	س
17	8	9	3	8	ص

حيث س هي مجموعة الذكور ، ص مجموعة الإناث .

والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة
إحصائية أم لا عند مستوى دلالة 0.05

6- القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين
من الأفراد في اختبار للاستيعاب .

12	11	15	15	7	س
6	7	8	2	7	ص

حيث س هي مجموعة الذكور ، ص مجموعة الإناث .

والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

7- قمت بتطبيق اختبار على عينة قوامها 10 من الطلاب ثم تم تدريبهم على طريقة الاختبار لمدة أسبوعين وتم إجراء الاختبار مرة أخرى والجدول التالي يوضح درجات الطلاب في الاختبارين :

28	16	18	17	22	15	32	25	27	30	درجات الاختبار الأول
14	10	12	26	9	24	16	28	18	25	درجات الاختبار الثاني

والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

8- القيم التالية تعبر عن الدرجات التي حصل عليها مجموعتين من الأفراد في اختبار للذكاء .

20	19	13	48	19	32	22	17	35	س
-	-	7	2	14	10	9	3	11	ص

حيث س هي مجموعة الذكور ، ص مجموعة الإناث .
والمطلوب :

حساب قيمة "ت" بالطريقة المناسبة مع بيان إذا كانت دالة إحصائية أم لا عند مستوى دلالة 0.05

الفصل التاسع اختبار كا²

مقدمه :

أولاً : الطريقة العامة لحساب كا²

ثانياً : تحديد مدى دلالة كا² من عدمه .

ثالثاً : الطريقة العامة لحساب كا² من الجدول التكرارى 2×1 .

رابعاً : الطريقة المختصرة لحساب كا² من الجدول التكرارى 2×1 .

خامساً : الطريقة العامة لحساب كا² من الجدول التكرارى 1×ن .

سادساً : الطريقة العامة لحساب كا² من الجدول التكرارى 2×2 .

سابعاً : الطريقة المختصرة لحساب كا² من الجدول التكرارى 2×2 .

ثامناً : الطريقة العامة لحساب كا² من الجدول التكرارى ن×ن .

ثالثاً : حساب كا² لدلالة فروق النسب المرتبطة .

مقدمه :

ترجع النشأة الأولى لاختبار كا² إلى البحث الذي نشره كارل بيرسون في أوائل القرن العشرين وهي تعد من أهم اختبارات الدلالة الإحصائية وأكثرها شيوعاً لأنها لا تعتمد على شكل التوزيع ولذا فهي تعد من المقاييس اللابارامترية أي مقاييس التوزيعات الحرة ولأنها تحسب لكل خلية من خلايا أى جدول تكرارى ثم تجميع القيم الجزئية للحصول على القيمة الكارلية لـ كا² .
وتستخدم كا² لحساب دلالة فروق التكرار أو البيانات العددية التي يمكن تحويلها إلى تكرار مثل النسب والاحتمال .

الطريقة العامة لحساب كا²

$$كا^2 = \frac{(تو - تم)^2}{تو}$$

حيث :

تو : هو التكرار الواقعى الذي يحدث بالفعل والموجود بالجدول .
تم : هو التكرار المتوقع حدوثه ويختلف حسابه باختلاف نوع الجدول المطلوب حساب كا² منه .

تحديد مدى دلالة كا² من عدمه

فى جميع الحالات نخرج من الحسابات بقيمة كا² المحسوبة نقارنها بقيمة كا² الجدولية كالتالى :

- إذا كانت $كا^2$ المحسوبة $< كا^2$ الجدولية فإن $كا^2$ تكون دالة إحصائية .
- إذا كانت $كا^2$ المحسوبة $> كا^2$ الجدولية فإن $كا^2$ ليست دالة إحصائية .

حالات حساب $كا^2$ من الجداول المختلفة :

1- الحالة الأولى : الطريقة العامة لحساب $كا^2$ من الجدول التكراري 2×1 :

يتكون الجدول 2×1 من صف واحد وعمودين دون خلايا المجموع إن وجدت بالجدول .
ولحساب قيمة $كا^2$ فى هذا الجدول تحسب من القانون العام :

$$كا^2 = \frac{(تو - تم)^2}{تو}$$

حيث $تم$ هنا تساوى متوسط التكرارات الواقعية الموجودة بالجدول .

مثال :

الجدول التالي يوضح آراء 80 شخص فى استبيان دار حول رفض أو قبول قضية الزواج العرفي .

الرأي	موافق	غير موافق	مج
التكرار	60	20	80

والمطلوب حساب قيمة χ^2 مع بيان مدى دلالتها إحصائياً عند

مستوى دلالة 0.05 ؟

الحل :

حساب التكرار المتوقع (تم) :

$$20 + 60$$

$$40 = \frac{\quad}{2} = \text{تم}$$

حساب χ^2 المحسوبة :

نكون الجدول التالي :

$\frac{(ت_و - ت_م)^2}{ت_م}$	$(ت_و - ت_م)^2$	ت_و - ت_م	ت_م	ت_و
10	400	20	40	6
10	400	20-	40	20
20	مجموع	-	-	-

من الجدول مباشرة فإن مجموع العمود الأخير يعطينا قيمة χ^2

كا² المحسوبة = 20 .

حساب χ^2 الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

$$\text{درجة الحرية} = \text{عدد الأعمدة} - 1 = 2 - 1 = 1$$

$$\text{مستوى الدلالة} = 0.05 .$$

بالبحث فى جداول كا² عند درجة حرية = 1 ومستوى دلالة 0.05

$$\text{نجد قيمة كا}^2 \text{ الجدولية} = 3.841 .$$

تحديد مدى دلالة كا² :

نقارن قيمة كا² المحسوبة بقيمة كا² الجدولية نجد أن :

$$\text{قيمة كا}^2 \text{ المحسوبة} = 20 < \text{قيمة كا}^2 \text{ الجدولية} = 3.841$$

لذا فان دالة إحصائية عند مستوى دلالة 0.05 .

2- الحالة الثانية : الطريقة المختصرة لحساب كا² من الجدول

التكراري 2×1 :

لحساب قيمة كا² فى هذا الجدول بالطريقة المختصرة فان قيمة كا²

من العلاقة :

$$\text{كا}^2 = \frac{(t_1 - t_2)^2}{t_1 + t_2}$$

حيث ت₁ هو التكرار الأكبر و ت₂ هي التكرار الأصغر .

مثال :

الجدول التالي يوضح آراء 80 شخص فى استبيان دار حول

رفض أو قبول قضية الزواج العرفي .

الرأي	موافق	غير موافق	مج
التكرار	60	20	80

والمطلوب حساب قيمة كا² بالطريقة المختصرة مع بيان مدى دلالتها إحصائياً عند مستوى دلالة 0.05 ؟

الحل :

حساب كا² المحسوبة :

$$20 = \frac{1600}{80} = \frac{2(20 - 60)^2}{20 + 60} = كا^2$$

حساب كا² الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

$$درجة الحرية = عدد الأعمدة - 1 = 2 - 1 = 1$$

$$مستوى الدلالة = 0.05 .$$

بالبحث في جداول كا² عند درجة حرية = 1 ومستوى دلالة 0.05

$$نجد قيمة كا² الجدولية = 3.841 .$$

تحديد مدى دلالة كا² :

نقارن قيمة كا² المحسوبة بقيمة كا² الجدولية نجد أن

$$قيمة كا² المحسوبة = 20 < قيمة كا² الجدولية = 3.841$$

لذا فإن دالة إحصائية عند مستوى دلالة 0.05 .

3- الحالة الثالثة : الطريقة العامة لحساب χ^2 من الجدول

التكراري $1 \times n$:

يتكون الجدول $1 \times n$ من صف واحد وعدد (ن) عمود دون خلايا المجموع إن وجدت بالجدول .

ولحساب قيمة χ^2 في هذا الجدول تحسب من القانون العام :

$$\chi^2 = \frac{(تو - تم)^2}{تو}$$

حيث $ت م$ هنا تساوى متوسط التكرارات الواقعية الموجودة بالجدول.

مثال :

الجدول التالي يوضح آراء 30 شخص في استبيان دار حول قضية الزواج العرفي .

الرأي	موافق	لا أدرى	معارض	مجم
التكرار	12	2	16	30

والمطلوب حساب قيمة χ^2 مع بيان مدى دلالتها إحصائياً عند مستوى دلالة 0.05 ؟

الحل :

حساب التكرار المتوقع (تم) :

$$10 = \frac{16 + 2 + 12}{3} = \text{تم}$$

حساب ك² المحسوبة :

نكون الجدول التالي :

$\frac{(تو - تم)^2}{تم}$	$(تو - تم)^2$	تو - تم	تم	تو
0.4	4	2	10	12
6.4	64	8-	10	2
3.6	36	6	10	16
10.4	مجموع	-	-	-

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة ك²

$$\text{ك}^2 \text{ المحسوبة} = 10.4 .$$

حساب ك² الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

$$\text{درجة الحرية} = \text{عدد الأعمدة} - 1 = 3 - 1 = 2$$

$$\text{مستوى الدلالة} = 0.05 .$$

بالبحث فى جداول كا² عند درجة حرية = 2 ومستوى دلالة 0.05
نجد قيمة كا² الجدولية = 5.991 .

تحديد مدى دلالة كا² :

نقارن قيمة كا² المحسوبة بقيمة كا² الجدولية نجد أن
قيمة كا² المحسوبة = 10.4 < قيمة كا² الجدولية = 5.991
لذا فان كا² دالة إحصائية عند مستوى دلالة 0.05 .

4- الحالة الرابعة : الطريقة العامة لحساب كا² من الجدول
التكراري 2×2 :

يتكون الجدول 2×2 من صفين وعمودين دون خلايا المجموع إن
وجدت بالجدول .
ولحساب قيمة كا² فى هذا الجدول تحسب من القانون العام :

$$\text{كا}^2 = \frac{(ت_و - ت_م)^2}{ت_و}$$

وتحسب ت_م لكل خلية فى هذا الجدول على حده من العلاقة :
مجموع الصف × مجموع العمود

$$ت_م = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلى}}$$

مثال :

الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين .

النوع الفكرة	ذكور	إناث	المجموع
مؤيد	35	37	72
معارض	14	34	48
المجموع	49	71	120

والمطلوب حساب قيمة χ^2 مع بيان مدى دلالتها إحصائياً عند مستوى دلالة 0.05 ؟

الحل :

حساب التكرار المتوقع (تم) :

$$29.4 = \frac{49 \times 72}{120} = \text{تم للخلية الأولى (35)}$$

$$42.6 = \frac{71 \times 72}{120} = \text{تم للخلية الثانية (37)}$$

$$19.6 = \frac{49 \times 48}{120} = \text{تم للخلية الثالثة (14)}$$

$$28.4 = \frac{71 \times 48}{120} = \text{تم للخلية الرابعة (34)}$$

حساب كا² المحسوبة :

نكون الجدول التالي :

$\frac{(تو - تم)^2}{تم}$	$(تو - تم)^2$	تو - تم	تم	تو
1.06	31.36	5.6	29.4	35
0.74	31.6	5.6-	42.6	37
1.6	31.36	5.6-	19.6	14
1.1	31.36	5.6	28.4	34
4.5	مجموع	-	-	-

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة كا²

$$\text{كا}^2 \text{ المحسوبة} = 4.5 .$$

حساب كا² الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

$$\text{درجة الحرية} = (\text{عدد الصفوف} - 1) \times (\text{عدد الأعمدة} - 1)$$

$$1 = 1 \times 1 = (1 - 2) \times (1 - 2) =$$

$$\text{مستوى الدلالة} = 0.05 .$$

بالبحث فى جداول كا² عند درجة حرية = 1 ومستوى دلالة 0.05
نجد قيمة كا² الجدولية = 3.841
تحديد مدى دلالة كا² :

نقارن قيمة كا² المحسوبة بقيمة كا² الجدولية نجد أن :
قيمة كا² المحسوبة = 4.5 < قيمة كا² الجدولية = 3.841
لذا فان كا² دالة إحصائية عند مستوى دلالة 0.05 .

5- الحالة الخامسة : الطريقة المختصرة لحساب كا² من الجدول

التكراري 2×2 :

يتكون الجدول 2×2 من صفين وعمودين دون خلايا المجموع إن
وجدت بالجدول .

ولحساب قيمة كا² فى هذا الجدول بالطريقة المختصرة نطبق
القانون التالي :

$$\text{كا}^2 = \text{فاى}^2 \times \text{ن}$$

حيث :

فاى : هو معامل ارتباط فاى والذى يحسب من العلاقة :

$$\text{فاى} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\sqrt{\text{هـ} \times \text{و} \times \text{ز} \times \text{ح}}}$$

حيث أ ، ب ، ج ، د ، هـ ، و ، ز ، ح ، ن

هم خلايا الجدول الرباعي الخلايا كما بالشكل التالي :

النوع الفكرة	ذكور	إناث	المجموع
أ	ب	ح	
جـ	د	ز	
هـ	و	ن	

مثال :

الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين .

النوع الفكرة	ذكور	إناث	المجموع
مؤيد	35	37	72
معارض	14	34	48
المجموع	49	71	120

والمطلوب حساب قيمة χ^2 مع بيان مدى دلالتها إحصائياً عند

مستوى دلالة 0.05 ؟

الحل :

حساب معامل فاي :

نعوض فى العلاقة :

$$\frac{أ \times د - ب \times ج}{\sqrt{هـ \times و \times ز \times ح}} = \text{فاى}$$

$$\frac{14 \times 37 - 34 \times 35}{\sqrt{72 \times 48 \times 71 \times 49}} = \text{فاى}$$

$$0.19 = \text{فاى}$$

حساب كا²:

$$\text{كا}^2 = \text{فاى}^2 \times \text{ن}$$

$$4.33 = 120 \times (0.19)^2 = \text{كا}^2$$

حساب كا² الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

$$\text{درجة الحرية} = (\text{عدد الصفوف} - 1) \times (\text{عدد الأعمدة} - 1)$$

$$1 = 1 \times 1 = (1 - 2) \times (1 - 2) =$$

$$\text{مستوى الدلالة} = 0.05 .$$

بالبحث فى جداول كا² عند درجة حرية = 1 ومستوى دلالة 0.05

$$\text{نجد قيمة كا}^2 \text{ الجدولية} = 3.841$$

تحديد مدى دلالة كا² :

نقارن قيمة كا² المحسوبة بقيمة كا² الجدولية نجد أن :
قيمة كا² المحسوبة = 4.33 < قيمة كا² الجدولية = 3.841
لذا فان كا² دالة إحصائية عند مستوى دلالة 0.05 .

6- الحالة السادسة : الطريقة العامة لحساب كا² من الجدول

التكراري ن×ن :

يتكون الجدول ن×ن من عدد (ن) من الصفوف وعدد (ن) من الأعمدة دون خلايا المجموع إن وجدت بالجدول .
ولحساب قيمة كا² في هذا الجدول تحسب من القانون العام :

$$\text{كا}^2 = \frac{(ت_و - ت_م)^2}{ت_و}$$

وتحسب ت_م لكل خلية في هذا الجدول على حده من العلاقة :

$$\text{ت}_م = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلي}}$$

مثال :

الجدول التالي يوضح العلاقة بين المتغيرين النوع وتأييد برنامج تليفزيوني معين .

المجموع	أرفض جداً	أرفض نوعاً ما	لا أدري	موافق نوعاً ما	موافق جداً	الفكرة
						النوع
88	5	28	13	37	5	ذكور
53	5	20	8	17	3	إناث
141	10	48	21	54	8	المجموع

والمطلوب حساب قيمة χ^2 مع بيان مدى دلالتها إحصائياً عند مستوى دلالة 0.05 ؟

الحل :

حساب التكرار المتوقع (تم) :

$$5 = \frac{8 \times 88}{141} = \text{تم للخلية الأولى (5)}$$

$$33.7 = \frac{54 \times 88}{141} = \text{تم للخلية الثانية (37)}$$

$$13.1 = \frac{21 \times 88}{141} = \text{تم للخلية الثالثة (13)}$$

$$29.95 = \frac{48 \times 88}{141} = \text{تم للخلية الرابعة (28)}$$

$$6.24 = \frac{10 \times 88}{141} = \text{تم للخلية الخامسة (5)}$$

$$3 = \frac{8 \times 53}{141} = \text{تم للخلية السادسة (3)}$$

$$20.29 = \frac{54 \times 53}{141} = \text{تم للخلية السابعة (17)}$$

$$7.89 = \frac{21 \times 53}{141} = \text{تم للخلية الثامنة (8)}$$

$$18 = \frac{48 \times 53}{141} = \text{تم للخلية التاسعة (28)}$$

$$3.75 = \frac{10 \times 53}{141} = \text{تم للخلية العاشرة (5)}$$

حساب كا² المحسوبة :

نكون الجدول التالي :

$\frac{(ت_و - ت_م)^2}{ت_م}$	$(ت_و - ت_م)^2$	ت_و - ت_م	ت_م	ت_و
-----------------------------	-----------------	-----------	-----	-----

0	0	5	37	5
0.32	10.9	3.3	33.7	37
0	0.01	0.1-	13.1	13
0.13	3.8	1.59-	29.95	28
0.24	1.5	1.24-	6.24	5
0	0	0	3	3
0.53	10.8	3.29-	20.29	17
0	0.01	0.11	7.89	8
0.22	4	2	18	20
0.42	1.56	1.25	3.75	5
1.86	مجموع	-	-	-

من الجدول مباشرة فان مجموع العمود الأخير يعطينا قيمة χ^2

كما المحسوبة = 1.86 .

حساب χ^2 الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

درجة الحرية = (عدد الصفوف - 1) × (عدد الأعمدة - 1)

$$4 = 4 \times 1 = (1 - 5) \times (1 - 2) =$$

مستوى الدلالة = 0.05 .

بالبحث فى جداول χ^2 عند درجة حرية = 4 ومستوى دلالة 0.05

نجد قيمة χ^2 الجدولية = 9.488

تحديد مدى دلالة χ^2 :

نقارن قيمة χ^2 المحسوبة بقيمة χ^2 الجدولية نجد أن :

قيمة كا² المحسوبة = 1.86 > قيمة كا² الجدولية = 9.488
لذا فان كا² ليست دالة إحصائية عند مستوى دلالة 0.05 .

7- الحالة السابعة : حساب كا² لدلالة فروق النسب المرتبطة
نحسب قيمة كا² لدلالة فروق النسب المرتبطة بالجدول الرباعي
الخلايا 2×2 من العلاقة :

$$كا^2 = \frac{(ب - ج)^2}{ب + ج}$$

حيث أن ب ، ج هم خلايا بالجدول الرباعي كما بالشكل :

ب	أ
د	ج

مثال :

احسب قيمة كا² لدلالة فروق النسب المرتبطة التالية مع بيان مدى دلالتها إحصائيا عند مستوى دلالة 0.05 .

مج	إناث	ذكور	النوع الفكرة
40	15	25	مؤيد

60	55	5	معارض
100	70	30	مج

الحل :

حساب قيمة كا² المحسوبة :

$$\frac{2(5 - 15)}{5 + 15} = كا^2$$

$$5 = كا^2$$

حساب كا² الجدولية :

لحسابها يلزم حساب درجة الحرية ومستوى الدلالة :

$$\text{درجة الحرية} = (\text{عدد الصفوف} - 1) \times (\text{عدد الأعمدة} - 1)$$

$$1 = 1 \times 1 = (1 - 2) \times (1 - 2) =$$

$$\text{مستوى الدلالة} = 0.05 .$$

بالبحث فى جداول كا² عند درجة حرية = 1 ومستوى دلالة 0.05

$$\text{نجد قيمة كا² الجدولية} = 3.841$$

تحديد مدى دلالة كا² :

نقارن قيمة كا² المحسوبة بقيمة كا² الجدولية نجد أن :

$$\text{قيمة كا² المحسوبة} = 5 < \text{قيمة كا² الجدولية} = 3.841$$

لذا فان χ^2 دالة إحصائية عند مستوى دلالة 0.05 .

تمارين

1- من الجدول الرباعي التالي :

س	ص	نعم	لا	مج
مؤيد		25	15	40
معارض		23	27	50
مج		48	42	90

احسب قيمة كا² في كل من الحالات التالية :

- بالقانون العام
 - بالطريقة المختصرة
- ثم بين مدى دلالتها إحصائيا عند مستوى دلالة 0.05 .

2- احسب كا² من الجدول التالي :

ثم بين مدى دلالتها إحصائيا عند مستوى 0.05

ذكور	4	6	8	3	6
إناث	2	10	25	9	14

3- من الجدول التالي :

مج	إناث	ذكور	الجنس
			الإيجابية
54	22	32	موافق
24	10	14	معارض
12	8	4	محايد
90	40	50	مج

احسب قيمة كا²

ثم بين مدى دلالتها إحصائيا عند مستوى دلالة 0.05 .

4- احسب كا² لدلالة فروق النسب المرتبطة التالية مع بيان دلالتها الإحصائية .

10	30
27	23

الفصل العاشر معاملات الارتباط - الانحدار

- أولاً : الارتباط ومعناه .
- ثانياً : أنواع الارتباط .
- ثالثاً : معامل الاقتران .
- رابعاً : معامل فاي .
- خامساً : معامل التوافق .
- سادساً : معامل ارتباط بيرسون .
- سابعاً : معامل ارتباط الرتب لسبيرمان .
- ثامناً : معنى الانحدار .
- تاسعاً : معادلة خط انحدار ص/س .
- عاشراً : معادلة خط انحدار س/ص .

الارتباط ومعناه :

تركز عدد من البحوث الاجتماعية على تحليل العلاقة بين أكثر من متغير حيث يهتم الباحث بتحديد كيف وإلى أي مدى يرتبط متغيرات أو أكثر، والإحصاءات المستخدمة في التحليلات ثنائية المتغير، فالمنطق متشابه إلى حد كبير وإن كانت الإحصاءات المستخدمة في دراسة العلاقات متعددة المتغير تتسم بدرجة كبيرة من التعقيد.

وعند تحليل العلاقة بين متغيرين يهتم الباحث بالإجابة عن ثلاثة تساؤلات هل ترتبط هذه المتغيرات ؟ وما هو اتجاه وشكل الارتباط الموجود ؟ هل هناك احتمال أن يكون الارتباط الذي تمت ملاحظته بين حالات العينة أحد خصائص المجتمع البحثي أم أن هذا الارتباط هو نتاج لصغر حجم العينة التي قد تكون غير ممثلة للمجتمع البحثي ؟

يمكن تحديد الارتباط بين متغيرين من خلال استخدام مجموعة من الإحصاءات تعرف باسم معاملات الارتباط ومعامل الارتباط هو رقم يلخص التحسن في تخمين القيم على متغير واحد لأي حالة على أساس معرفة قيم المتغير الثاني، فكلما ارتفع المعامل قوي الارتباط ، ومن ثم تحسنت قردتنا التنبؤية أو التفسيرية. وتتراوح معاملات الارتباط بين صفر وواحد (أو -1)، وتشير القيم التي تقترب من 1 إلى وجود ارتباط قوي نسبياً أما تلك التي تقترب من صفر فتشير إلى ارتباط ضعيف نسبياً. ويتطلب كل مستوى قياس

أنواع مختلفة من الحسابات وبالتالي فكل من هذه المستويات اختبارات ارتباط مختلفة.

إضافة إلى حجم الارتباط يهتم الباحث بمعرفة اتجاه العلاقة بين المتغيرين فهل هي علاقة طردية أو عكسية، وتجدر الإشارة هنا إلى أن مفهوم الاتجاه ليس له معنى على مستوى القياس الأسمى، حيث إن الأرقام على هذا المستوى من القياس مجرد عناوين للفئات، وبالتالي لا تتغير إشارات معاملات الارتباط الاسمية فكلها موجبة وتشير إلى مدى قوة الارتباط ، أما على مستوى قياس الفترة فإن الإشارات تتغير ولها دلالات هندسية على درجة عالية نسبياً من التعقيد.

وأخيراً يهتم الباحث باختبارات الدلالة الإحصائية وهي الاختبارات التي توضح احتمال أن تكون العلاقات التي يلاحظها الباحث نتاج التحيز في عملية الاختبار بدلاً من أن تعكس علاقات موجودة فعلاً في مجمع البحث.

أنواع الارتباط :

بالطبع عرفنا أن قيمة معامل الارتباط محصورة في الفترة المغلقة [-1 ، 1] وتتحدد نوعية الارتباط من الجدول التالي :

نوع الارتباط	قيمة معامل الارتباط
ارتباط طردى تام	1+
ارتباط طردى قوى	من 0.7 إلى أقل من 1+

ارتباط طردى متوسط	من 0.4 إلى أقل من 0.7
ارتباط طردى ضعيف	من صفر إلى أقل من 0.4
الارتباط منعدم	صفر
ارتباط عكسي تام	1-
ارتباط عكسي قوى	من -0.7 إلى أقل من -1
ارتباط عكسي متوسط	من -0.04 إلى أقل من -0.7
ارتباط عكسي ضعيف	من صفر إلى أقل من -0.4

طرق حساب الارتباط :

1- معامل الاقتران :

يستخدم معامل الاقتران لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهما صفات والجدول المزدوج الذي يمثل العلاقة بينهما مكون من (4) خلايا فقط دون خلايا المجموع نستخدم القانون التالي لمعامل الاقتران :

$$\text{معامل الاقتران} = \frac{\text{أ} \times \text{د} - \text{ب} \times \text{ج}}{\text{أ} \times \text{د} + \text{ب} \times \text{ج}}$$

حيث أ ، ب ، ج ، د هم الخلايا الأربع للجدول رباعى الخلايا كما بالشكل :

أ	ب
ج	د

مثال :

قام أحد الباحثين بعمل بحث عن نسب المدخنين من النوعين الذكور والإناث فحصل على بيانات الجدول التالي :

النوع التدخين	ذكور	إناث	مج
يدخن	25	15	40
لا يدخن	5	55	60
مج	30	70	100

والمطلوب حساب قيمة معامل الارتباط بالطريقة المناسبة مع بيان نوع هذا الارتباط ؟

الحل :

الجدول مكون من أربعة خلايا فقط والمتغيران صفات لذا نستخدم معامل الاقتران :

$$\text{معامل الاقتران} = \frac{أ \times د - ب \times ج}{\dots}$$

$$أ \times د + ب \times ج$$

$$\frac{1300}{1450} = \frac{5 \times 15 - 55 \times 25}{5 \times 15 + 55 \times 25} = \text{معامل الاقتران}$$

$$0.89 = \text{معامل الاقتران}$$

تحديد نوع الارتباط :

ارتباط طردى قوى .

2- معامل فاي :

يستخدم معامل فاي لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهما صفات أيضاً والجدول المزدوج الذي يمثل العلاقة بينهم مكون من (4) خلايا فقط دون خلايا المجموع نستخدم القانون التالي لحساب لمعامل فاي :

$$\text{معامل فاي} = \frac{أ \times د - ب \times ج}{\sqrt{ح \times ز \times و \times ه}}$$

حيث أ ، ب ، ج ، د ، ه ، و ، ز ، ح

هم خلايا الجدول الرباعي الخلايا كما بالشكل التالي :

المجموع	إناث	ذكور	النوع
			الفكرة
ح	ب	أ	مؤيد
ز	د	ج	معارض
ن	و	هـ	المجموع

والسؤال الآن : متى يستخدم معامل الاقتران ومتى يستخدم معامل فاي رغم تشابههما في الشروط ؟
يستخدم معامل فاي إذا كنا نريد استخدام جميع خلايا الجدول أو إذا كنا نريد الحصول على القيمة الأقل لمعامل الارتباط أو الأدق أما بخلاف ذلك نستخدم معمل الاقتران .

مثال :

قام أحد الباحثين بعمل بحث عن نسب المدخنين من النوعين الذكور والإناث فحصل على بيانات الجدول التالي :

مج	إناث	ذكور	النوع
			التدخين
40	15	25	يدخن

60	55	5	لا يدخن
100	70	30	مج

والمطلوب حساب قيمة معامل الارتباط بالطريقة المناسبة للحصول على القيمة الأقل والأدق لمعامل الارتباط مع بيان نوع هذا الارتباط ؟

الحل :

الجدول مكون من أربعة خلايا فقط والمتغيران صفات والمطلوب الحصول على القيمة الأقل والأدق لمعامل الارتباط لذا نستخدم معامل فاي :

$$\text{معامل فاي} = \frac{أ \times د - ب \times ج}{\sqrt{هـ \times و \times ز \times ح}}$$

$$\text{معامل فاي} = \frac{5 \times 15 - 55 \times 25}{\sqrt{40 \times 60 \times 70 \times 30}}$$

$$\frac{1300}{\text{معامل فاي}}$$

2245

معامل فاي = 0.58

تحديد نوع الارتباط :

ارتباط طردي متوسط .

التعليق :

نلاحظ أن قيمة معامل الاقتران أكبر من قيمة معامل فاي لحساب قيمة الارتباط لنفس المثال حيث أن معامل فاي أدق من معامل الاقتران لأنه يستخدم جميع خلايا الجدول .

3- معامل التوافق :

يستخدم معامل التوافق لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهما صفات أيضاً والجدول المزدوج الذي يمثل العلاقة بينهم يزيد عدد خلاياه عن (4) خلايا دون خلايا المجموع ونستخدم القانون التالي لحساب قيمة معامل التوافق :

$$\text{معامل التوافق} = \sqrt{\frac{ج - ا}{ج}}$$

حيث تحسب (ج) من العلاقة :
مربع الخلية

$$\text{ج} = \text{مج} \div \text{مجموع صف الخلية} \times \text{مجموع عمود الخلية}$$

مثال :

قام أحد الباحثين بعمل بحث عن المدخنين ومدى تأثرهم بمشاهدة برنامج خمسة لصحتك فحصل على بيانات الجدول التالي :

مج	لا يدخن	يدخن	التدخين
			مشاهدة البرنامج
178	116	62	دائماً يشاهد البرنامج
193	176	17	غالباً يشاهد البرنامج
78	73	5	أحياناً يشاهد البرنامج
23	20	3	لا يشاهد البرنامج
472	385	87	مج

والمطلوب حساب قيمة معامل الارتباط بالطريقة المناسبة مع بيان نوع هذا الارتباط ؟

الحل :

الجدول تزيد عدد خلاياه عن أربعة خلايا والمتغيران صفات لذا نستخدم معامل التوافق :

$$\text{ج} - 1$$

- 303 -

$$\frac{\quad}{\quad} = \text{معامل التوافق} \\ \rightarrow$$

حيث تحسب (جـ) من العلاقة :

مربع الخلية

$$\text{جـ} = \text{مجـ} \frac{\quad}{\quad} \\ \text{مجموع صف الخلية} \times \text{مجموع عمود الخلية}$$

$$\frac{^2(5)}{87 \times 78} + \frac{^2(17)}{87 \times 193} + \frac{^2(62)}{87 \times 178} = \text{جـ}$$

$$\frac{^2(176)}{385 \times 193} + \frac{^2(116)}{385 \times 178} + \frac{^2(3)}{87 \times 23} +$$

$$\frac{^2(20)}{385 \times 23} + \frac{^2(73)}{385 \times 78} +$$

$$+ 0.196 + 0.005 + 0.004 + 0.017 + 0.248 = \text{جـ} \\ 1.11 = 0.045 + 0.178 + 0.417$$

$$\sqrt{\frac{1 - 1.11}{1.11}} = \text{معامل التوافق}$$

معامل التوافق = 0.32

تحديد نوع الارتباط :

ارتباط طردي ضعيف .

4- معامل ارتباط بيرسون :

يستخدم معامل ارتباط بيرسون لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهما متغيرات كمية ويشترط تساوى عدد حالات كلا من المتغيرين ونستخدم القانون التالي لحساب قيمة معامل ارتباط بيرسون :

ر : هو معامل ارتباط بيرسون ويحسب من العلاقة :

ن مجـ (س×ص) - مجـ س × مجـ ص

$$r = \frac{\text{ن مجـ (س×ص) - مجـ س × مجـ ص}}{\sqrt{[\text{ن مجـ س}^2 - (\text{مجـ س})^2] \times [\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

مثال :

الجدول التالي يوضح درجات مجموعة من الطلاب فى اختبار تم إجراؤه على نفس الطلاب مرتين متتاليتين والمطلوب حساب قيمة معامل الارتباط لبيرسون بين درجات الاختبارين ؟

2	8	9	5	3	درجة الاختبار الأول
3	4	7	6	4	درجة الاختبار الأول

الحل :

نفترض أن درجات الاختبار الأول هي "س" ودرجات الاختبار الثاني هي "ص" ثم نكون الجدول التالي :

ص ²	س ²	س×ص	ص	س
16	9	12	4	3
36	25	30	6	5
49	81	63	7	9
16	64	32	4	8
9	4	6	3	2
126	183	143	24	27

حساب معامل الارتباط لبيرسون :

$$n \text{ مـجـ } (س \times ص) - \text{ مـجـ } س \times \text{ مـجـ } ص$$

$$r = \frac{[n \text{ مـجـ } س^2 - (\text{ مـجـ } س)^2] \times [n \text{ مـجـ } ص^2 - (\text{ مـجـ } ص)^2]}{\sqrt{[n \text{ مـجـ } س^2 - (\text{ مـجـ } س)^2] \times [n \text{ مـجـ } ص^2 - (\text{ مـجـ } ص)^2]}}$$

نعوض في المعادلة السابقة :

$$24 \times 27 - 143 \times 5$$

$$r = \frac{24 \times 27 - 143 \times 5}{\sqrt{[24^2 - 126 \times 5] \times [27^2 - 183 \times 5]}}$$

$$r = 0.668$$

تحديد نوع الارتباط :
ارتباط طردي متوسط .

5- معامل ارتباط الرتب لسبيرمان :
يستخدم معامل ارتباط الرتب لسبيرمان لحساب قيمة معامل الارتباط عندما يكون المتغيران المراد قياس الارتباط بينهما متغيرات كمية ويشترط تساوى عدد حالات كلا من المتغيرين أيضاً ونستخدم القانون التالى لحساب قيمة معامل ارتباط الرتب لسبيرمان:

$$r = \frac{6 \text{ مج ف}^2}{n(n-1)} - 1$$

حيث :

ر : معامل ارتباط الرتب لسبيرمان

ف = رتب المتغير الأول - رتب المتغير الثانى

ن : عدد الحالات

مثال :

الجدول التالى يوضح درجات مجموعة من الطلاب فى اختبار تم إجراؤه على نفس الطلاب مرتين متتاليتين والمطلوب حساب قيمة معامل ارتباط الرتب لسبيرمان بين درجات الاختبارين ؟

2	8	9	5	3	درجة الاختبار الأول
3	4	7	6	4	درجة الاختبار الأول

الحل :

نفترض أن درجات الاختبار الأول هي "س" ودرجات الاختبار

الثاني هي "ص" ثم نكون الجدول التالي :

مع ملاحظة أنه إذا تم ترتيب قيم س تصاعدي لابد من ترتيب قيم

ص تصاعدي والعكس بالعكس .

وهنا سوف نرتب القيم تصاعدي .

مع ملاحظة أنه إذا تساوى عدان أو أكثر فى القيمة يأخذ كل

منهم متوسط ترتيبهم .

فمثلاً المتغير ص يوجد به رقمان متساويان هما (4،4) وترتيبهما

(2،3) إذا يأخذ كل منهم متوسط الترتيب $2.5 = 2/5 = 2/(3+2)$

.

س	ص	رتب س	رتب ص	ف	ف ²
3	4	2	2.5	0.5-	0.25
5	6	3	4	1-	1
9	7	5	5	0	0
8	4	4	2.5	1.5	2.25
2	3	1	1	0	0
مج					
3.5					

حساب معامل ارتباط الرتب لسبيرمان :

$$r = 1 - \frac{6 \text{ مج ف}^2}{n(n^2 - 1)}$$

$$\frac{3.5 \times 6}{(1 - 25) 5} - 1 = r$$

$$\frac{21}{24 \times 5} - 1 = r$$

$$0.825 = 0.175 - 1 = r$$

تحديد نوع الارتباط :

ارتباط طردي قوى .

معنى الانحدار :

يهدف الانحدار إلى الإفادة من الارتباط في التنبؤ ، فإذا علمنا معامل ارتباط درجات اختبار الحساب بدرجات اختبار الجبر ، وعلمنا درجة أى طالب فى اختبار الحساب فإننا نستطيع أن نتنبأ بدرجةه فى الجبر وإذا علمنا درجة أى طالب آخر فى اختبار الجبر فإننا نستطيع أن نتنبأ بدرجةه فى الحساب .

وقد سمي هذا المفهوم الإحصائي بالانحدار لأنه ينحدر فى تقديره الدرجات المختلفة نحو المتوسط ولذا تسمى معادلات الانحدار أحياناً بمعادلات خطوط المتوسطات .

حساب الانحدار :

تعتمد معادلات الانحدار معاملات الارتباط وعلى الانحرافات المعيارية وعلى المتوسطات فهى بذلك تستعين بأهم المقاييس الإحصائية فى حسابها لهذا التنبؤ .

أولاً : معادلة خط انحدار ص/س :

تتلخص معادلة خط انحدار ص على س فى الصورة التالية :

$$ص = ر \times \frac{ع ص}{ع س} + (س م - س) + م ص$$

حيث :

ر = معامل ارتباط بيرسون ويحسب من العلاقة :

$$r = \frac{\sum (س \times ص) - \sum س \times \sum ص}{\sqrt{[\sum س^2 - \frac{(\sum س)^2}{ن}] \times [\sum ص^2 - \frac{(\sum ص)^2}{ن}]}}$$

ع ص = الانحراف المعياري لقيم ص ويحسب من العلاقة :

$$ع ص = \sqrt{\frac{\sum ص^2 - \frac{(\sum ص)^2}{ن}}{ن}}$$

ع س = الانحراف المعياري لقيم س ويحسب من العلاقة :

$$ع س = \sqrt{\frac{\sum س^2 - \frac{(\sum س)^2}{ن}}{ن}}$$

م س = متوسط قيم المتغير س

م ص = متوسط قيم المتغير ص

مثال :

الجدول التالي يوضح درجات خمس طلاب في اختبارين الأول س والثاني ص والمطلوب حساب معادلة خط انحدار ص/س ثم حساب قيمة ص عندما س = 10 .

الأفراد	أ	ب	ج	د	هـ
س	2	3	7	18	20
ص	5	7	6	12	10

الحل :

حساب معامل ارتباط بيرسون :

نكون الجدول التالي :

س	ص	س × ص	س ²	ص ²
2	5	10	4	25
3	7	21	9	49
7	6	42	49	36
18	12	216	324	144
20	10	200	400	100
50	40	489	786	354

$$r = \frac{n \sum (س \times ص) - \sum س \times \sum ص}{\sqrt{[n \sum س^2 - (\sum س)^2] \times [n \sum ص^2 - (\sum ص)^2]}}$$

$$r = \frac{50 \times 489 - 20 \times 40}{\sqrt{[50 \times 786 - (20)^2] \times [50 \times 354 - (40)^2]}}$$

$$40 \times 50 - 489 \times 5$$

$$\sqrt{\frac{40 \times 50 - 489 \times 5}{[{}^2(40) - 354 \times 5] \times [{}^2(50) - 786 \times 5]}} = r$$

$$r = 0.9$$

حساب المتوسطات :

$$10 = \frac{50}{5} = \frac{\text{مجموع س}}{\text{ن}} = \text{م س}$$

$$8 = \frac{40}{5} = \frac{\text{مجموع ص}}{\text{ن}} = \text{م ص}$$

حساب الانحراف المعياري :

نكون الجدول التالي :

ح ² ص	ح ص	ح ² س	ح س	ص	س
9	3-	64	8-	5	2

1	1-	49	7-	7	3
4	2-	9	3-	6	7
16	4	64	8	12	18
4	2	100	10	10	20
34		286			

$$7.56 = \frac{286}{5} \sqrt{\quad} = \frac{\text{مجـ ح}^2 \text{س}}{\text{ن}} \sqrt{\quad} = \text{ع س}$$

$$2.61 = \frac{34}{5} \sqrt{\quad} = \frac{\text{مجـ ح}^2 \text{ص}}{\text{ن}} \sqrt{\quad} = \text{ع ص}$$

حساب معادلة خط انحدار ص/س :

$$\text{ص} = \text{ر} \times \frac{\text{ع ص}}{\text{ع س}} + (\text{س} - \text{م س})$$

$$\text{ص} = 0.9 \times \frac{2.61}{7.56} + (\text{س} - 10)$$

$$\text{ص} = 0.31 + (\text{س} - 10)$$

$$ص = 0.31 س - 3.1 + 8$$

معادلة خط انحدار ص/س هي

$$ص = 0.31 س + 4.9$$

عندما س = 10 نستطيع التنبؤ بقيمة ص كالتالى :

$$ص = 8 = 4.9 + 10 \times 0.31$$

ثانياً : معادلة خط انحدار س/ص :

تتلخص معادلة خط انحدار س على ص فى الصورة التالية :

$$س = ر \times \frac{ع س}{ع ص} + (ص - م ص)$$

حيث :

ر = معامل ارتباط بيرسون ويحسب من العلاقة :

$$ر = \frac{ن مج (س \times ص) - مج س \times مج ص}{\sqrt{[ن مج س^2 - (مج س)^2] \times [ن مج ص^2 - (مج ص)^2]}}$$

ع ص = الانحراف المعياري لقيم ص ويحسب من العلاقة

$$ع ص = \sqrt{\frac{\text{مجم ح}^2 ص}{ن}}$$

ع س = الانحراف المعياري لقيم س ويحسب من العلاقة

$$ع س = \sqrt{\frac{\text{مجم ح}^2 س}{ن}}$$

م س = متوسط قيم المتغير س

م ص = متوسط قيم المتغير ص

مثال :

الجدول التالي يوضح درجات خمس طلاب في اختبارين الأول س والثاني ص والمطلوب حساب معادلة خط انحدار س/ص ثم حساب قيمة س عندما س = 8 .

الأفراد	أ	ب	ج	د	هـ
س	2	3	7	18	20
ص	5	7	6	12	10

الحل :

حساب معامل ارتباط بيرسون :

نكون الجدول التالي :

ص ²	س ²	س × ص	ص	س
25	4	10	5	2
49	9	21	7	3
36	49	42	6	7
144	324	216	12	18
100	400	200	10	20
354	786	489	40	50

ن مجـ (س×ص) - مجـ س × مجـ ص

$$= \frac{[\text{ن مجـ س}^2 - (\text{مجـ س})^2] \times [\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}{\sqrt{}}$$

$$40 \times 50 - 489 \times 5$$

$$= \frac{[\text{ن مجـ س}^2 - (\text{مجـ س})^2] \times [\text{ن مجـ ص}^2 - (\text{مجـ ص})^2]}{\sqrt{}}$$

$$= 0.9$$

حساب المتوسطات :

$$10 = \frac{50}{5} = \frac{\text{مجا س}}{\text{ن}} = \text{م س}$$

$$8 = \frac{40}{5} = \frac{\text{مجا ص}}{\text{ن}} = \text{م ص}$$

حساب الانحراف المعياري :

نكون الجدول التالي :

ح ² ص	ح ص	ح ² س	ح س	ص	س
9	3-	64	8-	5	2
1	1-	49	7-	7	3
4	2-	9	3-	6	7
16	4	64	8	12	18
4	2	100	10	10	20
34		286			

$$7.56 = \frac{\sqrt{286}}{5} = \frac{\sqrt{\text{مجا ح}^2 \text{ س}}}{\text{ن}} = \text{ع س}$$

$$2.61 = \frac{\sqrt{34}}{5} = \frac{\sqrt{\text{مجا ح}^2 \text{ ص}}}{\text{ن}} = \text{ع ص}$$

حساب معادلة خط انحدار س/ص :

$$\text{س} = \text{ر} \times \frac{\text{ع س}}{\text{ع ص}} + (\text{ص م} - \text{ص})$$

$$\text{س} = 0.9 \times \frac{7.56}{2.61} + (\text{ص} - 8) + 10$$

$$\text{س} = 2.6 (\text{ص} - 8) + 10$$

$$\text{س} = 2.6 - 20.8 + 10$$

معادلة خط انحدار س/ص هي

$$\text{س} = 2.6 \text{ ص} - 10.8$$

عندما ص = 8 نستطيع التنبؤ بقيمة س كالتالي :

$$\text{س} = 2.6 \times 8 - 10.8 = 10$$

تارين

1- احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط .

الأفراد	أ	ب	ج	د	هـ
علم النفس س	7	9	14	5	15
الصحة النفسية ص	11	13	15	6	10

2- احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط .

الأفراد	1	2	3	4	5	6	7	8	9	10
س	جيد	مقبول	جيد	ممتاز	جيد جدا	مقبول	جيد جدا	جيد جدا	ضعيف جدا	ضعيف
ص	جيد جدا	ممتاز	مقبول	مقبول	مقبول	جيد	جيد	ممتاز	ضعيف	ضعيف جدا

3- احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط .

10	9	8	7	6	5	4	3	2	1	الأفراد
6	7	5	3	4	2	10	11	9	8	درجات الإحصاء س
7	8	6	4	5	3	11	12	10	9	درجات علم النفس ص

4- احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع هذا الارتباط .

7	19	13	12	11	س
2	10	8	14	6	ص

5- من الجدول الرباعي التالي :

مج	لا	نعم	س ص
40	15	25	مؤيد
50	27	23	معارض
90	42	48	مج

احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع الارتباط ؟

6- احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع الارتباط ؟

6	3	8	6	4	ذكور
14	9	25	10	2	إناث

7- من الجدول التالي :

مج	إناث	ذكور	الجنس
			الإجابة
54	22	32	موافق
24	10	14	معارض
12	8	4	محايد
90	40	50	مج

احسب قيمة معامل الارتباط بالطريقة المناسبة ثم حدد نوع الارتباط ؟

8- احسب قيمة معامل الارتباط بالطريقة الأكثر دقة والأقل قيمة ثم حدد نوع الارتباط ؟

10	30
27	23

9- احسب معادلة خط انحدار ص/س

ثم احسب قيمة ص عندما س = 10

3	5	6	9	7	س
9	7	6	3	5	ص

10- احسب معادلة خط انحدار س/ص

ثم احسب قيمة س عندما ص = 10

3	5	6	9	7	س
9	7	6	3	5	ص

11- احسب معادلة خط انحدار ص/س

ثم احسب قيمة ص عندما س = 20

25	23	22	24	21	س
11	14	12	12	15	ص

12- احسب معادلة خط انحدار س/ص

ثم احسب قيمة س عندما ص = 10

25	23	22	24	21	س
11	14	12	12	15	ص

13- احسب معادلة خط انحدار ص/س

ثم احسب قيمة ص عندما س = 10

5	2	4	3	6	س
5	3	5	4	8	ص

14- احسب معادلة خط انحدار س/ص

ثم احسب قيمة س عندما ص = 20

5	2	4	3	6	س
5	3	5	4	8	ص

الفصل الحادي عشر الثبات والصدق

- أولاً : معنى الثبات .
- ثانياً : طرق حساب معامل الثبات .
 - طريقة إعادة الاختبار .
 - طريقة التجزئة النصفية .
- ثالثاً : معنى الصدق .
- رابعاً : قياس الصدق .
 - طريقة المقارنة الطرفية .

معنى الثبات :

إذا أجرى اختبار ما على مجموعة من الأفراد ورصدت درجات كل فرد في هذا الاختبار ثم أعيد إجراء نفس هذا الاختبار على نفس هذه المجموعة ورصدت أيضا درجات كل فرد ودلت النتائج على أن الدرجات التي حصل عليها الطلاب في المرة الأولى لتطبيق الاختبار هي نفس الدرجات التي حصل عليها هؤلاء الطلاب في المرة الثانية ، نستنتج من ذلك أن نتائج الاختبار ثابتة تماماً لأن نتائج القياس لم تتغير في المرة الثانية بل ظلت كما كانت قائمة في المرة الأولى .

حساب الثبات :

حساب معامل الارتباط هو خير طريقة لمقارنة هذه الدرجات التي حصل عليها الطلاب في الاختبارين .

ويحسب معمل الثبات من العلاقة التالية :

$$\text{معامل الثبات} = \frac{r^2}{r + 1}$$

حيث :

ر : هو معامل ارتباط بيرسون ويحسب من العلاقة :

$$r = \frac{n \text{ مـجـ (س} \times \text{ص) - مـجـ س} \times \text{مـجـ ص}}{\sqrt{[n \text{ مـجـ س}^2 - (\text{مـجـ س})^2] \times [n \text{ مـجـ ص}^2 - (\text{مـجـ ص})^2]}}$$

طرق حساب معامل الثبات :

1- طريقة إعادة الاختبار :

تقوم فكرة هذه الطريقة على إجراء الاختبار على مجموعة من الأفراد ثم إعادة إجراء نفس الاختبار على نفس مجموعة الأفراد بعد مضي فترة زمنية وهكذا يحصل كل فرد على درجة فى الإجراء الأول للاختبار وعلى درجة أخرى فى الإجراء الثانى للاختبار ، وعندما نرصد هذه الدرجات ونحسب معامل ارتباط درجات المرة الأولى بدرجات المرة الثانية فأتنا نحصل بذلك على معامل ثبات الاختبار .

مثال :

الجدول التالى يوضح درجات مجموعة من الطلاب فى اختبار تم إجراؤه على نفس الطلاب مرتين والمطلوب حساب قيمة معامل ثبات الاختبار ؟

2	8	9	5	3	درجة الاختبار الأول
3	4	7	6	4	درجة الاختبار الأول

الحل :

نفترض أن درجات الاختبار الأول هى "س" ودرجات الاختبار الثانى هى "ص" ثم نكون الجدول التالى :

ص ²	س ²	س×ص	ص	س
----------------	----------------	-----	---	---

16	9	12	4	3
36	25	30	6	5
49	81	63	7	9
16	64	32	4	8
9	4	6	3	2
126	183	143	24	27

حساب معامل الارتباط لبرسون :

$$n \text{ مجـ (س} \times \text{ص) - مجـ س} \times \text{مجـ ص}$$

$$r = \frac{[n \text{ مجـ س}^2 - (\text{مجـ س})^2] \times [n \text{ مجـ ص}^2 - (\text{مجـ ص})^2]}{\sqrt{[n \text{ مجـ س}^2 - (\text{مجـ س})^2] \times [n \text{ مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

نعوض في المعادلة السابقة :

$$24 \times 27 - 143 \times 5$$

$$r = \frac{24 \times 27 - 143 \times 5}{\sqrt{[24^2 - 126 \times 5] \times [27^2 - 183 \times 5]}}$$

$$r = 0.668$$

$$\frac{0.668 \times 2}{0.668 + 1} = \text{معامل الثبات}$$

معامل الثبات = 0.8

2- طريقة التجزئة النصفية :

تعتمد هذه الطريقة على تجزئة الاختبار إلى جزأين فقط بحيث يتكون الجزء الأول من الدرجات الفردية للاختبار ويتكون الجزء الثاني من الدرجات الزوجية للاختبار .

مثال :

الأسئلة								الأفراد
8	7	6	5	4	3	2	1	
0	0	0	1	1	1	1	1	1
0	0	1	1	0	1	1	1	2
0	0	0	1	1	0	1	1	3
1	1	1	1	0	1	1	1	4
0	0	1	0	0	1	1	1	5
1	1	0	0	1	1	1	1	6
0	0	1	1	0	1	1	1	7
0	1	1	1	1	1	1	1	8
0	0	0	0	1	1	1	1	9
1	1	1	1	1	1	1	1	10

الجدول السابق يوضح درجات عشرة طلاب في اختبار تم تقسيمه إلى ثماني أسئلة والمطلوب حساب قيمة معامل الثبات لدرجات الأسئلة الفردية والزوجية باستخدام طريقة التجزئة النصفية ؟

الحل :

نقوم بتجميع درجات الأسئلة الفردية على حدة ونسميها "س" ودرجات الأسئلة الزوجية على حدة ونسميها "ص" لكل طالب منفرداً ونضعها في الجدول التالي :

ص	ص	ص×ص	ص	س
---	---	-----	---	---

			الدرجات الزوجية	الدرجات الفردية
4	9	6	2	3
9	9	9	3	3
4	4	4	2	2
9	16	12	3	4
4	4	4	2	2
9	9	9	3	3
4	9	6	2	3
9	16	12	3	4
4	4	4	2	2
16	16	16	4	4
72	96	82	26	30

حساب معامل الارتباط لبيرسون :

$$r = \frac{N \text{ مجـ (س} \times \text{ص) - مجـ س} \times \text{مجـ ص}}{\sqrt{[N \text{ مجـ س}^2 - (\text{مجـ س})^2] \times [N \text{ مجـ ص}^2 - (\text{مجـ ص})^2]}}$$

$$r = \frac{26 \times 30 - 82 \times 10}{\sqrt{[26^2 - 82 \times 10] \times [30^2 - 96 \times 10]}}$$

نعوض في المعادلة السابقة :

$$26 \times 30 - 82 \times 10$$

$$r = \frac{26 \times 30 - 82 \times 10}{\sqrt{[26^2 - 82 \times 10] \times [30^2 - 96 \times 10]}}$$

$$r = 0.78$$

$$\frac{0.78 \times 2}{0.78 + 1} = \text{معامل الثبات}$$

$$0.88 = \text{معامل الثبات}$$

معنى الصدق :

الاختبار الصادق يقيس ما وضع لقياسه فاختبار الذكاء الذى يقيس الذكاء فعلاً اختبار صادق مثله فى ذلك كمثل المتر فى قياسه للأطوال والكيلو فى قياسه للأوزان والساعة فى قياسها للزمن وتختلف الاختبارات فى مستويات صدقها تبعاً لاقترابها أو ابتعادها من تقدير تلك الصفة التى تهدف إلى قياسها فاختبار الذكاء الذى يصل فى قياسه لتلك القدرة إلى مستوى 0.8 أصدق فى هذا القياس من أى اختبار آخر للذكاء لا يصل إلى هذا المستوى أى أنه أصدق مثلاً من الاختبار الذى يصل فى قياسه للذكاء إلى مستوى 0.5 .

ويحسب مستوى صدق الاختبار بمقارنة نتائجه بنتائج مقياس آخر دقيق لتلك الصفة ويسمى هذا المقياس بالميزان .

قياس الصدق : طريقة المقارنة الطرفية

تقوم هذه الطريقة على مقارنة متوسط درجات الأقوياء فى الميزان بمتوسط درجات الضعاف فى نفس ذلك الميزان بالنسبة لتوزيع درجات الاختبار ولذا سميت بالمقارنة الطرفية لاعتمادها على الطرف القوى الذى نسميه بأصحاب الميزان القوى والطرف الضعيف الذى نسميه أصحاب الميزان الضعيف .

ولحساب الدلالة الإحصائية للفرق بين أصحاب المستوى القوى والضعيف نستعين بالنسبة الحرجة :

$$\frac{1م - 2م}{\sqrt{\frac{ع^2}{2ن} + \frac{ع^2}{1ن}}} = \text{النسبة الحرجة}$$

حيث :

1م = متوسط درجات أصحاب الميزان الضعيف

2م = متوسط درجات أصحاب الميزان القوى

ع²₁ = تباين درجات أصحاب المستوى الضعيف

ع²₂ = تباين درجات أصحاب المستوى القوى

ن₁ = مجموع تكرارات أصحاب الميزان الضعيف = مج ك₁

ن₂ = مجموع تكرارات أصحاب الميزان القوى = مج ك₂

ويحسب المتوسط في البيانات المبوبة من العلاقة :

$$م = \frac{\text{مج (س × ك)}}{\text{مج ك}}$$

حيث "س" هي مركز الفئة وتحسب من العلاقة :

• س = (بداية الفئة الأولى + نهاية الفئة) / 2

• ك : هو التكرار

ويحسب التباين من العلاقة :

$$ع^2 = ل^2 \times \frac{مج - (ح^2 \times ك)}{مج ك} - \left[\frac{مج - (ح \times ك)}{مج ك} \right]^2$$

حيث :

ح = الانحراف ويحسب عن طريق وضع صفر أمام الفئة ذات اكبر تكرار ثم من أسفل (1 ، 2 ، 3 ،) ومن أعلى (-1 ، -2 ، -3 ،) .
ل = طول الفئة = الفرق بين بدايتي أى فئتين متتاليتين .

تحديد مدى دلالة النسبة الحرجة وصدق الاختبار من عدمه (3)

- إذا كانت النسبة الحرجة > 1.96 يكون الاختبار غير صادق عند مستوى دلالة 0.05 .
 - $1.96 >$ النسبة الحرجة > 2.58 يكون الاختبار صادق عن مستوى دلالة 1.96 .
 - إذا كانت النسبة الحرجة < 2.58 يكون الاختبار صادق عند مستوى دلالة 0.01 .
- بالطبع المقارنة بالقيمتين (1.96 ، 2.58) قيم ثابتة لا تتغير .

مثال :

الجدول التالي يوضح العلاقة بين فئات وتكرارات أصحاب مستوى الميزان القوى والضعيف لعدد من الطلاب فى اختبار للذكاء ، والمطلوب حساب قيمة معامل الصدق (النسبة الحرجة) وتحديد صدق الاختبار من عدمه عند مستوى دلالة 0.05 ؟

الفئات	16-14	19-17	22-20	25-23	28-26	31-29
تكرار الميزان الضعيف	4	3	8	0	0	0
تكرار الميزان القوى	0	0	0	5	7	9

الحل :

نكون الجدول التالى :

ف	ك ₁	ك ₂	س	ك ₁ ×س	ك ₂ ×س	ح	ح×ك ₁	ح×ك ₂	ح×ك ₁ ²	ح×ك ₂ ²
16-14	4	0	15	60	0	2-	8-	0	16	0
19-17	3	0	18	54	0	1-	3-	0	3	0
22-20	8	0	21	168	0	0	0	0	0	0
25-23	0	5	24	0	120	1	0	5	0	5
28-26	0	7	27	0	189	2	0	14	0	28
31-29	0	9	30	0	270	3	0	27	0	81
مجموع	15	21	-	282	579	-	11-	46	19	109

حساب المتوسط لأصحاب الميزان الضعيف :

مجـ (س × ك₁)

$$\frac{\text{مجـ ك}_1}{\text{مجـ ك}_1} = 1\text{م}$$

$$18.8 = \frac{282}{15} = 1\text{م}$$

حساب المتوسط لأصحاب الميزان القوى :

$$\frac{\text{مجـ (س} \times \text{ك}_2)}{\text{مجـ ك}_2} = 2\text{م}$$

$$27.5 = \frac{579}{21} = 1\text{م}$$

حساب طول الفئة :

ل = الفرقة بين بدايتي أى فئتين متتاليتين

$$ل = 14 - 17 = 3$$

حساب التباین لأصحاب الميزان الضعيف :

$$\left\{ \left[\frac{\text{مجـ (ح} \times \text{ك}_1)}{\text{مجـ (ح} \times \text{ك}_1)} \right] - \frac{\text{مجـ (ح} \times \text{ك}_2)}{-337} \right\} \times 2\text{ل} = 1^2\text{ع}$$

مجـ ك₁

مجـ ك₁

$$\left\{ \left[\frac{11^2}{15} \right] - \frac{19}{15} \right\} \times^2(3) = 1^2 \text{ع}$$

$$3.68 = 1^2 \text{ع}$$

حساب التباين لأصحاب الميزان القوى :

$$\left\{ {}^2 \left[\frac{\text{مجـ (ح} \times \text{ك}_2)}{\text{مجـ ك}_2} \right] - \frac{\text{مجـ (ح} \times \text{ك}_2)}{\text{مجـ ك}_2} \right\} \times {}^2 \text{ل} = 2^2 \text{ع}$$

$$\left\{ \left[\frac{46}{21} \right] - \frac{109}{21} \right\} \times^2(3) = 1^2 \text{ع}$$

$$33.29 = 1^2 \text{ع}$$

حساب قيمة ن₁ ، ن₂ :

$$ن_1 = \text{مج ك}_1 = 15$$

$$ن_2 = \text{مج ك}_2 = 21$$

حساب قيمة النسبة الحرجة :

$$\frac{1\mu - 2\mu}{\sqrt{\frac{\frac{2^2 \epsilon^2}{2\text{ن}} + \frac{1^2 \epsilon^2}{1\text{ن}}}} = \text{النسبة الحرجة}$$

$$\frac{18.8 - 27.5}{\sqrt{\frac{33.29}{21} + \frac{3.68}{15}}} = \text{النسبة الحرجة}$$

$$\text{النسبة الحرجة} = 6.4 .$$

تحديد صدق الاختبار :

قيمة النسبة الحرجة (6.4) < 1.96 عند مستوى دلالة 0.05 لذا
فان الاختبار صادق .

تمارين

1- قمت بتطبيق اختبار على مجموعة من الطلاب فى مادة
الإحصاء الاجتماعى مرتين مختلفتين ، وحصلت على
الدرجات التالية فى الاختبارين .

4	2	4	3	3	2	4	2	3	3	س
4	2	3	2	3	2	3	2	3	2	ص

والمطلوب :

حساب قيمة معامل الثبات بطريقة إعادة الاختبار .

2- فيما يلي درجات (5) طلاب في اختبارين س ، ص .

22	16	5	4	3	س
11	12	6	8	3	ص

والمطلوب :

حساب قيمة معامل الثبات بطريقة إعادة الاختبار .

3- فيما يلي درجات (5) طلاب فى اختبار تضمن 10 أسئلة :

الأسئلة										الأفراد
10	9	8	7	6	5	4	3	2	1	
3	5	4	4	5	5	2	3	1	4	1
1	3	1	1	2	4	1	5	3	2	2
1	5	2	4	1	3	1	2	1	5	3
1	2	4	1	1	1	2	1	1	2	4
1	1	1	4	3	1	1	2	2	2	5

والمطلوب :

حساب قيمة معامل الثبات بطريقة التجزئة النصفية .

4- من الجدول التالى احسب معامل الصدق بطريقة المقارنة

الطرفية وبين مدى صدق الاختبار من عدمه .

34-30	29-25	24-20	19-15	14-10	9-5	الفئات
0	0	4	6	3	2	تكرار الميزان الضعيف
9	6	7	3	0	0	تكرار الميزان القوى

5- من الجدول التالى احسب معامل الصدق بطريقة المقارنة

الطرفية وبين مدى صدق الاختبار من عدمه .

35-31	30-26	25-21	20-16	15-11	10-6	الفئات
0	0	0	5	6	3	تكرار الميزان الضعيف
6	7	6	5	0	0	تكرار الميزان القوى

الجدول الإحصائية

جدول كا²

درجة الحرية	مستوى الدلالة أو الثقة		
	0.05	0.01	0.001
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.13
9	16.92	21.67	27.88
10	18.31	23.21	29.59
11	19.68	24.73	31.26
12	21.03	26.22	32.91
13	22.36	27.69	34.53
14	23.69	29.14	36.12
15	25.00	30.58	37.70
16	26.30	32.00	39.25
17	27.59	33.41	40.79
18	28.87	34.81	42.31
19	30.14	36.19	43.82
20	31.41	37.57	45.32
21	32.67	38.93	46.80

22	33.92	40.29	48.27
23	35.17	41.64	49.73
24	36.42	42.98	51.18
25	37.65	44.31	52.62
26	38.89	45.64	54.05
27	40.11	46.96	55.48
28	41.34	48.28	56.89
29	42.56	49.59	58.30
30	43.77	50.89	59.70
31	44.99	52.19	61.10
32	46.19	53.49	62.49
33	47.40	54.78	63.87
34	48.60	56.06	65.25
35	49.80	57.34	66.62
36	51.00	58.62	67.99
37	52.19	59.89	69.35
38	53.38	61.16	70.71
39	54.57	62.43	72.06
40	55.76	63.69	73.41
41	56.94	64.95	74.75
42	58.12	66.21	76.09
43	59.30	67.46	77.42
44	60.48	68.71	78.75
45	61.66	69.96	80.08
46	62.83	71.20	81.40

47	64.00	72.44	82.72
48	65.17	73.68	84.03
49	66.34	74.92	85.35
50	67.51	76.15	86.66
51	68.67	77.39	87.97
52	69.83	78.62	89.27
53	70.99	79.84	90.57
54	72.15	81.07	91.88
55	73.31	82.29	93.17
56	74.47	83.52	94.47
57	75.62	84.73	95.75
58	76.78	85.95	97.03
59	77.93	87.17	98.34
60	79.08	88.38	99.62
61	80.23	89.59	100.88
62	81.38	90.80	102.15
63	82.53	92.01	103.46
64	83.68	93.22	104.72
65	84.82	94.42	105.97
66	85.97	95.63	107.26
67	87.11	96.83	108.54
68	88.25	98.03	109.79
69	89.39	99.23	111.06
70	90.53	100.42	112.31
71	91.67	101.62	113.56

72	92.81	102.82	114.84
73	93.95	104.01	116.08
74	95.08	105.20	117.35
75	96.22	106.39	118.60
76	97.35	107.58	119.85
77	98.49	108.77	121.11
78	99.62	109.96	122.36
79	100.75	111.15	123.60
80	101.88	112.33	124.84
81	103.01	113.51	126.09
82	104.14	114.70	127.33
83	105.27	115.88	128.57
84	106.40	117.06	129.80
85	107.52	118.24	131.04
86	108.65	119.41	132.28
87	109.77	120.59	133.51
88	110.90	121.77	134.74
89	112.02	122.94	135.96
90	113.15	124.12	137.19
91	114.27	125.29	138.45
92	115.39	126.46	139.66
93	116.51	127.63	140.90
94	117.63	128.80	142.12
95	118.75	129.97	143.32
96	119.87	131.14	144.55

97	120.99	132.31	145.78
98	122.11	133.47	146.99
99	123.23	134.64	148.21
100	124.34	135.81	149

جدول ت

درجة الحرية	مستوى الدلالة							
طرف واحد	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001	0.00005
طرفين	0.2	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001
2	1.89	2.92	4.30	9.92	14.09	31.60	44.70	100.14
3	1.64	2.35	3.18	5.84	7.45	12.92	16.33	28.01
4	1.53	2.13	2.78	4.60	5.60	8.61	10.31	15.53
5	1.48	2.02	2.57	4.03	4.77	6.87	7.98	11.18
6	1.44	1.94	2.45	3.71	4.32	5.96	6.79	9.08
7	1.41	1.89	2.36	3.50	4.03	5.41	6.08	7.89
8	1.40	1.86	2.31	3.36	3.83	5.04	5.62	7.12
9	1.38	1.83	2.26	3.25	3.69	4.78	5.29	6.59
10	1.37	1.81	2.23	3.17	3.58	4.59	5.05	6.21
11	1.36	1.80	2.20	3.11	3.50	4.44	4.86	5.92
12	1.36	1.78	2.18	3.05	3.43	4.32	4.72	5.70
13	1.35	1.77	2.16	3.01	3.37	4.22	4.60	5.51
14	1.35	1.76	2.14	2.98	3.33	4.14	4.50	5.36
15	1.34	1.75	2.13	2.95	3.29	4.07	4.42	5.24
16	1.34	1.75	2.12	2.92	3.25	4.01	4.35	5.13
17	1.33	1.74	2.11	2.90	3.22	3.97	4.29	5.04
18	1.33	1.73	2.10	2.88	3.20	3.92	4.23	4.97
19	1.33	1.73	2.09	2.86	3.17	3.88	4.19	4.90
20	1.33	1.72	2.09	2.85	3.15	3.85	4.15	4.84
21	1.32	1.72	2.08	2.83	3.14	3.82	4.11	4.78
22	1.32	1.72	2.07	2.82	3.12	3.79	4.08	4.74
23	1.32	1.71	2.07	2.81	3.10	3.77	4.05	4.69

24	1.32	1.71	2.06	2.80	3.09	3.75	4.02	4.65
25	1.32	1.71	2.06	2.79	3.08	3.73	4.00	4.62
26	1.31	1.71	2.06	2.78	3.07	3.71	3.97	4.59
27	1.31	1.70	2.05	2.77	3.06	3.69	3.95	4.56
28	1.31	1.70	2.05	2.76	3.05	3.67	3.93	4.53
29	1.31	1.70	2.05	2.76	3.04	3.66	3.92	4.51
30	1.31	1.70	2.04	2.75	3.03	3.65	3.90	4.48
35	1.31	1.69	2.03	2.72	3.00	3.59	3.84	4.39
40	1.30	1.68	2.02	2.70	2.97	3.55	3.79	4.32
45	1.30	1.68	2.01	2.69	2.95	3.52	3.75	4.27
50	1.30	1.68	2.01	2.68	2.94	3.50	3.72	4.23
55	1.30	1.67	2.00	2.67	2.92	3.48	3.70	4.20
60	1.30	1.67	2.00	2.66	2.91	3.46	3.68	4.17
65	1.29	1.67	2.00	2.65	2.91	3.45	3.66	4.15
70	1.29	1.67	1.99	2.65	2.90	3.43	3.65	4.13
75	1.29	1.67	1.99	2.64	2.89	3.42	3.64	4.11
80	1.29	1.66	1.99	2.64	2.89	3.42	3.63	4.10
85	1.29	1.66	1.99	2.63	2.88	3.41	3.62	4.08
90	1.29	1.66	1.99	2.63	2.88	3.40	3.61	4.07
95	1.29	1.66	1.99	2.63	2.87	3.40	3.60	4.06
100	1.29	1.66	1.98	2.63	2.87	3.39	3.60	4.05
200	1.29	1.65	1.97	2.60	2.84	3.34	3.54	3.97
500	1.28	1.65	1.96	2.59	2.82	3.31	3.50	3.92
1000	1.28	1.65	1.96	2.58	2.81	3.30	3.49	3.91
∞	1.28	1.64	1.96	2.58	2.81	3.29	3.48	3.89

جدول ف

درجة حرية التباين الصغير	درجة حرية التباين الكبير								
	1	2	3	4	5	6	8	12	∞
1	161	200	216	225	230	234	239	244	254
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.5
3	10.1	9.6	9.3	9.1	9.0	8.9	8.8	8.7	8.5
4	7.7	6.9	6.6	6.4	6.3	6.2	6.0	5.9	5.6
5	6.6	5.8	5.4	5.2	5.1	5.0	4.8	4.7	4.4
6	6.0	5.1	4.8	4.5	4.4	4.3	4.2	4.0	3.7
7	5.6	4.7	4.4	4.1	4.0	3.9	3.7	3.6	3.2
8	5.3	4.5	4.1	3.8	3.7	3.6	3.4	3.3	2.9
9	5.1	4.3	3.9	3.6	3.5	3.4	3.2	3.1	2.7
10	5.0	4.1	3.7	3.5	3.3	3.2	3.1	2.9	2.5
11	4.8	4.0	3.6	3.4	3.2	3.1	3.0	2.8	2.4
12	4.8	3.9	3.5	3.3	3.1	3.0	2.9	2.7	2.3
13	4.7	3.8	3.4	3.2	3.0	2.9	2.8	2.6	2.2
14	4.6	3.7	3.3	3.1	3.0	2.9	2.7	2.5	2.1
15	4.5	3.7	3.3	3.1	2.9	2.8	2.6	2.5	2.1
16	4.5	3.6	3.2	3.0	2.9	2.7	2.6	2.4	2.0
17	4.5	3.6	3.2	3.0	2.8	2.7	2.6	2.4	2.0
18	4.4	3.6	3.2	2.9	2.8	2.7	2.5	2.3	1.9
19	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.9
20	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.8
21	4.3	3.5	3.1	2.8	2.7	2.6	2.4	2.3	1.8
22	4.3	3.4	3.1	2.8	2.7	2.6	2.4	2.2	1.8

23	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.8
24	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.7
25	4.2	3.4	3.0	2.8	2.6	2.5	2.3	2.2	1.7
26	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.2	1.7
27	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.1	1.7
28	4.2	3.3	3.0	2.7	2.6	2.4	2.3	2.1	1.7
29	4.2	3.3	2.9	2.7	2.5	2.4	2.3	2.1	1.6
30	4.2	3.3	2.9	2.7	2.5	2.4	2.3	2.1	1.6
40	4.1	3.2	2.8	2.6	2.5	2.3	2.2	2.0	1.5
60	4.0	3.2	2.8	2.5	2.4	2.3	2.1	1.9	1.4
120	3.9	3.1	2.7	2.5	2.3	2.2	2.0	1.8	1.3
∞	3.8	3.0	2.6	2.4	2.2	2.1	1.9	1.8	1.0

أهم المراجع

- 1- اعتماد علام ، يسرى رسلان ، أساسيات الإحصاء الاجتماعي ، دار الثقافة للنشر والتوزيع .
- 2- أنور عطية العدل ، مبادئ الإحصاء الاجتماعي ، دار المعرفة الجامعية ، 1987 .
- 3- حسن محمد حسن ، أساسيات الإحصاء وتطبيقاته ، دار المعرفة الجامعية ، 1992 .
- 4- حسن محمد حسن ، مبادئ الإحصاء الاجتماعي ، دار المعرفة الجامعية ، 2000 .
- 5- خليفة عبد السميع خليفة ، الإحصاء التربوي ، مكتبة الأنجلو المصرية .
- 6- عبد الله عبد الحليم وآخرون ، الإحصاء مفاهيم أساسية ، 2003 .
- 7- غريب محمد سيد أحمد ، الإحصاء والقياس فى البحث الاجتماعي ، دار المعرفة الجامعية ، 1989 .
- 8- غريب محمد سيد أحمد ، ناجى بدر إبراهيم ، الإحصاء والقياس فى البحث الاجتماعي ، دار المعرفة الجامعية ، 1997 .
- 9- فاروق عبد العظيم وآخرون ، مبادئ الإحصاء ، دار المعرفة الجامعية .

- 10- فتحى عبد العزيز أبو راضى ، مبادئ الإحصاء الاجتماعى ،
دار المعرفة الجامعية .
- 11- محمد بهجت كشك ، مبادئ الإحصاء الاجتماعى ، دار
المعرفة الجامعية ، 1996 .
- 12- مصطفى زايد ، الإحصاء ووصف البيانات ، 1989 .

13- <http://www.mohp.gov.eg/Sec/Heducation/tadrib/5.doc>

14- http://www.arab-api.org/course13/c13_4.htm

15- http://dentarab.com/site/index.php?page=show_det&id=178

الفهرس

رقم الصفحة	الفصل
9	<u>الفصل الأول</u> علم الإحصاء تعريفه وأهميته
37	<u>الفصل الثاني</u> المفاهيم الإحصائية
73	<u>الفصل الثالث</u> العينات
123	<u>الفصل الرابع</u> تبويب وعرض البيانات
155	<u>الفصل الخامس</u> مقاييس النزعة المركزية
193	<u>الفصل السادس</u> مقاييس التشتت
213	<u>الفصل السابع</u> تحليل التباين
233	<u>الفصل الثامن</u> اختبار "ت"
271	<u>الفصل التاسع</u> اختبار "كا ² "
295	<u>الفصل العاشر</u>

	معاملات الارتباط - الاتحاد
327	<u>الفصل الحادي عشر</u> الثبات والصدق
345	الجدول الإحصائية
357	أهم المراجع