

الجامعة التقنية الشمالية

المعهد التقني نينوى

حقيبة تعليمية لمادة الاحصاء

محمد اسامه احمد الكاتب

المستوى الاول

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 وتفرغها 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 محددة وواضحة ومختصره ودقيقة .

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

- 1- التعريف 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 في بحثه في جداول تكرارية وأيضاً عرض هذه البيانات بالرسم البياني بطرقه المختلفة.
- 5- القدرة على وصف وتحليل البيانات من خلال مقاييس النزعة المركزية المختلفة مثل الوسط الحسابي والوسيط والمنوال 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 إحصائيا .
- 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 وأهميته

- أولاً : تعريف علم الإحصاء .
- ثانياً : أهمية علم الإحصاء .
- ثالثاً : تطور علم الإحصاء .
- رابعاً : علاقة علم الإحصاء بالعلوم الاجتماعية .

أولاً : 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 وعرضها بيانياً ووضعها فى جداول تلخيصية بطريقة تسهل تحليلها بهدف معرفة اتجاهات هذه الظواهر وعلاقات بعضها ببعض (2) .

ولقد كان الهدف الرئيسى من علم الإحصاء قديماً 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 والظواهر موضوع الدراسة .

ولكن مع تطور المجتمعات وتشابه جوانب الحياة الاقتصادية والاجتماعية الحديثة بها ، لم يعد مجرد توفير البيانات الكمية عن المتغيرات والظواهر موضوع الدراسة يفي بحاجات متخذى القرارات وصانعى السياسة العامة إلى تكوين صورة متكاملة الجوانب عن مجتمعهم والمجتمعات المحيطة به . فقام العلماء بتحديث نظريات علم الإحصاء وأساليبه وأدواته لى يعين الباحثين وغيرهم على استخلاص استنتاجات معينة من البيانات الكمية التى 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 -/العينه - وتعميم تلك الاستنتاجات على المجتمع الذى 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 النظرية اللازمة للقياس واتخاذ القرار فى كافة الميادين الاقتصادية والاجتماعية والسياسية والعسكرية)(3)

ثانياً: أهمية علم الإحصاء

لقد أصبح لعلم الإحصاء أهمية 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 أهمية الإحصاء بالنسبة لدارس علم الاجتماع أو علم النفس معتقداً أن الإحصاء موضوع يدخل فى صميم تخصص التجارئين 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 محدده 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) أهمية 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 التي تسفر عن تطبيق أداة إحصائية أو أكثر ليست نتائج قطعية أو غير قابلة للتحصيل والمراجعة . فإذا كانت الأدوات الإحصائية تستطيع أن تعين المرء على وصف البيانات وتصميم التجارب وعلى اختبار العلاقات بين الأشياء والوقائع التي يهتم بها إلا أن ذلك لا يلغى بصيرته السوسولوجية وخبرته المهنية .

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

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

ومما يعكس أهميته علم الإحصاء أنها تستخدم في توجيه عمله جمع البيانات وفي تفسير العلاقات التي تعكسها تلك البيانات . ومن ابرز المجالات التي تستخدم فيها المعالجات الإحصائية إجراء المقارنة بين عديد من الأشياء في كثير من المناسبات . ويمكننا القول أن الحياة الإنسانية سلسله من المواقف التي يتخذ فيها الفرد قراره بناء على ما تسفر عنه المقارنة التي يجريها بين عديد من الاحتمالات وهذه المقارنة في جوهرها عملية إحصائية تقترن بالقياس والتقييم والتقدير . فنجاح الإنسان في حياته يتحدد وفق مقياس معين في 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 نزل إلى السوق لشراء سلعة معينة ، فى موسم التنزيلات ، نهتم وبطريقه لا شعورية بحساب ثمن هذه السلعة بالنسبة إلى إجمالي النقود التي فى حوزتنا ونقدر ما إذا كان الباقي من هذه النقود وسوف يكفينا 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 التي 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 حربيه وضريبية ، وامتدت بعد ذلك لتشمل إحصاءات حجم السكان والماليد 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 عائلته برونلي Bernoulli وفرديريك جاوس F.gauss وكيثليه 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 (1517-1630) Keplr وجاليليو Galilio (1642-1564) 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 Pascal,B. (1623 1662) عالم الرياضيات والفيزياء والفيلسوف الفرنسى - 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 هيجيتير دافعا للكثيرين لدراسة النظريات والمشاكل المتعلقة بمباريات الصدفة ومنهم برنوللى (1654 - 1705) ودي موافر De Moivre (1667 - 1754) واربوثنوت 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 فى الوقت الحاضر ذات معان متعددة فمنها يفهم جمع المعلومات التى تبين الحالة فى 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 معانى :

(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 علي معلومات ذات قيمة من تلك البيانات الرقمية فإنها يجب أن تخضع للتحليل الاحصائي 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 في جمع ، وتصنيف وتحليل البيانات الرقمية.

ج- للإشارة إلى صفة أو خاصية للعينة تحت الدراسة.

والقاموس الحديث لعلم الاجتماع الذي 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 إلي 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).

رابعا : 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 إلى (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 إلى الانتحار بمثابة مصدر
لمعرفه 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 علي 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 في

التصويت الانتخابي وعن مستوياتهم التعليمية لا تنعقد المقارنة بينها إلا باستخدام المقاييس الإحصائية التي تكشف عن قوة الارتباط بين الميل للتصويت في 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 في القياس النفسى . ويعد علم النفس التجريبي وعلم النفس الاكلينيكي وعلم نفس الفروق الفردية من المجالات التي تعتمد اعتمادا جوهريا على المنهج الاحصائي في تناولها لموضوعات الدراسة .

ومن يقرأ مرجعا في القياس النفسى يجد أن علماء النفس يذهبون إلى أن كل شيء في مجال علمهم قابل للقياس تقريبا فوجد لديهم مقاييس للذكاء وللشخصية وللعواطف والميول وللاضطرابات النفسية والأمراض العقلية وكل مقياس من هذه المقاييس يخضع ، في واقع الأمر لأساليب إحصائية صارمة تحدد مدى ثباته وصدقه في قياس ما صمم لقياسه ويستخدم في المقارنة بين النتائج التي يتم التوصل إليها من دراسة عينه محددة من الأفراد وتلك التي يتم التوصل إليها من دراسة عينه أخرى (14)

وقد ظهر اهتمام كبير بتطبيق النظريات والطرق الإحصائية في العلوم الاجتماعية ، فقد أوضح كيتيليه (1796-1874) عالم الفلك الاجتماعي البلجيكي إمكان استخدام الاحتمالات والإحصاء لوصف وتفسير الظواهر الاجتماعية والاقتصادية وقدم مساهمات هامة في الطرق الإحصائية في تنظيم وإدارة الإحصاءات الرسمية - وقدم كذلك طريقه عامه للقياس في الانثروبولوجيا - وقد ساهم عالم النفس الانجليزي جالتون Galton (1822-1911) في تطبيق الطرق الإحصائية في علم النفس ، ووضع أساس علم القياس النفسى (psychometrics) وبدأ دراسة موضوع الارتباط والانحدار الذي اهتم به وطوره بعد ذلك عالم الإحصاء الانجليزي 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 Gosset (1876-1937) 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 فمذ مطلع القرن العشرين كان الاهتمام منصباً 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 Pearson,E.s وكذلك نيمان Neyman - ويعد الثلاثي فيشر - بيرسون - نيمان مؤسس منهج الاستقراء الإحصائي والذي يعرف حالياً بالاتجاه الكلاسيكي . وهو يعتمد علي المعلومات المتاحة من العينة فقط .

وقد ظهر في هذه الفترة اتجاه جديد يعرف بالاستقراء البيزياني Bayesianinference وذلك بجهود كل من جفريز jeffreys ورافري Ramsey وديفتني Definetti وجود Good وسافج Savage ولندلي lindley وآخرون O ويعتمد الاستقراء هنا على بيانات العينة بالإضافة إلى

المعلومات المسبقة **Prior. Information** وشهدت هذه الفترة أيضا عملا مكثفا كان فيها الاهتمام منصبا على صنع القرارات ، مما أدى إلى نشوء وظيفة حديثة للإحصاء تحت اسم نظرية القرارات الإحصائية **Decision theory** **Statistical** ويرجع ذلك إلى أعمال والد **Wald (1939)** ونيومان **Neuman** ومورجنسترن **Morgenstern** .

وقد صاحب هذا التطور الكبير في النظريات الإحصائية بداية ظهور مجموعة من التخصصات المختلفة تهتم بمجالات وأهداف خاصة - وقد بلغ هذا التطور قدرا هائلا يكاد يظهرها وكأنها علوما مستقلة . ومن هذه التخصصات بحوث العمليات **Operations Research** والإحصاء السكاني **Demography** ومراقبة الجودة **Quality control** والاقتصاد القياسي **Econometrics** ونظرا 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 الحيوى والاجتماع الرياضى والقياس الاجتماعى وعلم النفس الرياضى والقياس النفسى والقياس التربوى والاقتصاد الرياضى والتاريخ الاقتصادى الجديد أو القياس التاريخى (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 و قوانين .

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 الاجتماعي . وقد كان الاعتقاد السائد أن السلوك الاجتماعي والعلاقات الإنسانية التي تربط بين الأفراد في المجتمع إنما تخضع لنظم وقوالب 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 في العلوم الطبيعية التي تستخدم وحدات كمية لوصف ظواهرها والتعبير عنها بمعادلات رياضية والتنبؤ بها بتوافر شروط معينة .
- عدم استجابة البيئة الاجتماعية موضوع الدراسة للغايات التي يقصدها الباحث وعدم تمكن الباحث من السيطرة على كثير من العوامل التي تلعب دورا كبيرا في سير الحوادث وارتباط بعضها ببعض الآخر .

والمزايا التي يجنيها الباحث من الطرق الإحصائية يمكن تلخيصها فيما

يلي:-

- تساعد الباحث على إعطاء أوصاف على جانب كبير من الدقة العملية .
- فهدف العلم الوصول إلى أوصاف الظواهر و مميزاتا 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 التي حصل عليها الباحث. فمجرد الوصول إلى نتائج دون التحقق من ثباتها لا يكفي عادة كأساس يعتمد عليه في تفسير الحقائق وتحقيق الفروض.
- تساعد الإحصاء على تلخيص النتائج في شكل ملائم مفهوم فمجرد ذكر الدرجات لا يكفي للمقارنة بين الجنسين بل إن حساب متوسطى الدرجات قد سهل مهمة المقارنة كثيرا فالبيانات التي 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 الذي يصل إليه .

• تمكن الباحث من التنبؤ بالنتائج التي 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 فان 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 دارسة ، ذلك لان حقيقته قد تختلف كثيرا عما يبدوا عليه ، فهو يحتاج إلى ضبط في 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 في نواحي القياس والوصف بدرجة قد تكون كبيرة أو صغيرة حسب الطرق التي يستخدمها الباحث . وطرق الضبط الاحصائي خير وسيلة تعين الباحث على استبعاد هذه العوامل الشخصية .

إلا انه ينبغي أن يفهم من ذلك أن الإحصاء هو كل شيء في البحوث العلمية. فالإحصاء في يد من لا يجيد تطبيقها واستخدامها استخدام الخبير الفنى ، لا تفيد كثيرا . فهي مرحلة تالية لاكتشاف المشكلة وتحديدها ، وهي تتطلب عادة فروض علمية يتوقعها الباحث بناءً على دراساته السابقة وملاحظاته العديدة ، وهي تتطلب كذلك في 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 التي يستخدمها الباحث في كل خطوه من خطوات البحث ابتداء من المرحلة التمهيديّة 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 ممكنه . فالباحث عادة ما يجد نفسه أمام مجموعة كبيرة من البيانات الخام التي 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 مستخدما في ذلك الجانب الوصفي من الإحصاء حيث يمكنه أن يطبق هنا مجموعة من المقاييس الإحصائية التي لا تتعدى حد الوصف مثل مقاييس النزعة المركزية ومقاييس التشتت ومقاييس الارتباط والانحدار ... 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 في مجال العلوم الاجتماعية والتي يتضح منها بجلاء مدى ما تقدمه الإحصاء للباحث فهناك كلمة تحذير لابد أن

يعيها كل من يفكر في استخدام الأساليب الإحصائية ألا وهي أن التطبيق غير الصحيح للأسلوب الإحصائي ربما يؤدي إلى نتائج غير صحيحة ومضللة كما أن استخدام الأساليب الإحصائية يجب ألا يكون غاية في حد ذاته بل انه وسيلة الهدف منها هو تبصير الباحث بما هو بصدد القيام به وتبسيط وتوضيح خطوات البحث العلمي . (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 التي لم يعد أي فرض في غنى عنها ، ونحن نعيش عصر الثورة التكنولوجية وتهيمن على حياتنا لغة الأرقام. (18)

الفصل الرابع

تبويب وعرض البيانات

أولاً : العرض الجدولي 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 الاستمارات 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 طالباً 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 تصاعدياً 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 :

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

تبويب البيانات في جدول تكرارى 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 وفي هذه الحالة لا نعرف إلى أي فئة ينتمي 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 التالي .

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

ويعاب على هذه الطريقة أنها لا تصلح في حالة البيانات التي تحتوي 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 = $3.3 \times \text{لو (ن)} = 3.3 \times \text{لو (50)}$

$5.6 = 1.699 \times 3.3 =$

• نقرب عدد الفئات لأقرب رقم صحيح فتكون

عدد الفئات = 7

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

• نقرب طول الفئة لأقرب رقم صحيح فتصبح

طول الفئة = 10

• نختار بداية الفئة الأولى وهو أصغر رقم = 20

• نبدأ فى بناء الجدول كالتالى :

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

تبويب البيانات فى الجدول التكرارى المتجمع الصاعد :

ويقصد بالتكرار المتجمع الصاعد هو تجميع تكرار كل فئة على جميع التكرارات السابقة لها بحيث يكون مجموع التكرار التصاعدى للفئة الأخيرة مساوى 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 فى 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 التي تمثل التكرار.
- 6- إعادة كتابة الجدول بالأرقام .

مثال :

الجدول التالي يوضح البيانات التي حصل باحث في دراسة بين النوع و مشاهدة البرامج التعليمية لمجموعة من طلاب الصف الثالث الثانوي على النحو التالي :

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 ومشاهدة 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 :

النوع	ذكور	إناث
يشاهد		
لا يشاهد		

5- وضع العلامات .

النوع	ذكور	إناث
يشاهد	///	////
لا يشاهد	///	/ ///

6- إعادة كتابة الجدول بالأرقام .

النوع	ذكور	إناث	مج
يشاهد	5	4	9
لا يشاهد	5	6	11
مج	10	10	20

ثانياً : العرض البياني للبيانات الإحصائية

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

أولاً : العرض البياني للبيانات الغير مبوبة :

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

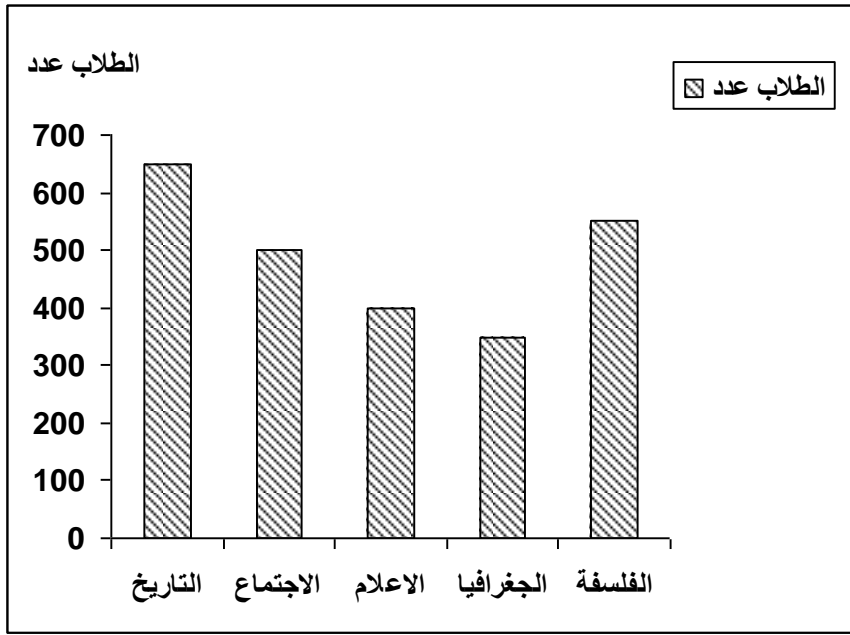
(1) طريقة الأعمدة البيانية البسيطة :

وفى 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 التالى 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



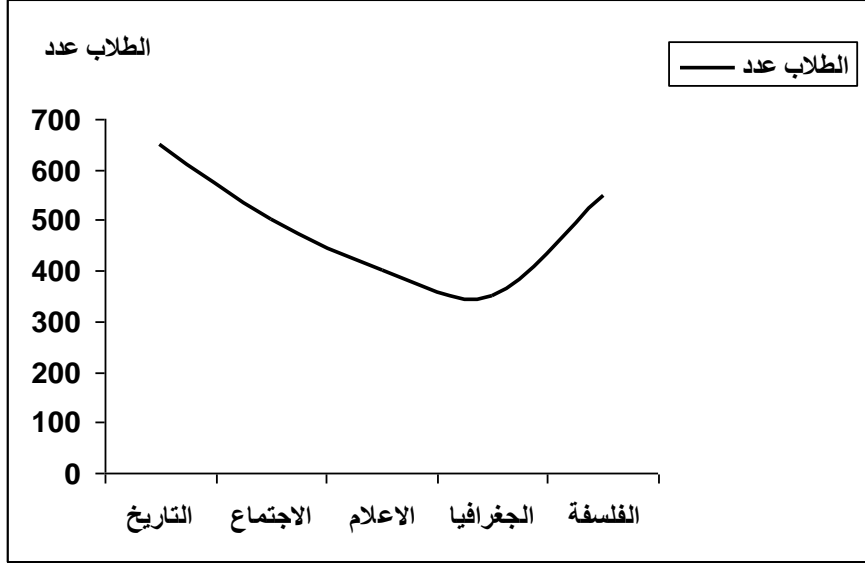
(2) طريقة المنحنى البيانى البسيط :

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 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



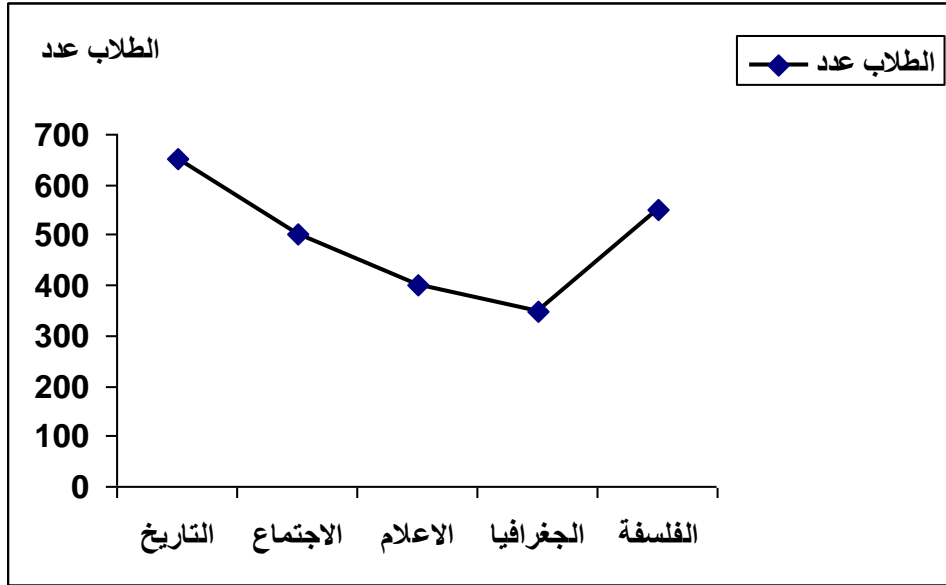
(3) طريقة الخط البياني المنكسر :

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 :

التكرار الفعلى للجزء

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

مجموع التكرارات

مثال :

الجدول التالى 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 :

$$\text{نحسب مجموع التكرارات} = 550+350+400+500+650$$

$$\text{مجموع التكرارات} = 2450$$

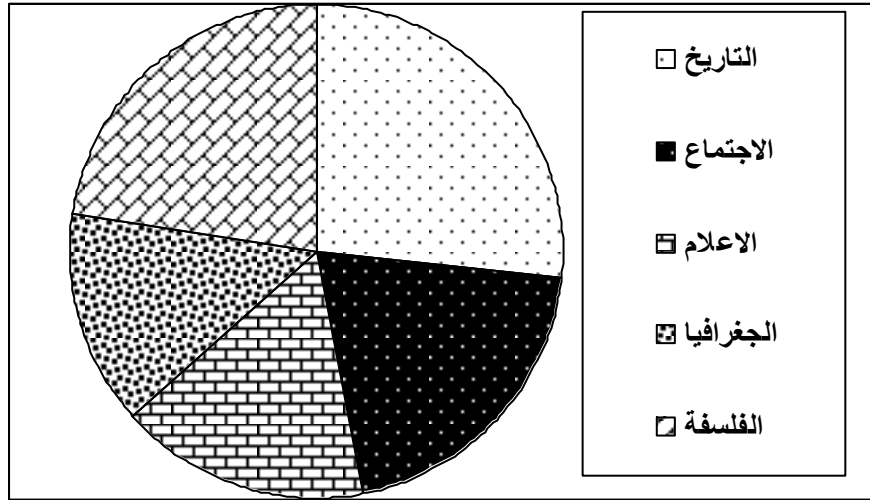
$$\text{زاوية قطاع التاريخ} = \frac{650}{2450} \times 360 = 95.5^\circ$$

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

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

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

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



(5) طريقة الأعمدة البيانية المتلاصقة :

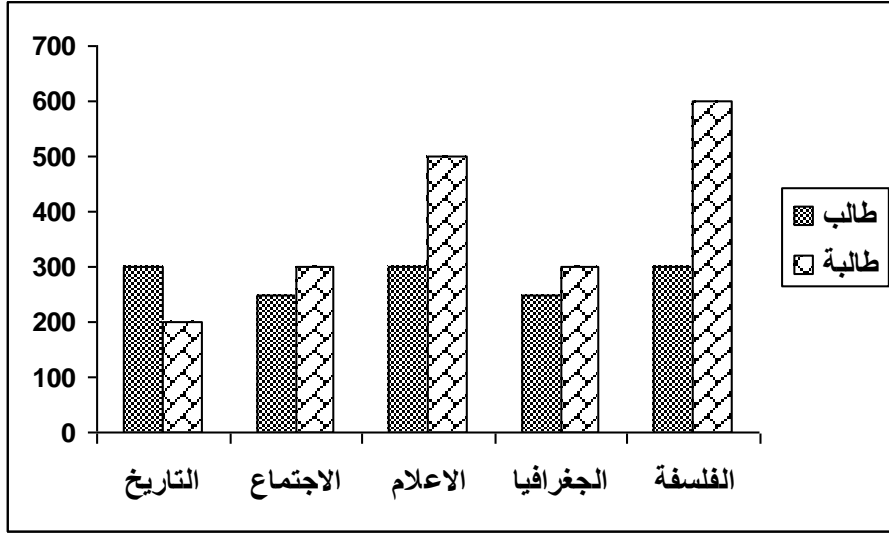
تسمى هذه الطريقة أيضا بطريقة الأعمدة البيانية المتجاورة وهي تشبه طريقة العمدة البيانية البسيطة ولكن يتم رسم عدد من الأعمدة متلاصقة يمثل كل منهم احد قيم المتغير .

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
طالبة	200	300	500	300	600

الحل :



(6) طريقة الأعمدة البيانية المجزأة :

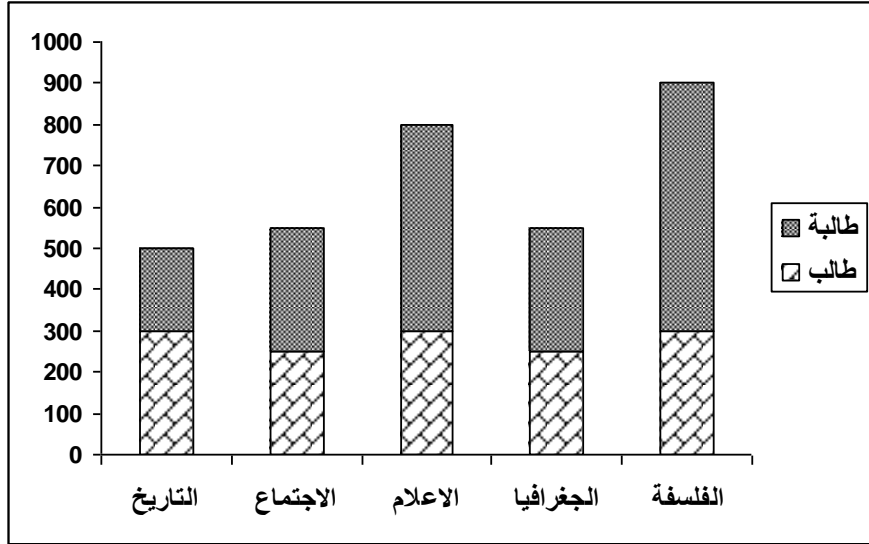
هذه الطريقة تشبه طريقة الأعمدة البيانية البسيطة ولكن يتم رسم عمود يمثل القيمة الأولى للمتغير ثم يليه أو يرتفعه عمود بباقي قيمة المتغير وتكون بادية العمود الثانى هى نهاية العمود الأول .

مثال :

الجدول التالي 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 :

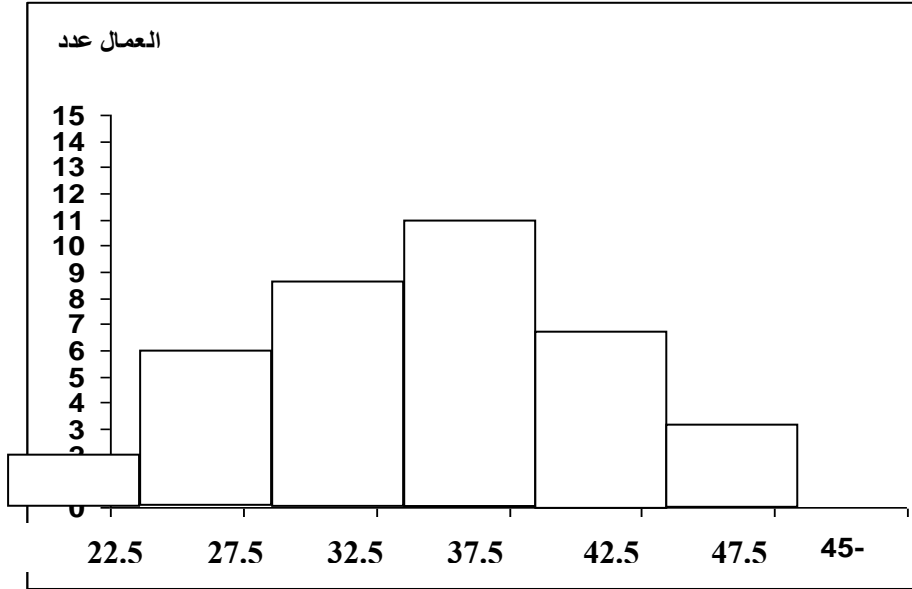
اعرض لهذا الجدول بيانياً باستخدام المدرج التكراري ؟

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 الأولى وفئة لاحقة للفئة الأخيرة وتكرار كل منهما صفر ، ثم نوصل كل نقطتين متتاليتين بخط مستقيم بالمسطرة .

ملحوظة :

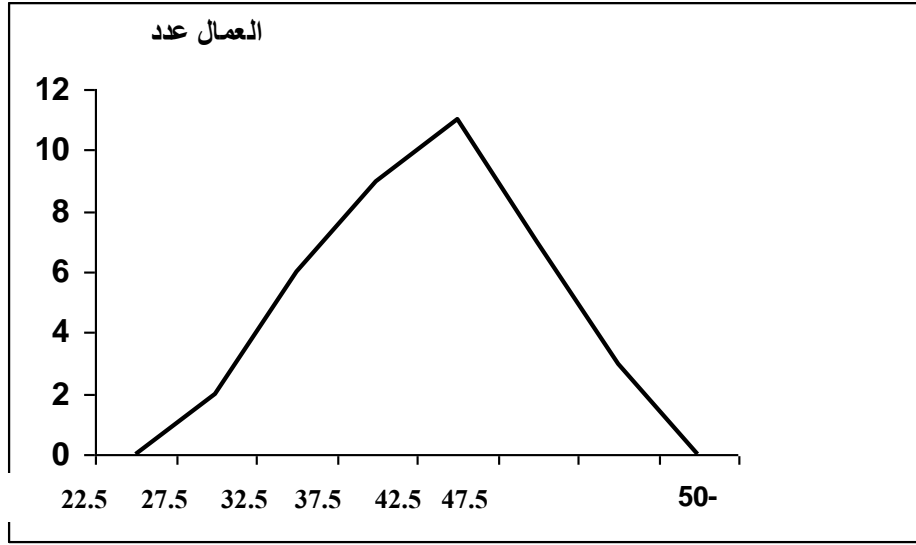
مساحة الشكل تحت المدرج التكرارى = مساحة الشكل تحت المضلع التكرارى .

مثال :

اعرض لهذا الجدول بيانياً باستخدام المضلع التكرارى ؟

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

الحل :



(3) المنحنى التكرارى :

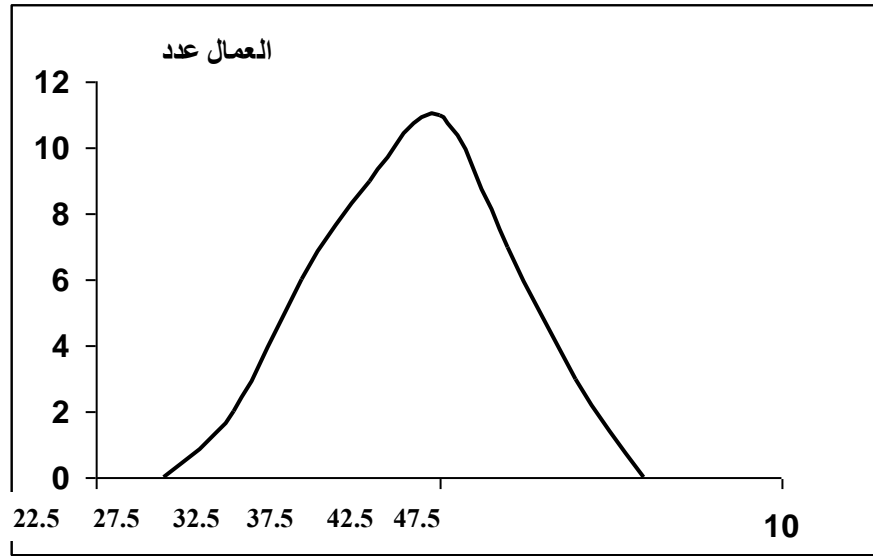
بعد رصد النقاط كما فى الطريقة السابقة نوصل كل نقطتين متتاليتين بمنحنى باليد

مثال :

اعرض لهذا الجدول بيانياً باستخدام المنحنى التكرارى ؟

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

الحل :



الفصل الخامس

مقاييس النزعة المركزية

- أولاً : الوسط الحسابى .
- 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 أيضاً 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:

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

حيث :-

$$\bar{X} = \text{الوسط الحسابي}$$

$$\sum = \text{مجموع}$$

$$X_i = \text{القيم}$$

$$n = \text{عدد القيم}$$

مثال :-

احسب الوسط الحسابي لدرجات 8 طلاب في مادة الإحصاء والتي كان بياناتهم كالتالي :

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

الحل :

$$\begin{aligned}\bar{X} &= \frac{\sum_{i=1}^8 X_i}{8} = \frac{X_1 + X_2 + X_3 + \dots + X_8}{8} \\ &= \frac{9+8+8+7+6+5+3+2}{8} = \frac{48}{8} = 6\end{aligned}$$

ثانياً : الوسيط

يعرف الوسيط على أنه القيمة التي تتوسط مجموعة من القيم إذا رتبنا 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: $(\frac{n+1}{2})$

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 = $4 = (\frac{n+1}{2})$ ، 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 :

$$\{(\frac{n}{2}), (\frac{n+1}{2})\}$$

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 =

$$\left\{ \left(\frac{8}{2} \right), \left(\frac{8+1}{2} \right) \right\} = \{4, 5\}$$

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

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

ثالثاً : المنوال :

المنوال هو القيمة الأكثر شيوعاً أو تكراراً .

حساب المنوال من البيانات الغير مبنوية

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 :-

المنوال = 8	7 - 8 - 9 - 8 - 10 - 8 - 12
المنوال = 10	10 - 12 - 15 - 10 - 12 - 10
المنوال = 15 ، 16	15 - 16 - 15 - 20 - 16 - 30
المنوال = لا يوجد	20 - 30 - 40 - 140 - 50 - 60

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 :

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

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

$$20 = 10 - 30 = \text{المنوال}$$

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 الغير مبوبة

المدى = أكبر قيمة - أصغر قيمة

مثال :

احسب المدى للبيانات التالية :

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

الحل :

نرتب القيم أولاً : (350-300-250-200-150-110-100-95-90-80)
المدى = 350 - 80 = 270

ثانياً : التباين والانحراف المعياري

يعتبر الانحراف المعياري من اهم المقاييس نظراً لدقته وان صيغته جاءت لايجاد وسيلة للتخلص من الاشارات السالبة ويعد من اكثر المقاييس انتشاراً من بقية المقاييس الاخرى يرمز للانحراف المعياري بالرمز S.

بينما يرمز للتباين بالرمز S^2 .

أي أنه إذا تم حساب أحدهما فيمكن حساب الآخر لأن الانحراف المعياري هو جذر التباين .

الانحراف المعياري للبيانات الغير مبوبة

باستخدام الطريقة المختصرة "طريقة الانحرافات" كالتالي

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

مثال :

احسب الانحراف المعياري للقيم التالية (9,10,12,13,15,19) .

الحل :

حساب الوسط الحسابي

$$\begin{aligned}\bar{X} &= \frac{\sum_{i=1}^6 X_i}{6} = \frac{X_1 + X_2 + X_3 + \dots + X_6}{6} \\ &= \frac{19+13+15+12+10+9}{6} = \frac{78}{6} = 13\end{aligned}$$

$$\begin{aligned}
S &= \sqrt{\frac{\sum_{i=1}^6 (X_i - \bar{X})^2}{6}} = \sqrt{\frac{(X_1 - 13)^2 + (X_2 - 13)^2 + \dots + (X_6 - 13)^2}{6}} \\
&= \sqrt{\frac{(19 - 13)^2 + (13 - 13)^2 + (15 - 13)^2 + (12 - 13)^2 + (10 - 13)^2 + (9 - 13)^2}{6}} \\
&= \sqrt{\frac{66}{6}} = \sqrt{11} = 3.3
\end{aligned}$$

ثالثاً : معامل الاختلاف المعياري للبيانات الغير مبوبة

تكون صيغة معامل الاختلاف المعياري كالتالي

الانحراف المعياري

$$\text{معامل الاختلاف المعياري} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} * 100$$

مثال :

إذا كان الوسط الحسابي لأوزان الطلبة 75 كغم والوسط الحسابي لأعمار نفس الطلبة 24 سنة والانحراف المعياري للأوزان 5 كغم والانحراف المعياري للأعمار 4 سنوات والمطلوب إيجاد معامل الاختلاف المعياري للأوزان والأعمار؟

الحل :

$$\text{معامل الاختلاف المعياري للأوزان} = 100 * \frac{5}{75} = 6.6$$

$$\text{معامل الاختلاف المعياري للأعمار} = 100 * \frac{4}{24} = 16.6$$

من خلال النتائج نلاحظ ان تشتت الطلاب من ناحية العمر اكبر من الوزن.

الدرجة المعيارية للبيانات الغير مبوبة

وهي العلاقة التي تعني ان الانحرافات عن الوسط الحسابي مقاسة بوحدة الانحراف وتكون صيغة الدرجة المعيارية كالتالي

$$\frac{\text{الدرجة} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}} = \text{الدرجة المعيارية}$$

مثال :

حصل طالب على 60 درجة في مادة الاحصاء و 70 درجة في مادة الرياضيات فأذا علمت ان الوسط الحسابي لدرجات الاحصاء 50 والرياضيات 65 وان الانحراف المعياري للاحصاء هو 2 وللرياضيات هو 2.5 في المادتين يكون مستواه اعلى؟

الحل :

$$\text{الدرجة المعيارية للاحصاء} = \frac{50-60}{2} = \frac{10}{2} = 5$$

$$\text{الدرجة المعيارية للرياضيات} = \frac{65-70}{2.5} = \frac{5}{2.5} = 2$$

من خلال النتائج نلاحظ ان مستوى الطالب في الاحصاء اعلى من الرياضيات.

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 -0.4 إلى أقل من -0.7
ارتباط عكسي ضعيف	من صفر إلى أقل من -0.4

طرق حساب الارتباط :

معامل الارتباط الخطى البسيط " لبيرسون " Pearson

هو مقياس يعين درجة الارتباط بين الظواهر وصيغته كالتالي

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{[\sum_{i=1}^n (x_i - \bar{x})^2] [\sum_{i=1}^n (y_i - \bar{y})^2]}}$$

مثال

فيما يلي المساحة المنزرعة بالأعلاف الخضراء بالآلاف هكتار، وإجمالي إنتاج اللحوم بالآلاف طن، خلال الفترة من 1995 حتى عام 2002.

السنة	1995	1996	1997	1998	1999	2000	2001	2002
المساحة	305	313	297	289	233	214	240	217
الكمية	592	603	662	607	635	699	719	747

والمطلوب: حساب معامل الارتباط بين المساحة والكمية، وما هو مدلوله ؟

الحل

بفرض أن (x) هي المساحة المنزرعة، (y) هي الكمية، ولحساب معامل الارتباط بين (y, x) يتم تطبيق المعادلة (6-2)، وذلك على النحو التالي:
حساب الوسط الحسابي لكل من المساحة، والكمية (\bar{y}, \bar{x}) .

$$\bar{x} = \frac{\sum x}{n} = \frac{2108}{8} = 263.5, \quad \bar{y} = \frac{\sum y}{n} = \frac{5264}{8} = 658$$

حساب المجاميع

x	y	$x - \bar{x}$	$(x - \bar{x})^2$	$y - \bar{y}$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
305	592	41.5	1722.25	-66	4356	-2739
313	603	49.5	2450.25	-55	3025	-2722.5
297	662	33.5	1122.25	4	16	134
289	607	25.5	650.25	-51	2601	-1300.5
233	635	-30.5	930.25	-23	529	701.5
214	699	-49.5	2450.25	41	1681	-2029.5
240	719	-23.5	552.25	61	3721	-1433.5
217	747	-46.5	2162.25	89	7921	-4138.5
2108	5264	0	12040	0	23850	-13528

$$\sum (x - \bar{x})^2 = 12040, \quad \sum (y - \bar{y})^2 = 23850,$$

$$\sum (x - \bar{x})(y - \bar{y}) = -13528$$

إذا معامل الارتباط قيمته هي:

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2} \sqrt{\sum(y - \bar{y})^2}} = \frac{-13528}{\sqrt{12040} \sqrt{23850}}$$

$$= \frac{-13528}{(109.727)(154.434)} = \frac{-13528}{16945.619} = -0.798$$

يوجد ارتباط عكسي قوي بين المساحة المنزرعة، وكمية إنتاج الحوم.

معامل ارتباط الرتب (اسبيرمان) Spearman

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبيين، ومثال على ذلك قياس العلاقة بين تقديرات الطلبة في مادتين، أو العلاقة بين درجة تفضيل المستهلك لسلعة معينة، ومستوى الدخل، فإنه يمكن استخدام طريقة "بيرسون" السابقة في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الأصلية، ويطلق على هذا المعامل "معامل ارتباط اسبيرمان" Spearman، ويعبر عنه بالمعادلة التالية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)} \quad (4-6)$$

حيث أن d هي الفرق بين رتب مستويات المتغير الأول X ، ورتب مستويات المتغير الثاني Y ، أي أن: $d = R_x - R_y$.

مثال (2-6)

فيما يلي تقديرات 10 طلاب في مادتي الإحصاء، والاقتصاد:

تقديرات إحصاء	أ	ج+	د	د+	ب+	ج+	أ+	ب	ب+	ب+
تقديرات اقتصاد	أ+	د	ج	ج	أ	ب	ب+	ب	ج	ب

والمطلوب:

1- احسب معامل الارتباط بين تقديرات الطلبة في المقررين.

2- وما هو مدلوله؟

الحل

1- بفرض أن X هي تقديرات الإحصاء، Y هي تقديرات الاقتصاد، يمكن حساب معامل الارتباط بينهما

باستخدام المعادلة (4-6)، وذلك بإتباع الآتي:

الرتب	1	2	3	4	5	6	7	8	9	10
تقديرات إحصاء	أ+	أ	ب+	ب+	ب+	ب	ج+	ج+	د	د
رتب X	1	2	$(3+4+5)/3=4$			6	$(7+8)/2=7.5$		9	10
تقديرات اقتصاد	أ+	أ	ب+	ب	ب	ب	ج-	ج-	ج-	د
رتب Y	1	2	3	$(4+5+6)/3=5$			$(7+8+9)/3=8$			10

• إذا يمكن حساب المجموع: $\sum d^2$ كما يلي:

x	y	رتب x	رتب y	d	d^2
أ	أ	2	1	1	1
ج	د	7.5	10	-2.5	6.25
د	ج	10	8	2	4
د	ج	9	8	1	1
ب	أ	4	2	2	1
ج	ب	7.5	5	2.5	6.25
أ	ب	1	3	-2	4
ب	ب	6	5	1	1
ب	ج	4	8	-4	16
ب	ب	4	5	-1	1
					44.5

$$\sum d^2 = 44.5$$

• معامل الارتباط هو:

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

$$= 1 - \frac{6(44.5)}{10(10^2 - 1)} = 1 - \frac{267}{990}$$

$$= 1 - 0.2697 = 0.7303$$

2- مدلول معامل الارتباط :

بما أن $r = 0.703$ ، ويدل ذلك على وجود ارتباط طردي قوي بين تقديرات الطالب في مادة

الإحصاء ، ومادة الاقتصاد .

ملحوظة: - يمكن استخدام صيغة معامل ارتباط "اسبيرمان" في حساب الارتباط بين متغيرين كميين، حيث يتم استخدام رتب القيم التي يأخذها المتغير، ونترك للطالب القيام بحساب معامل ارتباط الرتب بين المساحة والكمية في مثال (5-1) السابق، وعليه أن يقوم بتفسير النتيجة: (معاونة : $\sum d^2 = 148$)

معامل الاقتران : Coefficient of Association

يستخدم معامل الاقتران لقياس العلاقة بين ظاهرتين تنقسم كل منهما إلى قسمين (أو صفتين) فقط، وتكون البيانات موضوعة في جدول مزدوج يتكون من أربع خلايا 2×2 فقط. ويسمى الجدول في هذه الحالة "جدول الاقتران (2×2) " حيث توضع إحدى الظاهرتين أفقياً والأخرى رأسياً ويكون الشكل العام لجدول الاقتران 2×2 كما يلي :

جدول الاقتران 2×2

الخاصة (2)	الخاصة (1)	الظاهرة الثانية الظاهرة الأولى
B	A	الخاصية (1)
D	C	الخاصية (2)

نلاحظ أن الجدول أعلاه يتكون من أربع خلايا، حيث تمثل الحروف، A,B,C,D تكرارات هذه الخلايا. فالحرف A يمثل تكرار أو عدد المفردات التي تجمع بين الخاصية رقم (1) لكل من الظاهرتين، وهكذا مع باقي التكرارات. فمثلاً إذا كانت الظاهرتان هما : الانتماء السياسي، وما إذا كان الشخص مؤهلاً أم لا فالجدول التالي يمثل تصنيف 100 شخص حسب هاتين الظاهرتين.

التأهيل	مؤهل	غير مؤهل	الانتماء السياسي
له انتماء سياسي	50	12	
ليس له انتماء سياسي	20	18	

في هذا الجدول نجد أن 50 شخصاً مؤهلون ولهم انتماء سياسي، 12 شخصاً غير مؤهلين ولهم انتماء سياسي 20 شخصاً مؤهلون وليس لهم انتماء سياسي، وأخيراً 18 شخصاً غير مؤهلين وليس لهم انتماء سياسي. ويأخذ قانون معامل الاقتران الشكل التالي:

$$Ass.Coeff = \frac{AD - BC}{AD + BC} = \text{معامل الاقتران}$$

ومن أهم الملاحظات على معامل الاقتران ما يلي :

- 1- أن قيمة معامل الاقتران تنحصر ما بين +1، -1
- 2- إذا كان التكرار A أو التكرار D مساوياً للصفر فإن قيمة معامل الاقتران تكون مساوية لسالب واحد (-1) وبالتالي فالارتباط عكسي تام بمعنى أن العلاقة بين المتغيرين علاقة عكسية تامة.
- 3- أما إذا كان التكرار B والتكرار C مساوياً للصفر فإن قيمة معامل الاقتران تكون مساوية لموجب واحد (+1) أو بالتالي فإن الارتباط طردي تام.
- 3- إذا كان حاصل الضرب AD يساوي حاصل الضرب BC فإن قيمة معامل الاقتران تكون مساوية للصفر، مما يعني عدم وجود علاقة (أو اقتران) بين الظاهرتين
- 4- كلما اقتربت قيمة معامل الاقتران من الواحد الصحيح (سواء بالموجب أو السالب) كلما دل ذلك على وجود علاقة قوية بين الظاهرتين، وبالعكس كلما اقتربت قيمته من الصفر كلما دل ذلك على ضعف العلاقة بين الظاهرتين.

مثال (1): احسب معامل الاقتران بين الانتماء السياسي والمؤهل حسب الجدول أعلاه.

الحل :

جدول الاقتران بين الانتماء السياسي والمؤهل – كما سبق وذكرنا – يأخذ الشكل التالي

التأهيل	مؤهل	غير مؤهل
له انتماء	A ₅₀	B ₁₂
ليس له انتماء	C ₂₀	D ₁₈

وبالتعويض في قانون معامل الاقتران نحصل على :

$$\begin{aligned}
 Ass.Coeff &= \frac{AD - BC}{AD + BC} \\
 &= \frac{50 \times 18 - 12 \times 20}{50 \times 18 + 12 \times 20} \\
 &= \frac{660}{1140} \\
 &= 0.58
 \end{aligned}$$

وهذا يعني أن هناك علاقة اقتران طردية فوق المتوسط بين الانتماء السياسي للشخص والمؤهل. فكلما زادت درجة تأهيل الشخص العلمي كلما زاد انتماءه السياسي.

ولعل السؤال المهم هنا هو هل هذا الارتباط أو الاقتران حقيقي بين المتغيرين أم أن المتغيرين في واقع الأمر مستقلان ؟ للإجابة على هذا التساؤل لابد للباحث من استخدام ما يسمى " باختبار الاستقلال " .

معامل التوافق : Contingency Coefficient

إذا كان الجدول يتكون من أكثر من أربع خلايا بمعنى أن إحدى الظاهرتين أو كليهما تنقسم إلى أكثر من صفتين (أو قسمين) فإن المعامل الذي يحسب في هذه الحالة لقياس العلاقة بين الظاهرتين يسمى معامل التوافق. فمثلاً إذا كان المطلوب دراسة العلاقة بين المستوى التعليمي والذي ينقسم عادة إلى أكثر من قسمين (أمي - ابتدائي - متوسط..الخ) ودرجة الوعي السياسي لدى الفرد (عالية - متوسطة - متدنية) فإن المعامل الإحصائي الذي يقيس العلاقة بينهما هو ما يسمى بمعامل التوافق وهي الحالة العامة لدراسة العلاقة بين ظاهرتين مقارنة بالحالة الخاصة السابقة التي تحسب فيها معامل الاقتران الذي يقتصر على دراسة المتغيرات التي تتكون من صفتين - كما سبق وأشرنا آنفاً.

خطوات حساب معامل التوافق :

- 1- تربيعة التكرارات الموجودة بكل خلية من خلايا الجدول
- 2- قسمة مربع التكرار بكل خلية على حاصل ضرب مجموع الصف في مجموع العمود الذي به الخلية.
- 3- نكرر الخطوة (2) لكل الخلايا ونرمز للمجموع بالرمز C.
- 4- حساب معامل التوافق والذي يأخذ الشكل التالي :

$$Cont.Coeff. = \sqrt{\frac{C-1}{C}} = \text{معامل التوافق}$$

أي أن معامل التوافق هو الجذر التربيعي لخارج قسمة C-1 على C حيث C هي مجموع خوارج قسمة مربعات التكرارات بكل خلية على مجموع الصف في مجموع العمود الذي به كل خلية.

ملاحظات مهمة على معامل التوافق :

- 1- معامل التوافق لا يمكن أن يكون بالسالب، فهو إذن لا يحدد نوع أو اتجاه العلاقة بين الظاهرتين.
- 2- الحد الأدنى لمعامل التوافق يساوي الصفر.
- 3- الحد الأعلى لمعامل التوافق يحسب من العلاقة التالية :

$$\frac{\sqrt{4 \times \frac{(\text{عدد الصفوف} - 1) \times (\text{عدد الأعمدة} - 1)}{\text{عدد الصفوف} \times \text{عدد الأعمدة}}}}{2} = \text{الحد الأعلى لمعامل التوافق}$$

4- كلما اقتربت قيمة معامل التوافق من الحد الأعلى كلما كانت العلاقة قوية بين الظاهرتين، وكلما كانت قيمته قريبة من الصفر كلما كانت العلاقة ضعيفة بين الظاهرتين.

مثال (3):

الجدول التالي يمثل توزيع مائة ناخب أمريكي حسب مستوى التعليم والانتماء الحزبي والمطلوب قياس العلاقة بين هذين المتغيرين :

المجموع	ديموقراطي	جمهوري	الانتماء الحزبي المستوى التعليمي
18	8	10	شهادة متوسطة
34	14	20	شهادة ابتدائية
48	18	30	أمي
100	40	60	المجموع

الحل :

1- يمكن اختصار خطوات الحل بحساب C في خطوة واحدة كما يلي :

$$\begin{aligned}C &= \frac{10 \times 10}{60 \times 18} + \frac{8 \times 8}{40 \times 18} \\ &+ \frac{20 \times 20}{60 \times 34} + \frac{14 \times 14}{40 \times 34} \\ &+ \frac{30 \times 30}{60 \times 48} + \frac{18 \times 18}{40 \times 48} \\ &= 0.093 + 0.088 + 0.196 + 0.144 + 0.1313 + 0.169 \\ C &= 1.003\end{aligned}$$

2- معامل التوافق يساوي

$$\begin{aligned}Cont.Coeff. &= \sqrt{\frac{C-1}{C}} \\ &= \sqrt{\frac{1.003-1}{1.003}} \\ &= \sqrt{\frac{0.003}{1.003}} \\ &= 0.05\end{aligned}$$

3- الحد الأعلى لمعامل التوافق يساوي

$$\begin{aligned}&= \sqrt[4]{\frac{(3-1)(2-1)}{3 \times 2}} \\ &= \sqrt[4]{\frac{2 \times 1}{3 \times 2}} = \sqrt{\frac{2}{6}} \\ &= \sqrt[4]{\frac{1}{3}} \\ &= 0.76\end{aligned}$$

وواضح أن العلاقة ضعيفة بين المستوى التعليمي والانتماء الحزبي للعينة محل الدراسة حيث أنها تساوي 0.05 فقط مقارنة بالحد بالأعلى وهو 0.76

Simple Regression الانحدار الخطي البسيط

إن الغرض من استخدام أسلوب تحليل الانحدار الخطي البسيط، هو دراسة وتحليل أثر متغير كمي على متغير كمي آخر، ومن الأمثلة على ذلك ما يلي:

- دراسة أثر كمية السماد على إنتاجية الدونم.
- دراسة أثر الإنتاج على التكلفة.
- دراسة أثر كمية البروتين التي يتناولها الأبقار على الزيادة في الوزن.
- أثر الدخل على الإنفاق الاستهلاكي.

وهكذا هناك أمثلة في كثير من النواحي الاقتصادية، والزراعية، والتجارية، والعلوم السلوكية، وغيرها من المجالات الأخرى.

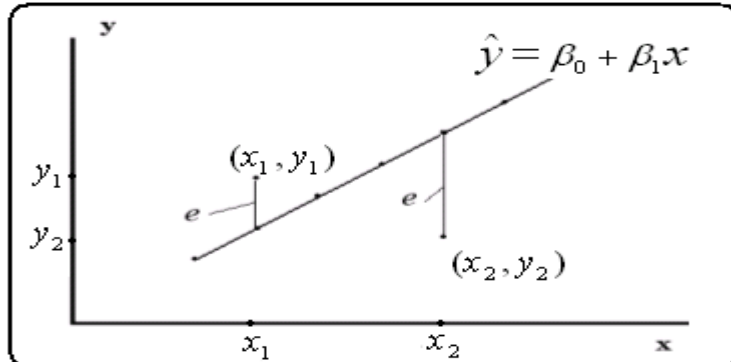
1/3/6 نموذج الانحدار الخطي

في تحليل الانحدار البسيط، نجد أن الباحث يهتم بدراسة أثر أحد المتغيرين ويسمى بالمتغير المستقل أو المتنبأ منه، على المتغير الثاني ويسمى بالمتغير التابع أو المتنبأ به، ومن ثم يمكن عرض نموذج الانحدار الخطي في شكل معادلة خطية من الدرجة الأولى، تعكس المتغير التابع كدالة في المتغير المستقل كما يلي:

$$y = \beta_0 + \beta_1 x + e \quad (5-6)$$

حيث أن:

- y : هو المتغير التابع (الذي يتأثر)
- x : هو المتغير المستقل (الذي يؤثر)
- β_0 : هو الجزء المقطوع من المحور الرأسي y ، وهو يعكس قيمة المتغير التابع في حالة انعدام قيمة المتغير المستقل $x = 0$
- β_1 : ميل الخط المستقيم $(\beta_0 + \beta_1 x)$ ، ويعكس مقدار التغير في y إذا تغيرت x بوحدة واحدة.
- e : هو الخطأ العشوائي، والذي يعبر عن الفرق بين القيمة الفعلية y ، والقيمة المقدرة $\hat{y} = \beta_0 + \beta_1 x$ ، أي أن: $e = y - (\beta_0 + \beta_1 x)$ ، ويمكن توضيح هذا الخطأ على الشكل التالي لنقط الانتشار.



2/3/6 تقدير نموذج الانحدار الخطي البسيط

يمكن تقدير معاملات الانحدار (β_0, β_1) في النموذج (5-6) باستخدام طريقة المربعات الصغرى، وهذا التقدير هو الذي يجعل مجموع مربعات الأخطاء العشوائية $\sum e^2 = \sum (y - (\beta_0 + \beta_1 x))^2$ أقل ما يمكن، ويحسب هذا التقدير بالمعادلة التالية:

$$\hat{\beta}_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}, \quad (6-1)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

حيث أن \bar{x} هو الوسط الحسابي لقيم x ، \bar{y} هو الوسط الحسابي لقيم y ، وتكون القيمة المقدره للمتغير التابع هو: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ ، ويطلق على هذا التقدير "تقدير معادلة انحدار y على x ".

مثال (3-6)

فيما يلي بيانات عن كمية البروتين اليومي بالجرام التي يحتاجها العجل الرضيع، ومقدار الزيادة في وزن العجل بالكيلوجرام، وذلك لعينة من العجول الرضيعة حجمها 10.

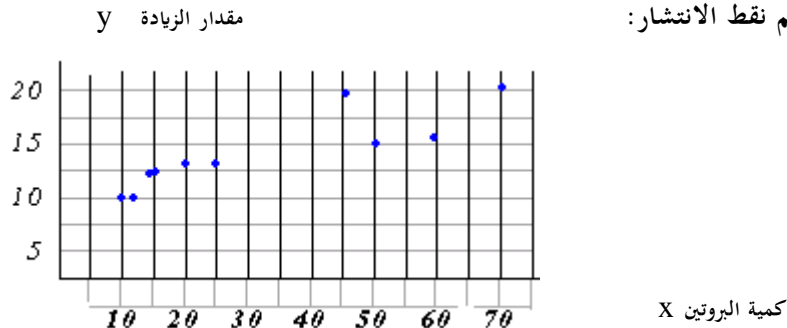
كمية البروتين	10	11	14	15	20	25	46	50	59	70
الزيادة في الوزن	10	10	12	12	13	13	19	15	16	20

والمطلوب :

- 1- ارسم نقط الانتشار، وما هو توقعاتك لشكل العلاقة ؟
- 2- قدر معادلة انحدار الوزن على كمية البروتين.
- 3- فسر معادلة الانحدار.
- 4- ما هو مقدار الزيادة في الوزن عند إعطاء العجل 50 جرام من البروتين ؟ وما هو مقدار الخطأ العشوائي؟
- 5- ارسم معادلة الانحدار على نقط الانتشار في المطلوب (1).

الحل

1- رسم نقط الانتشار:



من المتوقع أن يكون لكمية البروتين أثر طردي (إيجابي) على مقدار الزيادة في الوزن.

2- تقدير معادلة الانحدار.

بفرض أن x هي كمية البروتين، y هي مقدار الزيادة في الوزن، يمكن تطبيق المعادلتين في (6-6)، ومن ثم يتم حساب المجاميع التالية:

كمية البروتين x	الزيادة في الوزن y	$x y$	x^2	المجموع المطلوبة
10	10	100	100	$\sum x = 320$
11	10	110	121	$\sum y = 140$
14	12	168	196	$\sum xy = 5111$
15	12	180	225	$\sum x^2 = 14664$
20	13	260	400	
25	13	325	625	
46	19	874	2116	
50	15	750	2500	إذا الوسط الحسابي:
59	16	944	3481	$\bar{x} = \frac{\sum x}{n} = \frac{320}{10} = 32$
70	20	1400	4900	$\bar{y} = \frac{\sum y}{n} = \frac{140}{10} = 14$
320	140	5111	14664	

• بتطبيق المعادلة الأولى في (6-6) يمكن حساب $\hat{\beta}_1$ كما يلي:

$$\hat{\beta}_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{(10)(5111) - (320)(140)}{(10)(14664) - (320)^2}$$

$$= \frac{6310}{44240} = 0.1426$$

• بتطبيق المعادلة الثانية في (6-6) يمكن حساب $\hat{\beta}_0$ كما يلي:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 14 - (0.1426)(32) = 9.4368$$

• إذا معادلة الانحدار المقدرة، هي:

$$\hat{y} = 9.44 + 0.143x$$

3- تفسير المعادلة:

- الثابت $\hat{\beta}_0 = 9.44$: يدل على أنه في حالة عدم استخدام البروتين قى التغذية، فإن الوزن يزيد 9.44 كجم.
- معامل الانحدار $\hat{\beta}_1 = 0.143$: يدل على أنه كلما زادت كمية البروتين جرام واحد، حدث زيادة في وزن العجل بمقدار 0.143 كجم، أى زيادة مقدارها 143 جرام.

4- مقدار الزيادة في الوزن عند $x = 50$ هو:

$$\hat{y} = 9.44 + 0.143(50) = 16.59$$

وأما ومقدار الخطأ العشوائي هو:

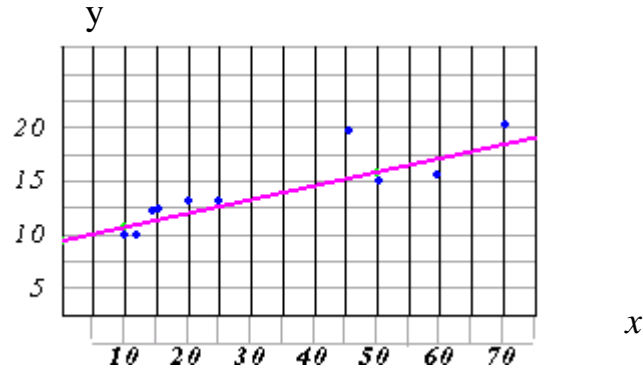
$$\hat{e}_{x=50} = y_{x=50} - \hat{y}_{x=50} = 15 - 16.59 = -1.59$$

5- رسم معادلة الانحدار على نقط الانتشار.

يمكن رسم معادلة خط مستقيم إذا علم نقطتين على الخط المستقيم.

x	50	10
\hat{y}	16.59	10.87

إذا معادلة الانحدار هي:



السلاسل الزمنية

هي مجموعة من القيم المشاهدة في ظاهرة معينة ولمدة من الزمن في فترات زمنية متشابهة فمثلاً مجموعة القيم الدالة على صادرات العراق السنوية من عام 1980 ولمدة خمس سنوات. وتهتم الكثير من الدراسات الاقتصادية والاجتماعية بدراسة السلسلة الزمنية حيث يمكن معرفة طبيعة التغيرات التي تطرأ على قيم الظاهرة مع الزمن وتحديد الاسباب والنتائج وتفسير العلاقات بينها والتنبؤ بما سيحدث من تغيير على قيم الظاهرة في المستقبل على ضوء الماضي، ان دراسة السلسلة الزمنية تقتضي تحليل العوامل المؤثرة فيها وقد وجد ان السلاسل الزمنية تتأثر بالعوامل التالية:

1-الاتجاه العام

2- التغيرات الموسمية

3- التغيرات الدورية

4- التغيرات العرضية

الاتجاه العام

هو العامل الاكثر تأثيراً على قيم الظاهرة في المدة الطويلة فمثلاً عدد السكان خلال 40 عام الماضية تمثل سلسلة زمنية منتشرة بدرجة كبيرة بالاتجاه العام نظراً لميل جميع الاعداد المكون لهذه السلسلة الزمنية من التزايد بصورة عامة ويكون الاتجاه العام موجباً وقد يكون الاتجاه العام سالباً عند تناقص قيم الظاهرة بمرور الزمن مثل الوفيات الناتجة عن الاصابة باحد الامراض ويكون الاتجاه العام مستقيماً اذا كان التغيير في قيمة الظاهرة ثابتاً وقد يكون غير مستقيم اذا كان التغيير غير ثابت في قيمة الظاهرة، وهناك عدة طرق لتعيين خط الاتجاه العام ومن هذه الطرق استخدام معادلة خط الانحدار

التغيرات الموسمية

هي التغيرات الدورية التي يتكرر حدوثها بانتظام وقد تكون يومية او شهرية او فصلية وتتكون نتيجة تأثيرات المناخ او عادات اجتماعية او قياسات دينية مثل الزيادة في المبيعات لدى المحلات في الايام التي تسبق الاعياد والتغيرات الموسمية.