عرض البيانات و تحليلها

د.فضيل دليو

جامعة قسنطينة 3
كلية علوم الإعلام والاتصال

جامعة قسنطينة 3

كلية علوم الإعلام والاتصال والسمعي البصري

عرض البيانات وتحليلها

أ. د. فضيل دليو

مطبوعة لطلبة *السنة الثانية: إعلام واتصال*

1440–1441 هـ – 2018–2019 م

السداسي: الثالث

عنوان الوحدة : وحدة تعليم منهجية

المادة: عرض البيانات وتحليلها

أهداف التعليم

التعرف على كيفيات عرض وتحليل البيانات.

المعارف المسبقة المطلوبة

التعرف على أهم مناهج البحث العلمي عموما وفي الإعلام والاتصال خصوصا، وكذا على أهم خطوات البناء النظري للبحث العلمي.

محتوى المادة:

مقدمة: أنواع البيانات

1 -طرق عرض البيانات (في متن الكتابة، في الجداول، بالتمثيل البياني)

2 -تحليل البيانات وتفسيرها (المفهوم ، الأنواع)

3 -طرق تحليل البيانات (أساليب الإحصاء الاجتماعي): المقاييس والمعاملات الإحصائية

طريقة التقييم

علامة الأعمال الموجهة 50% + الامتحان 50%.

المراجع

فضيل دليو : تقنيات تحليل البيانات في العلوم الاجتماعية والإعلامية، دار الثقافة للنشر والتوزيع، الأردن، 2010.

- Couratier Claire et Miquel Christian (2007): Les études qualitatives: théorie, applications, méthodologie pratique. Paris: L'Harmattan.
- MAURICE Angers (1997): Initiation pratique a la méthodologie des sciences humaines. Alger: Casbah.
- Miles, M., & Huberman, A. M. (1994): Qualitative data analysis: An expanded sourcebook (2nd ed.). Thousand Oaks, CA: Sage.
- MUCCHIELLI, R. (1970): Le questionnaire dans l'enquête psycho-sociale. Paris : Ed. Sociales Françaises.

المحتويات

تقديم

بسم الله الرحمن الرحيم والصلاة والسلام على أشرف المرسلين، وبعد..

في إطار التفاعل مع رغبة الجامعة في تغطية المواد الدراسية بمطبوعات جامعية تستدرك بعض النقص الملاحظ في المراجع المتخصصة، يسرنا أن نقدم لطلبة السنة الثانية (علوم الإعلام والاتصال) مطبوعة جامعية تجميعية تستجيب في خطوطها العريضة لمحتويات البرنامج الرسمي لمادة "عرض البيانات وتحليلها".

تعتبر هذه المادة تكملة للمفاهيم الأساسية المستخدمة في البحث، المناهج ومجالات الدراسة وكذا أدوات جمع البيانات، لأنه من المهم جدا تمكين الطلبة من مواصلة التعرف على معالجة البيانات باستخدام الأساليب الإحصائية المختلفة: **كيفية جمع وتصنيف البيانات وقياسها وإيجاد العلاقات بينها...**

وبذلك يكون الطلبة قد تعرفوا على مكونات العملية البحثية المتعلقة بمعالجة البيانات وتحليلها وتفسيرها باستخدام أهم الأساليب الإحصائية الوصفية والكمية، والتي حاولنا توصيفها ودعم شرحها بأمثلة واقعية ومن صلب إحدى التخصصات الجامعية في العلوم الاجتماعية: علوم الإعلام والاتصال. كما حاولنا مراعاة البساطة والوضوح في عرض ما رصدناه وجمعناه مما تيسر لنا من مراجع متخصصة حول موضوع مهم لذاته وللعديد من العلوم الاجتماعية والإنسانية في بعدها التطبيقي الذي أصبحت تستعمله.

من هذين المنظورين النظامي الخاص والنفعي العام تحدد أهمية وفائدة هذا العمل التجميعي المتواضع، الذي نرجو أن يكون في متناول فهم واستيعاب الطلبة الجامعيين الموجه إليهم، والله ولي التوفيق.

قسنطينة (الجزائر) 1440 -1441هـ - 2018-2019 م

1. طرق عرض البيانات وجدولتها

مقدمة: أنواع البيانات ومصادرها

قد تكون البيانات في الدراسات الاجتماعية والإنسانية وصفية أو كمية:

– بيانات وصفية: هي التي تتعلق بصفات معينة، وتصنف على أساس الاختلاف في نوع الصفة: ذكر–أنثى، هاو–محترف، أستاذ–طالب...

– بيانات كمية: هي التي تتعلق بمقدار وجود تلك الصفات أو كميتها، وتصنف إلى فئات متجانسة: 1–5، 6–10...

وتجمع هذه البيانات من مصادر غير مباشرة (ثانوية) ومباشرة (أولية).

– في الحالة الأولى (غير المباشرة) يكتفي الباحث بنقلها من مصادر ثانوية منشورة مثل السجلات والدوريات والكتب... وليس له أي دور في في تجهيزها وتبويبها وتصنيفها. وهي مفيدة في اقتصاد الوقت والجهد والمال، ولكنها قد لا تكون متطابقة مع الدراسة المعنية أو لا تكون وحدتها الإحصائية المستعملة متطابقة وخطة هذه الدراسة.

– أما إذا قام الباحث بإعداد وتجهيز البيانات بصفته الشخصية أو عن طريق أعوانه ودون الاعتماد على ما نشر من قبل، فإن المصدر حينئذ يسمى مصدرا أوليا مباشرا. وعادة ما يتم هذا الأمر في الدراسات الميدانية ووفقا لطبيعة الدراسة وخطتها.

1.1– عرض البيانات

بعد جمع البيانات لا بد من ترتيبها وتصنيفها بشكل يظهر العلاقات بينها. ويتم عرضها بعدة طرق أهمها ثلاث:

1.1.1– عرض البيانات ضمن سير الكتابة للموضوع

إن هذه الطريقة معقولة فيما لو كانت تتألف من عدد محدود من الأرقام، إلا أن الإحصاءات في أغلب الأحيان تتألف من أعداد كثيرة يصعب ذكرها كاملة في مضمون الكتابة.

2.1.1– عرض البيانات بواسطة الجداول

يتم ذلك بتصنيف المعلومات وترتيبها وفقا لبعض خواصها، وأهم أساليب الترتيب هي:

الترتيب التاريخي، الترتيب الهجائي، الترتيب الكمي، والترتيب الجغرافي... وطريقة العرض الجدولي تمتاز بالدقة، ولذلك فهي أهم أسلوب متبع لعرض المعلومات، وما يؤخذ على هذه الطريقة عدم إعطاء فكرة سريعة بمجرد إلقاء نظرة واحدة إلى الجداول. ومثال ذلك: جدول يمثل عدد طلبة قسم علوم الإعلام والاتصال خلال السنوات الدراسية 1999 – 2004.

جدول (1) عدد طلبة قسم علوم الإعلام والاتصال

السنة	1999–2000	2000–2001	2001–2002	2002–2003	2003–2004
عدد الطلبة	60	210	350	470	600

المصدر: إعداد شخصي مبني على معطيات افتراضية

3.1.1 – عرض المعلومات بواسطة التمثيل البياني

1.3.1.1– مفهوم التمثيل البياني وخصائصه

يستعمل التمثيل البياني للدلالة على أية عملية تترجم بواسطتها المعطيات في شكل تمثيل بصري. وكما هو معروف، فإن أهمية التمثيل البياني ترجع أساسا إلى خصائصه الذاتية، والتي تتمثل أساسا فيما يمتاز به من:

– كونه مشوق وفي متناول فهم مختلف أنواع الجمهور، ومن ثم فمن السهل أن يجذب إليه الانتباه ويعلق بذاكرة الجمهور.

– مساعدته على الملاحظة المباشرة والسريعة لأمور غير منتظرة، وكذا على تفهم الظاهرة المدروسة بمجرد النظر إليه.

2.3.1.1– أهم العناصر الشكلية للتمثيل البياني

تختلف الأشكال والرسوم البيانية باختلاف البيانات المراد عرضها. إن عدم وجود نماذج تمثيلية بيانية صالحة للتعميم عند فحص وتحليل وتلخيص أية معطيات، لا يمنع من إمكانية إعطاء بعض التوجيهات القاعدية والمبادئ الأولية التي تساعد على إعداد تمثيل بياني جيد.

بداية وبما أن أي تمثيل بياني يعتبر نتاجا لتركيب مجموعة من العناصر منظمة تبعا لمقاييس تقنية محددة، فإننا سنتعرض لأهمها فيما يأتي:

– العناوين الرئيسة والفرعية

إن العنوان الرئيس لأي تمثيل بياني يوضع عادة فوقه وفي الوسط. كما يجب، على العموم، أن يحتوي هذا العنوان على معلومات تخص ماهية بيانات الرسم، كيفية تصنيفها ومكانها وزمانها عند اللزوم... أما العناوين الفرعية، فالعرف هو وضعها تحت العناوين الرئيسة وببنط أصغر شريطة أن تحتوي على معلومات إضافية ضرورية، وإلا يستغنى عنها.

– تمايز التمثيلات البيانية

إن التفرقة بين الأشكال (الأعمدة أو الدوائر) أو الخطوط قد تتم بواسطة الألوان، التظليلات أو مختلف أنواع الخطوط (خط متواصل، متقطع، بالنقاط، بالنقاط والمطات، الخ). ولكنه من الناحية العملية، تفضل ا لاستعانة بالرسوم المتمايزة سواء في الهياكل أو الخطوط البيانية وذلك لتسهيل استنساخها.

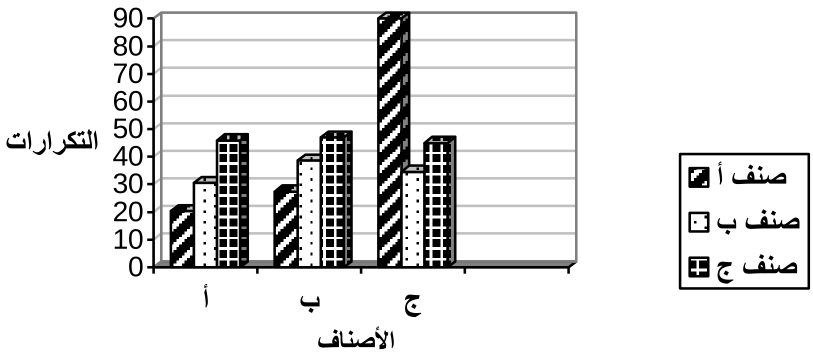
وتجدر الإشارة إلى ضرورة وضع تعريف مناسب لكل شكل أو منحنى داخل حيز التمثيل البياني أو في حيز مستقل مغلق (حيز أو مربع المفاتيح) في إحدى زواياه أو في حقل الملاحظات في أسفل الرسم، مع احترام التوازن الفضائي لكل ذلك، سواء تعلق الأمر بالأشكال أو الخطوط والمنحنيات، ومع الاحتراز أيضا من الكتابة فوق الرسوم أو الأشكال نفسها.

– مصدر التمثيل البياني

بعد الانتهاء من رسم التمثيل البياني، يجب ذكر المصدر الذي ينقل منه التمثيل البياني سواء كان كتابا أو منشورات هيئة متخصصة، مثل الديوان الوطني للإحصاء أو إحصاءات الأمم المتحدة، الخ، وذلك في أسفل الرسم. وأما إذا كان التمثيل البياني أصيلا، أي من إعداد الباحث نفسه، فمن المفيد ذكر مصادر المعطيات الممثلة فيه.

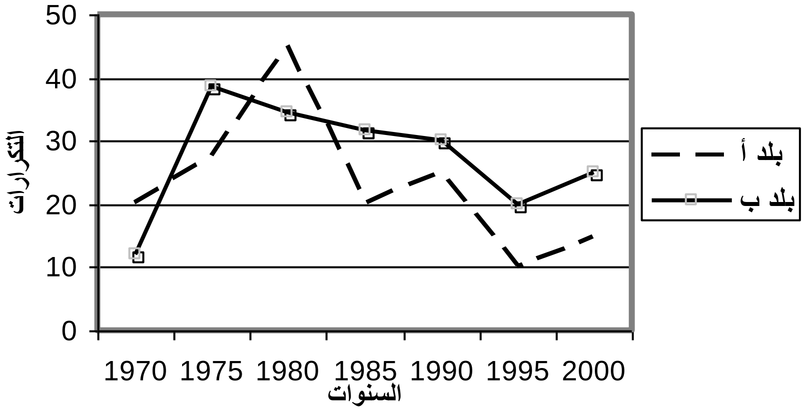
وتجدر الإشارة تباعا إلى تشخيص بياني لبعض الرسوم والأشكال البيانية النموذجية:

الشكل (1) : الأعمدة البيانية البسيطة



المصدر: إعداد شخصي

الشكل(2): الخطوط البيانية المعبرة عن تطور البطالة في بلدين



المصدر: إعداد شخصي

3.3.1.1- أهم طرق العرض البياني

من المعروف أن التمثيلات متعددة ومتنوعة ولكن معظمها يعتبر مجرد فروع تقوم على مبادئ أساسية للتمثيل. وقبل عرضها تجدر الإشارة إلى بعض مستلزمات تفضيل استعمال رسم أو شكل بياني على غيره من الرسوم أو الأشكال:

- طبيعة المعطيات التي يراد تمثيلها بيانيا،
- الأهداف المتوخاة من التمثيل البياني المختار،

- الجمهور الموجه إليه التمثيل البياني.

ويعتبر تحليل التوزيعات التكرارية أحد المبادئ الأساسية لأنه لا يمكن أن تكون المعالجة الإحصائية مناسبة إلا بعد معرفة كيفية توزيع المتغيرات. ولذا فمن المهم جدا معرفة المعطيات ليس فقط فيما يخص نظم العلاقات التي تربط بينها بل حقيقة طبيعة تغيرها الداخلي أيضا. وبناء عليه سنستعرض أهم طرق العرض البياني للمتغيرات:

- **الأعمدة أو المستطيلات البيانية**

تعتبر الأعمدة البيانية من أهم الطرق المستعملة في المقارنة بين أصناف متغير واحد أو بين عدة متغيرات في الزمان والمكان. فقد توضع هذه الأعمدة في شكل رأسي وحينئذ تسمى أعمدة أو أفقي وتسمى حينئذ مستطيلات ولكنه غالبا ما يوحد تصنيفها تحت اسم أعمدة (Bars).

وتذكر بعض المراجع سبعة أنواع من الرسوم البيانية بالأعمدة هي: -التمثيل بالأعمدة البسيطة، -التمثيل بالأعمدة المتعددة ، -التمثيل بالأعمدة المركبة ، -التمثيل النسبي بالأعمدة المركبة، -التمثيل بالأعمدة المزدوجة، -التمثيل بالأعمدة البسيطة ذات القيم السلبية والإيجابية ، - التمثيل بالنجوم أو التمثيل القطبي.

مع ملاحظة أن التمثيل البياني بواسطة الأعمدة لا يمكن أن يكون واضحا عندما تكون بيانات المتغير المدروس مستمرة، أي عندما تكون القيم متجاورة بحيث يمكن دائما إيجاد عنصر وسط، لأن كل قيمة من مدى السلسلة (طول الفئة) يمكن أن تكون قيمة للصفة، أي ممثلة لها. ومن ثم فالأنسب حينئذ استعمال التمثيل بالمدرجات التكرارية. لأن الأعمدة تعكس أصلا البيانات المتقطعة، فكل عمود يقابل صفة أو نوع أو متغير أو قيمة... وليس مجموعة قيم.

- **الدوائر**

في كثير من الأحيان تستعمل الدوائر أو أنصافها كأحد الطرق في التمثيل البياني لكميات متغير أو متغيرات الظاهرة المدروسة، أي في تمثيل أجزاء من ظاهرة كلية.

ويتمثل هذا الرسم البياني في دائرة تقسم إلى قطاعات يمثل كل واحد منها جزء محددا من إجمالي قيم الظاهرة أو متغيرا مقارنا مع متغيرات أخرى، بحيث تكون مساحة هذه القطاعات متناسبة بشكل دقيق مع كمية القيم أو المتغير الذي تمثله. وهذا يعني أن زوايا هذه القطاعات وأقواسها تكون أيضا كذلك.

وتجدر الإشارة إلى أن هذا النوع من التمثيل البياني مثله مثل الأعمدة النسبية مفيد في مقارنة نسب القيم المطلقة لمختلف الأصناف الممثلة.

– الخطوط البيانية

تعرضنا سابقا للمعالجة البيانية للمعلومات الكمية في أوقات محددة من الزمن، وكأننا بصدد أخذ صورة آنية لها. ولكن الباحث في العلوم الاجتماعية قد يحتاج إلى قيم معرفة تطور المتغيرات عبر الزمن (والتي تدعى بالسلاسل الزمنية) وتمثيلاتها البيانية. وفي هذه الحالة، لا يسعه إلا اللجوء إلى الرسوم البيانية الخطية المعروفة بفعاليتها في تمثيل السلاسل الزمنية أو ضعف أو شدة التغيرات في الظواهر الاجتماعية من فترة زمنية إلى أخرى.

فإذا أخذنا بعين الاعتبار طبيعة المعطيات، فإن هذا النوع من الرسوم البيانية يكون مناس بل عندما نكون بصدد معالجة عدد كبير من القيم المتوالية بيانيا، وعلى العكس من ذلك، يفضل تجنب استعماله عندما تكون القيم المراد تمثيلها قليلة.

كما أن هذا النوع من التمثيل الخطي مناسبا عندما نكون بصدد مقارنة عدة سلاسل أو متغيرات خلال نفس الفترة الزمنية المأخوذة، ولا يكون مناسبا عندما تكون المعلومة المراد إبرازها تشير إلى فروق قيمية في فترات زمنية مختلفة، أو عندما يكون عدم انتظام المعطيات حادا وعنيفا.

– تمثيلات بيانية أخرى

وهناك تمثيلات بيانية أخرى تعكس التوزيعات التكرارية والمتساوية (الجذوع والأوراق، الرسم البياني الصندوقي...)، السلاسل الزمنية (التمثيل النسبي بالخطوط المركبة، على شكل Z...)، الاقتران (الرسم البياني النقطي، بالحدين الأعلى والأدنى، بالكثافات...)، التمثيلات الإيقونية، الخرائطية، المختلطة...[*].

---

[*] لمزيد من التفصيل أنظر: فضيل دليو، التمثيلات البيانية في العلوم الاجتماعية، ديوان المطبوعات الجامعية، الجزائر، 2001.

2.1- تحضير البيانات

بعد أن يتم إدخال البيانات إلى الحاسوب تكون هذه الأخيرة جاهزة للتحليل ولاستخراج النتائج منها، وهي المرحلة التي يتم فيها تلخيص البيانات التي جمعت للتعبير عنها بشكل مختصر وبطريقة منظمة. أما فيما يتعلق بالأساليب الإحصائية الواجب استخدامها فان ذلك يعتمد على أهداف الدراسة وطبيعتها والفرضيات التي تم وضعها من قبل الباحث. ففي حالة الدراسات الوصفية فإن الباحث يلجأ إلى استخدام الأساليب الإحصائية الوصفي ة التي تصف الظاهرة أو المشكلة موضوع البحث، ولا تتطرق بتعمق إلى دراسة الظاهرة والتعرف على بعض نواحي العلاقة أو الارتباط بمتغيرات الدراسة. ومن أهم الأساليب الإحصائية الوصفي ة التكرار والوسط الحسابي والمنوال والوسيط والانحراف المعياري والتشتت.

أما في حال كون الدراسة ارتباطية وتهدف إلى التعمق في دراسة الظاهرة ووضع فرضيات وفحصها بشكل إحصائي فإن الباحث يلجأ إلى استخدام مقاييس الارتباط مثل اختبار سبيرمان وبيرسون والانحدار. كما قد يلجأ إلى بعض مقاييس الاختلاف مثل اختبار "ت" (t. test)...

ويتم في هذا الجزء التعرض للعمليات الأولية لتحليل البيانات والتي توصف أحيانا بعملية تحضير البيانات: تبويب البيانات، تكوين الجداول التكرارية والنسبية، وبشكل مختصر حيث أن مثل هذه الأساليب تكون مشروحة بشكل مفصل في كتب الإحصاء المتخصصة.

1.2.1- تكوين الجداول الإحصائية (الفئات والتكرارات والنسب)

إن العرض الجدولي للبيانات التي تفرغ من الاستمارات تتطلب تكوين جداول تكرارية إحصائية. ويستلزم تكوين هذه الجداول معرفة تحديد طول الفئات وتعيين حدودها وكيفية تفريغ البيانات ورصد تكراراتها.

- **تحديد طول الفئة أو مداها** (قبل معرفة حدودها): وهي، على العموم، القيمة الواقعة بين الحد الأدنى والحد الأعلى للفئة . وهناك ثلاث طرق (طريقة تقليدية وطريقتان إحصائيتان) لتحديد طول الفئة:

1 - الرجوع إلى الدراسات السابقة للاستفادة منها وتطبيق تقسيمها للفئات على بيانات دراستنا، أي اتباع الشائع في تقسيم البيانات المشابهة لبياناتنا.

2 - طريقة الإحصائي "ستورجز" (H. A. Sturges)، وتتمثل في المعادلة الآتية:

طول الفئة (ف)= المدى (الفرق بين أكبر قيمة وأصغر قيمة)
3.322+1 (لغ ن)

أو $$C = \frac{R-(X\max - X\min)}{1+3.322LogN}$$

حيث ن = عدد القيم، أي مجموع التكرارات.

وتجدر الإشارة هنا إلى أنه يفضل تطبيق هذه المعادلة عندما يكون عدد الأرقام المراد تبويبها في شكل فئات (أي عدد الفئات المقترح) محصورا بين 100 و1000. كما يجب إخضاع هذه المعادلة لاختبار المقارنة بين الوسط الحسابي (الذي سنراه لاحقا) المستخرج من البيانات المبوبة (بيانات الجدول التكراري) والمحسوب من البيانات غير المبوبة (الخام)، ويجب أن لا يكون الاختلاف جوهريا، وإلا عدل طول الفئة.

3- وهناك معادلة "يول" (Yule) التي لا تحسب عدد الفئات بالخوارزم (لغ.) بل بالجذر التربيعي لعدد المفردات: $2.5\sqrt[4]{N}$ ، أي أن طول الفئة يساوي "*المدى على* $2.5\sqrt[4]{N}$".

- **كيفية كتابة حدود الفئات:** هناك عدة طرق لكتابة حدود الفئات يراعى فيها دقتها وعدم تداخلها، إذ لا يصح مثلا كتابة الفئات بالطريقة الآتية:

5 – 10
10 – 15
15 – 20
20 – 25

لأن القيمة 10 مثلا متضمنة في الفئتين الأولى والثانية معا، وكذا القيمة 15 في الفئتين الثانية والثالثة وهكذا. بل يجب كتابة حدود الفئات بإحدى الطرق الآتية:

أولا) إذا كانت البيانات **غير مستمرة** (لا يمكن تجزئة وحدة قياسها الأولية: عدد الأطفال، عدد الطلبة، نوعهم –ذ، إ–...):

- دون استعمال المجالات:

5 – 9
10 – 14
15 – 19
20 – 24 (في حالة الحرص على تساوي الفئات)

- باستعمال طريقة المجالات بالنوعين الآتيين:

1- المجال [أ، ب] مجال مغلق من الطرفين تنتمي إليه جميع القيم الواقعة وسط المجال المغلق [ ].

2- المجال]أ ، ب[مجال مفتوح من الطرفين تنتمي إليه جميع القيم الواقعة بين القيمتين: "أ" و "ب"، ولا تنتمي إليه القيمتان: أ و ب.

وفي المثال السابق تكتب صيغة المجالات كالآتي:

[9 – 5]		] 10 – 4 [
[14 – 10]	أو	] 15 – 9 [
[19 – 15]		] 20 – 14 [
[24 – 20]		] 25 – 19 [

ثانيا) إذا كانت البيانات **مستمرة** نلجأ إلى استخدام طريقة المجالات بالنوعين الآتيين:

1- المجال] أ ، ب] مجال نصف مفتوح من الطرف الأول، تنتمي إليه جميع القيم الواقعة وسط المجال المفتوح من الطرف الأول باستثناء أول قيمة، أي القيمة أ.

2- المجال [أ ، ب [مجال نصف مفتوح من الطرف الثاني، تنتمي إليه جميع القيم الواقعة وسط المجال المفتوح الآخر باستثناء آخر قيمة، أي القيمة ب.

[9 – 4 [		] 10 – 5]
[14 – 9 [	أو	] 15 – 10]
[19 – 14 [		] 20 – 15]
[24 – 19 [		] 25 – 20]

مع ملاحظة أن حدود الفئات في الحالتين غير متطابقة (في الحالة الأولى تتوقف عند 24 وفي الحالة الثانية قد تنتهي عند 24.9).

أو طريقة المطات (–) المعادلة لعبارة "وأقل من":

5 –	أو	5 وأقل من 10
10 –		10 وأقل من 15
15 –		15 وأقل من 20
20 –		20 وأقل من 25
25 –		25 وأقل من 30

وهكذا...

ومعلوم أنه يفضل أن يكون طول مختلف فئات البيانات المجمعة متساويا إلا إذا اقتضت ضرورة البيانات وخصائص الظاهرة المدروسة غير ذلك، سواء في الفئتين الأولى والأخيرة أو في بعض الفئات الوسطى.

– تعيين الحد الحقيقي للفئة، مداها ومركزها:

<u>تختلف حدود الفئات باختلاف طبيعة البيانات</u> (متصلة / منقطعة):

– البيانات المتصلة: الحد الأعلى للفئة السابقة هو نفسه الحد الأدنى للفئة اللاحقة:

5 وأقل من 10

10 وأقل من 15

15 وأقل من 20

20 وأقل من 25

– البيانات المنقطعة (غير المستمرة): تكون حدود فئاتها غير متقاطعة، وتسمى الحدود الظاهرة. ولتحويلها إلى حدود متصلة نطرح ونضيف 0.5 في الحدين على التوالي. وتسمى الحدود المتحصل عليها بالحدود الفعلية للفئات:

بيانات غير مستمرة	حدودها الفعلية
5 – 9	[4.5 – 9.5 [
10 – 14	[9.5 – 14.5 [
15 – 19	[14.5 – 19.5 [
20 – 24	[19.5 – 24.5 [

ومن ثم **فمدى الفئة** في الجدول المتصل= الحد الأعلى– الحد الأدنى، أي 10–5 = 5 أو 9.5– 4.5 = 5

أما في الجدول المتقطع فإن مدى الفئة= الحد الأعلى– الحد الأدنى+ 1، أي (9–5)+1=5، أو الفرق بين الحد الأدنى للفئة الأولى والحد الأدنى للفئة الثانية (أو بين كل فئتين متتاليتين)،أي:10–5 = 5، و15–10=5...

وأما مركزها فهو يقع في منتصف المسافة بين حدي الفئة، أي منتصف الحد الأدنى زائدا الحد الأعلى، أي: $\frac{10+5}{2}=\frac{15}{2}=7.5$ (في مثال البيانات المتصلة) و $\frac{9+5}{2}$ أو $\frac{9.5+4.5}{2}=\frac{14}{2}=7$ (في مثال البيانات المنفصلة).

– **كيفية تفريغ البيانات وحساب التكرارات** : بعد تحديد طول الفئات وتعيين حدودها نعد جدولا بثلاثة أعمدة، نكتب في العمود الأول: الفئات، وفي العمود الثاني: تفريغ البيانات بالقيم أو بالخطوط، وفي العمود الثالث: التكرارات بالأرقام. ثم نقوم بتفريغ البيانات في الجدول بعد ترتيبها تصاعديا أو تنازليا:

– مثال (1): إذا حدد الباحث المستوى التعليمي في المستويات الآتية: ابتدائي، إكمالي، ثانوي، جامعي. وكانت لدينا ثلاثون استمارة وكل استمارة قد دون بها المستوى التعليمي المطابق لحالة المبحوث وكانت بيانات الاستمارات كالآتي:

1- ابتدائي	6- إكمالي	11- ابتدائي	16- ابتدائي	21- ابتدائي	26- ابتدائي
2- جامعي	7- إكمالي	12- جامعي	17- ثانوي	22- إكمالي	27- ابتدائي
3- ثانوي	8- إكمالي	13- إكمالي	18-جامعي	23-جامعي	28-إكمالي
4- ثانوي	9- ابتدائي	14- ابتدائي	19- ثانوي	24- إكمالي	29- ثانوي
5- إكمالي	10- ابتدائي	15- ابتدائي	20- إكمالي	25- ابتدائي	30- ابتدائي

فإنه يمكن تكوين جدول التفريغ التكراري على الشكل الآتي:

جدول (2) التفريغ التكراري للمستوى التعليمي

الفئات	تفريغ البيانات بالخطوط	التكرارات
ابتدائي	///// ///// //	12
إكمالي	///// ////	9
ثانوي	/////	5
جامعي	////	4
المجموع		30

وقد تمت عملية التفريغ عن طريق وضع علامة (/) أمام المستوى التعليمي المطابق للاستمارة مع الاستمرار في ذلك حتى نقل كل المستويات التعليمية الموجودة في الاستمارات الإحصائية إلى شكل خطوط في الجدول . وكلما تجمعت خمسة خطوط ضمت في شكل حزمة (/////).

وبالطبع فإن عدد خطوط الجدول لا بد وأن يساوي عدد الاستمارات.

– مثال (2): البيانات الآتية تمثل علامات خمسين (50) طالبا في مادة إعلامية، والمطلوب تفريغها في جدول توزيع تكراري:

29	28	25	26	26	27	25
30	33	34	32	31	30	34
32	31	33	30	34	31	31
39	35	31	33	31	32	30
38	37	35	39	36	37	38
38	42	37	44	36	40	35
44	49	40	45	41	39	43
						48

نقوم أولا بترتيب البيانات السابقة ترتيبا تصاعديا كما يأتي:

29	28	27	26	26	25	25
31	31	31	30	30	30	30
33	32	32	32	31	31	31
35	35	34	34	34	33	33
38	37	37	37	36	36	35
40	40	39	39	39	38	38
48	45	44	44	43	42	41
						49

نلاحظ من خلال هذا الترتيب أن أصغر قيمة هي 25 وأكبر قيمة هي 49 ، ثم نأخذ فئات متساوية ولنفرض أن عددها 5 في مثالنا هذا، لنقرر طول الفئة وذلك بقسمة المدى (هنا : 49 – 25 = 24) على عدد الفئات(24/5 = 4.8)مع تقريب الجواب إلى أعلى إلى أقرب عدد صحيح أي (5).

أو لنقرر انطلاقا، تأسيا ببعض الدراسات السابقة، أن مدى الفئة هو 5 مثلا.

أما إذا حسبنا مدى الفئات وفق معادلة "ستورجس" فسنحصل على مدى أقل بقليل:

$$\text{طول الفئة (ف)} = \frac{\text{المدى (الفرق بين أكبر قيمة وأصغر قيمة)}}{1+3.322\ (\text{لغ ن})}$$

$$= \frac{49-25}{1+3.322\ (\text{لغ ن})} = \frac{24}{1+3.322(1.698)} = 3.270 = 4$$

وسنقوم بتفريغ البيانات بالقيم وبالخطوط تبعا للافتراضين الأولين (عدد الفئات 5 وطولها 5) كما هو مبين أدناه في جدول التفريغ التكراري (3: أ، ب) لعلامات الطلبة:

جدول (3: أ، ب) التفريغ التكراري لعلامات الطلبة بالقيم والخطوط

الفئات	تفريغ البيانات بالقيم	التكرارات
25 وأقل من 30	25، 25، 26، 26، 27، 28، 29	7
30 وأقل من 35	30،30،30،30، 31،31، 31، 31، 31، 32، 32، 32، 33، 33، 33، 34، 34، 34	19
35 وأقل من 40	35، 35، 35، 36، 36، 37، 37، 37، 38، 38، 38، 39، 39، 39،	14
40 وأقل من 45	40، 40، 41، 42، 43، 44، 44	7
45 وأقل من 50	45، 48، 49	3
المجموع		50

الفئات	تفريغ البيانات بالخطوط	التكرارات
] 30 – 25 [	𝍸 //	7
] 35 – 30 [	𝍸 𝍸 𝍸 ////	19
] 40 – 35 [	𝍸 𝍸 ////	14
] 45 – 40 [	𝍸 //	7
] 50 – 45 [	///	3
المجموع		50

– تحويل الأرقام إلى نسب مئوية في الجداول البسيطة

إذا أردنا تحويل البيانات الرقمية المجدولة إلى نسب مئوية، فإننا نقوم بتقسيم العدد الجزئي على المجموع الكلي مضروبا في مائة (نسبة فئة الابتدائي في الجدول الموالي = $\frac{12\times100}{30} = \%40$).

وبعد تحويل الأرقام إلى نسب مئوية تجمع هذه النسب بعضها مع بعض لتكون مائة.

– المثال الأول: جدول (4) التفريغ التكراري والمئوي للمستوى التعليمي

الفئات	تفريغ البيانات بالخطوط	التكرار	التكرار المئوي %
ابتدائي	𝍸 𝍸 //	12	40
إكمالي	𝍸 ////	9	30
ثانوي	𝍸	5	16.67
جامعي	////	4	13.33
المجموع		30	100

– المثال الثاني: جدول (5) التفريغ التكراري والنسبي لعلامات الطلبة

الفئات	تفريغ البيانات بالخطوط	التكرار	التكرار النسبي %
[25 – 30[	卌 //	7	14
[30 – 35[	卌 卌 卌 ////	19	38
[35 – 40[	卌 卌 ////	14	28
[40 – 45[	卌 //	7	14
[45 – 50[	///	3	6
المجموع		50	100

2.2.1– البيانات المتجمعة وحسابها

إذا رجعنا إلى المثال الثاني وطرحنا السؤال الآتي: ما هو عدد الطلبة الذين تقل علاماتهم عن 30 نقطة؟ فالجواب بطبيعة الحال هو 7. أما عدد الطلبة الذين تقل علاماتهم عن 35 نقطة فهو 26 طالبا (7+ 19 = 26) وهكذا... نسمي هذا النوع من التكرارات بالتكرار المتجمع الصاعد (ك. ت. ص.).

أما إذا طرحنا السؤال بشكل آخر وقلنا ما هو عدد الطلبة الذين تساوي علاماتهم 25 نقطة فأكثر؟ بما أن أقل علامة هي 25، فإن عدد الطلبة الذين تساوي علاماتهم 25 نقطة فأكثر هو 50 طالبا. أما عدد الطلبة الذين تساوي علاماتهم 30 نقطة فأكثر فهو 50 طالبا (50–7=43)، وهكذا... نسمي هذا النوع من التكرارات بالتكرار المتجمع النازل أو الهابط (ك. ت. ن.). والجدول (6) الآتي يوضح ذلك:

الفئات	التكرار	ك. ت. ص	ك. ت. ن
[25 – 30[	7	7	50
[30 – 35[	19	26	43
[35 – 40[	14	40	24
[40 – 45[	7	47	10
[45 – 50[	3	50	3
المجموع	50	–	–

- **تحويل التكرارات المطلقة للبيانات المتجمعة إلى نسب مئوية**

تسمى التكرارات السابقة بالتكرارات المطلقة، ويمكن التعبير عنها بشكل نسب مئوية كما هو مبين في الجدول (7) الآتي:

جدول (7) التكرارات المتجمعة النسبية

الفئات	التكرار %	ك. ت. ص %	ك. ت. ن %
[25 – 30 [	14	14	100
[30 – 35 [	38	52	86
[35 – 40 [	28	80	48
[40 – 45 [	14	94	20
[45 – 50 [	6	100	6
المجموع	100	–	–

3.2.1– أنواع الجداول

يمكن تصنيف الجداول إلى: جداول بسيطة، جداول مزدوجة، جداول مركبة.

- **الجداول المبسطة:** هي الجداول التي تصنف فيها البيانات بموجب متغير واحد: كالعمر أو النوع أو المستوى التعليمي أو الدخل أو أي متغير آخر. ويبين الجدول (8) توزيع مائتي طالب من كلية العلوم الاجتماعية بموجب النوع، وهو نموذج لجدول بسيط.

جدول(8) يبين توزيع 200 طالب من كلية العلوم الاجتماعية بموجب النوع

النوع	العدد	النسبة المئوية
ذكور	150	75 %
إناث	50	25 %
المجموع	200	100 %

- **الجداول المزدوجة:** هي الجداول التي تصنف فيها البيانات بموجب متغيرين وتفيد هذه الجداول في التعبير عن العلاقات بين المتغيرات التي يراد توضيحها. ويمثل الجدول (9) نموذجا لجدول مزدوج يوزع طلاب كلية العلوم الاجتماعية بموجب النوع والتخصص.

ففي المثال السابق تعلق السؤال الأول ب النوع (ذكر، أنثى) والسؤال الثاني بالتخصص، فنجد أن نتائج كل سؤال تعطي معلومة منفصلة ولا تظهر لنا مثلا، كم عدد الإناث اللواتي يدرسن في قسم التاريخ؟، وهل هناك اختلاف واضح تبعا للتخصصات بين الذكور والإناث؟. أما عند استخدام الجداول المزدوجة لهذين السؤالين فإن القارئ يستطيع معرفة عدد كل الإناث والذكور في كل تخصص.

جدول (9) يوضح العلاقة بين النوع والتخصص

النوع / التخصص	ذكور		إناث		المجموع	
	ك	%	ك	%	ك	%
تاريخ						
فلسفة						
إعلام واتصال						
علم اجتماع						
علم النفس						
علم المكتبات						
المجموع						

- **الجداول المركبة:** هي الجداول التي تصنف فيها البيانات بموجب ثلاثة متغيرات أو أكثر، مثل الجدول (10).

الجدول (10) يوضح توزيع طلاب العلوم الاجتماعية بموجب النوع والتخصص والسنة الدراسية ومكان الولادة (ريف، مدينة)

التخصص / النوع / السنة الدراسية / مكان الولادة		تاريخ		فلسفة		إعلام واتصال		الخ.	المجموع	
		ذ	أ	ذ	أ	ذ	أ		ذ	أ
أولى	مدينة									
	ريف									
ثانية	مدينة									
	ريف									
ثالثة	مدينة									
	ريف									
المجموع										

4.2.1- شروط إعداد الجداول

تبدأ عملية تبويب البيانات، كما مر معنا، في جداول إحصائية بعد الانتهاء من عملية التفريغ وإحصاء الاستجابات. ويعتبر الجدول من الأساليب الإحصائية التي تساعد على فهم وتكوين فكرة سريعة ومختصرة عن المعلومات وعن طبيعتها ونظامها والعلاقات القائمة بينها، لهذا يفضل عند استعمال الجداول الابتعاد عن الجداول المعقدة وتلك التي تدمج في جدول واحد عدة مقارنات إلا للضرورة، لأن الجداول كالفقرة المكتوبة يجب أن تكون صياغتها واضحة وبسيطة وتربط الحقائق بعضها مع بعض بتسلسل منطقي. استنادا إلى هذا تبدأ عملية تبويب البيانات في جداول بسيطة مرورا بالجداول المزدوجة حتى تصل إلى الجداول المركبة. ولكن من الضروري عند إعداد مثل هذه الجداول مراعاة ما يأتي:

1 -تأخذ الجداول أرقاما متسلسلة خلال البحث سواء في ذلك الجداول التي توضع في صلب البحث أو تلك التي توضع ضمن ملاحق، ويدون رقم على رأس كل جدول، يتبعه عنوان الجدول الذي يشرح محتوياته بدقة، وإذا كان العنوان طويلا فيكتب على شكل هرم مقلوب. أما مصدر الجدول فيوضع في أسفله.

2 -تكتب عناوين موضوعات الأعمدة والصفوف بشكل مختصر ودقيق مع توضيح الوحدات المستعملة، كما ترتب درجات الصفة أو القيم ترتيبا تنازليا أو تصاعديا، بموجب الدرجة، أو القيمة، أو الزمان، أو الترتيب الأبجدي.

3 -يفضل أن يكون حجم الجدول متناسبا مع حجم الصفحة، وعندما يتطلب الجدول أكثر من صفحة، يشار بعبارة تابع جدول رقم كذا، وهو رقم الجدول الذي على الصفحة السابقة، ويفضل تكرار عنوان الجدول على الصفحة التالية كما يعاد كتابة عناوين الأعمدة والصفوف.

4 -تترك عادة الأماكن المخصصة للبيانات المجهولة بيضاء عند وجودها.

2. مفهوم تحليل البيانات وتفسيرها

إن تصنيف المعطيات وتنظيمها تبعا لأهداف الدراسة ليس هو الهدف النهائي لعملية البحث، إذ يجب أن تليها عملية استنتاج الملاحظات والنتائج العامة من كتلة المعطيات المصنفة والمنظمة. إن الجداول التصنيفية بحاجة إلى دراسة وتحليل مختلف عناصرها وأبعادها للوصول إلى نتائج مرتبطة أساسا بأهداف الدراسة وبفرضياتها الأولية.

ومن ثم يمكن القول أن **التحليل** يهدف إلى **تشخيص** و**توضيح** مختلف الخصائص المرتبطة بمتغيرات الدراسة والتي يمكن استنتاجها من الجداول التصنيفية. كما يهدف في المقام الثاني إلى **تحديد المعاني والأبعاد** الاجتماعية لهذه الخصائص، وهي العملية التي عادة ما تسمى **بالتأويل** في البحوث الكيفية أو **التفسير** في البحوث الكمية.

وكما هو ملاحظ تبقى متغيرات الدراسة هي العناصر المركزية لهذه المرحلة الأخيرة من مراحل البحث، ويبقى التحليل فيها يرتكز مباشرة على الجداول بأرقامها وإحصائياتها الرقمية المطلقة والنسبية.

إن نتائج هذا التحليل الاجتماعي يمكن التعبير عنها تبعا لـ"ج.أ. ديفيس" (J. A. DAVIS) بالمكتشفات/ النتائج (Findings) والتي هي –في تقديره– عبارة عن تقريرات حول الخصائص الإحصائية للمعطيات المحصل عليها عبر مختلف عمليات البحث. ومن أمثلة ذلك: "إن الأفراد الذين يحصلون على نقاط عالية في اختبارات الذكاء غالبا ما يحصلون على نقاط عالية في مقاييس تجنب الأحكام المسبقة".

أما التفسير أو التأويل، فإنه يتميز بخلاف ذلك، بطابع كيفي أكثر حدة. فهو ينطلق من "المكتشفات" ليقيم نتائج مرتبطة بمتغيرات الدراسة. وفي المثال السابق يمكن استنتاج ما يلي: "ضعف الذكاء يؤدي إلى معاداة الغرباء (SIERRA BRAVO, Ramón: 1979, 421).

وفي الأخير يمكن القول أن التحليل والتفسير هما مرحلتان أساسيتان في عملية التنظير التي قد تتوج عملية البحث. فـ"التحليل" يحول المعطيات الرقمية للجداول إلى تقريرات ذات طابع إحصائي يعتمدها "التفسير" في عمليته التنظيرية. وتتخلل هذه المهمة ثلاث عمليات:

- تتمثل الأولى في تحويل المعطيات الإحصائية الخاصة بكل متغير إلى تقريرات أو مفاهيم نظرية.

- وتعمل الثانية على تحويل المعطيات الإحصائية الخاصة بالعلاقات بين المتغيرات إلى اقتراحات أو تقريرات نظرية متعلقة بهذه العلاقات لتقارن بالفرضيات المطروحة.

- أما الثالثة فتتمثل في إدراج هذه الاقتراحات في إطار نظرية علمية تكون في الغالب قائمة.

– التفسير السببي وأنواعه

من المعروف في العلوم الاجتماعية أن التحليل أكثر شيوعا من التفسير، ولكنهما عادة ما يكونان متلازمين.

إن التفسير في هذا الإطار يعني تحديد سبب أو لماذا حدث شيء ما، ولكن تحديد سبب ظاهرة ما يعني تحديد القاعدة العامة التي تتحكم في وجودها ووجود أمثالها. وبهذا المعنى، فالتفسير عملية منطقية توصلنا إلى نتائج محددة بواسطة مقدمات محدد..

والتفسير عدة أنواع، أهمها على التوالي ثلاثة: الاستنتاجي، الاحتمالي، الوظيفي:

– الاستنتاجي: الإنسان حر / المتغير "س" إنسان" ⇐ س" حر.

– الاحتمالي: حوالي 95 % من العرب مسلمين / "س" عربي" ⇐ س" غالبا ما يكون مسلما ولكنه قد يكون غير ذلك.

– الوظيفي: يفسر حدوث الظاهرة بالأفعال أو الأنشطة المنجزة بغية تحقيق أهدافها، ويمكن تشخيصة كما يأتي: أن الوظيفة "و" في نظام أو نسق "ن" بتنظيم "ت" تيسر القيام بالنشاط "ش" في الظروف "ظ".

– أنواع التحليل

إن التحليل الاجتماعي بمعناه الواسع ذو أبعاد ومستويات مختلفة، فهو يتراوح بين أبسط مستوى (قراءة الجداول التصنيفية) وأعقده (تطبيق المعادلات الرياضية المعقدة).

ومن أهم أنواع التحليل ما يأتي:

1– التحليل الإحصائي.

2– التحليل الزمني (الديناميكي).

3- التحليل المركب (متعدد المتغيرات).

4- التحليل الدائري.

5- التحليل السوسيومتري.

تعتبر الأنواع الثلاثة الأولى استنتاجية أو سببية تستلزم توفر الشروط الآتية:

- أن تكون المتغيرات متلازمة في تغيرها، وهو ما يمكن معرفته بواسطة التحليل الإحصائي.

- أن يكون أحد المتغيرين أسبق من الناحية الزمنية من الآخر، ويقوم بتحديد ذلك التحليل الزمني.

- عدم وجود متغيرات مستقلة أخرى (وسيطة، دخيلة...) تتحكم في المتغير التابع، وهي من مهام التحليل المركب.

– أما النوع الرابع (التحليل الدائري)، فهو يستعمل عندما تطبق استمارتان على نفس المجموعة في فترتين زمنيتين مختلفتين، وذلك لمتابعة التغيرات التي قد تطرأ على المجموعة بين الفترتين ولاختبار "ثبات" البيانات.

– وأخيرا، فإن التحليل السوسيومتري مفيد جدا لدراسة البنية الاجتماعية ونظام الاتصال داخل الجماعات الصغيرة (دليو، 2005).

– التحليل الإحصائي والمقاييس

من المعروف أن عمليات التحليل الإحصائي نوعية ولا تصلح في جميع الحالات، لأن استعمال إحداها يخضع لتوفر شروط بعينها، ومن بينها مستوى القياس وطبيعته: اسمي، ترتيبي، مسافتي... الخ. وإذا أردنا إتقان ذلك يجب الرجوع بالتفصيل إلى متغيرات الاتجاه وخاصة مقاييسها بسلاليمها: سلم ليكرت، قوتمان، ثورستون...[*]

– بعض الأشكال والعمليات الإحصائية

إن عملية تصنيف المعطيات عادة ما تتم عن طريق الجدولة التي تعتبر نقطة انطلاق التحليل الإحصائي. ويعتبر "توزيع التكرارات" أبسط وأكثر أشكالها التصنيفية ذيوعا. إنها عبارة عن قائمة بجميع أصناف أو قيم المتغير التي تتميز بها جماعة ما، متبوعة بتكراراتها المناسبة.

---

[*] لمزيد من التفصيل أنظر: دليو (1998، 159-168).

وهي عادة ما تُجدْول في شكل تكرارات عادية (فردية ومجموعة) أو نسبية (%) أو متجمعة صاعدة أو نازلة (عادية ونسبية):

المتغيرات	التكرارات (ك)	التكرارات المتجمعة	التكرارات المتجمعة النسبية
1-	أرقام	أرقام متجمعة	نسب مئوية
2-	//	//	//
الخ-	//	//	//

وبالطبع يمكن تمثيل هذه التوزيعات التكرارية بيانيا في شكل مدرجات، مضلعات، منحنيات، دوائر، الخ.

بالإضافة إلى الجدولة والتوزيعات التكرارية هناك عمليات إحصائية أخرى تساعد على التعريف بخصائص المتغيرات ومنها: المتوسطات، مقاييس النزعة المركزية/ التشتت، معاملات التغير/ الارتباط، الخ.

ونظرا لكون الجدولة من أهم العمليات الخاصة بتصنيف البيانات، فإن قراءتها تتطلب على العموم الأخذ بعين الاعتبار القواعد العامة الآتية:

- المراجعة الدقيقة لمعطيات الجدول الاسمية والرقمية للتأكد من صحتها،
- التأكد من الهيكلة المناسبة للجدول من حيث طبيعة شكله ومتغيراته،
- القراءة المتأنية لعنوان الجدول المعني بالوصف ولأي ملاحظة أو تفسير مرتبطين به،
- اعتماد الوحدات المستعملة في الجدول (نسب مئوية أو غيرها)،
- اعتماد متوسطات الجدول ومجاميعه،
- الانتباه إلى مختلف التباينات الموجودة بين معطيات الجدول والعلاقة بينها،
- الانتباه إلى المفارقات والأمور غير العادية التي قد تلاحظ على نتائج الجدول.

وفيما يأتي نموذج تطبيقي للاستئناس به في كيفية قراءة ووصف الجداول التصنيفية:

– انطلاقا من الجدول الموالي الوارد في دراسة لـ"واليس" و"روبرتس" (SIERRA BRAVO, R.: 1979, 414-416) المطلوب قراءته ووصفه.

نسبة الأميين حسب السن، النوع (الجنس) واللون. 1952.

السن	البيض			غير البيض			اللونان معا		
	ذ	إ	مجموع	ذ	إ	مجموع	ذ	إ	مجموع
[14- 25[	1.2	0.5	0.8	7.2	1.4	3.9	1.8	0.6	1.2
[25- 35[	0.8	0.6	0.7	9.7	3.8	6.4	1.6	0.9	1.2
[35- 45[	1.2	0.5	0.8	7.5	5.9	6.6	1.7	1.0	1.3
[45- 55[	2.2	1.4	1.8	12.8	10.4	11.5	3.2	2.3	2.7
[55- 65[	3.6	3.4	3.5	19.4	16.9	18.1	4.7	4.4	4.5
65 فأكثر	5.6	4.4	5.0	35.8	31.2	33.3	7.6	6.2	6.9
14 فأكثر	2.1	1.5	1.8	12.7	8.2	10.2	3.0	2.1	2.5

المصدر: Statistical Abstract: 1955، 115

ملاحظة: عدد أفراد العينة: 25000 فردا يبلغ عمرهم 14 سنة أو يزيد، والأرقام عبارة عن متوسطات تقريبية تتماشى والاتجاه العام.

الجواب: تبعا لمعطيات الجدول السابق نطبق الخطوات الآتية:

1- يشير عنوان هذا الجدول والملاحظة المرفقة أعلاه إلى وجود أربعة متغيرات، أحدها تابع: نسبة الأمية، وثلاثة مستقلة: السن، النوع واللون أو العرق؛ وإلى أن المعطيات شملت عينة كبيرة نسبيا (25.000) من الذين بلغوا 14 سنة أو أكثر وذلك عام 1952.

2- إن الوحدات التي تشير إليها أرقام الجدول هي نسب مئوية تعبر عن متوسط عدد الأميين من بين كل 100 ساكن.

3- إن المتوسطات الكلية للجدول توجد في العمود الأيسر لكل مجموعة عمرية، وهي بالنسبة لجميع المجموعات العمرية (14 فأكثر) تساوي 2.50%، وهو متوسط نسبة الأميين حسب هذه العينة.

4- إن الفروق أو التباين بين مختلف المتوسطات عال، فهو يتراوح بين 35.8 % بالنسبة للذكور (65 سنة فأكثر) و 0.5 % بالنسبة للإناث (14-24 سنة) و(35-44 سنة).

5- أما بالنسبة لطبيعة علاقة المتوسطات بكل مقياس من مقاييس التصنيف: السن، النوع واللون، أي المتغيرات المستقلة، فيلاحظ ما يأتي:

- السن: يلاحظ من المجاميع أن نسبة الأمية هي تقريبا نفسها بالنسبة لثلاث مجموعات (14-44)، بعدها تزداد نسبة الأمية تدريجيا حتى تبلغ خمسة أضعاف الرقم الأولي عند مجموعة 60 سنة فأكثر (1.2 / 6.9).

- النوع: إن مقارنة عمود يْ الرجال والنساء بالنسبة للعرقين (اللونين) تشير إلى انخفاض محسوس في نسبة الأمية عند النساء، وخاصة مجموعة الفئة العمرية الأولى منهن (14-24).

- العرق: بمقارنة النسب المئوية الخاصة بالبيض وغير البيض، يلاحظ أيضا اختلاف كبير بلغ حده الأقصى (05 عند البيض / 33.3 عند غيرهم) عند مجموعة 65 فأكثر.

6- إن المقارنة بين مختلف الزيادات في نسب الأمية مقرونة بعمر العينة تشير إلى أنها أصغر عند البيض ومتقاربة بالنسبة للجنس (ذكور وإناث). وهذا يعني أن الارتفاع في نسبة الأمية عند غير البيض أكبر عند كبار السن وبأن الاختلاف بين الجنسين لا علاقة له بالسن. أما فيما يخص الجنس، فإن الدراسة المفصلة للجدول تشير إلى أن الزيادة في نسبة الذكور مقارنة مع الإناث هي أدنى عند البيض (2.5 − 1.5 =1) منها عند غيرهم (12.7-8.2 = 4.5).

7- أخيرا وفيما يخص المفارقات والأمور غير العادية، فإن أهم ما يمكن ملاحظته هو الانخفاض الشاذ في نسبة أمية فئة (25-34) من الذكور البيض (مقارنة مع سابقتها: 1.2 / 0.8)، عوض الارتفاع النسبي كما حدث عند النساء وغير البيض في باقي الفئات السابقة واللاحقة. وقد يكون ذلك راجع إلى كون هذه الفئة أكثر الفئات التي زودت الجيش الأمريكي بالجنود أثناء الحرب العالمية الثانية. وهذا يؤدي بنا إلى التساؤل عما إذا كان الجيش قد قام بتعليم عدد كبير من الجنود الأميين وإذا كان هذا التعليم قد شمل البيض أكثر من غيرهم.

3. طرق التحليل الإحصائي للبيانات

تمهيد

تعتبر الأساليب الإحصائية لوصف متغير واحد أول خطوة يجب أن يقوم بها الباحث عند تحليله البيانات تحليلا إحصائيا، وتستعمل لوصف كل متغير على حدة، وهي تلخص أهم صفات البيانات الرقمية في عدد واحد يرمز به ويدل عليها، يوضح هذا العدد ميل البيانات إلى التجمع في المركز أو ميلها إلى التشتت عن المركز، لهذا فالأساليب الإحصائية التي تصف متغيرا واحدا وصفا كاملا تتضمن: مقاييس النزعة المركزية ومقاييس التشتت:

1.3− مقاييس النزعة المركزية

بيّنا أن التوزيع التكراري بأنواعه المختلفة يهدف إلى تبويب البيانات الرقمية في صورة مناسبة موجزة توضح أهم معالمها الرئيسة، كما أوضحنا أهمية تمثيله بيانيا. لكن الدراسة الإحصائية لا تكتفي بمثل هذا الإيجاز في التبويب والعرض البياني بل تمضي إلى ما هو أعمق منهما، وذلك حينما تحاول أن تلخص أهم صفات تلك البيانات الرقمية في عدد واحد يرمز لها ويدل عليها، وقد يوضح هذا العدد نزعتها للتجمع أو التمركز حول قيمة معينة أو نزعتها للتشتت عنها.

وسنتناول هنا المقاييس الإحصائية المختلفة التي نعتمد عليها في معرفتنا لتمركز تلك البيانات دون دراسة تشتتها مسايرة للبرنامج الرسمي.

وتتلخص أهم مقاييس النزعة المركزية في المتوسط (أو الوسط)، والوسيط، والمنوال. وسيقصر تحليلنا الإحصائي عليها، وذلك لأنها أكثر تلك المقاييس فائدة وشيوعا.

1.1.3− المتوسط الحسابي (Arithmetic Mean)

المتوسط أكثر المقاييس الإحصائية انتشارا وذيوعا بين الناس لسهولته وفائدته التي تضفى عليه أهمية كبرى في حياتنا اليومية. فكثيرا ما يتحدث الناس عن متوسطات الأسعار في الشهر أو العام، ومتوسطات الأعمار واختلافها من جيل إلى جيل، ومن بلد إلى آخر، ومتوسطات الدخل الشهري والسنوي، وغير ذلك من الأمور العملية التي تتصل من قريب بحياتنا اليومية.

والناس في حسابهم لهذه المتوسطات وفي حديثهم عنها لا يستعينو ن إلا بالمتوسط الحسابي رغم أن هناك متوسطين آخرين كما سبق أن أشرنا إلى ذلك.

هذا وتختلف طرق حساب المتوسط الحسابي تبعا لمدى تبويب البيانات العددية التي تبدأ بها عمليات حساب المقاييس الإحصائية المختلفة.

وسنتناول في تحليلنا لطرق حساب المتوسط الحسابي، طريقة الدرجات الخام وطريقة التكرار وطريقة الفئات والطريقة المختصرة السريعة في حساب هذا المتوسط ثم نمر إلى حساب متوسط المتوسطات أو ما يسمى بالمتوسط الوزني.

– حساب المتوسط من الدرجات الخام (بيانات غير مبوبة)

المتوسط الحسابي للدرجتين 3، 5 هو 4 . وقد حصلنا على هذه النتيجة بأن جمعنا هاتين الدرجتين أي 3+5 = 8 ثم قسمنا حاصل الجمع على عدد الدرجات وهو 2 فأصبحت النتيجة مساوية $\frac{8}{2}$ = 4 أو $\frac{5+3}{2}$ = 4.

فالمتوسط الحسابي للدرجات الآتية: 12، 14، 25، 16، 15، 13، 17، 11، 18، 19.

يحسب بجمع هذه الدرجات ثم بقسمة الناتج على عددها، وبما أن مجموعها هو:

12+ 14+ 25+ 16+ 15+ 13+ 17+ 11+ 18+ 19 = 160 ،

وعددها هو 10.

إذن فالمتوسط الحسابي لهذه الدرجات = 160 ÷ 10 = 16

ويمكن أن نلخص هذه العمليات الحسابية في الصورة الآتية:

$$\text{المتوسط} = \frac{\text{مجموع الدرجات}}{\text{عددها}}$$

أي أن المتوسط: $\overline{\text{س}} = \frac{\text{مج س}}{\text{ن}} = \overline{X} = \frac{\sum x}{N}$

حيث أن: مج = المجموع = $\sum$

س = الدرجة = x

ن = عدد الدرجات = N

هذا ومن أهم مزايا هذه الطريقة دقتها الحسابية لخلوها من العمليات المختصرة التقريبية، ومن أهم عيوبها أنها تستغرق وقتا طويلا وخاصة عندما يزداد عدد الدرجات.

– حساب المتوسط من تكرار الدرجات (بيانات غير مبوبة):

عندما يزداد عدد الدرجات زيادة تبطئ من حساب المتوسط بالطريقة السابقة فإننا نلجأ إلى حساب تكرار هذه الدرجات تمهيدا لحساب المتوسط. والجدول (12) يوضح هذه الطريقة:

جدول (12) حساب المتوسط من تكرار الدرجات

الدرجة	التكرار	التكرار × الدرجة
س	ك	ك × س
3	2	6 = 2 ×3
4	1	4 = 1×4
5	6	30 = 6×5
6	3	18 = 3×6
المجموع	مج ك = ن= 12	مج (ك× س) = 58

وتتلخص خطوات حساب المتوسط في معرفة مجموع الدرجات، وهذا يساوي مجموع تكرار كل درجة في قيمتها وهو في مثالنا هذا 58، وبما أن مجموع التكرارات يساوي 12 إذن فالمتوسط يساوي 58 ÷ 12 = 4.83. ويمكن أن نلخص هذه العمليات في الصورة الآتية:

المتوسط = مجموع نواتج ضرب تكرار كل درجة في قيمتها / مجموع التكرارات

$$\bar{س} = \frac{مج\ (ك \times س)}{ن} = \bar{X}=\frac{\sum fx}{\sum f}$$

حيث يدل الرمز "ت" على التكرار = f.

وحيث تدل الرموز الأخرى على نفس ما دلت عليه في المعادلة السابقة.

هذا ومن أهم مزايا هذه الطريقة دقتها الحسابية وسرعة إجرائها وخاصة بالنسبة لطريقة الدرجات الخام، لكنها مع كل ذلك قد تستغرق من الفرد وقتا طويلا إذا كان المدى بين أكبر درجة وأصغر درجة كبيرا، كأن تكون مثلا أكبر درجة تفوق 100 وأصغر درجة 1 .

– المتوسط الحسابي "المرجح" (weighted)

إذا كنا بصدد حساب المتوسط من "تكرار" الدرجات واعتبرنا هذا الأخير بمثابة الأهمية النسبية للدرجات أو معاملاتها (في حالة نقاط مواد دراسية مثلا) أو أوزانها، نصف هذا المتوسط الحسابي بالمرجح، لأننا نعتبر التكرارات أوزانا مرجحة للدرجات (النقاط، الخيارات، الرتب،

الأفضليات...) أو معدلة لها. وبالطبع لا يختلف حسابه عن حساب المتوسط الحسابي من تكرار الدرجات.

ومن مجالات تطبيقه: تكرارات نقاط المواد التعليمية ذات المعاملات المختلفة، وتكرارات أفضليات أو خيارات مبحوث عندما يجيب على سؤال ترتيبي من قبيل: ما هي القنوات الفضائية التي تشاهدها؟ رتبها من واحد إلى عشرة. فتكرارات الرتبة الأولى ترجح بضربها في عشرة وتكرارات الرتبة الثانية في تسعة... والأخيرة في واحد. ثم تحسب النسب المئوية والمتوسطات على أساسها، أي من مجموعها.

- حساب المتوسطات من فئات الدرجات (بيانات مبوبة)

تعتمد طريقة حساب المتوسط من فئات الدرجات على منتصف الفئة لأنه يدل عليها ويلخصها في قيمة كما بينا من قبل.

وهكذا تصبح القيمة العددية لمنتصف الفئة ممثلة للدرجة التي تدل عليها كل فئة. فإذا كان منتصف الفئة الأولى في الجدول (19) هو 3 وامتدت حدودها من 2 إلى 4 وكان تكرارها 2 فإننا نلجأ في حسابنا لمجموع درجات هذه الفئة الأولى إلى ضرب تكرارها في منتصفه ا أي 2 × 3 = 6، ونكتفي بهذا الناتج على أنه يساوي تقريبا المجموع الذي نبحث عنه. وهكذا نستمر في حسابنا لمجموع درجات كل فئة بالطريقة نفسها حتى ننتهي من جدول التوزيع التكراري لفئات الدرجات، ثم نجمع هذه النواتج لنحصل بذلك على المجموع الكلي للدرجات. وعندما نقسم هذا المجموع على عدد الدرجات فإننا نحصل على المتوسط. والجدول(13) يوضح هذه الطريقة.

وهكذا نرى أن متوسط درجات هذا الجدول = $\frac{141}{17}$ = 8.29

ويمكن أن نلخص هذه العملية في الصورة الآتية:

$$\text{المتوسط} = \frac{\text{مجموع نواتج ضرب تكرار كل فئة في منتصفها}}{\text{مجموع التكرارات}}$$

أي أن : $\overline{\text{س}} = \frac{\text{مج ك ص}}{\text{مج ك}} = \overline{X} = \frac{\sum fxi}{\sum f}$

حيث يدل الرمز "ص" على منتصف الفئة X_i= .

هذا وبالرغم من السرعة التي تتميز بها هذه الطريقة عن الطريقتين السابقتين إلا أنها تتأثر بالتقريب الذي ينشأ من تلخيص جميع درجات كل فئة في منتصفها.

جدول (13) حساب المتوسط من فئات الدرجات لبيانات متقطعة

فئات الدرجات	التكرار	منتصف الفئة	التكرار × منتصف الفئة
	ك	ص	ك × ص
2–4	2	3	2 × 3 = 6
5–7	5	6	5 × 6 = 30
8–10	6	9	6 × 9 = 54
11–13	3	12	12 × 3 = 36
14–16	1	15	1 × 15 = 15
المجموع	مج ك=17= ن		مج (ك × ص) =141

- **حساب المتوسط بالطريقة المختصرة**

تهدف هذه الطريقة إلى اختصار وتبسيط العمليات الحسابية الطويلة التي ظهرت بوضوح في الطريقة السابقة، ولكن يشترط فيها تطبيقها على جداول التوزيعات التكرارية المنتظمة فقط.

وهي تعتمد في حسابها للمتوسط على:

1– افتراض أن منتصفات الفئات تتزايد تزايدا يساوي واحدا صحيحا. أي أن المنتصفات يتلو بعضها بعضا بالطريقة الآتية: 1، 2، 3، 4، 5، 6...

بدلا من الطريقة السابقة التي كانت تتزايد بها منتصفات الفئات تزايدا يساوي مدى كل فئة، أي بمعدل 3 درجات. أي أنها كانت تتزايد بالطريقة الآتية: 3 ، 6 ، 9 ، 12 ، 15.

2– هذا وتمضى هذه الطريقة في تبسيطها للعمليات الحسابية فتفترض مركزا لهذه المنتصفات يساوي صفرا ويقع بالقرب من منتصف التوزيع التكراري حيث تبدأ منه منتصفات الفئات الفرضية تزيد في كل خطوة واحدا صحيحا في اقترابها من النهاية الكبرى للتوزيع، وتنقص في كل خطوة من الخطوات واحدا صحيحا في اقترابها من النهاية الصغرى للتوزيع.

أي أننا نتخذ بدء التدريج في منتصف التوزيع بدلا من أوله، والمقارنة الآتية في الجدول (20) توضح هذه الفكرة:

جدول (14) مقارنة بين نوعين من أنواع التدريج

التدرج الذي يبدأ من أوله	0	1	2	3	4	5	6
التدريج الذي يبدأ من منتصفه	3−	2−	1−	0	1+	2+	3+

ونستطيع أن نلاحظ في وضوح مدى تناقص القيمة العددية للتدريج الثاني عن التدريج الأول في المثال السابق.

3− نقوم بحساب المجموع الفرضي للتكرارات بضرب التكرار في المنتصف الفرضي.

هذا وسنستعين بهذه الوسائل المختصرة في حسابنا للمتوسط من فئات الدرجات في الجدول (15).

جدول (15) حساب المتوسط من فئات الدرجات بالطريقة المختصرة

الفئات	التكرار	المنتصف الفرضي للفئة	التكرار × المنتصف الفرضي
	ك	ض	ك × ض
2−4 5−7	2 5	− 2 − 1	− 4 − 5 − 9
8−10	6	0	0
11−13 14−16	3 1	+ 1 + 2	+ 3 + 2 + 5
	17		− 4

يدل العمود الأول في الجدول السابق على فئات الدرجات، وقد وضعنا خطا فوق الفئة التي تمتد أطرافها من 8 إلى 10 وخطا تحتها لأننا فرضنا أنها تقع في نصف التوزيع ثم فرضنا أن منتصف هذه الفئة يساوي صفرا كما هو مبين بالعمود الثالث وحسبنا تدريج منتصفات الفئات التي تسبقها وتمتد منها إلى النهاية الصغرى للتوزيع على أساس تناقصها التدريجي الذي يساوى −1 لكل خطوة، وهكذا يمتد التدريج بالطريقة الآتية: −1 ، −2 .

حسبنا منتصفات الفئات التي تليها وتمتد منها إلى النهاية الكبرى للتوزيع على أساس تزايدها التدريجي الذي يساوي +1 لكل خطوة، وهكذا يمتد تدريجها بالطريقة الآتية: +1 ، +2 .

هذا ويدل العمود الثاني على تكرار فئات الدرجات، أما العمود الرابع فيدل على نواتج ضرب التكرار في المنتصفات الفرضية للفئات. وقد سجلنا مجموع الأعداد السالبة في أسفلها وإلى يسارها، وسجلنا أيضا مجموع الأعداد الموجبة في أسفلها وإلى يسارها ليسهل علينا حساب المجموع الكلي لنواتج ضرب التكرار في المنتصفات الفرضية للفئات.

4- نقوم بحساب المتوسط الفرضي الذي هو يساوي ناتج قسمة المجموع الفرضي لنواتج ضرب التكرار في المنتصفات الفرضية لكل فئة على مجموع التكرارات.

أي أن: المتوسط الفرضي = $\frac{\text{مج}(\text{ك} \times \text{ض})}{\text{مج ك}} = \frac{\sum fu}{\sum f}$

حيث تدل "ض" على المنتصفات الفرضية للفئات = u .

وهو في مثالنا يساوي $\frac{-4}{17} = -0.235$

5- نصحح التقدير الفرضي الخاص بمدى الفئة فهو في الحقيقة لا يساوي واحدا صحيحا كما فرضنا، ولكنه يساوي 3 ، إذن فعلينا أن نضرب هذا الناتج في 3 لنصحح هذا التقدير الفرضي.

أي $3 \times (-0.235) = -0.705$

6- ثم نصحح كذلك تقديرنا الفرضي الخاص بمنتصف الفئة التي بدأ منها تدريج المنتصفات: فمنتصف الفئة 8 – 10 التي بدأ منها التدريج الفرضي مساويا للصفر وحقيقته 9 ، إذن فعلينا أن نبدأ حسابنا من 9 حتى نصحح هذا الفرض الأخير، وذلك بإضافته إلى النتيجة السابقة أي أن المتوسط الحقيقي بالطريقة المختصرة يحسب كما يأتي:

المتوسط الحقيقي = $3(-0.235) + 9 = -0.705 + 9 = 8.29$

وهذا هو المتوسط نفسه الذي حصلنا عليه في الطريقة السابقة التي كانت تعتمد على المنتصفات الحقيقية للفئات وعلى تكرار كل فئة.

وهكذا يمكن أن نلخص هذه الخطوات في المعادلة الآتية

المتوسط الحقيقي=(مدى الفئة×المتوسط الفرضي)+منتصف الفئة التي بدأ منها تدريج المنتصفات

=مدى الفئة(مجموع نواتج ضرب التكرار في المنتصفات الفرضية للفئات)+منتصف الفئة الحقيقي التي بدأ منها التدريج

مجموع التكرارات

$$\bar{س} = ف \times \frac{(\text{مج ك ض})}{\text{مج ك}} + ص \quad = \quad \bar{X} = A + \frac{\sum fu}{\sum f}.C$$

حيث تدل:

"ف" على مدى الفئة = C

"ك" على التكرار

"ص" على المنتصف الحقيقي للفئة التي بدأ منها التدريج = A.

- **حساب المتوسط في الجداول التكرارية المفتوحة**

رأينا أن حساب متوسط البيانات المبوبة يعتمد على مراكز الفئات، ولكن هذه الأخيرة قد يصعب تحديدها عندما تكون حدود بعض الفئات (الأولى و/أو الأخيرة) مفتوحة:

جدول (16): بيانات حجم المشاهدة اليومية

فئات	تكرارات
[ربع ساعة-نصف ساعة[	15
[نصف ساعة- 1ساعة[	22
1 ساعة	10
]ساعة- ساعتان[	03
المجموع	50

في هذه الحالة يستبدل المتوسط بالوسيط أو المنوال، أو يلجأ إلى الاستغناء عن بيانات الفئة المفتوحة (يحسب المتوسط من دون اعتماد بياناتها)، ولكن الأفضل استخدام العلاقة العامة التقريبية بين مقاييس النزعة المركزية الثلاثة في تقدير قيمة المتوسط الحسابي، فيحسب المنوال والوسيط ثم توجد قيمة المتوسط من المعادلة الآتية:

$$\bar{س} = \frac{3 \text{ وسيط} - \text{المنوال}}{2}$$

- **متوسط المتوسطات أو المتوسط الوزني**

إذا كان متوسط مجموعة ما من الدرجات يساوي 4 وكان متوسط مجموعة أخرى مساويا 6 فقد يتبادر إلى الدهن أن متوسط المجموعتين يحسب بجمعهما وقسمتهما على اثنين، كما يأتي:

$$\frac{6+4}{2} = \frac{10}{2} = 5$$

ولن تكون هذه الإجابة صحيحة إلا إذا كان عدد درجات المجموعة الأولى مساويا لعدد درجات المجموعة الثانية، ولنضرب لذلك مثالا آخر:

المجموعة الأولى تتكون من: 3 ، 4 ، 5

ومتوسطها = $\frac{5+4+3}{3} = \frac{12}{3} = 4$

المجموعة الثانية تتكون من: 5 ، 6 ، 7

ومتوسطها = $\frac{7+6+5}{3} = \frac{18}{3} = 6$

ومتوسط المتوسطين أو المتوسط العام للمجموعتين يحسب بالطريقة المألوفة وذلك بجمع درجات المجموعتين ثم بقسمة الناتج على عدد درجات المجموعتين.

أي أن المتوسط العام = $\frac{(7+6+5)+(5+4+3)}{3+3} = \frac{18+12}{6} = 5$

أي أنه في هذه الحالة فقط = $\frac{6+4}{2} = 5$

حيث يدل الرقم 4 على متوسط المجموعة الأولى، ويدل الرقم 6 على متوسط المجموعة الثانية، ويدل الرقم 2 على عدد المتوسطات و هو في هذه الحالة 2 فقط.

- أما عندما لا يكون عدد درجات المجموعة الأولى مساويا لعدد درجات المجموعة الثانية فإن متوسط المتوسطات يحسب بالطريقة التالية:

المجموعة الأولى تتكون من: 2 ، 3 ، 4 ، 5 ، 6

ومتوسطها = $\frac{6+5+4+3+2}{5} = \frac{20}{5} = 4$

والمجموعة الثانية تتكون من: 5 ، 6 ، 7

ومتوسطها = $\frac{7+6+5}{3} = \frac{18}{3} = 6$

وقد يتبادر إلى الدهن أن متوسط الاثنين = $\frac{6+4}{2} = 5$

وعندما نحسب متوسط المتوسطين بالطريقة الجامعة التي اتبعت في حساب المتوسط العام نحصل على:

المتوسط العام = $\frac{(7+6+5)+(6+5+4+3+2)}{3+5} = \frac{18+20}{8} = \frac{38}{8} = 4.75$

والاختلاف بين هذا المتوسط الأخير 4.75 والمتوسط الذي حسبناه أولا وهو 5 نتج عن اختلاف عدد درجات المجموعة الأولى عن المجموعة الثانية ويمكن أن نلخص هذه الطريقة في المعادلة التالية: متوسط المتوسطات=

= مجموع درجات المجموعة الأولى + مجموع درجات المجموعة الثانية
عدد درجات المجموعة الأولى + عدد درجات المجموعة الثانية

وهناك طريقة أخرى لحساب متوسط المتوسطات:

بما أن المتوسط = $\frac{\text{مجموع الدرجات}}{\text{عددها}}$ ،

فإن مجموع الدرجات = المتوسط × عدد الدرجات (الناتج في المقام).

وهكذا يمكن أن نكتب معادلة متوسط المتوسطات في صورة أبسط من الصورة السابقة إذا عوضنا عن مجموع الدرجات بما يساويه.

م م =(متوسط المجموعة الأولى×عدد درجاتها)+(متوسط المجموعة الثانية×عدد درجاتها)
عدد درجات المجموعة الأولى + عدد درجات المجموعة الثانية

أي أن متوسط المتوسطات = $\frac{(م1 \times ن1) + (م2 \times ن2)}{ن1 + ن2}$

حيث أن: م1 = متوسط المجموعة الأولى

ن1 = عدد درجات المجموعة الأولى وهو يساوي أيضا عدد أفراد المجموعة الأولى

م2 = متوسط المجموعة الثانية

ن2 = عدد درجات المجموعة الثانية وهو يساوي أيضا عدد أفراد المجموعة الثانية.

وباستخدام هذه المعادلة الأخيرة يمكن أن نستخرج متوسط المتوسطات، وذلك بمعرفة:

م1 = 4 ، ن1 = 5

م2 = 6 ، ن2 = 3

متوسط المتوسطات = $\frac{(3\times6)+(5\times4)}{3+5} = \frac{18+20}{8} = \frac{38}{8} = 4,75$

وهذه النتيجة هي النتيجة نفسها التي حصلنا عليها بالطريقة المطولة السابقة.

ويسمى أحيانا متوسط المتوسطات بالمتوسط الوزني، وذلك لأننا نضرب المتوسط الأول في عدد درجاته، أي أننا نريد وزنه، وكذلك نضرب المتوسط الثاني في عدد درجاته أي أننا نريد وزنه.

وليست هذه الطريقة قاصرة على حساب متوسط متوسطين، بل يمكن أن تمتد لأكثر من ذلك، ولنضرب المثل التالي الذي يهدف إلى حساب متوسط المتوسطات الأربعة الآتية:

م1 = 7 ن1 = 7

م2 = 8 ن2 = 25

م3 = 6 ن3 = 35

م4 = 11 ن4 = 33

المتوسط الوزني= $\frac{(33\times11)+(35\times6)+(25\times7)+(7\times7)}{33+35+25+7} = \frac{363+210+200+49}{100} = 8.22$

- **الخواص الإحصائية للمتوسط**

تتلخص أهم الخواص الإحصائية للمتوسط الحسابي فيما يلي:

أ- مجموع الانحرافات: مجموع الانحرافات عن المتوسط يساوي صفرا. والانحراف هو مدى بعد أو قرب أية درجة ما عن المتوسط.

فمتوسط الدرجات التالية: 1، 4، 7 ، 9، 13 ، 17، 19

يحسب بجمعها وقسمة المجموع على عددها أي $\frac{70}{7} = 10$

ويحسب انحراف كل درجة عن المتوسط بطرح المتوسط منها:

الانحراف = الدرجة – المتوسط

وهكذا نرى أن انحراف الدرجة 1 = 1 – 10 = –9

وانحراف الدرجة 4 = 4 – 10= –6

وعندما نستمر في حسابنا لهذه الانحرافات نصل إلى الدرجة الأخيرة حيث نرى أن:

انحراف الدرجة 19 = 19 – 10 = 9

الجدول (17) يوضح الدرجات وانحرافاتها عن المتوسط

الدرجة	الانحراف: الدرجة –المتوسط
1	−9
4	−6
7	−3
9	−1
	−19
13	+3
17	+7
19	+9
	+19
مج = 70	مج = 0

وهكذا نرى أن مجموع الانحرافات السالبة يساوى −19 ومجموع الانحرافات الموجبة يساوي +19 والمجموع الكلي للانحرافات يساوى صفرا.

ولهذه الخاصية أهمية كبرى في حساب المتوسط بالطريقة المختصرة كما سبق أن بينا ذلك في تحليلنا لتلك الطريقة، وذلك عندما فرضنا متوسطا تخمينيا وحسبنا مجموع الانحرافات بالنسبة لذلك المتوسط التخميني، ثم صحّحنا هذا المجموع ليصبح مساويا للصفر في حسابنا للمتوسط الحقيقي. وتعتمد الطريقة العامة لحساب المتوسط على هذه الخاصية أيضا، فلو فرضنا أن "م" متوسط الدرجات "س_1 ، س_2" ينحرفان انحرافا سالبا عن هذا المتوسط وأن س_3، س_4 ينحرفان انحرافا موجبا عن هذا المتوسط فإن مجموع الانحرافات السالبة = مجموع الانحرافات الموجبة.

أي أن: $(\text{م}-\text{س}_1)+(\text{م}-\text{س}_2)=(\text{س}_3-\text{م})+(\text{س}_4-\text{م})$

$$\text{م}+\text{م}+\text{م}+\text{م}=\text{س}_1+\text{س}_2+\text{س}_3+\text{س}_4$$

$$4\,\text{م}=\text{س}_1+\text{س}_2+\text{س}_3+\text{س}_4$$

$$\text{م}=\frac{\text{س}_1+\text{س}_2+\text{س}_3+\text{س}_4}{4}$$

$$\text{المتوسط} = \frac{\text{مجموع الدرجات}}{\text{عددها}} \quad ، \quad \overline{\text{س}} = \frac{\text{مج س}}{\text{ن}}$$

وهذه هي المعادلة العامة التي تستخدم في حساب المتوسط من الأرقام الخام.

والمتوسط بهذا المعنى هو مركز الثقل أو مركز الاتزان الذي تتعادل بالنسبة له جميع القوى أو جميع فروق هذه القوى أو الانحرافات.

ب – الدرجات المتطرفة: يتأثر المتوسط بالدرجات القريبة منه تأثرا قليلا، ويتأثر بالدرجات البعيدة عنه تأثرا كبيرا.

فمتوسط الدرجات الآتية (2، 3، 4، 5، 6) يحسب بجمعها وقسمة الناتج على عددها، أي أن:

$$\overline{\text{س}} = \frac{6+5+4+3+2}{5} = \frac{20}{5} = 4$$

وإذا أضفنا إلى تلك الدرجات درجة قريبة من المتوسط ولتكن 5 ثم حسبنا المتوسط بعد ذلك لوجدنا أن:

$$\overline{\text{س}} = \frac{6+5+5+4+3+2}{6} = \frac{25}{6} = \frac{1}{6} + 4$$

$$(\text{أي } \frac{1}{6} + \frac{24}{6} = \frac{25}{6} = 4{,}16)$$

أي أن زيادة المتوسط الجديد عن المتوسط القديم تساوي $\frac{1}{6}$

وإذا أضفنا إلى تلك الدرجات 16 بدلا من إضافة 5 ثم حسبنا المتوسط بعد الإضافة لوجدنا أن:

$$\overline{\text{س}} = \frac{16+6+5+4+3+2}{6} = \frac{36}{6} = 6$$

أي أن زيادة المتوسط الجديد عن المتوسط القديم تساوي اثنين. وهذا الفرق الأخير أكبر من الفرق السابق لأن 16 تبعد عن المتوسط 4 أكثر مما تبعد 5 عن ذلك المتوسط نفسه، أي أنها أكثر تطرفا منها.

وهذه الخاصية توضح أهم عيوب المتوسط الحسابي، أي أن القيم المتطرفة في التوزيع تؤثر تأثيرا قويا على المتوسط، وقد تجعله أحيانا غير صالح كمقياس من مقاييس النزعة المركزية، لأنه في تلك الحالة يعطينا صورة خاطئة عن حقيقة تجميع البيانات العددية.

ج- عدد الدرجات: يتأثر المتوسط بعدد الدرجات، ويميل إلى الاستقرار كلما كان هذا العدد كبيرا، فعندما يكون العدد 100 مثلا فإن تأثر المتوسط بأية درجة يحسب على أنه أجزاء من مائة لأن هذه المائة تمثل مقام الكسر الذي نحسب منه المتوسط. وعندما يكون العدد 1000 مثلا فإن تأثر المتوسط بأية درجة يحسب على أنه أجزاء من الألف، وهكذا نرى أنه كلما زاد عدد الدرجات، زاد تبعا لذلك ميل المتوسط إلى الاستقرار وقل ميله للتغير والتذبذب.

د- جمع المتوسطات: كما رأينا في مبحث المتوسط الوزني فإن المتوسطات لا تجمع إلا عندما يتساوى عدد درجات المجموعات أي عدد أفراد كل جماعة لأن كل فرد يحصل على درجة، والجدول (18) يوضح ذلك:

جدول (18) جمع المتوسطات

درجات المجموعة الأولى	درجات المجموعة الثانية	مجموع الدرجات
6	4	10 = 4 + 6
9	8	17 = 8 + 9
11	9	20 = 9 + 11
16	12	28 = 12 + 16
23	22	45 = 22 + 23
مج = 65	مج = 55	مج = 120
المتوسط:5/65 =13	المتوسط: 5/55=11	المتوسط: 5/120 = 24

ومن هذا نرى أن: 13 + 11 = 24 .

أي أن: متوسط المجموعة الأولى + متوسط المجموعة الثانية = متوسط مجموع درجات المجموعتين.

هـ- طرح المتوسطات: تطرح المتوسطات عندما يتساوى عدد درجات المجموعات والجدول (19) يوضح هذه الفكرة.

جدول (19) طرح المتوسطات

درجات المجموعة الأولى	درجات المجموعة الثانية	فرق الدرجات
6	4	2 = 4 − 6
9	8	1 = 8 − 9
11	9	2 = 9 − 11
16	12	4 = 12 −16
23	22	1= 22 −23
مج = 65	مج = 55	مج = 10
المتوسط = 13	المتوسط = 11	المتوسط = 2

ومن هذا نرى أن: 13 –11 = 2

أي أن :

متوسط المجموعة الأولى– متوسط المجموعة الثانية= متوسط فرق درجات المجموعتين.

فوائد المتوسط:

تتلخص أهم الفوائد العملية التطبيقية للمتوسط في كونه يعتمد:

1– كمعيار قياسي: تعتمد المعايير الحيوية المختلفة على المتوسط. ولهذا يقاس ذكاء الفرد بالنسبة لمتوسط ذكاء جيله وأقرانه، ومدى انحرافه عن هذا المعيار زيادة ونقصانا. وينسب وزنه وطوله وحجمه إلى معايير أقرانه أيضا. ولهذا تصنع الملابس المختلفة لتناسب متوسطات أطوال وأحجام كل عمر من أعمار الإنسان وبما أن هذه المعايير تختلف في بعض نواحيها من بيئة لأخرى، لذلك نرى أن لكل بيئة معاييرها الخاصة بها. ومن هذا نرى خطأ نسبة الفرد إلى معايير غير معايير بيئته.

2– للمقارنة: تستخدم المتوسطات أحيانا لمقارنة مجموعة من الأفراد بمجموعة أخرى. كمثل مقارنة متوسط درجات فصل دراسي ما في امتحان للحساب بمتوسط درجات قسم دراسي ما في امتحان للغة أجنبية بمتوسط درجات قسم دراسي آخر بالنسبة للامتحان نفسه. هذا ولا تصح هذه المقارنة إلا إذا كانت المجموعات متجانسة وتقبل خ صائصها مثل تلك المقارنات. ومن أمثلة المقارنات الخاطئة ما يقوم منها على مقارنة متوسط أعمار الناس في بيئة صناعية أغلبها من الشبان بمتوسط أعمار الناس في بيئة زراعية قد يكون أغلبها من الأطفال والشيوخ، ولهذا تعتمد

شركات التأمين على دراسة متوسطات الأعمار بالنسبة لكل مهنة، وكل عمر، حتى تكون نتائجها صحيحة.

2.1.3– الوسيط (Median)

الوسيط هو النقطة التي تقع تماما في منتصف توزيع الدرجات بعد ترتيبها ترتيبا تصاعديا أو تنازليا بحيث يسبقها نصف عدد الدرجات ويتلوها النصف الآخر. ويحسب الوسيط بعدة طرق وذلك تبعا لمدى تبويب البيانات:

أولا) حساب الوسيط من الدرجات الخام:

ويعتمد حساب الوسيط اعتمادا كبيرا على عدد الدرجات ونوعها فرديا كان أم زوجيا. ولهذا تختلف طريقة حساب الوسيط تبعا لاختلاف هذا العدد من حيث كونه فرديا أو زوجيا.

– حساب الوسيط عندما يكون عدد الدرجات فرديا:

عندما نحسب الوسيط للدرجات الآتية:

17 ، 2 ، 5 ، 7 ، 10 ، 9 ، 8 .

فإننا نرتبها أولا ترتيبا تصاعديا كما يأتي:

2 ، 5 ، 7 ، 8 ، 9 ، 10 ، 17 .

ثم نبحث بعد ذلك عن النقطة التي تتصف هذه الدرجات، فنرى أنها تقع تماما عن د الدرجة 8 لأن عدد الدرجات التي تسبقها 3 وهي 2، 5، 7، وعدد الدرجات التي تليها 3 أيضا وهي 9، 10، 17 .

ويمكن أن نصل إلى معرفة ترتيب هذه النقطة وذلك بقسمة عدد الدرجات على 2 مع التقريب، أي $\frac{7}{2}$ =3.5 وعندما نقرب هذا الناتج إلى أقرب عدد صحيح نصل إلى أنه يساوي 4 .

وهكذا نستطيع أن نحسب ترتيب الدرجات لنصل إلى الدرجة التي ترتيبها الرابع بالنسبة لتدريج تلك الدرجات، فنرى أن العدد 2 ترتيبه الأول، والعدد 5 ترتيبه الثاني، والعدد 7 ترتيبه الثالث، والعدد 8 ترتيبه الرابع. أي أن الوسيط هو 8 .

ونستطيع أيضا أن نحسب ترتيب الدرجات من الطرف الآخر لتدريجها فنرى أن العدد 17 ترتيبه الأول، والعدد 10 ترتيبه الثاني، والعدد 9 ترتيبه الثالث، والعدد 8 ترتيبه الرابع. أي أن الوسيط هو 8 .

وتتلخص طريقة حساب وسيط الدرجات عندما يكون عددها فرديا في قسمة عدد الدرجات على 2 لتصنيفها، ثم يقرب الناتج إلى أقرب عدد صحيح لمعرفة ترتيب الوسيط، ثم يبحث عن الدرجة التي تقابل هذا الترتيب.

وبما أننا في هذه الحالة نقرب الناتج دائما إلى أقرب عدد صحيح، إذن ففي مقدورنا أن نستغني عن هذا التقريب بإضافة واحد صحيح إلى عدد الدرجات حتى يصبح زوجيا. ويصبح الناتج بذلك عددا صحيحا.

أي أن ترتيب الوسيط = $\frac{\text{عدد الدرجات} + 1}{2} = \frac{\text{ن} + 1}{2}$

حيث يدل الرمز "ن" على عدد الدرجات، بحيث يكون هذا العدد فرديا.

وعندما نحسب الوسيط للدرجات الآتية:

2 ، 5 ، 7 ، 8، 9 ، 10 ، 17 .

نتبع الخطوات الآتية:

أ) عدد الدرجات = 7

ب) ترتيب الوسيط = $\frac{7+1}{2} = 4$

ج) إذن الدرجة الوسطي لتدريج هذه الدرجات هي 8 .

- **حساب الوسيط عندما يكون عدد الدرجات زوجيا:**

عندما نحسب الوسيط للدرجات الآتية:

16، 13، 2 ، 9 ، 10 ، 11 .

نرتبها أولا ترتيبا تصاعديا: 2، 9، 10، 11، 13، 16. ثم نقسم عدد الدرجات الذي يساوي في مثالنا هذا 6 على 2 أي $\frac{6}{2} = 3$ لنعرف بذلك ترتيب الوسيط.

فإذا بدأنا نحسب ترتيب الدرجات من الطرف الأول لتدريج الدرجات أي من 2 لنصل إلى الدرجة التي ترتيبها الثالث فإننا نرى أن هذه الدرجة هي 10، وأما إذا بدأنا من الطرف الثاني، أي من 16، لنصل إلى الدرجة التي ترتيبها الثالث نرى أن هذه الدرجة هي 11 .

وهكذا نرى أن الوسيط يقع بين 10 ، 11 أي 10.5 وهذا يساوى متوسط الدرجتين: 10، 11.

أي $\frac{10+11}{2} = \frac{21}{2} = 10.5$

وهكذا تتلخص خطوات الوسيط لتلك الدرجات في:

1 – عدد الدرجات = 6

2 – ترتيب الوسيط = $\frac{ن}{2} = \frac{6}{2}$ = 3

3 – الدرجة التي ترتيبها الثالث من الطرف الأول لتدريج الدرجات هي: 10

4 – الدرجة التي ترتيبها الثالث من الطرف الثاني لتدريج الدرجات هي: 11

الوسيط = $\frac{11+10}{2}$ = 10.5

وبالطريقة نفسها يمكن حساب الوسيط للدرجات الآتية:

12، 13 ، 15 ، 20 ، 24 ، 25 ، 27 ، 30 .

وذلك بمعرفة ترتيب الوسيط $\frac{ن}{2} = \frac{8}{2}$ = 4

الوسيط = $\frac{24+20}{2}$ = 22

ملاحظة: في حالة كون البيانات غير مستمرة لن نتمكن من تحديد الوسيط تحديدا عمليا بل بالتقريب نحو الأعلى أو الأسفل. فيكون في المثال الأول الفرد 10 أو 11 لأن الفرد 10.5 لا وجود له.

ثانيا) حساب الوسيط من تكرار الدرجات الخام:

لحساب الوسيط من تكرار الدرجات الخام نحسب مجموع التكرارات والحدود الحقيقية للدرجات. فإذا كانت لدينا البيانات الموجودة في الجدول (20) وأردنا حساب وسيطها، نتبع الخطوات الآتية:

الجدول (20) حساب الوسيط من تكرار الدرجات الخام

الدرجة	التكرار
12	4
13	3
14	1
15	2
المجموع	10

1 – نحدد رتبة الوسيط بقسمة مجموع التكرارات على 2 :

– بما أن عدد الدرجات (مجموع التكرارات) = 10

– إذن فترتيب الوسيط = $\frac{10}{2}$ = 5

2– نحدد مكان رتبة الوسيط ضمن التكرارات ومقدار امتدادها، وذلك إما ابتداء من الطرف الأول للتوزيع (12) أو من طرفه الثاني (15):

1.2– من الطرف الأول للتوزيع، أي من الدرجة 12:

بما أن الدرجة الأولى في التوزيع 12 وتكرارها 4 إذن فالوسيط يليها ولا يقع في نطاق هذه الدرجة لأن ترتيبه 5 . وهذا يزيد على تكرار الدرجة الأولى الذي يساوي 4 بواحد صحيح، إذن فامتداد الوسيط في الدرجة الثانية يساوي الثلث الأول من نطاقها ($\frac{1}{3}$) لأن تكرار الدرجة الثانية "3"، والوسيط يمتد درجة واحدة من الطرف العلوي لهذه الثلاثة أي نطاقها.

– نحدد الحدود الحقيقية لدرجة الوسيط ومداها:

بما أننا نستطيع أن نعلم الحدود الحقيقية للدرجة 13 أي أن نعلم تماما حدها الحقيقي الأول، لذلك يسهل علينا حساب الوسيط. وحدود هذه الدرجة هي 12.5 – 13.5 كما سبق أن بينا ذلك في تحليلنا للحدود الحقيقية للفئات. وقد عاملنا هنا هذه الدرجة أي 13 على أنها فئة مداها واحد صحيح (أي أن للدرجة الواحدة حدين: ±0.5).

– نحسب الوسيط بإضافة الحد الحقيقي الأول لدرجة الوسيط إلى مقدار امتداده فيها:

أي أن الوسيط = $12.5 + (\frac{1}{3}) = 12.5 + 0.33 = 12.83$

إذن، الوسيط = الحد الأدنى الحقيقي لدرجة الوسيط + مقدار امتداده في هذه الدرجة.

2.2– من الطرف الأخير للتوزيع أي من الدرجة 15 (كمراجعة لنتيجة الطريقة السابقة)، ونتبع الخطوات نفسها:

– بما أن ترتيب الوسيط = 5 ، وبما أن تكرار الدرجة الأخيرة 15 هو 2، وتكرار الدرجة التي تسبقها هو 1، فالتكرار المتجمع حتى الدرجة 14 هو 3، وهذا ينقص 2 عن ترتيب الوسيط. إذن فالوسيط يقع في ثلثي ($\frac{2}{3}$) تكرار الدرجة السابقة لها وهي 13 وتكرارها 3 .

– وبما أن الحد الحقيقي الأعلى للدرجة 13 هو 13.5، وترتيب الوسيط ينقص عن هذا الحد بقيمة عددية مقدارها الثلثين ($\frac{2}{3}$).

إذن فحساب الوسيط يتم بطرح هذا المقدار من الحد الحقيقي الأعلى لدرجة الوسيط:

أي أن الوسيط = الحد الأعلى الحقيقي لدرجة الوسيط – مقدار امتداده في هذه الدرجة.

الوسيط = $13.5 - \frac{2}{3} = 13.5 - 0.67 = 12.83$

وهذه هي النتيجة نفسها التي حصلنا عليها بالطريقة الأولى.

ثالثا) حساب الوسيط من فئات الدرجات:

لحساب الوسيط من فئات الدرجات نحسب التكرار المتجمع الصاعد، والتكرار المتجمع النازل والحدود الحقيقية لفئات الدرجات.

وسنبين أولا طريقة حساب الوسيط من التكرار المتجمع الصاعد ثم من النازل:

الجدول رقم (21) يبين الفئات وحدودها الحقيقية وتكرارها المتجمع الصاعد والنازل

الفئات	الحدود الحقيقية	التكرار	ك. المتجمع الصاعد	ك. المتجمع النازل
18 − 17	18.5 – 16.5	1	1	37
20 – 19	20.5 – 18.5	5	6	36
22 – 21	22.5 − 20.5	8	14	31
24 – 23	24.5 − 22.5	8	22	23
26 – 25	26.5 – 24.5	5	27	15
28 – 27	28.5 − 26.5	6	33	10
30 – 29	30.5 − 28.5	0	33	4
32 – 31	32.5 – 30.5	1	34	4
34 – 33	34.5 − 32.5	0	34	3
36 – 35	36.5 − 34.5	2	36	3
38 − 37	38.5 − 36.5	1	37	1
		مج=37		

أ- حساب الوسيط من التكرار المتجمع الصاعد

لحساب الوسيط من التكرار المتجمع الصاعد تتبع الخطوات الآتية:

1- نحدد رتبة الوسيط: بما أن مجموع التكرارات= 37 ، فرتبة الوسيط = $\frac{37}{2} = 18.5$

2- نحدد مكان رتبة الوسيط وقيمة امتدادها ، وذلك بالبحث عن قيمة رتبته في التكرار المتجمع الصاعد أو القيمة التي تكبرها مباشرة ، فنجد أنها تقع في الفئة التي تمتد أطرافها من 23 إلى 24 لأن التكرار المتجمع الصاعد للفئة التي تسبقه يساوي 14، أي أقل منها. فهو يمتد في الفئة 23 – 24 بقيمة مقدارها زيادة فرق ترتيب الوسيط عن التكرار المتجمع للفئة السابقة التي تمتد من 21 إلى 22. وهذا الفرق = 18.5 − 14 = 4.5

3- نحدد مقدار ونسبة امتداد الوسيط في هذه الفئة: بما أن تكرار الفئة التي يقع فيها الوسيط يساوي 8 ، إذن فنسبة امتداد الوسيط لهذا التكرار تساوي $\frac{4.5}{8}=0.56$. لكن مدى هذه الفئة يساوي 2، إذن فمقدار هذا الامتداد يساوي 0.56 × 2 = 1.12

4- نحدد الحد الأول أو الأدنى (في حالة البيانات المتصلة) أو الحد الأدنى الحقيقي (في حالة البيانات المنفصلة) لفئة الوسيط وهو هنا يساوي 22.5

5- نحسب الوسيط بإضافة مقدار الامتداد إلى هذا الحد، أي أن الوسيط (ط) = 22.5 + 1.12 = 23.62 = 23.6 بالتقريب.

ويمكن أن نلخص هذه الخطوات في المعادلة الآتية:

ط=الحد الأدنى لفئة الوسيط+نصف عدد الدرجات–ك.م.ص. للفئة قبل الوسيطية× مدى الفئة

تكرار فئة الوسيط

أي أن: الوسيط $= ل1 + \frac{(ن/2 - ك1)}{ك} \times ف \;=\; Me = L1 + \frac{(N/2 - \sum f1)}{fMe} . C$

حيث أن: ل$_1$ = الحد الأدنى لفئة الوسيط = L

ن = عدد الدرجات = N = مج ك = $\sum f$

ك$_1$ = التكرار المتجمع للفئة قبل الوسيطية = $\sum f_1$

ك = تكرار فئة الوسيط = f Me

ف = مدى فئة الوسيط = C

وبتطبيق هذه المعادلة نحصل على:

ل$_1$ = 22.5 ن = 37 ك$_1$ = 14 ك = 8 ف = 2

أي أن الوسيط= $22.5 + \frac{4.5}{8} \times 2 = 22.5 + \frac{37 \div 2 - 14}{8} \times 2 = 22.5 + 1.12 = 23.62$

ب- حساب الوسيط من التكرار المتجمع النازل : لحساب الوسيط من التكرار المتجمع النازل نتبع الخطوات الآتية:

1- نحدد رتبة الوسيط: بما أن مجموع التكرارات= 37 ، فرتبة الوسيط = $\frac{37}{2} = 18.5$

2- نحدد فئة الوسيط وهي: 23 - 24.

3− نحدد الفئة التي تقع قبل فئة الوسيط (من أسفل إلى أعلى) وهي 25 – 26 وتكرارها المتجمع 15 .

4− زيادة رتبة الوسيط عن التكرار المتجمع للفئة التي تلي فئته = 18.5 – 15 = 3.5

5− تكرار فئة الوسيط يساوي 8

6− نسبة امتداد الوسيط في هذا التكرار = 3.5 ÷ 8 = 0.44

7− مدى فئة الوسيط=2، مما يجعل مقدار الامتداد= 0.44×2=0.88

8− الحد الحقيقي (لأن البيانات منفصلة) الأخير أو الأعلى لهذه لفئة الوسيط = 24.5

9− نحسب الوسيط بطرح مقدار الامتداد من الحد الحقيقي الأعلى لفئة الوسيط ، أي أنه يساوي 24.5 − 0.88 = 23.62

وهذه هي النتيجة نفسها التي حصلنا عليها بالطريقة السابقة التي اعتمدت على التكرار المتجمع الصاعد. ويمكن أن نلخص هذه الخطوات في المعادلة الآتية:

ط=الحد الأعلى لفئة الوسيط− نصف عدد الدرجات−ك.م.ن. للفئة بعد الوسيطية×مدى الفئة

تكرار فئة الوسيط

أي أن:

$$\text{الوسيط} = \text{ل}2 - \frac{(\text{ن}/2 - \text{ك}2)}{\text{ك}} \times \text{ف} = Me = L2 - \frac{(N/2 - \sum f2)}{fMe} \,.\, C$$

حيث ل$_2$ = الحد الأعلى لفئة الوسيط = L2

ن = عدد الدرجات = N = مج ك = $\sum f$

ك$_2$= التكرار المتجمع للفئة الموالية لفئة الوسيط = $\sum f2$

ك = تكرار فئة الوسيط = f Me

ف = مدى فئة الوسيط = C

وبتطبيق هذه المعادلة نحصل على:

ل$_2$ = 24.5 ن = 37 ك$_2$= 15 ك = 8 ف = 2

أي أن :

$$\text{الوسيط} = 24.5 - \frac{37 \div 2 - 15}{8} \times 2 = 24.5 - \frac{3.5}{8} \times 2 = 24.5 - 0.88 = 23.62$$

ج- حساب الوسيط الذي يقع ترتيبه على حدود الفئات : في بعض الحالات يصعب على الباحث حساب الوسيط بالطريقتين السابقتين وذلك عندما يقع ترتيبه على الحد الحقيقي القائم بين فئتين متتاليتين. وحينها يكون الوسيط مساويا للحد الحقيقي الأعلى أو الأدنى للفئة المعينة:

الجدول (22 أ) يوضح حساب الوسيط الذي يقع ترتيبه على حدود الفئات

الفئات	الحدود الحقيقية	التكرار	ك. المتجمع الصاعد	ك. المتجمع النازل
24 – 20	24.5 – 19.5	2	2	68
29 – 25	29.5 – 24.5	7	9	66
34 – 30	34.5 – 29.5	10	19	59
39 – 35	39.5 – 34.5	15	<u>34</u>	49
44 – 40	44.5 – 39.5	18	52	<u>34</u>
49 – 45	49.5 – 44.5	8	60	16
54 – 50	54.5 – 49.5	3	63	48
59 – 55	59.5 – 54.5	5	68	5
		مج=68		

لحساب الوسيط في هذه الحالة نتبع الخطوات الآتية:

1- ترتيب الوسيط = $\frac{68}{2}$ = 34

2-التكرار المتجمع الصاعد يدل على أن الوسيط يقع في الفئة 35-39

3- وبما أن التكرار المتجمع لهذه الفئة يساوي ترتيب الوسيط: 34

إذن فالوسيط يساوي الحد الأعلى الحقيقي لهذه الفئة أي 39.5

وإذا حسبنا الوسيط من التكرار المتجمع النازل نجد أن:

1-التكرار المتجمع النازل يدل على أن الوسيط يقع في الفئة 40-44

2- وبما أن التكرار المتجمع لهذه الفئة يساوي ترتيب الوسيط: 34

إذن فالوسيط يساوي الحد الأدنى الحقيقي لهذه الفئة أي 39.5

وهكذا نرى أن الوسيط في الحالتين يساوي 39.5 أي أن عملية حسابه صحيحة.

الجدول (22 ب) يوضح أيضا حساب الوسيط الذي يقع ترتيبه على حدود الفئات

الفئات	التكرارات	ك م ص	ك م ن
4 – 5	2	2	12
6 – 7	4	<u>6</u>	10
8 – 9	3	9	<u>6</u>
10–11	3	12	3
المجموع	12		

لحساب الوسيط في هذه الحالة نتبع الخطوات الآتية:

1- ترتيب الوسيط = $\frac{12}{2}$ = 6

2- التكرار المتجمع الصاعد يدل على أن الوسيط يقع في الفئة 6-7.

3- وبما أن التكرار المتجمع لهذه الفئة يساوي ترتيب الوسيط: 6.

إذن فالوسيط يساوي الحد الأعلى الحقيقي لفئة الوسيط (7-8) أي 7.5

وإذا حسبنا الوسيط من التكرار المتجمع النازل نجد أن :

1- التكرار المتجمع النازل يدل على أن الوسيط يقع في الفئة 8-9

2- وبما أن التكرار المتجمع لهذه الفئة يساوي ترتيب الوسيط: 6

3-إذن فالوسيط يساوي الحد الأدنى الحقيقي لفئة الوسيط (8-9) أي 7.5

وهكذا نرى أن الوسيط في الحالتين يساوي 7.5 .

د- حساب الوسيط الذي يقع في فئة لا تكرار لها : عندما يقع الوسيط في فئة تكرارها يساوي صفرا، فإننا نجد صعوبة في الاستعانة بالطرق السابقة لحساب الوسيط . ولذلك نعتمد في تحديده على منتصف الفئة الصفرية بعد رصد المتجمعين الصاعد والنازل.

الجدول (23) يوضح حساب الوسيط الذي يقع في فئة تكرارها يساوي صفرا

الفئات	الحدود الحقيقية	التكرار	ك. المتجمع الصاعد	ك. المتجمع النازل
7 – 5	7.5 – 4.5	1	1	34
10 – 8	10.5 – 7.5	7	8	33
13 – 11	13.5 – 10.5	9	17	26
16 – 14	16.5 – 13.5	0	17	17
19 – 17	19.5 – 16.5	6	23	17
22 – 20	22.5 – 19.5	7	30	11
25 – 23	25.5 – 22.5	2	32	4
28 – 26	28.5 – 25.5	2	34	2
		مج=34		

لحساب الوسيط في هذه الحالة نتبع الخطوات الآتية:

1- ترتيب الوسيط = $\frac{34}{2} = 17$

2-وبما أن التكرار المتجمع الصاعد يصل إلى 17 عند الفئة 11-13 ثم يظل كما هو في الفئة الموالية لأن تكرارها يساوي صفرا.

إذن فالوسيط يقع في نهاية الفئة التي تمتد من 11 إلى 13 أي عند 13.5

3- وبما أن التكرار المتجمع النازل يصل في تطوره من أسفل إلى أعلى إلى 17 عند الفئة 17 - 19 ثم يظل ثابتا في الفئة الموالية لأن تكرارها يساوي صفرا.

إذن فالوسيط يقع في بداية الفئة التي تمتد حدودها من 17 إلى 19 أي عند 16.5 .

4- أي أن ترتيب الوسيط بهذا المعنى يقع في الفئة التي تمتد بين 13.5 و 16.5 والتي يساوي تكرارها صفرا.

5- إذن فمنتصف هذه الفئة يدل على ترتيب الوسيط.

أي أن الوسيط = $\frac{16.5+13.5}{2} = \frac{30}{2} = 15$

- الخواص الإحصائية للوسيط

1- عند مقارنة مجموع الانحرافات المطلقة بالنسبة للمتوسط والوسيط نجد أن مجموع الانحرافات المطلقة عن الوسيط أصغر من مجموع الانحرافات المطلقة عن المتوسط، أي أن

الوسيط يتوسط توزيع الدرجات أكثر مما يتوسطها المتوسط، ولذا فإن الوسيط في أي توزيع تكراري عادي يقع بين المتوسط والمنوال.

2- يتأثر الوسيط بالدرجات الوسطى أكثر مما يتأثر بالدرجات المتطرفة في التوزيع التكراري، لأنه تابع لجميع وحدات المجموعة التي ينتمي إليها بالترتيب وليس بالدرجات. وذلك على عكس المتوسط الذي يتأثر بالدرجات المتطرفة أكثر. ولذا يصلح الوسيط كمقياس للنزعة المركزية أكثر من المتوسط عندما تكون أطراف التوزيع متراكمة، متجمعة وغير مستوية. كأن يلتوي التوزيع التكراري فتكثر فيه الأصفار والأعداد الصغيرة التي تقوم عند طرفه الأول أو تكثر فيه الأعداد الكبيرة التي تقوم عند طرفه الثاني.

ولتوضيح هذه الخاصية نحسب الوسيط والمتوسط من الدرجات الآتية:

25 20 15 10 5

فنجد أن الوسيط= 15 (القيمة الوسطى: الثالثة) والمتوسط= 15 (75÷5)

ثم نعلو بالطرف الأخير علوا كبيرا فنحول الـ 25 إلى 50 ثم نحسب بعد ذلك الوسيط والمتوسط للدرجات في صورتها الجديدة:

50 20 15 10 5

فنجد أن الوسيط = 15 والمتوسط = 20 (100÷5)

وهكذا نرى أن الوسيط لم يتغير في كلتا الحالتين، أي أنه لم يتأثر بما حدث في الطرف الأخير من تغير. وأن المتوسط تغير من 15 إلى 20 نتيجة لتغير الطرف الأخير للدرجات السابقة.

فالوسيط بهذا المعنى أكثر ثبوتا واستقرارا من المتوسط بالنسبة للأطراف، أو أن المتوسط أكثر حساسية من الوسيط بالنسبة لأطراف التوزيع.

أما عندما نغير الدرجة أو الدرجات الوسطى فإننا بذلك نغير قيمة الوسيط تغييرا كبيرا، ولا يكاد يصيب المتوسط من هذا التغيير إلا اختلافا بسيطا. ولنوضح هذه الفكرة بتغيير الدرجة الوسطى في المثال السابق من 15 إلى 10 فتصبح:

25 20 10 10 5

ونجد أن الوسيط = 10

والمتوسط = 14 (70÷5)

وإذا غيرنا الدرجة الوسطى من 10 إلى 20 فإننا نرى تغير الوسيط أكثر من تغير المتوسط، كما يبدو ذلك في المثال الموالي:

25 20 20 10 5

الوسيط = 20

والمتوسط = 16 (80÷5)

وهكذا نرى أن المتوسط أكثر تأثرا من الوسيط بالدرجات المتطرفة وأن الوسيط أكثر تأثرا من المتوسط بالدرجات الوسطى.

إن هذه الخاصية تحدد الأهمية النسبية لكل من المتوسط والوسيط، والميادين والمجالات التي يستخدم فيها كل منهما.

فوائد الوسيط : يصلح الوسيط للميادين نفسها التي يصلح لها المتوسط، أي للمعايير والمقارنة وخاصة عندما يكون التوزيع التكراري للدرجات ملتويا أي مرتفعا من أحد طرفيه (نحو الطرف الأول= التواء موجب، ونحو الطرف الثاني= التواء سالب)، وذلك بخلاف التوزيع الاعتدالي الذي يفضل فيه اعتماد المتوسط لقياس النزعة المركزية. والجداول: 24–26 توضح التوزيعات الثلاثة. مع ملاحظة أن مقاييس النزعة المركزية الثلاثة (متوسط، وسيط، منوال) تكون متساوية في حالة التوزيع الاعتدالي.

جدول(24): توزيع تكراري ملتو التواء موجبا

الدرجة	التكرار
2	7
3	13
4	20
5	10
6	9
7	4
8	1
المجموع	64

جدول(25): توزيع تكراري ملتو التواء سالبا

الدرجة	التكرار
2	1
3	4
4	9
5	10
6	20
7	13
8	7
المجموع	64

جدول(26):توزيع تكراري اعتدالي

الدرجة	التكرار
2	1
3	6
4	15
5	20
6	15
7	6
8	1
المجموع	64

3.1.3- المنوال (Mode)

يدل المنوال على أكثر الدرجات شيوعا، أو بمعنى أدق هو النقطة التي تدل على أكثر درجات التوزيع تكرارا.

– حساب المنوال من تكرار الدرجات

يمكن معرفة المنوال بسهولة عندما نقارن تكرار الدرجات لنبحث عن أكبرها.

الجدول (27) حساب المنوال من تكرار الدرجات

الدرجة	التكرار
12	3
13	7
14	10
15	8
16	6
17	2
المجموع	26

وهكذا نرى أن المنوال تعبر عنه أكبر الدرجات تكرارا وهي الدرجة 14 لأن تكرارها يساوي 10 وهذه العشرة هي أكبر تكرارات هذا الجدول.

– حساب المنوال من فئات الدرجات

لحساب المنوال من فئات الدرجات نبحث أيضا عن أكبر تكرار ثم نحدد الفئة التي تقابله. وبما أن الفئات تمتد إلى أكثر من درجة فهي لا تدل على نقطة المنوال دلالة دقيقة، ولذلك نستعين بمنتصف الفئة للدلالة على منوال التوزيع. والجدول الموالي يوضح ذلك:

جدول (28) حساب المنوال من فئات الدرجات

الفئات	منتصف الفئات	التكرار
13 – 11	12	11
16 – 14	15	3
19 – 17	18	9
22 – 20	21	13
25 – 23	24	11
28 – 26	27	3
المجموع		50

وهكذا نرى أن أكبر تكرار بهذا التوزيع هو 13 وهو تكرار الفئة التي تمتد حدودها من 20 إلى 22 وبما أن منتصف هذه الفئة يساوي 21 إذن فالدرجة التي تدل على المنوال هي 21 .

– حساب المنوال من الوسيط والمتوسط

تواجه الباحث أحيانا صعوبات شتى في حساب المنوال. وخاصة عندما يكثر عدد الفئات التي تحتوي على أكبر تكرار، كأن يدل الجدول السابق على فئة أخرى تكرارها 13 مثل تكرار الفئة 20–22 التي دل منتصفها المساوي لـ21 على المنوال.

والطريقة الإحصائية لحساب المنوال تعتمد على الوسيط والمتوسط، والمعادلة الآتية توضح علاقة هذه المقاييس الثلاثة.

المنوال = ثلاثة أمثال الوسيط – ضعف المتوسط

أي أن المنوال = (3 × الوسيط) – (2 × المتوسط)

و= 3 ط – 2 $\overline{\text{س}}$

حيث يدل الرمز "و" على المنوال

و الرمز "ط" على الوسيط

و الرمز "$\overline{\text{س}}$" على المتوسط

وعندما نستخدم هذه المعادلة في حساب المنوال للجدول السابق، علينا أن نستخرج أولا المتوسط والوسيط بالطريقة التالية التي يبينها الجدول (29).

جدول (29) حساب المنوال من الوسيط والمتوسط

الفئات	الحدود الحقيقية للفئات	منتصفات الفئات	ك	ص×ك	المتجمع الصاعد
11– 13	10.5 – 13.5	12	11	132	11
14– 16	13.5 – 16.5	15	3	45	14
17– 19	16.5 – 19.5	18	9	162	23
20– 22	**19.5 – 22.5**	**21**	**13**	273	36
23– 25	22.5 – 25.5	24	11	264	47
26– 28	25.5 – 28.5	27	3	81	50
المجموع			50	957	

$$\text{المتوسط} = \frac{\text{منتصفات الفئات} \times \text{تكراراتها}}{\text{مجموع التكرارات}} = \frac{957}{50} = 19.14$$

المتوسط = <u>19.14</u>

$$\text{الوسيط} = ل + \frac{(ق/2 - ت\ ق)}{ت} \times ف \quad = 19.5 + \frac{50/2 - 23}{13} \times 3$$

$$= 19.5 + \frac{25 - 23}{13} \times 3 = 19.5 + \frac{2}{13} \times 3 = 19.5 + (0.15 \times 3) = 19.95$$

الوسيط = <u>19.95</u>

<u>المنوال</u> = 3 ط – 2 $\overline{س}$ = 3 × 19.95 – 2 × 19.14 = 59.85 – 38.28 = <u>21.57</u>

– **حساب المنوال من تكرار الفئات المتجاورة** (طريقة كارل بيرسون)

يمكن حساب المنوال بطريقة أكثر دقة عن طريق معادلة "بيرسون" الرياضية، حيث يستعان بتكرار الفئة المنوالية وبتكرار الفئة السابقة لها والموالية لها أيضا. وتقوم هذه الفكرة على الاستفادة من الارتفاع التكراري الذي يسبق الفئة المنوالية ويؤدي إليها، والانخفاض التكراري الذي يعقبها ويتأثر بها، أي الإفادة من الفروق بين التكرارات الثلاثة.

فلو لاحظنا تكرار الفئة 17–19 التي تسبق الفئة المنوالية (20–22) لوجدناه مساويا 9، ويرتفع هذا التكرار في الفئة المنوالية (20–22) ليصل إلى 13. ولو لاحظنا تكرار الفئة (23–25)، التي تأتي بعد الفئة المنوالية، لوجدنا أنه يساوي 11 وهذا يمثل انخفاضا في التكرار بعدما ارتفع في الفئة المنوالية.

وتتلخص طريقة حساب المنوال في الخطوات الموالية:

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{تكرار الفئة المنوالية} - \text{تكرار الفئة السابقة للفئة المنوالية}}{(\text{ك. ف. م.} - \text{ك. ف. السابقة}) + (\text{ك. ف. م.} - \text{ك. ف. اللاحقة})} \times \text{مدى الفئة}$$

$$\text{المنوال} = ل + \frac{(ك\ م - ك\ ق)}{(ك\ م - ك\ ق) + (ك\ م - ك\ ب)} \times ف = \quad Mo = L + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2}\right) C \times$$

حيث: ل = الحد الأدنى للفئة المنوالية = L

ك م (تكرار الفئة المنوالية) – ك ق (تكرار الفئة قبل المنوالية)= Δ_1

ك م (تكرار الفئة المنوالية) – ك ب (تكرار الفئة بعد المنوالية)= Δ_2

ف = مدى الفئة = C

وهكذا يمكن أن نحسب المنوال للتوزيع التكراري للجدول (22) بالطريقة الموالية:

ل = 19.5 ك م = 13 ك ق = 9 ك ب = 11 ف = 3

إذن المنوال = $19.5 + \frac{13-9}{(13-11)+(13-9)} \times 3 = 19.5 + (\frac{4}{4+2} \times 3) =$

$= 19.5 + \frac{4}{6} \times 3 = 19.5 + \frac{12}{6} = 19.5 + 2 = 21.5$

<u>المنوال = 21.5</u>

وهذه هي القيمة نفسها التي حصلنا عليها بالطريقة السابقة التي اعتمدت على الوسيط والمتوسط في حساب المنوال.

ومن أهم مميزات طريقة تكرار الفئات المتجاورة دقتها وعدم اعتمادها على الوسيط والمتوسط. ولهذه الخاصية الأخيرة أهميتها في حساب الالتواء.

وهناك معادلة أخرى لحساب المنوال شبيهة بها، وهي تنص على أنه يساوي:

الحد الأدنى للفئة المنوالية + (مدى الفئة × $\frac{\text{تكرار الفئة بعد المنوالية}}{\text{تكرار الفئة بعد المنوالية + تكرار الفئة قبل المنوالية}}$)

وبناء على المعطيات السابقة نفسها فالمنوال =

$19.5 + 3 \times \frac{11}{11+9} = 19.5 + 3 \times \frac{11}{20} = 19.5 + 3 \times 0.55 = 19.5 + 1.65 = 21.15$

وهي النتيجة نفسها تقريبا.

كما يمكن حسابه أيضا بطريقة الفروق بين التكرارات، كما يلي:

المنوال= الحد الأدنى للفئة المنوالية + مدى الفئة × $\frac{\text{الفرق الأدنى}}{\text{مجموع الفرقين}}$

ومن المثال السابق يصبح المنوال = $19.5 + 3 \times \frac{4}{6} = 19.5 + 3 \times 0.66 = 19.5 + 2 = 21.5$

وهي قيمة منوالية تقريبية شبيهة بسابقاتها. كما يمكن إيجاد المنوال عن طريق الرسوم البيانية (المنحنى أو المدرج التكراري) باستعمال قيم الفئات الثلاث (المنوالية، ما قبلها وما بعدها).

والملاحظ أن قيمة المنوال تقريبية في كل الحالات لأنه أقل المتوسطات دقة، يليه في ذلك الوسيط، فالوسط الحسابي، والذي يعتبر أكثر الثلاثة دقة واستخداما في العلوم الاجتماعية (غريب سيد أحمد:2005، 132).

الخواص الإحصائية للمنوال

1- لا يتأثر المنوال بالدرجات المتطرفة ولا بالدرجات الوسطى في التوزيع التكراري، وإنما يتأثر بالتكرار نفسه عندما يبلغ نهايته العظمى بالنسبة لدرجة ما أو لفئة ما من الدرجات (لأنه هو الذي يحدد قيمة المنوال). فهو من هذه الناحية أكثر ثباتا واستقرارا من المتوسط والوسيط.

2- يتأثر المنوال بعدد فئات التوزيع وبمدى الفئة ، فكلما قل هذا العدد زاد تبعا لذلك مدى الفئات وارتفع تكرار الفئة المنوالية. وكلما كثر هذا العدد بالنسبة لنفس التوزيع السابق قل تبعا لذلك مدى الفئات وانخفض تكرار الفئة المنوالية.

3-عندما تتعدد قمم التوزيع التكراري (أكبر التكرارات) تتعدد أيضا قيم المنوال. فإذا كان للتوزيع قمتان كان لكل قمة منهما منوال. والمثال المبين بالجدول (30) يوضح ذلك:

جدول (30) يعكس توزيعا تكراريا ذي قمّتين

الدرجة	1	2	3	4	5	6	7	8	9	10	المجموع
التكرار	2	4	9	4	5	3	9	6	2	6	50

ويبلغ التكرار في هذا التوزيع نهايته العظمى 9 عند الدرجة 3 ثم يعود ليصل إلى هذه النهاية ثانية عند الدرجة 7 . أي أن له منوالا عند الدرجة 3 ومنوالا آخر عند الدرجة 7 . وقد لا يوجد المنوال أصلا في حالة كون التكرارات متساوية القيمة، أي كلها 5 في مثالنا السابق.

فوائد المنوال

يصلح المنوال للميادين نفسها التي صلح لها المتوسط والوسيط، أي في المعايير والمقارنة.

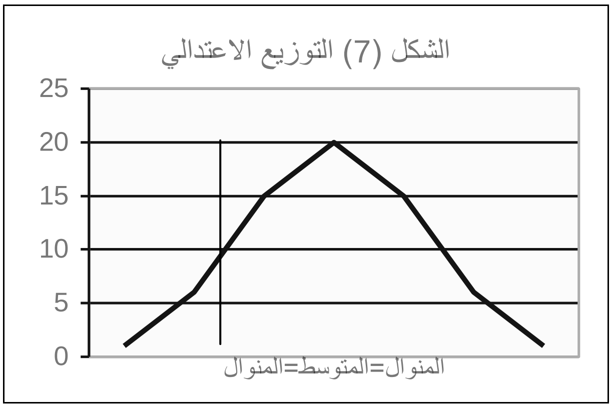
وله أهميته في النواحي التربوية والنفسية وخاصة عندما يراد معرفة العمر المنوالي لمراحل التعليم المختلفة. فمثلا العمر المنوالي لتلاميذ السنة الأولى الابتدائية هو 6 سنوات ونسبة الذكاء المنوالية هي مثلا 100 .

كما تصلح دلالته على الدرجة الأكثر شيوعا لمعالجة المشاكل التي تهدف إلى معرفة تركيز الظاهرة وموقعها، وخاصة في النواحي الصناعية والتجارية. فتاجر المجلات والكتب يعتمد في رواج بضاعته على المقاييس الأكثر شيوعا أو على المقاييس المنوالية.

وبما أن عملية حساب المنوال سهلة وسريعة، لذلك يمكن أحيانا تقدير قيمة المنوال بمجرد النظر لشكل التوزيع التكراري، ومن ثم تقدير النزعة المركزية تقديرا مبدئيا.

4.1.2- العلاقة بين مقاييس النزعة المركزية

1- تنطبق جميع مقاييس النزعة المركزية على بعضها وتتساوى جميعها في التوزيع التكراري الاعتدالي، كما حصل ذلك في الجدول (26) وشكله الموالي:



حيث نرى أن:

المتوسط = (مجموع نواتج ضرب كل قيمة في تكرارها على مجموع التكرارات) = 320 ÷ 64 = 5

الوسيط = (من الطرف الأول للتوزيع)

= الحد الأدنى الحقيقي لدرجة الوسيط + مقدار امتداده في هذه الدرجة

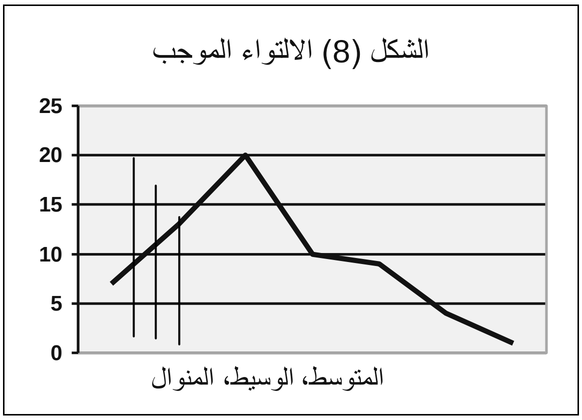
بما أن درجة الوسيط= مجموع التكرارات على 2، أي= $\frac{64}{2}$ = 32،

فالوسيط = 4.5 + $\frac{10}{20}$ $= 4.5 + 0.5$ = 5

المنوال = (القيمة التي تقابلها أكبر التكرارات) = 5

2- عندما يكون التوزيع التكراري ملتويا التواء موجبا (بسيطا) يمتد الطرف الطويل للمنحنى البياني إلى الجهة اليمنى ويصبح ترتيب مقاييس النزعة المركزية كما يأتي: 1-المتوسط ، 2- الوسيط، 3-المنوال (ومقدارها في المثال الموالي= 4.26 ، 4.1 ، 4)

وهو ما يؤكده حساب جميع مقاييس النزعة المركزية للتوزيع التكراري الموجب الالتواء من الجدول (24) السابق وشكله الموالي:



حيث أن:

المتوسط= $\frac{273}{64} = 4.26$

الوسيط (من الطرف الأول للتوزيع) = الحد الأدنى الحقيقي لدرجة الوسيط + مقدار امتداده في هذه الدرجة

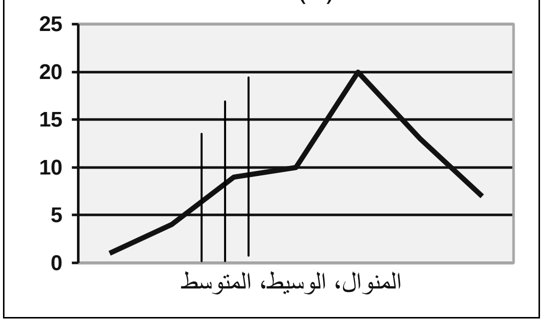
بما أن درجة الوسيط= مجموع التكرارات على 2، أي = $\frac{64}{2}$ = 32 ، فالوسيط =

$$3.5 + \frac{12}{20} = 3.5 + 0.6 = 4.1$$

المنوال= (الدرجة التي تقابلها أكبر التكرارات) = 4

3- عندما يكون التوزيع التكراري ملتويا التواء سالبا (بسيطا) يمتد الطرف الطويل للمنحنى البياني إلى الجهة اليسرى ويصبح ترتيب مقاييس النزعة المركزية كما يأتي: 1-المنوال ، 2- الوسيط، 3-المتوسط (ومقدارها في المثال الموالي= 6 ، 5.9 ، 5.73)

وهو ما يؤكده حساب جميع مقاييس النزعة المركزية للتوزيع التكراري السالب الالتواء من الجدول (25) السابق وشكله الموالي:



الشكل (9) الالتواء السالب

حيث أن:

المنوال (الدرجة التي تقابلها أكبر التكرارات) = 6

الوسيط (من الطرف الأول للتوزيع) = $5.5 + \frac{8}{20} = 5.5 + 0.4 = 5.9$

المتوسط = (نواتج ضرب كل قيمة في تكرارها على مجموع التكرارات) = $\frac{367}{64} = 5.73$

– وبالطبع يمكن اتخاذ هذه الخصائص كمقياس للالتواء، فإذا كان المتوسط والمنوال متساويين كان التوزيع متماثلا، بينما تزيد شدة الالتواء بازدياد الفرق بينهما، والعكس بالعكس.

ويمكن قياس الالتواء في حالة عدم تطابق مقاييس النزعة المركزية بطريقة "بيرسون" التي تعتمد على المتوسط والمنوال والانحراف المعياري تبعا للمعادلة الموالية:

$$\text{الالتواء} = \frac{\text{المتوسط} - \text{المنوال}}{\text{الانحراف المعياري}}$$

وتتراوح قيمته بين –3 في الالتواء السالب و + 3 في الالتواء الموجب. و هو يعتبر من مقاييس التشتت كما سنرى لاحقا.

– ونذكّر في الأخير أنه يمكن حساب مقاييس النزعة المركزية من بعضها البعض بحكم العلاقة الخطية (التقريبية العامة) التي تربطها، فلقد مر معنا أنه يمكن إيجاد المنوال من الوسيط والمتوسط، وكذا المتوسط من المنوال والوسيط. فإذا كان:

المنوال = 3 ط – 2 $\bar{\text{س}}$ ، ومنه:

المنوال= المتوسط الحسابي – 3 (المتوسط الحسابي – الوسيط).

فإن: $\bar{\text{س}}$ – المنوال= 3 ($\bar{\text{س}}$ –الوسيط) أو $\bar{\text{س}} = \frac{3\ \text{وسيط} - \text{المنوال}}{2}$

كما يمكن حساب الوسيط من المتوسط والمنوال باستنتاجه من المعادلة الأولى الخاصة بالمنوال:

المنوال = 3 ط − 2 $\overline{س}$ ، ومنه:

$$\text{الوسيط} = \frac{\text{المنوال} + 2\overline{س}}{3}$$

− من **المثال** الأخير العاكس للالتواء الموجب، يمكن اعتماد هذه المعادلة في حساب المنوال والمتوسط والوسيط للوصول إلى النتائج نفسها تقريبا:

$$\text{المنوال} = 3(4.1) - 2(4.26) = 12.1 - 8.52 = 3.58 \ (\approx 4)$$

$$\text{المتوسط} = \frac{12.30-4}{2} = \frac{8.30}{2} = 4.15 \ (\approx 4{,}26)$$

$$\text{الوسيط} = \frac{2(4.26)+4}{3} = \frac{12.52}{3} = 4.17 \ (\approx 4{,}1)$$

والأمر نفسه بالنسبة للالتواء السالب:

$$\text{المنوال} = 3(5.9) - 2(5.73) = 17.7 - 11.46 = 6{,}24 \ (\approx 6)$$

$$\text{المتوسط} = \frac{17.7-6}{2} = \frac{11.7}{2} = 5.85 \ (\approx 5{,}73)$$

$$\text{الوسيط} = \frac{2(5.73)+6}{3} = \frac{17.46}{3} = 5.82 \ (\approx 5.9)$$

2.3. مقاييس التشتت (Measures of Dispersion)

قمنا فيما سبق بتلخيص بيانات الظاهرة موضوع الدراسة في صورة قيمة واحدة تدل على نزوع قيم هذه الظاهرة نحو مركز واحد أو قيمة ممثلة لمجموعة القيم، وذلك من خلال مقاييس النزعة المركزية (المتوسط، الوسيط والمنوال). لكن قيم هذه المقاييس لا تكفي لإعطاء صورة كاملة عن مدى تجانس أو اختلاف قيم الظاهرة المدروسة. ولذلك وجب قياس مدى تشتت قيمها، لأنه قد يكون كبيرا إذا كانت قيمها بعيدة عن بعضها البعض وصغيرا إذا كانت قريبة من بعضها البعض. ومعلوم أنه كلما زاد التشتت كلما قلت أهمية مقياس المتوسط كقيمة ممثلة لمجموعة القيم والعكس بالعكس.

وإذا كانت القيمة المركزية أو الممثلة لمجموعة من القيم مهمة فلكي تكون عملية أكثر لا بد أن تكون مصحوبة بقيمة أخرى تعكس مدى تقارب أو تباعد قيم الظاهرة عن بعضها البعض وحول المتوسطات. ولتوضيح ذلك نضرب المثال الآتي:

لدينا مجموعتان متساويتان عددا ومجموعا (05 قيم مجموعها 110)، ووسيطهما ومتوسطهما الحسابيين متساويان أيضا (22)، ولكن درجة تجانسهما أو تشتتهما تختلف اختلافا واضحا:

المجموعة الأولى: 20 21 22 23 24

المجموعة الثانية: 10 17 22 22 39

المتوسط الحسابي للمجموعة الأولى(مجموع القيم على عددها)=110÷5 =22

المتوسط الحسابي للمجموعة الثانية = 110 ÷ 5 = 22

وسيط المجموعة الأولى (القيمة الوسطى في الترتيب) = 22

وسيط المجموعة الثانية = 22

إلا أن الملاحظ أن الاختلاف واضح بين توزيع قيم المجموعتين، فقيم المجموعة الأولى تبدو، حتى بالعين المجردة، متقاربة أكثر من قيم المجموعة الثانية. ولإثبات ذلك إحصائيا تستعمل مقاييس التشتت المطلق والنسبي، ومنها: المدى، المئينيات...، الانحراف الربيعي، الانحراف المتوسط والانحراف المعياري (كمقاييس مطلقة)، ومعامل الاختلاف ومقياس الالتواء لـ"بيرسون" (كمقياسين نسبيين). وفيما يلي عرض مبسط لها:

أولا) المدى (Range)

يعبر المدى عن الفرق بين أكبر قيمة وأصغر قيمة بين قيم الظاهرة المدروسة، أي أن المدى = أكبر قيمة – أصغر قيمة. وهو في المثال السابق يساوي:

في المجموعة الأولى: 24 – 20 = 4

في المجموعة الثانية: 39 – 10 = 29

وعليه يمكن القول أن المجموعة الأولى أقل تشتتا وأكثر تجانسا من المجموعة الثانية.

والأمر نفسه تقريبا ينطبق على المدى في التوزيعات التكرارية المفيّئة، حيث يكون المدى حينها يساوي الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى (في البيانات المتصلة). أما في البيانات المنفصلة فيساوي الفرق   +1 أو الحد الأعلى الحقيقي للفئة الأخيرة ناقص الحد الأدنى الحقيقي للفئة الأولى.

ومثال الحالتين، ما يلي:

الجدول (35– أ: بيانات متصلة)

الفئات	التكرارات
– 11	11
– 14	3
– 17	9
– 20	13
– 23	11
28 – 26	3
المجموع	50

المدى = 28 – 11 = 17

الجدول (35–ب: بيانات منفصلة)

الفئات	التكرارات
4 – 2	2
7 – 5	5
10 – 8	6
13 – 11	3
16 – 14	1
المجموع	50

المدى = (16 – 2)+1 = 15   أو   15=1.5–16.5

ولكن الملاحظ أن هذا المقياس لا يأخذ بعين الاعتبار كل قيم الظاهرة المدروسة بل يعتمد في حسابه على القيمتين (أو الحدين) المتطرفتين فقط، وهو أمر غير كاف لمعرفة مدى التشتت بدقة، إذ قد تكون إحداهما (أو كلاهما) متطرفة بينما القيم الأخرى للمجموعة نفسها متجمعة أو قريبة من بعضها البعض، كما قد تكون لمجموعتين مختلفتي التجانس (عددا وتوزيعا) القيم المتطرفة نفسها. لذلك تستعمل مقاييس أخرى للتشتت تأخذ بعين الاعتبار كل قيم الظاهرة، وأهمها –كما مر معنا– الانحراف الربيعي، الانحراف المتوسط والانحراف المعياري. ولا يلجأ إلى المدى إلا في حالة القياس التقريبي السريع غير المبالي بالدقة أو إذا كان للقيمتين المتطرفتين أهمية خاصة، كتوزيعات درجات الحرارة أو معرفة ذروة توزيع الصحف وحضيضها... دون معدلاتها.

ثانيا) المئينيات والعشيرات والربيعات (Percentiles, Decimals & Quartiles)

هي عبارة عن أجزاء من المائة وكلها مقاييس تعطينا صورة عن تموقع قيم التوزيع وعن مدى تقاربها أو تباعدها عن بعضها. ويسميها البعض بشبيهات الوسيط المعبر عن 50% من القيم مقابل باقي النسب المئوية التي تعبر عنها شبيهاته.

– **المئينيات**

هي الأجزاء الناتجة عن تقسيم مساحة منحنى التوزيع التكراري إلى مائة جزء متساوٍ.

فالمئين الأول (P_1) في الترتيب التصاعدي هو القيمة التي يسبقها 1% من القيم ويليها 99% منها، والمئين الثاني (P_2) هو القيمة التي يسبقها 2% من القيم ويليها 98% منها... والمئين الستين (P_{60}) هو القيمة التي يسبقها 60% من القيم ويليها 40% منها وهكذا.

ويتم حساب المئين من **البيانات غير المبوبة** بإتباع الخطوات الآتية:

– نرتب القيم ترتيبا تصاعديا

– نحدد رتبة المئين عن طريق المعادلة الآتية:

$\frac{م}{100}\times(ن+1)$ أو $\frac{P}{100}\times(n+1)$، أي: رقم المئين على مائة في عدد القيم زائد واحدا.

– نحدد موقع المئين ثم المئين الذي تعكسه القيمة المقابلة له.

مثال: أوجد المئين العشرين للنقاط التالية: 12، 14، 25، 16، 15، 13، 17، 11، 18، 19.

الحل: – نرتب النقاط تصاعديا:

النقاط: 11، 12، 13، 14، 15، 16، 17، 18، 19، 25

رتبها: (1)، (2)، (3)، (4)، (5)، (6)، (7)، (8)، (9)، (10)

- نجد ترتيب المئين العشرين بتطبيق المعادلة: $\frac{م}{100}\times(ن+1)$

$$م_{20} = \frac{20}{100}\times(10+1) = 0.2\times11 = 2.2$$

وهو ترتيب يقع بين الرتبة الثانية والثالثة وتقابله النقطتان: 12 و 13

- قيمة المئين العشرين = $\frac{12+13}{2} = 12.5$

ومعناه أن 20% من النقاط تقل عن 12.5 و 80% منها تزيد عن 12.5

- أما **حساب المئين من البيانات المبوبة** فتتم تقريبا بالطرق نفسها المتبعة في حساب الوسيط، لأن الوسيط نفسه عبارة عن المئين 50:

وباعتماد الطريقة الأولى المعتمدة سابقا في حساب الوسيط (من التكرار المتجمع الصاعد مثلا)، نقوم أولا بتشكيل جدولا تكراريا صاعدا، ثم نحدد رتبة المئين كالآتي: $\frac{م}{100}\times مجك$ أو $\frac{P}{100}\times\sum f$

- نحدد موقع ترتيب المئين ثم الفئة المئينية وهي الموالية للفئة المقابلة له، إذا كانت البيانات منفصلة. وأما إذا كانت البيانات متصلة فرتبة المئين تقع بين حدي الفئتين المقابلة والموالية.

- نحسب المئين من المعادلة "الشبه–وسيطية" الآتية:

$$\text{الحد الأعلى للفئة المئينية} + \frac{\text{ترتيب المئين} - \text{ك.م.ص. للفئة قبل المئينية}}{\text{ك.م.ص. اللاحق} - \text{ت.م.ص. السابق}} \times \text{طول الفئة}$$

أي أن: المئين= $ل1 + \frac{ر - ك.م.ص1}{ك.م.ص2 - ك.م.ص1} \times ف$ $=$ $P = L1 - \frac{V - \sum f1}{\sum f2 - \sum f1}.C$

مثال: المطلوب إيجاد المئين الثلاثين من بيانات الجدول التالي:

الجدول (36)

الفئات	التكرارات
11 – 13	11
14 – 16	3
17 – 19	9
20 – 22	13
23 – 25	11
26 – 28	3
المجموع	50

الحل: نشكل جدولا تكراريا متجمعا صاعدا:

الجدول (37)

ت.م.ص	التكرارات	الفئات
11	11	13 – 11
14	3	16 – 14
23	9	19 – 17 فئة م$_{30}$
36	13	22 – 20
47	11	25 – 23
50	3	28 – 26
	50	المجموع

– ترتيب م$_{30}$ = $\frac{30}{100} \times 50 = 0.3\times 50 = 15$

– نحدد الفئة المئينية، وهي الفئة الثالثة (17–19)

– نقوم بتحديد خصائصها: حدها الأعلى: 19، طولها: 2، ت.م.ص. السابق= 14 وت.م.ص اللاحق= 36.

– نطبق معادلتها: م$_{30}$ = ل1 + $\frac{\text{ر} - \text{ك.م.ص}1}{\text{ك.م.ص}2 - \text{ك.م.ص}1} \times$ ف

م$_{30}$ = $19 + \frac{15-14}{36-14} \times 2 = 19 + \frac{1}{22} \times 2 = 19 + 0.09 = 19.09$

ومعناه أن 30% من البيانات تقل عن 19.09 و 80% منها تزيد عن 19.09

– أما مثال ذلك من البيانات المتصلة فسنعرضه فيما يلي من خلال تمرين أنموذجي أورده "عزام صبري" (2006، 139–140):

البيانات التالية تمثل أطوال 40 طالبا موزعين كما يلي:

الجدول (38)

172–169	–166	–163	–160	–157	فئات الأطوال
7	9	12	7	5	عدد الطلبة

المطلوب: إيجاد المئين الأول (P_1) والثلاثين (P_{30}) والتسعين (P_{90})

الحل: نشكل أولا جدولا متجمعا صاعدا

الجدول (39)

فئات الأطوال	عدد الطلبة	نهاية الفئات العليا	ك. م. ص.
157–	5	157>	0
160–	7	160>	5
163–	12	163>	12
166–	9	166>	24
169–172	7	169>	33
		172>	40
المجموع	40		

أولا) لإيجاد المئين الأول نستخرج ترتيبه: $م_1 = \frac{1}{100} \times 40 = \frac{4}{10} = 0.4$

وبما أن الترتيب أقل من التكرار المتجمع الصاعد للفئة الأولى (5) فلا نستطيع حل السؤال إلا إذا أضفنا فئة سابقة وتكرارها بالطبع صفر لأن ترتيب أي مئين لا بد أن يكون له تكرار متجمع صاعد سابق وآخر لاحق.

الفئة المئينية $م_1$=157 وأقل من 160 وحدها الأدنى 157

طول الفئة المئينية 160–157=3

التكرار السابق لهذه الفئة = 0 والتكرار اللاحق لها= 5

إذن $م_1 = ل1 + \frac{ر - ك.م.ص1}{ك.م.ص2 - ك.م.ص1} \times ف$

$م_1 = 157 + \frac{0.4 - 0}{5 - 0} \times 3 = 157 + \frac{0.4}{5} \times 3 = 157 + 0.24 = 157.24$

ثانيا) لإيجاد المئين الثلاثين نستخرج ترتيبه: $م_{30} = \frac{30}{100} \times 40 = \frac{120}{10} = 12$

نلاحظ أن ترتيب المئين الثلاثين جاء مطابقا لأحد التكرارات المتجمعة الصاعدة وهو 12، ولذلك فهو يساوي نهاية الفئة المناظرة لهذا التكرار م. ص. (12)، أي 163.

ثالثا) لإيجاد المئين التسعين نستخرج ترتيبه: $م_{90} = \frac{90}{100} \times 40 = \frac{360}{10} = 36$

الفئة المئينية= 169 وأقل من 172 وحدها الأدنى= 169

طول الفئة كما مر معنا = 3

ك. م. ص. السابق = 33

ك. م. ص. اللاحق = 40

إذن م$_{90}$ = $ل1 + \frac{ر - ك.م.ص1}{ك.م.ص2 - ك.م.ص1} \times ف$

م$_1$ = $169 + \frac{36-33}{40-33} \times 3 = 169 + \frac{3}{7} \times 3 = 169 + 1.28 = 170.28$

- **العشيرات**

وتعني تقسيم منحنى التوزيعات التكرارية إلى عشرة أقسام متساوية، يسمى كل منها عشيرا. فمثلا العشير الثاني هو القيمة التي يسبقها عشري القيم ($\frac{2}{10}$) ويليها ($\frac{8}{10}$) وذلك بعد ترتيب القيم ترتيبا تصاعديا. وبالطبع فالعشير الخامس هو الوسيط وأن هناك تسعة عشيرات فقط.

ويتم حساب العشير من **البيانات غير المبوبة** بإتباع الخطوات الآتية:

- نرتب القيم ترتيبا تصاعديا
- نحدد رتبة العشير عن طريق المعادلة الآتية:

$\frac{ش}{100} \times (ن+1)$ أو $(n+1) \times \frac{D}{100}$، أي: رقم العشير على مائة في عدد القيم زائد واحدا.

- نحدد الترتيب الأدنى والترتيب الأعلى لترتيب العشير
- نحدد القيمتين المقابلتين للترتيبين، ليكون العشير هو الوسط الحسابي للقيمتين المقابلتين.

مثال: أوجد العشير الثالث للنقاط التالية: 12، 14، 25، 16، 15، 13، 17، 11، 18، 19. مع تفسير النتيجة.

الحل:

- نرتب النقاط تصاعديا:

النقاط: 11، 12، 13، 14، 15، 16، 17، 18، 19، 25

رتبها: (1)، (2)، (3)، (4)، (5)، (6)، (7)، (8)، (9)، (10)

- نجد ترتيب العشير الثالث بتطبيق المعادلة: $\frac{ش}{100} \times (ن+1)$

ش$_3$ = $\frac{30}{100} \times (10+1) = 0.3 \times 11 = 3.3$

وهو ترتيب يقع بين الرتبة الثالثة والرابعة وتقابله النقطتان: 13 و14

إذن العشير الثالث = $\frac{14+13}{2}=13.5$

ومعناه أن30% من النقاط تقل عن 13.5 و 70% منها تزيد عن 13.5

– أما **حساب المئين من البيانات المبوبة** فتتم تقريبا بالطرق نفسها المتبعة في حساب الوسيط والمئين...

– **الربيعات**

وتعني تقسيم منحنى التوزيع التكراري إلى أربعة أجزاء متساوية. ويوجد ثلاثة ربيعات: الأول أو الأدنى (P_{25} أو Q_1) هو القيمة التي تسبقها ربع القيم وتليها ثلاثة أرباعها، أما الربيع الثاني أو الوسيط (P_{50} أو Q_2) فهو القيمة التي تسبقها نصف القيم و يليها النصف الآخر، وأما الربيع الثالث أو الأعلى (P_{75} أو Q_3) فهو القيمة التي تسبقها ثلاثة رباع القيم وتليها ربعها.

أما كيفية حسابها فتتم بالطريقة نفسها التي وجدنا بها المئينيات والعشيرات.

ثالثا) الانحراف الربيعي (Quartile Deviation)

للتخفيف من عيب المدى المطلق وتجنب القيم المتطرفة والشاذة يلجأ لقياس التشتت إلى الانحراف الربيعي أو ما يسمى بنصف المدى الربيعي، وذلك بالاعتماد على النصف المتوسط لمجموعة القيم من خلال قيمتي الربع الأعلى (الأخير) والربع الأدنى (الأول).

– وللحصول على الانحراف الربيعي (نصف المدى الربيعي) من **البيانات غير المبوبة** يحسب المدى بين الربيع الأدنى (ر 1) والربيع الأعلى (ر 3) ثم يقسم على اثنين، أي أنه يعتبر معدل اختلاف الربيعين عن الوسيط في التوزيع التكراري: الانحراف الربيعي (ن ر) = $\frac{ر_3 - ر_1}{2}$

ومعلوم أن الربيع الأدنى يقع في نهاية الربع الأول (أي في نهاية 25 %) من مجموعة القيم بعد ترتيبها تصاعديا والربيع الأعلى يقع في نهاية الربع الثالث (أي في نهاية 75%). ولحساب الربيعيين تحدد رتبتهما ثم قيمة نسبتهما المئوية، ومثال ذلك من الجدول الآتي:

الجدول (40)

الدرجة	التكرار	ك م ص
1	2	2
2	4	6
3	9	15
4	4	19
5	5	24
6	3	27
7	9	36
8	6	42
9	2	44
10	6	50
المجموع	50	

ترتيب $ر_1$ = $\frac{\text{مج ك}}{4} = \frac{50}{4} = 12.5$

ترتيب $ر_3$ = $\frac{\text{مج ك}}{4} \times 3 = \frac{50}{4} \times 3 = 37.5$ أو = $50 - 12.5 = 37.5$

قيمة $ر_1$ (الوسيطية) = الحد الأدنى الحقيقي لدرجة الربيع+ مقدار امتداده فيها

$= 2.5 + \frac{6.5}{9} = 2.5 + 0.72 = 3.22$

قيمة $ر_3$ = $7.5 + \frac{1.5}{6} = 7.5 + 0.25 = 7.75$

الانحراف الربيعي = $\frac{ر_3 - ر_1}{2} = \frac{7.75 - 3.22}{2} = \frac{4.53}{2} = 2.26$

ومن المعلوم أنه كلما كان الانحراف الربيعي أكبر كان التشتت أكثر.

– أما **حسابه من التوزيعات التكرارية للفئات** (البيانات المبوبة) فيكون كالآتي:

نقوم أولا بتحديد رتبة الربيعين الأعلى والأدنى ثم نحسب الربيعين بنفس طريقة إيجاد الوسيط (باستعمال طول الفئات، حدودها الدنيا وتكراراتها المتجمعة الصاعدة)، لنقسم بعد ذلك الفرق بينهما على اثنين. أي أن:

$ر_1$ أو $ر_3$ = الحد الأدنى للفئة الربيعية + $\frac{\text{رتبة الربيع - ك م ص للفئة السابقة}}{\text{تكرار الفئة الربيعية}}$ × *طول الفئة*

وانحرافهما يساوي طبعا $\frac{ر_3 - ر_1}{2}$

ومثال ذلك من جدول التوزيع التكراري التالي:

الجدول (41)

الفئات	التكرارات	ت.م.ص
11 – 13	11	11
14 – 16	3	14 فئة $ر_1$
17 – 19	9	23
20 – 22	13	36
23 – 25	11	47 فئة $ر_3$
26 – 28	3	50
المجموع	50	

ترتيب $ر_1$ = $\frac{مج\ ك}{4} = \frac{50}{4} = 12.5$

ترتيب $ر_3$ = $\frac{مج\ ك}{4}\times 3 = \frac{50}{4}\times 3 = 37.5$ أو = 50− 12.5= 37.5

قيمة $ر_1$ = 14+ $2\times\frac{11-12.5}{14}$ = 14+ $2\times\frac{1.5}{14}$

= 14+ 2×0.10 = 14+0.2 = 14.2

قيمة $ر_3$ = 23+ $2\times\frac{36-37.5}{11}$ = 23+ $2\times\frac{1.5}{11}$

= 23+ 2×0.13 = 23+ 0.27 = 23.27

نصف المدى الربيعي= $\frac{ر_3 - ر_1}{2}$ = $\frac{14.2-23.27}{2}$ = $\frac{9.07}{2}$ = 4.53

وإذا كان هذا الانحراف أكثر استقرارا من المدى فهو لا يأخذ بعين الاعتبار التباين بين القيم المركزية، كما يستثني القيم المتطرفة. ولتفادي ذلك يلجأ عادة إلى مقياسي الانحراف المتوسط والانحراف المعياري، اللذين يعتمدان بخلاف سابقيهما على قياس التشتت بين القيم عن إحدى متوسطاتها وليس عن بعضها البعض.

رابعا) الانحراف المتوسط (Mean Deviation)

وهو يستعمل كل قيم المجموعة بحساب متوسط انحرافات كل منها عن أي مقياس من مقاييس النزعة المركزية، ولكننا نستعمل في الغالب الوسط الحسابي لأنه مناسب أكثر.

ولأن من خصائص المتوسط أن مجموع انحرافاته يساوي صفرا، بسبب تساوي الفروق الإيجابية والسلبية، فللحصول على التشتت حوله يجب التخلص من الإشارات السلبية. وهي عملية غير مضرة لأن الذي يهم هنا هو مقدار البعد عن المتوسط بغض النظر عن النقص والزيادة. ويمكن فعل ذلك بإحدى طريقتين: إما بتجاهل إشارات الفروق مع اعتماد القيم المطلقة فقط، أو تربيع هذه الفروق. والانحراف المتوسط يعتمد الطريقة الأولى بخلاف الانحراف المعياري الذي يعتمد الطريقة الثانية.

ويعرف الانحراف المتوسط باسم المتوسط الحسابي للفروق المطلقة $|ف|$ لكل قيمة عن المتوسط، ولحسابه نقسم مجموع هذه الفروق المطلقة (أي بعد إهمال إشاراتها السالبة) على عدد القيم.

- وهو في **حالة القيم غير المبوبة** يساوي:

$$\text{الانحراف المتوسط} = \frac{مج|س - \overline{س}|}{ن} \quad \text{أو} \quad \frac{مج|ف|}{ن}$$

حيث س= مختلف القيم، و $\overline{س}$= الوسط الحسابي و ن= عدد القيم

ومنه فإن متوسط القيم التالية: 12، 14، 25، 16، 15، 13، 17، 11، 18، 19 هو 16. وإذا حذفنا 16 من كل رقم من هذه الأرقام، مع تجاهل سلبية الإشارات وجمعنا النتائج ثم قسمناها على 10 (عدد هذه الأرقام)، فسنحصل على 3 ، وذلك بإتباع الخطوات التالية:

بما أن الوسط الحسابي للقيم غير المبوبة = مجموع القيم على عددها، أي:

$$16 = \frac{160}{10} = \frac{19+18+11+17+13+15+16+25+14+12}{10}$$

$$\text{فإن الانحراف المتوسط} = \frac{3+2+5+1+3+1+0+9+2+4}{10} = \frac{30}{10} = 3$$

- أما **الانحراف المتوسط للقيم المبوبة** فيساوي:

مجموع الفرق بين مراكز الفئات والمتوسط × التكرارات ÷ مجموع التكرارات، أي:

$$\frac{\text{مج}|\text{م} - \overline{\text{س}}| \times \text{ك}}{\text{مج ك}}$$

مثال: أوجد الانحراف المتوسط لبيانات جدول التوزيع التكراري التالي رقم 42:

الفئات	التكرارات
4−2	2
7−5	5
10−8	6
13−11	3
16−14	1
مجموع	17

خطوات الحل:

أولا) المتوسط الحسابي = $\frac{\text{مج ك م}}{\text{مج ك}}$ = $\overline{X} = \frac{\sum fu}{\sum f}$ = $\frac{141}{17} = 8.29$

ثانيا) الفرق بيم مراكز الفئات والمتوسط الحسابي مع إهمال الإشارات

ثالثا) الانحراف المتوسط = $\frac{\text{مج}|\text{م} - \overline{\text{س}}| \times \text{ك}}{\text{مج ك}} = \frac{44.13}{17} = 2.59$

الجدول (43)

الفئات	ك	مراكز الفئات	ك×م	$\|\text{م} - \overline{\text{س}}\|$	$\|\text{م} - \overline{\text{س}}\| \times$ ك
4−2	2	3	6	5.29	10.58
7−5	5	6	30	2.29	11.45
10−8	6	9	54	0.71	4.26
13−11	3	12	36	3.71	11.13
16−14	1	15	15	6.71	6.71
مجموع	17		141		44.13

ولكن نظرا لطول وصعوبة حساب هذا المقياس عادة ما يلجأ المتخصصون إلى استعمال الانحراف المعياري:

خامسا) الانحراف المعياري (Standard Deviation)

إن الانحراف المعياري، في **حالة البيانات غير المبوبة**، عبارة عن الجذر التربيعي الموجب لمتوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي، أي:

ع = الجذر التربيعي لـ: $\frac{\text{مج}(س - \overline{س})^2}{ن}$ أو $S = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N}}$

أو أن تكون القسمة في المعادلة السابقة على ن−1 (N-1) إذا كان حجم العينة كبيرا، أي يزيد عن 30.

كما أن هناك طريقة ثانية (مباشرة أكثر) لحسابه، ومعادلتها:

ع= الجذر التربيعي لـ: $\frac{\text{مج س}^2}{ن} - \left(\frac{\text{مج س}}{ن}\right)^2 = S = \sqrt{\frac{\sum_{i=1}^{N} X_i^2}{N} - \left(\frac{\sum_{i=1}^{N} X_1}{N}\right)^2}$

<u>مثال</u>: أوجد الانحراف المعياري للبيانات (غير المبوبة) التالية:

12، 14، 25، 16، 15، 13، 17، 11، 18، 19.

<u>الحل</u> <u>بالمعادلة الأولى</u>:

أولا) المتوسط الحسابي= $\frac{19 + 18 + 11 + 17 + 13 + 15 + 16 + 25 + 14 + 12}{10} =$

$\frac{160}{10} = 16$

ثانيا) انحراف القيم عن وسطها الحسابي= −4، −2، +9، صفر، −1، −3، +1، −5، +2، +3

ثالثا) مجموع مربعات القيم = 16 + 4 + 81 + صفر + 1 + 9 + 1 + 25 + 4 + 9 = 150

رابعا) متوسطها = $\frac{\text{مج}(س - \overline{س})^2}{ن} = \frac{150}{10} = 15$

خامسا) الانحراف المعياري = الجذر التربيعي لـ: 15 = 3.872

<u>الحل بالمعادلة الثانية</u>:

أولا) متوسط مجموع مربعات القيم =

$\frac{361 + 324 + 121 + 289 + 169 + 225 + 256 + 625 + 196 + 144}{10} = \frac{2710}{10} = 271$

ثانيا) مربع متوسط مجموع القيم =

$$256 = \left(\frac{160}{10}\right)^2 = \left(\frac{19+18+11+17+13+15+16+25+14+12}{10}\right)^2$$

ثالثا) الفرق بينهما = 271− 256= 15

رابعا) الانحراف المعياري = الجذر التربيعي لـ: 15 = 3.872 (وهي النتيجة ذاتها)

– **أما في حالة البيانات المبوبة:** فلانحراف المعياري (ع) =

الجذر التربيعي لـ: $\frac{مج(س-\overline{س})^2 \times ك}{مج\ ك}$ أو $s = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2 \times f}{\sum f}}$

وتكون القسمة في المعادلة السابقة على مج ك−1 $(\sum f - 1)$ إذا كان حجم العينة كبيرا، أي يزيد عن 30.

وله كذلك طريقة ثانية لحسابه: ع = الجذر التربيعي لـ:

$\frac{مج\ س^2 \times ك}{مج\ ك} - \left(\frac{مج س \times ك}{مج\ ك}\right)^2$ أو $s = \sqrt{\frac{\sum_{i=1}^{N} X_i^2}{\sum f} - \left(\frac{\sum_{i=1}^{N} X_1}{\sum f}\right)^2}$

مثال: أوجد الانحراف المعياري لبيانات الجدول التالي (مع ملاحظة أن فئاته متصلة ومتساوية الطول لتيسير استخدامه):

الجدول (44)

الفئات	التكرارات
−2	2
−4	5
−6	6
−8	3
12−10	1
مجموع	17

الحل: بالمعادلة الأولى (غير المباشرة):

الانحراف المعياري= الجذر التربيعي لـ:

$$\frac{\text{مج}\left(\text{س}-\overline{\text{س}}\right)^2\times\text{ك}}{\text{مج ك}} \quad \text{أو} \quad s=\sqrt{\frac{\sum_{i=1}^{N}(X_i-\overline{X})^2\times f}{\sum f}}$$

الجدول (45)

الفئات	ك	مراكز الفئات(س)	ك×س	$\left(\text{س}-\overline{\text{س}}\right)$	$\left(\text{س}-\overline{\text{س}}\right)^2$	$\left(\text{س}-\overline{\text{س}}\right)^2 \times$ ك
2−	2	3	6	3.52−	12.39	24.78
4−	5	5	25	1.52−	2.31	11.55
6−	6	7	42	0.48+	0.23	1.38
8−	3	9	27	2.48+	6.15	18.45
10−12	1	11	11	4.48+	20.07	20.07
المجموع	17		111			76.23

أولا) المتوسط الحسابي $\overline{\text{س}} = \frac{\text{مج ك س}}{\text{مج ك}} = \frac{111}{17} = 6.52$

ثانيا) مجموع مربع فروق مراكز الفئات عن متوسطها في التكرارات

$\left(\text{س}-\overline{\text{س}}\right)^2 \times$ ك = 76.23

ثالثا) متوسط هذا المجموع= $\frac{\text{مج}\left(\text{س}-\overline{\text{س}}\right)^2\times\text{ك}}{\text{مج ك}} = \frac{76.23}{17} = 4.48$

رابعا) الانحراف المعياري = الجذر التربيعي لـ (4.48)= 2.1، أي:

$$\frac{\text{مج}\left(\text{س}-\overline{\text{س}}\right)^2\times\text{ك}}{\text{مج ك}} = 2.1$$

– الحل بالمعادلة الثانية (المباشرة): ع = الجذر التربيعي لـ: $\frac{\text{مج س}^2\times\text{ك}}{\text{مج ك}} - \left(\frac{\text{مجس}\times\text{ك}}{\text{مج ك}}\right)^2$ أو

$$s=\sqrt{\frac{\sum_{i=1}^{N}X_i^2}{\sum f}-\left(\frac{\sum_{i=1}^{N}X_1}{\sum f}\right)^2}$$

الجدول (46)

الفئات	ك	مراكز الفئات (س)	ك×س	س^2× ك
2−	2	3	6	18
4−	5	5	25	125
6−	6	7	42	294
8−	3	9	27	243
10−12	1	11	11	121
المجموع	17		111	801

أولا) متوسط مجموع نواتج ضرب مربع مراكز الفئات في التكرارات، أي:

$$\frac{\text{مج س}^2\times\text{ك}}{\text{مج ك}} = \frac{801}{17} = 47.11$$

ثانيا) مربع متوسط مجموع التكرارات في مراكز الفئات، أي:

$$\left(\frac{\text{مجس}\times\text{ك}}{\text{مج ك}}\right)^2 = \left(\frac{111}{17}\right)^2 = 42.63$$

ثالثا) الفرق بينهما = 47.11− 42.63 = 4.48

رابعا) الانحراف المعياري = الجذر التربيعي لـ (4.48) = 2.1، أي:

$$\frac{\text{مج س}^2\times\text{ك}}{\text{مج ك}} - \left(\frac{\text{مجس}\times\text{ك}}{\text{مج ك}}\right)^2 = 2.1$$ (وهي نتيجة الطريقة الأولى نفسها)

وتجدر الإشارة في الأخير إلى أنه كلما كان الانحراف المعياري أكبر كان التشتت المطلق أكثر للقيم، لأن مقدار انحرافاتها عن الوسط الحسابي يكون بالطبع أكبر.

سادسا) معامل الاختلاف (Coefficient of Variation)

بما أن المقاييس السابقة تخص التشتت المطلق والتعبير عنها يكون بوحدات القيم الأصلية لمجتمع الدراسة، فإن مقارنة التشتت بين مجموعتين من القيم يعبر عنهما بوحدات مختلفة يتطلب **مقاييس تشتت نسبية**، والتي عادة ما يلجأ إليها أيضا عند مقارنة التشتت بين مجموعتين يختلف متوسطهما الحسابي اختلافا كبيرا.

ومن هذه المقاييس معامل الاختلاف، وهو يساوي ناتج قسمة الانحراف المعياري على المتوسط الحسابي مضروبا في مائة، أي يعبر عن الانحراف المعياري كنسبة مئوية من المتوسط الحسابي. ومعادلته:

$$خ = \frac{ع}{\bar{س}} \times 100$$

- مثال (حالة فرق كبير بين متوسطين حسابيين): لنفرض أن المتوسط الحسابي والانحراف المعياري لكل من نقاط الطالبات ونقاط الطلبة في مادة لغة أجنبية بلغا ما يلي:

الطالبات: $\bar{س} = 15$ ع= 2.5

الطلبة: $\bar{س} = 10$ ع= 2.1

فإن قياس التشتت النسبي بين نقاط المجموعتين يمكن حسابه عن طريق معامل الاختلاف:

بالنسبة للطالبات $خ = \frac{2.5}{15} \times 100 = 16.66\%$

بالنسبة للطلبة $خ = \frac{2.1}{10} \times 100 = 21\%$

وهو يبين أن درجة اختلاف نقاط الطالبات أقل من درجة اختلاف نقاط الطلبة، ولذلك فهي تعكس تشتتا نسبيا أقل.

- مثال (حالة مجموعتين من القيم يعبر عنهما بوحدات مختلفة): لنفرض أن لدينا المتوسط الحسابي والانحراف المعياري لتغطية موضوع ما في الصحف (المعبر عنها هنا بوحدة "السم") وتغطيته في القنوات التلفزيونية (المعبر عنها هنا بوحدة الدقائق) في بلد ما:

السم: $\bar{س} = 50$ ع= 12

الدقيقة: $\bar{س} = 15$ ع= 3

بالنسبة للسنتم $خ = \frac{12}{50} \times 100 = 24\%$

بالنسبة للدقيقة $خ = \frac{3}{15} \times 100 = 20\%$

وهو يدل على أن عرض المادة الصحافية المكتوبة كان أكثر تشتتا من المادة السمعية البصرية.

سابعا) مقياس الالتواء (بيرسون)

ذكرنا سابقا أنه عندما يكوم توزيع القيم غير متماثل يقال عنه أنه ملتوٍ، إما التواء موجبا (المتوسط أكبر من المنوال) أو سالبا (المتوسط أصغر من المنوال)، وأن شدة الالتواء تزيد بازدياد الفرق بينهما. ولقياس تشتت قيم مثل هذه التوزيعات وضع "كارل بيرسون" (K. Pearson) معادلة إحصائية تعتمد على الفرق بين المتوسط ($\overline{س}$) والمنوال (من) وعلى الانحراف المعياري (ع) كمقسوم عليه (أو كأساس لقياس التشتت):

$$الالتواء = \frac{\overline{س} - من}{ع}$$

وتتراوح قيمة هذا المقياس النسبي بين –3 في الالتواء السالب و +3 في الالتواء الموجب. ولكنها نادرا ما تكون مساوية لـ "صفر" (لندرة وجود التوزيع المتماثل) أو لـ ±1 (لشذوذ هاتين القيمتين). ولكون هذا المقياس يعبر عن نسبة فهو مجرد كسابقه من طبيعة وحدات القيم أيضا.

4. الأساليب الإحصائية لدراسة العلاقات

تمهيد

إن أهداف البحث العلمي لظاهرة ما لا تنحصر في حدود الوصف المجرد للظاهرة ولعناصرها في واقعها الراهن من خلال مناهج المسح ودراسة الحالة وتحليل المضمون.. بل قد تتجاوز ذلك إلى البحث في العلاقات السببية وأحيانا إلى التجريب لضبط التأثيرات ضبطا تجريبيا.

ولذلك فإن الباحث بهدف الوصول إلى التعميمات النظرية يستعين بأي من المناهج المنتمية إلى الدراسات الوصفية أو التفسيرية أو التجريبية، وذلك من خلال الوصف والتفسير بالمقارنة المنهجية في دراسة السببية أو دراسة العلاقات الارتباطية أو التجريب.

فالباحث إذن لا يكتفي بدراسة ووصف متغير واحد كما يحدث ذلك في مقاييس النزعة المركزية ومقاييس التشتت، ذلك لأن العلم يهدف إلى دراسة العلاقة بين الظواهر والحقائق. وهناك عدة أساليب إحصائية لدراسة العلاقات بين المتغيرات: معاملا ت الارتباط، اختبار كا ^{2}، دلالة النسب المئوية... ولكن قبل عرض أهمها تجدر الإشارة إلى كيفية صياغة الفروض والعلاقة بين المتغيرات وإلى طرق البحث عن العلاقات السببية:

1.4- طرق البحث عن العلاقات السببية

وهي الأساليب المتبعة للبحث في أسباب حدوث الظاهرة من خلال المقارنة والإجابة على السؤال لماذا؟ و يمكن تحديد التصميمات المنهجية في البحث عن العلاقات السببية المقارنة من خلال الآتي:

1.1.4- طريقة الاتفاق

تعني أن تكرار وجود متغير ما في أكثر من جماعة تحدث فيه الظاهرة محل الدراسة يجعلنا نفسر الظاهرة بوجود المتغير المشترك.

والمثال على ذلك هو اختيارات جماعة من الأطفال متباينة السمات، وتتفق في المشاهدة (الوقت الذي تقضيه في المشاهدة / محتوى البرنامج) وفي السلوك العدواني فإنه يمكن أن نعزي هذا السلوك المشترك إلى خصائص المشاهدة التلفزيونية.

ويمكن توضيح ذلك من خلال مثال آخر يخص العلاقة بين انخفاض الدخل (كمتغير مستقل: م. م.) والعزوف عن قراءة الصحف (كمتغير تابع: م. ت.):

- جماعة 1 (جامعيون) انخفاض الدخل (م م1) (م ت1) العزوف عن قراءة الصحف
- جماعة 2 (أميون) انخفاض الدخل (م م1) (م ت1) العزوف عن قراءة الصحف
- جماعة 3 (ثانوين) انخفاض الدخل (م م1) (م ت1) العزوف عن قراءة الصحف

ويحسم بناء العلاقة السببية في هذه الحالة الاختلاف في الخصائص الأخرى، ك المستوى التعليمي أو المرحلة العمرية على سبيل المثال ، مع الاتفاق في انخفاض الدخل، بحيث يمكن أن يقرر أن العزوف عن قراءة الصحف يعود إلى انخفاض مستوى الدخل.

2.1.4- طريقة الاختلاف

إن هذه الطريقة هي عكس الطريقة السابقة ، فإذا ما اتفقت الجماعات في كل المتغيرات واختلفت في متغير واحد، فإن الاختلاف في هذا المتغير يمكن أن يفسر سبب حدوث الظاهرة. أي أن غياب المتغير في إحدى الجماعات هو السبب في حدوث الظاهرة.

- جماعة 1 (م م1) ارتفاع مستوى التعليم (م ت1) قراءة الصحف
- جماعة 2 (م م1) ارتفاع مستوى التعليم (م ت1) قراءة الصحف
- جماعة 3 (م م1) انتشار الأمية (م ت2) عدم قراءة الصحف

وفي هذه الحالة يمكن تفسير عدم قراءة الصحف باعتبارها نتيجة لانتشار الأمية.

وهذان التصميمان يكملان بعضهما بعض ا بحيث يمكن أن يضيفهما الباحث معا كتصميم مشترك يبحث بداية في عناصر الاتفاق ثم عناصر الاختلاف لتقرير العلاقة السببية، وعادة ما يقوم الباحث بذلك منهجيا عندما يكشف عن عناصر الاتفاق لأغراض الضبط أو العزل، أو العكس وهو التصميم السائد في دراسات السببية المقارنة ويتفق تماما مع الدراسات الإعلامية وطبيعتها التي تتسم بتعدد المتغيرات وتداخلها. وفي البحث عن أي من المتغيرات، فإنه يبحث أيضا في المتغيرات العكسية في الوقت نفسه، وهذا هو التصميم المشترك للاتفاق والاختلاف.

3.1.4- طريقة التلازم في التغيير (التباين المشترك)

يعتمد هذا التصميم على القياس الكمي للعلاقة السببية، وملاحظة التغير في المتغير المستقل وكذلك التغير في المتغير التابع ، من خلال تطور التغير أو العلاقة الارتباطية بينهما. بحيث يمكن تفسير العلاقة السببية على أساس وجود هذا الارتباط أو غيابه.

- الجماعة (1) م م + م ت+
- الجماعة (2) م م ++ م ت++
- الجماعة (3) م م +++ م ت+++

مثل تفسير التباين في اكتساب المعرفة بسبب التباين في كثافة المشاهدة التلفزيونية.

- الجماعة (1) +1 ساعة كسب معرفي +
- الجماعة (2) +2 ساعتين كسب معرفي ++
- الجماعة (3) +3 ساعات كسب معرفي +++

أو ملاحظة التغير في النتائج على مدى فترات زمنية بتأثير التغير في المثير في هذه الفترات للجماعة الواحدة نفسها.

- الجماعة (1) +1 ساعة كسب معرفي +
- الجماعة (1) +2 ساعتين كسب معرفي ++
- الجماعة (1) +3 ساعات كسب معرفي +++

ومما يجدر ملاحظته في تطبيق التصميم المنهجي المناسب هو زيادة جهد الباحث ودقته، في الكشف عن المتغيرات الأصلية، الدخيلة أو الزائفة، وتفسير العلاقات السببية. ويستعين الباحث في تقرير ذلك بالمقارنة بين المتوسطات والفروق الإحصائية بين نتائج حركة المتغيرات المثبتة أو النافية لوجود العلاقة السببية.

وهي ميزة تفتقدها الدراسات الارتباطية التي تقف عند حدود تقرير العلاقة ومدى الارتباط (كما سنرى لاحقا)، ذلك أن الارتباط لا يعني أن حركة المتغير "س" هي السبب في حركة المتغير "ع" لأن العلاقة قد تكون بفعل متغير ثالث يؤثر في الاثنين معا: فالعلاقة الموجبة بين مستوى التعليم وتوزيع الصحف قد ترجع إلى ارتفاع مستوى الدخل مثلا. ولذلك نقول أننا نلاحظ *قياسيا* وجود *اتجاه واضح* (وليس علاقة سببية) بين تغير المستوى التعليمي وتوزيع الصحف.

2.4- بعض معاملات الارتباط

1.2.4- الارتباط

بعد عرض كيفية وصف متغير واحد أو ظاهرة واحدة من حيث إيجاد متوسطات هذا المتغير بقيمه المختلفة وتقديم طرق البحث عن العلاقة السببية، سنتطرق فيما يأتي إلى دراسة العلاقة بين متغيرين أو ظاهرتين من خلال معاملات الارتباط. ولكن قبل ذلك تجدر الإشارة إلى مفهوم الارتباط، أشكاله، مقاييسه وأنواعه:

- مفهوم الارتباط وأنواعه

تستخدم معاملات الارتباط في الكشف عن طبيعة العلاقة بين متغيرين: هل هي طردية (موجبة)، بمعنى أنه كلما زاد المتغير المستقل زاد معه المتغير التابع، وكلما نقص المتغير المستقل نقص معه المتغير التابع، مثل الزيادة في انتظام حضور الطلبة التي تؤدي إلى الزيادة في درجة تحصيلهم، أو النقص في التركيز الذي يؤدي إلى النقص في الفهم. أم هي علاقة عكسية (سالبة) بمعنى إذا زاد المتغير المستقل نقص المتغير التابع والعكس بالعكس، مثل الزيادة في منحة الطلبة التي تؤدي إلى النقص في احتجاجاتهم...

وعندما نعبر عدديا عن نوع هذه العلاقة في مجال العلوم الإنسانية والاجتماعية فإن هذه العلاقة تقع دائما بين (+1) و (-1)، وذلك لأن العلاقة التامة أو الارتباط الكامل الموجب (الطردي) أو السالب (العكسي)، نادرا ما يصادفنا في الدراسات الاجتماعية والإنسانية: فإذا رتبنا مثلا قيم المتغير "س" ترتيبا تصاعديا ووجدنا أن قيم المتغير "ص" المناظرة لها مرتبة ترتيبا تصاعديا أيضا نستنتج وجود ارتباط طردي تام بين المتغيرين "س" و"ص". أما إذا رتبنا قيم المتغير "س" ترتيبا تصاعديا ووجدنا أن قيم المتغير "ص" المناظرة لها مرتبة ترتيبا تنازليا نستنتج وجود ارتباط عكسي تام بين المتغيرين "س" و"ص". وفي الحالات الأخرى يتراوح معامل الارتباط (شدة وضعفا) كما أسلفنا بين (+1) و(-1). وفيما يأتي وصف لأهم مستويات قيم الارتباط:

إذا كانت قيمة الارتباط:

= صفرا (ر= 0) العلاقة بين المتغيرين منعدمة،

من > أو< من 0 إلى ±] 0.2 درجة ارتباط منخفضة جدا

من ± 0.2 إلى ±] 0.4 درجة ارتباط منخفضة أو ضعيفة

من ± 0.4 إلى ±] 0.7 درجة ارتباط متوسطة

من ± 0.7 إلى ±1   درجة ارتباط عالية أو قوية

– مقاييس الارتباط

عند دراسة الارتباط يجب أن نتناول المقاييس التالية:

معادلة التقدير أو الانحدار

تصف معادلة التقدير أو الانحدار العلاقة بين ظاهرتين، وكما يستدل من تسميتها فإن الأهداف الرئيسة لهذه المعادلة هي تقدير القيم التي يمكن أن يأخذها المتغير التابع بدلالة المتغير المستقل.

الخطأ المعياري للتقدير

يبين هذا المقياس مقدار انحراف القيم الفعلية للمتغير التابع عن القيم المقابلة لها والمحسوسة باستخدام معادلة الانحدار. ويستخدم هذا المقياس كمعيار لمقدار الثقة التي يمكن للباحث أن يضعها في القيم المقدرة (المحسوسة) حول منحنى الانحدار. وبذلك فهو كالانحراف المعياري، الذي يقيس التشتت، يعطينا فكرة حول مقدار الاعتماد أو الثقة بالتقديرات.

معامل الارتباط

يستخدم هذا المقياس لمعرفة درجة أو شدة العلاقة بين المتغيرين أو الظاهرتين المدروستين.

– أنواع الارتباط:

يمكن التمييز بين أنواع الارتباط الرئيسة الآتية:

1– الارتباط البسيط المستقيم

يتحقق هذا النوع من الارتباط إذا توفر الشرطان التاليان:

أ – أن تكون العلاقة موضوع الدراسة مقتصرة على ظاهرتين أو متغيرين اثنين فقط.

ب– أن تكون العلاقة موضوع الدراسة مما يمكن تمثيلها بيانيا بخط مستقيم أحسن تمثيل.

2 – الارتباط البسيط غير المستقيم

يتحقق هذا النوع من الارتباط إذا توفر الشرطان التاليان:

أ –أن تكون العلاقة موضوع الدراسة مقتصرة على ظاهرتين أو متغيرين اثنين فقط.

ب أن تكون العلاقة موضوع الدراسة مما لا يمكن تمثيلها بيانيا بخط مستقيم أي تحتوي على التواء أو أكثر ولكن مع وجود اتجاه عام للارتباط.

3 – الارتباط المتعدد

يتحقق هذا النوع من الارتباط عند القيام بدراسة العلاقة بين ظاهرة وظاهرتين أو أكثر في أن واحد، حيث يعتبر في هذه الحالة أحد المتغيرات هو المتغير التابع، وباقي المتغيرات الأخرى هي المتغيرات المستقلة.

مثال :3 متغيرات مستقلة: ارتفاع الدخل والمستوى التعليمي والسن، ومتغير تابع: زيادة قراءة الصحف.

4 – الارتباط الجزئي

يتحقق هذا الشكل من الارتباط عند القيام بدراسة الارتباط بين ظاهرة وظاهرتين أو أكثر منفردة، أي الارتباط الجزئي بين متغير تابع ومتغير مستقل مع معاملة المتغيرات الأخرى معاملة الثوابت (أي اعتبارها متغيرات ثابتة).

ويمكن قياس هذه الارتباطات عن طريق معادلة التقدير أو الخطأ المعياري أو الانحدار أو معاملات الارتباط، التي تعنينا هنا...

ولمعاملات الارتباط عدة أشكال منها:

- معامل ارتباط الرتب لـ"سبيرمان"
- معاملات ارتباط بيرسون
- معامل التوافق
- معامل فاي

ولكننا سنقتصر هنا على عرض معاملات "سبيرمان" و"بيرسون":

ك2أولا) معامل ارتباط الرتب لـ"سبيرمان"

يستخدم هذا المعامل لدراسة الارتباط بين البيانات النوعية، أي تلك التي لا يمكن قياسها كميا والتي يكون العدد فيها صغيرا. ويعتمد في حسابه على إعطاء المتغيرات رتبا لتحل محل القياس العددي.

ولقياس الارتباط بين المتغيرين "س" و"ع" نرتب كلا منهما حسب أفضليته (ترتيبا تنازليا أو تصاعديا) ثم نحسب الفرق (ف) (D) بين كل رتبتين متقابلتين، فنجد أن مجموع الفروق يساوي صفرا (مجـ ف=0) ($\Sigma D = 0$)، وبحساب مربعات هذه الفروق (ف^2) (D^2) يمكن إيجاد معامل ارتباط الرتب باستخدام العلاقة:

$$r_s = 1 - \frac{6\Sigma D^2}{N(N^2-1)} \qquad ر = 1 - \frac{6\ مجـ ف^2}{ن(ن^2-1)}$$

حيث: ر (r_s) هو معامل ارتباط "سبيرمان"

ف (D) الفرق بين رتب القيم

ن (N) عدد القيم

<u>مثال 1</u>: احسب معامل "سبيرمان" للارتباط بالرتب بين التقديرات الموالية لطالب وطالبة في سيرتهما العلمية الرباعية.

الجدول (50)

	بكالوريا	ليسانس	ماجستير	دكتوراه
تقديرات الطالبة "س"	جيد	ممتاز	مقبول	جيد جدا
تقديرات الطالب "ع"	قريب من الجيد	جيد	مقبول	ممتاز

<u>الحل</u>:

نرتب تقديرات الطالبين بحيث يعطى أحسن تقدير بالنسبة للطالبة "س" (ممتاز) الرتبة الأولى أي رقم 1، ثم يليه من حيث الأهمية التقدير جيد جدا الذي يأخذ الرتبة الثانية أي رقم 2، ثم يأتي التقدير جيد الذي يأخذ الرقم 3، ثم التقدير مقبول رقم 4.

ونرتب تقديرات الطالب "ع" بحيث يعطى أحسن تقدير وهو ممتاز الرقم 1، ثم تليه بقية التقديرات إلى حد تقدير مقبول الذي يأخذ الرقم 4.

كما يمكن أن يكون الترتيب بصورة عكسية أي نبدأ بأقل تقدير فنعطيه رقم 1، ثم الذي أحسن منه وهكذا حتى نصل إلى أحسن تقدير فنعطيه الرقم 4، إذ أن التقدير قد يكون تصاعديا أو تنازليا. ثم نجد الفرق بين رتب تقديرات الطالبين كما هو موضح في الجدول الموالي الذي نحسب منع معامل "سبيرمان" (ر) (r_s).

الجدول (51)

تقديرات "س"	تقديرات "ع"	رتب "س"	رتب "ع"	ف D	ف D^2
جيد	قريب من الجيد	3	3	0	0
ممتاز	جيد	1	2	1−	1
مقبول	مقبول	4	4	0	0
جيد جدا	ممتاز	2	1	1	1
				$\Sigma D = 0$	$\Sigma D^2 = 2$

$$r_s = 1 - \frac{6\Sigma D^2}{N(N^2-1)} = 1 - \frac{6\times2}{4(16-1)} = 1 - \frac{12}{60} = 1 - 0.2 = \mathbf{0.8}$$

(وهي علاقة موجبة "قوية")

مثال 2: أراد باحث أن يكشف عن العلاقة بين العمر والذكاء لدى مجموعة مكونة من ستة طلبة وكانت درجاتهم على هذين المتغيرين كالآتي:

الجدول (52)

الطلبة "ق"	العمر "س"	الذكاء "ع"	رتبة س	رتبة ع	ف	ف²
1	25	9	2	4	2−	4
2	15	10	3	3	0	0
3	30	12	1	1	0	0
4	10	8	5	5	0	0
5	8	11	6	2	4+	16
6	12	7	4	6	2−	4
					مج ف= 0	مج ف² =24

الحل: $ر = 1 - \frac{6\times24}{6(36-1)} = 1 - \frac{144}{6\times35} = 1 - \frac{144}{210} = 1 - 0.686 = 0.31$

(وهي علاقة موجبة "منخفضة" بين العمر والذكاء لدى هذه المجموعة).

مثال 3: احسب معامل سبيرمان للارتباط بالرتب بين المعدلات الموالية (من مائة) لعشرة طلبة في امتحان شهادة البكالوريا.

الجدول (53)

معدل الطالب في السنة الثالثة ثانوي (س)	75	82	65	90	77	60	55	87	91	73
معدل الطالب في شهادة البكالوريا (ع)	69	85	55	90	80	50	57	88	89	71

الحل: نرتب المعدلات (س) بحيث نعطى الرتبة الأولى (1) لأعلى معدل ثم الذي يليه، الخ. ونرتب المعدلات (ع) ثم نجد الفرق بين رتب معدلي كل طالب كما هو موضح في الجدول الآتي. وأخيرا نحسب معامل سبيرمان "ر" (r_s):

الجدول (54)

رتبة س	رتبة ع	ف	ف^2
6	7	1−	1
4	4	0	0
8	9	1−	1
2	1	1	1
5	5	0	0
9	10	1−	1
10	8	2	4
3	3	0	0
1	2	1−	1
7	6	1	1
		مج ف= 0	مج ف^2 =10

ر $=1-\frac{6\times10}{10\times99} = 0.94$ (ارتباط موجب "قوي")

ملاحظة: في المثالين السابقين (2، 3) لم تظهر درجات أو معدلات متساوية في أي من البيانات (س) أو (ع). أما إذا ظهرت بيانات متساوية (tied data) فيكون تعيين رتب هذه البيانات

(بعد ترتيبها) بأخذ الوسط الحسابي لرتب كل مجموعة من البيانات المتساوية واعتباره رتبة كل بيان في هذه المجموعة. ومثال عملية الترتيب في حالة التساوي ما يأتي:

– عين رتب البيانات الآتية تنازليا: 60، 65، 60، 70، 60، 72، 70

الحل: نكتب البيانات تنازليا ونعين الرتب كما سيأتي:

72، 70، 70، 65، 60، 60، 60

(1)، (2)، (3)، (4)، (5)، (6)، (7)

نجد الوسط الحسابي لرتبتي العلامتين: 70، 70. وهو 2.5 أي $\left(\frac{3+2}{2}\right)$

نجد الوسط الحسابي لرتب العلامات: 60، 60، 60. وهو 6 أي $\left(\frac{7+6+5}{2}\right)$

فتكون رتب العلامات كما هو مبين في الجدول الآتي:

الجدول (55)

العلامة	60	65	60	70	60	72	70
الرتبة	6	4	6	2.5	6	1	2.5

حساب معامل سبيرمان لجدول مزدوج (جدول الانتشار):

يمكن حساب معامل ارتباط الرتب للجداول التكرارية المزدوجة بالطريقة نفسها التي تعاملنا بها مع البيانات المبوبة، وذلك بترتيب فئات المتغيرين "س" و"ع" تصاعديا أو تنازليا ثم إتباع طريق أقطار الفروق المتساوية أو طريقة أقطار المجاميع المتساوية. ولكن قبل تفصيل الكلام عن الطريقتين تجدر الإشارة إلى **كيفية تكوين جدول مزدوج** من خلال المثال الآتي:

– ولتكن لدينا البيانات الآتية عن عدد غيابات عمال المؤسسة "س" (x) والوحدات المنتجة "ع" (y) في فروعها العشرين:

س(x): 4، 9، 2، 4، 10، 10، 8، 2، 1، 6، 2، 4، 8، 10، 6، 2، 5، 3، 8، 5

ع(y): 10، 10، 8، 2، 2، 4، 11، 6، 6، 8، 4، 5، 9، 10، 2، 2، 6، 3، 4، 8

المطلوب: تفريغ هذه البيانات في جدول توزيع تكراري مزدوج على أساس طول فئة الغيابات (2) وطول فئة الوحدات (2) والحد الأدنى للفئة الأولى بالنسبة للغيابات (1) والحد الأدنى للفئة الأولى بالنسبة للوحدات (2). فيكون الجدول كما سيأتي:

جدول(56أ) تكرارات المتغير "س"(x) جدول (56ب) تكرارات المتغير "ع"(y)

ك	التفريغ بالقيم	الفئات
5	1، 2، 2، 2، 2	]3 − 1]
4	3، 4، 4، 4	]5 − 3]
4	5، 6، 6، 6	]7 − 5]
3	8، 8، 8	]9 − 7]
4	9، 10، 10، 10	]11 − 9]

ك	التفريغ بالقيم	الفئات
5	2، 2، 2، 2، 3	]4 − 2]
4	4، 4، 4، 5	]6 − 4]
3	6، 6، 6	]8 − 6]
4	8، 8، 8، 9	]10 − 8]
4	10،10، 10، 11	]12 − 10]

الجدول (56 ج) يعكس التوزيع التكراري المزدوج

	]12−10]	]10−8]	]8−6]	]6−4]	]4−2]	الوحدات المنتجة	عدد الغيابات
fx	11	9	7	5	3	y / x	
5					4	2	]3 − 1]
4				3		4	]5 − 3]
4			2			6	]7 − 5]
3		3				8	]9 − 7]
4	3					10	]11 − 9]
Σf20		4	3	4	5	Fy	

ملاحظة: يمثل السطر الأول من الجدول مراكز فئات "ع" والعمود الأول مراكز فئات "س". أما الأرقام داخل الجدول فتمثل التكرارات المشتركة بين المتغيرين "س، ع" (x,y). فمثلا العدد 4 الذي يقع في تقاطع العمود الثالث والسطر الثالث يدل على عدد غيابات العمال والوحدات المنتجة التي تتراوح أرقامها بين حدي الفئة الأولى من الغيابات والفئة الأولى من الوحدات المنتجة (وهو هنا الرقم 2 فقط لأن الرقم 1 "السيني" يقع خارج حدود فئة "العِينات"). وأما "ك س"(fx) و"ك ع" (fy) فتمثل تكرارات المتغيرين "س،ع" (x,y)، و(Σf) مجموع كل منهما.

حساب معامل "سبيرمان": ويحسب بطريقتين: أقطار الفروق المتساوية ومجاميع الأقطار المتساوية، وسنكتفي هنا بعرض الطريقة الأولى:

- طريق أقطار الفروق المتساوية:

من خلال الجدول (57) المصنف لعلامات الطلبة في مادتين نقوم بالخطوات الموالية والموضحة في الجدول (58):

1- نرتب فئات المتغير تصاعديا ثم ونعطي لها الأرقام 1، 2، 3، 4، 5. كما نرتب فئات المتغير تصاعديا ونعطي لها الأرقام 1، 2، 3، 4.

2- نأخذ القطر الرئيس للجدول، وهو الذي يمر بخانات تكون متساوية الترتيب للمتغيرين، ونلاحظ أن الفرق بين رتبتي أي خانة من الخانات التي يمر بها القطر الرئيس يساوي صفرا. فبالنسبة لجدولنا نجد أن التكرارات 2، 5، 5، 4. ومجموعها يساوي 16 تقع على القطر الرئيس وبالتالي نجد الفرق بين رتبتي أي خانة من خانات هذه التكرارات يساوي صفرا.

3- نأخذ الأقطار المتوازية للقطر الرئيس، فتكون لدينا أقطار فوقه وأقطار تحته، فمثلا نجد أن التكرارات 1، 7، 3، 2 تقع على القطر الذي يوجد أعلى القطر الرئيس ومجموعها 13 والفرق بين رتبتي أي تكرار يساوي 1. أما القطر الذي يقع تحت القطر الرئيس وتكراراته هي 3، 4، 4 ومجموعها 11 والفرق بين رتبتي أي خانة يساوي -1.

4- نكون جدولا تكراريا قيمه هي الفروق (D) وتكراراته هي مجاميع تكرارات الأقطار السابقة التي تقابل الفروق.

5- نكون جدولين آخرين لرتب المتغيرين والتكرارات هي التكرارات الهامشية المقابلة لكل متغير أي رتب X تقابلها التكرارات f_x ورتب Y تقابلها التكرارات f_y. نحسب الانحرافات المعيارية لكل من الجداول الثلاثة السابقة، ويكون معامل الارتباط في هذه الحالة يساوي:

$$r_S = \frac{\sigma_x^2 + \sigma^2_y - \sigma^2_D}{2\sigma_x . \sigma_y}$$

حيث أن: σ^2_x تباين المتغير X

σ^2_y تباين المتغير Y

σ^2_D تباين الفروق D

الجدول (57) جدول مزدوج يوضح التكرارات المشتركة

20 − 16	16 − 12	12 − 8	8 − 4	]4 − 0]	الإحصاء / المنهجية
			1	2	]5 − 0[
		7	5	3	10 − 5
	3	5	4		15 − 10
2	4	4			20 − 15

الجدول (58) يوضح الأقطار والرتب

	رتب X	1	2	3	4	5	
رتب Y	X / Y	[0-4 [	[4-8 [	[8–12 [	[12-16[	[16-20[	f_y
1	[0 - 5 [	2	1				3
2	[5 - 10 [	3	5	7			15
3	[10 -15 [		4	5	3		12
4	[15 -20 [			4	4	2	10
	f_x	5	10	16	7	2	$\Sigma f=40$

أما جداول المتغيرات الثلاثة D, Y, X المستنتجة من الجدول (58) فهي:

الجدول (58ب) الجدول (58 ج) الجدول (58 أ)

جدول المتغير X

F^2_x	$f_x x$	f_x	رتبX
5	5	5	1
40	20	10	2
144	48	16	3
112	28	7	4
50	10	2	5
351	111	40	مجموع

جدول المتغير Y

F^2_y	$F_y y$	F_y	رتبY
3	3	3	1
60	30	15	2
108	36	12	3
160	40	10	4
331	109	40	مجموع

جدول الفروق D

F^2_D	$F_D D$	F_D	فروقD
13	13	13	1
0	0	16	0
11	11−	11	1−
24	2	40	مجموع

1− تباين المتغير X

$$\sigma^2_x = \frac{\Sigma f_x x^2}{\Sigma f_x} - \left(\frac{\Sigma f_x x}{\Sigma f_x}\right)^2 = \frac{351}{40} - \left(\frac{111}{40}\right)^2 = 1.075 \approx 1.08 \Rightarrow \sigma_x \approx 1.04$$

2− تباين المتغير Y

$$\sigma^2_{y} = \frac{\Sigma f_y y^2}{\Sigma f_y} - \left(\frac{\Sigma f_y y}{\Sigma f_y}\right)^2 = \frac{331}{40} - \left(\frac{109}{40}\right)^2 = 0.849 \approx 0.85 \Rightarrow \sigma_x \approx 0.92$$

3- تباين المتغير D

$$\sigma^2_{D} = \frac{\Sigma f_D D^2}{\Sigma f_D} - \left(\frac{\Sigma f_D D}{\Sigma f_D}\right)^2 = \frac{24}{40} - \left(\frac{2}{40}\right)^2 = 0.5975 \approx 0.60$$

نطبق صيغة سبيرمان:

$$rs = \frac{\sigma_x{}^2 + \sigma^2{}_y - \sigma^2{}_D}{2\sigma_x.\sigma_y} = \frac{1.08 + 0.85 - 0.60}{2(1.04)(0.92)} = \frac{1.33}{1.91} = 0.69$$

ثانيا) معاملات ارتباط "بيرسون"

تتفادى معاملات ارتباط "بيرسون" العيوب الموجودة في معامل ارتباط الرتب والمتعلقة باعتماده على الرتب في حسابه لا على القيم نفسها. ومعاملات ارتباط "بيرسون" هي:

- معامل ارتباط "بيرسون" عن طريق الانحراف (معقد الاستعمال)
- معامل ارتباط "بيرسون" عن طريق القيم الخام (نادر الاستعمال)
- معامل ارتباط "بيرسون" عن طريق القيم المبوبة (شائع الاستعمال). وسنفصل الكلام عنه فيما يأتي:

حساب معامل ارتباط "بيرسون" من بيانات مبوبة

تتعلق العملية هنا بمتغيرين، لذلك ستكون لدينا فئات تتعلق بالمتغير الأول (س) وفئات تتعلق بالمتغير الثاني (ع)، وسنرمز لتكرارات فئات المتغير الأول بـ (ك $_{س}$) (f_x) ولتكرارات فئات المتغير الثاني بـ (ك$_{ع}$) (f_y)، أما التكرارات المشتركة فنرمز لها بـ(ك) (f) .

ولحساب معامل ارتباط بيرسون لجدول مزدوج أي في توزيع تكراري لمتغيرين، فإننا نستخدم العلاقة المباشرة الآتية:

$$r = \frac{N\Sigma\Sigma fXY - \Sigma f_x X . \Sigma f_y Y}{\sqrt{[N\Sigma f_x X - (\Sigma f_x X)^2][N\Sigma f_y Y - (\Sigma f_y Y)^2]}}$$

$$ر = \frac{\text{ن مج ك س ع} - (\text{مج ك}_{س}\text{ س})(\text{مج ك}_{ع}\text{ ع})}{\sqrt{\left[\text{ن مج ك}_{س}\text{ س} - (\text{مج ك}_{س}\text{ س})^2\right]\left[\text{ن مج ك}_{ع}\text{ ع} - (\text{مج ك}_{ع}\text{ ع})^2\right]}}$$

حيث $N = \Sigma f$ (مجـ ك = ن)، تسمى f_x (ك $_{س}$) التكرارات الهامشية للمتغير س x، وكذلك f_y (ك $_{ع}$) تسمى التكرارات الهامشية للمتغير ع y، وينتج عن ذلك أن:

مجـ ك = مجـ ك$_{س}$ = مجـ ك $_{ع}$: $\Sigma f = \Sigma f_x = \Sigma f_y$.

والصيغة السابقة يمكن كتابتها بالشكل الموالي بشرط أن يكون مدى فئات المتغير الأول متساويا ومدى فئات الثاني:

$$ر = \frac{ن\ مجـ ك ح خ - (مجـ ك_{س}\ ح)(مجـ ك_{ع}\ خ)}{\sqrt{\left[ن\ مجـ ك_{س}\ ح - (مجـ ك_{س}\ ح)^2\right]\left[ن\ مجـ ك_{ع}\ خ - (مجـ ك_{ع}\ خ)^2\right]}}$$

$$r = \frac{N\Sigma\Sigma fUV - \Sigma f_x U \Sigma f_y V}{\sqrt{[N\Sigma f_x U - (\Sigma f_x U)^2][N\Sigma f_y V - (\Sigma f_y V)^2]}}$$

حيث: U هو الانحراف و V هو الاختلاف.

$$U = \frac{X_i - A}{C_x} = \frac{d_x}{c_x} \quad و \quad V = \frac{Y_i - A'}{C_y} = \frac{dy}{C_y}$$

و C_x هو مدى فئات المتغير X والمتغير Y ،

و A هو المتوسط الفرضي للمتغير X، أما 'A فهو الوسط الفرضي للمتغيرY.

كما يجب التنبيه بأنه ليس شرط أن تتساوى كل من C_y و C_x، وأيضا ليس شرط أن يتساوى كل من A و 'A .

مثال: لدينا الجدول المزدوج الآتي الذي يمثل علامات أربعين طالبا في مقياسي الإحصاء (X) والرياضيات (Y).

الجدول رقم (59)

الإحصاء / الرياضيات	[0-4 [	[4-8[	[8–12 [	[12-16[	[16-20 [
[0 - 5 [	2	1			
[5 - 10 [	3	5	7		
[10 - 15 [		4	5	3	
[15 - 20 [			4	4	2

3.4- بعض مقاييس الدلالة الإحصائية

عندما يستخرج الباحث في العلوم الإنسانية نتيجته فإنه يكون في حالة شك من صدقيتها وثباتها ودلالتها الإحصائية، لذلك يلجأ إلى مقاييس الدلالة الإحصائية لتبيان إلى أي حد يستطيع الباحث أن يتأكد من ثبات نتائجه وإلى أي حد يستطيع إرجاعها إلى عامل الصدفة وحده.
وسنتناول هنا مقياس كثير الاستخدام هو مقياس كا ^{2} ثم طريقة الكشف عن العلاقات من خلال النسب المئوية.

ولكن تجدر الإشارة قبل دلك إلى بعض قواعد اختبار الفروض الإحصائية ومنها تحديد مستوى الدلالة الإحصائية (مستوى المعنوية) الضروري لمقياس كا ^{2} وغيره من مقاييس الدلالة الإحصائية:

1.3.4- بعض قواعد اختبار الفروض الإحصائية

يذكر "فتحي عبد العزيز أبو راضي" (1998، 281-288) أنه يمكن حصر قواعد اختبار الفروض الإحصائية في أربع: وضع الفروض (الفرض الصفري أو فرض العدم والفرض البديل)، تحديد مستوى الدلالة الإحصائية، تحديد التوزيع النظري أو الاحتمالي للإحصائية المختبرة، وأخيرا حساب قيمة إحصائية الاختبار واتخاذ قرار قبول الفرض الصفري أو رفضه. وسنكتفي هنا بوصف العنصرين الأولين:

- التعبير عن الفرضيات رياضيا والحكم على مدى صحتها، وذلك باللجوء إلى مقياس نظري مجدول أو متوقع مسبقا مقابل المقياس المحسوب من البيانات المشاهدة من العينة. فمثلا إذا افتراضنا أن متوسط علامات طلبة العلاقات العامة مساو لمتوسط علامات طلبة الصحافة المكتوبة، فإنه يمكن التعبير عن هذا الفرض بالقول أن الفرق بين متوسط العلامات في التخصصين يساوي صفرا أو منعدم. ويسمى حينها بالفرض الصفري ويرمز له بالرمز (H_0). فإذا جاءت نتائج اختبار العينة مؤيدة له قبلناه وإلا رفضناه، وحينها نضطر إلى قبول الفرض البديل والذي يرمز له بالرمز (H_1). وفي مثالنا السابق سيقر نص الفرض البديل بوجود فرق بين متوسطي علامات التخصصين.

- يتم تحديد مستوى الدلالة الإحصائية لمعرفة نسبة الخطأ أو الاحتمال المسموح به لقبول الفروض الإحصائية أو رفضها. ولتحديدها يجب الأخذ بعين الاعتبار نسبة احتمال الخطأ في قبول أو رفض الفرضيات. فمثلا إذا قررنا قبول حدوث الخطأ (في الفرضين " H_1، H_0 " أو في البيانات

النظرية أو المتوقعة والعينية أو المشاهدة للصدفة مثلا) في خمس مرات كل مائة مرة، فإن هذا يعني أنه في عدد كبير من التجارب نتوقع أن نرفض الفرضية الصفرية وهي في الواقع صحيحة في 5% من المرات، وبذلك يكون الحد الأقصى الذي قررنا قبوله لاحتمال وقوع الخطأ هو 0.05، ويرمز لهذا الاحتمال بالرمز ($\propto$) وتسمى قيمة الاحتمال ($\propto$) بمستوى الدلالة الإحصائية. وقد جرت العادة على اعتماد مستويي 0.05 و 0.01 لقبول صحة الفرضية الصفرية.

2.3.4- مقياس مربع كاي (اختبار كا^2)

عندما تكون بيانات المتغيرات وصفية (الحالة الزواجية، التعليمية...) أو بعضها وصفيا وبعضها الآخر كميا يستعمل اختبار كا^2. ولهذا الاختبار عدة استخدامات منها:

أولا) اختبار مدى معنوية الفرق بين التوزيع المشاهد والتوزيع المتوقع:

إذا كان الفرق بين التوزيع المشاهد (المحسوب من عينة معينة) والتوزيع المتوقع (النظري) غير معنوي، نستنتج في هذه الحالة أن الاختلاف الظاهري بين التوزيعين وهو اختلاف غير حقيقي يرجع إلى أخطاء المعاينة.

- ولاستخدام هذا الاختبار نقوم بحساب كا ^{2} كما يأتي:

$$\text{كا}^2 \text{ المحسوبة} = \text{مجـ} \frac{(\text{التكرار المشاهد} - \text{التكرار المتوقع})^2}{\text{التكرار المتوقع}}$$

$$= \text{مجـ} \frac{(\text{س} - \text{ت})^2}{\text{ت}}$$

ثم نقارن كا ^{2} المحسوبة بـ كا ^{2} الجدولية (من الجداول الإحصائية) بدرجات حرية (ن- 1) وعند مستوى معنوية (دلالة) معين (غالبا 0.05)، ثم نحسب الفرق بين المشاهد والمتوقع. فإذا كانت كا^2 المحسوبة أكبر من كا ^{2} الجدولية يكون الفرق معنويا وذا دلالة إحصائية (نرفض الفرضية الصفرية ونقبل الفرضية البديلة =H_1). أما إذا كانت كا ^{2} المحسوبة أقل من كا ^{2} الجدولية فالفرق يكون غير معنوي (نقبل الفرضية صفرية= H_0).

ومن الملاحظ أن كا ^{2} لا تساوي صفرا إلا إذا كان كل مربع من المربعات التي بها تكرار مشاهد = صفرا، وهذا يعني تساوي التكرار المشاهد والمتوقع لكل مربع أو خلية. ومن جانب آخر فإن قيمة كا^2 المحسوبة تكون كبيرة إذا كان الفرق بين التكرارات المشاهدة والمتوقعة كبيرا.

مثال: إذا أردنا معرفة ما إذا كان هناك فرق بين الطلبة في تفضيلهم للتخصصات المختلفة في قسم علوم الإعلام والاتصال، وسحبنا عينة عشوائية حجمها 720 طالبا وطالبة من الجذع المشترك وكان التوزيع التكراري لتفضيلاتهم كما يأتي:

الجدول (60)

التخصص المفضل	صحافة مكتوبة	سمعي-بصري	مسرح	اتصال مؤسسة	المجموع
التكرار	210	190	170	150	720

فهل هذه البيانات تتفق مع الافتراض القائل بعدم وجود فروق أصلية بين تفضيلات الطلبة وذلك عند مستوى معنوية 0.05 ؟

الحل:

الجدول (61)

التخصص المفضل	صحافة مكتوبة	سمعي-بصري	مسرح	اتصال مؤسسة	المجموع
التكرار المشاهد	210	190	170	150	720
التكرار المتوقع	180	180	180	180	720

الفرضية الصفرية (H_0) = لا يوجد فرق (معنوي) ذو دلالة إحصائية بين التكرار المشاهد والمتوقع.

الفرضية البديلة (H_1) = يوجد فرق ذو دلالة إحصائية بين التكرار المشاهد والمتوقع.

التكرار المتوقع: 720 طالبا ÷ 4 تخصصات = 180 طالبا في كل تخصص (تبعا لافتراض تعادل أو تساوي التوزيع).

- كا^2 المحسوبة = مجـ $\frac{(ش-ت)^2}{ت}$

$$= \frac{(180-210)^2}{180} + \frac{(180-190)^2}{180} + \frac{(180-170)^2}{180} + \frac{(180-150)^2}{180}$$

$$= \frac{900}{180} + \frac{100}{180} + \frac{100}{180} + \frac{900}{180} = 5 + 0.56 + 0.56 + 5 = 11.12$$

كا^2 المستخرجة من الجداول الإحصائية بمستوى دلالة 0.05، وبدرجات حرية (ن−1) أي (4−1=3) =7.81

ونظرا لأن قيمة كا ^{2} المحسوبة أكبر من كا ^{2} المستخرجة من الجداول نرفض الفرض الصفري ونقبل الفرض البديل بوجود فرق معنوي بين التكرارات المشاهدة والمتوقعة.

ثانيا) اختبار العلاقة بين خاصيتين لجدول مزدوج دون علم مسبق بالتكرار المتوقع:

يتم اختبار العلاقة بتطبيق قانون كا ^{2} نفسه بعد حساب التكرار المتوقع: فإذا كان لدينا جدولا به "ل" صفا، و"م" عمودا، ونريد اختبار العلاقة بين الصفة أو الخاصية التي قسمت على أساسها الصفوف، والخاصية التي قسمت على أساسها الأعمدة، فإننا نستخدم اختبار كا ^{2} لهذا الغرض. مع ملاحظة أن درجات الحرية في مثل هذا الجدول هي (ل− 1) (م−1)، وسوف نوضح استخدامات كا^2 في حالتين:

الحالة الأولى: إذا كان الجدول التكراري المزدوج يتكون من أربع خلايا أي 2×2 وفي هذه الحالة يسمى بجدول **الاقتران**.

الحالة الثانية: إذا كان الجدول التكراري المزدوج يتكون من أكثر من أربع خلايا مثل 2×3 أو 3×3 أو أكبر من ذلك. وفي هذه الحالة يسمى بجدول **التوافق**.

أولا: إذا كان الجدول يتكون من أربع خلايا (جدول اقتران:2×2):

- حساب كا^{2}بعد حساب التكرار المتوقع:

مثال: إذا كنا نريد معرفة هل توجد علاقة بين مشاهدة أفلام العنف والسلوك العدواني أم لا، فإننا نأخذ مجموعة من الذين تم تعرضهم لمشاهد العنف وأخرى لم تتعرض لذلك، ثم نلاحظ طبيعة سلوك كل مجموعة منهما.

فإذا كان لدينا مجموعة من الأشخاص الذين شاهدوا أفلام عنف حجمها 460 منهم 7 أفراد صدرت عنهم سلوكات عنيفة، ولدينا مجموعة أخرى من الأشخاص لم يتعرضوا لمشاهد العنف حجمها 340 منهم 13 صدرت عنهم سلوكات عنيفة،

فالمطلوب: معرفة هل توجد علاقة بين مشاهدة العنف وممارسته أم لا وذلك عند مستوى الدلالة 0.05

الجدول (62)

السلوك العدواني / المشاهدة	لم يعتدوا	اعتدوا	المجموع
شاهدوا	453	7	460
لم يشاهدوا	327	13	340
المجموع	780	20	800

<u>الحل</u>: نحسب التكرارات المتوقعة لكل خلية من الخلايا الأربع:

التكرار المتوقع للخلية الأولى= $\frac{\text{مجموع. الصف} \times \text{مجموع. العمود}}{\text{المجموع. الكلي}}$

$= \frac{\text{مج.ل}_1 \times \text{مج. م}_1}{\text{مج.ك}} = \frac{780 \times 460}{800} = 448.5$

التكرار المتوقع للخلية الثانية = $\frac{\text{مج.ل}_1 \times \text{مج. م}_1}{\text{مج.ك}} = \frac{20 \times 460}{800} = 11.5$

التكرار المتوقع للخلية الثالثة = $\frac{\text{مج.ل}_1 \times \text{مج. م}_1}{\text{مج.ك}} = \frac{780 \times 340}{800} = 331.5$

التكرار المتوقع للخلية الرابعة = $\frac{\text{مج.ل}_1 \times \text{مج. م}_1}{\text{مج.ك}} = \frac{20 \times 340}{800} = 8.5$

الجدول (63)

السلوك العدواني / المشاهدة	لم يعتدوا	اعتدوا	المجموع
شاهدوا	453 (448.5)	7 (11.5)	460
لم يشاهدوا	327 (331.5)	13 (8.5)	340
المجموع	780	20	800

وبمعرفة التكرار المتوقع لإحدى الخلايا يمكن الحصول على التكرار المتوقع للخلايا الثلاث الأخرى عن طريق طرح التكرار المتوقع للخلية من مجموع الصف أو مجموع العمود لمعرفة التكرار المتوقع للخلية التي تشترك مع الخلية الأولى في الصف أو في العمود.

ثم نحسب كا^2 المحسوبة كالآتي:

كا^2 المحسوبة = مج $\frac{(\text{ن}-\text{ت})^2}{\text{ت}} = \frac{(448.5-453)^2}{448.5} + \frac{(11.5-7)^2}{11.5} + \frac{(331.5-327)^2}{8.50}$

$+ \frac{(8.50-13)^2}{8.50} = 0.05 + 1.76 + 0,060 + 2.37 = 4.25$

كا^2 الجدولية بدرجات حرية (ل-1) (م-1)= (2-1) (2-1)= 1، عند مستوى دلالة 0.05

كا^2 (0.05، 1) = 3.84

ونظرا لأن كا ^{2} المحسوبة أكبر من كا ^{2} المستخرجة من الجداول الإحصائية، فإننا نرفض الفرض الصفري ونقبل الفرض البديل، وهو أنه يوجد فروق ذات دلالة معنوية بين الفئتين من حيث القيام بسلوك عدواني. أي توجد علاقة بين مشاهدة العنف والسلوك العدواني.

- حساب كا^2 دون حساب التكرار المتوقع:

في جدول الاقتران هناك طريقة أخرى مباشرة لحساب كا^2 من البيانات دون الحصول على التكرار المتوقع للخلايا وذلك وفقا للصيغة الآتية:

$$\text{كا}^2 = \frac{\text{ن}\,(\text{أ}\times\text{د} - \text{ب}\times\text{ج})^2}{\text{مج.ل}_1 \times \text{مج.ل}_2 \times \text{مج. م}_1 \times \text{مج. م}_2}$$

حيث ن هو المجموع الكلي

أ : التكرار المشاهد في الخلية الأولى

ب: التكرار المشاهد في الخلية الثانية

ج: التكرار المشاهد في الخلية الثالثة

د : التكرار المشاهد في الخلية الرابعة

مج ل$_1$: مجموع الصف الأول

مج ل$_2$: مجموع الصف الثاني

مج م$_1$: مجموع العمود الأول

مج م$_2$: مجموع العمود الثاني

ويمكن توضيح ذلك من خلال جدول الاقتران الآتي:

الجدول (64)

السلوك العدواني / المشاهدة	لم يعتدوا	اعتدوا	المجموع
شاهدوا	أ	ب	مج ل^1
لم يشاهدوا	ج	د	مج ل^2
شاهدوا	مج م^1	مج م^2	ن

وبتطبيق ذلك على المثال السابق يتضح أن:

كا^2 المحسوبة = $\frac{800[(327\times7)-(13\times453)]^2}{20\times780\times340\times460}$

$$= \frac{800(2289-5889)^2}{2439840000} = \frac{10368000000}{2439840000} = 4.25$$

وهي النتيجة السابقة نفسها.

<u>ثانيا) الجدول المزدوج يتضمن أكثر من أربع خلايا (جدول توافق: 2× 3 أو 3 × 3...):</u>

إذا كان لدينا ظاهرتان وكانت إحداهما تتضمن أكثر من متغيرين، فإن الجدول التكراري المزدوج في هذه الحالة يسمى بجدول التوافق حيث يتضمن أكثر من أربع خلايا.

<u>مثال</u>: إذا أردنا معرفة العلاقة بين المستوى التعليمي (ابتدائي، متوسط، عالي) والعمل (عامل، عاطل)، تبعا لتكرارات الجدول المزدوج سداسي الخلايا الآتي:

الجدول (65)

العمل / التعليم	عامل	عاطل	المجموع
ابتدائي	أ 35	ب 45	80
متوسط	ج 120	د 60	180
عالي	هـ 85	و 15	100
المجموع	240	120	360

المطلوب: هل توجد علاقة بين الظاهرتين أم لا عند مستوى الدلالة 0.05 ؟

<u>الحل</u>:

الفرضية الصفرية (السلبية)= H_0 = لا يوجد فرق ذو دلالة إحصائية بين الظاهرتين.

الفرضية البديلة (الإيجابية)= H_1 = يوجد فرق ذو دلالة إحصائية بين الظاهرتين.

نحسب التكرار المتوقع لكل خلية= <u>مجـ الصف الذي به الخلية×مجـ العمود الذي به الخلية</u>

المجموع الكلي للتكرارات

التكرار المتوقع للخلية أ $= \frac{240 \times 80}{360} = 53.3$

التكرار المتوقع للخلية ب $= \frac{120 \times 80}{360} = 26.7$

التكرار المتوقع للخلية ج $= \frac{240 \times 180}{360} = 120$

التكرار المتوقع للخلية د $= \frac{120 \times 180}{360} = 60$

التكرار المتوقع للخلية هـ $= \frac{240 \times 100}{360} = 66.7$

التكرار المتوقع للخلية و $= \frac{120 \times 100}{360} = 33.3$

كا^2 المحسوبة = مجـ $\frac{(ق-ت)^2}{ت} = \frac{(53.3-35)^2}{53.3} + \frac{(26.7-54)^2}{26.7} + \frac{(120-120)^2}{120} + \frac{(60-60)^2}{60} +$
$\frac{(66.7-85)^2}{66.7} + \frac{(33.3-15)^2}{33.3} =$ 6.28 + 12.54 + صفر + صفر + 5.02 + 10.06 = 33.9

كا^2 الجدولية بدرجات حرية (ل−1) (م−1)= (2−1) (3−1)= 1×2 =2، عند مستوى دلالة 0.05 كا^2 (0.05، 2) = 5.99

ولأن كا^2 المحسوبة أكبر من كا^2 المستخرجة من الجداول الإحصائية، فإننا نرفض الفرض الصفري ونقبل الفرض البديل القائل بوجود علاقة بين الظاهرتين.

3.3.4− اختبار الفرق بين المتوسطات (اختبار ستيودنت:"ت" T):

لقياس مدى معنوية الفرق بين متوسطي عينتين معروف الانحراف المعياري لكل منهما، قدّم "و. س. جوسي" (W.S. Gosset) هذه المعادلة $t = \frac{\overline{X} - \mu}{s / \sqrt{N-1}}$ ، التي عرفت لاحقا باسم توزيع "ت" لـ "ستيودنت" (Student)، وهو الاسم المستعار الذي كان يكتب به.

إنه تصميم إحصائي مبني على أساس المقارنة بين متوسطي عينتين وذلك في ضوء تقدير الخطأ المعياري للفرق بين المتوسطين. أي أن:

اختبار "ت" = الفرق بين متوسطي العينتين على الخطأ المعياري لهذا الفرق.

ويكون الانحراف المعياري "ع" (الخطأ المعياري) للفرق بين المتوسطين عبارة عن الجذر التربيعي للتباين (ع^2): $ع_{\overline{س}_1 - \overline{س}_2} = \sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}$

وبناء عليه يرمز للقيمة المعيارية لـ "ت" في حالة العينات الصغيرة (أقل من 30 مفردة) كالآتي: ت= $\frac{\overline{س}_1 - \overline{س}_2}{\sqrt{\frac{ع_1^2}{ن_1-1} + \frac{ع_2^2}{ن_2-1}}}$ أو كما مر معنا: $t = \frac{\overline{X} - \mu}{s / \sqrt{N-1}}$

وهذه القيمة المعيارية لها توزيع احتمالي "ت" بدرجات الحرية ($ن_1+ن_2 - 2$).

كم يمكن صياغة قانون "ت" كما يلي: $\frac{\overline{س}_1 - \overline{س}_2}{\sqrt{\frac{ن_1 ع_1^2 + ن_2 ع_2^2}{ن_1+ن_2-2}\left(\frac{1}{ن_1} + \frac{1}{ن_2}\right)}}$

أما في حالة العينات الكبيرة (>30) فحساب "ت" يختلف قليلا، لأن العينات الصغيرة عادة ما تعطي معلومات عن المجتمع أقل دقة، وصيغته:

ت= $\frac{\overline{س}_1 - \overline{س}_2}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$ والتي تختزل عند تساوي عدد مفردات المجتمَعين هكذا: = $\frac{\overline{س}_1 - \overline{س}_2}{\sqrt{\frac{ع_1^2 + ع_2^2}{ن_1}}}$

كما يمكن يستخدم اختبار "ت" عندما لا تكون متغيرات الدراسة كمية، أي في حالة استحالة إيجاد المتوسط والانحراف المعياري للعينتين. ويكون ذلك عن طريق النسب المئوية وباستخدام القانون التالي: ت= $\frac{ب_1 - ب_2}{\sqrt{ل(1ل)\left(\frac{1}{ن_1} + \frac{1}{ن_2}\right)}}$

حيث ب = النسبة، ن= حجم العينة

ل= حجم الخطأ المعياري= $\frac{ن_1 ر_1 + ن_2 ر_2}{ن_1 + ن_2}$

وأما عن أهم ميزات اختبار "ت" فتتمثل في كونه يُستخدم مع جميع الحالات مهما كان حجم العينة. ولكنه كاختبار إحصائي خاص بالعينات يفترض في أغلب الأحيان أن تكون بيانات العينة موزعة توزيعا معتدلا.

ولأن بيانات العينة، مهما كانت متماثلة، لن تعكس تماما خصائص المجتمع المستخرجة منه، فإنه توجد بعض الفروق بين خصائص العينة ومعالم مجتمعها والتي يمكن تقديرها على أساس احتمالات أخطاء معينة (درجات أو مستويات المعنوية والثقة).

ولأن توزيع قيم "ت" يعتمد على حجم العينة وبالتالي على درجات الحرية، فإنه يلزم عمل توزيع احتمالي لكل درجة من درجات الحرية. وبالطبع هناك جداول تبين قيم "ت" عند مستويات معنوية مختلفة.

فمثلا، قد نلجأ عند اختبار مدى فاعلية عامل معين إلى سحب عينتين، الأولى قبل تأثير هذا العامل والثانية بعده. ثم نختبر ما إذا كان الفرق بين متوسطي العينتين القبلية والبعدية فرقا دالا (Significant) إحصائيا أو لا. فإذا كان الفرق دالا نستنتج فاعلية العامل وإلا فعدم فاعليته لأن الفرق قد يعود للصدفة المطلقة أو لخطأ المعاينة.

– مثال1 (حجم عيني كبير> 30 وعدد مفردات المجتمعين متساو):

العينة الأولى: ن (حجم العينة)= 100، $\overline{س}$ = 29، ع= 5

العينة الثانية: ن = 100، $\overline{س}$ = 28، ع= 6

المطلوب: اختبار مدى صحة الفرض الصفري (H_0) القائل أن متوسطي المجتمعين متساويان ($\overline{س}_1=\overline{س}_2$)، مقابل الفرض البديل ($H_1$) القائل أن متوسط مجتمع العينة الأولى أكبر ($س_1>س_2$) أو أن الفرق بينهما دال، وذلك عند مستوى المعنوية (الدلالة) 0.05

الحل: ت= $$\frac{\overline{س}_1-\overline{س}_2}{\sqrt{\frac{ع_1^2}{ن_1}+\frac{ع_2^2}{ن_2}}} = \frac{29-28}{\sqrt{\frac{(5)^2}{100}+\frac{(6)^2}{100}}} = \frac{1}{\sqrt{\frac{25}{100}+\frac{36}{100}}} = \frac{1}{\sqrt{0.61}} = \frac{1}{0.78} = 1.28$$

أو بتطبيق المعادلة الخاصة بتساوي مفردات المجتمعين كما هو في هذا المثال:

$$\frac{\overline{س}_1-\overline{س}_2}{\sqrt{\frac{ع_1^2+ع_2^2}{ن_1}}} = \frac{1}{\sqrt{\frac{25+36}{100}}} = \frac{1}{\sqrt{0.41}} = \frac{1}{0.78} = 1.28$$

(وهي النتيجة نفسها بالطريقة المختزلة بسبب تساوي عدد أفراد المجتمعين=100)

كيفية إيجاد المنطقة الحرجة وحساب معنوية الفرق:

– نحسب درجة الحرية ($ن_1+ن_2-2$ = 100+100−2= 198)

– فإذا اعتمدنا مستوى المعنوية أو الثقة المحدد سلفا بـ 0.05 (وهو الشائع) واستعنا بالاختبار ثنائي الطرف من جدول توزيع "ت" الكائن بالملحق، لوجدنا أنه بالنسبة لدرجة الحرية 198 (أي أكثر

من 120 في الجدول) نحتاج إلى "ت" = 1.96 أو أكثر ليكون الفرق دالا عند مستوى 0.05. بينما لا نحتاج في الاختبار أحادي الطرف سوى إلى ت=1.65 أو أكثر ليكون الفرق دالا.

وبما أن قيمة "ت" المحسوبة = 1.28 في مثالنا ، وهي أقل من قيمة "ت" الجدولية في الاختبارين، فهي إذن لا تقع في منطقة رفض الفرضية (المنطقة الحرجة)، وعليه لا يمكن رفض فرضيتنا الصفرية. ونستنتج من ذلك أن الفرق $\bar{س}_1 - \bar{س}_2$ هو فرق يرجع للصدفة أو لأخطاء المعاينة وليس فرقا جوهريا عند مستوى الدلالة 0.05

–كما يمكن أن نستنتج ذلك بطريقة أخرى: بما أن القيمة 1.28 تقع في الجدول تحت مستوى الثقة 0.10 و0.20 (في الاختبارين أحادي وثنائي الطرف على التوالي)، فهذا يعني أيضا أن الفرق ليس معنويا، لأن الفرق المعنوي في مثالنا يكون عندما يقع اختبار "ت" في مستوى المعنوية: 0.05

– مثال 2 (حجم عيني كبير> 30 وعدم تساوي عدد مفردات المجتمعين):

العينة الأولى: ن (حجم العينة)= 60، $\bar{س}$ = 13.87، ع= 4.4

العينة الثانية: ن = 70، $\bar{س}$ = 14.97، ع= 4.28

المطلوب: مدى معنوية الفرق بين العينتين، عند مستوى الثقة 0.05

الحل:

$$ت = \frac{\bar{س}_1 - \bar{س}_2}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}} = \frac{14.97 - 13.87}{\sqrt{\frac{(4.4)^2}{60} + \frac{(4.28)^2}{70}}} = \frac{1.1}{\sqrt{\frac{19.36}{60} + \frac{18.31}{70}}}$$

$$= \frac{1.1}{\sqrt{0.26 + 0.32}} = \frac{1.1}{\sqrt{0.58}} = \frac{1.1}{0.76} = 1.44$$

– درجة الحرية= 60+70−2= 128

– القيمة المقابلة لدرجة الحرية عند مستوى0.05 في الجدول:

"ت" = 1.98

ما دامت قيمة "ت" المحسوبة (1.44) أقل من قيمتها الجدولية فالفرق غير دال إحصائيا لذلك نقبل الفرضية الصفرية، أي تساوي المتوسطين.

–كما يمكن أن نستنتج ذلك بطريقة أخرى: ما دامت القيمة 1.44 تقع في الجدول عند مستوى الثقة 0.1 (في الاختبار أحادي الطرف)، فهذا يعني أيضا أن الفرق ليس معنويا، لأن المناسب في مثالنا هو مستوى الثقة 0.05

– وتجدر الإشارة في الأخير إلى أن هناك من يلجأ إلى اختبار "ت" للكشف عن مدى معنوية الفرق بين الوسط الحسابي للعينة والوسط الحسابي لمجتمع الدراسة للتأكد من صدق التمثيل العيني العشوائي (إحسان محمد حسن وعبد الحسين زيني: 1982، 175–176). فيعتمد على الفرضية الصفرية القائلة بعدم وجود فرق معنوي بين العينة ومجتمع الدراسة على جميع مستويات الثقة: 95% – 99%.

ولاختبار هذا الفرق يطبق المعادلة التالية:

$$\text{"ت"} = \frac{\overline{\text{س}}_{\text{ع}} - \overline{\text{س}}_{\text{م}}}{\text{ع}_{\text{ع}}\,\text{ن}_{\text{م}}}$$

حيث: "$\text{ع}_{\text{ع}}\,\text{ن}_{\text{م}}$" هي الوسط الحسابي للخطأ المعياري للعينة = $\frac{\text{ع}_{\text{م}}}{\sqrt{\text{ن}_{\text{ع}}}}$

$\overline{\text{س}}_{\text{ع}}$ = الوسط الحسابي للعينة، $\overline{\text{س}}_{\text{م}}$ = الوسط الحسابي لمجتمع الدراسة

4.3.4– الكشف عن العلاقة من خلال النسب المئوية

تعتمد الكثير من البحوث، خاصة التي تتطرق لمجالات قياس الرأي العام والاتجاهات، على النسب المئوية.

كما أن كثيرا من النتائج التي يتم عرضها في بعض هذه البحوث لا تكون إلا على صورة نسب مئوية لمن أجابوا بنعم على سؤال ما في إحدى المجموعات ولمن أجابوا بنعم على السؤال نفسه في مجموعة أخرى: فقد تكون المقارنة مثلا بين النسب المئوية للذكور والنسب المئوية للإناث فيما يختصص بمتغير من المتغيرات. وأحيانا تكون المقارنة داخل المجموعة الواحدة بين من أجاب بنعم على السؤال الأول في إحدى الاستمارات ومن أجاب بنعم على السؤال الثاني في الاستمارة نفسها، ويكون الهدف في البحث معرفة الدلالة بين النسبتين.

وفي حالة المقارنة بين النسب في مجموعتين يكون حساب الدلالة الإحصائية للنسب غير المرتبطة، وفي حالة المقارنة بين النسب داخل مجموعة واحدة يكون حساب الدلالة الإحصائية للنسب المرتبطة.

أولا) حساب الدلالة الإحصائية للنسب غير المرتبطة (النسبة الحرجة: Critical Ratio):

ويتم حسابها بثلاث كيفيات مختلفة الخطوات. ويمكن تشخيص ذلك من خلال المثال الآتي:

طبقت استمارة على مجموعتين من الطلبة إحداهما من الناجحين وكان عددهم 100 والأخرى من الراسبين وكان عددهم 50. فأجاب 45 من الناجحين بنعم على سؤال يتعلق بكفاءة طرق التدريس، كما أجاب 20 من الراسبين بنعم على السؤال نفسه. فهل هناك فرقا له دلالة إحصائية بين من أجابوا بنعم في المجموعتين على هذا السؤال؟

- **الطريقة الأولى:** وتتمثل خطواتها ومعادلاتها فيما يأتي:

1- نحسب النسبة المئوية لمن أجابوا بنعم في المجوعتين على النحو الآتي:

- النسبة المئوية لمن أجابوا بنعم على السؤال من الناجحين: = $\frac{45}{100} \times 100 = 45$ %
- النسبة المئوية لمن أجابوا بنعم على السؤال من الراسبين: = $\frac{20}{50} \times 100 = 40$ %

2- نحصل على النسبة المئوية 1 (أي P 1) حسب القانون الآتي:

$$P1 = \frac{\text{ن}1 \times \text{النسبة. المئوية. للمجموعة}1 + \text{ن}2 \times \text{. النسبة. المئوية. للمجموعة}2}{\text{ن}1 + \text{ن}2}$$

وهي في المثال= $\frac{40 \times 50 + 45 \times 100}{50 + 100} = \frac{6500}{150} = 43.3$ % تقرب إلى 43%

3- نحصل على النسبة المئوية 2 (أي P2) حسب القانون الآتي:

$$P2 = 100\% - P1$$

$$P2 = 100 - 43.3 = 56.7 \%$$ تقرب إلى 57 %

4- نحسب الدلالة الإحصائية لـ P 1 P2 حسب القانون الآتي:

$$P1\ P2 = \sqrt{P1P2\left[\frac{1}{N^1} + \frac{1}{N^2}\right]}$$

بحيث أن ن$_1$ (N_1) = عدد أفراد المجموعة الأولى، ون$_2$ (N_2) = عدد أفراد المجموعة الثانية

وبتطبيق القانون على المثال السابق نحصل على:

$$P1\ \ P2=\sqrt{43\times 57\left[\frac{1}{100}+\frac{1}{50}\right]}$$

$$P1\ \ P2=\sqrt{2451(0.02+0.01)}$$

$$P1\ \ P2=\sqrt{73.53}$$

$$P1\ \ P2=8.57$$

5- يتم بعد ذلك حساب الفرق بين النسبة المئوية "أ" والنسبة المئوية "ب" وبتطبيق ذلك على المثال السابق (أ، ب) تكون النتيجة كالآتي:

الفرق بين النسبتين المئويتين أ، ب من الخطوة 1 = 45-40 = 5.

6- يتم بعد ذلك قسمة الناتج من الفرق بين النسبتين المئويتين (في الخطوة رقم 5) على الناتج في P1 P2 (الخطوة رقم 4) للحصول على النسبة الحرجة (Critical Ratio : CR) وذلك حسب القانون الآتي:

النسبة الحرجة (ن.ح.) أو $CR = \dfrac{\text{الفرق بين النسبتين أ،ب}}{P1P2}$

وفي المثال السابق نجد أن قيمة CR كما يلي: $CR = \dfrac{5}{8.57} = 0.58$

7- تعتبر نتيجة الخطوة السابقة (0.58) المتدنية غير دالة عند المستويين (0.01 و 0.05)، لأن القاعدة تقول أن:

إذا بلغت النسبة الحرجة 1.96 يكون هناك فرق معنوي له دلالة عند مستوى 0.05، وإذا بلغت 2.58 يكون هناك أيضا فرق معنوي عند مستوى 0.01 ، أما ما دون ذلك فليس له فرق معنوي، وأن الفرق الظاهر يرجع إلى الصدفة وظروف إجراء البحث.

- **الطريقة الثانية**، وتتمثل في حساب النسبة الحرجة (ن.ح.) (CR) تبعا للقانون الآتي:

$$CR=\frac{P1-P2}{\sqrt{\frac{P1(100-P2)}{N_1}+\frac{P2(100-P1)}{N_2}}}$$

وبتطبيق هذه المعادلة على المثال السابق نحصل على ما يأتي:

$$CR=\frac{45-40}{\sqrt{\frac{45(100-40)}{45}+\frac{40(100-45)}{20}}}=\frac{5}{\sqrt{\frac{45(55)}{45}+\frac{40(60)}{20}}}=\frac{5}{\sqrt{55+120}}$$

$$CR = \frac{5}{\sqrt{175}} = \frac{5}{13.22} = 0.38$$

وهي غير دالة إحصائيا حسب الخطوة رقم 7 في الطريقة الأولى.

- **الطريقة الثالثة** وتتمثل خطواتها ومعادلاتها فيما يأتي:

1- نسبة من أجاب بنعم من الناجحين: $= \frac{45}{100} \times 100 = 45 \% = 0.45$

2- نسبة من أجاب بنعم من الراسبين: $= \frac{20}{50} \times 100 = 40 \% = 0.40$

3- عدد من أجاب بنعم من الراسبين والناجحين على المجموع الكلي=

$$\frac{\text{عدد نعم في}(1) + \text{عدد نعم في}(2)}{ن_1 + ن_2} = \frac{45+20}{100+50} = \frac{65}{150} = 0.43$$

4- الفرق بين واحد صحيح والنسبة الكلية = $1 - 0.43 = 0.57$

5- الخطأ المعياري لتقدير التباين: $\sqrt{\frac{20 \times 45 \times 0.57}{50 \times 100}} = \sqrt{\frac{513}{5000}} = \sqrt{0.1026} = 0.32$

6- القيمة الناتجة: $\frac{\%(2) - \%(1)}{\text{الخطأ المعياري}} = \frac{0.40 - 0.45}{0.32} = \frac{0.05}{0.32} = 0.16$

وهي غير دالة حسب الخطوة رقم 7 في الطريقة الأولى.

ثانيا) حساب الدلالة للنسبة المئوية المرتبطة

عند حساب دلالة النسب المئوية داخل المجموعة الواحدة بالنسبة لمتغير من المتغيرات تتبع عدة خطوات يمكن تشخيصها من خلال المثال الآتي:

أجابت مجموعة متكونة من 250 طالبا على السؤالين الآتيين الواردين في إحدى الاستمارات:

س(1): هل تتغيب أحيانا عن الدراسة،

أجاب 150 بنعم

أجاب 100 بلا

س(2): هل تخاف من الامتحانات الشفوية؟

أجاب 125 بنعم

أجاب 125 بلا

الحل:

1- يتم وضع نتائج السؤالين في مجاميع أعمدة وصفوف الجدول التالي للتبسيط:

الجدول (66)

س1 / س2	لا	نعم	مج
نعم	25	100	125
لا	75	50	125
مج	100	150	250

وقد تم توزيع النتائج الداخلية في المربعات من مجاميع الأعمدة والصفوف (إجابات س1 مجاميعه عمودية في أسفل الجدول وإجابات س2 أفقية على يساره) كالآتي:

أ- طرح مجموع العمود الأول من مجموع الصف الأول للحصول على القيمة الأولى بالصف الأول: 125- 100 = 25

ب- طرح القيمة التي تم الحصول عليها من الخطوة السابقة من مجموع الصف الأول للحصول على من أجابوا بنعم على السؤالين: 125- 25 = 100

ج- طرح القيمة الناتجة في الخطوة الأولى من مجموع العمود الأول للحصول على من أجابوا بلا على السؤال الأول وبلا على السؤال الثاني: 100 - 25 = 75

د- طرح القيمة الناتجة في الخطوة الثانية من مجموع العمود الثاني للحصول على من أجابوا بنعم على السؤال الأول وأجابوا بلا على السؤال الثاني: 150 - 25 = 50

2- يتم حساب النسبة المئوية للنتائج الواردة في الجدول السابق بالكيفية الآتية:

الجدول (67)

س1 / س2	لا	نعم	مج
نعم	10% (ب)	40% (أ)	50 %
لا	30% (د)	20% (ج)	50 %
مج	40%	60%	100 %

3- يتم حساب معامل ارتباط فاي (Ph C) من الجدول السابق وقيمته في المثال السابق = 0.41 . مع العلم أن معامل فاي يعتمد في حسابه على التكرارات الموجودة في جدول الانتشار (المزدوج: 2 × 2، الذي مر معنا في معامل الارتباط) ومعادلته كالآتي: = $\frac{AD-BC}{\sqrt{EFGH}}$

4- يتم حساب النسب المئوية للإجابات كما يأتي:

أ- النسبة المئوية (1) لمن أجاب بنعم على السؤال الأول = $\frac{150}{250}$ × 100 = 60 %

ب- النسبة المئوية (2) لمن أجاب بنعم على السؤال الثاني= $\frac{125}{250}$ × 100 = 50 %

5- يتم عمل تقدير للنسبة بحساب المتوسط للنسبة (1)، (2) في الخطوة السابقة كالآتي:

متوسط النسبة = 60 + 50 = 110 – 2 = 55 (النسبة أ).

6- يتم طرح متوسط النسبة من 100 = 100– 55 = 45 (النسبة ب).

7- يتم حساب الفرق بين النسبتين (1)، (2) في الخطوة رقم (4) = 60–50 = 10.

8- تطبق معادلة النسبة المئوية الآتية:

دلالة النسبة المئوية = $CP = \frac{P1-P2}{\sqrt{\frac{2(P1\times P2)}{N}\times 1-CR}}$

$$CP = \frac{0.60-0.50}{\sqrt{\frac{2(55\times 45)}{250}\times 1-0.41}} = \frac{0.10}{\sqrt{19.8\times 0.59}} = \frac{0.10}{0.034} = 2.94$$

الفرق يكون دالا عند 0.05 لو بلغت قيمته 1.96 ، كما يكون دالا عند 0.01 لو بلغت قيمته 2.58 فما فوق. ويكون غير دال إذا كانت قيمته دون ذلك.

- وتجدر الإشارة في الأخير إلى أنه يمكن حساب النسبة الحرجة (لإيجاد الفروق وتوضيح مدى دلالة الفرق بين نتائج مجموعتين) أيضا بمعرفة الوسط الحسابي والانحراف المعياري للمجموعتين، ودون استعمال النسب المئوية. ويكون ذلك بتطبيق قانون النسبة الحرجة (ن.ح.) التالي: ن. ح. = $\frac{م_1 - م_2}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$

حيث $م_1$ = الوسط الحسابي للمجموعة الأولى

$م_2$ = الوسط الحسابي للمجموعة الثانية

$ع_1^2$ = مربع الانحراف المعياري للمجموعة الأولى

ع $_2^2$= مربع الانحراف المعياري للمجموعة الثانية

ن $_1$ = عدد حالات المجموعة الأولى

ن $_2$ = عدد حالات المجموعة الثانية

مثال: احسب الفرق بين بيانات مجموعتي الجدول (68) الموالي:

الفئات	ك المجموعة 1	ك المجموعة 2
2−	2	3
4−	5	6
6−	6	2
8−	3	9
12−10	1	7
مجموع	17	27

الحل:

− جدول المجموعة 1 (رقم 69)

الفئات	ك	مراكز الفئات (س)	ك×س	$(س-\overline{س})$	$(س-\overline{س})^2$	$(س-\overline{س})^2$ × ك
2−	2	3	6	3.52−	12.39	24.78
4−	5	5	25	1.52−	2.31	11.55
6−	6	7	42	0.48+	0.23	1.38
8−	3	9	27	2.48+	6.15	18.45
12−10	1	11	11	4.48+	20.07	20.07
المجموع	17		111			76.23

− جدول المجموعة 2 (رقم 70)

الفئات	ك	مراكز الفئات (س)	ك×س	$(س-\overline{س})$	$(س-\overline{س})^2$	$(س-\overline{س})^2$ × ك
2−	3	3	9	4.81−	23.13	69.39
4−	6	5	30	2.81−	7.89	47.34
6−	2	7	14	0.81−	0.65	1.30
8−	9	9	81	1.19+	1.41	12.69
12−10	7	11	77	3.19+	10.17	71.19
المجموع	27		211			201.91

$$\overline{س}_{م1} = \frac{مج\ ك\ س}{مج\ ك} = \frac{111}{17} = 6.52$$

$$\overline{س}_{م2} = \frac{مج\ ك\ س}{مج\ ك} = \frac{211}{27} = 7.81$$

$$ع_1 = \sqrt{\frac{مج\left(س - \overline{س}\right)^2 \times ك}{مج\ ك}} = \sqrt{\frac{76.23}{17}} = \sqrt{4.48} = 2.1$$

$$ع_2 = \sqrt{\frac{مج\left(س - \overline{س}\right)^2 \times ك}{مج\ ك}} = \sqrt{\frac{201.91}{27}} = \sqrt{7.47} = 2.73$$

$$ن.ح = \frac{6.52 - 7.81}{\sqrt{\frac{2.1^2}{17} + \frac{2.73^2}{27}}} = \frac{-1.29}{\sqrt{0.52}} = -1.79$$ (غير دالة إحصائيا كما مر معنا)

قائمة المصادر

- **باللغة العربية**

- إبراهيم علي إبراهيم عبد ربه (2005): الإحصاء الوصفي، الإسكندرية، الدار الجامعية.
- أبو صالح محمد صبحي وعدنان محمد عوض (1983): مقدمة في الإحصاء، الأردن..
- أحمد الأشقر (1999): مقدمة في الإحصاء، مفاهيم وطرائق، عمان، مكتبة دار الثقافة.
- أحمد عبادة سرحان (1962): مقدمة في الإحصاء الاجتماعي، مصر، دار المعارف.
- أحمد عبد المنعم حسن (1986): أصول البحث الاجتماعي العلمي، بيروت، دار الطليعة.
- إحسان محمد الحسن (1996): الأسس العلمية لمناهج البحث الاجتماعي ، القاهرة، المكتبة الأكاديمية.
- بشير معمرية (2002): القياس النفسي وتصميم الاختبارات النفسية ، باتنة (الجزائر)، شركة باتنيت للمعلوماتية والخدمات المكتبية والنشر.
- جلال عبد الوهاب (1984): العلاقات الإنسانية والإعلام، الكويت، ذات السلاسل.
- جامعة مؤتة (1992): "*مجلة مؤتة للبحوث والدراسات*"، المجلد السابع، العدد 3، الأردن.
- ليلى داود (2000): البحث العلمي في العلوم الاجتماعية والنفسية، جامعة دمشق.
- مدني دسوقي مصطفى (1977): مبادئ علم الإحصاء، القاهرة، دار النهضة العربية، ط.5.
- محمد الجوهري و عبد الله الخريجي (1997): طرق البحث الاجتماعي، مصر، دار المعرفة الجامعية.
- محمد علي الأطرقجي(د.س.): الوسائل التطبيقية في الطرق الإحصائية ،بيروت، دار الطليعة للطباعة والنشر.
- محمد عبد الحميد (2000): البحث العلمي في الدراسات الإعلامية، سوريا، عالم الكتب.
- محمد عبيدات (1999): منهجية البحث العلمي، جامعة الأردن.
- محمود محمد صفوت (1962): مراحل البحث الإحصائي، القاهرة، مكتبة الانجلو المصرية.
- محمود السيد أبو النيل (1987): الإحصاء النفسي والاجتماعي والتربوي، بيروت، دار النهضة العربية.

- مختار محمد الهانسي (1982): مقدمة في طرق الإحصاء الاجتماعي ، بيروت، دار النهضة العربية.
- معين أمين السيد (د.ت.): المعين في الإحصاء، الجزائر، دار العلوم للنشر والتوزيع.
- موساوي عبد النور (2004): الإحصاء الوصفي، قسنطينة، منشورات جامعة منتوري قسنطينة.
- رجاء وحيد دويدري(2000): البحث العلمي: أساسياته النظرية وممارسته العملية ، لبنان/دمشق، دارالفكر المعاصر.
- سامي محمد ملحم (2000): القياس والتقويم في التربية وعلم النفس، عمان، دار المسيرة للنشر.
- شريف شطيبي (2002): محاضرات في الإحصاء الوصفي، جامعة منتوري قسنطينة.
- عاطف عدلي العبد وزكي أحمد عزمي (2002): الأسلوب الإحصائي واستخداماته في بحوث الرأي العام والإعلام (الدراسات الميدانية–تحليل المحتوى–العينات)، القاهرة، دار الفكر العربي.
- عباس محمد عوض (1984): علم النفس الإحصائي، بيروت، الدار الجامعية.
- عبد العزيز القوصي وآخرون (1956): الإحصاء في التربية وعلم النفس، القاهرة، النهضة.
- عزام صبري (2006): الإحصاء الوصفي، عمان، عالم الكتب الحديث– جدارا للكتاب العالمي.
- غريب سيد أحمد (2005): الإحصاء في البحوث الاجتماعية والإعلامية ، مصر، دار المعرفة الجامعية.
- فتحي عبد العزيز أبو راضي (1998): الطرق الإحصائية في العلوم الاجتماعية، بيروت، دار النهضة العربية للطباعة والنشر.
- فضيل دليو (1997): أسس البحث وتقنياته في العلوم الاجتماعية، قسنطينة، ديوان المطبوعات الجامعية.
- فضيل دليو (2001): التمثيلات البيانية في العلوم الاجتماعية ، الجزائر، ديوان المطبوعات الجامعية.
- فضيل دليو (يناير 2005): التحليل الإحصائي للبيانات الكمية والكيفية، مجلة الباحث الاجتماعي (جامعة قسنطينة)، العدد 7. في: https://www.researchgate.net/publication/363040680_althlyl_alahsayy_llbyanat_alkmyt_walkyfyt

- فضيل دليو (2014): "معايير الصدق والثبات في البحوث الكمية والكيفية""، مجلة العلوم الاجتماعية، جامعة سطيف2، العدد19، ديسمبر 2014.

- باللغات الأخرى

- AZAFRA, M.-J. (1999) : Cuestionarios, Madrid, C.I.S..
- BLALOCK, Hubert M. (1966): Estadística social, México, f.c.e.
- BUGUEDA, J.(1974): Manual de técnicas de investigación social, Madrid, I.E.P.
- CHAMBERS, J. M. & al. (1983): Graphical Methods for Data Analysis , Boston, Duxbury Press.
- COMBESSIE, J-C. (1999) : La méthode en sociologie, Paris, la découverte.
- CORTADA, Nuria & Carrol, J. M. (1996): Estadística aplicada, Buenos Aires.
- FERRANDO, García y otros (1989): El análisis de la realidad social, Madrid, Alianza Ed.
- Goetz & LeCompte. in : Iraima V. Martínez M. : Criterios de fiabilidad y validez en la Investigación Cualitativa con Enfoque Etnográfico. http://www.articuloz.com/monografias-articulos/criterios-de-fiabilidad-y-validez-en-la-investigacion-cualitativa-con-enfoque-etnografico-2076198.html. Posteado: 31/03/2010
- KöNIG, Rene (1973): Tratado de sociología empírica (Tome 1), Madrid, Tecnos.
- MAURICE Angers (1997) : Initiation pratique a la méthodologie des sciences humaines, Alger, Casbah.
- MAYNTZ, Renate y Otros (1980): Introducción a los métodos de sociología empírica, Madrid, Alianza Universidad.
- MUCCHIELLI, R. (1970) : Le questionnaire dans l'enquête psycho-sociale, Paris, Ed. Sociales Françaises.
- PARDINAS, Felipe (1979): Metodología y técnicas de investigación en ciencias sociales, México.
- Pérez Serrano 1988 in: Iraima V. Martínez M. : Criterios de fiabilidad y validez en la Investigación Cualitativa con Enfoque Etnográfico. http://www.articuloz.com/monografias-articulos/criterios-de-fiabilidad-y-validez-en-la-investigacion-cualitativa-con-enfoque-etnografico-2076198.html. Posteado: 31/03/2010
- PONS, Ignacio (1993): Programación de la investigación social (Serie Cuadernos Metodológicos, N°8), Madrid, C.I.S.
- PUJADAS MUÑOZ, J. J. (1992) : El método biográfico, Madrid, C.I.S.
- RUIS-MAYA PEREZ, Luis (Ed.) (1990): Metodología estadística para el análisis de datos cualitativos, Madrid, C.I.S.
- SIERRA BRAVO, Ramón (1979): Técnicas de investigación social, Madrid, Paraninfo.