د . جويدة عميرة

التحليل الإحصائي في البحوث الإجتماعية

جوانا
للنشر والتوزيع

الناشر : [القاهرة] : جوانا للنشر والتوزيع ،
الوصف : 182 ص. : إيض. ؛ 24 سم .

المقدمة:

إن بداية ظهور علم الإحصاء ونشأته كانت مع ظهور التعدادات والإحصاءات التي كانت تهم الحكومات والمؤسسات الإدارية في الدول والممالك.و كانت متمثلة في إحصاء السكان والجنود والحيوانات والأملاك و الممتلكات والغلال والمزارع.

ثم بدأ الإحصاء يتطور تدريجيا عبر العصور حتى نهاية القرن التاسع عشر وبداية القرن العشرين. أين شهد قفزة نوعية، إذ تطور تطورا هائلا بفضل جهود العديد من علماء الرياضيات، وأصبح اليوم علم مشترك لكافة الفروع والتخصصات العلمية الدقيقة منها والاجتماعية، ذلك لأنه العلم الذي يحاول تحليل و استخراج النتائج الرياضية من خلال دراسة الظواهر المختلفة.

إن الهدف من هذا الكتاب هو توضيح مختلف المراحل التي يمر بها البحث الإحصائي في العلوم الاجتماعية، وعليه فقد تضمن سبعة فصول:

الفصل الأول: خصصناه للجانب المنهجي وكمدخل للإحصاء الوصفي حيث بلورنا فيه تعريف الإحصاء ونشأته. كما صغنا فيه مصادر جمع البيانات الإحصائية وأساليب جمعها، كما تطرقنا إلى طرق اختبار العينات.

الفصل الثاني: ارتأينا فيه الحديث عن التوزيعات التكرارية باختلافها البسيطة و المزدوجة و المفتوحة،...الخ. وللاستفادة أكثر من هذا الفصل لاحظنا أنه لا بد من التطرق إلى كيفية مراجعة البيانات الإحصائية قبل تبويبها في الجداول الإحصائية.

الفصل الثالث: تطرقنا فيه إلى عرض البيانات الإحصائية في الجداول التكرارية وكيفية مقارنتها وقراءتها إحصائيا من جداول مزدوجة بسيطة ذات متغيرين مستقل وتابع، و جداول مركبة مضاعفة التقاطع.

كما تعرضنا في هذا الفصل إلى العرض البياني للبيانات الإحصائية في مختلف أنواع المتغيرات الكمية المتصلة، والمنفصلة وكذا المتغيرات الكيفية.

الفصل الرابع: تعرضنا فيه إلى مختلف مقاييس النزعة المركزية من متوسطات إحصائية منوال، وسيط، الربيعيات، المئينات، العشيرات،...الخ.

الفصل الخامس: شمل هو الآخر مثل الفصل الرابع أنواع المقاييس الإحصائية المستعملة في حالة الجداول البسيطة أي ذات متغير واحد، لكن في هذا الفصل تعرضنا إلى مقاييس التشتت أو التبعثر مثل المدى، الانحراف، المتوسط، التباين، والانحراف المعياري...الخ.

الفصل السادس: وفيه حددنا أهم معاملات الارتباط المستعملة في حالة الجداول المزدوجة، أي ذات متغيرين سواء كان كلاهما متغيرين الكميين معا، كمعامل بيرسون، والائتلاف أو كلاهما كيفيين معا كمامل كارل بارسون،أو أحدوهما كيفي والثاني كمي كمعامل الاقتران، التوافق، لامدا، فاي، سبيرمان،...الخ.
كما توقفنا في هذا الفصل على أشكال الانتشار، وكيفية رسم خطوط الانحدار في جداول بسيطة ومركبة.

الفصل السابع: وكان هذا الفصل فصلا ختاميا حيث رأينا أنه من واجبنا التعرض في مثل هذا الكتاب إلى معامل استقلالية الظواهر كا^2.

وفي الأخير نأمل بهذا الجهد المتواضع أن نكون قد أفدنا طلابنا بالقليل مما تعلمناه في كيفية معالجة أي بحث إحصائي في البحوث الاجتماعية.

الفصـــل الأول

ماهية علم الإحصاء

أولا: مدخل للإحصاء

1- تعريف علم الإحصاء

2- تاريخ نشأة علم الإحصاء

ثانيا: مصادر جمع البيانات الإحصائية

1 - البيانات الجاهزة

2 - البيانات الميدانية

ثالثا: أساليب جمع البيانات الإحصائية

1- المسح أو الحصر الشامل

2- المسح بالعينة

رابعا: طرق اختيار العينة

1- العينات العشوائية

2- العينات الغير عشوائية

أولا: مدخل للإحصاء.

١- تعريف علم الإحصاء.

يلعب الإحصاء دورا بارزا في البحوث الاجتماعية و الإنسانية، إذ يعد الأداة الرئيسية التي يستخدمها المخططون في تخطيطهم.

فالإحصاء غايته استقاء و تسجيل و تفسير المعلومات الكمية، التي يؤثر فيها عددا من الأسباب، فهو عملية جمع المشاهدات المحدودة أو غير المحدودة لدراستها و معرفة خصائصها و صفاتها و الإلمام بنتائجها، و ترتيب هذه النتائج بأشكال عددية حتى يتمكن المخططون من أخذ القرارات في النهاية.

فالإحصاء ما هو إلا أسلوب علمي ينظم التفكير، يعتمد على جمع الحقائق لظاهر معينة تكون محل الدراسة، و تسجيل الأعداد على هيئة بيانات و استخراج المقاييس اللازمة لإظهار النتائج و الاحتمالات بصورة سليمة و واضحة، و الكشف عن المؤثرات التي تتحكم بهذه النتائج و ايجاد العلاقة بينهما، لتوضيح الطرق الأنسب لحل المشكلة المراد دراستها.

ولقد ظهرت أهمية الإحصاء منذ نهاية القرن الثامن عشر مع تقدم مختلف العلوم في مختلف المجالات. حيث توجه بعض الباحثون في وضع القوانين و الاحتمالات و الخوض في ضبط العلاقات بين الأسباب و المسببات في مختلف الظواهر.

أ- الإحصاء لغويا:

كان استخدام الإحصاء عند نشوءه محصورا على الأعمال الخاصة بشؤون الدولة، كما يدل ذلك الأصل اللغوي لاسم هذا العلم statistique الذي هو مشتق من الكلمة اليونانية status أو الإيطالية statista و تعني كلاهما الدولة السياسية. و معناه مجموعة الحقائق

الخاصة بشؤون الدولة. فمن الطبيعي إذا أن تكون الدولة أول من اهتم بجمع البيانات و ذلك لإدارة شؤون البلاد خاصة في معرفة عدد السكان لأغراض حربية و ضرائبية منا اشرنا اليها سابقا.

ب- الإحصاء في اللغة العربية:

الإحصاء في اللغة العربية يقصد به العد الشامل أو تعداد الأشياء و تصنيفها كعد عدد الواردات، عدد الطلبة...الخ.

ت- الإحصاء كعلم:

هو مجموعة من المناهج و الطرق العلمية المطبقة لاتخاذ أنسب حل لأي إشكال مطروح. و بعبارة أخرى الإحصاء هو مجموعة من المناهج و الطرق العلمية التي من خلالها يمكن للباحث جمع، تنظيم و تلخيص و كذا تحليل المعطيات، مما يسمح له في الأخير إعطاء استنتاج عام للبحث، و بالتالي أخذ موقف معين من الموضوع المدروس.

و كخلاصة الإحصاء هو فرع من فروع الرياضيات يهتم بتصميم التجارب على العينات، بتحليل البيان الإحصائي، و القيام باستقراءات حول مجتمع من القياسات بدءا من المعلومات التي تحويها العينة، فهو بهذا يهتم بتطوير و استخدام أساليب التصميم، التحليل و القيام باستقراء.

٢- تاريخ نشأة الإحصاء.

تطور علم الإحصاء عبر سنوات طويلة، و ثم ذلك بفضل جهود كثيرة من عدة باحثين اشتغلوا في حقول مختلفة، و من دول عديدة كجاوس gauss من ألمانيا، لابلاس la place من فرنسا، كتليه quetlet من بلجيكا، جلتون galton من انجلترا... إلخ.

ان فكرة الإحصاء قديمة ترجع إلى البعيد في تاريخ الإنسانية، و قام بتطبيق هذه الفكرة واستخدامها في التدابير السياسة قدماء الصيين منذ أربعة آلف سنة. حيث كانوا يلجئون لجداول تعدادية في زراعتهم، كانت أقرب ما يكون مفهومها لمفهوم الإحصاء في عصرنا الحاضر. كما عرف قدماء المصريين و الإغريق و الرومان و العرب الإحصاء في أوج حضارتهم، حيث أدخلوه في أعمالهم لاسيما العسكرية و الزراعية و السكانية. فمثلا نجد قدماء المصريين قاموا بتعداد سكانهم و ثروتهم واستخدموا النتائج في تنظيم مشروع بناء الأهرام، و كذلك عمل رمسيس الثاني تعدادا للسكان تمهيدا لعملية إقطاع الأراضي و توزيعها على السكان بطريقة عادلة.

لقد كان تطور الإحصاء في القديم تطورا بطيئا حتى القرن العشرين. أين شهد تطورا و تقدما سريعا و هائلا، فصدرت عنه عدة نظريات و قوانين. و كان هذا التطور ملازما لتطور نظرية الاحتمالات في الرياضيات، إذ لم يأخذ الإحصاء هذا الشكل الجديد إلا عبر إسهامات العديد من الخبراء كبسيلولي ليكا pacioli luca، كبلر kipler، جاليليو galilio ما بين ١٤٩٣- ١٩٤٢. اللذين قاموا بتطوير نماذج الاحتمالات عبر التاريخ الحقيقي لنظرية الاحتمالات. كما لا ننسى إسهامات باسكال بلاز B.pascal، فرمات format، دانيال برنولي D.bernouli و فريدريك جاوس F.gauss و لابلاس la place ما بين سنة ١٧٠٠- ١٨٢٠.

لقد ظهر الاهتمام الكبير بتطبيق النظريات و الطرق الإحصائية في مجال العلوم الاجتماعية و الإنسانية على يد كيتليه عالم الفلك الاجتماعي، الذي بين إمكانية استخدام الاحتمالات و الإحصاء لوصف و تفسير الظواهر الاجتماعية و الاقتصادية، و قدم مساهمات هامة في الطرق الإحصائية في تنظيم إدارة الإحصاءات الرسمية، كما قدم طريقة القياس في الأنتروبولوجيا.

فحين قدم جالتون نظريته في تطبيق الطرق الإحصائية في علم النفس و وضع بذلك أساس علم القياس النفسي، واهتم كذلك كارل برسون K.pearsan و سبيرمان بدارسة الارتباط و الانحدار بين المتغيرات و كذا وضعوا أساس التحليل العاملي.

و من جهة أخرى اهتم فيشر fischer بالتقديرات و وضع بذالك نظريته حولها. شأنه شأن كارل برسون و نيمان neyman اللذان وضعا نظريتهما حول اختبار الفروض، لهذا يعد كل من الثلاثي (فيشر و كارل برسون و نيمان) من مؤسسي المنهج الاستقرائي في الإحصاء الذي يعرف حاليا بالاتجاه الكلاسيكي الذي يعتمد على البيانات المجمعة من العينات فقط.

و هكذا ارتبط علم الإحصاء بالعلوم المختلفة الأخرى زيادة عن علم الاجتماع و علم النفس، كعلم الفلك، علم الاقتصاد، علم السياسة، التجارة... الخ. و منذ هذا التاريخ أخذ هذا العلم سبله نحو الأبحاث المتنوعة و التخصصات المتشبعة. كالإحصاء الحيوي، الاجتماع الرياضي، علم النفس الرياضي، القياس النفسي، القياس التربوي، علم الاقتصاد الإحصائي، الإحصاء السكاني...الخ.

و أصبح مجال استعماله واسعا عند السياسيون لاختبار آراء شعوبهم، و المخططون لوضع مشاريعهم، و الأطباء لمعرفة أنجع أدويتهم...الخ.

ثانيا: مصادر جمع البيانات الإحصائية.

ذكرنا في تعرف الإحصاء سابقا أن عمل الباحث يبدأ بجمع المعطيات الإحصائية، وهي الخطوة الأولى التي يقوم بها لأنها أساس بحثه. و منه فيمكن أن نعتبر هذه المرحلة من أهم المراحل التي يمر بها البحث الإحصائي، باعتبار أنه ليس لنتائج التحليل الإحصائي أية قيمة إذا لم تكن البيانات الإحصائية التي قام بتحليلها قد جمعت بشكل صحيح. و لذلك فمهما بذل من جهد و عناية في استخدام أحسن الأساليب الإحصائية لا يمكن أن تعوضه عن عدم صحة جمع البيانات التي اتخذت أساسا للدراسة الإحصائية. و يمكن الحصول على هذه البيانات من مصدرين أساسيين هما:

١- البيانات الجاهزة / المصادر الغير مباشرة.

من خلال هذه البيانات يستطيع الباحث التعامل مع مادة سبق جمعها عن ظاهرة ما، و باستطاعته الرجوع إليها و أخذ المعلومات المراد التحقق منها، كسجلات التلاميذ في

الثانوية و الجامعات، ملفات المرضى في المستشفيات، التقارير الرسمية التي تصدرها المؤسسات الصحية، التعليمية، الاقتصادية، الحكومية...الخ. و التي تحتوي على بيانات تتعلق بالسكان و نوعهم و حجمهم، مهنتهم و مستواهم التعليمي...الخ.

فمن خلال هذه البيانات يستطيع الباحث التعامل مع مادة سبق جمعها إلا أن هذا النوع من البيانات يعتريه نقصا كبيرا بحيث تجمع هذه البيانات عادة لأغراض مختلفة تماما عن أغراض الدراسة المراد البحث فيها.
لذلك فإنها لا تعطي المعلومات المطلوبة و الكافية لغرض الدراسة، كما أنه ليس لدى الباحث أية رقابة كافية على مدى صحة المعلومات.

فمثلا إذا كنا بصدد دراسة أسباب وفيات الأمهات، فمصدر الدراسة في هذه الحالة هو ملفات النساء المتوفيات. تلك الملفات تحتوي على بيانات سبق جمعها من طرف الأطباء. فنجد فيها بيانات حول سنهن، مدة حملهن، رتبة المولود، و تفصيلا شاملا عن التشخيص الطبي لسبب الوفاة...الخ. إلا أن الباحث في العوم الاجتماعية يحتاج معرفة بالإضافة إلى هذه المعلومات بيانات أخرى حول المحيط الاقتصادي والاجتماعي للأم المتوفية، مثلا مستواها التعليمي، مستواها المعيشي، دخل الأسرة، مكان السكن...الخ. و التي قد لا تتوفر في الغالب بهذه الملفات.

٢- البيانات الميدانية / المصادر مباشرة.

كثيرا ما يجد الباحث في العلوم الاجتماعية والإنسانية بأن البيانات الإدارية غير ملائمة لغرض دراسته لأنها غير مكتملة و لا تجيبه عن تساؤلاته فيلجا في هــذه الحالة إلى جمع البيانات بنفسه من ميدان بحثه فهي بيانات غير شاملة في الغالب لأنها تخص مجموعة صغيرة فقط من أفراد المجتمع الإحصائي: فيحصل على بيانات من مصدرها الأصلي ذلك عن طريق الاتصال بمفردات وحدته الإحصائية إما مباشرة من خلال توجيه

الأسئلة على مجتمع بحثه، و إما عن طريق المقابلة الشخصية للبحوث أو عن طريق إرسال الاستمارات التي تحوي على مجموعة من الأسئلة التي تخدم أهداف البحث.

لهذا فهناك عدة طرق يتم من خلالها جمع البيانات الإحصائية، هذه الطرق تختلف باختلاف موضوع البحث، و الإمكانيات المادية و البشرية المتاحة لدى الباحث.

ثالثا: أساليب جمع البيانات الإحصائية.

هناك أسلوبان أو طريقتين لجمع البيانات

١-الأسلوب الأول / المسح أو الحصر الشامل.

ويقصد به إدخال كل مفردات المجتمع الإحصائي المعني بالدراسة دون استبعاد أي فرد منه. تستعمل غالبا هذه الطريقة في المجتمعات الإحصائية مجهولة المعالم والتي تتطلب جمع بيانات شاملة عن كل فرد من أفراد مجتمع الدراسة حتى يتمكن من تحديد خصائصه و معالمه بكل دقة. إلا أن مثل هذا الأسلوب لجمع البيانات يحتاج إلى وقت وجهد كبيرين.

٢_ الأسلوب الثاني/ المسح بالعينة.

هو عملية جمع البيانات عن جزء ممثل للمجتمع الإحصائي يعرف اصطلاحا بالمجتمع الإحصائي المرجعي، فالعينة إذا هي جزءا من السكان المعنيين بالدراسة و البيانات التي تجمع عنها تنطبق على ذلك المجتمع، و هي تمثل نسبة مئوية منه. و قد نأخذ بأسلوب العينة في الأبحاث حتى نتغلب على الصعوبات و عيوب الحصر الشامل حيث يهدف هذا الأسلوب بدراسة عدد محدود من الأفراد المعنيين بالدراسة بصورة أفضل من خلال جمع معلومات دقيقة و كثيرة عن كل فرد و بالتالي يمكن التحصل على نتائج ذات دقة أفضل و بتكاليف أقل و وقت أقصر.

رابعا: طرق اختيار العينة.

يمكن تصنيف العينات إلى عينات عشوائية و غير عشوائية.

١- العينات العشوائية.

فالعشوائية هي التي يكون إطار معاينتها محددا و موجودا. و هي تلك العينات التي تسحب بأسلوب عشوائي لهذا فكل قوانين الاحتمالات تطبق عليها، و فيها يكون لكل فرد من أفراد الوحدة الإحصائية فرصة للظهور ضمنها ومن أنواعها نجد العينة العشوائية البسيطة، العينة المنتظمة، العينة الطبقية و العينة العنقودية...الخ.

٢- العينات الغير عشوائية.

و يدعى هذا النوع من العينات أحيانا بالعينات الشخصية ويتم الحصول على البيانات في هذا النوع من العينات بطريقة غير عشوائية، وبالتالي لا يمكن تطبيق قوانين الاحتمالات و الاستدلال الإحصائي عليها. و فيها يتعمد الباحث أن تكون عينته متكونة من وحدات معينة لاعتقاده أنها تمثل المجتمع الإحصائي الأصلي، لهذا عند استعمال هذا النوع من العينات أن نكون لى قدر كبير من التحفظ على النتائج المستخلصة في تعميمها. ومن أنواعها نجد عينة الحصص، عينة الكرة الثلجية و العينة العرضية...الخ.

الفصل الثاني

البيانات في علم الإحصاء

أولا: مراجعة و تبويب البيانات الإحصائية

ثانيا: التوزيعات التكرارية

١- بعض التعريفات الأساسية في الإحصاء

٢- تبويب التوزيع التكراري البسيط

٣- تبويب التوزيع التكراري المزدوج

٤- التوزيع التكراري المنتظم و غير المنتظم

٥- التوزيع التكراري المفتوح و المغلق

٦- التوزيع التكراري المتجمع

٧ - تمرينات محلولة حول الفصل الأول و الثاني.

أولا: - مراجعة و تبويب البيانات الإحصائية.

بعد الانتهاء من عملية جمع البيانات من الميدان تأتي المرحلة الثانية من مراحل البحث الإحصائي، و هي عملية فرز و مراجعة و تدقيق الاستمارات لتأكد من أن كل سؤال في الاستمارة قد أجيب عليه. كذلك لفحص ما إذا كانت المعلومات التي أذلي بها المبحوث خاطئة أو ناقصة أو متناقضة، ثم يقوم الباحث بإرجاعها إلى الميدان لتصحيحها أو إلغائها إن تعذر له ذلك.

و بعد الانتهاء من هذه المرحلة تأتي المرحلة الثانية و هي مرحلة الترميز المسبق للاستمارات. و فيها يقوم الباحث بوضع رموز عددية في الخانة المتروكة في هامش الاستمارة، ذلك الترميز يمثل رقم احتمال الإجابة في السؤال و كمثال على ذلك ترميز السؤالين التالين في الاستمارة:

س١- ما هو مستواك التعليمي؟

١- أمي ☐ ٢- ابتدائي ☐ ٣- متوسط ☒ ٤- ثانوي ☐ ٥- جامعي ☐ | ٣

س٢ - ما هي حالتك المدنية ؟

١- متزوج☐ ٢- أعزب ☒ ٣- أرمل ☐ ٤- مطلق ☐ | ٢

و بعد الانتهاء من هذه العملية أي ترميز الأسئلة. سؤالا بعد الآخر نقوم بإعطاء أرقاما تسلسلية للاستمارات.

فالباحث بعد أن يجمع بيانات موضوع دراسته يستحيل عليه أن يستوعب هذه البيانات على ماهي ليه دون أن يضعها في صورة مبسطة يسهل معها دراستها. فيقوم بترميز الأسئلة ثم يقوم بتفريغها في جدول التفريغ البياني. بحيث يقسم هذا الجدول إلى صفوف و أعمدة فيتبث على العمود الأول أرقام الاستمارات أو المبحوثين و في الأعمدة المتبقية أرقام الأسئلة كما هو مبين في الجدول التالي:

رقم السؤال / رقم الاستمارة	س ١	س ٢	س ٣	س ٤		س ن
١						
٢						
٣						
ن						

إذا كان حجم العينة صغيرا يتم تفريغها يدويا على الجدول السابق، أما إذا كان حجمها كبيرا فيمكن الاستعانة بالآلات التي تعتمد على نظام البطاقات المثقبة سابقا و الأقراص الممغنطة و الأشرطة حاليا، مثلا باستعمال حزمة البرامج الجاهزة أو ما يسمى بالحقبة الإحصائية للعلوم الاجتماعية SPSS و هذا طبعا بعد تنقيح و تصحيح و إعطاء نماذج الإجابات.

ثانيا: التوزيعات التكرارية.

التوزيع التكراري هو عبارة عن توزيع البيانات المأخوذة عن ظاهرة معينة على الفئات. بحيث تقع كل مفردة في فئة واحدة فقط تكون متجانسة فيما بينها. أي مع المفردات الأخرى في نفس الفئة. أما إذا كان عدد البيانات صغيرا فإنه يمكننا بناء جدول تكراري بترتيب البيانات إما تصاعديا أم تنازليا حتى نصل إلى أعلى قيمة للبيان.

ولوضع البيانات في جدول تكراري، نرسم جدولا ذو ثلاث أعمدة، يمثل العمود الأول المتغير المدروس، العمود الثاني لوضع جدولة البيانات و العمود الثالث لوضع التكرارات.

ثم نعود إلى البيانات الأصلية من جدول التفريغ البياني و نأخذها واحدة تلوى الأخرى في ترتيب تصاعدي أم تنازلي في العمود الأول، و نضع علامة بالعمود الثاني في الجدول لكل مفردة أمام الفئة التي تنتمي إليها المفردة، و منعا لاختلاط العلامات ببعضها البعض، و تفاديا لصعوبة عدها عند الانتهاء من وضعها يستحسن أن نضعها في شكل مجموعات. كل منها متكونة من خمسة علامات أربعة عمودية و الخامسة تقطع الأربعة جميعها، فتصبح العلامات على شكل حزم IIII أو على شكل مربع و الخامسة تقطع المربع في النصف كل ———————————— ⧅

خط في المربع يدل على مفردة. و بهذا يسهل عد المجموعات في النهاية.

و عند الانتهاء من عملية وضع العلامات / الجدولة يقوم الباحث بملأ العمود الأخير بعدد العلامات الموجودة في كل فئة. و في الأخير نقوم بجمعها كلها لنحصل على المجموع العام و هو التكرار الكلي الذي يمثل مجموع العينة.

١-بعض التعريفات الأساسية في الإحصاء.

- **المتغير:** المتغير هو عبارة عن صفة أو ميزة بإمكانها أن تأخذ على الأقل قيمتين أي تتغير من حالة إلى أخرى، و المتغير عبارة عن رمز مثل السن، المستوى التعليمي...الخ. و الذي يمكن أن يأخذ عدة قيم و التي تسمى بمجال المتغير.

و المتغير نظريا يمكنه أن يكون كمي و يأخذ جميع القيم بين حدي مدى التغير و يسمى هذا النوع متغيرات كمية مستمرة أو متصلة variables continues. مثل متغير السن، الوزن، الطول...الخ. فسن المبحوث يمكن أن يكون ١٩ سنة ٨ أشهر و ٤ أيام و طوله متر و ٧٥ سنتم.

و على العكس من ذلك يوجد نوع آخر من المتغيرات الكمية لا يمكن بطبيعتها أن تأخذ جميع القيم التي تقع في حدود مدى التغير. كعدد الأطفال في الأسرة الذي يأخذ القيم التالية: طفل، طفلين، ثلاث أطفال... الخ. أو عدد الغرف في البيت التي تأخذ القيم التالية: غرفة، غرفتين، ثلاث غرف... الخ.

أي أنه لا يمكن أن نتحصل على طفل و ربع في الأسرة مثلا، أو طفلين و نصف... الخ وهذا النوع من المتغيرات يدعى بالمتغيرات الكمية المنفصلة.

كما يوجد نوع آخر من المتغيرات غير قابلة للقياس أي لا تخضع للتكميم تسمى بالمتغيرات الكيفية variables qualitatives.

أما من ناحية علاقة المتغيرات ببعضها البعض فهي تنقسم عادة إلى متغيرات حرة / مستقلة و هي المسببة في وجود الظاهرة، و متغيرات تابعة و هي نتيجة المتغيرات المستقلة أو هي الظاهرة المدروسة نفسها، وقد يكون تأثير المتغيرات المستقلة بالمتغيرات التابعة تأثيرا مباشرا أو تأثير بواسطة. و عليه فقد يكون لدينا متغيرات آخرى و هي عبارة عن متغيرات وسيطة حاملة لفعالية المتغيرات المستقلة نحو المتغيرات التابعة و تسمى بالمتغيرات الرائزة variables testes.

– **التكرارت:** يرمز لها برمز **ت** أو **ك** أو f_i و هي عدد المرات التي ظهرت فيها المفردات المصنفة ضمن مجال فئة معينة.

– **الفئة:** هي مجموعة من قيم المفردات المصنفة و المحصورة بين الحدين أدنى و أعلى. يسمان بحدي الفئة. و حد الفئة يختلف عن حد مجال التوزيع التكراري إذ أن الحد الأدنى له هو الحد الأول للفئة الأولى و حده الأعلى هو آخر حد للفئة الأخيرة.

– **مجال الفئة:** هو الفرق بين أعلى و أدنى حد في الفئة.

– **عدد الفئات:** إذا كانت العينة أكثر من ١٠٠٠ وحدة إحصائية فنستخدم قاعدة سترجس sturges **و هي:**

عدد الفئات = ١ + ٣,٣٢٢ لغ **ن**

حيث **ن** هو عدد العينة / حجم العينة

و على الباحث بعد ذلك تحديد طول الفئات عن طريق المعادلة التالية:

$$\text{طول الفئة} = \frac{\text{الفرق بين أكبر و أصغر قيمة في البيانات}}{\text{عدد الفئات}}$$

و بذلك تكون الفئات متساوية الطول و إذا حدث و نتج عن المعادلة السابقة كسورا عشرية فيجب تقريبها إلى رقم صحيح. مثلا نحصل على طول فئة يساوي ٥,٧٣ فلا بد في هذه الحالة من تقريبها إلى طول فئة مقدارها ٦. أما إذا وجدناها تساوي مثلا ٥،٤٣ فنقربها الى رقم ٥.

هذا يعني أن عملية التقريب تعتمد على العدد الذي يأتي بعد الفاصلة، فإذا كان أكبر من ٥ نقرب الرقم إلى الرقم اللاحق له و إذا كان أقل من ٥ نحتفظ بالرقم الصحيح الذي هو قبل الفاصلة.

كما وضع ستر جس صيغة أخرى في مثل هذه الحالة و هي:

$$\text{عدد الفئات} = \frac{\text{١+ ١٠ لغ ن}}{\text{٣}}$$

مثال: إذا كانت لدينا عينة تتألف من ٥٠ تكرار و أردنا تصنيفها بهذه الصيغة، نبحث أولا عن اللوغاريتم العشري للعدد ٥٠ الذي يساوي ١,٦٩٩ و باستبدال الرموز بقيمتها نجد:

$$\text{عدد الفئات} = \text{١} + \frac{\text{١٠ لغ٥٠}}{\text{٣}} = \text{١} + \frac{1.699 \times 10}{3} = \text{٧ فئات}$$

أما إذا كان حجم العينة أقل من ١٠٠٠ وحدة فحينئذ على الباحث استعمال قاعد **بول** pule و صياغتها كالتالي:

$$\text{عدد الفئات} = \text{٢,٥} \sqrt{\sqrt{\text{ن}}}$$

حيث ن دائما تبقى تمثل حجم العينة.

ثم نقوم بالبحث عن طول الفئة كما في الطريقة السابقة.

و بما أن العدد الأمثل للفئات لا يرتبط فقط بحجم مفردات العينة فقط، بل كذلك بالمدى المطلق. و هو الفرق بين أعلى قيمة و أقلها في سلسلة البيانات الإحصائية. و الذي نختاره لهذه العينة، لهذا كلما كان صغيرا كلما زاد عدد الفئات و العكس صحيح.

فمثلا لدينا عينة متكون من ٣٠٠ وحدة إحصائية أصغر سن وجدناه عندها هو ٢٠ سنة و أكبر سن ٩٥ سنة.

فإن المدى المطلق لهذه للعينة هو ٩٥ – ٢٠ = ٧٥ سنة

فإذا أردنا أن نصنفها في ٨ فئات فإن مجال كل فئة سيكون مساويا إلى:

$\frac{75}{8}$ = ٩,٣٧ سنة أي تقريبا ٩ سنة. و بضرب هذا العدد أي مجال الفئات في عددها

٨ × ٩ = ٧٢ سنة

و هذا الرقم المتحصل عليه لا يغطي المدى المطلق لكل الفئات الذي يساوي ٧٥ سنة لذلك نختار تصنيف أخر و ليكن مثلا عدد الفئات متكون من ٧.

فمجال الفئة سيكون $\frac{75}{7}$ = ١٠,٧١ ≈ ١١ سنة و بضرب هذا العدد بعدد الفئات نجد:

١١ × ٧ = ٧٧ سنة.

أي أن ٧٧ سنة أكبر من ٧٥ سنة الذي يمثل المدى المطلق لذلك نأخذ هذا التصنيف. و بالتالي يكون تصنيف بيانات هذه العينة هو٧ فئات ذات الطول مجال كل فئة ١١ سنة كالتالي ٢٠– ٣٠، ٣١ – ٤١، ٤٢ – ٥٢،...الخ

و هنا يجب أن نفرق بين ما يعرف بالحدود الفعلية لكل فئة limits réelles de chaque classe و هي الأصغر قيمة نظرية و أكبرها لمتغيرات هذه الفئة بحيث تكون نهاية الفئة السابقة هي بداية الفئة التالية لها. فالحدود الفعلية لفئة ٢٠ – ٣٠ هي ١٩,٥ – ٢٩,٥ أي أن ١٩,٥ هو الحد الفعلي الأدنى للفئة و ٢٩,٥ هو الحد الفعلي الأعلى للفئة.

- **مركز الفئة:** مركز الفئة هو مجموع حدي كل فئة تقسيم اثنين

كمثال $\frac{30 + 20}{2} =$ ٢٥ سنة

أما مركز الفئة الفعلي للفئة ١٩,٥ – ٢٩,٥ فهو $\frac{29.5+19.5}{2}=$ ٢٤,٥

٢- تبويب التوزيع التكراري البسيط.

المثال الأول: التوزيع التكراري في حالة البيانات الكيفية.

لدينا البيانات التالية التي تبين توزيع ٤٠ مبحوث حسب مستواهم التعليمي المطلوب تلخيصها و تبويبها في جدول توزيع تكراري بسيط.

جامعي، ابتدائي، ابتدائي، أمي، ابتدائي، جامعي، ثانوي، ابتدائي، أمي، ثانوي، متوسط، ابتدائي، متوسط، متوسط، جامعي، ابتدائي، أمي، متوسط، أمي، ثانوي، أمي، متوسط، أمي، ابتدائي، ثانوي، أمي، ثانوي، ثانوي، أمي، ابتدائي، جامعين ابتدائي، جامعي، ابتدائي، أمي، أمي، ابتدائي، ثانوي، أمي، ابتدائي.

المستوى التعليمي	الجدولة	التكرارات
أمي	I HHH HHH	١١
ابتدائي	II HHH HHH	١٢
متوسط	HHH	٥
ثانوي	II HHH	٧
جامعي	HHH	٥
المجموع		٤٠

المثال الثاني: التوزيع التكراري في حالة البيانات الكمية

البيانات التالية تبين سن ٨٠ مسن في دار العجزة و المطلوب تبويبها في جدول توزيع تكراري.

٦٨	٨٤	٧٥	٨٢	٦٨	٩٠	٦٢	٨٨	٧٦	٩٣
٧٣	٧٩	٨٨	٧٣	٦٠	٩٣	٧١	٥٩	٨٥	٧٥
٦١	٦٥	٧٥	٨٧	٧٤	٦٢	٩٥	٧٨	٦٣	٧٢
٦٦	٧٨	٨٢	٧٥	٩٤	٧٧	٦٩	٧٤	٦٨	٦٠
٩٦	٧٨	٨٩	٦١	٧٥	٩٥	٦٠	٧٩	٨٣	٧١
٧٩	٦٢	٦٧	٩٧	٧٨	٨٥	٧٦	٦٥	٧١	٧٥
٦٥	٨٠	٧٣	٥٧	٨٨	٧٨	٦٢	٧٦	٥٣	٧٤
٨٦	٦٧	٧٣	٨١	٧٢	٦٣	٧٦	٧٥	٨٥	٧٧

أول شيء نقوم به قبل التبويب هو ترتيب البيانات إما تصاعديا إما تنازليا، وليكن مثلا تصاعديا.

٥٣	٥٧	٥٩	٦٠	٦٠	٦٠	٦١	٦١	٦٢	٦٢
٦٢	٦٢	٦٣	٦٣	٦٥	٦٥	٦٥	٦٦	٦٧	٦٧
٦٨	٦٨	٦٨	٦٩	٧١	٧١	٧١	٧٢	٧٢	٧٣
٧٣	٧٣	٧٣	٧٤	٧٤	٧٤	٧٥	٧٥	٧٥	٧٥
٧٥	٧٥	٧٥	٧٦	٧٦	٧٦	٧٦	٧٧	٧٧	٧٨
٧٨	٧٨	٧٨	٧٨	٧٩	٧٩	٧٩	٨٠	٨١	٨٢
٨٢	٨٣	٨٤	٨٥	٨٥	٨٥	٨٦	٨٧	٨٨	٨٨
٨٨	٨٩	٩٠	٩٣	٩٣	٩٤	٩٥	٩٥	٩٦	٩٧

المدى المطلق = ٩٧ – ٥٣ = ٤٤

فإذا أردنا أن نصنفها إلى ٦ فئات مثلا فإن مجال كل فئة سيكون

$\frac{44}{6}$ = ٧,٣٣ ≈ ٧ و بضرب هذا العدد بعدد الفئات نجد ٧ × ٦ = ٤٢

فإن ٤٢ سنة لا تغطي المدى المطلق لكل الفئات الذي يساوي ٤٤، لهذا سوف نختار عدد فئات متكون من ٧ لأن ٧ × ٧ = ٤٩ و هو أكبر من المدى المطلق لكل الفئات.

فئات السن	الجدولة	التكرارات
٥٠ – ٥٦	|	١
٥٧ – ٦٣	𝍸 𝍸 |||	١٣
٦٤ – ٧٠	𝍸 𝍸	١٠
٧١ – ٧٧	𝍸 𝍸 𝍸 𝍸 𝍸	٢٥
٧٨ –٨٤	𝍸 𝍸 𝍸	١٤
٨٥ –٩١	𝍸 𝍸	١٠
٩٢–٩٨	𝍸 ||	٧
المجموع		٨٠

ملاحظة: هناك بعض المتغيرات أين عدد فئاتها و أطوالها متفق عليه عالميا كسن خصوبة المرأة الذي يتراوح ما بين ١٩ – ٤٩ سنة.
و الذي يكون مقسما في فئات خماسية عددها ٧. وكذا تقسم وفيات الأطفال الرضع أين الفئة الأولى مداها ما بين ٠ – ٧ أيام، الفئة الثانية ٨ –27 يوما،و الفئة الثالثة ٢٨ – ٣٦٥ يوما.

٣- تبويب التوزيع التكراري المزدوج.

إلى جانب عملية تفريغ البيانات السابق الإشارة إليها أي ذات متغير واحد فإنه يمكن تفريغ البيانات في جداول مزدوجة حيث تؤدي العملية السابقة إلى جدول بسيط ذو متغير واحد فقط.

و لإجراء عملية تفريغ البيانات في جدول مزدوج أي يربط بين متغيرين اثنين نقوم برسم جدول ذي أعمدة و صفوف، يوضع على العمود الأول صفات المتغير الأول بينما يختص العمود الأفقي بوضع قيم المتغير الثاني، ثم نقوم بعملية تفريغ البيانات بطريقة الحزم التي أشرنا إليها سابقا.

في مثل هذه الحالة لا يمكن اختصار البيانات بواسطة التوزيع التكراري العادي، حيث يكون لدينا نوعان من البيانات، فإذا حاولنا اختصار كل منها في توزيع خاص بها فسوف يكون لدينا توزيعان مختلفان منفصلان عن بعضها. الأمر الذي لا ساعدنا في دراسة العلاقة بين المتغيرين، لهذا نحاول تبويبها في جدول واحد هو جدول توزيع تكراري مزدوج.

فمثلا لدينا البيانات التالية حول سن المرأة المنجبة و مستواها التعليمي و نريد ربط العلاقة بينهما.

السن	المستوى التعليمي	السن	المستوى التعليمي	السن	المستوى التعليمي	السن	المستوى التعليمي	السن	المستوى التعليمي
١٩	ثانوي	٢٥	أمي	٣٠	ابتدائي	٢٨	أمي	٣٦	جامعي
٢٠	متوسط	٣٠	ابتدائي	٢٧	ثانوي	١٦	أمي	٢٧	ثانوي
١٦	ابتدائي	٣٢	متوسط	٣٠	أمي	٢٣	متوسط	١٨	ثانوي
٢٠	متوسط	٣٥	متوسط	٣٣	جامعي	٢٨	متوسط	١٥	ابتدائي
٢٧	جامعي	٢٠	متوسط	٣٥	أمي	١٩	ثانوي	٢٨	جامعي
٢٥	ثانوي	٢٦	متوسط	٢٢	ثانوي	٣٠	جامعي	٤٠	متوسط
		٣٥	جامعي	٤٥	أمي	٤٩	أمي		

عملية تفريغ البيانات في هذا الجدول المزدوج تكون على الشكل التالي:

المستوى / السن	أمي	ابتدائي	متوسط	ثانوي	جامعي	المجموع
١٥ – ١٩	I	II		III		٦
٢٠ – ٢٤			III	I		٤
٢٥ – ٢٩	II		III	III	II	١٠
٣٠ – ٣٤	I	II	I		II	٦

٣٥ – ٣٩	I		I		II	٤
٤٠ – ٤٤			I			١
٤٥ – ٤٩	II					٢
المجموع	٧	٤	٩	٧	٦	٣٣

٣- التوزيع التكراري المنتظم والغير المنتظم.

في بعض الأحيان تكون البيانات منفصلة في جزء و مجملة في جزء آخر من المجموعة. و في مثل هذه الحالة لا يمكن عمل فئات متساوي، لأن بعض الظواهر يصبح معها استخدام الفئات غير المنتظمة أكثر ملاءمة لعرض الظاهرة، مثل ذلك دراسة ظاهرة وفيات الأطفال الرضع.

و بعبارة أخرى إذا كان مدى الفئات في التوزيع التكراري متساويا في جميع الفئات فإن التوزيع يكون منتظما والعكس صحيح، أي عندما لا تكون أطوال الفئات متساوية في جميع الفئات فنقول أنه توزيع غير منتظم.

مثال عن التوزيع التكراري المنتظم.

الجدول التالي يبين توزيع سن مجموعة من النساء المستعملات لوسائل منع الحمل.

السن	التكرارات
١٥ – ١٩	١٠
٢٠ – ٢٤	١٦
٢٥ – ٢٩	١٥
٣٠ – ٣٤	٣٠
٣٥ – ٣٩	١٨
٤٠ – ٤٤	١٥
٤٥ – ٤٩	١٠
المجموع	١١٤

فطول الفئة في هذا الجدول هي ٥ في جميع الفئات و بالتالي فالجدول التكراري منتظم.

مثال عن التوزيع التكراري الغير منتظم.

الجدول التالي يبين توزيع وفيات الأطفال الرضع

فئات السن	التكرارات
٠ – ٧ أيام	١٨
٨ – ٢٧ يوما	١٩
٢٨ – ٣٦٥ يوما	٣٣
المجموع	٧٠

في هذا الجدول نلاحظ أن فئاته غير متساوية الطول. فالفئة الأول طولها ٧ أيام والفئة الثانية طولها ٢٠ يوما و الأخيرة طولها ٣٣٨ يوما.

و في هذه الحالة أي حالة التوزيعات التكرارية الغير متساوية نحتاج إلى تعديل التكرارات أي نقوم بحساب التكرار المعدل وفق العلاقة التالية:

$$\text{التكرار المعدل} = \frac{\text{تكرار الفئة الأصلية}}{\text{طول الفئة}}$$

و لتوضيح هذا نقوم بتعديل تكرارات الفئات السابقة في الجدول التالي.

فئات السن	التكرار الأصلي	طول الفئة	التكرار المعدل
٠ – ٧ أيام	١٨	٧	٢,٥٧
٨ – ٢٧ يوما	١٩	٢٠	٠,٩٥
٢٨ – ٣٦٥ يوما	٣٣	٣٣٨	٠,٠٩
المجموع	٧٠		

ملاحظة: في حالة الجداول الغير منتظمة نقوم إذا بتعديل التكرارات ثم نعمل بها في جميع المقاييس الإحصائية، ماعدا في حالة المتوسط الحسابي لأنه ما يهمنا فيه هو مركز الفئة وليس طولها.

٥- التوزيع التكراري المفتوح و المغلق.

التوزيع التكراري نوعان:

الأول: مغلق و هو ما عرف و حدد حدا فئتيه العليا و السفلى معا. و في هذا النوع من الجداول تكون العمليات الحسابية أسهل، لأننا نحتاج في الغالب إلى مراكز الفئات و يمكن حسابها منه، و مثل ذلك الجدولين السابقين.

ففي الجدول الذي يبين توزيع سن النساء المستعملات لوسائل الحمل منع الحمل، نجد أن الحد الأدنى للفئة الأولى في الجدول معين و معلوم و قدر بـــ ١٥ سنة، كذلك الحد الأعلى في الفئة الأخيرة للجدول محدد و قدر بـــ ٤٩ سنة.

الثاني: مفتوح و هو ما لم يتعين و يحدد حده الأعلى أو الأدنى أو كلاهما معا.

المثال الأول: جدول مفتوح من طرفه الأدنى.

الجدول التالي يبين مدة إقامة ٧٠ مريض بالمستشفى.

التكرارات	المدة
٣	أقل من ٥ أيام
٧	٦ – ١٠
١٨	١١ – ١٥
١٥	١٦ – ٢٠
١٢	٢١ – ٢٥
١٥	٢٦ – ٣٠
٧٠	المجموع

المثال الثاني: جدول مفتوح من طرفه الأعلى أين تكون نهاية الفئة الأخيرة للجدول غير محددة كمايلي:

التكرارات	المدة
٣	١ – ٥
٧	٦ – ١٠
١٨	١١ – ١٥
١٥	١٦ – ٢٠
١٢	٢١ – ٢٥
١٥	٢٦ يوما فأكثر
٧٠	المجموع

المثال الثالث: جدول مفتوح من كلا طرفيه أي أن الفئة الأولى غير محددة و نهاية الفئة الأخيرة غير محددة هي الأخرى كما هو موضح فيما يلي:

التكرارات	المدة
٣	أقل من ٥ أيام
٧	٦ – ١٠
١٨	١١ – ١٥
١٥	١٦ – ٢٠
١٢	٢١ – ٢٥
١٥	أكثر من ٢٦ يوما
٧٠	المجموع

٦ – التوزيع التكراري المتجمع.

في بعض الأحيان نحتاج إلى معرفة عدد المفردات التي تساوي أو تزيد عن قيمة معينة. لهذا فالتوزيع التكراري البسيط لا يمكنه أن يجيبنا بسهولة و بشكل مباشر، و لكي نتمكن من الإجابة على تلك الأسئلة، لا بد من وضع التوزيع التكراري في شكل جديد سنطلق عليه التوزيع التكراري المتجمع.

و الفكرة الأساسية في التوزيع التكراري المتجمع هي تجميع التكرارات أمام الحد الأعلى لكل فئة. و في هذه الحالة يكون التوزيع التكراري متجمعا صاعدا حيث أن التكرارات في صعود مستمر. أو تجميع التكرارات أمام الحد الأدنى لكل فئة ابتداء من أسفل التوزيع و في هذه الحالة يكون التوزيع هابط/نازل حيث أن التكرارات في نزول مستمر.

مثال:الجدول التالي يوضح توزيع ١٢٠ مدخن حسب فئات السن ؟

- ما هو عدد المدخنين الأقل من ٧٩ سنة ؟
- ما هو السن الذي يقل عنه ربع $\frac{1}{4}$ أفراد العينة ؟
- ما هو عدد المدخنين الذين يتراوح سنهم ما بين ٨٠ – ٩١ سنة ؟
- ما نوع هذا التوزيع ؟

التوزيع التكراري البسيط		التوزيع التكراري المتجمع			
		التوزيع التكراري المتجمع الصاعد		التوزيع التكراري المتجمع النازل	
فئات السن	التكرارات	الحدود العليا للفئات	التكرار الصاعد	الحدود الدنيا للفئات	التكرار النازل
٤٤- ٤٩	٢	أقل من ٥٠ سنة	٢	أكثر من ٤٤ سنة	١٢٠
٥٠- ٥٥	٥	أقل من ٥٦ سنة	٧	أكثر من ٥٠ سنة	١١٨
٥٦- ٦١	١٢	أقل من ٦٢ سنة	١٩	أكثر من ٥٦ سنة	١١٣
٦٢- ٦٧	١٠	أقل من ٦٨ سنة	٢٩	أكثر من ٦٢ سنة	١٠١
٦٨- ٧٣	١٣	أقل من ٧٤ سنة	٤٢	أكثر من ٦٨ سنة	٩١
٧٤- ٧٩	٥٩	أقل من ٨٠ سنة	١٠١	أكثر من ٧٤ سنة	٧٨
٨٠- ٨٥	٧	أقل من ٨٦ سنة	١٠٨	أكثر من ٨٠ سنة	١٩
٨٦- ٩١	٥	أقل من ٩٢ سنة	١١٣	أكثر من ٨٦ سنة	١٢
٩٢- ٩٧	٢	أقل من ٩٨ سنة	١١٥	أكثر من ٩٢ سنة	٧
٩٨- ١٠٣	٥	أقل من ١٠٤ سنة	١٢٠	أكثر من ٩٨ سنة	٥
المجموع	١٢٠				

الإجابة:

- عدد المدخنين الأقل من ٧٩ سنة هو ١٠١ مدخن.
- السن الذي يقل عنه $\frac{1}{4}$ أفراد العينة هو ٦٧ سنة.
- عدد المدخنين الذين يتراوح عمرهم ما بين ٨٠- ٩١ سنة هو ١٢ مدخنا.
- هذا التوزيع هو توزيع تكراري منتظم مغلق و مبوب.

٧ - تمرينات محلولة حول الفصل الأول و الثاني.

التمرين الأول:

يبين الجدول التالي توزيع ١٥٧ امرأة حسب الوسيلة التي تستعملها لتنظيم نسلها.

الوسيلة المستعملة	ك
الوسائل الكيميائية	٩
الوسائل الهرمونية	٧٩
الوسائل الميكانيكية	٤١
الوسائل الطبيعية	٢٨
المجموع	١٥٧

السؤال: ما نوع التوزيع و المتغير؟

الجواب: هذا التوزيع توزيع تكراري بسيط ذو متغير واحد فقط. و هو الوسيلة المستعملة لتنظيم النسل و هو متغير كيفي.

التمرين الثاني:

فيما يلي بيانات حول عدد الأطفال الأحياء و سن ٣٠ امرأة متزوجة في سن الإنجاب ١٩– ٤٩ سنة.

المطلوب:

١- عرض هذه البيانات؟

٢-ما هو نوع الجدول و المتغير؟

السن	عدد الأطفال	السن	عدد الأطفال	السن	عدد الأطفال	السن	عدد الأطفال
١٩	–	٢٦	٢	٢٧	٢	٣٣	٤
١٨	١	١٥	–	٢٣	١	٣٧	٥
٢٥	١	١٦	–	٢٥	١	٤٠	٦
٤٠	٦	٢٠	١	٢٨	٢	٢٥	١
٤٩	٧	١٩	١	٢٩	٣	٢٨	٢
٤٥	٦	٢٠	١	١٥	–	٣٠	٢
٤٠	٣	٢١	٢	١٦	–		
٢٠	١	١٩	١	٢٨	١		

الحل:

عدد الأطفال / السن	٠ – ٢	٣ – ٥	٦ – ٨	المجموع
١٥ – ١٩	IIII III			٨
٢٠ – ٢٤	IIII			٥
٢٥ – ٢٩	IIII III	I		٩
٣٠ – ٣٤	I	I		٢
٣٥ – ٣٩		I		١

٤٠ – ٤٤		I	II	٣
٤٥ – ٤٩			II	٢
المجموع	٢٢	٤	٤	٣٠

٣

الجدول هو جدول تكراري مزدوج ذو متغيرين الأول مستقل و المتمثل في سن المرأة، و هو متغير كمي متصل، و الثاني متغير تابع متمثل في عدد الأطفال الأحياء، و هو متغير كمي منفصل.

التمرين الثالث:

الجدول التالي يبين توزيع وفيات الأطفال الرضع في مستشفى ما؟

السن	ك
٠ – ٧ أيام	٢
٨ – ٢٨ يوم	٥
٢٨ – ٣٦٥ يوم	١٢
المجموع	٧٠

السؤال: ما نوع الجدول و المتغير؟

الحل: الجدول بسيط ذو متغير واحد مبوب و مغلق و غير منتظم إذ أن الفئة الأولى طولها ٧ أيام.الفئة الثانية ٢٠ يوما.الفئة الثالثة ٣٣٧ يوما.

الفصل الثالث

أساليب عرض البيانات الإحصائية

أولا - العرض ألجدولي

١ - الجدول المزدوج

٢ - الجدول المضاعف التقاطع

٣- شروط عرض الجدول الإحصائي.

ثانيا- العرض البياني

١- الرسومات البيانية في حالة المتغيرات الكمية

- الخط البياني
- السلاسل الزمنية
- المدرج التكراري
- المضلع التكراري
- المنحنى التكراري

٢ الرسومات البيانية في حالة المتغيرات الكيفية

- المستطيل البياني
- الرسومات الدائرية و أجزاؤها
- الأعمدة البيانية

ثالثا- تمرينات محلولة حول الفصل الثالث

أولا – العرض ألجدولي.

بعد جمع البيانات الإحصائية و تبويبها كما رأينا تأتي مرحلة أخرى في البحث الإحصائي و هي عملية عرض هذه البيانات. إذ أن عرضها يتيح للقارئ أخذ فكرة سريعة عن الظاهرة المدروسة دون تعب أو جهد. و هناك طريقتين لعرض البيانات الإحصائية. إما بالعرض ألجدولي أو العرض البياني.

١- الجدول المزدوج:

يبنى الجدول الإحصائي بطريقة منظمة و ممنهجة. فقد يكون الجدول الإحصائي مشتملا على متغيرين فقط فيسمى بجدول بسيط التقاطع، أو مشتملا على ثلاث متغيرات فيسمى بجدول إحصائي مضاعف التقاطع. يوضع أحد المتغيرات عموديا عادة ما يكون المتغير المستقل و المتغير الثاني أفقيا وهو المتغير التابع.

وإذا كانت المتغيرات تشتمل على حد أدنى من القيم فالجدول التكراري المزدوج سيكون مشتملا على أربعة خانات مركزية هي خانات التقاطع، و أربعة خانات للتوزيع الهامشي خانتين أفقيتين للمتغير المستقل و خانتين عموديتين للمتغير التابع و خانة واحدة للمجموع العام. و هكذا يكون عدد الخانات في المجموع ٩ خانات في حالة أبسط جدول مزدوج.

إن الجدول الإحصائي لا بد أن يشتمل على كل التقاطعات الممكنة و التوزيعات الهامشية لأنها كلها ضرورية للتأويل الإحصائي، ثم أن لهذه الطريقة فائدة عملية إذ تسمح للباحث أن يكون ملما بكل المعطيات الضرورية للتحليل و بعد إتمام هذه العملية تأتي مرحلة أخرى هي:

- **قراءة الجدول إحصائيا:**

بعد بناء الجدول حسب الهندسة المذكورة يبقى الجدول غير قابل للقراءة إلا إذا نسب. و تعتبر هذه العملية من الناحية المعرفية أعقد عملية في التحليل المتعدد المتغيرات، لأنها هي التي تحدد نهائيا الطريقة الوحيدة لقراءة الجداول المناسبة للفرضيات.

و توجد ثلاث إمكانيات للتنسيب:

١- أفقيا ٢- عموديا ٣- بالنسبة للمجموع العام

لكن لا يهمنا عند تحليل الفرضيات التي تربط العلاقة بين متغيرين احدهم مستقل و الثاني تابع أن نقوم بالتنسيب انطلاقا من المجموع العام لأن ذلك يمحو نتائج التقاطع، فيبقى لنا إمكانية التنسيب الأفقي و العمودي. و يتم التنسيب الأفقي إذا كان المتغير المستقل عموديا و العكس صحيح فيكون التنسيب عموديا إذا كان المتغير المستقل أفقيا.

القاعدة: تنسيب الجدول يكون عكس وضع المتغير المستقل في الجدول

في حالة التنسيب الأفقي فان المقارنة تكون في نفس اتجاه المتغير المستقل أي بين خانة تقاطع (أج) و (ب ج) في المثال أدناه و تنطلق المقارنة في هذا الجدول انطلاقا من التوزيع الهامشي لنسب المتغير التابع أي بين ج(أب) و ك(أب) و هذا ما يسمى ب**الاتجاه العام للجدول** كما هو موضح في الجدول التالي:

المتغير التابع / المتغير المستقل	ج	ك	المجموع
أ	أج	أك	أ(ج ك)
ب	ب ج	ب ك	ب(ج ك)
المجموع	ج(أب)	ك(أب)	(أب)(ج ك)

مثال: كيفية تنسيب الجدول التكراري البسيط التقاطع.من الجدول التالي الذي يربط علاقة بين مكان إقامة المرأة بمدى استعمالها لوسائل منع الحمل.

استعمال وسائل منع الحمل / مكان الإقامة	تستعمل		لا تستعمل		المجموع	
	ك	%	ك	%	ك	%
حظر	٢٥	٧١,٤٣	١٠	٢٨,٥٧	٣٥	١٠٠
ريف	٥	٣٣,٣٣	٣٠	٦٦,٦٦	٤٥	١٠٠
المجموع	٤٠	٥٠	٤٠	٥٠	٨٠	١٠٠

٢- الجدول المضاعف التقاطع.

يبدأ التحليل المتعدد المتغيرات لما يدخل على الجدول ذو التقاطع البسيط الذي انتهى بتبيان علاقة أولية بين متغيرين (متغير مستقل و متغير تابع) متغير جديد يسمى المتغير الرائز، و هو أيضا متغير نسبي افتراضي جديد يرمي إلى اختبار العلاقة الأولى و إلى تكميم إمكانية تأثيرها سلبيا أم إيجابيا بدخوله في الجدول.

نظريا يدخل على هندسة الجدول التكراري البسيط التقاطع تغير بحيث يصبح المتغير الرائز وسيط بين المتغير المستقل و المتغير التابع في العلاقة الأولى. بحيث تنقسم البنية الأولى للجدول إلى بنيتين على الأقل بسبب تضاعف خانات الجدول ذو التقاطع البسيط.

فإذا كان الجدول المضاعف التقاطع يكتفي في أبسط أشكاله أي أن كل من المتغير المستقل، الرائز و التابع يأخذون قيمتين فقط. فسيكون مشتملا بالضرورة على ٨ خانات مركزية ذات التقاطع الثلاثي (م م، م ر، م ت). كما أن هذا الجدول كذلك يشترط فيه كي يكون مكتملا أن يشمل على الأربع خانات للجدول البسيط الأول و ٨ خانات هامشية للمجاميع المتغيرات الثلاث و خانة واحدة للمجموع العام.

أما عن تنسيبه و قراءته و مقارنته فيتم طبقا للقاعدة المذكورة سابقا في الجدول التكراري البسيط التقاطع. كما هو موضح في الجدول التالي:

م ت / م م + م ر		م ت١	م ت٢	المجموع
م م١	م ر١	م١ ر١ ت١	م١ ر١ ت٢	م١ ر١ ت١ ت٢
	م ر٢	م١ ر٢ ت١	م١ ر٢ ت٢	م١ ر٢ ت١ ت٢
م م١(م ر١ م ر٢)		م١ ت١	م١ ت٢	م١ ت١
م م٢	م ر١	م٢ ر١ ت١	م٢ ر١ ت٢	م٢ ر١ ت١ ت٢
	م ر٢	م٢ ر٢ ت١	م٢ ر٢ ت٢	م٢ ر٢ ت١ ت٢
م م٢(م ر١ م ر٢)		م٢ ت١(ر١ ر٢)	م٢ ت٢(ر١ ر٢)	م٢ ت١ ت٢
المجموع العام		ت١(م١ م٢)	ت م(م١ م٢)	م١(ت١ت٢)م٢(ت١ ت٢)

مثال: كيفية تنسيب الجدول المضاعف التقاطع. نقف دائما عند المثال السابق و هو ربط العلاقة بين مكان إقامة المرآة و مدى استعمالها لوسائل منع الحمل، إلا أننا نضيف متغيرا آخر رائزا و هو مدى موافقتها على استعمال وسائل منع الحمل. فسيكون الجدول الجديد كمايلي:

الاستعمال لوسائل منع الحمل / مكان الإقامة و مدى الموافقة		تستعمل		لا تستعمل		المجموع	
		ك	%	ك	%	ك	%
حضر	موافقة	٢٠	٩٠,٩١	٢	٩,٠٩	٢٢	١٠٠
	غير موافقة	٥	٣٨,٤٦	٨	٦١,٥٤	١٣	١٠٠
مجموع الحضريات		٢٥	٧١,٤٣	١٠	٢٨,٥٧	٣٥	١٠٠
ريف	موافقة	١٠	٦٦,٦٦	٥	٣٣,٣٣	١٥	١٠٠
	غير موافقة	٥	١٦,٦٧	٢٥	٨٣,٣٣	٣٠	١٠٠
مجموع الريفيات		١٥	٣٣,٣٣	٣٠	٦٦,٦٦	٤٥	١٠٠
المجموع		٤٠	٥٠	٤٠	٥٠	٨٠	١٠٠

إن النتائج المحتملة الممكن التوصل إليها من خلال إدخال متغيرا رائزا على الجدول البسيط التقاطع هي:

- إما زوال العلاقة الأولى بين المتغير المستقل و التابع تماما.
- إما تخفيض قيمة العلاقة الأولى بين المتغيرين.
- إما عدم تغير العلاقة الأولى بين المتغيرين في الجدول البسيط التقاطع.
- و إما ظهور علاقة جديدة.

و في مثالنا السابق لاحظنا في الجدول البسيط التقاطع، أنة أبرز علاقة أولية مفادها أن المرأة كلما كانت تقطن في منطقة حضرية إلا و زادت نسبة استعمالها لوسائل منع الحمل حيث قدرت بـــ ٧١,٤٣% و العكس صحيح أي أنها كلما سكنت أو كانت تقطن منطقة ريفية كلما قلت نسبة استعمالها لوسائل منع الحمل حيث قدرت بنسبة ٦٦,٦٦%.

و عندما أدخلنا متغيرا ثالثا رائزا ألا و هو موقف المرآة من استعمال هذه الوسائل، عثرنا على علاقة جديدة مفاده أن:

الحضريات الموافقات على استعمال وسائل منع الحمل هن اللواتي تستعملنها فعلا بكثرة حيث بلغت نسبتهن ٩٠,٩١%. فحين اللواتي لا توافقن منهن على استعمالها لا تستعملنها فعلا بنسبة ٦١,٥٤%.

و نفس الملاحظة تذكر على اللواتي تقطن في مناطق الريفية فالموافقات منهن تستعملن وسائل منع الحمل بكثرة بنسبة ٦٦,٦٦% أما اللواتي لا توافقن عليها فلا تستعملنها بنسبة ٨٣.٣٣%.

و بهذا فنلاحظ ظهور علاقة جديدة مفادها أن موقف المرآة من استعمال وسائل منع الحمل هو المحدد على مدى إقبالها على استعمال وسائل منع الحمل و ليس مكان سكناها أو إقامتها.

٣- شروط عرض الجدول الإحصائي.

هناك مجموعة من الشروط يجب توافرها في الجدول الإحصائي منها:

- أن يكون للجدول عنوانا كاملا مختصرا معبرا عن ما يحويه الجدول من بيانات.
- لا بد أن يكون لكل جدول رقم بحيث يسهل عملية الرجوع إليه مرة أخرى.
- أن يذكر اسم المتغيرين على الصف و العمود.
- أن تحدد الوحدات المستخدمة في الجدول مثلا:أيام، سنة، شهر،كلغ...الخ.
- إذا كانت هناك أي ملاحظات أو تفسيرات، مثلا في جدول لا يضم كل أفراد العينة كون بعضهم فقط معنيين بالمتغير المبوب، ففي هذه الحالة لا بد من وضع توضيح في الهامش. ففي المثال السابق أي في حالة الجدول البسيط المتقاطع كانت العينة متكونة من ٨٠ امرأة ٤٠ امرأة تستعمل وسائل منع الحمل و ٤٠ امرأة لا تستعمل وسائل منع

الحمل. و مثلا أردنا أن نضع جدول توزيع تكراري آخر لتبيين الوسائل المستعملة لتنظيم النسل. ففي هذه الحالة فالعينة ستكون ٤٠ امرأة مستعملة لوسائل منع الحمل فقط.
لهذا لا بد من توضيح سبب انخفاض حجم العينة في هامش الصفحة أين توضع الإحالات المرجعية على أن يشار إلى هذه التفسيرات بدون أرقام لأن الأرقام توضع فقط للمراجع المقتبس منها، و إنما في هذه الحالة نرمز لها مثلا برمز نجمة *.

ثانيا- العرض البياني.

من بين مهمات البحث الإحصائي عرض البيانات بشكل رسوم بيانية بعد تنظيمها في جداول معينة و مرتبة حسب الفرضيات التي افترضها الباحث في بداية بحثه، ذلك لأن الجداول التكرارية وحدها غير كافية لإعطاء معلومات سهلة التحليل و واضحة المعالم إلا بعد عرضها في رسوم بإمكانها أن تزيح الغموض عن التوزيع.

فالعرض البياني يبين لنا تطور الظاهرة المدروسة في زمن معين و في مكان معين بمجرد الملاحظة الأولى. فهي أحسن وسائل الإيضاح و أكثرها جاذبية في التعبير عن تغير و تطور الظاهرة المدروسة. فقد يجد بعض الناس صعوبة في فهم أو تتبع مجموعة من الأرقام في التوزيعات التكرارية، إذ لا يستهويهم العرض بالأرقام، لهذا نلجأ إلى استخدام الرسوم التوضيحية.

إلا أن هذه الرسومات البيانية المستخدمة تختلف باختلاف نوع البيانات أي المتغيرات و كذا نوع الجدول التكراري المراد عرضه. فأساليب العرض البياني اذا تتعدد و تتنوع حسب:

- طبيعة أو نوعية البيانات سواء أكانت مبوبة أو غير مبوبة.
- الهدف من العرض البياني.
- حسب نوع المتغير لأن لكل نوع من أنواع المتغيرات رسومات توضيحية خاصة.

* و يمكن عرض أهمية العرض البياني فيما يلي:

- الإفصاح عن خصائص الظاهرة المدروسة بصورة سريعة.
- إمكانية إجراء مقارنات بين توزيعات مختلفة كما في المدرج التكراري المتقابل و الأعمدة البيانية المزدوجة.
- استخلاص بعض المؤشرات الإحصائية في التوزيع التكراري دون استخدام الصيغ الرياضية.

* أما عن الشروط التي لا بد توفرها في الرسم البياني فهي:

- رسم محورين متعامدين متجانسين م س م ع، المحور م س يكون أفقيا يمثل الظاهرة المدروسة و المحور م ع يكون عموديا يمثل التكرارات. و يطلق على نقطة تقاطعهما نقطة الأصل أو نقطة الصفر، و تكون قيم المتغير المدروس (س) على يمين هذه النقطة موجبة و على يسارها سالبة. أما المحور م س فتكون القيم موجبة فوق نقطة الصفر و سالبة تحتها.
- لا بد من وضع عنوان و رقم لكل رسم يستحسن أن يكون قبل الرسم البياني نفسه.
- إذا كان الرسم يتطلب مفتاحا فيوضع هذا المفتاح.
- لا بد من ذكر سلم الرسم و لا يشترط مطلقا أن يعبر في الرسم عن كل وحدة في القيم بمسافة طولها ١ سنتمتر، بل قد نضطر في كثير من الأحيان إلى التعبير عن كل واحدة بجزء من السنتمتر، أو أكثر من السنتمتر، فاختيار الوحدات يتوقف على حيز الحجم الذي نرسم فيه، و القيم التي نريد تمثيلها.

و نعرض فيما يلي أهم الأشكال البيانية على أن نعقب ذلك بكيفية رسم كل واحد منها.

١- الرسومات البيانية في حالات المتغيرات الكمية.

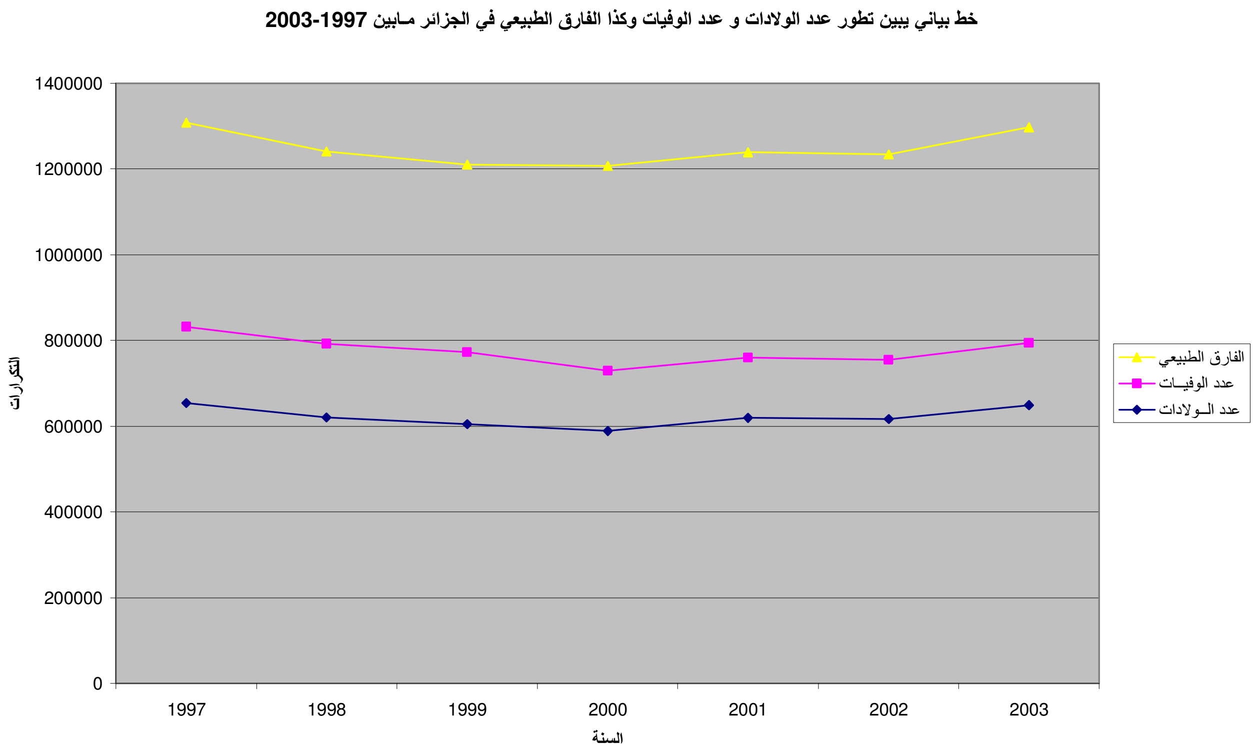
- **الخط البياني:**

يستخدم الخط البياني للتعبير عن التغير أو التطور لظاهرة ما خلال فترة زمنية محددة، أو لبيان العلاقة بين متغيرين أو أكثر، بحيث نمثل على المحور م س الأفقي الزمن و هو

(المتغير المستقل) و على المحور م ع العمودي / الرأسي قيم الظاهرة (المتغير التابع) طبعا مع الاختيار الأنسب لمقياس الرسم.

مثال: الجدول التالي يبين تطور عدد الولادات و الوفيات و الفارق الطبيعي بالألف في الجزائر ما بين ١٩٩٧ إلى ٢٠٠٣ و المطلوب تمثيلها بخط بياني.

	عدد الـــــولادات	عدد الوفيـــــات	الفارق الطبيعي
١٩٩٧	٦٥٤٠٠٠	١٧٨٠٠٠	٤٧٦٠٠٠
١٩٩٨	٦٢٠٠٠٠	١٧٢٠٠٠	٤٤٩٠٠٠
١٩٩٩	٦٠٥٠٠٠	١٦٨٠٠٠	٤٣٧٠٠٠
٢٠٠٠	٥٨٩٠٠٠	١٤٠٠٠٠	٤٧٨٠٠٠
٢٠٠١	٦١٩٠٠٠	١٤١٠٠٠	٤٧٩٠٠٠
٢٠٠٢	٦١٧٠٠٠	١٣٨٠٠٠	٤٧٩٠٠٠
٢٠٠٣	٦٤٩٠٠٠	١٤٥٠٠٠	٥٠٣٥٠٠

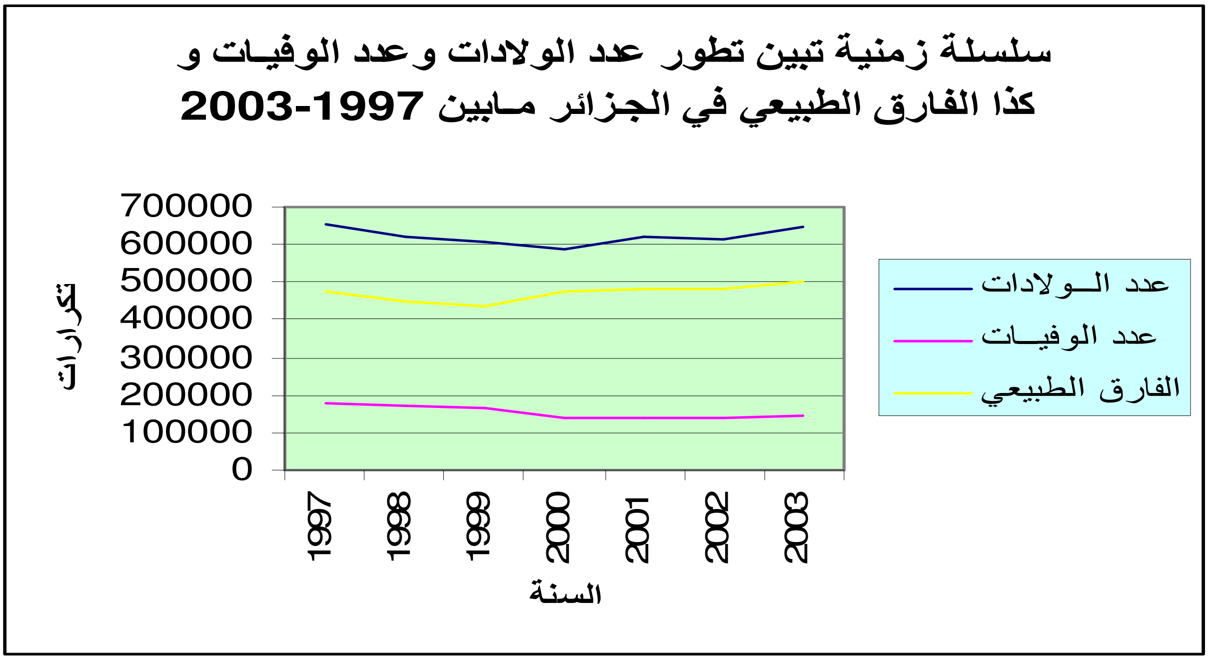


لرسم الخط البياني نقوم بوصل الثنائيات المركبة (س، ع) في المثال الساب(السنوات، التكرارات) بخطوط منكسرة وبهذا يختلف الخط البياني عن السلاسل الزمنية، إذ أن هذه الأخيرة نقوم بوصل النقاط فيها بخطوط منحنية أي بدون استعمال المسطرة.

السلاسل الزمنية: -

نستعمل السلاسل الزمنية في حالة المتغيرات الكمية المتصلة أي يكون تغيرها مع الزمن تغيرا متصلا، والسلسلة الزمنية تساعد بذلك في إظهار الاتجاه العام للظاهرة موضوع الدراسة، وفي هذا النوع من الرسومات تظهر قيمة الظاهرة في كل فترة بنقطة معينة على ارتفاع يمثل هذه القيمة ثم نصل هذه النقاط ببعضها البعض حسب تسلسلها الزمني بخط منحني. وبهذا تكون السلسلة الزمنية مشابهة للخطوط البيانية فقط هذه الأخيرة تكون متصلة بخطوط منكسرة.

ولنطبق على الجدول السابق الخاص بتطور وفيات والولادات والفارق الطبيعي في الجزائر ما بين ١٩٩٧- ٢٠٠٣ سلسلة زمنية.



نلاحظ من الرسم خط منعرج بين الصعود أحيانا و الهبوط أحيانا أخرى والسبب في ذلك هو أن الظاهرة الإحصائية لا تخضع في تغيراتها لقانون ثابت أو لعلاقة ثابتة. بل أنها تخضع لكثير من العوامل والمؤثرات، وهذه المؤثرات بعضها يحدث بشيء من الانتظام فيعطينا الاتجاه العام لنمو الظاهرة وبعضها يحدث بصورة عشوائية فيتسبب في هبوط وصعود السلسلة الزمانية.

المدرج التكـــــــراري: -

يستعمل في حالة المتغيرات الكمية المتصلة، وفي الجداول التكرارية المبوبة المغلقة، وهو عبارة عن مجموعة من المستطيلات موضوعة جنبا إلى جنب قاعدة كل منها تساوي طول الفئة التي يمثلها وارتفاعها يساوي التكرار المقابل للفئة، والمقارنة بين التكرارات تتم على أساس مقارنة مساحة المستطيلات حيث أن مجموع مساحة المستطيلات تعبر عن التكرار الكلي/ العينة، وتختلف طريقة العمل في حالة الفئات المتساوية عنها في حالة الفئات غير المتساوية.

وفي هذا الشكل نرسم محورين متعامدين متجانسين، حيث يأخذ المحور الأفقي عادة لتمثيل الفئات والمحور الرأسي لتمثيل التكرارات، ثم نقسم المحور الأفقي إلى أقسام متساوية بمقياس رسم مناسب بحيث يكفي لتمثيل جميع الفئات ونقوم بتدريج المحور الرأسي حسب مقياس رسم مناسب، ثم نرسم على كل فئة مستطيلا يتناسب مع التكرار الخاص بالفئة وتمتد قاعدة المستطيل على المحور الأفقي من أول الفئة إلى آخرها فنحصل بذلك على شكل مستطيلات متلاصقة.

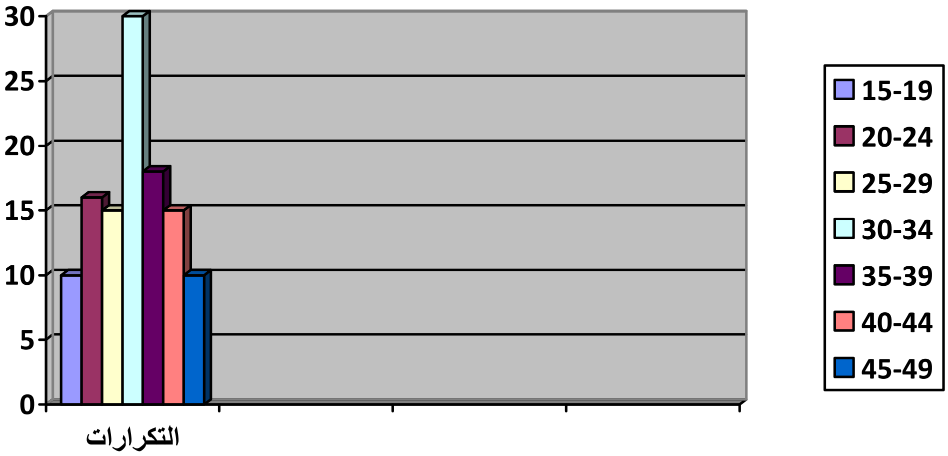
أ- المدرج التكراري في حالة الجداول التكرارية المنتظمـــة.

نرسم محورين متعامدين متجانس بحيث يمثل المحور العمودي التكرارات والمحور الأفقي على الفئات، ثم نمثل كل فئة بمستطيل قاعدته طول الفئة وارتفاعها تكرارها فنحصل على المدرج التكراري كما هو موضح في المثال التالي.

مثل: تمثيل التوزيع التكراري التالي الذي يبين توزيع مجموعة من النساء المستعملات لوسائل منع الحمل بمدرج تكراري.

الســــــن	التكـــــــرارات
١٥ – ١٩	١٠
٢٠ – ٢٤	١٦
٢٥ – ٢٩	١٥
٣٠ – ٣٤	٣٠
٣٥ – ٣٩	١٨
٤٠ – ٤٤	١٥
٤٥ – ٤٩	١٠
المجموع	١١٤

مدرج تكراري يبين توزيع مجموعة من النساء المستعملات لوسائل منع الحمل حسب السن



ب- المدرج التكراري في حالة التوزيعات التكرارية الغير المنتظمــــة:

وفيه نمثل كل فئة بمستطيل قاعدته طول الفئة وارتفاعه تكرارها المعدل وللحصول على **<u>التكرار المعدل</u>** نقسم التكرارات الأصلية على طول كل فئة المقابلة لها حتى تكون مساحات المستطيلات معبرة عن التكرارات التي تمثلها.

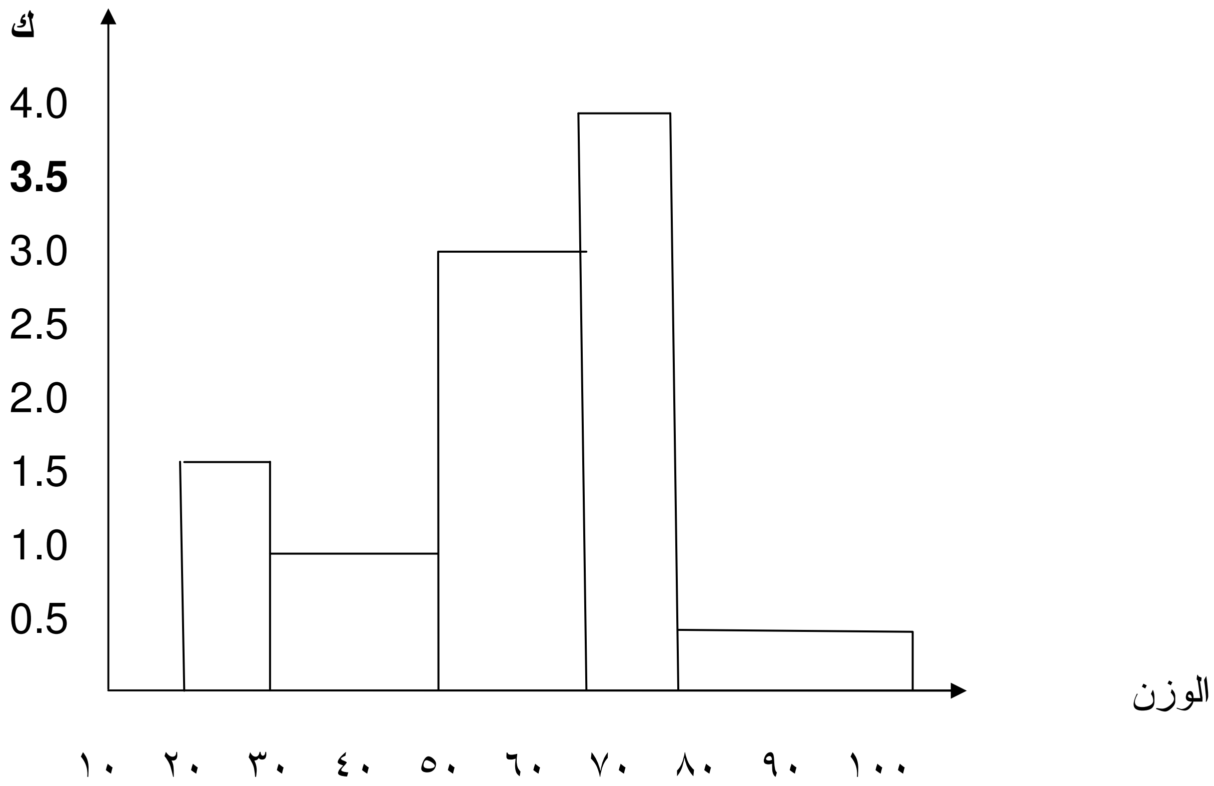
مثال: الجدول التالي يبين توزيع ١٥٠ مصاب بمرض السرطان حسب وزنهم والمطلوب تمثيلهم على مدرج تكراري.

الفئـــات الـــوزن	التكـــرارات
١٠ – ٢٠	١٥
٢٠ – ٤٠	٢٠
٤٠ – ٦٠	٦٠
٦٠ – ٧٠	٤٠
٧٠ – ١٠٠	١٥
المجموع	١٥٠

نلاحظ من خلال هذا الجدول أن فئاته غير متساوية لهذا فلابد من تعديل التكرارات كما نلاحظ:

الـــــــوزن	ك	طول الفئـــة	التكرار المعدل
١٠ – ٢٠	٢٠	10	١،٥
٢٠ – ٤٠	٢٠	٢٠	١
٤٠ – ٦٠	٦٠	٢٠	٣
٦٠ – ٧٠	٤٠	١٠	٤
٧٠ – ١٠٠	١٥	٣٠	٠،٥
المجمـــــوع	١١٥		

مدرج تكراري يبين توزيع أفراد العينة حسب وزنهــــــــم



جـ - المدرج التكراري في حالة توزيعين مختلفيـــــــن.

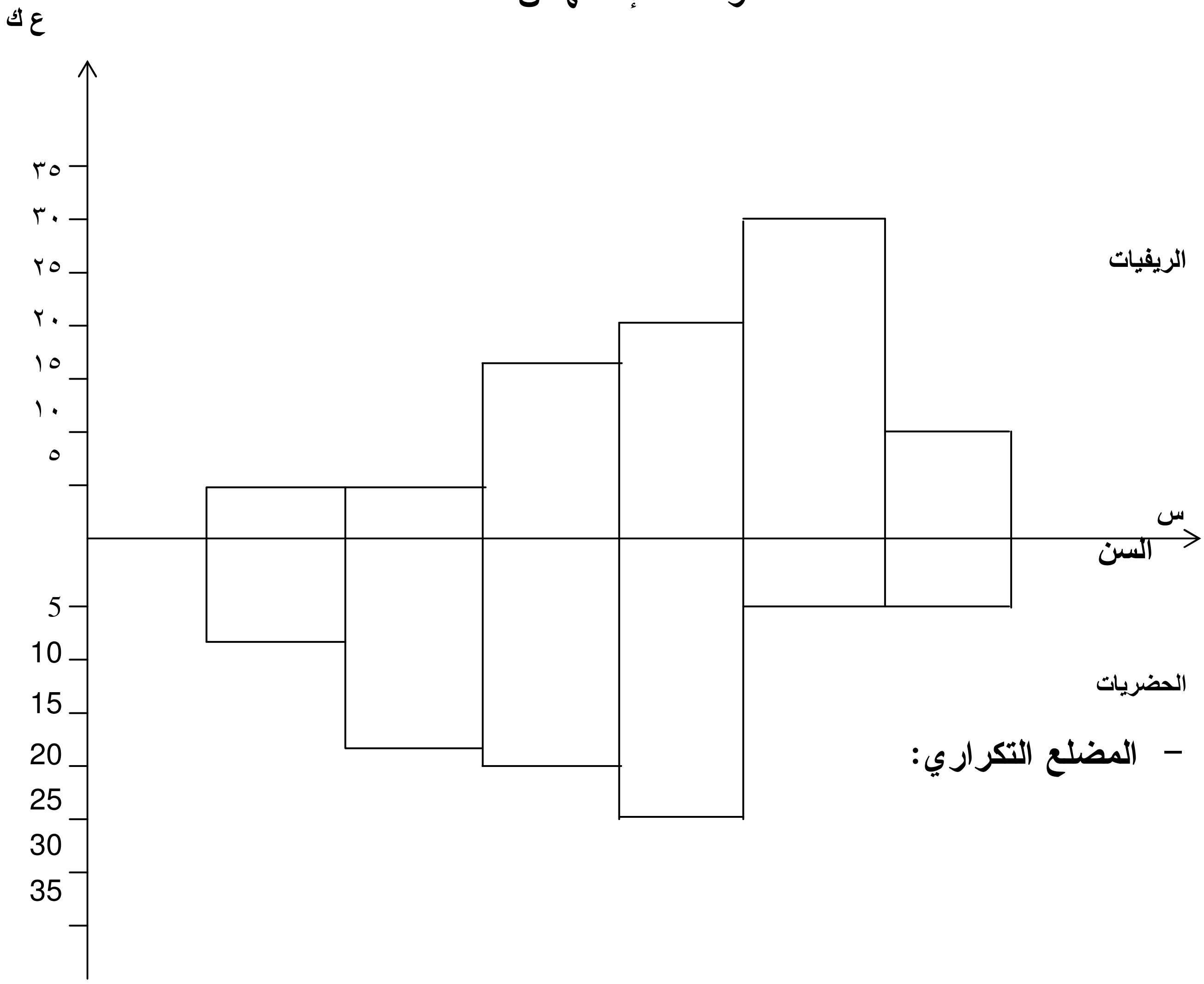
يمكن استخدام المدرج التكراري للمقارنة بين توزيعين مختلفين شريطة أن تكون العينة متساوية في كل التوزيعين وذلك باستخدام جهتي المحور الأفقي م س بحيث يرسم أحد المدرجين فوقه والآخر تحته.

مثـــال: الجدول التالي يبين لنا توزيع عينة من النساء المشتغلات في الحرف اليدوية حسب سنهن ومنطقة سكنهن والمطلوب تمثيلهن برسم مدرج تكراري.

الســـن	الريفيـــــات	الحضريـــات
٢٠ – ٢٥	٣	٨
٢٥ – ٣٠	٤	١٩
٣٠ – ٣٥	16	٢٠
٣٥ – ٤٠	٢٠	٢٥
٤٠ – ٤٥	٣٠	٤

٤٥ – ٥٠	٧	٤
المجـــــموع	٨٠	٨٠

مدرج تكراري يبين توزيع عينة من النساء المشتغلات في الحرف اليدوية حسب سنهن ومنطقة إقامتهــن



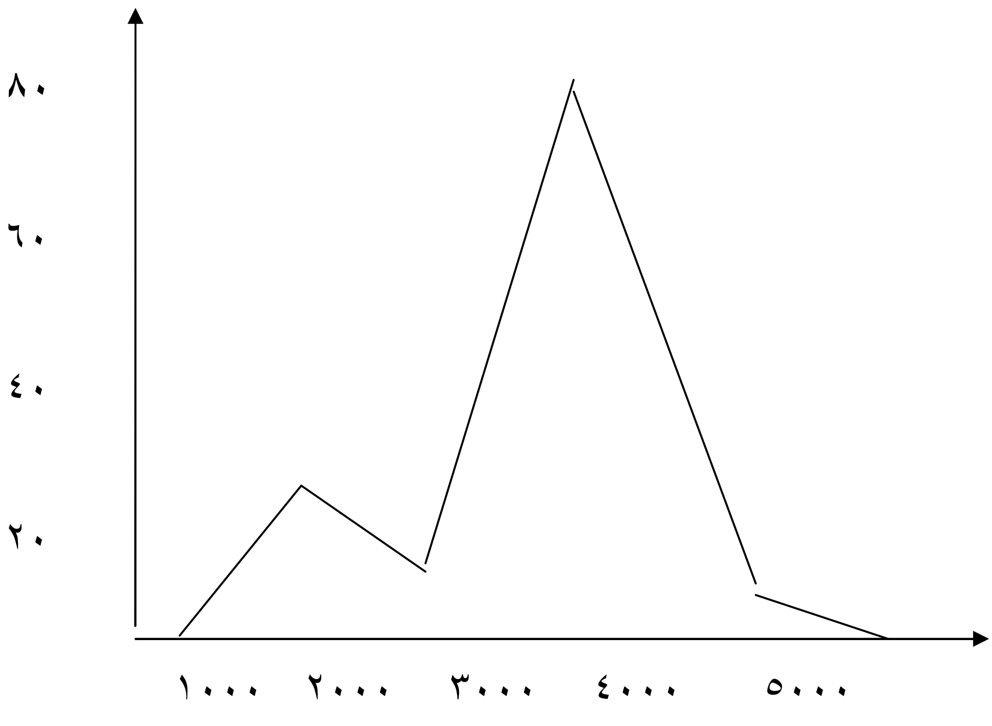
- **المضلع التكراري:**

في هذا الشكل نقوم بتقسيم المحورين م س، م ع، كما في حالة المدرج التكراري تماما، ثم نحدد مراكز منصفات الفئات على المحور الأفقي ونرمز بنقطة في المستوي إحداثياتها الأفقية هي مراكز الفئات وإحداثياتها العمودية هي التكرارات المناظرة لها. ثم نصل بين هذه النقاط بمستقيمات أي أننا نستعمل المسطرة، فنحصل على المضلع التكراري وهذا باعتبار أن التكرارات في كل فئة تتجمع أو ترتكز عند مركز الفئة. ويستحسن غلق المضلع التكراري مع المحور الأفقي، وذلك بأن نفترض وجود فئة وهمية قبل الفئة الأولى بالجدول تساوي بقية الفئات في الطول وتكرارها يساوي صفرا. وفئة وهمية أخرى بعد آخر فئة في الجدول متساوية هي الأخرى في الطول لبقية الفئات وتكرارها يساوي صفرا كذلك.

فإذا رسمنا النقطتين الممثلتين اللتين تقعان على المحور م س عند مركز الفئتين وثم إيصالها بطرفي المضلع التكراري يتم قفله كما هو مبين في المثال التالي الذي يوضع وزن ١٤٠مولود عند ولادتهم.

التكـــــرارات	الوزن عنـــد الولادة
٢٥	١٠٠٠ – ٢٠٠٠
٢٠	٢٠٠٠ – ٣٠٠٠
٨٠	٣٠٠٠ – ٤٠٠٠
١٥	٤٠٠٠ – ٥٠٠٠
١٤٠	المجموع

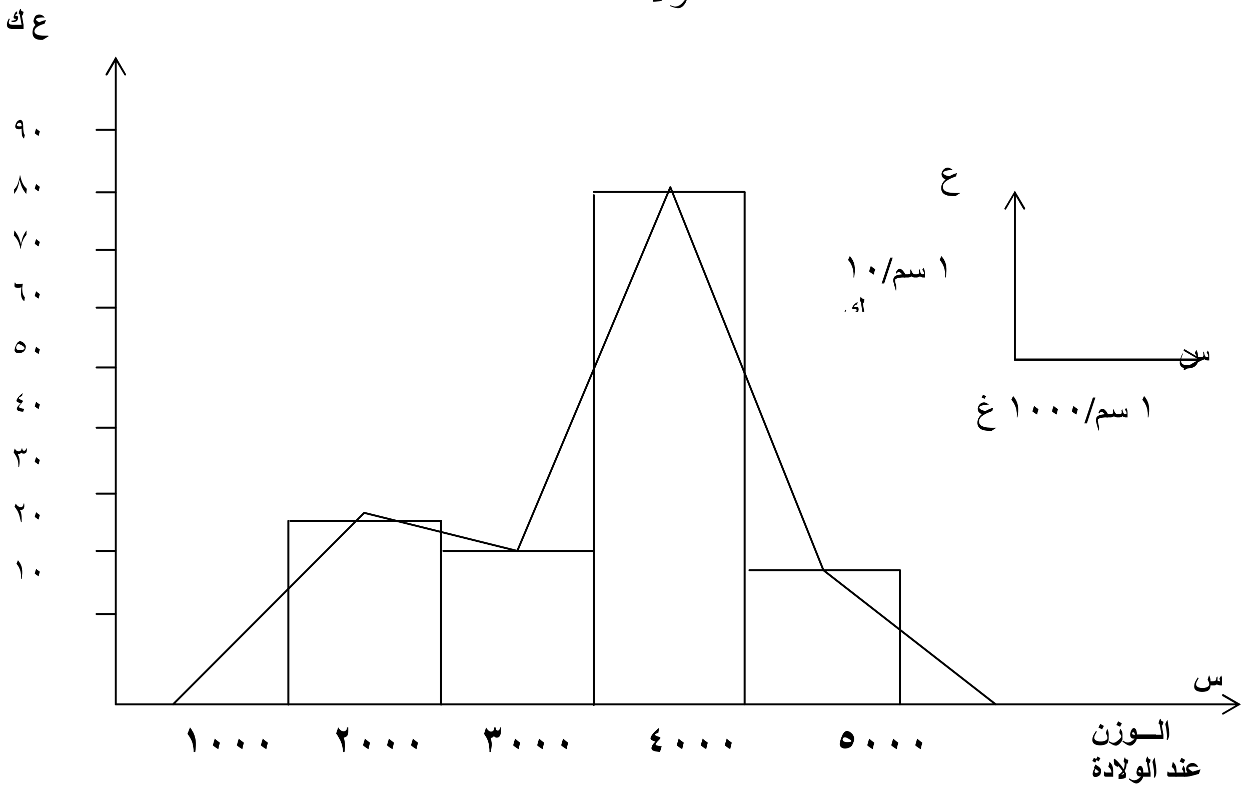
مضلع تكراري يبين وزن ١٤٠مولود عند ولادتهم.



كما يمكننا رسم المضلع التكراري إبتداءا من المدرج التكراري وذلك بأخذ منتصفات القواعد العليا للمستطيلات في المدرج التكراري ونصلها بمستقيمات منكسرة فنحصل على مضلع تكراري كما هو مبين في الرسم التالي:

مدرج ومضلع تكراري يبين توزيع ١٤٠ مولود حسب وزنهم

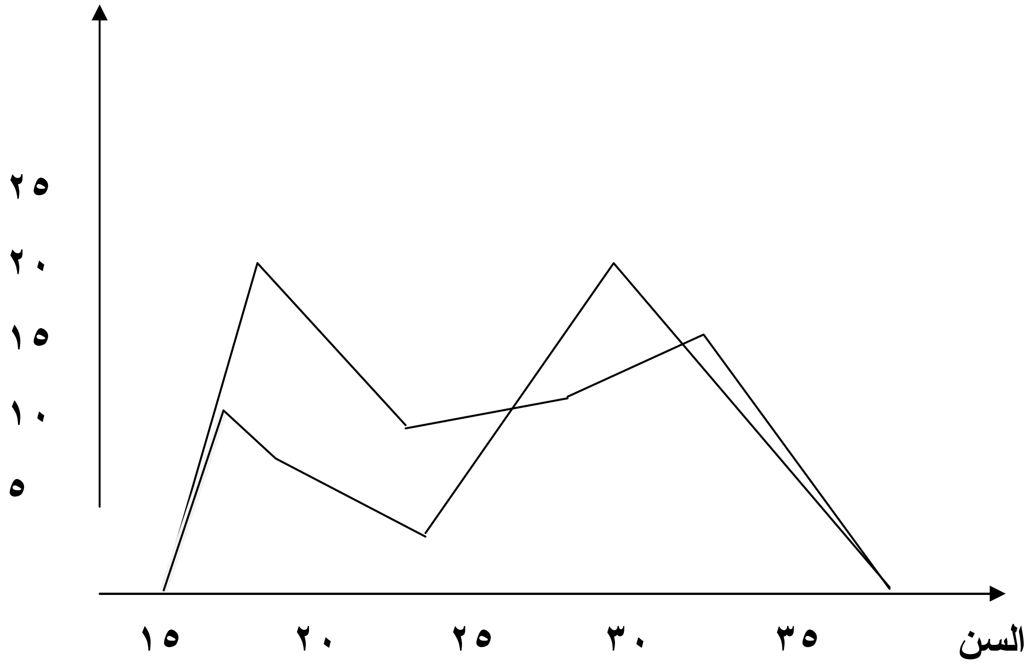
عند الــــــــولادة



كما يمكن رسم المضلع التكراري للمقارنة بين توزيعين مختلفين حيث تطبق نفس الطريقة السابقة في حالة توزيع تكراري الأحادي، لكن في هذه المرة تطبق مرتين وكمثال على ذلك لدينا الجدول التالي الذي يبين توزيع ١١٤ فرد حسب سنهم ورأيهم في ظاهرة التدخين أي هل هو مضر بالصحة أم لا.

الـــــرأي / الســـــن	نعم	لا
١٥ – ٢٠	٢٠	١٣
٢٠ – ٢٥	١٢	٨
٢٥ – ٣٠	١٤	٥
٣٠ – ٣٥	١٨	٢٤
المجمـــــــوع	٦٤	٥٠

مضلع تكراري يبين توزيع أفراد عينة حسب سنهم ورأيهم في ظاهرة التدخين



- **المنحنى التكـــراري:**

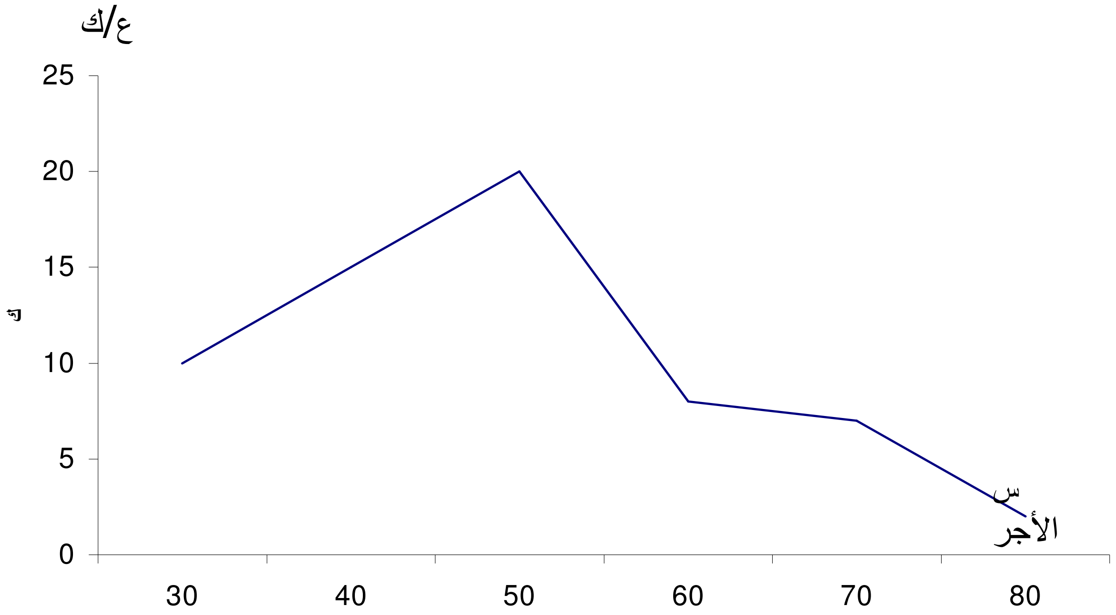
هناك طريقة أخرى لتمثيل التوزيعات التكرارية المغلقة ذات المتغيرات الكمية المتصلة في شكل هندسي واضح. وذلك برسم المنحنى التكراري الذي نحصل عليه بتمديد خطوط المضلع التكراري المنكسرة. ولرسم المنحنى التكراري نمثل بنقاط في المستوي إحداثياتها الأفقية هي مراكز الفئات وإحداثياتها العمودية هي التكرارات المناظرة لها، ثم نمدد هذه النقاط بخط منحني. كما أن المنحنى التكراري لا يغلق مع المحور السيني الأفقي كما هو الحال في المضلع التكراري.

مثـــال: يبين التوزيع التالي إيجار سكن الذي تدفعه ٦٢ أسرة في الأسبوع اليورو € والمطلوب تمثيله بالمنحنى التكـــراري.

فئات الإيـــجار	التكـــــــرارات
٣٠ – ٤٠	١٠
٤٠ – ٥٠	١٥
٥٠ – ٦٠	٢٠
٦٠ – ٧٠	٨

٧٠ – ٨٠	٧
٨٠ – ٩٠	٢
المجمـــــوع	٦٢

منحنى تكراري يبين توزيع ٦٢ أسرة حسب الإيجار السكن الذي تدفعه في الأسبوع



يختلف شكل المنحنى التكراري باختلاف نوع البيانات التي تزيد تمثيلها وكذا الأعراض الدراسة العملية، ويتم تصنف المنحنيات حسب مدى الإللتوائها فمنها المتماثلة والملتوية، وكذا حسب مدى تفرطحها، فمنها المفرطحة والمدببة وكذا حسب صيغتها الرياضية من أهمها التوزيع الطبيعي توزيع (ت) وتوزيع (كا^{2})... إلخ.

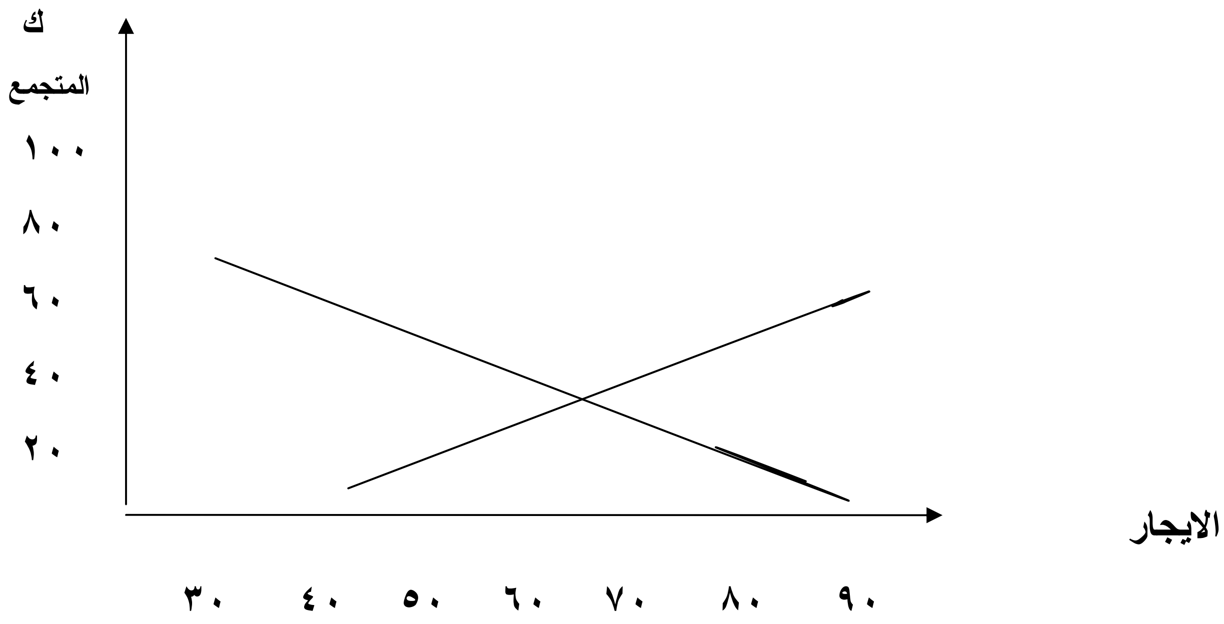
– المنحنى التكراري المتجمـــــع:

نستخدم المنحنى التكراري المتجمع النازل والصاعد لتمثيل التكرارات المتجمعة الصاعدة والنازلة بيانيا. ويرسم بوضع نقاط إحداثياتها الأفقية تمثل الحدود العليا أم الدنيا للفئات(الحد الأعلى في المنحنى المتجمع الصاعد والحد الأدنى للفئة في حالة رسم المنحنى المتجمع النازل)، أما إحداثياتها العمودية أو الرأسية فتمثل التكرار المتجمع

النازل أم الصاعد، ثم نقوم بإيصال هذه النقاط باليد وليس بخطوط مستقيمة، كما هو مبين في المثال التالي:

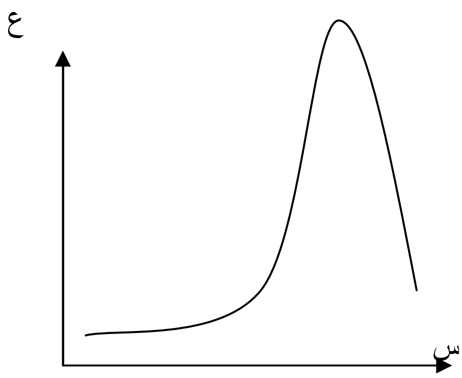
فئات الأجر	ك	التكرار المتجمع الصاعد	التكرار المتجمع النازل
٣٠ – ٤٠	١٠	١٠	٦٢
٤٠ – ٥٠	١٥	٢٥	٥٢
٥٠ – ٦٠	٢٠	٤٥	٣٧
٦٠ – ٧٠	٨	٥٣	١٧
٧٠ – ٨٠	٧	٦٠	٩
٨٠ – ٩٠	٢	٦٢	٢
المجمـــــوع	٦٢		

منحنى تكراري متجمع نازل وصاعد يمثل توزيع ٦٢ أسرة حسب الإيجار الأسبوعي للسكن

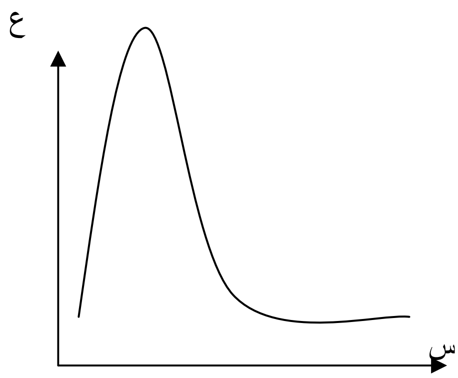


أنواع المنحنيـــــــات: -

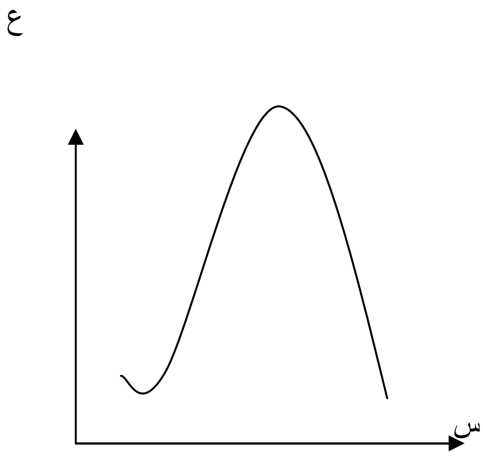
١/- من حيث الالتواء والتماثـــــل:



منحنى سالب الالتواء
منحنى ملتوي نحو
اليسار

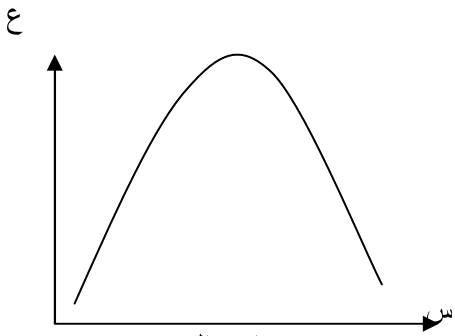


منحنى موجب الالتواء
منحنى ملتوي نحو
اليمين

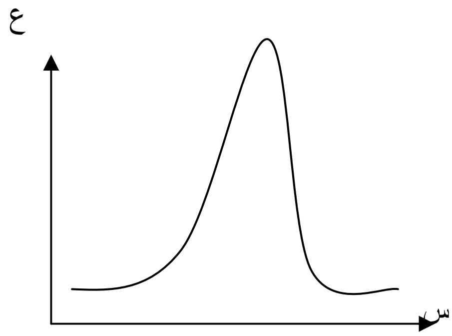


منحنى متماثل
منحنى متماثل حول الوسط

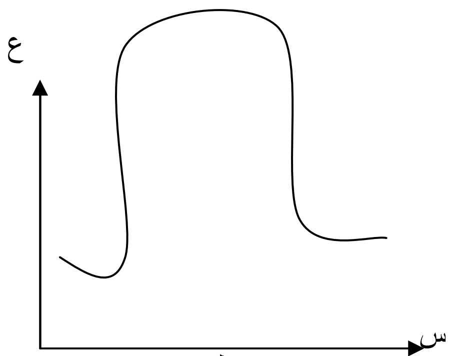
٢/- من حيث التدبيـــــب:



منحنى اعتدالـــــي

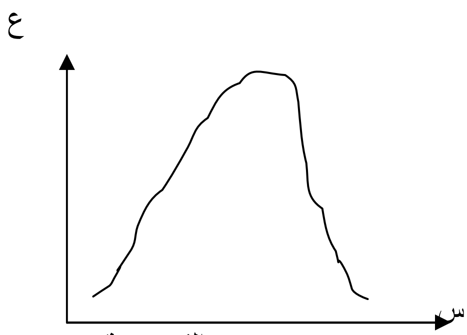


منحنى مدبب

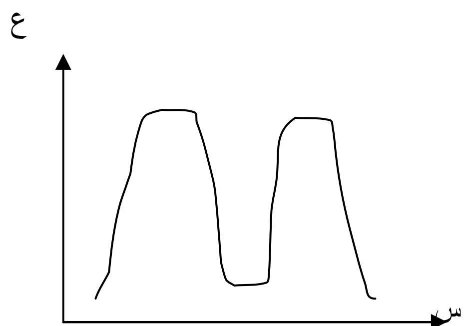


منحنى مفرطـــــح

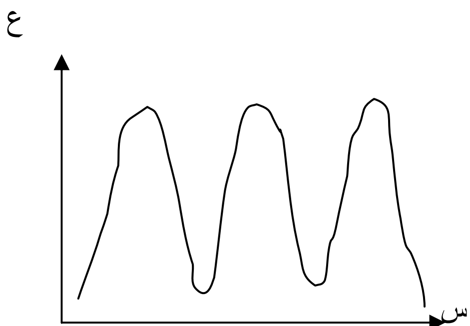
٣/- من حيث عدد القيـــــم:



منحنى وحيد القمـــــة



منحنى ثنائي القمـــــة



منحنى متعدد القمم

٢- الرسومات البيانية في حالة المتغيرات الكيفيـــــة:

- المستطيل البيانـــــي:

يستعمل المستطيل البياني عندما يرغب الباحث تمثيل ظاهرة كيفية أو كمية لكن منفصلة، وهو أكثر الأشكال استعمالا وسهولة، وفيه يقوم برسم مستطيل عادة طوله ١٠ سم، ثم نقسمه إلى أقسام متساوية في الارتفاع وقاعدة كل منها عبارة عن القيمة المراد إظهارها على الرسم، والتي تحسب وفق القانون التالي:

$$\text{قاعدة المستطيل كله} \times \frac{\text{البيان الجزئي}}{\text{البيان الكلي}}$$

مثـــال: بلغت وفيات الأطفال الرضع في مستشفى ما ١٠٠ حالة موزعين حسب الأسباب التالية والمطلوب تمثيلها برسم مستطيل بياني.

الأسبـــــاب	ك
اختناقات تنفسية	٢٥
الإصابة بالإسهال	٤٠
ولادة قبل الأوان	٢٠
أسباب متصلة بالقلب والأوعية	١٥
المجمـــــــــوع	١٠٠

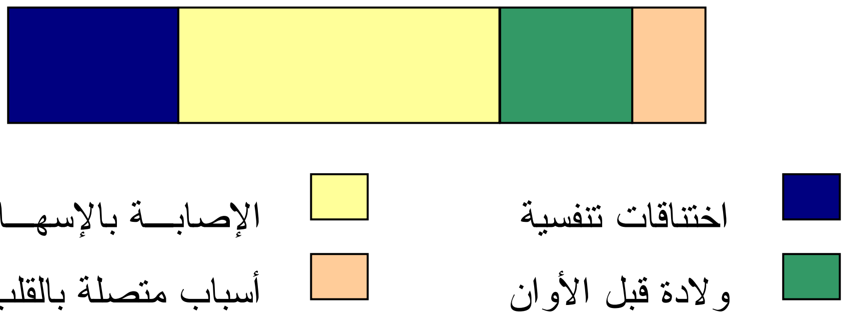
١- حساب قاعدة الاختناقات التنفسية = $\frac{٢٥}{١٠٠} \times ١٠ = ٢,٥$

٢- حساب قاعدة الإصابـــة بالإسهال = $\frac{٤٠}{١٠٠} \times ١٠ = ٤$

٣- حساب قاعدة الولادة قبـــل الأوان = $\frac{٢٠}{١٠٠} \times ١٠ = ٢$

٤- حساب قاعدة الأسباب المتصلة بالقلب والأوعية = $\frac{١٥}{١٠٠} \times ١٠ = ١,٥$

مستطيل بياني يبين توزيع أسباب ١٠٠ طفل متوفى بالمستشفى



- الرسومات الدائرية وأجزائها:

هي شكل من الرسومات البيانية المساحية تستخدم خاصة في حالة البيانات الكيفية التي يوجد فيها تفاوت كبير بين أرقامها، أين لا يمكن تمثيلها بمستطيل بياني أو أعمدة بيانية.
وهي تستخدم عادة للمقارنة بين المكونات المختلفة لظاهرة معينة ببعضها البعض، كما تستخدم للمقارنة بين ظاهرتين أو أكثر، أو مقارنة ظاهرة واحدة بنوعيها خلال فترات زمنية متفاوتة لإظهار الاختلاف بين المجموع الكلي لقيم الظاهرة أو من ظاهرة إلى أخرى.
فهذا النوع من الرسومات من أفضل الطرق لتمثيل البيانات ذات الصفة المشتركة ونستطيع من خلالها أن نقارن أجزاء الدائرة ببعضها البعض.

أ- تمثيل البيانات بالدائـــــرة:

إذا كانت لدينا بيانات كيفية مقسمة إلى قيم، فيمكن تمثيلها بالمساحة الكلية للدائرة، حيث نقسم هذه الأخيرة إلى قطاعات تتلاقى في المركز بحيث تكون متناسبة مع المقادير الجزئية التي تكون في مجموعها المساحة الكلية للدائرة، ونميز هذه القطاعات عليها بألوان مختلفة.

أما عن خطوات تمثيل البيانات بدائرة نسبية فيكون كما يلي:

- تحويل البيانات الفرعية/ الجزئية إلى نسب مئوية من المجموع الكلي للبيان.
- بما أن الزاوية المركزية /الكلية في الدائرة تقدر بـــ ٣٦٠ ° فإن ١% من مساحة الدائرة يمثله قطاع زاوي قدره ٣,٦ ° لذلك يمكن تمثيل أجزاء المجموع الكلي بقطاعات مساحية كل منها عبارة عن النسبة المئوية للبيان الجزئي× ٣,٦. كما هو موضح في المثال التالي:

مثال: لدينا الجدول التالي الذي يبين توزيع ٢٥٠ فرد حسب حالتهم المدنية والمطلوب تمثيلهم بدائرة نسبية.

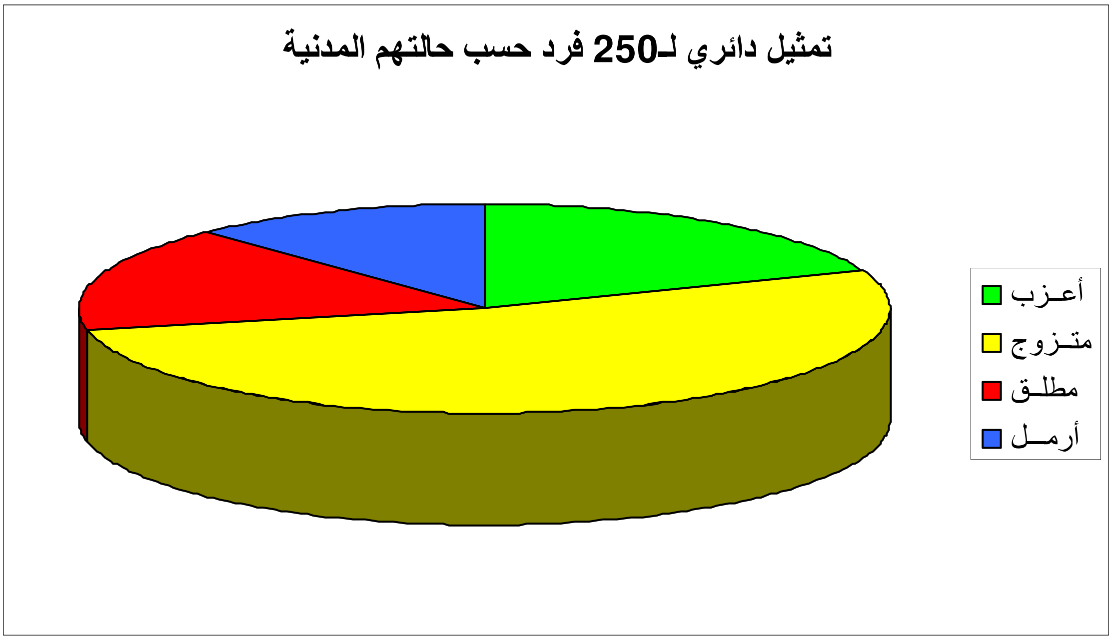
الحالة المدنية للفرد	ك
أعـــزب	٤٨
متــــزوج	١٣١
مطلـــــق	٤١
أرمـــــــل	٣٠
المجـــموع	٢٥٠

إذن فالخطوة الأولى لرسم الدائرة هي تحويل القيم الفرعية إلى نسب مئوية من المجموع الكلي مثلا:

$$\frac{٤٨}{٢٥٠} \times ١٠٠ = ١٩,٢$$

ثم تحويل قيمة كل نسبة بيان جزئي إلى قطاع زاوي ذلك بضربه × ٣,٦° مثلا ١٩,٢ × ٣,٦° = ٦٩,١٢ كما هو مبين في الجدول التالي:

الحالة المدنية	ك	النسبة	تحويل النسب إلى درجات مئوية
أعزب	٤٨	١٩،٢	٦٩،١٢ = ٣،٦ × ١٩,٢°
متزوج	131	52،4	52.4١٨٨،٦٤ = ٣،٦ ×
مطلق	٤١	١٦،٤	٥٩،٠٤ = ٣،٦ × ١٦,٤
أرمل	٣٠	١٢،٠٠	٤٣،٢٠ = ٣،٦ × ١٢,٠٠
المجموع	٢٥٠	١٠٠	٣٦٠ = ٣،٦ × ١٠٠



ب- تمثيل البيانات بنصف دائرة نسبيــــــــة.

إن شروط استعمال نصف دائرة مشابهة لشروط استعمال الدائرة النسبية الكلية ونفس الخطوات تتبع الفرق المتواجد بينهما فقط يمكن في أن النصف دائرة قطاعها الزاوي الكلي = ١٨٠° أي أن ١% من مساحة الدائرة يقابلها قطاع زاوي مقداره ١,٨° كما هو مبين في المثال التالي:

الحالة المدنية	ك	النسبة	تحويل النسب إلى درجات مئوية
أعزب	٤٨	١٩،٢	١٩،٢ × ١،٨ = ٣٤,٥٦
متزوج	131	52,4	52,4× ١،٨ = ٩٤,٣٢
مطلق	٤١	١٦،٤	١٦،٤ × ١،٨ = ٢٩,٥٢
أرمل	٣٠	١٢،٠٠	١٢،٠ × ١،٨ = ٢١,٦٠
المجموع	٢٥٠	١٠٠	١٠٠ × ١،٨ = ١٨٠

ج- تمثيل البيانات بربع دائـــــــــــرة:

التمثيل بربع دائرة هو الآخر يتبع نفس الشروط والخطوات المتبعة في الدائرة النسبية أو نصف الدائرة إلا أننا في هذه الحالة نجد أن مساحة الربع دائرة الكلية ١٠٠% يقابلها قطاع زاوي قدره ٩٠°. إذ نجد كل ١% من مساحة الدائرة يقابلها قطاع زاوي في هذه الحالة مقدر بـــ ٠,٩°، كما هو مبين في المثال التالي:

الحالة المدنية	ك	النسبة	تحويل النسب إلى درجات مئوية
أعزب	٤٨	١٩،٢	١٩،٢ × ٠،٩ = ١٧،٢٨
متزوج	131	52,4	52,4× ٠،٩ = ٤٧،١٦
مطلق	٤١	١٦،٤	١٦،٤ × ٠،٩ = ١٤،٧٦
أرمل	٣٠	١٢،٠٠	١٢،٠٠ × ٠،٩ = ١٠،٨٠
المجموع	٢٥٠	١٠٠	١٠٠،٠٠ × ٠،٩ = ٩٠،٠٠

- الأعمدة البيانيـــــــــة:

تستعمل أشكال الأعمدة البيانية باختلافها في حالة المتغيرات الكمية المنفصلة أو المتغيرات الكيفية، وتستخدم للمقارنة بين قيم ظاهرة واحدة أو عدة ظواهر.

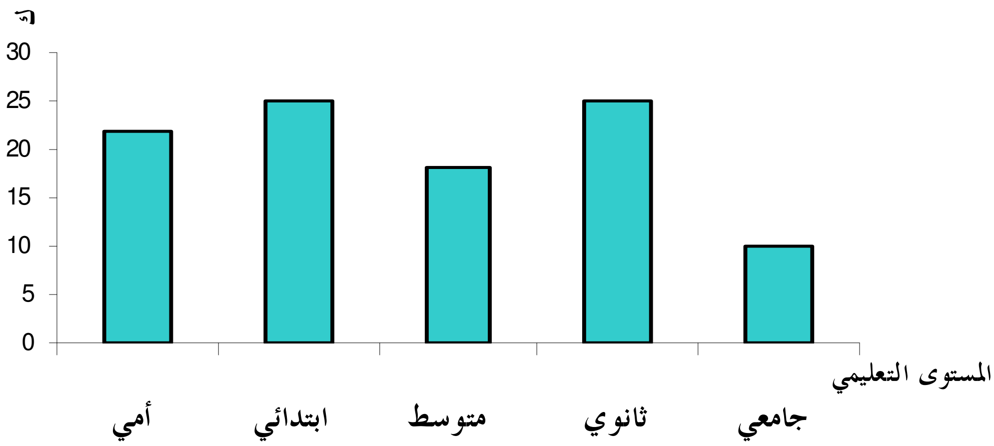
والأعمدة البيانية ما هي إلا أعمدة مستطيلات رأسية تتناسب ارتفاعاتها مع الأعداد التي تمثلها الأعمدة وتكون قواعدها متساوية في العرض، ويؤخذ المحور م س الأفقي لتمثيل المتغير أو الظاهرة المدروسة بينما نتبث على المحور العمودي م ع القيم المختلفة وهناك عدة أنواع من الأعمدة البيانية أهمها ما يلي:

أ- **الأعمدة البيانية البسيطـــــة:**

يستعمل هذا النوع من الأعمدة لتمثيل متغير واحد فقط، فهي تبين اختلاف قيم الظاهرة من زيادة ونقصان وفي محل هذه الحالة نقوم برسم محورين متعامدين متجانسين م س، م ع نثبت على المحور م س قيم المتغير وعلى المحور م ع التكرارات المقابلة لها، كما هو مبين في المثال التالي، الذي يبين توزيع مجموعة من المبحوثات حسب مستواهن التعليمي

المستوى التعليمي	ك
أمـــــي	٢٢
ابتدائـــــي	٢٥
متوســـط	١٨
ثانـــــوي	٢٥
جامعـــــي	١٠
المجمـــــــوع	١٠٠

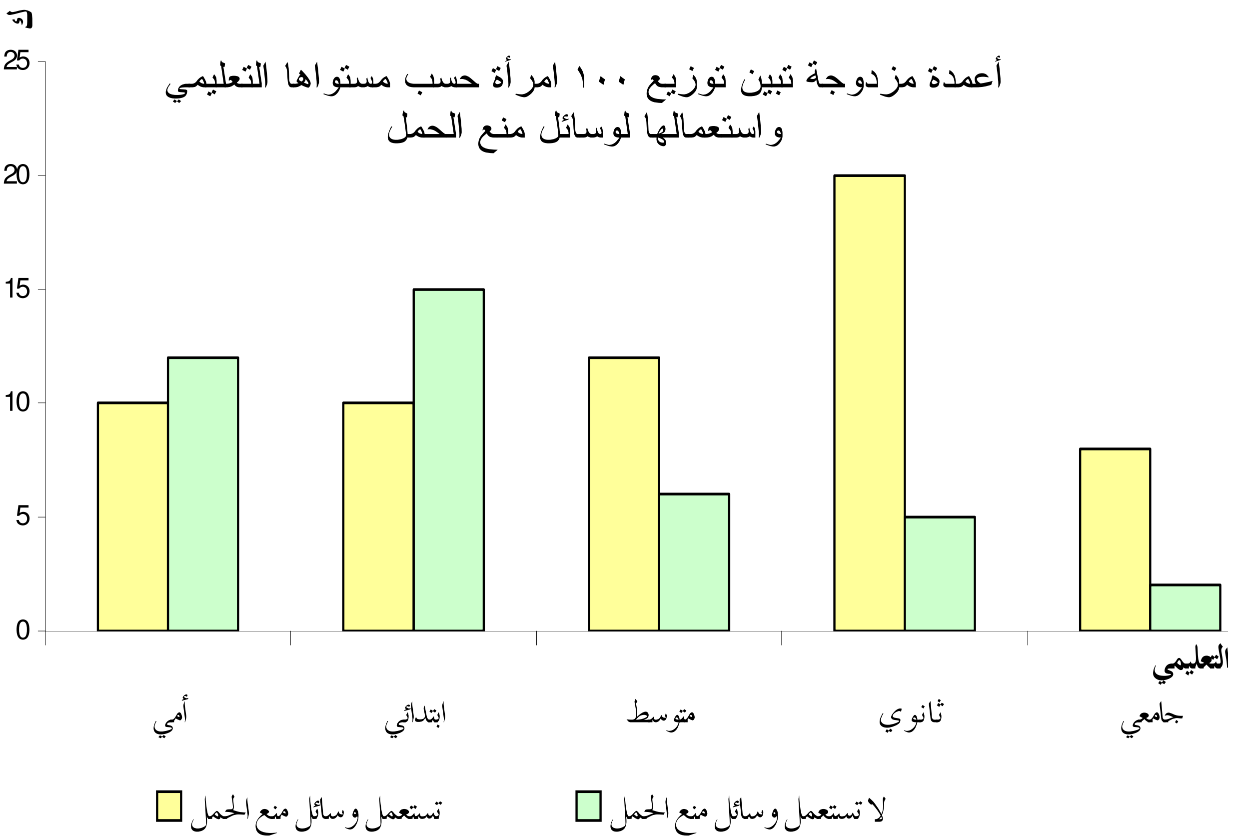
أعمدة بيانية تبين توزيع مجموعة من النســــاء حسب مستواهن التعليمـــي



ب ــ- الأعمدة البيانية المزدوجة :

يمكن للأعمدة البيانية أن تستخدم للمقارنة بين توزيعين أو أكثر بحيث يتلاصق كل عمودين متفقين في فئة معينة، على أن يستخدم لكل ظاهرة لون مختلف كما هو مبين في المثال التالي، الذي يوضح توزيع ١٠٠ امرأة حسب مستواهن التعليمي ومدى استعمالهن لوسائل منع الحمل.

المستوى التعليمي	تستعمل وسائل منع الحمل	لا تستعمل وسائل منع الحمل	المجموع
أمي	١٠	١٢	٢٢
ابتدائي	١٠	١٥	٢٥
متوسط	١٢	٦	١٨
ثانوي	٢٠	٥	٢٥
جامعي	٨	٢	١٠
المجموع	٦٠	٤٠	١٠٠



ج- الأعمدة البيانية المركبة النسبيـــــة:

في هذا النوع من الأعمدة نقوم بمقارنة ظاهرتين في نفس العمود على السواء، وذلك بتحويل نسب كل قيمة في الظاهرتين. وبالتالي يمكن مقارنة كميات هذه الظاهرة من ناحية النسب ونشير هنا أنه لا يمكن مقارنة كل عمود مع آخر، وإنما يمكن مقارنة الجزيئات ببعضها البعض من نفس العمود، أو مقارنة الجزء المعين من إحدى الظاهرتين من عمود بالآخر مثيله في العمود الثاني، ذلك بمعرفة الفرق بين نسبتيهما بالنسبة للمجموع العام للجدول، كما هو موضح في المثال التالي:

الجدول التالي يوضح تطور عدد الولادات والوفيات في منطقة ما بين ١٩٩٥- ٢٠٠٥

السنة	عدد الولادات	عدد الوفيات
١٩٩٥	٢٠٠	٧٥
١٩٩٦	١٨٥	٩٠
١٩٩٧	٢٢٥	١٠٠
١٩٩٨	٢٢٥	٨٥
١٩٩٩	٢٤٠	٨٠
٢٠٠٠	١٩٥	١٠٠
٢٠٠١	٢١٠	١١٠
٢٠٠٢	٢٢٥	١٠٥
٢٠٠٣	٢٥٠	٩٥
٢٠٠٤	٢٣٠	١١٠
٢٠٠٥	٢٣٥	١٠٠

المطلوب تمثيل العدد الإجمالي للولادات والوفيات المسجلة ما بين ١٩٩٥ و٢٠٠٥ بنسب مئوية لمجموع الأحداث لكل سنة في رسم بياني.

الحل: تمثيل النسب المئوية لكل حدث يكون كالتالي:

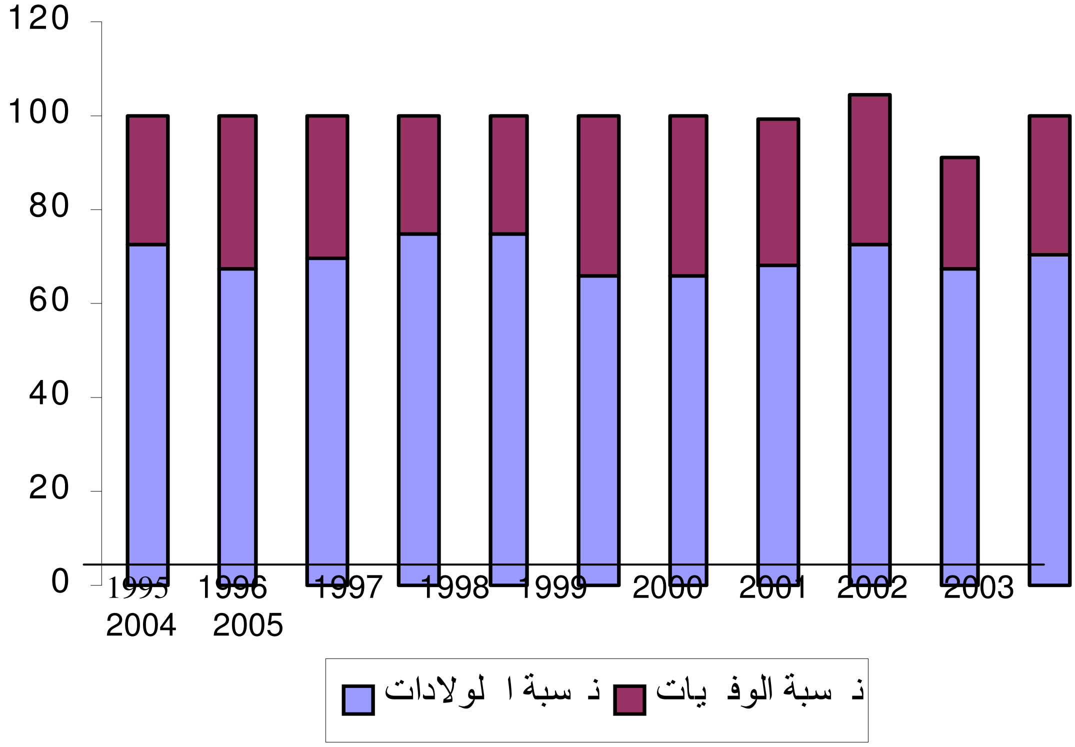
النسبة المئوية للولادات في ١٩٩٥ = $\frac{٢٠٠}{٢٠٠ + ٧٥}$ = ٧٢,٧%

النسبة المئوية للوفيات في ١٩٩٥ = ١٠٠% - ٧٢,٧% = ٢٧,٣%

وبقية السنوات الأخرى نوفيها في الجدول الموالي:

نسبة الوفيات	نسبة الولادات	لسنة
٢٧,٣	٧٢,٧	١٩٩٥
٣٢,٧	٦٧,٣	١٩٩٦
٣٠,٨	٦٩,٣	١٩٩٧
٢٥,٤	٧٤,٦	١٩٩٨
٢٥,٠	٧٥,٠٠	١٩٩٩
٣٣,٩	٦٦,١٠	٢٠٠٠
٣٤,٤	٦٥,٦٠	٢٠٠١
٣١,٤	٦٨,٢٠	٢٠٠٢
٣١,٨	٧٢,٥٠	٢٠٠٣
٢٣,٤	٦٧,٦٠	٢٠٠٤
٢٩,٩	٧٠,١٠	٢٠٠٥

أعمدة بيانية مركبة ومجزئة تبين تطور ظاهرة الولادات والوفيات ما بين ١٩٩٥-٢٠٠٥



ثالثا- تمرينات محلولة حول الفصل الأول

التمرين الأول:

من بين المتغيرات التالية عين المتصلة و المنفصلة منها.

- عدد الأسهم أو الحصص المباعة في البورصة.

هو متغير كمي منفصل.

- عدد الأطفال في الأسرة.

هو متغير كمي منفصل.

- عدد الحجر في البيت.

هو متغير كمي منفصل.

- سن العجزة الشيوخ في دار العجزة.

متغير كمي متصل.

- طول ١٠ رضع في المستشفى.

متغير كمي متصل.

- نوع المهنة.

متغير كيفي

التمرين الثاني:

صمم جدولا ذو تقاطع مضاعف من عينة تساوي ١١٥٠ امرأة متزوجة. ابحث عن التقاطع بين نوع اسرتها و قيمته نووية و ممتدة و عمل المرأة المقسم التي تعمل و لا تعمل – و عدد أطفالها الذي قيمته، ١- ٣ أطفال و ٤- ٦ أطفال و هو المتغير التابع.

مع العلم أن عدد المقيمات في الأسر النووية اللواتي تعملن يقدر بـــ ٣٠٠ امرأة.

عدد المقيمات في الأسر الممتدة اللواتي تعملن يقدر بـــ ٢٠٠ امرأة.

عدد المقيمات في الأسر الممتدة اللواتي لا تعملن يقدر بـــ ٤٥٠ امرأة.

عدد المقيمات في الأسر النووية اللواتي لا تعملن يقدر بـــ ٢٠٠ امرأة.

و إذا كان عدد اللواتي لهن مابين ١- ٣ أطفال يقدر بـــ ٦٥٠ امرأة تنتمي ٦٠٪ منهن إلى اللواتي تعملن و تنتمي نصفهن إلى المقيمات في الأسر النووية و أن نصف اللواتي لهن مابين ١- ٣ أطفال هن مقيمات في أسر نووية مهما كان وضعهن (تعمل، لا تعمل).

الاسئلة:

١- حاول تركيب الجدول مكملا كل تقاطعاته.

٢- قراءة الجدول.

عدد الأطفال / نوع الأسرة و مدى العمل		١- ٣ أطفال		٤- ٦ أطفال		المجموع	
		ك	%	ك	%	ك	%
نووية	تعمل	١٩٥	٦٥	١٠٥	٣٥	٣٠٠	١٠٠
	لا تعمل	١٣٠	٦٥	٧٠	٣٥	٢٠٠	١٠٠
مجموع اللواتي تقطن في أسر نووية		٣٢٥	٦٥	١٧٥	٣٥	٥٠٠	١٠٠
ممتدة	تعمل	١٩٥	٩٧,٥	٥	٢,٥	٢٠٠	١٠٠
	لا تعمل	١٣٠	٢٨,٨٩	٣٢٠	٧١,١١	٤٥٠	١٠٠
مجموع اللواتي تقطن في أسر الممتدة		٣٢٥	٥٠	٣٢٥	٥٠	٦٥٠	١٠٠
المجموع		٦٥٠	٥٦,٥٢	٥٠٠	٤٣,٤٨	١١٥٠	١٠٠

نلاحظ من خلال قراءتنا لهذا الجدول أن الاتجاه العام له يتجه نحو اللواتي لهن مابين ١- ٣ أطفال نسبة ٥٦,٦٢٪ مقابل ٤٣,٤٨٪ لهن مابين ٤- ٦ أطفال.

و عند إدخالنا للمتغير المستقل ألا و هو الأصل نوع الأسرة لمعرفة تأثيره على عدد أطفالها نلاحظ أن نسبة كبيرة من اللواتي تقطن في أسر نووية لهن مابين ١- ٣ أطفال نسبة ٦٥٪ مقابل ٥٠٪ فقط من اللواتي تقطن في أسر ممتدة.

و عند إدخالنا للمتغير الثالث و هو المتغير الرائز و المتمثل في عمل المرأة لتفكيك أكثر العلاقة الأولى، نلاحظ أن اللواتي تقطن في أسر نووية لم يطرأ عليهن تغير يذكر بإدخال متغير عمل المرأة إذ بقيت نسبت ٦٥٪ منهن لهن مابين ١- ٣ أطفال سواء كانت تعمل أم لا.

فحين اللواتي تقطن في الأسر الممتدة فنلاحظ أن العلاقة الأولى قد تغيرت، إذ أن اللواتي تعملن منهن ٩٧,٥٪ لهن مابين ١– ٣ أطفال مقابل ٧١,١١٪ منهن و اللواتي لا تعملن لهن مابين ٤– ٦ أطفال.

التمرين الثالث:

أراد باحث القيام بدراسة أسباب تواجد المسنين في دور العجزة، فوقف عند هذه البيانات الممثلة لأعمارهم:

٦٧، ٨٣، ٦٤، ٧٤، ٩٢، ١٠٢، ٥٨، ٩٥، ٨٢، ٧٩، ١٠٧، ٦٩، ٨٢، ٨٤، ٧٨، ٦٨، ١٠١، ٨٧، ٥٢، ٨٣، ٨٠، ٨٥، ١٠٤، ٩٠، ٧٢، ٨٧، ٧٢، ٨٥، ١٠٥، ٨٢، ٧٩، ٩٩، ٨٤، ٧٦، ١٠٣، ٦٤، ٨٧، ٩٧، ٦٢، ٨٩، ١٠٨، ٩٢، ٩١، ٩٦، ٧٠، ٧٣، ٩٩، ٨٠.

الأسئلة:

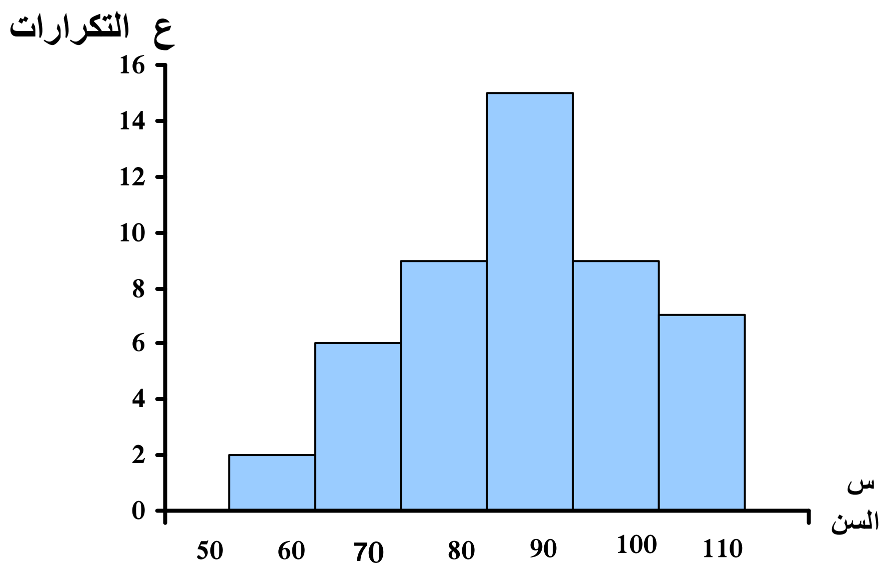
قم بتبويب هذه البيانات على جدول توزيع تكراري ثم مثلها بيانيا بمدرج تكراري و مضلع تكراري، على أن تبدأ الفئة الأولى بـــ ٥٠ و عدد الفئات ٦ و طول الفئة ١٠.

الحل: ١– تفريغ البيانات

السن	الجدولة	التكرارت
٥٠ – ٦٠	II	٢
٦٠ – ٧٠	~~IIII~~ I	٦
٧٠ – ٨٠	IIII IIII	٩
٨٠ – ٩٠	~~IIII~~ ~~IIII~~ IIII IIII	١٥
٩٠ – ١٠٠	~~IIII~~ IIII	٩
١٠٠ – ١١٠	~~IIII~~ II	٧
المجموع		٤٨

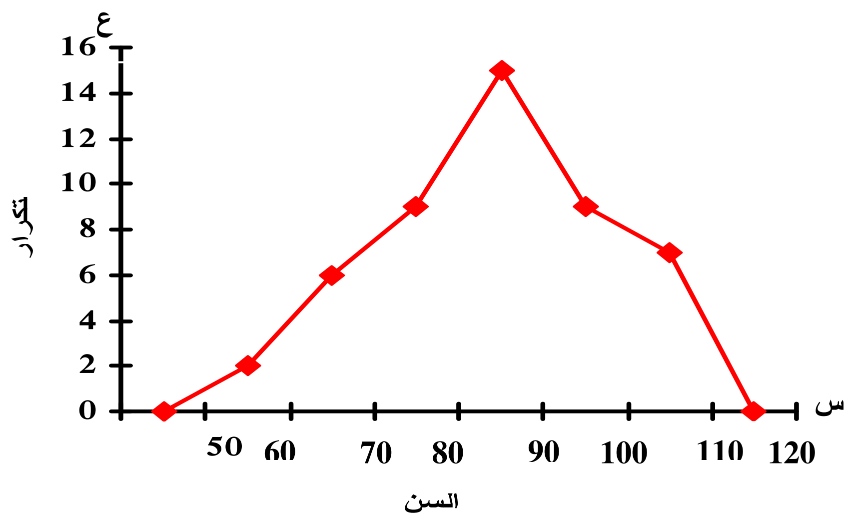
٢– رسم المدرج التكراري

مدرج تكراري يبين توزيع سن مجموعة من المسنين في دار العجزة



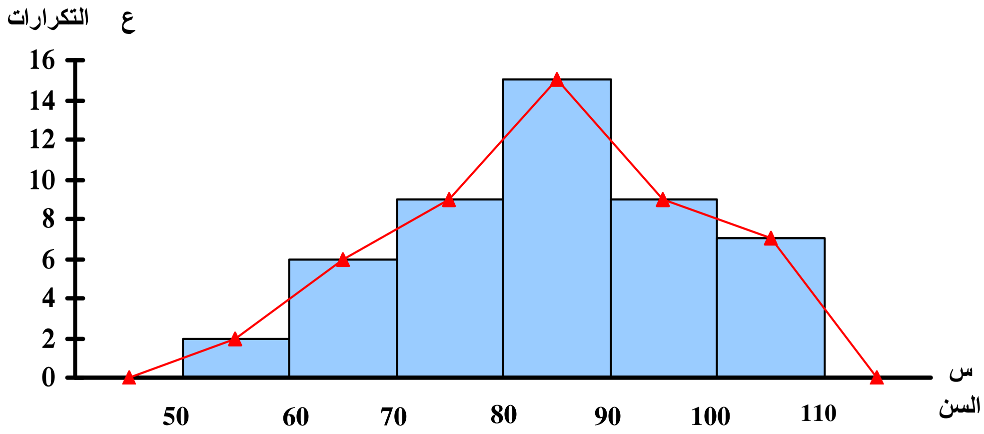
٢- رسم المضلع التكراري

مضلع تكرار يبين توزيع سن مجموعة من المسنين في دار العجزة



٣- رسم المضلع و المتدرج التكراري على نفس المعلم

مضلع و مدرج تكراري يبين توزيع سن مجموعة من المسنين في دار العجزة.



التمرين الرابع:

الجدول التالي يبين توزيع مجموعة من العمال في مؤسسة ما حسب مستواهم التعليمي و المطلوب تمثيلهم بدائرة نسبية، و بالأعمدة البيانية و حلاقات.

المستوى التعليمي	التكرارات
أمي	٤٥
ابتدائي	٤٠
متوسط	٦٠
ثانوي	٣٥
جامعي	٢٠
المجموع	٢٠٠

الحل:

أ- التمثيل بالدائرة النسبية.

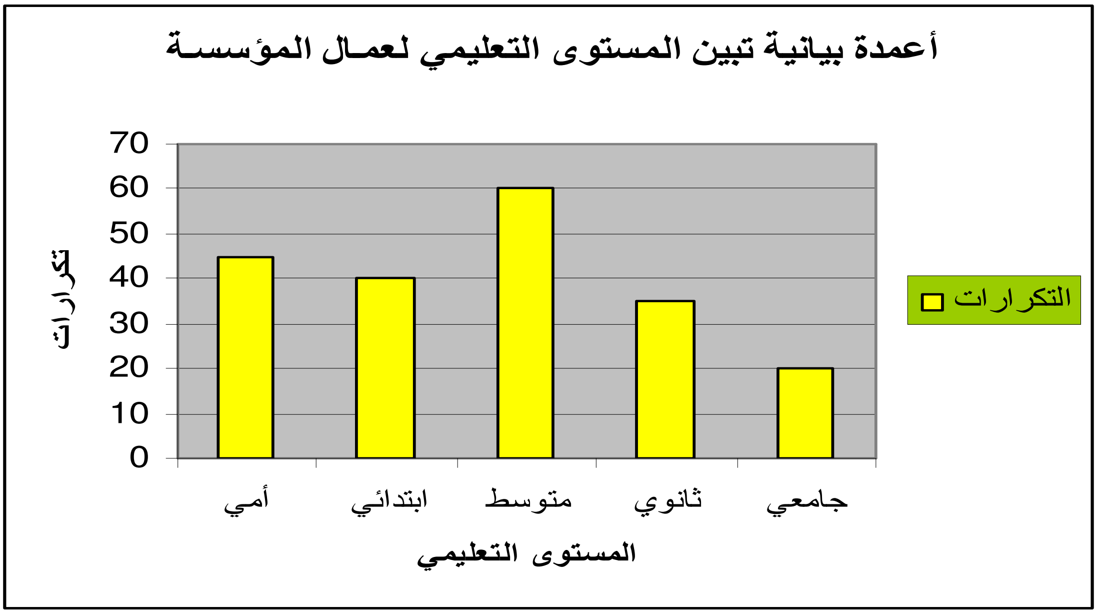
أولا: نقوم بتحويل القيم إلى نسب مئوية.

ثانيا: نقوم بتحويل النسب المئوية إلى دراجات مئوية كما هو مبين في الجدول التالي.

المستوى التعليمي	التكرارات	٪	x ٣,٦
أمي	٤٥	٢٢,٥	٨١
ابتدائي	٤٠	٢٠	٧٢
متوسط	٦٠	٣٠	١٠٨
ثانوي	٣٥	١٧,٥	٦٣
جامعي	٢٠	١٠	٣٦
المجموع	٢٠٠	١٠٠	٣٦٠

أ- التمثيل بالدائرة النسبية

ب- التمثيل بالأعمدة.


أعمدة بيانية تبين المستوى التعليمي لعمـال المؤسسـة
70
60
50
40
30
20
10
0
تكرارات
التكرارات
أمي
ابتدائي
متوسط
ثانوي
جامعي
المستوى التعليمي

الفصـــل الرابـــع

مقـــاييس النزعة المركزيـــة

أولا- المتوسطات الإحصائيـــة

١- المتوسط الحسابـــي

٢- الوسط الفرضـــي

٢- المتوسط المرجـــح

٣- المتوسط الهندســي

٤- المتوسط الهندسي المرجع

٥- الوسط التوافقـــي

٦- الوسط التربيعـــي

ثانيا: الوســـيـط

ثالثا: المئيـــنـــات

رابعا: الربيعـــات

خامسا: العشيـــرات

سادسا: المنـــوال

سابعا: العلاقة بين مقايس النزعة المركزية

ثامنا: خواص مقايس النزعة المركزيـــة

أولا – المتوسطات الإحصائيـــة.

إن وصف الجداول التكرارية وتحديد شكل المنحنيات التي تمثلها، وكذا المقارنة بين المنحنيات لمختلف المتغيرات في الظاهرة المراد دراستها، يتم باستخدام مقاييس خاصة وهي مقاييس النزعة المركزية، مقاس التبعثر والتشتت تم مقاييس الالتواء و التفرطح.

و كلمة نزعة مركزية تعني التمركز و التكثف حول رقم معين، حيث يلاحظ عامة أن قيم الظاهرة تميل دائما إلى التمركز عند قيمة معينة في التوزيع التكراري. لهذا فهناك مقاييس إحصائية تحاول تلخيص البيانات/قيم الظاهرة في رقم واحد يدل عليها ويرمز لها. هذا الرقم يوضح نزعة البيانات إلى التمركز حول فئة معينة، وكلما ابتعدنا عن هذه القيمة فإن عدد البيانات يبدأ في التناقص. ولهذا فهناك من الطرق ما توصلنا إلى تحديد هذا الرقم جبريا أو بالرسم وسوف نعرض فيما يلي أهمها:

١ – المتوسط الحسابـــــي:

هو أشهر مقاييس النزعة المركزية الذي يقاس بجمع قيم كل عناصر المجموعة ثم قسمة النتيجة على عدد عناصر المجموعة، بمعنى آخر مجموع القياسات الخاصة بظاهرة معينة على عدد القياسات. ويعتبر المتوسط الحسابي أهم مقاييس النزعة المركزية وأكثرها استخداما، كما أنه يسهل حسابه.

أ- المتوسط الحسابي في حالة البيانات الغير مبوبة.

إن متوسط عدد من القيم هو خارج قسمة مجموع هذه القيم على عددها ويرمز للمتوسط الحسابي برمز سَ، 'X ويمكن أن يكتب جبريا كالآتي:

المتوسط الحسابي = $\dfrac{\text{مجموع المفـردات}}{\text{عـــدد المفــردات}}$

وبالرموز: $\bar{س} = \dfrac{\text{مج س}}{ن}$ $\qquad X' = \dfrac{\sum X}{N}$

فلنرمز لأي مفردة من مفردات مجموعة من القيم بالرمز س وإلى عدد القيم بالرمز ن فعندها يكون $س_١$، $س_٢$، $س_٣$....$س_ن$، هي القيم المرغوب في حساب متوسطها، الحسابي وهكذا فان:

$$\bar{س} = \frac{س_١ + س_٢ + س_٣ + ...س_ن}{ن}$$

مثـــال:

حساب المتوسط الحسابي للقيم التالية ٨، ٣، ٥، ١٢، ١٠ هو كالتالي:

$$\bar{س} = \frac{\text{مج س}}{ن} = \frac{٨ + ٥ + ٣ + ١٢ + ١٠}{٥} \quad \frac{٣٨}{٥} = ٧،٦$$

في المتوسط الحسابي لابد أن يكون مجموع انحرافات القيم عن متوسطها الحسابي مساوية لصفرا.

وبالتطبيق على المثال السابق نجد:

(٨ - ٧،٦) + (٣ - ٧،٦) + (٥ - ٧،٦) + (١٢ - ٧،٦) + (١٠ - ٧،٦) = ٠

(-٠،٤) + ٤،٦ + ٢،٦ - ٤،٤ - ٢،٤ =(-٧،٢) + ٧،٢ = ٠

ب- المتوسط الحسابي في حالة البيانات المبوبة.

أما إذا كانت القيم/البيانات موزعة في جدول توزيع تكراري يحتوي على فئات فنحن هنالا نعرف قيمة كل حالة وإنما نعرف أن كل حالة تنحصر قيمتها بين رقمين هما الحد الأدنى والأعلى للفئة، لذلك علينا افتراض قيمة متساوية لكل أفراد الفئة الواحدة، وذلك

بأن نعطي كل حالة في الفئة قيمة هي مركز تلك الفئة ويحسب المتوسط الحسابي وفق القانون التالي:

$$\bar{س} = \frac{مج\ س \times ك}{مج\ ك} \qquad \bar{X} = \frac{\sum Fi \times Ci}{\sum Fi}$$

ويلاحظ أن هذه القاعدة تطبق فقط في حالة الجداول التكرارية المقفلة، أما الجداول التكرارية المفتوحة أي التي لا تتضمن الحد الأعلى أو الأدنى أو كلاهما في التوزيع التكراري فإنه من المستحيل التوصل إلى مراكز هذه الفئات لإيجاد المتوسط الحسابي و سوف نرى لاحقا كيف نصل إلى المتوسط الحسابي في مثل هذا النوع من الجداول.

مثـــال: الجدول التالي يبين توزيع ٦٠ مولودا حيا حسب مدة حمله بالأسبوع. والمطلوب حساب متوسط مدة الحمل لهؤلاء المواليد عن طريق المتوسط الحسابي.

مدة الحمـــــــل	عدد المواليـــــــــد
٢٨ – ٣٠	٣
٣١ – ٣٣	٥
٣٤ – ٣٦	٣٠
٣٧ – ٣٩	٢٠
٤٠ – ٤٢	٢
المجموع	٦٠

لحساب المتوسط الحسابي في حالة البيانات المبوبة لابد من حساب منتصفات الفئات والتي تساوي

$$\frac{\text{الحد الأدنى للفئة + الحد الأعلى لها}}{٢}$$

مثلا $\frac{٢٨ + ٣٠}{٢}$ = ٢٩ أسبوعـــا.

ثم نقوم بحساب مجموع كل تكرار مضروبا في مركز الفئة المقابل له أي مجموع ك ×س كما هو مبين في الجدول التالي:

مدة الحمل	عدد المواليد ك/	س	ك × س
٢٨ – ٣٠	٣	٢٩	٨٧
٣١ – ٣٣	٥	٣٢	١٦٠
٣٤ – ٣٦	٣٠	٣٥	١٠٥٠
٣٧ – ٣٩	٢٠	٣٨	٧٦٠
٤٠ – ٤٢	٢	٤١	٨٢
المجموع	٦٠		٢١٣٩

$\bar{س} = \frac{مج\ ك×س}{مج\ ك} = \frac{٢١٣٩}{٦٠} =$ ٣٥،٦٥ أسبوعا

أي بالتقريب ٣٦ أسبوعا الموافقة لتسعة أشهر، أي أن الأغلبية من هؤلاء الأطفال ولدوا في أوانهم.

٢- الوسط الفرضـــــــــي:

أ- الوسط الفرضي في حالة البيانات الغير مبوبة.

إذا كانت القيم كبيرة فيمكن اختيار قيمة منها ككمية/ أو قيمة ثابتة ونقوم بطرح القيم الأخرى منها بحيث نحصل على قيم متغير جديد يقسم على عدد القيم ويضاف إلى القيمة المختارة، فنحصل بذلك على المتوسط الحسابي.

الوسط الفرضي = الوسط الفرض + $\frac{\text{مجموع انحرافات القيم عن وسطها الفرضي}}{\text{عدد القيـــــم}}$

أي $\bar{\text{س}} = \text{أ} + \frac{\text{مج ح}}{\text{ن}}$

مثـــال: إيجاد المتوسط الحسابي للقيم التالية: ١٥٠٠٠، ١٨٦٠٠، ١٧٨٠٠، ١٤٥٠٠، ١٩٠٠٠. في هذه الحالة نفترض أن القيمة المختارة للمتوسط هي ١٤٥٠٠ ثم نطرح هذا الرقم من بقية القيم الأخرى فتحصل على ما يلي:

(١٥٠٠٠ − ١٤٥٠٠) + (١٨٦٠٠ − ١٤٥٠٠) + (١٧٨٠٠− ١٤٥٠٠) + (١٤٥٠٠ − ١٤٥٠٠) + (١٩٠٠٠ − ١٤٥٠٠) =

٥٠٠ + ٤١٠٠ + ٣٣٠٠ + ٠ + ٤٥٠٠ = ١٢٤٠٠.

ثم نقسم على عدد المفردات أو القيم هذه النتيجة فنحصل على٥ /١٢٤٠٠ = ٢٤٨٠.

وبهذا يصبح المتوسط الحسابي = ٢٤٨٠ + ١٤٥٠٠ = ١٦٩٨٥٠.

و بالطريقة العادية نجد المتوسط الحسابي في هذا المثال يساوي:

$\frac{\text{مج س}}{\text{ن}}$ $\frac{١٥٠٠ + ١٨٦٠٠ + ١٧٨٠٠ + ١٤٥٠٠ + ١٩٠٠٠}{٥}$ $\frac{٨٤٩٠٠}{٥}$

= ١٦٩٨٥٠.

وبهذا نرى أننا تحصلنا على نفس النتيجة في كلا الحالتين ويكون مجموع الانحراف عن الوسط دائما يساوي صفرا.

(١٦٩٨٠ − ١٥٠٠٠) + (١٦٩٨٠ − ١٨٦٠٠) + (١٦٩٨٠ − ١٧٨٠٠) + (١٦٩٨٠ −
(١٤٥٠٠ + (١٦٩٨٠ − ١٩٠٠٠) = + ١٩٨٠ − ١٦٢٠ − ٨٢٠ + ٢٤٨٠ − ٢٠٢٠ =

+ ٤٤٦٠ − ٤٤٦٠ = ٠.

ب- الوسط الفرضـــي في حالة البيانات المبوبة:

يمكن حساب المتوسط الحسابي بطريقة أخرى أسهل إذا كان الجدول هو توزيع تكراري منتظم أي تساوت فيه أطوال الفئات، بحيث يصبح طول الفئة هو القاسم المشترك وذلك لتبسيط العملية الحسابية، خاصة في حالة الجداول ذات الفئات العديدة أي التوزيعات التكرارية الطويلة.

وفي هذه الحالة نفترض أن الوسط الفرضي يساوي مركز فئة معينة، بحيث نحسب انحرافات بقية الفئات عن هذا الوسط الفرضي ويستحسن أن يكون هذا الوسط الفرضي المختار في وسط الجدول مقابل أكبر تكرار كلما أمكن ذلك وهذا لتسهيل العمليات الحسابية وبهذا فلوسط الفرضي في حالة البيانات المبوبة تساوي:

$$\text{الوسط الفرضي} = \text{أ} + \frac{\text{مجموع حوا صل ضرب التكرار} \times \text{انحرافات الفئات عن الوسط الفرضي}}{\text{مجموع التكرارات}}$$

$$\bar{\text{س}} = \text{أ} + \frac{\text{مج ك} \times \text{ح}}{\text{مج ك}}$$

وبتطبيق هذا القانون على المثال السابق نجد الوسط الفرضي يكون مساويا إلى:

مدة الحمل	ك	س	ح	ك × ح
٢٨ − ٣٠	٣	٢٩	− ٦	− ١٨
٣١ − ٣٣	٥	٣٢	− ٣	− ١٥

٣٤ – ٣٦	٣٠	٣٥	٠	٠
٣٧ – ٣٩	٢٠	٣٨	+ ٣	٦٠
٤٠ – ٤٢	٢	٤١	+ ٦	١٢
المجمـــوع	٦٠		٠	٣٩

$$\bar{\text{س}} = \text{أ} + \frac{\text{مج ك} \times \text{ح}}{\text{مج ك}} = ٣٥ + \frac{٣٩}{٦٠} = ٣٥،٦٥ \text{ أسبوعا}$$

٣- الوسط الحسابي المرجـــــح.

أ – الوسط المرجح في حالة البيانات الغير مبوبة.

وهذا يستخدم عندما نفاضل بين مفردات توزيع أي ظاهرة في الأهمية، أي إذا علمنا أن بعض مفرداتها تكتسي أهمية بالنسبة للبحث على أخرى، حيث يرجح كل مفردة أو قيمة بأهميتها ويحسب وفق القانون التالي:

$$\text{الوسط المرجـــــح} = \frac{\text{مج س} \times \text{و}}{\text{مج و}}$$

مثـــال: من تحقيق رصده فيه ١٠٠ عائلة وجد ٢٠ عائلة لديها ٤ أطفال ٤٠ لديهم ٥ أطفال، ٣٠ لديهم ٦ أطفال والباقي لديهم ٧ أطفال

إذن الوسط المرجح للأطفال هؤلاء الأسر هو:

$$\frac{\text{مج س×و}}{\text{مج و}}$$

$$\frac{(\text{٤} \times \text{٢٠}) + (\text{٥} \times \text{٤٠}) + (\text{٦} \times \text{٣٠}) + \text{٧} \times \text{١٠}}{\text{١٠٠}} = \frac{\text{٥٣٠}}{\text{١٠٠}} = \text{٥،٣ طفل لكل عائلة.}$$

٤- الوسط الحسابي بالطريقة المختصرة:

يمكننا كذلك حساب المتوسط الحسابي بطريقة مختصرة سهلة ولإيجاده نقوم مثل في الوسط الفرضي بحساب مراكز الفئات س ثم تحديد وسط فرضي أ والذي يكون مقابل لأكبر تكرارا من مراكز الفئات ثم نجد انحراف مراكز الفئات عن الوسط الفرضي ح وفي الأخير نجد الانحرافات المختصـــــرة.

$$\text{حَ} = \frac{\text{ح}}{\text{ل}} \qquad \frac{\text{الانحراف مركز الفئة عن الوسط الفرضي}}{\textbf{طول الفئة}}$$

ثم نجد حاصل ضرب مجموع الانحرافات المختصرة في التكرار المقابل لها.

وبهذا الوسط الحسابي بالطريقة المختصرة يساوي:

$$\text{سَ} = \text{أ} + \frac{\text{مج حَ} \times \text{ك}}{\text{مج ك}} \times \text{ل}$$

وبتطبيق هذه المعادلة على المثال السابق أعلاه نجد أن متوسط مدة حمل هؤلاء المواليد بطريقته المختصرة كما يلي:

مدة الحمل	ك	س	ل	ح	حَ	حَ × ك
٢٨ – ٣٠	٣	٢٩	٢	– ٦	– ٣	– ٩
٣١ – ٣٣	٥	٣٢	٢	– ٣	– ١،٥	– ٧،٥

٣٤ – ٣٦	٣٠	٣٥	٢	٠	٠	٠
٣٧ – ٣٩	٢٠	٣٨	٢	+ ٣	١،٥	٣٠
٤٠ – ٤٢	٢	٤١	٢	+ ٦	٣	٦
المجمـوع	٦٠			٠	٠	١٩،٥

$$\bar{س} = أ + \frac{مج\ حَ \times ك}{مج\ ك} \times ل$$

$$٣٥ + \frac{١٩،٥}{٦٠} \times ٢ = ٣٥،٦٥ \text{ أسبوعا}$$

١ – ٨– المتوسط الهندســـــــــي:

يستخدم المتوسط الهندسي في حالة المفردات التي تزيد بنسب ثابتة كما دراسة النمو السكاني والاقتصادي ... إلخ حيث يعتبر أنسب المتوسطات في حالة معدلات التغير ويحسب وفق المعادلة التاليـــــة:

أ/– في حالة البيانات غير المبوبـــــــة:

$$\text{لو هـ} = \frac{١}{ن}\ \text{مج}\ \ \text{لو س}$$

مثـــــــال: بلغ عدد سكان الجزائر في سنة ١٩٩٠ ٢٥,٠٢٢,٠٠٠ نسمة وفي سنة ٢٠٠٤ حوالي ٣٢١٠٠٠٠٠ نسمة. فعدد السكان في منتصف هذه الفترة أي في عام ١٩٩٧ كان:

$$\text{هـ} = \sqrt{(٢٥٠٢٢٠٠٠)\ (٣٢١٠٠٠٠٠)} = ٢٨,٣٢٠,٠٠٠$$

ب/- في حالة البيانات المبوبــــــــة:

يحسب المتوسط الهندسي عندما يكون لدينا جدول مبوب في فئات وفق العلاقة التاليـــة:

$$\text{هـ} = \sqrt[\text{ن}]{\text{س}_1^{\text{ك}_1} . \text{س}_2^{\text{ك}_2} \ldots\ldots \text{س}_\text{ن}^{\text{ك}_\text{ن}}}$$

أي أن **لو هـ** = $\dfrac{\text{مج ك لو س}}{\text{مج ك}}$

مثـــــال: أحسب الوسط الهندسي للجدول التالي الذي يوضح معدل وفيات الأطفال الرضـــــــع

لـــ ٨٠ دولة مثـــــلا.

معدل وفيات الأطفال الرضع	ك	س	لو س	ك × لو س
١٠ – ٢٠	٥	١٥	١،١٨	٥،٨٨
٢٠ – ٣٠	١٥	٢٥	١،٣٩	٢٠،٩٧
٣٠ – ٤٠	٢٠	٣٥	١،٥٤	٣٠،٨٨
٤٠ – ٥٠	١٨	٤٥	١،٦٥	٢٩،٧٦
٥٠ – ٦٠	٢٢	٥٥	١،٧٤	٣٨،٢٩
المجمــــــــــوع	٨٠			١٢٥،٧٨

لو هـ = $\dfrac{\text{مج ك لو س}}{\text{مج ك}} = \dfrac{١٢٥،٧٨}{٨٠} = ١،٥٧$ وهو متوسط التغير في معدل وفيات الرضع لـــ ٨٠ دولــــــــة.

- 1 -9المتوسط الهندسي المرجـــــــح:

ويحسب وفق الصيغة التاليـــــــة:

$$هـ = \sqrt[\sum و]{س_1^{و_1} \cdot س_2^{و_2} \ldots\ldots س_ن^{و_ن}}$$

أي أن **لو هـ** = $\frac{مج\ و\ لو\ س}{مج\ و}$

مثــــــال: أوجد الوسط الهندسي المرجع للبيانات التاليـــــة:

السنــة	عدد الـــــــولادات (س)	عدد الوفيات المزايد الجدد (و)
٢٠٠٠	١٧٠,٠٠٠	٩٥
٢٠٠١	١٥٠,٠٠٠	٦٠
٢٠٠٢	١٨٠,٠٠٠	٦٥
٢٠٠٣	١٦٠,٠٠٠	٨٠
٢٠٠٤	١٧٥,٠٠٠	٦٣

الحـــــل:

س	و	لو س	و لو س
١٧٠,٠٠٠	٩٥	٥,٢٣	٤٩٦,٨٩
١٥٠,٠٠٠	٦٠	٥,١٨	٣١٠,٥٦
١٨٠,٠٠٠	٦٥	٥,٢٥	٣٤١,٥٩
١٦٠,٠٠٠	٨٠	٢,٥٠	٤١٦,٣٣
١٧٥,٠٠٠	٦٣	٥,٢٤	٣٣٠,٣١
	٣٦٣		**١٨٩٥,٦ ٨**

لو هـ = $\frac{مج\ و\ لو\ س}{مج\ و} = \frac{١٨٩٥،٦٨}{٣٦٣} = ٥،٢٢$

١- ١٠- الوسط التوافقـــــــــي:

هو عبارة عن وسط يحسب معدلات زمنية وهو مقلوب الوسط الحسابي لمقلوبات قيم المتغير المدروس. ويعتبر أفضل المتوسطات لتمثيل المعدلات في التغير والسرعات وكذلك الأسعار والأثمان ... الخ.

ويحسب وفق الصيغ التاليــــــــة:

أ/- في حالة البيانات غير المبوبـــــــة:

إذا كانت لدينا مجموعة من القيم عددها ن و هي $س_١$ $س_٢$ ... $س_ن$ فأول خطوة نقوم بها هي إيجاد مقلوب هذه القيم $\frac{١}{س_١}$ أي $\frac{١}{س}$ $\frac{١}{س_ن}$ ......

وبعد هذا نقوم بتطبيق الصيغـــة التاليـــــة:

$$ق = \frac{ن}{\frac{١}{س_١} + \frac{١}{س_٢} + \frac{١}{س_ن}} = \frac{ن}{\sum \frac{١}{س_ر}}$$

مثـــــال: لدينا البيانات التاليـــة ٦، ٨، ١٣، ١٥ والمطلوب حساب متوسطها التوافقي.

$$ق = \frac{٤}{٠,١٧ + ٠,١٢ + ٠,٠٨ + ٠,٠٧}$$

$= ٩،١٠$

ب/- في حالة البيانات المبوبـــــــة:

في هذه الحالة يتوجب علينا معرفة مراكز الفئات (س) لأننا سوف نقوم بحساب مقلوبات ~~هذه~~ المراكز أي $\frac{١}{س}$ ثم نقوم بضرب النتيجة في التكرار المقابل لها كما سنوضحه في المثال اللاحق عن المعادلة المطبقة لإيجاده فهي نحو التالي:

$$ق = \frac{مج\ ك}{مج\ \frac{ك}{س}}$$

مثـــــــال: الجدول التالي يوضح تطور عدد الوفيات حسب الفئات العمريـــــة.

الفئـــات	**ك**	**س**	$\frac{١}{س}$	$\frac{ك \times ١}{س}$
٠ – ٥	١٥	٢،٥	٠،٤	٩
٥ – ١٠	١٣	٧،٥	٠،١٣	١،٦٩
١٠ – ١٥	٢٠	١٢،٥	٠،٠٨	١،٦
١٥ – ٢٠	٣٥	١٧،٥	٠،٠٦	٢،١٠
٢٠ – ٢٥	٥	٢٢،٥	٠،٠٤	٠،٢
المجمـــــوع	٨٨			**١١،٥٩**

$$ق = \frac{مج\ ك}{مج\frac{ك}{س}} = \frac{٨٨}{١١،٥٢} = ٧،٦٤$$

– 1 -11 الوسط ألتربيعي:

الوسط ألتربيعي هو الجذر ألتربيعي للوسط الحسابي لمربعات قيم المتغير المدروس.

أ/– في حالة البيانات غير المبوبــــــــة:

$$ت = \sqrt{\frac{\sum س^{٢}}{ن}}$$

مثـــــــال: إيجاد الوسط ألتربيعي للقيم التاليــــــة: ٥، ٨، ١٧، ٢٠

$$\sqrt{\frac{(٥)^{٢} + (٨)^{٢} + (١٧)^{٢} + (٢٠)^{٢}}{٤}} = \sqrt{\frac{٢٥ + ٦٤ + ٢٨٩ + ٤٠٠}{٤}} = \sqrt{\frac{٧٧٨}{٤}}$$

ت =   =   =

=   ١٩٤،٥ = ١٣،٩٥

ب/- في حالة البيانات المبوبـــــــة:

$$\text{ت} = \sqrt{\frac{\text{مج س}^{٢}\text{ ك}}{\text{مج ك}}}$$

مثـــــال: أوجد الوسط ألتربيعي للقيم التاليـــــــة:

الفئات	ك	س	$س^{٢}$	$س^{٢} \times ك$
١٠ – ٢٠	٣٠	١٥	٢٢٥	٦٧٥٠
٢٠ – ٣٠	٢٥	٢٥	٦٢٥	١٥٦٢٥
٣٠ – ٤٠	١٣	٣٥	١٢٢٥	١٥٩٢٥
٤٠ – ٥٠	٣٥	٤٥	٢٠٢٥	٧٠٨٧٥
٥٠ – ٦٠	١٠	٥٥	٣٠٢٥	٣٠٢٥٠
المجمـــــوع	**١١٣**			**١٣٩٤٢٥**

$$\text{ت} = \sqrt{\frac{١٣٩٤٢٥}{١١٣}} = \sqrt{\frac{\text{مج س}^{٢}\text{ ك}}{\sum \text{ك}}} = \sqrt{١٢٣٣،٨٥} = ٣٥،١٣$$

١٢– ١– العلاقة بين المتوسطـــــــــات:

لو قارنا بين مختلف أنواع المتوسطات السابقـــة الذكر نجد أن:

المتوسط ألتربيعي > المتوسط الحسابي > المتوسط الهندسي > الوسط التوافقي

أي أن ت > س > هـــ > ق، وهذه المتراجحة تبقى صحيحة مهما اختلف التوزيع التكراري في بياناتـــه.

مثـــــال: أوجد المتوسط بمختلف الطرق للجدول التالي الذي يبين مدّة زواج ١٠٠ امرأة.

مدة الـــــــزواج	ك
١ – ٥	١٥
٥ – ٩	٣٢
٩ – ١٣	١٨
١٣ – ١٧	٢٢
١٧ – ٢١	١٣
المجمـــــوع	**١٠٠**

١/- **حساب الوسط ألتربيعي:** $ث = \sqrt{\frac{\sum س^{٢}.ك}{\sum ك}}$

٢/- **المتوسط الحسابي:** $\bar{س} = \frac{\sum ك\,س}{\sum ك}$

٣/- **الوسط الهندسي:** $هـ = \sqrt{س_{١}^{ك١}.س_{٢}^{ك٢} ... س_{ن}^{كن}}$ = لو هـ = $\frac{١}{ن}\sum$ ك لو س

٤/- **الوسط التوافقي:** $ق = \frac{\sum ك}{\sum \frac{ك}{س}}$

$ك.\frac{١}{س}$	$\frac{١}{س}$	ك.س²	س²	ك.س	س	ك	مدة الزواج
٥	٠،٣٣	١٣٥	٩	٤٥	٣	١٥	١ – ٥
٤،٥٧	٠،١٤	١٥٦٨	٤٩	٢٢٤	٧	٣٢	٥ – ٩
١،٦٤	٠،٠٩	٢١٧٨	١٢١	١٩٨	١١	١٨	٩ – ١٣
١،٤٧	٠،٠٧	٤٩٥٠	٢٢٥	٣٣٠	١٥	٢٢	١٣ – ١٧
٠،٦٨	٠،٠٥	٤٦٩٣	٣٦١	٢٤٧	١٩	١٣	١٧ – ٢١
١٣،٣٦		١٣٥٢٤		١٠٤٤		١٠٠	المجموع

بالتعويض في القوانين السابقة نجد ت = $\sqrt{\frac{١٣٥٢٤}{١٠٠}}$ = ١١،٦٣ سنـــة.

سَ = $\frac{١٠،٤٤}{١٠٠}$ = ١٠،٤٤ سنـــــة.

ق = $\frac{١٠٠}{١٣،٣٦}$ = ٧،٤٨ سنـــــة.

2 - الوسيـــــــــــط:

الوسيط هو القيمة التي تقسم مجموعة من القيم إلى قسمين متساويين أي أن تكون القيم الأكبر منها مساوية للقيم الأصغر منها بحيث تكون قيم المتغير مرتبة تصاعديا أو تنازليا، بمعنى آخر هو القيمة التي تقع في منتصف سلسلة مرتبة ترتيبا تصاعديا أو تنازليا بحيث يكون عدد المفردات الأصغر منها مساوية لعدد المفردات الأكبر منهـــــا.

2 -1 - الوسيط في حالة البيانات الغير مبوبـــــــة:

نقوم بترتيب البيانات تصاعديا أو تنازليا ثم نأخذ القيمة التي تقع في الوسط تماما إذا كان مجموع عدد البيانات فرديا، أما إذا كان مجموع عدد أفراد البيانات زوجيا فإن الوسيط هو القيمة الناتجة عن جمع القيمتين اللتين تقعان في الوسط وقسمة النتيجة على اثنيـــــن.

مثـــــــال ١: في حالة عدد البيانات فرديـــــا.

إذا كانت أعمار تسعة أفراد هي ١٥، ٣٢، ٤٤، ٢٨، ٣٣، ١٧، ٢٣، ٣٠، ١٨ فوسيط أعمارهم هو:

أولا: نقوم بترتيب القيم ترتيبا تصاعديا أو تنازليا، ثم نأخذ القيمة التي تقسم السلسلة إلى قسمين متساويين كما يلي: ١٥، ١٧، ٢٣، ٢٨، ٣٠، ٣٢، ٣٥، ٤٤.

فالوسيط هو السن ٢٨ سنة لأنه يقسم التوزيع إلى جزئيين متساويين ٤ أفراد قبله و٤ أفراد بعده.

وعلى هذا يصبح ترتيب الوسيط عندما يكون عدد القيم فرديا عبارة عن محصلة قسمة عدد الحالات / البيانات + ١ تقسيم ٢، ثم إذا حدث وتحصلنا على كسورا عشرية فسنقوم بتقريب الناتج إلى أقرب عدد صحيح لمعرفة ترتيب الوسيط.

$\frac{٩ + ١}{٢}$ $\frac{ن + ١}{٢}$

وفي المثال السابق يصبح ترتيب الوسيط هو ـــــــ ـــــــ = ٥ =

وبالفعل أخذ الوسيط الرتبة الخامسة في السلسلة المرتبة تصاعديا السابقة.

مثـــــــال ٢: في حالة وجود بيانات زوجية العـــــــدد

أما إذا كان العدد زوجيا فسوف تصبح هناك قيمتين للوسيط وعلينا أن نأخذ متوسطها.

ومثال ذلك إذا زدنا قيمة أخرى للمثال السابق وليكن فردا عاشرا سنه يبلغ ٣١ سنة فسيصبح الوسيط حينها ١٥، ١٧، ١٨، ٢٣، ٢٨، ٣٠، ٣١، ٣٢، ٣٣، ٤٤.

لهذا نحسبه أو بالأحرى نجده عن طريق أخذ المتوسط لهذين الوسطين أي:

و = $\frac{٢٨ + ٣٠}{٢}$ = ٢٩ سنــــــة.

أما عن رتبته فتساوي ـــــــ $\frac{١ + ١٠}{٢}$ = $\frac{١١}{٢}$ = ٥،٥

وبالفعل ٢٩ سنة أخذ المرتبة ٥،٥ في السلسلة المرتبة تصاعديا.

أما عن مرتبة القيم التي تدخل في حساب قيمة الوسيط فهي $\frac{ن}{٢}$ = $\frac{١٠}{٢}$ = ٥

و $\frac{١٠}{٢}$ + ١ = ـــ =5.5 ≈ ٦ وبالفعل ٢٨ أخذ الرتبة ٥ و ٣٠ المرتبة السادسة.

- 2 - 2حساب الوسيط في حالة البيانات المبوبـــــــة:

أما إذا كانت المعطيات مرتبة في جدول تكراري ذو فئات فحساب الوسيط يكون أولا بحساب رتبته.

حيث أن رتبة الوسيط تساوي رتبة الوسيط في حالة البيانات المبوبـــــة = $\frac{ن}{٢}$

ثم نقوم بالبحث عن هذه الرتبة أما في التكرار المجتمع الصاعد أن النازل أم التكرار المجتمع النسبي الصاعد أم النازل ... الخ، لهذا فهناك عدة طرق الحساب الوسيط نذكر منها.

٢ – ٢ – ١ الطريقة الأولـــــى:

مثـــــــال:لدينا الجدول التالي الذي يبين وزن ١٠٠ مولود جديد والمطلوب حساب وسيط تلك الأوزان.

الوزن عند الولادة بالغرام	ك
١٠٠٠ – ١٥٠٠	٣
١٥٠٠ – ٢٠٠٠	٦
٢٠٠٠ – ٢٥٠٠	١٢
٢٥٠٠ – ٣٠٠٠	١٨
٣٠٠٠ – ٣٥٠٠	٢٥
٣٥٠٠ – ٤٠٠٠	١٨
٤٠٠٠ – ٤٥٠٠	١٠
٤٥٠٠ – ٥٠٠٠	٨
المجمـــوع	١٠٠

الطريقة الأولـــــى:

$$\text{الوسيـــط} = \text{الحد الأدنى للفئة الوسطية} \quad \frac{\text{رتبة الوسيط} - \text{التكرار المجتمع الصاعد للفئة قبل الوسطية}}{\text{تكرار الفئة الوسطية}} \times \text{طول الفئة الوسطية}$$

الفئـــــات	ك	ك المجتمع الصاعد	التكرار النسبي	التكرار المجتمع النسبي	ك المجتمع النازل	ك المجتمع النسبي النازل
١٠٠٠ – ١٥٠٠	٣	٣	٠،٠٣	٠،٠٣	١٠٠	١
١٥٠٠ – ٢٠٠٠	٦	٩	٠،٠٦	٠،٠٩	٩٧	٠،٩٧
٢٠٠٠ – ٢٥٠٠	١٢	٢١	٠،١٢	٠،٢١	٩١	٠،٩١
٢٥٠٠ – ٣٠٠٠	١٨	٣٩	٠،١٨	٠،٣٩	٧٩	٠،٧٩
٣٠٠٠ – ٣٥٠٠	٢٥	٦٤	٠،٢٥	٠،٦٤	٦١	٠،٦١
٣٥٠٠ – ٤٠٠٠	١٨	٨٢	٠،١٨	٠،٨٢	٣٦	٠،٣٦
٤٠٠٠ – ٤٥٠٠	١٠	٩٢	٠،١٠	٠،٩٢	١٨	٠،١٨
٤٥٠٠ – ٥٠٠٠	٨	١٠٠	٠،٠٨	١	٨	٠،٠٨
المجمـــــوع	١٠٠					

أول شيء نقوم به لحساب الوسط هو إيجاد رتبته وفي هذا المثال نجدها تقدر بـــ

$$\frac{\text{ن}}{٢} = \frac{١٠٠}{٢} = ٥٠$$

إذا ألقينا نظرة على التكرار المجتمع الصاعد نجد هذه الرتبة ضمتها فئة وزن ٣٠٠٠ – ٣٥٠٠ لوجود ٦٤ طفل، إذ أن التكرار المجتمع قبل هذه الفئة ٣٩ طفل في فئة ٢٥٠٠ – ٣٠٠٠ غ أي أن ترتيب الوسط لا يغطي هذه الفئة بل يقع في الفئة ٣٠٠٠ – ٣٥٠٠ وبتطبيق القانون السابق لإيجاد الوسيط نجد أن

و = حـو + ((ن/٢ - ن س) / ك و) × ل = ٣٠٠٠ + ((٥٠ - ٣٩) / 50) × ٥٠٠ =

٣٠٠٠ + (٥٠٠٠ × ٠،٤٤) = ٣٢٢٠ غ.

٢ – ٢ - 2 الطريقة الثانيـــــــــة:

وفي هذه الطريقة نقوم باستخدام التكرار المجتمع النازل بدل التكرار المجتمع الصاعد ويصبح بذلك القانون كالآتي:

الوسيط = **الحد الأعلى للفئة الوسطية** – (**رتبة الوسيط – التكرار المجتمع النازل للفئة بعد الوسطي** / **تكرار الفئة الوسطية**) **× طول الفئة**

وبتطبيق القانون على المثال السابق نجد ويساوي

و = ٣٥٠٠ – ((٥٠ – ٣٦) / ٢٥) × ٥٠٠ = ٣٥٠٠ – ٢٨٠ = ٣٢٢٠ غ

٢ – ٢ - 3 الطريقة الثالثــــــــة:

وهذه الطريقة تسمى بالطريقة النسبية لهذا فعند استعمالها لا بد من حساب التكرار النسبي وقانونها كالآتي :

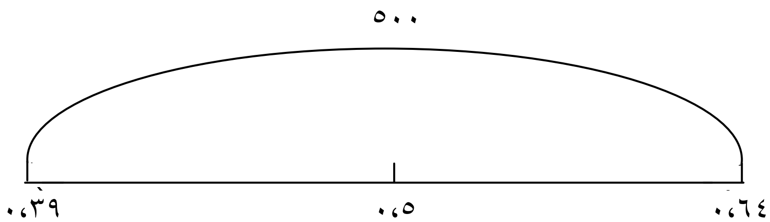
الوسيط = قيمة الحد الأدنى للفئة الوسطية + (مجال الفئة الوسطية / التكرارات النسبية للفئة الوسطية) (٠،٥ – التكرارات النسبية المجتمعة الصاعدة للفئة قبل الوسطية)

وبتعويض هذه المعادلة على المثال السابق نجد:

و = ٣٠٠٠ + (٥٠٠ / ٠،٢٥) (٠،٥ – ٠،٣٩) = ٣٠٠٠ + ٢٠٠٠ (٠،١١) = ٣٢٢٠ غ

٢ – ٢ – ٤ الطريقة الرابعـــــة:

وهي طريقة بيانية نسبيا وهي كالتالـــــي:



و = ٣٠٠٠ س

٠،٥٠ – ٠،٣٩ = ٠،١١

لدينـــا ٠،٢٥ ⟵ ٥٠٠غ

٠،١١ ⟵ س

س = (٠،١١ × ٥٠٠٠) / ٠،٢٥ = ٢٢٠ غ.

و = ٣٠٠٠ + ٢٢٠ = ٣٢٢٠ غ.

٢ – ٢ – ٥ الطريقة الخامســـــة:

وهي طريقة نسبية كذلك وتحسب وفق القانون التالــــــي:

الوسيط = الحد الأدنى للفئة الوسطية + ((٠،٥ – التكرار المجتمع الصاعد النسبي السابق للفئة الوسطية) / التكرار النسبي للفئة الوسطية) × طول الفئة الوسطية

و = ٣٠٠٠ + $\frac{٠،٠٥ - ٠،٣٩}{٠،٢٥}$ × ٥٠٠ = ٣٠٠٠ + ٢٢٠ = ٣٢٢٠ غ

٢ – ٢ – ٦ الطريقة السادســـة:

وهي طريقة تقريبية كذلك ويحسب الوسيط فيها كما يلـــي:

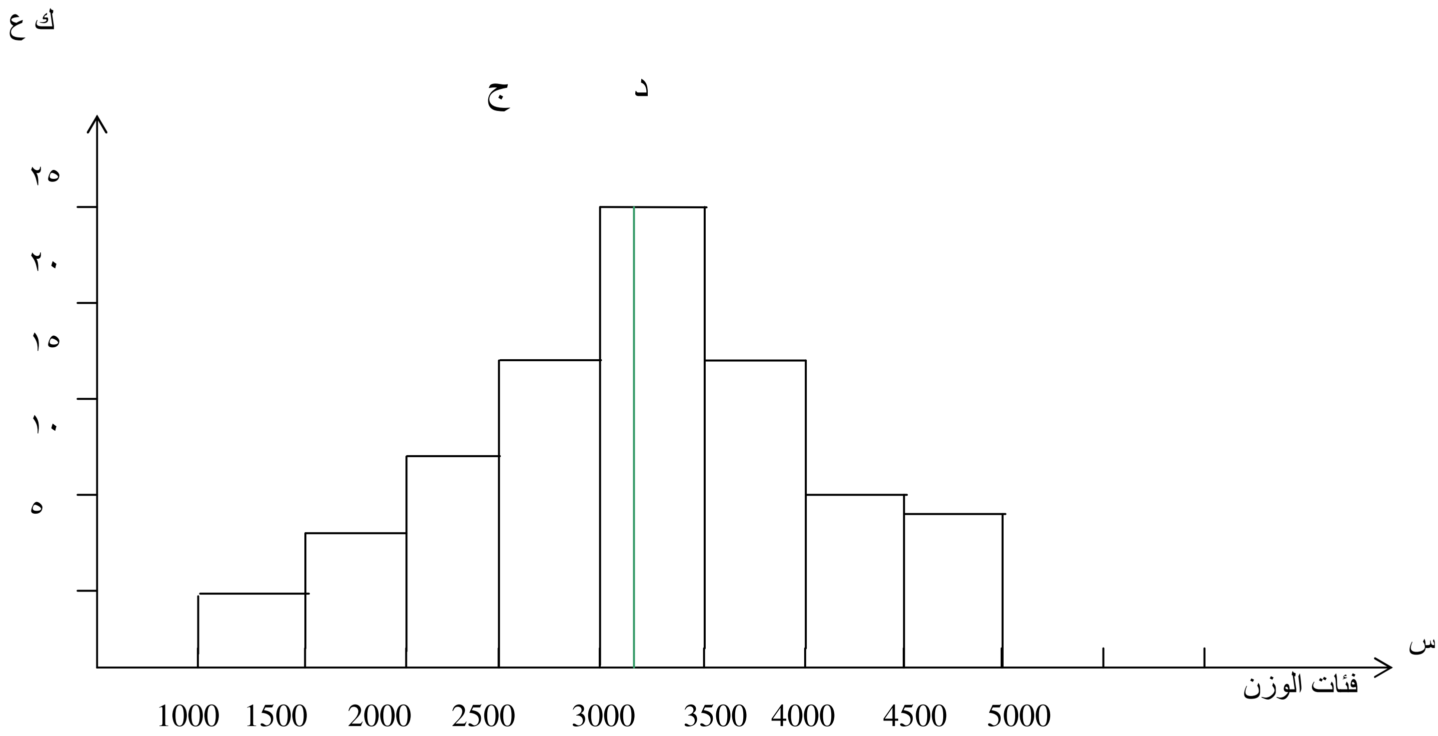
إن التكرارات المجتمعة قبل الفئة الوسطية تساوي ٣٩.

إذن نأخذ ١١ طفل من ذوي وزن ٣٠٠٠ – ٣٥٠٠ والبالغ عددهم ٢٥ طفل ويصبح الوسيط بذلك يقدر بـــ:

و = ٣٠٠٠ + $\frac{١١}{٢٥}$ (٣٥٠٠ – ٣٠٠٠) = ٣٠٠٠ + (٠،٤٤) (٥٠٠) = ٣٢٢٠ع.

٢ – ٢ – ٧ الطريقة السابعـــــة:

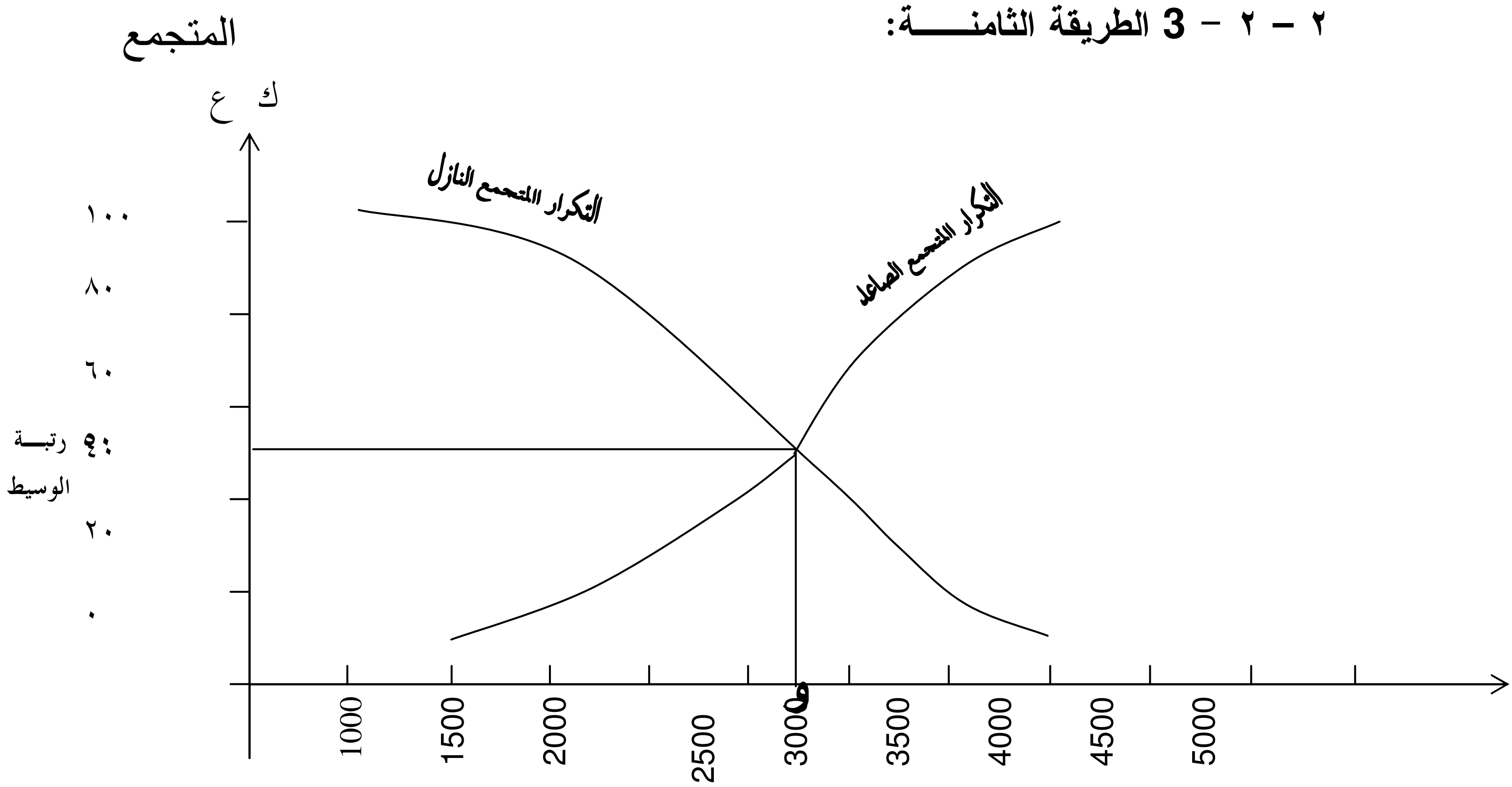
وهي عن طريقة رسم المدرج التكراري للبيانات كما يلـــــي:



لدينا من الرسم أو = (أب) = (٥٠٠) = ٢٠٠

أو = ٣٠٠٠ + ٢٢٠ = ٣٢٢٠ غ.

٢ – ٢ – 3 الطريقة الثامنـــــة:



كما توجد طرق أخرى لإيجاد الوسيط لكن هذه أهمها.

٢– ٣– ترتيب الوسيط عند الفئة التي لا تكرار لهـــــا:

إذا صادفنا ترتيب الوسيط عند فئة منعدمة التكرار فمن الصعب الاستعانة بإحدى القوانين السالفة الذكر لهذا نقوم بحسابه كما يلـــــــي:

مثـــــــال: لدينا الجدول التالي الذي يبين توزيع مجموعة من الأفراد حسب سنهم والمطلوب إيجاد الوسيط أي وسيط سنهم.

السـن	ك
١٠ – ٢٠	٨
٢٠ – ٣٠	١٨
٣٠ – ٤٠	٣٤
٤٠ – ٥٠	٠
٥٠ – ٦٠	٢٠
٦٠ – ٧٠	٨
٧٠ – ٨٠	١٦
٨٠ – ٩٠	١٦
المجمـــوع	**١٢٠**

لحساب الوسيط في هذه الحالة نقوم بحساب التكرار المجتمع الصاعد والنازل وطبعا نقوم بحساب رتبة الوسيط التي تساوي $\frac{ن}{٢}$ = $\frac{١٢٠}{٢}$ = ٦٠

السن	ك	ك الصاعد	ك الضابط
١٠ – ٢٠	٨	٨	١٢٠
٢٠ – ٣٠	١٨	٢٦	١١٢
٣٠ – ٤٠	٣٤	٦٠	٩٤
٤٠ – ٥٠	٠	٦٠	٦٠
٥٠ – ٦٠	٢٠	٨٠	٦٠
٦٠ – ٧٠	٨	٨٨	٤٠
٧٠ – ٨٠	١٦	١٠٤	٣٢
٨٠ – ٩٠	١٦	١٢٠	١٦
المجموع	١٢٠		

نلاحظ من هذا الجدول أن التكرار المتجمع الصاعد يصل إلى ٦٠ عند فئة ٣٠– ٤٠ وهي الفئة الوسيطة بما أن رتبة الوسيط هي ٦٠ إذن فتوجد ضمنها ثم يظل كما هو في الفئة التي تليها لأن تكرارها متساويا لصفر، إذن فالوسيط يقع في نهاية الفئة ٣٠– ٤٠. وإذا ما نظرنا إلى التكرار المتجمع النازل فنجد أن رتبة الوسيط ٦٠ عند الفئة التي تمتد أطرافها من ٥٠– ٦٠ ثم يظل ثابتا في الفئة التي تسبقها، إذن فالوسيط يقع في بداية الفئة ٥٠– ٦٠ أي أن ترتيب الوسيط بهذا يقع بين ٤٠– ٥٠ وهي حدود الفئة المنعدمة التكرار لهذا فالوسيط في مثل هذه الحالة مركز الفئة التي لاتكرار لها أي أن الوسيط = $\frac{٤٠+٥٠}{٢}$

= ٤٥

٤- المئينـــــــــات:

هو عبارة عن تقسيم مساحة المنحنى لتوزيع تكراري إلى مئة جزء متساوي فالمئين الأول $م_{١}$ هو القيمة التي يسبقها ١% من البيانات ويليها ٩٩% من القيم وهذا بافتراض أن البيانات مرتبة مسبقا ترتيبا تصاعديا والمئين الأربعون ٤٠ $م_{٤٠}$ هو القيمة التي يسبقها ٤٠% من البيانات ويليها ٦٠% وبالطبع بالافتراض أن هذه البيانات/القيم مرتبة ترتيبا تصاعديا.

المئين في حالة البيانات المبوبـــــــة:

ويحسب وفق المعادلة التالية في حالة التكرار المجتمع الصاعد

$$\text{المئين} = \text{الحد الأدنى للفئة المئينة} + \frac{\text{ترتيب المئين} - \text{التكرار المتجمع الصاعد السابق للفئة المئينة}}{\text{تكرار الفئة المئينة}} \times \text{طول الفئة}$$

مثـــــال: أحسب المئين لتوزيع ١٢٠ فرد حسب سنهم المطلوب حساب $م_{٣٠}$

الســن	ك	ك الصاعد	ك النازل
١٠ – ٢٠	٨	٨	١٢٠
٢٠ – ٣٠	١٨	٢٦	١١٢
٣٠ – ٤٠	٣٤	٦٠	٩٤
٤٠ – ٥٠	١٨	٧٨	٦٠
٥٠ – ٦٠	٢	٨٠	٤٢
٦٠ – ٧٠	٨	٨٨	٤٠
٧٠ – ٨٠	١٦	١٠٤	٣٢
٨٠ – ٩٠	١٦	١٢٠	١٦
المجمـــــوع	١٢٠		

٣٠

لإيجاد المئين نستخرج ترتيب م٣٠ = $\frac{\quad}{١٢٠}$ × ١٢٠ = ٣٦.

إذا م٣٠ = ٣٠ + $\frac{٣٦ - ٢٦}{٦٠ - ٢٦}$ × ١٠ = ٣٢،٩٤

٤ – **الربيعيات:**

يقصد بالربيعيات في التحليل الإحصائي هي تقسيم مساحة منحنى التوزيع التكراري إلى أربعة أجزاء متساوية، حيث يوجد ثلاث ربيعيات مرتبة من اليسار إلى اليمين وهـــــي:

- الربيع الأول أو الأدنى م٢٥.
- الربيع الثاني أو الوسيط م٥٠.
- الربيع الثالث أو الأعلى م٧٥.

ففي بعض الأحيان نحتاج إلى معرفة القيمة إلي تقع عند ربع السلسلة الإحصائية الصاعد أم النازلة لمجموعة من البيانات مثلما نعمل في الوسيط أين كنا نبحث عن القيمة التي تقيم السلسلة الإحصائية إلى جزأين متساويتين.

تعريف الربيع الأدنـــى: هو القيمة التي يقل عنها ربيع القيم ويزيد عنها¾ القيم ويمكن إيجاده بعد ترتيب البيانات تصاعديا أم تنازليا:

في حالة البيانات المبوبـــــــة:

رتبة الربيع الأدنى ر١ = $\frac{ن}{٤}$

أما قيمته فتحسب وفق المعادلة التاليـــــة:

قيمة ر١ =

الحد الأدنى لفئة الربيع الأول + $\frac{\text{ترتيب الربيع الأول – التكرار المتجمع الصاعد السابق للفئة الربيعية}}{\text{تكرار االفئة الربيعية}}$ × طول الفئة الربيعية

وبتطبيق القانون على الجدول أعلاه نجد ر١ تقدر بما يلي:

- رتبة ر١ = $\frac{ن}{٤}$ = $\frac{١٢٠}{٤}$ = ٣٠

ر١ = ٣٠ + $\frac{٣٠ - ٢٦}{٦٠ - ٢٦}$ × ١٠ = ٣٠ + $\frac{٤}{٣٤}$ × ١٠ = ٣١،١٨ سنة

الربيع الأعلى/ الثالـــــــث:

وهو القيمة التي يسبقها ثلاث أرباع القيم 3/4 ويليها الربع أي ربع القيم ويحسب كما يلي:

إن رتبة ر٣ في هذه الحالة تحسب وفق المعادلة التالية ر٣ = $\frac{٣ن}{٤}$ وهناك حالتين لإيجاده مثـــل ر١.

في حالة التكرار المتجمع الصاعـــــد:

ر٣ = الحد الأدنى لفئة الربيع الثالث + $\frac{\text{ترتيب الربيع الثالث – التكرار المتجمع الصاعد السابق لفئة الربيع الثالث}}{\text{تكرار فئة الربيع الثالث}}$ × طول الفئة الربيعية

وبالتطبيق على المثال أعلاه نجد رتبة ر٣ = $\frac{٣ \times ١٢٠}{٤}$ = ٩٠

إذن الفئة الربيعية الثالثة هي: ٧٠ – ٨٠.

وبالتعويض في القانون نجد: ر٣ = ٧٠ + $\frac{٩٠ - ٨٨}{١٢٠ - ٨٨}$ × ١٠ = ٧٠،٠٩

أي أن ٧٥% من هؤلاء الأفراد يقل سنهم عن ٧٠،٠٩ سنة وأن ٢٥% منهم يزيد عن ٧٠،٠٩ سنة.

5 - العشيـــــــــــرات:

نقصد به تقسيم مساحة المنحنى لتوزيع التكراري إلى عشرة أقسام متساوية وكل قسم نسميه عشير، فمثلا العشير الرابع هو القيمة التي يسبقها ١٠/٤ من البيانات ويليها ١٠/٦ منها طبقا بافتراض دائما ترتيب البيانات تصاعديا.

في حالة البيانات المبوبة: ويحسب كما يلي:

$$\textbf{رتبة العشير} = \frac{\text{العشير المطلوب} \quad \frac{\text{مجموع}}{\text{ك}}}{١٠}$$

أما قيمته فتحسب وفق المعادلة التاليـــــــة:

$$\textbf{قيمة العشير العشيرية} = \textbf{الحد الأدنى لفئة العشير} + \frac{\text{رتبة العشير} - \text{التكرار المتجمع السابق } \textbf{للفئة العشرية}}{\textbf{تكرار الفئة العشرية}} \times \textbf{طول الفئة}$$

وبتطبيق القانون على المثال السابق نجد العشير الثالث مثلا يقدر بـــ:

$$\text{رتبة العشير الثالث} = \text{م}_{٣} = \frac{٣}{١٠} \times ١٢٠ = ٣٦$$

$$\text{م}_{٣} = ٣٠ + \frac{٣٦ - ٢٦}{٦٠ - ٢٦} \times ١٠ = ٣٢،٩٤$$

٦ - المنـــــــــوال:

المنوال لمجموعة من القيم هو القيمة الأكثر انتشارا أو الأكثر تكرارا وشيوعا بين القيم وهذا هو الأساس الذي بناءا عليه يعتبر المنوال وسطا ممثلا للقيم التي حسب لأجلها وعلى ذلك فإن تحديده يتوقف على تكرار القيم المجموعة الإحصائية ويتم إيجاد المنوال

حسابيا وعن طريق الرسم إلا أن قيمة المنوال قد لا توجد، وحتى وأن وجدت في بعض الأحيان قد لا تكون واحدة في السلسلة الإحصائية أو الجدول التكراري.

أ/ في حالة البيانات غير المبوبـــــــــــة:

إذا لم يتكرر أي من القيم فلا يوجد منوالا.

مثال ١: لدينا البيانات التالية: ٣، ٥، ٨، ١٠، ١٢، ١٥، ١٦.

لا يوجد منوال لهذه القيم حيث أيا من القيم لم يتكرر أكثر من مرة.

مثال ٢: إذا تكرر أحد القيم فيكون منوال واحد فقط مثلا في السلسلة التالية ٢، ٢، ٥، ٧، ٩، ٩، ٩، ١٠، ١٠، ١١، ١٢، ١٨.

فالمنوال هو رقم ٩ لأنه هو الذي تكرر كثيرا ٣ مرات عن القيم الأخرى

مثال ٣: كما يمكن أن نأخذ السلسلة الإحصائية منوالين فأكثر كما هو مبين في المثال التالي: ٢، ٣، ٤، ٤، ٤، ٥، ٥، ٧، ٧، ٧، ٩.

إذا هذه السلسلة لديها منوالين هما ٤، ٧ ويسمى المنوال في هذه الحالة بالمنوال المزدوج.

ب/ في حالة البيانات المبوبـــــــــــة:

إن طريقة حساب المنوال في هذه الحالة تختلف عن كيفية استخراجه من البيانات المبوبة وهناك عدة طرق لحسابه من بينها ما يلي:

مثال: أوجد المنوال السن الأكثر شيوعا لعينة متكونة من ١٧٠ امرأة موزعة حسب سنها عند أول زواج.

السـن عنـد أول زواج	ك
١٥ – ١٨	١٠
١٨ – ٢١	٢٠
٢١ – ٢٤	٢٥
٢٤ – ٢٧	٣٠
٢٧ – ٣٠	٤٠
٣٠ – ٣٣	٢٠
٣٣ – ٣٦	١٨
٣٦ – ٣٩	٥
٣٩ – ٤٢	٢
المجمـــــوع	**١٧٠**

الطريقة الأولـــــــــى:

باعتبار مركز الفئة ذات أكبر تكرار في الجدول التكراري، ومن الجدول السابق نلاحظ أن المنوال هذا التوزيع هو مركز الفئة ٢٧ – ٣٠ وهو بذلك: $\frac{٢٧ + ٣٠}{٢}$ = ٢٨،٥ سنة، وذلك لأن تكرارها هو أكبر تكرار في الجدول.

إذن قدر بـــ ٤٠ امرأة متزوجة ما بين سن ٢٧ – ٣٠ سنة، والواضح أن هذه الطريقة هي تقريبية فهي تفترض تماثل التوزيع على جانبي مركز الفئة المنوالية.

الطريقة الثانيـــــــــة:

وهي طريقة حسابية نستخدم من خلالها الصيغة التاليـــــــة:

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{تكرار الفئة بعد المنوالية}}{\text{تكرار الفئة بعد المنوالية} + \textbf{تكرار الفئة قبل المنوالية}} \times \text{طول الفئة المنوالية}$$

$$\textbf{مو} = \text{أ} + \frac{\text{ك}_{٢}}{\text{ك}_{٢} + \text{ك}_{١}} \times \text{ل}$$

وبتطبيق هذا القانون على المثال السابق نجد المنوال يساوي:

$$\text{مو} = ٢٧ + \frac{٢٠}{٢٠ + ٣٠} \times ٣ = ٢٧،٩٩ \text{ سنة}$$

الطريقة الثالثــــــــة:

وهي طريقة الفروق وواضع هذه الطريقة هو كار ل برسون وهي لا تختلف كثيرا عن سابقتها فهي تهتم بالفروق بين التكرارات أكثر مما تهتم بالتكرارات نفسها ذلك لأن الخطوة الأولى في هذه الطرقة تنحصر في إيجاد الفرق بين التكرارات الفئة المنوالية

والفئتين اللتين حولها، ووضع المنوال يتحدد في هذه الطرقة بالفرق بين تكرار الفئة المنوالية والتي حولها وبتالي يحسب المنوال بهذه الطريقة وفق لصيغة التالية:

المنوال =

$$\text{الحد الأدنى للفئة المنوالية} + \frac{\text{تكرار الفئة المنوالية} - \text{تكرار الفئة قبل المنوالية}}{(\text{تكرار الفئة المنوالية} - \text{تكرار الفئة قبل المنوالية}) + (\text{تكرار الفئة المنوالية} - \text{تكرار الفئة بعد المنوالية})} \times \text{طول الفئة المنوالية}$$

$$\text{مو} = \text{أ} + \frac{\text{ف}١}{\text{ف}١ + \text{ف}٢} \times \text{ل}$$

وبتطبيق قانون برسون على المثال نجد مو $= ٢٧ + \frac{٤٠ - ٣٠}{(٤٠-٣٠) + (٤٠-٢٠)} \times ٣ =$ ٢٧،٩٩ سنة

نلاحظ أن المنوال محسوبا بكلتا الطريقتين الثانية والثالثة متساوي إلا أنه مع الطريقة الأولى لا يساويه وهذا ليس بغريب حيث أن الطريقة الأولى كما قلنا تقريبية وعدم الدقة أو التفاوت في النتيجة يرتفع كلما كان مدى الفئة طويلا.

الطريقة الرابعـــــــــــة:

وهي طريقة الرافعة إذ أنها تنظر إلى الفئة المنوالية على أنها تمثل رافعة تتجاوزها قوة يعبر عنها تكرار الفئة قبل المنوالية ومقاومة يعبر عنها بتكرار الفئة بعد المنوالية وعلى هذا يمكن تحديد موضع المنوال أو قيمته عند نقطة ارتكازها/ الرافعة، وبالنظر إلى المثال السابق يصبح المنوال يقدر بما يلي:

لما كان قانون الرافعة يساوي:

القوة × ذراع = المقاومة × ذراع

فإن التكرار السابق للفئة المنوالية × س = التكرار اللاحق للفئة المنوالية × (١٠ - س)

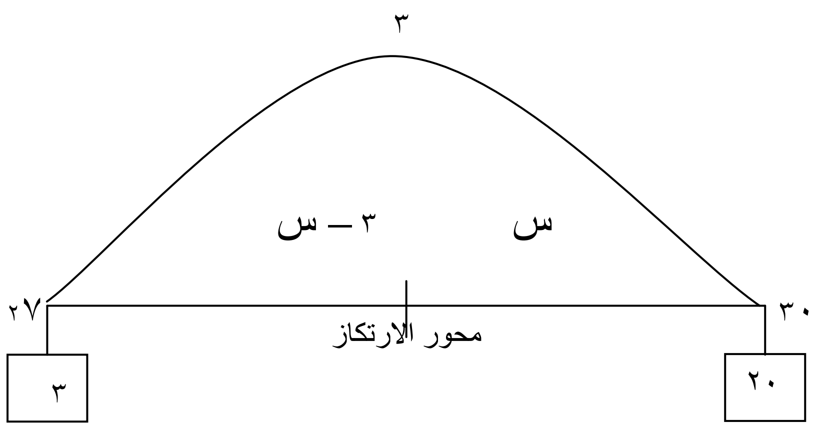
بالتعويض نجد: ٣٠ × س = ٢٠ × ٣ – س

٣٠ س = ٦٠ – ٢٠ س

٣٠ س + ٢٠ = ٦٠

٥٠ س = ٦٠ س = ٦٠ / ٥٠ = ١،٢

إذا فالمنوال = ٢٧ + ١،٢ = ٢٨،٢ سنة



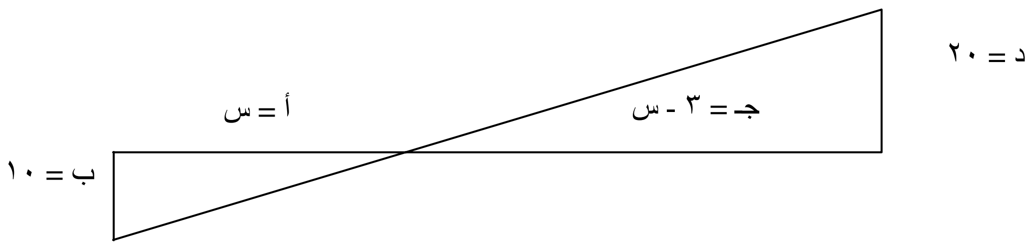
الطريقة الخامســـــة:

وهي طريقة المحاور، ونستعملها لإيجاد المنوال بفرض أن قيمة المنوال تقع على بعد س من بداية الفئة وعلى بعد ٣ – س من نهايتها بحيث يصبح المنوال = ٣٠ + س وبما أن: أ / ب = جـ / د

س / ١٠ = (٣– س) / ٢٠ أي: ٢٠ س = ٣٠ – ١٠ س ⇦ ٢٠ س + ١٠ س = ٣٠

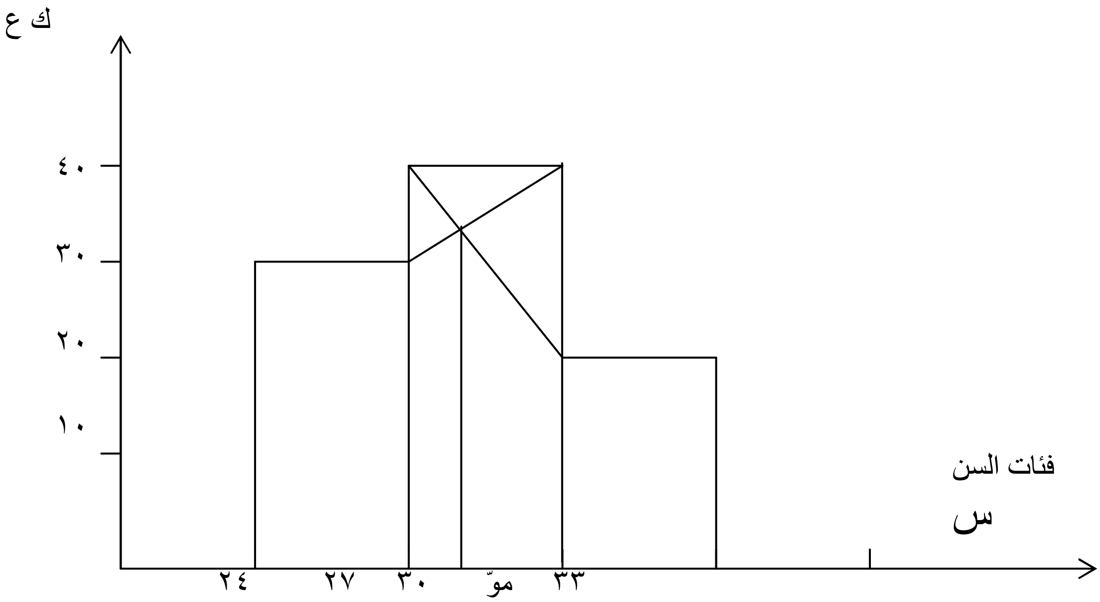
⇦ ٣٠ س = ٣٠ ⇦ س = ٣٠ / ٣٠ = ١

المنوال = ٢٧ + ١ = ٢٨ سنة.



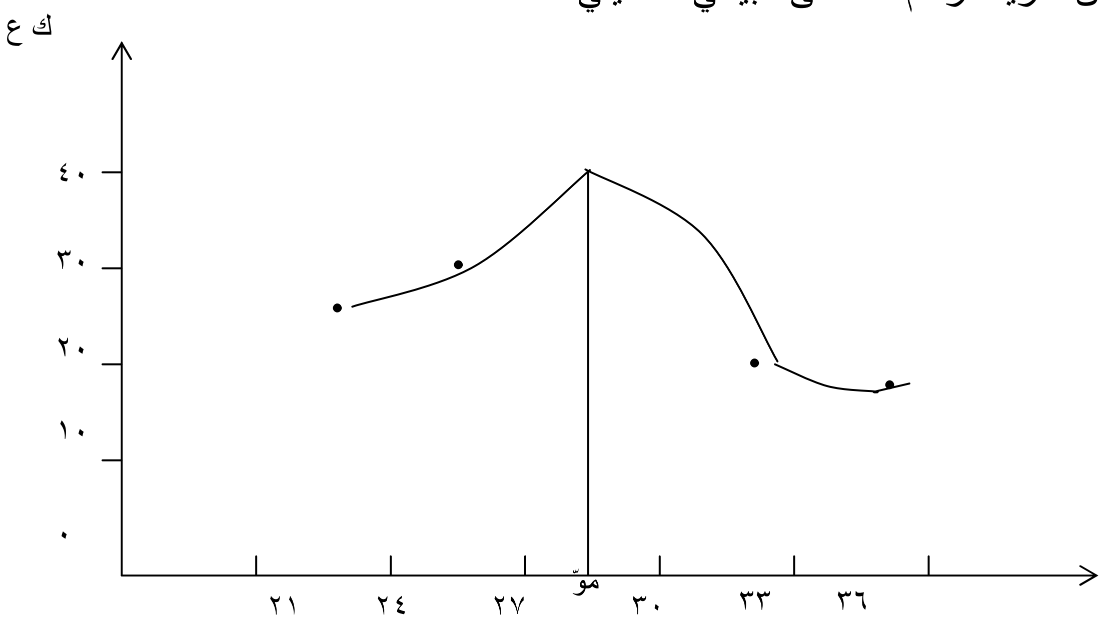
الطريقة السادســـــة:

وهي طريقة تستعمل عن طريقة الرسم المدرج التكراري كما يلي:



الطريقة السابعـــــة:

عن طريقة رسم المنحنى البياني كما يلي:



إيجاد المنوال في حالة البيانات الكمية المفصلـــــــــة:

إن أكبر تكرار يقابل المتغيرات المنفصلة في جدول تكراري هو المنوال كما نلاحظه من هذا الجدول الذي يبين عدد الأبناء لمجموعة من العائلات.

عدد الأبناء في العائلة	عدد العائلات
٠	٣
١	٧
٢	١١
٣	١٤
٤	٢٠
٥	١٦
٦	١٢
٧	٧
المجموع	**٩٠**

وفي هذا الجدول ذوي أن المنوال فيه موجود عند القيمة ٤ أي أربعة أطفال في العائلة وعلى ذلك يكون المنوال ٤ مباشرة دون إجراء عمليات حسابية كما هو في حالة البيانات الكمية المتصلة فبمجرد ملاحظة التكرارات فأكبرها في التوزيع مثل ما هو في هذا الجدول ٢٠ عائلة القيمة التي تقابلها هو المنوال.

إيجاد المنوال من البيانات الكيفيــــــة:

استجوبت عينة من عمال مؤسسة ما فكان توزيعهم حسب حالتهم المدينة كالآتي:

الحالــة المدنيــــــة	ك
أعزب	٣٠
متزوج	٣٠
مطلق	١٢
أرمل	١٠
المجمــــــوع	**٨٢**

المنوال في هذا التوزيع هو القيمة التي تقابل أكبر تكرار وهم إذن العمال العزاب والمتزوجين إذن فهذا التوزيع يحتوي على منوال مزدوج.

المنوال في التوزيعات التكرارية الغير منتظمة:

إيجاد المنوال في التوزيعات التكرارية الغير منتظمة لابد من تعديل التكرارات ثم نطبق أي من الصيغ السالفة الذكر، كما هو مبين في المثال التالي:

مثـــــال: أوجد المنوال من الجدول التالي الذي يبين مدة أقدميه العمال في مهنتهم في مؤسسة ما.

الفئات	ك	س	ك / س	الفرق
١٠ – ١٢	٥	٢	٢،٥	
١٢ – ١٦	١٢	٤	٣	
١٦ – ٢٠	١٧	٤	٤،٢٥	
٢٠ – ٢٥	٤٠	٥	٨	٣،٧٥
٢٥ – ٣٥	٢١	١٠	٢،١	٥،٩
المجمـــــوع	**٩٥**			

يتضح من التوزيع هذا أن الفئة المنوالية هي: ٢٠ – ٢٥ لأنها مقابلة لأكبر تكرار معدل والمقدر بـــ ٨

ولإيجاد المنوال يجب إعداد جدول مثل ما سبق لتعديل التكرارات ويصبح بعدها المنوال:

$$\text{مو} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{ك}٢}{\text{ك}١ + \text{ك}٢} \times \text{ل}$$

$$\text{مو} = ٢٠ + \frac{٣،٧٥}{٣،٧٥ + ٥،٩} \times ٥ = ٢١،٩٤ \text{ سنة}$$

٧- العلاقة بين مقاييس النزعة المركزيـــــــــة:

تكمن الإحصائيين من إيجاد علاقة تقريبية بين المتوسطات الثلاثة: المتوسط الحسابي، الوسيط –المنوال- وتستخدم هذه العلاقة عندما يتعذر استخراج إحداها كما يحدث عندما يراد إيجاد المتوسط الحسابي من الجدول التكرارية المفتوحة أين لا يمكنا معرفة مركز الفئة الأولى والأخيرة في الجدول مثلا.

وسوف نعرض فيما يلي طبيعة العلاقة التقريبية في الحالات التاليـــــــة:

١- إذا كان التوزيع التكراري متماثلا تماما:

يكون التوزيع التكراري متماثلا تماما إذا كان المنحنى الذي يمثله يتماثل حول المحور الرأسي، أين يقسم المنحنى إلى جزئيين متطابقين، وهذا النوع من التوزيعات تنقص فيه التكرارات بالتماثل على جانبي أكبر تكرار فيه وهذا ما يسمى بالمنحنى الإعتدالي.

أحسب المتوسط الحسابي، الوسيط والمنوال وكذا قسم برسم المنحنى التكراري للجدول التالــــــي:

الفئـــــــات	ك
١٠ - ١٥	٢
١٥ – ٢٠	١٥
٢٠ – ٢٥	٢٠
٢٥ – ٣٠	٢٦
٣٠ – ٣٥	٢٠
٣٥ – ٤٠	١٥
٤٠ – ٤٥	٢
المجمـــــــــوع	١٠٠

الحـــــــــل:

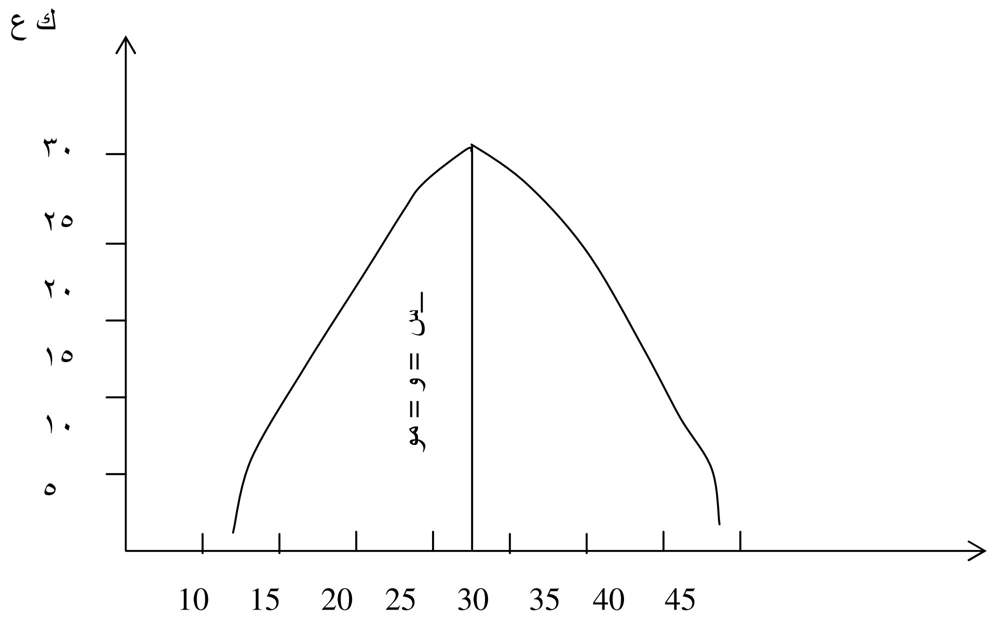
الفئات	ك	س	ك × س	ك م الصاعد	ك م نازل
١٠ – ١٥	٢	١٢،٥	٢٥	٢	١٠٠
١٥ – ٢٠	١٥	١٧،٥	٢٦٢،٥	١٧	٩٨
٢٠ – ٢٥	٢٠	٢٢،٥	٤٥٠	٣٧	٨٣
٢٥ – ٣٠	٢٦	٢٧،٥	٧١٥	٦٣	٦٣
٣٠ – ٣٥	٢٠	٣٢،٥	٦٥٠	٨٣	٣٧
٣٥ – ٤٠	١٥	٣٧،٥	٥٦٢،٥	٩٨	١٧
٤٠ – ٤٥	٢	٤٢،٥	٨٥	١٠٠	٢
المجموع	**١٠٠**		**٢٧٥٠**		

$$\text{س} = \frac{\text{مج ك س}}{\text{مج ك}} = \frac{٢٧٥٠}{١٠٠} = ٢٧،٥$$

$$\text{الوسيط} = \text{حـ و} + \frac{\frac{\text{ن}}{٢} - \text{ن س}}{\text{ك و}} \times \text{ل} = ٢٥ + \frac{٥٠ - ٣٧}{٢٦} \times ٥ = ٢٧،٥$$

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{\text{ف}_1}{\text{ف}_1 + \text{ف}_2} \times \text{ل} = ٢٥ + \frac{٢٠}{٢٠+٢٠} \times ٥ = ٢٧،٥$$

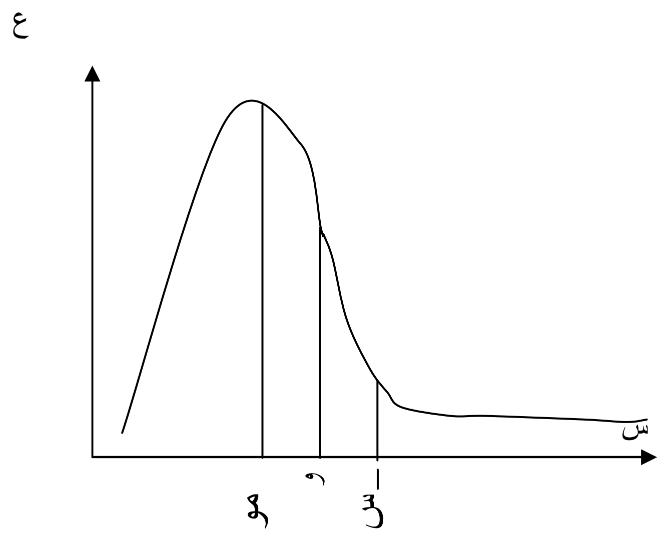
وفي هذه الحالة نجد أن المنحنى يتخذ شكلا اعتدا ليا كما وجدنا س = مو = و = ٢٧،٥ معنى هذا أن جميع مقاييس النزعة المركزية تنطبق على بعضها البعض وتتساوى جميعا في التوزيع التكراري المعتدل.



إذا كانت التوزيعات التكرارية قريبة من التماثل أو ملتوية بطريقة غير شديدة لاحظ أن بعض التكرارات لا تمثل توزيعا معتدلا حول قيمة متوسطة بل تتوزع بطريقة غير معتدلة في الجدول، وفي هذا النوع من الجداول تختلف قيم كل من الوسط/المتوسط الوسط والمنوال حسب نوع الالتواء الموجب والسالب.

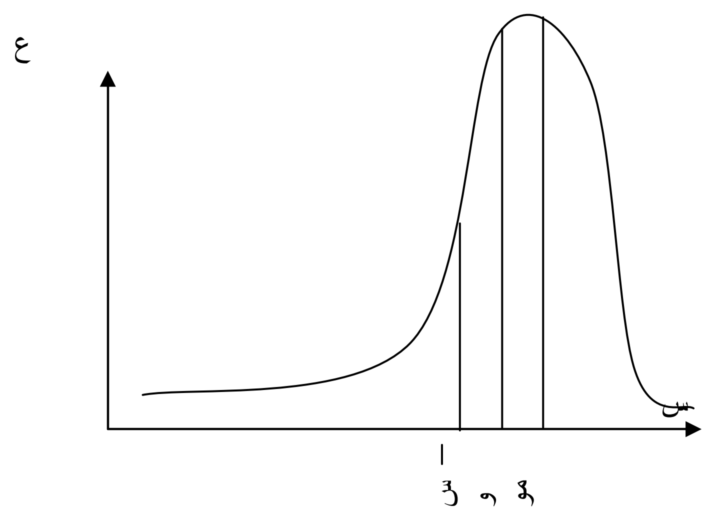
أ/- عندما يكون التوزيع التكراري ملتويا التواء موجب ٥ > ٠

في هذه الحالة يمتد الطرف الطويل للمنحنى إلى الجهة اليمنى ويصبح ترتيب مقاييس النزعة المركزية كما يلي المتوسط، الوسط،ثم المنوال أي أن س، و، مو.



ب/- عندما يكون التوزيع التكراري ملتويا التواء سالب ٥ < ٠

في هذه الحالة يمتد الطرف الطويل إلى الجهة اليسرى ويصبح حينها ترتيب مقاييس النزعة المركزية كما يلي المنوال، الوسيط، المتوسط- أي أن مو، و، سَ.



٨ - خواص مقاييس النزعة المركزية الثـــــــلاث:

٨-١- الوسط الحسابي أو المتوسط الحسابي:

- إن المجموع الجبري لانحرافات مجموعة من القيم عن وسطها الحسابي يساوي دائما صفرا.
- $\sum$ ح = ٠.
- يستغل في حساب المتوسط الحسابي جميع قيم التوزيع لذلك فهو أدق المتوسطات الثلاث وأكثر ثباتا.
- يستخرج المتوسط الحسابي كذلك لاستعماله في معاملات أخرى كمقاس التشتت إلا أن عيوب تكمن في المتغيرات.
- أنه لا يمكن استعماله في البيانات المتغيرات الكيفية.
- لا يمكن حسابه في الجداول التكرارية المفتوحة.

- يتأثر بالقيم الشاذة، أي الكبيرة جدا أو الصغيرة جدا عن بقية قيم الجدول التكراري.

٨-٢- الوسيط أو الأوســـــط:

- إن قيمة الوسيط على العكس من الوسط الحسابي لا تتأثر بالقيم الشاذة لهذا في حالة وجود مثل هذه القيم يفضل استخدام الوسيط.
- يمكن حساب الوسيط من الجداول التكرارية- المفتوحة لأننا في حسابه لا نحتاج لمعرفة مراكز الفئات مثل حالة المتوسط الحسابي.
- يمكن إيجاد قيمة الوسيط من خلال الرسم البياني كالمنحني المتجمع الصاعد والنازل.
- إذا كان البحث يهتم بمعرفة ما إذا كانت قيمة معينة تقع في النصف العلوي أو السفلي من التوزيع.

٨-٣- المنـــــــــــــوال:

- يعتبر المنوال المقياس الوحيد الذي يمكن استعماله لإيجاد تمركز المتغيرات الكيفية.
- ويمتاز المنوال بسهولة حسابه وسرعته، ولما كان المنوال يدل على الدرجة الأكثر شيوعا لذلك فهو يصلح لمعالجة المشاكل التي تهدف إلى معرفة تركيز الظاهرة وموقعها وخاصة في النواحي الديموغرافية، إلا أن حساب المنوال يكون بطرق تقريبية وعموما يفضل المنوال في الحالات التالية.
- إذا أريد الحصول على معامل مركزي في أقصى وقت ممكن دون الاهتمام كثيرا بالدقة في حسابه إذا كان هدف الباحث معرفة القيمة التي يتفق عليها أغلب أفراد المجموعة / العينة.

تمرينات

التمرين لأول:

أحسب متوسط أوزان ١٠٠ طالب التالي.بالطريقة العادية و عن طريق الوسط الفرضي.

الوزن بالكلغ	التكرارات (ك)
٦٠ – ٦٢	٥
٦٣ – ٦٥	١٨
٦٦ – ٦٨	٤٢
٦٩ – ٧١	٢٧
٧٢ – ٧٤	٨
المجموع	١٠٠

الحل: إيجاد مراكز الفئات ثم مجموع مراكز الفئات مضروبا في التكرارات المقابلة لها.كما هو مبين في الجدول ثم الانحرافات عن مراكز الفئات.

الوزن بالكلغ	ك	س	ك × س	ح	ك × ح
٦٠ – ٦٢	٥	٦١	٣٠٥	– ٦	– ٣٠
٦٣ – ٦٥	١٨	٦٤	١١٥٢	– ٣	– ٥٤
٦٦ – ٦٨	٤٢	٦٧	٢٨١٤	٠	٠
٦٩ – ٧١	٢٧	٧٠	١٨٩٠	٣+	٨١
٧٢ – ٧٤	٨	٧٣	٥٨٤	٦+	٤٨
المجموع	١٠٠		٦٧٤٥	٠	٤٥

$$\overline{\text{س}} = \frac{\text{مج}\ (\text{ك} \times \text{س})}{\text{مج}\ (\text{س})} = \frac{6745}{100} = \text{٦٧,٤٥ كلغ}$$

عن طريق الوسط الفرضي س = أ + $\dfrac{\text{مج}\ (\text{ك} \times \text{ح})}{\text{مج}\ (\text{ك})}$

$$= \text{٦٧} + \frac{6745}{100} = \text{٦٧,٤٥ كلغ}$$

التمرين الثاني:

حدد متوسط مدة أقدميه السكن لـــ ٧٠ عائلة في منطقة معينة، موزعين حسب الفئات التالية:

المدة	ك
٥٠ – ٦٠	٨
٦٠ – ٧٠	١٠
٧٠ – ٨٠	١٦
٨٠ – ٩٠	١٥
٩٠ – ١٠٠	١٠
١٠٠ – ١٢٠	٨
١٢٠ – ١٨٠	٣
المجموع	**٧٠**

<u>الحل:</u> حساب مراكز الفئات ثم ضربها في التكرار المقابل لها كما يلي:

المدة	ك	س	ك × س
٥٠ – ٦٠	٨	٥٥	٤٤٠
٦٠ – ٧٠	١٠	٦٥	٦٥٠
٧٠ – ٨٠	١٦	٧٥	١٢٠٠
٨٠ – ٩٠	١٥	٨٥	١٢٧٥
٩٠ – ١٠٠	١٠	٩٥	٩٥٠
١٠٠ – ١٢٠	٨	١١٠	٨٨٠
١٢٠ – ١٨٠	٣	١٥٠	٤٥٠
المجموع	٧٠		**٥٨٤٥**

<u>ملاحظة:</u> نلاحظ أن الفئات في هذا الجدول غير متساوية و رغم هذا فلم نقم بتعديل التكرارات، لأنه ما يهمنا هو مركز الفئة و بالتالي سيكون متوسط مدة إقامة هؤلاء العائلات في تلك المنطقة هو كالتالي بالطريقة العادية:

$$\overline{\text{س}} = \frac{\text{مج}(\text{ك} \times \text{س})}{\text{مج}(\text{س})} = \frac{5845}{70} = \text{٨٣,٥ سنة}$$

المتوسط الحسابي المرجح = $\frac{\text{مج ك × و}}{\text{مج و}}$

مثـــــــال: لدينا ثلاث مصالح للتوليد في ثلاث مستشفيات.

المصلحة الأولى: أ/- عدد النساء اللواتي وضعن عددهن ٢٠ عدد المتوفيات و٠٥.
المصلحة الثانية: ب/- عدد النساء اللواتي وضعن عددهن ٢٥ عدد المتوفيات و٠٣.
المصلحة الثالثة: ج/- عدد النساء اللواتي وضعن عددهن ٣٠ عدد المتوفيات و٠٤.
والمطلوب حساب المتوسط الحسابي المرجح.

$$\bar{\text{س}} = \frac{\text{مج ك و}}{\text{مج و}}$$

المصالـــــح	ك	و	ك × و
أ	٢٠	٥	١٠٠
ب	٢٥	٣	٧٥
ج	٣٠	٤	١٢٠
المجمـــــوع		**١٢**	**٢٩٥**

$$\bar{\text{س}} = \frac{\text{مج ك و}}{\text{مجك}} = \frac{٢٩٥}{١٢} = ٢٤،٥٨$$

التمرين الثالث:

أوجد وسيط مدة زواج ٢٠٠ زوج عن طريق التكرار المتجمع الصاعد.

المدة	ك
٠ – ٤	٣١
٤ – ٨	٣٦
٨ –١٢	٤٠
١٢ – ١٦	٦٠
١٦ – ٢٠	٢٠
٢٠ – ٢٤	١٣
المجموع	٢٠٠

الحل:

المدة	ك	ك المتجمع الصاعد
٠ – ٤	٣١	٣١
٤ – ٨	٣٦	٦٧
٨ –١٢	٤٠	١٠٧
١٢ – ١٦	٦٠	١٦٧
١٦ – ٢٠	٢٠	١٨٧
٢٠ – ٢٤	١٣	٢٠٠
المجموع	٢٠٠	

$$\text{و} = \text{د و} + \frac{\frac{\text{ن}}{2} - (\text{ن})(\text{س})}{\text{ك}_{\text{و}}} \times \text{ف}$$

١- حساب رتبة الوسيط $\frac{\text{ن}}{2} = \frac{200}{2} = ١٠٠$

∴ الفئة الوسيطة ٨ – ١٢ لأنها تحتوي على تكرار ١٠٧

∴ و = ٨ + $\frac{100 - 67}{40}$ × ٤ = ١١,٣ سنة

التمرين الرابع:

الجدول التالي يبين لنا توزيع مجموعة من المرضى في المستشفى مصابين بمرض السرطان حسب سنهم.و المطلوب حساب منوال السن عند هؤلاء المصابين حسب طريقة الفروق **لكارل بارسون.**

السن	ك
٢٠- ٣٠	٣٢
٣٠ – ٤٠	٤٠
٤٠ – ٥٠	٥٠
٥٠ – ٦٠	٣٠
٦٠ – ٧٠	٢٥
٧٠ – ٨٠	١٥
٨٠ – ٩٠	٢٠
المجموع	**٢١٢**

مو = أ+ $\frac{\Delta_1}{\Delta_1+\Delta_2}$ × ل

نلاحظ أن أكبر تكرار في هذا التوزيع هو ٥٠ إذا الفئة المنوالية هي ٤٠ – ٥٠ و بتطبيق القانون نجد مو = ٤٠ + $\frac{50-40}{(50-40)(50-30)}$ × ١٠ = ٤٣,٣٣ سنة

الفصـــل الخامـــــــس

مقاييس التشتت و التبعثر

أولا: المـــــــــدى

ثانيا: الانحراف المتوســط

ثالثا: التباين والانحراف المعياري

رابعا: معامل الاختـــــــلاف

خامسا: مقاس الالتواء والتفرطح

سادسا: العلاقة بين المتوسطات الإحصائية ومقاييس التشتت

مقاييس التشتــــــت:

تبين لنا مقاييس النزعة المركزية كما لاحظنا القيمة المركزية للتوزيع التكراري دون أن تظهر لنا كيف تتوزع وتنتشر قيم المتغير على هذه الكمية المركزية. فهي تلخص وتوصف مجموعة/ عينة بقيمة واحدة ولذلك فمعرفتها واجبة عند القيام بمقارنات بين قيم مجموعات مختلفة مثلا متوسط النساء المستعملات لوسائل منع الحمل مع متوسط النساء اللواتي لا تستعملن وسائل منع الحمل.

كذلك متوسط بين وفاة الذكور بسبب ما مقارنة بمتوسط سن وفيات الإناث بنفس السبب ... إلخ، لكن هل يكفي أن نأخذ هذه المقاييس لوصف قيم المجموعة وصفا كاملا. طبعا لا، لأننا نحتاج إلى معرفة تباعد القيم عن بعضها البعض أي وصف درجة اختلاف مفردات/قيم كل من المجموعتين عن بعضها البعض أو بعبارة أخرى نصف درجة تشتتها.

فمثلا لدينا مجموعتان واحدة تنظم النسل والأخرى لا تنظم النسل حسب الطبقة الاجتماعية التي تنتمي إليها.

الطبقـــــــة	تنظـــــم	لا تنظـــــم
طبقة فقيرة	–	٢٤
طبقة متوسطة	٢٥	٢٥
طبقة مرتفعة	٢٥	٦٠

نلاحظ أن المتوسط الحسابي لكل من المجموعتين = ٢٥ فهل يمكننا أن نقول أن المجموعتين متجانستين حسب المستوى المعيشي أو الطبقي.

إن مجرد النظر لهذه البيانات ندرك أن قيم المجموعة الأولى وهي النساء اللواتي تنظمن نسلهن مبعثرة وغير متجانسة فيما بينها بينما المجموعة الثانية متقاربة جدا أي اللواتي لا تنظمن نسلهن لهذا في هذه الحالة لا بد من توضيح كيف تتوزع هذه القيم حول أوساطها، فهل هذه القيم تكون قريبة من حول المراكز أو بعيدة عنه ومعرفة درجة التباعد هذا سيكشف لنا جهة أخرى مهمة من خواص التوزيع، وهو ما يعبر عنه

بالتشتت. فمقياس التشتت يعطي فكرة سليمة عن صلاحية الوسط لتمثيل البيانات فالباحث يحتاج عادة إلى استخدام قيمة تعبر عن مدى تباعد القيم أو تقاربها في المجموعات التي يشملها البحث.

ويقصد بالتشتت هو تباعد القيم عن بعضها لكن هذا بدوره يحمل بطياته عدة تساؤلات لعدم تجانس البيانات في بعض أوقاته لذا اتفق على نقطة ثابتة لقياس التباعد أو التقارب عن هذه النقطة وهو المتوسط الحسابي خير ممثل لهذه النقطة حيث أن غالبية النقاط تكون قريبة نحو هذه النقطة وقد يكون هذا البعد كبيرا أي أن البيانات مبعثرة والتفاوت كبيرا بينـــــها. البعد قليلا أي أن البيانات/القيم غير مبعثرة والتفاوت قليلا بينها، أو قد يكون هذا البعد متساوي أي لا يوجد تشتت أي تجانس الوحدات ومن أهم مقاييس التشتت تقوم بقياس مدى تجانس التوزيعات الإحصائية أو عدم تجانسها.

١-المــــــــدى:

أ/- المدى في حالة البيانات الغبر مبوبــــة:

وفيه نقوم بترتيب البيانات تصاعديا أم تنازليا ثم نقوم نقوم بحسابه وفق الصيغة التاليـــــة:

المـــــــدى = أكبر قيمـــــة – أصغر قيمـــــة

مثــــــال: لدينا عائلتين موزعتين حسب سن أفرادها

- العائلة الأولى: ١٠، ٨، ١٥، ٣٥، ٣٧
- العائلة الثانية: ٢، ٨، ٤٣، ٤٠، ٥

- المدى عند العائلة الأولى هو: ٨، ١٠، ١٥، ٣٥، ٣٧ ‏ ‏ ‏ ٣٧ – ٨ = ٢٩
- المدى عند العائلة الثانية هو: ٢، ٥، ٨، ٤٠، ٤٣ ‏ ‏ ‏ ٤٣ – ٢ = ٤١

وعلى ضوء هذه القيم المتحصل عليها نلاحظ أن العائلة الثانية أكثر تشتتا من العائلة الأولى، هذا يعني أن مفردات الثانية أكثر تبعثرا من مفردات العائلة الأولى من ناحية سنهم.

وقد نحصل على بيانات تكون فيها بعض القيم متطرفة قيمة كبيرة جدا أو صغيرة جدا، لذا فإن قيمة المدى تتأثر بها ويكون البعد كبيرا لذا ينصح بحرف هذه القيم المتطرفة عن طريقة استخدام إحدى الصبغ التالية:

المدى المئين = المئين الأعلى – المئين الأدنى = المئين ٩٩ – المئين الأول = $م_{٩٩} - م_{١}$

نصف المدى المئيني = $\frac{\text{المدى المئين}}{٢} = \frac{م_{٩٩} - م_{١}}{٢}$

المدى العشيري = العشير التاسع – العشير الأول، $ع_{٩} - ع_{١}$

نصف المدى العشيري = $\frac{ع_{٩} - ع_{١}}{٢}$

المدى الربيعي = الربيع الأعلى – الربيع الأدنى $ر_{٧٥} - ر_{٢٥}$

نصف المدى الربيعي = $\frac{\text{الربيع الأعلى – الربيع الأدنى}}{٢} = \frac{ر_{٧٥} - ر_{٢٥}}{٢}$

مثـــــال: لدينا البيانات التالية التي بين علامات ٨ طلبة في مادة علم الإجتماع العام والمطلوب حساب، المدى المطلق ونصف المدى الربيعي، ١٥، ١٣، ١٠، ٨، ٦، ١٦، ١١، ١٢

الحـــــل:

أولا الترتيب التصاعدي للعلامات ٦، ٨، ١٠، ١١، ١٢، ١٣، ١٥، ١٦

المدى المطلق = أكبر قيمة – أصغر قيمة = ١٦ – ٦ = ١٠

إيجاد نصف المدى الربيعي:

– رتبة الربيع الأدنى ر٢٥ $= \frac{٢٥}{١٠٠}$ (٨ – ١٠.) $= \frac{٢٢٥}{}$ = ٢،٢٥

– رتبة الربيع الأعلى ر ٧٥ $= \frac{٧٥}{١٠٠}$ (٨ – ١) = ٦،٧٥

القيم المناظرة للربيع الأدنى م٢٥ هي الربيع الثاني والثالث من ٨، ١٠

قيمة الربيع الأدنى $= \frac{٨ + ١٠}{٢}$ = ٩

القيم المناظرة للربيع الأعلى ر٧٥ هي الربيع السادس والسابع ١٣، ١٥

قيمة الربيع الأعلى = $\frac{١٣+١٥}{٢}$

٥

وعليه فإن المدى الربيعي=١٤ – ٩ =٥ أما نصف المدى الربيعي ——

٢

= ٢،٥

ب/- المدى في حالة البيانات المبوبــة:

أ)- نصف المدى الربيعي/ الإنحراف الربيعــــــي:

ويحسب وفق الصيغة الرياضية التالية:

$$\frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{٢} = \frac{ر_{٧٥} - ر_{٢٥}}{٢}$$

مثـــــال:

أوجد نصف المدى الربيعي للجدول التالي الذي يبين توزيع مجموعة من الأشخاص حسب مدة ترسيمهم في العمل.

الفئـــــات	ك	س	ك الصاعد
٠ — ٥	٣	٢،٥	٣
٥ — ١٠	٨	٧،٥	١١
١٠ — ١٥	٢٠	١٢،٥	٣١
١٥ — ٢٠	١٣	١٧،٥	٤٤
٢٠ — ٢٥	٦	٢٢،٥	٥٠
٢٥ — ٣٠	١٠	٢٧،٥	٦٠
المجمـــــــوع	٦٠		

_المدى المطلق = الحد الأعلى للفئة العليا – الحد الأدنى للفئة الدنيا

= ٣٠ — ٠ = ٣٠

_ المدى الفعلي المطلق = ٣٠،٥ — (- ٠،٥) = ٣١

_ إيجاد نصف المدى الربيعي.

أولا لابد من إيجاد الربيع الأعلى ر٧٥

رتبة الربيع الأعلى = (٧٥ / ١٠٠) × (٦٠) = ٤٥ إذن فموقعه في الفئة ٢٠– ٢٥

إذن: ر٧٥ = ٢٠ + ((٤٥ — ٤٤) / (٦٠ _ ٤٤)) × ١٠ = ٢٠،٣١ ______

ثانيا إيجاد الربيع الأدنى ر٢٥:

١- رتبة الربيع الأدنى = (٢٥ / ١٠٠) (٦٠) = ١٥ إذن فموقعه في الفئة ١٠– ١٥

إذن: ر٢٥ = ١٠ + ((١٥ — ١١) / (٤٤ _ ١١)) × ٥ = ١٠ + (٤ / ٣٣) × ٥ = ١٠،٦١

بتطبيق النتائج على القانون السابق ذكره نجد أن نصف المدى الربيعي = $\frac{٢٠،٣١ - ١٠،٦١}{٢}$

= ٤،٨٥

2 – الإنحراف المتوســــط:

الإنحراف المتوسط هو مقياس من مقاييس التشتيت يقيس درجة الإنحراف عن المتوسط الحسابــــــي.

أ_ في حالة البيانات غير المبوبـــــة:

الإنحراف المتوسط = $\frac{\Sigma|ح_ر|}{ن} = \frac{\Sigma|س_ر - \bar{س}|}{ن}$

مثـــــال: أوجد الإنحراف المتوسط للقيم التالية التي تبين مدة استعمال وسائل منع الحمل لخمسة نساء في سن الإنجاب ٥، ١٠، ٦، ٣، ١٥

الحل: حساب $\bar{س}$ = $\frac{٥ + ١٠ + ٦ + ٣ + ١٥}{٥}$ = ٧،٨ سنة

$ح_١ = |س_١ - \bar{س}|$ = ٥ – ٧،٨ = ٢،٨

$ح_٢ = |س_٢ - \bar{س}|$ = ١٠ – ٧،٨ = ٢،٢

$ح_٣ = |س_٣ - \bar{س}|$ = ٦ – ٧،٨ = ١،٨

$ح_٤ = |س_٤ - \bar{س}|$ = ٣ – ٧،٨ = ٤،٨

$ح_٥ = |س_٥ - \bar{س}|$ = ١٥ – ٧،٨ = ٧،٢

= $\Sigma|\Sigma س - \bar{س}|$ = ١٨،٨

الإنحراف المتوسط هنا أم = $\frac{|\Sigma ح|}{ن}$ = $\frac{١٨،٨}{٥}$ = ٣،٧٦

ب)- في حالة البيانات المبوبـــــة:

لإيجاد الإنحراف المتوسط في حالة البيانات المبوبة نتبع القانون التالية:

$$\text{الإنحراف المتوسط. أ. م} = \frac{\Sigma\ |\text{ح}| \times \text{ك}}{\Sigma\ \text{ك}}$$

مثـــــال: أوجد الإنحراف المتوسط للبيانات التالية:

الفئات	ك	س	ك. س	احا	احا ك
٠ – ٥	٣	٢،٥	٧،٥	١٣,٤٢	٤٠,٢٦
٥ – ١٠	٨	٧،٥	٦٠	٨,٤٢	٦٧,٣٧
١٠ – ١٥	٢٠	١٢،٥	٢٥٠	٣,٤٢	٦٨,٤
١٥ – ٢٠	١٣	١٧،٥	٢٢٧،٥	١,٥٨	٢٠,٥٤
٢٠ – ٢٥	٦	٢٢،٥	١٣٥	٦,٥٨	٣٩,٤٨
٢٥ – ٣٠	١٠	٢٧،٥	٢٧٥	١١,٥٨	١١٥,٥
المجمـــوع	**٦٠**		**٩٥٥**		**٣٥١,٨٤**

$$\bar{\text{س}} = \frac{\Sigma\ \text{ك س}}{\Sigma\ \text{ك}} = \frac{٩٥٥}{٦٠} = ١٥,٩٢$$

$$\textbf{أم} = \frac{}{} = \frac{٣٥١,٨٤}{٦٠} = \textbf{٥,٨٦}$$

٣/- التباين والإنحراف المعيـــــاري:

التباين هو مجموع مربعات الإنحرافات عن وسطها الحسابي مقسوما على حجم العينة وإنما الإنحراف المعياري فهو الجذر التربيعي للتباين.

أ)- في حالة البيانات غير المبوبـــــة:

التباين : $\text{ع}^{٢} = \frac{\Sigma\ \text{ح}^{٢}}{\text{ن}}$ أما الإنحراف المعياري: $\text{ع} = \sqrt{\frac{\Sigma\ \text{ح}^{٢}}{\text{ن}}}$

مثـــــال: أوجد التباين والإنحراف المعياري للتوزيعين التاليين الذان يبين تنظيم نسل مجموعة من النساء حسب الطبقة الإجتماعية.

الطبقـــــة	تنظيم النســـل	لا تنظيم
فقيــــــرة	٠	٢٤
متوسطـــــة	٢٥	٢٤
غنية/ مرتفعة	٥٠	٢٦
المجمــــوع	٧٥	٧٥

أولا حساب التباين والإنحراف المعياري للنساء المنظمات للنسل حسب الطبقة التي تنتمين إليها.

الطبقـــــة	ك	ح	ح٢
طبقة فقيرة	٠	– ٢٥	٦٢٥
طبقة متوسطة	٢٥	٠	٠
طبقةغنية/ مرتفعة	٥٠	+ ٢٥	٦٢٥
المجمــــوع	٧٥	٠	١٢٥٠

$$\bar{\text{س}} = \frac{\sum \text{س}}{\text{ن}} = \frac{٧٥}{٣} = ٢٥$$

حساب التباين $\text{ع}^{٢} = \frac{\sum \text{ح}^{٢}}{\text{ن}} = \frac{١٢٥٠}{٣} = ٤١٦،٦٧$

حساب الإنحراف المعياري $\text{ع} = \sqrt{\frac{\sum \text{ح}^{٢}}{\text{ن}}} = \sqrt{\frac{١٢٥٠}{٣}} = \sqrt{٤١٦،٦٧} = ٢٠،٤١$

ثانيــــا: حساب التباين والإنحراف المعياري للنساء الغير منظمات للنسل حسب الطبقة الإجتماعية المنتميات إليها.

الطبقـــــة	ك	ح	ح٢
طبقة فقيرة	٢٤	– ١	١
طبقة متوسطة	٢٥	٠	٠
طبقةغنية/ مرتفعة	٢٦	+ ١	١
المجمــــوع	٧٥		٢

$$\bar{س} = \frac{٧٥}{٣} = ٢٥$$

حساب التباين: $ع^{٢} = \frac{\Sigma ح^{٢}}{ن} = \frac{٢}{٣} = ٠،٦٧$

حساب الإنحراف المعياري $ع = \sqrt{\frac{\Sigma ح^{2}}{ن}} = \sqrt{\frac{٢}{٣}} = \sqrt{٠،٦٧} = ٠،٨٢$

ومن هنا نلاحظ أن المجموعة الثانية وهي النساء غير المنظمات للنسل أكثر انسجاما من المجموعة الأولى ومن النساء المنظمات للنسل وذلك لأن انحرافها كبير جدا عن الوسط وتقدر بـــ ٢٠،٤١ أما المجموعة الثانية فتقريبا انحرافها عن الوسط منعدما إذ قدرت بـــ ٠،٨٢

ب)- في حالة البيانات المبوبـــــة:

التباين $ع^{٢} = \frac{\Sigma ح^{٢} \times ك}{\Sigma ك}$

أما الإنحراف المعياري $ع = \sqrt{\frac{\Sigma ح^{٢} \times ك}{\Sigma ك}}$

مثـــــال: أوجد الإنحراف المعياري والتباين للجدول التالي:

الفئـــــــات	ك
١٥ – ٢٥	٣٤٣
٢٥ – ٣٥	٣٨٠
٣٥ – ٤٥	٥٣٠
٤٥ – ٥٥	٥٤٧
٥٥ – ٦٥	٤٦٤
٦٥ – ٧٥	٢٧١
٧٥ – ٨٥	١١٤
المجمـــــــــوع	**٢٦٤٩**

الحـــــل:

لفئـــات	ك	س	ك. س	ح	ح٢	ح٢ × ك
١٥ – ٢٥	٣٤٣	٢٠	٦٨٦٠	– ٢٦،٣	٦٩١،٦٩	٢٣٧٢٤٩،٦٧
٢٥ – ٣٥	٣٨٠	٣٠	١١٤٠٠	– ١٦،٣	٢٦٥،٦٩	١٠٠٩٦٢،٢
٣٥ – ٤٥	٥٣٠	٤٠	٢١٢٠٠	– ٦،٣	٣٩،٦٩	٢١٠٣٥،٧
٤٥ – ٥٥	٥٤٧	٥٠	٢٧٣٥٠	٣،٧	١٩،٦٩	٧٤٨٨،٤٣
٥٥ – ٦٥	٤٦٤	٦٠	٢٧٨٤٠	١٣،٧	١٨٧،٦٩	٨٧٠٨٨،١٦
٦٥ – ٧٥	٢٧١	٧٠	١٨٩٧٠	٢٣،٧	٥٦١،٦٩	١٥٢٢١٧،٩٩
٧٥ – ٨٥	١١٤	٨٠	٩١٢٠	٣٣،٧	١١٣٥،٦٩	١٢٩٤٦٨،٦٦
المجمــوع	**٢٦٤٩**		**١٢٢٧٤٠**			**٧٣٥٥١٠،٨١**

المتوسط الحسابي $\bar{س} = \frac{\sum س ك}{\sum ك} = \frac{١٢٢٧٤٠}{٢٦٤٩} = ٤٦،٣$

حساب التباين $ع^{٢} = \frac{\sum ح^{٢}. ك}{\sum ك} = \frac{٧٣٥٥١٠،٨١}{٢٦٤٩} = ٢٧٧،٦٦$

حساب الإنحراف المعياري $ع = \sqrt{\frac{\sum ح^{٢}. ك}{\sum ك}} = \sqrt{\frac{٧٣٥٥١٠،٨١}{٢٦٤٩}} = ١٦،٦٦$

ويعتبر الإنحراف المعياري من أهم مقاييس التشتت المستخدمة في العلوم الإجتماعية كما أن الإنحراف المعياري يستعمل في عدة مؤثرات أخرى كالإرتباط وفي تحديد أشكال التوزيعات الإحصائية والإحتمالات... إلخ.

٤ _ معامل الإختـــلاف:

عبارة عن النسبة بين الإنحراف المعياري والمتوسط الحسابي ويعتبر من مقاييس التشتت النسبية، ويستعمل خاصة في المقارنة بين توزيعات إحصائية غير متجانسة ويحسب وفق الصيغة التالية: $\frac{ع}{\bar{س}}$ وفي المثال السابق أعلاه نجده يقدر ب $\frac{١٦،٦٦}{٤٦،٣} = ٠،٣٦$

٥_ مقاييس الإلتواء والتفرطـــــح:

هذا النوع من المقاييس يبين لنا ويحدد نوع شكل المنحنيات الممثلة للتوزيعات من حيث درجة تدبب قمة المنحنى أو إنبساطها ومن حيث تماثل الخط المنحنى حول محور القيمة المتوسطة للتوزيع وهذا الوصف يتم إستخدام إما معامل الإلتواء الذي يصف التوزيعات التكرارية من حيث التماثل ومقاييس التفرطح والذي يقوم بوصف التوزيعات التكرارية من حيث تدبب قمة المنحنى الذي يمثلها أو تفرطحها وإنبساطها.

٥_١_معامل الإلتـــــــواء:

ويحسب وفق القانوق التالـــي:

$$\text{معامل الإلتواء} = \frac{\textbf{الوسط الحسابي – المنوال}}{\text{الإنحراف المعياري}} \quad \text{أي:}\quad ى = \frac{\bar{س} - مو}{ع}$$

إذا كان سَ > مو فإن الإلتواء يكون موجبا والعكس صحيح.

مثال: حساب معام الإلتواء للجدول السابق وجدنا: سَ = ٤٦،٣ ع = ١٦،٦٦
مو = ٥٠،٣٣

$$ى = \frac{٤٦،٣ - ٥٠،٣٣}{١٦،٦٦} = ٠،٢٤$$ أي أن الإلتواء ضئيل جدا وسالب

$$\text{أما معامل التفرطح ت} = \frac{\text{العزم الرابع}}{\text{مربع التباين}}$$

٥_٢_العلاقة بين المتوسطات الإحصائية ومقاس التشتـــــت:

– نصف المدى الربيعي = ٠،٨٤٥٠ من الإنحراف المتوسط

- نصف المدى الربيعي = ٠،٦٧٤٥ من الإنحراف المعياري
- الإنحراف المتوسط = ٠،٧٩٧٩ من الإنحراف المعياري
- الإنحراف المعياري = ١،٤٨٣٠ من نصف المدى الربيعي
- الإنحراف المعياري = ١،٢٥٣٠ من الإنحراف المتوسط
- الإنحراف المتوسط = ١،١٨٢٠ من نصف المدى الربيعي

_ معامل الإختلاف الربيعي = ر٣ – ر١ × ٢و

_ الإنحراف المتوسط = $\frac{٢}{٣}$ الإنحراف المعياري

_ الإنحراف الربيعي = $\frac{٤}{٥}$ الإنحراف المعياري

_ الإنحراف المتوسط عن الوسط الحسابي = $\frac{\sum (س - \bar{س})}{ن}$ بيانات غير مبوبة

= $\frac{\sum (س - \bar{س})\, ك}{\sum ك}$ بيانات مبوبة

_ الإنحراف المتوسط عن الوسيط = $\frac{\sum (س - و)}{ن}$ بيانات غير مبوبة

= $\frac{\sum (س - و)\, ك}{ك}$ بيانات مبوبة

_ الإنحراف المتوسط عن المنوال = $\frac{\sum (س - مو)}{ن}$ بيانات غير مبوبة

= $\frac{\sum (س - مو)\, ك}{\sum ك}$ بيانات مبوبة

_ معامل الإختلاف المتوسط = $\frac{أم - \bar{س}}{\bar{س}}$

$\frac{أم - مو}{مو}$

$\frac{ع}{س}$

_ معامل الإختلاف = × ١٠٠ إذا كان التوزيع التكراري مقفولا/ مغلقا

_ معامل الإختلاف = × ١٠٠ إذا كان التوزيع التكراري مفتوحا

تمارينات

التمرين الأول:

أوجد الانحراف المعياري للقيم التالية ٥، ٩، ١٢، ١٤.

القيم	ح	ح٢
٥	٥- ١٠= - ٥	٢٥
٩	٩- ١٠= - ١	١
١٢	١٢- ١٠= +٢	٤
١٤	١٤- ١٠= +٤	١٦
المجموع	٠	**٤٦**

$$\overline{س} = \frac{مج(س)}{ن} = \frac{5+9+12+14}{4} = ١٠$$

$$ع = \sqrt{\frac{مج(ح^2)}{ن}} = \sqrt{\frac{46}{4}} = ٣,٣٦$$

التمرين الثاني:

البيانات التالية تمثل درجات أو علامات طالب لمادتي الإحصاء و الرياضيات، و المطلوب حساب الانحراف المعياري للقيم المتغيرين و أيهما أكثر تشتتا؟ حساب التباين و معامل الاختلاف؟

الإحصاء / الرياضيات	٤٠ – ٥٠	٥٠ – ٦٠	٦٠ – ٧٠	٧٠ – ٨٠	٨٠ – ٩٠	٩٠ – ١٠٠	المجموع
٤٠ – ٥٠	٣	٥	٤	–	–	–	١٢
٥٠ – ٦٠	٣	٦	٦	٢	–	–	١٧
٦٠ – ٧٠	١	٤ –	٩	٥	٢	–	٢١
٧٠ – ٨٠	–	–	٥	١٠	٨	١	٢٤
٨٠ – ٩٠	–	–	١	٤	٦	٥	١٦
٩٠ – ١٠٠	–	–	–	٢	٤	٤	١٠
المجموع	**٧**	**١٥**	**٢٥**	**٢٣**	**٢٠**	**١٠**	**١٠٠**

<u>الحل:</u>

أولا حساب التباين و الانحراف المعياري و كذا معامل الاختلاف لدرجات الرياضيات.

درجات الرياضيات	ك	س	ك×س	ح	ح²	ك×ح²
٤٠ – ٥٠	١٢	٤٥	٥٤٠	– ٢٤,٥	٦٠٠,٢٥	٧٢٠٣
٥٠ – ٦٠	١٧	٥٥	٩٣٥	– ١٤,٥	٢١٠,٢٥	٣٥٧٤,٢٥
٦٠ – ٧٠	٢١	٦٥	١٣٦٥	– ٤,٥	٢٠,٢٥	٤٢٤,٢٥
٧٠ – ٨٠	٢٤	٧٥	١٨٠٠	٥,٥	٣٠,٢٥	٧٢٦
٨٠ – ٩٠	١٦	٨٥	١٣٦٠	١٥,٥	٢٤٠,٢٥	٣٨٤٤
٩٠ – ١٠٠	١٠	٩٥	٩٥٠	٢٥,٥	٦٥٠,٢٥	٦٥٠٢,٤
المجموع	**١٠٠**		**٦٩٥٠**			**٢٢٢٧٥**

المتوسط الحسابي لدرجات الرياضيات $\overline{\text{س}} = \frac{\text{مج}(\text{ك} \times \text{س})}{\text{مج}(\text{س})} = \frac{6950}{100} = ٦٩,٥$

التباين في نقاط / درجات الرياضيات $= \frac{\text{مج}(\text{ك} \times \text{ح}^2)}{\text{مج}(\text{ك})} = \frac{22275}{100} = ٢٢٢,٧٥$

الانحراف المعياري لدرجات الرياضيات $ع = \sqrt{\frac{مج(ك \times ح^2)}{مج(ك)}} = \sqrt{\frac{22275}{100}} = ١٤,٩٢$

معامل الاختلاف لدرجات الرياضيات $= \frac{(ع)(س)}{س} \times ١٠٠ = \frac{14.92}{69.5} \times ١٠٠ =$
٢١,٤٦٪

ثانيا: حساب التباين، الانحراف المعياري، و كذا معادل الاختلاف لدرجات الإحصاء.

درجات الرياضيات	ك	س	ك×س	ح	ح٢	ك×ح٢
٤٠- ٥٠	١٢	٤٥	٣١٥	- ٢٦,٤	٦٩٦,٩٦	٤٨٧٨,٧٢
٥٠- ٦٠	١٧	٥٥	٨٢٥	- ١٦,٤	٢٦٨,٩٦	٤٠٣٤,٤
٦٠- ٧٠	٢١	٦٥	١٦٢٥	- ٦,٤	٤٠,٩٦	١٠٢,٤
٧٠- ٨٠	٢٤	٧٥	١٧٢٥	٣,٦	١٢,٩٦	٢٩٨,٠٨
٨٠- ٩٠	١٦	٨٥	١٧٠٠	١٣,٦	١٨٤,٩٦	٣٦٩٩,٢
٩٠- ١٠٠	١٠	٩٥	٩٥٠	٢٣,٦	٥٥٦,٩٦	٥٥٦٩,٦
المجموع	١٠٠		٧١٤٠			١٩٥٠٤

المتوسط الحسابي لدرجات الإحصاء $= \overline{س} = \frac{مج(ك \times س)}{مج(س)} = \frac{7140}{100} = ٧١,٤$

التباين في نقاط / درجات الإحصاء $= \frac{مج(ك \times ح^2)}{مج(ك)} = \frac{19504}{100} = ١٩٥,٠٤$

الانحراف المعياري لدرجات الإحصاء $ع = \sqrt{\frac{مج(ك \times ح^2)}{مج(ك)}} = \sqrt{\frac{22275}{100}} = ١٣,٩٦$

معامل الاختلاف لدرجات الإحصاء = $\frac{ع_ع}{\bar{س}_ع}$ = $\frac{13.63}{71.4}$ × ١٠٠ = ١٩,٠٩٪

و من هنا نستنتج أن درجات الرياضيات هي أكثر تشتتا و كذا اختلافا و تغيرا من درجة الإحصاء.

الفصـــــــل الســـــــادس

تحليل الارتباط والانحدار

أولا: قياس الارتباط بين متغيرين – مستقل وتابع – كميين معا

١)– معامل ارتباط كارل بيرسون

٢)– معامل الائتلاف

ثانيا: قياس الارتباط بين متغيرين –مستقل وتابع كيفيين معا

أو أحدهما كيفي والثاني كمي

١)– معامل ارتباط سيبرمان

٢)– معامل ارتباط الاقتران

٣)– معامل ارتباط فآي

٤)– معامل التوافـــق

٥)– معامل ارتباط لامد

ثالثــا: أشكال الانتشار

رابعــا: الرسم البياني لخطوط الانحدار

الارتباط والانحدار:

كما تتبعنا في المقاس السابقة أنها تقيس متغير واحد فقط، كسن، المدة ... إلخ أما إذا كانت البيانات تتعلق بسلوك ظاهرتين أو متغيرين أو أكثر كالعدد الحقيقي للأطفال والعدد المتالي للأطفال في الأسرة أو عدد الأطفال الأحياء بالمستوى التعليمي للأم ... إلخ لمعرفة ما إذا كانت هناك علاقة بينهما وتحديد مقدارها ونوعها في حالة وجودها أي هل هي طردية بمعنى أن تغير الظاهرتين في اتجاه واحد وبالتالي إذا زادت قيمة إحدى الظاهرتين تميل الثانية أي قيم الظاهرة الثانية إلى الزيادة وإذا نقصت قيمة الظاهرة الأولى تميل قيمة الظاهرة الأخرى إلى النقصان، أو أن هذه العلاقة عكسية بمعنى أن تغير الظاهرتين في اتجاه مضاد وبالتالي فإذا زادت قيمة إحدى الظاهرتين تميل قيمة الظاهرة الأخرى إلى النقصان والعكس بالعكس.

وقد تنشأ مسألة تحليل الارتباط في كل مرة يتساءل الباحث فيها عنها إذا كانت هناك علاقة بين القيم التي يأخذها أحد المتغيرين أي المتغير المستقل والقيم التي يأخذها المتغير الثاني وهو المتغير التابع.

ومن الطبيعي أن الباحث في غالب الأحيان لا يكتفي بمعرفة ما إذا كانت هناك علاقة بين هذه المتغيرات بل أنه يرغب في معظم الأحيان تحديد مقدار هذه العلاقة ونوعها، في حالة وجودها. ونشير هنا أن وجود العلاقة أو الارتباط بين ظاهرتين أو متغيرين لا يعني بالضرورة وجود علاقة سببية أي أن إحدى الظاهرتين نتيجة الظاهرة الأخرى بل قد تكون نتيجة لعوامل خارجية خارجة عن نطاق المتغيرين موضوع الدراسة.

إن هذا الوصف الدقيق لنوعية العلاقة بين المتغيرات يدخل ضمن مجال الإحصاء الاجتماعي أي الوقوف على طبيعة العلاقة بين أكثر من متغير واحد وعن طريق هذا التفسير العددي يتسنى للباحث أن يصدر تنبؤات على الدراسة التي هو بصدد القيام بها

ويطلق على المعامل الذي يصف نوع العلاقة وقيمتها من المتغيرين معاملا الارتباط وقيمته تنحصر ما بين –١، +١.

وتعتبر دراسة الارتباط الإحصائي بين المتغيرات بالغ الأهمية في البحوث الاجتماعية لأنه يعطينا معيارا نستطيع من خلاله تقدير قيمة الفرضيات التي وضعناها أثناء الدراسة بحيث يمكننا من خلال دراسة الارتباط أن نثبت هذه الفرضيات أو ننفيها نفيا باتا.

وعند دراستنا للعلاقات بين المتغيرات المختلفة نجد أن بعضها يرتبط بأكثر من متغير واحد، مثل جداول الارتباط المركبة السالفة الذكر، في هذه الحالة أما أن نقوم بدراسة علاقة المتغير بجميع المتغيرات المرتبطة به دفعة واحدة وهذا ما يسمى بالارتباط الكلي، أو ندرس علاقة المتغير المستقل بالتابع فقط وهذا ما يسمى بالارتباط الجزئي.

وبصفة عامة فإن دراجات الارتباط أو مستوياته يمكن تحديدها على ضوء قيم معامل الارتباط التالية.

قيمة معامل الارتباط	مدى الحكم عليـــــــــه
+ ١	علاقة طردية كاملة بين المتغيرين
من + ٠،٧ + ٠،٩٩	علاقة طردية قوية بين المتغيرين
من + ٠،٤ + ٦	علاقة طردية متوسطة بين المتغيرين
من + ٠،٢ + ٠،٤	علاقة طردية ضعيفة قليلا بين المتغيرين
أقل من + ٠،٢	علاقة طردية ضعيفة للغاية بين المتغيرين
٠	لا توجد علاقة بين المتغيرين
أكثر من – ٠،٢	علاقة عكسية ضعيفة للغاية بين المتغيرين
من – ٠،٢ – ٠،٤	علاقة عكسية ضعيفة نوعا ما بين المتغيرين
من – ٠،٤ – ٠،٦	علاقة عكسية متوسطة بين المتغيرين
من – ٠،٧ – ٠،٩٩	علاقة عكسية قوية بين المتغيرين
– ١	علاقة عكسية كاملة بين المتغيرين

١ـ قياس الارتباط بين متغيرين مستقل وتابع كميين معا:

١ـ١- معامل ارتباط كارل بيرسون:

نفرض أنه لدينا متغيرين المتغير الأول مستقل س والمتغير الثاني تابع ع ولدينا ن من الأزواج القيم (س، ع) حيث أن قيم س = س $_{١}$، س $_{٢}$، س $_{٣}$ ... س $_{ن}$ قيم ع هي: ع $_{١}$، ع$_{٢}$، ع$_{٣}$ ... ع $_{ن}$ وهما متغيرين كميين معا.

فالعلاقة بين المتغيرين س، ع ستتوقف على قيم كل من المتغيرين ولذلك فلابد أن نستخدم هذه القيم لقياس علاقة الارتباط بينهما.

والارتباط يقيس العلاقة بين التغير في قيم س والتغير في قيم ع ولذلك فلابد لنا من إيجاد التغير في قيم كل من س، ع وأفضل طريقة لقياس هذا التغير هي إيجاد الفرق بين قيم كل متغير ووسطه في هذه الحالة.

ومعامل ارتباط كارل بيرسون يسد نقصا كبيرا في حالة استعمالنا لمعامل سبرمان عند وجود متغيرين كميين قابلين للقياس لكن في جداول غير مبوبة.

حيث يهتم هذا الأخير بالرتب وليس بالقيم نفسها، وحساب الارتباط بالرتب تقل دقته من حسابه على أساس القيم فزيادة القيم أو نقصها لا تتغير من قيمة المعامل على أساس الرتب مادامت هذه الزيادة أو النقصان لا تغير وضع القيمة بالنسبة لمجموعة بينما يتأثر معامل الارتباط بيرسون بأي تغير في أي قيم من قيم المتغيرين

أ- معامل كارل بيرسون في حالة البيانات غير المبوبـــــة:

ويحسب وفق القانون التالي:

$$ر = \frac{١}{ن} \left[\frac{(س - \bar{س})(ع - \bar{ع})}{ع_{س}\, ع_{ع}} \right]$$

أحسب معامل كارل برسون للبيانات التاليـــــــــة:

	س	ع
	١	٧
	٣	١٩
	٤	٢٥
	٧	٤٣
	١٠	٦١
المجمـــــــوع	**٢٥**	**١١٥**

س	ع	س – سَ	ع – عَ	(س – سَ)^{٢}	(ع – عَ)^{٢}	(س – سَ)(ع – عَ)
١	٧	– ٤	– ٢٤	١٦	٥٧٦	٩٦
٣	١٩	– ٢	– ١٢	٤	١٤٤	٢٤
٤	٢٥	– ١	– ٦	١	٣٦	٦
٧	٤٣	٢	١٢	٤	١٤٤	٢٤
١٠	٦١	٥	٣٠	٢٥	٩٠٠	١٥٠
٢٥	**١٥٥**			**٥٠**	**١٨٠٠**	**٣٠٠**

$$\text{سَ} = \frac{\Sigma\,\text{س}}{\text{ن}} = \frac{٢٥}{٥} = ٥$$

$$\text{عَ} = \frac{\Sigma\,\text{ع}}{\text{ن}} = \frac{١٥٥}{٥} = \mathbf{٣١}$$

$$\text{ع}_{\text{س}} = \sqrt{\frac{\text{ح}^{٢}_{\text{س}}}{\text{ن}}} = \sqrt{\frac{٥٠}{٥}} = ٣،١٦$$

$$\text{ع}_{\text{ع}} = \sqrt{\frac{\text{ح}^{٢}_{\text{ع}}}{\text{ن}}} = \sqrt{\frac{١٨٠٠}{٥}} = ١٩،٠٣$$

بتعويض هذه المحاصيل في قانون ر نجد أن:

$$\text{ر} = \frac{1}{\text{ن}} \cdot \frac{\Sigma|(\text{س} - \text{سَ})(\text{ع} - \text{عَ})|}{\text{ع}_{\text{س}}\,\text{ع}_{\text{ع}}}$$

$$\text{ر} = \frac{1}{٥} \cdot \frac{٣٠٠}{(١٩،٠٣)(٣،١٦)} = ٠،٩٩$$

هذا يعني أن الارتباط قوي جدا وموجب/ طردي قريب من أن يكون كاملا بين المتغيريـــــن.

ب)- في حالة البيانات المبوبـــــة:

إذا كان لدينا مجموعتين كبيرتين من قيم ظاهرتين أو متغيرين كميين معا يتجاوز تكرارها الكلي ٥٠ وأردنا دراسة، الارتباط بين قيم المتغيرين فلابد من تبويب هذه البيانات في فئات ومن ثم أعداد جداول تكرارية مزدوجة من النوع المعروف ويحسب حينها معامل الارتباط بيرسون وفق العلاقة التالية.

$$\text{ر} = \frac{\sum \text{ك} (\text{س} - \text{و})(\text{ع} - \text{وَ}) - \text{ن}\, \text{ح}_{\text{س}}\, \text{ح}_{\text{ع}}}{\text{ن}\, \text{ع}_{\text{س}}\, \text{ع}_{\text{ع}}}$$

مثـــــال: لنفرض أننا نرغب في إجراء دراسة تحليلية للارتباط بين عمر العامل وأجره الأسبوعي معطى بالا وروا ولإجراء هذه الدراسة أخذنا عينة عشوائية تتكون من ٩٠ عاملا وعاملة واعتبرنا العمر هو المتغير المستقل والأجر هو المتغير التابع وبعد أن حصلنا على البيانات التالية وهي قيم س، ع المبوبة في الجدول التالي والذي سوف نقوم بحساب معامل كارل برسون لـــه.

ع / س	٥٠ – ٦٠	٦٠ – ٧٠	٧٠ – ٨٠	٨٠ – ٩٠	٩٠ – ١٠٠	المجموع
٢٠ – ٣٠	٣	٣	١	–	–	٧
٣٠ – ٤٠	–	٤	٦	٥	–	١٥
٤٠ – ٥٠	–	١	١٥	٥	٩	٣٠
٥٠ – ٦٠	–	٤	٦	٨	٢	٢٠
٦٠ – ٧٠	٢	٤	١٠	٢	–	١٨
المجموع	٥	١٦	٣٨	٢٠	١١	٩٠

الحـــــل:

١)- حساب $ح_س$ و $ع_س$:

فئات العمر	$ك_س$	$س_س$	$ح_س$	$ح_س \times ك$	$ح^٢_س \times ك$
٢٠ – ٣٠	٧	٢٥	– ٢٠	– ١٤٠	٢٨٠٠
٣٠ – ٤٠	١٥	٣٥	– ١٠	– ١٥٠	١٥٠٠
٤٠ – ٥٠	٣٠	٤٥	٠	٠	٠
٥٠ – ٦٠	٢٠	٥٥	+ ١٠	٢٠٠	٢٠٠٠
٦٠ – ٧٠	١٨	٦٥	+ ٢٠	٣٦٠	٧٢٠٠
المجمـــوع	٩٠			٢٧٠	١٣٥٠٠

من الجدول نجد أن $ح_س = \frac{ح_س \times ك_س}{ك_س} = \frac{٢٧٠}{٩٠} = ٣$

$$ع_س = \sqrt{\frac{\Sigma ح^٢_س \times ك_س}{\Sigma ك_س} - (ح_س)^٢} \; ع_س = \sqrt{\frac{١٣٥٠٠}{٩٠} - (٣)^٢} = ١١،٨٧$$

٢)- حساب $ح_ع$، و $ع_ع$:

فئات الأجر	$ك_ع$	$س_ع$	$ح_ع$	$ح_ع \times ك$	$ح^٢_ع \times ك$
٥٠ – ٦٠	٥	٥٥	– ٢٠	– ١٠٠	٢٠٠٠
٦٠ – ٧٠	١٦	٦٥	– ١٠	– ١٦٠	١٦٠٠
٧٠ – ٨٠	٣٨	٧٥	٠	٠	٠
٨٠ – ٩٠	٢٠	٨٥	+ ١٠	٢٠٠	٤٠٠٠
٩٠ – ١٠٠	١١	٩٥	+ ٢٠	٢٢٠	٤٤٠٠
المجمـــــوع	٩٠			١٦٠	١٢٠٠٠

من الجدول نجد كذلك أن $ح_ع = \frac{ح_ع \times ك_ع}{ك_ع} = \frac{١٦٠}{٩٠} = ١،٧٨$

$$ع_ع = \sqrt{\frac{\Sigma ح^٢_ع \times ك_ع}{ك_ع \Sigma} - (ح_ع)^٢} = \sqrt{\frac{١٢٠٠٠}{٩٠} - (١،٧٨)^٢} = ١١،٤١$$

ثم بعد إيجاد قيم ح $_{س}$، ح $_{ع}$، ع $_{س}$، ع $_{ع}$ يجري تنظيم جدول ارتباط على النحو التالي لإيجاد Σ ك (س – و) (ع – وَ).

	ح ع	– ٢٠	– ١٠	٠	١٠	٢٠		
ح س	ع س	–٥٠ ٦٠	–٦٠ ٧٠	–٧٠ ٨٠	–٨٠ ٩٠	–٩٠ ١٠٠	ك س	مجموع حواصل الضرب
– ٢٠	٢٠ – ٣٠	٣ ١٢٠٠	٣ ٦٠٠	0 1	– –	–	٧	١٨٠٠
– ١٠	٣٠ – ٤٠	–	٤ ٤٠٠	٦ ٠	٥ – ٥٠٠	–	١٥	– ١٠٠
٠	٤٠ – ٥٠	–	١ ٠	١٥ ٠	٥ ٠	٩ ٠	٣٠	٠
١٠	٥٠ – ٦٠	–	٤ – ٤٠٠	٦ ٠	٨ ٨٠٠	٢ ٤٠٠	٢٠	٨٠٠
٢٠	٦٠ – ٧٠	٢ – ٨٠٠	٤ – ٨٠٠	١٠ ٠	٢ ٤٠٠	–	١٨	– ١٢٠٠
		٥	١٦	٣٨ ٠	٢٠ –	١١	٩٠	
	مجموع حواصل الضرب	٤٠٠	– ٢٠٠		٧٠٠	٤٠٠		١٣٠٠

١٢٠٠ هي حاصل ضرب تكرار الفئة × ح $_{س}$ × ح $_{ع}$ أي ٣ × (– ٢٠) (– ٢٠) = ١٢٠٠

والآن بمعلومة المحاصيل يمكن تطبيق القانون الخاص بمعامل كارل بيرســــون

ر = (Σ ك (س - و)(ع - وَ) – ن ح $_{س}$ ح $_{ع}$) / (ن ع $_{س}$ ع $_{ع}$) = (٩٠ (٣)(١،٧٨)_١٣٠٠) / (٩٠ (١١،٨٧)(١١،٤١)) = ٠،٠٦٧

هذا يعني وجود علاقة طردية ضعيفة بين المتغيرين س، ع أي أن سن العامل ليس هو المحدد لأجره، فقد تكون هناك أسباب أخرى أو عوامل أخرى كالأقدمية في المهنة والشهادة لمتحصل عليها ... إلخ.

٢- معامل الائتلاف:

يوجد طريقة أخرى لقياس العلاقة بين ظاهرتين كميتين متغيرين كميين وذلك باستخدام معادلة وضعها بول هذه المعادلة تعطي لنا مقدار العلاقة بين المتغيرين وهذه المعادلة تسمى بمعامل الائتلاف ويرمز له بالرمز ل.

وتحسب وفق الصيغة التاليـــــــة:

$$\text{ل} = \frac{\sqrt{\text{ج ب}} - \sqrt{\text{أ د}}}{\sqrt{\text{أ د}} + \sqrt{\text{ج ب}}}$$

حيـــــث أن:

أ = عدد الحالات التي يكون فيها كل من المتغيرين المستقل والتابع فوق المتوسط الحسابي.

د = عدد الحالات التي يكون فيها كل من المتغيرين س ع تحت المتوسط الحسابي.

ج = عدد الحالات التي يكون فيها كل من قيم المتغير المستقل فوق المتوسط الحسابي وقيم عَ تحته.

ب = عدد الحالات التي يكون فيها كل قيم المتغير المستقل تحت المتوسط الحسابي وقيم المتغير التابع ع فوقه.

ولحساب هذا المعامل إذن نحتاج إلى حساب المتوسط الحسابي لكل من المتغير المستقل والتابع ثم نقوم بتحديد كل من أ، ب، ج، د وهذا ما سنلاحظه في المثال التالي.

مثـــــال: الجدول التالي يبين لنا توزيع نسبة الوفيات وعدد الإصابات بمرض ما خلال ستة سنوات

السنــــوات	نسبة الوفـــــاة		عدد المصابيــن
٢٠٠٠	١٦	ك	٣٥٠
٢٠٠١	٢٠	ب	٦٠٠
٢٠٠٢	١٣	ك	٢٦٠
٢٠٠٣	٣٥	أ	٥٠٠
٢٠٠٤	١٣	ك	٤٢٠
٢٠٠٥	٤٠	أ	٥٢٠
المجمـــــوع			**٢٦٥٠**

$$\bar{س} = \frac{\Sigma س}{ن} = \frac{١٣٧}{٦} = ٢٢،٨٣$$

$$\bar{ع} = \frac{\Sigma ع}{ن} = \frac{٢٦٥٠}{٦} = ٤٤١،٦٦$$

أ = ٢، ب = ١، ج = ٠، ك = ٣

وبتطبيق القيم التالية في القانون نجــــــد:

$$\frac{\sqrt{ج ب} - \sqrt{أ د}}{\sqrt{أ د} + \sqrt{ج ب}} = \frac{\sqrt{١ \times ٠} - \sqrt{٢ \times ٣}}{\sqrt{٢ \times ٣} + \sqrt{١ \times ٠}} = \frac{٢،٤٥}{٣،٤٥} = ٠،٧١$$

وبهذا نلاحظ أن درجة الائتلاف بين نسبة المرض والوفاة قوية معنى هذا أن هذا المرض خطير ومميت.

٢ـ قياس الارتباط بين متغيرين مستقل وتابع ترتيبيين معا:

١- معامل ارتباط سيبرمان:

في كثير من الأحيان لا يستطيع الباحث تحديد قيم المتغير أثناء تغيره، ويصبح من السهل بالنسبة إليه ترتيب مراحل تغيره كأن يحدد أيهما الأول وأيهما الثاني وأيهما الأخير وفي هذه الحالة يستطيع الباحث ترتيب القيم المتعلقة بكل متغير أو ظاهرة وإيجاد العلاقة بين رتب المتغير الأول ورتب المتغير الثاني فإذا كانت هذه الرتب متفقة تماما كان الارتباط موجبا بصفة كاملة (+١) وإذا كان أحد المتغيرين أخذ ترتيبا تنازليا والثاني تصاعديا كان الارتباط بين المتغيرين سالبا (- ١) بصفة كاملة وطريقة إيجاد معامل سيبرمان تقوم على أساس أنه كلما كان الفرق بين رتب القيم المتقابلة، في المتغيرين كبيرا كلما قلت قيمة معامل الارتباط بين المتغيرين والعكس صحيح، لهذا كانت الخطوة الأولى في هذه الطريقة تشمل على إيجاد الفرق بين رتب القيم المتقابلة سوءا كانت مرتبة تصاعديا أم تنازليا معا، لكل من المتغير المستقل والتابع، وينتج عن هذه الفروق، قيم موجبة وأخرى سالبة الإشارة بحيث أن مجموع الفروق الموجبة يساوي مجموع الفروق السالبة، لهذا لإيجاد معامل ارتباط سيبرمان بين المتغيرين س،ع علينا من تربيع الفروق حتى تتحصل على الإشارات الموجبة فقط ويحسب معامل الارتباط لسبيرمان وفق العلاقة التاليـــــــــة:

$$ر = ١ - \frac{٦ \sum ف^{٢}}{ن(ن^{٢} - ١)}$$

مثـــــال: الجدول التالي يبين تقدير ١٠ طلاب في مادتي الإحصاء والاقتصاد السكاني والمطلوب إيجاد درجة الارتباط بين النتائج المتحصل عليها في كلا المادتين.

رقـــم الطالـــب	تقدير الإحصاء س	تقدير الاقتصاد السكاني ع	رتب س	رتب ع	ف	$ف^{٢}$
١	ضعيف جدا	جيد	١	٦	٥ -	
٢	مقبول	جيد	٥،٥	٦	٠،٥ -	٢٥
٣	ممتاز	جيد جدا	١٠	٨،٥	١،٥	٠،٢٥
٤	مقبول	مقبول	٥،٥	٣،٥	٢	٢،٢٥

٥	ضعيف	جيد	٢،٥	٦	– ٣،٥	٤
٦	جيد جدا	مقبول	٩	٣،٥	٥،٥	١٢،٢٥
٧	جيد	ممتاز	٨	١٠	– ٢	٣٠،٢٥
٨	ضعيف	ضعيف جدا	٢،٥	١	١،٥	٤
٩	مقبول	ضعيف	٥،٥	٢	٣،٥	٢،٢٥
١٠	مقبول	جيد جدا	٥،٥	٨،٥	– ٣	٩
المجموع			٥٥	٥٥	٠	١٠١،٥

$$ر = ١ - \frac{٦ \Sigma ف^{٢}}{ن(ن^{٢} - ١)} = ١ - \frac{٦ (١٠١،٥)}{١٠ (١٠^{٢} - ١)} = \frac{٦٠٩}{٩٩٠} - = ٠،٣٩$$

أي أن هناك ارتباط طردي ضعيف نوعا مابين درجات الطلاب العشر في مادتين الإحصاء والاقتصاد السكاني بمعنى آخر بإمكان الطالب أن يكون قد رسب في مادة لكن نجح في المادة الأخرى.

ملاحظـــــة: لاحظنا في هذا المثال أن كثير من الطلاب تحصلوا على نفس التقدير في هذه الحالة أخذنا متوسط رتبة التقديرين فمثلا فيما يخص ترتيب قيم المتغير المستقل آلا وهو تقدير الإحصاء نلاحظ أن تقدير ضعيف جدا أخذ المرتبة الأولى ثم بعد ذلك وجدنا تقدير ضعيف لكل من الطالبين رقم ٥ و٨، ففي هذه الحالة افترضنا أن الطالب ٥ أخذ الرتبة ٢ والطالب ٨ أخذ الرتبة ٣ ثم أعطتنا متوسط الرتبتين كرتبة لتقدير ضعيف أي $\frac{٢ + ٣}{} = \frac{٥}{} = ٢،٥$

ونفس الشيء فيما يخص التقدير الموالي له بما أننا قمنا بالترتيب التصاعدي للقيم وكانت بذلك تقدير مقبول هذا التقدير الذي اشترك فيه الطالب٢،٤، ٩ والعاشر وبنفس الطريقة نقوم بإعطاء رتب لهؤلاء الطلبة المشتركين في هذا التقدير ونحسب متوسطهم مثلا. نلاحظ أن الرتبة الأخيرة لتقدير ضعيف كانت ٣ إذا نواصل الترتيب فالطالب رقم ٢ يأخذ الرتبة ٤، الطالب ٣ يأخذ رتبة ٥، الطالب ٩ يأخذ رتبة ٦، الطالب ١٠ يأخذ رتبة ٧ ثم نحسب متوسط هذه الرتب أي: $\frac{٤ + ٥ + ٦ + ٧}{٤} = \frac{٢٢}{٤} = ٥،٥$

وللتأكد من صحة وضع الترتيب المقابل لكل قيمة في رتب المتغيرين هناك وسيلة للتأكد من ذلك وهي أن يكون مجموع رتب المتغير المستقل تساوي مجموع رتب المتغير التابع ويساوي النتيجة المتحصل عليها من القانون التالي: $\frac{ن (ن + ١)}{٢}$

وبتطبيق القانون على المثال السابق نجد عدد القيم ن = ١٠

$$\frac{ن (ن + ١)}{٢} = \frac{١٠ (١٠ + ١)}{٢} = ٥٥$$

كذلك نجد مجموع رتب س = مجموع، رتب ع = ٥٥

٣_ المعاملات المستعملة في حالة متغيرات كيفية أو كيفية مع كميـــــة:

أ_ معامل ارتباط الاقتران:

واضع هذا القانون يونع ونستعمله في حالة وجود توزيع تكراري مزدوج بسيط التقاطع أي أن كل من المتغير المستقل والتابع يحتويان على قيمتين فقط أي يكون الجدول مشتملا على ٤ خانات للتقاطع فقط وهو يستعمل عادة في حساب درجة العلاقة بين المتغيرات الكيفية. ويحسب وفق القانون التالي:

$$ن = \frac{أ د \_ ب ج}{أ د + ب ج}$$

مثـــال: أوجد درجة الاقتران بين المتغيرين المستقل متمثل في عمل المرأة والتابع في مدى استعمالها لوسائل منع الحمل.

الإستعمال لوسائل منع الحمـــــل / العمـــــل	تستعمـــــل	لا تستعمـــــل	المجمـــــوع
تعمـــــل	٢٠ أ	٥ **ب**	٢٥
لا تعمـــــل	١٠ **ج**	١٥ **د**	٢٥
المجمـــــوع	**٣٠**	**٢٠**	**٥٠**

$$\text{ن} = \frac{\text{أ د \_ ب ج}}{\text{أ د + ب ج}} = \frac{(٢٠)(١٥) - (٥)(١٠)}{(٢٠)(١٥) + (٥)(١٠)} = \frac{٣٠٠ - ٥٠}{٣٠٠ + ٥٠} = ٠،٧١$$

إذن فهناك اقتران قوي موجب بين عمل المرأة ومدى استعمالها لوسائل منع الحمل أي كلما كانت المرأة عاملة كلما زاد إقبالها على استعمال وسائل منع الحمل.

ب – معامل فآي:

يعتبر هذا المعامل حالة خاصة من الحالات التي يستخدم فيها معامل التوافق فهو لا يستخدم إلاّ في الحالات التي يقسم فيها كل من المتغيرين إلى قيمتين فقط ومن أمثالها الصفات ذكر، أنثى ... إلخ، لكن عند استخدامه نستعمل نسب مئوية فإذا أراد الباحث معرفة أثر الرأي في ظاهرة ما حسب الجنس يمكن استعمال معامل فآي لأن الرأي مقسم إلى نعم لا مثلا، والجنس، ذكر، أنثى.

ويحسب معامل فآي وفق الصفة التاليـــــــة:

$$\varphi = \frac{\text{ب ج – أ د}}{\sqrt{\text{هـ ى هـَ ىَ}}}$$

مثـــال: أحسب معامل فآي للجدول التالي الذي يربط علاقة بين مكان سكن المبحوث برأيه في ظاهرة التدخين

الـــــرأي / مكان الإقامة	مضـــــــر	غير مضـــــــر	المجمــــوع
حضر	٢٠	٢٦	٤٦
ريف	٢٨	٢٩	٥٧
المجمـــــــوع	٤٨	٥٥	١٠٣

لحساب معامل فآي لابد من تحويل القيم إلى نسب مئوية بالنسبة لمجموع العام للعينة، وتصبح كما يلي في الجدول التالي:

الـــرأي / مكان الإقامة	مضــر	غير مضــر	المجمـــــوع
خطر	١٩،٤٢ أ	٢٥،٢٤ ب	٤٤،٦٦ هـَ
	٢٧،١٨ ج	٢٨،١٥ د	٥٥،٣٤ يَ
المجمــــــوع	٤٦،٦٠ هــ	٥٣،٣٩ ي	١٠٠

$$\varphi = \frac{\text{أ ك} - \text{ب ج}}{\sqrt{\text{هـ ي هـَ يَ}}} \quad \frac{(١٩،٤٢)(٢٨،١٥) - (٢٥،٢٤)(٢٧،١٨)}{\sqrt{٤٦،٦ \times ٥٣،٣٩ \times ٤٤،٦٦ \times ٥٥،٣٤}}$$

$$\varphi = \frac{٥٤٦،٦٧ - ٦٨٦،٠٢}{\sqrt{٦١٤٨٩٨٨،٤}} = -٠،٠٠٦ =$$

أي أن هناك علاقة عكسية ضعيفة جدا تقريبا منعدمة، بين مكان الإقامة للمبحوث ورأيه فيما كان التدخين مفر بالصحة.

ج - معامل التوافـــــــــق:

إذا كانت البيانات المطلوب دراسة العلاقة بينهما عبارة عن بيانات وصفية لكلا المتغيرين أو وصفية لأحدهما وكمية للمتغير الآخر –أي المستقل والتابع- وكانت مقدمة لنا على شكل أكثر من قيمتين لأحد المتغيرين أي لدينا على الأقل ستة خانات للتقاطع فإننا نستعمل معامل التوافق.

ولحساب معامل التوافق نتبع الخطوات التاليـــــة:

١. نربع تكرار كل فئة.

٢. نقسم مربع تكرار كل فئة على حاصل ضرب مجموع التكرارات الأفقية والرأسية للصف والعمود ثم نجمع خوارج القسمة ولنفرض أن المجموع = ج

وبحسب معامل التوافق وفق القانون التالي:

$$ق = \sqrt{\frac{ج - ١}{ج}}$$

مثـــــال: في دراسة حول انتشار الأمية في منطقة ما وقفنا عند هذا الجدول الذي يبين توزيع مجموعة من الأفراد حسب سنهم ومستواهم التعليمي.المطلوب: حساب درجة التوافق بين المتغيرين

المستوى التعليمي / السن	أمـــــي	ابتدائـــي	متوسـط	ثانـــوي	جامعـي	المجموع
١٠ – ١٥	–	٤	٢	–	–	٦
١٥ – ٢٠	٦	١	١٠	١	–	١٨
٢٠ – ٢٥	١	٦	٤	٣	١	١٥
٢٥ – ٣٠	–	١	١	٢	٣	٧
٣٠ – ٣٥	–	–	١	٢	١	٤
المجـــموع	٧	١٢	١٨	٨	٥	٥٠

وبهذا تنحصر عملية الحصول على معامل التوافق في إيجاد مربع تكرار كل خلية تقاطع مقسوما على حاصل ضرب تكرار العمود × تكرار الصف كما يلي:

$$ج_{١} = \frac{١}{٦}\left[\frac{(٠)^{٢}}{٧} + \frac{(٤)^{٢}}{١٢} + \frac{(٢)^{٢}}{١٨} + \frac{(٠)^{٢}}{٨} + \frac{(٠)^{٢}}{٥}\right] = ٠،٢٥$$

$$ج_{٢} = \frac{١}{١٨}\left[\frac{(٦)^{٢}}{٧} + \frac{(١)^{٢}}{١٢} + \frac{(١٠)^{٢}}{١٨} + \frac{(١)^{٢}}{٨} + \frac{(٠)^{٢}}{٥}\right] = ٠،٦٠$$

$$ج_{٣} = \frac{١}{١٥}\left[\frac{(١)^{٢}}{٧} + \frac{(٦)^{٢}}{١٢} + \frac{(٤)^{٢}}{١٨} + \frac{(٣)^{٢}}{٨} + \frac{(١)^{٢}}{٥}\right] = ٠،٣٥$$

$$ج_{٤} = \frac{١}{٧}\left[\frac{(٠)^{٢}}{٧} + \frac{(١)^{٢}}{١٢} + \frac{(١)^{٢}}{١٨} + \frac{(٢)^{٢}}{٨} + \frac{(٣)^{٢}}{٥}\right] = ٠،٣٤$$

$$ج_{٥} = \frac{١}{٤}\left[\frac{(٠)^{٢}}{٧} + \frac{(٠)^{٢}}{١٢} + \frac{(١)^{٢}}{١٨} + \frac{(٢)^{٢}}{٨} + \frac{(١)^{٢}}{٥}\right] = ٠،١٨$$

مجموع: $ج = ج_{١} + ج_{٢} + ج_{٣} + ج_{٤} + ج_{٥} = ٠،٢٥ + ٠،٦٠ + ٠،٣٥ + ٠،٣٤ + ٠،١٨ = ١،٧٢$

وبالتعويض بقيمة ج في القانون نجد أن معامل التوافق يقدر في هذا التوزيع بـــ:

$$ق = \sqrt{\frac{ج - ١}{ج}} = \sqrt{\frac{١،٧٢ -}{١}} = ٠،٦٤$$

وبهذا نلاحظ أن درجة التوافق بين السن والمستوى التعليمي للمبحوث متوسطة.

د- معامل ارتباط لامـــــــد:

ويستخدم هذا المعامل في حالة وجود متغيرين إما كيفيين معا أو أحدهما كيفي والثاني كمي وواضع هذا القانون العالم جوتمان عام ١٩٤١.

ويحسب وفق الصيغة التالية:

$$ل_{س ع} = \frac{\sum ك - ك_{ع}}{ن - ك_{ع}}$$

حيث أن ك هو تكرار الفئة المنوالية لكل فئة من فئات المتغير المستقل ك $_{ع}$ هو عبارة عن الفئة المنوالية للتوزيع الهامشي لمتغير التابع ع.

ملاحظـــــة: إن كل من معامل التوافق ولامدا تنحصر قيمتهما بين ٠، +١.

مثـــال: الجدول التالي يربط علاقة بين الطبقة الاجتماعية التي تنتمي إليها المبحوثة بسنها عند أول زواج والمطلوب إيجاد درجة العلاقة الإرتباطية بين المتغير ين حسب معامل لامدا.

الطبقــة / الســن	طبقة فقيـــرة	طبقة متوسطــة	طبقة غنيـــة	المجمـــوع
٢٠ – ٢٥	٥٠	٣٠	١٠	٩٠
٢٥ – ٣٠	١٥	٤٠	٢٠	٧٥
٣٠ – ٣٥	١٣	١٨	٣٥	٦٦
المجمـــوع	٧٨	٨٨	٦٥	**٢٣١**

الحـــــل: نضع المتغير التابع ع والمتغير المستقل الطبقة الاجتماعية (س)

$$\text{إذن: ل}_{\text{ع س}} = \frac{\text{مج ك ـ ك}_{\text{ع}}}{\text{ن ـ ك}_{\text{ع}}} = \frac{١٢٥ - ٩٠}{٢٣١ - ٩٠} \quad \frac{٣٥}{١٤١} = ٠{,}٢٥$$

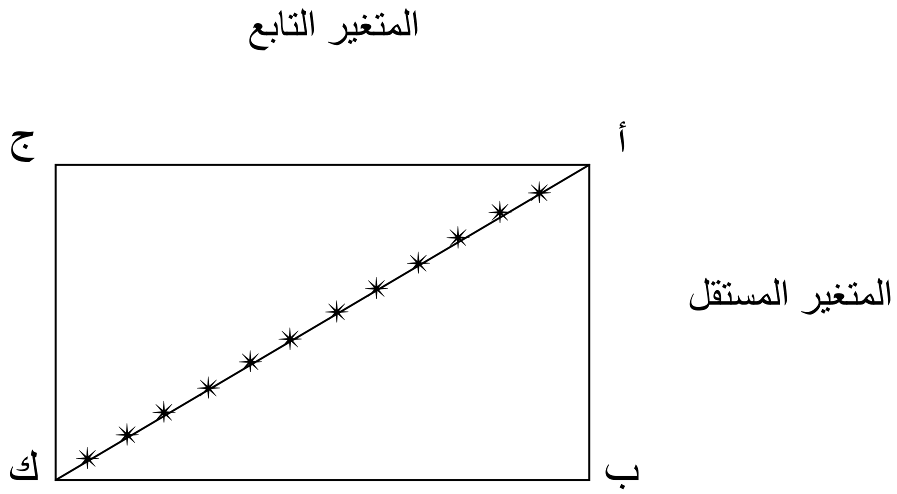
ر _ أشكال الانتشار:

ولتوضيح كل ما جاء من أنواع وقيمة الارتباط نفترض إحدى الحالات التي يكون فيها الارتباط.

أ)- موجب تماما أي الارتباط كلي ر = +١

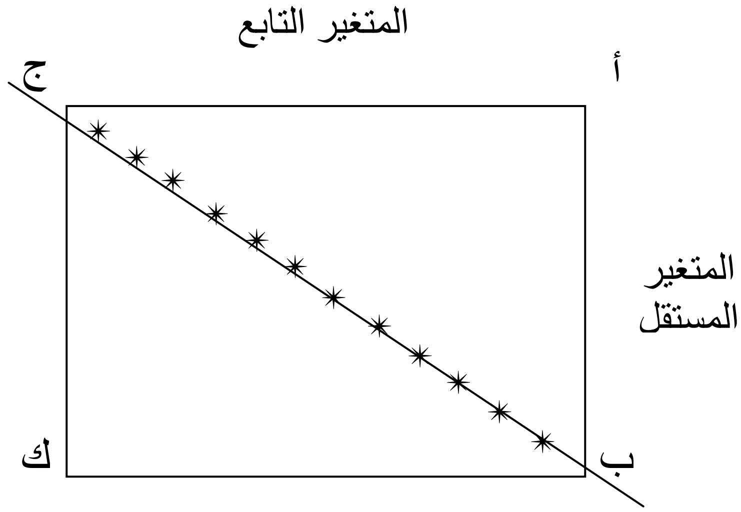
عندما يكون الارتباط يقدر بـــ +١ بين المتغير المستقل والمتغير التابع فإننا نلاحظ أن جميع التكرارات تكون متجمعة في خط مستقيم هو قطر الشكل أ أين يصل الركن أ والركن ك وذلك لأن جميع القيم الصغيرة في أحد المتغيرين تتبعها قيم صغيرة في

المتغير الآخر وكلما زادت وكبرت القيمة في أحد المتغيرين زادت وكبرت القيمة المقابلة لها في المتغير الآخر كما يوضحه الشكل – أ –

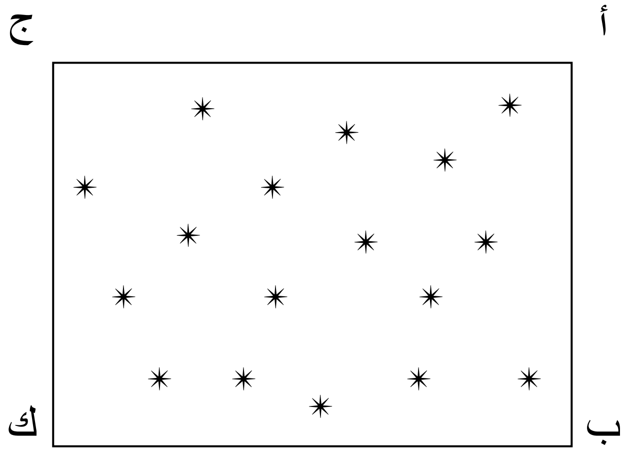


ب)- سالب تماما أي الارتباط كلي سالب ر = - ١

أما إذا كانت العلاقة تامة سالبة (- ١) تجمعت التكرارات في القطر الذي يربط بين الركنين ج ب في تخطيط الانتشار وذلك لأن القيم الصغيرة في أحد المتغيرين تقابلها القيم الكبيرة في المتغير الثاني. الشكل – ب –



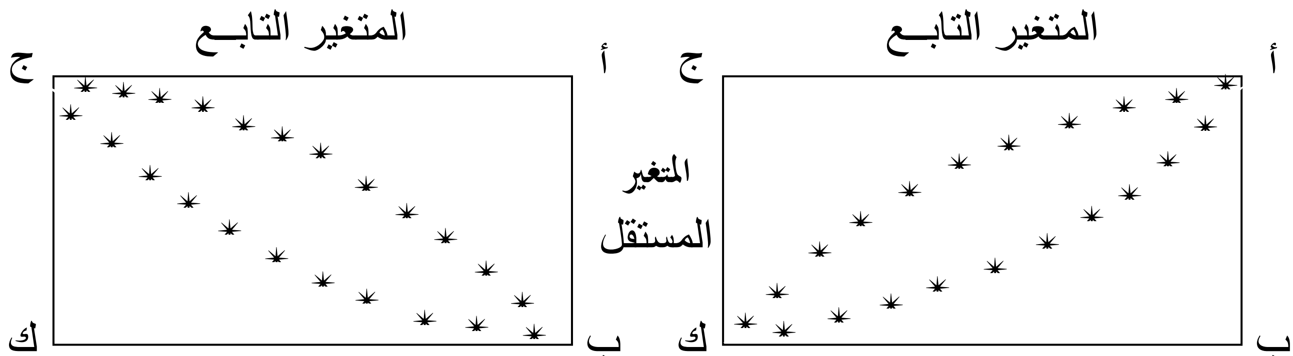
ج)- إما إذا كانت العلاقة صفرية أي منعدمة بين المتغيرين فإن نقط التكرارات تكون موزعة على الشكل دون أن تجد إتجاه عام لها في تجميعا كما هو الحال في الشكل – ج –



د)- وفي حالة العلاقة الجزئية سواء كانت موجبة مثل الشكل (ك) أو سالبة مثل الشكل (و) نجد انتشار التكرارات يتخذ اتجاها عاما إلاّ أن هذا الاتجاه يتبع عادة شكل بيضاوي ⬭ وكلما اتسع الشكل قلت قيمة معامل الارتباط بين المتغيرين وكلما ضاق زادة قيمة معامل الارتباط حتى تصل أقصاها عندما يصبح الشكل خطا محددا.

وكأن مجرد ملاحظة التوزيع في تخطيط الانتشار يفيد في معرفة العلاقة بين المتغيرين ونوعها، لكن الإحصاء الوصفي/ الاجتماعي لا يكتفي بهذا أي ملاحظة التوزيع ووصف العلاقة وصفا تقريبيا بل يهدف إلى قياس عددي لهذه العلاقة بإحدى المعاملات المذكورة سابقا.

الشكل –د- الشكل –و-



ز ـ الرسم البياني لخطوط الانحدار:

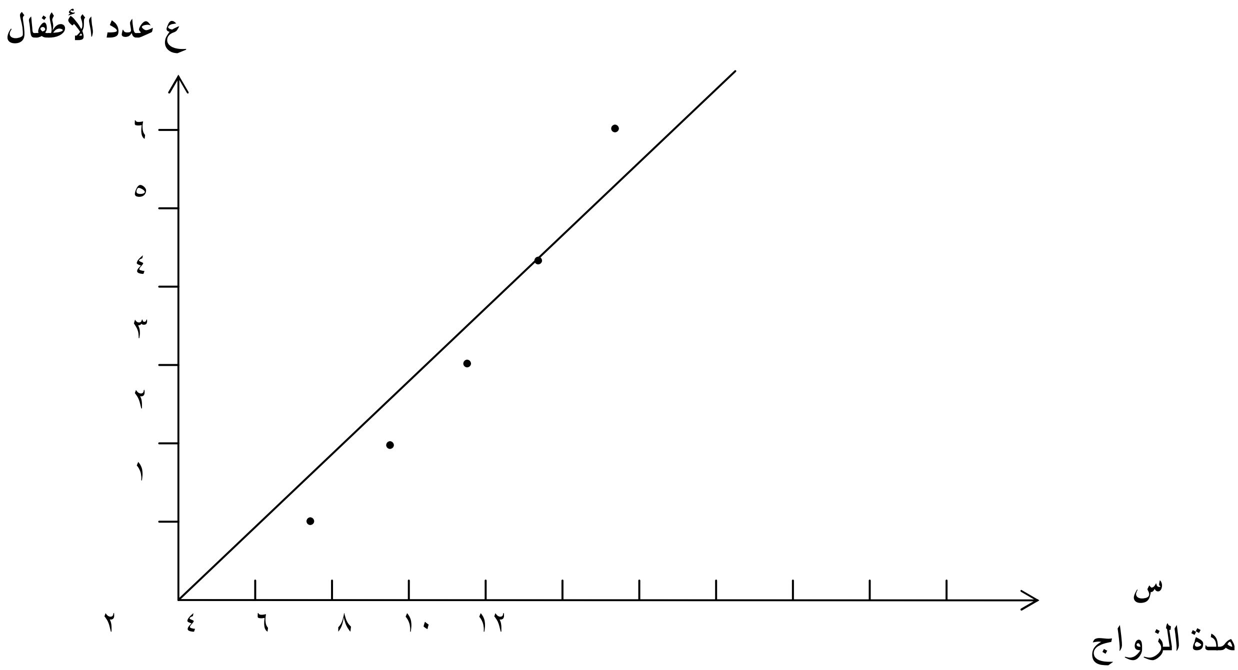
تعتبر طريقة الرسم البياني في كشف نوع العلاقة الإرتباطية بين متغيرين من أسهل الطرق وأوضحها وهي تتلخص بتمثيل القيم المتناظرة للظاهرتين في مستوى إحداثيتي م س، م ع، بحيث يتم تمثيل كل قيمتين متناظرتين بنقطة في المستوي ثم نصل بينها في حالة ما إذا وجد اتجاه عام لها يوحي بوجود علاقة وهذه النقاط تنتشر على المستوي بأشكال مختلفة مثلما لاحظنا أعلاه.

في حالة بيانات غير مبوبـــــة:

مثـــــال: الجدول التالي يبين لنا مدّة الزواج وعدد الأطفال لــ ستة نساء والمطلوب تمثيلها بخط انحدار $\frac{س}{ع}$

مدة الـــــزواج	عدد الأطـــــفال
٢	٠
٤	١
٦	٢
٨	٣
١٠	٤
١٢	٦

فلو مثلنا هذه القيم على أساس أن مدة الزواج هي الظاهرة المؤثرة السببية أي كمتغير مستقل وعدد الأطفال هي ظاهرة تابعة فإننا نخصص محور السينات لقيم مدة الزواج $\frac{س}{ع}$ والمحور م ع لعدد الأطفال وبهذا سيكون خط الانحدار هو



خطوط الانحدار في التوزيعات التكرارية المزدوجـــــة:

في مثالنا السابق عن العلاقة بين مدة الزواج وعدد الأطفال كانت كل قيمة من المتغير المستقل (س) تقابل قيمة واحدة فقط من قيم المتغير التابع (ع) غير أنه في كثير من الحالات تكون القيمة الواحدة للمتغير المستقل تقابل عدة قيم من المتغير التابع أي في حالة الجداول التكرارية المزدوجة حيث أنه مثلا قيمة س$_{1}$ تقابل ع$_{1}$، ع$_{2}$، ع$_{3}$ ... ع$_{n}$.

فلو أردنا في مثل هذه الحالة تمثيل العلاقة بين س، ع فإن أسهل طريقة يمكن إتباعها هي أن نرصد كل قيمة من قيم المتغير التابع ع مع المتوسط الحسابي لجميع قيم المتغير المستقل (س) المرتبطة بهاو بالمثل نرصد كل قيمة من قيم المتغير المستقل س مع المتوسط الحسابي لجميع قيم المتغير ع المرتبطة بها وبالتالي تكون معادلة الانحدار في هذه الحالة.

$$\frac{س}{ص} = (س - \bar{س}) = ر \frac{ع_{س}}{ع_{ع}} (ع - \bar{ع})$$

$$\frac{ص}{س} = (ص - \bar{ص}) = ر \frac{ع_{ص}}{ع_{ع}} (ع - \bar{ع})$$

أما معامل الانحدار: $\frac{ع}{س} = م = ر \frac{ع_{ع}}{ع_{س}}$

أما معامل الانحدار: $\frac{س}{ع} = م = ر \frac{ع_{س}}{ع_{ع}}$

مثـــال:

بتطبيق هذه القوانين على الجدول السابق المستعمل في حساب معامل كارل برسون نجد:

معادلة: $\frac{س}{ع} = (س - \bar{س}) = ر \frac{ع}{س} (ع - \bar{ع})$

$$(س - ٤٨) = ٠،٠٦٧ \left[\frac{١٢،٢٥}{٩،٤٩}\right] (ع - ٧٦،٧٨)$$

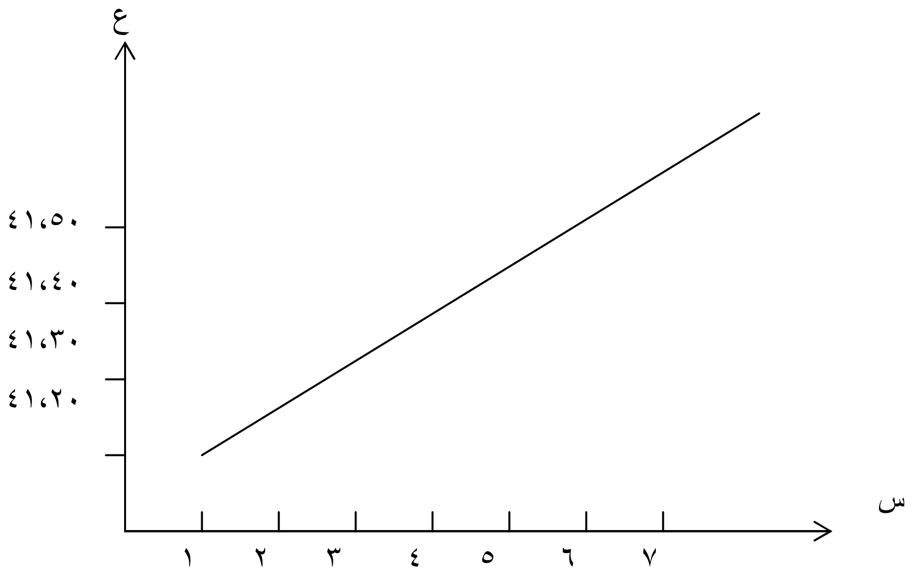
$$س - ٤٨ = ٠،٠٩ (ع - ٧٦،٧٨)$$

$$س - ٤٨ = ٠،٠٩ ع - ٦،٩١$$

$$س = ٠،٠٩ ع - ٦،٩١ + ٤٨$$

$$س = ٠،٠٩ ع + ٤١،٠٩$$

س	١	٢	٣	٤	٥	٦
ع	٤١، ١٨	٤١، ٢٧	٤١، ٣٦	٤١، ٤١	٤١، ٥٤	٤١، ٦٣



أما معامل الانحدار $\frac{س}{ع}$ = م = ر $\frac{ع_س}{ع_ع}$ = ٠،٦٧ × $\frac{١٢،٢٥}{٩،٤٩}$ = ٠،٠٨٦

وبنفس الطريقة نقوم بحساب معادلة انحدار $\frac{ع}{س}$ ومعامل انحدار $\frac{ع}{س}$ بتعويض قيمهم في المعادلة فقط.

التمرين الأول:

الجدول التالي يربط بين الطبقة الاجتماعية التي ينتمي إليها الطالب و موقفه من ظاهرة التدخين في الجامعة.

الرأي / الطبقة	موافق	غير موافق	حيادي	المجموع
الطبقة الفقيرة	٥	٤	١٦	٢٥
الطبقة المتوسطة	١٨	١٦	١٢	٤٦
الطبقة الغنية	٢٨	٢٥	١٤	٦٧
المجموع	**٥١**	**٤٥**	**٤٢**	**١٣٨**

ما هو شكل انتشار هذه الجداول التكرارية؟

الحل:

نلاحظ في هذا الجدول أن قيم المتغير المستقل تميل إلى الارتفاع و العكس اتخذته قسم المتغير التابع إذ أخذت في الانخفاض، و هذا يدل على أن الارتباط سيكون عكسي بين المتغيرين (الطبقة الاجتماعية و الرأي) هذا من جهة و من جهة أخرى نلاحظ أن الخط المستقيم المائل الذي يمر على أكبر قيم للتقاطع ينحدر من أعلى اليسار إلى أسفل اليمين دليل على وجود ارتباط عكسي قوي بين المتغيرين.

التمرين الثاني:

أراد باحث في علم النفس الاجتماع دراسة تأثير المعالجة من تعاطي المخدرات.

و نتائجها.فوقف عند هذه الجدول.

النتيجة / العلاج	شفيوا	لم يشفو	المجموع
عولجوا	١١٥	٢٥	١٤٠
لم يعالجوا	٣٥	١٧٥	٢١٠
المجموع	١٥٠	٢٠٠	٣٥٠

ما هي العلاقة بين العلاج و نتيجته؟

الحل:

نلاحظ أن هذا الجدول يحتوي على متغيرين فقط و كلا المتغيرين يحتويان على قيمتين فقط إذا فأحسن معامل لقياس درجة الارتباط بينهما هو معامل الاقتران:

$$ن = \frac{(أ)(ك)-(ب)(ج)}{(أ)(ك)+(ب)(ج)} = \frac{(35)(25)-(175)(115)}{(35)(25)+(175)(115)} = \frac{19250}{21000} = ٠,٩١$$

أي هناك اقتران قوي بين العلاج و نتيجته أي أنه كلما تقدم المرضى / متعاطي المخدرات الرائز عن العلاج كلما كانت نتيجة ذلك حسنة أي التخلي عنها و العكس صحيح.

التمرين الثالث:

الجدول التالي يبين لنا سنة الولادات المتوئمة و عدد الولادات في مستشفى حسب الأشهر.

الشهور	نسبة الولادات الموئمة٪	عدد الولادات
جانفي	٥	٢٠٠
فيفري	١٢	١٠٠
مارس	٣	٢٥٠
أفريل	٨	٢٠٠
ماي	٢	٣٠٠
جوان	٤	٣١٠
جويلية	٣	٢٠٠
أوت	١	٢١٠
سبتمبر	٤	٣٦٠
أكتوبر	٥	٤٠٠
نوفمبر	٧	٣٢٠
ديسمبر	٨	٤٠٠

ما هي درجة العلاقة بين نسبة الولادات المتوئمة و عدد الولادات في هذا المستشفى؟

<u>الحل:</u> يمكن حل هذا التمرين بمعامل الاختلاف.

$$\overline{س} = \frac{مج(س)}{ن} = ٥,١٧$$

$$\overline{ص} = \frac{مج(ص)}{ن} = ٢٧٠,٨٣$$

أ = ٢، ب = ٤، ج = ٢، د = ٤

$$ل = \frac{\sqrt{(4)(2)} - \sqrt{(2)(4)}}{\sqrt{(4)(2)} + \sqrt{(4)(2)}} = \frac{0}{5.66} = 0$$

أي أن درجة الاختلاف بين نسبة الولادات المتوئمة و عدد الولادات منعدمة، فلا توجد علاقة في توزيعهما عبر أشهر السنة.

التمرين الرابع:

أحسب معامل ارتباط الرتب لسبران للبيانات التالية:

س	٧٥	٧٥	٦٥	٨٢	٧٥	٨٢	٩٠	٦٥
ع	٦٩	٥٦	٥٦	٥٠	٨٨	٦٩	٨٩	٨٩

<u>الحل:</u>

س	ع	رتب س	رتب ع	ف	ف٢
٧٥	٦٩	٤	٤,٥	– ٠,٥	٠,٢٥
٧٥	٥٦	٤	٢,٥	– ١,٥	٢,٢٥
٦٥	٥٦	١,٥	٢,٥	– ١,٠	١
٨٢	٥٠	٦,٥	١	٥,٥	٣٠,٢٥
٧٥	٨٨	٤	٦	– ٢	٤
٨٢	٦٩	٦,٥	٤,٥	٢	٤
٩٠	٨٩	٨	٧,٥	٠,٥	٠,٢٥
٦٥	٨٩	١٥	٧,٥	– ٦	٣٥
المجموع		**٣٦**	**٣٦**	**٠**	**٧٨**

ر = $\frac{1-6\ \text{مج}(ف)}{ن(ن^2-1)} = \frac{(78)1-1}{(63)8} = \frac{468-1}{504}$ = ١ – ٠,٩٣ = ٠,٠٧

يعني هذا أن هناك ارتباط ضعيف للغاية تقريبا منعدما بين المتغيرين س،ص.

للتأكد من صحة وضع الترتيب المقابل لكل قيمة لرتب المتغيرين تقوم الرتب في كلا المتغيرين متساوي = ٣٦.

و يساوي النتيجة المتحصل عليها من القانون التالي $\frac{ن(ن+1)}{2}$ و بالتطبيق $\frac{8(8+1)}{2}$ = ٣٦

الترتيب صحيح في الجدول.

التمرين الخامس:

الجدول التالي يوضح تصرفات الأمهات اتجاه أطفالهن المصابين بإسهال حسب المستوى التعليمي.

و المطلوب حساب قيمة الارتباط بين تصرف الأم و مستواها التعليمي.

التصرف \ المستوى التعليمي	أمي	ابتدائي	متوسط	ثانوي+جامعي	المجموع
كشف طبي	٦	٢٥	٢٥	١٨	٧٤
دواء	١	٤	٥	٦	١٦
محاليل تقليدية	٥	٨	٨	٥	٢٦
محاليل تقليدية+كشف طبي	٢	١٢	١٣	١١	٣٨
محاليل + دواء	١	٧	٥	٣	١٦
المجموع	١,٥	٥٦	٥٦	٤٣	١٧٠

الحل: حساب معامل التوافق **ق** = $\sqrt{\frac{ج - 1}{ج}}$

ج١ = $\frac{1}{74}\left[\frac{(6)^2}{15}+\frac{(26)^2}{56}+\frac{(25)^2}{56}+\frac{(18)^2}{43}\right] = \frac{1}{74}(7.53+11.16+11.16+2.4) =$ ٠,٤٣

ج٢ = $\frac{1}{16}\left[\frac{(11)^2}{15}+\frac{(4)^2}{56}+\frac{(5)^2}{56}+\frac{(6)^2}{43}\right] = \frac{1}{16}(0.84+0.45+0.28+0.07) =$ ٠,١٠

ج٣ = $\frac{1}{26}\left[\frac{(5)^2}{15}+\frac{(8)^2}{56}+\frac{(8)^2}{56}+\frac{(5)^2}{43}\right] = \frac{1}{26}(0.58+1.14+1.14+1.67) =$ ٠,١٧

ج٤ = $\frac{1}{38}\left[\frac{(2)^2}{15}+\frac{(12)^2}{56}+\frac{(13)^2}{56}+\frac{(11)^2}{43}\right] = \frac{1}{38}(2.81+3.02+2.75+0.27) =$ ٠,٢٣

ج٥ = $\frac{1}{16}\left[\frac{(1)^2}{15}+\frac{(7)^2}{56}+\frac{(5)^2}{56}+\frac{(3)^2}{43}\right] = \frac{1}{16}(0.21+0.45+0.87+0.07) =$ ٠,١

مجموع ج = ٠,٤٣ + ٠,١ + ٠,١٧ + ٠,٢٣ + ٠,١ = ١,٠٣

∴ **ق** = $\sqrt{\frac{ج - 1}{ج}} = \sqrt{\frac{ج - 1}{ج}} = \sqrt{\frac{903}{1.03}} =$ ٠,١٧

هذا يعني أن هناك توافق ضعيف جدا بين تصرف الأم اتجاه الرضيع عند الإجابة بالإسهال و مستواها التعليمي أي مهما اختلف مستواها التعليمي إلا و أخذته إلى الطبيب.

الفصـــــل الســــــابع

إختبـــــــار كــــــا[٢]

أولا: إختبار كــــــا[٢]

ثانيا: تصحيح يانـــس

١ـ اختبار كــا[٢]:

إن الباحث الاجتماعي عند معرفته بتوزيع كا[٢]، ويضعه تحت تصرفه فله جانب كبير من الأهمية وتستعمل هذه الأداة بصورة رئيسية لاختبار الفرضيات التي تقوم على أساس مقارنة مجموعة من التكرارات النظرية مع مجموعة من التكرارات الفعلية لتقيم الفرق بينهما لمعرفة ما إذا كان هذا الفرق فرقا ظاهريا نتيجة قوى الحظ والصدفة أم أنه فرق حقيقي نتيجة قوى أخرى غير قوى الحظ والصدفة فإذا وجد أن هذا الفرق كان فرقا ظاهريا بمستوى دلالة معين نقبل فرضية العدم أما إذا وجد أن هناك فرقا حقيقيا بمستوى دلالة معين يرفض فرضية العدم، وهذا وأن إحدى مزايا هذا الاختبار الرئيسية أنه لا يتضمن أية افتراضات حول شكل توزيع المجتمع الإحصائي، وعليه فيمكن اختبار العلاقة بين المستوى التعليمي للمرأة وعدد أبنائها، أو العلاقة بين العقم والطلاق ... إلخ.

ومن أجل اختبار استقلالية الظواهر نقوم بتصنيف البيانات في جدول خاص كما سوف نرى في المثال التالي.

مثـــــال: الجدول التالي يربط علاقة بين الحالة المدنية للشخص ومدى ادخاره ونريد معرفة هل ادخار المبحوث مستقل عن كونه متزوج أم لا.

الادخار / الحالة المدنيـــة	مدخـــــــر	غير مدخـــــر	المجمـــوع
متزوج	٩٥٠	٦٢٥	١٥٧٥
أعزب	٢٥٠	١٢٥	٣٧٥
المجموع	**١٢٠٠**	**٧٥٠**	**١٩٥٠**

1- **نقوم بتحديد فرضية العدم:** وفي هذا المثال ليس هناك علاقة بين الحالة المدنية والإيجار للشخص المبحوث.

٢- **تحديد مستوى الدلالة:** ونسبته في الغالب ١% أو ٥% في العلوم الاجتماعية.

٣- **تحديد التكرارات النظرية:** وذلك بأن نعد جدول توافق نظري حيث نفترض أن المجاميع فيه متساوية ويحسب وفق القانون التالي.

$$ كَ = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلي/ العينة}} $$

وبتطبيق هذا القانون على الجدول أعلاه نجد:

$$ كَ\ \text{المدخرون المتزوجون} = \frac{١٩٧٥ \times ١٢٠٠}{١٩٥٠} = ٩٦٩،٢٣ $$

$$ كَ\ \text{المدخرون العزاب} = \frac{٣٧٥ \times ١٢٠٠}{١٩٥٠} = ٢٣٠،٧٧ $$

$$ كَ\ \text{الغير مدخرون المتزوجون} = \frac{١٥٧٥ \times ٧٥٠}{١٩٥٠} = ٦٠٧،٧٧ $$

$$ كَ\ \text{الغير مدخرون العـــزاب} = \frac{٧٥٠ \times ٣٧٥}{١٩٥٠} = ١٤٤،٢٣ $$

4- **حساب كا^٢ المحسوبة:** وتحسب وفق القانون التالي:

$$ كا^٢ = \text{مج} \frac{(ك - كَ)^٢}{كَ} $$

ك	كَ	ك – كَ	(ك – كَ)^٢	(ك – كَ)^٢ / كَ
٩٥٠	٩٦٩،٢٣	– ١٩،٢٣	٣٦٩،٧٩	٠،٣٨
٢٥٠	٢٣٠،٧٧	١٩،٢٣	٣٦٩،٧٩	١،٦٠
٦٢٥	٦٠٥،٧٧	١٩،٢٣	٣٦٩،٧٩	٠،٦١
١٢٥	١٤٤،٢٣	١٩،٢٣	٣٦٩،٧٩	٢،٥٦
المجمـــوع				**٥،١٥**

٥- **الخطوة المتبقية** هي معرفة درجة الحرية ثم الكشف في جدول كا^{٢} علما إذا كانت قيمة كا٢ لهذه القيمة لدرجة الحرية ذات دلالة إحصائية عند نسبة ٥% أو ١% مثلا.

ودرجة الحرية تحسب وفق القانون التالي: (عدد الصفوف - ١) (عدد الأعمدة - ١)

أي (ن$_{١}$- ١) (ن$_{٢}$ - ١) = (٢- ١) (٢- ١) = ١

٦- عند درجة حرية كا^{٢} ١ الجدولة نجدها من الجدول الموضوع في نهاية الكتاب = ٣،٨٤١ عند نسبة ٥%.

٧- **القرار والمقارنـــــة:** إذا قارنا بين قيمة كا^{٢} المحسوبة = ٥،١٥ مع كا^{٢} الجد ولية من أجل مستوى دلالة ٥% والتي قدرت بـــ ٣،٨٤١ لوجدنا أن كا^{٢} الجد ولية < من كا^{٢} المحسوبة، ومنه نستنتج أن الفروق بين التكرارات النظرية والتكرارات الفعلية هي فروق جوهرية وبالتالي نرفض فرضية العدم ونقول أن هناك العلاقة بين الحالة المدنية للمبحوث وادخاره الشهري.

ملاحظـــــة:

- إذا كانت كا^{٢} الجد ولية > من كا^{٢} المحسوبة نستنتج أن هناك فروق بين التكرارات النظرية والتكرارات الفعلية وهي فروق ظاهرية راجعة للصدفة وبالتالي نقبل فرضية العدم والاستقلال أي لا توجد علاقة بين المتغيرين.
- أما إذا كانت كا^{٢} الجد ولية < من كا^{٢} المحسوبة نستنتج أن هناك فروق جوهرية ونرفض فرضية العدم أو الاستقلال ونقول أن هناك علاقة بين المتغيرين.

تصحيح يانـــــس Yets:

يقوم هذا التصحيح على أساس إنقاص (٠،٥) من الفرق بين التكرارات النظرية والتكرارات الفعلية الشيء الذي يؤدي إلى زيادة احتمال أن يكون كا^{٢} المحسوبة ناتجة لقوى الحظ والصدفة. لكن هذا الإنقاص شرط أن يكون بالقيمة المطلقة ويستخدم كا^{٢}

المصحح إذا كان الفرق بين التكرارات النظرية والفعلية ك – كَ < ٥ ويصبح بذلك قانون اختبار كا^{٢} المصحح كالتالي:

$$\text{كا}^{٢} = \frac{(\text{ا ك – كَ ا – ٠،٥})^{٢}}{\text{ك}}$$

وبتطبيق هذا القانون على المثال السابق نجد كا^{٢} المصحح يقدر بـــ:

ك	كَ	ك – كَ	ا ك – كَ ا – ٠،٥	(ا ك – كَ ا – ٠،٥)^{٢}	(ا ك – كَ ا – ٠،٥)^{٢} / ك
٩٥٠	٩٦٩،٢٣	– ١٩،٢٣	١٨،٧٣	٣٥٠،٨١	٠،٣٦
٢٥٠	٢٣٠،٧٧	١٩،٢٣	١٨،٧٣	٣٥٠،٨١	١،٥٢
٦٢٥	٦٠٥،٧٧	١٩،٢٣	١٨،٧٣	٣٥٠،٨١	٠،٥٨
١٢٥	١٤٤،٢٣	١٩،٢٣	١٨،٧٣	٣٥٠،٨١	٢،٤٨
المجموع					٤،٨٩

نفس الملاحظة تذكر بعد تصحيح كا^{٢} أين وجدنا كا^{٢} المحسوبة> كا^{٢} الجد ولية أي أن الفروق بين التكرارات النظرية والفعلية هي فروق جوهرية وبالتالي نرفض فرضية العدم، أي توجد علاقة بين الحالة المدنية للشخص ومدى ادخاره.

حالة خاصـــــــــة:

وهو نموذج خاص من اختبار كا^{٢} ونطلق عليه فحص انحراف نتائج الملاحظة التجريبية عن المتوقعة أي في حالة وجود صف واحد.

مثـــــــال: هذا الجدول أخد من دراسة ميدانية لباحث يبحث عن العلاقة بين تغيب العمال حسب مكان عملهم أثناء الورديات.

مكان العمل	عامل بالبلدية	في قطاع البناء	في القطاع الصناعي	في قطاع الخدمات	في القطاع التقليدي	تاجر	فلاح	المجموع
عـدد الأيـام	٦٩	٣٨	٣٣	٩٠	١١٥	١٣٨	١٤٥	٦٢٨

هنا نفترض أن النتائج تتوزع باحتمالات متساوية وهذه الحالة موجودة في اختبار كا^{٢} أي تكون عدد الأيام المغيبة متساوية لحالات مكان العمل أي $\frac{٦٢٨}{٧}$ = ٨٩،٧١ وهذا ما نسميه بالتكرارات النظرية كَ.

ك	$\bar{ك}$	$(ك - \bar{ك})$	$(ك - \bar{ك})^{٢}$	$\frac{(ك - \bar{ك})^{٢}}{\bar{ك}}$
٦٩	٨٩،٧١	– ٢٠،٧١	٤٢٨،٩٠	٤،٧٨
٣٨	٨٩،٧١	– ٥١،٧١	٢٦٧٣،٩٢	٢٩،٨٠
٣٣	٨٩،٧١	– ٥٦،٧١	٣٢١٦،٠٢	٣٥،٨٥
٩٠	٨٩،٧١	٠،٢٩	٠،٠٨٤	٠،٠٠٠٩
١١٥	٨٩،٧١	٢٥،٢٩	٦٣٩،٥٨	٧،١٣
١٣٨	٨٩،٧١	٤٨،٢٩	٢٣٣١،٩٢	٢٥،٩٩
١٤٥	٨٩،٧١	٥٥،٢٩	٣٠٥٦،٩٨	٨٩،٧١
				١٩٣،٢٦

درجة الحرية = (عدد الأعمدة – ١) = (٧ – ١) = ٦.

كا^{٢} الجدولة عند درجة حرية ٦ نسبة دلالة ٥ % مثلا تقدربـــ ١٢،٥٩٢

كا^{٢} المحسوبة > كا^{٢} الجدولة فهناك فروق جوهرية ونرفض فرضية العدم أي أن هناك علاقة بين نوعية العمل وعدد الأيام التي تم الغياب فيها.

كما يمكننا استعمال تصبح ينس في هذا المثال لزيادة احتمال أن تكون $كا^{٢}$ المحسوبة ناتجة عن قوى الحظ والصدفة.

جدول $كا^{2}$

درجة الحرية \ α	0,90	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
١	0,0158	0,455	1,074	1,642	2,706	3,841	5,412	6,635	10,827
٢	0,211	1,386	2,408	3,219	4,605	5,991	7,824	9,210	13,815
٣	0,584	2,366	3,665	4,642	6,251	7,815	9,837	11,345	16,366
٤	1,064	3,357	4,878	5,989	7,779	9,488	11,668	13,277	18,467
٥	1,610	4,351	6,064	7,289	9,236	11,070	13,388	15,086	20,515
٦	2,204	5,348	7,231	8,558	10,645	12,592	15,033	16,812	22,457
٧	2,833	6,346	8,383	9,803	12,017	14,067	16,622	18,475	24,322
٨	3,490	7,344	9,524	11,030	13,362	15,507	18,168	20,090	26,125
٩	4,168	8,343	10,656	12,242	14,684	16,919	19,679	21,666	27,877
١٠	4,865	9,342	11,781	13,442	15,987	18,307	21,161	23,209	29,588
1١	5,578	10,341	12,899	14,631	17,275	19,675	22,618	24,725	31,264
١٢	6,304	11,340	14,011	15,812	18,549	21,026	24,054	26,217	32,909
١٣	7,042	12,340	15,119	16,985	19,812	22,362	25,472	27,688	34,528
١٤	7,790	13,339	16,222	18,151	21,064	23,685	26,873	29,141	36,123
١٥	8,547	14,339	17,322	19,311	22,307	24,996	28,259	30,578	37,697
١٦	9,312	15,338	18,418	20,465	23,542	26,296	29,633	32,000	39,252
١٧	10,085	16,338	19,511	21,615	24,769	27,587	30,995	33,409	40,790
١٨	10,865	17,338	20,601	22,760	25,989	28,869	32,346	34,805	42,312
١٩	11,651	18,338	21,689	23,900	27,204	30,144	33,687	36,191	43,820
٢٠	12,443	19,337	22,775	25,038	28,412	31,410	35,020	37,566	45,315
٢١	13.240	20,337	23,858	26,171	29,615	32,671	36,343	38,932	46,797
٢٢	14,041	21,337	24,939	27,301	30,813	33,924	37,659	40,289	48,268
٢٣	14,848	22,337	26,018	28,429	32,007	35,172	38,968	41,638	49,728
٢٤	15,659	23,337	27,096	29,553	33,196	36,415	40,270	42,980	51,179
٢٥	16,473	24,337	28,172	30,675	34,382	37,652	41,566	44,314	52,620
٢٦	17,292	25,336	29,246	31,795	35,563	38,885	42,856	45,642	54,052
٢٧	18,114	26,336	30,319	32,912	36,741	40,113	44,140	46,963	55,476
٢٨	18,939	27,336	31,391	34,027	37,916	41,337	45,419	48,278	56,893
٢٩	19,768	28,336	32,461	35,139	39,087	42,557	46,693	49,693	58,302
٣٠	20,599	26,336	33530	36,250	40,256	43,773	47,962	47,962	59,703

. مطالعة ممتعة : محمد عموش .▲

تمرينات:

التمرين الأول:

ليكن الجدول التالي الذي يوضح توزيع مجموعة متكونة من ١٤١ حسب الجنس والرأي في موقف ما.والمطلوب إيجاد كا^{٢} عند نسبة ٥٪.

الرأي / الجنس	موافق جدا	موافق	لا أدري	أرفض	أرفض جدا	المجموع
ذكور	٥	٣٧	١٣	٢٨	٥	٨٨
إناث	٣	١٧	٨	٢٠	٥	٥٢
المجموع	٨	٥٤	٢١	٤٨	١٠	١٤١

الحل:

أولا:حساب $\overline{ك}$ النظرية فيما يخص الذكور:

موافق جدا = $\frac{8 \times 88}{141}$ =٤,٩٩

موافق= $\frac{54 \times 88}{141}$ =٣٣,٧٠

لا أدري: $\frac{21 \times 88}{141}$ =١٣,١١

ارفض: $\frac{48 \times 88}{141}$ =٢٩,٩٦

ارفض جدا: $\frac{10 \times 88}{141}$ =٦,٢٤

ثانيا: حساب $\overline{\text{ك}}$ النظرية فيما يخص الإناث:

موافق جدا = $\frac{8 \times 53}{141}$ = ٣,٠١

موافق= $\frac{54 \times 53}{141}$ = ٢٠,٣٠

لا أدري= $\frac{21 \times 53}{141}$ = ٧,٩٠

أرفض = $\frac{48 \times 53}{141}$ = ١٨,٠٤

أرفض جدا = $\frac{10 \times 53}{141}$ = ٣,٧٦

ك	$\overline{\text{ك}}$	ك - $\overline{\text{ك}}$	|ك - $\overline{\text{ك}}$| - ٠,٥	(|ك - $\overline{\text{ك}}$| - ٠,٥)^2	$\frac{\left(\left(\text{ك} - \overline{\text{ك}}\right) - 0.5\right)^2}{\overline{\text{ك}}}$
٥	٤,٩٩	٠,٠١	-٠,٤٩	٠,٢٤	٠,٠٥
٣٧	٣٣,٧٠	٣,٣	٢,٨	٧,٨٤	٠,٢٣
١٣	١٣,١٠	-٠,١١	٠,٤	٠,١٦	٠,٠١
٢٨	٢٩,٩٦	-١,٩٦	١,٤٦	٢,١٣	٠,٠٧
٥	٦,٢٤	-١,٢٤	٠,٧٤	٠,٥٥	٠,٠٩
٣	٣,٠١	-٠,٠١	٠,٤٩	٠,٢٤	٠,٠٨
١٧	٢٠,٣١	٣,٣	٢,٨	٧,٨٤	٠,٣٩
٨	٧,٩٠	٠,١	٠,٤	٠,١٦	٠,٠٢
٢٠	١٨,٠٤	١,٩٦	١,٤٦	٢,١٣	٠,١٢
٥	٣,٧٦	١,٢٤	٠,٧٤	٠,٥٥	٠,١٥

حساب درجة الحرية (ن- ١)(ن$_2$ – ١) = (٢ – ١)(٥ – ١) = ٤

كا^2 الجدولية عند نسبة ٥٪ = ٩,٤٨

كا^2 المحسوبة = $\frac{\left(\left(\text{ك} - \overline{\text{ك}}\right) - 0.5\right)^2}{\overline{\text{ك}}}$ = ١,٢١

إذن كا^2 المحسوبة ١,٢١ < كا^2 الجدولية ٩,٤٨

أي أن الفروق ظاهرية بين التكرارات الفعلية و النظرية و بالتالي راجع ذلك للصدفة أي نقبل فرضية الاستقلال فلا توجد علاقة بين النسب و الموقف.

التمرين الثاني:

لنفرض أن طالبا أراد معرفة رأي ٤٣ امرأة عاملة في مؤسسة معينة و هي عينة عشوائية.و عند سؤال كل واحدة منهن حول مدى موافقتها على تنظيم النسل وجد ٣٠ امرأة توافقن على ذلك و ١٣ لا توافق.

و السؤال هل يمكن أن تكون هذه التكرارات ناشئة من مجتمع إحصائي حيث الآراء مقسمة حول هذا الموضوع بشكل متساوي.

الحل:

إن فرصة العدم في هذه الحالة هي أن الآراء مقسمة ٥٠٪ توافق و ٥٠٪ لا توافقن، علما أن هذه الفرضية قد أخذت بصورة تحكيمية (أي بتساوي) إن التكرارات النظرية أو المتوقعة $\overline{ك}$ هي $\frac{43}{2}$ = ٢١,٥

ك	$\overline{ك}$	ك - $\overline{ك}$	| ك - $\overline{ك}$ | - ٠,٥	(| ك - $\overline{ك}$ | - ٠,٥)^{٢}	$\frac{\left(\left(ك - \overline{ك}\right) - 0.5\right)^2}{\overline{ك}}$
٣٠	٢١,٥	٨,٥	٨	٦٤	٢,٩٨
١٣	٢١,٥	- ٨,٥	٨	٦٤	٤,٩٢
٤٣					**٥,٦٩**

درجة الحرية = (عدد الأعمدة - ١) = (٢- ١) =١

كا^{٢} الجدولية عند درجة حرية ١ نسبة ٥٪ = ٣,٨٤١

∴ كا^{٢} المحسوبة ٥,٦٩ > كا^{٢} الجدولية ٣,٨٤١

و بهذا فإننا لا نستطيع الثقة بالفرضية القائلة بانقسام آراء هؤلاء النساء بنسبة ٥٠٪ توافقن على وسائل منع الحمل و ٥٠٪ منهن لا توافقن عليه.

التمرين الثالث:

لنفرض أن إعارة الكتب بمكتبة علم الاجتماع في الأسبوع كانت كالأتي:

الأيام	عدد الكتب
السبت	٦٠
الأحد	٤٠
الاثنين	٥٠
الثلاثاء	٨٠
الأربعاء	٩٠
الخميس	١٠٠
المجموع	٤٢٠

فهل يوجد فرق بين عدد الكتب المعارة في أيام الأسبوع في هذه المكتبة عند نسبة ٥٪.

__الحل:__

__أولا:__ حساب درجة الحرية(عدد الصفوف – ١) =(٦– ١)=٥

__ثانيا:__ حساب التكرارات النظرية $\overline{ك} = \frac{420}{6} = ٧٠$

ك	$\overline{ك}$	ك – $\overline{ك}$	| ك – $\overline{ك}$ | – ٠,٥	(| ك – $\overline{ك}$ | – ٠,٥)^{٢}	$\frac{\left(\left(ك - \overline{ك}\right) - 0.5\right)^2}{\overline{ك}}$
٦٠	٧٠	– ١٠	٩,٥	٩٠,٢٥	١,٢٩
٤٠	٧٠	– ٣٠	٢٩,٥	٨٧٠,٢٥	١٢,٤٣
٥٠	٧٠	– ٢٠	١٩,٥	٣٨٠,٢٥	٥,٤٣
٨٠	٧٠	١٠	٩,٥	٩٠,٢٥	١,٢٩
٩٠	٧٠	٢٠	١٩,٥	٣٨٠,٥	٥,٤٣
١٠٠	٧٠	٣٠	٢٩,٥	٨٧٠,٢٥	٣٨,٣٢

كا^{٢} الجدولية عند درجة حرية ٦ و نسبة ٥٪ = ١١,٠٧

إذن كا^2 الجدولية < من كا^2 المحسوبة ١١,٠٧ < ٣٨,٣٢

إذن هناك فرق جوهري بين التكرارات الفعلية و النظرية فنرفض فرضية العدم أي هناك علاقة بين المتغيرين.فلا يوجد استقلال بين عدد الكتب المعارة في اليوم حسب أيام الأسبوع.

التمرين الرابع:

لدينا الجدول التالي الذي يمثل العلاقة بين الأصل الجغرافي و الحالة المدنية للأفراد.

اختبر فرضية الاستقلال أي أنه ليس هناك علاقة بين الأصل الجغرافي و الحالة المدنية عند نسبة ١٪.

الحالة المدنية / الأصل	أعزب	مطلق	متزوج	المجموع
ريفي	٤	٣	٤٢	٤٩
حضري	١٢	–	٢٣	٣٥
المجموع	١٦	٣	٦٥	٨٤

<u>الحل:</u>

أولا _ حساب $\bar{ك}$ النظرية بالنسبة للريفين

أعزب = $\frac{49 \times 16}{84}$ = ٩,٣٣

مطلق = $\frac{49 \times 3}{84}$ = ١,٧٥

متزوج = $\frac{49 \times 55}{84}$ = ٣٧,٩١

ثانيا _ حساب $\bar{ك}$ النظرية بالنسبة للحضريين

أعزب = $\frac{35 \times 16}{84}$ = ٦,٦٦

مطلق = $\frac{35 \times 3}{84}$ = ١,٢٥

متزوج = $\frac{35 \times 65}{84}$ = ٢٧,٠٨

ثالثا _ حساب أن درجة الحرية = (ن١ – ١)(ن٢ – ١) = (٢ – ١)(٣ – ١) =٢

ك	$\overline{ك}$	ك – $\overline{ك}$	ا ك – $\overline{ك}$ ا – ٠,٥	(ا ك – $\overline{ك}$ ا – ٠,٥)^{٢}	$\frac{\left((ك - \overline{ك}) - 0.5\right)^2}{\overline{ك}}$
٤	٩,٣٣	– ٥,٣٣	٤,٨٣	٢٣,٣٣	٢,٥
٣	١,٧٥	١,٢٥	٠,٧٥	٠,٥٦	٠,٣٢
٤٢	٣٧,٩١	٤,٠٩	٣,٥٧	١٢,٨٩	٠,٣٤
١٢	٦,٦٦	٥,٣٤	٤,٨٤	٢٣,٤٢	٣,٥٢
–	١,٢٥	– ١,٢٥	٠,٧٥	٠,٥٦	٠,٤٥
٢٣	٢٧,٠٨	– ٤,٠٨	٣,٥٨	١٢,٨٢	٠,٤٧
					٧,٦

عند درجة حرية = ٢ و نسبة دلالة ١٪ كا^{٢} الجدولية = ٩,٢١٠

∴ كا^{٢} الجدولية ٩,٢١٠ > كا^{٢} المحسوبة ٧,٦

∴ فالفروق جوهرية و بالتالي نرفض فرضية العدم أي توجد علاقة بين الأصل الجغرافي و الحالة المدنية للشخص.

قائمة المراجع:

١- المراجع باللغة العربية:

١- أبوعبانة فتحي محمد. مدخل إلى التحليل الإحصائي في الجغرافيا البشرية، دار النهضة العربية، بيروت، ١٩٨٦.

٢- جلاطو جيلالي. الإحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الجزائر، ١٩٩٩.

٣- حليمي عبدالقادر. مدخل إلى الإحصاء، ديوان المطبوعات الجامعية، الجزائر، ١٩٩٤.

٤- خاطر أحمد مصطفى وآخرون. التحليل الإحصائي للبحوث في الخدمة الاجتماعية، المكتب الجامعي الحديث، الإسكندرية، ١٩٩٨.

٥- زايد مصطفى. الإحصاء ووصف البيانات، مطبعة الشريف، السعودية، ط٢، ١٩٨٨.

٦- عوض منصور وآخرون. أساسيات علم الإحصاء الوصفي، دار صفاء للنشر والتوزيع، عمان، ط١، ١٩٩٩.

٧- غريب محمد سيد أحمد. الإحصاء و القياس في البحث الاجتماعي، ١٩٩١.

٨- عوض منصور، هبري عزام، مبادئ الإحصاء، دار صفاء للطباعة والنشر والتوزيع، عمان، الطبعة١، ٢٠٠٠.

٩- كنجو أنيس. الإحصاء، مؤسسة الرسالة، بيروت، الجزء الأول، ١٩٨٢.

١٠- مهدلي محمد محمود. تطبيقات علم الإحصاء في العلوم الاجتماعية، الاسكندرية، ٢٠٠٢.

١١- كلاس محمد. محاضرات في الإحصاء التطبيقي، ديوان المطبوعات الجامعية، الجزائر، ١٩٩٣.

٢- المراجع باللغة الفرنسية:

1- Hamdani Hocine. Statistique descriptive et expression graphique, OPU, Alger, 1988.

2- Murray R. Spiegel. Théorie et applications de la statistique, imprimerie Louis- jean, Paris, 1972.

3- Vandeschcrick christophe, jwantelet jean- marie. De la statistique descriptive aux mesures des inégalités université catholique de Pouvain la neuve, Belgique, 2003.