وزارة التعليم العالي والبحث العلمي
كلية العلوم الاقتصادية والتجارية وعلوم التسيير
قسم العلوم الاقتصادية

مطبوعة

محاضرات في الإحصاء الوصفي

مطبوعة موجهة لطلبة السنة الأولى ل.م.د

إعداد الدكتور تومي عبد الرحمان
الأستاذ الدكتور مغاري عبد الرحمان

السنة الجامعية 2018 ـ 2019

بسم الله الرحمن الرحيم

يعتبر الاحصاء الوصفي من العلوم التي لا يمكن الاستغناء عنها في أي بحث أكاديمي كان، سواء تعلق بمعالجة ظاهرة تنتمي الى علم من علوم المادة، أو كانت تنتمي إلى العلوم الاجتماعية، لما يتضمنه من آليات وقواعد أساسية في منهجية البحث العلمي.

ومع التطور التكنولوجي الهائل، برزت الى حقل المعرفة العلمية برامج تعالج المعطيات الإحصائية للظاهرة محل الدراسة، مثل: (برنامج spss) أو (برنامج stata) أو (برنامج sas) ...الخ. بداية من عملية الجمع، مرورا بالتنظيم والترتيب، الى التحليل والحصول على المعلومات التي تحقق أهداف البحث.

بناء على ما تقدم ونظرا لكوننا ندرس هذه المادة العلمية منذ حوالي 15 سنة، رأينا من الواجب أن نضع بين أيدي طلبتنا الأعزاء وأساتذتنا الأفاضل (في مختلف تخصصات العلوم الاقتصادية) خلاصة هذه التجربة في شكل مطبوعة تحمل خمسة فصول موزعة على النحو التالي.

الفصل الأول: يتضمن مفاهيم وترتيب وتنظيم وعرض للبيانات الاحصائية مدعما بأمثلة وتمارين. أما الفصل الثاني فقد جاء تحت عنوان ، مقاييس النزعة المركزية ، حيث نبرز فيه كل الطرق الاحصائية لقياس النزعة المركزية بأسلوب سلس سهل وبأمثلة واقعية ، لينتهي بمجموعة من التمارين.

أما الفصل الثالث، يأتي تحت عنوان مقاييس التشتت والشكل، ليعالج معطيات الظاهرة محل الدراسة من حيث ضعف أو قوة تشتتها وقياس قيمة الانحراف ونوع الالتواء...الخ.

بينما جاء الفصل الرابع ليعالج العلاقة بين المتغيرات العشوائية، المستقل منها والتابع، ثم تحديد قوة الارتباط، لينتهي بصلاحية النموذج، مع تدعيم الفصل بتمارين محلولة وغير محلولة.

تختم المذكرة بالفصل الخامس الموسوم بالأرقام القياسية ، مع مختلف صيغ الحساب. وقد تعمدت في إضافته لتعميم الفائدة ، على الرغم من عدم ادراج الارقام القياسية في بعض البرامج المقررة.

وكما هو الحال، يبقى عملنا جهد مقل، فإن أصبنا فذلك فضل من الله ونعمة منه وإن كان غير ذلك، نلتمس العذر لكل من أطلع على نقاط ضعفنا.

نرجو في الختام أن يجد كل مهتم بالإحصاء الوصفي ضالته والحمد لله أولا وآخرا.

الإحصاء الوصفي (الايضاحي، البياني)

شيء من التاريخ

ذكر الإحصاء في التاريخ البابلي (3000 سنة قبل الميلاد)، كما ذكر في التاريخ الصيني (2000 سنة ق. م)، وفي مصر القديمة (1700 سنة ق . م)، وقد نبه القرءان الكريم إلى الفرق الواضح بين مصطلح "العد " أي الحساب، ومصطلح "الإحصاء " في قوله تعالى (وآتاكم من كل ما سألتموه وإن تعدوا نعمة الله لا تحصوها إن الإنسان لظلوم كفار) الآية 34 من سورة ابراهيم. مما يدل على أن الإحصاء هو مرحلة تأتي بعد العد، تحتاج إلى عمليات تعرف الآن بالجمع والترتيب والجدولة والرسوم البيانية، والطرق الاحصائية الرياضية، والتحليل ثم النتائج وما يتبعها من تفسير وقرارات، وهي مراحل يمكن أن نطلق عليها اصطلاحا بالمنهج الاحصائي.

يعتبر علم الإحصاء بمثابة المرجع الذي لا يستغنى عنه أي علم آخر، لما يتضمنه من آليات (طرق) تساعد الباحث على تحليل الظاهرة محل الدراسة، لذلك نجد استعمالاته في العلوم الطبيعية تماما مثلما يستعمل في العلوم الانسانية (منها الاقتصاد).

فلا الصناعة، ولا الفلاحة ولا قطاع الخدمات يمكنها الاستغناء عنه، وبعبارة مبسطة، هو علم متميز بمنهجه، يقدم خدماته للعلوم التجريبية مثلما يقدمها للعلوم الانسانية (علم الاجتماع، النفس، الاقتصاد، التاريخ، الثقافة، السياسة ...الخ) .

يعرف هذا المنهج بطرقه الاحصائية المتعددة، بداية من مرحلة الوصف، مرورا بالتنبؤ، إلى مرحلة اتخاذ القرار. وعلى هذا الأساس، فإن الجزء المتعلق بجمع المعطيات وتلخيصها نسميه عادة "الإحصاء الوصفي ".

I ـ مصطلحات أساسية

01 ـ المجتمع الإحصائي

هو أي مجتمع (بشر، حيوان، طيور، نبات، أشياء مادية...الخ) يكون محل الدراسة، بحيث يشترك جميع أفراده في صفة أو أكثر.

مثال: طلبة السنة الأولى اقتصاد في الجزائر، يشتركون في أكثر من صفة يمكن أن تكون موضوع دراسة (المستوى العلمي، السن، القامة، الوزن، الجنسية، الحالة العائلية، الدخل، التحصيل العلمي، ظاهرة الغياب عن المحاضرة، ...الخ)، يرمز للمجتمع الإحصائي عادة بالرمز(N).

02 ـ العينة الاحصائية

هي جزء اخترناه من المجتمع محل الدراسة، نلجأ إليها حينما يتعذر علينا حصر كل المعطيات (البيانات) المتعلقة بالظاهرة (الصفة) المراد دراستها، يرمز لها عادة بالرمز(n) ونميز فيها نوعين أساسيين:

2– 1 ـ عينات احتمالية

يقصد بها أن كل وحدة من وحدات العينة لها الفرصة لأن تكون جزء من الدراسة، بناء على عمليات السحب.

هذا النوع من العينات ينقسم إلى ثلاثة (03) أقسام.

2 ـ 1 ـ 1 ـ العينات العشوائية: تتصف بتجانس قيمها (بياناتها) وتساو في فرص احتمالية السحب بالنسبة لوحداتها، سواء كان هذا السحب بالطريقة الكلاسيكية العشوائية، أو بالأرقام العشوائية عن طريق الكومبيوتر.

2 ـ 1 ـ 2 ـ العينة المنتظمة: إذ أن نقطة الاختيار في البداية تكون بشكل عشوائي، ثم إتمام باقي العملية بشكل منتظم.

2 ـ 1 ـ 3 ـ العينة الطبقية: تستعمل في مجتمع إحصائي متنوع، مثل فئات أجور العاملين، وفئات الأعمار للسكان، أو كأنك تقول لدينا 100 طالب(ة) منهم 60 % إناث يتوزعن على التخصصات التالية: 25 % اقتصاد، 25 % إدارة، 40 % محاسبة وهكذا ...

2– 2 ـ عينات غير احتمالية

في حالة عدم توفر مصادر المعطيات لدى الباحث، أو ضعف البرامج الاحصائية، أو عدم قدرة الباحث على الوصول إلى العينات الاحتمالية، يلجأ إلى مثل هذا النوع، كالاستبيان والمقابلة وسبر الآراء ...الخ . غالبا ما نجد في مثل هذا النوع ثلاثة آليات نوجزها على النحو التالي:

2 ـ 2 ـ 1 ـ طريقة القطعة: إذ تركز أساسا على الاستبيان والمقابلة لمجموعة محددة في مكان ـ ما ـ

2 ـ 2 ـ 2 ـ طريقة الحصة: وهي مقابلة لأشخاص نعتقد أنهم أفضل الوحدات الاحصائية للبحث، غير أن هذه الطريقة لا تصلح عند تعميم النتائج.

2 ـ 2 ـ 3 ـ طريقة الملائمة: أو ما يسمى بعينة "الفرصة" أو العينة الانتزاعية، وهي تعني أنك تقابل شخص ليدلك على شخص آخر، وهكذا ... إلى أن تصل إلى الأخير، لذلك تسمى أحيانا بطريقة كرة الثلج.

والخلاصة، أن العينات غير الاحتمالية لا تعطي صورة واقعية عن الدراسة، كما أن العينة بشكل عام أخذت لتمثل المجتمع الإحصائي أحسن تمثيل، وإلا تعتبر متحيزة لجزء ـ ما ـ من ذلك المجتمع.

03 ـ الوحدة الاحصائية

تسمى أيضا الفرد، وهي أي عنصر من العناصر المشكلة للمجتمع أو العينة محل الدراسة، مثل مجتمع المؤسسات الصغيرة والمتوسطة، إن كل مؤسسة تعتبر فردا أو عنصرا أو هي وحدة إحصائية، يرمز لها عادة بالرمز n_i.

04 ـ الصفة

كما تسمى أيضا بالخاصية، هي قابلة للتغير في القيمة التي تحملها من وحدة إحصائية إلى أخرى، ويلاحظ أن معظم الصفات البيولوجية، هي صفات متغيرة عشوائيا، لذلك نسمي الصفة المتغيرة بالمتغير العشوائي، خذ على سبيل المثال: قامة الإنسان، هي صفة بيولوجية تتغير من شخص إلى آخر (الوزن، الحجم، المساحة، اللون، الحالة العائلية ...الخ).

يرمز لها عادة بالرمز x أو أي رمز آخر . يمكن أن تخضع أية صفة متغيرة للدراسة الاحصائية، ولذلك تسمى أحيانا بالظاهرة المدروسة.

4 ـ 1 ـ الصفة بين الوسم أو العد أو القياس:

كل وحدة إحصائية من أفراد المجتمع أو العينة تحمل صفة على الأقل مشتركة مع باقي الوحدات، قابلة للدراسة.

مثال: أخذنا عينة من طلبة السنة الأولى اقتصاد (n = 10) لدراسة العلاقة بين الوزن X والطول L، تلاحظ أن أي طالب n_i يشترك مع باقي الطلاب في صفتي الوزن والطول . عادة نميز في الصفة (المتغير) محل الدراسة صنفين أساسيين :

4 ـ 1 ـ 1 ـ صفات كيفية غير قابلة للقياس (بأية وسيلة قياس متداولة عالميا)

وهي كثيرة في حياتنا العملية، كاللون، الجنس، الحالة العائلية، العلامات التجارية ...الخ . فهي كلها غير قابلة للقياس، لذلك يلجأ في البحوث العلمية إلى وسمها (تعليمها) قبل حسابها وإخضاعها إلى طرق التحليل. هذا الصنف من المتغيرات (الصفات) ينقسم بدوره إلى قسمين.

4 ـ 1 ـ 1 ـ 1 ـ صفات كيفية اسمية

وهي كذلك لأنها غير قابلة للترتيب المنطقي أو التنظيم (بمعنى، حتى لو أنت رتبتها يكون ترتيبك لها بدون معنى).

مثال: عينة من طلبة كلية الطاقة ($n = 10$) يحملون الجنسيات التالية .

الجنسية	الجزائرية	التونسية	الليبية	المغربية	الموريتانية	النيجرية
ع . الطلبة	02	01	03	01	02	01

لا يوجد أساس منطقي يمكن أن ترتب الجنسية على ضوئه.

تلاحظ بأن الجنسية صفة غير قابلة للقياس، قمنا بوسمها، مرة جزائرية وأخرى تونسية ...وهكذا

4 ـ 1 ـ 1 ـ 2 ـ صفات كيفية رتيبة (ترتيبية)

أي أنها صفات قابلة للترتيب المنطقي، وهي غير قابلة للقياس كما ذكرنا سابقا.

مثال: عينة تتكون من ($n = 20$) طالب تحصلوا في مادة الإحصاء للسداسي الأول على الدرجات التالية :

xi	ممتاز جدا	مقبول	حسن	ممتاز	جيد	ضعيف	جيد جدا	$\sum ni$
ni	07	03	02	03	01	02	02	20

يمكن ترتيب هذه الدرجات بشكل منطقي ومتعارف عليه في الأوساط العلمية على النحو التالي:

xi	ممتاز جدا	ممتاز	جيد جدا	جيد	حسن	مقبول	ضعيف	$\sum ni$
ni	07	03	02	01	02	03	02	20

4 – 1 – 2 ـ صفات كمية

وهي صفات يمكن قياسها وعدها في نفس الوقت . نميز فيها نوعان :

4 – 1 – 2 – 1 ـ صفات كمية منقطعة (منفصلة)

وهي كل صفة (متغير) قابلة للعد، تأخذ قيم صحيحة (غير كسرية)، وتتغير بوحدات كاملة، لذلك تسمى أحيانا بالصفات الوثابة، مثل عدد الفنادق في الجزائر العاصمة، لا نقول عددها 125.4 بل الأصح 125 فندق، أو عدد أفراد الأسرة لا نقول يوجد بها 6.12 فردا، بل الأصح 6 أفراد.

4 – 1 – 2 – 2 ـ صفات كمية مستمرة

وهي كل صفة (متغير) قابلة للقياس والعد، حيث تكون قيمها المشاهدة (المسجلة) على شكل أعداد كسرية ضمن مجال معين (على شكل فئات).

مثل: قامة الإنسان ضمن مجال { 170 — 180 } سم . سوف نجد في الحياة العملية ملايين البشر الذين تتراوح أطوالهم ضمن هذا المجال . خد على سبيل المثال ظاهرة (صفة) العمر من صفر سنة إلى خمس سنوات، سوف تكتشف من له سنة وأربعة أشهر الملايين من الأطفال، وعليه فأنت أمام أعداد كسرية لا نهاية لها.

ملاحظة هامة

تعتبر التعاريف السابقة نسبية، فإذا أخذنا على سبيل المثال متغير السن كظاهرة محل الدراسة، نظريا هو متغير كمي مستمر، لكن في الواقع العملي يقاس ضمن أفضل حالات اقرب يوم، كما أن المتغير الكمي المنقطع، إذا كان عدد وحداته الاحصائية أكثر من ثلاثين ($n > 30$)، يفضل تنظيمها على شكل فئات.

05 ـ القيمة الإحصائية

لقد رأينا بأن الصفة محل الدراسة يمكن أن نسميها بالظاهرة أو المتغير، حينما نخضعها إلى العد أو القياس أو الوسم (حسب طبيعتها) نحصل على قيم نسميها بالقيم الاحصائية، أو البيانات أو المشاهدات، وسلسلة البيانات المتحصل عليها تسمى عادة بالسلسلة الاحصائية، وهي عموما تسجل في جدول يطلق عليه "جدول المعطيات".

مثال : نريد دراسة علامة الطلبة (X) من خلال العينة التالية، وذلك بعد تصحيح الأستاذ لأوراق امتحان الإحصاء للسداسي الأول، فكانت النتائج على النحو التالي:

جدول رقم 01: يبين علامات لعينة من طلبة السنة الأولى إحصاء تتعلق بالسداسي الأول

الطلبة (ni)	$n1$	$n2$	$n3$	$n4$	$n5$	$\sum ni$
العلامة (xi)	12	10	17	05	02	05

المصدر: معطيات افتراضية من طرفنا

ـ الصفة محل الدراسة (المتغير): علامة الطلبة (X)

ـ العينة محل الدراسة: $ni = 05$

ـ الوحدة الإحصائية: هي أي طالب من مجموع الطلبة $\sum ni = n1\ ,n2,\ n3,\ \ n4,\ n5$

ـ القيمة الإحصائية: هي مثلا $x3 = 17$

ـ جدول المعطيات كما ترى يحتاج إلى ترقيم متسلسل، وعنوان للإيضاح ومصدر المعطيات.

06 ـ الكيفيات ومجال التغير

نسمي الطرق الممكنة للمتغير محل الدراسة بالكيفيات، ومختلف هذه الكيفيات يطلق عليها "منطقة" أو مجال التغير.

مثال: نريد دراسة صفة (متغير) الجنس البشري، حينها تكون كيفيات المتغير محل الدراسة هي:

ـ الذكر يرمز له بالرمز (M) والأنثى بالرمز (F)

ـ وبالتالي مجال المتغير هو $[\ ;F]$

مثال آخر: نريد دراسة متغير الحالة الاجتماعية، هذه الصفة لها الكيفيات (متزوج، أعزب، مطلق، أرمل)، وبالتالي مجال المتغير محل الدراسة هو $[m\ ,\ c\ ,\ d\ ,\ v\ ]$.

07 ـ أمثلة محلولة

المثال الأول

يهمنا دراسة ظاهرة الخصوبة في علاقتها مع بعض المؤشرات السوسيواقتصادية الجزائرية، لعينة من الولايات على النحو التالي.

(%)

البيان / الولايات	مؤشر الخصوبة	الفلاحة	التعليم	البدو	وفاة الأطفال
سطيف	80.20	47.00	52.00	10.20	20.50
قالمة	83.70	70.00	45.00	15.00	18.00
تلمسان	90.40	45.70	60.70	29.00	24.00
تيزي و ز و	60.50	60.50	70.40	35.00	21.00
ورقلة	94.20	25.90	40.00	25.00	22.00

السؤال

حدد كل من المتغير محل الدراسة، نوع المتغير، المجتمع، العينة، الوحدة الإحصائية، القيمة الإحصائية، الكيفيات، مجال التغير.

الجواب

- المتغير محل الدراسة: متغير الخصوبة
- نوع المتغير: كمي مستمر لأن القيم كسرية غير صحيحة
- المجتمع: $N = 48$ ولاية
- العينة: $n = 05$ ولايات
- الوحدة الإحصائية: أي ولاية من الولايات الخمس، ولتكن تيزي و ز و (n_4)
- القيمة الإحصائية: مثل $xi = 70.40$
- الكيفيات: أربعة، وهي كل من الفلاحة، التعليم، البدو، وفاة الأطفال.
- مجال التغير: هو هذه الكيفيات الأربعة

المثال الثاني

نريد دراسة متوسط سعر المواد الزراعية الأساسية وتطوراتها خلال السداسي الأول لعام (2016،2015) في السوق العالمي حسب الجدول التالي:

الرقم	المواد الزراعية	متوسط السعر خلال السداسي الأول (2015)	متوسط السعر خلال السداسي الثاني(2016)	نسبة التطور %	التغيرات $ / الطن
01	القمح اللين	209.10	350.65	+ 36	+ 141.55
02	القمح الصلب	325.53	409.42	+ 26	+ 83.89
03	الذرة	270.86	321.14	+ 19	+ 50.28
04	قهوة أرابيكا	4911.48	6116.66	+ 25	+ 1205.18
05	زيت ع الشمس	580.49	706.41	+ 22	+ 125.92
06	سكر أبيض	732.60	701.70	- 04	- 30.90

السؤال: حدد كل من: المتغير محل الدراسة، نوعه، المجتمع، العينة، الوحدة الإحصائية، القيمة الإحصائية، الكيفيات، ومجال التغير.

الجواب

ـ المتغير محل الدراسة: متوسط سعر المواد الزراعية الأساسية

ـ نوع المتغير: كمي مستمر

ـ المجتمع: كل المواد الزراعية الأساسية $NI = N1 ,N2 ,N3 ,,,, Np$

ـ العينة: $n_i = 06$ وهي كل المواد الزراعية الموجودة في الجدول

ـ الوحدة الإحصائية: مثل قهوة أرابيكا (n_4)

ـ القيمة الإحصائية: مثل 409.42 $ / طن = x_2

ـ الكيفيات: عدد كيفيات المتغير محل الدراسة أربعة، وهي كل من متوسط السعر للسداسي الأول من السنتين (2015،2016)، نسبة التطور، التغيرات.

المثال الثالث

ما يهمنا هو دراسة متغير "الحالة المدنية" يتعلق بعشرين (20) شخص أخذوا كعينة، على اعتبار الرموز التالية:

أعزب (c)، متزوج (m)، أرمل (v)، مطلق (d) حيث كانت القيم على النحو التالي:

m , m , d, c , c , m , c , c , c , m , c ; m , v , m , v , d , c , c , c , m

السؤال: نفس الأسئلة السابقة.

الجواب

ـ المتغير محل الدراسة: الحالة المدنية

ـ المجتمع المستهدف من الدراسة: هو المجتمع البشري لتلك المنطقة التي أجريت فيها الدراسة

ـ العينة محل الدراسة: 20 شخص

ـ نوع المتغير: متغير كيفي إسمي (غير رتبي)

ـ الكيفيات: هي كل من أعزب، متزوج، أرمل، مطلق

ـ مجال المتغير هو $[c, m, v, d]$

ـ الوحدة الإحصائية: أي شخص من ضمن 20 شخص (n_i)

ـ القيمة الإحصائية: لا يمكن الحصول عليها من خلال القياس أو الحساب، بل نحصل عليها من خلال وسم (تعليم) كل شخص بالحالة المدنية التي تناسبه (هل هو أعزب أم متزوج ...الخ).

نفرض أن المتغير محل الدراسة ـ الحالة المدنية ـ هو (X)

وبالتالي القيم الإحصائية الرمزية هي: $x1, x2, x3, x4,,,,,,,,,, xn$ وعل هذا الأساس يمكن الحصول عليها بناء على معطيات المثال بالترتيب من اليسار إلى اليمين كما يلي:

$x1 = m, x2 = m, \quad x3 = d, \quad x4 = c, x5 = c, x6 = m, x7 = c,,,,$

وهكذا إلى غاية : $x_{20} = m$

تمرين :

ما يهمنا دراسة تدفق المغتربين الجزائريين نحو وطنهم خلال 35 يوما، ابتداء من 05 جويلية إلى 09 أوت من عام 2015 . حيث كان عدد الوافدين خلال كل يوم على النحو التالي.

1160، 980، 850، 700، 320، 500، 950، 1000، 1100، 1120، 890، 980، 290، 500، 700، 850، 950، 1000، 350، 890، 850، 850، 490، 700، 1250، 1250، 1200، 1300، 600، 700، 700، 400، 800، 890، 850 .

السؤال: حدد ما يلي:

الظاهرة محل الدراسة (المتغير)، نوعها، المجتمع، العينة، الوحدة الإحصائية، القيمة الإحصائية، كيفيات المتغير، مجال المتغير، ما هو العدد الذي تكرر أكثر من غيره، ما هي نسبته ؟

08 ـ التوزيع

يعبر عن العلاقة بين قيم الصفة (المتغير) محل الدراسة وتكرارها في المجتمع (أو العينة) محل الدراسة، حيث يمكننا من حساب الوزن النسبي لأية قيمة من القيم بالنسبة إلى مجموع كل قيم وحدات المجتمع (أو العينة).

مثال: نفرض لدينا عينة من الطلبة ($n_i = 10$) تبين معدلات السداسي الأول حسب الجدول التالي:

.

ni رقم الطالب	$n1$	$n2$	$n3$	$n4$	$n5$	$n6$	$n7$	$n8$	$n9$	$n10$	10
xi المعدل	10	12	12	10	09	14	12	13	09	12	$\sum_{i=1}^{i=10} ni$

من خلال الجدول ترى معي بأن المعدل السائد لهذه العينة هو 12 / 20 حيث تكرر 04 مرات وبالتالي

وزنه النسبي هو: $\frac{4}{10} \times 100 = 40\ \%$

09 ـ السلسلة الإحصائية

هي مجموع معطيات رقمية مرتبة حسب معيار أو عدة معايير (الطول أو العمر بالنسبة للأشخاص ورقم الأعمال بالنسبة للمؤسسات على سبيل المثال)، فيما إذا كان الزمن (السنة، الفصل، الشهر) أحد هذه المعايير، عندها نكون أمام دراسة السلاسل الزمنية أو التاريخية.

مثال: عينة تتكون من عشرة (10) طلبة، كانت القامة عندهم على النحو التالي: 160 سم ، 165، 170، 174، 180، 173، 160، 175، 170، 179 سم. إن معطيات الطول لهذه العينة من الطلبة تشكل سلسلة إحصائية، يمكن ترتيبها تصاعديا أو تنازليا، كما يمكن تنظيمها في جدول.

II ـ جمع وتبويب وعرض البيانات الإحصائية

01 ـ جمع البيانات الإحصائية

يقصد بجمع المعطيات (المشاهدات أو القيم أو البيانات) الإحصائية، هو الحصول على بيانات رقمية أو وصفية تتميز بالدقة، حول الظاهرة المراد دراستها، ولتحقيق هذا الهدف، على الباحث أو الدارس التوجه إلى مصادر هذه المعطيات.

1 ـ 1 ـ مصادر جمع البيانات: تتفق كتب الإحصاء في معظمها على مصدرين:

1 ـ 1 ـ 1 ـ مصادر تاريخية: وهي تشمل البيانات المتوفرة في الكتب، المجلات، الجرائد، الوثائق الرسمية، مواقع الهيئات الوطنية والدولية.

يمكن أن نسمي هذا النوع من المصادر، بالمصادر غير المباشرة، حيث توفر على الباحث مشقة الجمع من الميدان، وما يترتب عن ذلك من جهد ووقت وتكاليف.

1 ـ 1 ـ 2 ـ مصادر ميدانية: أي أن البيانات تأخذ من الميدان مباشرة، سواء أكان ذلك عن طريق المشاهدة والتسجيل، مثلما يحصل في التجارب الطبية والفيزيائية، وعند دورة الإنتاج في المعامل، أو عن طريق المقابلات الشخصية أو بواسطة الاستبيان والمراسلات.

1 ـ 2 ـ طرق جمع المعطيات الإحصائية

1 ـ 2 ـ 1 ـ طريقة الحصر الشامل

أي أن الباحث يمكنه جمع كل معطيات الظاهرة من جميع وحدات المجتمع محل الدراسة.

مثال :

نريد دراسة ظاهرة تطور التعليم الجامعي، وذلك بالتركيز على الأساتذة اللذين تتوفر فيهم الكفاءة العلمية من درجة أستاذ التعليم العالي وأستاذ محاضر قسم(أ). بدون شك يمكن عد كل الأساتذة على مستوى 48 ولاية بالعودة إلى وزارة التعليم العالي، ذلك لأن العدد يمكن التحكم فيه.

1 – 2 – 2 ـ طريقة العينة

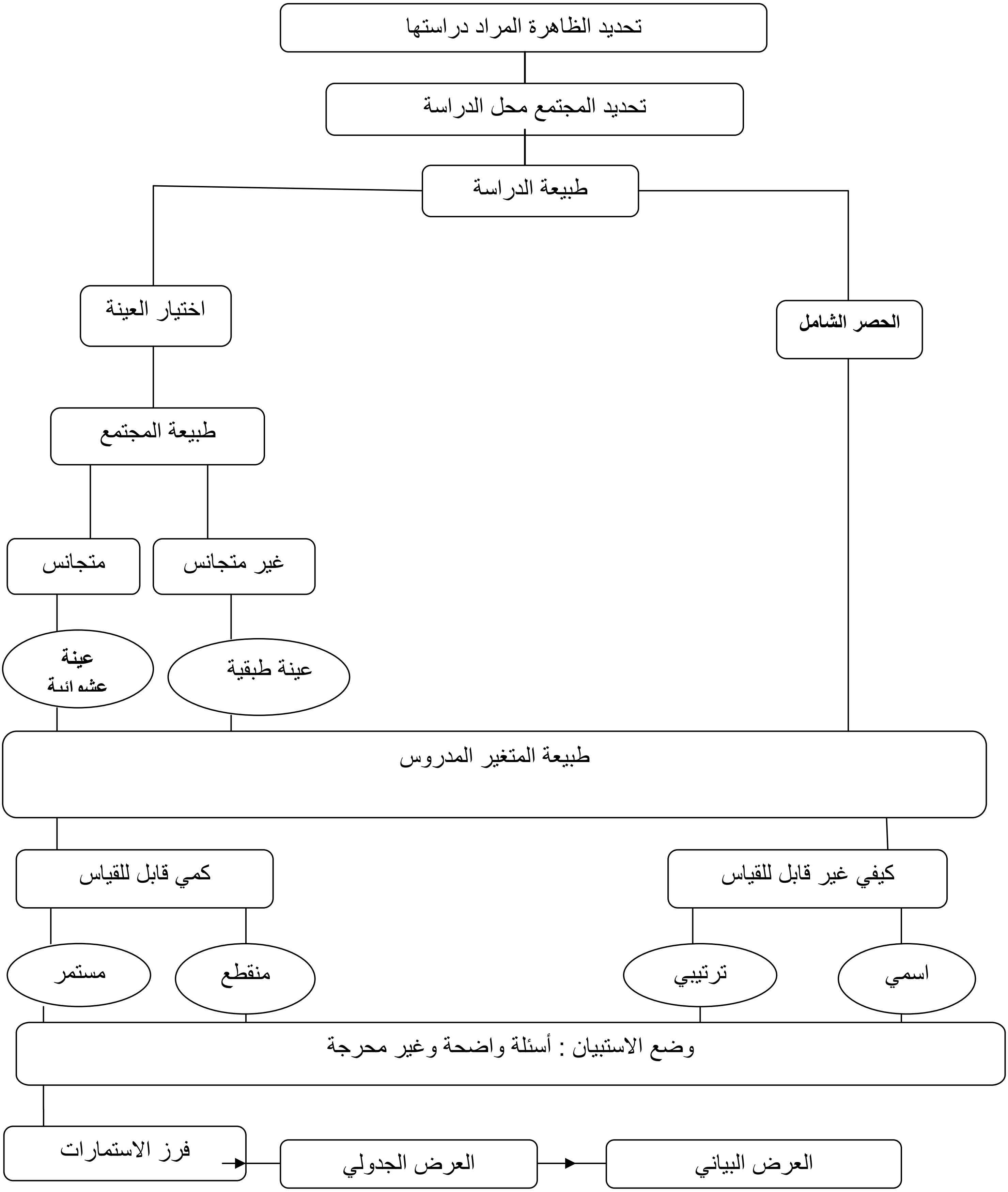
نظرا لصعوبة الحصر الشامل لكل قيم الظاهرة محل الدراسة، كثيرا ما يلجأ الباحث إلى استعمال العينة.

02 ـ تنظيم المعطيات الاحصائية

تأتي هذه الخطوة بعد تحديد الظاهرة، والمجتمع المدروس والطريقة التي تجمع بها المعطيات، ونوع العينة، وطبيعة المتغير (كيفي، كمي) ونموذج الاستبيان...الخ. يقوم الباحث بفرز الاستبيانات ثم يفرغ تلك المعطيات في جدول يطلق عليه **"جدول التفريغ"** وهو أسلوب من أساليب التبويب يساعد على تجنب الخطأ في الحساب.

تنقل البيانات بعد تفريغها إلى جداول نهائية مناسبة، تراعى فيها الدقة وطبيعة المتغير المدروس.

مخطط رقم 01 : مختلف المراحل التي يقوم بها الباحث قبل عملية العرض الجدولي



2 – 1 – الجداول الإحصائية: نميز بين نوعين من الجداول

2 – 1 – 1 – الجداول البسيطة

عادة ما تستخدم هذه الجداول في حالة العينات الصغيرة، أي عندما تكون $n < 30$ وهي في الغالب تتكون من ثلاثة أعمدة، حيث يخصص العمود الأول للأرقام، أو الأعداد والثاني للقيم (البيانات) والثالث للتكرارات.

مثال: نفرض لدينا عينة تتكون من (n = 20) أسرة لأحد أحياء الصفيح بالجزائر العاصمة، نريد دراسة متوسط عدد أفراد الأسرة لهذه العينة من خلال المعطيات المتوفرة حسب الجدول التالي:

الرقم (n_i)	ع الأسر(F_I)	ع أفراد الأسرة(x_i)
01 02 03 04	01 10 07 02	02 03 04 05
————	20	$\sum F_I$

تلاحظ من خلال الجدول بأن مجموع التكرارات يساوي تماما حجم العينة: $\sum F_I =$

2 – 1 – 2 – جداول المعطيات المبوبة

تعرف بجداول التوزيع التكراري، يلجأ إليها الباحث عند استخدام العينات الكبيرة ($n > 30$) وهي العينات الشائعة الاستعمال في دراسة العلوم الاجتماعية (مثل الاقتصاد). يختلف عدد الأعمدة فيها والأسطر حسب أهداف الاستعمال.

مثال: أخذنا عينة تتكون من (n = 50) طالب(ة) مسجلين جدد للعام الدراسي(2018/ 2019) بكلية العلوم الاقتصادية، نريد دراسة المعدل الذي تحصلوا عليه في شهادة البكالوريا بناء على الملفات المقدمة أثناء التسجيل.

المعدل xi	التكرار F_I	ت. صاعد $Fi\uparrow$	التكرار النسبي fi	نسبي مئوي f_i %
10 – 12	20	20	$\frac{20}{50} = 0.40$	40
12 – 14	12	32	$\frac{12}{50} = 0.24$	24
14 – 16	08	40	$\frac{08}{50} = 0.16$	16
16 – 18	07	47	$\frac{07}{50} = 0.14$	14
18 – 20	03	50	$\frac{03}{50} = 0.06$	60
$\sum F_I$	50	——	$\sum fi = 1$	$\sum fi\,\% = 100\,\%$

تلاحظ معي أن هذا الجدول متعدد الأهداف، نأخذ على سبيل المثال.

– المعدل الذي يساوي 10 وأقل من 12 . هذا المجال يمثل 20 طالب (ة) وهو السائد إذ أن وزنه النسبي يمثل 0.4 وهو أعلى وزن بالنسبة لباقي المعدلات، حيث يستحوذ على 40 % من مجموع الطلبة، أو من باقي مجموع المجالات.

03 – عرض البيانات الاحصائية

يتوقف عرض البيانات الاحصائية بإحدى الطريقتين، إما عن طريق الجداول كما سبق وإما بواسطة مختلف الرسومات البيانية، وهذا يتوقف على طبيعة ونوع المتغير كما يأتي.

3 – 1 – الرسم البياني لمتغير كيفي إسمي

هذا النوع من المتغيرات يقبل أي رسم من ثلاثة رسومات بيانية ممكنة هي: القطاعي والأنبوبي والعمود المجزأ.

مثال: نريد دراسة سوق السيارات في الجزائر قصد التعرف على العلامة (X) الأكثر استحواذا في السوق، فكانت المعطيات على النحو التالي:

العلامة (X)	التكرار $f_i\%$
رونو بوجو سيتروان علامات أخرى	24 31 14 31
$\sum_{i=0}^{n} f_i$	100

بعد العرض الجدولي، والذي يمكن أن يعدل إلى جدول يحتوي على التكرارات النسبية كما رأينا سابقا، ننتقل إلى العرض البياني.

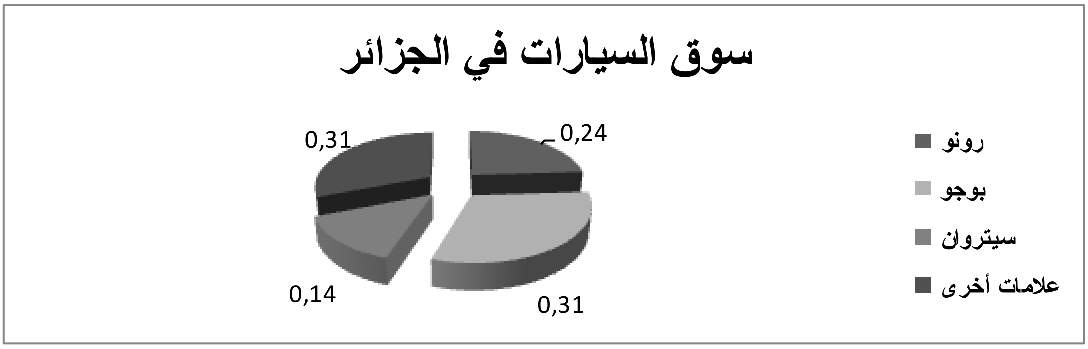
3 – 1 – 1 – الرسم البياني القطاعي (الدائري)

يمكن الحصول عليه، إما باستعمال درجات الدائرة أو بالاستناد إلى التكرار النسبي المئوي المتوفر في

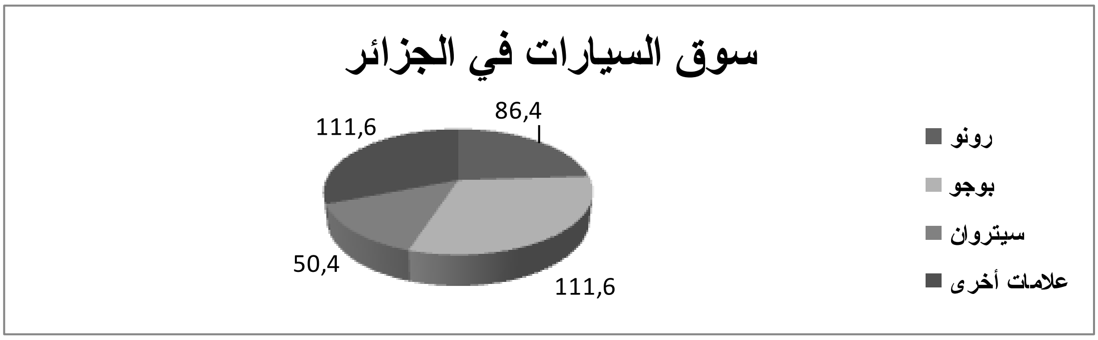
الجدول أعلاه. وعلى هذا الأساس علينا تطوير الجدول السابق إلى جدول يحتوي على الدرجات الموافقة

xi	fi التكرار النسبي	الدرجات الموافقة (°o)
رونو	$\frac{24}{100} = 0.24$	$0.24 \times 380 = 86.4$
بوجو	0.31	111.6
سيتروان	0.14	50.4
علامات أخرى	0.31	111.6
Σ	1.00	380

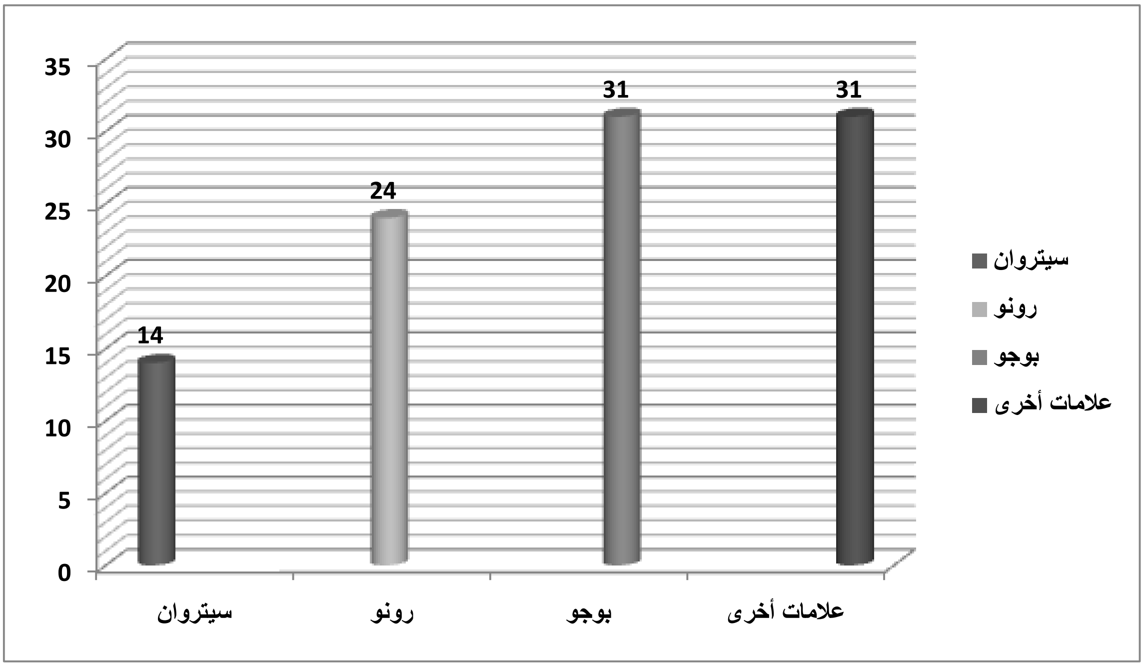
رسم بياني قطاعي باستعمال التكرارات النسبية



رسم بياني قطاعي باستعمال درجات الدائرة



3 – 1 – 2 – الرسم البياني الأنبوبي



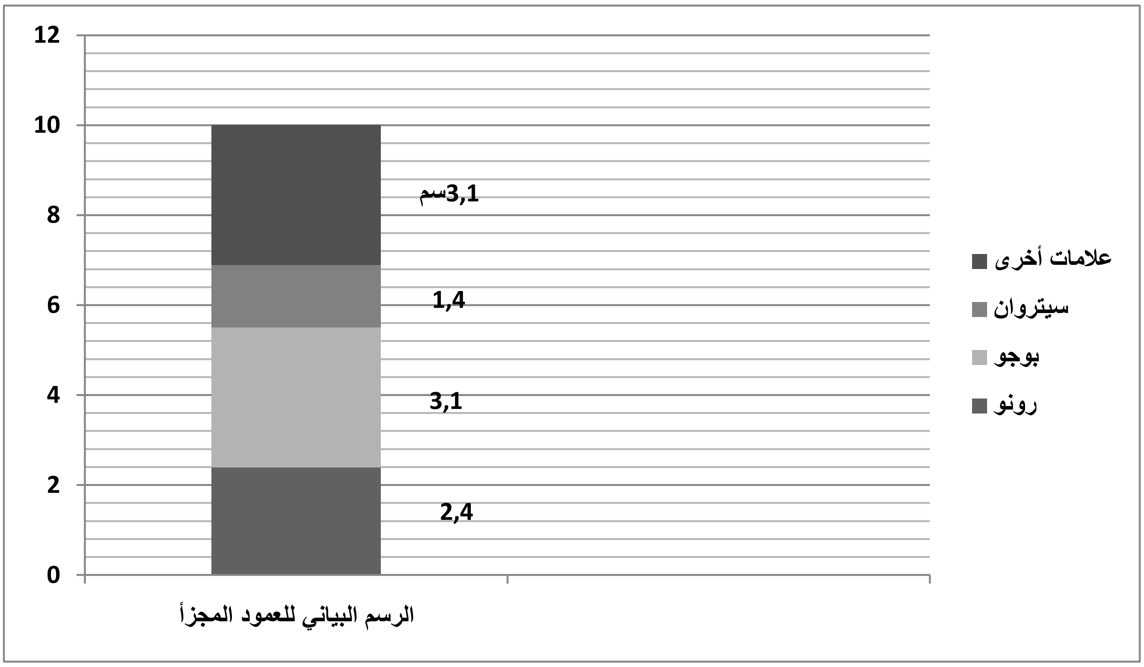
تلاحظ من خلال الرسم البياني الأنبوبي، أننا اعتمدنا فيه على التكرارات النسبية المئوية، كما يمكن رسمه بناء على التكرارات الطبيعية.

3 – 1 – 3 – الرسم البياني للعمود المجزأ

حيث نفرض أن طول المستطيل مثلا هو 10 سم ثم نبحث عن الطول الموافق لكل علامة من علامات السيارات بواسطة الجدول المساعد على النحو التالي.

xi	fi	الطول الموافق (سم)
رونو	0.24	$0.24 \times 10 = 2.4$
بوجو	0.31	$0.31 \times 10 = 3.1$
سيتروان	0.14	$0.14 \times 10 = 1.4$
علامات أخرى	0.31	$0.31 \times 10 = 3.1$
Σ	1.00	10

بعد الحصول على المعطيات اللازمة نقوم برسم العمود المجزأ بواسطة الكومبيوتر كما يلي.



3 – 2 ـ الرسم البياني لمتغير كيفي رتبي

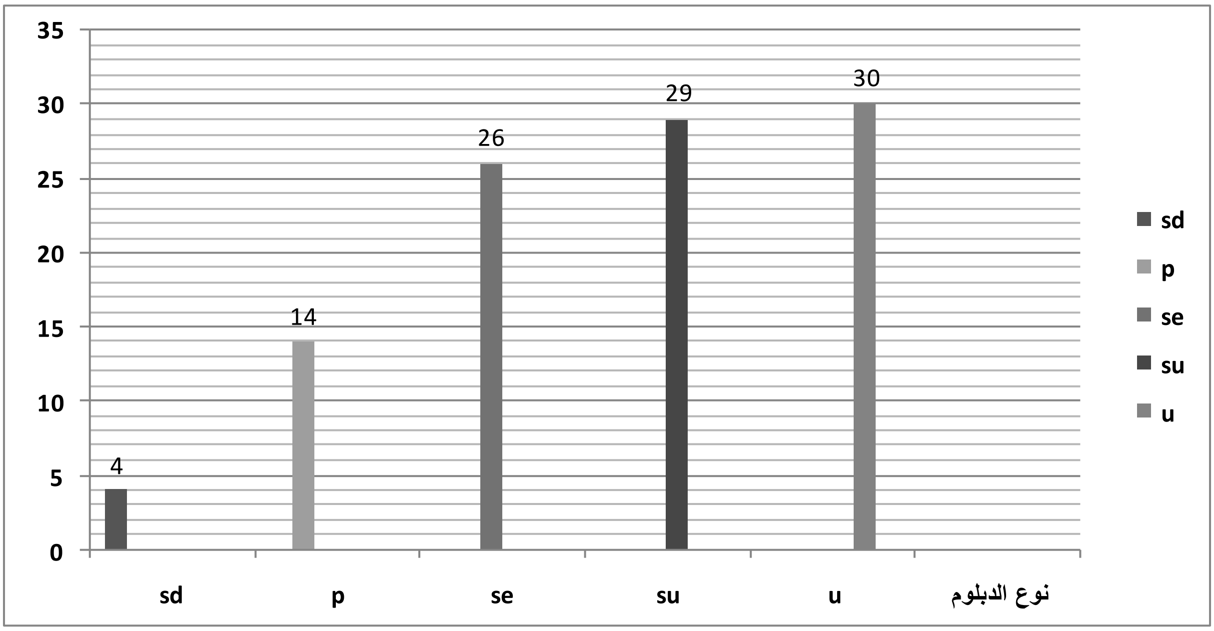
يمكن معالجة معطيات هذا النوع من المتغيرات بيانيا، تماما مثلما مر بنا عند معالجة المتغير الكيفي الإسمي، كما يمكن معالجته عن طريق أعمدة التكرار المتجمع الصاعد .

مثال: استجوبنا عينة تتكون من 30 شخص حول آخر دبلوم تحصلوا عليه، فكانت المعطيات بعد تنظيمها في الجدول على النحو التالي:

xi	F_I	$Fi\uparrow$	fi
Sd	04	$4 \quad +0 = \quad 4$	$\frac{04}{30} = 0.133$
P	10	$4 + 10 = \ 14$	0.33
Se	12	$14 + 12 = \ 26$	0.40
Su	03	$26 \quad +3 \ = 29$	0.10
u	01	$29 + \ 1 \ = 30$	0.03
$\sum$	30	$---$	$\cong 01$

حيث أن sd تعبر بدون شهادة، p ابتدائي، se ثانوي، su عالي غير جامعي، u جامعي

رسم بياني لمتغير كيفي رتبي بواسطة التكرار المتجمع الصاعد



3 – 3 – الرسم البياني لمتغير كمي منقطع

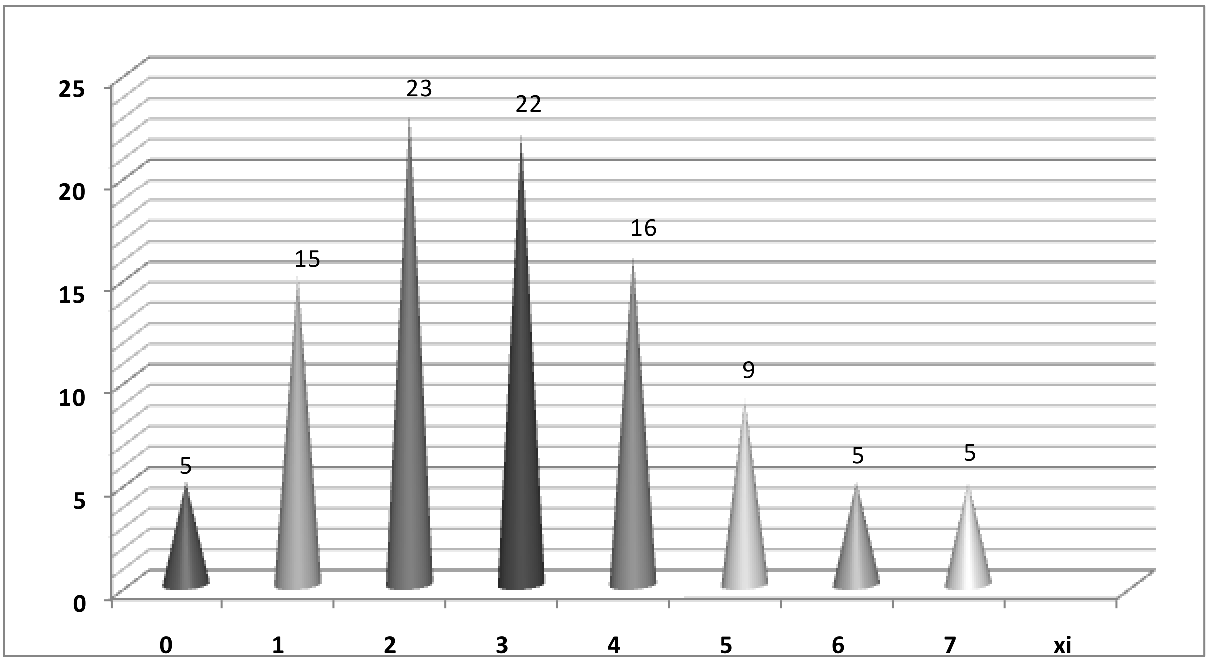
المتغير الكمي المنقطع يتميز بطبيعة عددية، أي أنه قابل للعد وغير قابل للقياس، ويمكن ترجمته بيانيا عن طريق الأعمدة البسيطة أو بواسطة التكرار المتجمع الصاعد أو النازل.

مثال: سجل مسئول مبيعات لمنتوج "حقيبة اليد النسائية" مستوى الطلب اليومي خلال مدة 100 يوم مفتوحة ومتتابعة، قدمت المعطيات في الجدول التالي.

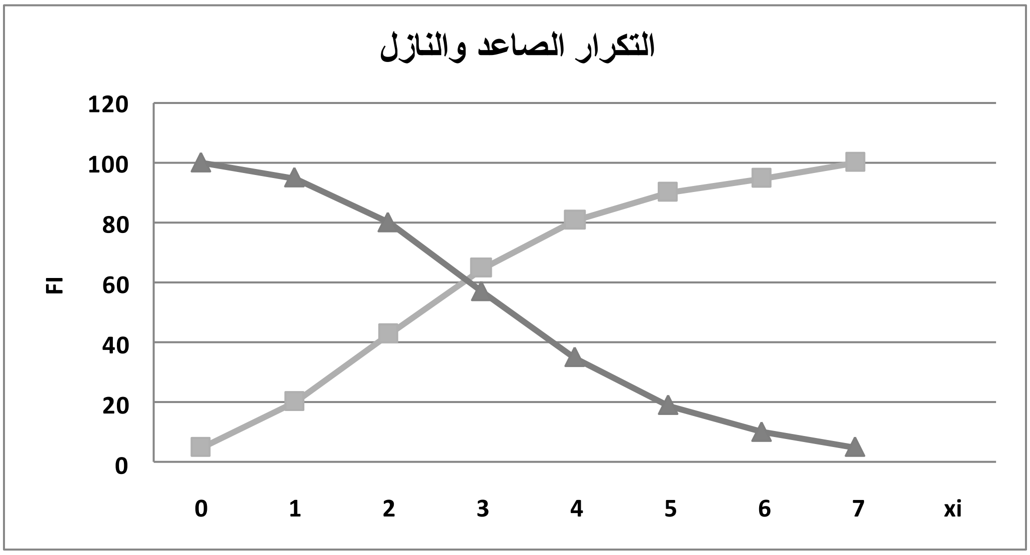
عدد الحقائب المطلوبة / يوم (xi)	عدد الأيام التي بيعت فيها الحقائب (F_I)	ت . متجمع صاعد $Fi\uparrow$	ت. متجمع نازل $Fi\downarrow$
0	05	05	100
1	15	20	$100-5=95$
2	23	43	$95-15=80$
3	22	65	$80-23=57$
4	16	81	$57-22=35$
5	09	90	$35-16=19$
6	05	95	$19-09=10$
7 من الحقائب فما فوق	05	100	$10-05=05$
$\sum F_I$	100	——	——

وعلى هذا الأساس يمكن تقديم العروض البيانية التالية.

3 – 3 – 1 – العرض البياني للأعمدة البسيطة



3 ـ 3 ـ 2 ـ العرض البياني للتكرار المتجمع الصاعد والنازل



3 ـ 4 ـ الرسم البياني لمتغير كمي مستمر

إذا كان المتغير الكمي المنقطع هو قيمة إحصائية تنتمي إلى مجموعة الأعداد الصحيحة، فإن المتغير الكمي المستمر هو قيمة إحصائية تنتمي إلى مجموعة الأعداد الكسرية، وهذا يعني أن القيم المشاهدة (المسجلة) للظاهرة محل الدراسة هي قيم ناجمة عن القياس (الطول، الحجم، الوزن، النمو الاقتصادي، الزمن ...الخ)، ونظرا لكون الأعداد كسرية لذلك يتطلب معالجتها من خلال تقسيمها إلى فئات، بحيث كل فئة تعبر عن مجموعة من القيم. وعلى هذا الأساس فإن العرض البياني المناسب لهذا المتغير يأخذ واحدا من الرسومات التالية :المدرج التكراري، المضلع التكراري، المنحنى المتجمع الصاعد والمنحنى المتجمع النازل، المنحنى البياني.

مثال: سلسلة إحصائية تتكون من 51 نوع من الأشجار أخذت كعينة مقاسة ب: سم

10.8 سم، 08.9، 11.7، 13.2، 07.9، 15.2، 06.6، 07.6، 06.0، 09.8، 05.7، 11.5
04.2، 03.2 ، 04.2، 12.0، 02.3، 04.9، 16.0، 10.4، 06.9، 07.8، 10.1، 80.0
11.3، 08.9، 10.0، 05.6، 05.8، 11.8، 11.1، 13.7، 06.1، 09.4، 18.5، 04.8
08.7، 09.5، 05.8، 40.0، 06.7، 07.0، 12.0، 06.8، 08.6، 09.6، 03.5، 04.5
05.8، 12.7، 06.5 سم.

قبل تقديم الرسومات البيانية المذكورة أعلاه، علينا القيام ببعض الخطوات الأساسية لضمان الحصول على جدول تكراري مستوف للشروط، يمكننا في الأخير من تقديم الرسومات البيانية اللازمة.

3 – 4 – 1 ـ حساب المدى العام

هو الفرق بين أعلى قيمة إحصائية وأدنى قيمة، ويعطى بالعلاقة التالية:

$$E = Xmax - Xmini = (18.5 - 2.3) = 16.2$$

3 – 4 – 2 حساب عدد الفئات

نذكر أنه من الأفضل اختيار خمس فئات على الأقل تتميز بطول متساو، حتى تكون مفيدة عند تحليل الظاهرة المدروسة، وهي بالمناسبة تعتبر قاعدة عامة. لحساب عدد الفئات اللازمة، نختار واحدة من المعادلتين .

ـ معادلة Yule

تستعمل عادة، عندما يكون عدد الوحدات الاحصائية أقل من 1000 وحدة إحصائية، وتعطى على النحو التالي: $N.de\ classes = 2.5\sqrt[4]{n}$، يمكن إدخال تعديلات على هذه العلاقة .

نفرض G هو عدد الفئات التي نبحث عنها.

$G = 2.5\sqrt[4]{n}$

ندخل اللوغاريتم العشري على طرفي المعادلة فيكون: $\log G = \log[2.5.\sqrt[4]{n}]$

$$= 2.5\left[\frac{1}{4}\log n\right]$$

$$G = 2.5 \times 10^{\frac{\log n}{4}}$$

ـ معادلة Striges

تستعمل هذه المعادلة حينما يكون عدد الوحدات الإحصائية المشاهدة أكثر من 1000 وحدة إحصائية. تعطى بالعلاقة التالية. $N.de\ classes = 1 + 3.322\ logn$

ملاحظة: في كلا المعادلتين نستخدم اللوغاريتم العشري.

يتضح من خلال المعادلتين أن معادلة **Yule** هي الأنسب، وبالتالي يكون عدد الفئات هو:

$$G = 2.5\left[10^{\frac{log51}{4}}\right] = 2.5\left[10^{\frac{1.7075}{4}}\right] = 2.5\ [10^{0.4268}] = 2.5\ [2.6717] = 6.67$$

$$G \cong 07$$

3 – 4 – 3 ـ طول الفئة: يمكن الحصول عليه من خلال تقسيم المدى العام على عدد الفئات. أي أن

$$k = \frac{E}{G} = \frac{16.2}{07} = 2.31 \text{ سم}$$

ملاحظة هامة: لتحديد طول الفئة بدقة نرجع إلى الوحدات الإحصائية حيث أعطيت برقم واحد بعد الفاصلة، وعليه يكون $k =$ 2.3 سم، وهو مشترك لكل الفئات.

3 ـ 4 ـ 4 ـ تشكيل الفئات

ـ **الفئة الأولى:** أي فئة تتكون من حدين (حد أدنى وحد أعلى). أما الحد الأدنى فهو القيمة الإحصائية الأقل في السلسلة، حيث نجد في مثالنا هذا القيمة 2.3سم، أما حدها الأعلى فهو ناجم عن إضافة طول الفئة الى الحد الأدنى.

بمعنى :نفرض أن الفئة هي $[a - b[$ حيث $a = 2.3$ وبالتالي $b = a + k$

وعليه، يكون $b = 2.3 + 2.3 = 4.6$ ، وهكذا نكون قد تحصلنا على نموذج الفئة الأولى. $[2.3 - 4.6[$.

ـ **الفئة الثانية:** هي عبارة عن الحد الأعلى للفئة الأولى نضيف له طول الفئات. أي أنها $[4.6 - 6.9[$، وهكذا دواليك حتى الفئة السابعة كما هو مبين في الجدول التالي.

الفئات (x_i)	التكرارات(F_I)	$f_i = F_I/\sum F_I$	$F\uparrow$	$F\downarrow$
$[2.3 - 4.6[$	07	0.14	07	51
$[4.6 - 6.9[$	15	0.29	22	44
$[6.9 - 9.2[$	09	0.18	31	29
$[9.2 - 11.5[$	12	0.23	43	20
$[11.5 - 13.8[$	05	0.10	48	08
$[13.8 - 16.1[$	02	0.04	50	03
$[16.1 - 18.4[$	01	0.02	51	01
$\sum F_I$	51	1.00	--------	--------

كما يمكن الحصول على التكرار النسبي المئوي الصاعد والنازل

$\%Fi\uparrow$	$\%Fi\downarrow$
14	100
43	86
61	57
84	39
94	15
98	06
100	02
-------	-------

ملاحظة: من خلال المثال السابق، يمكن الحصول على طول الفئة في الحالات التالية.

3-4 - 3 - 1 - حالة توفر الفئة

طول الفئة = الفرق بين الحد الأعلى والحد الأدنى

مثال: من معطيات المثال السابق، نأخذ مثلا الفئة الرابعة، طولها هو $k4 = (11.5 - 9.2) = 2.3$

3-4 - 3 - 2 - حالة توفر مراكز الفئات فقط

طول الفئة = الفرق بين مركزي فئتين متتاليتين

مثال: توفرت لدينا معطيات إحصائية بالشكل التالي.

الفئات (x_i)	مراكز الفئات $(\acute{x_i})$
$04 - 08$	$\acute{x}1 = \frac{4+8}{2} = 06$
$08 - 12$	$\acute{x}2 = \frac{8+12}{2} = 10$
$12 - 16$	$\acute{x}3 = 14$
$16 - 20$	$\acute{x}4 = 18$

بعد حساب مراكز الفئات، يمكن بسهولة حساب أي طول لأية فئة كانت، وليكن طول الفئة الثالثة $k3 = \acute{x}3 - \acute{x}2 = (14 - 10 = 04)$ وهكذا بالنسبة لباقي الفئات الأخرى.

من هذا المثال نستنتج بأن مركز الفئة (أية فئة) هو: $\acute{x}_i = \frac{a_i+b_i}{2}$ حيث ai هو الحد الأدنى للفئة ، bi هو الحد الأعلى للفئة.

سؤال: من خلال هذه المعادلة، استنتج كل من: $a_i \; ; \; b_i$.

ملاحظة

في الجداول التكرارية التي يكون فيها مجال الفئة الأولى أو الأخيرة مفتوحا، لا يمكن معرفة طول تلك الفئة إلا بالاستناد على الفئة التي تليها، أو التي قبلها.

بعد معالجة المعطيات الإحصائية في جدول تكراري، يمكن أن نمر إلى الرسومات البيانية الممكنة لهذا النوع من المتغيرات.

3 – 4 – 5 – المدرج التكراري

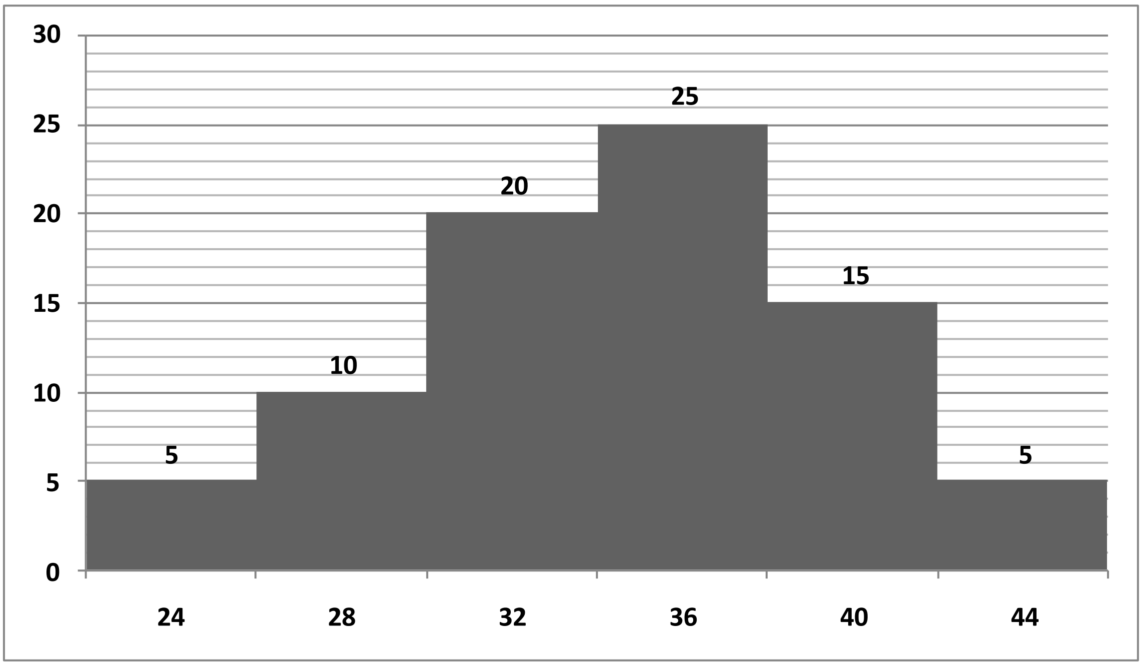
المدرج يترجم التكرارات المعبرة عن الفئة، بواسطة مستطيلات قاعدتها طول الفئة المناسبة، وارتفاعها عدد التكرارات الطبيعية (n_i) المقابلة لتلك الفئة.

مثال: الجدول التالي يوضح توزيع أجور عينة من العمال في أحد فنادق الجزائر العاصمة.

أجور العاملين (xi)	F_I	$\acute{x}i$	$F\uparrow$	$F\downarrow$	fi
$20-24$	05	22	05	80	0.062
$24-28$	10	26	15	75	0.125
$28-32$	20	30	35	65	0.250
$32-36$	25	34	60	45	0.312
$36-40$	15	38	75	20	0.187
$40-44$	05	42	80	05	0.062
$\sum_{i=1}^{6} F_I$	80	-------------	----------	-------	0.998 $\cong 1.00$

نتيجة : $01 \geq \sum fi \geq 0$ كما لا يجوز جمع كل من التكرار الصاعد والنازل وكذلك مراكز الفئات .

رسم بياني للمدرج التكراري، حيث وضعنا على المحور الأفقي الفئات وعلى العمودي التكرارات



3 – 4 – 6 ـ المضلع التكراري

خلافا للمدرج التكراري، فهو يبنى على أساس وضع مراكز الفئات على المحور الأفقي، بينما نضع التكرارات الطبيعية على المحور العمودي.

من خلال معطيات المثال السابق يمكن الحصول على المضلع التكراري كما هو موضح في الرسم

3– 4 – 7 – منحنيات المتجمع الصاعد والنازل

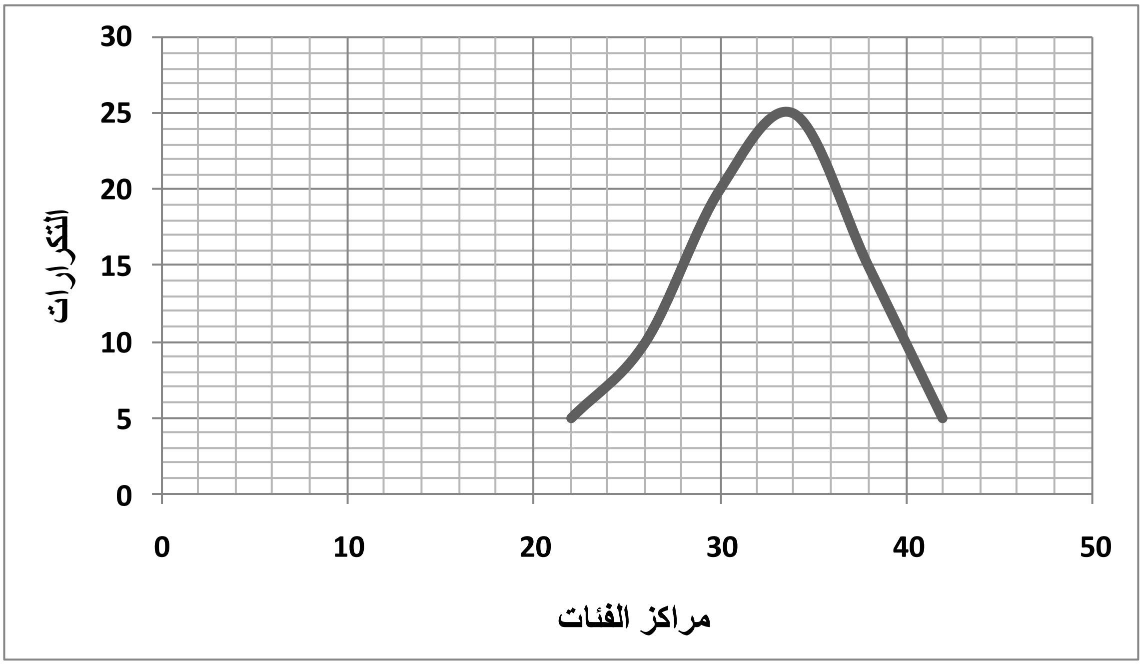
زيادة عن منحنيات التكرارات الصاعدة والنازلة التي رأيناها في المتغير الكمي المنقطع، يمكن رسمها بواسطة التكرارات النسبية الصاعدة والنازلة.

سؤال: المطلوب رسم هذه المنحنيات من خلال معطيات المثال السابق.

3 – 4 – 8 – المنحنى التكراري

هو نفسه المضلع التكراري مع تحوير طفيف في الرسم ليتحول إلى منحنى. هذا باليد، أما بواسطة الكومبيوتر وباستعمال تقنية الايكسال وفقا لنفس المعطيات، فيكون على النحو التالي.

انطلاقا من معطيات المثال السابق نقدم المنحنى البياني المناسب.



.

3 – 4 – 9 – حالة خاصة : أطوال الفئات الغير متساوية

مثال: لدينا عينة تمثل نوع من المصابيح المستوردة ($n = 100$) قمنا بدراسة العمر الزمني لها، فكانت المعطيات على النحو التالي.

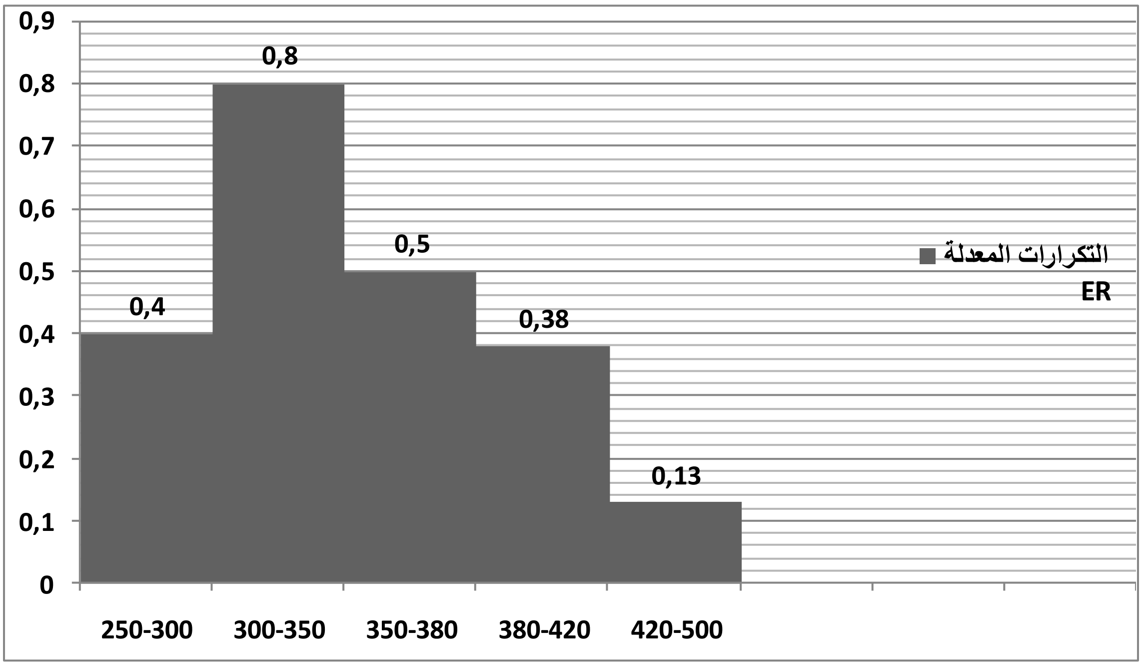
ع. الزمني xi	$250-300$	$300-350$	$350-380$	$380-420$	$420-500$
المصابيح F_I	20	40	15	15	10

نلاحظ بأن طول الفئات غير متساو (غير موحد)، وهو ما يدفعنا إلى تعديل التكرارات (لأن قاعدة المقارنة غير ثابتة) حتى يكون هناك تناسبا بين طول الفئة والتكرار المقابل لها. نرمز إلى التكرار المعدل بالرمزER

نحصل عليه من خلال المعادلة: $\boldsymbol{ER} = \frac{F_I}{k_i}$وبناء على هذا الأساس نحصل على الجدول التالي.

xi	F_I	ki	$ER = \frac{F_I}{K_I}$	$\acute{x}i$
$250-300$	20	50	0.40	275
$300-350$	40	50	0.80	325
$350-380$	15	30	0.50	365
$380-420$	15	40	0.38	400
$420-500$	10	80	0.13	460
$\sum F_I$	100	-------	-------	-------

بعد إجراء التعديلات اللازمة على التكرارات، حينها يمكن رسم المدرج التكراري، حيث نضع على المحور الأفقي الفئات، وعلى المحور العمودي التكرارات المعدلة (ER).



سؤال: هل يمكن إنجاز مضلع تكراري لفئات غير متساوية الطول.

الجواب: لا يمكن إنجاز ذلك إلا بعد أن نقسم المدرج إلى فئات متساوية الطول (فئات جزئية)، ثم نمرر خط منكسر يربط منتصف قمة كل مستطيل، مع إضافة فئتين خياليتين من اليمين واليسار، بحيث يقطع المضلع المحور الأفقي عبر منتصف الفئتين الخياليتين.

مثال: لدينا 140 تعاونية فلاحية موزعة حسب المساحة (هكتار) على النحو التالي.

المساحة (هك) xi	F_I	ki	ER	$F\uparrow$	$F\downarrow$
$0 \quad -5$	10	05	10	10	140
$5-10$	30	05	30	40	130
$10-20$	40	10	$40.\frac{5}{10} \quad =20$	80	100
$20-40$	60	20	$60\,.\frac{5}{20} \quad =15$	140	60
$\sum F_I$	140	----	----	----	----

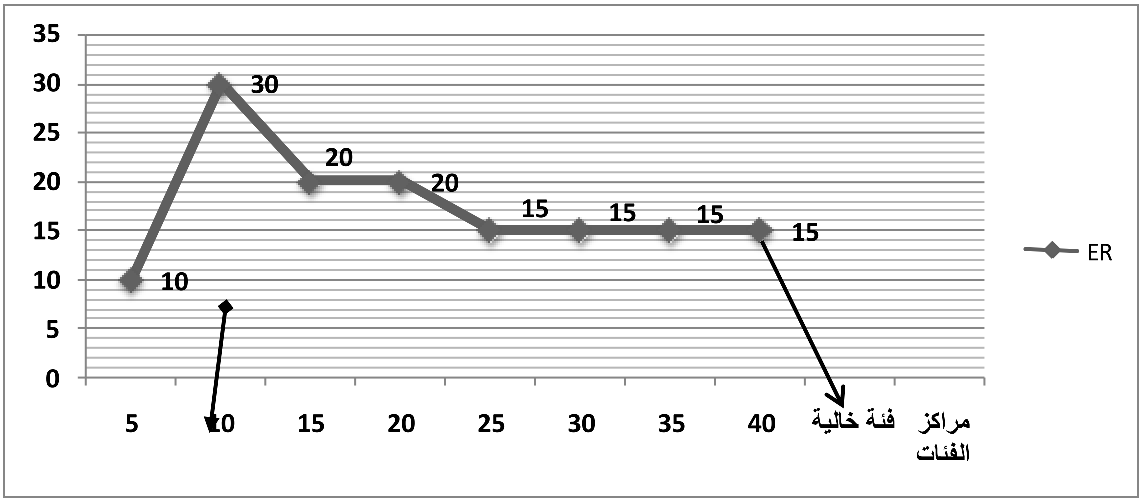
نظرا لكون الفئتان الأولى والثانية متساويتي الطول، كما أنهما يحوزان على أقصر طول، فلا داعي لتعديل التكرارات الخاصة بهما.

وبالتالي نلجأ إلى تعديل البقية، بحيث يكون أساس التعديل هو أحد الطولين القصيرين ($k1$, $k2$)، وليكن مثلا $k1$.

$$ER3 = \frac{n3 \times k1}{k3} = \frac{40 \times 5}{10} = 20$$

$$ER4 = \frac{n4 \times k1}{k4} = \frac{60 \times 5}{20} = 15$$

وبناء على هذا الأساس يكون العرض البياني للمضلع التكراري على النحو التالي:



تلاحظ معي أننا قسمنا مساحات التعاونيات إلى قطع متساوية الطول، بحيث كل واحدة تساوي 05 هكتار، وإلا لما استطعنا إنجاز المضلع التكراري.

ملاحظة: يمكن رسم المضلع التكراري على نفس الرسم البياني للمدرج التكراري المعدل غلى فئات متساوية الطول. كما يمكن تحوير هذا المضلع إلى منحنى بياني كما رأينا سابقا، بحيث تكون فواصله مراكز الفئات وتراتبه هي التكرارات المعدلة.

سؤال: أرسم العرض البياني للتكرار المتجمع ($Fi\uparrow\ ,\ Fi\downarrow$).

III – تمارين الفصل الأول

01 – تمارين محلولة

التمرين الأول: أجب عن كل الأسئلة التالية

- كيف نسمي مجموع الأشخاص أو الأشياء محل الدراسة الإحصائية.
- كيف يطلق على كل عنصر من المجتمع محل الدراسة .
- كيف نسمي الصفات القابلة للقياس. والقياسات.
- كيف يطلق على سلسلة المشاهدات محل الدراسة.

التمرين الثاني: نهتم بدراسة التدفق السنوي لنهر النيل بين 1871 م و 1970 .المطلوب تحديد كل من:

- المتغير محل الدراسة، نوعه
- المجتمع، طبيعته
- الوحدة الإحصائية، القيمة الإحصائية، شكل سلسلة إحصائية افتراضية .

التمرين الثالث: قال تعالى (وإن تعدوا نعمة الله لا تحصوها)

- كم مرة ورد مصطلح العد والإحصاء بصفة متلازمة في القرءان الكريم، أذكر رقم الآية والسورة.
- هل يوجد فرق بين مصطلح العد ومصطلح الإحصاء ؟. قدم أمثلة .

التمرين الرابع: أجب عن الأسئلة التالية

- فيما يبحث وإلى ماذا يهدف علم الإحصاء.
- ماهي الخطوات التي تلزم علم الإحصاء المرور بها
- ما هي في نظرك الأسباب التي تجعلنا في كثير من الأحيان نلجأ إلى أسلوب المعاينة (اختيار العينة).
- عرف كل من المجتمع، العينة
- متى نحصل على البيانات (المشاهدات، القيم) النوعية.

ـ عرف كل من: التدرج الإسمي والتدرج الترتيبي

ـ عرف البيانات الكمية العددية

ـ قدم شرحا لطرق جمع البيانات الإحصائية

ـ تصنف طرق المعاينة، إلى الطرق العشوائية (الاحتمالية) والطرق غير العشوائية، والمطلوب:

1- شرح طرق اختيار العينة غير العشوائية

2- تعريف العينة العشوائية

3- نفرض لدينا حجم عينة مختارة (n = 25) من مجتمع إحصائي حجمه (N = 500)، ما هو احتمال سحب كل مفردة (وحدة) إحصائية

4 - ما هو الفرق بين العينة العشوائية البسيطة، والعينة العشوائية المنتظمة

5 - أردنا اختيار عينة حجمها (n = 200) من مجموعة من بطاقات التسجيل في كلية الاقتصاد التي سجل فيها (N = 3000) طالب(ة) وذلك من أجل دراسة الأخطاء الواردة في البطاقات.

ـ فسر كيف نقوم بتعيين مفردات (وحدات)إحصائية لعينة منتظمة

ـ ما هي العينة العشوائية الطبقية ؟. عرف العينة العشوائية المنتظمة.

ـ كيف نقوم بتحديد حجم العينات الجزئية المتناسبة من كل طبقة (شريحة، فئة)

ـ هناك عدة طرق لجمع البيانات، أهمها: المقابلة الشخصية، الملاحظة المباشرة، الاستبانة (الاستبيان). اشرح بالتفصيل كل طريقة على حدى.

ـ ماذا تعرف عن برنامج SPSS.

ـ ماذا يقصد بعملية الترميز.

التمرين الخامس: اختبر نفسك بعبارة (نعم) أو (لا)

ـ الإحصاء هو مجموعة من الطرق العلمية

ـ المجتمع الإحصائي أقل من العينة

ـ معاينة شاملة تعتبر بمثابة تعداد

ـ جدول وحدات وقيم إحصائية هو مجموعة من المعطيات لصفة نوعية

ـ نستطيع بصفة دائمة حساب التكرارات المتجمعة للمتغيرات النوعية الاسمية

ـ السلسلة الاحصائية، هي في كل الحالات منظمة (مرتبة)

ـ مجموع التكرارات النسبية دائما تساوي الواحد(01)الصحيح

ـ المضلع لتكرارات نسبية يتطابق مع مضلع التكرارات المطلقة وفق عامل المضاعف

ـ التكرارات المطلقة لمتغير إحصائي كمي تحسب دائما على أساس مئوي (%)

ـ فقط المتغيرات الاحصائية المستمرة يمكن تجميعها على شكل فئات

ـ ما هو الفرق بين مبادئ السلسلة الإحصائية، والتوزيع الإحصائي

التمرين السادس: معاينة تتعلق بالحالة الزوجية لعمال مؤسسة، كانت على النحو التالي

الكيفيات ($les\ modalités)$	التكرارات (F_I)
أعزب	172
متزوج	270
أرمل	19
مطلق	39

السؤال:

ـ ما هي طبيعة الصفة "الحالة الزوجية"، عرف المتغير الإحصائي المشترك مع هذه الصفة.

ـ أحسب التكرارات الكلية (المطلقة، الطبيعية) والتكرارات النسبية والنسبية المئوية لهذا التوزيع.

ـ قدم العروض البيانية المناسبة لهذا المتغير

التمرين السابع: توزيع تقدير الدرجات لدفعة من الطلبة عند نهاية سنة التخرج، كانت على النحو التالي

الكيفيات	التكرارات
جيد جدا	04
جيد	11
قريب من الجيد	45
مقبول	110

ـ ما هي طبيعة صفة "التقدير المحصل". عرف المتغير الإحصائي المشترك مع هذه الصفة.

ـ أحسب التكرارات الكلية والتكرارات النسبية لهذا التوزيع.

ـ قدم العروض البيانية المناسبة.

تمرين رقم 08: لدينا السلسلتين الإحصائيتين التاليتين: $xi\ ; 2\ ,4\ ,7\ ,9$

$yi\ \ ; 14\ ,11\ ,10\ ,08$ المطلوب حساب العبارات التالية.

$\sum_{i=1}^{i=4}(xi+yi)$ ، $\sum_{i=1}^{i=4} yi$ ، $\sum_{i=1}^{i=4} xi$ ثم حساب الآتي عند $i = 4$

$\sum 2xi$ ، $2[\sum xi]$ ، $\sum 5(xi+3)^2$ ، $[\sum yi]+[\sum xi]$

02 ـ تمارين للأعمال الموجهة

أولا: للتدقيق والاستيعاب

الأهداف: يمكن إجمال أهداف الإحصاء الوصفي في النقاط التالية.

ـ تجميع وترتيب وتبويب وتنظيم وتلخيص، تحليل المعطيات باستمرار، من خلال السلسلة الاحصائية مع توضيح الخاصية المدروسة.

ـ اقتراح فرضيات نسبية للمجتمع محل الدراسة من خلال العينة.

أدوات الاستعمال: من أجل تحقيق الأهداف التي ذكرناها سابقا، لابد من الاستعانة بأدوات التحليل التي نوجز ذكرها على النحو التالي.

ـ مختلف الجداول التكرارية

ـ الرسومات البيانية (أعمدة، دائرة، مدرج، مضلع ...الخ)

ـ المؤشرات (المتوسطات، التشتت، الارتباط والانحدار، السلاسل الزمنية ...الخ)

الطرق: بواسطة الإحصاء الوصفي الكلاسيكي (الوحيد، الثنائي المتغير)، أو طرق الإعلام الآلي (طريقة: SPSS أو A D D ...الخ).

ثانيا: تذكير وتطبيق

تختلف العروض البيانية للظاهرة محل الدراسة، حسب نوعها.

المتغير النوعي (الكيفي) الغير قابل للقياس، يترجم بيانيا من خلال: القطاعات الدائرية، الأنابيب المستطيلة، العمود المستطيل المجزأ

المتغير الكمي المنقطع القابل للعد، يترجم من خلال الأعمدة البسيطة، حيث نضع على المحور الأفقي القيم المسجلة (x_i) وعلى المحور العمودي التكرارات الطبيعية (F_I).

المتغير الكمي المستمر: بعد إدراج المعطيات في جدول إحصائي يتضمن كل من القيم المسجلة، التكرارات الطبيعية ،مراكز الفئات، التكرارات المتجمعة الصاعدة، التكرارات المتجمعة النازلة، حينها يمكن القيام برسم واحد من الأشكال التالية :

ـ المدرج التكراري (الفئات على المحور الأفقي، التكرارات الطبيعية على المحور العمودي)

ـ المضلع التكراري (مراكز الفئات على المحور الأفقي ن التكرارات الطبيعية على المحور العمودي)

ـ المنحنى البياني يمكن رسمه باليد مع تحوير في المضلع التكراري وجعله على شكل منحنى بياني، أو من خلال الكومبيوتر المتوفر على أسلوب (EXEL).

03 ـ تمارين غير محلولة

التمرين الأول: يهمنا دراسة متغير "الحالة المدنية " نسميه (X) والسلسلة الإحصائية للقيم المحددة من قبل المتغير تتعلق بعشرين (20) شخص حسب الكيفيات التالية: (أعزب: C)، (متزوج: M)، (أرمل :V)، (مطلق: D) وموزعة على النحو التالي:

$$m\,m\,d\,c\,c\,m\,c\,c\,c\,m\,c\,m\,v\,m\,v\,d\,c\,c\,c\,m$$

بحيث تقرأ من اليمين إلى اليسار.

ـ حدد (ي) كل من: الصفة محل الدراسة، طبيعتها، نوعها، المجتمع، العينة، نطاق المتغير.

ـ حدد (ي) كل من: $x3\,,x7\,,x11\,,x8\,,x15$.

ـ انجز جدولا إحصائيا يتكون من: $\boldsymbol{x_i}\ \ ,\boldsymbol{F_I}\ \ \ ,\boldsymbol{f_i}$

ـ انجز(ي) الرسومات البيانية المناسبة.

التمرين الثاني: سألنا خمسون (50) شخصا حول آخر شهادة متحصل عليها ،على أساس أن كيفيات المتغير هي: (بدون شهادة: SD)، (متوسط :P)، (ثانوي: SE)، (عالي غير جامعي: SU)، (جامعي: U)، أما السلسلة الإحصائية فكانت على التوالي:

$sd\,,sd\,,sd\,,sd\,,p\,,p\,,p\,,p\,,p\,,p\,,p\,,p\,,p\,,p\,,p\,,su\,,su\,,su\,,su\,,su\,,su\,,su\,,su\,,$
$se\,,se\,,se\,,se\,,se\,,se\,,se\,,se\,,se\,,se\,,se\,,se\,,se\,,se\,,u\,,u\,,u\,,u\,,u\,,u\,,u\,,u\,,u\,,$
$u\,,u\,,u$.

المطلوب ما يلي:

ـ الصفة محل الدراسة، طبيعتها، نوعها، المجتمع، العينة، نطاق المتغير .

ـ تحديد كل من: $x5\,,x10\,,x11\,,x13\,,x16\,,x23\,,x34\,,x47$.

ـ إنجاز جدولا إحصائيا يتكون من: $\boldsymbol{x_i}\,,\boldsymbol{F_I}\ ,\ \boldsymbol{f_i}\ \ \boldsymbol{f_\uparrow}\ \ ,\boldsymbol{F_\uparrow}$

ـ إنجاز كل الرسومات البيانية الممكنة.

التمرين الثالث: حي يتكون من خمسون (50) عائلة . المتغير (X) يقدم عدد الأشخاص لكل أسرة (x_i) على النحو التالي:

$1\,,1\,,1\,,1\,,1\,,2\,,2\,,2\,,2\,,2\,,2\,,2\,,2\,,2\,,3\,,3\,,3\,,3\,,3\,,3\,,$
$3\,,3\,,3\,,3\,,3\,,3\,,3\,,3\,,3\,,4\,,4\,,4\,,4\,,4\,,4\,,4\,,4\,,4\,,4\,,5\,,$
$5\,,5\,,5\,,5\,,5\,,6\,,6\,,6\,,8\,,8\,.$

العمل المطلوب:

ـ حدد(ي): الصفة محل الدراسة، طبيعتها، نوعها، المجتمع، العينة، نطاق المتغير.

ـ حدد(ي) كل من: $x5 , x10 , x15 , x20 , x25 , x30 , x35 , x49$.

ـ أنجز (ي) الجدول الإحصائي الذي يتضمن كل عناصر السؤال الثالث من الثاني السابق.

ـ أنجز(ي) كل من الرسم البياني للأعمدة البسيطة، الرسم البياني للتكرار النسبي الصاعد (القيم على المحور الأفقي والتكرار النسبي الصاعد على المحور العمودي).

التمرين الرابع: قمنا بقياس قامة خمسون (50) طالب(ة)، فكانت القيم المشاهدة كما يلي:

157 156 156 156 156 156 155 155 154 154 154 153 153 152 152 152

162 162 162 161 160 160 161 160 160 160 159 159 158 158 157 157

171 171 170 169 169 168 168 168 167 166 165 164 164 164 164 163

171 171

المطلوب:

ـ تحديد الصفة محل الدراسة، طبيعتها، نوعها، المجتمع، العينة، الوحدة الاحصائية، القيمة الإحصائية .

ـ تحديد الصفة محل الدراسة ، طبيعتها ، نوعها ، المجتمع ، العينة ، الوحدة الاحصائية ، القيمة الإحصائية .

ـ إنجاز جدول إحصائي يتكون من : $x_i\ ,\ \ F_I\ ,\ F_{\uparrow}\ ,\ F_{\downarrow}\ ,\ f_i\ ,\ f_{i\,\%}$

ـ أحسب كثافة التكرارات (hi) إذا علمت أنها تعطى بالعلاقة : $\boldsymbol{h_i = \frac{F_I}{K_I}}$.

ـ أحسب كثافة التكرار النسبي (di) إذا علمت أنها تعطى بالعلاقة : $\boldsymbol{di = \frac{fi}{ki}}$

ـ استنتج التكرار الأكثر كثافة (الطبيعي والنسبي).

ـ أنجز كل الرسومات البيانية الممكنة.

الحل النموذجي لسلسلة تمارين الفصل الأول

جواب التمرين الأول

ـ نسمي مجموع الأشخاص أو الأشياء محل الدراسة: المجتمع الإحصائي

ـ كل شخص أو شيء من المجتمع محل الدراسة يسمى: فرد أو وحدة إحصائية.

ـ نسمي الصفات القابلة للقياس: المتغيرات، بينما نسمي القياسات بالمشاهدات.

ـ يطلق على سلسلة المشاهدات محل الدراسة: سلسلة إحصائية، وهي بصفة عامة تدرج ضمن جدول معطيات.

جواب التمرين الثاني

ـ المتغير المدروس: التدفق السنوي، أما نوعه فهو متغير كمي مستمر(الماء يقاس بالحجم)

ـ المجتمع: 100 سنة (1871 ــ 1970)، أما طبيعته فهو متجانس (السنوات)

ـ الوحدة الإحصائية: مثلا 1900 (خذ أية سنة).

ـ القيمة الإحصائية: مثلا 750 مليون لتر مكعب من الماء الذي تدفق خلال سنة ـ ما ـ

ـ تشكيل سلسلة إحصائية افتراضية.

السنوات (ni)	1871	1872	1873	1874	1875	1876	1877	1878	.....
حجم الماء المتدفق (xi)	60مليون ل^3	80	120	50	90	115	130	140	.....

جواب التمرين الثالث

ـ ورد هذا الجزء من الآية في القرءان الكريم: اربع مرات

ـ الآية الأولى: قوله تعالى " وإن تعدوا نعمة الله لا تحصوها إن الله لغفور رحيم" الآية 18 من سورة النحل.

ـ الآية الثانية: قوله تعالى " وأتاكم من كل ما سألتموه وإن تعدوا نعمة الله لا تحصوها إن الإنسان لظلوم كفار" الآية 34 من سورة إبراهيم.

ـ الآية الثالثة: قوله تعالى "لقد أحصاهم وعدهم عدا " الآية 94 من سورة مريم

ـ الآية الرابعة: قوله تعالى " ليعلم أن قد ابلغوا رسالات ربهم وأحاط بما لديهم وأحصى كل شيء عددا " الآية 28 من سورة الجن.

ـ نعم يوجد فرق بين مصطلح "العد" ومصطلح " الإحصاء"، إذ أن العد يعبر عن معطيات ظاهرة متوفرة للجميع لا تتصف بالسرية في غالب الأحيان، كما أنها تعبر عن مدخلات أساسية تتعرض للترتيب والتنظيم والتحليل والتفسير(جداول، رسومات بيانية، طرق إحصائية رياضية مختلفة) من أجل الحصول على نتائج (تسمى معلومات ذات قيمة، وغالبا ما تكون سرية للغاية) تساعدنا على بلورة القرار تجاه الظاهرة محل الدراسة.

وعليه، فإن الانتقال من المعطيات إلى المعلومات (المخرجات) يمر حتما بعدة مراحل تسمى بالمنهج المستخدم في علم الإحصاء. هذا ما يجعلنا نستنتج بأن هناك فرق واضح بين المصطلحين.

مثال: حينما تجمع معطيات حول البطالة في الجزائر، يمكن معرفة الاتجاه العام من خلال الرسومات البيانية، ولكن تكتشف فيما بعد بأن البطالة كمتوسط تختلف من ولاية إلى أخرى، ومن جهة إلى أخرى وبين سكان المدن وسكان الريف، وبين الهضاب والوسط والجنوب والشمال، وبين الفئات العمرية في سن العمل، وبين أصحاب الشهادات وغيرهم، وبين قطاع وآخر...الخ.

جواب التمرين الرابع

ـ يبحث علم الإحصاء في طرائق جمع البيانات وتحليلها وتفسيرها، وذلك من خلال الرسومات البيانية والعلاقات الرياضية، كما يهدف إلى وصف متغير أو مجموعة من المتغيرات (صفات)، من خلال دراسة عينة تساعد على التوصل إلى قرارات مناسبة تعمم على المجتمع الذي أخذت منه هذه العينة.

ـ أما الخطوات التي تلزم علم الإحصاء المرور بها، هي اربع خطوات أساسية تتمثل في :

الخطوة الأولى: تحديد المشكلة (هدف الدراسة) بوضوح ودقة، من أجل ضمان نتائج دقيقة.

الخطوة الثانية: تحديد الأداة التي تستخدم في جمع المعطيات (الاستبانة).

الخطوة الثالثة: تحديد العينة التي ستجمع منها المعطيات (البيانات) وطرائق جمعها.

الخطوة الرابعة: ترميز البيانات وتحويلها إلى أرقام حتى يتسنى إدخالها إلى الحاسوب والتعامل معها بسهولة، وبالتالي إجراء التحاليل الاحصائية حسب هدف البحث.

فائدة

الركائز الأساسية لعلم الإحصاء هي: المتغيرات، اختيار العينة، تصميم الاستبانة (الاستبيان).

ـ الأسباب التي تجعلنا نلجأ إلى أسلوب العينة هي: تجانس وحدات المجتمع (مثل المواد السائلة إذ لا يوجد مبرر لإجراء الفحص لكل أفراد المجتمع)، كذلك عوامل الوقت والجهد والكلفة، كما أن الوحدات المستخدمة في الاختبار يمكن أن تكون عرضة للتلف عند فحص كل أفراد المجتمع (بيض، مصابيح الإنارة، الفرامل ...الخ)، وأخيرا تعذر حصر كل أفراد المجتمع لأسباب عملية مثل: رأي المستهلك حول سلعة ما.

ـ المجتمع الإحصائي هو كل العناصر أو الأفراد الذين تنصب عليهم الدراسة الاحصائية، أو هم كل الأفراد الذين تتعلق بهم مشكلة البحث (طلاب جامعة، سكان ولاية، صنف من الطيور، الصادرات والواردات لبلد ...الخ). أما العينة فهي ذلك الجزء من أفراد المجتمع محل الدراسة، إذ أن حجمها يتحدد بعدد أفرادها، وهي في العادة ما تجري عليها الدراسة.

ـ نحصل على هذا النوع من البيانات، عندما تكون الصفة محل الدراسة هي صفة (سمة) نوعية، والتي تصنف حسب الأنواع أو الأصناف، وليس بقيم عددية، مثل: تصنيف الجنس إلى ذكر وأنثى، وتصنيف أقسام الكلية إلى: الاقتصاد، التجارة، التسيير.

ـ التدرج الإسمي

هو مقياس يصنف فقط عناصر الظاهرة التي تختلف في النوعية، لا في الكمية، وكثيرا ما نستخدم الأعداد لتحديد هوية المفردات (الوحدات الاحصائية) مثل تصنيف الجنس، حيث نجعل صفر(0) للذكر وواحد (1) للأنثى. لاحظ أن (0، 1) لا يدلان على قيم عددية هنا وبالتالي لا يخضعان للعمليات

الحسابية، لأنه بكل بساطة يمكن أن نعين عددين بدلهما مثل (7، 8) ليدلا على نوع الجنس. هناك في الحياة العملية ما لا يحصى من الأمثلة على ذلك. مثل الحالة الاجتماعية (أعزب، متزوج)، ونوع العمل(إداري أكاديمي، عمل آخر).

ـ التدرج الترتيبي

زيادة على خواص التدرج الاسمي، فإنه يسمح بالمفاضلة، أي بترتيب العناصر حسب سلم معين متفق عليه.

أمثلة:

المثال الأول: الرتب الأكاديمية: أستاذ التعليم العالي ① أستاذ محاضر أ ② أستاذ محاضر ب ③ أستاذ مسا عد أ ④ أستاذ مساعد ب ⑤.

المثال الثاني: تقديرات الطلاب: ممتاز ⑤ جيد جدا ④ جيد ③ مقبول ② راسب ①

المثال الثالث: درجة التأييد للإجابة على سؤال: موافق بشدة ⑤ موافق ④ متردد ③ لا أوافق ② لا أوافق بشدة ①. تجدر الإشارة إلى أن هذا المقياس لا يحدد الفرق بدقة بين قيم أفراد العينة محل الدراسة.

ـ البيانات الكمية: هي بيانات لصفة (سمة) محل الدراسة قابلة للعد أو القياس، وعليه فإن البيانات (المشاهدات) المتحصل عليها، تتكون من مجموعة من الأعداد أو الأرقام تسمى بيانات (قيم) كمية أو عددية، مثل علامات الطلبة، كميات السلع، درجات الحرارة، أجور العاملين .

ـ جمع البيانات الاحصائية يتم وفق طريقتين: طريقة المسح الشامل، حيث تجمع البيانات (المشاهدات، القيم الاحصائية) من جميع مفردات (الوحدات الاحصائية) المجتمع محل الدراسة، دون استبعاد أية مفردة، أما طريقة العينة، فهي تعني اختيار عينة تمثل المجتمع محل الدراسة أحسن تمثيل، حتى يتسنى تعميم النتائج وهي في غاية الدقة والصدق.

ـ شرح أصناف العينة

أولا: كيف نختار عينة غير عشوائية

يتم ذلك بطريقة انتقائية، إذ أنها لا تمثل المجتمع محل الدراسة تمثيلا دقيقا، وإنما تتم وفق اختيار الباحث، وبالتالي لا تكون فرصة مواتية للتساوي بين أفراد المجتمع عند الظهور في العينة. إن الهدف من هذا النوع من العينات هو الحصول على نتائج استطلاعية فقط، وبالتالي **لا يمكن استخدام أساليب الإحصاء التحليلي، لأنها تقتصر على العينات العشوائية فقط**. وعلى هذا الأساس تم تقسم العينات الغير عشوائية إلى ثلاثة أنواع.

.

01 ـ العينات العرضية

تحدث عندما يتم جمع البيانات من المواطنين لمعرفة اتجاهاتهم نحو استهلاك سلعة ــ ما ــ مثلا، أو من العمال في مصنع كبير لإبداء رأيهم حول إدارة المصنع، أو النظم الرقابية بغرض الحصول على معطيات أو مؤشرات بأقل كلفة أو جهد ممكن.

02 ــ العينات الطبقية غير العرضية

يمكن الحصول عليها بتقسيم المجتمع محل الدراسة إلى شرائح متجانسة في وحداتها الإحصائية، مثل: تقسيم عمال مصنع ــ ما ــ إلى شريحة الإداريين وشريحة العمال أو إلى إناث وذكور، أو تقسيمهم حسب المستوى التعليمي وهكذا.

وعلى هذا الأساس تراعى نسبة المجموعات الفرعية في الدراسة، وبالتالي فالعينة لكل شريحة تأخذ بطريقة غير عشوائية، إذ يلجأ الباحث إلى اختيار من يصادفهم فقط.

03 ــ العينة الغرضية

والتي تستخدم لغرض معين ومحدد، كأن ندرس تكاليف صناعة ــ ما ــ وهو ما يتطلب تعاونا من المستجوب لتوفير المعطيات المراد الحصول عليها.

ثانيا: كيف نختار عينة عشوائية

من شروط العينة العشوائية أن تكون وحداتها متجانسة في القيم الإحصائية (استبعاد القيم المتطرفة)، كما أن للوحدات الاحصائية نفس الفرصة عند اختيارها من المجتمع محل الدراسة، بمعنى يكون احتمال سحب أية وحدة معروفا ومتساويا، ويمكن حسابه، لذلك تسمى بالعينة الاحتمالية.

مثال: لدينا مجتمع حجمه $N = 500$ وحدة إحصائية، نريد سحب عينة عشوائية تتكون من $n = 25$ وحدة إحصائية. وحدة إحصائية. إن احتمال سحب وحدة (مفردة) هو : $\frac{25}{500} \times 100 = 5\%$ بمعنى أن كل وحدة يمكن سحبها 5 % من مجموع 500 وحدة ليس 1 % فقط. لأنه من شروط السحب ، أن تسحب الوحدة (لا على التعيين) المرة الأولى ، وتسجل قيمتها ثم تعاد إلى افراد المجتمع ، ثم يتم السحب مرة ثانية وتسجل قيمة الوحدة المسحوبة ثم تعاد إلى مجموع الوحدات ، ثم يتم السحب مرة ثالثة ، وتسجل قيمة الوحدة المسحوبة ثم تعاد إلى المجموع مرة أخرى ... وهكذا دواليك، خمس وعشرون (25) مرة، أي بعدد أفراد العينة المختارة للدراسة.

أنواع العينة العشوائية (الاحتمالية)

01 ــ العينة العشوائية البسيطة

إذ يمكن الحصول عليها بالطريقة التقليدية كما مر بنا، أو عن طريق الجداول الإحصائية الجاهزة (جداول الأعداد العشوائية)، أو عن طريق الحاسب الآلي.

02 ــ العينة المنتظمة

وهي شكل من أشكال العينة البسيطة، وتعرف بأنها العينة التي تأخذ بحيث يتم إضافة رقم معين وبشكل منتظم، من خلال قائمة كاملة لوحدات المجتمع مرتبة عشوائيا. تعتبر العينة المنتظمة بديلا عن العينة العشوائية البسيطة للأسباب التالية:

2 - 1 - هي أكثر سهولة في التنفيذ مقارنة بالعينة العشوائية البسيطة.

2 – 2 – العينة العشوائية البسيطة، ليس من السهل تعيينها من قبل شخص غير مدرب.

مثال: لدينا مجتمع الطلبة $N = 3000$ طالب (ة) وحجم العينة محل الدراسة $n = 200$ نريد معرفة الأخطاء الواردة في بطاقات التسجيل، وبالتالي نقوم بتعيين وحدات العينة المنتظمة على النحو التالي:

نحسب طول الفترة: $\frac{\mathbf{3000}}{\mathbf{200}} = \mathbf{15}$ ← مجال اختيار الأرقام، ثم نقوم باختيار رقما عشوائيا من 1 إلى 15 وليكن مثلا الرقم 8 وهو رقم البطاقة الأولى.

نضيف 15 إلى الرقم 8 فيكون: $15 + 8 = 23$ وهو رقم البطاقة الثانية.

بنفس الطريقة: $15 + 23 = 38$ رقم البطاقة الثالثة، ثم $15 + 38 = 53$ رقم البطاقة الرابعة...

إن آخر بطاقة مسحوبة هي رقم 2993 (حينها يكون عدد البطاقات بعدد أفراد العينة).

ملاحظة: إذا لم يكن طول الفترة عددا صحيحا، نقرب إلى العدد الصحيح.

03 – العينة الطبقية العشوائية

هي العينة التي تأخذ من خلال تقسيم وحدات المجتمع إلى طبقات (شرائح) أو فئات متجانسة، عندها يمكن اختيار عينة عشوائية بسيطة أو منتظمة، أو كليهما.

مثال: نريد دراسة ظاهرة غياب الطلبة عن محاضرة الرياضيات لقسم الاقتصاد على ضوء المعطيات التالية.

السنة الأولى ليسانس	السنة الثانية	السنة الثالثة
500 طالب (ة)	400	280

ولتحقيق هذا الهدف، نختار عينة تتكون من 150 طالب(ة).

السؤال: ما هو حجم العينة لكل شريحة (مستوى)؟. لدينا $N = 1180$ (مجموع الطلبة)، نريد عينة $n = 150$.

حجم عينة السنة الأولى هو : $n_1 = \frac{150}{1180} \times 500 \cong 64$

حجم عينة السنة الثانية هو : $n_2 = \frac{150}{1180} \times 400 = 50$

حجم عينة السنة الثالثة هو $n_3 = n - (n_1 + \ _2) = 150 - (114) = 36$

ـ شرح طرق جمع البيانات

01 ـ طريقة المقابلة الشخصية

وهي مقابلة أفراد العينة والتحدث إليهم عن الظاهرة محل الدراسة، وعليه، فإن كمية المعطيات التي نحصل عليها تكون دقيقة إلى حد ما، خلافا لتحليلها الذي يتصف بالصعوبة. لذلك علينا أن ننتبه إلى تدوين البيانات أثناء المقابلة، لأن أي خطأ في التدوين يؤدي بالضرورة إلى خطأ في النتائج.

02 ـ طريقة الملاحظة المباشرة

تستعمل هذه الطريقة في غياب وحدات العينة، من أمثلة ذلك: الوقوف على نقطة تقاطع الطرق، وعدد السيارات التي تمر من خلال هذا التقاطع من الساعة الثامنة، إلى الساعة التاسعة صباحا، لحصر كثافة السير في وقت ذهاب الموظفين إلى أعمالهم.

أو مراقبة تصرف مجموعة من الأطفال أثناء اللعب وتدوين الملاحظات بهدف التعرف على سلوكيات الأطفال في بعض المواقف.

03 ـ الاستبانة (الاستبيان)

هو وسيلة لجمع البيانات اللازمة للتحقق من فرضيات المشكلة قيد الدراسة، أو الإجابة على أسئلة البحث، وحتى نضمن دقة النتائج وصحتها، علينا الالتزام بالشروط التالية:

3 ـ 1 ـ الأسئلة تكون بسيطة ومفهومة للجميع، ولا تكون غامضة، مثل: كم عدد الأطفال اللذين تقل أعمارهم عن 10 سنوات لديك.

3 ـ 2 ـ الابتعاد عن الأسئلة التي توحي بالإجابة، مثل: ألا تعتقد بأن أسلوب هذا الكتاب مبسط للدارس: نعم ..لا . المجيب يذهب للإجابة الأولى، وكأن الباحث يريد الإجابة التي يتوقعها هو نفسه.

3 ـ 3 ـ تحديد الكميات أو الوحدات حينما تكون الإجابة أرقاما مثل: كم تحتاج من كمية الماء للشرب يوميا. أحدهم يجيب لتر من الماء، والآخر 4 كؤوس، ... لذلك تعاد الصياغة على النحو التالي: كم لتر من الماء تشرب يوميا.

3 ـ 4 ـ الأسئلة تكون واضحة، ومباشرة لا تحتاج للتفكير بعمق.

3 ـ 5 ـ الاستبيان يكون قصيرا قدر الإمكان.

3 ـ 6 ـ يوزع الاستبيان على مجموعة صغيرة للتجريب والتعديل لبعض الأخطاء قبل التطبيق النهائي.

3 ـ 7 ـ الاستبيان يكون صادقا وثابتا، تجنبا للنتائج المضللة، والتمكن بالتالي من تعميمه، ويكون القرار المتخذ بعد التحليل صالحا.

ـ برنامج $spss$ هو برنامج يساعد على إدخال ومعالجة المعطيات الاحصائية للظاهرة محل الدراسة وتقديم النتائج بدقة وبسرعة فائقة . يمكن تنزيله من الانترنت بأية لغة، والتعامل معه بسهولة.

ـ الترميز معناه الانتقال من الاستبيان إلى برنامج SPSS، وبمعنى آخر تحويل إجابات كل سؤال إلى أرقام.

ملاحظة هامة: حسب مفهوم برنامج SPSS فإن:

01 ـ الأشخاص المستجوبون يطلق عليهم اسم حالات (*Des Cases*).

02 ـ كل سؤال (فقرة) في الاستبيان هو بمثابة: متغير (V*ariable*).

03 ـ إجابات الأشخاص عن الأسئلة (الفقرات): تسمى قيم المتغيرات (المشاهدات أو البيانات).

جواب التمرين الخامس

ـ نعم الإحصاء هو مجموعة من الطرق العلمية

ـ العكس تماما، بل العينة هي جزء من المجتمع

ـ نعم، المعاينة الشاملة تعتبر بمثابة تعداد

ـ خطأ، لا يمكن أن نعتبر جدول من الوحدات والقيم الاحصائية هو بمثابة مجموعة من المعطيات لصفة نوعية، قد يكون لصفة كمية .

ـ من الخطأ القول أننا نستطيع حساب التكرارات المتجمعة للمتغير الكيفي الإسمي.

ـ غير صحيح، ليس شرطا أن تكون السلسلة الاحصائية في كل الحالات منظمة (مرتبة)

ـ نعم، وهو كذلك، مجموع التكرارات النسبية يساوي الواحد الصحيح.

ـ نعم، مضلع التكرارات النسبية يتطابق مع مضلع التكرارات المطلقة (الطبيعية) وفق عامل المضاعف.

ـ خطأ، التكرارات المطلقة لمتغير كمي لا تحسب على أساس مئوي (%).

ـ خطأ، بل أيضا المتغيرات الكمية المنقطعة يمكن تجميعها على شكل فئات.

ـ الفرق بين السلسلة الاحصائية والتوزيع الاحصائي، يكمن في كون السلسلة الاحصائية يمكن أن تكون قيمها غير مرتبة، خلافا للتوزيع الإحصائي الذي هو دائما تكون قيم الظاهرة فيه مرتبة.

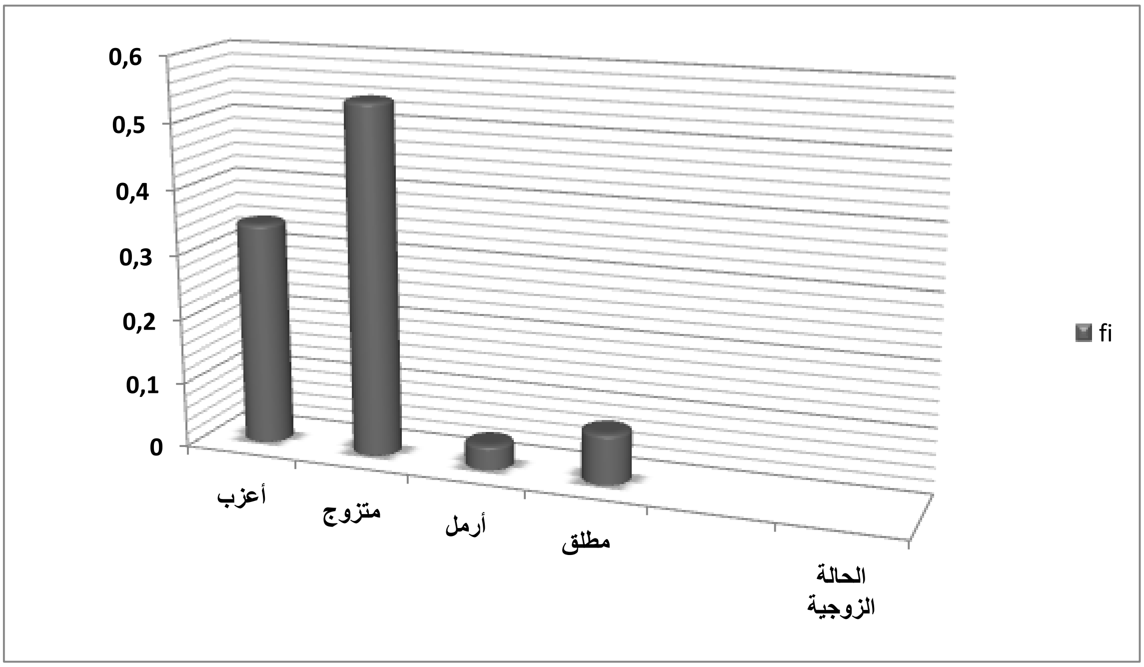
جواب التمرين السادس

ـ طبيعة الصفة محل الدراسة "الحالة الزوجية" هي طبيعة كيفية إسمية، أما المتغير الإحصائي المشترك مع هذه الصفة هو كل الكيفيات الواردة في الجدول (أعزب، متزوج ...الخ) والتي تعتبر بمثابة عناوين ليس إلا، على اعتبار أنها غير قابلة للترتيب أو التصنيف.

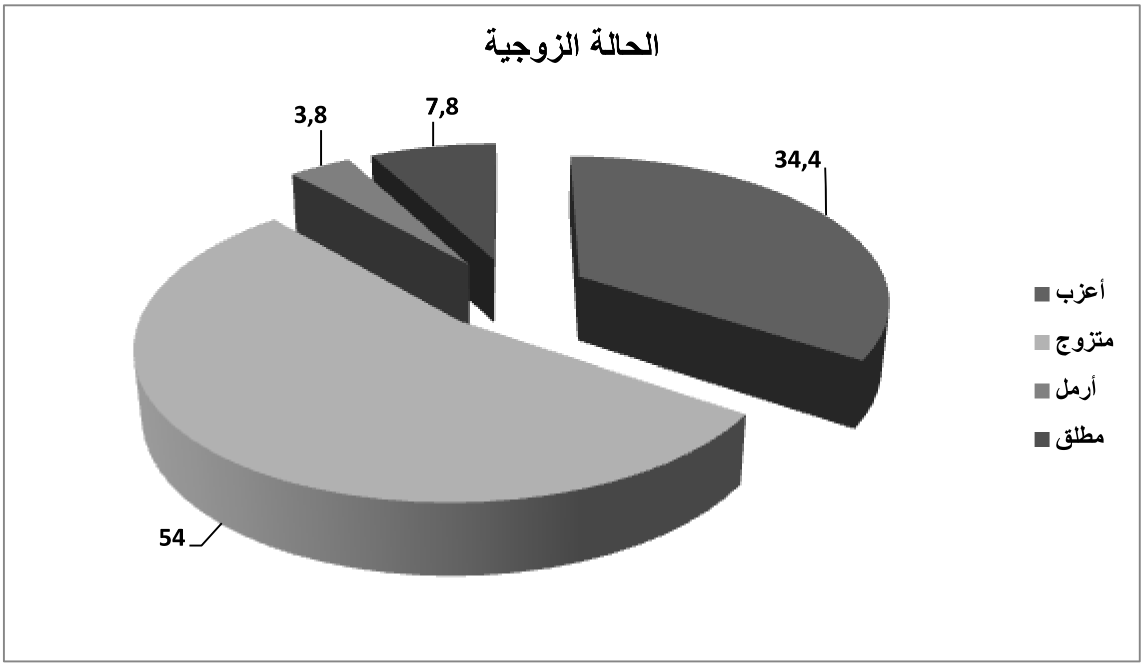
ـ حساب التكرارات المطلقة (الطبيعية) والتكرارات النسبية والنسبية المئوية كما هو وارد في الجدول التالي:

الحالة الزوجية xi	التكرارات المطلقة	تكرارات نسبية fi	$fi\%$
أعزب	172	0.344	34.40
متزوج	270	0.540	54.00
أرمل	19	0.038	03.80
مطلق	39	0.078	07.80
$\sum F_I$	500	1.000	% 100.00

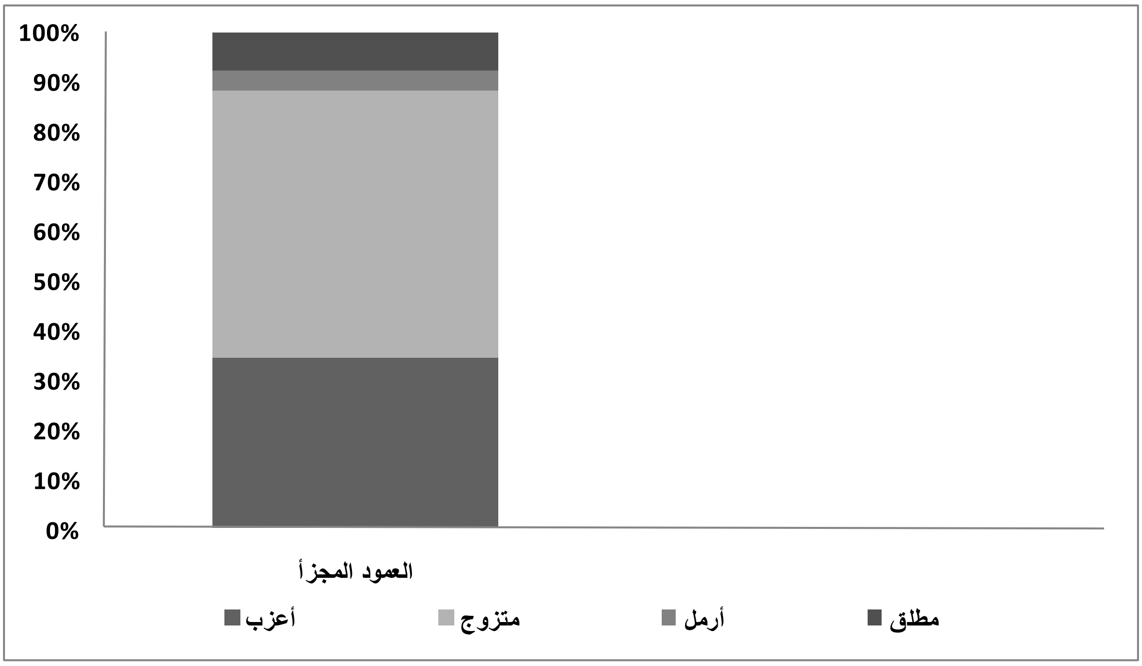
ـ الرسم البياني الأنبوبي بناء على أساس التكرار النسبي



ـ الرسم البياني القطاعي على أساس النسبة المئوية



ـ الرسم البياني للعمود المجزأ بناء على أساس النسبة المئوية



جواب التمرين السابع

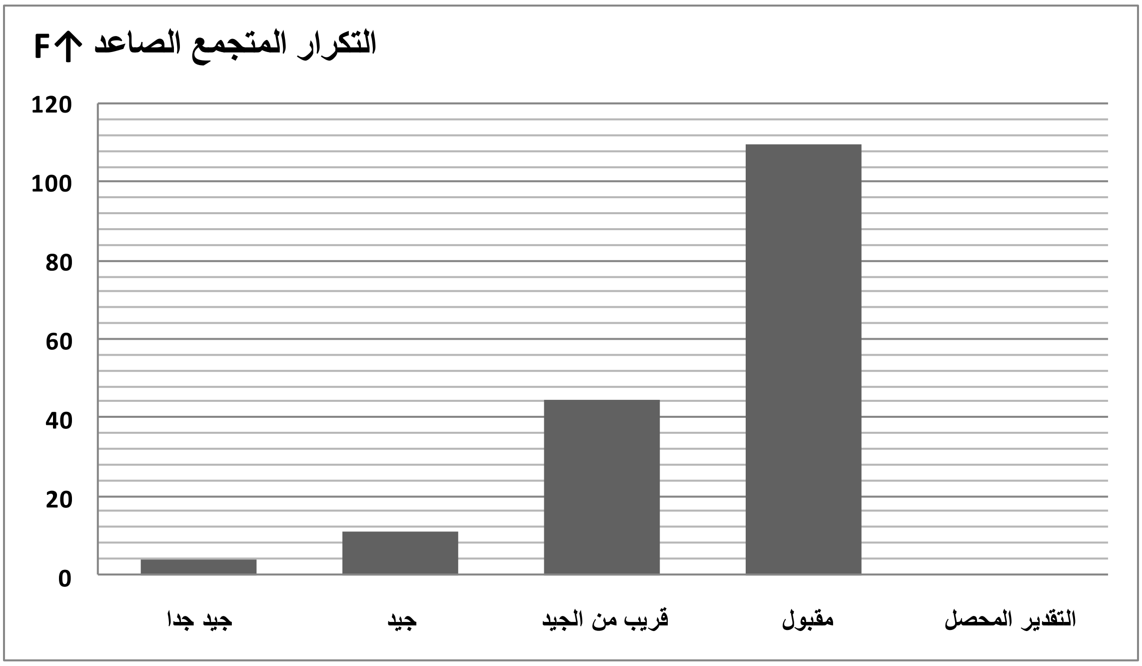
ـ طبيعة الصفة محل الدراسة " التقدير المتحصل" هي طبيعة كيفية رتيبة، أما المتغير الإحصائي المشترك مع هذه الصفة، هو الكيفيات الواردة في الجدول وبشكل مرتب ترتيبا منطقيا، وفي حالات أخرى، سلم القياس يعبر عن النقاط المحصلة في الامتحان، وعليه، فإن المتغير الكيفي الرتبي وفق هذا التصنيف، تعد الكيفيات فيه مرتبة وليست اعتباطية.

ـ المتغير الكيفي الرتبي يسمح بحساب كل من: التكرارات المطلقة (الطبيعية)، التكرار المتجمع الصاعد، التكرار النسبي، التكرار النسبي الصاعد كما هو موضح في الجدول التالي.

الدرجة	تكرار مطلق	ت م صاعد $Fi\uparrow$	تكرار نسبي fi	ت ن صاعد $fi\uparrow$
جيد جدا	04	04	0.024	0.024
جيد	11	15	0.065	0.088
قريب من الجيد	45	60	0.265	0.353
مقبول	110	170	0.647	1.000
المجموع	170	------------------	1.000	------------------

العرض البياني لهذا المتغير الإحصائي يكون على النحو التالي

ـ زيادة عن العرض البياني الذي مر بنا والمتعلق بالمتغير الكيفي الإسمي، يمكن إضافة رسم بياني من خلال التكرار المتجمع الصاعد.



ـ حساب رمز الجمع ($\sum_{i=0}^{n}$)

$$\sum_{i=1}^{i=4} xi = x1 + x2 + x3 + x4 = 2 + 4 + 7 + 9 = 22$$

$$\sum_{i=1}^{i=4} yi = y1 + y2 + y3 + y4 = 8 + 10 + 11 + 14 = 43$$

$$\sum_{i=1}^{i=4} (xi + yi) = (x1 + y1) + (x2 + y2) + (x3 + y3) + (x4 + y4)$$

$$(2 + 8) + (4 + 10) + (7 + 11) + (9 + 14) = 65$$

$$\sum_{i=1}^{i=4} xi + \sum_{i=1}^{i=4} yi = (x1 + x2 + x3 + x4) + (y1 + y2 + y3 + y4) =$$
$$22 + 43 = 65$$

وهذا يعني أن: $\sum_{i=1}^{i=4}(xi + yi) = \sum_{i=1}^{i=4} yi + \sum_{i=1}^{i=4} xi$

$$2\left(\sum_{i=1}^{i=4} xi\right) = (22) . 2 = 44$$

$$\sum_{i=1}^{i=4} 2\,xi = (2.9) + (2.7) + (2.4) + (2.2) + 8 + 4 = 18 + 14 = 44$$

يمكن التأكد من خلال القاعدة:

$$a.\left(\sum_{i=1}^{i=4} xi\right) = \left(\sum_{i=1}^{i=4} a.xi\right)$$

$$\sum_{i=1}^{i=4} 5.(xi + 3)^2 = \sum_{i=1}^{i=4} 5.(xi^2 + 6xi + 9) = \sum_{i=1}^{i=4}(5xi^2 + 30xi + 45) =$$

$$\sum_{i=1}^{i=4} 5xi^2 + \sum_{i=1}^{i=4} 30xi + \sum_{i=1}^{i=4} 45 = 5.\sum_{i=1}^{i=4} xi^2 + 30.\sum_{i=1}^{i=4} xi + \sum_{i=1}^{i=4} 45 =$$

$$= (5 \times 150) + (30 \times 22) + (4 \times 45) = 750 + 660 + 180 = 1590$$

الفصل الثاني

مقاييس النزعة المركزية

مقدمة

لقد مر بنا في الفصل الأول ، التمكن من التعرف على الظاهرة محل الدراسة بكل ما تحمله من خصائص ، والمعالجة الأولية للمعطيات كالجمع والترتيب والعرض الجدولي والرسومات البيانية الموافقة ، وهي كلها خطوات أساسية تسمح للباحث بالتعرف على الظاهرة من حيث الوصف الهندسي وتحديد طبيعة التوزيع .

تبقى هذه الخطوات غير كافية، في غياب تحليل للمعطيات من الناحية الحسابية التي تزودنا بالأعداد، ومن تم التنبؤ واتخاذ القرار المناسب. من هذا المنطلق تأتي أهمية الوصف الحسابي الكمي، حيث يستند بالأساس على جملة من المقاييس نميز فيها نوعين:

الأول: مقاييس النزعة المركزية (مقاييس الوضع).

الثاني: مقاييس التشتت.

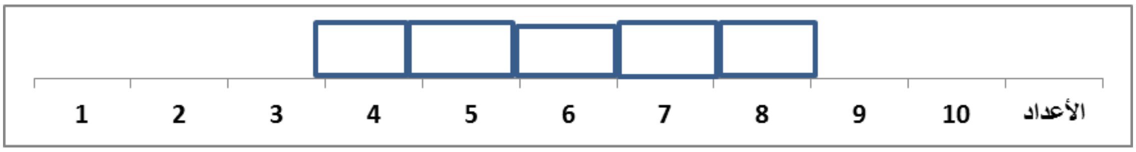
ما هي مقاييس النزعة المركزية

مثال : لدينا السلسلتين الاحصائيتين التاليتين : $x_i\ ;4\ ;\ 5\ ;\ 6\ ;\ 7\ ;8$ و $y_i\ 2\ ;4\ ;6\ ;8\ ;10$ ، ذعنا نجري العمليات التالية.

ـ المدى العام للسلسلة x_i هو $l_1 = (8-4) = 4$ والمدى العام للسلسلة y_i هو $l_2 = 8$ معناه ، مدى السلسلة الثانية ضعف مدى السلسلة الأولى.

ـ الفرق المشترك بين كل زوج متتابع في السلسلة الأولى ، هو واحد (1) ، بينما الفرق المتتابع في السلسلة الثانية هو اثنان (2).

ـ الوسط الحسابي للسلسلة الأولى : $\bar{x} = \frac{30}{5} = 6$ والوسط الحسابي للسلسلة الثانية ايضا يساوي 6



ـ يلاحظ أن القيمة الإحصائية (6) تتجمع حولها باقي القيم بشكل مركز في السلسلة الأولى بخلاف السلسلة الثانية.

ـ تميل القيم الإحصائية إلى التمركز حول قيمة معينة وهي (6) في مثالنا، حيث كلما ابتعدنا عن هذه القيمة، تكون القيم الأخرى في تناقص.

ـ القيمة (6) في السلسلتين تعبر عن مركز توزيع إحصائي.

بناء على هذا الأساس، فإننا كثيرا ما نبحث عن عدد واحد يعبر عن القيم الموجودة في جداول العرض، وهو ما يطلق عليه اسم <u>المتوسط</u>، الذي يشترط فيه ما يلي:

ــ التعريف الدقيق للمقياس، بناؤه على جميع المشاهدات، ومن السهل فهمه وتفسيره، كما يجب أن يكون حسابه سهل وسريع، ولا يتأثر كثيرا بالقيم الشاذة.

ولتحديد هذا الوسط (الموضع ،المركز) نستعمل جملة من القياسات.

المقياس	الرمز
الوسط الحسابي	$\overline{X}$
الوسيط	ME
شبيهات الوسيط (الربيعيات ، العشيرات ، المئينات)	(pi) , (Di) , (Q1,Q3)
المنوال	MO
الوسط الهندسي	MG
الوسط التوافقي	MH
الوسط التربيعي	MQ

اــ الوسط الحسابي ــ الوسط ــ

تعريف : الوسط الحسابي لمجموعة من القيم الاحصائية ، هو قيمة لو استبدلت بكل مفردة من مفردات العينة، لكان مجموع قيم المفردات الجديدة مساويا تماما لمجموع قيم المفردات الأصلية.

01 ــ الوسط الحسابي لبيانات غير مبوبة

إذا كانت لدينا القيم المشاهدة هي : $x1\ ,x2\ ,x3\ ,x4\ ,\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots xn$

يمكن حينها حساب الوسط الحسابي لهذه لقيم بواحدة من ثلاثة طرق.

1 ــ 1 ــ الطريقة المباشرة:

$$\overline{x} = \frac{x1+x2+x3+x4,\ldots\ldots\ldots\ldots\ldots,+xn}{n} = \frac{\sum_{i=1}^{n} xi}{n}$$

1 ــ 2 ــ الطريقة المختصرة: نتبع الخطوات التالية

ــ نأخذ إحدى القيم كوسط فرضي و لتكن A مثلا.

ــ لا توجد أية قيود على اختيار هذا الوسط الفرضي

ــ يفضل اللجوء إلى وسط فرضي، يكون من خلاله مجموع الانحرافات أصغر ما يمكن

ــ نقوم بحساب انحراف كل قيمة عن وسطها الفرضي كما يلي :

$$di = (\ xi - A)$$

ـ حساب الوسط الحسابي على النحو التالي :

$$\bar{x} = A + \frac{\sum_{i=1}^{n} di}{n}$$

1 ـ 3 ـ الطريقة الأكثر اختصارا

ـ بعد حساب انحرافات كل قيمة عن وسطها الفرضي $d_i = (x_i - A)$ نقوم بتقسيم d_i على ثابت وليكن مثلا C بمعنى $\acute{d} = \frac{d_i}{c}$ ، ومنه نقوم بحساب الوسط الحسابي وفق الطريقة المختصرة كما يلي

$$\bar{x} = A + \frac{\acute{d_\imath}}{n} \times C$$

حيث أن الثابتC يختار من ضمن أعداد الانحرافات حتى تقلص إلى أقل قيمة، كما سنرى في المثال الموالي.

مثال1 : واردات السلع والخدمات لبلد ـ ما ـ خلال خمس سنوات متتالية كانت على النحو التالي : 40 مليار $ أمريكي ، 45 ، 50 ، 55 ، 60 مليار $ أمريكي. أوجد الوسط الحسابي لقيمة الواردات بالطرق الثلاثة .

الحل : تلاحظ أن العينة $n < 30$ وبالتالي لا تحتاج إلى تبويب (على شكل فئات). وعليه ، يكون الحل كما يلي:

xi	$di = (xi - A) = xi - 50$	$\acute{d_\imath} = \frac{di}{c}$ $, c = 10$
40 45 50 55 60	$d1 = 40 - 50 = -10$ $d2 = 45 - 50 = -5$ $d3 = 50 - 50 = 0$ $d4 = 55 - 50 = +5$ $d5 = 60 - 50 = +10$	-1 -0.5 0 $+0.5$ $+1$
$\sum_{i=1}^{i=5} xi = 250$	$\sum_{i=1}^{i=5} di = 0$	$\sum_{i=1}^{i=5} \acute{d_\imath} = 0$
الطريقة المطولة $\bar{x} = \frac{\sum_{i=1}^{i=5} xi}{n} = \frac{250}{5} = 50$	الطريقة المختصرة $\bar{x} = +\frac{\sum_{i=1}^{i=5} di}{n} =$ $50 + \frac{0}{5} = 50$	الطريقة المركزة $\bar{x} = A + \frac{\sum_{i=1}^{i=5} \acute{d_\imath}}{n} =$ $50 + \frac{0}{5} = 50$

تمرين : لدينا عينة تتكون من خمسة (05) أسر ، كان انفاقها اليومي على التوالي : 1000 د ج ، 1300 ، 1500 ، 2000 ، 2500 د ج .أحسب متوسط انفاق الأسر في اليوم بالطرق الثلاثة.

02 ـ الوسط الحسابي لبيانات مبوبة

نميز في هذه الحالة بين نوعين من البيانات (بيانات كمية منقطعة وبيانات كمية مستمرة)

2 ـ 1 ـ بيانات كمية منقطعة

في الدراسات الاحصائية ، يكون الباحث أحيانا أمام قيم احصائية تتميز بأهمية خاصة ، مما يكسبها وزنا يختلف من حالة إلى أخرى . أمام هذه الوضعية نلجأ إلى استعمال طريقة الوسط الحسابي المرجح ، والتي تعطى بالعلاقة

$$\bar{x}_p = \frac{x_1F_1+x_2F_2+x_3F_3+\cdots+x_nF_m}{\sum F_i}$$

مثال: تحصل أحد الطلبة في السداسي الأول من السنة الأولى على النقاط الواردة في الجدول

المواد	رياضيات	إحصاء	اقتصاد سياسي	محاسبة	فرنسية	تجارة	انجليزية
العلامة xi	12	14	07	09	11	10	12
المعامل F_i	03	02	02	04	02	01	02

السؤال : ما هو معدل السداسي للطالب.

يحتم على الباحث أن يعامل هذه المفردات(البيانات) بمعاملات ترجيح مختلفة، توضح أهمية كل مفردة عن الأخرى، بمعنى كل قيمة لها وزن مناظر لها. أما المواد في هذا المثال، فهي الحالات الممكنة للمتغير محل الدراسة.

$$\bar{x}_p = \frac{\sum_{i=1}^{7} x_iF_i}{\sum F_i}$$

وعلى هذا الأساس يكون الوسط الحسابي المرجح:

$$\frac{(3.12)+(2.14)+(2.7)+(4.9)+(2.11)+(1.10)+(2.12)}{3+\ 2+\ 2+\ 4+\ 2+\ 1+\ 2} = \frac{170}{16} = 10.62$$

مثال آخر: أردنا دراسة سعر البطاطس، فاخترنا لذلك خمسة (05) أسواق، على ضوئها كانت المعطيات المبينة في الجدول.

رقم السوق	01	02	03	04	05	$\sum F_I$
سعر كغ/ دج (x_i)	25	30	35	40	45	------------
الكمية المشترات (F_i) بلكغ	05	03	02	01	0.5	11.5

السؤال : ما هو السعر المرجح للكلغ وفق هذه العينة من الأسواق.

الجواب : $\bar{x}p = \frac{(5.25)+(3.30)+(2.35)+(1.40)+(0.5.45)}{11.5} = \frac{347.5}{11.5} = 30.22$

وهكذا فإن السعر المرجح للكلغ الواحد هو 30.22 د ج

فائدة :

يلاحظ من خلال نتائج المثالين السابقين أن قيمة الوسط الحسابي متجانسة مع باقي القيم الأخرى، بمعنى غير بعيدة (متطرفة)، وفي نفس الوقت ليس شرطا أن تكون النتيجة إحدى القيم الموجودة في العينة محل الدراسة.
وعلى هذا الأساس، إذا كانت النتيجة بعيدة عن التجانس معناه أن هذه العينة لا يليق أن نطبق عليها الوسط الحسابي لأنه غير صالح للاستعمال، وبالتالي يجب معالجة المتغير محل الدراسة بمقياس آخر.

2 ـ 2 ـ بيانات كمية مستمرة

يعالج الوسط الحسابي قيم المتغير الكمي المستمر ، حيث يلجأ الباحث إلى توزيع القيم على شكل فئات للحفاظ على التجانس من جهة ، ومن جهة أخرى لتعيين قيمة وسطية لكل فئة ، تسمى بمركز الفئة ، حيث تسهل عملية الحساب.

للحصول على الوسط الحسابي، نتبع إحدى الطرق الأربعة التي سوف نوضحها من خلال المثال التالي:

نريد قياس متوسط الأجر لعينة من أجور العاملين في القطاع الخاص .   القيمة : 1000 دج

فئات الأجر xi	15-10	20-15	25-20	30-25	35-30	40-35	42-40	50-45	$\sum F_i$
عدد العمال F_i	10	13	25	16	12	08	04	02	90

سوف نستعرض في الجدول الموالي كل من الطريقة المباشرة وطريقة الوسط الفرضي (المختصرة)

xi	F_i	$\acute{x}i$	$F_i\acute{x}_i$	$\acute{x}_i - A$	$F_i.(\acute{x}_i - A)$
15-10	10	12.5	125	18-	180-
20-15	13	17.5	227.5	13-	169-
25-20	25	22.5	562.5	08-	200-
30-25	16	27.5	440	03-	48-
35-30	12	32.5	390	02+	24+
40-35	08	37.5	300	07+	56+
45-40	04	42.5	170	12+	48+

50-45	02	47.5	095	17+	34+
المجموع	90	--------------	2310	--------------	-435

2 – 2 – 1 ـ الطريقة المباشرة

نقوم بحساب الوسط الحسابي من خلال حساب مركز الفئة الذي يمكن الحصول عليه وفق العلاقة $\acute{x}_i = \frac{a+b}{2}$

حيث a هو الحد الأدنى للفئة و b هو الحد الأعلى لها. بناء على المعطيات السابقة يكون الوسط الحسابي كما يلي.

$$\bar{x} = \frac{\sum \acute{x}_i F_i}{\sum F_i} = \frac{2310}{90} = 25{,}66 \times 1000 = 25660$$

بما أن فئات الأجر تحسب بالآلاف ، معناه أن متوسط الأجر لهذه العينة من أجور العاملين هو ما يعادل 25660 دج / الشهر.

2 – 2 – 2 ـ أما طريقة الوسط الفرضي (المختصرة) فقد مرت بنا سابقا، لذلك لا داعي لتكرار شرح الطريقة.

أخذنا القيمة (A = 30,5) كوسط فرضي ،وبناء على النتائج الموجودة في الجدول أعلاه نجد

$$\bar{x} = A + \frac{\sum(\acute{x}_i - A)}{\sum F_i} = 30{,}5 + \left[\frac{-435}{90}\right].1000 = 25660$$

2 – 2 – 3 ـ طريقة مركز الفئة المفروض

حيث نختار مركز فئة وليكن $\acute{x}_1 = 12{,}5$ نجعله أساس الترتيب التصاعدي لمراكز الفئات المتبقية ، وعلى هذا الأساس تكون لدينا مراكز فئات جديدة نرمز لها بالرمز y_i ثم نبني الجدول الموافق الموضح أدناه.

2 – 2 – 4 ـ الطريقة المركزة

لقد سبق وأن تعاملنا معها ، لذلك ننبه فقط إلى الثابت (C) حيث يكون في المتغير الكمي المستمر هو طول الفئة (K) ، أما إذا كان طول الفئة غير ثابت (متساو) نلجأ إلى تقسيم الفروق على القاسم المشترك.

حساب النتائج

لحساب نتائج طريقة مركز الفئة المفروض والطريقة المركزة نستعين بالجدول التالي، حيث $k = 5$

xi	F_i	$\acute{x}i$	مركز الفئة المفروض yi	$F_i y_i$	$F_i(\acute{x}_i - A)$	$\frac{F_i(\acute{x}_i - A)}{k}$
15-10	10	12.5	0	0	180-	36-
20-15	13	17.5	1	13	169-	33.8-
25-20	25	22.5	2	50	200-	40-
30-25	16	27.5	3	48	48-	9.6-
35-30	12	32.5	4	48	24+	4.8+
40-35	08	37.5	5	40	56+	11.2+
45-40	04	42.5	6	24	48+	9.6+
50-45	02	47.5	7	14	34	6.8+
$\sum F_i$	90		28	237	435-	87-

ـ طريقة مركز الفئة المفروض

$$\bar{x} = \left[\frac{\sum y_i F_i}{\sum F_i}\right] .k + \acute{x}_1 = \left[\frac{237}{90}\right] .5 + 12{,}5 = 25{,}66 \times 1000 = 25660$$

ـ طريقة الوسط الفرضي المركز

تذكر بأننا فرضنا $A = 30{,}50$ وطول الفئة ثابت (متساو) ،حيث $k = 5$ ومنه ، يمكن حساب نتيجة الوسط الحسابي .

$$\overline{\boldsymbol{x}} = A + \frac{\sum F_i(\frac{\acute{x}_i - A}{k})}{\sum F_i} .K = 30{,}5 + \left[\frac{-87}{90}\right] .5 = 25660$$

2 ـ 2 ـ 5 ـ حالة خاصة: المتوسط المشترك

في حالة وجود سلسلتين أو أكثر محل الدراسة لظاهرة معينة، يمكن دمج هذه السلاسل لتكوين متوسط مشترك.

مثال: نريد معرفة متوسط مجموع النقاط لطلبة الإحصاء اللذين يتوزعون على (03) مجموعات، حسب المعطيات التالية.

المجموعات	متوسط درجات الطلاب (x_i)	عدد الطلاب لكل مجموعة (F_i)
A	80	20
B	70	30
C	60	40

السؤال : أوجد متوسط درجات الطلاب لمادة الإحصاء في المجموعات الثلاثة.

الحل :

$$\bar{x}_{(A,B,\ c)} = \frac{F_1\bar{x}_1 + F_2\bar{x}_2 + F_3\bar{x}_3}{F_1 + F_2 + F_3}$$

$$= \frac{(20\times 80) + (30\times 70) + (40\times 60)}{20+30+40} = 67.8$$

03 ـ مزايا وعيوب الوسط الحسابي

يتصف الوسط الحسابي ببعض المزايا والعيوب نوردها على النحو التالي.

ـ يتميز بسهولة حسابه ودقة قياسه، خاصة إذا كانت القيم موزعة توزيعا طبيعيا (متماثل أو شبه متماثل)، كما يعتبر غير متحيز، بمعنى نتيجته تعبر إلى حد ما عن متوسط قيم أفراد المجتمع محل الدراسة.

ـ من خلال معادلته الرياضية يمكن استنتاج مجموع قيم افراد المجتمع أو العينة

$$\bar{x} = \frac{\sum xi}{n} \Rightarrow \sum xi = \bar{x}\,.\,n$$

ـ من عيوبه أنه يتأثر بالقيم المتطرفة، ذلك لأنه يضم كل أفراد العينة في حسابه

ـ يتميز الوسط الحسابي بخواص الجمع والضرب. إذا كانت لدينا قيمة ثابتة مثل A يمكن إجراء العمليات التالية

$$\bar{x} = \frac{\sum(xi + A)}{n} = \bar{x} + A$$

$$\bar{x} = \frac{\sum(xi.A)}{n} = \bar{x}.A$$

ـ مجموع انحرافات قيم المتغير الاحصائي عن وسطها تساوي صفرا ، بمعنى $\sum(xi - \bar{x}) = 0$

نفرض أن d_i هو الانحراف عن الوسط الحسابي ، فيكون $di = \sum(xi - \bar{x}) = 0$

ـ انحراف القيمة الأولى عن الوسط $d1 = x1 - \bar{x}$

ـ انحراف القيمة الثانية عن الوسط $d2 = x2 - \bar{x}$

... ...

... ...

ـ انحراف القيمة الأخيرة عن الوسط $dn = xn - \bar{x}$

مجموع الانحرافات عن الوسط (1)..... $\sum_{i=1}^{n} di = \sum_{i=1}^{n} xi - n\bar{x}$

لدينا : $\bar{x} = \frac{\sum x_i}{n}$ بالتعويض في المعادلة 1 يكون

$$\sum di = \sum xi - n\left(\frac{\sum xi}{n}\right) = 0 \Leftrightarrow \sum di = \sum xi - \sum xi = 0$$

ـ الوسط الحسابي يستجيب بشكل جيد لشروط قاعدة ($yule$)

ـ مجموع مربعات الانحرافات $\sum(xi - \bar{x})^2$ بالنسبة للوسط الحسابي ، تكون أصغر ما يمكن ، مقارنة بأية قيمة أخرى (أكبر كانت أو أقل من قيمة الوسط الحسابي)

مثال : أوزان خمسة طلبة كانت على النحو التالي: 70 ، 85 ، 72 ، 65 ، 90 كلغ

المطلوب : إثبات أن مجموع مربعات الانحرافات بالنسبة للوسط الحسابي هي أصغر ما يمكن؟.

xi	$di = (xi - \bar{x})$	di^2	$di(xi - 73)$	di^2
70	6.4- =76.4-70	40,96	3- =73-70	09
85	8.6 =76.4-85	73,96	12 =73-85	144
72	4.4- =76.4-72	19,36	1- =73-72	01
65	11.4- =76.4-65	129,96	8- =73-65	64
90	13.6 =76.4-90	184,96	17 =73-90	289
382	$\sum d_i = 0$	$\boxed{449,2}$	$\sum di = 16$	$\boxed{507}$

الوسط الحسابي هو: $\bar{x} = 76.4$ وهكذا نستنتج من خلال هذا الجدول مايلي:

ـ مجموع الانحرافات عن الوسط الحسابي تساوي الصفر(0) ، بينما مجموع الانحرافات عن أية قيمة (في هذا المثال فرضنا القيمة 73) تختلف عن الصفر.

ـ مجموع مربع الانحرافات عن الوسط الحسابي ، هي قيمة أقل تماما من مجموع مربع الانحرافات عن أية قيمة تختلف عن الوسط الحسابي.

$$\sum(xi-\bar{x})^2 < \sum(xi-a)^2 \Leftrightarrow 449.2 < 507$$

ـ إذا كان (a , b) ثابتان حقيقيان لكل القيم الحقيقية $x1\ ,\ x2\ ,\ x3,,,,,,xi\ ,,,xn$ فإن

$$\sum_{i=1}^{n}(axi+b) = a\sum_{i=1}^{n}xi + nb = a\left[\frac{1}{n}\sum_{i=1}^{n}xi\right] + b$$

$$\overline{ax+b} = a\bar{x} + b$$

ـ لا يمكن حساب الوسط الحسابي إذا كان الجدول في حالة بيانات مبوبة مفتوحا من أحد طرفيه، بمعنى لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة.

10 ـ لا يمكن استخدامه في حالة الظواهر الوصفية غير الرقمية ، بمعنى لا نستطيع تحديده للبيانات مثل : ممتاز ، جيد جدا ، جيد ، مقبول ، ضعيف.

II ـ الوسيط

يعرف على أنه القيمة الاحصائية التي تتوسط مجموع القيم المشاهدة ، بغض النظر عن ترتيبها ترتيبا تصاعديا أو تنازليا ، أو هو القيمة التي يكون عدد القيم الأقل منها مساويا لعدد القيم الأكبر منها. يرمز له عادة بالرمز.ME

01 ـ الوسيط لبيانات غير مبوبة (بيانات خام).

أمثلة

المثال الأول : السلسلة الاحصائية التالية ، تعبر عن مردود الهكتار الواحد من إنتاج البطاطس بالقنطار لخمس سنوات متتالية : 25 قنطار/ هكتار ، 15 ، 20 ، 29 ، 12 قنطار/ هكتار. ما هي القيمة المتوسطة لمردود الهكتار خلال المدة محل الدراسة.

الحل :

ـ نرتب القيم ترتيبا تصاعديا : 12 ، 15 ، 20 ، 25 ، 29

ـ نلاحظ أن مجموع وحدات العينة محل الدراسة فرديا : بمعنى أنه يتخذ الصيغة $(2k+1)$، في هذه الحالة نتبع الخطوات التالية.

ـ نعين رتبة الوسيط وفق المعادلة $t = \frac{n+1}{2}$

ـ نحدد قيمة الوسيط، وهي القيمة التي ترتيبها $\frac{n+1}{2}$

ـ رتبة الوسيط: $t = \frac{n+1}{2} = \frac{5+1}{2} = 3$

ـ ومنه قيمة الوسيط التي رتبتها 3 أي ME = 20 وبالتالي متوسط مردود الهكتار خلال المدة محل الدراسة هو 20 قنطار للهكتار الواحد.

المثال الثاني: درجات الحرارة المسجلة خلال الستة شهور من نهاية السنة الماضية كانت على النحو التالي: 33، 35، 29 ، 31 ، 26 ، 20. نفس السؤال السابق.

الحل :

نرتب المفردات ترتيبا تصاعديا: 20، 26، 29، 31، 33، 35

ـ نلاحظ أن مجموع وحدات العينة زوجيا ، أي يتخذ الشكل (2 k)

ـ رتبة الوسيط:

$t1 = \frac{n}{2} + 1 = \frac{6}{2} + 1 = 4$ ، $t2 = \frac{n}{2} = \frac{6}{2} = 3$

ومنه يكون الوسيط هو : $ME = \frac{t1+t2}{2} = \frac{31+29}{2} = 30$

المثال الثالث: عينة تتكون من إحدى عشر(11) أسرة تسكن بصفيح العاصمة، نريد معرفة متوسط أفراد الأسرة حسب معطيات الدفتر العائلي: 2 ، 3 ، 5 ، 6 ، 7 ، 9 ، 12 ، 12 ، 14 ، 16 ، 16

الحل:

بعد إجراء العمليات، نجد me = 09 وهذا يعني أن 50 % من حجم هذه العينة عدد أفراد أسرها في المتوسط ≤ 09 أفراد، و 50 % منها ≤ 09 أفراد.

المثال الرابع: عينة تتكون من اثنى عشر(12) عاملا تبين كشوف رواتبهم كما يلي: 15 ، 17 ، 20 ، 26 ، 26 ، 30 ، 35 ، 40 ، 42 ، 46 ، 50 ، 60 ألف دينار / الشهر، نريد معرفة متوسط الراتب لهذه العينة.

الحل:

بعد إجراء العمليات نجد me = 32,5 وبما أن الظاهرة تتعلق بالرواتب، معناه أن متوسط الراتب لهذه العينة من العمال هو : $me = 32.5 \times 1000 = 32500$ دج / الشهر

بمعنى هناك نصف العمال يتقاضون راتبا ≥ 32500 دج / الشهر، وهناك النصف الآخر يتقاضى راتبا

≤ 32500دج / الشهر.

02 ـ الوسيط لبيانات مبوبة: نميز بين حالة المتغير الكمي المنقطع، والمتغير الكمي المستمر

2 ـ 1 ـ حالة متغير كمي منقطع: لحساب الوسيط نتبع الخطوات التالية.

ـ إنجاز التكرار المتجمع الصاعد أو النازل

ـ حساب رتبة الوسيط (t) بغض النظر عن كون أفراد السلسلة (2k+1) أو (2k).

ـ نعيين قيمة الوسيط (me) من القيم (x_i) حسب رتبته.

مثال 1: عينة من طلبة السنة الأولى اقتصاد ، تتكون من 200 طالب(ة) ، سجلنا مدى الالتزام بوقت محاضرة الاحصاء ، فكانت المعطيات حسب الجدول التالي.

الطلبة(F_i)	50	20	35	30	40	25	200
الوقت الضائع بالدقائق (x_i)	1	2	3	4	5	6	$\sum F_i$

أوجد الوسيط الذي يعبر عن الوقت الضائع ؟.

الحل : ننجز جدولا يتضمن التكرار المتجمع الصاعد والنازل

xi	F_i	$F\uparrow$	$F\downarrow$
1	50	50	200
2	20	70	150
(3)	35	(105)	(130)
4	30	135	95
5	40	175	65
6	25	200	25
$\sum F_i$	200	------------------------	-----------------------

ـ رتبة الوسيط: $t = \frac{\sum ni}{2} = \frac{200}{2} = 100$

ـ تحديد الوسيط من الجدول، وهو القيمة الموافقة للترتيب سواء الصاعد أو النازل، وذلك بالبحث عن القيمة 100 بالنسبة للصاعد نجدها عند العدد 105 وفي التكرار النازل نجدها عند العدد 130 وعلى هذا الأساس نجد

Me = 30 أي أن معدل التخلف عن المحاضرة ثلاثة دقائق. وبالتالي يوجد 50 % من الطلبة الغائبون ، مدة غيابهم ≥ 03 دقائق .والباقي $\leq$ 03 دقائق.

مثال 2 : عينة تتكون من 500 حاج(ة) يشكلون السلسلة الاحصائية التالية.

السن xi	20	25	30	35	40	45	50	55	60	70+
عدد الحجاج F_i	10	30	90	110	60	70	40	35	40	15

أحسب الوسيط الذي يحدد متوسط سن الحجاج ؟.

الحل: ننجز جدولا نرتب فيه كل من ($F\uparrow$,$F\downarrow$) ثم نختار أحدهما لتعيين الوسيط بعد حساب رتبته

xi	F_i	$F\uparrow$	$F\downarrow$
20	10	10	500
25	30	40	490
30	90	130	460
35	110	240	370
40	60	300	260
45	70	370	200
50	40	410	130
55	35	445	90
60	40	485	55
70+	15	500	15
$\sum F_i$	500	-----------------------	----------------------

ـ رتبة الوسيط $t = \frac{\sum F_i}{2} = \frac{500}{2} = 250$

ـ على هذا الأساس يكون الوسيط $me = 40$

ـ تفسير النتيجة: 50 % من الحجاج أعمارهم $\geq$ 40 سنة، والباقي أعمارهم $\leq$ 40 سنة.

2 ـ 2 ـ حالة متغير كمي مستمر: للوصول إلى حساب الوسيط نتبع إحدى الطرق الثلاثة التالية

2 ـ 2 ـ 1 ـ بواسطة التكرار المتجمع الصاعد $(F_{\uparrow})$: حيث نقوم بما يلي:

ـ حساب رتبة الوسيط: $\mathbf{t} = \frac{\sum F_i}{2}$

ـ تحديد الفئة الوسيطة وفقا للرتبة التي نتحصل عليها

ـ يحسب الوسيط بالعلاقة الرياضية التالية.

$$me = A(me) + \frac{\frac{\sum F_I}{2} - F_{\uparrow}(F-1)}{F(me)} \times k\,(me)$$

حيث: $A\,(me)$ هو الحد الأدنى للفئة الوسيطة، $F_{\uparrow}\,(F-1)$ هو التكرار المتجمع الصاعد ما قبل الفئة الوسيطة ، $F(me)$ تكرار الفئة الوسيطة الطبيعي ، $k(me)$ طول الفئة الوسيطة.

مثال : عينة من الطلبة تحمل قياسات الطول التالية

متغير الطول(سم) xi	الطلبة F_i	$F \uparrow$
160-155	03	03
165-160	07	10
170-165	09	19
$\{$ 175-170 $\}$	$\{$ 16 $\}$	$\{$ 35 $\}$
180-175	11	46
185-180	04	50
$\sum F_i$	50	---------------------------

ـ رتبة الوسيط: $t = \frac{\sum F_i}{2} = \frac{50}{2} = 25$

ـ الفئة الوسيطية : هي التي تقابل رتبة الوسيط ، وفي الجدول هي الفئة $[170 - 175[$

وبالتالي : الحد الأدنى للفئة الوسيطة هو $A(me) = 170$ ، أما التكرار المتجمع الصاعد ما قبل الفئة الوسيطة فهو 19 ، تكرار الفئة الوسيطة الطبيعي 16 ، وأخيرا طول الفئة الوسيطة هو 05 .

بناء على هذه النتائج يمكن حساب الوسيط بالعلاقة الرياضية التالية

$$me = 170 + \left[\frac{25 - 19}{16}\right] \times 5 = 171{,}875 \cong 1{,}72$$

تفسير النتيجة: يوجد 50 % من طلبة العينة أطوال قاماتهم $\geq$ 1,72 مترا ، كما يوجد 50 % من الطلبة أطوال قاماتهم $\leq$ 1,72 مترا.

2 ـ 2 ـ 2 ـ بواسطة التكرار المتجمع النازل (F↓): نتبع نفس الخطوات التي مرت بنا مع التكرار المتجمع الصاعد.

ـ حساب رتبة الوسيط

ـ تحديد الفئة الوسيطة وفقا للرتبة

ـ حساب me وفقا للمعادلة التالية.

$$me = B\,(me) - \left[\frac{\frac{\sum F_i}{2} - F_{\downarrow}(F+1)}{F(me)}\right] \times k\,(me)$$

حيث $B\,(me)$ الحد الأعلى للفئة الوسيطة، (t) رتبة الوسيط، $F_{\downarrow}(F+1)$ التكرار المتجمع النازل ما بعد الفئة الوسيطة، $F(me)$ تكرار الفئة الوسيطة الطبيعي، $K\,(me)$ طول الفئة الوسيطة.

مثال : بناء على معطيات المثال السابق ، أحسب متوسط طول قامة الطلبة محل الدراسة.

xi	F_i	$F\downarrow$
160-155	03	50
165-160	07	47
170-165	09	40
{175-170 }	{16}	{31 }
180-175	11	15
185-180	04	04
$\sum F_i$	50	------------------------------

ـ نحسب رتبة الوسيط : $t = \frac{50}{2} = 25$

ـ تعيين الفئة الوسيطة من خلال الرتبة، وهي $[170 - 175[$

ـ وعليه ، الحد الأعلى للفئة الوسيطة $B(me) = 175$

ـ التكرار المتجمع النازل ما بعد الفئة الوسيطة هو (15)

ـ تكرار الفئة الوسيطة الطبيعي (16)

ـ طول الفئة الوسيطة (05) وهكذا يمكن الحصول على الوسيط وفق المعادلة الرياضية التالية

$$me = 175 - \left[\frac{25 - 15}{16}\right] \times 5 = 171.875 \cong 1{,}72$$

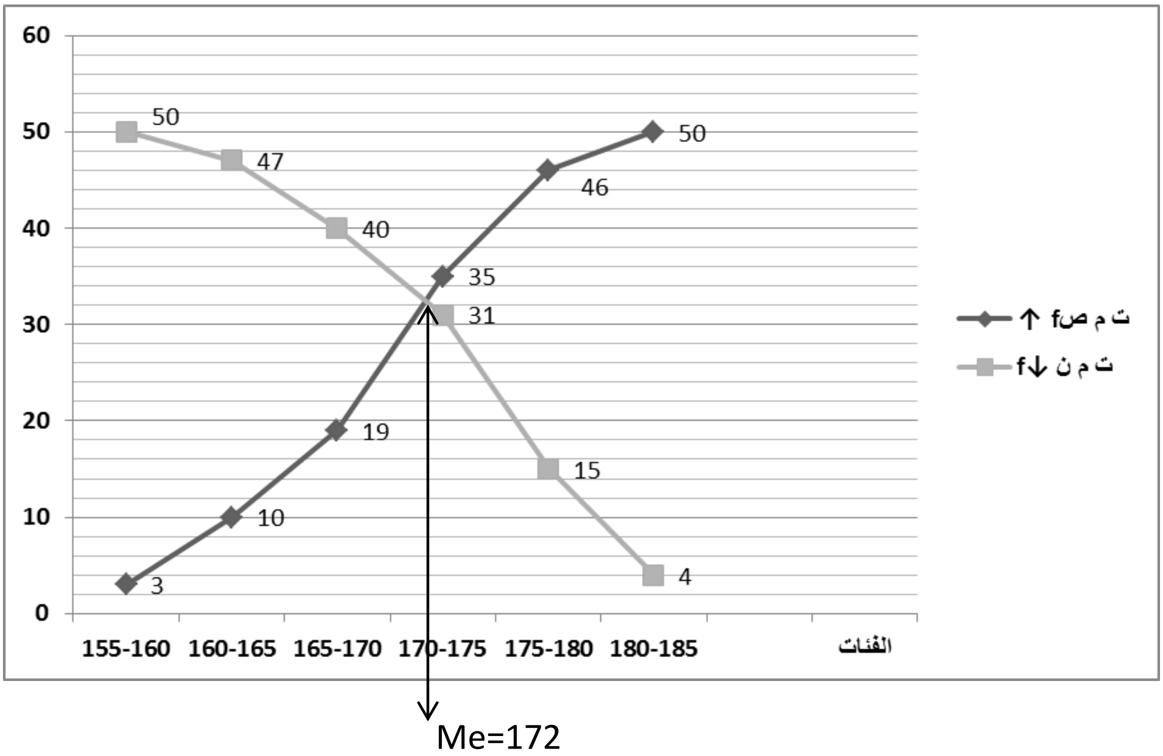
2 ـ 2 ـ 3 ـ الوسيط بيانيا

ـ نقوم برسم معلم متعامد، بحيث نضع في محوره العمودي التكرارات $(F_{\downarrow}^{\uparrow})$ ، وفي المحور الأفقي نضع الفئات.

ـ نقوم برسم منحنى التكرار المتجمع الصاعد، ومنحنى التكرار المتجمع النازل.

ـ نقطة تقاطع المنحنيين نسقطها على محور السينات، فنحصل على الوسيط (me)

مثال: من خلال معطيات المثال السابق استنتج الوسيط.



03 ـ مميزات وعيوب الوسيط: يمكن تلخيص مزايا وعيوب الوسيط من خلال الجدول التالي

	المميزات
01	لا يتأثر الوسيط بالقيم المتطرفة
02	سهولة حسابه، بغض النظر عن كون بياناته مبوبة أو غير مبوبة
03	يمكن حسابه من خلال جداول تكرارية مفتوحة
04	يفضل حسابه في التوزيعات الملتوية
05	يصلح للتعبير عن متوسط الصفات، بمعنى يمكن حسابه للظواهر الغير رقمية التي يمكن ترتيبها، مثل درجات الطلاب أو الحالة الاقتصادية والاجتماعية على اعتبار أنها: عالية جدا، عالية، متوسطة، متدنية، معدومة.
06	مجموع الانحرافات المطلقة عن الوسيط ، أقل من مجموع الانحرافات المطلقة عن أية قيمة أخرى خلافا للوسيط.
07	يقسم المدرج التكراري إلى مساحتين متساويتين
08	يستعمل خاصة في دراسة متغيرات: الأجور، الأسعار، أخطاء القياس، الوفيات، المدة المتوسطة للحياة...الخ
	العيوب

01	يعتمد الوسيط على ترتيب البيانات غير المبوبة، حيث أن هذه العملية تكون أحيانا مجهدة
02	الوسيط لا يستخدم الكثير من القيم المتاحة، بمعنى لا يعتمد في حسابه على كل قيم المتغير محل الدراسة
03	يعتمد الوسيط على قيمة واحدة، أو قيمتين في المنتصف فقط، ولا يأخذ بعين الاعتبار القيم الباقية

III- شبيهات الوسيط (الربيعات ، العشيرات ، المئينات):

إذا كان الوسيط يقسم القيم إلى مجموعتين من حيث العدد ، فإن المقاييس الشبيهة بالوسيط تفيد نفس الغرض ، وتسمى بمقاييس الموضع ، وهي عادة ما تستعمل لمعرفة قيمة الوحدة الاحصائية التي تمثل نسبة معينة من المجتمع محل الدراسة ، كما أنها تتمتع بنفس الخصائص التي رأيناها في مقياس الوسيط.

01 ـ الربيعات : مقياس من مقاييس الموضع ، وهو عبارة عن مجموعة من القيم تقسم التكرار الكلي بنسب معينة ، ونميز فيه ثلاثة قيم ، تقسم التكرار الكلي إلى أربعة أجزاء.

1 ـ 1 ـ الربيعي الأول (الأدنى): وهي القيمة التي يسبقها ربع القيم الأصغر منها.

1ـ 2 ـ الربيعي الثاني (الوسيط) وهي القيمة التي يسبقها نصف القيم الأصغر منها.

1 ـ 3 ـ الربيعي الثالث (الأعلى) وهي القيمة التي يسبقها ¾ القيم الأصغر منها.

الجدول التالي يوضح ترتيب الربيع ، حيث (n) هي حجم العينة أو مجموع التكرارات

الربيعات / البيانات	الربيع الأول $Q1$	الربيع الثاني $Q2$	الربيع الثالث Q3
الغير المبوبة	$t1 = \frac{n+1}{4}$	$t2 = \frac{2(n+1)}{4}$	$t3 = \frac{3(n+1)}{4}$
المبوبة	$t1 = \frac{n}{4}$	$t2 = \frac{2n}{4} = \frac{1}{2}n$	$t3 = \frac{3n}{4}$

1 ـ 4 ـ القوانين المستخدمة في حساب الربيع

1 ـ 4 ـ 1 ـ حالة البيانات غير المبوبة

مثال : أحسب الربيع الأدنى والأعلى لدرجات عينة تتكون من تسعة (09) طلاب كانت على النحو التالي

10 ; 9 ; 12 ; 8 ; 4 ; 5 ; 3 ; 7 ; 6x_i ;

خطوات الحل:

ـ نرتب البيانات ترتيبا تصاعديا : 3 ، 4 ، 5 ، 6 ، 7 ، 8 ، 9 ، 10 ، 12

ـ حساب ترتيب الربيع (Q_1) حيث أن : $t_1 = \frac{n+1}{4} = \frac{9+1}{4} = 2{,}5$

$$Q1 = 4 + \frac{1}{2}\,(5-4)$$

ومنه:

$$Q_1 = 4,5$$

ـ ترتيب الربيع الثالث(Q_3)حيث أن $tQ_3 = \frac{3(n+1)}{4} = 7.5$

ـ قيمة الربيع: تلاحظ معي بأن قيمة الربيع الثالث تقع بين القيمة السابعة والقيمة الثامنة (في نصف المسافة بينهما). وعلى هذا الأساس تعالج تماما مثل الربيع الأول حيث:

قيمة الربيع الثالث = القيمة السابعة +2/1 (القيمة الثامنة – القيمة السابعة)، أي أن:

$$Q3 = 9 + \frac{1}{2}(10 - 9) = 9,5$$

1 – 4 – 2 ـ حالة البيانات المبوبة

مثال: أحسب الربيع (الأول والثالث) من المعطيات الموجودة في الجدول أدناه

xi	F_I
5-1	07
10-6	10
15-11	16
20-16	32
25-21	24
30-26	18
35-31	10
40-36	05
45-41	01
$\sum F_I$	123

الحل : نستعين بالجدول التالي الذي يحتوي على التكرار المتجمع الصاعد

xi	F_I	$F\uparrow$
05-01	07	07
10-06	10	17
{15-11}	{16}	{33}
20-16	32	65
25-21	24	89
{30-26}	{18}	{107}
35-31	10	117

40-36 45-41	05 01	122 123
$\sum F_I$	123	---------------------------

ـ ترتيب الربيع الأدنى $t_1 = \frac{\sum F_i}{4} = \frac{123}{4} = 30.75$

ـ تحديد الفئة الربيعية: حسب الترتيب هي $[11-15[$

ـ قيمة الربيع الأدنى: $Q_1 = A + \frac{\frac{\sum F_I}{4} - F\uparrow(Fi-1)}{F_I(Q_1)} \times K$

$$Q1 = 11 + \frac{30,75 - 17}{16} \times 4 = 14,44$$

ـ ترتيب الربيع الأعلى: $t3 = \frac{3(ni)}{4} = \frac{3(123)}{4} = 92,25$

ـ تحديد الفئة الربيعية: من خلال الترتيب يتضح بأنها الفئة $[26-30[$

ـ حساب الربيعي الثالث (الأعلى) $Q3 = 26 + \frac{92,25-89}{18} \times 4 = 26,72$

ملاحظة: الربيع الثاني يساوي تماما الوسيط، لذلك لا داعي لحسابه، إلا في حالة الطلب.

02 ـ القوانين المستخدمة في حساب العشير والمئين

هي نفسها القوانين المستخدمة في الربيع ، مع مراعاة الفرق في حساب الرتبة (الترتيب) ، بحيث إذا كنا بصدد حساب العشير تكون معادلة الترتيب هي :

2ـ 1 ـ في حالة بيانات غير مبوبة:

ـ ترتيب العشير $t_d = \frac{i(F+1)}{10}$ ، أما المئين فيكون ترتيبه بشكل عام $t_p = \frac{i(F+1)}{100}$

مثال : من المثال السابق الخاص ببيانات غير مبوبة أحسب كل من العشير السابع (7) والمئين الخامس والعشرون (25).

ـ رتبة العشير $t_{d_7} = \frac{7(9+1)}{10} = 7$ ومنه يمكن تعيين قيمته والتي تتوافق مع الرتبة إذ نجد $d_7 = 9$

ـ أما رتبة المئين فهي $t_{p25} = \frac{25(9+1)}{1} = \frac{250}{100} = 2,5$ ومنه $p_{25} = 4 + \frac{1}{2}(5-4) = 4,5$

نستنتج من خلال ما سبق أن $Q_1 = P_{25} = 4{,}5$ يمكن استنتاج مساواة أخرى غير هذه.

IV – المنوال

هو من أهم مقاييس النزعة المركزية، ويأتي في المرتبة الثالثة بعد الوسط الحسابي والوسيط. غالبا ما يستعمل لمعالجة الظواهر ذات المتغيرات الكيفية.

يعرف بأنه القيمة الاحصائية الأكثر شيوعا (انتشارا) أو تكرارا ، يرمز له بالرمز (m_o).نميز فيه اربعة أنواع .

01 – توزيع تكراري وحيد المنوال : بمعنى أن السلسلة الاحصائية لا تحتوي إلا على قيمة إحصائية واحدة شائعة.

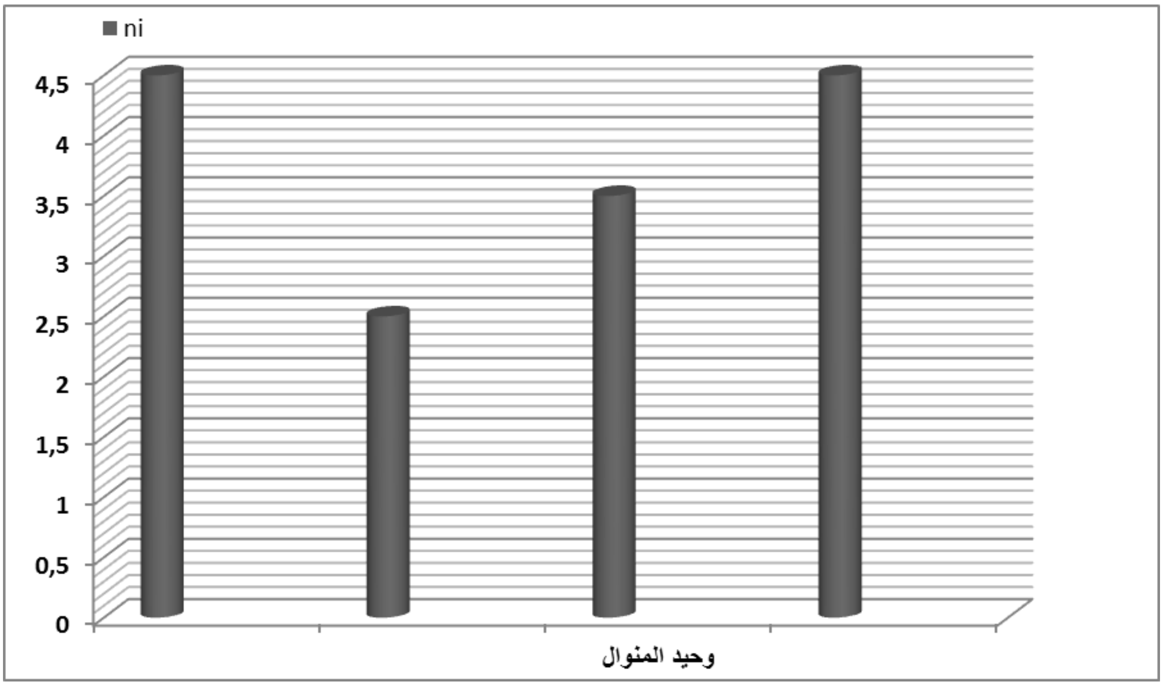
مثال : لتكن لدينا القيم الاحصائية : 2 ، 4 ، 3 ، 7 ، 4 ، 1 ، 4 ، نلاحظ أن القيمة اربعة (4) هي الوحيدة الشائعة ، وعليه يكون المنوال : $mo = 4$.

02 – توزيع تكراري ثنائي المنوال: بمعنى يحتوي التوزيع على قيمتين شائعتين دون غيرهما.

مثال : لتكن لدينا القيم الاحصائية : 2 ، 4 ، 2 ، 7 ، 4 ، 1 ، نلاحظ أن القيمتين 2 و 4 هما الشائعتين ، وبالتالي المنوال هو : $m1 = 2, m2 = 4$.

03 – توزيع تكراري متعدد المنوال : أي أن السلسلة الاحصائية تحتوي على أكثر من قيمتين شائعتين ، في هذه الحالة لا يكون للمنوال أي مذلول إحصائي يذكر ، وبالتالي لا يستعمل كمقياس من مقاييس النزعة المركزية إلا في بعض الحالات الاستثنائية.

مثال : الرسم البياني التالي يوضح : توزيع تكراري وحيد المنوال



04 ـ حساب المنوال

4 ـ 1 ـ حالة بيانات غير مبوبة

مثال: طالب تحصل على النقاط التالية تتعلق بثماني مواد: 07، 12، 09، 14، 16، 10، 09، 13.

المنوال : هو القيمة الاحصائية الأكثر شيوعا من غيرها ، وبالتالي $mo = 09$.

مثال آخر : لدينا السلسلتين $(A\,;B)$ تترجمان رقم المبيعات لمنتو جين خلال أسبوع ، بالدينار الجزائري

$$A\ ;\ 5000\,,7520\,,9480\,,3900\,,7520\,,8000\,,3900$$

$$B\,;11000\,,16000\,,21000\,,12000\,,11000\,,21000\,,12000$$

نلاحظ من خلال معطيات السلسلتين أن:

السلسلة (A) تتميز بثنائي المنوال: $m_o \ = 3900{,}7520$

السلسلة (B) تتميز بتعدد المنوال: $m_o(B) = 11000\,,21000\,,12000$

4 ـ2 ـ حالة بيانات مبوبة

لحساب المنوال في حالة البيانات المبوبة ، نراعي ما يلي.

4 – 2 – 1 – حالة متغير كمي منقطع

مثال : كان معدل محصول مزرعة نموذجية في إنتاج الطماطم للهكتار الواحد ، خلال خمس سنوات على النحو التالي

05	04	03	02	01	السنة
5,3	04	4,2	05	3,5	طن/ هك

أحسب المنوال لمعطيات هذا الجدول ؟.

بدون شك أننا ننظر إلى الهكتار الأكثر إنتاجا خلال الفترة محل الدراسة ، بمعنى $m_o = 5{,}3$

مثال آخر: كان معدل الانتحار في أوساط الشباب خلال السداسي الأول من سنة 2014 في الجزائر كما يلي:

06	05	04	03	02	01	ni
03	08	05	10	07	04	xi

تلاحظ أن : $mo = 10$

مثال آخر : ما هي قيمة المنوال بناء على معطيات الجدول التالي.

14	04	18	09	12	05	ni
09	03	08	05	02	07	xi

المنوال هو القيمة الأكثر تكرار $m_o = 80$

4 – 2 – 2 – حالة متغير كمي مستمر

إذا كنا أمام توزيع تكراري مرتب على أساس الفئات نقوم بما يلي

ـ نحدد الفئة المنوالية : وهي الفئة التي تكون أمام أكثر تكرار طبيعي كان أو معدل.

ـ نحسب المنوال بطريقة $Pearson$ أو ما يسمى بقانون الفروق. $m_o = A_O + \frac{\Delta1}{\Delta1+\Delta2} \times K_{mo}$

حيث أن : A_O هو الحد الأدنى للفئة المنوالية ، K_{mo} طول الفئة المنوالية ، $\Delta1$ هو الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها ، $\Delta2$ هو الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.

مثال : التوزيع التالي ، يبين سن عينة من الحجاج الجزائريين لسنة ـ ما ـ والمطلوب حساب المنوال

xi	F_I
25-20	50
30-25	70
35-30	100
40-35	250
45-40	200
50-45	170
$\sum F_I$	840

ـ الفئة المنوالية : هي الفئة التي تقابل أكثر تكرار $[35-40[$

ـ الحد الأدنى للفئة المنوالية $A_{O=35}$ ، $\Delta1 = (250-100) = 150$ ،

$\Delta2 = (250-200) = 50$ ، $K_{MO} = 05$ وبناء على هذه المعطيات، يمكن حساب المنوال

$$mo = 35 + \left(\frac{150}{150+50}\right) \times 5 = 38{,}75 \cong 39$$ سنة

فائدة : مركز الفئة المنوالية: $\acute{x} = \frac{35+40}{2} = 37.5$ يسمى بالمنوال التقريبي، وفي حالة التوزيع المتماثل فإن: $\bar{x} = m_e = m_o$

4 ـ 2 ـ 3 ـ حالة متغير كمي مستمر لفئات ذات أطوال غير متساوية

للحصول على قيمته نتبع الخطوات التالية:

ـ نقوم بتعديل التكرارات وفق القاعدة التي مرت بنا $ER = \frac{F_i}{K_i}$

ـ نستخرج الفئة المنوالية ذات التكرار المعدل والأكثر شيوعا

ـ نحسب المنوال بالعلاقة

$$m_o = A + \left(\frac{\acute{\Delta}1}{\acute{\Delta}1+\acute{\Delta}2}\right) \times K_{mo}$$

حيث أن :

$$\acute{\Delta}1 = ER_{mo} - ER_{(MO-1)}$$

$$\Delta 2 = ER_{mo} - ER_{(mo+1)}$$

A_{mo} الحد الأدنى للفئة المنوالية ذات التكرار المعدل

K_{mo} طول الفئة المنوالية

مثال : في الشهر الماضي ، سجل مدير واجهة لبيع تجهيزات ولوازم الاعلام الآلي سعر البيع بالدينار الجزائر ، كما سجل عدد الوحدات المباعة الموضحين في الجدول التالي

سعر البيع xi	عدد الوحدات المباعة F_i	طول الفئة k_i	التكرار المعدل
150-50	80	100	0,8
250-150	160	100	1,6
350-250	720	100	7,2
[400-350]	1680	50	[33,6]
500-400	2720	100	27,2
650-500	1760	150	11,73
750-650	640	100	6,4
850-750	160	100	1,6
950-850	80	100	0,8
$\sum F_I$	8000	-----------------------	-----------------------

من خلال الجدول المنجز يمكن تحديد الفئة المنوالية بسهولة ، وهي التي تقابل أكثر تكرار عدلناه ، وليس التكرار الطبيعي ، وهو $ER = 33{,}6$ ، وهكذا تكون الفئة المنوالية هي $[350 - 400[$ ، وبالتالي يمكن حساب المنوال وفق العلاقة السابقة .

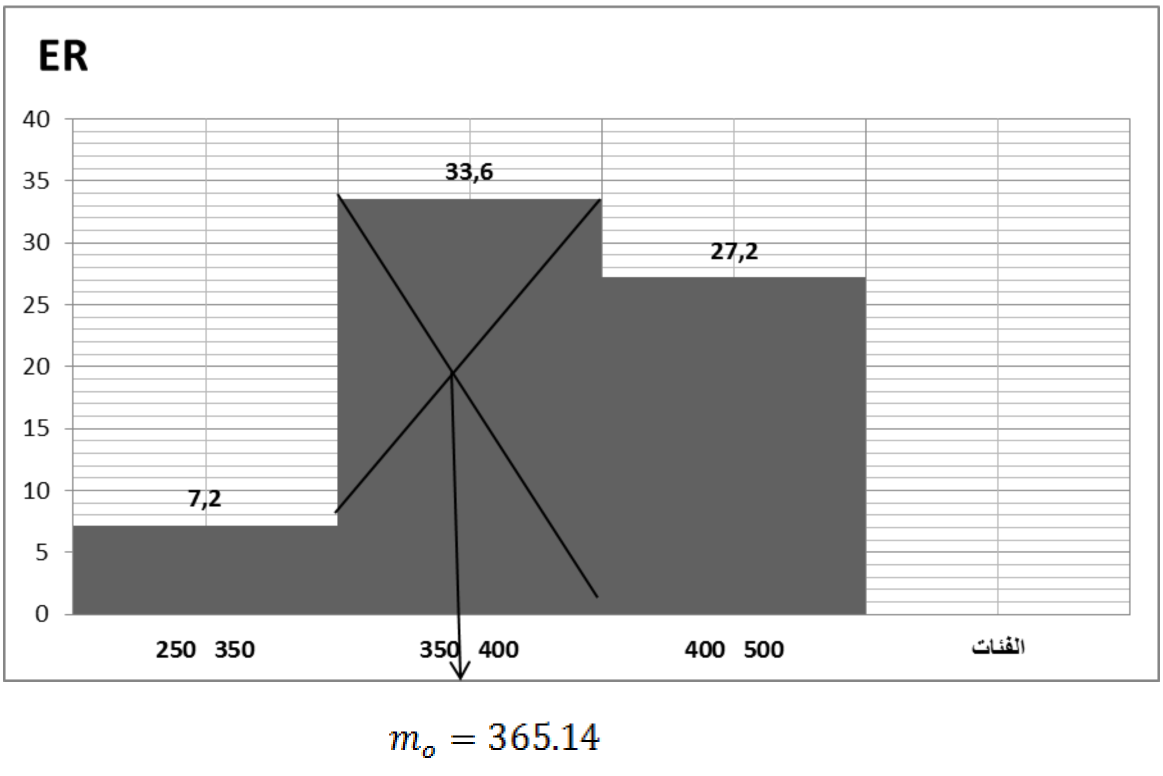
$$m_o = A + \left(\frac{33{,}6 - 7{,}2}{(33{,}6 - 7{,}2) + (33{,}6 + 27{,}2)}\right) \times 50 = 365{,}14$$

سؤال : ما هو المنوال التقريبي ؟.

المنوال التقريبي هو: $\acute{x}4 = \frac{350+400}{2} = 375$

4 ـ 3 ـ حساب المنوال بالطريقة البيانية

من خلال المثال السابق ، يمكن حساب المنوال بالطريقة البيانية ، إذ نكتفي بالفئة المنوالية ، والفئة التي قبلها والفئة التي بعدها، ثم نقوم برسم هذه الفئات حسب حجم كل فئة (قيمة التكرار المعدل الذي يوافقها).



$m_o = 365.14$

ملاحظة : العلاقة بين الوسط الحسابي والوسيط والمنوال تأخذ واحدة من ثلاثة حالات

$\bar{x} = \mathrm{m_e} = \mathrm{m_o}$ ⟸ توزيع تكراري متناظر

$\bar{x} > \mathrm{m_e} > \mathrm{m_o}$ ⟸ توزيع تكراري غير متناظر من اليمين

$\bar{x} < m_e < m_o$ ⟸ توزيع تكراري غير متناظر من اليسار

05 ـ مميزات وعيوب المنوال

المميزات	
01	يمكن حساب المنوال من الجداول المفتوحة
02	لا يتأثر بالقيم المتطرفة
03	من أفضل المتوسطات لتمثيل المعطيات الغير رقمية
04	سهولة حسابه بالمعادلة أو الرسم
العيوب	
01	أسوأ مقياس من مقاييس النزعة المركزية، لأنه يتأثر بطول الفئة وعدد الفئات
02	من غير المستحسن استخدامه من خلال معطيات كثيرة الالتواء
03	لا يأخذ أثناء حسابه كل المعطيات
04	من الصعب الاعتماد عليه، لأن قيمته تتفاوت عند حسابها لنفس المعطيات بعدة طرق
05	أحيانا لا يوجد منوال لأن معطياته لا تتكرر ، وأخرى تجد أكثر من منوال

عموما ،يعتبر الوسط الحسابي أكثر مقاييس النزعة المركزية استخداما، أما المنوال فهو أقل مقاييس النزعة المركزية استخداما.

V ـ الوسط الهندسي

الوسط الهندسي يرمز له عادة بالرمز MG ويعرف على أنه الجذر النوني لحاصل ضرب مجموعة من القيم عددها n إذ يعطى بالعلاقة $G = \sqrt[n]{(x1).(x2).(x3)\ldots\ldots\ldots.xn}$ حيث أن

G : هو الوسط الهندسي ، $x1 \;,\; x2 \;,\; x3 \;,\ldots\ldots\ldots, xn$ هي القيم المشاهدة.

يستخدم الوسط الهندسي في دراسة الظواهر التي تزيد مفرداتها بنسبة ثابتة ، دراسة النمو السكاني ، نمو الحيوانات ، الحشرات ، الطيور والبكتيريا ، دراسة النمو الاقتصادي ، متوسط التضخم ، الأرقام القياسية لمناسب الأسعار ، معدل التغير في المبيعات ، معدلات الفائدة ، متوسط السرعة.

01 ـ الوسط الهندسي لبيانات غير مبوبة

لتكن قيم المتغير الاحصائي محل الدراسة هي : $x1 \;,\; x2 \;,\; x3 \;,\ldots\ldots, xn$

الوسط الهندسي هو $M_G = \sqrt[N]{(x1).(x2).(x3).,\ldots,.(xn)}$ ندخل اللوغاريتم العشري على طرفي المعادلة فيكون

$$logM_G = \frac{1}{n}\sum(logxi) \Leftrightarrow \frac{1}{n}[logx1 + logx2+,\ldots,+logxn] \Leftrightarrow$$

$$\boldsymbol{M_G = 10^{\frac{\sum logxi}{n}}}$$

مثال : كان معدل النمو الاقتصادي للفترة (2010-2006) على النحو التالي

السنة	06	07	08	09	10
معدل النمو الاقتصادي xi	2,0	3,0	2,4	2,4	3,3

المصدر : **الديوان الوطني للإحصاء**

السؤال : ما هو متوسط النمو الاقتصادي للفترة محل الدراسة ؟.

الحل :

$$M_G = \sqrt[5]{(102)(103)(102{,}4)(102{,}4)(103{,}3)}$$

$$logM_G = \frac{1}{5}\{log102 + log103 + log102{,}4 + log102{,}4 + log103{,}3\}$$

$$logM_G = \frac{1}{5}\{2 + 2{,}01 + 2{,}01 + 2{,}01 + 2{,}01\} = \frac{10{,}04}{5} = 2{,}008$$

$$M_G = \mathbf{10^{2{,}008}} \Rightarrow M_G = 101{,}85 \cong 102\%$$

أي أن المعدل المتوسط للنمو الاقتصادي خلال الفترة محل الدراسة هو 2 %.

مثال آخر : مؤشر الأسعار العام للفترة (2010 - 2006) كان على النحو التالي

السنة	2006	07	08	09	10
المؤشر العام xi	2,31	3,68	4,86	5,74	3,91

المصدر : الديوان الوطني للإحصاء

المطلوب: حساب متوسط الأسعار العام للفترة محل الدراسة ؟ .

$$log\ \ M_G = [log102{,}31 + log103{,}68 + log104{,}86 + log105{,}74 + log103{,}91]$$

$$= \frac{1}{5}[2{,}01 + 2{,}01 + 2{,}02 + 2{,}02 + 2{,}02]$$

$$= \frac{1}{5}(10{,}08) = 2{,}016$$

$$\Rightarrow \mathrm{M_G} = 10^{2{,}016} \Rightarrow\ \mathrm{M_G} = 103{,}75\%$$

أي أن متوسط معدل الأسعار العام للفترة محل الدراسة هو *3,75* %

02 ـ الوسط الهندسي لبيانات مبوبة

يعطى بالعلاقة الرياضية التالية

$$M_G = \sqrt[\sum ni]{x1^{n1}.x2^{n2}.x3^{n3},\dots,.xn^{nk}}$$

$$\log M_G = \frac{1}{\sum ni}\ .\sum ni.logxi \Leftrightarrow \log M_G = \frac{\sum ni.logxi}{\sum ni} \Leftrightarrow$$

$$\boldsymbol{m_g = 10^{\frac{\sum ni.logxi}{\sum ni}}}$$

حيث أن

$$M_G = \frac{1}{\sum}[n1logx1 + n2logx2 + n3logx3, \dots, +nklogxk]$$

مثال : نسبة التغير في الأجر القاعدي خلال الفترة (2000 ـــ 2011) كانت على النحو التالي

الفترة	2004-2000	07-05	09-08	11-10
نسبة التغير xi	% 40	27	30	25

المطلوب: حساب متوسط التغير في الأجر القاعدي خلال الفترة محل الدراسة

$$M_G = \sqrt[12]{(140^5).(127^3).(130^2).(125^2)}$$

$$\log M_G = \frac{1}{12}[5log140 + 3log127 + 2log130 + 2log125]$$

$$log\ M_G = \frac{1}{12}[10{,}73 + 6{,}31 + 4{,}23 + 4{,}19]$$

$$logM_G = \frac{1}{12}(25{,}46) \Leftrightarrow logM_G = 2{,}12 \Rightarrow M_G = 10^{2{,}12} = 131{,}82\%$$

وهذا يعني أن متوسط نسبة الزيادة في التغير للأجر القاعدي هو $31{,}82$ % .

03 ـ مزايا وعيوب الوسط الهندسي

نلخص أهم مزايا وعيوب الوسط الهندسي في الجدول التالي

<table>
<tr><th colspan="3">المميزات</th></tr>
<tr><td>01</td><td colspan="2">تعريفه واضح ومحدد وقيمته جيدة</td></tr>
<tr><td>02</td><td colspan="2">جميع القيم المشاهدة تدخل في حسابه</td></tr>
<tr><td rowspan="2">03</td><td colspan="2">أقل تأثرا بالقيم المتطرفة من الوسط الحسابي
مثال : قارن بين الوسط الحسابي والوسط الهندسي بناء على القيم التالية : 2 ، 4 ، 1000</td></tr>
<tr><td>الوسط الحسابي
$\bar{x} = \frac{2+4+1000}{3} = \frac{1006}{3}$
$= 335{,}3$</td><td>الوسط الهندسي
$M_G = \sqrt[3]{(2).(4).(1000)} = \sqrt[3]{8000}$
$= 20$</td></tr>
</table>

	لا حظ تأثير القيمة المتطرفة 1000 على نتيجة حساب كل من الوسط الحسابي والوسط الهندسي، وهو كما ترى فرق شاسع.
04	الوسط الهندسي أكثر ملائمة من الوسط الحسابي في الحصول على وسط النسب المختلفة
05	يصلح الوسط الهندسي لتمثيل البيانات، حتى وإن اتصفت باختلافات كبيرة
06	هو أهم مقياس لقياس معدلات النمو والمواليد والتغير في مستوى الأسعار
العيوب	
01	لحسابه، يحتاج إلى عمليات معقدة نوعا ما
02	لا يمكن حسابه في حالة وجود قيم مشاهدة صفرية أو سالبة
03	لا يمكن حسابه من جداول مفتوحة
04	لا يمكن الحصول عليه من خلال الرسومات البيانية

VI- الوسط التوافقي

يستعمل عادة في تحديد متوسط العلاقة بين متغيرين تربطهما علاقة عكسية، مثل السرعة والزمن، العرض والطلب، سعر الصرف بين عملتين، الاستثمار والبطالة...الخ.

يرمز له عادة بالرمز (H) ويعرف على أنه مقلوب الوسط الحسابي لمقلوبات القيم المشاهدة.

01 – الوسط التوافقي لبيانات غير مبوبة

إذا كان لدينا (n) من القيم المشاهدة: $x1 \;,\; x2 \;,\; x3 \;,\dots\dots, xn$ فإن الوسط التوافقي لهذه القيم هو

$$\boldsymbol{H = \frac{n}{\sum \frac{1}{xi}} = \frac{n}{\frac{1}{x1}+\frac{1}{x2}+\frac{1}{x3}+,\dots,+\frac{1}{xn}}}$$

مثال : مؤسسة النقل العمومي بين الولايات ، أرادت دراسة متوسط السرعة ، فاختارت لذلك اربعة (04) حافلات بسرعة محددة على النحو التالي

الحافلات	1	2	3	4
السرعة كم/ سا xi	70	80	100	120

المطلوب : حساب متوسط السرعة لهذه العينة من الحافلات.

$$H = \frac{n}{\sum \frac{1}{xi}} = \frac{4}{\frac{1}{70} + \frac{1}{80} + \frac{1}{100} + \frac{1}{120}} = 88{,}88\ km/h$$

مثال آخر: تبين معطيات معدل صرف الدينار الجزائري مقابل الدولار للفترة (2009 – 2011) على النحو التالي

السنة	2009	2010	2011
د ج / $	72,65	74,41	74,00

ما هو متوسط معدل سعر الصرف للفترة محل الدراسة.

$$h = \frac{3}{0{,}0406} = 73{,}89\ DA/\$$$

02 – الوسط التوافقي لبيانات مبوبة

مثال: مؤسسة اقتصادية تنتج مادة السكر، حيث تتوفر على عشرة(10) آلات إنتاج موزعة حسب الكمية المنتجة في الساعة (طن/ساعة)، كما في الجدول التالي.

معدل الانتاج المتوسط xi	70	60	40	$\sum F_I$
عدد الآلات F_I	02	04	04	10

المطلوب : متوسط الانتاج لمجموع الآلات ؟.

للحصول على الوسط التوافقي لبيانات مبوبة، نأخذ بعين الاعتبار التكرارات الواردة في الجدول، لذلك تعطى العلاقة نظريا بالشكل التالي

$$H = \frac{\sum F_I}{\sum \frac{xk}{x1}} = \frac{F_1}{\frac{F_1}{x_1} + \frac{F_2}{x_2} + \frac{F_3}{x_3} + ,,,,, + \frac{F_K}{x_k}}$$

وبناء على هذا الأساس يكون حساب المعدل المتوسط للإنتاج الكلي هو:

$$H = \frac{10}{\frac{04}{40} + \frac{04}{60} + \frac{02}{70}} = 51{,}21 \;\; t/h$$

مثال آخر : مؤسسة اقتصادية تمارس نشاط النقل للبضائع بين ميناء الجزائر والمنطقة التجارية بالحميز ، تريد هذه المؤسسة معرفة متوسط السرعة لعينة تتكون من اثنان وعشرون (22) شاحنة موزعة على النحو التالي

السرعة المتوسطة xi (كلم/سا)	40	50	60	70	$\sum F_i$
عدد الشاحنات F_i	05	07	04	06	22

السؤال : ماهو متوسط سرعة عينة الشاحنات محل الدراسة.

$$H = 52{,}70 \; k_m/h$$

مثال آخر : عينة تتكون من مائة (100) تاجر تمارس تجارة الجملة ، أجرينا دراسة حول المحفزات الضريبية والتصريح الضريبي ، فكانت معطيات العينة على النحو التالي

المحفز الضريبي xi %	10-05	15-10	20-15	25-20	30-25	$\sum F_i$
عدد التجار المصرحين F_i	05	15	20	25	35	100

المطلوب: حساب كل من المتوسطات: $\bar{X}, M_G \;, H$

لتسهيل الحصول على هذه المتوسطات، نستعين بالجدول التالي

xi	F_i	$\acute{x}_i$	$F_i/\acute{x}_i$	$F_i.\acute{x}$	$log\acute{x}_i$	$F_i.log\acute{x}_i$
10-05	05	7,5	0,66	37,5	0,87	4,35

15-10 20-15 25-20 30-25	15 20 25 35	12,5 17,5 22,5 27,5	1,20 1,14 1,11 1,27	187,5 350 562,5 962,5	1,09 1,24 1,35 1,44	16,35 24,80 33,75 50,40
$\sum F_i$	100	----------	5,38	2100	----------	129,65

ـ الوسط الحسابي(%) : $\bar{x} = \frac{\sum F_i \acute{x}_i}{\sum F_i} = \frac{2100}{100} = 21$

ـ الوسط الهندسي (%): $m_g = 19,79$

ـ الوسط التوافقي (%): **H = 18,58**

استنتاج هام: من خلال النتائج تلاحظ معي أن : $\bar{x} > m_g > h$ وهذا يعني أن التوزيع غير طبيعي ، بمعنى غير منتظم (غ متناظر أو غ متماثل)

03 ـ مزايا وعيوب الوسط التوافقي

مثله مثل المتوسطات التي سبقته، فهو لا يخلو من مزايا وعيوب، لذلك سوف نوجزها في الجدول التالي

<table>
<tr><td colspan="3">المميزات</td></tr>
<tr><td>01</td><td colspan="2">هو المقياس الملائم في حالة متغيرين تربطهما علاقة عكسية ، مثل السرعة بالنسبة للزمن ، الانتاج بالنسبة للزمن...الخ</td></tr>
<tr><td>02</td><td colspan="2">عند قياسه يأخذ بعين الاعتبار كل قيم العينة</td></tr>
<tr><td rowspan="2">03</td><td colspan="2">أقل تأثرا بالقيم المتطرفة ، مقارنة بالوسط الحسابي ، قارن على سبيل المثال بين الوسطين بالنسبة للقيم التالية: 04 ، 08 ، 90 .</td></tr>
<tr><td>الوسط الحسابي
$\bar{x} = \frac{\sum xi}{n} = \frac{102}{3} = 34$</td><td>الوسط التوافقي
$H = \frac{n}{\sum \frac{1}{xi}} = \frac{3}{\frac{1}{4} + \frac{1}{8} + \frac{1}{90}} \cong 08$</td></tr>
</table>

العيوب	
عملياته الحسابية معقدة نوعا ما	01
لا يمكن حسابه في حالة وجود قيم صفرية أو سالبة	02
لا يمكن حسابه من جداول مفتوحة	03
لا يمكن حسابه بيانيا	04

VII– الوسط التربيعي

يعرف على أنه الجذر التربيعي لمتوسط مربعات قيم المتغير الاحصائي . يرمز له عادة بالرمز (MQ) يستخدم بكثرة في التطبيقات الطبيعية.كما يعطي أهمية متزايدة للقيم الشاذة للمتغير الاحصائي ، كما يتعامل مع القيم السالبة ، خلافا للمتوسطات السابقة.

01 – الوسط التربيعي لبيانات غير مبوبة : لحسابه نطبق العلاقة الرياضية التالية

$$MQ = \sqrt{\frac{\sum xi^2}{n}}$$

مثال : أوجد الوسط التربيعي للقيم التالية : 02 ، 05 ، 03 ، 06

$$MQ = \sqrt{\frac{2^2+5^2+3^2+6^2}{4}} = \sqrt{\frac{74}{4}} = 4{,}30$$

02 – الوسط التربيعي لبيانات مبوبة

2 – 1 – الوسط التربيعي لمتغير كمي منقطع

مثال : سجل النمو الاقتصادي لعينة من المؤسسات الانتاجية خلال فترة الدراسة البيانات التالية

70	10	40	20	المؤسسات الانتاجية F_i
$\sum F_i$	1-	2	3	النمو الاقتصادي xi %

المطلوب : حساب متوسط النمو لهذه العينة من المؤسسات الانتاجية.

الحل : في مثل هذه الحالات ، نأخذ بعين الاعتبار ترجيح النمو لكل مجموعة من المؤسسات ، وعليه نطبق الوسط التربيعي المرجح $M_Q = \sqrt{\frac{\sum F_i x_i^2}{\sum F_i}}$ وحتى تكون العمليات الحسابية سهلة، نستعين بالجدول التالي

xi	F_i	xi^2	$F_i x_i^2$
03 02 01-	20 40 10	09 04 01	180 160 10
$\sum F_i$	70	----------------------	350

ومنه يكون الوسط التربيعي

$$MQ = \sqrt{\frac{350}{70}} = \sqrt{5} = 2{,}236$$

2 – 2 ـ الوسط التربيعي لمتغير كمي مستمر

يعالج تماما مثل الوسط التربيعي المرجح ، والاختلاف يكمن فقط في أخذ قيم مراكز الفئات بدل القيم الأصلية.

مثال : أجرينا دراسة على رواتب العاملين من خلال عينة تتكون من خمسون (50) عاملا لإحدى المؤسسات كانت موزعة على النحو التالي

xi فئات الأجر (1000) دج	العمال (F_i)
20-15 25-20 30-25 35-30 40-35 45-40 50-45 +50	10 14 10 06 04 03 02 01
$\sum F_i$	50

لحساب المتوسط التربيعي نستعين بجدول نبرز فيه مراكز الفئات على النحو التالي

xi	$\acute{x}i$	$\acute{x}i^2$	F_i	$F_i \acute{x}_i^{\ 2}$
20-15	17,5	306,25	10	3062,5
25-20	22,5	506,25	14	7087,5
30-25	27,5	756,25	10	7562,5
35-30	32,5	1056,25	06	6337,5
40-35	37,5	1406,25	04	5625
45-40	42,5	1806,25	03	5418,75
50-45	47,5	2256,25	02	4512,5
+50	52,5	2756,25	01	2756,25
$\sum F_i$	----------------	---------------	50	42362,5

وعلى هذا الأساس نحسب الآن الوسط التربيعي وفق العلاقة الرياضية التالية

$$M_Q = \sqrt{\frac{\sum F_i x_i^{\ 2}}{\sum F_i}}$$

$$M_Q = \sqrt{\frac{42362,5}{50}} = \sqrt{847,25} = 29,10$$

وعليه ، يكون الوسط التربيعي لفئات الأجر هو(29.10) ×(1000) = 29100 دج

فائدة

المتوسطات الحسابية المرجحة والتوافقية المرجحة والهندسية العادية (غير المرجحة) تستعمل لحساب المؤشرات المركبة للأسعار أو الكميات.وفي نفس السياق نذكر بالعلاقة الموجودة بين المتوسطات حسب حالة التوزيع.

ـ إذا كان التوزيع غير طبيعي فإن: $H < G < \bar{X} < Q$

ـ أما إذا كان طبيعيا فإن المتوسطات الأربعة متساوية: $H = G = \bar{X} = Q$.

وبصورة عامة تستعمل المتوسطات السابقة وفق المعادلة $f(x) = x^{\propto}$ والتي تعرف وفق الجدول التالي

$\propto$	1-	0	1	2
نوع الوسط	H	G	$\bar{X}$	Q

03 ـ تمارين محلولة

التمرين الأول : أختبر نفسك في ستة (06) أسئلة بعبارة (نعم) أم (لا)

ـ الوسيط هو قيمة المتغير الكمي التي تربط بين منحنى المتجمع الصاعد ومنحنى المتجمع النازل.

ـ معامل التغير هو عدد دون أهمية

ـ الربيعات ،العشيرات ، المئينات ،هي من عائلة الكميات

ـ المنوال المطلق لتوزيع ـ ما ـ يتطابق مع قيمة المتغير المحصور بين قيمتين مختلفتين

ـ المتغير الكمي المنقطع الآتي: (1، 2، 3، 3، 3، 4، 5، 6) يقبل كوسيط القيمة (3.5)

ـ توجد علاقة بين مختلف المتوسطات

التمرين الثاني (حالة عدم الترجيح)

كانت معاينة اوراق ثمان (08) طلبة تتعلق باختبار مقيم على أساس خمسون (50) نقطة ، تمثل السلسلة الاحصائية التالية: 25 ، 43 ، 35 ، 20 ، 32 ، 30 ، 35 ، 24

المطلوب : حساب وشرح الخلاصات الاحصائية التالية

ـ المنوال، الوسيط، متى يتحقق الوسيط حينما نضيف القيمة 45 إلى السلسلة.

ـ الوسط الحسابي ، التربيعي ، الهندسي ، التوافقي.

التمرين الثالث (حالة المرجح المنقطع)

مؤسسة صناعية تبيع مختلف الآلات، يهمنا دراسة عدد الآلات المباعة في اليوم، وبناء على هذا الأساس، يعرف المتغير محل الدراسة بالمتغير(X)، حيث رصدنا المعطيات المتعلقة بالبيع لمدة 600 يوم مفتوحة في الجدول التالي

عدد أيام البيع	عدد الآلات المباعة في اليوم
98	00
232	01
119	02
85	03
50	04
16	05

المطلوب:

ـ احسب التكرار النسبي للمبيعات والتكرار النسبي الصاعد ، العرض البياني للتكرارات النسبية

ـ قدم تعريف وحساب كل من: المنوال، الوسيط، الوسط الحسابي.

التمرين الرابع (حالة المرجح المستمر)

دراسة تتعلق بعينة من زوار احد المتاحف، نريد من خلالها معرفة أثر صفة العمر والتي تبدأ من خمسة عشر(15) سنة فما فوق حسب الجدول التالي.

عدد الزائرين	العمر (سنة)
096	من 15 إلى أقل من 25 سنة
118	25 ــــ 35
138	35 ــــ 50
101	50 ــــ 65
047	65 سنة فما فوق

ـ عرف المتغير محل الدراسة المشترك مع صفة " العمر" ثم أحسب التكرار النسبي والتكرارات الصاعدة والنازلة، عرض بياني للمدرج التكراري.

ـ عرف المنوال، أحسب قيمته، عرف الوسيط، أحسبه بيانيا وحسابيا، احسب المتوسط الحسابي.

الحل النموذجي

حل التمرين الأول

1ـ صحيح 2 ـ صحيح 3 ـ صحيح 4 ـ صحيح 5 ـ خطأ 6 ـ صحيح

حل التمرين الثاني (حالة عدم الترجيح)

نحن أمام عينة ضعيفة العدد، لذلك لا يتطلب ترتيب المعطيات في جدول إحصائي بل يكفي ترتيبها غلى النحو التالي: $x1 = 20 \quad x2 = 24 \quad x3 = 25 \quad x4 = 30$

$$x5 = 32 \quad x6 = 35 \quad x7 = 35 \quad x8 = 43$$

المنوال: مادمنا أمام عدد قليل من المعطيات يبقى المنوال ليس له تأثير كبير، إذ نجد القيمة 35 هي الوحيدة التي تكررت مرتين، وبالتالي يمكن اعتبارها كنموذج للمنوال، لكن لا تكسب أهمية كبيرة.

الوسيط: هو قيمة تكون أكبر من الجزء الأول لسلسلة القيم المرتبة ترتيبا تصاعديا، وأقل من الجزء الثاني لسلسلة القيم، لذلك علينا ترتيب المعطيات كما يلي: 20، 24، 25، 30، 32، 35، 35، 43.

نلاحظ أن عدد الوحدات زوجي ، وبالتالي لا توجد قيمة مباشرة تقسم العينة إلى قسمين متساويين في الوحدات، لذلك نحن في حاجة إلى قيمة تقسم القيمتين (30 و 32) أي القيمة 31 ولكنها غير متضمنة في السلسلة الاحصائية ، وعلى هذا الأساس نحن في حاجة إلى قيمة أخرى مثل القيمة 45 بحيث إذا اضفناها إلى السلسلة تكون على النحو التالي: 20 ، 24 ، 25 ، 30 ، 32 ، 35 ، 35 ، 43 ، 45 وتصبح العينة تساوي تسعة بدل ثمانية، وبالتالي يكون الوسيط هو القيمة 32 التي تقسم المعطيات إلى قسمين متساويين في التوزيع.

الوسط الحسابي

$$\bar{x} = \frac{\sum xi}{n} = \frac{244}{8} = 30{,}5$$

الوسط التربيعي

$$Q^2 = \frac{\sum xi^2}{n} = \frac{20^2+24^2+25^2+30^2+32^2+35^2+35^2+43^2+}{8} = \frac{7824}{8} = 978$$

$$MQ = \sqrt{Q^2} = \sqrt{978} = 31{,}273$$

الوسط الهندسي

$$MG = \sqrt[8]{20.24.25.30.32.35.35.43} = 29{,}7086$$

الوسط التوافقي

$$MH = \frac{n}{\sum \frac{1}{xi}} = \frac{8}{\frac{1}{20}+\frac{1}{24}+\frac{1}{25}+\frac{1}{30}+\frac{1}{32}+\frac{1}{35}+\frac{1}{35}+\frac{1}{43}}$$

$$= \frac{8}{0{,}2766486711} = 28{,}9175$$

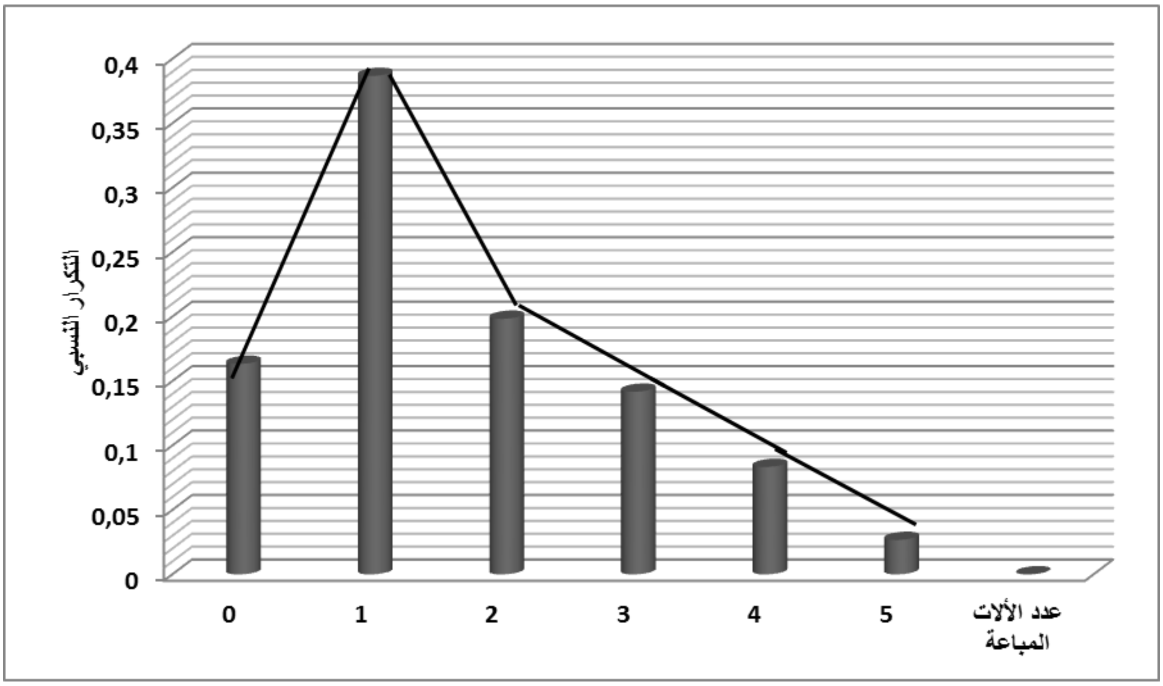
نستنتج من خلال حل التمرين أن: $H < G < \bar{X} < Q$ مما يعني أن توزيع القيم غير متناظر (متماثل).

حل التمرين الثالث (حالة المرجح المنقطع)

ـ الجدول التالي يقدم حساب التكرار النسبي والتكرار النسبي الصاعد

xi	F_i	fi	fi ↑
00	98	0,1633	0,1633
mo 01	232	0,3867	0,5500
02	119	0,1983	0,7483
03	85	0,1417	0,8900
04	50	0,0833	0,9733
05	16	0,0267	1,0000
$\sum F_i$	600	1,0000	------------------------

أما العرض البياني فيكون على النحو التالي



ـ المنوال: هو القيمة التي تقابل أكثر تكرار طبيعي أو نسبي، وبناء على هذا الأساس نجد بأن أكثر تكرار في الجدول هو 232 وبالتالي قيمة المنوال هي: MO = 1

ـ الوسيط: حينما تقوم بحساب الرتبة من خلال التكرار الطبيعي أو التكرار النسبي، فسوف تجد بأن التكرار النسبي الصاعد الموافق هو 0.55 و الذي يتوافق تماما مع القيمة 1 = *ME*

ـ الوسط الحسابي: $\overline{x} = \sum_{i=1}^{i=n} fi.xi = 1{,}675$ تلاحظ معي أننا في حسابنا للوسط الحسابي وظفنا التكرار النسبي بدلا من التكرارات الطبيعية.

حل التمرين الرابع (حالة المرجح المستمر)

صفة "العمر" مشتركة مع المتغير الاحصائي المستمر والذي نسميه X ، حيث جمعت قيمه على شكل فئات ، والعرض البياني للمتغير X يتعين علينا الحساب المسبق لكل من :

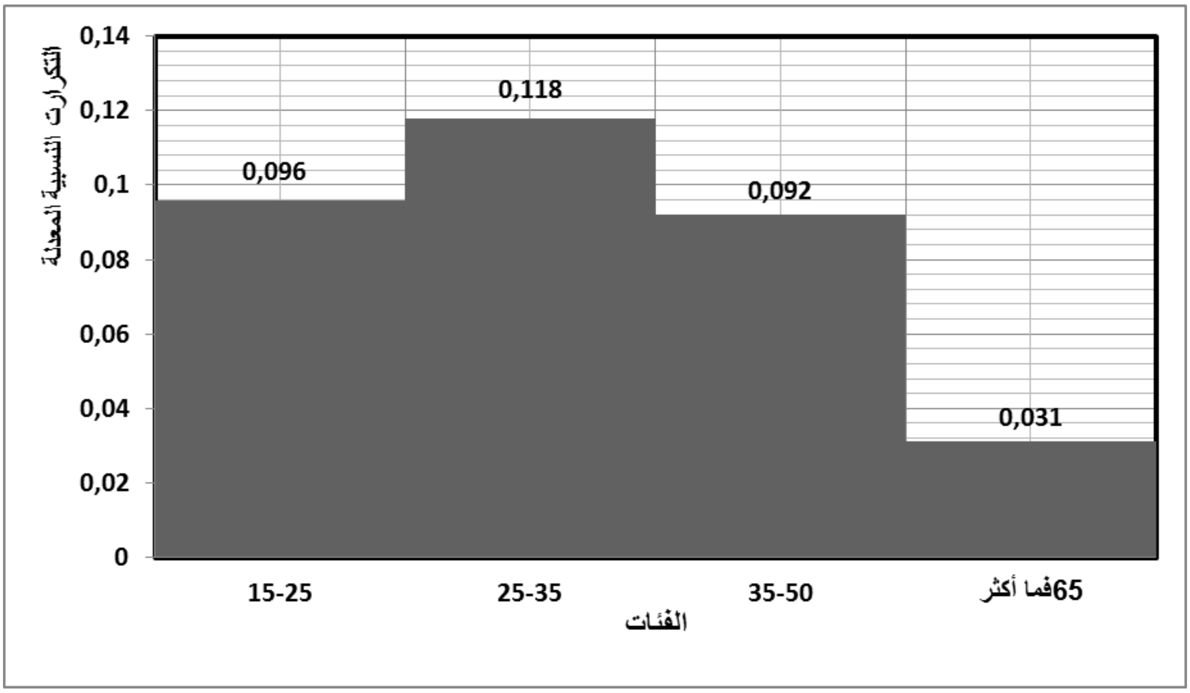
ـ التكرارات النسبية لتشكيل المدرج التكراري

ـ التكرارات الصاعدة والنازلة لرسم منحنيات التكرارات المتجمعة

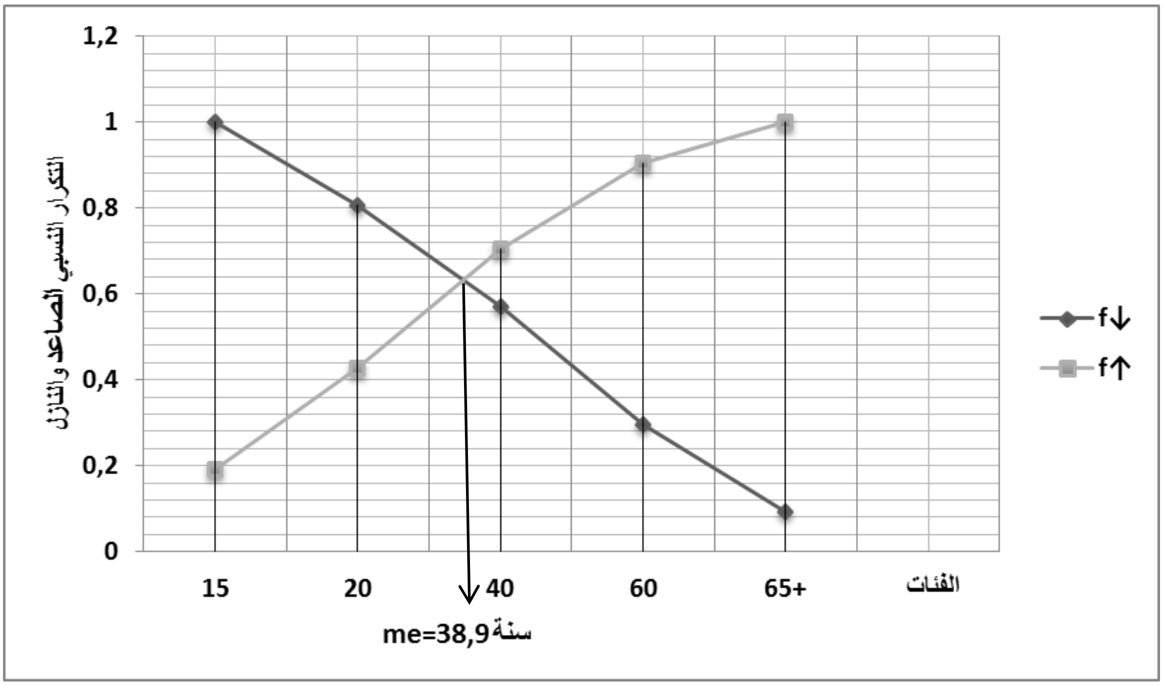
و الملاحظ في هذا التمرين أن أطوال الفئات غير متساوية، مما يحتم علينا تعديل التكرارات وفق القاعدة التي مرت بنا $ER = \frac{F_i}{k_i}$ أو بطريقة أخرى تكتشفها من خلال هذا الجدول.

العمر xi	F_i	fi	ki	طول الفئة المعدل $\frac{ki}{5}$ (ER)	ت نسبي معدل $\frac{f_i}{ER}$	ت نسبي صاعد $fi\uparrow$	ت نسبي نازل $fi\downarrow$
25-15	096	0,192	10	02	0,096	0,192	1,000
35-25	118	0,236	10	02	0,118	0,428	0,808
50-35	138	0,276	15	03	0,092	0,704	0,572
65-50	101	0,202	15	03	0,067	0,906	0,296
65 فما فوق	047	0,094	15	03	0,031	1,000	0,094
$\sum F_i$	500	1,000	--------	--------------------	---------	---------	--------

وهكذا من خلال التكرار النسبي المعدل، يمكن رسم المدرج التكراري كما يلي



ـ أما حساب الوسيط بيانيا، فسيتم من خلال معرفة نقطة التقاطع بين التكرارات النسبية المتجمعة الصاعدة والتكرارات النسبية المتجمعة النازلة، حيث نسقطها على المحور الأفقي لمعرفة هذه القيمة.



ـ حساب المنوال: لحساب المنوال المتعلق بمتغير كمي مستمر لفئات ذات أطوال غير متساوية، نطبق العلاقة

$$MO = A_{MO} + \frac{\acute{\Delta}1}{\acute{\Delta}1+\acute{\Delta}2} . \boldsymbol{k_{mo}}$$

وبناء على هذا الأساس نقوم بحساب هذه المجاهيل.

ـ الفئة المنوالية : هي الفئة التي تكون أمام أكبر تكرار (سواء كان طبيعي أو نسبي) ، وحسب جدول المعطيات، نجد التكرار النسبي المعدل الأكبر تقابله الفئة $[25 - 35[$ وعليه يكون

$$mo = 25 + \frac{0{,}022}{0{,}022+0{,}026} . 10 = 25 + 4{,}58 = 29{,}58$$

ـ حساب الوسيط: نتبع نفس الخطوات التي رأيناها في الدروس السابقة، فقط في هذا التمرين نتعامل مع التكرار النسبي (كأنه تكرار صاعد طبيعي).

ـ رتبة الوسيط:

$$\boldsymbol{t_{me}} = \frac{\sum fi}{2} = \frac{1}{2} = 0{,}5$$

ـ نحدد الفئة الوسيطية بالبحث عن هذه القيمة في التكرار النسبي الصاعد ، وبما أنها غير موجودة نبحث عن العدد الذي تتضمنه من الأعلى نحو الأسفل فنجد *0,704* وبالتالي الفئة التي تقابل هذه القيمة هي $[35 - 50[$

ـ نطبق العلاقة الرياضية :

$$\boldsymbol{m_e = a + \frac{\frac{\sum fi}{2} - f\uparrow(n-1)}{fi_{me}} . k_{me}} =$$

$$= 35 + \frac{0{,}5 - 0{,}428}{0{,}276} . 15 = 38{,}91$$

معنى هذا أن هناك 50 % من زائري المتحف أعمارهم $\geq$ 38.91 سنة، والنصف الآخر من عينة المجتمع محل الدراسة أعمارهم ≤ 38.91 سنة.

ـ حساب الوسط الحسابي : يتعين علينا حساب مراكز الفئات وبعض المجاميع الأخرى ، لذلك نستعين بالجدول التالي.

xi	x́i	F_i	$F_i \acute{x}_i$
25-15	20	096	1920
35-25	30	118	3540
50-35	42,5	138	5865
65-50	57,5	101	5807,5
+65	72,5	047	3407,5
$\sum$	----------------------	500	20540

تلاحظ في المجال المفتوح للفئة الأخيرة ، استعنا بطول الفئة التي سبقتها ، أما لو كانت الفئة الاولى سوف نستعين بالفئة التي تليها مباشرة.

$$\bar{x} = \frac{\sum F_i.\acute{x}_i}{\sum F_i} = \frac{20540}{500} = 41,08$$

من خلال نتائج مختلف المتوسطات، نجد بأن $\bar{x} > m_e > m_o$ و هذا يعني أن التوزيع غير متناظر (متماثل)، لكن المهم أن يكون الوسيط في مثل هذه الحالات دائما قيمته تتوسط قيم المتوسطات الأخرى(الوسط الحسابي والمنوال).

04 – تمارين غير محلولة

التمرين الأول

المعطيات التالية تتعلق بكلفة المبيت ليوم واحد في أحد الفنادق من ثلاثة (03) نجوم بالجزائر العاصمة ، وذلك من خلال فواتير الزبائن : 2500 دج ، 2300 ، 2700 ، 2200 ، 2900 ، 3000 دج

المطلوب : حساب كل من الوسط الحسابي ، الوسيط ، المنوال ، الربيع الأدنى ، الربيع الوسط ، الربيع الأعلى ، العشير الثالث ، المئين الخامس. ماذا يمكن أن تستنتج؟.

التمرين الثاني

نفرض أن معطيات التمرين السابق تتعلق بستة (06) زبائن حيث قضى كل واحد عدة ليالي بالترتيب : 03 ليالي ، 02 ، 06 ، 10 ، 05 ، 01 .

المطلوب : حساب كل من الوسط الحسابي ، الوسيط ، المنوال ، الربيعات ، العشير السادس ، المئين الخمسين ، استنتج العلاقة بين المتوسطات.

التمرين الثالث

نفرض أن عدد العاملين في الفندق الوارد في التمرين السابق هو خمسون (50) عاملا، حيث يتقاضون رواتبهم بالدينار الجزائري والمبينة في الجدول أسفله.

xi. 1000	20-15	25-20	30-25	35-30	40-35	45-40	50-45	+50
F_i	10	14	15	06	04	03	02	01

أحسب كل من : الوسط الحسابي ، الوسيط ، المنوال ، قارن بين النتائج ، ثم استنتج معادلة من الشكل
$mo = me = \bar{x}$ أو $(\bar{x} - mo) = 3(\bar{x} - me)$

التمرين الرابع

خلال الأربعة سنوات الماضية سجل نمو الناتج الداخلي الخام المعدلات السنوية التالية: 6.9 % ، 7.0 ، 4.3 ، 3.5 %. ما هو المعدل المتوسط للنمو الاقتصادي السنوي.

التمرين الخامس

ثلاثة مجموعات من العمال استفادت من الزيادة الأجرية خلال الفترة (2000 ــ 2009) وفقا للمعطيات التالية : السنوات الأربعة الأولى كانت الزيادة بنسبة 60 % أما السنتين المواليتين فقد كانت بنسبة 20 % ، بينما اقتصرت السنة السابعة على معدل زيادة قدره 10 % ، أما السنوات المتبقية فقد كانت70 %

السؤال : ماهو معدل النمو السنوي المتوسط للزيادة المحققة.

التمرين السادس

مؤشر الأسعار العام للفترة (2005 ـ 2011) كان على النحو التالي

11	10	9	8	7	6	2005
4.5	1.8	3.0	2.9	2.5	2.0	% 1.7

السؤال : حساب متوسط الأسعار العام للفترة

التمرين السابع

نريد معرفة متوسط السرعة لعينة تتكون من أربعة (04) حافلات تسير في خطوط بين المدن كما هو مبين في الجدول التالي

04	03	02	01	الحافلات
140	120	90	70 كلم/ ساعة	السرعة القصوى

التمرين الثامن

تبين معطيات البنك المركزي بأن معدلات صرف الدينار بالنسبة للدولار الامريكي خلال السنوات الخمسة الماضية كانت على النحو التالي

11	10	09	08	07	السنة
73	74.9	74.0	74.4	75.0	د ج / 1 $

السؤال: ما هو متوسط سعر صرف الدينار الجزائري للفترة محل الدراسة

التمرين التاسع

مخطط التنمية للفترة (2005 – 2009) يبين معدلات الاستثمار وأثرها على انخفاض معدل البطالة

معدل الاستثمار	12 % من *pib*	14	16	20	25
معدل البطالة	15	13	12.5	11	10

ما هو متوسط أثر معدل الاستثمار على انخفاض البطالة.

التمرين العاشر

عينة تتكون من مائة (100) طالب(ة) أجرينا عليها دراسة تتعلق بالتحصيل العلمي من خلال اختبار في مادة الاحصاء ، حيث كانت العلامة الكاملة 40 ، بعد تصحيح الأوراق تحصلنا على المعطيات التالية

العلامة	10-05	15-10	20-15	25-20	30-25	$\sum ni$
عدد الطلبة	05	15	20	25	35	100

السؤال : أحسب كل من الوسط الحسابي ، التوافقي ، الهندسي ، ثم حدد(ي) العلاقة التي تربط بينهم على شكل متراجحة ، هل التوزيع طبيعي. لماذا.

الفصل الثالث

مقاييس التشتت و الشكل

مقدمة

على الرغم من أهمية مقاييس النزعة المركزية في شتى البحوث والدراسات إلا أنها تعتبر غير كافية لوصف القيم المشاهدة للصفة محل الدراسة، فهي قاصرة عن توضيح طبيعتها وكيفية توزيع مفرداتها. كما أنها تقتصر على حساب القيمة المركزية فقط.

لذلك تعد مقاييس التشتت والشكل من المقاييس الاحصائية الهامة، لأنها تعتبر تكملة لما قبلها، نظرا لكونها تظهر توزيع وانتشار قيم المتغير الاحصائي حول القيمة المركزية.

تعريف التشتت

هو مدى الاقتراب أو الابتعاد للقيم المشاهدة حول وسطها، بمعنى إذا كانت القيم مركزة حول وسطها فإن قيمة التشتت تكون صغيرة، أما إذا كانت القيم مبعثرة وبعيدة حول وسطها تكون قيمة التشتت كبيرة

مثال: لدينا السلاسل الاحصائية التالية

السلسلة A	5 ، 4 ، 3 ، 2 ، 1	$\bar{x} = 3$
السلسلة B	1 ، 2 ، 1 ، 6 ، 5	$\bar{x} = 3$
السلسلة C	3 ، 3 ، 3 ، 3 ، 3	$\bar{x} = 3$

ملاحظات

ـ نلاحظ في السلسلة الاحصائية C أن الوسط الحسابي يتساوى مع كل قيم السلسلة، بينما في السلسلة A نجد تباعد القيم عن وسطها، أما في السلسلة B فإن هذا التباعد للقيم عن وسطها كان بشكل أكبر.وهكذا يجد الباحث نفسه مضطرا لاستعمال مقاييس أخرى تكمل مقاييس التوسط، لتحديد الانحرافات أو البعثرة بين قيم الوحدات الاحصائية محل الدراسة.

مثال آخر: السلسلتين التاليتين تعبران عن علامات عشرة (10) طلبة لمادتي الرياضيات والاحصاء

رياضيات	00	07	07	07	07	07	07	07	10	11
احصاء	03	04	05	06	07	07	08	09	10	11

بعد الحساب سوف تجد بأن الخصائص الثلاثة للنزعة المركزية متساوية: $\bar{x} = m_e = m_o = 07$ بينما لو أردنا حساب مجال الدراسة، والذي يسمى عادة بالمدى العام لوجدناه على النحو التالي.

$$E_M = (11 - 0) = 11$$

$$E_S = (11 - 3) = 08$$

معنى ذلك، أن تشتت سلسلة علامات الرياضيات أكبر من تشتت علامات الاحصاء.وهكذا نستنتج بأن مقاييس التشتت التي سوف نتعرف عليها بالتدرج (من المهم إلى الأهم) تلعب دور التعرف على مدى تجانس المجموعات.
نميز في مقاييس التشتت قسمين، يسمى الأول بمقاييس التشتت المطلقة، أما الثاني فيطلق عليه بمقاييس التشتت النسبية.

مقاييس التشتت	
مقاييس التشتت المطلقة	مقاييس التشتت النسبية
01 ـ المدى	01 ـ معامل الاختلاف
02 ـ متوسط الانحراف المطلق	02 ـ الدرجات المعيارية
03 ـ الانحراف الربيعي	03 ـ الخطأ المعياري
04 ـ التباين	
05 ـ الانحراف المعياري أو القياسي	

I ـ المدى

يسمى بالمدى العام، أو المدى المطلق و يرمز له عادة بالرمز E . كما يعرف على أنه الفرق بين أكبر قيمة وأصغر قيمة (حالة بيانات خام) أو الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى (حالة بيانات مبوبة).

يحسب بالعلاقة $E = x_{max} - x_{mini}$ ويفيد في معرفة التشتت للمجتمعات الصغيرة المتجانسة، مثل رقابة خطوط الانتاج، الارصاد الجوي ...الخ، كما يعتبر من أسهل مقاييس التشتت عند حسابه. غير أنه لا يخلو من عيوب، فهو أقل مقاييس التشتت دقة، وتعتبر قيمته مضللة، كونه يعتمد على قيمتين فقط عند حسابه، كما لا يمكن حسابه في الجداول المفتوحة نظرا لغياب الحد الأعلى للفئة الأخيرة، أو الحد الأدنى للفئة الأولى.وأخيرا لا يمكن استخدام هذا المقياس للمقارنة بين توزيعات تختلف في وحدات قياسها، مثل متغير محل الدراسة يقاس بالكيلوغرام والمتغير الثاني يقاس بالمتر.

01 ـ المدى العام لبيانات خام (غير مبوبة)

مثال: لدينا السلسلتين الاحصائيتين التاليتين

10	08	06	04	02	A
08	07	06	05	04	B

السؤال: أوجد كل من الوسط الحسابي والمدى العام للتوزيعين.

يلاحظ أن:

$$\overline{x}_A = \overline{x}_B = \frac{30}{5} = 06$$

بينما

$$E_A \neq E_B$$

$$E_A = (10 - 02) = 08$$

$$E_B = \ (08 - 04) = 4$$

تحليل: بالرغم من أن وسط كل مجموعة من المجموعتين هو 06 غير أن مدى قيم التوزيع الأول هو ضعف مدى قيم التوزيع الثاني، وهذا يعني أن قيم السلسلة A تتغير حول وسطها ضعف تغير قيم السلسلة B حول وسطها.

02 ـ المدى العام لبيانات مبوبة

2 ـ 1 ـ متغير كمي منقطع

مثال: لدينا عينة تتكون من ثمان وعشرون(28) أسرة ريفية موزعة حسب عدد الأولاد لكل أسرة

xi الأسرة	01	02	03	04	05	06	07	$\sum F_i$
F_i عدد الابناء	00	04	02	05	07	03	01	22

السؤال: أحسب كل من الوسط الحسابي، المدى العام

$$\overline{x} = \frac{\sum ni.xi}{\sum ni} = \frac{94}{22} = 4{,}27 \cong 04$$

تلاحظ أن الوسط الحسابي الأنسب هو المرجح، أما المدى العام فهو: $E_x = (07 - 01) = 06$ معنى هذا أن تشتت الأسر في عدد الأبناء غير كبير، بمعنى قريب إلى التجانس، لأنه غير بعيد عن وسطه.

2 ـ 2 ـ متغير كمي مستمر

مثال: لتكن لدينا المعطيات المتعلقة بتوزيعين إحصائيين على النحو التالي.

$x1$	F_1	$x2$	F_2
08-02	04	< 08	04
12-08	06	12-08	06
16-12	05	16-12	05
20-16	10	20-16	10
24-20	02	> 20	02

المطلوب: المقارنة بين التوزيعين

التوزيع الأول

$$E_{x1} = (24 - 02) = 22$$

أما التوزيع الثاني

$$E_{x2} = (24 - 04) = 20$$

وهذا يعني أن التوزيع الثاني أكثر تجانسا من الأول.

II ـ متوسط الانحراف المطلق

إذا كان لدينا توزيع إحصائي $(x_i \ , \ n_i)$ و $i \in \{1, \ldots.., r\}$ نسمي متوسط انحراف مطلق، العدد الذي نرمز له بالرمز $E_{\bar{x}}$ والمعرف بواسطة المعادلة $E_{\bar{x}} = \frac{1}{n}\sum[x_i - \bar{x}]$. يستعمل متوسط الانحراف المطلق كمؤشر للتشتت.

ـ متوسط الانحراف المطلق يعبر عن البعد المتوسط لقيم المتغير الاحصائي، محل الدراسة عن قيمة مركزية مثل الوسط الحسابي.

ـ أما الانحراف، فهو الفرق بين قيم المتغير الاحصائي والقيمة المركزية $(\bar{x})$ أو (m_e) أو (m_o).

ـ متوسط الانحراف المطلق، سواء عن الوسط الحسابي أو الوسيط أو المنوال يعتبر أقل تأثرا بالقيم الشاذة، كما يعتمد في حسابه على كل القيم المشاهدة، إلا أن ما يعاب عليه هو إهمال إشارة الانحرافات واعتبارها مطلقة مما يصعب من معالجته رياضيا والتعرف على خصائصه، كما لا يمكن استخدامه من جداول تكرارية مفتوحة.

مثال: أخذنا عينة تتكون من تسعة (09) طلبة للسنة الأولى اقتصاد، ومن خلال أوراق اختبار الاحصاء تحصلنا على المشاهدات التالية (العلامة /20).

09	08	07	06	05	04	03	02	01	رقم الطالب
07	09	07	12	15	09	10	04	07	العلامة

المطلوب: حساب كل من الوسط الحسابي، الوسيط، المنوال، ثم حساب متوسط الانحراف عن كل هذه المتوسطات.

الحل: لحساب المتوسطات الثلاثة، نستعين بالجدول التالي

الرقم	x_i(قيم غير مرتبة)	ترتيب القيم
01	07	04
02	04	07
03	10	07
04	09	07
05	15	09
06	12	09
07	07	10
08	09	12
09	07	15
$\sum x_i$	80	----------------------------

أولا: حساب كل من الوسط الحسابي، الوسيط، المنوال

الوسط الحسابي

$$\bar{x} = \frac{\sum xi}{n} = \frac{80}{9} = 8,9$$

الوسيط

ـ رتبة الوسيط

$$t_e = \frac{n+1}{2} = \frac{10}{2} = 05$$

وبالتالي قيمته هي ME = 09 من العمود الذي رتبت فيه القيم.

المنوال

هو القيمة الأكثر شيوعا، وبالتالي: MO = 07

ثانيا: حساب متوسط الانحراف المطلق عن الوسط الحسابي، الوسيط ، المنوال

xi	$\lvert xi-\bar{x}\rvert$	$\lvert xi-me\rvert$	$\lvert x_i-m_o\rvert$
07 04 10 09 15 12 07 09 07	1,9 4,9 1,1 0,1 6,1 3,1 1,9 0,1 1,9	02 05 01 00 06 03 02 00 02	00 03 03 02 08 05 00 02 00
$\bar{x}=8{,}9$	$\sum=21{,}1$	$\sum=21$	$\sum=23$
$\bar{x}=\frac{\sum x_i}{n}$	$E_{\bar{x}}=\frac{\sum\lvert xi-\bar{x}\rvert}{n}=\frac{21{,}1}{9}=2{,}34$	$E_{\bar{x}}(me)=\frac{\sum\lvert xi-me\rvert}{n}=\frac{21}{9}=2{,}33$	$E_{\bar{x}}(mo)=\frac{\sum\lvert xi-mo\rvert}{n}=\frac{23}{9}=2{,}56$

مثال آخر: اردنا دراسة متوسط سعر بطاطس ولاية الوادي في خمسة أسواق مختارة بالجزائر العاصمة، فكانت المعطيات على النحو التالي.

رقم السوق	01	02	03	04	05	المجموع
سعر كلغ (x_i)	55	60	65	70	75 دج	-----------
الكمية المشترات (F_i)	4.0	3.5	03	2.5	02 كلغ	15

السؤال: أحسب متوسط الانحراف المطلق عن الوسط الحسابي

الحل: نبحث بداية عن الوسط الحسابي المرجح ثم نقوم بحساب الانحراف المطلق، ولأجل ذلك نستعين بالجدول التالي.

xi	F_i	$x_i.F_i$	$\lvert xi-\bar{x}\rvert$	$F_i.\lvert x_i-\bar{x}\rvert$
55	4.0	220	8.33	33.32
60	3.5	210	3.33	11.65
65	3.0	195	1.67	05.01
70	2.5	175	6.67	16.67
75	2.0	150	11.67	23.34
$\sum$	15	950	----------	89.99

الوسط الحسابي المرجح: $\bar{x} = \frac{\sum x_i.F_i}{\sum F_i} = \frac{950}{15} = 63.33$

متوسط الانحراف المطلق عن الوسط الحسابي: $E_{\bar{x}} = \frac{\sum F_i.|x_i - \bar{x}|}{\sum F_i} = \frac{89.99}{15} = 5.99 \cong 6$

مثال آخر: اجرينا دراسة حول ظاهرة تخلف الطلبة عن محاضرة الاحصاء في وقتها، فكان اختيارنا على عينة عشوائية تتكون من عشرون (20) طالب (ة) سجلوا عند الدخول خلال الخمسة عشرة دقيقة الأولى من الحصة، فكانت المعطيات وفق الجدول التالي

مراكز فئات الوقت الضائع $\acute{x}_i$	02	04	06	08	10	$\sum F_i$
عدد الطلبة المتخلفون F_i	01	03	08	06	02	20

المطلوب: حساب متوسط الانحراف عن الوسيط.

الحل: لا يمكن أن نجد متوسط الانحراف دون حساب الوسيط، كما لا يمكن حساب الوسيط دون ايجاد الفئة الوسيطية، وبالتالي لا بد من تحويل مراكز الفئات إلى فئات. لأجل ذلك علينا استنتاج طول الفئة.

ـ طول الفئة: مركز الفئة الثاني ـ مركز الفئة الأول = (4) ـ (2) = 02

ـ نصف طول الفئة = طول الفئة / 2 = 01 ومنه البحث عن الحد الأدنى للفئة.

ـ الحد الأدنى للفئة الأولى = مركز الفئة الأولى ـ ½ طول الفئة : (2) ـ (1) = 01

ـ الحد الأعلى للفئة الأولى = مركز الفئة الأولى + ½ طول الفئة : (2)+(1)= 03

وعلى هذا الأساس ننجز جدول الفئات كما يلي

$\sum F_i$	11-9	9-7	7-5	5-3	3-1	xi
	10	08	06	04	02	$\acute{x}_i$
20	02	06	08	03	01	F_i

وهكذا يمكن حساب الوسيط ومتوسط الانحراف له، بعد الاستعانة بالجدول التالي

xi	$\acute{x}_i$	F_i	$F\uparrow$	$\lvert \acute{x}_i - me \rvert$	$F_i.\lvert \acute{x}_i - m_e \rvert$
03-01	02	01	01	4,5	4,5
05-03	04	03	04	2,5	7,5
07-05	06	08	12	0,5	4,0
09-07	08	06	18	1,5	9,0
11-09	10	02	20	3,5	7,0
$\sum$	-------------	20	-------	-----------------	32

ـ ترتيب الوسيط: $t_{me} = \frac{\sum F_i}{2} = 10$

ـ قانون الوسيط: $M_E = A_{me} + \left[\frac{\frac{\sum F_i}{2} - F\uparrow(F-1)}{F_{me}}\right].k_{me}$

ـ تحديد الفئة الوسيطية: هي $[5-7[$

ـ حساب الوسيط: $ME = 5 + \frac{10-4}{8} \times 2 = 6,5$

متوسط الانحراف

$$E_{\bar{x}}(me) = \frac{F_i.\lvert \acute{x}_i - m_e \rvert}{\sum F_i} = \frac{32}{20} = 1,6$$

III ـ الانحراف الربيعي

01 ـ المدى الربيعي : يعرف على أنه الفرق بين الربيع الثالث والربيع الأول، ويرمز له عادة بالرمز IQ سهل الحساب، يستعمل في المقارنة بين توزيعين أو أكثر، هو أفضل من المدى العام، كونه لا يتأثر بالقيم المتطرفة، يضم 50 % من المجتمع محل الدراسة، غير أن استعمالاته محدودة.

يعطى بالعلاقة التالية: $IQ = Q_3 - Q_1$

02 ـ نصف المدى الربيعي

يمثل متوسط الفرق بين الربيع الثالث والربيع الأول، ونظرا لكون المدى العام يتأثر بالقيم المتطرفة لجأ الاحصائيون إلى استخدام ما يشبه المدى في قياس التشتت، وبالتالي فالهدف هو التغلب على أهم عيوب المدى العام (المطلق).يعطى بالقاعدة: $\frac{Q_3 - Q_1}{2} = \frac{IQ}{2}$ ويفسر على أن 50 % من القيم الاحصائية تبعد في المتوسط عن الوسيط بأقل من $\frac{IQ}{2}$.

مثال: أحسب الانحراف الربيعي للبيانات التالية:

$$X;\ 12\ ,\ 08\ ,\ 03\ ,\ 05\ ,\ 04\ ,\ 06\ ,\ 07\ ,\ 09\ ,\ 10$$

الحل:

ـ نقوم بترتيب البيانات ترتيبا تصاعديا ثم نبحث عن ترتيب الربيع الأول (عدد الوحدات فردي) والثالث، بعدها نقوم بالحساب فنجد ما يلي:

$t_{q1} = 2.5\ ,\ t_{q3} = 7.5$ ومنه تحديد قيم الربيعين: $Q_1 = 4.5\ ,\ Q_3 = 9.5$ بعد حساب هذه النتائج يمكن الآن حساب الانحراف الربيعي.

ـ المدى الربيعي

$$IQ = Q_3 - Q_1 = (9.5) - (4.5) = 05$$

ـ نصف المدى الربيعي

$$\frac{IQ}{2} = \frac{Q_3 - Q_1}{2} = \frac{5}{2} = 2.5$$

وتفسر النتيجة على أن 50 % من القيم المشاهدة تبتعد في المتوسط عن الوسيط بأقل من 2.5 .

مثال آخر: الاختبار الشفوي في مادة الاحصاء لعينة من الطلبة تتكون من 100 طالب(ة) مكننا من الحصول على المعطيات التالية (العلامة من 50)

xi نقاط الطلبة	F_i الطلبة	$F\uparrow$
10-05	10	10
15-10	12	22
20-15	20	42
25-20	29	71
30-25	11	82
35-30	08	90
40-35	06	96
45-40	04	100
$\sum F_i$	100	--------------------------------

المطلوب: احسب الانحراف الربيعي

ـ حساب Q_3

رتبته: $75 = \frac{300}{4} = \frac{(\sum F_i)}{4} = t_{q3}$ نبحث عن فئة الربيع الثالث من خلال $F\uparrow$ وهي تتوافق تماما مع $F\uparrow 82$ أي الفئة $[25-30[$، وعلى هذا الأساس نقوم بحساب الربيع الثالث مثلما رأينا في دروسنا السابقة.

$$Q_3 = 25 + \left[\frac{75-71}{11}\right].5 = 26.81$$

ـ حساب Q_1

بنفس الطريقة نجد الربيع الأول (الأدنى)، حيث $t_{q1} = \frac{\sum F_i}{4} = 25$ ومنه $Q_1 = 15,75$

ـ المدى الربيعي

$$IQ = 11,06$$

ـ نصف المدى الربيعي

$$\frac{IQ}{2} = 5,53$$

ـ تفسير النتيجة: معنى ذلك أن 50 % من نقاط الطلبة الشفوية تبتعد في المتوسط عن الوسيط بأقل من 5.53

03 ـ النسبة بين المدى الربيعي والمدى العام

يستعمل هذا المقياس، لقياس تشتت 50 % من القيم الاحصائية التي تقع حول القيمة المركزية للتوزيع محل الدراسة، ويعطى بالعلاقة:

$$R = \left[\frac{Q_3 - Q_1}{E}\right] \times 100$$

ملاحظة

ـ إذا كان: $R = 50\,\%$ معنى ذلك أن التوزيع الاحصائي للمعطيات حول مركزها متناظر

ـ إذا كان: $R < 50\,\%$ معنى ذلك أن التوزيع الاحصائي للمعطيات حول مركزها قليل التشتت

ـ إذا كان: $R > 50\,\%$ معنى ذلك أن التوزيع الاحصائي للمعطيات حول مركزها قوي التشتت

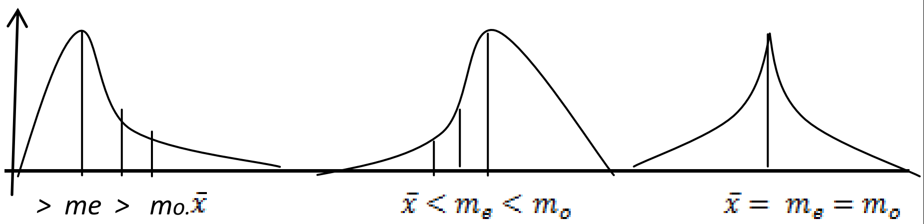
مثال: من نتائج المثال السابق، أحسب (R).

الحل:

$$R = \left[\frac{11,06}{40}\right] \times 100 = \frac{1106}{40} = 27,65\,\%$$

وهذا يدل على أن المدى الربيعي لا يمثل سوى 27.65 % من المدى العام ، أي أن 50 % من قيم الوحدات الاحصائية محل الدراسة التي تقع حول المركز قليلة التشتت بالنسبة لمركزها.

04 ـ مزايا وعيوب الانحراف الربيعي

	مميزات الانحراف الربيعي
01	يمكن التعرف على تماثل التوزيع من عدمه، من خلال الربيعات: *Q1*، *Q2*، *Q3* كما يلي: ـ في حالة التوزيع المتماثل: $Q_3 - Q_2 = Q_2 - Q_1$ لا يوجد التواء ـ في حالة الالتواء الموجب (ناحية اليمين): $Q_3 - Q_2 > Q_2 - Q_1$ ـ في حالة الالتواء السالب (ناحية اليسار): $Q_3 - Q_2 < Q_2 - Q_1$  1ـ لا يوجد التواء 2 ـ التواء على اليسار 3 ـ التواء على اليمين
02	الانحراف الربيعي لا يتأثر بالقيم المتطرفة
03	يمكن حساب الانحراف الربيعي من الجداول التكرارية المفتوحة والمغلقة
	عيوب الانحراف الربيعي
01	يعتمد الانحراف الربيعي على قيمتين ويهمل باقي القيم
02	تتوقف جودة الانحراف الربيعي على درجة تركز البيانات عند الربيعين الأدنى والأعلى، أما إذا كانت ثغرات بينهما، يصبح المقياس غير ملائم لقياس التشتت المطلق.

IV ـ التباين والانحراف المعياري أو القياسي

01 ـ التباين

إذا كان (X) متغير إحصائي للتوزيع $(x_i \ , \ n_i)$ وعندنا $i \in \{1, \dots, r\}$ نسمي تباين الرمز ² للمجتمع والرمز s_x^2 للعينة . ويعرف بأنه متوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي.

1 ـ 1 ـ حالة بيانات خام (غير مبوبة) ونميز فيها طريقتين

ـ طريقة الانحرافات عن المتوسط الحسابي:

$$s_x^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

تلاحظ أننا أقسمنا على$(n-1)$ لأننا تعاملنا مع عينة، أما لو كان المجتمع ككل سوف نقسم على (N) والتي تسمى بدرجات الحرية، وأسباب ذلك هي:

الأول

المعروف في التحليل الاحصائي أن كل قيمة إحصائية مستقلة عن الأخرى، لكن حينما نحسب المتوسط الحسابي من أجل استخراج قيمة الانحرافات، يلاحظ أن أحد هذه الانحرافات يمكن الحصول عليه بالطرق الحسابية وفق قاعدة مجموع الانحرافات عن الوسط الحسابي يساوي صفرا، وعلى هذا الأساس يتم طرح واحد (1) من عدد الانحرافات الكلية لكي نحصل على انحرافات مستقلة، لذلك يكون عدد الانحرافات المستقلة هو

(n-1) وهو ما يسمى بدرجات الحرية.

الثاني

بما أننا نعتمد في دراستنا على العينة لتعميم نتائجها على المجتمع محل الدراسة، علينا أن ندرك بأنه مهما كان سحب العينة دقيقا، فإن احتمال وجود قيم متطرفة مثلما هي في المجتمع وارد، ويترتب على ذلك احتمال أن تكون قيمة انحرافات افراد العينة عن الوسط أو مربعاتها أقل من مثيلتها في المجتمع، ولقد وجد student (1908) أنه يمكن تعويض النقص في قيمة التباين للعينة عندما نقسم على (n-1) بدلا من n.

ـ طريقة البيانات الأصلية

$$s_x^2 = \frac{\sum x_i^2 - \frac{(\sum xi)^2}{n}}{n-1}$$

1 ـ 2 ـ حالة بيانات مبوبة، نميز فيها طريقتين

ـ طريقة الانحرافات عن المتوسط الحسابي

$$s_x^2 = \frac{\sum F_i.(x_i - \bar{x})^2}{\sum(n-1)}$$

نحن هنا أمام متغير كمي منقطع (مرجح) أما إذا كنا أمام متغير كمي مستمر سوف نقتصر على مركز الفئة، بمعنى بدل x_i نضع مكانه $(\acute{x}_i)$.

ـ طريقة البيانات الأصلية

$$s_x^2 = \frac{\sum F_i.x_i - \frac{(\sum F_i.x_i)^2}{n}}{n-1}$$

مثال: الجدول التالي يوضح الأجور الشهرية لعينة من العمال لمؤسسة ـ ما ـ حسب الجدول التالي

الأجور $10^3 \times$ (د ج)	30	55	60	65	75	70	80	80	85	90
العمال	01	02	03	04	05	06	07	08	09	10

المطلوب: حساب التباين

الحل: نلاحظ بأن البيانات غير مبوبة، بمعنى غير مرتبة على أساس التكرارات، وبالتالي نختار واحدة من الطريقتين، ولتكن طريقة الانحرافات عن المتوسط الحسابي، وعلى هذا الأساس نستعين بالجدول التالي.

الأجور (x_i)	$\bar{x} = 69 , (x_i - \bar{x})$	$(x_i - \bar{x})^2$
90	21	441
85	16	256
80	11	121
80	11	121
70	01	01
75	06	36
65	04-	16
60	09-	81
55	14-	196
30	39-	1521
$\sum xi = 690$	$\sum_{i=1}^{i=10} (x_i - \bar{x}) = 0$	$\sum (x_i - \bar{x})^2 = 2790$

$$s_x^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{2790}{10-1} = 310$$

1 – 3 ـ خصائص التباين

يتميز التباين بجملة من الخصائص ، نستعرضها في الجدول التالي

01	إذا كان كل من X و Y متغيرين مستقلين، وكان المتغير Z هو مجموعهما، أي $Z = X_I + Y_I$ فإن: $S_Z^2 = S_X^2 + S_Y^2$
02	إذا كانت لدينا مجموعتان من القيم المشاهدة، وقد تحصلنا على تباين كل من المجموعة الأولى والثانية، فإن التباين المشترك لجميع القيم هو: $$S_t^2 = \frac{\sum(xi_1 - \bar{x}_1)^2 + \sum(xi_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}$$
03	يمكن استعماله في حساب احصاءات أخرى (الانحراف المعياري، الخطأ القياسي، معامل الاختلاف ...الخ)
04	كل أفراد العينة يدخلون في حساب التباين، وهذا ما يجعله يتأثر بالقيم المتطرفة أكثر من غيره من المقاييس، لأن تربيع الانحرافات يعطي وزنا أكبر للقيم المتطرفة.

02 ـ الانحراف المعياري أو القياسي

قد نجد خلط في بعض المراجع الاحصائية بين مصطلح الانحراف المعياري والذي يخصص كقاعدة عند معالجة كل بيانات المجتمع محل الدراسة، وبين مصطلح الانحراف القياسي، والذي يخصص كقاعدة لمعالجة بيانات العينة التي أخذناها من المجتمع. بناء على هذا التوضيح، سوف نكتفي بدراسة الانحراف القياسي، على اعتبار أن العينة هي الوسيلة الأكثر استعمالا في بحوث العلوم الاجتماعية.

يعرف الانحراف القياسي ببساطة، على أنه الجذر التربيعي للتباين، ويعطى بالعلاقة التالية

$$s_x = \sqrt{s_x^2}$$

أو نعوض بما يساوي ما تحت الجذر كما يلي.

2- 1- حالة بيانات غير مبوبة

$$S_x = \sqrt{\frac{\sum xi - (\sum xi)^2}{n-1}} \quad \text{أو} \quad S_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

حيث (n) هو حجم العينة

2- 2 - حالة بيانات مبوبة

$$S_x = \sqrt{\frac{\sum F_i . x_i - \frac{(F_i x_i)^2}{\sum F_i}}{\sum F_i - 1}} \quad \text{أو} \quad S_x = \sqrt{\frac{\sum F_i . (x_i - \bar{x})^2}{\sum F_i - 1}}$$

حيث $(\sum F_i)$ هو مجموع التكرارات الطبيعية

مثال: حالة البيانات غير المبوبة

أحسب قيمة الانحراف القياسي بالطريقتين وقارن بين النتائج للمعطيات التالية: 04، 03، 05، 07، 08، 02، 12، 09، 13.

الحل: بداية نقوم بحساب الوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{i=9} xi}{n} = \frac{63}{9} = 07$$

بعد هذا نقوم بحساب التباين بالطريقتين استنادا إلى الجدول المساعد التالي.

$\sum (x_i - \bar{x})^2$	$\sum (x_i - \bar{x})$	$\sum xi^2$	xi	الرقم
09	03-	16	04	01
16	04-	09	03	02
04	02-	25	05	03
00	00	49	07	04
01	1+	64	08	05
25	5-	04	02	06
25	5+	144	12	07
04	2+	81	09	08
36	6+	169	13	09
120	00	561	63	$\sum x_i$

ـ **الطريقة المباشرة** (طريقة الانحرافات عن الوسط الحسابي):

$$s_x^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{120}{9-1} = 15$$

ومنه نقوم بحساب الانحراف القياسي:

$$s_x = \sqrt{s_x^2} = \sqrt{15} = 3.87$$

ـ **الطريقة الموسعة** (طريقة البيانات الأصلية):

$$s_x^2 = \frac{\sum xi^2 - (\sum xi)^2/n}{n-1} = \frac{561 - (63)^2/9}{9-1} = \frac{561-441}{8} = 15$$

ومنه نجد الانحراف القياسي: $s_x = 3.87$

مثال آخر: عينة تتكون من ستة (06) طلبة، سجلوا خلال السداسي الأول من السنة، الغيابات التالية:

06	05	04	03	02	01	رقم الطالب
08	03	07	05	02	00	عدد الغياب

المطلوب: حساب كل من التباين والانحراف القياسي مع تفسير النتيجة.

الحل: للحصول على نتائج المقاييس المرغوبة نستعين بالجدول التالي

$(x_i - \bar{x})^2$	$x_i - \bar{x}$	xi^2	xi	رقم الطالب
17.30	4.16-	00	00	01
4.66	2.16-	04	02	02
0.70	0.84	25	05	03
8.06	2.84	49	07	04
1.12	1.06-	09	03	05
14.74	3.84	64	08	06
46.58	$\cong 00$	151	25	$\sum$

ـ الوسط الحسابي

$$\overline{x} = \frac{\sum xi}{n} = \frac{25}{6} = 4,16$$

ـ التباين

$$s_x^2 = \frac{\sum(x_i-\overline{x})^2}{n-1} = \frac{46.58}{6-1} = 9.31$$

ـ تفسير النتيجة: أي أن تشتت القيم عن وسطها ($\overline{x}$) بلغ (9.31)، وهي كما ترى أكثر من الضعف، وهذا راجع بالأساس إلى مربع تمييز الصفة (الغياب)، وهو ما يعاب على مثل هذا المقياس. لذلك لجأ الاحصائيون إلى استعمال الانحراف القياسي.

ـ الانحراف القياسي

$$s_x = \sqrt{s_x^2} = \sqrt{9.31} = 03.05$$

الانحراف القياسي لهذه العينة بلغ (03.05) كقيمة إحصائية بعيدة عن وسطها، وبمعنى آخر، إذا كانت أمثلية الغيابات هي 04 غيابات لكل طالب (ة) في المتوسط خلال السداسي، فإن هناك زيادة أو نقصان تقدر بثلاثة غيابات.

مثال: حالة بيانات مبوبة (كمي منقطع أو مستمر)

الجدول التالي يبين توزيع الأجور لعينة من العمال حسب رواتبهم الشهرية الاجمالية

$xi.10^3$	F_i	$\acute{x}i$	$(\acute{x}_i - \overline{x})$	$(\acute{x}_i - \overline{x})^2$	$F_i.(\acute{x}_i - \overline{x})^2$
50-40	10	45	18.49-	341.88	3418.8
60-50	15	55	08.49-	72.08	1081.2
70-60	30	65	01.51+	02.28	68.40
80-70	12	75	11.51+	132.48	1589.76
90-80	06	85	21.51	462.68	2776.08
$\sum$	73	-------	------------	-------------	8934.24

ـ الوسط الحسابي

$$\overline{x} = \frac{\sum F_i.\acute{x}_i}{\sum F_i} = \frac{4635}{73} = 63.49$$

ـ التباين

$$s_x^2 = \frac{\sum ni.(x_i - \bar{x})^2}{(\sum ni) - 1} = \frac{8934.49}{73 - 1} = \frac{8934.49}{72} = 124.09$$

ـ الانحراف القياسي

$$s_x = \sqrt{s_x^2} = \sqrt{124.09} = 11.13$$

سؤال: فسر النتائج المتحصل عليها.

2 ـ 3 ـ تصحيح *sheppard´*

إن استعمال مركز الفئة كقيمة مساوية لكل القيم التي تتكرر في هذه الفئة، يؤدي إلى فروق بسيطة في قيمة الوسط الحسابي. أما في حالة حساب التباين، فقد وجد أن استعمال مركز الفئة يؤدي إلى زيادة قيمة التباين، بسبب عدم دقة تمثيل مركز الفئة للقيم المتكررة فيه. ومثل هذا الفرق في قيمة التباين لا يكون كبيرا أو لا يحتاج إلى تصحيح بيانات الجدول إذا كانت تلك البيانات موزعة توزيعا متماثلا، وكان مدى الفئة لا يزيد عن 20/1 من المدى الكلي. أما إذا كان العكس، يمكن تصحيح قيمة التباين باستعمال طريقة "شبرد" على النحو التالي: $s_{corrigé}^2 = s_x^2 - \frac{(k)^2}{12}$ وبتطبيق هذا التصحيح على ظاهرة أجور العاملين في المثال السابق (لأن مدى الفئة يمثل 20/4) وبالتالي فان التباين المصحح هو

$$s_{cor}^2 = 124.09 - \frac{(10)^2}{12} = 115,76$$

ومن تم نحصل على الانحراف القياسي

$$s_x = \sqrt{115.76} = 10.75$$

2 ـ 4 ـ خصائص الانحراف القياسي

01	أهم مقاييس التشتت وأكثرها استخداما، يستعمل في حساب معامل الارتباط والتوزيعات الاحصائية والاحتمالية، غير أنه لا يمكن استخدامه في مقارنة تشتت عينتين تختلفان في وحدات قياسهما مثل الكلغ والمتر.
02	لا يمكن حساب الانحراف القياسي من التوزيعات التكرارية المفتوحة، وفي هذه الحالة يفضل استخدام الانحراف الربيعي لدراسة التشتت.
03	يمكن قياس الحدود التي تنحرف بها قيم المتغير عن وسطها، ويستفاد من هذه الخاصية في تحديد نسب عدد الوحدات الاحصائية لتوزيع إحصائي متماثل، أو قريب من التماثل حسب الحالات التالية: ـ إذا كان $\{\bar{x} \mp 0,67S_x\}$ فإن المجال يحتوي على 50 % من العينة محل الدراسة. ـ إذا كان $\{\bar{x} \mp S_x\}$ فإن المجال يحتوي على 68 % من العينة محل الدراسة.

ـ إذا كان $\{\bar{x} \mp 2 S_x\}$ فإن المجال يحتوي على 95 % من العينة محل الدراسة.

ـ إذا كان $\{\bar{x} \mp 3 S_x\}$ فإن المجال يحتوي على 99 % من العينة محل الدراسة.

مثال: احسب الانحراف القياسي للبيانات التالية: 02، 19، 19، 19، 18، 18، 17، 16 ، 16، 16، 15، 15، 14، 14، 14، 13، 13، 12، 11، 21، ثم احسب عدد القيم التي تنحصر بين: $\bar{x} \mp S_x$ ثم $\bar{x} \mp 2S_x$ ثم $\bar{x} \mp 3 S_x$.

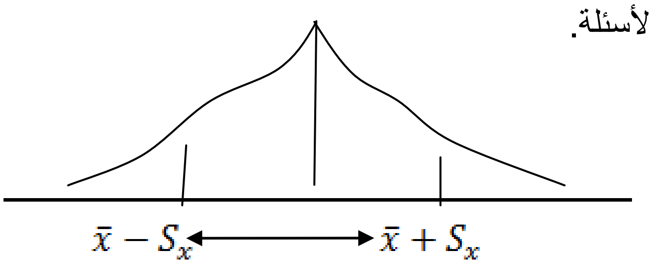
الحل:

$\sum x_i = 320$ و $\bar{x} = 16$ و $S_x = 2,8$

نقوم بحساب: $\bar{x} + S_x = 16 + 2,8 = 18,8$

نقوم بحساب: $\bar{x} - S_x = 16 - 2,8 = 13,2$

أنظر الآن في قيم التوزيع سوف تجد 14 قيمة محصورة بين هاتين النتيجتين من أصل 20 قيمة، أي أن 70 % من هذه القيم موجودة ضمن المجال $[\bar{x} \pm S_x]$. بنفس الطريقة تجد الجواب على باقي الأسئلة.



V ـ التشتت النسبي (العلاقة بين التشتت المطلق ومقاييس النزعة المركزية)

01 ـ معامل الاختلاف

هو طريقة تستخدم للمقارنة بين تشتت توزيعات غير متجانسة. يمكن حسابه من الجداول التكرارية المفتوحة والمغلقة عى حد سواء.

1 ـ 1 ـ معامل الاختلاف للجداول التكرارية المفتوحة

لحساب معامل الاختلاف من جداول مفتوحة، سواء كانت الفئة الأولى أو الأخيرة، نختار طريقة من الطريقتين، وذلك باتباع الخطوات التالية .

الطريقة الأولى: معامل الاختلاف الربيعي

ـ استخراج الربيع Q_1

ـ استخراج الربيع $Q3$

ـ استخدام معادلة معامل الاختلاف الربيعي كما يلي.

$$C.Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

الطريقة الثانية: العلاقة بين نصف المدى الربيعي و الوسيط

ـ حساب نصف المدى الربيعي

$$\frac{Q_3 - Q_1}{2}$$

حساب الوسيط ME

ـ استخدام المعادلة

$$C.Q.D = \frac{Q_3 - Q_1}{2} \cdot \frac{1}{ME} \times 100$$

$$= \frac{Q_3 - Q_1}{2ME} \times 100$$

1 ـ 2 ـ معامل الاختلاف للجداول التكرارية المغلقة

يمكن حساب معامل الاختلاف من الجداول التكرارية المغلقة بواحدة من اربعة طرق

الطريقة الأولى: العلاقة بين الانحراف القياسي والوسط الحسابي

ـ نقوم بحساب $\bar{x}$ (بأية طريقة من الطرق المعروفة)

ـ نقوم بحساب الانحراف المعياري (مجتمع) أو الانحراف القياسي (عينة) بأية طريقة من الطرق المعروفة.

ـ نستخدم معادلة معامل الاختلاف على النحو التالي

$$c.v = \frac{s_x}{\bar{x}} \times 100 / \bar{x} \neq 0$$

الطريقة الثانية: العلاقة بين متوسط الانحراف المطلق والوسط الحسابي

ـ نقوم بحساب $\bar{x}$ (بأية طريقة)

ـ نقوم بحساب متوسط الانحراف المطلق

ـ نستخدم المعادلة $$c.m.d = \frac{E_x}{\bar{x}} \times 100$$

الطريقة الثالثة: العلاقة بين متوسط الانحراف المطلق والوسيط

ـ نقوم بحساب متوسط الانحراف المطلق ثم الوسيط

ـ نستخدم المعادلة : $$c.m.d = \frac{E_x}{m_e} \times 100$$

الطريقة الرابعة: معامل الاختلاف بناء على المدى العام

$$c.m.d = \frac{x_{max} - x_{min}}{x_{max} + x_{min}} \times 100$$

أمثلة محلولة

المثال الأول

نريد دراسة معامل الاختلاف المتعلق بنقاط السداسي الأول لعينة من طلبة السنة الأولى اقتصاد، وفق معطيات الجدول التالي.

x_i	F_i
25 20	02
30 25	04
35 30	00
40 35	01
45 40	08
50 45	10
55 50	15
60 55	15
65 60	10
70 65	03
75 70	01
80 75	01
$\sum Fi$	**70**

حل المثال: الملاحظ للوهلة الأولى، أن المجال مغلق من الطرفين، لذلك يمكن أن نختار واحدة من الطريقتين 1 أو 2 ولتكن الطريقة الأولى.

ـ عند حساب كل من S_x و $\bar{x}$ سوف نجد النتيجة كما يلي

$$s_x = 11.19 \text{ و } \bar{x} = 51.60$$

$$c.v = \frac{s_x}{\bar{x}} \times 100 = \frac{11.19}{51.60} \times 100 = 21.68\,\%$$

ـ تلاحظ معي من خلال النتائج أن نسبة معامل الاختلاف ليست كبيرة مقارنة بوسطها الحسابي، وبالتالي يعتبر التشتت هنا ليس بالشكل الذي يطرح مخاوف.

المثال الثاني

نريد دراسة معامل الاختلاف لعينة من الطلبة، حيث تتعلق هذه الدراسة بظاهرة الوزن ن فكانت المعطيات المشاهدة على النحو التالي : 70 كلغ، 85، 60، 75، 55 كلغ .

حل المثال

$s_x = 11.65$ و $\bar{x} = 69$

و بالتالي

$$c.v = \frac{s_x}{\bar{x}} \times 100 = \frac{11.65}{69} \times 100 = 16.88\,\%$$

02 ـ الدرجات المعيارية

يرمز لها عادة بالرمز $\boldsymbol{z}$. إن أهم وظيفة تقوم بها، هي المقارنة بين مفردتين لمجموعتين مختلفتين على الأقل، لذلك يتطلب تحويل قيم مفردات كل مجموعة إلى وحدات قياسية، حتى يكون للمقارنة معنى ودلالة، وذلك من خلال الوسط الحسابي والانحراف المعياري (إذا كان مجتمع).

2 ـ 1 ـ خصائص الدرجات المعيارية : للدرجات المعيارية خصائص نوجزها على النحو التالي.

ـ المتوسط الحسابي لأي توزيع تكراري يساوي صفر (0)

ـ الانحراف المعياري (S) يساوي الواحد (1)

ـ التباين (S^2) يساوي الواحد (1)

ـ $Z_I \sim (0.1)$ أي أن الدرجات المعيارية تتوزع طبيعيا، بوسط حسابي يساوي الصفر وتباين يساوي الواحد.

2 ـ 2 ـ الصيغة الرياضية: يعطى بالصيغة التالية

$$\boldsymbol{z_i = \frac{x_i - \bar{x}}{s}}$$

حيث $\bar{x}$ تمثل الوسط الحسابي لأية مجموعة محل الدراسة.

03 ـ تمارين محلولة

التمرين الأول

نريد معرفة مدى استيعاب طالب في السنة الأولى بكلية الاقتصاد، من خلال عينة من المواد كنا قد اختبرناه فيها، وذلك حسب ما هو مبين في الجدول التالي.

المواد المختبر فيها	x_i علامة الطالب من 20	$\bar{x}$ للمواد الممتحن فيها	$\|x_i - \bar{x}\|$ الانحراف عن المتوسط	S الانحراف المعياري	Z الدرجة المعيارية
الاحصاء الوصفي	17	08	09	05	**1.8**
منهجية البحث العلمي	15	12	03	02	**1.5**
الاقتصاد الجزئي	17	15	02	02	**1.0**
فرنسية	12	10	02	04	**0.5**

تلاحظ من خلال نتائج الجدول، أن استيعاب الطالب في مادة الاحصاء الوصفي كان أفضل من بقية المواد .

التمرين الثاني

x_i	F_i
30 20	03
40 30	61
50 40	132
60 50	153
70 60	140
80 70	51

90 80	02
$\sum Fi$	542

أجريت امتحانات للمجموعة الأولى من طلبة السنة الأولى لكلية الاقتصاد في مختلف المواد للسداسي الأول، فكانت الدرجات النهائية كما هي موضحة في الجدول الموالي .

السؤال :

إذا كان أحد الطلبة قد تحصل على مجموع 70 درجة، فما هي الدرجة المعيارية لهذا الطالب.

حل التمرين

$\bar{x} = 54.7$ و $S = 11.9$ وبالتالي، فإن الدرجة المعيارية هي:

$$z = \frac{x_i - \bar{x}}{s} = \frac{70 - 54.7}{11.9} = 1.3$$

04 ـ الخطأ المعياري للمتوسطات

يقصد بالخطأ المعياري للمتوسطات، بأنه عبارة عن تقدير للانحراف المعياري للمتوسطات الحسابية المحسوبة من عدد من العينات العشوائية الكبيرة الحجم ($n \geq 30$) المأخوذة من مجتمع واحد متجانس.

يعطى بالعلاقة التالية

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

بمعنى قسمة الانحراف المعياري على جذر حجم العينة. تدل نتيجة العلاقة على مدى تبعثر المتوسطات حول متوسط المجتمع، وبالتالي فإن قيمة الخطأ المعياري هي مقياس لدرجة الاعتماد على متوسط العينة.

و على هذا الأساس، فإن المتوسط الذي تكون فيه قيمة الخطأ المعياري له صغيرة، يكون أكثر اعتمادا عليه من المتوسط الذي تكون قيمة الخطأ المعياري له كبيرة.

بمعنى آخر إذا كانت قيمة الخطأ المعياري صغيرة دلت على تركز المتوسطات حول متوسط المجتمع.

05 ـ أمثلة محلولة

المثال الأول

المعطيات التالية تمثل سعر إحدى المواد من خلال ثمان أسواق مختلفة ، والمطلوب حساب الخطأ المعياري:

x_i : 02، 06، 08، 11، 13، 13، 13، 14

حل المثال: $\bar{x} = 10$ و $s = 4$ **وبالتالي** $S_{-} = \frac{s}{\sqrt{n}} = \frac{4}{\sqrt{8}} = 1.41$

المثال الثاني

أحسب الخطأ المعياري لأجور عينة من العمال تتكون من 100 عامل في إحدى المؤسسات محل الدراسة، من خلال معطيات الجدول التالي.

x_i	33 30	36 33	39 36	42 39	45 42	48 45
n_i	02	04	26	47	15	06

حل المثال: $\bar{x} = 43.11$ و $s = 2.955$ ، ومنه $s_{\bar{x}} = 0.2955$

VI ـ العزوم

يلجأ الباحث إلى العزوم، لأن كل من مقاييس النزعة المركزية والتشتت أحيانا لا تكفي لوصف التوزيعات التكرارية والمقارنة بين مختلفها. تستخدم العزوم عادة من أجل قياس كل من التواء التوزيع وتفرطح قمة التوزيع.

01 ـ أنواع العزوم

نميز في العزوم ثلاثة أنواع، وهي كل من: **العزوم الصفرية، العزوم حول قيمة ثابتة (a)، العزوم حول الوسط الحسابي.**

02 ـ أهمية العزوم

تكمن أهمية العزوم المركزية في تحديد شكل التوزيع إن كان متماثلا أو ملتوي، هذا من جهة، ومن جهة أخرى تحدد درجة تد بب أو تفرطح أو اعتدال التوزيع محل الدراسة.

03 ـ القوانين التي تطبق بواسطتها العزوم

3 ـ 1 ـ العزوم الصفرية

3 ـ 1 ـ 1 ـ العزوم الصفرية لبيانات غير مبوبة: نقدمها موجزة في الجدول التالي

العزم الصفري الأول M_1^0	العزم الصفري الثاني M_2^0	العزم الصفري الثالث M_3^0	العزم الصفري الرابع M_4^0

$M_4^0 = \frac{\sum(x_i - 0)^4}{n}$ $M_4^0 = \frac{\sum x_i^4}{n}$	$M_3^0 = \frac{\sum(x_i - 0)^3}{n}$ $M_3^0 = \frac{\sum x_i^3}{n}$	$M_2^0 = \frac{\sum(x_i - 0)^2}{n}$ $M_2^0 = \frac{\sum x_i^2}{n}$	$M_1^0 = \frac{\sum(x_i - 0)^1}{n}$ $M_1^0 = \bar{x}$

مثال: نفرض توفر لدينا المعطيات التالية: $X_I : 01 \ , \ 02 \ , \ 03$ والمطلوب حساب العزوم الصفرية الأربعة

حل المثال

$M_3^0 = \frac{36}{3} = 12$ و $M_2^0 = \frac{14}{3} = 4.6$ و $M_1^0 = \frac{6}{3} = 2$

$M_4^0 = \frac{98}{3} = 32.6$ و

3 – 1 – 2 – العزوم الصفرية لبيانات مبوبة

العزم الصفري الأول	العزم الصفري الثاني	العزم الصفري الثالث	العزم الصفري الرابع
$M_1^0 = \frac{\sum n_i . x_i}{\sum n_i}$ $M_1^0 = \bar{x}$	$M_2^0 = \frac{\sum n_i . x_i^2}{\sum n_i}$	$M_3^0 = \frac{\sum n_i . x_i^3}{\sum n_i}$	$M_4^0 = \frac{\sum n_i . x_i^4}{\sum n_i}$

3 – 2 – العزوم حول قيمة ثابتة (a) وتسمى بالعزوم حول وسطها الفرضي، وهي لا تختلف عن سابقها وتحسب بنفس الطريقة .

3 – 2 – 1 – العزوم حول قيمة ثابتة لبيانات غير مبوبة

العزم الأول حول	العزم الثاني حول	العزم الثالث حول	العزم الرابع حول

الوسط الفرضي $\acute{M}_1$	الوسط الفرضي $\acute{M}_2$	الوسط الفرضي $\acute{M}_3$	الوسط الفرضي $\acute{M}_4$
$\acute{M}_1=\frac{\sum(x_i-A)}{n}$	$\acute{M}_2=\frac{\sum(x_i-A)^2}{n}$	$\acute{M}_3=\frac{\sum(x_i-A)^3}{n}$	$\acute{M}_4=\frac{\sum(x_i-A)^4}{n}$

3 ـ 2 ـ 2 ـ العزوم حول قيمة ثابتة لبيانات مبوبة، حيث يكون الفرق بينها وبين البيانات غير المبوبة، أننا نأخذ بعين الاعتبار التكرارات، كما يوضحه الجدول التالي

$\acute{M}_1$	$\acute{M}_2$	$\acute{M}_3$	$\acute{M}_4$
$=\frac{\sum ni(x_i-A)}{\sum ni}$	$=\frac{\sum ni(x_i-A)^2}{\sum ni}$	$=\frac{\sum ni(x_i-A)^3}{\sum ni}$	$=\frac{\sum ni(x_i-A)^4}{n}$

3 ـ 3 ـ العزوم حول الوسط الحسابي (حول القيمة المركزية) تسمى أحيانا بالعزوم المركزية، وهي كما تعرف بعنوانها، تعبر عن الفرق بين أية قيمة، والقيمة المركزية لها.

3 ـ 3 ـ 1 ـ العزوم حول الوسط الحسابي لبيانات غير مبوبة

العزم المركزي M_1	العزم المركزي M_2	M_3	M_4
$=\frac{\sum(x_i-\bar{x})}{n}$	$=\frac{\sum(x_i-\bar{x})^2}{n}$	$=\frac{\sum(x_i-\bar{x})^3}{n}$	$=\frac{\sum(x_i-\bar{x})^4}{n}$

3 ـ 3 ـ 2 ـ العزوم المركزية لبيانات مبوبة ، تحسب قيمها حسب العلاقات المبينة في الجدول الآتي، وهي لا تختلف عن سابقاتها إلا من حيث التكرارات .

M_1	M_2	M_3	M_4
$=\frac{\sum ni(x_i-\bar{x})^1}{n}$	$=\frac{\sum ni(x_i-\bar{x})^2}{n}$	$=\frac{\sum ni(x_i-\bar{x})^3}{n}$	$=\frac{\sum ni(x_i-\bar{x})^4}{n}$

مثال: إذا كانت العزوم حول الوسط الفرضي 35 هي :

1 ـ العزم الأول حول الوسط الفرضي هو $\acute{M}_1 = -1.9$

2 ـ العزم الثاني حول الوسط الفرضي هو $\acute{M}_2 = 181$

3 ـ العزم الثالث حول الوسط الفرضي هو $\acute{M}_3 = -970$

4 ـ العزم الرابع حول الوسط الفرضي هو $\acute{M}_4 = 98500$

المطلوب: حساب كل من

1 ـ المتوسط الحسابي. 2 ـ العزوم المركزية (الثاني، الثالث، الرابع) . 3 ـ التباين . 4 ـ الانحراف المعياري.
5 ـ العزم الصفري الثاني.

حل المثال

1 ـ يحسب الوسط الحسابي كما يلي: $\bar{x} = A + \acute{M}_1$ وبالتالي $\bar{x} = 35 - 1.9 = 33.1$

وعليه $\bar{x} = M_1^0 = 33.1$

2 ـ العزوم المركزية (الثاني، الثالث، الرابع)

العزم المركزي الثالث M_3	العزم المركزي الثاني M_2
$M_3 = \acute{M}_3 - 3\acute{M}_1\acute{M}_2 + 2(\acute{M}_1)^3$ $M_3 = -970 - 3(-1.9)(181) + 2(-1.9)^3 = 75.418$ $M_3 = 75.418$	$M_2 = \acute{M}_2 - (\acute{M}_1)^2$ $M_2 = 181 - (-1.9)^2 = 177.39$ $M_2 = S^2 = 177.39$ $= 13.32 S = \sqrt{s^2} = \sqrt{177.39}$ أي أن الانحراف القياسي هو 13.32
العزم المركزي الرابع M_4	
$M_4 = \acute{M}_4 - 4\acute{M}_1\acute{M}_3 + 6(\acute{M}_1)^2\acute{M}_2 - 3(\acute{M}_1)^4$ $= 95009.364 M_4 = 98500 - 4(-1.9)(-970) + 6(-1.9)^2(181) - 3(-1.9)^4$	

5 ـ العزم الصفري الثاني: $M_2^0 = M_2 + (M_1^0)^2 = 177.39 + (33.1)^2 = 1273$

IIV ـ الالتواء

يقصد به انحراف التوزيع التكراري لقيم الظاهرة محل الدراسة عن التماثل، وبمعنى آخر هو بعد منحنى القيم عن التماثل (التناظر). يستعمل الالتواء عادة في الدراسات الاحصائية لثلاثة أسباب أساسية.

ـ لتحديد شكل التوزيع الاحصائي

ـ للمقارنة بين توزيعين من حيث شكل التوزيع

ـ لأن المتوسطات والتشتت مقاييس غير كافية لوصف التوزيعات التكرارية ومقارنتها، فقد نجد توزيعان مثلا لهما نفس المتوسط ونفس درجة التشتت، ومع هذا يختلفان من حيث شكل التوزيع.

01 ـ تفسير قيمة الالتواء (معامل الالتواء)

يعطى لقيمة الالتواء بعد حسابه ثلاثة تفاسير نوجزها على النحو التالي

ـ معامل الالتواء مساويا للصفر، معنى ذلك أن توزيع السلسلة الاحصائية لقيم الظاهرة هو توزيع متماثل.

ـ معامل الالتواء موجب (+)، هذا يدل على وجود التواء موجب

ـ معامل الالتواء سالب (-)، هذا يدل على وجود التواء سالب

02 ـ شكل التوزيع: يمكن العودة إلى مميزات الانحراف الربيعي للإطلاع على هذه الأشكال

<table>
<tr><th colspan="2">التوزيع الطبيعي معناه</th></tr>
<tr><td colspan="2">ـ لا يوجد التواء
ـ القيم موزعة بتماثل حول قيمة وسطية معينة
ـ معامل الالتواء يساوي الصفر
ـ الوسط الحسابي = الوسيط = المنوال (لهم نفس القيمة ويقعوا في نقطة واحدة)</td></tr>
<tr><th>التوزيع الملتوي ناحية اليمين معناه</th><th>التوزيع الملتوي ناحية اليسار معناه</th></tr>
<tr><td>ـ التواء موجب
ـ ذيل التوزيع جهة اليمين</td><td>ـ التواء سالب
ـ ذيل التوزيع موجود على اليسار</td></tr>
</table>

ـ القيم المتطرفة (الكبيرة) تتجه ناحية اليمين ـ الوسط الحسابي > الوسيط ـ الوسط الحسابي > المنوال ـ عدد أكبر من القيم كلها أقل من المتوسط، ونجدها على يساره	ـ القيم المتطرفة الصغيرة موجودة على اليسار ـ الوسط الحسابي < الوسيط ـ الوسط الحسابي < المنوال ـ معظم قيم الظاهرة محل الدراسة تتمركز جهة اليمين وتكون أكبر من قيمة المتوسط الحسابي

03 ـ القوانين المستخدمة لحساب معامل الالتواء

لعل أشهر القوانين المستخدمة في هذا المجال، هي طرق كل من *Karl Pearson* و *Bowley's.* نرمز لمعامل الالتواء بالرمز (sk)، وعلى هذا الأساس نقوم بحسابه وفق الطرق التالية.

3 ـ 1 ـ طريقة بيرسون

الطريقة المطلقة	**الطريقة النسبية**
$sk = \bar{x} - m_o$	$sk = \dfrac{\bar{x} - m_o}{s}$ $sk = \dfrac{3(\bar{x} - m_e)}{s}$

3 ـ 2 ـ طريقة با ولي

الطريقة المطلقة	الطريقة النسبية
$sk = Q_3 + Q_1 - 2m_e$	$sk = \dfrac{Q_3 + Q_1 - 2m_e}{Q_3 - Q_1}$

4 ـ ملاحظات هامة: نسجل في هذا الإطار نوعان من الملاحظات

4 ـ 1 ـ ملاحظات عن الالتواء تتعلق بالمعادلات

ـ التوزيعات الملتوية التواء معتدلا تعني: $\bar{x} - m_o = 3(\bar{\ } - m_e)$ حيث أن المتوسطات الثلاثة المشهورة متساوية (الوسط الحسابي، الوسيط، المنوال).

ـ التوزيعات المتماثلة تعني:

$$[Q_3 - Q_2] - [Q_2 - Q_1] = 0$$

أو نقول

$$[Q_3 - m_e] - [m_e - Q_1] = 0$$

ـ التوزيعات الملتوية تعني:

$$[Q_3 - Q_2] - [Q_2 - Q_1] \neq 0$$

أو نقول

$$[Q_3 - m_e] - [m_e - Q_1] \neq 0$$

4 ـ 2 ـ ملاحظات عن الالتواء بواسطة المقارنات

ـ إذا كان لدينا $[Q_3 - m_e] > [m_e - Q_1]$ معناه معامل الالتواء موجبا، والتوزيع ملتويا نحو اليمين.

ـ إذا كان لدينا $[Q_3 - m_e] < [m_e - Q_1]$ معناه معامل الالتواء سالبا، والتوزيع ملتويا نحو اليسار

ـ إذا كان لدينا $[Q_3 - m_e] = [m_e - Q_1]$ معناه معامل الالتواء يساوي صفرا، والتوزيع يكون معتدلا (متماثلا).

05 ـ أمثلة محلولة

المثال الأول: ليكن لدينا التوزيع الاحصائي التالي، والمطلوب حساب معامل الالتواء

$\acute{x}_i$	07	06	05	04	03	02	01	00	$\sum Fi$
F_i	00	01	04	08	19	29	27	12	100

حل المثال

تلاحظ معي من خلال معطيات الجدول أن القيم تمثل مراكز للفئات، وبالتالي يجب التعامل بحذر مع الحل.

يمكن تطبيق العلاقة التالية:

$$sk = \frac{\overline{x} - m_o}{s}$$

1 ـ الوسط الحسابي: باتباع طريقة الوسط الفرضي نجد أن $\bar{x} = 2 + \frac{0}{100} = 2$

2 ـ المنوال: بإتباع القاعدة: $m_o = a + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times k$ نجد بأن المنوال هو $_o = 1.67$

3 ـ بإتباع طريقة الانحراف المعياري تجد بأن التباين هو: $s^2 = 1.78$ وبالتالي يكون الانحراف المعياري هو: $s = \sqrt{1.78} = 1.334$

4 ـ ومنه يمكن حساب معامل الالتواء $\boldsymbol{sk} = \frac{\bar{x} - m_o}{s} = \frac{2-1.7}{1.3}$

ومنه $sk = 0.23$ وهو كما ترى من النتيجة التواء موجب، وبالتالي ذيل التوزيع متجه نحو اليمين، وأن الوسط الحسابي أكبر من المنوال...الخ.

المثال الثاني

أحسب معامل الالتواء مستعينا في حسابه على الطريقة

$\boldsymbol{sk} = \frac{3(\bar{x} - m_e)}{s}$ من خلال معطيات هذا الجدول.

x_i		F_i
110	100	04
120	110	16
130	120	36
140	130	52
150	140	64
160	150	40
170	160	32
180	170	11
$\sum$ Fi		255

حل المثال

1 ـ الوسط الحسابي: $\bar{x} = 143$

2 ـ الوسيط: $m_e = 143.05$

3 ـ الانحراف المعياري:

$$s = 16.10$$

4 ـ وعلى هذا الأساس نجد معامل الالتواء.

$$sk = -0.0093$$

وبما أن الالتواء سالب، معنى ذلك أن ذيل التوزيع متجه نحو اليسار وأن الوسط الحسابي هو أقل من الوسيط...الخ.

فائدة

إذا كان توزيع البيانات ملتويا، هذا يدل على وجود قيم متطرفة على اليمين أو اليسار (موجبة أو سالبة)، وهي تؤثر لا محالة على الوسط الحسابي والانحراف المعياري لذلك على الباحث التفكير في مقاييس أخرى لقياس النزعة المركزية والتشتت، بدلا من الوسط الحسابي والانحراف المعياري.

الفصل الرابع
الارتباط والانحدار

تعريف

يقصد بالارتباط هو تلك العلاقة التي تربط بين متغيرين أو أكثر محل الدراسة، مثل العلاقة بين متغير الدخل ومتغير الاستهلاك، أو العلاقة بين الطلب والعرض والأسعار، القلق والتحصيل المدرسي، الادخار والاستثمار أو الاستثمار والبطالة... الخ.

01 ـ المتغير التابع والمتغير المستقل

من الملاحظ في دراسة العلاقة بين مختلف الظواهر (إن وجدت) ومدى قوتها من ضعفها، أن هناك متغير هو المؤثر، نسميه أحيانا بالمتغير المستقل، مثل تأثير الدخل على زيادة الانفاق (الدخل مؤثر، وبالتالي مستقل)، والآخر متغير متأثر نسميه بالمتغير التابع.

02 ـ قوة العلاقة بين المتغيرين

تأخذ العلاقة بين متغيرين أو أكثر عدة مستويات، نوجزها في الجدول التالي.

1 +	0.8 +	0.5 +	00	0.5 -	0.8 -	1 -

أي أن قيمة الارتباط لا تخرج عن هذا المجال

3 ـ تفسير قيمة الارتباط

من خلال مجال القيم الذي مر بنا نسجل ما يلي.

ـ ارتباط يساوي الصفر، يدل على انعدام العلاقة بين المتغيرين محل الدراسة

ـ ارتباط (0 إلى 0.4) يعبر عن علاقة طردية ضعيفة، ومن (0.5 إلى 0.9) يعبر عن علاقة طردية تزداد قوتها كلما اقتربت من الواحد.

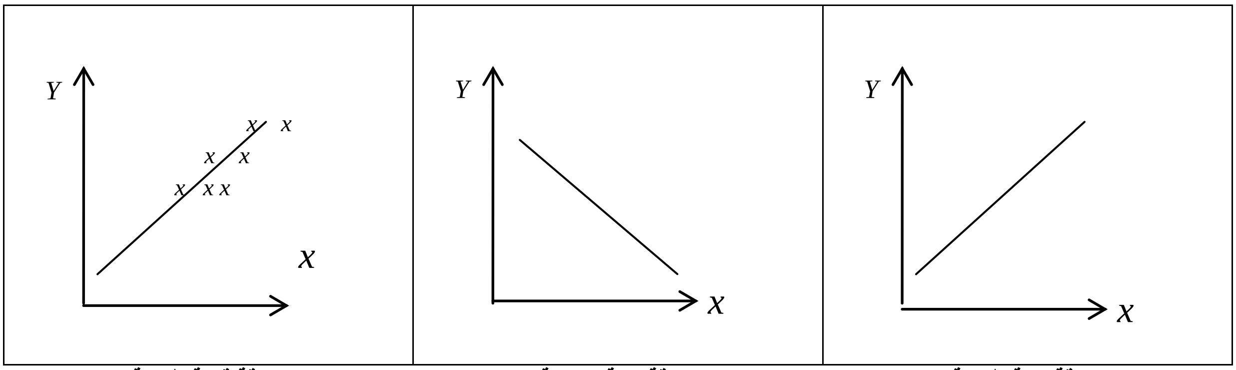
ـ ارتباط يساوي (+1) يعبر عن علاقة طردية تامة

ـ العكس صحيح بالنسبة ليسار المجال، فالارتباط يدل على وجود علاقة عكسية، تزداد قوته كلما اقتربنا من (-1)

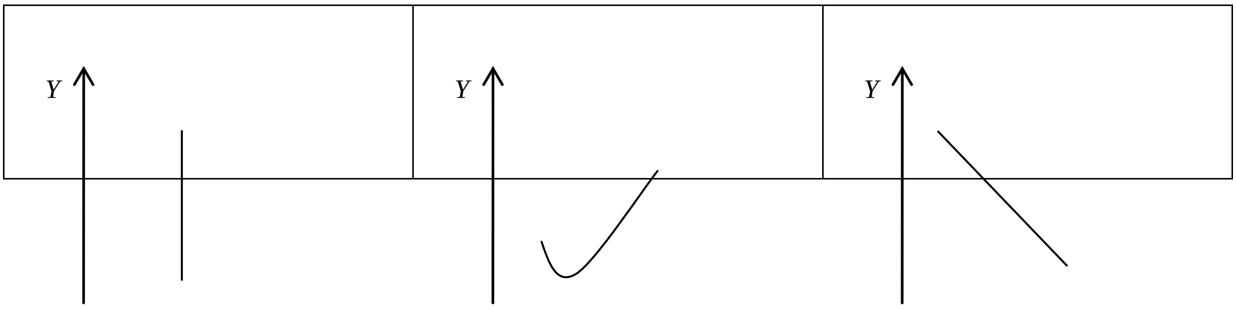
ـ ارتباط (-1) يدل على علاقة عكسية تامة بين المتغيرين محل الدراسة.

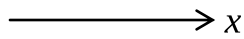
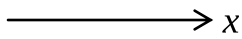
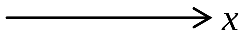
4 ـ نوع العلاقة بين المتغيرين

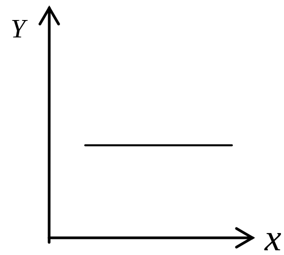
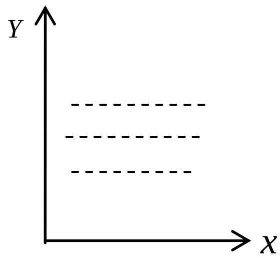
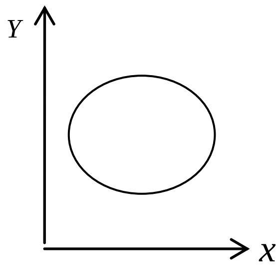
الذي يحدد نوع العلاقة هي سحابة النقاط (لوحة الانتشار) التي ترسم من خلال قيم المتغيرين، في محورين متعامدين أحدهما أفقي، عادة ما يخصص لقيم المتغير المستقل، والآخر عمودي يخصص للمتغير التابع . وعلى هذا الأساس يستنتج من الرسم البياني نوع العلاقة من كونها خطية أم غير ذلك، وكونها طردية أم عكسية أم غير موجودة أصلا، كما تفسرها الأشكال التالية.



علاقة تامة طردية — علاقة تامة عكسية — علاقة قوية طردية



		
علاقة قوية عكسية	علاقة غير خطية	علاقة ضعيفة

		
علاقة ضعيفة	لا توجد علاقة	علاقة غير خطية

5 ـ معادلة خط الانحدار

تعطى عادة على الشكل التالي: $f(x) = y = a + bx_i$

ـ حيث نعتبر x_i متغير مستقل مؤثر.

ـ كما نعتبر y_i متغير تابع متأثر.

ـ نعتبر a ثابت ، أما b فهو نسبة التغير في خط الانحدار للمعادلة، حينما يتغير x_i

ـ لحساب الثوابت واستنتاج المعادلة، نقوم بإتباع الخطوات التالية.

الخطوة الأولى: حساب نسبة التغير في خط الانحدار وفق المعادلة التالية $b = \frac{\sum x_i \quad _i - n . \overline{xy}}{\sum x_i^2 - n . \overline{x}^{\,2}}$

الخطوة الثانية: حساب الثابت وفق المعادلة التالية $a = \bar{y} - b\bar{x}$

ملاحظة: لتسهيل عمليات الحساب، الأفضل الاستعانة بجدول مساعد كما سيأتي في المثال.

المثال الأول: المعطيات التالية تبين العلاقة بين سعر الغرفة في فندق وعدد السياح الوافدين إله.

سعر الغرفة: وحدة نقدية (x_i)	02	03	1.5	04	06	08
عدد السياح (y_i)	40	35	45	30	25	15

المطلوب: رسم السحابة، حساب الثوابت $(a\,,b)$، استنتاج معادلة خط الانحدار، قدر عدد السياح إذا كان سعر الغرفة 3.5 وحدة نقدية.

حل المثال:

ـ نقوم بحساب المتوسط الحسابي لكلا المتغيرين

$\bar{x} = \frac{\sum x_i}{n} = \frac{24.5}{6} = 4$ و $\bar{y} = \frac{\sum y_i}{n} = \frac{190}{6} = 31.66$ وبالتالي يكون

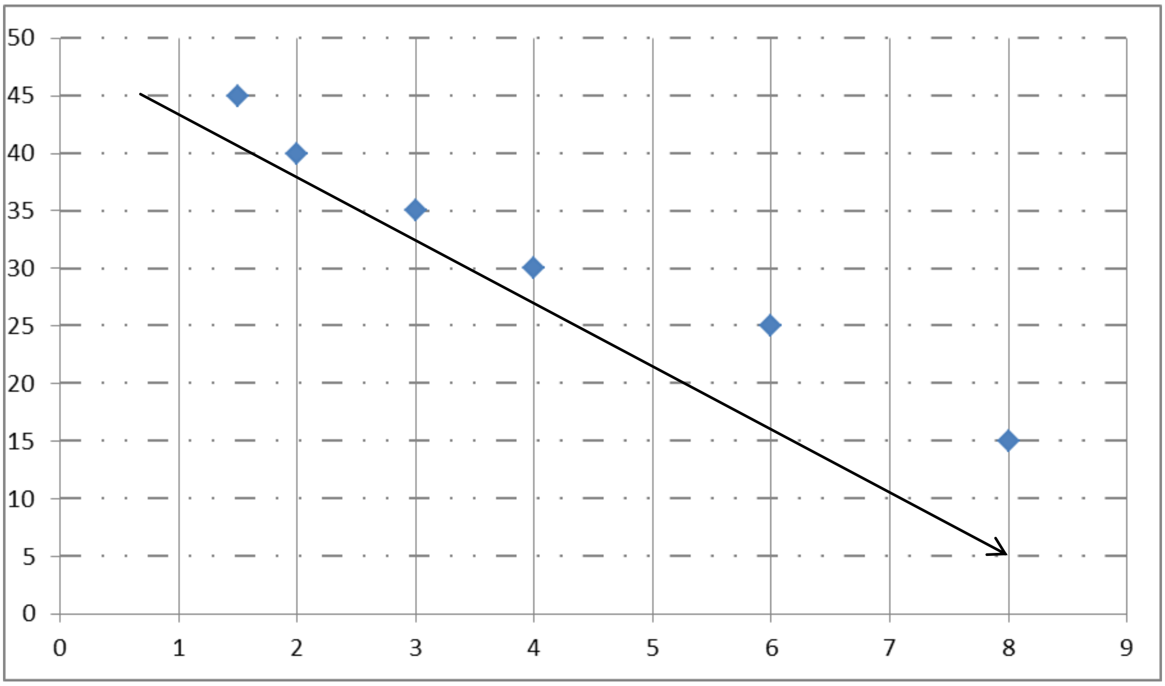
$16(\bar{x})^2 =$ و $\bar{x}.\bar{y} = 126.64$

ـ انجاز الجدول المساعد

x_i	x_i^2	y_i	$x_i y_i$
02	04	40	80
03	09	35	105
1.5	2.25	45	67.5
04	16	30	120
06	36	25	150
08	64	15	120
24.5	131.25	190	642.5

وبناء على نتائج الجدول يمكن حساب المطلوب

ـ رسم السحابة



تلاحظ من خلال الشكل أن العلاقة بين المتغيرين هي علاقة عكسية، أي كلما كان سعر الغرفة (المؤثر) منخفضا كان الاقبال أكثر، والعكس صحيحا.

ـ حساب الثوابت $(a \;\; , \;\; b)$

$$b = \frac{x_i y_i - n\overline{xy}}{\sum x_i^2 - n\bar{x}^2} = \frac{642.5 - 6(126.64)}{131.25 - 6(16)} = \frac{-117.34}{35.25} = -3.32$$

$$a = \bar{y} - b\bar{x} = 31.66 - (-3.32)(4) = 44.94$$

ـ استنتاج المعادلة:

$$y_i = 44.94 - 3.32x_i$$

ـ عدد السواح إذا كان سعر الغرفة 3.5 وحدة نقدية: وذلك بالتعويض في المعادلة التي توصلنا إليها نجد $y_{3.5} = 44.94 - 3.32(3.5) = 33.32$ أي أن عدد السياح في حدود 33 سائح.

06 ـ ملاحظات هامة

من خلال الشكل الذي مر بنا تلاحظ معي أن هناك قيم قريبة من الخط المستقيم، وبعضها بعيدة عنه.

ـ كان الأولى أن تكون تلك النقاط كلها على استقامة واحة، أي على الخط، ولما كانت بعيدة عنه جهة اليمين أو جهة اليسار معناه هناك أخطاء، أي انحراف القيم عن خط الانحدار ولذلك فإن شكل الانتشار النموذجي عادة ما يأخذ المعادلة التالية $y_i = a + bx_i + e_i$ حيث e_i تمثل تلك الاخطاء (المتغيرات العشوائية).وللحصول على أحسن خط رياضيا يجب أن يكون مجموع الأخطاء مساويا للصفر، لذلك نستعمل المربعات الصغرى للتقليص إلى ابعد حد من تلك الفرو قات.

ـ تلاحظ من خلال المثال السابق أننا استعملنا متغير مستقل واحد، وهذا يعني أننا نتعامل مع نموذج خطي بسيط (تحليل الانحدار البسيط)، أما إذا كانت العلاقة بين متغير تا بع واحد وعدة متغيرات مستقلة، نكون أمام نموذج خطي متعدد (تحليل الانحدار المتعدد).

ـ إذا كانت العلاقة خطية في هذا المثال، يمكن أن تكون غير خطية حين نتعامل مع معطيات أخرى

ـ من خلا ل معادلة الانحدار التي تحصلنا عليها استطعنا أن نتنبأ بعدد الوافدين إلى الفندق عند سعر معين للغرفة، هذا يعني أن تحليل الانحدار هو أحد الأساليب الاحصائية الأساسية التي تستعمل في التنبؤ بسلوك الظواهر الاقتصادية.

ـ المفهوم الاقتصادي لمعامل الانحدار (b) يكمن في كونه مقياس كمي، قيمته تزيد أو تقل عن الواحد، كما أنه من خلال الاشارة التي يحملها تعرف نوع العلاقة بين المتغيرين (طر دية أم عكسية) وفي مثالنا السابق يتأكد بأن العلاقة عكسية بين المتغيرين بسبب الاشارة السالبة لمعامل الانحدار. كما يشترط وجود علاقة دالية بين المتغيرين محل الدراسة.

ـ لكل علاقة انحدارية علاقة ارتباطية، ولكن يختلفان في كيفية الحساب والتحليل.

7ـ معامل الارتباط

يلجأ الباحث إلى حساب معامل الارتباط بين متغيرين أو أكثر حينما يتعذر عليه معرفة الصفة محل الدراسة المستقلة (المؤثرة) من الصفة الأخرى المتأثرة، نظرا للتداخل الموجود بينهما. وعلى هذا الأساس يفترض مرة أن x_i هو المستقل ويقوم بحساب معامل الانحدار(b) كما رأينا في معادلة خط الانحدار، ومرة أخرى يفترض أن y_i هو المتغير المستقل، ويقوم بحساب معامل الانحدار $(\acute{b})$.

ـ يرمز إلى معامل الارتباط بالرمز (r)، حيث يكون بسيطا إذا كانت الدراسة تتعلق بمتغيرين فقط، أما إذا كانت تتعلق بين ظاهرة وعدة ظواهر أخرى يسمى بالارتباط المتعدد.

8ـ حساب معامل الارتباط

إذا كنا أما م تطبيق طرق بيرسون و سيبرمان، ننبه الباحث إلى أن معامل الارتباط لبيرسون لا يصلح إلا في حالة البيانات الكمية فقط، أما معامل ارتباط سيبرمان، فهو يصلح استعماله للبيانات الكمية والبيانات الوصفية بصفة أخص. يستنتج من العلاقة

$$\sqrt{b.\acute{b}} = r_{(x\,,y)}$$

نعوض بما يساويها ونختار علاقة.

8 – 1– علاقة التعريف

$$r_{(x\,,y)} = \sqrt{\frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} \cdot \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum y_i^2 - n\bar{y}^2}}$$

أي أن $(\acute{b})$ تعبر عن معامل الانحدار للعلاقة $x_i = \acute{a} + \acute{b}y_i + e_i(y_i) =$ مما يدل على افتراضنا بأن y_i هو المتغير المستقل والآخر (x_i) متغير تابع .

8 – 2 – العلاقة الموسعة

$$r(x\,,y) = \frac{\sum x_i y_i - n\overline{xy}}{\sqrt{\sum x_i^2 - n\,\bar{x}^2 . \sum y_i^2 - n\bar{y}^2}}$$

لو تأملت إلى ما تحت الجذر في العلاقتين لوجدته عبارة عن الانحراف القياسي الذي درسناه في مقاييس التشتت.وأن صورة الكسر في العلاقة الثانية تعبر عن التزاوج بين المتغيرين، تسمى بتغاير $(x\ ,\ y)$ لذلك يمكن التعويض بما يساويهما فنجد.

$$\boldsymbol{r(x\,,y)} = \frac{\boldsymbol{cov(x\,,y)}}{\boldsymbol{sd_x.sd_y}} = \frac{\text{تغاير}\ \boldsymbol{(x\,,y)}}{\boldsymbol{sd_x.sd_y}} \quad \text{مع} \quad -\boldsymbol{1} \leq \boldsymbol{r_{x,y}} \leq \boldsymbol{1}$$

09 – أمثلة محلولة

المثال الأول: الجدول التالي يبين عدد الطائرات المستعملة (x_i) والربح المحقق (y_i) لخمس شركات طيران

210	100	60	20	10	x_i

y_i	10	10	30	40	110

المطلوب: معرفة قوة الارتباط بين المتغيرين محل الدراسة.

حل المثال: لتحقيق غاية الحل نستعين بالجدول التالي

x_i	y_i	x_iy_i	x_i^2	y_i^2
10	10	100	100	100
20	10	200	400	100
60	30	1800	3600	900
100	40	4000	10000	1600
210	110	23100	44100	12100
400	200	29200	58200	14800

$\sum x_i^2 = 58200 \bar{x} = \frac{\sum x_i}{n} = \frac{400}{5} = 80$ و $\bar{y} = \frac{\sum y_i}{n} = \frac{200}{5} = 40$ و

$\bar{y}^2 = 1600$ و $\bar{x}^2 = 6400$ و$\sum y_i^2 = 14800$ من خلال هذه المعطيات يمكن حساب

معامل الارتباط وفق القواعد التي مرت بنا.

1 ـ معادلة خط الانحدار $b = \frac{\sum x_i \quad _i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{13200}{26200} = 0.50$ ومنه $a = \bar{y} - b\bar{x} = 0$

وعليه تكون المعادلة: $y_i = 0.5x_i$

2 ـ لحساب معامل الارتباط علينا افتراض أن $\mathrm{y_i}$ هو المتغير المستقل، وبالتالي

$\acute{b} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum y_i^2 - n\bar{y}^2} = \frac{13200}{6800} = 1.94$ ومنه $\acute{a} = \bar{x} - \acute{b}\bar{y} = 2.4$

ومنه $r(x,y) = \sqrt{b\acute{b}} = \sqrt{0.97} = 0.98$

وهو كما تلاحظ معامل الارتباط قريب من الواحد الصحيح الموجب، مما يدل على ارتباط قوي جدا وطردي بين المتغيرين محل الدراسة، بمعنى كلما زاد عدد الطائرات زادت الأرباح المحققة.

3 ـ حساب الربح المقدر حينما يصل عدد الطائرات إلى 250 طائرة. نعوض في معادلة خط الانحدار الأولى نحصل على الربح المتوقع. $y_{250} = 0.5(250) = 125$

المثال الثاني: من أجل دراسة العلاقة بين سلسلة إحصائية X وسلسلة إحصائية Y نقدم المعطيات التالية.

$\sum_{i=1}^{i=10} y_i^2 = 430$ و $\sum_{i=1}^{i=10} x_i^2 = 565$ و $\sum_{i=1}^{i=10} y_i\ 64 =$ و $\sum_{i=1}^{i=10} x_i = 73$

$\sum_{i=1}^{i=10} x_i y_i = 465$

والمطلوب هو:

ـ حساب الوسط الحسابي والتباين لكلا المتغيرين ثم حساب تغاير (x, y)

ـ حساب معامل الارتباط الخطي.

ـ نستعمل قيمتين إضافيتين مع تقديم المعطيات التالية.

$\sum_{i=11}^{i=12} y_i^2 = 148$ و $\sum_{i=11}^{i=12} x_i^2 = 260$ و $\sum_{i=11}^{i=12} y_i = 14$ و $\sum_{i=11}^{i=12} x_i = 18$

$\sum_{i=11}^{i=12} x_i y_{i=196}$

المطلوب حساب معامل الارتباط الخطي الجديد.

حل المثال

ـ متوسط السلسلتين هو:

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{i=10} x_i = \frac{1}{10}.73 = 7.3$$

$$\bar{y} = \frac{1}{10}\sum_{i=1}^{i=10} y_i = \frac{1}{10}.64 = 6.4$$

ـ أما تباين السلسلتين فهو كالآتي.

$$s_x^2 = \frac{1}{10}\sum_{i=1}^{i=10} x_i^2 - \bar{x}^2 = \frac{1}{10}.565 - (7.3)^2 = 56.5 - 53.29 = 3.21$$

$$s_y^2 = \frac{1}{10}\sum_{i=1}^{i=10} y_i^2 - \bar{y}^2 = \frac{1}{10}430 - (6.4)^2 = 43 - 40.96 = 2.04$$

ـ حساب تغاير (x, y)

$$s_{x,y} = \frac{1}{10}\sum_{i=1}^{i=10} x_i y_i - \overline{xy} = \frac{1}{10}.465 - (7.3)(6.4) = 46.5 - 46.72 = -0.22$$

ـ معامل الارتباط الخطي للسلسلتين (x, y)

$$r(x,y) = \frac{s_{xy}}{s_x s_y} = \frac{-0.22}{\sqrt{3.21}.\sqrt{2.04}} = \frac{-0.22}{2.56} = -0.09$$

ـ المعطيات الاضافية تعالج على النحو التالي

$$\sum_{i=1}^{i=10} y_i + \sum_{i=11}^{i=12} y_i = \sum_{i=1}^{i=12} y_i = 78 \text{ و } \sum_{i=1}^{i=10} x_i + \sum_{i=11}^{i=12} x_i = \sum_{i=1}^{i=12} x_i = 91$$

$$\text{و } \sum_{i=1}^{i=10} x_i^2 + \sum_{i=11}^{i=12} x_i^2 = \sum_{i=1}^{i=12} x_i^2 = 825$$

$$\sum_{i=1}^{i=10} y_i^2 + \sum_{i=11}^{i=12} y_i^2 = \sum_{i=1}^{i=12} y_i^2 = 578$$

$$\sum_{i=1}^{i=10} x_i y_i + \sum_{i=11}^{i=12} x_i y_i = \sum_{i=1}^{i=12} x_i y_i = 661$$

وعلى أساس هذه النتائج تكون قيم متوسط السلاسل الاحصائية الجديدة على النحو التالي

$$\bar{y} = \frac{1}{12}\sum_{i=1}^{i=12} y_i = \frac{1}{12}78 = 6.5 \text{ و } \bar{x} = \frac{1}{12}\sum_{i=1}^{i=12} x_i = \frac{1}{12}91 = 7.58$$

ـ وبالتالي تباين السلاسل الاحصائية الجديد يكون كالتالي

$$s_x^2 = \frac{1}{12}\sum_{i=1}^{i=12} x_i^2 - \bar{x}^2 = \frac{1}{12}825 - (7.58)^2 = 68.75 - 57.46 = 11.29$$

$$s_y^2 = \frac{1}{12}\sum_{i=1}^{i=12} y_i^2 - \bar{y}^2 = \frac{1}{12}578 - (6.5)^2 = 48.17 - 42.25 = 5.92$$

ـ ونفس الشيء بالنسبة لتغاير (x, y)

$$s_{xy} = \frac{1}{12}\sum_{i=1}^{i=12} x_i y_i - \overline{xy} = \frac{1}{12}661 - (7.58)(6.5) = 55.08 - 49.27 = 5.81$$

ـ وبالتالي معامل الارتباط الخطي الجديد ما بين السلسلتين الاحصائيتين (X) و(Y) محل الدراسة يكون

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{5.81}{\sqrt{11.29}\sqrt{5.92}} = \frac{5.81}{8.16} = 0.71$$

المثال الثالث: الجدول الآتي يبين قيم إحصائية تتعلق السعر(Y) والكمية المتوفرة (X) من منتج في أحد الأسواق، حيث يعطى(Y) بالدينار الجزائري و(X) عدد المنتج من كمية السلعة.

x_i	02	04	06	12	15	20	22	24	28	30
y_i	98	90	84	72	64	62	55	54	45	40

المطلوب ما يلي:

1 ـ العرض البياني لسحابة القيم (النقاط)

2 ـ حساب الثوابت (a , b) لمعادلة خط الانحدار $y_i = a + bx_i$ التي تمثل الانحدار والقيمة المتوسطة.

3 ـ حساب معامل الارتباط الخطي.

4 ـ حساب الاخطاء (e_i) لكل قيم (x_i) للمتغير (X) المرافقة لتقديرات الثوابت (a , b) ثم أحسب تباين الأخطاء التي سجلتها (s_e^2) .

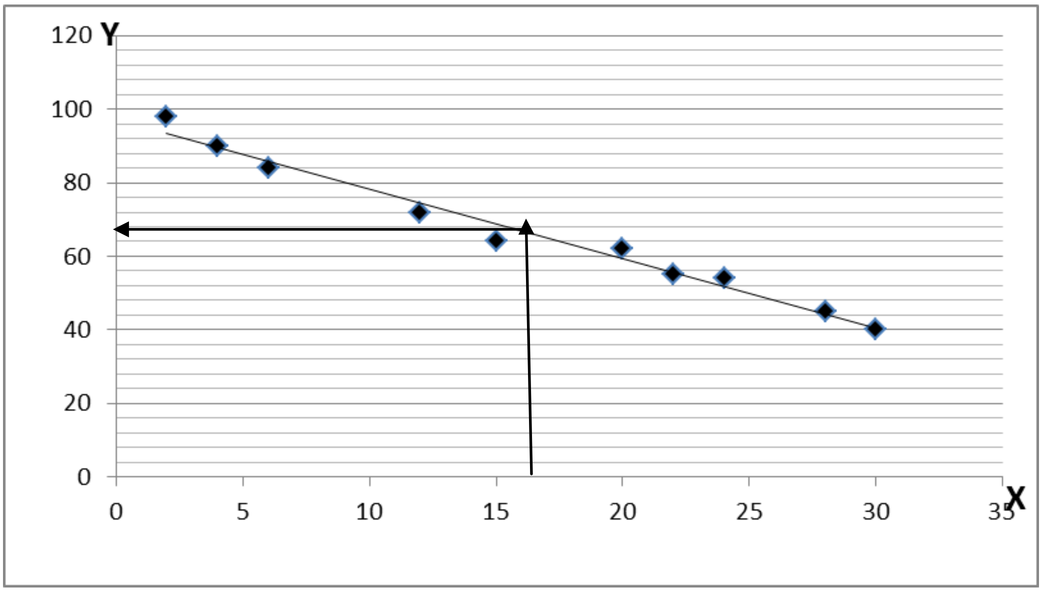
5 ـ ما هو السعر المنتظر إذا توفرت الكمية (05) ثم الكمية (26).

حل المثال

لتسهيل حل المثال يقتضي منا الاستعانة بالجدول الموالي الذي يبين مختلف الحسابات المساعدة للحل.

الرقم	x_i	y_i	x_i^2	y_i^2	x_iy_i
01	02	98	04	9604	196
02	04	90	16	8100	360
03	06	84	36	7056	504
04	12	72	144	5184	864
05	15	64	225	4096	960
06	20	62	400	3844	1240
07	22	55	484	3025	1210
08	24	54	576	2916	1296
09	28	45	784	2025	1260
10	30	40	900	1600	1200

1– العرض البياني لسحابة النقاط (القيم) وخط الانحدار مع الوسط الحسابي للمتغيرين (x, y)



من خلال الشكل يتضح أن العلاقة بين المتغيرين محل الدراسة، هي علاقة عكسية، بمعنى كلما زادت الكمية من المنتج المعروض أدت إلى انخفاض السعر. وبما أن نقاط القيم قريبة جدا من الخط المستقيم دليل على قوة الارتباط.

2 ـ حساب (a , b)، قبل حساب الثوابت علينا حساب المتوسط الحسابي لكلا المتغيرين.

$\bar{x} = 16.3$ ، وكذلك $s_x^2 = 91.21$ ومنه $s_x = 9.55$

$\bar{y} = 66.4$ ، وكذلك $s_y^2 = 336.04$ ومنه $s_y = 18.33$ ومنه $s_{xy} = -173.32$

$b = \frac{s_{xy}}{s_x^2} = \frac{-173.32}{91.21} = -1.9002$ ومنه $a = \bar{y} - b\bar{x} = 97.37$

3 ـ معامل الارتباط الخطي

$$r = \frac{s_{xy}}{s_x s_y} = \frac{-173.32}{(9.55)(18.33)} = -0.99$$

كما تلاحظ هو قريب من (1) بالقيمة المطلقة ويتوافق تماما مع الشكل البياني السابق لقرب سحابة النقاط من المستقيم، كما تدل الإشارة السالبة على العلاقة العكسية بين المتغيرين محل الدراسة.

4 ـ لحساب الأخطاء، يجب حساب النقاط التقديرية لمنحنى المربعات الصغرى . وعليه لابد من حساب كل زوج من خلال معادلة خط الانحدار $\acute{y}_i = -1.90.x_i + 97.37$

لنحسب مثلا من أجل i=1 سوف تجد

$$\acute{y}_1 = -1.90.(2) + 97.37 = 93.57$$

بعد هذا نقوم بحساب الفرق بين القيمة المشاهدة في السلسلة والقيمة التي تحصلنا عليها (93.57) كما يلي:

$$e_1 = y_1 - \acute{y}_1 = 98 - 93.57 = 4.43$$

وبهذه الطريقة نكمل باقي القيم كما هو موضح في الجدول الموالي.

الرقم	x_i	y_i	$\acute{y}_i$	e_i	e_i^2
01	02	98	93.57	4.43	19.60
02	04	90	89.77	0.23	0.05
03	06	84	85.97	1.97 -	3.89
04	12	72	74.57	2.57 -	6.61
05	15	64	68.87	4.87 -	23.72
06	20	62	59.37	2.63	6.92
07	22	55	55.57	0.57 -	0.32
08	24	54	51.77	2.23	4.98
09	28	45	44.17	0.83	0.69
10	30	40	40.37	0.37 -	0.13
المجموع	163	664	664	0.00	66.92

تأكد لنا بأن مجموع الأخطاء (e_i) مساوية للصفر، أي أن $\sum_{i=1}^{i=10} ei = 0$

من جهة أخرى نجد أن متوسط الأخطاء يساوي الصفر، أي أن $\bar{e} = \frac{1}{10} \sum_{i=1}^{i=10} e_i = 0$

وبالتعريف، فإن تباين الأخطاء هو:

$$s_e^2 = \frac{1}{n}\sum_{i=1}^{i=n}(e_i - \bar{e})^2$$

$$s_e^2 = \frac{1}{n}\sum_{i=1}^{i=n} e_i^2 = \frac{1}{10}\sum_{i=1}^{i=10} e_i^2 = \frac{1}{10}66.92 = 6.692$$

الآن علينا التأكد من أن المتوسط الحسابي للمتغير التابع هو نفسه المتغير الحسابي المقدر، أي هل $(\bar{y} = \bar{\acute{y}})$.

$\bar{y} = \frac{1}{10}\sum_{i=1}^{i=10} y_i = \frac{664}{10} = 66.4$ $\quad \bar{\acute{y}} = \frac{1}{10}\sum_{i=1}^{i=10} \acute{y} = \frac{664}{10} = 66.4$ تلاحظ معي لهما نفس النتيجة.

5 ـ للجواب على هذا السؤال بكل بساطة نعوض في معادلة خط الانحدار القيمة (x_i) بالكمية (5) فتكون النتيجة كما يلي:

$$\acute{y}_i = -1.90.(5) + 97.37 = 87.87$$

وبنفس الطريقة بالنسبة للكمية $x_i = 26$

$$\acute{y}_{26} = -1.90.(26) + 97.37 = 47.97$$

أي أن الكمية 26 تقتضي 47.97 دينار جزائري.

09 ـ معامل الارتباط الرتبي: يدرس العلاقة بين (x , y) في حالات محددة مثل:

ـ عدم معرفة توزيع المتغيرين قيد الدراسة

ـ بيانات كيفية قابلة للترتيب ، وحالة البيانات الكمية . وعلى هذا الأساس يمكن حساب معامل الارتباط (سبيرمان) وفق الصيغة التالية.

$$r_s = 1 - \frac{6\sum_{i=1}^{i=n} {d_i}^2}{n(n^2-1)}$$

مثال : تبين المعطيات التالية في الجدول أسفله الدرجة التقديرية لستة (06) طلبة في امتحاني الإحصاء والرياضيات ، حيث تمثل (y_i) درجة الاحصاء و (x_i) درجة الرياضيات

رقم الطالب	الدرجة التقديرية للرياضيات x_i	الدرجة التقديرية للإحصاء y_i

01	ضعيف	مقبول
02	ممتاز	جيد جدا
03	جيد	جيد
04	ضعيف جدا	ضعيف
05	مقبول	ضعيف جدا
06	جيد جدا	ممتاز

خطوات الحل

ـ أولا: نقوم بترتيب الدرجات التقديرية لكل مقياس ترتيبا تصاعديا. حسب الجدول التالي.

ترتيب درجات الرياضيات x_i		ترتيب درجات الاحصاء y_i	
الدرجات	الترتيب	الدرجات	الترتيب
ممتاز	01	ممتاز	01
جيد جدا	02	جيد جدا	02
جيد	03	جيد	03
مقبول	04	مقبول	04
ضعيف	05	ضعيف	05
ضعيف جدا	06	ضعيف جدا	06

ـ ثانيا: نقوم بحساب مربع الفروق وفق الجدول أسفله.

d_i^2	d_i	ترتيب x_i	ترتيب y_i	x_i	y_i
01	01	05	04	ضعيف	مقبول
01	01-	01	02	ممتاز	جيد جدا
00	00	03	03	جيد	جيد
01	01	06	05	ضعيف جدا	ضعيف
04	02-	04	06	مقبول	ضعيف جدا
01	01	02	01	جيد جدا	ممتاز
08				المجموع	

10 ـ تمارين

التمرين الأول

ـ مؤسسة إنتاجية تريد معرفة تقديرات حجم المبيعات (X) للسنتين 2016 و2017.لأجل ذلك عليها دراسة حجم البيع للفترة 2006 ـ 2016.النتائج تعطى في الجدول أسفله.

السنة	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
التاريخ t	01	02	03	04	05	06	07	08	09	10
البيع المنجز	3400	4000	3200	3700	3600	3100	3300	3500	4200	4100

1

ـ أحسب المعلمتين (a, b) لمعادلة الانحدار $x = bT + a$ حيث t تمثل السنة ما قبل (2005) وتعطى على النحو التالي.

$\sum_{t=1}^{t=10} x_t = 36100$ $\sum_{t=1}^{t=10} tx_i = 202300$ $\sum_{t=1}^{t=10} t^2 = 385$ $\sum_{t=1}^{t=10} t = 55$

2ـ ما هو الحجم المقدر لمبيعات المنتج (x) للسنتين 2016 و 2017 . أعرض بيانيا سحابة النقاط، خط الانحدار، النقاط المقدرة.

التمرين الثاني

أجب : (نعم) أم (لا)

ـ سحابة نقاط على شكل كرة ، هي مؤشر على وجود ارتباط خطي.

ـ العلاقة الآتية هي دائما صحيحة

$$s_e^2 = \frac{1}{n}\sum_{i=1}^{i=n} (e_i - \bar{e})^2 = \frac{1}{n}\sum_{i=1}^{i=n} e_i^2$$

ـ العلاقة الآتية تعني أنها دوما صحيحة

$$s_e^2 = s_y^2(1 - r^2)$$

ـ معامل الارتباط الخطي هو عدد غير مجزأ ، مبدأ المربعات الصغرى يهدف إلى تقليص مجموع مربع الأخطاء

ـ خط (بيان) المربعات الصغرى يمر في بعض الحالات على القيمة الوسطية ، حين يكون معامل الارتباط الخطي قريب من الصفر معناه وجود ارتباط خطي قوي

ـ يمكن حساب خط الانحدار y للمتغير المستقل x وخط الانحدار x للمتغير المستقل y

التمرين الثالث

البيانات التالية تمثل مردود الكميات المحصلة من الزيتون والمقدرة بالطن ، ويعبر عنها (y_i) الناجمة عن السماد المستخدم المقدر ب (كلغ) ويعبر عنه (xi)

x_i	01	02	05	03	02	01	04	18
y_i	03	04	06	03	05	04	08	33

إنجاز كل من : سحابة النقاط ، معادلة خط الانحدار بالنسبة ل (y_i) و (x_i) ، حساب معامل الارتباط (r) مع تفسير النتيجة ، حساب الأخطاء (e_i) لكل قيم (x_i) ثم تباين الأخطاء s_e^2 ثم أحسب ما هي الكمية المنتظر تحصيلها إذا ارتفع السماد المستخدم إلى 10 كلغ للشجرة الواحدة.

الفصل الرابع

الارتباط والانحدار

تعريف:

يقصد بالارتباط هو تلك العلاقة التي تربط بين متغيرين أو أكثر محل الدراسة، مثل العلاقة بين متغير الدخل ومتغير الاستهلاك، أو العلاقة بين الطلب والعرض والأسعار، القلق والتحصيل المدرسي، الادخار والاستثمار أو الاستثمار والبطالة... الخ.

01 ـ المتغير التابع والمتغير المستقل

من الملاحظ في دراسة العلاقة بين مختلف الظواهر (إن وجدت) ومدى قوتها من ضعفها ، أن هناك متغير هو المؤثر، نسميه أحيانا بالمتغير المستقل ، مثل تأثير الدخل على زيادة الانفاق (الدخل مؤثر، وبالتالي مستقل)، والآخر متغير متأثر نسميه بالمتغير التابع.

02 ـ قوة العلاقة بين المتغيرين

تأخذ العلاقة بين متغيرين أو أكثر عدة مستويات، نوجزها في الجدول التالي.

+ 1	+ 0.8	+ 0.5	00	- 0.5	- 0.8	- 1

أي أن قيمة الارتباط لا تخرج عن هذا المجال

3 ـ تفسير قيمة الارتباط: من خلال مجال القيم الذي مر بنا نسجل ما يلي.

ـ ارتباط يساوي الصفر، يدل على انعدام العلاقة بين المتغيرين محل الدراسة

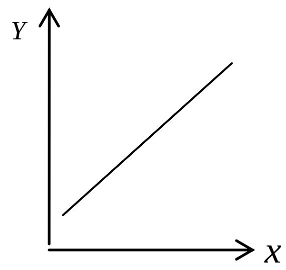
ـ ارتباط (0 إلى 0.4) يعبر عن علاقة طردية ضعيفة، ومن (0.5 إلى 0.9) يعبر عن علاقة طردية تزداد قوتها كلما اقتربت من الواحد.

ـ ارتباط يساوي (+1) يعبر عن علاقة طردية تامة

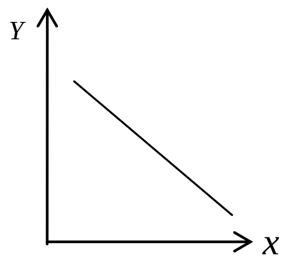
ـ العكس صحيح بالنسبة ليسار المجال، فالارتباط يدل على وجود علاقة عكسية، تزداد قوته كلما اقتربنا من (-1)

ـ ارتباط (-1) يدل على علاقة عكسية تامة بين المتغيرين محل الدراسة.

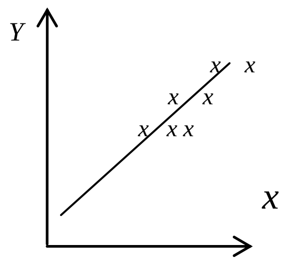
4 ـ نوع العلاقة بين المتغيرين: الذي يحدد نوع العلاقة هي سحابة النقاط (لوحة الانتشار) التي ترسم من خلال قيم المتغيرين، في محورين متعامدين أحدهما أفقي، عادة ما يخصص لقيم المتغير المستقل، والآخر عمودي يخصص للمتغير التابع. وعلى هذا الأساس يستنتج من الرسم البياني نوع العلاقة من كونها خطية أم غير ذلك، وكونها طردية أم عكسية أم غير موجودة أصلا، كما تفسرها الأشكال التالية.



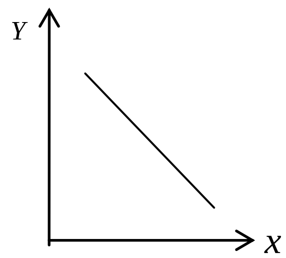
علاقة تامة طردية



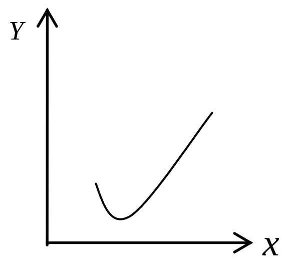
علاقة تامة عكسية



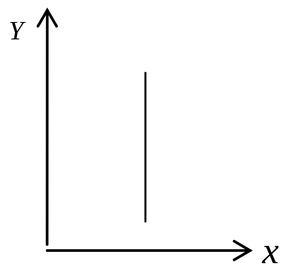
علاقة قوية طردية



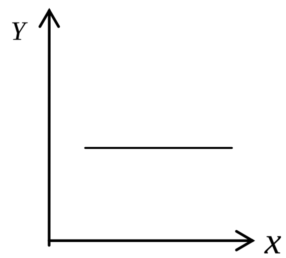
علاقة قوية عكسية



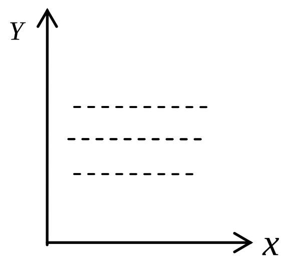
علاقة غير خطية



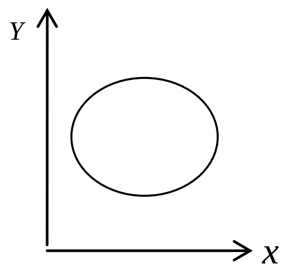
علاقة ضعيفة



علاقة ضعيفة



لا توجد علاقة



علاقة غير خطية

5 ـ معادلة خط الانحدار

تعطى عادة على الشكل التالي: $f(x) = y = a + bx_i$

ـ حيث نعتبر x_i متغير مستقل مؤثر، كما نعتبر y_i متغير تابع متأثر.

ـ نعتبر a ثابت ، أما b فهو نسبة التغير في خط الانحدار للمعادلة، حينما يتغير x_i

ـ لحساب الثوابت واستنتاج المعادلة، نقوم بإتباع الخطوات التالية.

الخطوة الأولى: حساب نسبة التغير في خط الانحدار وفق المعادلة التالية

$$b = \frac{\sum x_i y_i \; - \; . \; \overline{xy}}{\sum x_i^2 \; - \; n \, . \, \overline{x}^{\,2}}$$

الخطوة الثانية: حساب الثابت وفق المعادلة التالية $a = \bar{y} - b\bar{x}$

ملاحظة: لتسهيل عمليات الحساب، الأفضل الاستعانة بجدول مساعد كما سيأتي في المثال.

المثال الأول : المعطيات التالية تبين العلاقة بين سعر الغرفة بالفندق وعدد السياح الوافدين إله.

سعر الغرفة: وحدة نقدية (x_i)	02	03	1.5	04	06	08
عدد السياح (y_i)	40	35	45	30	25	15

المطلوب : رسم السحابة ، حساب الثوابت $(a\,,b)$، استنتاج معادلة خط الانحدار، قدر عدد السياح إذا كان سعر الغرفة 3.5 وحدة نقدية.

حل المثال

ـ نقوم بحساب المتوسط الحسابي لكلا المتغيرين

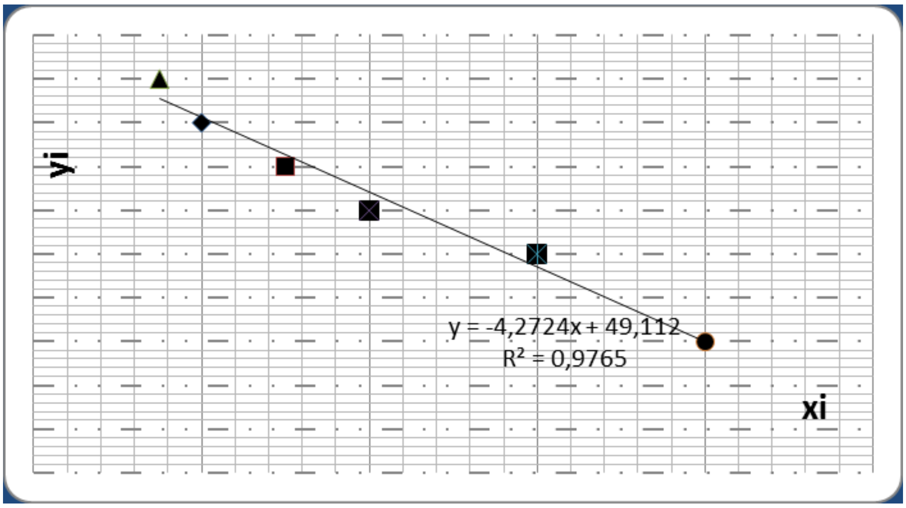
$\bar{x} = \frac{\sum x_i}{n} = \frac{24.5}{6} = 4$ و $\bar{y} = \frac{\sum y_i}{n} = \frac{190}{6} = 31.66$ وبالتالي يكون

$(\bar{x})^2 = 16$ و $\bar{x}.\bar{y} = 126.64$

ـ انجاز الجدول المساعد: وبناء على نتائج الجدول يمكن حساب المطلوب

x_i	2	3	1.5	4	6	8	24.5
${x_i}^2$	4	9	2.25	16	36	64	131.25
y_i	40	35	45	30	25	15	190
$x_i y_i$	80	105	67.5	120	150	120	642.5

ـ رسم سحابة النقاط المتحصل عليها من خلال الأزواج (x_i, y_i) مع خط الانحدار



تلاحظ من خلال الشكل أن العلاقة بين المتغيرين هي علاقة عكسية، أي كلما كان سعر الغرفة (المؤثر) منخفضا كان الاقبال أكثر، والعكس صحيحا.

ـ حساب الثوابت $(a \quad , \quad b)$

$$b = \frac{x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{642.5 - 6(126.64)}{131.25 - 6(16)} = \frac{-117.34}{35.25} = -3.32$$

$$a = \bar{y} - b\bar{x} = 31.66 - (-3.32)(4) = 44.94$$

ـ استنتاج المعادلة: $y_i = 44.94 - 3.32x_i$

ـ عدد السواح إذا كان سعر الغرفة 3.5 وحدة نقدية: وذلك بالتعويض في المعادلة التي توصلنا إليها نجد

$$y_{3.5} = 44.94 - 3.32(3.5) = 33.32$$

أي أن عدد السياح في حدود 33 سائح.

06 ـ ملاحظات هامة

من خلال الشكل الذي مر بنا تلاحظ معي أن هناك قيم قريبة من الخط المستقيم، وبعضها بعيدة عنه. كان الأولى أن تكون تلك النقاط كلها على استقامة واحة، أي على الخط، ولما كانت بعيدة عنه جهة اليمين أو جهة اليسار

معناه هناك أخطاء، أي انحراف القيم عن خط الانحدار ولذلك فإن شكل الانتشار النموذجي عادة ما يأخذ المعادلة التالية $y_i = a + bx_i + e_i$ حيث e_i تمثل تلك الاخطاء (المتغيرات العشوائية).

وللحصول على أحسن خط رياضيا يجب أن يكون مجموع الأخطاء مساويا للصفر، لذلك نستعمل المربعات الصغرى للتقليص إلى ابعد حد من تلك الفرو قات.

ـ تلاحظ من خلال المثال السابق أننا استعملنا متغير مستقل واحد، وهذا يعني أننا نتعامل مع نموذج خطي بسيط (تحليل الانحدار البسيط)، أما إذا كانت العلاقة بين متغير تا بع واحد وعدة متغيرات مستقلة، نكون أمام نموذج خطي متعدد (تحليل الانحدار المتعدد).

ـ إذا كانت العلاقة خطية في هذا المثال، يمكن أن تكون غير خطية حين نتعامل مع معطيات أخرى

ـ من خلا ل معادلة الانحدار التي تحصلنا عليها استطعنا أن نتنبأ بعدد الوافدين إلى الفندق عند سعر معين للغرفة، هذا يعني أن تحليل الانحدار هو أحد الأساليب الاحصائية الأساسية التي تستعمل في التنبؤ بسلوك الظواهر الاقتصادية.

ـ المفهوم الاقتصادي لمعامل الانحدار (b) يكمن في كونه مقياس كمي، قيمته تزيد أو تقل عن الواحد، كما أنه من خلال الاشارة التي يحملها تعرف نوع العلاقة بين المتغيرين (طر دية أم عكسية) وفي مثالنا السابق يتأكد بأن العلاقة عكسية بين المتغيرين بسبب الاشارة السالبة لمعامل الانحدار. كما يشترط وجود علاقة دالية بين المتغيرين محل الدراسة.

ـ لكل علاقة انحدارية علاقة ارتباطية، ولكن يختلفان في كيفية الحساب والتحليل.

07 ـ معامل الارتباط

يلجأ الباحث إلى حساب معامل الارتباط بين متغيرين أو أكثر حينما يتعذر عليه معرفة الصفة محل الدراسة المستقلة (المؤثرة) من الصفة الأخرى المتأثرة، نظرا للتداخل الموجود بينهما. وعلى هذا الأساس يفترض مرة أن x_i هو المستقل ويقوم بحساب معامل الانحدار(b) كما رأينا في معادلة خط الانحدار، ومرة أخرى يفترض أن y_i هو المتغير المستقل، ويقوم بحساب معامل الانحدار ($\acute{b}$).

ـ يرمز إلى معامل الارتباط بالرمز (r)، حيث يكون بسيطا إذا كانت الدراسة تتعلق بمتغيرين فقط، أما إذا كانت تتعلق بين ظاهرة وعدة ظواهر أخرى يسمى بالارتباط المتعدد.

08 ـ حساب معامل الارتباط

إذا كنا أما تطبيق طرق بيرسون و سبيرمان، ننبه الباحث إلى أن معامل الارتباط لبيرسون لا يصلح إلا في حالة البيانات الكمية فقط، أما معامل ارتباط سبيرمان ، فهو يصلح استعماله للبيانات الكمية والبيانات الوصفية بصفة أخص.

يستنتج من العلاقة $\sqrt{b.\acute{b}}$ = $r_{(x,y)}$ نعوض بما يساويها ونختار علاقة.

ـ علاقة التعريف

$$r_{(x,y)} = \sqrt{\frac{\sum x_i y_i - n\overline{xy}}{\sum x_i^2 - n\bar{x}^2} \cdot \frac{\sum x_i y_i - n\overline{xy}}{\sum y_i^2 - n\bar{y}^2}}$$

أي أن $(\acute{b})$ تعبر عن معامل الانحدار للعلاقة $x_i = \acute{a} + \acute{b}y_i + e_i(y_i)$ مما يدل على افتراضنا بأن y_i هو المتغير المستقل والآخر(x_i) متغير تابع .

ـ العلاقة الموسعة

$$r(x,y) = \frac{\sum x_i y_i - n\overline{xy}}{\sqrt{\sum x_i^2 - n\,\bar{x}^2 . \sum y_i^2 - n\bar{y}^2}}$$

لو تأملت إلى ما تحت الجذر في العلاقتين لوجدته عبارة عن الانحراف القياسي الذي درسناه في مقاييس التشتت. وأن صورة الكسر في العلاقة الثانية تعبر عن التزاوج بين المتغيرين، تسمى بتغاير $(x\ ,\ y)$ لذلك يمكن التعويض بما يساويهما فنجد.

$$r(x,y) = \frac{cov(x,y)}{sd_x . sd_y} = \frac{\text{تغاير}\ (x,y)}{sd_x . sd_y}$$

$$\text{مع} \quad -1 \leq r_{x,y} \leq 1$$

مثال: الجدول التالي يبين عدد الطائرات المستعملة (x_i) والربح المحقق (y_i) لخمس شركات طيران

x_i	10	20	60	100	210
y_i	10	10	30	40	110

المطلوب: معرفة قوة الارتباط بين المتغيرين محل الدراسة.

حل المثال: لتحقيق غاية الحل نستعين بالجدول التالي

x_i	10	20	60	100	210	**400**

y_i	10	10	30	40	110	**200**
$x_i y_i$	100	200	1800	4000	23100	**29200**
${x_i}^2$	100	400	3600	10000	44100	**58200**
${y_i}^2$	100	100	900	1600	12100	**14800**

$\bar{y} = \frac{\sum y_i}{n} = \frac{200}{5} = 40$

$\sum x_i^2 = 58200$ و $\frac{\sum x_i}{n} = \frac{400}{5} = 80$ و

من خلال هذه المعطيات يمكن حساب $\bar{y}^2 = 1600$ و $\bar{x}^2 = 6400$ و $\sum y_i^2 = 14800$

معامل الارتباط وفق القواعد التي مرت بنا.

1 ـ معادلة خط الانحدار $b = \frac{\sum x_i \quad _i - n\overline{xy}}{\sum x_i^2 - n\bar{x}^2} = \frac{13200}{26200} = 0.50$ ومنه $a = \bar{y} - b\bar{x} = 0$

وعليه تكون المعادلة: $y_i = 0.5x_i$

2 ـ لحساب معامل الارتباط علينا افتراض أن y هو المتغير المستقل، وبالتالي

$$\acute{b} = \frac{\sum x_i y_i - n\overline{xy}}{\sum y_i^2 - n\bar{y}^2} = \frac{13200}{6800} = 1.94$$

ومنه $\acute{a} = \bar{x} - \acute{b}\bar{y} = 2.4$

ومنه $r(x,y) = \sqrt{b\acute{b}} \quad = \quad \sqrt{0.97} = 0.98$

وهو كما تلاحظ معامل الارتباط قريب من الواحد الصحيح الموجب، مما يدل على ارتباط قوي جدا وطردي بين المتغيرين محل الدراسة، بمعنى كلما زاد عدد الطائرات زادت الأرباح المحققة.

3 ـ حساب الربح المقدر حينما يصل عدد الطائرات إلى 250 طائرة. نعوض في معادلة خط الانحدار الأولى نحصل على الربح المتوقع. $y_{250} = 0.5(250) = 125$

مثال آخر

من أجل دراسة العلاقة بين سلسلة إحصائية X وسلسلة إحصائية Y نقدم المعطيات التالية.

$\sum_{i=1}^{i=10} y_i^2 = 430$ و $\sum_{i=1}^{i=10} x_i^2 = 565$ و $\sum_{i=1}^{i=10} y_i$ 64= و $\sum_{i=1}^{i=10} x_i = 73$

$\sum_{i=1}^{i=10} x_i y_i = 465$

والمطلوب هو: حساب الوسط الحسابي والتباين لكلا المتغيرين ثم حساب تغاير (x,y)

ـ حساب معامل الارتباط الخطي.

ـ نستعمل قيمتين إضافيتين مع تقديم المعطيات التالية.

$\sum_{i=11}^{i=12} y_i^2 = 148$ و $\sum_{i=11}^{i=12} x_i^2 = 260$ و $\sum_{i=11}^{i=12} y_i = 14$ و $\sum_{i=11}^{i=12} x_i = 18$

$\sum_{i=11}^{i=12} x_i y_{i=196}$

المطلوب: حساب معامل الارتباط الخطي الجديد.

حل المثال

ـ متوسط السلسلتين هو: $\bar{x} = \frac{1}{10}\sum_{i=1}^{i=10} x_i = \frac{1}{10}.73 = 7.3$

$\bar{y} = \frac{1}{10}\sum_{i=1}^{i=10} y_i = \frac{1}{10}.64 = 6.4$

ـ أما تباين السلسلتين فهو كالآتي.

$$s_x^2 = \frac{1}{10}\sum_{i=1}^{i=10} x_i^2 - \bar{x}^2 = \frac{1}{10}.565 - (7.3)^2 = 56.5 - 53.29 = 3.21$$

$$s_y^2 = \frac{1}{10}\sum_{i=1}^{i=10} y_i^2 - \bar{y}^2 = \frac{1}{10}430 - (6.4)^2 = 43 - 40.96 = 2.04$$

ـ حساب تغاير (x,y)

$$s_{x,y} = \frac{1}{10}\sum_{i=1}^{i=10} x_i y_i - \overline{xy} = \frac{1}{10}.465 - (7.3)(6.4) = 46.5 - 46.72 = -0.22$$

ـ معامل الارتباط الخطي للسلسلتين (x,y)

$$r(x,y) = \frac{s_{xy}}{s_x s_y} = \frac{-0.22}{\sqrt{3.21}.\sqrt{2.04}} = \frac{-0.22}{2.56} = -0.09$$

ـ المعطيات الاضافية تعالج على النحو التالي

$\sum_{i=1}^{i=10} y_i + \sum_{i=11}^{i=12} y_i = \sum_{i=1}^{i=12} y_i = 78$ و $\sum_{i=1}^{i=10} x_i + \sum_{i=11}^{i=12} x_i = \sum_{i=1}^{i=12} x_i = 91$

$$\sum_{i=1}^{i=10} x_i^2 + \sum_{i=11}^{i=12} x_i^2 = \sum_{i=1}^{i=12} x_i^2 = 825$$

$$\sum_{i=1}^{i=10} y_i^2 + \sum_{i=11}^{i=12} y_i^2 = \sum_{i=1}^{i=12} y_i^2 = 578$$

$$\sum_{i=1}^{i=10} x_i y_i + \sum_{i=11}^{i=12} x_i y_i = \sum_{i=1}^{i=12} x_i y_i = 661$$

وعلى أساس هذه النتائج تكون قيم متوسط السلاسل الاحصائية الجديدة على النحو التالي

$$\bar{y} = \frac{1}{12}\sum_{i=1}^{i=12} y_i = \frac{1}{12}78 = 6.5 \text{ و } \bar{x} = \frac{1}{12}\sum_{i=1}^{i=12} x_i = \frac{1}{12}91 = 7.58$$

ـ وبالتالي تباين السلاسل الاحصائية الجديد يكون كالتالي

$$s_x^2 = \frac{1}{12}\sum_{i=1}^{i=12} x_i^2 - \bar{x}^2 = \frac{1}{12}825 - (7.58)^2 = 68.75 - 57.46 = 11.29$$

$$s_y^2 = \frac{1}{12}\sum_{i=1}^{i=12} y_i^2 - \bar{y}^2 = \frac{1}{12}578 - (6.5)^2 = 48.17 - 42.25 = 5.92$$

ـ ونفس الشيء بالنسبة لتغاير (x, y)

$$s_{xy} = \frac{1}{12}\sum_{i=1}^{i=12} x_i y_i - \overline{xy} = \frac{1}{12}661 - (7.58)(6.5) = 55.08 - 49.27 = 5.81$$

ـ وبالتالي معامل الارتباط الخطي الجديد ما بين السلسلتين الاحصائيتين (X) و(Y) محل الدراسة يكون

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{5.81}{\sqrt{11.29}\sqrt{5.92}} = \frac{5.81}{8.16} = 0.71$$

مثال آخر

الجدول الآتي يبين قيم إحصائية تتعلق بالسعر(Y) والكمية المتوفرة (X) من منتج في أحد الأسواق، حيث يعطى(Y) بالدينار الجزائري و(X) عدد المنتج من كمية السلعة.

x_i	02	04	06	12	15	20	22	24	28	30
y_i	98	90	84	72	64	62	55	54	45	40

المطلوب ما يلي:

1 ـ العرض البياني لسحابة القيم (النقاط)

2 ـ حساب الثوابت (a , b) لمعادلة خط الانحدار $y_i = a + bx_i$ التي تمثل الانحدار والقيمة المتوسطة.

3 ـ حساب معامل الارتباط الخطي.

4 ـ حساب الاخطاء (e_i) لكل قيم (x_i) للمتغير (X) المرافقة لتقديرات الثوابت (a , b) ثم أحسب تباين الأخطاء التي سجلتها (s_e^2) .

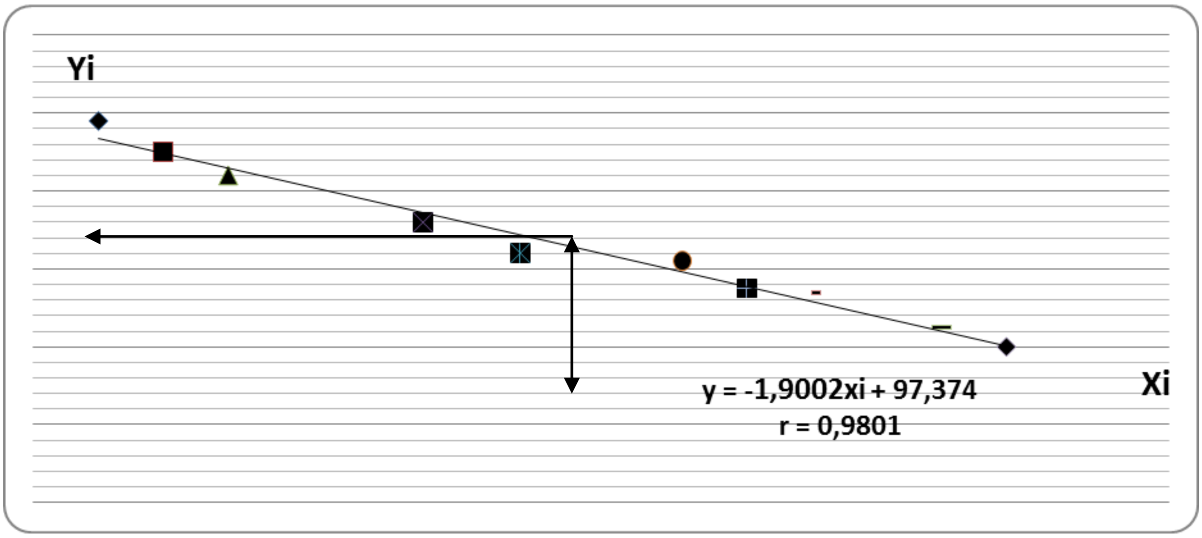
5 ـ ما هو السعر المنتظر إذا توفرت الكمية (05) ثم الكمية (26).

حل المثال

لتسهيل حل المثال يقتضي منا الاستعانة بالجدول الموالي الذي يبين مختلف الحسابات المساعدة للحل.

الرقم	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
01	02	98	04	9604	196
02	04	90	16	8100	360
03	06	84	36	7056	504
04	12	72	144	5184	864
05	15	64	225	4096	960
06	20	62	400	3844	1240
07	22	55	484	3025	1210
08	24	54	576	2916	1296
09	28	45	784	2025	1260
10	30	40	900	1600	1200

1ـ العرض البياني لسحابة النقاط (القيم) وخط الانحدار مع الوسط الحسابي للمتغيرين (x , y)



من خلال الشكل يتضح أن العلاقة بين المتغيرين محل الدراسة، هي علاقة عكسية، بمعنى كلما زادت الكمية من المنتج المعروض أدت إلى انخفاض السعر. وبما أن نقاط القيم قريبة جدا من الخط المستقيم دليل على قوة الارتباط.

2 ــ حساب (a , b)، قبل حساب الثوابت علينا حساب المتوسط الحسابي لكلا المتغيرين.

$\bar{x} = 16.3$ ، وكذلك $s_x^2 = 91.21$ ومنه $s_x = 9.55$

$\bar{y} = 66.4$ ، وكذلك $s_y^2 = 336.04$ ومنه $s_y = 18.33$ ومنه $_{xy} = -173.32$

$b = \frac{s_{xy}}{s_x^2} = \frac{-173.32}{91.21} = -1.9002$ ومنه $a = \bar{\ } - b\bar{x} = 97.37$

3 ــ معامل الارتباط الخطي

$$r = \frac{s_{xy}}{s_x s_y} = \frac{-173.32}{(9.55)(18.33)} = -0.99$$

كما تلاحظ هو قريب من (1) بالقيمة المطلقة ويتوافق تماما مع الشكل البياني السابق لقرب سحابة النقاط من المستقيم، كما تدل الإشارة السالبة على العلاقة العكسية بين المتغيرين محل الدراسة.

4 ـ لحساب الأخطاء، يجب حساب النقاط التقديرية لمنحنى المربعات الصغرى . وعليه لابد من حساب كل زوج من خلال معادلة خط الانحدار $\acute{y}_i = -1.90.x_i + 97.37$

لنحسب مثلا من أجل i=1 سوف تجد $\acute{y}_1 = -1.90.(2) + 97.37 = 93.57$ بعد هذا نقوم بحساب الفرق بين القيمة المشاهدة في السلسلة والقيمة التي تحصلنا عليها (93.57) كما يلي:

$e_1 = y_1 - \acute{y}_1 = 98 - 93.57 = 4.43$ وبهذه الطريقة نكمل باقي القيم كما هو موضح في الجدول أسفله.

الرقم	01	02	03	04	05	06	07	08	09	10	Σ
x_i	02	04	06	12	15	20	22	24	28	30	163
y_i	98	90	84	72	64	62	55	54	45	40	664
$\bar{y}_i$	93.57	89.77	85.97	74.57	68.87	59.37	55.57	51.77	44.17	40.37	664
e_i	4.43	0.23	-1.97	-2.57	-4.87	2.63	-0.57	2.23	0.83	-0.37	0.00
${e_i}^2$	19.60	0.05	3.89	6.61	23.72	6.92	0.32	4.98	0.69	0.13	66.92

تأكد لنا بأن مجموع الأخطاء (e_i) مساوية للصفر، أي أن $\sum_{i=1}^{i=10} \quad_i = 0$

من جهة أخرى نجد أن متوسط الأخطاء يساوي الصفر، أي أن $\bar{e} = \frac{1}{10} \sum_{i=1}^{i=10} e_i = 0$

وبالتعريف، فإن تباين الأخطاء هو:

$$s_e^2 = \frac{1}{n} \sum_{i=1}^{i=n} (e_i - \bar{e})^2$$

$$s_e^2 = \frac{1}{n}\sum_{i=1}^{i=n} e_i^2 = \frac{1}{10}\sum_{i=1}^{i=10} e_i^2 = \frac{1}{10} 66.92 = 6.692$$

الآن علينا التأكد من أن المتوسط الحسابي للمتغير التابع هو نفسه المتغير الحسابي المقدر، أي هل $(\bar{y} = \bar{\acute{y}})$.

$$\bar{y} = \frac{1}{10}\sum_{i=1}^{i=10} y_i = \frac{664}{10} = 66.4$$

$$\bar{\acute{y}} = \frac{1}{10}\sum_{i=1}^{i=10} \acute{y} = \frac{664}{10} = 66.4$$

تلاحظ معي لهما نفس النتيجة.

5 ـ للجوب على هذا السؤال بكل بساطة نعوض في معادلة خط الانحدار

القيمة (x_i) بالكمية (5) فتكون النتيجة كما يلي:

$$\acute{y}_i = -1.90.(5) + 97.37 = 87.87$$

وبنفس الطريقة بالنسبة للكمية $x_i = 26$

$$\acute{y}_{26} = -1.90.(26) + 97.37 = 47.97$$

أي أن الكمية 26 تقتضي 47.97 دينار جزائري.

09 ـ معامل الارتباط الرتبي

يدرس العلاقة بين (x, y) في حالات محددة مثل: عدم معرفة توزيع المتغيرين قيد الدراسة ، بيانات كيفية قابلة للترتيب ، وحالة البيانات الكمية. وعلى هذا الأساس يمكن حساب معامل الارتباط (سيبرمان) وفق الصيغة التالية

$$r_s = 1 - \frac{6\sum_{i=1}^{i=n} {d_i}^2}{n(n^2 - 1)}$$

مثال

تبين معطيات الجدول التالي، الدرجة التقديرية لستة (06) طلبة في امتحاني الإحصاء والرياضيات ، حيث تمثل (y_i) درجة الاحصاء و (x_i) درجة الرياضيات.

رقم الطالب	01	02	03	04	05	06
الدرجة التقديرية للرياضيات x_i	ضعيف	ممتاز	جيد	ضعيف جدا	مقبول	جيد جدا
الدرجة التقديرية للإحصاء y_i	مقبول	جيد جدا	جيد	ضعيف	ضعيف جدا	ممتاز

خطوات الحل :

ـ أولا نقوم بترتيب الدرجات التقديرية لكل مقياس ترتيبا تصاعديا. حسب الجدول التالي

ترتيب درجات الرياضيات x_i		**ترتيب درجات الاحصاء y_i**	
الدرجات	**الترتيب**	**الدرجات**	**الترتيب**
ممتاز	01	**ممتاز**	01
جيد جدا	02	**جيد جدا**	02
جيد	03	**جيد**	03
مقبول	04	**مقبول**	04
ضعيف	05	**ضعيف**	05
ضعيف جدا	06	**ضعيف جدا**	06

ـ نقوم بحساب مربع الفروق وفق الجدول أسفله.

y_i	x_i	y_i ترتيب	x_i ترتيب	d_i	d_i^2
مقبول	ضعيف	04	05	01	01
جيد جدا	ممتاز	02	01	01-	01
جيد	جيد	03	03	00	00
ضعيف	ضعيف جدا	05	06	01	01
ضعيف جدا	مقبول	06	04	02-	04
ممتاز	جيد جدا	01	02	01	01
المجموع					08

ـ بعد حساب مربع الفروق ، نقوم بحساب معامل الارتباط الرتبي

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6(8)}{6(36-1)} = 0.772$$

من خلال النتيجة ندرك بأن العلاقة بين المقياسين مقبولة وطردية .

10 ـ معامل الاقتران (r_s)

رأينا فيما سبق أن معامل الارتباط (r_{xy}) الخطي البسيط لصاحبه "بيرسون" يعطينا مستوى قوة الارتباط بين متغيرين (ظاهرتين محل الدراسة) في حالة البيانات الكمية.

أما معامل الارتباط الرتبي (r_s) لصاحبه "سبيرمان" فهو يقيس لنا مستوى قوة الارتباط بين ظاهرتين تتميزان بصفات لها خاصية الترتيب.

بينما معامل الاقتران (r_c) فإننا نستعمله في حالة وجود ظاهرتين محل الدراسة، كل منهما تتصف بصفتين (مجال التغير) فقط لا غير. مثل (طويل، قصير)، (متعلم، غ متعلم)، (خفيف، ثقيل)، (يدخن، لا يدخن)...الخ .وعلى هذا الأساس نفرض من البداية أن:

y_i \ x_i	x_1	x_2

ـ المتغير x_i له صفتين فقط هما : (x_1) و (x_2)

ـ المتغير $_i$ له صفتين فقط هما : (y_1) و (y_2)

y_1	A	B
y_2	C	D

بناء على هذا الأساس ، يمكن حساب معامل الاقتران وفق القاعدة التالية :

$$r_c = \frac{(AD)-(CB)}{(AD)+(CB)}$$

مثال: الجدول التالي يبين معطيات تتعلق بالأثر المتبادل بين العمل والزواج، لعينة تتكون من 36 شخص

العمل / الزواج	يعمل	لا يعمل
متزوج	15	06
غير متزوج	07	08

$$r_c = \frac{(15\times 8)-(7\times 6)}{(15\times 8)+(7\times 6)} = \frac{120-42}{120+42} = \frac{78}{162} = 0.48$$

هذا يعني أن معامل الاقتران يؤكد بأن هناك ارتباط طردي وضعيف بين العمل والزواج.

11 ـ معامل التوافق (r_a)

يستعمل في حالة بيانات وصفية للظاهرتين معا ، أو لأحدهما ، مع وجود أكثر من صفتين (مجال التغير).

يحسب وفق العلاقة التالية:

$$r_a = \sqrt{\frac{B-1}{B}}$$

مثال: لتكن لدينا معطيات الجدول التالية التي تخص العلاقة بين المستوى التعليمي من جهة، ونوع العمل من جهة ثانية ،لعينة تتكون من اربعة وأربعين(44) عاملا لإحدى المؤسسات.

المستوى التعليمي / نوع العمل	ابتدائي	ثانوي	جامعي	Σ

المطلوب: حساب معامل التوافق

غير ماهر	11	05	02	18
نصف ماهر	06	02	02	10
ماهر	02	05	09	16
$\sum$	19	12	13	44

حل المثال: لاحظ معي، يتكون الجدول من تسعة (09) أزواج، مثل الزوج $(x_2\ ,\ y_2)$ هو (2) وهكذا...

وعلى هذا الأساس نبدأ بحساب (B) كما يلي

$$B=\frac{11^2}{(18).(19)}+\frac{5^2}{18.12}+\frac{2^2}{18.13}+\frac{6^2}{10.19}+\frac{2^2}{10.12}+\frac{2^2}{10.13}+\frac{2^2}{16.19}+\frac{5^2}{16.12}+\frac{9^2}{16.13}=1.29$$

ومنه $r_a=\sqrt{\frac{B-1}{B}}=\sqrt{\frac{1.29-1}{1.29}}=\sqrt{\frac{0.29}{1.29}}=\sqrt{0.224}=0.47$

تلاحظ من خلال النتائج ، أن العلاقة بين المتغيرين محل الدراسة هي علاقة طردية وضعيفة.

12 ـ صلاحية النموذج

تتوقف دراسة صلاحية النموذج على: حساب معامل الارتباط (r_{xy}) و حساب الانحرافات القياسية المقدرة (الأخطاء).

وعلى هذا الأساس تكون معادلة خط الانحدار المقدرة

$$\underset{(\text{المقدرة})}{\hat{y}} = \underset{(SD_a)}{a} + \underset{(SD_b)\ (r_{xy})}{bx_i}$$

حيث

$$SD_b=\frac{SD_e}{\sqrt{\sum(x_i-\bar{x})^2}} \quad \text{مع} \quad SD_e=\sqrt{\frac{\sum(\hat{y}-y_i)}{n-2}}$$

و

$$SD_a=\bar{x}SD_b+\frac{SD_e}{\sqrt{n-1}}$$

مثال: نريد دراسة صلاحية النموذج لعينة من الطلبة المتعلق بمتغيري معدلي السداسي الأول والسداسي الثاني ، بناء على المعطيات الواردة في الجدول.

الطلبة	01	02	03	04	05	06	07	08	09	10
x_i م س1	06	05	08	08	07	06	10	04	09	07
y_i م س2	08	07	07	10	05	08	10	06	08	06

حل المثال: للوصول إلى الشكل النهائي للنموذج، علينا القيام بثلاثة خطوات على الأقل، وهي معادلتي خط الانحدار لكل من المتغيرين $(x\ ,y)$ ثم حساب قوة (r). لتحقيق ذلك ، نستعين بالجدول التالي.

x_i	y_i	$x_i y_i$	x_i^2	y_i^2	$\hat{y}_i$	e_i	e_i^2	$(x_i-\bar{x})^2$
06	08	48	36	64	07	-1	1	1
05	07	35	25	49	6.5	-0.5	0.25	4
08	07	56	64	49	08	1	1	1
08	10	80	64	100	08	-2	4	1
07	05	35	49	25	7.5	2.5	6.25	0
06	08	48	36	64	07	-1	1	1
10	10	100	100	100	09	-1	1	9
04	06	24	16	36	06	0	0	9
09	08	72	81	64	8.5	0.5	0.25	4
07	06	42	49	36	7.5	1.5	2.25	0
70	75	540	520	587			17	30

حيث : $e_i = (\hat{y}_i - y_i)$ ، أما $\hat{y}_i$ المقدرة فإننا نحصل على قيمها من خلال التعويض في معادلة خط الانحدار $y_i = a + bx_i$ بعد الحصول على المعلمات $(a\ ,b)$ وتشكيل المعادلة كما سيأتي لاحقا.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{70}{10} = 07$$

و $\bar{y} = \frac{\sum y_i}{n} = \frac{75}{10} = 7.5$ ومنه معامل الانحدار

$$b = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = 0.5 \quad \text{و} \quad a = \bar{x} - b\bar{y} = 4$$

وبالتالي يمكن الحصول على معادلة خط الانحدار كما يلي.

$$y_i = 4 + 0.5x_i$$

وهي تبين أن المتغير المستقل هو $_i$

بما أننا نريد حساب قوة الارتباط ، علينا حساب معلمات المعادلة الثانية ، على اساس أن y_i هو المتغير المستقل ،

بمعنى آخر تشكيل معادلة خط الانحدار $x_i = \acute{a} + \acute{b}y_i$

وعليه، نقوم بحساب معلمات المعادلة على النحو التالي.

$$\acute{b} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum y_i^2 - n\bar{y}^2} = 0.612$$

$$\acute{a} = \bar{x} - \acute{b}\bar{y} = 2.41$$

و بالتالي تشكيل المعادلة: $(y) = x_i = 2.41 + 0.612y_i$

من خلال هذه النتائج، يمكن حساب قوة الارتباط على اعتبار أن $r_{(,y)} = \sqrt{b\acute{b}}$

$r_{(x,y)} = \sqrt{(0.5)(0.612)} = \sqrt{0.306} = 0.55$ ،

وهي نتيجة تعبر عن وجود ارتباط بين المتغيرين محل الدراسة ، يتصف كونه ضعيف ، كما أن العلاقة بين المتغيرين طردية .

ـ تحديد الشكل النهائي للنموذج

1 ـ تقدير: $f(x) = \hat{y}_i$ حيث وجدنا معادلة خط الانحدار $y_i = 4 + 0.5x_i$ ، بناء على هذا الأساس نقوم بحساب الفروق (الأخطاء).

على سبيل المثال: $\hat{y}_1 = 4 + 0.5(6) = 7$ وأن $\hat{y}_2 = 4 + 0.5(5) = 6.5$ و $\hat{y}_3 = 8$ و $\hat{y}_4 = 8$ ... أخيرا $\hat{y}_{10} = 7.5$. وهي كما ترى نعوض في (x_i) بما يساويه من القيم المشاهدة (محل الدراسة)، وبهذه الكيفية أكملنا الجدول أعلاه.

2 ـ حساب الانحرافات القياسية المقدرة

$$SD_b = \frac{SD_e}{\sqrt{(x_i-\bar{x})^2}} = \frac{1.46}{\sqrt{30}} = \frac{1.46}{5.47} = 0.266 \text{ و } SD_e = \sqrt{\frac{\sum(\hat{y}_i-y_i)^2}{n-1}} = \sqrt{\frac{17}{10-2}} = \sqrt{\frac{17}{8}} = 1.46$$

$$SD_a = \bar{x}SD_b + \frac{SD_e}{\sqrt{n}} = 7(0.266) + \frac{1.46}{\sqrt{10}} = 1.92$$

وهكذا يمكن التعبير عن الشكل النهائي للنموذج كما يلي:

$$y_i = 4 + 0.5\,x_i$$

$$\hat{y}_i = SD_a + SD_b.(r_{x,y})$$

$$\hat{y}_i = 1.92 + 0.266 \quad (r = 0.55)$$

نلاحظ أن الانحرافات القياسية للمعلمات كبيرة نسبيا مقارنة مع قيم المعلمات (a, b) وبالتالي يعتبر تفسير المتغير المستقل x_i للمتغير y_i ضعيفا نسبيا.

13 ـ تمارين محلولة

التمرين الأول: مؤسسة إنتاجية تريد معرفة تقديرات حجم المبيعات (X) للسنتين 2016 و2017.لأجل ذلك عليها دراسة حجم البيع للفترة 2006 ـ 2015.المعطيات في الجدول أسفله.

السنة	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
التاريخ t	01	02	03	04	05	06	07	08	09	10
البيع المنجز x_t	3400	4000	3200	3700	3600	3100	3300	3500	4200	4100

1ـ أحسب المعلمتين (a, b) لمعادلة الانحدار $x = a + bT$ حيث t تمثل السنة ما قبل (2005) وتعطى على النحو التالي.

$\sum_{t=1}^{t=10} x_t = 36100$ ، $\sum_{t=1}^{t=10} tx_i = 202300$ ، $\sum_{t=1}^{t=10} t^2 = 385$ ، $\sum_{t=1}^{t=10} t = 55$

2ـ ما هو الحجم المقدر لمبيعات المنتج (x) للسنتين 2016 و 2017. أعرض بيانيا سحابة النقاط، خط الانحدار، النقاط المقدرة.

حل التمرين الأول

ـ متوسط البيع خلال الفترة هو

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{i=10} x_i = \frac{36100}{10} = 3610 \text{ (مباع)}$$

ـ متوسط تغير الفترة (T) هو

$$\bar{T} = \frac{1}{10}\sum_{t=1}^{t=10} t = \frac{55}{10} = 5.5 \text{ (سنة)}$$

ـ تباين المتغير (T) هو

$$S_T^2 = \frac{1}{10}\sum t^2 - \bar{T}^2 = \frac{385}{10} - (5.5)^2 = 8.25 \quad \text{(سنة)}^2$$

ـ التغاير بين المتغيرين (X, T) هو

$$S_{XT} = \frac{1}{10}\sum_{t=1}^{t=10} t.x_t - \bar{X}.\bar{T} = \frac{202300}{10} - (5.5 \times 3610) = 375 \text{ (بيع × سنة)}$$

الآن يمكن حساب المعلمتين (a , b) لخط الانحدار (X) بالنسبة ل(T)

$$a = \bar{X} - b\bar{T} = 3610 - 45.45(5.5) = 3360.02 \quad \text{و} \quad b = \frac{S_{X,T}}{S_T^2} = \frac{375}{8.25} = 45.45$$

ـ معادلة خط الانحدار هي :

$$x_i = 3360.02 + 45.45t$$

ـ تقديرات سنة 2016 معناه $t = 11$ وبالتالي

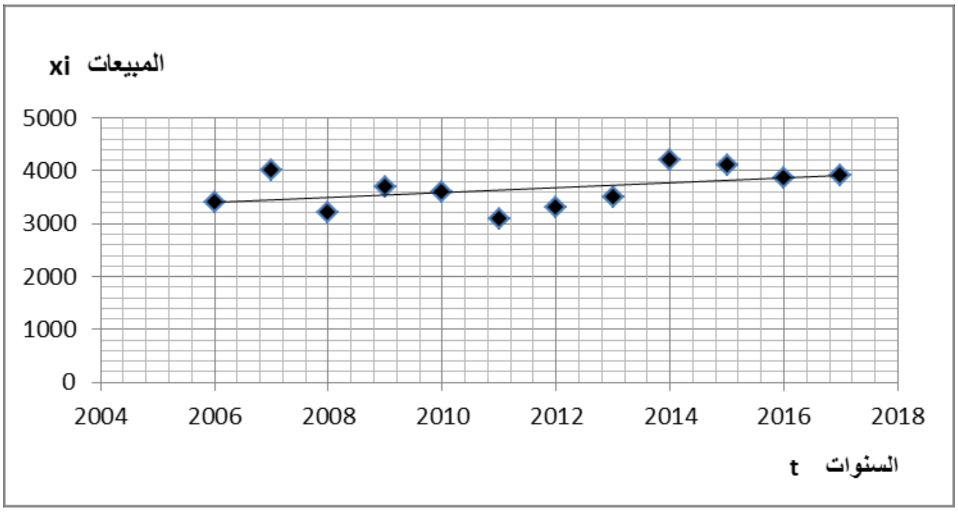
$$x_{11} = 3360.02 + 45.45.(11) \cong 3860 \text{ (مباع)}$$

ـ تقديرات سنة 2017 معناه $t = 12$ وبالتالي

$$x_{12} = 3360.02 + 45.45.(12) = 3905 \text{ (مباع)}$$

ـ رسم السحابة وخط الانحدار، بالإضافة إلى الكميات المقدرة للسنتين (2015 و 2016) كما هو مبين في الرسم البياني ادناه.

الرسم البياني: يبين تقدير المبيعات



التمرين الثاني ـ أجب: (نعم) أم (لا)

ـ سحابة نقاط على شكل كرة، هي مؤشر على وجود ارتباط خطي.

ـ العلاقة الآتية هي دائما صحيحة

$$s_e^2 = \frac{1}{n}\sum_{i=}^{i=n} (e_i - \bar{e})^2 = \frac{1}{n}\sum_{i=1}^{i=n} e_i^2$$

ـ العلاقة الآتية تعني أنها دوما صحيحة

$$s_e^2 = s_y^2\left(1 - r^2\right)$$

ـ معامل الارتباط الخطي هو عدد غير مجزأ

ـ مبدأ المربعات الصغرى يهدف إلى تقليص مجموع مربع الأخطاء

ـ خط (بيان) المربعات الصغرى يمر في بعض الحالات على القيمة الوسطية.

ـ حين يكون معامل الارتباط الخطي قريب من الصفر معناه وجود ارتباط خطي قوي

ـ يمكن حساب خط الانحدار y للمتغير المستقل x وخط الانحدار x للمتغير المستقل y

حل التمرين الثاني: الاجابة بنعم أم لا

الرقم	1	2	3	4	5	6	7	8
نعم		نعم	نعم		نعم			نعم
لا	لا			لا		لا	لا	

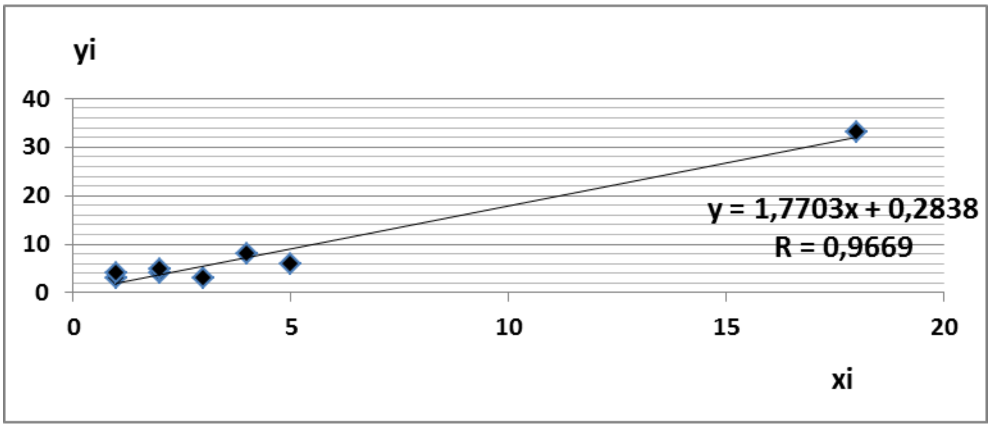
التمرين الثالث : البيانات التالية تمثل مردود الكميات المحصلة من الزيتون والمقدرة بالطن ، ويعبر عنها (y_i) الناجمة عن السماد المستخدم المقدر ب (كلغ) ويعبر عنه (xi)

x_i	01	02	05	03	02	01	04	18
y_i	03	04	06	03	05	04	08	33

إنجاز كل من : سحابة النقاط ، معادلة خط الانحدار بالنسبة ل (y_i) و (x_i) ، حساب معامل الارتباط (r) مع تفسير النتيجة ، حساب الأخطاء (e_i) لكل قيم (x_i) ، حساب الانحرافات القياسية المقدرة مع الشكل النهائي للنموذج (صلاحية النموذج) ثم أحسب ما هي الكمية المنتظر تحصيلها إذا ارتفع السماد المستخدم إلى 10 كلغ للشجرة الواحدة.

حل التمرين الثالث

ـ سحابة النقاط: يمكن رسمها من خلال قيم المتغيرين (x_i, y_i) كما هو موضح في الرسم البياني أسفله.



x_i	y_i	x_i^2	y_i^2	$x_i y_i$	$\hat{y}_i$	e_i	e_i^2	$(x_i - \bar{x})^2$

01	03	01	09	03	2.05	-0.95	0.90	12.25
02	04	04	16	08	3.82	-0.18	0.03	6.25
05	06	25	36	30	9.13	3.13	9.79	0.25
03	03	09	09	09	5.59	2.59	6.70	2.25
02	05	04	25	10	3.82	-1.18	1.39	6.25
01	04	01	16	04	2.05	-1.95	3.80	12.25
04	08	16	64	32	7.36	-0.64	0.40	0.25
18	33	324	1089	594	32.14	-0.86	0.73	182.25
36	66	384	1264	690			23.74	219.75

ـ معادلة خط الانحدار للمتغير المستقل (x_i) ، ومعادلة خط الانحدار للمتغير المستقل (y_i)، لأجل ذلك نستعين بالجدول التالي.

$\overline{xy} = (4.5).(8.25) = 37.13$ و $\bar{y} = \frac{\sum y_i}{n} = \frac{66}{8} = 8.25$ و $\bar{x} = \frac{\sum x_i}{n} = \frac{36}{8} = 4.5$

وبالتالي $\bar{y}^2 = 68$ و $\bar{x}^2 = 20.25$

$$b = \frac{\sum x_i y_i - n\overline{xy}}{\sum x_i^2 - n\bar{x}^2} = \frac{690 - 8(37.13)}{384 - 8(20.25)} = \frac{392.96}{222} = 1.77$$

$a = \bar{y} - b\bar{x} = 8.25 - 1.77(4.5) = 0.28$. بناء على هذا الأساس يمكن وضع معادلة خط الانحدار بالنسبة للمتغير (x_i) كما يلي :

$$f(x) = y_i = 0.28 + 1.77x_i$$

ـ بنفس الطريقة يمكن الحصول على معادلة خط الانحدار بالنسبة للمتغير(yi)

$$\acute{b} = \frac{\sum x_i y_i - n\overline{xy}}{\sum y_i^2 - n\bar{y}^2} = \frac{392.96}{1264 - 8(68)} = \frac{392.96}{720} = 0.55$$

ومنه $\acute{a} = \bar{y} - \acute{b}\bar{x} = 8.25 - 0.55(4.5) = 5.76$

وعليه، يمكن بناء معادلة خط الانحدار كما يلي.

$$f(y) = x_i = 5.76 + 0.55y_i$$

وبناء على النتائج السابقة، يمكن حساب معامل الارتباط $(r_{x,y})$ بإحدى الطريقتين حيث سبق أن تعرضنا لهما عند تطرقنا لمعامل الارتباط .

$$r_{(x,y)} = \sqrt{b.\acute{b}} = \sqrt{(1.77).(0.55)} = \sqrt{0.9735} = 0.98$$

ـ تفسير النتيجة : كما تلاحظ فإن قيمة معامل الارتباط بين المتغيرين محل الدراسة قوي جدا ، والعلاقة طردية (موجبة) .حيث يمكن استنتاجها من الرسم البياني الذي مر بنا ، مما يعني هناك علاقة عضوية بين زيادة السماد في التربة وارتفاع كمية محصول الزيتون.

ـ حساب الفروق (الأخطاء)

بما أن $e_i = (\hat{y} - y_i)$ هذا يعني البحث عن تقدير $\hat{y}_i$ من خلال معادلة خط الانحدار

$$y_i = 0.28 + 1.77x_i$$

وعلى هذا الأساس نقوم بالحساب على النحو التالي.

$\hat{y}_1 = 0.28 + 1.77(1) = 2.05$ و $\hat{y}_2 = 0.28 + 1.77(2) = 3.82$

و $\hat{y}_3 = 0.28 + 1.77(5) = 9.13$

و $\hat{y}_4 = 0.28 + 1.77(3) = 5.59$ ... وهكذا حتى نصل إلى

$\hat{y}_8 = 0.28 + 1.77(18) = 32.14$

أما باقي النتائج فهي موجودة بالجدول أعلاه.

ـ حساب الانحرافات القياسية المقدرة.

$$SD_e = \sqrt{\frac{\Sigma(\hat{y}_i - y_i)^2}{n-1}}$$

وبالتعويض بما يساوي المعادلة نجد

$$SD_e = \sqrt{\frac{23.74}{7}} = 1.84$$

$$SD_b = \frac{SD_e}{\sqrt{(x_i - \bar{x})^2}} = \frac{1.84}{\sqrt{219.75}} = \frac{1.84}{14.82} = 0.12$$

وكذلك $SD_a = \bar{x}S_b + \frac{SD_E}{\sqrt{n}}$ وبالتعويض

$$SD_a = 4.5(0.12) + \frac{1.84}{\sqrt{8}} = 0.54 + \frac{1.84}{2.82} = 1.19$$

ـ صلاحية النموذج:

لدينا معادلة خط الانحدار بالنسبة للمتغير المستقل (x_i):

$$f(x) = y_i \quad = \quad 0.28 \quad + \quad 1.77x_i$$

$$(\text{مقدرة}\ \hat{y}_i) \quad \underset{(1.19)}{SD_a} \quad + \quad \underset{(0.12)}{SD_b} \quad \underset{(0.98)}{(r_{x,y})}$$

يلاحظ أن قيم المعلمات المقدرة مقارنة بقيم معلمات معادلة خط الانحدار كبيرة نسبيا، مما يدل على عدم صلاحية النموذج. بمعنى تفسير المتغير المستقل للمتغير التابع ضعيفا.

التمرين الرابع

الجدول التالي يحتوي على معطيات تتعلق بمتغير أعمار الأزواج (xi) وأعمار الزوجات (yi) بالسنوات.

x_i	25	28	30	32	35	36	38	39	42	45
y_i	20	26	29	30	25	18	26	35	35	46

1 ـ ـ أحسب معادلة انحدار (y_i) بالنسبة ل (x_i)

2 أحسب معادلة انحدار (x_i) بالنسبة ل (y_i).

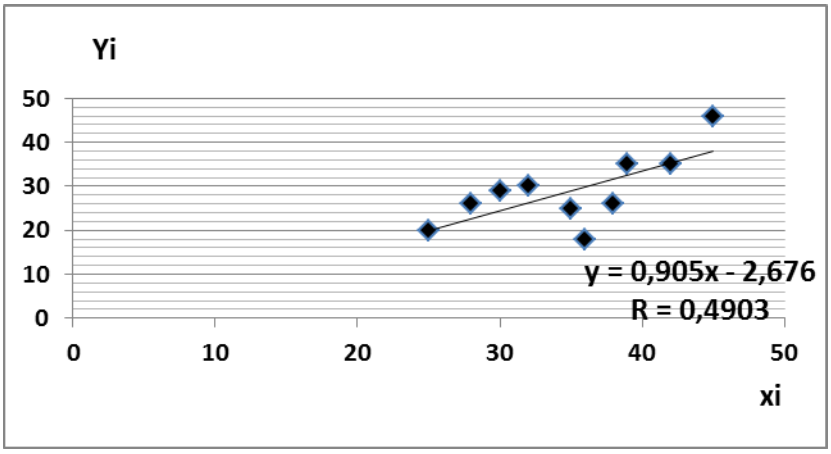
3 ـ أحسب العمر المناسب للأزواج (x_i) حينما يكون عمر زوجاتهم 25 سنة.

4 ـ أحسب العمر المناسب للزوجات (y_i) حينما يكون عمر أزواجهم 30 سنة.

5 ـ أحسب معامل الارتباط بين المتغيرين مع تفسير النتيجة.

حل التمرين

قبل الجواب عن الأسئلة الواردة ، ومن باب الإفادة نقوم برسم سحابة النقاط مع خط الانحدار من خلال تقنية الإكسيل.



من خلال الرسم البياني يتضح أن العلاقة بين المتغيرين طردية (موجبة).

$\sum xi$	$\sum yi$	$\sum x_i^2$	$\sum y_i^2$	$\sum x_i y_i$
350	290	358	598	324

ـ معادلة خط الانحدار (yi)

من المعلوم أن:

$$x_i - \bar{x} = r\frac{s_x}{s_y}(y_i - \bar{y})$$

من جهة ثانية:

$$b_{x,y} = \frac{\sum x_i y_i}{\sum y_i^2} \quad \text{لدينا} \quad ، \; r\frac{s_x}{s_y} = \frac{\sum xy}{\sum y^2}$$

$$b_{x,y} = \frac{324}{598} = 0.5572$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{290}{10} = 29$$ و

$$\bar{x} = \frac{\sum x_i}{n} = \frac{350}{10} = 35$$ و

بالتعويض في المعادلة $(y_i - \bar{y})$ $xi - \bar{x} = b_{x,y}$ معناه : $x_i - 35 = 0.5572(y_i - 29)$ مما يعني

$$x_i = 18.41 + 0.5572y_i$$

ومنه إذا كان $(yi) = 25$ فإن عمر الأزواج سيكون

$$x_{25} = 18.41 + 0.5572(25) = 32.7 \text{ (سنة)}$$

ـ بنفس الطريقة تحصل على معادلة خط الانحدار (xi)

لدينا المعادلة $y_i - \bar{y} = r\frac{y}{s_x}(x_i - \bar{x})$ ، ومن جهة أخرى لدينا

$$r\frac{s_y}{s_x} = \frac{\sum x_i y_i}{\sum x_i^2}$$

كما لدينا

$\bar{x} = \frac{\sum x_i}{n} = \frac{350}{10} = 35$ مع $b_{x,y} = \frac{324}{358} = 0.905$ وبالتالي $\sum x_i^2 = 358$ و $\sum xi\, yi = 324$

بالتعويض في المعادلة السابقة بما يساويها نجد

$$y_i - 29 = 0.905(x_i - 35)$$

ومنه

$y_i = -2.675 + 0.905(30) = 24.475$ نجد $x_i = 30$ ولأجل $y_i = -2.675 + 0.905x_i$

ـ أخيرا حساب معامل الارتباط : $r_{x,y} = 0.4903$. كما تلاحظ تعبر قيمة معامل الارتباط على وجود علاقة ضعيفة بين المتغيرين.

14 ـ تمارين غير محلولة

فترة العمل بالسنوات (x_i)	06	04	10	12	09	07	11

التمرين الأول: الجدول التالي يحتوي على معطيات تتعلق بالدخل المقابل للفترة التي قضتها في العمل عينة من عمال إحدى المؤسسات.

300	400	550	700	450	800	750	الدخل السنوي بألف دج (y_i)

المطلوب: أحسب الدخل السنوي لموظف قضى 15 سنة خدمة في وظيفته.

التمرين الثاني: البيانات الواردة في الجدول أدناه تخص التكاليف التي انفقت على صيانة آلة خلال عمرها الانتاجي محل الدراسة.

14	12	10	08	06	04	02	عمر الآلة بالسنوات (x_i)
1500	1200	900	700	400	250	200	تكاليف الصيانة بألف د ج (y_i)

المطلوب : أوجد كل من

ـ رسم سحابة النقاط، معادلة خط الانحدار، قوة معامل الارتباط، تفسير النتيجة

مقدمة

كل مقياس إ حصا ئي يظهر التغيرات التي تحد ث في متغير إحصائي أو عدة متغيرات ، خلال الزمان أو المكان أو خلال إحدى الخصائص ذات العلاقة ، يمكن أن نطلق عليه رقما قياسيا.

لذلك ، تعتبر الأرقام القياسية من المؤشرات الاحصائية التي يوظفها الباحث عند قياس تطور(تغير) سعر (سلعة أو خدمة) أو كمية أو قيمة وذلك بين فترتين زمنيتين مختلفتين ، تكون إحداهما أساس للقياس. أو بين مجموعتين أو مكانين.

قديما، كان المقصود بالأرقام القياسية، هو قياس التغير في الأسعار بين فترات زمنية متتالية، ومع تطور البحث العلمي، اتسع نطاق استعمالها، حيث أضحت تعني قياس التغير في الكميات والتغير في القيمة بين فترات زمنية متتالية.

وعلى هذا الأساس ، تعتبر الارقام القياسية في يومنا من أهم الآليات التي تستخدم في قياس تطور الأسعار أو قيمة العملة لبلد ــ ما ــ أو تطور مكونات الأسواق المالية محلية كانت أو عالمية ، كما تهتم بقياس تطور الأجور خاصة على المستوى الكلي...الخ.

Iــ تعريف الأرقام القياسية

الرقم القياسي: هو رقم نسبي، وظيفته تكمن في قياس التغير الذي يحصل بين قيمتين أصليتين من قيم الظاهرة محل الدراسة. إذا كان لدينا على سبيل المثال سلسلة من القيم، يمكن حساب التغير النسبي لكل قيمة من قيم الظاهرة، بالنسبة إلى قيمة معينة يستخدمها الباحث كأساس للمقارنة.

II ــ أنواع الأرقام القياسية

نميز في الأرقام القياسية ثلاثة أنواع:

ــ الأرقام القياسية للأسعار ، الأرقام القياسية للكميات ، الأرقام القياسية للقيم.

III ــ قواعد حساب الأرقام القياسية

نميز بين ثلاثة حالات من طرق حساب الأرقام القياسية :

ــ الرقم القياسي البسيط ، الرقم القياسي التجميعي البسيط ، الرقم القياسي المرجح.

01 ــ الرقم القياسي البسيط:

نكون هذا الرقم من معطيات تاريخية لسلاسل زمنية فردية ، هذه السلاسل التي تغطي إما فترة زمنية معينة أو تشمل مناطق مختلفة.

نقوم باختيار فترة زمنية أو مكان معين كأساس، شريطة أن تكون المفردة المختارة تساوي 100 وبالتالي المفردات الأخرى في السلسلة تعرض في صورة نسب مئوية من الأساس المختار. لذلك، نسمي الرقم القياسي البسيط أحيانا بمنسوب السعر أو منسوب الكمية أو منسوب القيمة.

1 – 1 – منسوب السعر: هو نسبة السعر لسلعة واحدة في فترة المقارنة إلى سعرها في فترة الأساس.

نفرض p_0 هي سعر السلعة في فترة الأساس و p_1 هي سعر نفس السلعة خلال فترة المقارنة، وعليه تكون القاعدة على النحو التالي:

$$p_{01} = \frac{p_1}{p_0} \times 100$$

1 – 2 – منسوب الكمية أو الحجم : يلجأ الباحث إلى هذه القاعدة ، حينما يكون اهتمامه منصب على مقارنة كميات أو حجوم السلعة (إنتاج ، استهلاك ، تصدير ، استيراد ...الخ).بنفس الطريقة يمكن حساب التغير الحال بين فترتين .

$$q_{01} = \frac{q_1}{q_0} \times 100$$

1 – 3 – منسوب القيمة: المعروف أن القيمة، هي ناتجة عن حاصل ضرب الكمية في السعر، وعليه تكون القاعدة على النحو التالي:

$$v_{01} = \frac{p_1 \cdot q_1}{p_0 \cdot q_0} \times 100 = \frac{v_1}{v_0} \times 100$$

1 – 4 – مناسيب السلسلة : إذا كان لدينا $p_1 ,p_2 ,p_3 , \ldots . . p_n$ تمثل الأسعار خلال الفترات المتتالية: 1, 2, 3...، إذن فإن: $p_{\frac{2}{1}} ,p_{\frac{3}{2}} ,p_{\frac{4}{3}} , \ldots$ تمثل مناسيب الأسعار لكل فترة زمنية ، مقارنة بالفترة الزمنية السابقة لها. نسمي أحيانا هذه المناسيب بمناسيب السلسلة.

ما دمنا نعالج الرقم القياسي البسيط، يمكن في هذه الحالة استخدام ما يسمى بمتوسط المناسيب ، وذلك بغرض الوصول إلى الرقم القياسي لمجموعة من السلع . نستعمل في هذا الأساس إما الوسط الحسابي أو الوسط الهندسي.

إذا كان لدينا N سلعة وكانت مناسيبها على التوالي: $p_1 ,p_2 ,p_3 , \ldots . . ، p_n$ مع الأخذ بعين الاعتبار أن:

$$p = \frac{p_1}{p_0} \times 100$$

بناء على هذا الأساس يمكن إيجاد متوسط المناسيب على النحو التالي:

$$p_{01} = \frac{\sum \frac{p_1}{p_0}}{N} \times 100$$

02 ـ الرقم القياسي التجميعي البسيط

أحيانا يرغب الباحث في معرفة أثر السعر لمجموعة من السلع في فترة زمنية معينة، بالنسبة إلى فترة زمنية أخرى. وفي هذه الحالة تكون كافة السلع لها نفس الأهمية عند الحساب.

تعطى علاقة الحساب على النحو التالي:

$$p_{01} = \frac{\sum p_1}{\sum p_0} \times 100$$

03 ـ الرقم القياسي المرجح

يستعمل احيانا للتغلب على بعض عيوب الطريقة السابقة وهذا ما يجعل الباحث مطالبا باستخدام معامل الترجيح الملائم، إذ غالبا ما يلجأ إلى استخدام كمية أو حجم السلعة المباعة (المنتجة) خلال فترة الأساس أو فترة المقارنة أو فترة نموذجية (التي تتضمن متوسط لعدد من السنوات).كأوزان تشير إلى أهمية السلع الداخلة في تكوين الرقم القياسي.

هناك أربع صيغ ممكنة يعتمد عليها الباحث وهي:

q_0 : كمية فترة الأساس

q_1 : كمية فترة المقارنة

q_m : كمية الفترة النموذجية

q_{0+1} : كميتي فترتي الأساس و المقارنة

3 ـ 1 ـ المنسوب المرجح: نميز فيه خمسة طرق:

الطريقة الأولى: استعمال القيمة الخاصة بسنة الأساس $(p_0 q_0)$ لترجيح مناسيب الأسعار. في هذه الحالة نطبق القاعدة التالية :

$$p_{01} = \frac{\sum \left[\left(\frac{p_1}{p_0}\right)(p_0 q_0)\right]}{\sum (p_0 q_0)} \times 100$$

الطريقة الثانية: يستعمل فيها الباحث القيمة الخاصة بسنة المقارنة $(p_1 q_1)$ لترجيح مناسيب الأسعار وتعطى العلاقة على النحو التالي:

$$p_{01} = \frac{\sum\left[\left(\frac{p_1}{p_0}\right)(p_1 q_1)\right]}{\sum(p_1 q_1)} \times 100$$

الطريقة الثالثة: يأخذ بعين الاعتبار أسعار سنة الأساس وكميات سنة المقارنة $(p_0\ q_1)$ لترجيح مناسيب الأسعار

وبناء على هذا الأساس تعطى العلاقة كما يلي:

$$p_{01} = \frac{\sum\left[\left(\frac{p_1}{p_0}\right)(p_0 q_1)\right]}{\sum(p_0 q_1)} \times 100$$

الطريقة الرابعة: يستعمل فيها أسعار سنة المقارنة وكميات سنة الأساس $(p_1\ q_0)$ لترجيح مناسيب الأسعار وتكون العلاقة على النحو التالي:

$$p_{01} = \frac{\sum\left[\left(\frac{p_1}{p_0}\right)(p_1 q_0)\right]}{\sum(p_1 q_0)} \times 100$$

الطريقة الخامسة: في هذه الطريقة نستعمل أسعار وكميات سنة معينة (نموذجية) تختلف عن سنة الأساس أو سنة المقارنة، أي نستعمل $(p_m\ q_m)$ وفق المعادلة التالية:

$$p_{01} = \frac{\sum\left[\left(\frac{p_1}{p_0}\right)(p_m q_m)\right]}{\sum(p_m q_m)} \times 100$$

ملاحظة: ما ينطبق على مناسيب الأسعار ينطبق أيضا على مناسيب الكميات ومناسيب القيم.

04 ـ الرقم القياسي التجميعي المرجح

نميز فيه بين عدة طرق تبعا للمنظرين الاقتصاديين اللذين وضعوا هذه القواعد، نختار منهم ما يلي:

4 ـ1 ـ رقم لاسبير : هو رقم قياسي تجميعي مرجح يستخدم سنة الأساس ويميز فيه مقياس يتعلق بالأسعار والثاني يتعلق بالكميات .

ـ أما الرقم القياسي التجميعي للأسعار: يفترض فيه ثبات أذواق المستهلكين واستمرارهم في استهلاك نفس كميات السلع حتى لو تغيرت أسعارها بالارتفاع أو الانخفاض.

ـ بينما في الرقم القياسي التجميعي للكميات: يفترض ثبات الأسعار في فترتي الأساس والمقارنة بض النظر عن تغير الكميات المستهلكة في الفترتين.

نكتفي بتقديم طريقة حساب الرقم القياسي للأسعار، على اعتبار أن الرقم القياسي للكميات هو نفس القاعدة مع وضع (q_{01}) بدلا من (p_{01}).

$$p_{01} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

ملاحظات تخص رقم لا سبير		
	المزايا	العيوب
01	يستعمل هذا الرقم للتغلب على عيوب الرقم التجميعي البسيط الذي مر بنا.	إذا كان الفارق كبيرا بين فترة الأساس وفترة المقارنة أو إذا تغيرت أهمية السلع النسبية، فإن هذا الرقم يفقد مصداقيته، حيث لا يكون معبرا تعبيرا دقيقا عن التغير في السعر
02	يصلح استعمال هذا الرقم في البلاد المتخلفة التي لا تتوفر على الامكانات الفنية اللازمة	عادة ما ينحاز هذا الرقم إلى أعلى نتيجة لكون معظم السلع مرنة في الطلب، بمعنى أن الكميات المطلوبة تتناقص كلما زاد السعر.
03	يمكن استعماله في المقارنة بين سلسلة من الأرقام القياسية المتتالية. حيث تكون الكميات مشتركة ويبقى أثر التغير في السعر فقط.	هذا الرقم ترجيحه يكون صغيرا في حالة السلع التي تتميز بسعر منخفض والعكس صحيحا، وهذا يعتبر خطأ فني يجب تلافيه.

4 – 2 – رقم باش : يهدف إلى حساب الثمن الذي سيدفع مقابل كميات سنة المقارنة ، نسبة إلى الثمن الذي سيدفع مقابل نفس هذه الكميات ولكن بسعر الأساس.

$$p_{01} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$$

ملاحظات تخص رقم باش		
	المزايا	العيوب
01	يتغلب على عيوب الرقم التجميعي البسيط	متحيز للأسفل
02	يأخذ أحدث الكميات المتداولة في السوق كترجيحات.	لا يستطيع أن يقارن بين سلسلة من الأرقام القياسية نظرا لتير الترجيحات من فترة إلى أخرى.
03	————————————	تتأخر نتائج الدراسة عند استخدام هذا الرقم ، نظرا لصعوبة الحصول على كميات المقارنة في وقت مبكر.

4 ـ 3 ـ رقم (مارشال ـ ادجو رث): يهدف هذا الرقم إلى إيجاد الثمن الذي يدفع مقابل كميات سنتي الأساس بسعر المقارنة ، نسبة إلى الثمن الذي يدفع مقابل نفس هذه الكميات ولكن بسعر الأساس.

$$p_{01} = \frac{\sum(q_0 + q_1)p_1}{\sum(q_0 + q_1)p_0} \times 100$$

فائدة : هذا الرقم يتفادى عيوب كل من رقم لا سبير ورقم باش.

4 ـ 4 ـ رقم فيشر (الرقم القياسي الأمثل)

وهو عبارة عن الوسط الهندسي لكل من رقمي لا سبير وباش. سمي بالأمثل لأنه يعكس الزمن والمعامل في نفس الوقت ، هذا بالإضافة إلى سهولة حسابه ، مما ينعكس إيجابا على دقة الحساب وتفاديه بالتالي لكل عيوب الارقام القياسية السابقة ، مما يجعل استخدامه مفضلا في كافة الأحوال.

$$p_{01} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

أو نقول : $p_{01} = \sqrt{l \times p}$

حيث : (L) و (P) هما قانوني لاسبير وباش على التوالي.

05 ـ الرقم القياسي لنفقة المعيشة

تعريف نفقة المعيشة: تعرف نفقة المعيشة على أنها ثمن الحصول على السلع والخدمات التي تستهلك فعلا، خلال فترة زمنية معينة.

أما مستوى المعيشة، يقصد به كمية السلع والخدمات المستهلكة في نفس الفترة. ونظرا لكون رفاهية السكان تتحقق حينما تزداد مداخيلهم ، وبالتالي زيادة الانفاق ، هذا يؤدي حتما إلى تغير في مستوى الأسعار بالزيادة ، نظرا لزيادة الطلب على السلع والخدمات ، كذلك أثر الزيادة في الأسعار يحصل حينما تتغير التركيبة السكانية بالزيادة ، لأن ذلك يولد طلب جديد على السلع والخدمات.

$$(\text{l'indice des frais de subsistance}) L_{fs} = \frac{\sum p_1 q_0}{p_0 q_0} \times 100$$

بناء على هذا الأساس، تقودنا دراسة الرقم القياسي لنفقة المعيشة، إلى دراسة الأجر الحقيقي والأجر النقدي.

5 ـ 1 ـ الأجر النقدي (الاسمي) : نظريا هو مقدار النقود التي يحصل عليه العاملين ، مقابل بيع مجهوداتهم الفكرية والعضلية للغير ، بمعنى الوحدات النقدية التي يحصل عليها الأفراد.

5 ـ 2 ـ الأجر الحقيقي: يقصد به معرفة مقدار السلع والخدمات التي يحصل عليها المستهلكون، مقابل إنفاق دخلهم النقدي.

06 ـ الرقم القياسي للأجور: وهو يستعمل خصيصا لقياس التطور الذي يحدث لأجور فئة عمالية معينة ولفترتين زمنيتين مختلفتين. من هذا المنطلق يمكن حساب رقم قياسي مرجح بطريقتين مختلفتين.

6 ـ 1 ـ الرقم القياسي للأجور استنادا إلى الفئة العمالية : يحسب بالطريقة التالية.

$$I_{S\ t1/t0} = \frac{\sum n_0^j \times \frac{S_1^j}{S_0^j}}{\sum n_0^j} \times 100$$

حيث:

- n_0^j تمثل عدد الأجراء للفئة j محل الدراسة
- S_0^J تمثل أجور الفئة j محل الدراسة في فترة الأساس ، S_1^J تمثل أجور الفئة j محل الدراسة في الفترة الحالية (أثناء الدراسة)

مثال : نفرض أن عدد عمال الصيانة لا حدى المؤسسات كان 10 عمال ، وأن متوسط الأجر لهذه الفئة كان 35000دج بحلول نهاية 2012 مع العلم أن متوسط الأجر لهذه الفئة سنة الأساس (2010) كان 25000دج.

المطلوب: حساب الرقم القياسي للأجور المتعلق بهذه الفئة.

حل المثال:

$$Is_{\ t1/t0} = \frac{\sum 10 \times \frac{35000}{25000}}{\sum 10} \times 100 = 140$$

تفسير النتيجة: أي أن أجور العاملين تغيرت بزيادة قدرها *40 %* = (*140 – 100*)

6 ـ 2 ـ الرقم القياسي للأجور استنادا إلى علاقة لاسبير

$$I_{L\ S\ t1/t0} = \frac{\sum \frac{S_1^j}{S_0^j} \times S_0^j \times n_0^j}{\sum S_0^j \times n_0^j} \times 100$$

ملاحظة : كل الرموز السابقة تستعمل في هذه الطريقة.

مثال : إذا كانت سنة الأساس هي : 2010 وأن الرقم القياسي لأسعار الاستهلاك في الجزائر العاصمة هو160 % مع نهاية شهر سبتمبر 2015 ، بينما سجل 160.6 % في السنة الموالية (2016).

المطلوب :حساب القدرة الشرائية للدينار الجزائري.

حل المثال

ـ القدرة الشرائية للدينار الجزائري سنة 2015 هي : $p(a) = \frac{1_{DA}}{160} \times 100 = 0.625$

تفسير النتيجة : واحد (1) دينار جزائري في سنة الأساس (2010) اصبح يساوي سنة 2015 فقط 0.625 دج ، أي أنه فقد خلال المدة (2010 ـ 2015) ما قيمته 0.375 دج.

ـ القدرة الشرائية للدينار الجزائري سنة 2016 هي : $p(a) = \frac{1_{DA}}{160.6} \times 100 = 0.622$. نفس التفسير السابق ، مع ملاحظة أن التدهور لقيمة العملة الوطنية زاد .

07 ـ الاعتبارات الواجب مراعاتها عند تكوين الارقام القياسية

7 ـ 1 ـ اختيار المعطيات التي يتكون منها الرقم القياسي :

ويتعلق ذلك ، باختيار السلع التي تدخل في تركيب الرقم القياسي واختيار المصادر التي نستقي منها أسعار السلع ، لأنه من الصعوبة بمكان ادخال جميع السلع في تركيب اي رقم قياسي ، نظرا لتعدد السلع من جهة وتعدد درجات السلعة الواحدة بين بلدان متقدمة وأخرى متخلفة.

لذلك يلجأ عادة الى اختيار عينة من السلع تتصف باستقطاب الجزء الأكبر من انفاق المجتمع.

7 ـ 2 ـ اختيار فترة الأساس : إذ يشترط مراعاة ما يلي:

ـ ان تكون فترة الأساس هادئة خالية من الاضطرابات ، حتى لا تكون قيم الظاهرة قيما شاذة .

ـ الا تكون فترة الأساس بعيدة جدا عن الفترات محل الاختبار ، إذ لا معنى لأسعار اليوم بأنها تضاعفت عشرات المرات عن اسعار 1850 مثلا. إلا إذا كانت الدراسة بغرض المقارنة. ومع هذا نتائج البحث تبقى غير معبرة تعبيرا دقيقا لعدة عوامل تدخل على مدار الفترة ، مثل التعديلات في التوضيب ، والأغلفة أو اختفائها تماما من السوق ...الخ.

7 ـ 3 ـ اختيار الصيغة

لقد مرت بنا عدة صيغ ، لذلك يجب اختيار الصيغة المناسبة لتكوين الرقم القياسي ، هل نأخذ بالأرقام القياسية البسيطة ، المرجحة ، أو الارقام القياسية للمناسيب وترجيحاتها ، لذلك يتطلب اختبار صحة الارقام القياسية من خلال اختبارات الانعكاس في الزمن واختبارات الانعكاس في المعامل.

08 ـ تعديل سعر سلعة ـ ما ـ

الهدف من تعديل سعر سلعة ، هو الحصول على التغيرات الحقيقية الخالية من أثر التضخم ، ولتحقيق هذه الغاية ، نستعمل في التعديل أحد الأرقام المتعارف عليها ، مثل الرقم القياسي لنفقات المعيشة أو الرقم القياسي لأسعار التجزئة أو اسعار الجملة ، إذ يمكننا الحصول على هذه الارقام من الديوان الوطني للإحصاء .

$$\text{السعر المعدل} = \frac{\text{سعر السلعة في سنة ما}}{\text{الرقم القياسي لأسعار الجملة أو نفقات المعيشة أو التجزئة}} \times 100$$

كما يمكن تقدير القوة الشرائية لأية سلعة.

$$\text{الرقم القياسي للقوة الشرائية لأية سلعة} = \frac{\text{السعر النسبي لهذه السلعة}}{\text{الرقم القياسي لأسعار السلع المماثلة}}$$

09 ـ الأرقام القياسية للتجارة الخارجية : تهدف الارقام القياسية للتجارة الخارجية الى توضيح التغيرات التي تطرأ على قيم وكميات وأسعار كل من الصادرات والواردات ، نميز فيها ما يلي :

9–1ـــ معدل التبادل الاجمالي = $\frac{\text{الرقم القياسي لكمية الواردات}}{\text{الرقم القياسي لكمية الصادرات}} \times 100$

9 ـ2ـ معدل التبادل الصافي = $\frac{\text{الرقم القياسي لأسعار الصادرات}}{\text{الرقم القياسي لأسعار الواردات}} \times 100$

9 ـ 3ـ معدل التبادل الداخلي = $\frac{\text{الرقم القياسي لقيمة الصادرات}}{\text{الرقم القياسي لقيمة الواردات}} \times 100$

IV- تمارين محلولة

التمرين الأول

المعطيات الواردة في الجدول أدناه ، تمثل أسعار 05 سلع مختلفة ، على اعتبار أن سنة الأساس هي سنة 2016 وأن سنة المقارنة هي 2018 .

السلع (دج)	سعر 2016	سعر 2018
السلعة 1....................	40	50
السلعة 2....................	60	60
السلعة 3....................	20	30
السلعة 4....................	50	70
السلعة 5....................	80	90
	$=250 \sum p_0$	$\sum p_1 = 300$

المطلوب حساب :

1 ـ أحسب الرقم القياسي بالطريقة التجميعية البسيطة.

2 ـ أحسب الرقم القياسي التجميعي باستخدام الوسط الحسابي .

حل التمرين

ـ الرقم القياسي بالطريقة التجميعية البسيطة $p_{01} = \frac{\sum p_1}{\sum p_0} \times 100$

$$p_{01} = \frac{300}{250} \times 100 = 120$$

معنى ذلك أن اسعار هذه العينة من السلع ارتفعت بزيادة قدرها 20 % عما كانت عليه سنة الأساس (2016).

ـ الرقم القياسي التجميعي باستخدام الوسط الحسابي : لحساب هذا الرقم القياسي نستعين بالجدول التالي.

الرقم القياسي التجميعي باستخدام الوسط الحسابي			
السلع	سعر الأساس $(P_0)2016$	سعر المقارنة $(P_1)2018$	$(P_1 / P_0) \times 100$
A	40	50	125
B	60	60	100
C	20	30	150
D	50	70	140
E	80	90	112
			627

$$p_{01} = \sum_{i=0}^{n} \frac{p_1}{p_0} \times 100/n = 627/5 = 125,4$$

معنى ذلك ، هناك زيادة بنسبة 25,4 %

التمرين الثاني

نفرض لديك عينة من السلع A B C D E F والمطلوب هو تكوين الرقم القياسي البسيط للبيانات الواردة في الجدول التالي ، على اعتبار أن سنة الأساس هي : 2010 تساوي 100 ، مع العلم أن كل من سنتي 2015 و 2018 هي سنة مقارنة.

السلع (دج)	P_0	P_1	P_2
A	250	255	240
B	225	220	250
C	190	200	230
D	300	300	260
E	110	120	150
F	215	200	240

الحل : نوجزه في الجدول الموالي ، وذلك بتطبيق قاعدة الحساب للرقم القياسي البسيط

سعر 2018	سعر 2015	سعر 2010	السلع (دج)
96	102	100	A
111,11	97,8	100	B
121	105,3	100	C
86,67	100	100	D
136,37	109	100	E
111,63	93	100	F
662,78	**607,10**	**600**	**المجموع النسبي**
110,50	**101,20**	**100**	**الرقم القياسي البسيط**

التفسير

من خلال النتائج يتضح أن هذه التشكيلة من السلع ارتفعت اسعارها سنة 2015 مقارنة بسنة الأساس 2010 في حدود 1,20 % بينما تواصل الارتفاع سنة 2018 لتسجل بذلك 10,50 % بالنسبة لسنة الأساس 2010.

التمرين الثالث

اليك المعطيات التالية المتعلقة بعينة من السلع ، حسب الجدول التالي :

p_1	q_1	p_0	q_0	السلع (دج)
30	7	40	6	A
50	5	45	4	B
40	1.5	90	0.5	C

المطلوب : حساب كل من رقم لاسبير ورقم باش

الحل : للحصول على نتائج السؤال ، نستعين بالجدول التالي

p_0q_1	p_1q_1	p_0q_0	p_1q_0	p_1	q_1	p_0	q_0	السلع
280	210	240	180	30	7	40	6	A
225	250	180	200	50	5	45	4	B
135	60	45	20	40	1.5	90	0.5	C
640	**520**	**465**	**400**					**المجموع**

01 – رقم لا سبير

$$p_{01} = \sum p_1 q_0 / \sum p_0 q_0 \times 100$$

$$p_{01} = 400/465 \times 100 = 86$$

02 – رقم باش

$$p_{01} = \sum p_1 q_1 / \sum p_0 q_1 \times 100$$

$$p_{01} = 520/640 \times 100 = 81,2$$

التمرين الرابع

أحسب نفقات المعيشة من الجدول التالي

مجموعات الاستهلاك	رقم المؤشر	الوزن النسبي
الطعام	352	48
الكهرباء و الغاز	320	10
الملابس	230	08
الايجار	160	12
أشياء أخرى (مستهلكة)	190	15

للجواب على السؤال ، نضيف فقط للجدول السابق خانة أخرى تتعلق بنتيجة ضرب رقم المؤشر بالوزن النسبي ثم نطبق قاعدة الحساب كما نرى لاحقا.

مجموعات الاستهلاك	رقم المؤشر(I)	الوزن النسبي (COF)	I × COF
الطعام	352	48	16.896
الكهرباء و الغاز	320	10	3.200
الملابس	230	08	1.840
الايجار	160	12	1.920
أشياء أخرى (مستهلكة)	190	15	2.850
المجموع		**93**	**26.706**

مؤشر كلفة المعيشة : (L_{fs})

$$L_{fs} = I \times COF/COF = 26.706/93 = 287{,}16$$

معنى ذلك ، أن كلفة المعيشة ارتفعت بمقدار 187,16 % عما كانت عليه في سنة الأساس .

التمرين الخامس:

أحسب رقم فيشر من المعطيات الواردة في الجدول أدناه

السلع	السعر (دج)		الكمية (كلغ)	
	2017	2018	2017	2018
A	8	20	50	60
B	2	6	15	10
C	1	2	20	25
D	2	5	10	08
E	1	5	40	30

الحل : للوصول الى نتائج السؤال نستعين بالجدول التالي

السلع	P_0	P_1	q_0	q_1	p_1q_0	p_0q_0	p_1q_1	p_0q_1
A	8	20	50	60	1000	400	1200	480
B	2	6	15	10	90	30	60	20
C	1	2	20	25	40	20	50	25
D	2	5	10	08	20	20	40	16
E	1	5	40	30	200	40	150	30
المجموع					1350	510	1500	571

فيشر المؤشر المثالي

$$p_{01} = \sqrt{\sum p_1 q_0 / \sum p_0 q_0 \times \sum p_1 q_1 / \sum p_0 q_1} \times 100$$

بالتعويض ما تحت الجذر نجد ما يلي:

$$p_{01} = \sqrt{1350/510 \times 1500/571} = \sqrt{2025000/291210} \times 100 = \underline{263,05}$$

تفسير النتيجة : معنى ذلك أن الأسعار ارتفعت في سنة 2018 بنسبة <u>163 %</u> عما كانت عليه في سنة 2017.

المراجع

I – باللغة الوطنية

01 – محمد محمد جبر المغربي ، الاحصاء التحليلي في البحوث الاقتصادية والاجتماعية ، المكتبة العصرية للنشر والتوزيع ، جمهورية مصر العربية، الطبعة الأولى ، 2011

02 – سعيد السيد علي اسماعيل ، مبادئ الاحصاء الوصفي والتطبيقي ، مؤسسة حورس للنشر والتوزيع ، الإسكندرية ، جمهورية مصر العربية ، 2006

03 – تيلولت سامية ، مبادئ في الاحصاء ، مطبعة دار الحديث ، القبة الجزائر ، الطبعة الثانية ،2009

04 – عبد الكريم بوحفص ، الاحصاء المطبق في العلوم الاجتماعية والانسانية، ديوان المطبوعات الجامعية ، الجزائر العاصمة ، الطبعة الثالثة منقحة ،2011

05 – جلاطو جيلالي ، الاحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية ، الجزائر العاصمة ، 1999.

06 – أنيس اسماعيل كنجو، الإحصاء والاحتمال، مكتبة العبيكان ،المملكة العربية السعودية ، الطبعة الأولى ، 2000

07 – عدنان ماجد وآخرون ، مبادئ الاحصاء والاحتمالات ، جامعة الملك فيصل ، المملكة العربية السعودية ، الطبعة الثانية ، 2007

08 – جلال الصياد ، عبد الحميد محمد ربيع، مبادئ الطرق الاحصائية ، مطبعة تهامة ، جدة ،المملكة العربية السعودية ،الطبعة الأولى 1983.

09 – احمد عبد السميع طبيه ، مبادئ الاحصاء ،دار البداية للنشر والتوزيع ، عمان الأردن، الطبعة الأولى ، 2008

10 – ابراهيم مراد الدعمة ،مازن حسن الباشا ، أساسيات في علم الاحصاء مع تطبيقات spss ، الطبعة الأولى، دار المناهج للنشر والتوزيع الأردن، 2013

11 – شرف الدين خليل ، الاحصاء الوصفي ، شبكة الأبحاث والدراسات الاقتصادية.

12 – عبد الرزاق عزوز، الكامل في الاحصاء ، دروس مفصلة ، تمارين ومسائل مع الحلول ، الجزء الأول ، ديوان المطبوعات الجامعية ، الجزائر، 2010

13 – عبد الناصر رويسات ، الاحصاء الوصفي ومدخل للاحتمالات "دروس وتمارين" ديوان المطبوعات الجامعية ، وهران ، الجزائر

II – باللغة الأجنبية

1- jean - Louis Monino, STATISTIQUE DESCRIPTIVE , Dunod , paris, 4e édition, 2010

2- Bernard Verlan, Geneviève Saint-Pierre, Statistiques et Probabilités, Berti Editions, Alger, 2008

3- Luc Albarello, Statistique Descriptive, Edition De Boeck Université, 2e édition, Bruxelles ,2007

4 - Etienne BRESSOUD, Jean-Claude KAHANE, Statistique Descriptive, 2e édition, France, paris, 2010
