(مبادئ الإحصاء)

د/ نصر الحسيني

12/2/2024

الفصل الأول ← مقدمة لدراسة علم الإحصاء :-

← علم الإحصاء من أهم علوم الرياضيات يدخل في جميع العلوم الطبيعية والتطبيقية الإجتماعية وهو من العلوم الواسعة والمهمة في كل مجالات الحياة.

تعريفات :- statistic

١- علم الإحصاء : هو العلم الذي يهتم بدراسة الظواهر المختلفة عن طريق جمع البيانات وتلخيصها وعرضها وترتيبها وتحليلها واختبار الفروض الإحصائية عنها وتفسيرها والنتائج ونشرها لذوي الأهمية لاتخاذ القرار اللازم.

الأهداف

- جمع البيانات ← من الظاهرة محل الدراسة
- عرض البيانات وتنظيمها ← عن طريق جداول تكرارية / بيانياً
- تحليل البيانات ← باستخدام بعض النظريات الرياضية
- تفسير النتائج واتخاذ القرار ← حدوث التنبؤ بهذه الظاهرة.

← يمكن تقسيم علم الإحصاء إلى :-

1. الإحصاء الوصفي.
2. الإحصاء الاستدلالي (التحليلي).

الوصفي

هو الذي يتعامل مع البيانات بالجمع والعرض والتلخيص بغرض إظهار وصف هذه البيانات ← جمع البيانات وعرض البيانات

الاستدلالي

هو الذي يهتم باستخدام البيانات المجمعة والمعروضة عن الظاهرة في التنبؤ والتحليل والتخطيط والتقدير لها في المستقبل

النتائج ← تحليل ← تفسير ← تنبؤ

الظاهرة:

هو حدث يتكون بعد يتكرر لعدة شروط وظروف

أنواع الظواهر ← منها ① ظواهر طبيعية الزلازل البراكين المطر

② ظواهر اجتماعية EX ← الزواج والطلاق.

③ ظواهر إقتصادية EX ← البطالة.

④ ظواهر سياسية EX ← الحروب

← المجتمع الإحصائي ← هو عبارة عن مجموعة محدودة "معدودة" أو غير محدودة "غير معدودة" من المفردات التي تشترك في خصائص معينة.

← المجتمع المحدود "الذي: هو مجتمع يحتوي على عدد معروف من العناصر مثل مجتمع طالبات جامعة الأزهر.

← المجتمع غير المحدود": هو مجتمع يحتوي على عدد غير معروف من العناصر مثل الأسماك في البحر، كرات الدم الحمراء في جسم الإنسان.

أنواعه ① مجتمع ذو متغير واحد.

② مجتمع ذو متغيرين.

③ مجتمع متعدد المتغيرات.

* العينة الإحصائية * صعبة وأحيانا مستحيلة

هو جزء من مجتمع ← عدد محدد من العناصر المختارة من المجتمع

* تعريف المعلمة ~~والإحصاءة~~: أي مقياس ~~[illegible]~~ رقمي يحسب من المجتمع

* تعريف الإحصاءة: أي مقياس رقمي يحسب من العينة.

← فيما يلي تقديرات مجموعة مكونة من ٢٠ الطلاب في إحدى المواد ، التقديرات :
جيد - جيد جداً - راسب - مقبول - مقبول - مقبول - ممتاز - جيد - جيد - مقبول
جيد جداً - جيد - مقبول - مقبول - مقبول - جيد - جيد - جيد - مقبول - جيد - راسب
- مقبول
بيانات وصفية
* أوجد جدول توزيع هذه البيانات

الصفات	العلامات	التكرار
ممتاز	ا	1
جيد جداً	اا	2
جيد	ااا ، ~~اااا~~	8
مقبول	اا ~~اااا~~	7
راسب	اا	2
مجموع التكرار		20

* جدول التوزيع البيانات يتكون من صفين
أول صف ← التقديرات ثاني صف ← عدد الطلاب

التقديرات	ممتاز	جيد جدا	جيد	مقبول	راسب
عدد الطلاب	1	2	8	7	2

٢ عرض البيانات الكمية المتقطعة : هي بيانات لمتغير يأخذ قيماً معينة ولا يوجد له أي قيمة بين قيمتين . "ما بتتكون من أعداد"
← يوضع في العمود الأول جميع القيم التي في البيانات
← ~ ~ الثاني العلامات
~ ~ الثالث التكرار
يتكون من ٣ أعمدة

Date :
No:

← البيانات التالية تمثل عدد الأطفال في مجموعة من الأسر ١ ، ٢ ، ٢ ، صفر ، ٣ ، ٣ ، ٥ ،
صفر ، ١ ، ١ ، ٢ ، ٣ ، ٢ ، ١ ، ٣ ، ٢ ، ٢ ، ١ ، ٢ ، ٣ ، ٢ ، ٢ ، ٢

① أوجد جدول توزيع البيانات

جدول تفريغ

بيانات	علامات	تكرار
صفر	॥	٢
١	𝍸	٥
٢	𝍸 ॥।	٨
٣	॥॥	٤
٥	।	١
المجموع		٢٠

② جدول تكراري توزيع البيانات

عدد الأولاد

~~التقديرات~~ البيانات	صفر	١	٢	٣	٥	
عدد ~~الأطفال~~ الأسر	٢	٥	٨	٤	١	#

ⓑ عرض البيانات الكمية المستمرة : لعمل جدول تفريغ البيانات نكون جدول مكون من ٣ أعمدة
يوضع في العمود الأول ما يسمى بالمجموعات (فئات)
~ ~ الثاني العلامات
~ ~ الثالث التكرار

⊛ طريقة عمل الفئات (مجموعات) :
① نحسب مدى البيانات ويساوي أكبر مشاهدة مطروحًا منها أصغر مشاهدة
② نقسم هذا المدى إلى أقسامًا متساوية يسمى كل قسم بالفئة ولا توجد قاعدة رياضية لتحديد عدد الأقسام وإنما يمكن أخذها بين عدد محصور بين (٥ ، ١٢)
(بالخبرة)

Subject:

③ نحسب طول الفئة وهو = $\frac{\text{المدى}}{\text{عدد الفئات}}$ مع مراعاة إذا كان خارج

القسمة عدد صحيح وكسر يقرب إلى العدد الصحيح التالي مهما كان الكسر صغير

على سبيل المثال: إذا كان المدى ٤٦ وعدد الفئات ٥

خارج القسمة = $\frac{٤٦}{٥}$ = ٩,٢

= ١٠

④ نكون الفئات فتكون الفئة الأولى بدايتها أصغر مشاهدة ونهايتها أصغر مشاهدة مضاف إليها طول الفئة مطروحًا منها ١ وبداية الفئة الأولى مضافًا إليها ١

مثال/ البيانات التالية تبين أعداد الآلات الصالحة في عينة من ٢٠ مصنع بإحدى المدن الصناعية ٤٧ ، ٤٠ ، ٥٢ ، ٧٧ ، ٩٠ ، ٦٢ ، ٨٥ ، ٩ ، ٤٤ ، ٦٥ ، ٨٨ ، ٧١ ، ٧٥ ، ٦١ ، ٦٧ ، ٥٥ ، ٦٨ ، ٧٢ ، ٩٥ ، ٦٦ ، ٦٩ ،

← تلخيص هذه البيانات في جدول توزيع تكراري مناسب

الحل

المدى = ٩٥ - ٤٠ = ٥٥

عدد الفئات = ٦

طول الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}}$ = $\frac{٥٥}{٦}$ = ٩,١٦٦٦ ≃ ١٠

جدول التفريغ

الفئات	العلامات	التكرار
40 - 49	|||	٣
50 - 59	||	٢
60 - 69	~~||||~~ ||	٧
70 - 79	||||	٤
80 - 89	||	٢
90 - 99	||	٢
المجموع		٢٥

فئات الآلات	40-49	50-59	60-69	70-79	...
عدد المصانع	3	2	7	4	2

* تكوين جداول التوزيعات التكرارية للمتغيرات الكمية المتصلة
- يحدد المدى و يقسم إلى فئات وتكون كل فئة عبارة عن فترة ويحدد الحد الأدنى و الحد الأعلى لكل فئة
- عدد الفئات لا يقل عن ٥ و لا يزيد عن ٢٠
- في حالة الفئات متساوية الطول نحصل على طول الفئة بقسمة المدى على عدد الفئات مع تقريب خارج القسمة

* مثال

البيانات التالية تمثل الوزن بالكيلوجرام لعينة من ٥٠ طالب في أحدى الكليات

٤٢، ٧٤، ٤٠، ٦٠، ٨٢، ١٠٨، ٤١، ٦١، ٧٥، ٨٣، ٦٣، ٥٣، ١٠٠،

٧٦، ٨٤، ٥٠، ٦٧، ٦٥، ٧٨، ٧٧، ٥٦، ٨١، ٩٥، ٨٤، ٨٣،

٢، ٨٧، ٦٨، ٩٤، ٨٠، ٦٩، ٩٤، ٨٠، ٧٩، ٧٩، ٥٤، ٧٣

٥٩، ٨١، ١٠٠، ٦٦، ٤٩، ٧٧، ٩٠، ٨٤، ٧٦، ٤٢، ٦٤، ٦٩، ٧٠

٨٠، ٧٢، ٥٠، ٧٩، ٥٢، ١٠٣، ٩٦، ٥١، ٨٦، ٧٨، ٩٤، ٧١

* كوني جدول توزيع تكراري مناسب

المدى = ١٠٨ - ٤٠ = ٦٨ كجم

عدد الفئات = ٧

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{٦٨}{٧} = ٩\frac{٥}{٧} \simeq ١٠$$

- جدول تعويع الوزنات

فئة الوزنات	العلامات	التكرار
40 les than 50	卌	5
50 — 60	卌 III	8
60 — 70	卌 卌	10
70 —	卌 卌 卌	15
80 —	卌 卌 III	13
90 —	卌	5
100 les than 110	IIII	4
المجموع		60

- جدول توزيع

فئة الوزن	
التكرار	

- التوزيعات التكرارية + الفئات غير متساوية الطول
+ التوزيعات التكرارية متساوية الطول يمكن تسميتها توزيعات تكرارية منتظمة ولكن الاعتبارات عملية قد تقتضي إعداد جداول تكرارية غير متساوية الطول (توزيعات تكرارية غير منتظمة)

مثلاً الجدول التالي يبين توزيع سكان جمهورية مصر العربية حسب السن

-

Date :
No.

غير مطالبين بية ←

السن بالسنة	أقل من 12	-12	65 and more	المجموع
عدد الأفراد - التكرار	12587927	23996171	1072082	36656189

* جداول التوزيعات التكرارية المتجمعة

~~جداول التوز~~ مما سبق يتضح لنا أن الجداول التكرارية لمتغير معين تبين عدد المفردات التى تنتمى لكل فئة وقد يكون المطلوب معرفة عدد المفردات التى تقل عن قيمة معينة أو عدد المفردات التى تساوى أو تزيد عن قيمة معينة والجداول التى تمدنا بهذه المعلومات تسمى جداول التوزيعات التكرارية المتجمعة وهى نوعين

1) جدول توزيع تكرارى متجمع صاعد

هو جدول يبين عدد المفردات التى تقل قيمتها عن الحد الأدنى لفئة معينة ويكون جدول التوزيع التكرارى الصاعد من عمودين أحدهما يبين حدود الفئات والآخر يبين التكرارات المتجمعة المناظرة لتلك الحدود

* الجدول التالى يبين التوزيع التكرارى المتجمع الصاعد لأوزان عينة من 60 طالب

حدود الفئات	التكرار المتجمع الصاعد
أقل من 40	0
أقل من 50	5
أقل من 60	13
أقل من 70	23
أقل من 80	38
أقل من 90	51
أقل من 100	56
أقل من 110	60

٣) جدول التوزيع التكراري المتجمع الهابط
هو جدول يبين عدد المفردات التي تساوي أو تزيد قيمتها عن الحد الأدنى لكل فئة ويكون من عمودين: يضا أحدهما يبين الحدود الدنيا للفئات والآخر يبين التكرارات المتجمعة النازلة المناظرة لتلك الحدود

التكرار المتجمع الهابط	حدود الفئات
60	40 فأكثر
55	50 فأكثر
47	60 فأكثر
37	70 فأكثر
22	80 فأكثر
9	90 فأكثر
4	100 فأكثر
0	110 فأكثر

* الجداول التكرارية المزدوجة ← تناولنا فيما سبق دراسة طريقة عمل جداول تكرارية لبيانات متغير واحد وصفي أو كمي وهي تسمى بالجداول البسيطة ولكن يحدث أن تكون هناك رغبة في تكوين جدول تكراري مزدوج لمتغيرين تم جمع بياناتهما لكل مفردة من مجموعة من المفردات وتكون البيانات في هذه الحالة على شكل أزواج مرتبة وكل زوج منها لمفردة واحدة وعلى ذلك غرار جدول التوافق والاقتران

* المثال - البيانات التالية تبين نوع ومؤهل عينة مكونة من 20 موظف في إحدى المؤسسات الإنتاجية

رقم الموظف	النوع	المؤهل
1	ذكر	

مثال 5

البيانات التالية تبين النوع والمؤهل لعينة مكونة من 20 موظف بإحدى المؤسسات الإنتاجية

رقم الموظف	النوع	المؤهل
1	ذكر	جامعي
2	أنثى	أقل من متوسط
3	ذكر	متوسط
4	ذكر	جامعي
5	أنثى	متوسط
6	ذكر	متوسط
7	ذكر	متوسط
8	أنثى	متوسط
9	ذكر	أقل من متوسط
10	ذكر	جامعي
11	أنثى	جامعي
12	ذكر	متوسط
13	أنثى	متوسط
14	ذكر	متوسط
15	ذكر	متوسط
16	أنثى	متوسط
17	ذكر	أقل من متوسط
18	أنثى	جامعي
19	ذكر	جامعي
20	ذكر	متوسط

لتكوين جدول تفريغ مزدوج لهذه البيانات نلاحظ أن المتغير الأول وهو النوع مقسم الى جزئين (ذكر- أنثى) وأن المتغير الثاني وهو المؤهل له ثلاثة أجزاء (أقل من متوسط - متوسط - جامعي) ولذلك نكون جدول تفريغ مزدوج به صفتين وثلاثة أعمدة (جدول توافق) - ولتفريغ بيانات الموظف

الأول (ذكر جامعي) نضع علامة في الخلية الناتجة من تقاطع الصف الأول (ذكر) مع العمود الثالث (جامعي) وهكذا نحصل على جدول تفريغ البيانات

جدول تفريغ بيانات النوع والمؤهل لعينة من 20 موظف

المؤهل / النوع	أقل من متوسط	متوسط	جامعي	المجموع
ذكر	//	~~////~~ //	////	13
انثى	/	////	//	7
المجموع	3	11	6	20

جدول توزيع تكراري(التوافق) للنوع والمؤهل لعينة من 20 موظف

المؤهل / النوع	أقل من متوسط	متوسط	جامعي	المجموع
ذكر	2	7	4	13
انثى	1	4	2	7
المجموع	3	11	6	20

جدول هامشي للمؤهل لعينة من 20 موظف

المؤهل	أقل من متوسط	متوسط	جامعي	المجموع
عدد الموظفين	3	11	6	20

جدول هامشي للنوع لعينة من 20 موظف

النوع	ذكر	انثى	المجموع
العدد	13	7	20

2-1-2) العرض البياني للبيانات الوصفية

البيانات : هي مجموعة من المشاهدات أو الملاحظات المأخوذة أثناء دراسة عينة وقد تكون بيانات رقمية (كمية) مثل أطوال وأوزان مجموعة من الطلاب
أو قد تكون غير رقمية (وصفية) مثل لون البشرة والشعر

المتغيرات العشوائية : هو مقدار له خصائص رقمية (كمية) وغير رقمية وتتغير قيمته من عنصر لآخر من عناصر المجتمع أو العينة .
مثل دراسة ظاهرة الوزن أو الطول أو قياس معدل الذكاء ~~المتغير~~

الثابت : يأخذ رقم واحد ولا يتغير
المتغير : لا يأخذ رقم واحد

المتغيرات العشوائية تنقسم إلى قسمين :
① المتغيرات العشوائية الرقمية : وهي أيضًا تنقسم إلى قسمين :
ⓐ متغير عشوائي رقمي متقطع (منفصل) :
اذا كانت كل القيم التي يأخذها المتغير رقمية وغير قابلة للعد بغض النظر عن أنها نهائية أو غير نهائية .
ⓑ متغير عشوائي رقمي مستمر (متصل) :
اذا كانت كل القيم التي يأخذها المتغير رقمية وغير قابلة للعد مثل أن سرعة السيارة تتراوح ما بين ٨٠ إلى ١٠٠ كم/س .

② المتغيرات العشوائية الغير رقمية .

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

طرق جمع البيانات: هناك طريقتين

① المسح الشامل أو الحصر الشامل وهنا يتم جمع البيانات من جميع عناصر المجتمع

مميزاتها: الدقة العالية والوضوح والتفصيل والمصداقية.

عيوبها: تحتاج وقت وجهد وتكاليف

② العينة: وتعرف بأنها ذلك الجزء الذي يتم اختياره من المجتمع بهدف تعميم نتائجه على المجتمع بشرط أن تكون ممثلة المجتمع أحسن تمثيل.

مميزاتها: توفير الوقت والجهد والتكاليف

عيوبها: أن الاختيار غير الصحيح للعينة يتسبب في الحصول على المعلومات غير صحيحة وأيضا يواجه الباحث صعوبة في اختيار حجم العينة.

أنواع العينات:

① عشوائية بسيطة: هي العينة التي يكون لكل فرد من أفراد المجتمع الاحصائي نفس فرصة للاختيار فيتم اختيار العينة البسيطة

② عشوائية طبقية: وتكونا طبقًا لحجم المجتمع

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

أنواع المجتمع الاحصائي :

① مجتمع احصائي صغير : اذا كان صغير حيث يكون حجمه أقل من أو يساوي ٢٠ مفردة نقوم بجمع هذه المفردات ونقوم باعطاء كل مفردة رقم بحيث يكون كل رقم على بطاقة بحيث تكون البطاقات متشابهة من حيث المظهر ثم نقوم بخلط هذه البطاقات ومن ثم سحب عينة بالحجم الذي نريده

② مجتمع احصائي كبير : نقوم بترتيب مفردات المجتمع من 1 إلى ن حيث (ن هي حجم المجتمع) ونختار من الأرقام
لنفرض أن لدينا مجتمع مكون من ٥٠٠ موظف وأردنا اختيار عينة مقدارها ٢٠ موظف فاننا نعطي الموظفين أرقام متسلسلة من 1 إلى ٥٠٠ ثم نختار منها ٢٠ رقم بطريقة عشوائية مراعاة عدم التكرار .

تابع أنواع العينات :

③ العينة الطبقية : وهي عينة التي يكون فيها المجتمع الاحصائي مقسم إلى طبقات أو فئات قد تكون على أساس جغرافي أو على أساس أخر ويحسب عدد أفراد العينة كما يلي :

$$\text{عدد أفراد العينة} = \frac{\text{عدد أفراد الطبقة}}{\text{عدد أفراد المجتمع الكلي}} \times \text{عدد أفراد العينة المراد اختيارها}$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

مثال مجتمع حجمه ٥٠٠٠ مفردة قسمت إلى الفئات التالية :

١ الفئة أ وتساوي 1000 مفردة
٢ " ب " 2000 "
٣ " جـ " 1500 "
٤ " د " 500 "

اذا أردنا سحب عينة عشوائية منه بحيث يكون عدد مفرداتها يساوي ١٪ من مجموع المفردات للمجتمع بحيث يكون جميع الفئات المجتمع ممثلة في العينة ؟

كيف يتم سحب هذه العينة ؟

الحل

الحجم الكلي للعينة = ١٪ = ٠,٠١ × ٥٠٠٠ = ٥٠ مفردة

حجم العينة المسحوبة من الفئة أ = $\frac{١٠٠٠}{٥٠٠٠}$ × ٥٠ = ١٠ مفردة

" " " " ب = $\frac{٢٠٠٠}{٥٠٠٠}$ × ٥٠ = ٢٠ مفردة

" " " " جـ = $\frac{١٥٠٠}{٥٠٠٠}$ × ٥٠ = ١٥ مفردة

" " " " د = $\frac{٥٠٠}{٥٠٠٠}$ × ٥٠ = ٥ مفردات

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

مصادر جمع البيانات :

① مصدر تاريخي : وهو ما يؤخذ من السجلات المحفوظة مثل سجلات المواليد والوفيات وغيرها

② مصدر ميداني : وهو عبارة عن البيانات المجموعة من أفراد المجتمع الاحصائي كله أو جزء منه سواء بالاتصال المباشر (بطريقة المقابلة) أو غير مباشرة مثل (البريد ، التليفون ، الطريقتين معًا)

ضع علامة (✓) أم (X) :

① العينة الاحصائية هي جزء من المجتمع الاحصائي ( )
② المعلمة تمثل خاصية من خواص المجتمع الاحصائي ( )
③ المجتمع هو الذي يخضع فعليًا للدراسة ( )

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

"مقاييس النزعة المركزية"

نزعة مركزية : الرغبة فى التمركز والتكثف نحو رقم معين .

من المقاييس :

1. الوسط الحسابى
2. الوسيط
3. الرباعين
4. الوسط الهندسى
5. الوسط التوافقى

الوسط الحسابى (المتوسط) (Mean) :

$$\text{الوسط الحسابى} = \frac{\text{مجموع قيم المشاهدات}}{\text{عدد المشاهدات}}$$

أ) اذا كانت البيانات غير مبوبة :

على سبيل المثال : اذا كانت أوزان عينة من ٥ طلاب بالكجم هى :
(٥٥ ، ٦٠ ، ٦٥ ، ٨٠ ، ٧٠) أوجدى الوسط الحسابى

$$\text{الوسط الحسابى} = \frac{٣٣٠}{٥} = ٦٦ \text{ كجم}$$

- اذا كانت X ترمز للظاهرة محل الدراسة وكان لدينا (n) قراءة من قيم الظاهرة ولتكن : $X_1 , X_2 , \ldots\ldots X_n$
فيكون الوسط الحسابى رمزه ($\overline{X}$)

$$\overline{X} = \frac{\Sigma X}{n} = \frac{\text{مجموع قيم } X}{\text{عددهم}}$$

حيث (ΣX) تعبر عن مجموع قيم الظاهرة .

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

ب) اذا كانت البيانات متكررة في جدول تكراري :

اذا كانت لدينا قيم المشاهدات من : $x_1, x_2, \ldots, x_k$

ولل مفردة تتكرر عدد من المرات وتكون تكراراتها المقابلة على الترتيب :

$f_1, f_2, f_3, \ldots f_k$

فان الوسط الحسابي يكون :

$$\bar{X} = \frac{x_1 f_1 + x_2 f_2 + \cdots x_k f_k}{f_1 + f_2 + \cdots f_k} = \frac{\Sigma xf}{\Sigma f}$$

مثال الجدول التالي يبين التوزيع التكراري لعدد الأطفال في عينة مكونة من 100 أسرة مأخوذة من احدى القرى المصرية والمطلوب ايجاد الوسط الحسابي لعدد الأطفال للأسرة .

x عدد الأطفال	1	2	3	4	5	6
f التكرار	4	15	50	25	4	2

X	f	Xf
1	4	4
2	15	30
3	50	150
4	25	100
5	4	20
6	2	12
مجموع	$\Sigma f = 100$	$\Sigma Xf = 316$

Date: / / Subject:

$$\bar{X} = \frac{\sum fx}{\sum f} = \frac{316}{100} = 3.16 \simeq 3$$

أى أن متوسط عدد الأطفال لكل أسرة = ٣ أطفال .

② ايجاد الوسط الحسابى لبيانات مبوبة :

فى البيانات لمتغير متصل المجدولة فى جدول توزيع تكرارى فان البيانات تفقد ذاتيتها بمعنى أنه بدلاً من معرفة كل قيمة بطريقة محددة فإنه يصبح فى الامكان فقط أن نعرف كم قيمة تقع فى كل فئة من فئات الجدول الخاص بالتوزيع التكرارى المعطى له وبسبب عدم معرفة القيم الحقيقية لكل مفردة فى داخل كل فئة من فئات الجدول فإنه لحساب الوسط الحسابى يفترض أن القيم الواقعة فى كل فئة تتركز عند مركز تلك الفئة فإذا كان لدينا جدول توزيع تكرارى له K من الفئات مراكزها هى: $X_1, X_2, \ldots, X_K$ وكانت تكراراتها: $f_1, f_2, \ldots, f_K$ فاننا نفترض أن f_1 القيمة الواقعة فى الفئة الأولى كل منها يساوى X_1 وهكذا .

فان الوسط الحسابى يحسب من العلاقة :

$$\bar{X} = \frac{\sum Xf}{\sum f} = \frac{\text{مجموع حاصل ضرب مركز الفئات × التكرار}}{\text{مجموع التكرارات}}$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

مثال الجدول التالى يبين التوزيع التكرارى لعينة من 50 عامل من العاملين بأحد المصانع حسب راتبهم الشهري بعشرات الجنيهات

المطلوب: حساب الوسط الحسابى للراتب الشهري .

فئات الراتب	30_	40_	50_	60_	70_	مجموع
عدد العمال	2	9	21	11	7	50

الحل

فئات الراتب	مراكز الفئات X	F	XF
30_	35	2	70
40_	45	9	405
50_	55	21	1155
60_	65	11	715
70_	75	7	525
مجموع		$\Sigma f = 50$	$\Sigma XF = 2870$

طول الفئة $= 10 \Leftarrow$ نصف طول الفئة $= 5$

مركز الفئة = الحد الأدنى للفئة + نصف طول الفئة

$= 40 + 5 = 45$ وهكذا

$$\therefore \bar{X} = \frac{\Sigma fx}{\Sigma f} = \frac{2870}{50} = 57,4$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

H.W

التوزيع التكراري النسبي والمئوي للوزن بالكجم لعينة من 60 طالب.

فئات الوزن	التكرار	التكرار النسبي	التكرار المئوي
40 -	5	$\frac{5}{60} = 0{,}083$	8,3%
50 -	8	0,133	13,3%
60 -	10	0,166	16,6%
70 -	15	0,250	25%
80 -	13	0,216	21,6%
90 -	5	0,083	8,3%
100 -	4	0,066	6,6%
المجموع	60	1	100%

٢٠-٢-٢٠٢٤م

١٠ شعبان ١٤٤٥هـ

الموارد الاقتصادية:

وسائل متاحة تستخدم لانتاج سلع وخدمات تشبع احتياجات الفرد ولابد لانتاج السلعة أنها تحتاج لوسائل الانتاج (الموارد الاقتصادية).

ملحوظة:

- المورد الاقتصادى له ثمن، أما المورد الذى ليس له ثمن لا يعد مورد اقتصادى ويطلق عليه مورد غير اقتصادى.

- محور موضوع علم الاقتصاد هو الموارد الاقتصادية

سؤال: لماذا نطلق على المورد أنه اقتصادى؟

لأنه يسهم فى تحديد الدخل

- كلما زاد المورد الاقتصادى زاد الدخل وكلما قل المورد الاقتصادى قل الدخل.

- العوامل التى تؤثر على الموارد من حيث كونه مورد طبيعى أم اقتصادى؟

① المعلومات الاضافية ② المستوى الفنى أو التقنى

③ الندرة النسبية للموارد

أنواع الموارد:

① من حيث أصلها تنقسم إلى:

أ موارد طبيعية

ب موارد بشرية: فكل جهد يبذله الانسان يسهم فى العملية الانتاجية وهى أساس العملية الانتاجية

جـ موارد تكنولوجية: رأس المال والتقنية مثل الآلات والمعدات وهى عبارة عن جهد الانسان فى فترة سابقة

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

H.W

التوزيع التكراري النسبي والمئوي للوزن بالكجم لعينة من 60 طالب .

فئات الوزن	التكرار	التكرار النسبي	التكرار المئوي
40 –	5	$\frac{5}{60} = 0,083$	8,3%
50 –	8	0,133	13,3%
60 –	10	0,166	16,6%
70 –	15	0,250	25%
80 –	13	0,216	21,6%
90 –	5	0,083	8,3%
100 –	4	0,066	6,6%
المجموع	60	1	100%

١٧ شعبان ١٤٤٥هـ "سكشن احصاء" ٢٧-٢-٢٠٢٤م

(١) فى دراسة عن الصحف اليومية الأكثر توزيعًا فى القاهرة تم سؤال ٢٥ شخص عن الصحيفة التى يفضلونها فكانت اجابتهم:

(الأهرام - الأخبار - الأخبار - المصرى اليوم - المصرى اليوم - الأخبار - الأهرام - الأخبار - اليوم السابع - اليوم السابع - المصرى اليوم - اليوم السابع - الأهرام - الوفد - الأهرام - اليوم السابع - الأهرام - الأخبار - الأهرام - المصرى اليوم - الوفد - الأهرام - المصرى اليوم - المصرى اليوم - المصرى اليوم - اليوم السابع)

المطلوب: وضع هذه البيانات فى جدول تكرارى.

الصحف	العلامات	التكرار
الأهرام	~~||||~~ ||	7
الأخبار	~~||||~~	5
المصرى اليوم	~~||||~~ |	6
اليوم السابع	~~||||~~	5
الوفد	||	2
المجموع		25

Date: / / Subject:

الجدول التكراري:

الصحف	الأهرام	الأخبار	المصري اليوم	اليوم السابع	الوفد
التكرار	7	5	6	5	2

③ فيما يلي الراتب الشهري 100 جنيه لعينة حجمها 27 موظف تم اختيارهم من بين موظفي احدى البيانات.

(182 - 211 - 187 - 188 - 203 - 200 - 195 - 185 -
205 - 206 - 205 - 115 - 191 - 203 - 199 - 218 -
202 - 197 - 214 - 195 - 193 - 202 - 208 - 196 -
200 - 203 - 204)

المطلوب: كون جدول التوزيع التكراري.

الحل

المدى $= 218 - 115 = 103$

طول الفئة $= \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{103}{6} = 17,2 \simeq 20$

الفئات	العلامات	التكرارات
115 _	/	1
135 _	_	0
155 _	_	0
175 _	$\cancel{||||}$ /	6
195 _	$\cancel{||||}$ $\cancel{||||}$ $\cancel{||||}$ ||||	19
215 _	/	1
المجموع		27

٣ البيانات التالية تمثل دخول 30 شخص لعينة ما:

(65 _ 79 _ 74 _ 81 _ 66 _ 62 _ 80 _ 79 _ 95 _ 81 _ 95 _
83 _ 64 _ 88 _ 68 _ 91 _ 78 _ 81 _ 60 _ 85 _ 88
84 _ 83 _ 63 _ 65 _ 72 _ 75 _ 73 _ 70 _ 80)

الحل

المدى = 95 _ 60 = 35

طول الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}} = \frac{35}{8} = 4,4 \simeq 5$

جدول التفريغ :

التكرار	العلامات	الفئات
4	||||	60 -
4	||||	65 -
4	||||	70 -
4	||||	75 -
8	~~||||~~ |||	80 -
3	|||	85 -
1	|	90 -
2	||	95 -
30		المجموع

الجدول التكراري المتجمع الصاعد :

التكرارات المتجمع الصاعد	الحدود العليا للفئات
0	أقل من 60
4	65 " "
8	70 " "
12	75 " "
16	80 " "
24	85 " "
27	90 " "
28	95 " "
30	100 " "

الجدول التكراري المتجمع الهابط.

حدود الفئات	التكرار المتجمع الهابط
أكثر من 60	30
65 " "	26
70 " "	22
75 " "	18
80 " "	14
85 " "	6
90 " "	3
95 " "	2
100 " "	0

H.W اذا كان لديك البيانات التالية عن الدخل اليومي لمجموعة من العاملين.

الفئات	40 -	45 -	50 -	55 - 60	المجموع
عدد العمال	10	25	20	10	65

المطلوب: تكوين الجدول التكراري المتجمع الصاعد والمتجمع الهابط

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

الانحراف الربيعي للبيانات المبوبة :

$$Q_1 = l_1 + \left[\frac{\frac{\Sigma f}{4} - f_1}{f_{Q_1}}\right] \cdot C$$

$l_1 \Leftarrow$ الحد الأدنى لفئة الربيعي الأول

$\Sigma \frac{f}{4} \Leftarrow$ ترتيب الربيعي الأول

$f_1 \Leftarrow$ التكرار المتجمع الصاعد للربيعي الأول

$f \Leftarrow$ تكرار الفئة للربيعي الأول من الجدول الأصلي

$C \Leftarrow$ طول الفئة .

$$Q_3 = l_3 + \left[\frac{\frac{3\Sigma f}{4} - f_3}{f_{Q_3}}\right] \cdot C$$

مثال

الجدول التالي يمثل رواتب 100 موظف .

فئات الرواتب	70-	80-	90-	100-	110-	120-	130-	مجموع
عدد الموظفين	5	7	21	33	18	13	3	100

المطلوب : أوجدي نصف المدى الربيعي لرواتب الموظفين

الحل

الحدود العليا للفئات	التكرار المتجمع الصاعد
أقل من 70	0
80 " "	5
90 " "	12
100 " "	33
110 " "	66
120 " "	84
130 " "	97
140 " "	100

$C = 10 \qquad L_1 = 90$

$\frac{\Sigma F}{4} = \frac{100}{4} = 25 \qquad , f_1 = 12$

$F_{Q_1} = 21$

$Q_1 = 90 + \left[\frac{25 - 12}{21}\right] \times 10 = 96,19 \Rightarrow 1$

$\frac{3 \Sigma F}{4} = 75 \quad , \quad L_3 = 110$

$F_3 = 66 \quad , F_{Q_3} = 18 \quad , C = 10$

$Q_3 = 110 + \left[\frac{75 - 66}{18}\right] \cdot 10 = 115 \rightarrow 2$

من ١، ٢

$\therefore Q = \frac{Q_3 - Q_1}{2} = \frac{115 - 96,19}{2} = 9,41$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

من أهم مميزات الانحراف الربيعي :
أنه يفضل استخدامه كمقياس في حالة وجود قيم شاذه
كما أنه سهل وبسيط

الانحراف المتوسط (M.D) :
هو مقياس من مقاييس التشتت يقيس بدقة الانحراف عن الوسط الحسابي .

$$M.D = \frac{\Sigma |x - \bar{x}|}{n}$$

أولاً : في حالة البيانات الغير مبوبة :
أوجدى الانحراف المتوسط للمشاهدات الآتية :

1, 3, 5, 9, 12

$$\bar{x} = \frac{1+3+5+9+12}{5} = \frac{30}{5} = 6$$

x	$x - \bar{x}$	$\lvert x - \bar{x} \rvert$
1	$1 - 6 = -5$	5
3	$3 - 6 = -3$	3
5	$5 - 6 = -1$	1
9	$9 - 6 = 3$	3
12	$12 - 6 = 6$	6

$$\Longrightarrow \Sigma |x - \bar{x}| = 18$$

$$\therefore M.D = \frac{18}{5} = 3{,}6$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

ثانيًا: في حالة البيانات المبوبة :

$$M.D = \frac{\Sigma F |X - \bar{X}|}{\Sigma F}$$

مثال

أوجدي الانحراف المتوسط للقيم التالية التي تمثل درجات ١٠٠ طالبة في مادة الاحصاء

الدرجات	40-	50-	60-	70-	80	المجموع
التكرار	10	20	50	5	15	100

الحل

مركز الفئة

الدرجات	X	التكرار	XF	$\|X - \bar{X}\|$	$F\|X - \bar{X}\|$
40-	45	10	450	19,5	195
50-	55	20	1100	9,5	190
60-	65	50	3250	0,5	25
70-	75	5	375	10,5	52,5
80-	85	15	1275	20,5	307,5
مجموع		100	6450		770

$$\bar{X} = \frac{\Sigma XF}{\Sigma F} = \frac{6450}{100} = 64,5$$

$$\therefore M.D = \frac{770}{100} = 7,7$$

Date: / / Subject:

خصائص الوسط الحسابى :

① مجموعات انحرافات قيم المشاهدات عن وسطها الحسابى = صفر

فمثلاً :

20 , 27 , 15 , 21 , 17

الحل

$$\bar{X} = \frac{\Sigma X}{n} = \frac{100}{5} = 20$$

$$\therefore \Sigma (X - \bar{X}) = (20-20) + (27-20) + (15-20) + (21-20) + (17-20) = 0$$

② يأخذ كل قيم المشاهدات فى الاعتبار .

③ أكثر المقاييس استخداماً فى التحليل الاحصائى

عيوبه :

① يتأثر بالقيم الشاذة (المتطرفة) وهى القيم الكبيرة جداً أو القيم الصغيرة جداً

② لا يمكن حسابه بيانياً ولا يمكن حسابه للبيانات الوصفية

③ يصعب حسابه للتوزيعات المفتوحة

الوسيط

هو القيمة التى تقسم مجموعة المشاهدات المرتبطة (تصاعدياً أو تنازلياً) إلى قسمين متساويين عدد المشاهدات على يمين الوسيط يساوى عدد المشاهدات على يسار الوسيط

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

ايجاد الوسيط للبيانات غير المبوبة :

في هذه الحالة نرتب البيانات ترتيبًا تصاعديًا أو تنازليًا فاذا كان عدد البيانات فردي فان الوسيط في هذه الحالة يكون المشاهدات التي في المنتصف وسوف نرمز له بالرمز (med) .

مثال

أوجدي الوسيط للبيانات التالية :

$6, 12, 9, 11, 14, 15, 16, 8, 10$

الحل

ترتيبهم : $6, 8, 9, 10, 11, 12, 14, 15, 16$

$n = 9$

$med = 11$

- احسبي الوسيط للقياسات التالية :

$5, 12, 7, 11, 8, 7, 10, 13$

الحل

ترتيبهم : $5, 7, 7, 8, 10, 11, 12, 13$

$n = 8$

$$(med) = \frac{8+10}{2} = 9$$

أما اذا كانت عدد البيانات كبيرة :

نفرض أن عددها (n) فاذا كان فردي فان رتبة الوسيط $= \frac{n+1}{2}$

أما اذا كان زوجي فان الوسيط هو متوسط الحدين اللذين رتبتهما : $\left(\frac{n}{2}, \frac{n}{2}+1\right)$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

إيجاد الوسيط لبيانات مبوبة :

① ترتيب الوسيط = $\frac{\Sigma f}{2}$ (أى نصف مجموع التكرارات).

ويمكن إيجاد الوسيط حسابياً أومن الرسم :

ولإيجاد الوسيط حسابياً نتبع الخطوات التالية :

② نكون جدول التوزيع التكرارى المتجمع الصاعد

③ نحدد الفئة الوسيطية وهى الفئة التى يكون التكرار المتجمع الصاعد المقابل للحد الأعلى لها لايقل عن ($\frac{\Sigma f}{2}$)

بينما يكون تكرار المتجمع الصاعد المقابل للحد الأعلى للفئة السابقة لها يكون أقل من ($\frac{\Sigma f}{2}$).

فيكون الوسيط هو احدى قيم هذه الفئة.

ونحسب الوسيط من العلاقة الآتية :

$$med = L - \left[\frac{\frac{\Sigma f}{2} - \Sigma f_1}{F_{md}} \right] \times C$$

خصائص الوسيط :

① يمكن حسابه لأى مجموعة من القيم وبعدد قليل من العمليات الحسابية

② لايتأثر بالقيم المتطرفة

③ يمكن إيجاده بيانياً كما يمكن إيجاده فى حالة التوزيعات المفتوحة

عيوبه :

① لايعتمد على جميع المشاهدات وقد تتغير مشاهدة أو أكثر دون التأثير على قيم الوسيط

② لايدخل فى التحاليل الاحصائية بعكس الوسط

$L \Leftarrow$ الحد الأدنى لفئة الوسيط

$\frac{\Sigma F}{2} \Leftarrow$ ترتيب الوسيط

$\Sigma F_1 \Leftarrow$ التكرار المتجمع الصاعد السابق لترتيب الوسيط

$F_{md} \Leftarrow$ تكرار فئة الوسيط في الجدول الأصلي.

$C \Leftarrow$ طول الفئة.

مثال

البيانات التالية تمثل الأجور الأسبوعية -- اجنيه لقسمين -- ١٠٠ عامل.

فئات الراتب	60-	70-	80-	90-	100-	المجموع
عدد العمال	6	12	47	25	10	100

أوجد الوسيط:

فئات الراتب	عدد العمال
60 -	6
70 -	12
80 -	47
90 -	25
100 -	10
مجموع	100

الحدود العليا للفئات	التكرار المتجمع الصاعد
أقل من 60	0
70 " "	6
80 " "	18
90 " "	65
100 " "	90
110 " "	100

$C = 10$ ، $\frac{\Sigma F}{2} = 50$ $L_1 = 80$ ، $\Sigma F_1 = 18$

$F_{md} = 47$

$$\therefore med = 80 - \left[\frac{50-18}{47}\right] \times 10 = 86{,}8$$

Date: / / Subject:

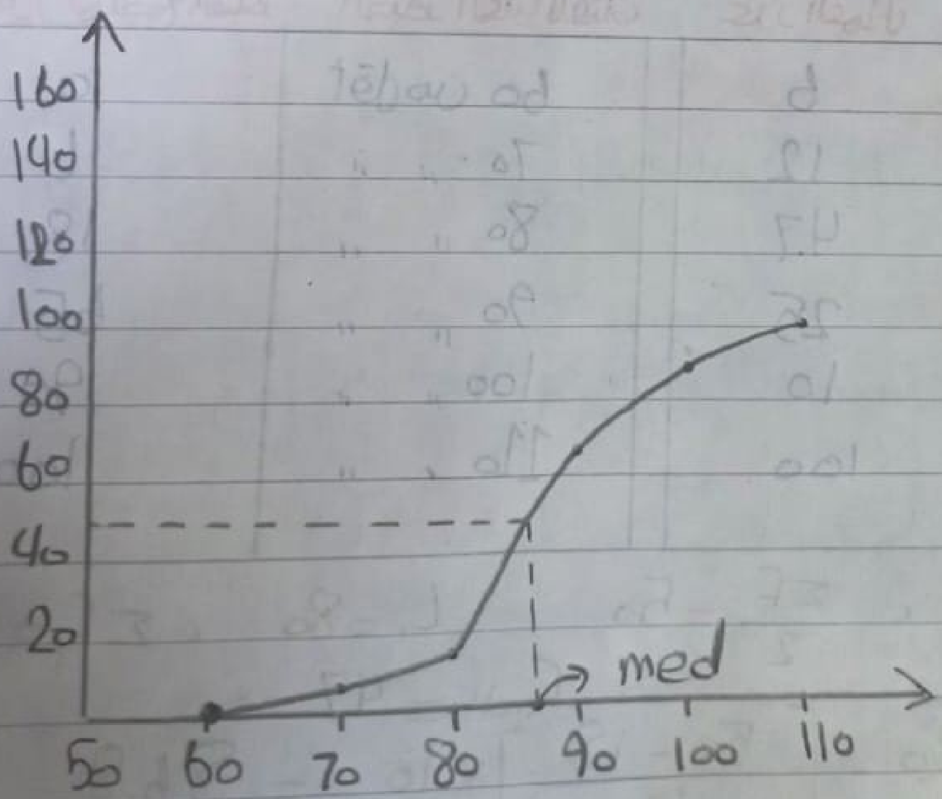
لإيجاد الوسيط بالرسم :

① نحسب ترتيب الوسيط = $\frac{\Sigma f}{2}$ ومن هذه النقطة نرسم نعين ترتيب الوسيط على المحور الرأسى ومن هذه النقطة نرسم موازى للمحور الأفقى يتقاطع مع المنحنى فى نقطة يكون الاحداثى السينى لهذه النقطة هى قيم الوسيط حيث يمكن ايجاد الوسط من المنحنى التكرارى المتجمع الهابط بطريقة مماثلة واذا رسمنا كلاً من المنحنى الصاعد والهابط فى شكل واحد وبنفس المحاور فان الاحداثى الأفقى لنقطة تقاطعهم يكون ممثلاً لقيمة الوسيط والاحداثى عند $\frac{\Sigma f}{2}$

مثال أوجد الوسيط للأجور فى المثال السابق من المنحنى التكرارى المتجمع الصاعد مرة والهابط مرة أخرى

الحل

أولاً من الصاعد :



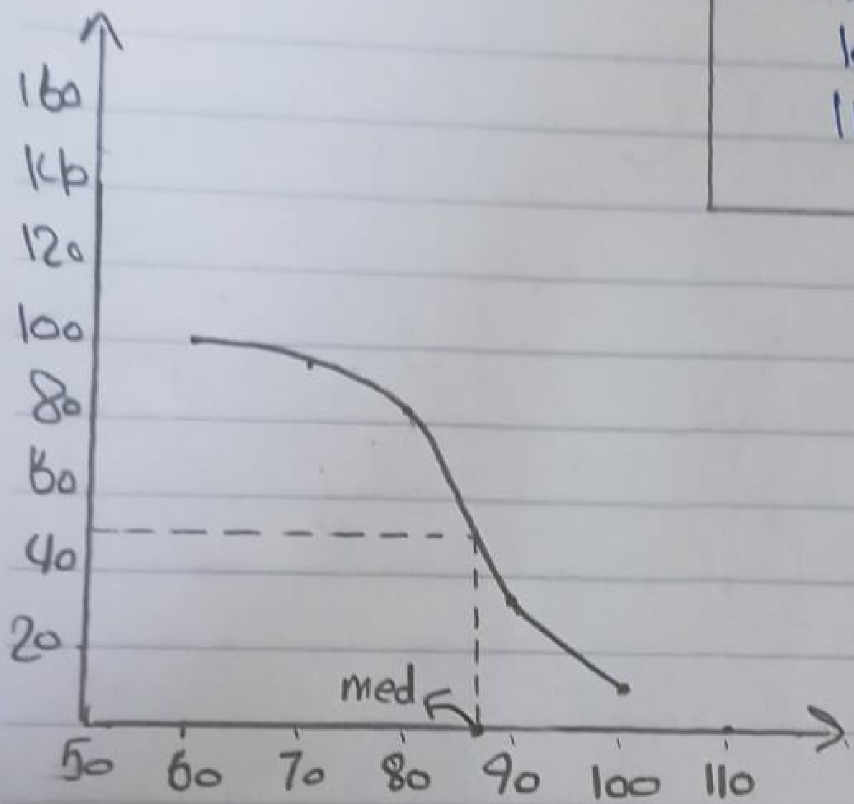
$\Rightarrow med = 86.8$

Elattar

Date: / / Subject:

Sat | Sun | Mon | Tue | Wed | Thu | Fri

الحدود الدنيا للفئات	التكرار المتجمع الهابط
أكثر من 60	100
70	94
80	82
90	35
100	10
110	0



$\Rightarrow$ med = 86,8

6

"الانحراف المعياري"

تعريفه:

يعرف بأنه الجذر التربيعي لمجموع مربعات انحرافات قيم المشاهدات عن وسطها الحسابي مقسوماً على حجم العينة ويرمز له بالرمز (S) وهو من أهم مقاييس التشتت وأكثرها استخداماً في التحليل الاحصائي.

① في حالة البيانات غير المبوبة.

$$S = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n}} \quad \text{أو} \quad \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2}$$

مثال: أوجدي الانحراف المعياري للقيم الآتية:

5, 14, 11, 7, 3

الحل

$$\bar{x} = \frac{\Sigma x}{n} = \frac{5+14+11+7+3}{5} = 8$$

x	$x-\bar{x}$	$(x-\bar{x})^2$
5	-3	9
14	6	36
11	3	9
7	-1	1
3	-5	25
$\Sigma x = 40$		$\Sigma (x-\bar{x})^2 = 80$

$$\therefore S = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n}}$$

$$= \sqrt{\frac{80}{5}} = 4$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

٢ في حالة البيانات المبوبة :

هناك طريقتين لإيجاد الانحراف المعياري في حالة البيانات المبوبة (الطريقة المباشرة ، طريقة الانحراف المتوسط).

$$S = \sqrt{\frac{\Sigma F x^2}{\Sigma F} - \bar{x}^2}$$

مثال الجدول التالى بين التوزيع التكراري لدرجات عينة من ١٠٠ طالبة في مادة مبادئ الاحصاء :

فئات الدرجات	0-	10-	20-	30-	40-	
التكرار	2	5	27	47	19	100

احسبي الانحراف المعياري.

	(مركز الفئة x)	f	xf	x^2	Fx^2
0	5	2	10	25	50
10	15	5	75	225	1,125
20	25	27	675	625	16,875
30	35	47	1645	1,225	57,575
40	45	19	855	2,025	38,475
		$\Sigma f = 100$	$\Sigma fx = 3260$		$\Sigma Fx^2 = 114,100$

$$\bar{x} = \frac{\Sigma Fx}{\Sigma F} = \frac{3260}{100} = 32,6$$

$$(\bar{x})^2 = (32,6)^2 = 1062,76$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

$$\therefore S = \sqrt{\frac{\Sigma f x^2}{\Sigma f} - \bar{x}^2}$$

$$= \sqrt{\frac{114100}{100} - 1062,76} = 8,85$$

معامل الاختلاف /معامل التغير

هو أحد مقاييس التشتت وليس له وحدة قياس ويستخدم لمقارنة التشتت النسبى أو مقياس درجة تجانس البيانات ونرمز له بالرمز $\boxed{C.V}$

وهناك عدة طرق لحساب معامل الاختلاف أهمها :

(1) معامل الاختلاف المعيارى (2) معامل الاختلاف الربيعى

معامل الاختلاف المعيارى $\Leftarrow$ $C.V_{I} = \frac{S}{\bar{x}} \cdot 100$

معامل الاختلاف الربيعى $\Leftarrow$ $C.V_{(II)} = \frac{Q_3 - Q_1}{Q_1 + Q_3} \cdot 100$

مثال

اذا كان لدينا المعلومات التالية عن سلعتين فأيهما أجود

البيانات	السلعة I	السلعة II
n	50	50
$\bar{x}$	52	48
S	8	7

Date: / / Subject:

Sat | Sun | Mon | Tue | Wed | Thu | Fri

الحل

نقوم بحساب معامل الاختلاف السلعتين :

$$C.V_{I} = \frac{S}{\bar{X}} \cdot 100 = \frac{8}{52} \times 100 = 15{,}4$$

$$C.V_{II} = \frac{S}{\bar{X}} \cdot 100 = \frac{7}{48} \times 100 = 14{,}6$$

ونلاحظ أنه كلما قل معامل الاختلاف زادت جودة السلعة.

∴ تكون السلعة الثانية هي الأجود

مثال اذا كان لدينا البيانات التالية عن العائد الاستثماري خلال السنوات الخمس الماضية لشركتين كما يلي :

H.W

الشركة A	الشركة B
10	12
2	5

- أي الشركتين أكثر استثماراً ؟

"المنوال"

تعريفه: هو المشاهدة الأكثر تكراراً في مجموعة المشاهدات.

- إيجاد المنوال لبيانات غير مبوبة:

عند حساب المنوال للبيانات غير المبوبة ننظر إلى البيانات فإذا وجدنا أن هناك مشاهدات مكررة أكثر من غيرها فيكون المنوال هو تلك المشاهدات
وإذا كان هناك مشاهدتين أو أكثر لهما نفس التكرار فإن المنوال في هذه الحالة لا يكون وحيداً.

مثال أوجد المنوال للبيانات التالية:

(6, 7, 5, 7, 10, 7, 8, 9, 8, 7)

الحل

بالتدقيق في البيانات نجد أن العدد 7 مكرر أربع مرات والعدد 8 ثلاث مرات
أي أن العدد 7 مكرر أكثر من غيره ∴ المنوال = 7

مثال أوجد المنوال للبيانات التالية:

(مقبول - جيد - مقبول - جيد - جيد - راسب - جيد - ممتاز - جيد)

المنوال = جيد

مثال أوجدى المنوال للبيانات التالية:
(5, 7, 9, 11, 5, 10, 5, 10, 5, 10, 12, 10)

المنوال = 10، 5
حيث أن العدد 10 مكرر ثلاث مرات والعدد 5 مكرر ثلاث مرات
∴ المجموعة لها منوالين هما: 10، 5

مثال أوجدى المنوال للبيانات التالية:
(6, 12, 7, 8, 9, 11, 14)
حيث أن مجموعة المشاهدات لا يوجد بها أى مشاهدة مكررة
فان المجموعة خالية من المنوال!!

ثانيًا: ايجاد المنوال للبيانات المبوبة:
كما عرفنا سابقًا أن البيانات المبوبة قد تكون بيانات متقطعة
أو بيانات مستمرة فاذا كانت البيانات متقطعة فان المنوال
هو المشاهدة الأكثر تكرارًا.

مثال البيانات التالية تمثل تقديرات مجموعة من الطلاب فى
احدى المقررات

التقدير	ممتاز	جيد جدًا	جيد	مقبول	راسب
عدد الطلاب	2	5	10	15	4

نلاحظ أن تقدير مقبول أكثر تكرارًا من البيانات الأخرى
∴ المنوال = مقبول.

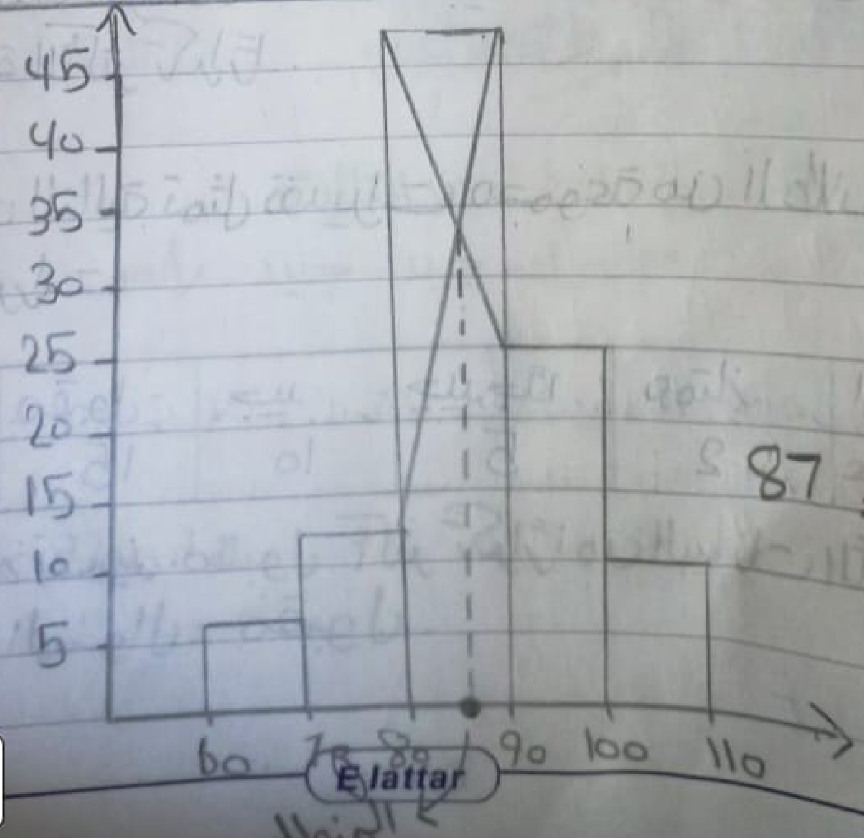
إيجاد المنوال بالرسم من المدرج التكراري :

المنوال هو القيمة الأكثر تكراراً من غيرها فإن المنوال يقع في الفئة المنوالية.

ولإيجاد المنوال من المدرج التكراري نصل نقطة الرأس الأيمن العلوي للمستطيل الواقع على يساره بخط مستقيم وكذلك نقطة الرأس الأيسر العلوي للمستطيل المرسوم فوق الفئة المنوالية بنقطة الرأس العلوي بالمستطيل الواقع على يمينه بخط مستقيم ثم نسقط عموداً من نقطة تقاطع هذين الخطين المستقيمين على المحور الأفقي فيقابله عند النقطة الممثلة للمنوال.

مثال أوجد المنوال للأجور الأسبوعية بالجنيه لعينة من 100 عامل للبيانات التالية :

فئات الأجور	-60	-70	-80	-90	-100	المجموع
عدد العمال	6	12	47	25	10	100



المنوال = 87

الفئة المنوالية: هي الفئة التي يقابلها أكبر تكرار (80 -)
وبرسم مستطيلات المدرج التكراري الممثلة لهذه الفئة والفئة السابقة والفئة اللاحقة بالرسم نجد أن المنوال = 87

إيجاد المنوال بالحساب (طريقة الرافعة).

من الجدول التوزيع التكراري يمكن تحديد بداية الفئة المنوالية وطولها (C) وتكرار الفئة السابقة لها (F_1) وتكرار الفئة التالية لها (F_2) وباستخدام هذه القيم يحسب المنوال من المعادلة الآتية:

$$mode = L_1 - \left(\frac{F_1}{F_1 + F_2}\right) \cdot C$$

$L_1 \Leftarrow$ الحد الأدنى للفئة المنوالية.

وبتطبيق الصيغة السابقة على المثال السابق لحساب المنوال

$$mode = 80 - \left(\frac{25}{12+25}\right) \cdot 10 = 86,67$$

عيوب هذه الطريقة:

يؤخذ على هذه الطريقة عدم دقتها لأنها تهمل أكبر تكرار في التوزيع وهو تكرار الفئة المنوالية

طريقة الفروق (طريقة بيرسون):

وتكون هذه الطريقة على أساس تلافي العيب الموجود في طريقة الرافعة وهو اهمال تكرار الفئة المنوالية عند حساب قيمة المنوال

وعليه فان قيمة المنوال تحسب من العلاقة التالية :

$$mode = L_1 + \left(\frac{f - f_1}{2f - f_1 - f_2} \right) . C$$

وبتطبيق تلك الصيغة على المثال السابق :

$$mode = 80 + \left(\frac{47 - 12}{2(47) - 12 - 25} \right) . 10 = 86{,}14$$

مميزات المنوال :

① مقياس سهل الحساب ولا يتأثر بالقيم المتطرفة

② مقياس يمكن إيجاده بيانيًا

③ يمكن إيجاده لجميع البيانات الوصفية والتوزيعات المفتوحة .

عيوبه :

① لا يأخذ كل القيم في الاعتبار

② قد يكون لبعض البيانات أكثر من منوال أو خالية من المنوال

③ يصعب التعامل معه احصائيًا .

الارتباط والانحدار .

أولاً : معامل الارتباط الخطي لبيرسون :

يستخدم معامل الارتباط لبيرسون والذى سوف يرمز له بالرمز (r_p)

ولقياس قوة الارتباط بين المتغيرين يعطى بالعلاقة الآتية :

$$r_p = \frac{n \sum xy - \sum x \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \cdot \sqrt{n \sum y^2 - (\sum y)^2}}$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

Σx : مجموع قيم x ، Σy : مجموع قيم y

Σxy : مجموع حاصل ضرب قيم x بالقيم المناظرة لقيم y

Σx^2 : مجموع مربعات قيم x

Σy^2 : مجموع مربعات قيم y .

مثال احسبي معامل الارتباط لبيرسون للقيم التالية :

x	10	11	14	11	13	15
y	21	23	29	23	27	31

الحل

x	y	xy	x^2	y^2
10	21	210	100	441
11	23	253	121	529
14	29	406	196	841
11	23	253	121	529
13	27	351	169	729
15	31	465	125	961
$\Sigma x = 74$	$\Sigma y = 154$	$\Sigma xy = 1938$	$\Sigma x^2 = 932$	$\Sigma y^2 = 4030$

$n = 6$

$$r_p = \frac{6(1938) - (74)(154)}{\sqrt{6(932)^2 - (74)^2} \cdot \sqrt{6(4030)^2 - (54)^2}} = 1$$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

خصائصه: $-1 \leq r_p \leq 1$

r_p	التعبير
0	لايوجد ارتباط بين متغيرين
أقل من 0,3	لايوجد ارتباط بين متغيرين
0,3 - 0,5	يوجد ارتباط ضعيف
0,5 - 0,7	ارتباط متوسط
0,7 - 0,9	يوجد ارتباط قوى
0,9 - 1	ارتباط قوى جداً
1	ارتباط تام

ثانياً: معامل ارتباط سبيرمان.

وقد عرفنا كيف نقيس قوة الارتباط بين متغيرين اذا كان المتغيرين يمثلان بيانات كمية باستخدام عامل ارتباط بيرسون وهذا العامل لايمكن حسابه للبيانات الوصفية

لذلك نشأت الحاجة للايجاد مقياس يصلح للبيانات الوصفية وهذا المقياس هو معامل الارتباط للسبيرمان ويطلق عليه احياناً (معامل ارتباط الرتب).

Elattar

مثال أوجدي رتب المشاهدات الآتية :

X	6	9	2	4	7	8

الحل

X	6	9	2	4	7	8
الرتب	3	6	1	2	4	5

مثال أوجدي رتب المشاهدات الآتية :

X	12	9	4	8	9	10	12	14
الرتب	6,5	3,5	1	2	3,5	5	6,5	8

مثال أوجدي رتب المشاهدات الآتية :

التقديرات	جيد	مقبول	راسب	مقبول	ممتاز
الرتب	4	2,5	1	2,5	5

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

" الأرقام القياسية وتطبيقاتها "

- هناك الكثير من العوامل التي دفعت المسئولين لبحث واستقصاء مقدار التغير في الأسعار ونفقات المعيشة
اما ليتمكنوا من تحديد الزيادة المناسبة في الرواتب والأجور التي تتفق مع الزيادة التي تطرأ على أسعار السلع وتكاليف المعيشة أو لدراسة كيفية معالجة هذه الزيادة.

- ولذلك نشأت الحاجة إلى إيجاد مقاييس احصائية تعبر بصورة واضحة ودقيقة عن مقدار هذه التغيرات من حيث الزيادة أو النقص في الأسعار أو في الكميات المنتجة أو المستهلكة أو قيمة الصادرات أو الواردات
وهذه النسب أو المقاييس تسمى بالأرقام القياسية.

تعريف الرقم القياسي:

يعرف بأنه رقم نسبي يقيس التغير في ظاهرة واحدة أو أكثر من ظاهرة من وقت لآخر أو من مكان لآخر.

تركيب الرقم القياسي:

عند تركيب الرقم القياسي للأسعار يجب تحديد طبيعة السلع الداخلة في فترتي المقارنة والأساس
من حيث نوع السلعة أو طريقة البيع إلخ ---
وحتى يكون الرقم القياسي معبرًا عن التغير في الظاهرة يجب أن تتصف فترة الأساس بالآتي:

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

① أن تكون قريبة نسبيًا من فترة المقارنة
② أن تكون فترة عادية تتميز بالاستقرار

ملحوظات :

- يرمز إلى سنة الأساس ⇐ 0
- " " " المقارنة ⇐ 1
- " " الكمية ⇐ Q
- " " السعر ⇐ P
- " " السعر سنة الأساس ⇐ P_0
- " " " " المقارنة ⇐ P_1
- " " الكمية سنة الأساس ⇐ Q_0
- " " " " المقارنة ⇐ Q_1

أنواع الأرقام القياسية :

① أرقام قياسية بسيطة ② أرقام قياسية مرجحة

- الأرقام القياسية البسيطة :

① المناسيب
② الرقم القياسي التجميعي البسيط
③ " " النسبي البسيط

① المناسيب :
هو من أبسط صور الأرقام القياسية فإذا تم نسبة قيمة ظاهرة في فترة المقارنة إلى قيمتها في فترة الأساس نحصل على ما يسمى بالمنسوب .

- منسوب السعر للسلعة = $\frac{\text{سعر السلعة x فترة المقارنة}}{\text{سعر السلعة x فترة الأساس}} \times 100$

- منسوب الكمية للسلعة = $\frac{\text{كمية السلعة x فترة المقارنة}}{\text{كمية السلعة x فترة الأساس}} \times 100$

- منسوب القيمة للسلعة = $\frac{\text{قيمة السلعة x فترة المقارنة}}{\text{قيمة السلعة x فترة الأساس}} \times 100$

مثال
الجدول التالى يبين أسعار البيع فى بلد ما الكميات المنتجة من القمح والأرز والشعير فى عامى 1980 ، 1990 .

السلعة	P_0 1980	P_1 1990	Q_0 1998	Q_1 1990
قمح	3	7	20	100
أرز	3,5	6	10	90
شعير	2	4	3	10

$\Sigma P_0 = 8,5 \quad , \Sigma P_1 = 17$

المطلوب : أوجدى المنسوب السعرى لكل سلعة .
② منسوب الكمية لكل سلعة

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

٣ احسبي الرقم القياسي التجميعي البسيط.

٤ " " " النسبي.

الحل

منسوب السعر لسلعة القمح:

$$= \frac{P_1}{P_0} \times 100 = \frac{7}{3} \times 100 = 233,20$$

منسوب الأرز $= \frac{6}{3,5} \times 100 = 171,43$

منسوب الشعير $= \frac{4}{2} \times 100 = 200$

منسوب الكمية للقمح $= \frac{Q_1}{Q_0} = \frac{100}{20} \times 100 = 500$

للأرز " " $= \frac{90}{10} \times 100 = 900$

للشعير " " $= \frac{10}{3} \times 100 = 333,3$

- حيث أن سنة الأساس دائماً تساوي 100 في المقارنة ولذلك بالنسبة لمنسوب سعر الكيلوجرام من القمح زاد بنسبة 133,2 %

- منسوب سعر الكيلوجرام من الأرز زاد بنسبة 71,43 %

- " " " " الشعير " " 100 %

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

② الرقم القياسي التجميعي البسيط

هو مجموع أسعار السلع في سنة المقارنة مقسومة على مجموع أسعار السلع في سنة الأساس ثم ضرب ناتج القسمة في ١٠٠.

$$\text{الرقم القياسي التجميعي البسيط} = \frac{\Sigma P_1}{\Sigma P_0} \times 100$$

اجابة المطلوب الثالث في المثال السابق :

$$\text{الرقم القياسي التجميعي البسيط} = \frac{17}{8,5} \times 100 = 200$$

③ الرقم القياسي النسبي البسيط

ويعرف بأنه متوسط مناسيب السعر لمجموعة من السلع وضرب الناتج x ١٠٠

$$\text{الرقم القياسي النسبي البسيط} = \frac{1}{N}\left(\Sigma \frac{P_1}{P_0}\right) \times 100$$

حيث (N) هو عدد السلع الداخلة في تركيب الرقم القياسي .

اجابة المطلوب الرابع في المثال السابق:

$$\text{الرقم القياسي النسبي البسيط} = \frac{1}{3}\left(\frac{7}{3} + \frac{6}{3,5} + \frac{4}{2}\right) \times 100$$

$$= 201,59$$

Date: / / Subject:

Sat | Sun | Mon | Tue | Wed | Thu | Fri

مثال

الجدول التالى يبين أسعار البيع فى بلد ما الكميات المنتجة من القمح والأرز والشعير والذرة فى عامى 2000 ، 2024 .

السلع	P_0 2000	P_1 2024
القمح	10	30
الأرز	15	35
الشعير	20	30
الذرة	10	20

المطلوب:

① المنسوب السعرى لكل سلعة

② احسبى الرقم القياسى التجميعى البسيط

③ " " " النسبى البسيط

Date ١٤/ ٤/ ٢٠٢٤ Subject:

Sat Sun Mon Tue Wed Thu Fri

" الأرقام القياسية المرجحة "

هى أربعة أرقام:

① الرقم القياسى التجميعى والنسبى المرجح للاسبير .

استخدم لاسبير الكميات المنتجة والمستهلكة فى سنة الأساس من السلع الداخلة فى تركيب الرقم القياسى كوزن لكل سلع

حيث أن:

$$\text{الرقم القياسى التجميعى المرجح للاسبير} = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100$$

$$\text{الرقم القياسى النسبى المرجح للاسبير} = \frac{\sum \frac{P_1}{P_0} \times Q_0}{\sum Q_0} \times 100$$

مثال الجدول التالى يمثل أسعار وكميات أربع سلع فى عامى : 1995 ، 1997 باعتبار أن سنة الأساس هى 1995

المطلوب:

إيجاد الأرقام القياسية المرجحة الأربعة .

السلع	P_0 1995	P_1 1997	Q_0 1995	Q_1 1997
A	6	24	18	36
B	16	32	14	28
C	26	18	10	20
D	8	10	12	14
المجموع	66	84	54	98

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

الحل

P_1Q_0	P_0Q_0	$\frac{P_1}{P_0}$	$\frac{P_1}{P_0}\cdot Q_0$
432	108	4	72
448	224	2	28
180	260	0,69	6,9
120	96	1,25	15
$\Sigma P_1Q_0 = 1180$	$\Sigma P_0Q_0 = 688$		$\Sigma \frac{P_1}{P_0}Q_0 = 121,9$

الرقم القياسي التجميعي المرجح للاسبير $= \frac{1180}{688} \times 100 = 171,5\ \%$

الرقم القياسي النسبي المرجح للاسبير $= \frac{121,9}{54} \times 100 = 225,7\ \%$

② الرقم القياسي التجميعي والنسبي المرجح لباش.
استخدم باش الكميات المنتجة أو المستهلكة في سنة المقارنة من السلع الداخلة في تركيب الرقم القياسي كوزن لكل سلع
حيث أن:

الرقم القياسي التجميعي المرجح لباش $= \frac{\Sigma P_1Q_1}{\Sigma P_0Q_1} \times 100$

الرقم القياسي النسبي المرجح لباش $= \frac{\Sigma \frac{P_1}{P_0}Q_1}{\Sigma Q_1} \times 100$

Elattar

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

من المثال السابق :

$P_1 Q_1$	$P_0 Q_1$	$\frac{P_1}{P_0}$	$\frac{P_1}{P_0} Q_1$
864	216	4	14,4
896	448	2	56
360	520	0,69	13,6
140	112	1,25	175
$\Sigma P_1 Q_1 = 2,260$	$\Sigma P_0 Q_1 = 1296$		$\Sigma \frac{P_1}{P_0} Q_1 = 231,3$

الرقم القياسى التجميعى المرجح لباش $= \frac{2,260}{1296} \times 100 = 174,4\ \%$

الرقم القياسى النسبى المرجح لباش $= \frac{231,3}{98} \times 100 = 236\ \%$

(٣) الرقم القياسى التجميعى والنسبى المرجح لإدجورث

استخدم ادجورث كل من كميات فترتى الأساس والمقارنة معًا كأوزان للترجيح باستخدام الوسط الحسابى أو الوسط الهندسى

حيث أن

رقم ادجورث باستخدام الوسط الحسابى $= \frac{\Sigma P_1 (Q_0 + Q_1)}{\Sigma P_0 (Q_0 + Q_1)} \times 100$

رقم ادجورث باستخدام الوسط الهندسى $= \frac{\Sigma P_1 \sqrt{Q_1 Q_0}}{\Sigma P_0 \sqrt{Q_1 Q_0}} \times 100$

Date: / / Subject: ...... Sat Sun Mon Tue Wed Thu Fri

من المثال السابق:

Q_1+Q_0	$P_1(Q_1+Q_0)$	$P_0(Q_1+Q_0)$	Q_0Q_1	$\sqrt{Q_0Q_1}$
54	1296	324	648	25,4
42	1344	672	392	19,7
30	540	780	200	14,14
26	260	208	168	12,4
	$\Sigma=3440$	$\Sigma=1984$	$\Sigma=1408$	$\Sigma=72,2$

رقم الجوبت باستخدام الوسط الحسابى $= \frac{3440}{1984} \times 100 = 173,4\%$

رقم الجوبت باستخدام الوسط الهندسى $= \frac{1618,52}{934,44} \times 100 = 173,2\%$

$P_1\sqrt{Q_1Q_0}$	$P_0\sqrt{Q_1Q_0}$
609,6	152,4
630,4	315,2
254,52	367,64
124	99,2
$\Sigma=1.618,52$	$\Sigma=934,44$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

٤ الرقم القياسي الأمثل لفشر.

قام فشر بتكوين رقم قياسي للأسعار يجمع بين رقمي لاسبير وباش وأطلق عليه الرقم القياسي الأمثل وهو عبارة عن الوسط الهندسي لرقمي لاسبير وباش

حيث أن:

$$\text{الرقم القياسي الأمثل} = \sqrt{\text{رقم لاسبير} \times \text{رقم باش}} \times 100$$

$$= \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}} \times 100$$

بالنسبة للمثال السابق:

$$\text{الرقم القياسي الأمثل} = \sqrt{1,74 \times 1,72} \times 100$$

$$= 172,9 \simeq 173\%$$

مثال

اذا كان الرقم القياسي للاسبير للأسعار يساوي 117,6 ورقم باش للأسعار يساوي 120,2. أوجد رقم القياسي الأمثل؟

(1)

Date: ٢ / ٤ / ٢٠٢٤ Subject:

"معامل الارتباط سبيرمان" ⇐ رمزه : r_s

شكل القانون:

$$r_s = 1 - \left(\frac{6 \sum d_i^2}{n(n^2-1)}\right)$$

$(d) \Leftarrow$ الفرق مابين x ، y .

مثال احسب معامل الارتباط لسبيرمان للمشاهدات التالية:

x	10	11	14	11	13	15
y	21	23	29	23	27	31

الحل

x	رتب x	y	رتب y	di	di^2
10	1	21	1	0	0
11	$\frac{2+3}{2} = 2,5$	23	2,5	0	0
14	5	29	5	0	0
11	2,5	23	2,5	0	0
13	4	27	4	0	0
15	6	31	6	0	0

$$r_s = 1 - \left(\frac{6 \sum d_i^2}{n(n^2-1)}\right) = 1 - \left(\frac{6(0)}{6(6^2-1)}\right) = 1$$

ارتباط تام

ملاحظات :

① معامل ارتباط بيرسون يصلح للبيانات الكمية والوصفية القابلة للترتيب

② معامل ارتباط سبيرمان تقريبي ولذلك لا يساوي معامل ارتباط بيرسون الا اذا كانت البيانات تامة الارتباط أي أن :

$r_p = r_s$

مثال في البيانات التالية تمثل درجات مجموعة من الطلاب في احدى الاختبارات في مادتي الرياضة والاحصاء.

الرياضة	7	5	6	9	10	8	7
الاحصاء	9	7	5	10	10	9	8

المطلوب:

احسبي معامل الارتباط لبيرسون وسبيرمان لمادتي الرياضة والاحصاء.

الحل

x	y	xy	x^2	y^2
7	9	63	49	81
5	7	35	25	49
6	5	30	36	25
9	10	90	81	100
10	10	100	100	100
8	9	72	64	81
7	8	56	49	64
$\Sigma x = 52$	$\Sigma y = 58$	$\Sigma xy = 446$	404	500

Σ

Elattar

$$r_p = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum(x^2) - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

$$= \frac{7(446) - (52)(58)}{\sqrt{7(404) - (2704)}\sqrt{7(500) - (3364)}}$$

r_p بيرسون $= 0{,}8162$

x	رتب x	y	رتب y	di	di^2
7	3,5	9	4,5	-1	1
5	1	7	2	-1	1
6	2	5	1	1	1
9	6	10	6,5	-0,5	0,25
10	7	10	6,5	0,5	0,25
8	5	9	4,5	0,5	0,25
7	3,5	8	3	0,5	0,25
					$\Sigma = 4$

$$r_s = 1 - \left(\frac{6\sum di^2}{n(n^2-1)}\right) = 1 - \left(\frac{6(4)}{7(48)}\right) = 0{,}9286$$

$\therefore$ المعاملان ليسوا نفس الناتج لأن $r_p \neq r_s$

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

معامل الاقتران : ⟸ يرمز له بالرمز (C.C)

هو معامل يستخدم لدراسة قوة الارتباط بين ظاهرتين وكل ظاهرة لها صفتين فقط وتكون لدينا بيانات على الصورة التالية :

الظاهرة الأولى / الظاهرة الثانية	الصفة الأولى للظاهرة الأولى	الصفة الثانية للظاهرة الأولى
الصفة الأولى للظاهرة الثانية	F_1	F_2
الصفة الثانية للظاهرة الثانية	F_3	F_4

$$C.C = \frac{F_1F_4 - F_2F_3}{F_1F_4 + F_2F_3}$$

مثال من دراسة العلاقة بين التدخين والتعليم في إحدى المؤسسات كانت لدينا البيانات التالية :

التدخين / التعليم	يدخن	لا يدخن
متعلم	12	8
غير متعلم	10	10

Date: / / Subject:

Sat Sun Mon Tue Wed Thu Fri

الحل

$$C.C = \frac{F_1F_4 - F_2F_3}{F_1F_4 + F_2F_3}$$

$$= \frac{12(10) - 8(10)}{12(10) + 8(10)} = 0,2$$

خط الانحدار :

عند دراسة العلاقة بين ظاهرتين:
قد تكون علاقة خطية لذلك قد تكون لدينا حاجة إلى صورة رياضية تفسر هذا الخط دون الرسم لأن العلاقة الرياضية تكون من الرسم بالاضافة أنه يمكن ايجاد قيمة عددية بطريقة سهلة ودقيقة من أحد المتغيرين اذا علمت القيمة الأخرى.

ميل الخط المستقيم ← $y = mX + C$ → الجزء المقطوع من المحور y

← خط الانحدار التابع

ومعنى ذالك :

- أنه يتحدد المستقيم تمامًا اذاتم معرفة m ، C من المشاهدات.

- وأهم طريقة لتحديد الثوابت هذه في المربعات الصغرى :
تتلخص في جعل مجموع مربعات الفرق بين القيم على الخط المهد والقيم الحقيقية أقل مايمكن لذالك سميت بطريقة المربعات الصغرى.

#