الفصل الأول:مقدمة في علم الإحصاء

الفصل الأول:مقدمـة في علم الإحصاء

علم الإحصاء(Statistics): ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع البيانات وتنظيمها وتلخيصها وعرضها وتحليلها وذلك للوصول إلى نتائج موثوقة لدعم اتخاذ قرارات سليمة على ضوء هذا التحليل.

ينقسم علم الإحصاء إلى قسمين أساسيين:

أ- إحصاء وصفي: طرق تنظيم المعلومات وفهمها.

ب- إحصاء استدلالي: يختص بعملية اتخاذ القرارات المناسبة بشأن المجتمع بناء لى المعلومات التي تم الحصول عليها.

الفصل الأول: مقدمـة في علم الإحصاء

المجتمع الإحصائي(Population): مجموعة ذات خصائص مشتركة من الأشياء ذات أهمية خاصة لدراسة علمية.

يقسم المجتمع الإحصائي إلى قسمين:

أ- محدود: وهو الذي يكون فيه عدد محدود من الأشياء أو الأفراد.

ب- غير محدود: وهو الذي يكون فيه عدد الأشياء غير منته.

العينة الإحصائية(Sample): جزء من المجتمع الإحصائي تختار بحيث تمثل جميع أفراد المجتمع تمثيلا جيدا.

مميزات العينة الإحصائية:

- ✓ استحالة فحص المجتمع بأكمله(فقدان أو تلف عناصر المجتمع).
- ✓ العينة أقل تكلفة و أكثر سرعة.
- ✓ العينة أكثر دقة.

المعلمة(Parameter): شيء يميز المجتمع الإحصائي كله.

الإحصاءة(Statistic): شيء يميز العينة الإحصائية.

الفصل الأول:مقدمة في علم الإحصاء

البيانات(Data): هي مجموعة المشاهدات المأخوذة أثناء دراسة معينة وقد تكون بيانات **وصفية** مثل المستوى التعليمي ولون الشعر أو بيانات **رقمية** مثل أطوال مجموعة من الطلاب و درجاتهم

المتغير(Variable): هو مقدار له خصائص رقمية أو غير رقمية تتغير قيمته من عنصر إلى آخر من عناصر المجتمع أو العينة.

مصادر جمع البيانات الإحصائية:

أ- تاريخي: وهو ما يؤخذ من السجلات المحفوظة.

ب- ميداني: وهو عبارة عن البيانات المجموعة من أفراد المجتمع كله أو جزء منه.

المحاضرة الثانية

العينات

انواع العينات

يمكن تقسيم العينات إلى قسمين رئيسين:

1. عينة عشوائية.
2. عينة غير عشوائية.

العينة العشوائية:

هي عملية اختيار مفردات البحث بطريقة تمنح تكافؤ الفرص لكل الوحدات (المفردات) وباحتمال معلوم للاختيار.

العينة الغير العشوائية:

يتضمن كل الطرق التي لا يتم اختيار مفرداتها عن طريق إعطاء فرص متكافئة لجميع المفردات للاختيار وباحتمال معلوم للاختيار.

أقسام العينة العشوائية

1. العينة العشوائية البسيطة.
2. العينة العشوائية الطبقية.
3. العينة العشوائية العنقودية.

العينة العشوائية البسيطة:

في هذه العملية يتم اختيار المفردات بطريقة فردية مباشرة من خلال عملية عشوائية وفيها تكون لكل الوحدات غير المختارة نفس الفرصة للاختيار مثل الوحدات المختارة.

عملية الاختيار العشوائي في العينة البسيطة:

1. يمكن استخدام طريقة القرعة إذا كان حجم مجتمع البحث صغيرا.
2. يمكن أن يتم عملية اختيار مفردات العينة باستخدام الحاسب الآلي ببرنامج معين.
3. يمكن أن يتم الاختيار العشوائي باستخدام جداول الأرقام العشوائية .
4. يمكن اختيار مفردات العينة بإتباع طريقة العينة العشوائية المنتظمة.

طريقة اختيار العينة العشوائية باستخدام جداول الارقام العشوائية

إذا اردنا اختيار عينة قوامها 5% من طلاب الجامعة البالغ عددهم 20000 طالب وطالبة باستخدام جداول الأرقام العشوائية.عددهم 1000 طالب.

1. الحصول على قوائم تضم كل طلاب الجامعة البالغ عددهم 20000 طالبا.
2. إعطاء رقم متسلسل لكل طالب من 1-20000.
3. تفتح أي صفحة من صفحات جدول الأرقام العشوائية ونضع اصبعنا على أي رقم في أول الجدول أو في منتصفه أو أي مكان وهكذا حتى نصل للعدد المنشود وهو 1000 طالب.

اختيار العينة العشوائية البسيطة بالطريقة المنتظمة

وهو عبارة عن طريقة اختيار الوحدات من قائمة بتطبيق فترات منتظمة للاختيار بحيث يتم اختيار المفردة التي تقع بعد عدد معين من المفردات مبتدئاً من مفردة عشوائية.

- تتمثل طريقة الاختيار تماما كالعينة البسيطة العشوائية في اختيار أول مفردة ومن ثم إذا افترضنا أن الرقم الأول كان 16 يتم الاختيار عن طريق مضاعفات الرقم 16 مثلا 16×2 16×3 16×4 وهكذا حتى الوصول إلى العدد المطلوب وهو 1000 طالب.

العينة العشوائية الطبقية

يلجأ إليها في حالة وجود مجتمعات تتميز بتباين نوعيات مفرداتها بحيث يمكن تقسيمها إلى مجموعات أو طبقات لكل مجموعة أو طبقة منها خصائص أو مميزات معينة تتميز بها عن بقية الطبقات الأخرى.

طريقة الاختيار:

1. تحديد حجم مجتمع البحث الكلي.
2. تقسيم مجتمع البحث إلى فئات متباينة حسب موضوع الدراسة بحيث تتجانس المفردات الداخلية لكل فئة أو طبقة.
3. تحديد عدد أفراد كل طبقة.
4. تحديد نسبة مفردات كل مجموعة أو طبقة بالنسبة للعدد الكلي لمفردات مجتمع البحث.
5. استخدام الأسلوب العشوائي البسيط أو المنتظم لاختيار المفردات في كل طبقة.

المحاضرة الثالثة

التوزيعات التكرارية

عبارة عن جداول لجميع القيم التي يمكن أن يأخذها المتغير موضع الدراسة و عدد التكرارات المناظرة لكل قيمة

مثال1

البيانات التالية تمثل عدد أيام الغياب 30 طالب في الصف الأول المتوسط بإحدى المدارس في الأسبوع الأول من شهر رمضان:

0 2 3 1 0 0 0 0 1 2 2 5

3 2 1 1 1 1 1 2 0 1 3 2

3 2 2 3

الجدول التكراري

	卌 |	
	卌 |||	
	卌 |||	
	卌 ||	
	|	

مثال 2

البيانات التالية تمثل درجات 50 طالبة في مقرر الإحصاء (لأقرب درجة صحيحة)

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

لتكوين الجدول التكراري نتبع الخطوات التالية:

1) نحدد عدد الفئات المناسب و يكون عادة بين 5-15 فئة تبعاً لعدد البيانات.

في مثالنا نختار عدد الفئات = 5

2) نحدد مدى البيانات = أكبر قيمة – أصغر قيمة.

المدى =
49-25=24

٣) نحدد طول الفئة $\Delta = \frac{\text{المدى}}{\text{عدد الفئات}}$

$$\Delta = \frac{24}{5} = 4.8 \approx 5$$

٤) نحدد الحد الأدنى لأول فئة و يكون $\geq$ أصغر قيمة للبيانات

الحد الأدنى= 25
الحد الأدنى الفعلي= 25 - 0.5 = 24.5

5) نحدد الحد الأعلى الفعلي لأول فئة وذلك بإضافة طول الفئة للحد الأدنى الفعلي

الحد الأعلى الفعلي= 24.5+5=29.5
الحد الأعلى= 29.5-0.5 = 29

6) نحدد الحدود الدنيا و العليا لكل فئة بإضافة طول الفئة لكل حد ، و نعين الحدود الفعلية بإضافة طول الفئة أيضاً لكل حد فعلي.

حدود الفئة	الحدود الفعلية للفئة
25 - 29	24.5 – 29.5
30 – 34	29.5 – 34.5
35 – 39	34.5 – 39.5
40 - 44	39.5 – 44.5
45 - 49	44.5 – 49.5

٧) نعين مراكز الفئات x_i باستخدام القانون

$$x_i = \frac{\text{الحد الأدنى} + \text{الحد الأعلى}}{2}$$

أو يمكن إيجاد مركز الفئة الأولى فقط ثم إضافة طول الفئة لإيجاد مركز الفئة التالية.

$$x_1 = \frac{29+25}{2} = 27$$
$$x_2 = 27 + 5 = 32$$

8) نكون الجدول التكراري.

التكرار f_i	العلامات	الحدود الفعلي للفئة x_i	الحدود الفعلية للفئة	حدود الفئة
7	~~IIII~~ II	27	24.5 – 29.5	25 - 29
18	~~IIII~~ ~~IIII~~ ~~IIII~~ III	32	29.5 – 34.5	30 – 34
13	~~IIII~~ ~~IIII~~ III	37	34.5 – 39.5	35 – 39
9	~~IIII~~ IIII	42	39.5 – 44.5	40 - 44
3	III	47	44.5 – 49.5	45 - 49
$\sum f_i=50$				

التكرار النسبي

التكرار النسبي للفئة هو نسبة تكرار الفئة
إلى مجموع التكرارات
فإذا كان تكرار الفئة هو f_i و مجموع التكرارات هو n فإن
التكرار النسبي p_i يعطى من العلاقة:

$$p_i = \frac{f_i}{n} = \frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}}$$

التكرار المئوي

التكرار المئوي للفئة هو حاصل ضرب تكرارها النسبي في
100

بالعودة للمثال السابق

التكرار المئوي %	التكرار النسبي p_i	فئة حدود الفئة	x_i فئة الحدو الفعلية ل	الحدود الفعلية للفئة	حدود الفئة
14	$\frac{7}{50} = 0.14$	7	27	24.5 – 29.5	25 - 29
36	$\frac{18}{50} = 0.36$	18	32	29.5 – 34.5	30 – 34
26	$\frac{13}{50} = 0.26$	13	37	34.5 – 39.5	35 – 39
18	$\frac{9}{50} = 0.18$	9	42	39.5 – 44.5	40 - 44
6	$\frac{3}{50} = 0.06$	3	47	44.5 – 49.5	45 - 49
100%	**1**	50			$\sum$

سؤال

ما هو عدد الطلاب الذين درجاتهم أقل من أو تساوي 39؟

7+18+13 =38

و هو ما يعرف بالتكرار المتجمع للفئة

التكرار المتجمع للفئة

مجموع تكرارات جميع القيم التي تكون أقل من أو تساوي الحد الأعلى الفعلي للفئة.

لإنشاء الجدول التوزيع التكراري المتجمع

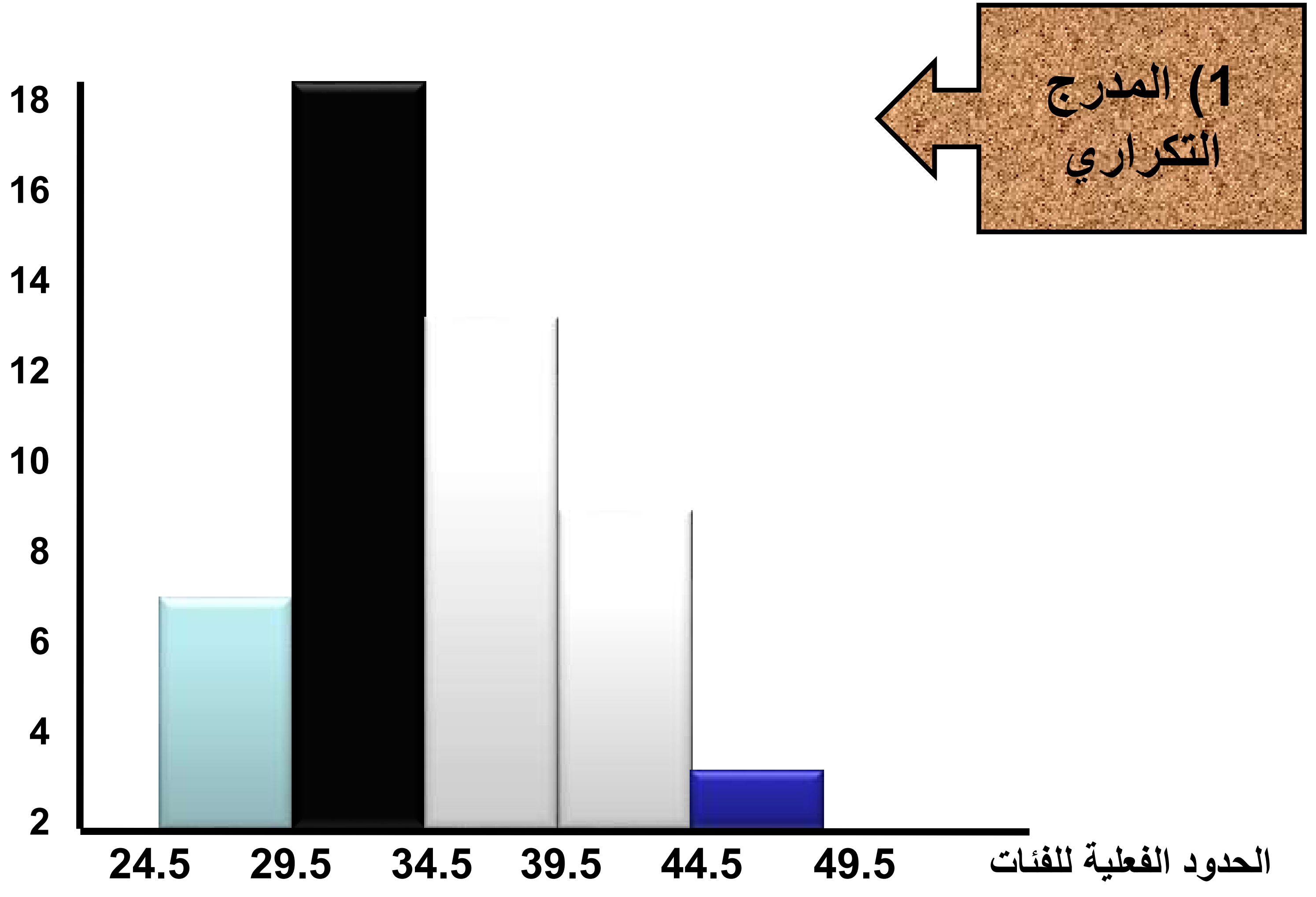
حدود الفئة	الحدود الفعلية للفئة	التكرار	أقل من الحد الأعلى الفعلي للفئة	التكرار المتجمع
25 - 29	24.5 – 29.5	7	أقل من 29.5	7
30 – 34	29.5 – 34.5	18	أقل من 34.5	25
35 – 39	34.5 – 39.5	13	أقل من 39.5	38
40 - 44	39.5 – 44.5	9	أقل من 44.5	47
45 - 49	44.5 – 49.5	3	أقل من 49.5	50
		$\sum f = 50$		

لاحظي أن التكرار المتجمع للفئة الأخيرة = مجموع التكرارات

بنفس الطريقة يمكن الحصول على التوزيع التكراري المتجمع النسبي و التوزيع التكراري المتجمع المئوي و لكن بالتجميع المتتالي للتكرار النسبي و المئوي للفئات على الترتيب.

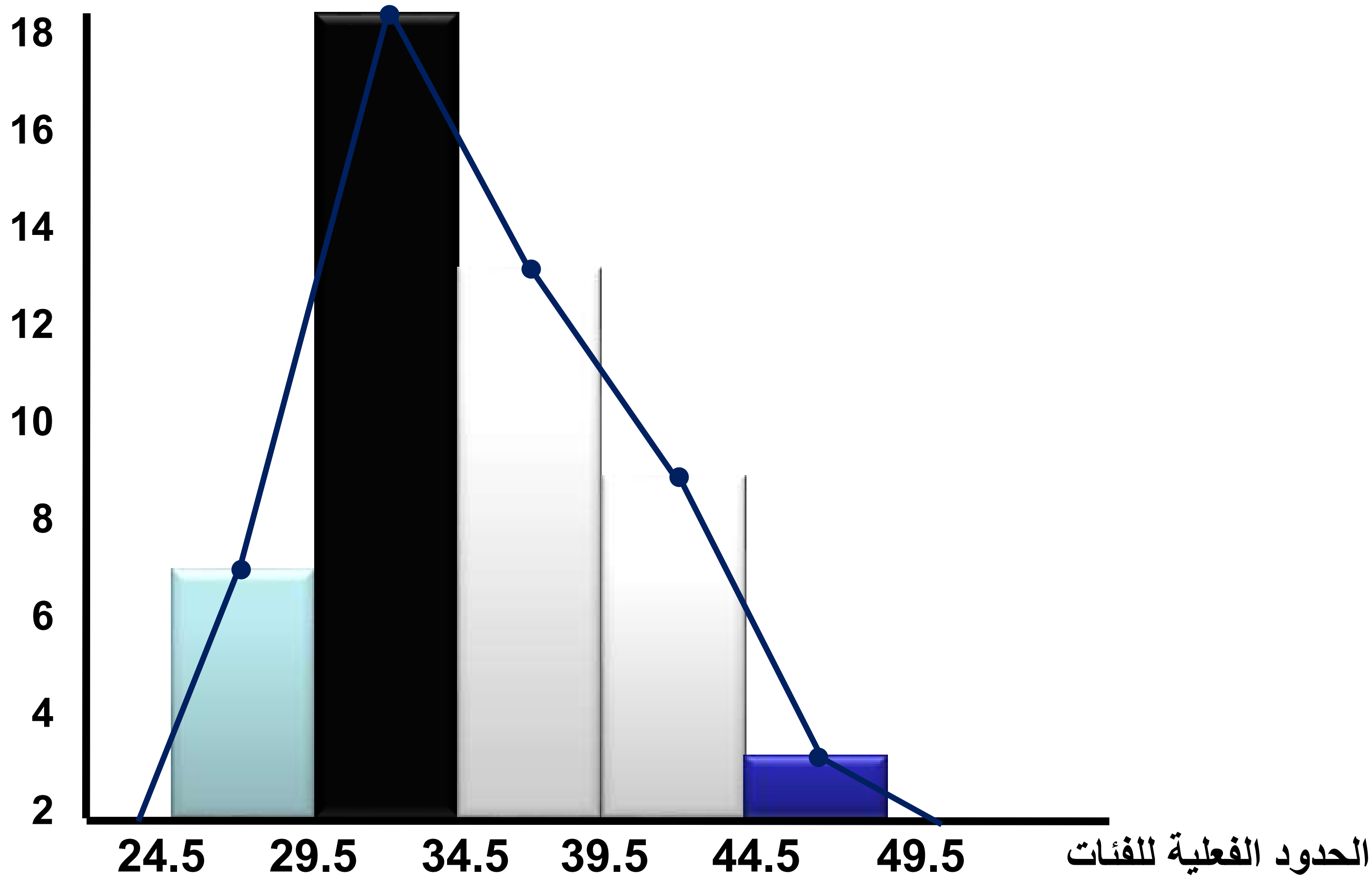
تمثيل التوزيعات التكرارية بيانياً

توجد ثلاث طرق لتمثيل التوزيعات التكرارية بيانياً:



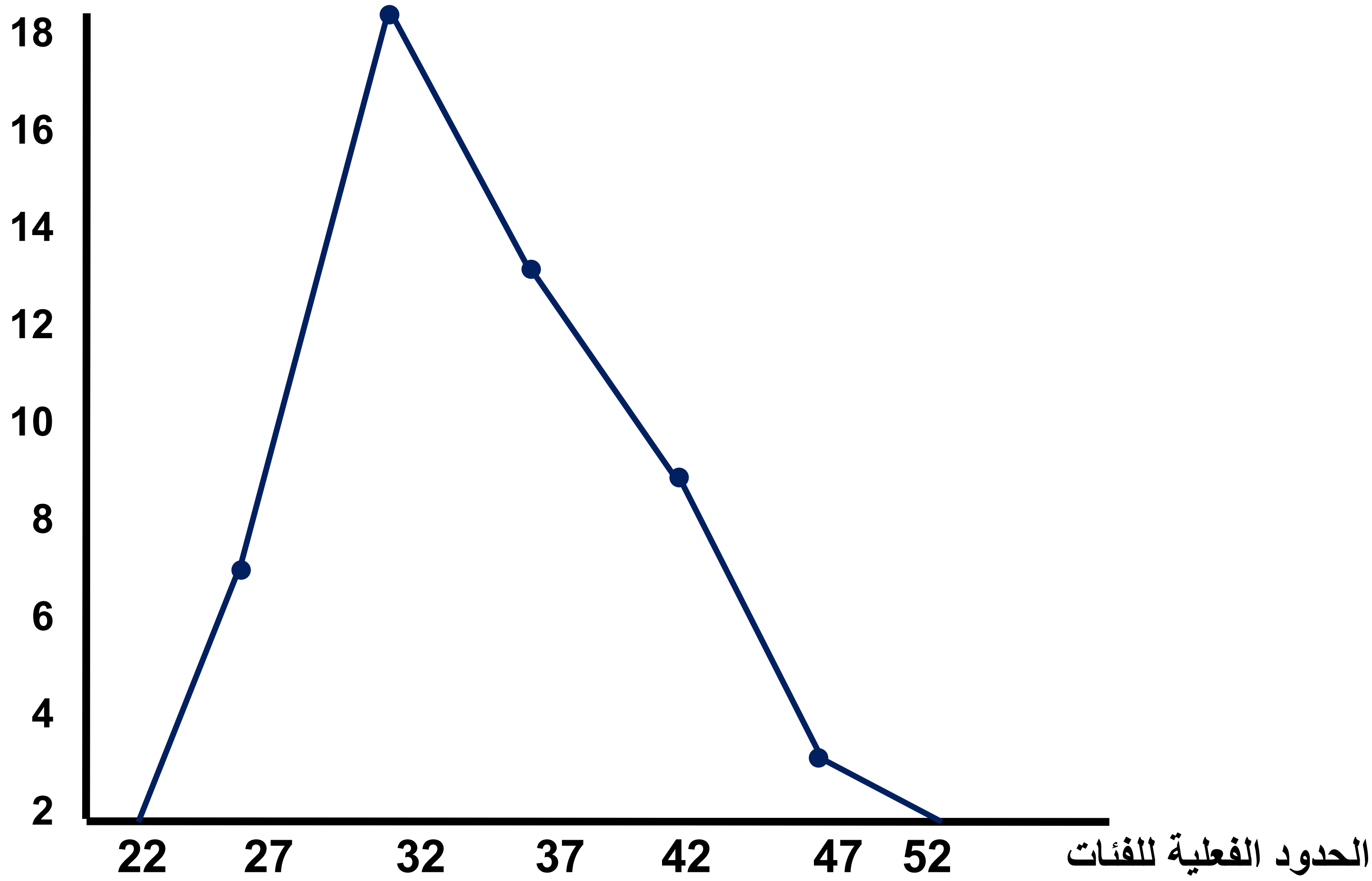
2) المضلع التكراري

و هو مضلع نحصل عليه بتنصيف الأضلاع العلوية للمستطيلات و نصل بينها .

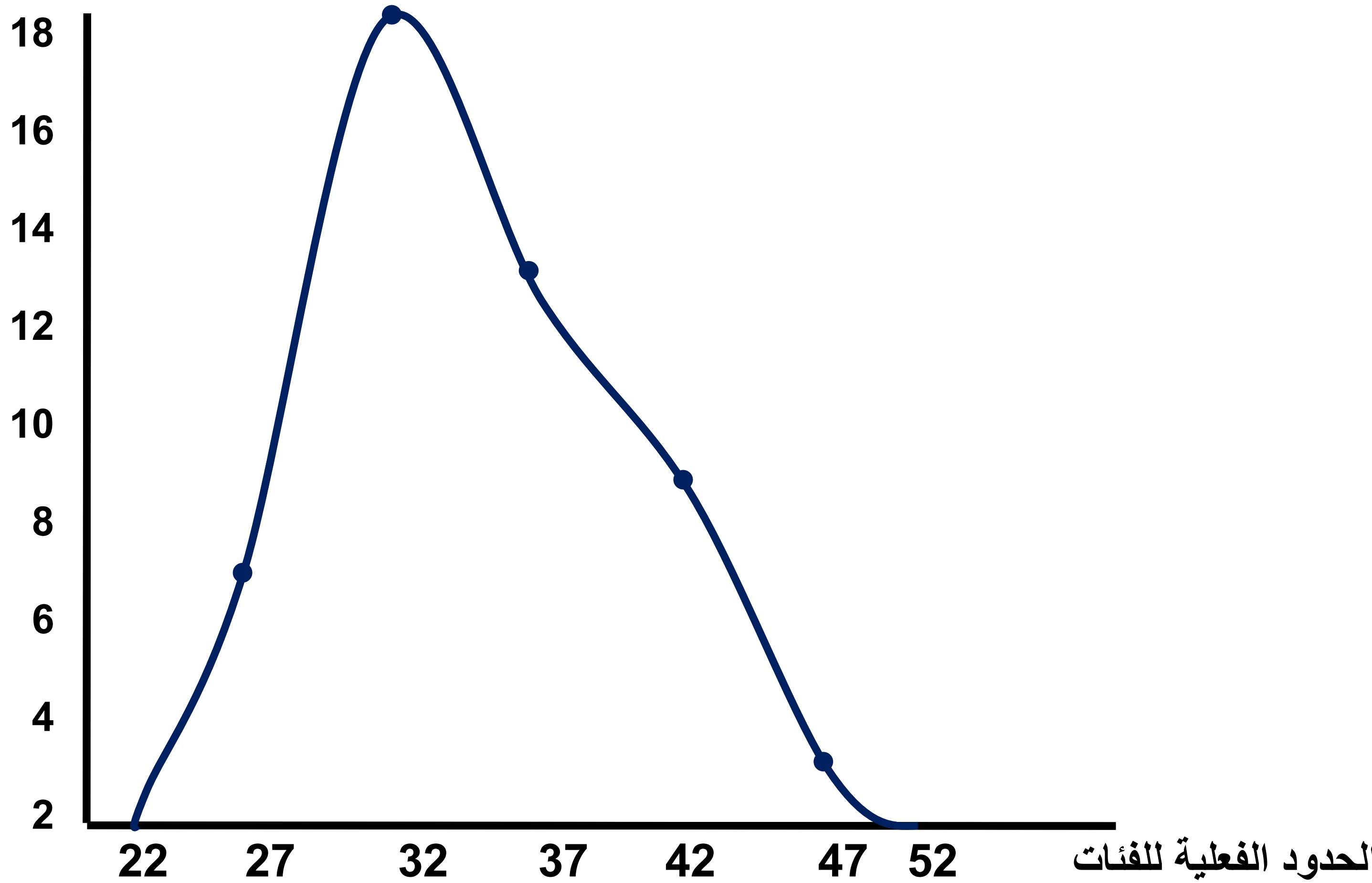


طريقة أخرى لرسم المضلع التكراري

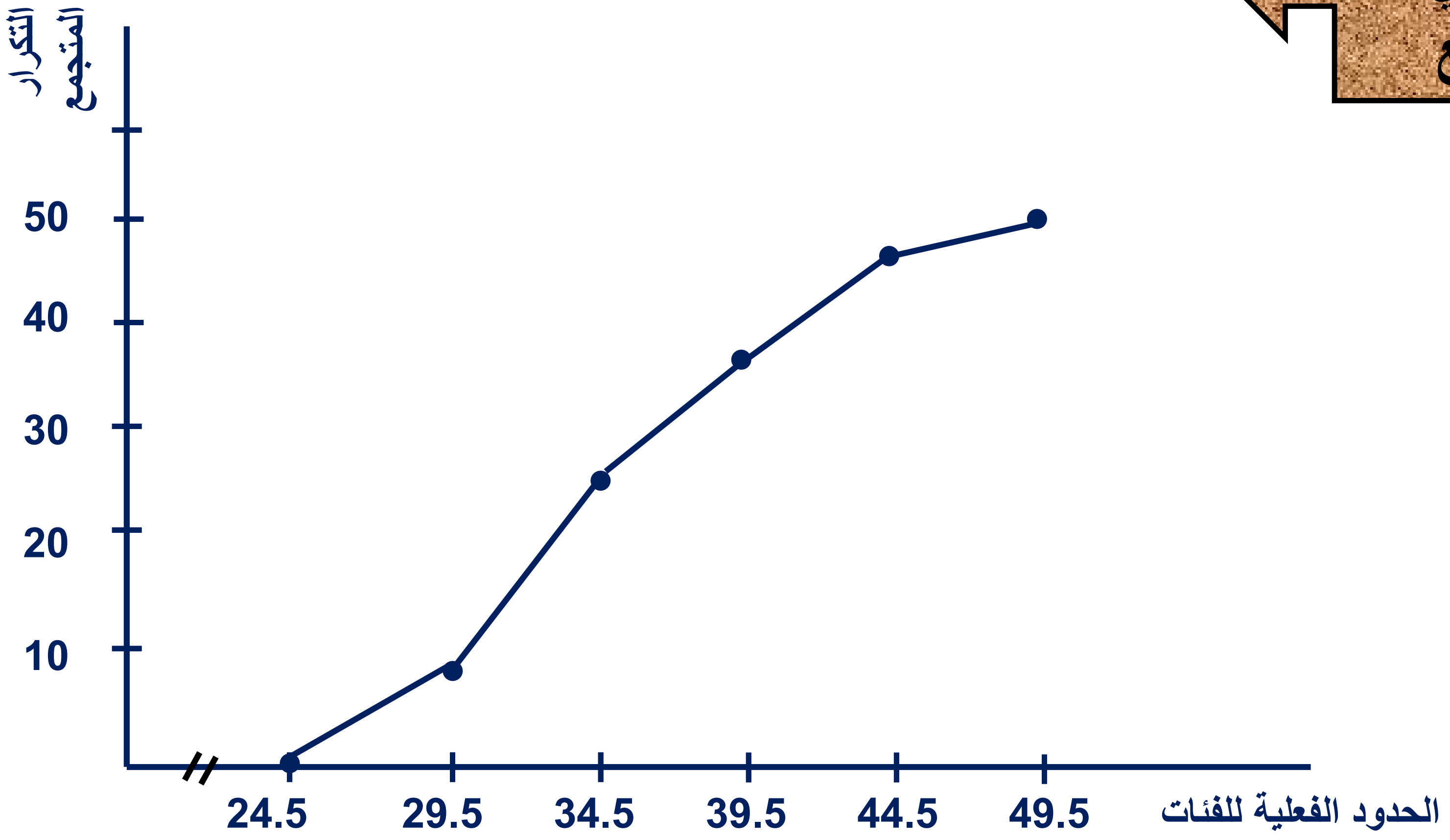
نحدد النقاط (الأزواج المرتبة) التي إحداثيها الأفقي مركز الفئة و إحداثيها العمودي تكرار الفئة ثم نصل بينها بقطع مستقيمة.



3) المنحنى التكراري



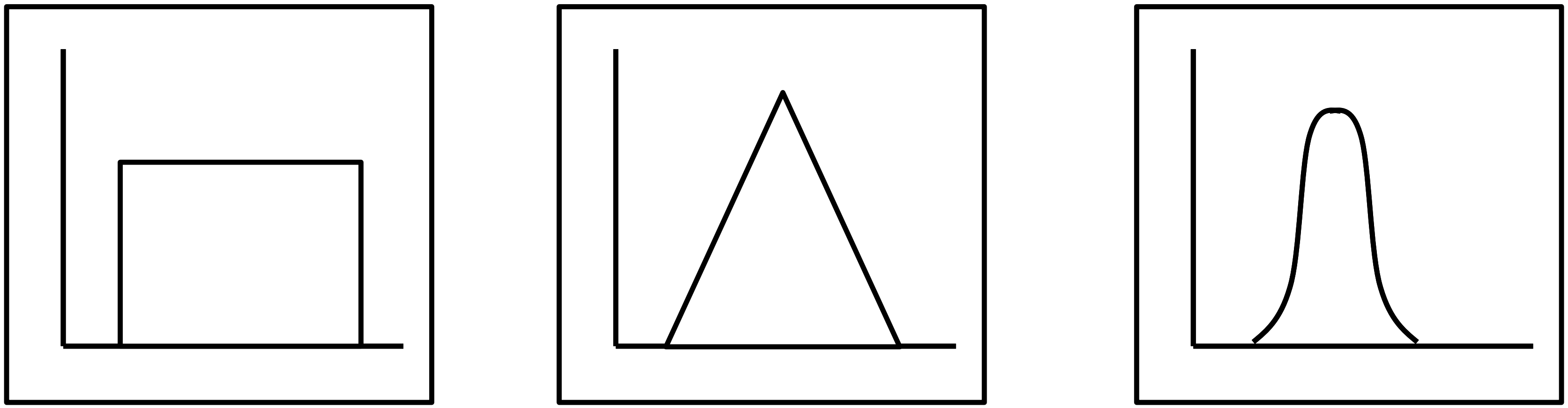
المضلع
التكراري
المتجمع



أشكال التوزيعات التكرارية

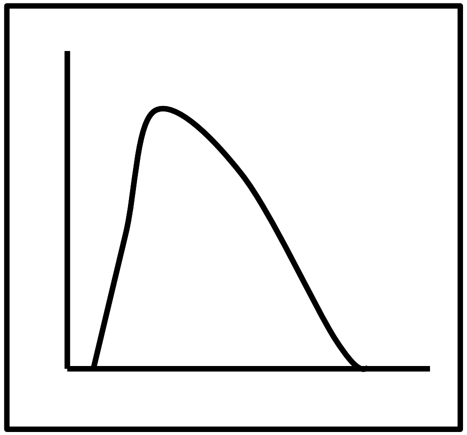
1) تماثل التوزيع

يكون التوزيع متماثل إذا وجد عمود يقسم التوزيع إلى قسمين منطبقين و هذا قليل التحقق على أرض الواقع و لكن تكثر التوزيعات المتماثلة تقريباً.

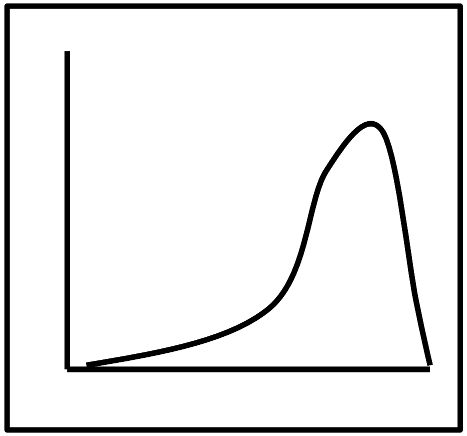


التوزيع الملتوي

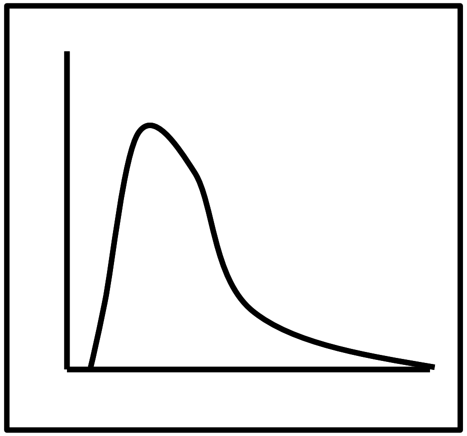
إذا كان عدم تماثل التوزيع واضحاً و ذلك عندما يمتد أحد طرفيه كثيراً لليمين أو لليسار أو كان علياً من جهة و منخفضاً في أخرى



توزيع معتدل الالتواء



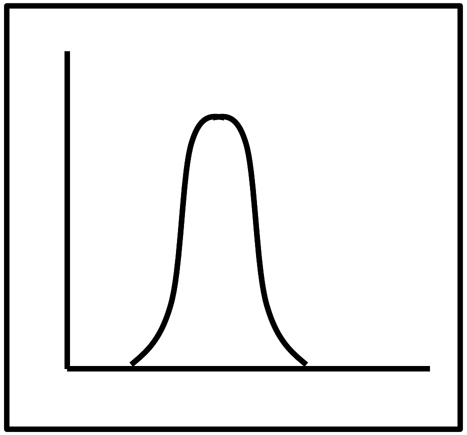
توزيع ملتو نحو اليسار
سالب الالتواء



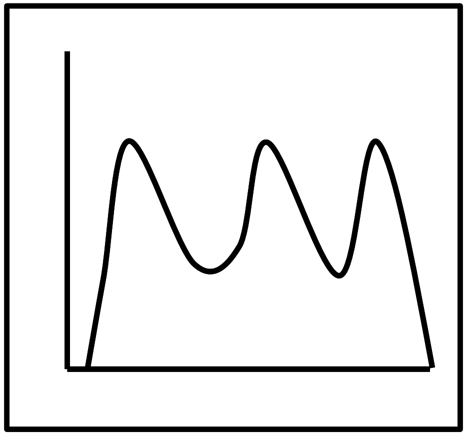
توزيع ملتو نحو اليمين
موجب الالتواء

2) منوال التوزيع

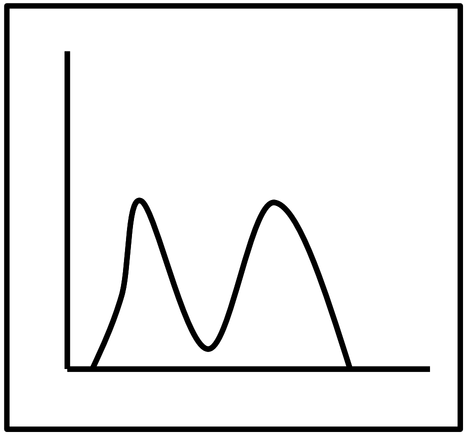
و هو القيمة التي تكررت أكثر من غيرها (يمثل القمة في التوزيع)



أحادي المنوال



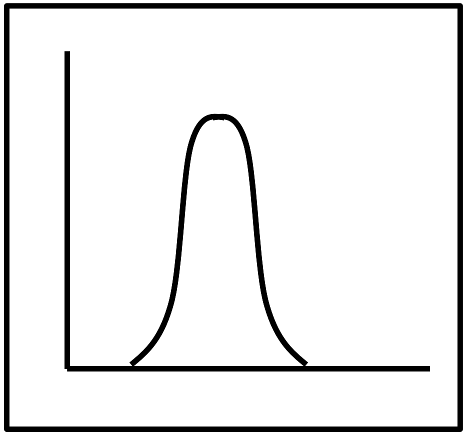
متعدد المنوال(ثلاثي)



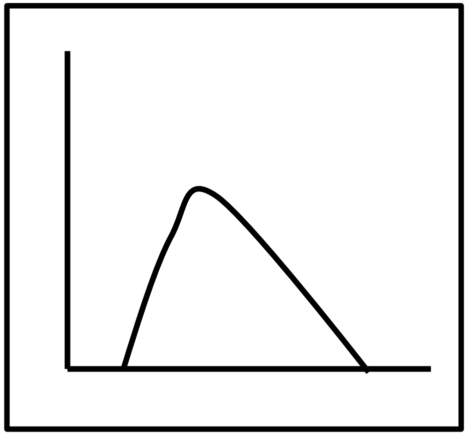
ثنائي المنوال

3) تفرطح التوزيع

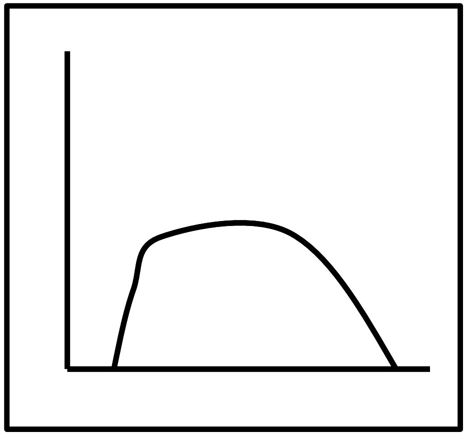
مدى علو قمة التوزيع التكراري



قليل التفرطح (مدبب)

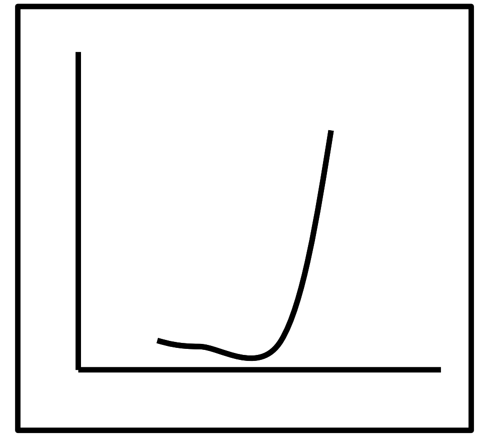


متوسط التفرطح



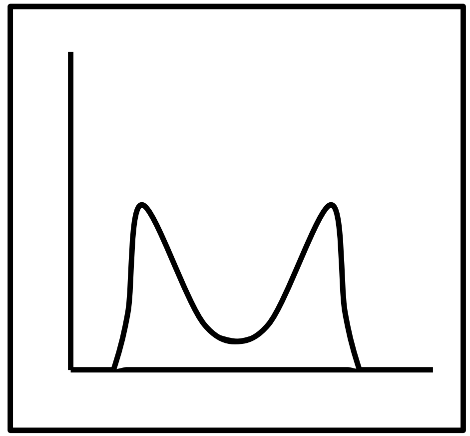
كبير التفرطح

4) بعض التوزيعات الشهيرة



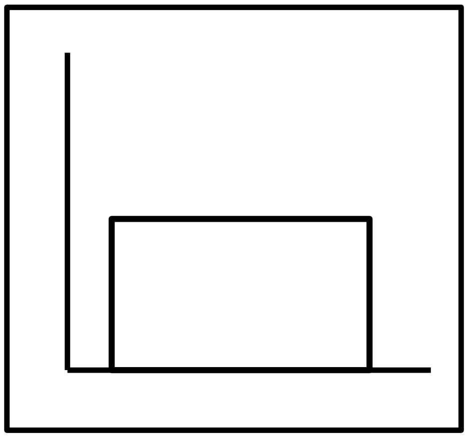
توزيع ل

- ملتو نحو اليسار
- له منوال في أقصى اليمين
- إذا كان ملتو نحو اليمين فقمته أقصى اليسار.



توزيع U

- متماثل
- ثنائي المنوال (له قمتان من اليمين و اليسار)



توزيع متجانس(مستطيل)

- متماثل
- ليس له منوال

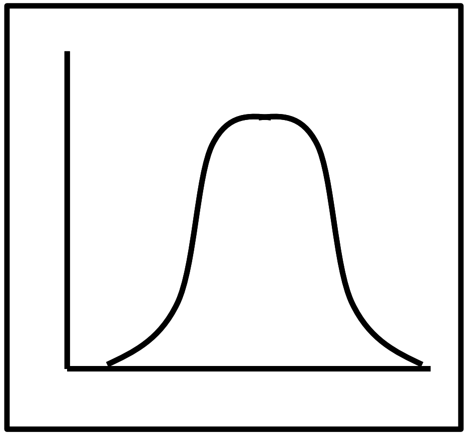
بعض الشروط ليكون المضلع تكرارياً

١) أن يكون مجموع التكرارات مساوياً لعدد البيانات.

$$\sum f = n$$

٢) أن لا يوجد أي تكرار سالب، أي يقع التوزيع بالكامل أعلى المحور الأفقي.

٣) أن يكون لتكرار كل فئة قيمة واحدة فقط.



توزيع الجرس

- أحادي المنوال
- متماثل حول نقطة المنوال

واجب

البيانات التالية تمثل الأجور اليومية بالريال ل 50 عامل في مصنع كوني الجدول التكراري

47	36	40	55	75	53	46	43	21	10
66	56	46	35	47	32	52	48	41	30
27	25	57	15	37	22	63	21	61	62
54	42	35	49	39	32	45	31	72	50
65	18	79	23	48	44	32	51	44	42

كوني جدول تكراري ذو 7 فئات ، ثم أوجدي كلاً من:

1. التكرار النسبي
2. التكرار المئوي
3. التكرار المتجمع الصاعد.
4. التكرار المتجمع النسبي
5. المضلع التكراري.
6. المضلع المتجمع الصاعد.

المحاضرة الرابعة

الفصل الرابع

عرض البيانات الإحصائية ووصفها

الإحصاء قديماً

مجرد جمع المعلومات و ترتيبها في جداول أو ابرازها في رسوم بيانية أو أشكال تصويرية.

الإحصاء الحديث

العلم الذي يبحث في جمع البيانات و تنظيمها و عرضها و تحليلها و استقراء النتائج و اتخاذ القرارات.

العلمي

جمع البيانات

تنظيم و عرض البيانات

استقراء النتائج و اتخاذ القرارات

جمع البيانات

- عملية الحصول على المعلومات أو قيم المشاهدات أو القياسات للتجارب التي يجريها الإحصائي.

و عرض المعلومات

- عملية وضع المعلومات في جداول منسقة و عرضها بطرق مناسبة كالأشكال الهندسية و الرسوم البيانية و غيرها.

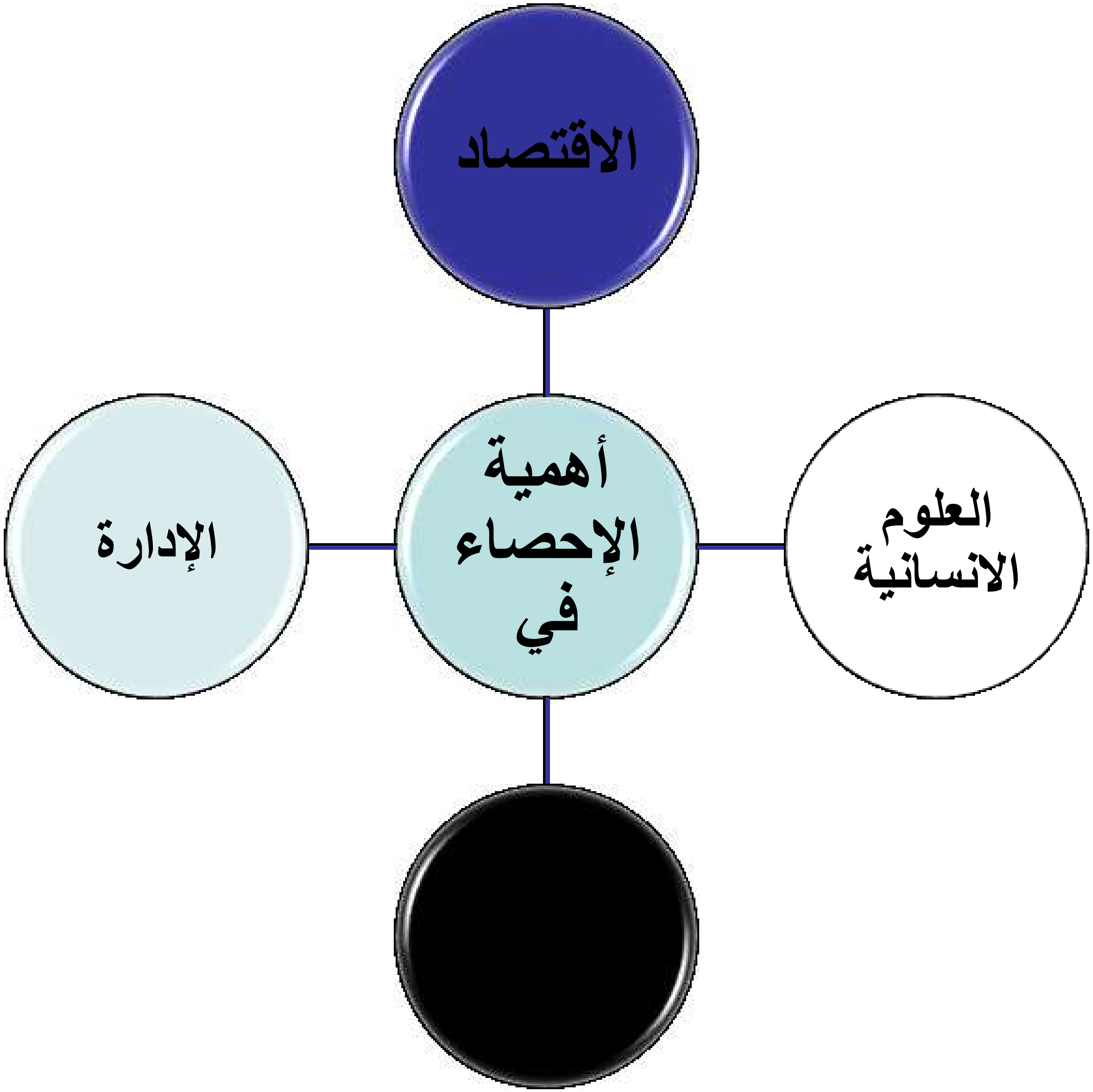
تحليل البيانات

- عملية إيجاد قيم لمقاييس و اقترانات معينة تحدد قيمها من البيانات موضع الدراسة.

الاستقراء و اتخاذ القرارات

- الاستنتاجات التي يصل إليها الباحث و تكون على شكل تقديرات أو تنبؤات أو تعميمات أو قرار برفض أو قبول الفرضيات الإحصائية.

- علم الإحصاء
  - إحصاء استقرائي
    - استقراء النتائج و اتخاذ القرارات
  - إحصاء وصفي
    - جمع و تبويب البيانات


الاقتصاد
أهمية الإحصاء في
الإدارة
العلوم الانسانية

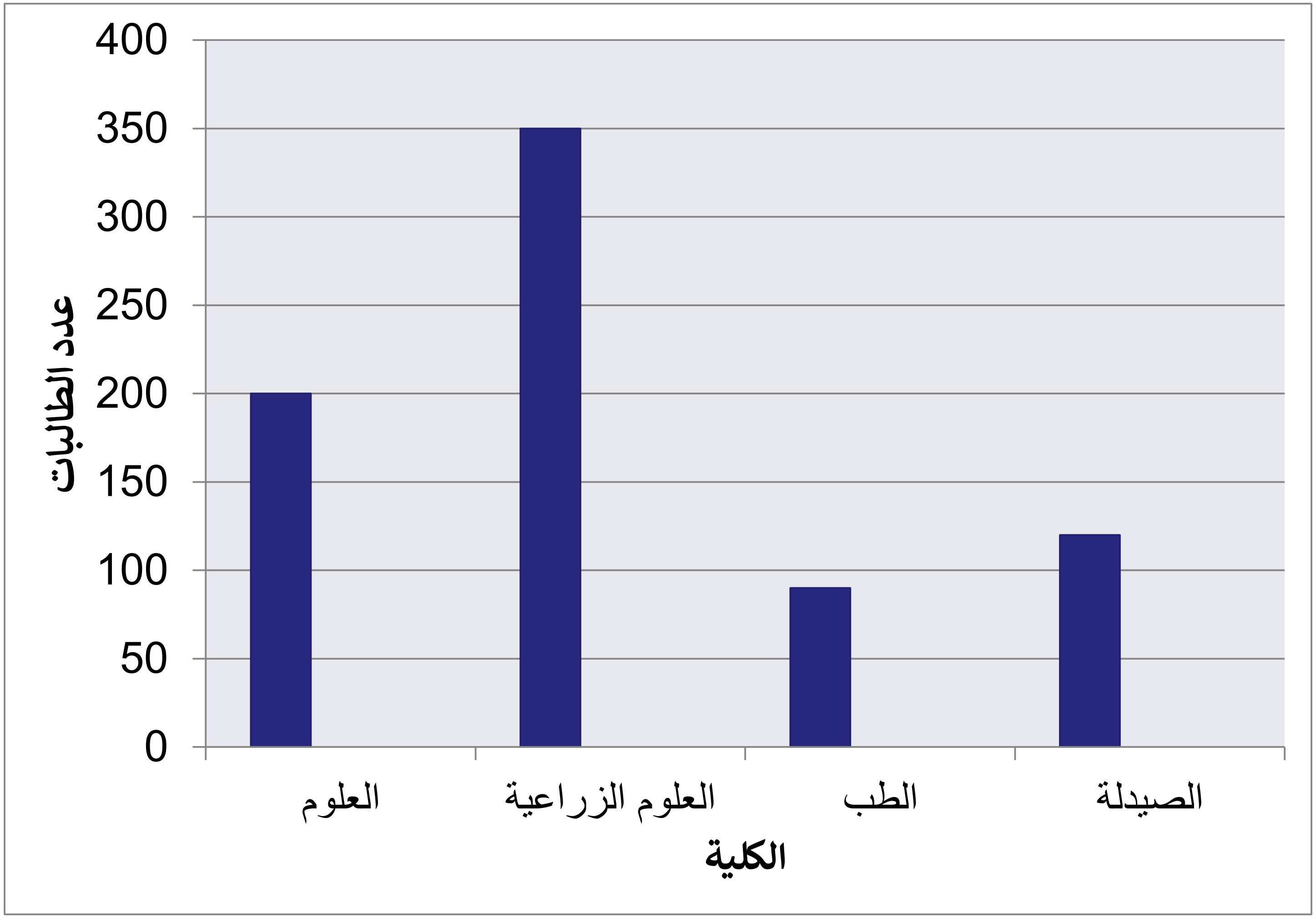
طرق عرض البيانات

1. طريقة الجدول

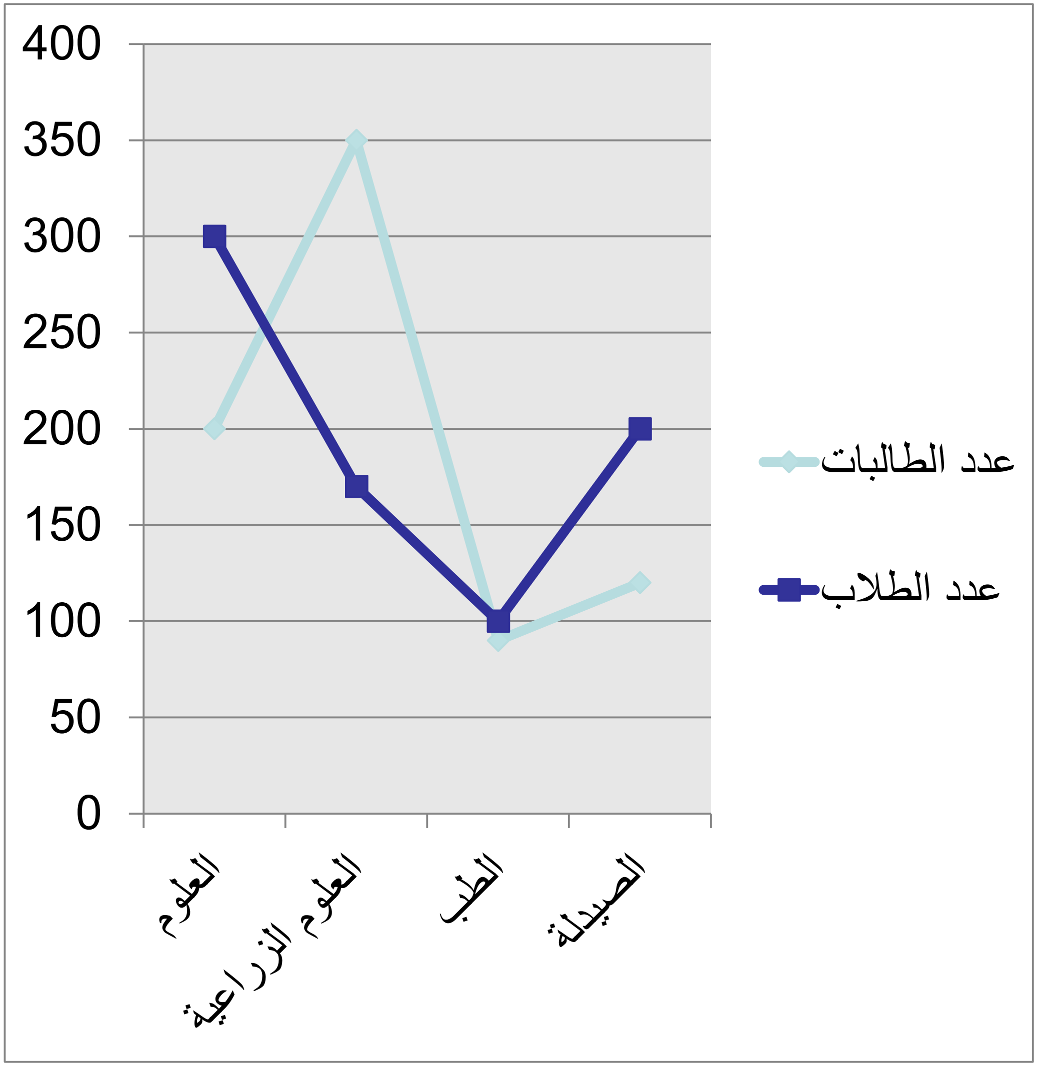
مثال عدد الطالبات المقبولات في جامعة الملك فيصل في إحدى السنوات

الكلية	عدد الطالبات
العلوم	200
العلوم الزراعية	350
الطب	90
الصيدلة	120
الدراسات التطبيقية	150
المجموع	910

2. طريقة الأعمدة أو المستطيلات

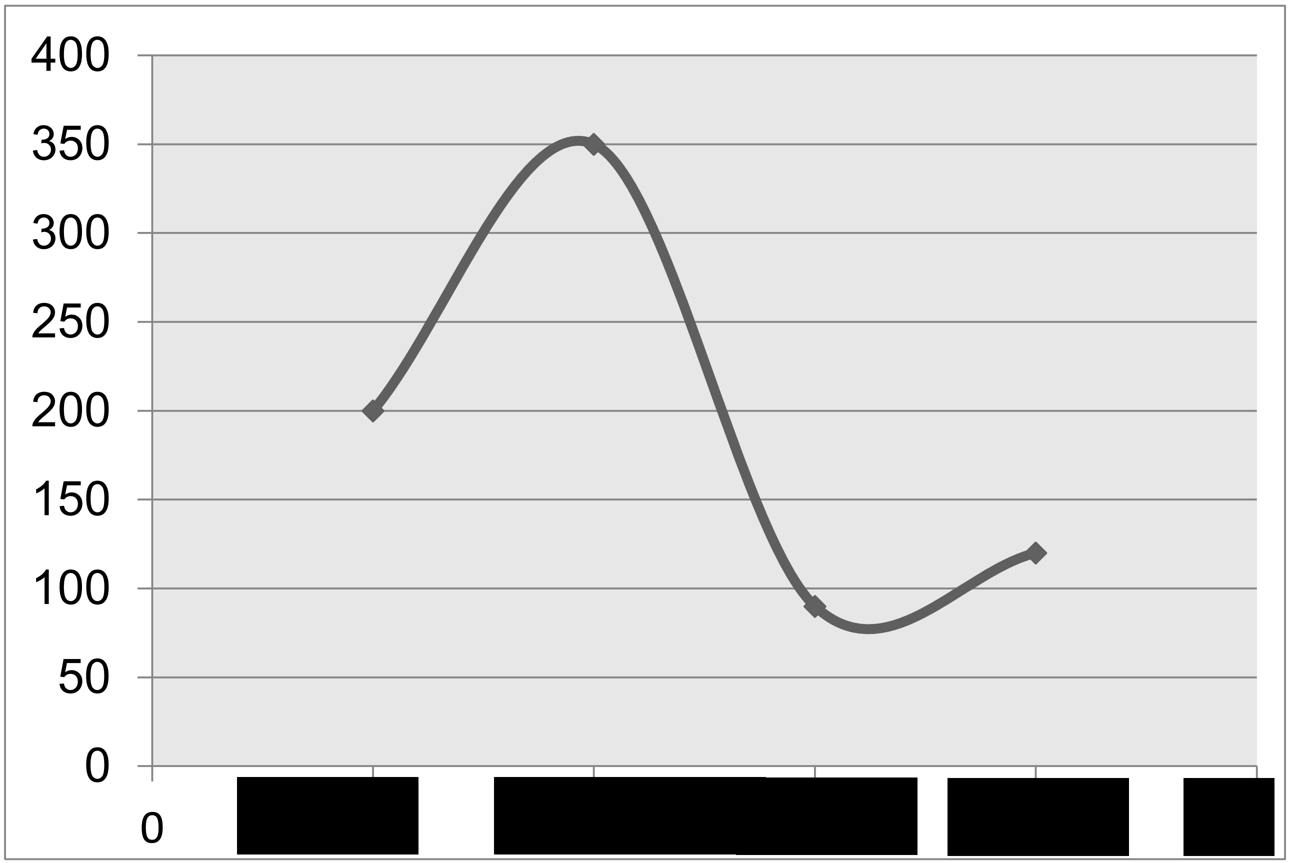


3. طريقة الخط المنكسر



الكلية	عدد الطالبات	عدد الطلاب
العلوم	200	300
العلوم الزراعية	350	170
الطب	90	100
الصيدلة	120	200
الدراسات التطبيقية	150	100
المجموع	910	870

4. طريقة الخط المنحني

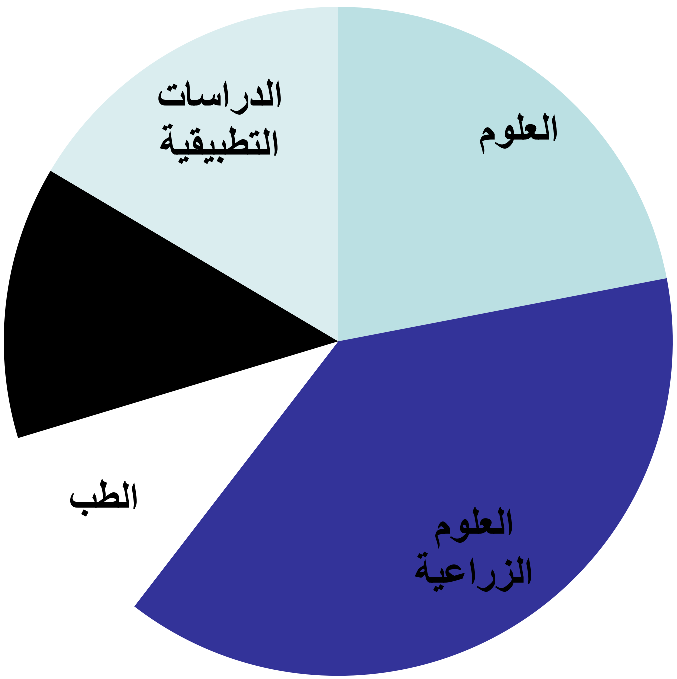


5. طريقة الدائرة

تستخدم هذه الطريقة لتقسيم الكل إلى أجزائه حيث تقسم الدائرة إلى قطاعات دائرية.

$$\text{قياس زاوية القطاع} = \frac{\text{تكرار البيانات}}{\text{العدد الكلي}} \times 360^\circ$$

مثال



الكلية	عدد الطالبات	زاوية القطاع
العلوم	200	العلوم
العلوم الزراعية	350	العلوم الزراعية
الطب	90	الطب
الصيدلة	120	الزراعية
الدراسات التطبيقية	150	الدراسات التطبيقية
المجموع	910	المجموع

المحاضرة الخامسة

جمع البيانات Collecting data

Objectives :

1) The student will learn to represent a given set of data in a frequency distribution table.

2) the student will learn to represent the data by frequency histogram and frequency polygon.

الأهداف :

1) سوف يتعلم الطالب تمثيل البيانات بجدول التوزيع التكراري

2) سوف يتعلم الطالب تمثيل البيانات بمدرج تكراري وبمضلع تكراري

Frequency التكرار

Key words :

المصطلحات :

Statistics : **الإحصاء**

Data : **البيانات**

Frequency : **التكرار**

Frequency distribution table : **جدول التوزيع التكراري**

Frequency histogram : **المدرج التكراري**

جدول التوزيع التكراريFrequency distribution table

نتائج إلقاء حجر نرد 20 مرة

Example

A dice was rolled 20 times , the results are given below :

3 2 4 2 2 3 2 6 5 5 1 1 2 3 5 2 1 4 1 6

We can represent the data in a frequency distribution table :

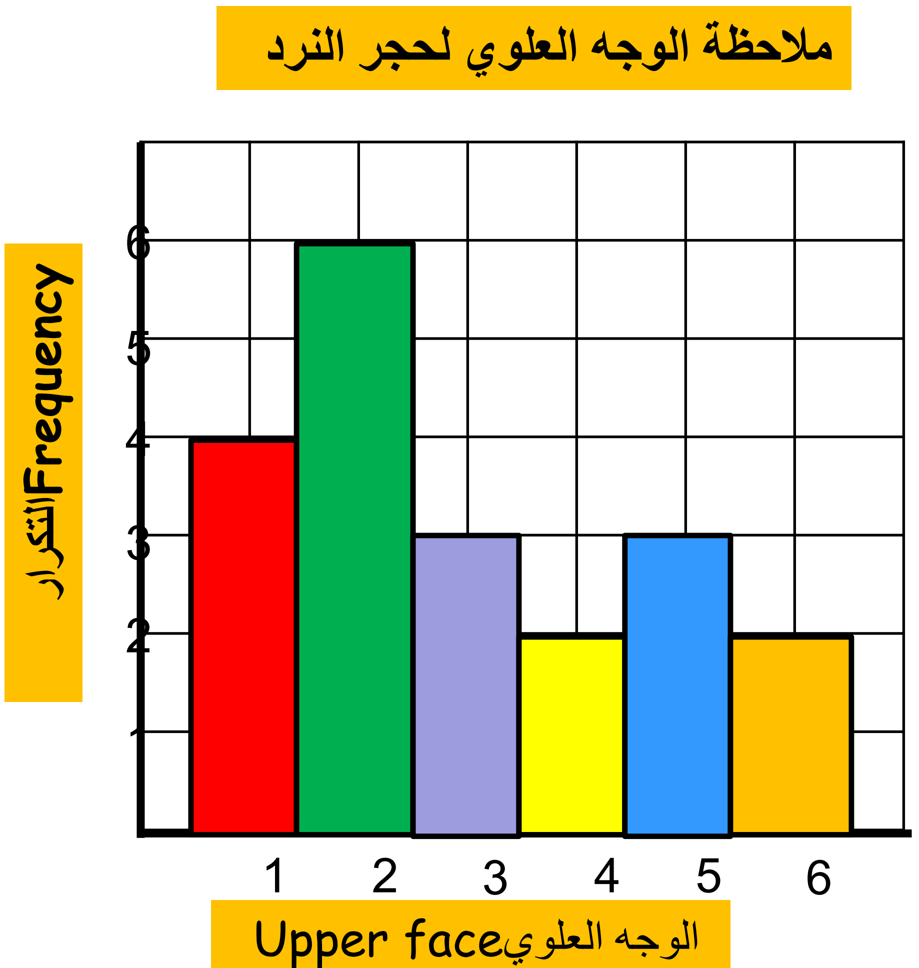
Upper face الوجه العلوي	Tally العلامات	Frequency التكرار
	////	
	~~////~~ /	
	///	
	///	
	//	
total		20

يمكن تمثيل البيانات بجدول التوزيع التكراري

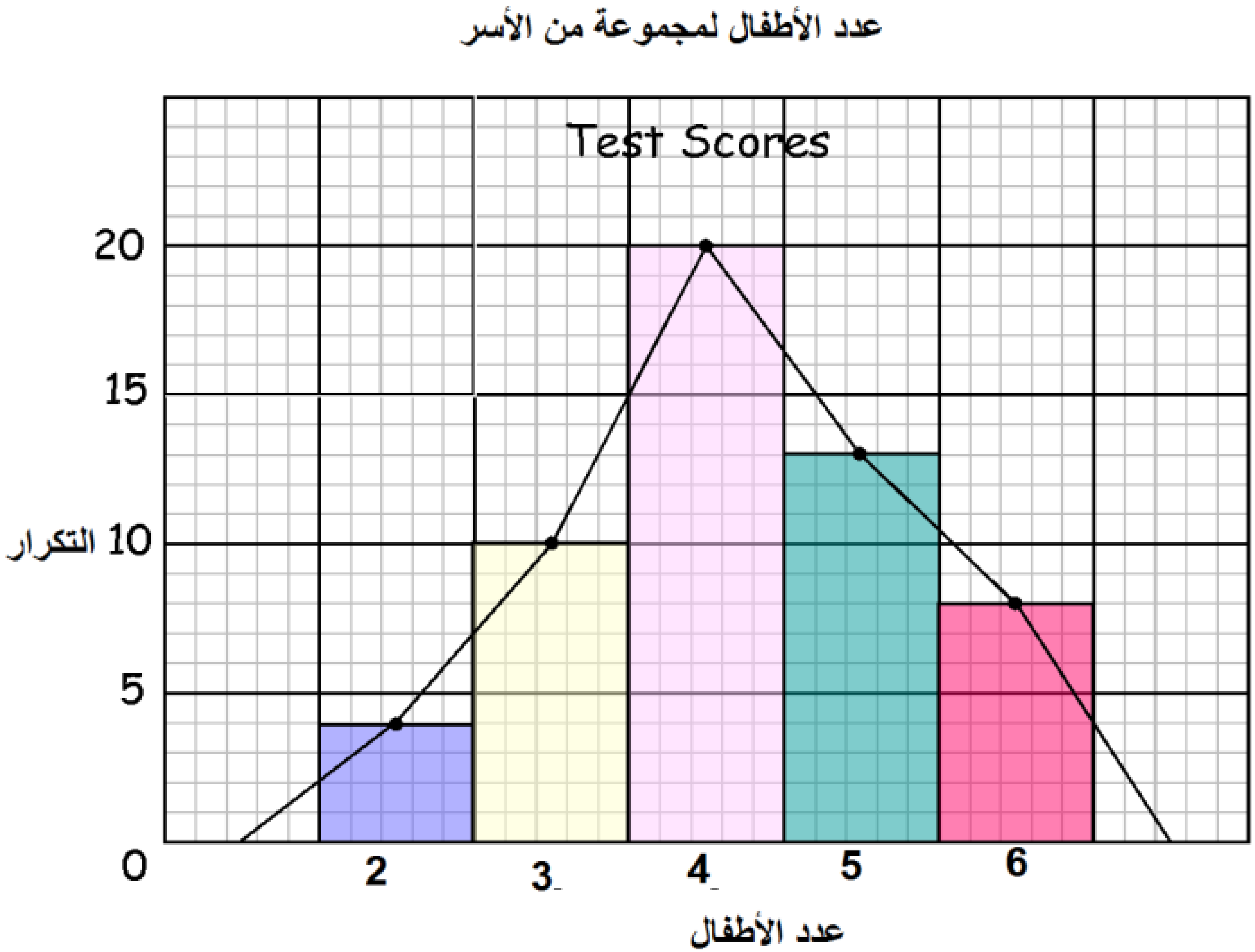
تمثيل البيانات بمدرج تكراري

Represent the data in a frequency histogram :

الوجه العلوي Upper face	التكرارFrequency
1	4
2	6
3	3
4	2
5	3
6	2
total	20



Frequency Polygons



ملاحظات على المدرج التكراري Histogram

This is a type of column graph where the columns are drawn next to each other. You should note the following features.

- Each axis is labelled.
- The graph has a title.
- The columns are centred on the scores.
- The first column begins one-half of a column width in from the vertical axis.

هو نوع من الرسم البياني بالأعمدة:

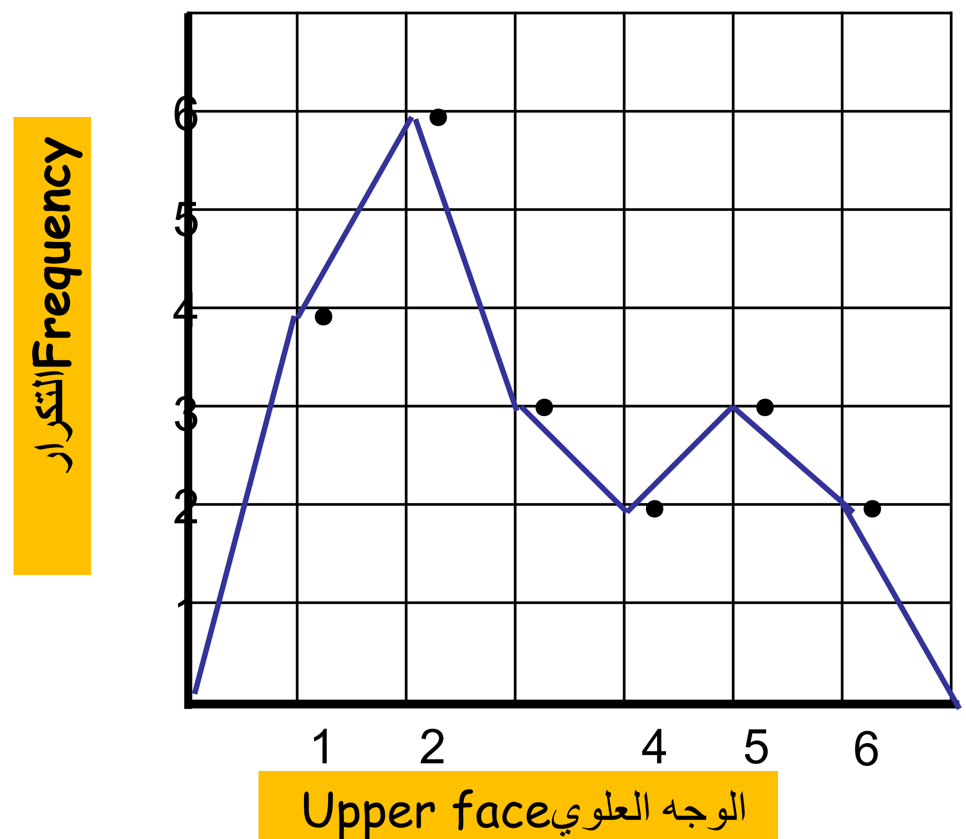
1. كل محور له عنوان
2. الرسم له عنوان
3. العمود يتوسط التدريج الأفقي
4. العمود الأول يبدأ بعد نصف وحدة من المحمور الرأسي

تمثيل البيانات بالمضلع التكراري

Represent the data in a frequency polygon :

الوجه العلوي Upper face	التكرار Frequency
1	4
2	6
3	3
4	2
5	3
6	2
total	20

ملاحظة الوجه العلوي لحجر النرد



ملاحظات على المضلع التكراري polygon

This is a type of line graph. Note that the axes are the same as for the histogram. You should note the following features.

- Each axis is labelled.
- The graph has a title.
- The dots showing the data are joined by straight lines.
- The first and last dot are connected to the horizontal axis as shown.
- The first score is one unit in from the vertical axis.

هو نوع من الرسم البياني بالأعمدة:

1. كل محور له عنوان
2. الرسم له عنوان
3. النقاط توصل مع بعضها بالمسطرة على شكل خطوط مستقيمة
4. القيمة الأولى توضع بعد وحدة واحدة من المحور الرأسي
5. توصل أول وآخر نقطة بالمحور الأفقي كما شاهدنا

المدرج التكراري والمضلع التكراري Frequency histogram & frequency polygon

The ages of a random sample of 25 students are given below : : أعمار عينة عشوائية من 25 طالب كما يلي

12 , 13 , 16 , 15 , 14 , 13 , 16 , 15 , 13 , 12 , 12 , 13 , 15 , 14 , 15 , 13 , 14 , 13 ,15 , 12 , 15 , 13 , 14 , 12 , 13

* Complete the frequency table to represent the data

* أكمل الجدول التكراري لتعرض هذه البيانات

* ارسم المدرج التكراري والمضلع التكراري

* Draw the frequency histogram and the frequency polygon

age	tally	(f)
total		

المحاضرة السادسة والسابعة

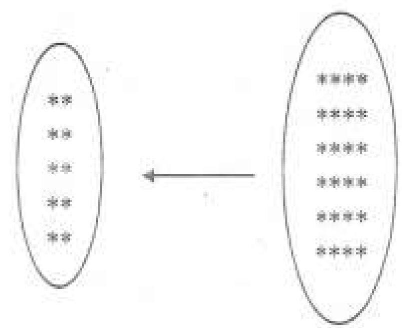
مقاييس النزعة المركزية والتشتت

الجلسة الأولى :
(مقاييس النزعة المركزية)

الجلسة الثانية:
(مقاييس التشتت)

تقديرات

معالم المجتمع Parametric		إحصاءات العينة Statistic	
المتوسط الحسابي(ميو)	μ	المتوسط الحسابي(إكس بار)	$\bar{x}$
التباين(سيجما تربيع)	σ^2	التباين	S^2
الانحراف المعياري(سيجما)	σ	الانحراف المعياري	S
الارتباط (رو)	ρ	الارتباط	r
معامل الانحدار (بيتا)	β	معامل الانحدار	b



نحسب هذه التقديرات من خلال دراسة جميع أفراد المجتمع (الحصر الشامل)

نحسب هذه التقديرات من خلال دراسة عينة ممثلة لأفراد المجتمع (طريقة المعاينة)

المحاضرة السادسة:
(مقاييس النزعة المركزية)

من المتوقع من الطالب في نهاية هذه الجلسة أن يكون قادرا على:

1. شرح معنى النزعة المركزية.
2. تعريف المتوسط الحسابي(الوسط الحسابي).
3. حساب المتوسط الحسابي من البيانات
4. تعريف الوسيط
5. حساب الوسيط من البيانات
6. تعريف المنوال.
7. حساب المنوال من البيانات

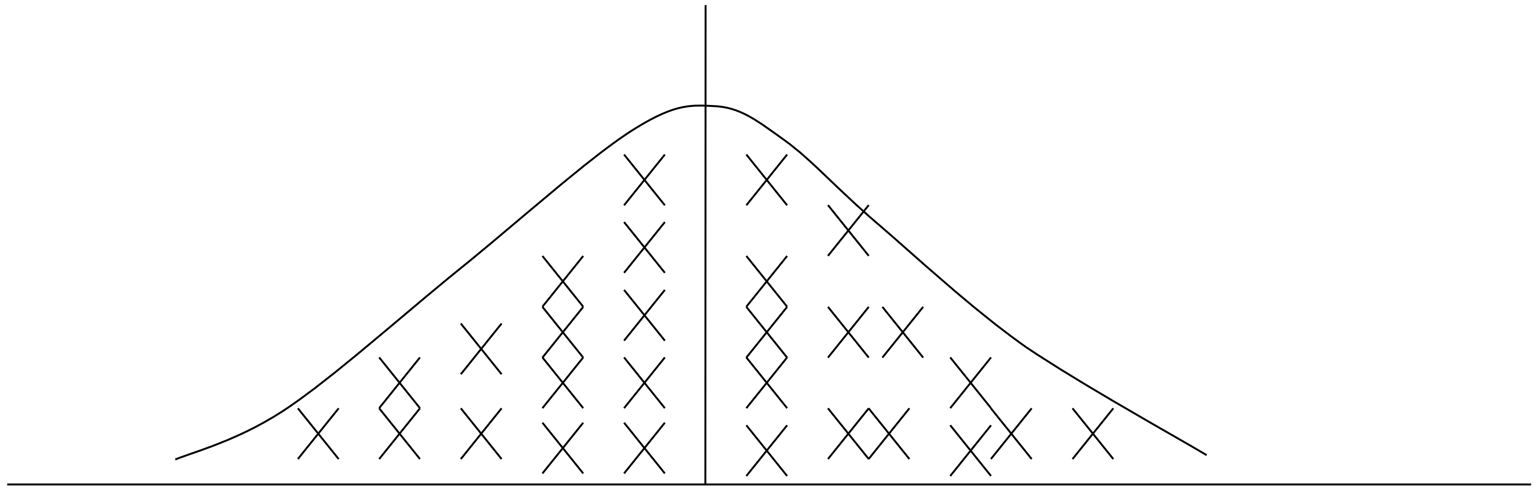
الجلسة الأولى :
(مقاييس النزعة المركزية)

عناصر الجلسة الأولى: **(75 دقيقة)**

1. مفهوم النزعة المركزية.
2. المتوسط الحسابي (الوسط الحسابي).
3. الوسيط.
4. المنوال.

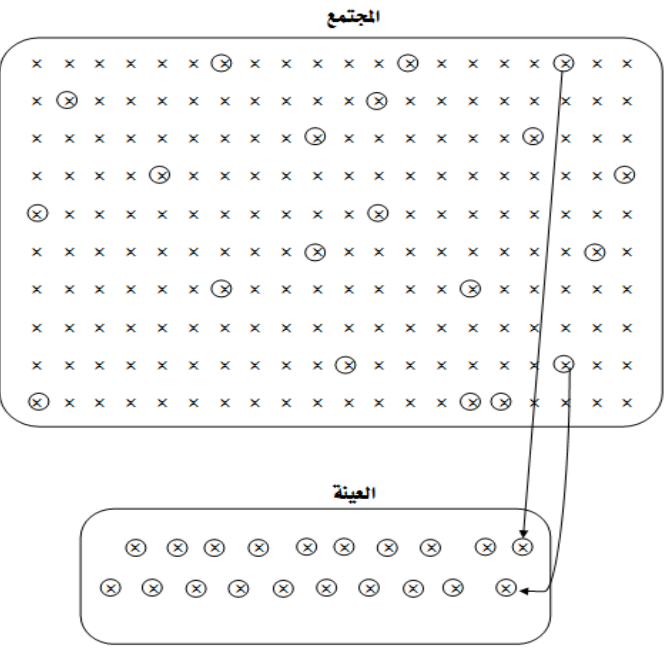
معنى النزعة المركزية :
هي ميل البيانات للتجمع حول المركز كما في الشكل التالي:

النزعة المركزية في حد ذاتها ظاهرة



المتوسط الحسابي

أهم مقاييس النزعة المركزية وأكثرها انتشارا، ويرمز له بالرموز الآتية:



الرمز	معناه	كيفية نطقه	طريقة حسابه
μ	المتوسط الحسابي للمجتمع (معلمة)	ميو	μ المتوسط
$\bar{x} =$	المتوسط الحسابي للعينة (تقدير) أو (إحصاءة)	إكس بار	$\bar{x}$ المتوسـ (تقدير)

مجموع

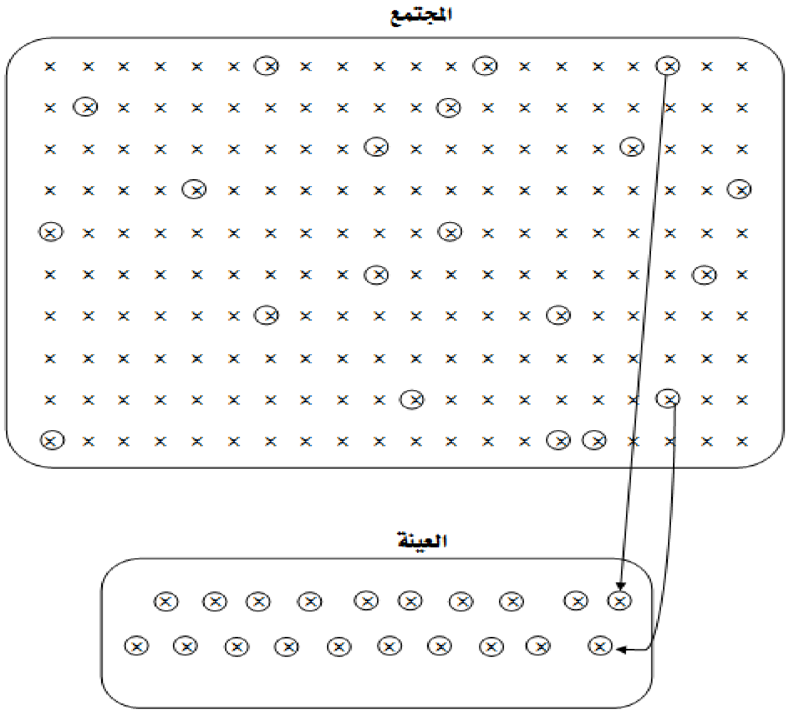
البيانات

$$\mu = \frac{\sum x}{N}$$

المتوسط الحسابي للمجتمع

حجم المجتمع

المعادلة تعني : أن المتوسط الحسابي للمجتمع يساوي مجموع البيانات تقسيم حجم المجتمع



المتوسط الحسابي

المتوسط الحسابي من البيانات الخام

مثال : احسب متوسط البيانات التالية:

3 , 5 , 1 , 4 , 2 , 3 , 5 , 1 , 4 , 2 , 3 , 5 , 1 , 4 , 2, 3 , 5 , 1 , 4 , 2

$$\bar{x} = \frac{3+5+1+4+2+3+5+1+4+2+3+5+1+4+2+3+5+1+4+2}{20}$$

$$\bar{x} = \frac{60}{20} = 3$$

50%

50%

الوسيط

رتب البيانات ترتيبا تصاعديا وحدد رتبة الوسيط

أ- إذا كان (n) فرديا ، فإن رتبة الوسيط هي : $\frac{n+1}{2}$.

ب- إذا كان (n) زوجيا، فإن رتبة الوسيط هي: $\frac{n}{2}$ و $\frac{n+1}{2}$.

الوسيط للأعداد الفردية هي: القيمة التي تقابل رتبته و قيمة

الوسيط لأعداد الزوجية هي: متوسط القيمتين اللتين تقابل رتبتي الوسيط.

الوسيط

مثال (1): أوجد قيمة الوسيط للبيانات التالية:

$$27\ ,\ 12\ ,\ 15\ ,\ 300\ ,\ 22\ ,\ 18\ ,\ 14\ ,\ 24\ ,\ 17$$

الحل: 1- ترتيب البيانات: $12\ ,\ 14\ ,\ 15\ ,\ 17\ ,\ 18\ ,\ 22\ ,\ 24\ ,\ 27\ ,\ 300$

الوسيط

2- نحدد موقع الوسيط (رتبته) كالتالي:

بما أن عدد البيانات : $n = 9$ وهو عدد فردي، إذن رتبة الوسيط هي:

$$\frac{n+1}{2} \quad \text{أي} \quad \frac{9+1}{2} = 5$$

3- إذن قيمة الوسيط هي القيمة الخامسة في البيانات المرتبة وهي:

$$\text{Med} = 18$$

الوسيط

مثال (2): أوجد قيمة الوسيط للبيانات التالية:

$$3 , 9 , 12 , 16 , 14 , 33 , 10 , 15 , 20 , 7$$

الحل: 1- ترتيب البيانات: $3 , 7 , 9 , 10 , 12 , 14 , 15 , 16 , 20 , 33$

القيمتين الوسيطتين

2- تحديد موقع الوسيط: عدد البيانات $n = 10$ ، بما أن عدد البيانات عدد زوجي، فإن

الوسيط له رتبتين: $\frac{10}{2} = 5$ و $\frac{10}{2} + 1 = 6$

3- نحدد قيمة الوسيط، وهي متوسط القيمتين الوسيطتين: الخامسة والسادسة، وهي:

$$\frac{12+14}{2} = 13$$

المنوال Mode

المنوال لمجموعة من القيم هو: القيمة التي تتكرر أكثر من غيرها ، أو القيمة الأكثر شيوعاً . وقد لا يكون للقيم منوال وقد يوجد أكثر من منوال واحد .

مـثال 1 - المجموعة 22,5,7,9,9,9,10,10,11,12,18
لـها منوال واحد وهو 9

مـــثال 2 – المجموعة 3,5,8,10,15,16
ليس لـها منوال

مثال 3 - المجموعة 2,3,4,4,4,5,5,7,7,7,9
لـها منوالان وهما 7,4 وتسمى مجموعة ذات منوالين

المحاضرة السابعة:
(مقاييس التشتت)

من المتوقع من الطالب في نهاية هذه الجلسة أن يكون قادرا على:

1. شرح معنى تشتت البيانات.
2. حساب المدى.
3. حساب التباين.
4. حساب الانحراف المعياري.
5. حساب التغاير.

الجلسة الثانية:
(مقاييس التشتت)

عناصر الجلسة الأولى: **(75 دقيقة)**

1. مفهوم التشتت.
2. المدى.
3. التباين.
4. الانحراف المعياري.
5. التغاير.

التشتت

المجموعة الأولى : , 14 , 12 , 10 , 8 , 6 , 5

المجموعة الثانية : , 18 , 15 , 10 , 5 , 2 , 1

19

احسب المتوسط لكل مجموعة.

ما الشي المختلف بين المجموعتين.

0 1 2 3 4 5 6 7 8 9 1 1 1 1 1 1 1 1 1 1 2

المدى

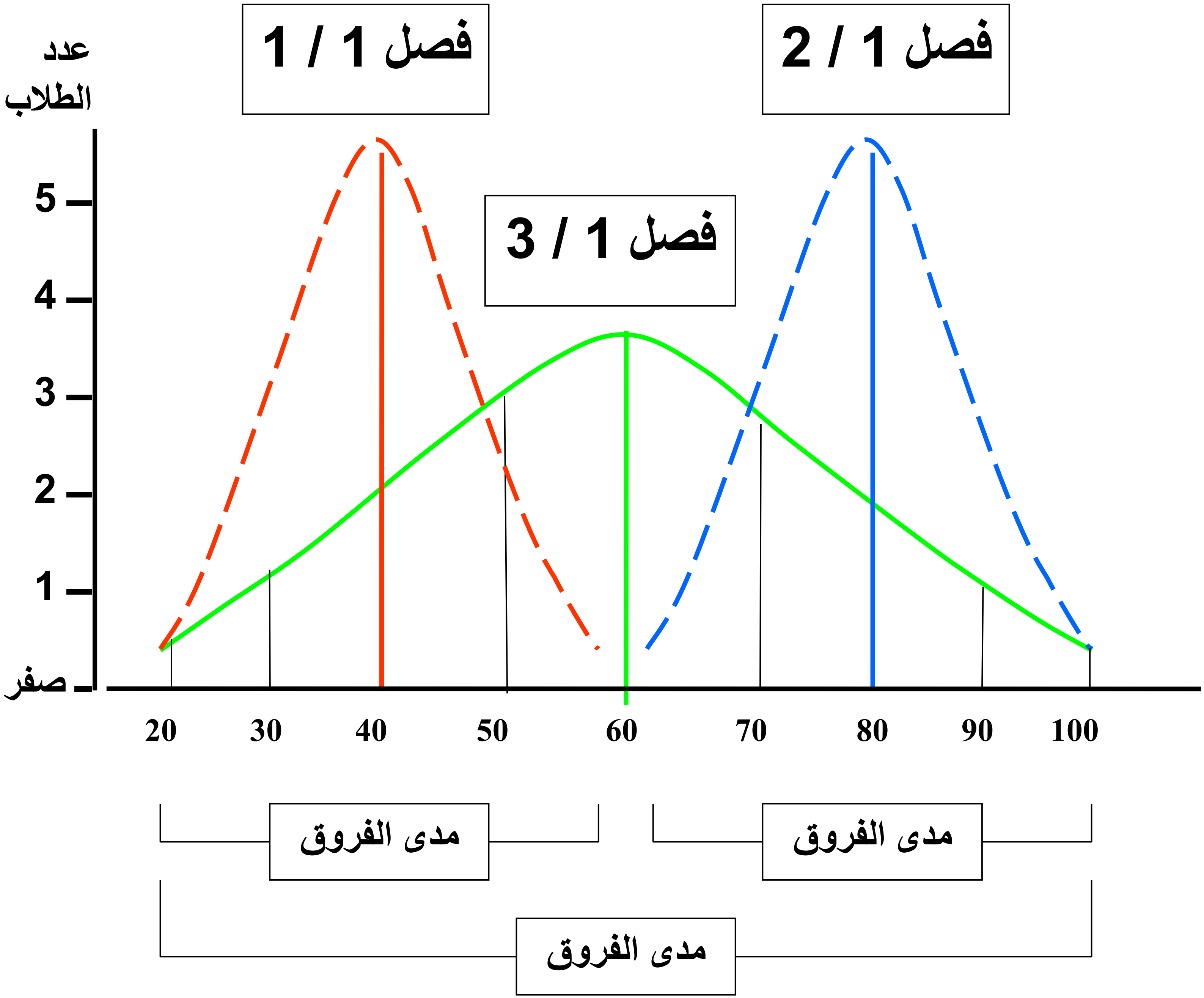
الفرق بين أكبر قيمة وأصغر قيمة في مجموعة ما

A : 5 , 6 , 8 , 10 , 12 , 14 , 15

Range = 15-5 =10

B : 1 , 2 , 5 , 10 , 15 , 18 , 19

Range = 19 -1 =18



عدد الطلاب
فصل 1 / 1
فصل 1 / 2
فصل 1 / 3
5
4
3
2
1
صفر
20
30
40
50
60
70
80
90
100
مدى الفروق
مدى الفروق
مدى الفروق

التباين

$$S_x^2 \qquad\qquad \sigma^2$$

هو مقياس لاختلاف البيانات وتشتتها، وهو متوسط مربعات انحرافات القيم عن وسطها الحسابي، ويرمز له بالرمز S^2 ويحسب من الصيغة الرياضية الآتية:

تباين متغير x

$$S_x^2 = \frac{\sum(x-\overline{x})^2}{n}$$

التباين

S_x^2 σ^2

تباين متغير x

$$S_x^2 = \frac{\sum (x-\overline{x})^2}{n}$$

مثال : أوجد تباين العينة الممثلة بالبيانات التالية:

x_1	x_2	x_2	x_4	x_5	x_6
5	8	4	7	4	2

2 3 4 $\boxed{5}$ 6 7 8

4

التباين

(الحل)

$$S_x^2 = \frac{\sum(x-\overline{x})^2}{n}$$

$$= \frac{(5-5)^2+(5-8)^2+(5-4)^2+(5-7)^2+(5-4)^2+(5-2)^2}{6}$$

$$= \frac{(0)^2+(-3)^2+(1)^2+(-2)^2+(1)^2+(3)^2}{6}$$

$$= \frac{0+9+1+4+1+9}{6} = \frac{24}{6} = 4$$

الانحراف المعياري

يعرّف الانحراف المعياري لعينة حجمها (n) مسحوبة من مجتمع ما بأنه الجذر التربيعي لتباين هذه البيانات وبالتالي فإن الانحراف المعياري للبيانات $x_1 , x_2 , \dots x_n$ والتي وسطها الحسابي $\bar{x}$

$$s_x^2$$

الانحراف المعياري

هل الانحراف المعياري دائما أقل من التباين؟

التباين	الانحراف المعياري
1	1
0.25	0.5
4	2

التغاير

هو التباين بين متغيرين ، أي الاختلاف بين بيانات المتغيرين S_{xy}

التباين داخل متغير x $S_x^2 = \frac{\sum(x-\bar{x})^2}{n}$ الانحراف المعياري $S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$

التباين داخل متغير y $S_y^2 = \frac{\sum(y-\bar{y})^2}{n}$ الانحراف المعياري $S_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}}$

التغاير S_{xy}: هو التباين بين متغير x ومتغير y

$$S_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{n}$$

التغاير

التباين حالة خاصة من التغاير:

مدى التباين من صفر إلى مالا نهاية $0 \leq \sigma_x^2 \leq \infty$

مدى التغاير من ما لا نهاية إلى ما لا نهاية $\infty \leq \sigma_{xy} \leq \infty$

مجموع التباين داخل كل متغير مع التباين بين المتغيرين يسمى (التباين الكلي)

التغاير

المتغير x : 15 , 14 , 12 , 10 , 8 , 6 , 5

المتغير y : 19 , 18 , 15 , 10 , 5 , 2 , 1

المطلوب:

- تباين المتغير X
- تباين المتغير y
- تغاير xy

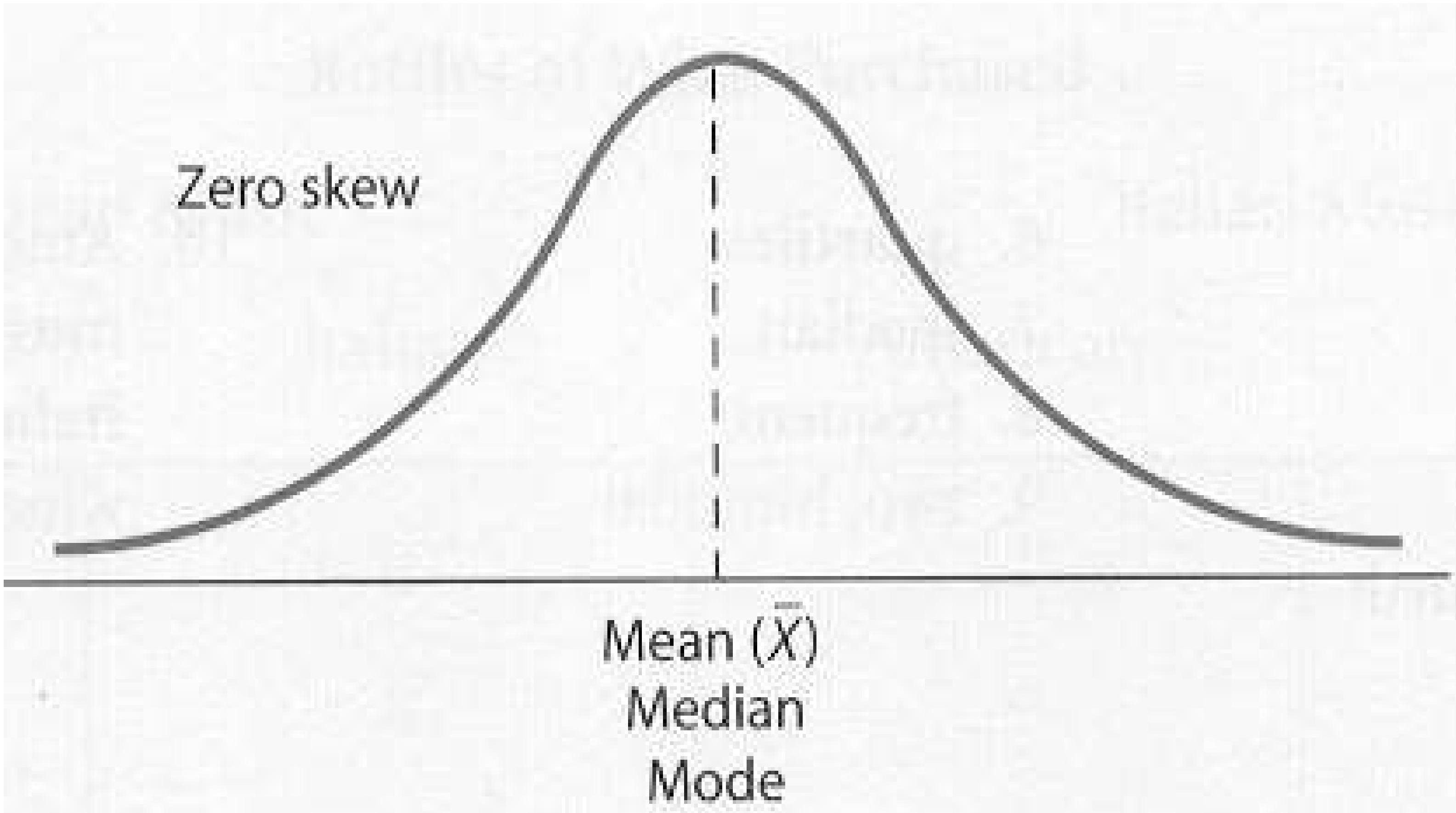
الالتواء

يبين الالتواء مدى انحراف التوزيع عن التوزيع المتماثل. ويكون التوزيع ملتويا عندما لا تتساوي كثافة التكرارات على جانبي المتوسط. فإذا كان ذيل التوزيع في الطرف الأيمن كان التوزيع ملتويا لليمين، وإذا كان ذيل التوزيع في الطرف الأيسر كان التوزيع ملتويا لليسار. وتستخدم عدة معادلات لقياس الالتواء

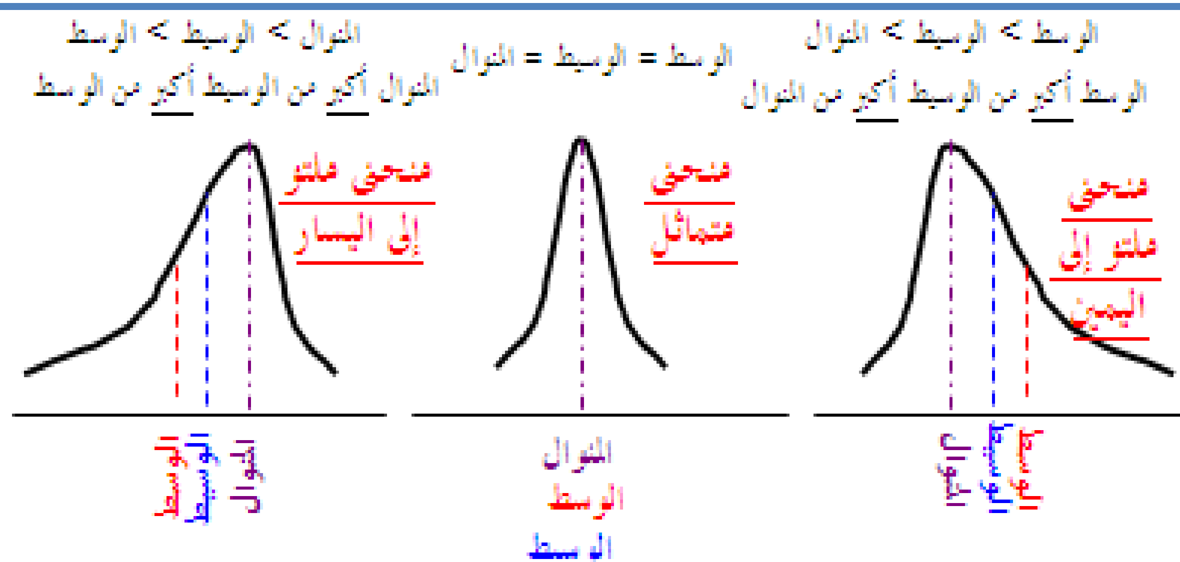
$$\text{معامل الالتواء} = \frac{\text{المتوسط} - \text{المنوال}}{\text{الانحراف المعياري}}$$

$$\text{معامل الالتواء} = \frac{3(\text{المتوسط} - \text{الوسيط})}{\text{الانحراف المعياري}}$$

العلاقة بين المتوسط والوسيط والمنوال والالتواء



العلاقة بين المتوسط والوسيط والمنوال والالتواء



في جميع المنحنيات المبينة ، لاحظ الآتي

المنوال هو القيمة المناظرة لأعلى نقطة في المنحنى

الوسيط يقع دائماً بين الوسط و المنوال

المحاضرة الثامنة والتاسعة

الإحصاء الحيوي
Biostatistics

الارتباط والانحدار الخطي البسيط
Correlation &
Simple Linear Regression

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

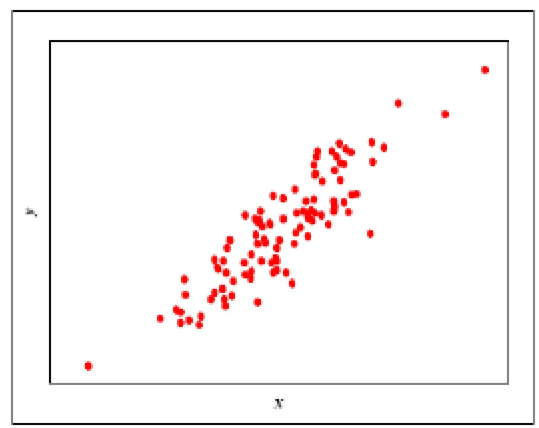
- الإرتباط هو تعيين طبيعة وقوة العلاقة بين متغيرين
- معامل الإرتباط correlation coeficient هو مؤشر هذه العلاقة
- اذا كان لدينا متغيران، المتغير X وهو متغير يتم تحديده من قبل الباحث أو الشخص ويسمى بالمتغير المستقل independent variable
- يرافق المتغيرX المتغير Y ويسمى بالمتغير التابع dependent variable وهو متغير تابع لأن نتيجته غير محددة وتعتمد على قيم المتغير المستقل.

الارتباط والانحدار الخطي البسيط & Correlation
Simple Linear Regression

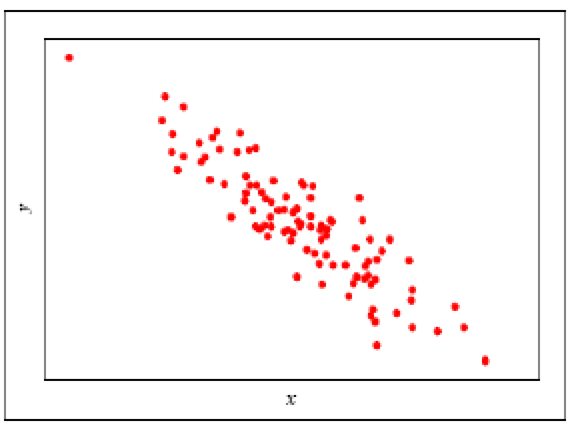
أنواع الارتباط

الارتباط الموجب (الطردي) Positive correlation : اذا تغير أحد المتغيرين X أو Y فإن المتغير الآخر يتبعه بنفس الاتجاه.

الإرتباط السالب (العكسي) Negative correlation : إذا تغير أحد المتغيرين X أو Y فإن المتغير الآخر بتبعه بالاتجاه المضاد.

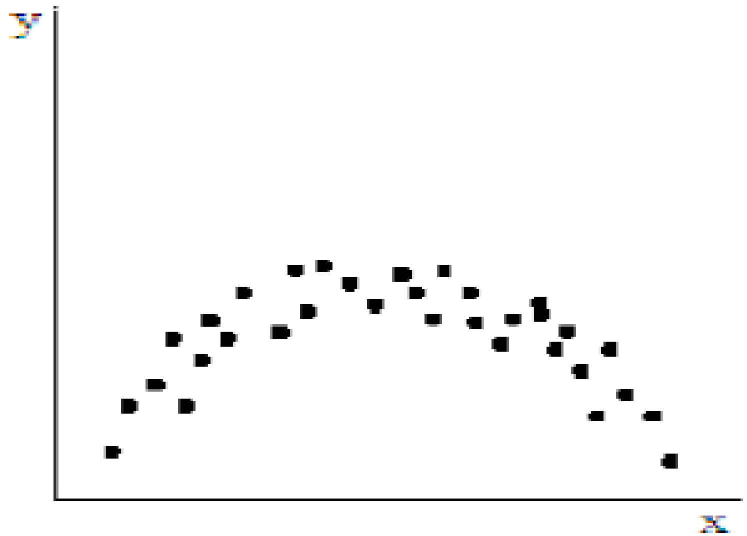


شكل الانتشار الخاص بالارتباط الموجب (الطردي)

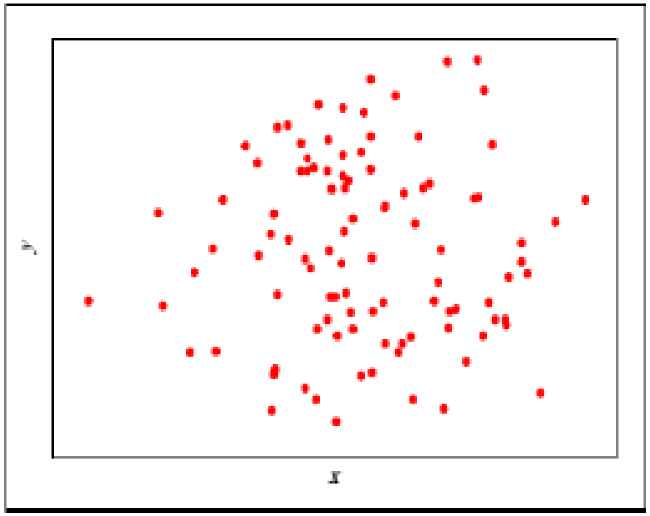


شكل الانتشار الخاص بالارتباط السالب (العكسي)

الارتباط والانحدار الخطي البسيط & Correlation
Simple Linear Regression

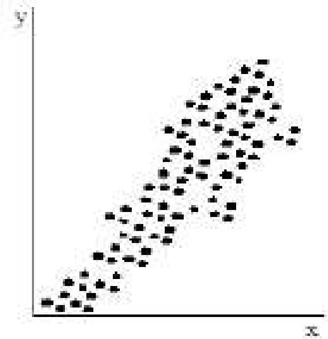


شكل الانتشار الخاص بالعلاقه الغير خطيه
بين متغيرين (ظاهرتين)

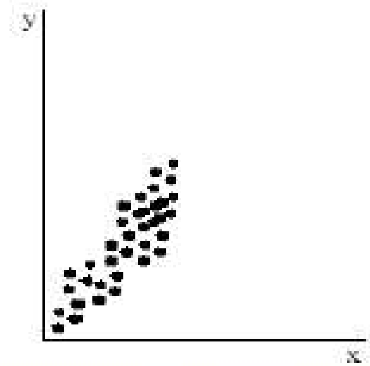


شكل الانتشار الخاص باستقلال
متغيرين (ظاهرتين)

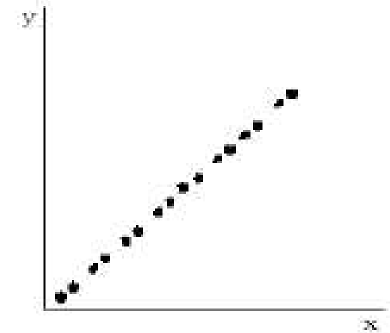
الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression



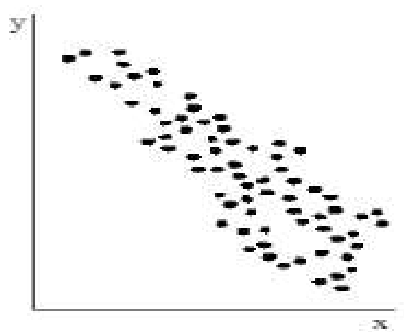
ارتباط طردي



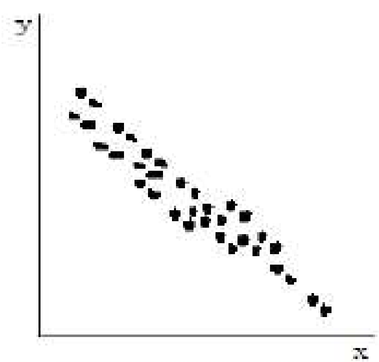
ارتباط طردي قوي



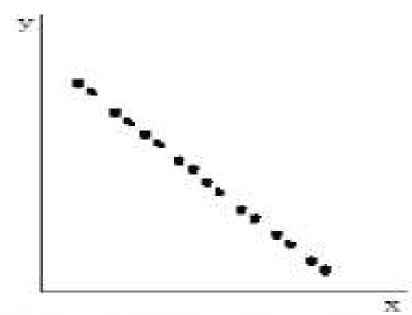
ارتباط طردي تام



ارتباط عكسي



ارتباط عكسي قوي



ارتباط عكسي تام

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

قياس الارتباط

- تستخدم معاملات الارتباط لقياس درجة الارتباط بين متغيرين(ظاهرتين).

معامل الارتباط :

- ويعرف معامل الارتباط والذي يرمز له بالرمز r عبارة عن مقياس رقمي يقيس قوة الارتباط بين متغيرين ، حيث تتراوح قيمته بين $(-1 \le r \le +1)$
- وتدل إشارة المعامل الموجبة على العلاقة الطردية ، بينما تدل إشارة المعامل السالبة على العلاقة العكسية .

الارتباط والانحدار الخطي البسيط & Correlation Simple Linear Regression

- الجدول التالي يوضح أنواع الارتباط واتجاه العلاقة وشكل الانتشار لكل نوع

قيمة معامل الارتباط	المعنى
+1	ارتباط طردي تام
من 0.70 إلى 0.99	ارتباط طردي قوي
من 0.50 إلى 0.69	ارتباط طردي متوسط
من 0.01 إلى 0.49	ارتباط طردي ضعيف
0	لا يوجد ارتباط

- وما قيل عن الارتباط الطردي ينطبق على الارتباط العكسي مع عكس الاشارة

معامل بيرسون للارتباط الخطي
Pearson Linear Correlation Coefficient

- معامل بيرسون للارتباط الخطي من أكثر معاملات الارتباط استخدامًا خاصة في العلوم الإنسانية و الاجتماعية .
- و مستوى القياس المطلوب عند تطبيق معامل بيرسون للارتباط هو أن يكون كلا المتغيرين مقياس فترة أو نسبي أو بمعنى اخر أن تكون بيانات كلا المتغيرين (الظاهرتين) بيانات كمية .

- حساب معامل بيرسون للارتباط الخطي :
- يمكن حساب معامل بيرسون بدلالة القراءات لبيانات المتغيرين باستخدام الصيغة التالية:

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

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

$\sum_{i=1}^{n} x_i y_i$: مجموع حاصل ضرب x في y

$\sum x$: مجموع قيم المتغير x

$\sum y$: مجموع قيم المتغير y

$\sum x^2$: مجموع مربعات قيم المتغير x

$\sum y^2$: مجموع مربعات قيم المتغير y

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

- مثال1 : سجلت ست قراءات تقريبية لحجم الإنتاج وحجم صادرات دواء معين باحدى الدول خلال عدة سنوات كما يلي:

حجم الصادرات (y)	2	2	2	1	1	1
حجم الإنتاج (x)	3	4	2	2	2	2

- الحل

	x	y	xy	x^2	y^2
	3	2	6	9	4
	4	2	8	16	4
	2	2	4	4	4
	2	1	2	4	1
	2	1	2	4	1
	2	1	2	4	1
$\sum$	15 $=\sum x$	9 $=\sum y$	24 $=\sum xy$	41 $=\sum x^2$	15 $=\sum y^2$

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

$$r_p = \frac{6(24) - (15)(9)}{\sqrt{((6\times 41) - 15^2)((6\times 15) - 9^2)}} =$$

$$\frac{144 - 135}{\sqrt{(246 - 225)(90 - 81)}} = \frac{9}{\sqrt{189}} = \frac{9}{13.75} = 0.65$$

علاقة طردية متوسطة

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

- مثال 2 البيانات التالية تمثل المتغيرين X,Y

Y	1	3	8	7	6	5	7	8	12	12
X	9	11	17	18	19	16	16	19	23	23

- احسب معامل الارتباط الخطي، ما مدى قوة العلاقة الخطية؟

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

	x	y	xy	x^2	y^2
	9	1	9	81	1
	11	3	33	121	9
	17	8	136	289	64
	18	7	126	324	49
	19	6	114	361	36
	16	5	80	256	25
	16	7	112	256	49
	19	8	152	361	64
	23	12	276	529	144
	23	12	276	529	144
$\sum$	171	69	1314	3107	585
	$=\sum x$	$=\sum y$	$=\sum xy$	$=\sum x^2$	$=\sum y^2$

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

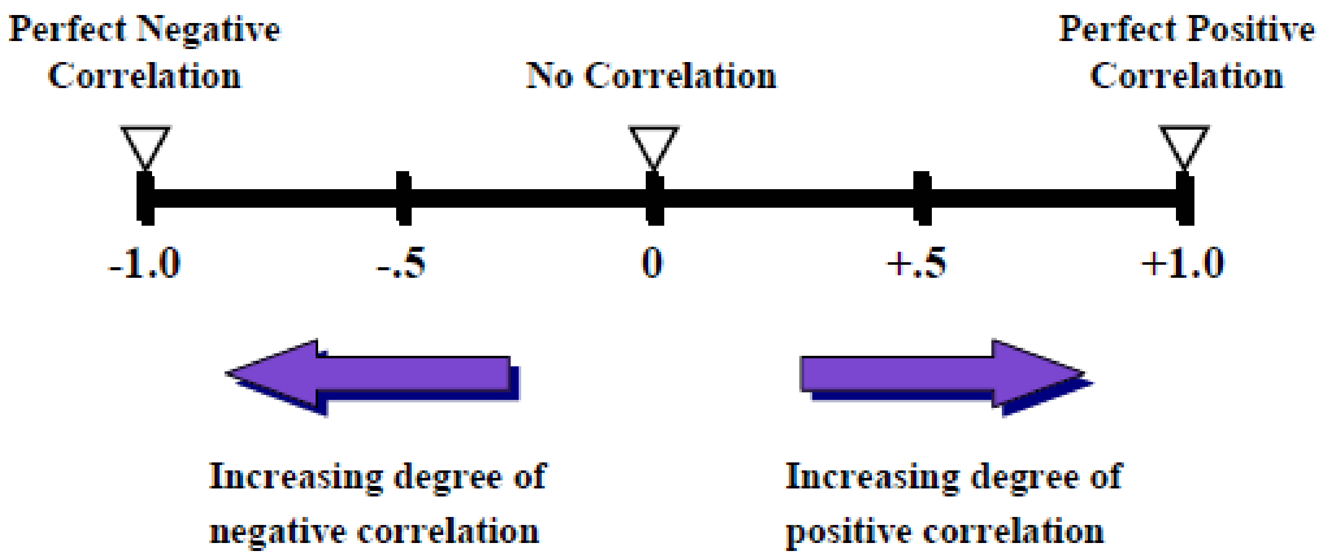
$$r_p = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

$$r_p = \frac{10(1314) - (171)(69)}{\sqrt{((10\times 3107) - 171^2)((10\times 585) - 69^2)}} =$$

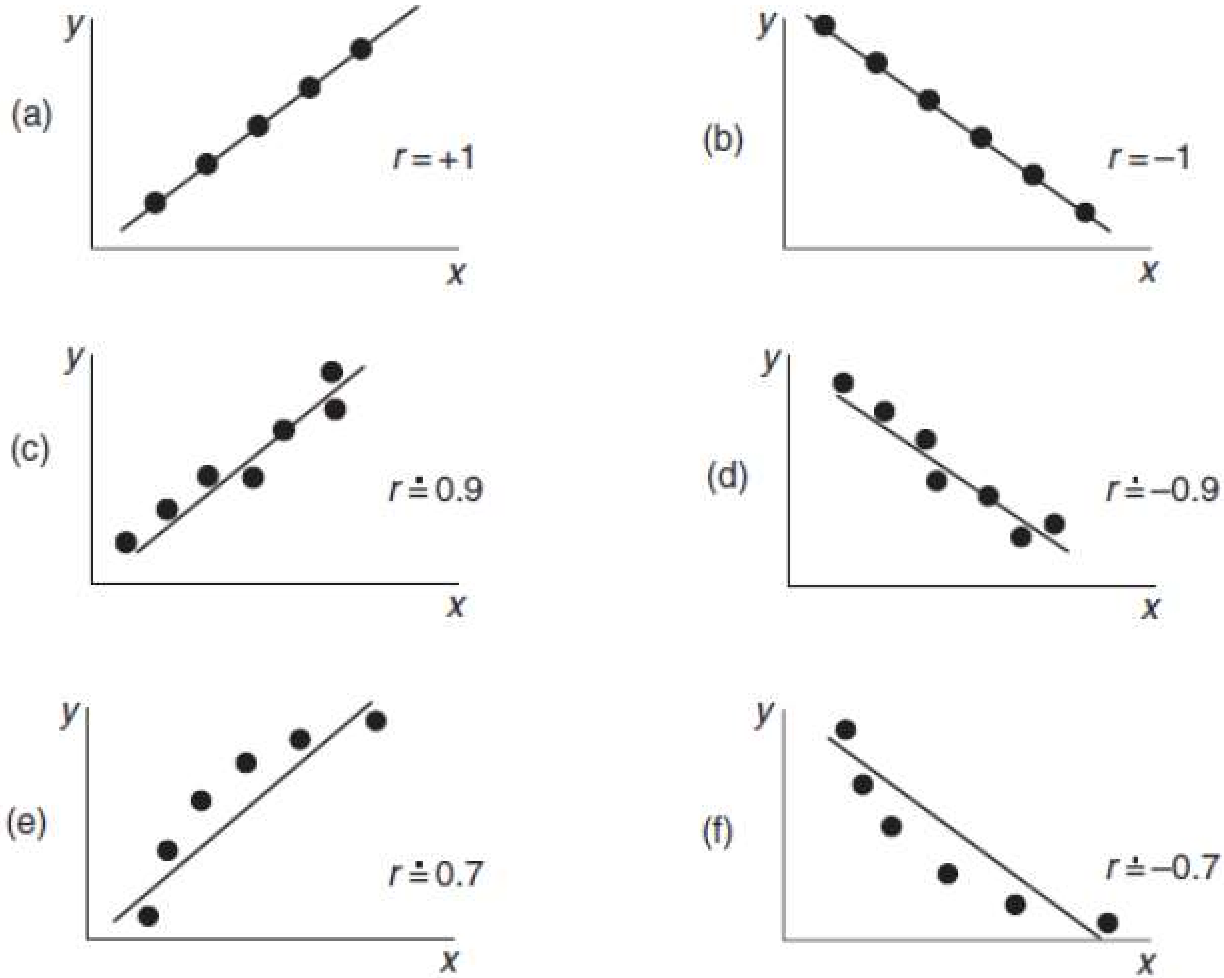
$$\frac{13140 - 11799}{\sqrt{(31070 - 29241)(5850 - 4761)}} = \frac{1341}{\sqrt{18291089}} = \frac{1341}{1411.30} = 0.95$$

<u>علاقة ارتباط طردية قوية</u>

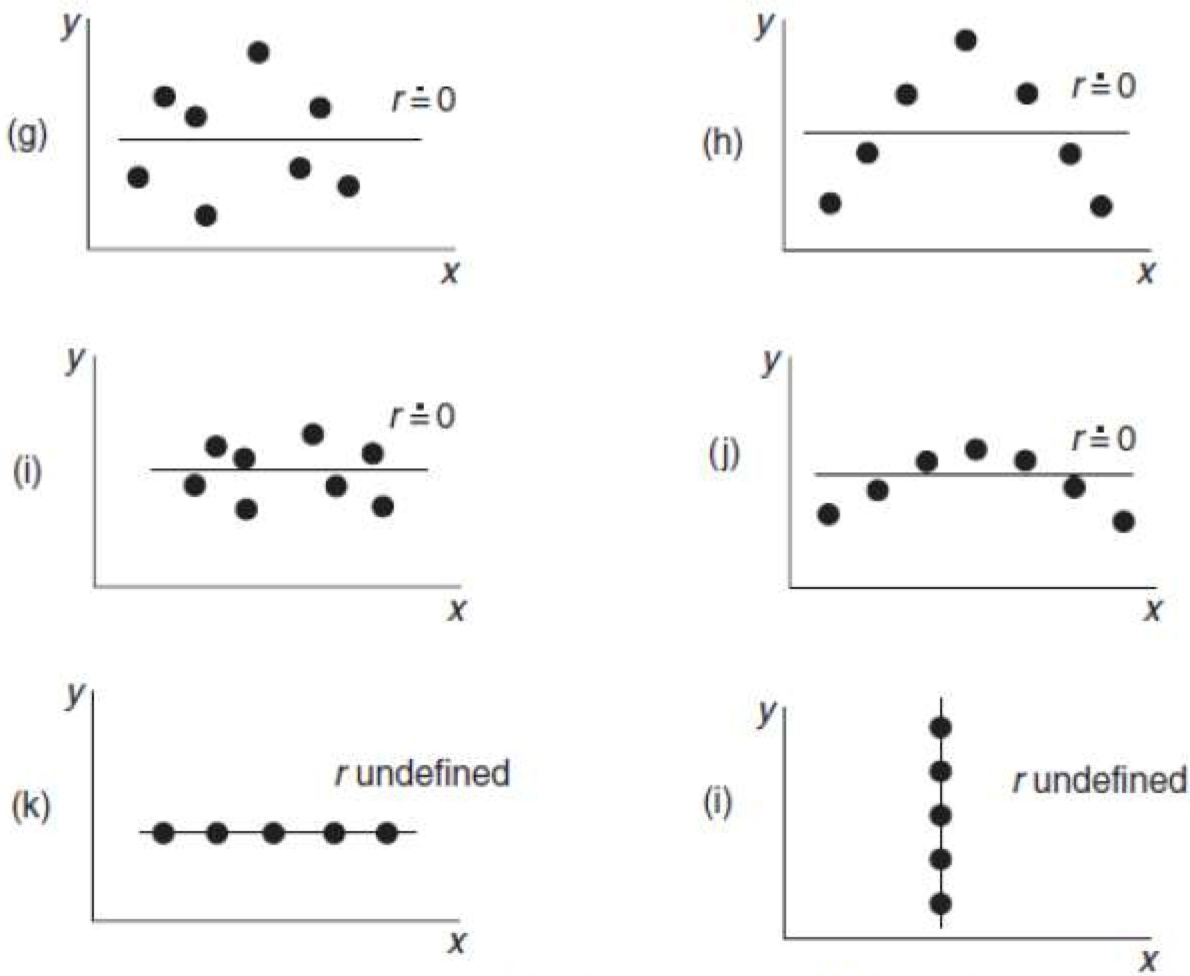
الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression



الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression



الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression



الانحدار الخطي البسيط Simple Linear Regression

- الانحدار هو أسلوب يمكن بواسطته تقدير قيمة أحد المتغيرين بمعلومية قيمة المتغير الآخر عن طريق معادلة الانحدار
- تهدف دراسة الانحدار التنبوء بقيمة متغير (Y)بمعرفة متغير آخر (X)
- ويعرف المتغير الأول بالمتغير التابع (dependent)ويرمز له Yفي حين يعرف المتغير الآخر بالمتغير المستقل (Independent)ويرمز له X
- لذا المتغير Xعرف بالمتغير المستقل في حين Yتتعين قيمتها تبعاً لقيمة Xلذا عرفت Yبالمتغير التابع)أي تبعاً لقيمة X
- كما أن الانحدار هنا بسيط لوجود متغيرين فقط تابع ومستقل، وعند ذكر كلمة الخط نعني بها خط الانحدار.

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

- الغرض من استخدام أسلوب تحليل الانحدار الخطي البسيط، هو دراسة وتحليل أثر متغير كمي على متغير كمي آخر، ومن الأمثلة النموذجية على تحليل الانحدار:
- اعتماد ضغط الدم Y على عمر الشخص X،
- أو اعتماد الوزن لحيوانات التجربة Y على معدل التغذية اليومي X هذا الارتباط والتابعية بين X و Y هي ما ندعوه بالانحدار أو الارتباط فنقول ارتباط Y ب X
- دراسة أثر كمية السماد على إنتاجية الدونم.
- دراسة أثر كمية البروتين التي يتناولها الأبقار على الزيادة في الوزن.

ويلاحظ من ذلك أن نموذج الانحدار يعتمد دائماً على علاقة السببية بمعنى ان يكون التغير في المتغير المستقل مسبب رئيسي للتغير في المتغير التابع.

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

- معادلة الخط المستقيم

$$\hat{y} = a + bx$$

حيث a : ثابت الانحدار أو الجزء المقطوع من محور y

b : ميل الخط المستقيم أو معامل انحدار Y على X (أو Y/X)

- وتحسب القيمتان a و b من العلاقتين التاليتين :

حيث:

$$a = \frac{\sum y - b\sum x}{n}$$

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

الارتباط والانحدار الخطي البسيط Correlation & Simple Linear Regression

لإيجاد قيمة مقدرة جديدة $\hat{y}_h$ نعوض بقيمة معلومة للمتغير المستقل ولتكن x_h في معادلة تقدير خط الانحدار Y/X

$$\hat{y} = a + bx$$

نعوض

الارتباط والانحدار الخطي البسيط Simple Linear Regression

لدراسة علاقة الاستهلاك المحلي (y)بالإنتاج (x)لمادة الإسفلت (بالمليون برميل) خلال عدة سنوات، أخذنا عشر قراءات تقريبية كما يلي :

y	6	8	9	8	7	6	5	6	5	5
x	10	13	15	14	9	7	6	6	5	5

أوجد معادلة الانحدار الخطي البسيط، وتوقع قيمة الاستهلاك عندما يصل إنتاج 16,000,000برميل .

- الحل :

	x	y	xy	x^2
	10	6	60	100
	13	8	104	169
	15	9	135	225
	14	8	112	196
	9	7	63	81
	7	6	42	49
	6	5	30	36
	6	6	36	36
	5	5	25	25
	5	5	25	25
$\sum$	90	65	632	942
	$=\sum x$	$=\sum y$	$=\sum xy$	$=\sum x^2$

$$b=\frac{n\sum xy-(\sum x)(\sum y)}{n\sum x^2-(\sum x)^2}=\frac{6320-(90)(65)}{9420-90^2}=\frac{6320-5850}{9420-8100}=\frac{470}{1320}=0.36$$

$$a=\frac{\sum y-b\sum x}{n}=\frac{65-(0.36\times 90)}{10}=3.26$$

∴معادلة خط الانحدار البسيط في هذه الحالة :

$$\hat{y}=3.26+0.36x$$

الارتباط والانحدار الخطي البسيط Simple Linear Regression

- ولتوقع قيمة الاستهلاك المحلي عندما يصل الإنتاج 16000000برميل، نحول وحدة هذه القيمة من برميل إلى مليون برميل بالقسمة على مليون أي أن القيمة المستخدمة في توقع الاستهلاك هي $x_h = 16$
- وبالتعويض في المعادلة السابقة نجد أن:

$$\hat{y}_h = a + bx_h$$
$$= 3.26 + 0.36(16) = 9.02$$

أي أن الاستهلاك قد يصل إلى **9.02** مليون برميل، أي ما يعادل **9020000** برميل خلال السنة.

المحاضرة العاشرة

الإحصاء الحيوي

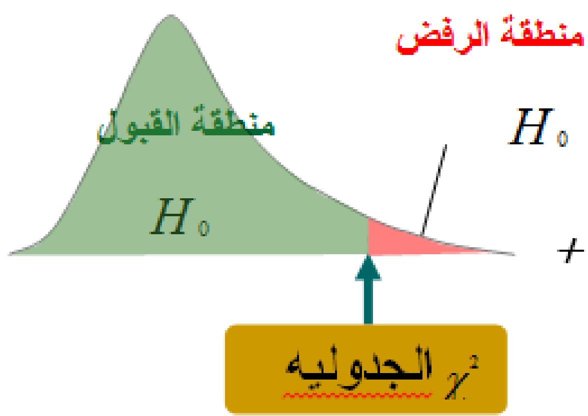
Biostatistics

اختبار كاي تربيع

Chi-Square test

اختبار كاي Chi-Square test

- يستخدم اختبار مربع كاي للبيانات الاسمية ، فالمتغيرات يجب أن تكون مصنفة و مقاسة بمقياس إسمي ، ويستخدم لموازنة التوزيعات التكرارية بالمتغيرات ، و يسمح بمعالجة البيانات النوعية التي تكون على شكل تكرارات لمجموعات أو أصناف معينة
- نتخذ القرار بناءً على قيمة إحصاء الاختبار مربع كاي (نحدد منطقة الرفض و منطقة القبول على الرسم)



Chi-Square test اختبار كاي

- مثال: يعتبر العنف من القضايا المهمة المنتشرة في المجتمع ، أخذت عينة من 300 شخص ذكور واناث بالغين ، وأجابوا عن أسئلة ما اذا كانوا مع قرارات الحكومة المتخذة حديثا لمعاقبة من يقومون بالعنف. صنفت الاجابات حسب الجدول التالي. هل تعطي العينة معلومات كافية وخلاصة بأن الاتجاهات والآراء لدى الجنسين تعتمد على بعضها (غير مستقلة)؟ $\alpha = .01$

Table 2 x 3 Contingency Table

	In favor	Against	No opinion	Row Total
Men (M)	93	70	12	175
Women (W)	87	32	6	125
Column Total	180	102	18	300

اختبار كاي Chi-Square test

- مجموع كل من الأعمدة لوحدها والصفوف لوحدها يسمى marginal frequencies
- مجموع كل التكرارات في الجدول يسمى Grand total

الفرضيات

العدمية H0 : الآراء والاتجاهات مستقلة عن بعضها

H_0: Gender and opinions of adults are independent

البديلة Ha : الآراء والاتجاهات غير مستقلة عن بعضها

H_a Gender and opinions of adults are dependent

نحسب التكرارات المتوقعة ، لكن قبل ذلك سنعطي رموزا لمدلولات الجدول

- F: An individual selected is in favor. أي الأفراد الذين مع القرارات
- A: An individual selected is against.أي الأفراد الذين ضد القرارات
- N: An individual selected is no opinion. أي الأفراد الذين ليس لهم رأي حول الموضوع
- M: An individual selected is man. الأشخاص الذكور
- W: An individual selected women. الأشخاص الإناث

Chi-Square test اختبار كاي

Using the marginal frequencies, we can list the following probabilities:

$$P(F) = \frac{180}{300}, P(A) = \frac{102}{300}, P(N) = \frac{18}{300}, P(M) = \frac{175}{300}, P(W) = \frac{125}{300}$$

$$P(F \text{ and } M) = P(F)P(M) = P(F)\ P(M) = \frac{180}{300} \times \frac{175}{300}$$

$$P(F \text{ and } W) = P(F)P(W) = P(F)\ P(W) = \frac{180}{300} \times \frac{125}{300}$$

$$P(A \text{ and } M) = P(N)P(E) = P(A)\ P(M) = \frac{102}{300} \times \frac{175}{300}$$

$$P(A \text{ and } W) = P(A)P(W) = P(A)\ P(W) = \frac{102}{300} \times \frac{125}{300}$$

$$P(N \text{ and } M) = P(N)P(E) = P(N)\ P(M) = \frac{18}{300} \times \frac{175}{300}$$

$$P(N \text{ and } W) = P(N)P(W) = P(N)\ P(W) = \frac{18}{300} \times \frac{125}{300}$$

Chi-Square test اختبار كاي

$$Expected\ frequency = \frac{(Row\ \ total)(Column\ total)}{Grand\ total} \qquad \frac{(180)(175)}{300} = 105$$

Table Observed and expected frequency

	In Favor (F)	**Against (A)**	**No Opinion (O)**	**Row Totals**
Men (M)	93 (105.00)	70 (59.50)	12 (10.50)	175
Women (W)	87 (75.00)	32 (42.50)	6 (7.50)	125
Column Totals	180	102	18	300

Chi-Square test اختبار كاي

- نحسب قيمة الاختبار **Compute the test statistic** حسب العلاقة التالية
- حيث أن O= Observed ، E= Expected

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} \ldots\ldots\ldots\ldots\ldots + \frac{(O_6 - E_6)^2}{E_6}$$

$$= \frac{(93 - 105.00)^2}{105.00} + \frac{(70 - 59.50)^2}{59.50} + \frac{(12 - 10.50)^2}{10.50}$$

$$+ \frac{(87 - 75.00)^2}{75.00} + \frac{(32 - 42.50)^2}{42.50} + \frac{(6 - 7.50)^2}{7.50}$$

$$= 1.371 + 1.853 + .214 + 1.920 + 2.594 + .300 = 8.252$$

اختبار كاي Chi-Square test

- تحديد درجة الحرية الخاصة باختبار كاي **number of degrees of freedom**
- **صيغة تحديد درجة الحرية الخاصة باختبار كاي هي**

$$df = (r - 1)(c - 1).$$

حيث r = row ، c = column

$$df = (2 - 1)(3 - 1) = 2 \text{ degrees of freedom.}$$

نوجد القيم الحرجة critical value من جدول كاي عند α = 0.01 ودرجة حرية 2

وتساوي $\chi^2_{0.01} = 9.21$

اختبار كاي Chi-Square test

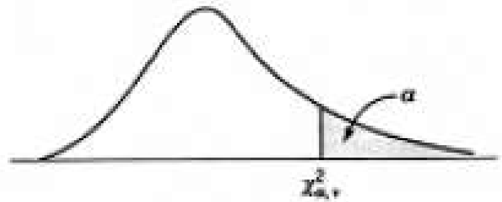
					α				
ν	0.995	0.990	0.975	0.950	0.500	0.050	0.025	0.010	0.005
1	0.00 +	0.00 +	0.00 +	0.00 +	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	1.39	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.38	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59

اختبار كاي Chi-Square test

- نقارن قيمة كاي المحسوبة χ^2 مع قيمة كاي المجدولة $\chi^2_{0.01}$
- وحيث أن قيمة كاي المحسوبة =8.252 وهي أقل من القيمة المجدولة لكاي 9.21 عند درجة حرية 2. بمعنى اخر لا تقع داخل منطقة الرفض.

- القرار: لايمكن رفض الفرضية العدمية عند مستوى دلالة 01.
- الاستنتاج : نستنتج أن الجنس والآراء للأشخاص البالغين لايعتمد كل منهما على الآخر أو أنهم مستقلين. أو ببساطة (ليس هنالك علاقة بين الجنس والآراء للأشخاص البالغين في الدراسة السابقة)

Percentage Points of the χ^2 Distribution*



	α								
ν	0.995	0.990	0.975	0.950	0.500	0.050	0.025	0.010	0.005
1	0.00 +	0.00 +	0.00 +	0.00 +	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	1.39	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.38	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	9.34	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	10.34	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	11.34	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	12.34	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	13.34	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	14.34	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	15.34	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	16.34	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	17.34	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	18.34	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	19.34	31.41	34.17	37.57	40.00
25	10.52	11.52	13.12	14.61	24.34	37.65	40.65	44.31	46.93
30	13.79	14.95	16.79	18.49	29.34	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	39.34	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	49.33	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	59.33	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	69.33	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	79.33	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	89.33	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	99.33	124.34	129.56	135.81	140.17

* ν = degrees of freedom.