المحاضره الاولى

عرض البيانات الإحصائية ووصفها

تعريف :

- يختص علم الإحصاء بجمع وتبويب وتنظيم وعرض وتلخيص وتحليل واستقراء النتائج من البيانات الرقمية لعمل استدلالات واتخاذ القرارات في ظل عدم التأكد.

وينقسم علم الإحصاء إلي قسمين:

أ- الإحصاء الوصفي: يشمل جمع وتبويب وتنظيم وعرض وتلخيص البيانات الرقمية.

ب- الإحصاء الاستقرائي (التحليلي): تحليل البيانات واستقراء النتائج واتخاذ القرارات

ينقسم علم الإحصاء إلي :

علم الإحصاء

<u>الإحصاء الوصفي</u>
جمع وتبويب وتنظيم وعرض
وتلخيص البيانات الرقمية.

<u>الإحصاء الاستقرائي (التحليلي)</u>
تحليل البيانات واستقراء
النتائج واتخاذ القرارات

طرق عرض البيانات:

أولا:

- طريقة الجداول: وهي وضع البيانات في جداول لعرض تغير ظاهرة مع الزمن أو مع مسميات كالبلدان والمدارس والجامعات وغيرها أو مع الزمن والمسميات معاً .وعند استعمال هذه الطريقة يجب مراعاة :

(١)عنوان الجدول .

(٢) الوحدات المستعملة .

(٣) مذكرات المصادر التي أخذت منها البيانات.

(٤) مذكرات تفسر القيم الشاذة أن وجدت.

مثال (١):

الجدول التالي يمثل عدد الطالبات المقبولات في جامعة الدمام في الفترة من 1420ه إلي 1427ه .

لاحظ انه يعرض تغير ظاهره مع الزمن ومستوفي شروط العرض بالجدول .

عدد الطالبات المقبولات في جامعة الدمام في الفترة من 1420ه إلي 1427ه .

جدول (١)

عدد الطالبات	السنة
350	1420
450	1421
500	1422
700	1423
920	1424
1080	1425
1450	1426
1800	1427

المصدر: جامعة الدمام شؤون الطالبات .

ثانيا:

- العرض البياني للبيانات الإحصائية:

١- طريقة المستطيلات أو الأعمدة: توضع المسميات أو الزمن علي المحور الأفقي أو العمودي ورسم مستطيل علي كل مسمي ويكون ارتفاع كل مستطيل ممثلا للقيمة المقابلة.
وتستعمل هذه الطريقة للمقارنة بين قيم الظواهر حسب الزمن أو المسميات وأيضا للمقارنة بين قيم الظواهر حسب المسميات علي مدي عدة سنوات.

ولأستعمال الطريقة يجب مراعاة الأتي:

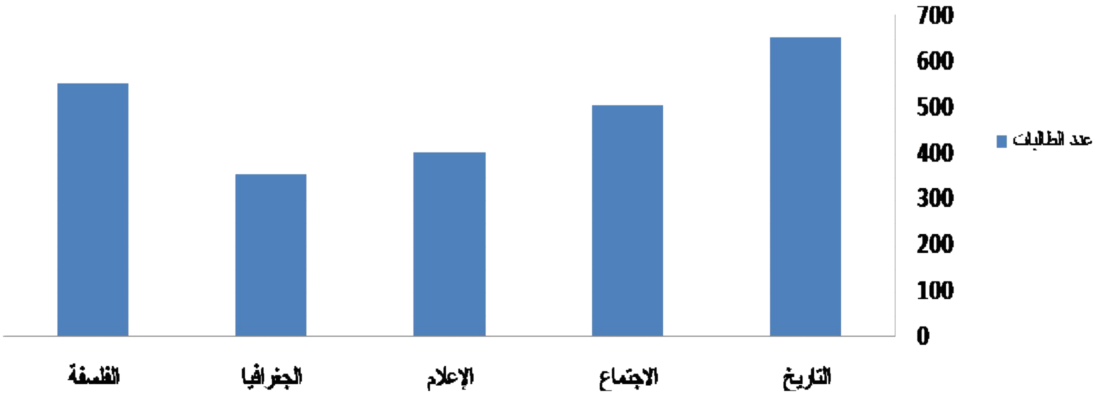
١. عنوان الشكل.
٢. الوحدات المستعملة .
٣. مذكرات المصادر التي أخذت منها الإشكال.

مثال (٢):

الجدول التالي يوضح أعداد الطالبات ببعض أقسام كلية الآداب جامعة الدمام والمطلوب عرض هذه البيانات باستخدام طريقة المستطيلات والأعمدة:

جدول (٢) :

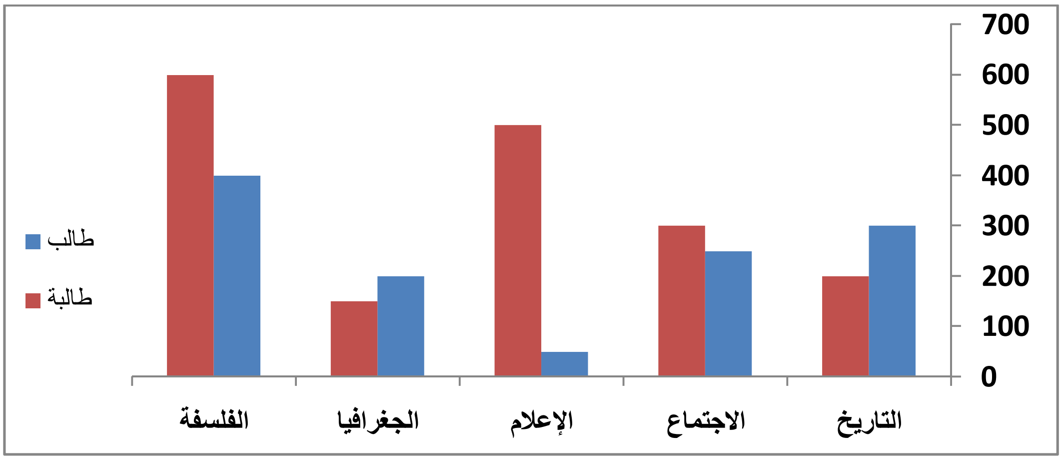
القسم	**التاريخ**	**الاجتماع**	**الإعلام**	**الجغرافيا**	**الفلسفة**
عدد الطلاب	650	500	400	350	550



مثال(٣):

الجدول التالي يوضح أعداد الطلاب و الطالبات ببعض أقسام كلية الآداب جامعة الدمام والمطلوب عرض هذه البيانات باستخدام طريقة المستطيلات والأعمدة:

القسم	**التاريخ**	**الاجتماع**	**الإعلام**	**الجغرافيا**	**الفلسفة**
عدد الطلاب	300	250	50	200	400
عدد الطالبات	200	300	500	150	600



٢/ طريقة الخط المنكسر:

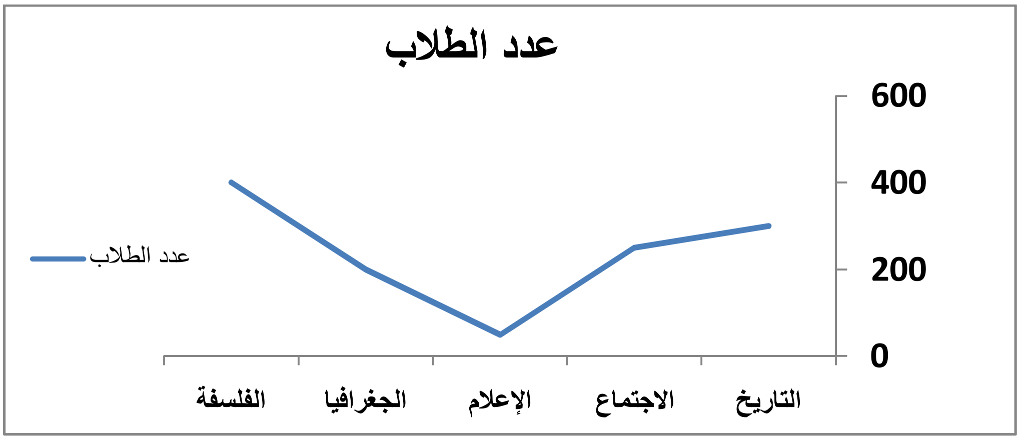
تستعمل هذه الطريقة البيانية الناتجة من:

- تغير ظاهرة مع الزمن (مثل تغير درجة حرارة مريض مع الزمن بالساعات).
- تغير ظاهرة مع مسميات(مثل تغير أعداد الطلاب في جامعة مع السنوات).
- تغير ظاهرتين معاً (مثل تغير أعداد الطلاب حسب الكليات علي مدي فترة زمنية محددة).

مثال (٤):

الجدول التالي يوضح أعداد الطلاب ببعض أقسام كلية الآداب جامعة الدمام والمطلوب عرض هذه البيانات باستخدام طريقة الخط المنكسر:

القسم	التاريخ	الاجتماع	الإعلام	الجغرافيا	الفلسفة
عدد الطلاب	300	250	50	200	400



٣/ طريقة الخط المحني:

وهي نفس طريقة الخط المنكسر ولكن بدون زوايا

مثال :

ارسمي الخط المنحني لبيانات - المثال (٤)

٤- طريقة الدائرة:

تستعمل هذه الطريقة بتقسيم المتغير أو الظاهرة الكلية إلي أجزاء ، حيث أ ن المجموع الكلي للظاهرة يساوي مساحة الدائرة 360 .ثم يتم تحويل كل جزء إلي قطاع من الدائرة ..

زاوية القطاع = (قيمة الجزء÷المجموع الكلي) × 360

تستعمل الدائرة لتمثيل متغير أو ظاهرة واحدة فقط.

مثال(٤)

أرسمي الدائرة لبيانات الجدول في المثال (٢) .

الحل:

زاوية القطاع = (قيمة الجزء÷المجموع الكلي) × 360

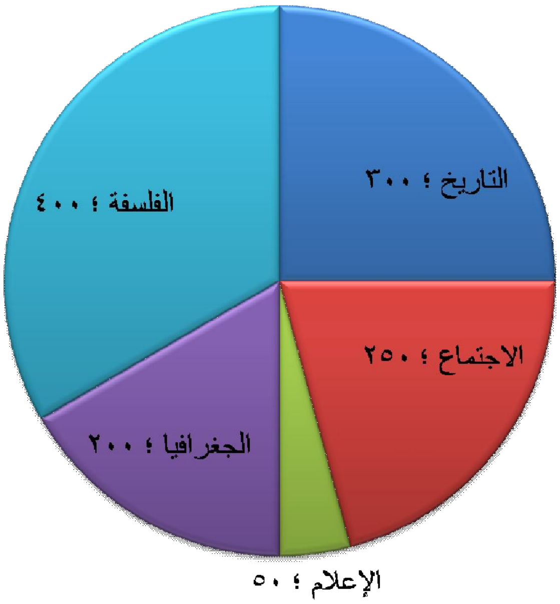
قسم التاريخ = (650÷ 2450)×360= 95.5

قسم الاجتماع = (500÷ 2450)×360 = 73

قسم الأعلام = (400÷2450)×360= 58.7

قسم الجغرافيا = (350÷2450)×360= 51

قسم الفلسفة = (550÷2450)×360= 80.8



المحاضره الثانيه

التوزيع التكراري

- ةالتوزيعات التكرارية هي أحدي طرق تنظيم البيانات الكثيرة بحيث لا تخسر هذه البيانات من أهميتها.
- والطريقة الأساسية لبناء التوزيع التكراري هي تقسيم مدي قيم البيانات إلي فئات وحصر عدد البيانات الواقعة ضمن كل فئة ..

مثال (١)

إذا كان لديك البيانات التالية : (7, 4, 8 , 3, 4 , 2 , 6 , 4 , 3 , 5 , 7 , 5, 4 , 5 , 6) كوني توزيع تكراري يعرض هذه البيانات.

الحل:

البيانات	التكرار
2	1
3	2
4	4
5	3
6	2
7	2
8	1

وعند بناء التوزيع التكراري يجب مراعاة:

١. يجب أن تكون الفئات منفصلة عن بعضها البعض.
٢. أن تكون الفئات متساوية في الطول .
٣. يجب أن تكون الفئات كافئه لاحتواء جميع البيانات.

ولشرح الخطوات المتبعة لبناء التوزيع التكراري نأخذ المثال التالي الخطوات :

مثال (٢) :

البيانات التالية تمثل كمية المبيعات لأربعين بائعاً بأحدي المحلات التجارية ألكبري:

28	27	26	6	24	3	21	20
21	7	7	25	5	23	13	10
15	8	17	13	12	15	18	10
12	19	13	9	5	36	4	13
16	9	30	17	6	14	8	12

ضعي البيانات جدول توزيع تكراري

خطوات أنشاء التوزيع التكراري:

١/حساب ألمدي = أعلي قيمة – أصغر قيمة .

36 - 3 = 33

٢/ طول الفئة (C)= المدي÷ عدد الفئات .

33 /5 = 6.6 ≈ 7 (نقرب لأعلي)

- ✓ إذا كان الناتج كسر نقرب إلي أعلي.
- ✓ (إذا كان ألمدي كبير أو عدد البيانات كبير فأنه يتم تقسيم البيانات إلي فئات يتراوح عددها من 5إلي 15فئة).

٣/ نعين الحد الادني للفئة الأولي 3 ثم نطرح منه نصف وحدة دقة لنعين الحد الادني الفعلي للفئة الأولي

3 - 0.5 = 2.5

٤/ نعين الحد الاعلي الفعلي للفئة الأولي ذللك بإضافة طول الفئة إلي الحد الأدنى الفعلي 9.5 =2.5+7

٥/ نعين الحدود العليا والدنيا الفعلية للفئات الباقية وذلك بإضافة طول الفئة لكل حد علي التوالي.

٦/ نفرغ البيانات علي الفئات.

٧/ نسجل مجموع تكرارات كل فئة إمامها في عمود التكرارات (نرمز لتكرار الفئة ب **f** ومجموع التكرارات **n**).

٨/ مركز الفئة (**X**) = $X_i = \frac{(L+U)}{2}$

$$X_1 = \frac{2.5+9.5}{2} = 12/2 = 6$$

الحل:

التكرار ( )	**مراكز الفئات ( )**	**الفئات الفعلية**	**الفئات**
12	(2.5+9.5) /2=6	2.5 - 9.5	3 – 9
13	(9.5+16.5)/2=13	9.5 - 16.5	10 – 16
8	20	16.5 - 23.5	17- 23
6	27	23.5 - 30.5	24 -30
1	34	30.5 - 37.5	31 – 37
40			**المجموع (n)**

التوزيع التكراري النسبي:

التكرار النسبي لكل فئة هو نسبة تكرار تلك الفئة إلي مجموع التكرارات (n)

حيثُ أن :

N: مجموع التكرارات $p = \frac{f}{n}$

f : التكرار

ونلاحظ أن مجموع التكرار النسبي= 1

أما إذا ضربنا كل تكرار نسبي ب 100% نحصل علي التوزيع التكراري المئوي

مثال(٧) :

أوجدي التوزيع التكرار النسبي والمئوي للمثال (٦) ..

الحل:

حدود الفئات	التكرار النسبي	التكرار المئوي
3 – 9	12/40 = 0.3	0.3*100 =30%
10 – 16	13/40 = 0.325	0.325*100=32.5%
17- 23	8/40 = 0.2	0.2*100=20%
24 -30	6/40 = 0.15	0.15*100=15%
31 – 37	1/40 = 0.025	0.025*100=2.5%

التوزيع التكراري المتجمع:

يتم وضع الحدود الفعلية للفئات مع التكرارات المتجمعة المقابلة لها .ونبدأ دائما بالحد الادني الفعلي للفئة الأولي ونعتبر تكراره المتجمع صفراً .

ويمكن أن نحصل علي التكرار المتجمع النسبي اذا استعملنا التكرارات النسبية بدلاً عن التكرارات.وأيضاً بالنسبة للتكرار المئوي.

مثال(٨)

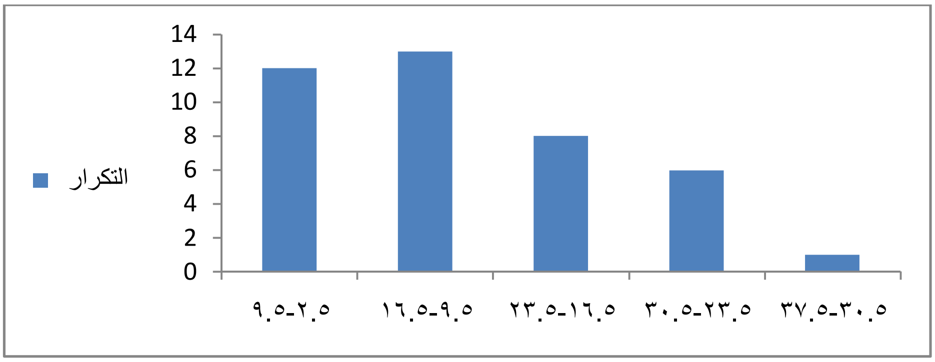
التوزيع التكراري المتجمع للمثال (٦)

الحل:

الحدود الفعلية للفئات	التكرار المتجمع
اقل من 2.5	0
اقل من 9.5	0+12=12
اقل من 16.5	12+13 =25
اقل من 23.5	25+8 =33
اقل من 30.5	33+6=39
اقل من 37.5	39+1=40

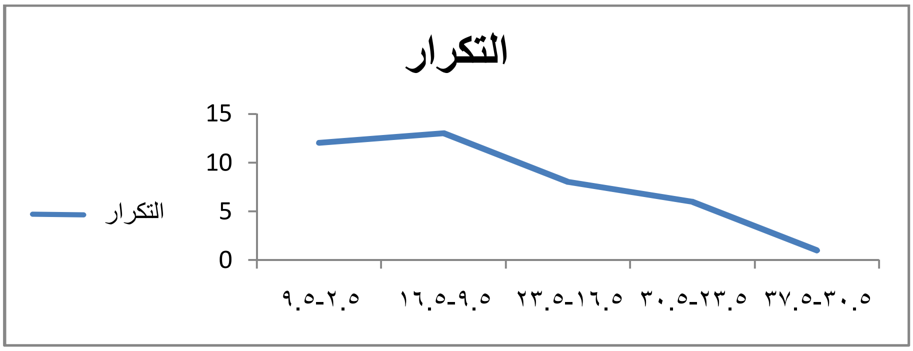
تمثيل التوزيعات التكرارية بيانيا:

(١) المدرج التكراري : هو عبارة عن تمثيل تكرار كل فئة من فئات التوزيع التكراري بمستطيل حدود قاعدته الحدود الفعلية لتلك الفئة ..



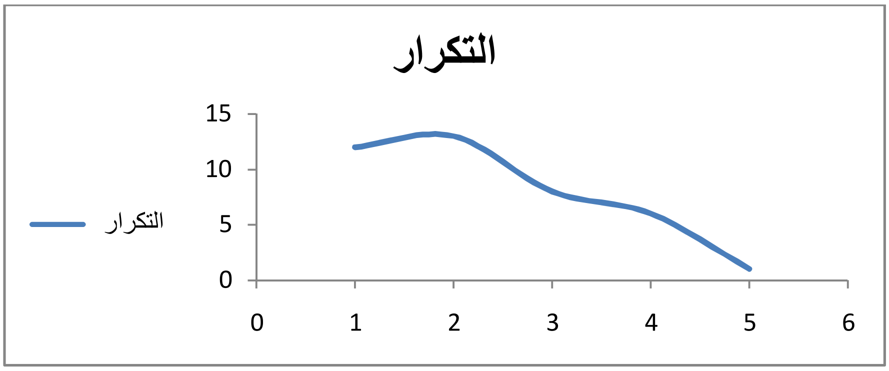
المضلع التكراري:

هو مضلع مغلق نحصل عليه بتنصيف الإضلاع العلوية للمستطيلات في المدرج التكراري ثم نوصل النقاط بعضها ببعض بخطوط مستقيمة، ولكي نغلق الخط المنكسر الذي حصلنا عليه نعتبر أن هناك فئتين متطرفتين واحدة إلي أقصي اليسار والثانية إلي أقصي اليمين تكرار كل منهما صفراً .وذلك بأخذ مركز كل من هاتين الفئتين.



المنحني التكراري :

هو تمهيد للمضلع التكراري نوصل المراكز بمنحني بدلاً من الخطوط المنكسرة



ويمكن استعمال الطرق الثلاثة السابقة لتمثيل التوزيع التكراري المتجمع بيانيا.

أشكال التوزيعات التكرارية:

- عند وصف البيانات توجد ثلاث خواص لابد من معرفتها وهي(الشكل والنزعة المركزية والتغير).
- يتم التعرف علي شكل التوزيع التكراري من التوزيع نفسه أومن مضلعة أو مدرجه التكراري أو العرض بالساق والورقة الخاص به.:

١/ التمييز بين التوزيعات المتماثلة والتوزيعات غير المتماثلة .

٢/ التمييز بين التوزيعات ذات المنوال الواحد والتوزيعات ذات العدة منوالات.

٣/ التمييز بين التوزيعات كبيرة التفرطح ومتوسطة التفرطح وقليلة التفرطح (مدبب).

المحاضرة الثالثة

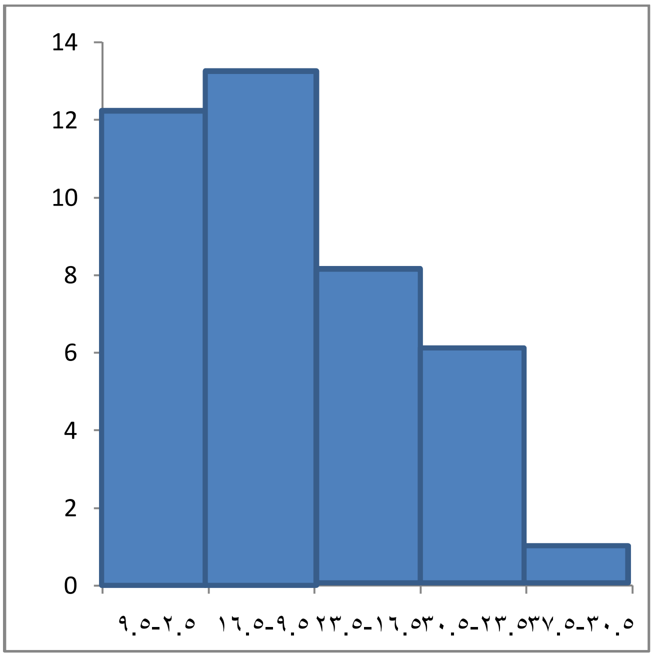
تمثيل التوزيعات التكرارية بيانيا

(١) المدرج التكراري :

- هو عبارة عن تمثيل تكرار كل فئة من فئات التوزيع التكراري بمستطيل حدود قاعدته الحدود الفعلية لتلك الفئة. وطول ارتفاعه يتناسب مع تكرارها.

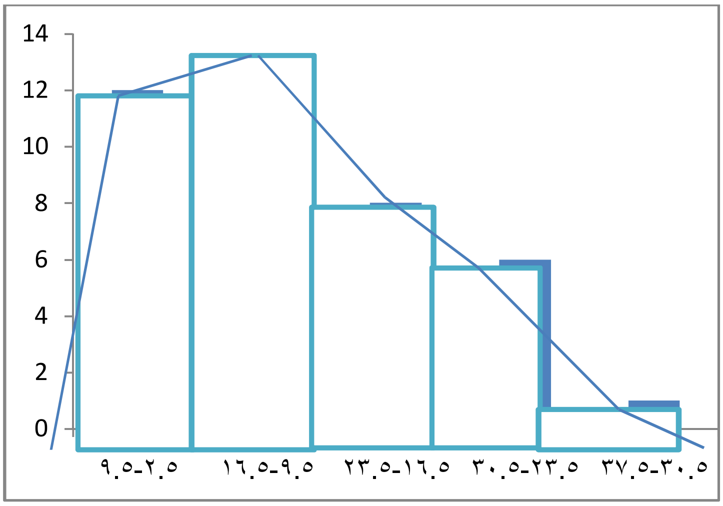
مثال (٩) : من الجدول رقم (٦) ارسم المدرج التكراري

الحل:



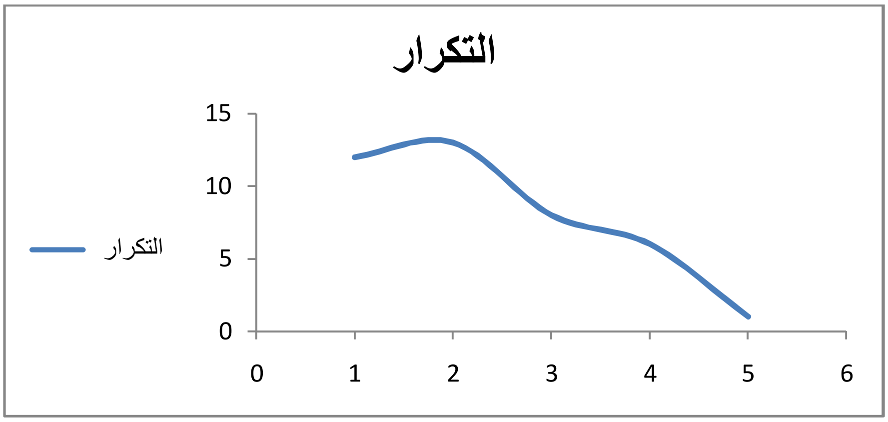
المضلع التكراري:

- هو مضلع مغلق نحصل عليه بتنصيف الإضلاع العلوية للمستطيلات في المدرج التكراري ثم نوصل النقاط بعضها ببعض بخطوط مستقيمة، ولكي نغلق الخط المنكسر الذي حصلنا عليه نعتبر أن هناك فئتين متطرفتين واحدة إلي أقصي اليسار والثانية إلي أقصي اليمين تكرار كل منهما صفراً .وذلك بأخذ مركز كل من هاتين الفئتين.



المنحني التكراري:

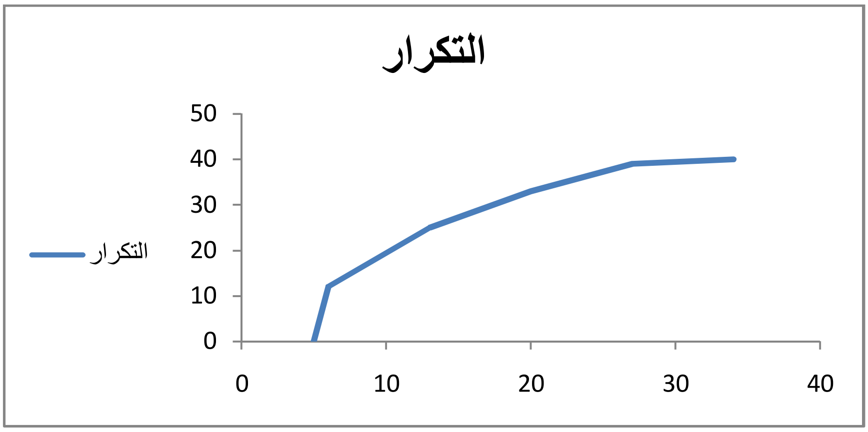
- هو تمهيد للمضلع التكراري نوصل المراكز بمنحني بدلاً من الخطوط المنكسرة



ويمكن استعمال الطرق الثلاثة السابقة لتمثيل التوزيع التكراري المتجمع بيانيا.

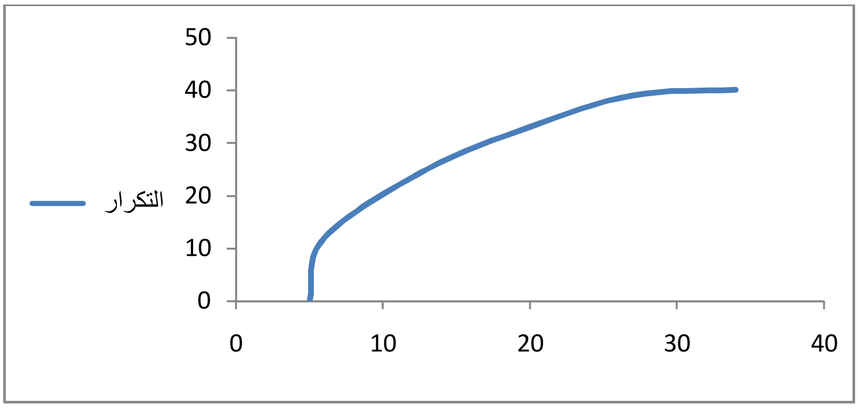
المضلع التكراري المتجمع:

- نحصل علي المضلع التكراري المتجمع برصد التكرار المتجمع لأي فئة مقابل الحد الاعلي الفعلي لها ثم نوصل هذه النقاط فيما بينها بخطوط مستقيمة.



المنحني التكرار المتجمع:

- ونحصل علية بتمهيد المضلع التكراري



- وبنفس الطريقة السابقة التي مثلنا بها التوزيع التكراري المتجمع بيانيا نستطيع تمثيل التوزيع التكراري المتجمع النسبي والتوزيع التكراري المتجمع المئوي وذلك باستعمال التكرارات المتجمعة النسبية والتكرارات المتجمعة المئوية علي المحور العمودي.

أشكال التوزيعات التكرارية:

- عند وصف البيانات توجد ثلاث خواص لابد من معرفتها وهي(الشكل والنزعة المركزية والتغير).

يتم التعرف علي شكل التوزيع التكراري من التوزيع نفسه أومن مضلعة أو مدرجه التكراري أو العرض بالساق والورقة الخاص به.

١/ التمييز بين التوزيعات المتماثلة والتوزيعات غير المتماثلة .
٢/ التمييز بين التوزيعات ذات المنوال الواحد والتوزيعات ذات العدة منوالات.
٣/ التمييز بين التوزيعات كبيرة التفرطح ومتوسطة التفرطح وقليلة التفرطح (مدبب).

المحاضره الرابعه

مقاييس النزعة المركزية

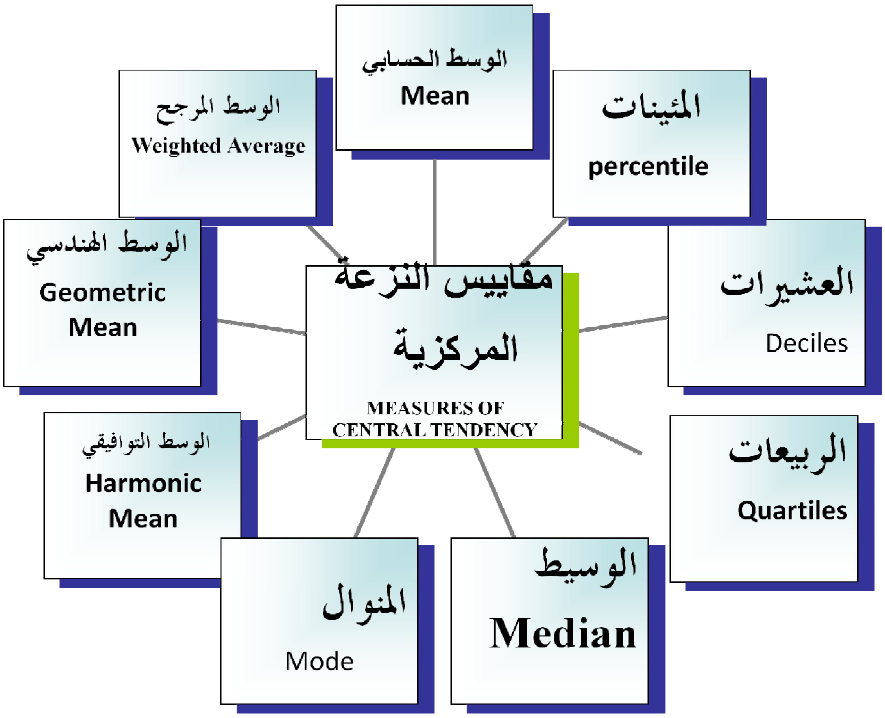
مقدمة:

- إن الهدف الأساسي من استخدام مقاييس النزعة المركزية ومقاييس التشتت هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها ومقدار تشتت البيانات حول هذا المركز (درجة تجانس البيانات) ومن خلال هذين المؤشرين يتمكن الباحث من فهم أبعاد الظاهرة قيد الدراسة.
- كل ظاهرة في الحياة العامة لها ميل للتجمع حول نقطة معينة ، ومن ثم إذا استطعنا تحديد هذه النقطة فإننا سنصل إلى قيمة متوسطة تتجمع حولها القيم.
- يسمى ذلك الميل إلى التجمع حول هذه القيمة بالنزعة المركزية - تسمى المقاييس المستخدمة - مقاييس النزعة المركزية ، وهي القيم التي تتوزع حولها القراءات

شروط المقياس الجيد:

- يحسب بطريقة سهلة لا تؤثر على دقة البيانات.
- يأخذ في الاعتبار جميع المفردات المطلوب حساب المقياس لها.
- يكون له معنى طبيعي مفهوم يستخدم في الحياة العامة .
- يعكس التغير في الظاهرة ، ولا يتغير بتغير طرق حسابه.
- يخضع للعمليات الجبرية خضوعا تاما .
- لا يتأثر بالقيم الشاذة أو المتطرفة.
- لا يتأثر باختلاف العينات ذات الحجم الواحد.

مقاييس النزعة المركزية بأشكالها المتعارف عليها في علم الإحصاء



مقاييس النزعة المركزية (الأساسية):

- الوسط الحسابي
- الوسيط
- الوسط الهندسي
- المنوال
- الوسط التوافقي

(أولا) الوسط الحسابي Arithmetic Mean

- الوسط الحسابي يعتبر من أهم مقاييس النزعة المركزية والأكثر استخداما في الإحصاء والحياة العملية إذ يستخدم عاده في الكثير من المقارنات بين الظواهر المختلفة.
- يعتبر الوسط أكثر المقاييس الإحصائية انتشارا وشيوعا بين الباحثين لسهولته وفائدته التي تضفي عليه أهمية كبرى في حياتنا اليومية فكثيرا ما يتحدث الأفراد عن متوسطات الأسعار في الشهر الأول أو العام الأول ومتوسطات الأعمار واختلافاتها من جيل إلى جيل ، ومن بلد إلى بلد آخر ومتوسطات الدخل الشهري والسنوي ، وغير ذلك من الأمور العملية التي تتصل من قريب بحياتنا اليومية .

الوسط الحسابي أو الوسط (يرمز له بالرمز $\overline{x}$ ويقرأ x bar) للمجموعة n من الأعداد $x_1, x_2, \ldots, x_n$ ؟

$$\overline{X} = \frac{x_1 + x_1 + \ldots + x_n}{n} = \frac{\sum_{j=1}^{n} X_j}{n} = \frac{\sum X_j}{n}$$

- أي أن الوسط الحسابي يساوي مجموع القراءات مقسوما علي عددها ، لاحظ أن المتوسط الحسابي يستخدم كل البيانات المتوفرة في طريقة حسابه.

مثال ١:

إذا كانت أوزان مجموعة من الطلبة بالكيلو جرام هي علي التوالي: 50, 60, 80, 70, 100

فإن الوسط الحسابي لأوزان الطلبة هو:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\overline{X} = \frac{50+60+80+70+100}{5} = \frac{360}{5} = 72$$

مثال ٢:

إذا كان الوسط الحسابي لعلامات عدد من الطلاب هو 56، ومجموع علاماتهم هو 2800. أوجد عدد هؤلاء الطلاب ؟

$$\overline{X} = \frac{\sum_{i=1} X_i}{n}$$

$$56 = \frac{2800}{n}$$

$$n = \frac{2800}{56}$$

$$n = 50$$

ملاحظه:

- الوسط الحسابي يعتبر من أهم مقاييس البيانات الكمية ولا يستخدم مع البيانات النوعية

الوسط الحسابي المرجح:

- إذا كان لدينا مجموعات ذات إعداد مختلفة من البيانات وعلم الوسط الحسابي لكل مجموعة. كيف نحصل علي الوسط الحسابي للمجموعات إذا دمجت معا؟

قاعده: N_1

لدينا مجموعة ذات N_1 من القيم ووسطها الحسابي $\overline{X_1}$.ومجموعة ثانية ذات N_2 من القيم ووسطها الحسابي . $\overline{X}_2$. فأن الوسط الحسابي للمجموعات ذات $N_1 + N_2$ من القيم الناتجة من دمج المجموعتين هو:

$$\overline{X} = \frac{\overline{X}_1 N_1 + \overline{X}_2 N_2}{N_1 + N_2}$$

وهذه القاعدة صحيحة لود مجنا أي عدد محدود من المجموعات مع بعضها البعض .

مثال ٣:

إذا كان عدد الطالبات في المدرج (A) 45 طالبة ومتوسط درجاتهم 63 درجة وعدد الطالبات في المدرج (B) 35 طالبة ومتوسط درجاتهم 74 درجة ، ما الوسط الحسابي للطالبات في المدرجين (A and B).

$$= \frac{(2835 + 2590)}{(80)} = \frac{5425}{80} = 67.8$$

$$= \frac{(2835 + 2590)}{(80)} = \frac{5425}{80} = 67.8$$

المحاضره الخامسه

تابع : مقاييس النزعه المركزيه

الوسط الهندسي:

إذا كان لدينا N من الإعداد الموجبة $x_1 , x_2 , \ldots , x_n$

فان وسطها الهندسي يعرف بالمعادلة: $G = \sqrt[N]{X_1 X_2 \ldots\ldots\ldots X_N}$

مثال ٥:

أوجد الوسط الهندسي للإعداد: 8 , 13 , 32 , 43 , 6

$$G = \sqrt[5]{(8*13*32*43*6)} = 15.37$$

مثال ٦:

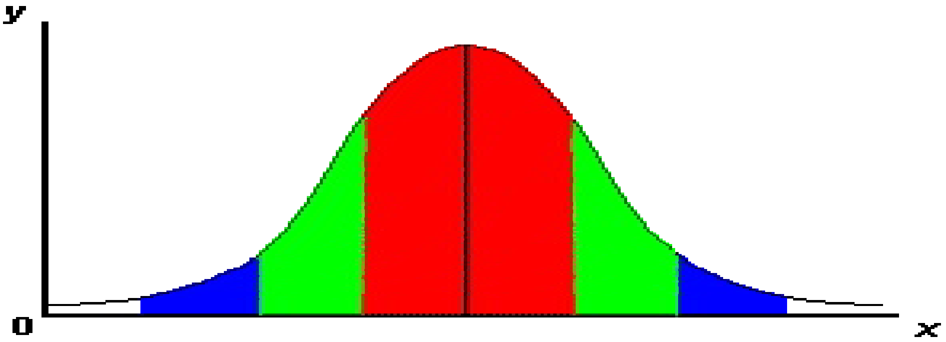
أحسب الوسط الهندسي للقيم التالية : 10 , 6 , 7 , 23 , 5 , 8 , 9 , 14

$$\sqrt[8]{48686400} =$$

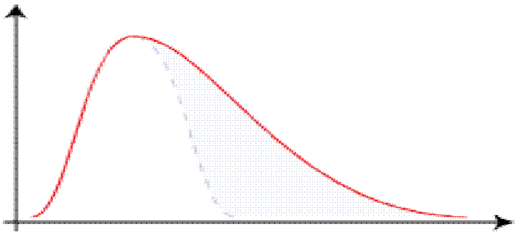
(ثانيا) الوسيط Median:

- هو مقياس أخر من مقاييس النزعة المركزية ويمثل ألمشاهده التي تكون التكرارات التي تسبقها تساوي التكرارات التي تليها. أوهو النقطة التي تقع تماما في منتصف توزيع الدرجات بحيث يسبقها نصف عدد الدرجات ويتلوها النصف الآخر أو هو القيمة التي تقسم البيانات إلى مجموعتين متساويتين وذلك بعد ترتيبها تصاعدياً أو تنازلياً . ومن خلال التعريف للوسيط نجد أنه يتأثر بالدرجات الوسطى أكثر مما يتأثر بالدرجات المتطرفة، وهو يصبح بهذه الصفة على نقيض المتوسط الذي يتأثر بالدرجات المتطرفة أكثر من تأثره بالدرجات الوسطى . ولذا يصلح الوسيط كمقياس للنزعة المركزية أكثر من المتوسط عندما تكون أطراف التوزيع متراكمة متجمعة غير مستوية كأي يلتوي التوزيع التكراري ناحية اليمين أو يلتوي ناحية ..

المنحنى التكراري المعتدل:

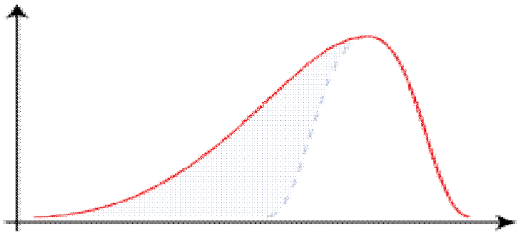


منحنى تكراري ملتوي ناحيه اليمين



Pos tive Skew

منحنى تكراري ملتوي ناحيه اليسار



Negative Skew

رتبه الوسيط $= \frac{n+1}{2}$

عدد القيم الفردي

عدد القيم الزوجي

في هذه الحاله الوسيط له رتبتان هما على التوالي:

$$\frac{n}{2} \ \& \ \frac{n}{2} + 1$$

مثال ٧:

أحسب الوسيط للقيم : 6, 5, 4, 3, 112

عدد البيانات (n) = 5 (فردي)

الحل:

نرتب البيانات 3 , 4 , 5 , 6 , 112

اذاً ترتيب الوسيط = $\frac{n+1}{2}$ $\frac{5+1}{2} = \frac{6}{2} = 3$

الوسيط = 5

مثال ٨:

احسب الوسيط للقيم: 1, 3, 6, 7, -8, -3

عدد البيانات (n) = 6 (زوجي)

الحل:

نرتب البيانات = 8- , 3- , 1 , 3 , 6 , 7

ترتيب الوسيط الأول = $\frac{n}{2}$ = $\frac{6}{2} = 3$

الوسيط الأول = ١

يتبع الحل:

ترتيب الوسيط الثاني = 4 (6/2)+1 = 3+1= (n/2) +1

الوسيط الثاني = ٣

الوسيط الكلي = ٢/ (الوسيط الاول + الوسيط الثاني) 1+3/2 =4/2 =2

(ثالثا) المنوال Mode :

هو عبارة عن القيمة الأكثر تكراراً أو شيوعاً في العينة أو يدل المنوال على أكثر الدرجات شيوعا ،أي هي النقطة التي تدل على أكثر درجات التوزيع تكرارا

المنوال = 2

مثال ٩:

احسب المنوال للقيم 2 , 11 , 2 , 4 , 3 , 2

اذا المنوال = 2

المنوال: أقل مقاييس النزعة المركزية تأثر بالقيم الشاذة

- لا يمكن اعتبار المنوال مقياسا للنزعة المركزية إن لم يكن هناك قيم مكررة .

مثال ١٠:

3 , 4 , 5 , 6 كل مشاهدة تكررت مرة واحده ولا يوجد مشاهدة تكررت أكثر من غيرها ... إذً لا يوجد منوال !

مثال ١١:

إن تساوت تكرارات البيانات

كل مشاهدة مكررة مرتين ولا يوجد قيمة مكررة أكثر من باقي القيم

اذا: 2 , 1 , 4 , 3 , 4 , 1 , 3 2 المنوال = لا يوجد منوال

- يمكن إيجاد أكثر من منوال واحد في البيانات

مثال ١٢:

أحسب المنوال للبيانات التالية : 5 , 1 , 4 , 2 , 1 , 2 , 5

المنوال = 1 , 2 , 5

المحاضره السادسه

مقاييس النزعة المركزية للتوزيع التكراري

الوسط الحسابي من التوزيع التكراري:

تعريف: اذا كان لدينا توزيع تكراري عدد فئاته h وكانت مراكز الفئات (اوالقيم) X1,X2,X3,...,Xn وكانت التكرارات المقابلة لها f1,f2,f3,...,fn

فان الوسط الحسابي يكون:

$$\overline{X} = \frac{X_1 f_1 + X_2 f_2 + \ldots + X_h f_h}{f_1 + f_2 + \ldots + f_h} = \frac{\sum_{i=1}^{h} X_i f_i}{\sum_{i=1}^{h} f_i} = \frac{\sum_{i=1}^{h} X_i f_i}{n}$$

حيث n تمثل مجموع التكرارات.

خطوات حساب الوسط الحسابي:

$$n = \sum_{i=1}^{h} f_i$$

١/ إيجاد مجموع التكرارات:

٢/ حساب مراكز الفئات: x_i

٣/ضرب مركز الفئة في التكرار المناظر له $x_i * f_i$

وحساب مجموع $x*f$

٤/ حساب الوسط الحسابي بتطبيق القاعدة

مثال٣:

الجدول التالي يعرض توزيع 40 تلميذ حسب أوزانهم. والمطلوب إيجاد الوسط الحسابي؟

42-44	40-42	38-40	36-38	34-36	32-34	فئات الوزن
1	**5**	**10**	**13**	**7**	**4**	**عدد التلاميذ**

التكرار* مراكز الفئات	مراكز الفئات	التكرار	الفئات
33*7 =231	33	7	32-34
35*4 =140	35	4	34-36
37*13 =481	37	13	36-38
39*10 = 390	39	10	38-40
41*5 = 205	41	5	40-42
43*1 = 43	43	1	42-44
= 1490		n = 40	

تطبيق القاعدة:

$$\overline{X} = \frac{\sum_{i=1}^{h} x_i * f_i}{\sum f_i} = \frac{1490}{40} = 37.3$$

الوسط الهندسي للتوزيع التكراري:

إذا كان لدينا توزيع تكراري عدد فئاته h وكانت مراكز الفئات أو(القيم) $x_1, x_2, \dots, x_n$

وكانت التكرارات المقابلة لها $f_1, f_2, \dots\dots\dots, f_h$ فان الوسط الهندسي:

$$G = \sqrt[N]{X_1^{f_1} X_2^{f_2} \dots\dots X_h^{f_h}}$$

$$N = \sum_{i=1}^{h} f_i$$

مثال ٤:

أوجدي الوسط الهندسي للمثال(3)

مراكز الفئات	التكرار	الفئات
33	7	32-34
35	4	34-36
37	13	36-38
39	10	38-40
41	5	40-42
43	1	42-44
	n = 40	

تطبيق القاعدة:

$$G = \sqrt[N]{x_1^{f_1} * x_2^{f_2} * x_3^{f_3} * \dots\dots\dots * x_N^{f_h}}$$

$$= \sqrt[40]{33^{7} * 35^{4} * 37^{13} * 39^{10}} * 41^{5} * 43^{1} =$$

مثال ٥:

الجدول التالي يوضح العلاقة بين فئات الدخل بأحد المصانع وعدد العمال .

أحسب:

١. الوسط الحسابي .
٢. الوسط الهندسي .

عدد العمال	فئات الدخل
20	20 - 30
40	30 - 40
100	40 - 50
30	50 - 60
10	60 - 70

أولاً : الوسط الحسابي:

	مركز الفئات	عدد العمال ( )	فئات الدخل
500	25	20	20 - 30
1400	35	40	30 - 40
4500	45	100	40 - 50
1650	55	30	50 - 60
650	65	10	60 - 70
8700		200	المجموع

$$\overline{X} = \frac{\sum_{i=1}^{h} x_i * f_i}{\sum f_i}$$

$$= \frac{8700}{200} = 43.5$$

ثانياً : الوسط الهندسي:

مركز الفئات	عدد العمال ( )	فئات الدخل
25	20	20 - 30
35	40	30 - 40
45	100	40 - 50
55	30	50 - 60
65	10	60 - 70

$$G = \sqrt[N]{x_1^{f_1} * x_2^{f_2} * ....x_N^{f_N}} =$$

$$\sqrt[200]{25^{20} * 35^{40}} * 45^{100} * 55^{30} * 65^{10} =$$

المحاضره السابعه

الوسيط

الوسيط للتوزيع التكراري:

- حساب الوسيط من الجدول التكراري (البيانات المبوبة)
- من جدول التكرار المتجمع الصاعد أو النازل أو الرسم البياني لكلاهم أو أحدهم وسنورد هنا تفصيلا كاملا لذلك .

تعريف:

الفئة الوسيطية هي أول فئة يزيد تكرارها المتجمع عن $\frac{n}{2}$ أو يساويه حيث n هي مجموع التكرارات .
ولإيجاد الوسيط للتوزيع التكراري ذي الفئات نفرض أن:

h: عدد الفئات .

n : مجموع التكرارات أي : $n=\sum_{i=1}^{h} f_i$

C : طول الفئة الوسيطية.

a: الحد الادني الفعلي للفئة الوسيطية

n1: التكرار المتجمع للفئة التي تسبق الفئة الوسيطية مباشرة.

b : الحد الادني الفعلي للفئة الوسيطية .

fm: تكرار الفئة الوسيطية .

القاعدة: $$M=\frac{(\frac{n}{2}-n_1)}{f_m}\times C$$

مثال ٦:

الجدول التالي يمثل الأجر اليومي للعامل بالريال في مائتين سوبر ماركت .المطلوب متوسط الأجر اليومي للعامل.

التكرار	الفئات
30	5 - 15
20	15 - 25
60	25 – 35
50	**35 - 45**
40	45- 55
200	المجموع

الفئة الوسيطية

f_m

التوزيع التكراري المتجمع:

التكرار المتجمع	حدود الفئات الفعلية
0	اقل من 5
30	اقل من 15
50	اقل من 25
110	اقل من 35
160	اقل من 45
200	اقل من 55

$n_1 \leftarrow$ 50

a → اقل من 35

الحل:

الفئة الوسطية هي: (35 – 45)

$$\frac{n}{2}=\frac{200}{2}=100$$

a = 35 fm = 50

n1= 50 C = 10

$$M=a+\left(\frac{(\frac{n}{2}-n_1)}{f_m}\right)\times C$$

$$=35+\left(\frac{(100-50)}{50}\right)\times 10=35+(1\times 10)=35+10=45$$

المنوال للتوزيع التكراري:

- الفئة المنوالية (أو الفئات)التي يقابلها اكبر تكرار تسمي الفئة المنوالية (الفئات المنوالية).
- مركز الفئة المنوالية يسمي المنوال التقريبي وإذا كان هناك عدة فئات منوالية فيكون لدينا عدة منوالات تقريبية

مثال ٧:

أوجد المنوال للمثال (6)

التكرار	الفئات
30	5 - 15
20	15 - 25
60	25 – 35
50	35 - 45
40	45- 55
200	المجموع

الفئة المنوالية → 25 – 35

الحل:

الفئة المنوالية = (35 – 25)

مركز الفئة المنوالية = $\frac{\text{الحدالأعلي}+\text{الحدالأدني}}{2}$

المنوال التقريبي = $\frac{25+35}{2}=\frac{60}{2}=30$

مقارنة بين صفات الوسط الحسابي والوسيط والمنوال:

١. الوسط الحسابي أكثر مقايي النزعة المركزية استعمالا لسهولة حسابه وسهولة تعريفه ويخضع للعمليات الحسابية.

٢. نستطيع معرفة مجموع القيم إذا علمنا الوسط الحسابي ومجموع التكرارات .(من التعريف هو مجموع البيانات مقسوماً علي عددها).

٣. بما أن الوسط الحسابي يعتمد علي جميع القيم فإن قيمته تتغير إذا خزفنا أو غيرنا في مفردات البيانات .

٤. مجموع انحرافات قيم البيانات عن وسطها الحسابي يساوي صفر:

$$\sum_{i=1}^{n}(\chi_i-\bar{x})=0;\sum_{i=1}^{h}(\chi_i-\bar{x})f_{\,i}$$

أذن الوسط الحسابي هو نقطة اتزان للمدرج التكراري وإذا أضفنا أي عدد من القيم المساوية للوسط الحسابي إلي البيانات فان هذه الإضافة لا تؤثر عليه أما إذا أضفنا مفردات تختلف قيمتها عن قيمة الوسط الحسابي فإن قيمته تتأثر.

٥. ومن أهم نواقصه تأثره الشديد بالقيم المتطرفة.

٦. أما الوسيط فهو سهل الحساب وسهل التعريف ولا يتأثر بالقيم المتطرفة ولا يعتمد علي جميع القيم دائماً ،فتغيير قيمه من القيم ربما يؤثر في قيمته أو لا يؤاثر.

٧. الوسط الحسابي أكثر ثبوتا من الوسيط.

٨. المنوال أقل مقاييس النزعة المركزية استعمالا وفي البيانات قليلة العدد عديم الفائدة ، أما في البيانات الكبيرة العدد فله معني معقول . وهو يتأثر بطريقة اختيار الفئات. وهو لا يتأثر بالقيم الشاذة.

٩. يفضل استعمال الوسط الحسابي إذا كان التوزيع متماثلاً .

١٠ يفضل استعمال الوسيط إذا كان التوزيع ملتوياً .

١١. يفضل استعمال المنوال إذا كان التوزيع متعدد المنوالات.

المحاضره الثامنه

مقاييس التشتت

مقاييس التشتت:

هي مقاييس عددية تستخدم لقياس اختلاف أو تشتت البيانات . والاختلاف أو التشتت لمجموعة من البيانات هو مقدار تفرق أو تباعد أو انتشار البيانات فيما بينها . فتشتت البيانات يكون صغيرا إذا كانت البيانات متقاربة فيما بينها والعكس بالعكس وأما البيانات المتساوية فلا اختلاف ولا تشتت فيها .ومقاييس التشتت تستخدم لوصف مجموعة البيانات وكذلك لمقارنة مجموعات البيانات المختلفة إذ أن مقاييس النزعة المركزية لا تكفي وحدها لوصف مجموعة البيانات أو مقارنة مجموعات البيانات المختلفة .ومن أشهر مقاييس التشتت نذكر :

- المدى : Range
- التباين : Variance
- الانحراف المعياري : Standard Deviation .
- الانحراف المتوسط : Deviation The Mean
- معامل الاختلاف (أو التغير): Coefficient of Variation

مقاييس التشتت للبيانات الأولية:

١. **المدى Range :** يعتبر المدى من أسهل مقاييس التشتت تعريًفا وحسابا ويعطينا فكرة سريعة عن مدى تفرق البيانات .ويعرف المدى لمجموعة من البيانات بالصيغة التالية :

$$\text{Range} = X_{\max} - X_{nin}$$

حيث أن :

$X_{\max}$ = أكبر قيمة (للبيانات المفردة) = الحد الاعلي للفئة العليا (للبيانات المبوبة)

$X_{\min}$ = أصغر قيمة (للبيانات المفردة = الحد الادني للفئة الدنيا (للبيانات المبوبة)

ألمدي = Upper – Lower

مثال ١:

المشاهدات التالية وهي عبارة عن أوزان(بالكيلوجرام) لمجموعة مكونة من سبعة أشخاص . أوجد المدى

.25 , 30 , 40 , 45 , 35 , 55 , 50

الحل:

المدى = اكبر قيمة – أصغر قيمة

55 – 25 = 30

٢. **التباين : Variance :** التباين للبيانات المفردة (غير المبوبة) إذا كانت x1, x2, ..., xn عينه حجمها n وكان متوسطها هو $\overline{X}$

فإن تباين العينة يعرف كما يلي :

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{(n-1)}$$

مثال٢:

أوجد التباين لمجموعة الأوزان (بالكيلوجرام) التالية: 8 , 5 , 3 ,2 ,7.

الحل:

$$S^2 = \frac{\sum (X_i - \overline{X})^2}{(n-1)}$$

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{7+2+3+5+8}{5} = \frac{15}{5} = 5$$

$$S^2 = \frac{(7-5)^2 + (2-5)^2 + (3-5)^2 + (5-5)^2 + (8-5)^2}{5-1}$$

$$= \frac{4+9+4+0+9}{4} = \frac{26}{4} = 6.5$$

٣. **الانحراف المعياري Standard Deviation** : هو الجذر الربيعي الموجب للتباين أي أن :

$$S = \sqrt{S^2} = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n-1}}$$

مثال٣:

أحسب الانحراف المعياري من المثال (2) ؟

الحل:

$$S^2 = 6.5$$

$$\therefore S = \sqrt{S^2} = \sqrt{6.5} = 2.5$$

٤. **الانحراف المتوسط: Deviation The Mean** : القيم المطلقة للانحراف هي مجموع الانحرافات المطلقة عن المتوسط بينما يعرف الانحراف المتوسط علي أنه انحراف قيم المفردات عن متوسطها الحسابي بغض النظر عن إشارات الانحرافات ويرمز له بالرمز M.D . ولما كان مجموع الانحرافات عن المتوسط يساوي صفرا إذا راعينا الإشارة التي تسبق الانحراف ولكن إذا أهملنا الإشارة نحصل علي مقياس آخر للتشتت هو مجموع القيم المطلقة للانحراف.

$$\sum_{i=1}^{n} \left| X_i - \overline{X} \right|$$

وبالقسمة علي n نحصل علي الانحراف المتوسط ويمكن حسابه من المعادلة :

$$M.D=\frac{\sum_{i=1}^{n}\left|X_i-\overline{X}\right|}{n}$$

مثال ٤:

أوجدي الانحراف المتوسط للمفردات 5 ,6 , 2 , 7, 5

$$M.D=\frac{\sum_{i=1}^{n}\left|X_i-\overline{X}\right|}{n}$$

$$\overline{X}=\frac{\sum_{i=1}^{n}X_i}{n}$$

$$\overline{X}=\frac{\sum_{i=1}^{n}X_i}{n}=\frac{5+7+2+6+5}{5}=\frac{25}{5}=5$$

$$M.D=\frac{\sum_{i=1}^{n}|5-5|+|7-5|+|2-5|+|6-5|+|5-5|}{5}$$

$$=\frac{0+2+3+1+0}{5}=\frac{6}{5}=1.2$$