مقاييس النزعة المركزية

عند التمعن في الظواهر التي حولنا نلاحظ على العموم أن القيم التي تأخذها الوحدات الإحصائية لهذه الظواهر قريبة من بعضها البعض، من خلال تجمعها حول قيمة مركزية معينة، وهناك عدد قليل من هذه القيم يبتعد عن هذا التجمع سواء من ناحية الكبر أو من ناحية الصغر، وكأن هذه القيمة المركزية تعكس قطب جاذبية تسمى هذه الظاهرة بظاهرة النزعة المركزية.

هناك عدة مقاييس للتعبير عن هذه الظاهرة من أهمها: المتوسط الحسابي، الوسيط والمنوال.

أولا- المتوسط الحسابي $\bar{\ }$ **:**

يعتبر المتوسط الحسابي من أهم مقاييس النزعة المركزية وأكثرها استخداما في النواحي التطبيقية، ويعرف عموما على أنه مجموع القيم مقسوما على عددها.

1- الطريقة المباشرة:

1-1- المتوسط الحسابي البسيط:

إذا كانت لدينا بيانات غير مبوبة – القيم غير مكررة- فإن:

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

حيث: X_i تمثل قيم المتغير الإحصائي المتقطع، و n تمثل عدد القيم.

مثال1:

ليكن لدينا علامات مجموعة من الطلبة في مقياس الإحصاء1:

12، 14، 16، 08، 07، 05، 11، 15، 06، 13.

المطلوب: حساب متوسط علامات الطلبة؟.

الحـــــــــل:

$$\bar{X} = \frac{12+14+16+08+07+05+11+15+06+13}{10} = \frac{107}{10} = 10{,}7$$

متوسط علامات الطلبة في مقياس الإحصاء1 هو 10,7.

2-1- المتوسط الحسابي المرجح:

إذا كانت لدينا بيانات مبوبة – هناك قيم مكررة- فإن:

$$\bar{X} = \frac{n_1X_1 + n_2X_2 + \cdots + n_kX_k}{n} = \frac{\sum_{i=1}^{k} n_iX_i}{n}$$

حيث: X_i تمثل قيم المتغير الإحصائي المتقطع، و n_i تمثل التكرارات المطلقة الموافقة لها و n تمثل مجموع التكرارات.

إذا إستخدمنا التكرارات النسبية فإن: $\bar{X} = f_1X_1 + f_2X_2 + \cdots + f_kX_K = \sum_{i=1}^{k} f_iX_i$

مثال2: البيانات التالية تمثل عدد الغرف في المسكن الواحد لعينة من 50 مسكن ببلدية سطيف.

عدد الغرف	عدد المساكن n_i	n_iX_i
1	1	1
2	8	16
3	13	39

52	13	4
30	6	5
24	4	6
21	3	7
16	2	8
199	50	المجموع $\sum n_i$

المطلوب:

حساب متوسط عدد الغرف في المسكن الواحد؟

الحـــــــل: $\overline{X} = \frac{\sum_{i=1}^{k} n_i X_i}{n} = \frac{\sum_{i=1}^{8} n_i X_i}{50} = \frac{199}{50} = 3{,}98$

متوسط عدد الغرف في المسكن الواحد هو 3,98.

2- الطريقة غير المباشرة:

عندما تكون لدينا بيانات كبيرة القيم فإننا نستخدم طريقة غير مباشرة في حساب المتوسط الحسابي (طريقة الانحراف عن متوسط فرضي)، الغرض منها هو تصغير قيم البيانات من أجل تسهيل عملية حساب المتوسط الحسابي.

2-1- المتوسط الحسابي البسيط:

إذا كانت لدينا بيانات غير مبوبة – القيم غير مكررة- فإننا نتبع الخطوات التالية:

أ- نفرض قيمة ثابتة $\boldsymbol{a}$ تكون قريبة من القيم الأصلية وتتوسطها.

ب- حساب قيم جديدة $\boldsymbol{i}$ حيث: $\boldsymbol{d_i = X_i - a}$.

ج- حساب المتوسط الحسابي للقيم الجديدة: $\boldsymbol{\overline{d} = \frac{\sum d_i}{n}}$,

د- حساب المتوسط الحسابي الأصلي حيث لدينا: $\boldsymbol{\overline{d} = \frac{\sum d_i}{n} = \frac{\sum (X_i - a)}{n} = \frac{\sum X_i - \sum a}{n} = \frac{\sum X_i}{n} -}$

$\boldsymbol{\frac{na}{a}}$

$$\boldsymbol{\overline{d} = \overline{X} - a \Rightarrow \overline{X} = \overline{d} + a}$$

مثال3: لتكن لدينا القيم التالية: 12، 10، 15، 8، 5 أحسب الوسط الحسابي عن طريق انحرافات القيم عن وسط فرضي.

الحل: نفرض أن $a = 9$، وعليه نحسب الانحرافات وفق الصيغة التالية: $\boldsymbol{d_i = X_i - a}$، فيتنتج لدينا:

$d_i = x_i - 9$	X_i
-4	5
-1	8
1	10

12	3
15	6
المجموع	**5**

$$\bar{d} = \frac{\sum d_i}{n} = \frac{5}{5} = 1$$

$$\bar{X} = \bar{d} + a = 1 + 9 = 10$$

كما أننا نحصل على نفس النتيجة بتطبيق العلاقة:

$$\bar{X} = \frac{\sum_{i=}^{5} X_i}{n} = \frac{50}{5} = 10$$

2-2- المتوسط الحسابي المرجح:

إذا كانت لدينا بيانات مبوبة – القيم مكررة- فإننا نتبع نفس الخطوات السابقة ولكن بوضع:

$$\bar{d} = \frac{\sum_{i=1}^{k} d_i n_i}{n}$$

كما أن: $\bar{X} = \bar{d} + a$

مثال4: أحسب المتوسط الحسابي بطريقة الانحراف عن متوسط فرضي نفترض أنه $a = 800$؟

X_i	n_i	$d_i = x_i - 800$	$d_i\, n_i$
500	10	-300	-3000
700	16	-100	-1600
800	02	0	0
850	18	50	900
900	16	100	1600
1000	10	200	2000
المجموع	72	/	-100

$$\bar{d} = \frac{\sum_{i=1}^{k} n_i d_i}{n} = \frac{\sum_{i=1}^{6} n_i d_i}{72} = \frac{-100}{72} = -1,39$$

حساب المتوسط الحسابي الأصلي: $\bar{X} = \bar{d} + a = -1,39 + 800 = 798,61$

إذن المتوسط الحسابي لهذه البيانات: 798,61.

ملاحظة: إذا كان المتغير الإحصائي مستمر فإننا نعوض $_i$ بمراكز الفئات C_i في كل المعادلات السابقة أي:

$$\bar{X} = \frac{n_1 c_1 + n_2 c_2 + \cdots + n_k c_k}{n} = \frac{\sum_{i=1}^{k} n_i c_i}{n}$$

مثال5:

البيانات التالية تمثل أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـ LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف:

أوزان الطلبة X_i	عدد الطلبة n_i	C_i	$X_i\, n_i$
]55 – 50]	2	52,5	105

287,5	57,5	5	$]60-55]$
750	62,5	12	$]65-60]$
1080	67,5	16	$]70-65]$
1015	72,5	14	$]75-70]$
620	77,5	8	$]80-75]$
247,5	82,5	3	$]85-80]$
4105	/	60	المجموع $\sum n_i$

المطلوب: حساب متوسط أوزان الطلبة؟

الحل: $\bar{X} = \frac{\sum_{i=1}^{7} n_i c_i}{n} = \frac{4105}{60} = 68{,}42$

متوسط أوزان الطلبة هو 68,42 كلغ.

3- خصائص المتوسط الحسابي:

أ- يتأثر المتوسط الحسابي بالقيم المتطرفة.

مثلا: تتكون أسرة من 5 أفراد، تبلغ أعمارهم على الترتيب: 55، 49، 13، 11، 10.

المتوسط الحسابي هو: $\bar{X} = \frac{55+49+13+11+10}{5} = \frac{138}{5} = 27{,}6$

نلاحظ أن النتيجة المحصل عليها لا تمثل أي قيمة من القيم المدروسة، وبالتالي نستنتج أنه لا يمكن استعمال المتوسط الحسابي في مثل هذه الحالات، لأنه من المفروض أن تكون قيم المتغير الإحصائي متمركزة حول النتيجة المحصل عليها.

ب- يستعمل المتوسط الحسابي في حالة المتغيرات الكمية أي القابلة للقياس.

ج- لا يمكن حساب المتوسط الحسابي في حالة البيانات المفتوحة (أقل من، أكثر من).

د- مجموع انحرافات القيم عن متوسطها يساوي دائما الصفر. $\sum(X_i - \bar{X}) = 0$

$$\sum(X_i - \bar{X}) = \sum X_i - \sum \bar{X}$$

$$\sum(X_i - \bar{X}) = \sum X_i - n\bar{X} = n\bar{X} - n\bar{X} = 0$$

حيث أنها تدل على أن المتوسط الحسابي يقع في مركز البيانات.

هـــ- مجموع مربع انحرافات القيم عن متوسطها الحسابي أقل من مجموع مربعات انحراف نفس القيم عن أي قيمة أخرى، أي: $\sum(X_i - \bar{X})^2 < \sum(X_i - X_\alpha)^2$

حيث أنها تدل على أن المتوسط الحسابي أقرب من جميع القيم.

مثلا إذا كانت لدينا القيم التالية: 24، 8، 6، 16، 14، 4، أثبت صحة الخاصيتين د ، هـــ؟

نفرض أن النقطة المختارة هي: $X_\alpha = 10$ ولدينا: $\bar{X} = \frac{\sum_{i=1}^{6} X_i}{n} = \frac{72}{6} = 12$

$(X_i - X_\alpha)^2$	$(X_i - X_\alpha)$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})$	القيم x_i
196	14	144	12	24

8	-4	16	-2	4
6	-6	36	-4	16
16	4	16	6	36
14	2	4	4	16
4	-8	64	-6	36
$\sum X_i = 72$	$\sum(X_i-\bar{X})=0$	$\sum(X_i-\bar{X})^2=280$	$\sum(X_i-X_\alpha)=12$	$\sum(X_i-X_\alpha)^2=304$

بالنسبة للخاصية الأولى نلاحظ من خلال الجدول أن: $\sum(\ _i-\bar{X})=0$ وهو المطلوب.
أما بالنسبة للخاصية الثانية فمن خلال مقارنة النتائج المتحصل عليها في الجدول أعلاه نجد: $280 < 304$
ينتج عن ذلك: $\sum(X_i-\bar{X})^2 < \sum(X_i-X_\alpha)^2$.

ثانيا- المنوال:

يعرف المنوال على أنه القيمة الأكثر شيوعا وانتشارا من غيرهم من قيم المجموعة المعطاة، ويرمز لها بـــ: M_o.

1- حساب المنوال في حالة سلسلة إحصائية:

قيمة المنوال للبيانات: 12، 16، 10، 12، 17، 9 هي: $M_o = 12$ لأنها الأكثر تكرارا من غيرها.
قيمة المنوال للبيانات: 10، 15، 18، 13، 10، 15، 16 هي: $M_o = 10$ و $_o = 15$.
البيانات التالية: 14، 16، 9، 17، 13 ليس لها منوالا.

2- حساب المنوال في حالة توزيع تكراري لمتغير إحصائي منفصل:

يستنتج مباشرة من جدول التوزيع التكراري، مع الإشارة إلى أنه يمكننا أن نجد أكثر من منوال، كما يمكننا ألا نجد ولا منوال.

مثال6: البيانات التالية تمثل عدد الغرف في المسكن الواحد لعينة من 35 مسكن ببلدية سطيف.

عدد الغرف (قيم المتغير) i	**عدد المساكن (التكرار)** n_i
1	3
2	8
3	13
4	5
5	6
المجموع	35

المنوال في هذا التوزيع هو: $M_o = 3$

الشرح:

عدد المساكن الأكثر شيوعا وانتشارا في هذا التوزيع هي التي تحتوي على 3 غرف.

3- حساب المنوال في حالة توزيع تكراري لمتغير إحصائي مستمر:

إذا كان لدينا جدول توزيع تكراري على شكل فئات فإننا نتبع الخطوات التالية لحساب المنوال:

- تحديد الفئة المنوالية: وهي الفئة التي تقابل أكبر تكرار عندما تكون أطوال الفئات متساوية، أو الفئة التي تقابل أكبر تكرار معدل عندما تكون أطوال الفئات غير متساوية.
- حساب المنوال بطريقة المد الداخلي:

$$M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times A_{M_o}$$

حيث: Lim_{M_o}: الحد الأدنى للفئة المنوالية، Δ_1: الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.

Δ_2: الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها، A_{M_o}: طول الفئة المنوالية.

مثال7: أحسب قيمة المنوال لبيانات المثال5؟

أوزان الطلبة X_i	عدد الطلبة n_i
$]55 - 50]$	2
$]60 - 55]$	5
$]65 - 60]$	12
$]70 - 65]$	16
$]75 - 70]$	14
$]80 - 75]$	8
$]85 - 80]$	3
المجموع $\sum n_i$	60

بما أن أطوال الفئات متساوية فإن الفئة المنوالية هي: $]70 - 65]$

وبالتالي فإن: $\Delta_1 = 16 - 12 = 4$ ، $\Delta_1 = 16 - 14 = 2$

ومنه :

$$M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times A_{M_o}$$

$$M_o = 65 + \left[\frac{4}{4+2}\right] \times 5 = 68{,}33$$

وزن الطلبة الأكثر شيوعا وانتشارا هو 68,33 كلغ.

4- تحديد المنوال بيانيا:

يحدد المنوال بيانيا بواسطة المدرج التكراري، وهذا بإتباع الخطوات التالية:

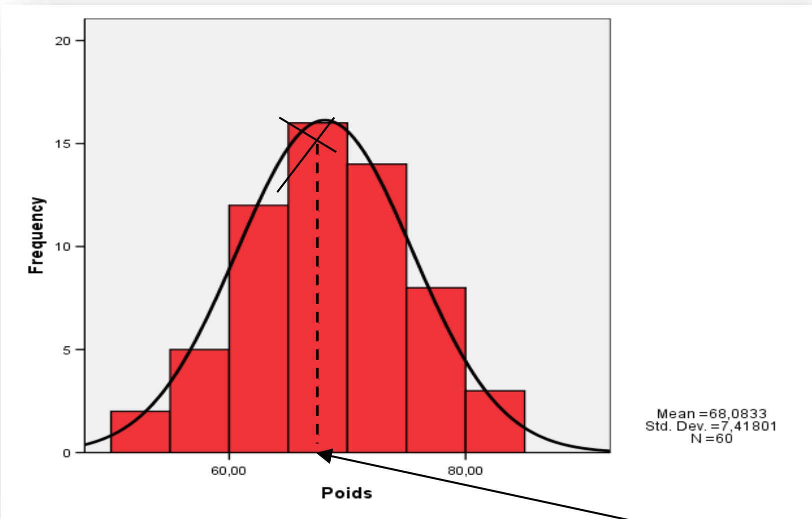
أ- نرسم المدرج التكراري للتوزيع.

ب- نصل بخط مستقيم رأس الحد الأعلى للفئة المنوالية برأس الحد الأعلى للفئة السابقة لها.

ج- نصل بخط مستقيم رأس الحد الأدنى للفئة المنوالية برأس الحد الأدنى للفئة اللاحقة لها.

د- من تقاطع الخطين السابقين نسقط عمودا على المحور الأفقي ونقطة تقاطعه مع المحور الأفقي تمثل تقديرا لقيمة المنوال بيانيا.

مثال8: حدد قيمة المنوال بيانيا للمثال7؟



Mo= 68,33

ثالثا- الوسيط:

الوسيط هو أحد مقاييس النزعة المركزية الذي يأخذ بعين الاعتبار رتبة القيم، ويعرف الوسيط على أنه القيمة التي تقسم البيانات إلى جزئين متساويين بحيث تكون قيم المتغير الإحصائي مرتبة ترتيبا تصاعديا أو تنازليا، ونرمز له بالرمز $\boldsymbol{M_e}$

1- حساب الوسيط في حالة سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

لحساب الوسيط في هذه الحالة نتبع الخطوات التالية:

أ- ترتيب البيانات ترتيبا تصاعديا.

ب- إذا كان عدد البيانات n عددا فرديا فإن الوسيط هو القيمة التي رتبتها $\frac{n+1}{2}$ أي: $M_e = X_{\left(\frac{n+1}{2}\right)}$

ج- إذا كان عدد البيانات n عددا زوجيا فإن الوسيط هو متوسط القيمة التي رتبتها $\frac{n}{2}$ والقيمة التي رتبتها $\frac{n}{2}+1$ أي:

$$M_e = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}$$

مثال9: أحسب الوسيط للسلسلتين الإحصائيتين التاليتين:

- السلسلة الأولى: (1، 2، 5، 4، 3، 6، 7، 7، 5، 3، 1، 9)
- السلسلة الثانية: (1، 0، 2,5، 3، 5، 5,5، 1، 6، 7,5)

الحل:

- السلسلة الأولى: (1، 1، 2، 3، 3، 4، 5، 5، 6، 7، 7، 9)

عدد البيانات زوجي أي: 12، ومنه:

$$M_e = \frac{\left(\frac{n}{2}\right) + X_{\left(\frac{n}{2}+1\right)}}{2} = \frac{X_6+X_7}{2} = \frac{4+5}{2} = 4.5$$

50% من البيانات أقل من 4,5، و50% من البيانات أكبر من 4,5.

- السلسلة الثانية: (0، 1، 1، 2,5، 3، 5، 5,5، 6، 7,5)

عدد البيانات فردي أي: 9، ومنه: $M_e = X_{\left(\frac{n+1}{2}\right)} = X_5 = 3$

50% من البيانات على الأكثر أقل من 3، و50% من البيانات على الأكثر أكبر من 3.

2- حساب الوسيط في حالة توزيع تكراري لمتغير إحصائي مستمر:

إذا كان لدينا جدول توزيع تكراري على شكل فئات فإننا نتبع الخطوات التالية لحساب الوسيط:

- تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$، أي: $N_{M_e}^{\uparrow} \geq \frac{n}{2}$

- حساب الوسيط بطريقة المد الداخلي:

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times A_{M_e}$$

حيث:

Lim_{M_e}: الحد الأدنى للفئة الوسيطية، $N_{M_e-1}^{\uparrow}$: التكرار المتجمع الصاعد المطلق للفئة قبل الفئة الوسيطية.

n_{M_e}: تكرار الفئة الوسيطية ، A_{M_e}: طول الفئة الوسيطية.

مثال10: بالعودة إلى المثال5، أحسب الوسيط واشرح النتيجة؟

أوزان الطلبة X_i	**عدد الطلبة n_i**	**$.N_i^{\uparrow}$**

$]55-50]$	2	2
$]60-55]$	5	7
$]65-60]$	12	19
$]70-65]$	16	35
$]75-70]$	14	49
$]80-75]$	8	57
$]85-80]$	3	60
المجموع $\sum n_i$	60	/

تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$ ، أي:

$$N_{M_e}^{\uparrow} \geq \left(\frac{n}{2} = \frac{60}{2} = 30\right)$$

ومنه الفئة الوسيطية هي: $]70-65]$

- حساب الوسيط بطريقة المد الداخلي:

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times A_{M_e} = 65 + \left[\frac{30-19}{16}\right] \times 5 = 68{,}44$$

الشرح: هناك 50% من الطلبة أوزانهم أقل من 68,44 كلغ و 50% من الطلبة أوزانهم أكبر من 68,44 كلغ.

ملاحظة: الوسيط بيانيا هو نقطة التقاطع بين المنحنى المتجمع الصاعد والنازل.

مثال11: حدد قيمة الوسيط بيانيا للمثال5؟