الفصل الرابع

مقاييس التمركز أو التوسط

Measures of Central Tendency

(٤ : ١) مقدمة

ان معظم القيم لمختلف الظواهر الطبيعية تتمركز عادة في الوسط أو قريبة منه . ومقاييس التمركز أو التوسط لأي مجموعة من البيانات التابعة لظاهرة ما ، هي تلك المقاييس التي تبحث في تقدير قيمة تتمركز حولها أغلبية هذه البيانات وان هذه القيمة المتوسطة أو المتمركزة هي رقم واحد يعبر عن أو يمثل جميع بيانات تلك المجموعة .

وأهم مقاييس التوسط هي :

(١) الوسط الحسابي (أو المتوسط) $\bar{y}$ — The Arithmetic Mean

(٢) الوسط الهندسي G — The Geometric Mean

(٣) الوسط التوافقي H — The Harmonic Mean

(٤) الوسط التربيعي Q — The Quadratic Mean

(٥) الوسيط Me — The Median

(٦) المنوال Mo — The Mode

هذا وسنشرح كيفية حساب كل مقياس من المقاييس أعلاه في حالتين :

١. حالة البيانات غير المبوبة

٢. حالة البيانات المبوبة

(٤ : ٢) الوسط الحسابي — The Arithmetic Mean

الوسط الحسابي أو المتوسط لقيم متغير ما هو القيمة الناتجة من قسمة مجموع تلك القيم على عددها ويرمز له بالرمز $\bar{y}$.

طرق حسابه :

(أ) من بيانات غير مبوبة :

تعريف (٤ : ١) :

اذا كان لدينا n من القيم أو المشاهدات : $y_1 , y_2 , ..., y_n$

فإن الوسط الحسابي لها هو : $\bar{y} = \frac{\sum^{n} y_i}{n}$

مثال (١) : البيانات التالية تمثل كمية المطر الساقطة سنويا (بالمليمترات) على مدينة الموصل خلال فترة خمس سنوات ٥٢٠ ، ٣٥٠ ، ٤٥٠ ، ٣٨٠ ، ٤٠٠
فما هو متوسط سقوط المطر خلال هذه الفترة ؟

الحـل :

$$\bar{y} = \frac{\sum y_i}{n} = \frac{520 + 350 + 450 + 380 + 400}{5}$$

$$= \frac{2100}{5}$$

$$= 420 \text{ mm.}$$

اي ان معدل سقوط الأمطار خلال تلك الفترة هو ٤٢٠ ملم .

مثال (٢) : أحسب الوسط الحسابي لأطوال نباتات القطن في جدول (٣ : ٥) قبل تبويبها .

الحـل :

$$\bar{y} = \frac{\sum y_i}{n} = \frac{80 + 84 + ... + 75}{80}$$

$$= \frac{6126}{80} = 76{\cdot}58 \quad \text{cm.}$$

اي ان معدل طول النبات هو ٧٦،٥٨ سم .

(ب) من بيانات مبوبة :

تعريف (٤ : ٢) :

إذا كانت $y_1 , y_2 , \ldots\ldots y_k$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها $f_1 , f_2 , \ldots\ldots f_k$ على التوالي ، فالوسط الحسابي هو =

$$\bar{y} = \frac{\sum_{i=1}^{k} f_i y_i}{\sum_{i=1}^{k} f_i}$$

خطوات إيجاد الوسط الحسابي في بيانات مبوبة هي كالآتي :

(١) تعيين مراكز الفئات y_i.

(٢) ضرب مركز كل فئة بمقدار تكرارها $(f_i y_i)$

(٣) قسمة مجموع (حاصل ضرب مركز كل فئة × تكرارها) على مجموع التكرارات .

مثال (٣) : إستخرج الوسط الحسابي لأطوال النباتات من جدول التوزيع التكراري (٣ : ٦).

الحـل : عين مركز الفئات ثم أضرب مركز كل فئة × تكرارها كما في الجدول التالي :

جدول (٤ : ١)

الفئـات	التكـرار	مركز الفئـات	التكرار × مركز الفئات
٣٦ – ٤٠	١	٣٥,٥	٣٥,٥
٤١ – ٥٠	٢	٤٥,٥	٩١,٠
٥١ – ٦٠	٥	٥٥,٥	٢٧٧,٥
٦١ – ٧٠	١٥	٦٥,٥	٩٨٢,٥
٧١ – ٨٠	٢٥	٧٥,٥	١٨٨٧,٥
٨١ – ٩٠	٢٠	٨٥,٥	١٧١٠,٠
٩١ – ١٠٠	١٢	٩٥,٥	١١٤٦,٠
	$\Sigma f_i = 80$		$\Sigma f_i x_i = 6130{\cdot}0$

$$\therefore \quad \bar{y} = \frac{\sum f_i y_i}{\sum f_i} = \frac{6130}{80} = 76{\cdot}62 \text{ cm.}$$

أي ان معدل طول النبات هو ٧٦,٦٢ سم .

لاحظ بأن هذا الرقم يختلف قليلا عن الوسط الحسابي لنفس البيانات قبل تبويبها ووضعها في جدول توزيع تكراري (٧٦,٥٨ سم) . ان الفرق هذا بين الرقمين يعود ألى فقدان المعلومات عن المفردات او المشاهدات بسبب وضعها في مجاميع فنحن نفرض بأن طول كل النباتات في فئة معينة مساوياً لمركز تلك الفئة .

مهم جداً // (٣) خواص الوسط الحسابي

(أ) مجموع إنحرافات القيم عن وسطها الحسابي تساوي صفرا

أي (للبيانات غير المبوبة) $\sum (y_i - \bar{y}) = 0$

او (للبيانات المبوبة) $\sum f_i(y_i - \bar{y}) = 0$

البرهان :

$$\begin{aligned}
\sum (y_i - \bar{y}) &= \sum y_i - \sum \bar{y} \\
&= \sum y_i - n\bar{y} \\
&= \Sigma y_i - \Sigma y_i \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\sum f_i(y_i - \bar{y}) &= \sum f_i y_i - \bar{y}\sum f_i \\
&= \sum f_i y_i - \left(\frac{\sum f_i y_i}{\sum f_i} \right) \sum f_i \\
&= \sum f_i y_i - \sum f_i y_i \\
&= 0
\end{aligned}$$

والجدولان التاليان يوضحان ذلك :

جدول (٤ : ٢)

$(y_i - \bar{y})$	y_i
0·4	8
− 4·6	3
− 2·6	5
4·4	12
2·4	10
$\sum(y_i - \bar{y}) = 0$	$\sum y_i = 38$ $\bar{y} = 7{\cdot}6$

جدول (٤ : ٣)

الفئات	التكرار f_i	مركز الفئات y_i	$f_i y_i$	$(y_i - \bar{y})$	$f_i(y_i - \bar{y})$
٦٢ – ٦٠	٥	٦١	٣٠٥	– ٦.٤٥	– ٣٢.٢٥
٦٥ – ٦٣	١٨	٦٤	١١٥٢	– ٣.٤٥	– ٦٢.١٠
٦٨ – ٦٦	$\boxed{٤٢}$	٦٧	٢٨١٤	$\boxed{– ٠.٤٥}$	– ١٨.٩٠
٧١ – ٦٩	٢٧	٧٠	١٨٩٠	+ ٢.٥٥	+ ٦٨.٨٥
٧٤ – ٧٢	٨	٧٣	٥٨٤	+ ٥.٥٥	+ ٤٤.٤٠
	$\sum f_i =$ ١٠٠		$\sum f_i y_i =$ ٦٧٤٥ $\bar{y} = \frac{\sum f_i y_i}{\sum f_i} =$ ٦٧.٤٥		$\sum f_i(y_i - \bar{y}) = 0$

(ب) مجموع مربعات الانحرافات عن الوسط الحسابي هي أقل ما يمكن أي أقل من مجموع مربعات الانحرافات عن أية قيمة غير الوسط الحسابي نفسه أي ان $\sum(y_i - \bar{y})^2$ أقل ما يمكن .

البرهان :

نفرض أن A هو أي قيمة أو وسط فرضي غير الوسط الحسابي فسنبرهن بأن $\sum(y_i - A)^2$ هي أكبر من قيمة $\sum(y_i - \bar{y})^2$: