مقاييس التمركز او التوسط

في العمليات الاحصائية وعند جمع البيانات يحصل تجمع وتراكم لعدد كبير من قيم المتغير حول قيمة معينة وهذا الميل للتجمع حول تلك القيمة يسمى بالنزعة المركزية للتوزيع البياني وتسمى تلك القيمة التي يحصل التجمع حولها بمقياس النزعة المركزية (القيمة المتوسطة).

إن مقاييس النزعة المركزية (القيم المتوسطة) هي قيم احصائية ذات أهمية كبيرة في وصف التوزيعات البيانية ومقارنتها وتلخيصها وتمثيلها ومن أهم مقاييس النزعة المركزية التي سنتناول دراستها وطرق استخراجها في هذه المرحلة هي:

1- الوسط الحسابي 2 - الوسيط 3 - المنوال 4 - الوسط الهندسي 5 - الوسط التوافقي.

ويتوقف اختيار ايجاد احدى (القيم المتوسطة) في العملية الاحصائية على طبيعة البيانات التي يدرسها الاحصائي وعلى الغاية التي ينشدها من دراسة تلك البيانات كما يتوقف على صلاحية ذلك المقياس لإظهار أفضل النتائج إذ أن لكل مقياس استخداماته في مجال العمليات الاحصائية. كما أن لكل مقياس مزايا وعيوباً في تمثيل التوزيعات الاحصائية.

(*) لقد وضع (يول) أسساً للمفاضلة بين مقاييس النزعة المركزية لاختيار نوع المقياس الذي يجب استخدامه في العملية الاحصائية لكي يكون ذلك المقياس جديراً بذلك التمثيل لتلك التوزيعات الاحصائية. وتتلخص هذه الأسس بما يلي:

1 - من المستحسن أن تكون القيمة المتوسطة (مقياس النزعة المركزية) قيمة موضوعية محدودة وليست مجرد تقرير ذاتي من الباحث. اذ يستحسن أن تكون طريقة حسابها طريقة رياضية لا يختلف فيها اثنان بسيطة وغير معقدة.

2- من الأفضل استخدام جميع قيم المتغير عند إيجاد القيمة المتوسطة لكي تكون ممثلة لمميزات التوزيع بأكمله.

3 - من الأفضل أن تكون القيمة المتوسطة من القيم التي يكون تأثرها بذبذبات العينة الاحصائية قليل بقدر الامكان لأن من النادر أن تكون القيم المتوسطة متساوية.

٤ - من المستحسن أن تحقق القيمة المتوسطة أهم الشروط وهي صلاحيتها للمعالجة الرياضية (أي إمكانية استخدامها في القوانين الاحصائية الأخرى مثل الانحراف المعياري) او معامل الاختلاف او الدرجة المعيارية الذي سوف يأتي ذكرها فيما بعد.

1 - الوسط الحسابي:

الوسط الحسابي لمجموعة من قيم المتغير هو خارج قسمة المجموع الجبري لهذه القيم على عددها ويرمز له $\underline{X}$، والمجموع قيم المتغير $\sum Xi$ وعدد قيم المتغير n

أ - الوسط الحسابي لقيم المتغير غير المبوبة في جدول تكراري:

يحسب الوسط الحسابي في هذه الحالة حسب القانون الآتي: نفرض ان القيم هي X1، X2، Xn .....

مثال (1) :

كان عدد الثمار في كل نبات و لأصناف مختلفة من الباذنجان كما يلي:

4، 8 ، 10 ، 12 ، 16 ، 18 ، 20 ، 24

المطلوب: جد الوسط الحسابي لعدد الثمار

الحل: $\underline{X} = \frac{\sum_{i=1}^{n} Xi}{n}$

$$\underline{X} = \frac{4 + 8 + 10 + 12 + 16 + 18 + 20 + 24}{8}$$

$\underline{X} = \frac{112}{8} = 14$ عدد الثمار في النبات الواحد

طريقة أخرى (طريقة الوسط الفرضي – الطريقة الانحرافات)

تستخدم هذه الطريقة في حالة كون قياسات العينة اعداد كبيرة يصعب التعامل معها عند ايجاد الوسط الحسابي خصوصا عند عدم توفر حاسبات تفي بالغرض مما يفضل اختزال هذه الاعداد الى اعداد اصغر يسهل التعامل معها . نفرض وسطاً حسابياً يكون قريب من الوسط الحسابي من نفس البيانات او خارج عنها ومن الأفضل اختياره من بين قيم المتغير وليكن A = 12 ثم نحسب انحرافات جميع قيم المتغير عن الوسط المفروض. ثم نجد انحرافات القيم عن الوسط الفرضي ويرمز له di وعليها فان الوسط الحسابي يكون كالاتي :

$$\underline{X} = A + \frac{\sum di}{n}$$

حيث ان:

تمثل الوسط الفرضي = A

تمثل انحرافات القيم عن الوسط الفرضي = di

و ان di= Xi - A

ملاحظة قيمة A قيمة ثابتة نختارها من ضمن القيم ، نعتقد انها تمثل وسط القيم او مركزها اي تكون قريبة من الوسط الحسابي.

كان عدد الثمار في كل نبات و لأصناف مختلفة من الباذنجان كما يلي:

4، 8 ، 10 ، 12 ، 16 ، 18 ، 20 ،24

المطلوب: جد الوسط الحسابي لعدد الثمار في النبات بطريقة الانحرافات؟

الحل:

1- نحدد قيمة A بحيث تكون قريبة من الوسط الحسابي وفي قيمة A تمثل 12 او اي قيمة اخرى نختارها
2- إيجاد قيم d من القانون التالي:

di= Xi – A

D1= 4-12= -8	D5= 16-12= 4
D2= 8-12= -4	D6= 18-12= 6
D3= 10-12= -2	D7= 20-12= 8
D4= 12-12= 0	D8= 24-12= 12
∑di= 16	

3- نجمع ∑di

∑di= d1+d2+d3+d4+d5+d6+d7+d8

∑di= 16

4- الآن يتم تطبيق القانون لأستخراج قيمة الوسط الحسابي.

$$\underline{X} = A + \frac{\sum di}{n}$$

$$\underline{X} = 12 + \frac{16}{8} = 14$$

ب- الوسط الحسابي لقيم المتغير المبوبة في جدول تكراري:

1- الطريقة المباشرة:

يستخرج الوسط الحسابي من القانون الآتي:

$$\underline{X} = \frac{\sum fi\, Xi}{\sum fi}$$

حيث ان :

الوسط الحسابي= $\underline{X}$

التكرارات= fi

مراكز الفئات= Xi

خطوات ايجاد الوسط الحسابي لبيانات مبوبة

1- تعيين مراكز الفئات.
2- ضرب مركز كل فئة في التكرار المقابل له.
3- قسمة مجموع (حاصل ضرب مركز كل فئة * تكرارها) على مجموع التكرارات.

مثال(2) : الجدول الاتي يبين توزيع عدد الثمار للنبات لفترة أسبوعين في لـ 200 صنف من أصناف الطماطة المختلفة و المزروعة في بيئة محمية.

المطلوب: جد الوسط الحسابي لعدد الثمار

عدد الثمار(ثمرة نبات)$^{-1}$	20-23	24-27	28-31	32-35	36-39
عدد الأصناف	40	30	60	50	20

الحل:

الفئات	التكرار (fi)	مراكز الفئات (Xi)	مركز الفئة * التكرار (fi Xi)
23-20	40	21.5	860
27-24	30	25.5	765
31-28	60	29.5	1770

35-32	50	33.5	1675
39-36	20	37.5	750
المجموع	200		5820

نطبق القانون

$$\underline{X} = \frac{\sum fi\ Xi}{\sum fi}$$

$\underline{X} = \frac{5820}{200} = 29.1$ الوسط الحسابي لعدد الثمار للصنف الواحد

طريقة الوسط الفرضي:

نفرض وسطاً فرضياً و الأفضل اختياره من بين مراكز الفئات و نحسب انحراف مركز كل فئة عن الوسط الفرضي ثم يستخرج الوسط الحسابي من القانون :

$$\boldsymbol{X = A + \frac{\sum fidi}{\sum fi}}$$

خطوات ايجاد الوسط الحسابي للبيانات مبوبة:

1 -نجد مراكز الفئات Xi

2 -نحدد وسطا فرضيا A

3 -نجد انحرافات المراكز عن الوسط الفرضي di من القانون التالي:

di = Xi - A

4 -نضرب كل انحراف في التكرار المقابل له fi di

5 -نجمع حاصل الضرب (الانحرافات * التكرارات)

6 -نطبق القانون

ملاحظة : اختيار الوسط الفرضي يؤدي الى تبسيط العمليات الحسابية ويكون اختياره حسب القواعد الاتية :

1 -ان يكون الوسط الفرضي مركز لأحدى الفئات

2 -ان يكون قريبا من الوسط الحسابي

3 -ان يكون امام اكبر تكرار (30)

- نرجع الى المثال السابق:

الفئات	التكرار (fi)	مراكز الفئات (Xi)	di = Xi - A	الانحرافات * التكرار (fi di)
23-20	40	21.5	21.5-30= -8.5	-340
27-24	30	25.5	25.5-30= -4.5	-135
31-28	60	29.5	29.5-30= -0.5	-30
35-32	50	33.5	33.5-30= 3.5	175
39-36	20	37.5	37.5-30= 7.5	150
المجموع	200			180-

الآن نطبق القانون:

$$X = A + \frac{\sum fidi}{\sum fi} = 30 + \frac{}{200} = 29.1$$

مزايا وعيوب الوسط الحسابي:

يتميز الوسط الحسابي بالمزايا التالية:

1 .أنه سهل الحساب .

2 .يأخذ في الاعتبار كل القيم .

3 .أنه أكثر المقاييس استخداما وفهما .

ومن عيوبه .

1 .أنه يتأثر بالقيم الشاذة والمتطرفة .

2 .يصعب حسابه في حالة البيانات الوصفية .

3 .يصعب حسابه في حالة الجداول التكرارية المفتوحة