وزارة التعليم العالي والبحث العلمي
جامعة تكريت
كلية الإدارة والاقتصاد
قسم ادارة الاعمال
المرحلة الاولى

محاضرات في مادة الاحصاء

مدرس المادة

أ.م.د هادي مران احمد

1446 هـ **2024 مـ**

الفصل الثالث

مقاييس النزعة المركزية (مقاييس التوسط)

مقدمة :

إن معظم القيم لمختلف الظواهر تتحرك عادة في الوسط أو بالقرب منه (ومقاييس) التوسط لأي مجموعة من البيانات التابعة لظاهرة ما ، هي تلك المقاييس التي تبحث في تقوية قيمة تتمركز حولها أغلبية هذه البيانات وأن هذه القيمة المتوسطة أو المتمركزة هي عبارة عن رقم واحد يعبر عن أو يمثل جميع بيانات لتلك المجموعة ، هذا العدد يميل لأن يقع في وسط تلك المجموعة من البيانات في حالة ترتيبها حسب صغرها أو كبرها لذا أطلق عليها مقاييس النزعة المركزية.

أهم مقاييس النزعة المركزية :

1- الوسط الحسابي.

2- الوسط الهندسي.

3- الوسط التوافقي.

4- الوسط التربيعي.

5- الوسيط.

6- المنوال.

1- الوسط الحسابي :

ويسمى في بعض الأحيان الوسط أو المتوسط أو المعدل الحسابي لقيم متغير ما ، وهو أهم مقاييس النزعة المركزية على الإطلاق لما يمتاز بخصائص جيدة بسهولة في الحساب ، وهو عبارة عن القيمة الناتجة من قسمة مجموع المشاهدات على عددها ويرمز له بالرمز $\overline{x}$.

طرق حساب الوسط الحسابي :

أ- في حالة البيانات الغير المبوبة :

الوسط الحسابي يكون مساوي إلى :

$$\bar{x} = \frac{x_1 + x_2 + \dots x_n}{n} = \frac{\sum_{i=1}^{n} xi}{n}$$

حيث أن $\bar{x}$ يمثل تقدير الوسط الحسابي لقياسات مفردات المجتمع الذي اختيرت منه العينة.

مثال:

البيانات التالية أوجد الوسط الحسابي لأوزان عمال لعينة حجمها 10عمال.

وزن العمال

(50.2, 52.9, 58.1, 59.2, 60.9, 61.9, 62.3, 63.2, 65.3, 68.3)

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{(50.2 + 52.9 + \dots\dots 68.3)}{10}$$

$$= \frac{601.9}{10}$$

$$= 60.19$$

يلاحظ تمركز قيمة $\bar{x}$ وسط هذه المجموعه لذا سمي مقياس نزعة مركزية

ب- حساب الوسط الحسابي في حالة البيانات المبوبة :

لتكن $x_1 , x_2 \dots , x_n$ تمثل مراكز الفئات لتوزيع تكراري عدد فئاته n وبأن $f_1 , f_2 , \dots , f_n$ تمثل التكرارات المقابلة لهذه الفئات يستخرج الوسط الحسابي ، سواء كانت أطوال الفئات متساوية أم لا حيث يعبر عن صيغة الوسط الحسابي للبيانات المبوبة بالشكل التالي :

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

حيث أن :

Xi= مركز الفئة.

$\sum_{i=1}^{n} f_i x_i$: تمثل حاصل ضرب مراكز الفئات بالتكرارات.

مثال 2 :

الآتي توزيع تكراري لدرجات الحرارة في مدينة معينة لمدة 95 يوماً يطلب إيجاد متوسط درجة الحرارة خلال هذه الفترة.

درجات الحرارة (الفئات)	عدد الأيام fi	مركز الفئة xi	حاصل ضرب fi xi
0 – 1	4	0.5	(4) (0.5) = 2
1 – 2	8	1.5	(8) (1.5) = 12
2 – 3	12	2.5	(12) (2.5) = 30
3 – 4	16	3.5	56
4 – 5	20	4.5	90
5 – 6	25	5.5	137.5
6 – 7	6	6.5	39
7 – 8	4	7.5	30
	95		396.5

مركز الفئة = $\frac{0+1}{2} = 0.5$

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{396.5}{95} = 4.174$$

تمرين : الآتي توزيع تكراري لأطوال عينة من الأشخاص قوامها (*69*) شخص يطلب إيجاد متوسط طول كل شخص في هذه العينة.

فئات الأطوال	عدد الأشخاصfi	مركز الفئة xi	fi xi
140 – 148	4	144	576
148 – 156	6	152	912
156 – 164	15	160	2400
164 – 172	20	168	3360
172 – 180	17	176	2992
180 – 188	7	184	1288

188 - 196	1	192	192
	69		11720

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{\mathbf{11720}}{\mathbf{69}} = 169.85$$

مزايا وعيوب الوسط الحسابي :

أ- المزايا :

1- بساطة وسهولة فكرته.

2- أن حسابه يستند إلى كافة البيانات المتاحة.

3- أن حسابه يخضع للعمليات الجبرية.

ب- العيوب :

1- لا يمكن إيجاد الوسط الحسابي وبيانات وصفية كالجنس واللون وصنف الدم ما عدا الحالات التي يتم فيها إعادة تقييس الصفات.

2- لا يمكن حساب الوسط الحسابي في حالة فقدان قيمة أو أكثر من قيم العينة.

3- أن قيمة الوسط الحسابي تتأثر بالقيم الشاذة (المتطرفة) لذا لا يمكن الاعتماد على هذا القياس في هذه الحالة.

4- لا يمكن إيجاده في حالة التوزيعات التكرارية المفتوحة من طرف أو طرفين.

5- لا يجوز الاعتماد على قيم الوسط الحسابي في حالة التوزيعات ذات التواء حاد سواء أكانت التواء سالب أو موجب بسبب انحياز قيمة الوسط الحسابي نحو جهة الالتقاء.

خصائص الوسط الحسابي :

1- مجموع انحرافات قيم المتغير x من وسطها الحسابي الذي احتسبت منه يساوي صفر :

أ- في حالة البيانات غير المبوبة.

$$\sum_{i=1}^{n}(x_i - \bar{x})$$

البرهان : نفرض ان

$$d_i = x_i - \bar{x} \qquad , i= 1,2,\ldots\ldots,n$$

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n}(x_i - \bar{x})$$ بإضافة المجموع للطرفين

$$= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} \longrightarrow \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \longrightarrow n\bar{x} = \sum_{i=1}^{n} x_i$$

نعوض عن $\sum \boldsymbol{xi}$ بما يساوي في المعادلة يكون :

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i = 0$$

ب- في حالة البيانات المبوبة :

$$\sum_{i=1}^{n} f_i (x_i - \bar{x}) = 0$$

البرهان :

$$\sum_{i=1}^{n} f_i x_i - \bar{x}\sum_{i=1}^{n} f_i \rightarrow \bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

$$\bar{x}\sum_{i=1}^{n} f_i = \sum_{i=1}^{n} f_i x_i$$

نعوض عن $\sum_{i=1}^{n} f_i x_i$ بما يساويها في المعادله اعلاه

$$\bar{x}\sum_{i=1}^{n} f_i - \bar{x}\sum_{i=1}^{n} f_i = 0$$

الوسط التوافقي Harmonic mean :

يعتبر الوسط التوافقي أحد مقاييس النزعة المركزية ذات الاستخدامات القليلة في التطبيقات الإحصائية ويرمز له بالرمز H نسبة إلى كلمة Hormonic.

أ- في حالة البيانات غير المبوبة :

يأخذ الوسط التوافقي الصيغة التالية بعد أخذ مقلوب القيم كما يلي :

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

حيث أن : $\frac{\mathbf{1}}{\boldsymbol{xi}}$ تمثل مقلوب قيم المشاهدات.

مثال 3 :

جد الوسط التوافقي للبيانات التالية علماً أنها تمثل عدد الموظفين في بعض أقسام الشركة العامة للنقل البري :

8 , 8 , 7 , 4 , 3 , 5 , 2 قيم المشاهدات

الحل :

$$\sum_{i=1}^{n} \frac{1}{x_i} = \frac{\mathbf{1}}{\mathbf{2}} + \frac{\mathbf{1}}{\mathbf{5}} + \frac{\mathbf{1}}{\mathbf{3}} + \frac{\mathbf{1}}{\mathbf{4}} + \frac{\mathbf{1}}{\mathbf{7}} + \frac{\mathbf{1}}{\mathbf{8}} + \frac{\mathbf{1}}{\mathbf{8}}$$

$= 0.5 + 0.2 + 0.33 + 0.25 + 0.14 + 0.13 + 0.13 = 1.68$

$H = \frac{\mathbf{7}}{\mathbf{1.68}} = 4.17$

ب- في حالة البيانات المبوبة :

يأخذ الوسط التوافقي الصيغة التالية بالاعتماد على مقلوب قيم مركز الفئة :

$$H = \frac{\sum_{i=1}^{m} f_i}{\sum_{i=1}^{m} \frac{f_i}{x_i}}$$

علماً أن fi تمثل التكرار xi مراكز الفئات.

مثال 4 :

البيانات التالية تمثل عدد العاملين f أحد المصانع التابعة لوزارة الصناعة.

المطلوب : احسب الوسط التوافقي ؟

فئات الأجر	عدد العاملين fi	مركز الفئة xi	$\frac{fi}{xi}$
50 – 60	8	55	$\frac{8}{55}$ = 0.145
60 – 70	10	65	$\frac{10}{165}$ = 0.154
70 – 80	16	75	$\frac{16}{75}$ = 0.213
80 – 90	14	85	0.165
90 – 100	10	95	0.105
100 – 110	5	105	0.048
110 - 120	2	115	0.017
	65		0.874

$$H = \frac{\sum_{i=1}^{m} f_i}{\sum_{i=1}^{m} \frac{f_i}{x_i}} = \frac{65}{0.874} = 74.741$$

عيوب الوسط التوافقي :

1- لا يمكن تحديد قيمة الوسط التوافقي إذا كانت إحدى قيم المتغير العشوائي مساوية للصفر أو إحدى مراكز فئات التوزيع التكراري مساوية للصفر.

2- لا يمكن إيجاد قيمته في حالة التوزيعات التكرارية المفتوحة من طرف واحد أو كلا الطرفين.

3- لا يمكن حساب قيمته في حالة فقدان قيمة أو أكثر من قيم العينة.

4- لا يمكن حساب قيمته في حالة فقدان البيانات النوعية.

3- الوسط الهندسي *Geometric mean* :

أ- في حالة البيانات غير المبوبة :

لو فرضنا لدينا القيم أو المشاهدات التالية $x_1 , x_2 , \dots , x_n$، فالوسط الهندسي هو عبارة عن الجذر النوني الحاصل ضرب تلك القيم ويرمز له بالحرف G نسبة إلى كلمة Geometric وكما يلي :

$$G = \sqrt[n]{x_1.x_2......x_n}$$

يرفع الجذر : $G = (x_1.x_2.......x_n)^{\frac{1}{n}}$

بأخذ لوغاريتم الطرفين :

$$\log G = \frac{1}{n}\log(x_1.x_2.....x_n)$$

Log G = $\frac{\boldsymbol{1}}{\boldsymbol{n}}[\mathbf{log}\,\boldsymbol{x}\mathbf{1} + \mathbf{log}\,\boldsymbol{x}\mathbf{2} + \cdots + \mathbf{log}\,\boldsymbol{n}]$

الصيغة اللوغارتمية للوسط الهندسي في حالة البيانات غير المبوبة $\log G = \dfrac{x\sum_{i=1}^{n}\log x_i}{n}$

مثال 5 :

القيم التالية تمثل عدد الوحدات المنتجة في أحد المصانع الإنتاجية.

المطلوب : أوجد الوسط الهندسي لهذه الوحدات :

Xi : 3 , 5 , 7 , 8 , 3 , 7 , 2

الحل :

- الطريقة الأولى :

G = $\sqrt[6]{(\mathbf{3})(\mathbf{5})(\mathbf{7})(\mathbf{3})(\ \mathbf{7})(\ \mathbf{2})}$ = 4.14

- الطريقة الثانية : باستخدام اللوغاريتم :

Log G = $\dfrac{\mathbf{X}\ \sum_{i=1}^{6}\mathbf{log}\,\boldsymbol{xi}}{\mathbf{6}} = \dfrac{\mathbf{log(3)+log(5)+log(7)+log(3)+log(7)+log\ (2)}}{\mathbf{6}}$

$$= \frac{0.477+0.698+0.903+0.477+0.845+0.301}{6}$$

$\text{Log } G = 0.617 \rightarrow G = 4.14$

ملاحظة : في المثال رقم 5 أعلاه في حالة احتساب الوسط الحسابي للبيانات فأنه يكون مساوي إلى $\bar{x} = 4.64$، ومن هذه النتيجة يتضح أن الوسط الحسابي لمجموعة من القيم أكبر من الوسط الهندسي لتلك القيم ، كما أنه إذا كان أحد القيم سالباً فلا يمكن إيجاد الوسط الهندسي لها.

ب- الوسط الهندسي في حالة البيانات المبوبة :

الوسط الهندسي في حالة البيانات المبوبة يمثل وسط هندسي مدون ، حيث أن التكرارات في الواقع تمثل الأوزان التي تبين أهمية كل فئة ، فلو كانت $x_1 , x_2 , \ldots , x_n$ تمثل مراكز الفئات في جدول التوزيع التكراري وتكراراتها هي $f_1 , f_2 , \ldots , f_m$ فأن الوسط الهندسي يكون مساوي للصيغة التالية :

$$G = \sqrt[\sum_{i=1}^{m} f_i]{\frac{f_1 , f_2 , \ldots . f_m}{x_1 . x_2 \ldots . x_m}}$$

ونظراً لصعوبة حل الصيغة أعلاه نأخذ لوغاريتم الطرفين بعد رفع الجذر كما يلي :

$$G = \left(\frac{f \, , , f \, , , , , f}{x_1 . x_2 \ldots . x_m} \right)^{\frac{1}{\sum_{i=1}^{m} f_i}}$$

$$\log G = \frac{1}{\sum_{i=1}^{m} f_i} \log \left(\frac{f \, , , f \, , , , , f}{x_1 . x_2 \ldots . x_m} \right)$$

$$= \frac{1}{\sum_{i=1}^{m} f_i} \left(f_1 \log x_1 + f_2 \log x_2 + \ldots + f_m \log x_m \right)$$

الصيغة النهائية باستخدام اللوغاريتم

$$\log G = \frac{\sum_{i=1}^{m} f_i \log x_i}{\sum_{i=1}^{m} f_i}$$

مثال // البيانات التالية تمثل عدد الموظفين المحالين على التقاعد والاستقالة هي إحدى الوزارات موزعين حسب الفئات العمرية لهم .

المطلوب : أوجد الوسط الهندسي لهذه البيانات :

الفئات العمرية	عدد العاملين fi	مركز الفئة xi	log xi	fi log xi
60 – 62	5	61	1.79	(5) (1.79) = 8.8910
63 – 65	18	64	1.81	(18) (1.81) = 32.5116
66 - 68	42	67	1.81	76.6962
69 – 71	27	70	1.85	49.8177
72 – 74	8	73	1.86	14.9064
	100			182.8229 $\rightarrow \sum_{i=1}^{m} \boldsymbol{fi \log xi}$

$$\log G = \frac{\sum_{i=1}^{m} f_i \log x_i}{\sum_{i=1}^{m} f_i} = \frac{\mathbf{182.8229}}{\mathbf{100}} = 1.828229 = 67.3$$

عيوب الوسط الهندسي :

1- لا يمكن استخدامه في حالة الجداول المفتوحة.

2- لا يمكن استخدامه إذا كانت ‘احدى قيم المجموعة صفر.

3- لا يمكن استخدامه إذا كانت قيم المجموعة كمية سالبة.

4- الوسط الحسابي الموزون (المرجح) :

في بعض الأحيان تكون إحدى المفردات أكثر أهمية من الاخرى لذلك يتوجب أخذ ذلك بنظر الاعتبار عند حساب الوسط الحسابي الموزون.

أ- في حالة البيانات غير المبوبة :

افترض أن x_1 , x_2 , ... , x_n تمثل قياسات عينة من المفردات قوامها n وأن w_1 , w_2 , ... , w_n أوزان هذه المفردات يعرف الوسط الحسابي المرجح على النحو التالي :

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

مثال 7 : لديك القيم التالية وأوزانها ، المطلوب إيجاد الوسط الحسابي الموزون :

Xi	Wi	xi wi
70	10	(70) (10) = 700
60	30	(60) (30) = 1800
75	10	(75) (10) = 750
55	50	(55) (50) = 2750
المجموع	100	6000

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} = \frac{\mathbf{6000}}{\mathbf{100}} = 60$$

ب- الوسط الحسابي الموزون في حالة البيانات المبوبة :

أما في حالة التوزيعات التكرارية فأن الوسط الحسابي الموزون يكون وفق الصيغة التالية :

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i f_i x_i}{\sum_{i=1}^{n} w_i f_i}$$

حيث أن fi تمثل التكرار

Xi = مركز الفئة

wi = الأوزان

مثال 8 : الآتي توزيع تكراري لإنتاج مصنع معين (طن) من سلعة معينة لأحد الأيام موزع حسب عدد المكائن العاملة في ذلك اليوم وعدد ساعات العمل المحددة لإشغال كل ماكنة حسب مواصفات المنشأ ، يطلب حساب متوسط انتاجية الماكنة الواحدة في هذا المصنع.

فئات الانتاج	عدد المكائن Fi	عدد الساعات الاوزان wi	مركز الفئة xi	wi fi	wi fi xi
2 – 4	4	6	3	(6) (4) = 24	(24) (3) = 72
4 – 6	5	5	5	(5) (5) = 25	(25) (5) = 125
6 – 8	6	6	7	(6) (6) = 36	(36) (7) = 252
8 – 10	3	4	9	(3) (4) = 12	(12) (9) = 108
10 – 12	2	4	11	(2) (4) = 8	(8) (11) = 88
	20			105	645

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i f_i x_i}{\sum_{i=1}^{n} w_i f_i} = \frac{\mathbf{645}}{\mathbf{105}} = 6.143$$