الفصل الثاني

مقاييس النزعة المركزية

٢- ١ **الوسط الحسابي** : ويسمى في بعض الاحيان (الوسط) او (المعدل الحسابي) وهو احد اهم مقاييس النزعة المركزية ويمكن ايجاده بسهولة للبيانات مبوبة و غير مبوبة .

أ) اذا كانت البيانات مبوبة : اذا كان لدينا n من القيم اي $x_1, x_2, \dots, x_n$ فأن الوسط الحسابي لها هي $x = \frac{\sum x_i}{n}$

ب) البيانات غير مبوبة : اذا كان لدينا $(x_1, x_2, \dots, x_n)$ تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها $(f_1, f_2, \dots, f_n)$ على التوالي فأن الوسط الحسابي لها $\bar{x} = \frac{\sum fi\, x_i}{\sum fi}$

مثال/ البيانات في جدول التوزيع التكراري تمثل اعمار موظفين المطلوب ايجاد متوسط اعمارهم

فئات	Fi	Xi	Fi xi
21 – 30	8	25.5	204
31 – 40	12	35.5	426
41 – 50	16	45.5	728
51 – 60	9	55.5	499.5
61 – 70	5	65.5	337.5
	$\sum fi = 50$		$\sum_{i=1}^{5} fixi = 2185$

$$\bar{x} = \frac{\sum_{i=1}^{n} fixi}{\sum fi} = \frac{\sum_{i=1}^{5} fixi}{\sum_{i=1}^{5} fi} = \frac{2185}{50} = 43.7 \approx 44$$

٢-٢. **الوسيط** : وهو احد مقاييس النزعة المركزية المهمة في التطبيقات الاحصائية ويعرف بأنه تلك القيمة التي تقسم مجموعة المتغير الى قسمين متساويين ويرمز له بـ Me

أ) اذا كانت البيانات غير مبوبة : اذا كان لدينا n من القيم $(x_1, x_2, \dots, x_n)$ نقوم اولا بترتيب البيانات تصاعديا او تنازليا ويقسم الى حالتين :

١) اذا كان n عدد فردي فان الوسيط هو $Me = \frac{n+1}{2}$

٢) اذا كان n عدد زوجي فأن الوسيط يكون بين $\left(\frac{n}{2}, \frac{n}{2} + 1\right)$

مثال١/ جد الوسيط للقيم الاتية $x = 84\,,80\,,82\,,76\,,87$

ترتيب البيانات تصاعديا $x = 76\,,80\,,82\,,84\,,87$

بما ان n عدد فردي فأن الوسيط $Me = 82$

مثال ٢/ ليكن $x = 5,4,8,7,3,12,9,2$ بما ان n=8 فنطبق القانون $\left(\frac{n}{2}, \frac{n}{2}+1\right)$

اولا ترتيب البيانات تنازليا $x = 12\,,9,8,7,5,4,3,2$

$$\left(\frac{8}{2}, \frac{8}{2}+1\right) = (4,5)$$

اي ان قيمة الوسيط هو الوسط الحسابي للعددين في المرتبة الرابعة والخامسة اي

$$Me = \frac{7+5}{2}$$

$= 6$

ب) البيانات المبوبة : وتقسم الى

١) البيانات مبوبة لمتغير متقطع :

نفرض وجود توزيع تكراري لبيانات متغير (متقطع) عدد فئاته (n) و $f_1\,, f_2\,, \dots, f_n$ تمثل التكرارات المقابلة لفئات هذا التوزيع ، وتفرض ان $F_1\,, F_2\,, \dots, F_n$ تمثل التكرارات المتجمعة الصاعدة المقابلة للحدود العليا ، ومن تعريف الوسيط فأن الوسيط لهذه البيانات يمثل ترتيب $\frac{\sum_{i=1}^{n} fi}{2}$ فاذا كان $F_k < \frac{\sum fi}{2} < F_{k+1}$ فأن فئة الوسيط تسلسلها $(k+1)$.

مثال/ التوزيع التكراري الاتي يمثل عينة من الاسر حسب افراد الاسرة يطلب حساب الوسيط لعدد افراد الاسرة

فئات	تكرار	تكرار مجتمع صاعد Fi
$2-4$	5	تكرار اقل او يساوي (4) ← 5
$5-7$	8	تكرار اقل او يساوي (7) ← 13
$8-10$	11	تكرار اقل او يساوي (10) ← 24
$11-13$	9	تكرار اقل او يساوي (13) ← 33
$14-16$	7	تكرار اقل او يساوي (16) ← 40
	$\sum fi = 40$	

نجد ان $\frac{\sum fi}{2} = 20$ وان ترتيب الوسيط ضمن التكرار المجتمع الصاعد هو $13 < 20 < 24$ وعليه فان $Me = \frac{8+10}{2} = 9$.

٢) البيانات المبوبة لمتغير مستمر

نفرض وجود توزيع تكراري لبيانات متغير (مستمر) عدد فئاته (n) وان $f_1, f_2, \dots, f_n$ تمثل التكرارات المقابلة لفئات التوزيع ونفرض ان $F_1, F_2, \dots, F_n$ تمثل التكرارات المجتمعة الصاعدة المقابلة للحدود العليا لفئات التوزيع وان $\frac{\sum_{i=1}^{n} fi}{2}$ يمثل ترتيب الوسيط في التوزيع فاذا كان

$$F_{k-1} < \frac{\sum_{i=1}^{n} fi}{2} < F_k$$

فان فئة الوسيط هي الفئة التي تسلسلها (k) ويتم حساب قيمة الوسيط حسب معطيات الصيغة التالية :

$$M_e = L_k + \frac{h_k}{f_k}\left(\frac{\sum_{i=1}^{n} fi}{2} - F_{k-1}\right)$$

حيث ان

L_k : الحد الادنى لفئة الوسيط

f_k : تكرار فئة الوسيط

h_k : طول فئة الوسيط

F_{k-1} : التكرار المجتمع الصاعد السابق لفئة قبل الوسيطية.

مثال/ التوزيع التكراري الاتي يمثل الدخل الشهري لعينة من الاسر قوامها (80) اسرة ، المطلوب ايجاد الوسيط للدخل الشهري للأسرة في هذه العينة

الفئات	التكرار (عدد الاسر) Fi	التكرار المجتمع الصاعد Fi
100−	12	تكرار مجتمع صاعد اقل من (120) ← 12
120−	21	تكرار مجتمع صاعد اقل من (140) ← 33
140−	28	تكرار مجتمع صاعد اقل من (160) ← 61
160−	11	تكرار مجتمع صاعد اقل من (180) ← 72
180−< 200	8	تكرار مجتمع صاعد اقل من (200) ← 80
	$\sum_{i=1}^{5} fi = 80$	

ترتيب الوسيط $\frac{\sum_{i=1}^{5} fi}{2} = 40$

نلاحظ ان $61 > 40 > 33$ وعليه فان فئة الوسيط هي (140−)

وبذلك فان $L_3 = 140$ ، $h_3 = 20$ ، $f_3 = 28$ ، $F_2 = 33$

$$Me = L_3 + \frac{h_3}{f_3}\left(\frac{\sum_{i=1}^{5} fi}{2} - F_2\right)$$

$$= 140 + \frac{20}{28}(40 - 33) = 140 + \frac{5}{7} + 7 = 145$$

٢−٣. المنوال The Mode

يعرف المنوال بأنه تلك القيمة التي تتكرر اكثر من غيرها من بين مجموعة من القيم او انها القيمة الشائعة من بين مجموعة من القيم .

١− المنوال لبيانات غير مبوبة

اذا كان لدينا $x_1, x_2, \dots, x_n$ قياسات عينة من المفردات قوامها (n) ونفرض ان x_i قيمة من قيم هذه المجموعة تتكرر اكثر من غيرها عندئذ فأن المنوال هذه المجموعة هو x_i (يرمز للمنوال بـ M_0) .

مثال/ جد المنوال للبيانات التالية : 3 ,1 ,3 ,1 ,3 ,9 ,3 ,5 ,4 ,1 , 3 , 2 واضح ان العدد (3) تكرار اربع مرات وهو اكبر من اي تكرار وعليه فان المنوال لهذه المجموعة هو (3)

٢- المنوال لبيانات مبوبة

أ) المنوال لبيانات مبوبة لمتغير متقطع

اذا كان لدينا $x_1, x_2, \dots, x_n$ تمثل مراكز فئات لتوزيع تكراري عدد فئاته (n) وان $f_1, f_2, \dots, f_n$ تمثل التكرارات المقابلة لفئات هذا التوزيع عندئذ فان المنوال لهذا التوزيع يمثل قيمة مركز الفئة التي تقابل اكبر تكرار في التوزيع .

مثال/ التوزيع الاتي يمثل (40) عائلة فلاحية حسب ملكيتها من عدد اشجار معينة يطلب ايجاد المنوال لهذا التوزيع

الفئات (عدد الاشجار)	60-74	75-89	90-104	105-119	120-134
التكرار (عدد العوائل) Fi	2	6	14	10	8

نلاحظ ان اكبر تكرار في هذا التوزيع هو العدد (14) المقابل للفئة الثالثة (90-104) وعليه فان المنوال هو مركز هذه الفئة اي 97 (اي $M_0 = 97$) .

ب) المنوال لبيانات مبوبة لمتغير مستمر

اذا كان لدينا $x_1, x_2, \dots, x_n$ مراكز فئات توزيع تكراري عدد فئاته (n) وان $f_1, f_2, \dots, f_n$ تمثل التكرارات المقابلة لهذه الفئات عندئذ نستخرج المنوال بهذه الصيغة :

$$M_0 = L_1 + \left(\frac{d_1}{d_1+d_2}\right) * w$$

ان فئة المنوال لهذا التوزيع هي الفئة التي تقابل اكبر تكرار علما ان

L_1 : يمثل الحد الادنى لفئة المنوال

d_1 : يمثل الفرق بين التكرار السابق لفئة المنوال وتكرار فئة المنوال

d_2 : يمثل الفرق بين التكرار اللاحق لفئة المنوال وتكرار فئة المنوال

w : يمثل طول الفئة

مثال/ التوزيع التكراري الاتي يمثل اعمار الموظفين في احدى الشركات المطلوب ايجاد العمر الشائع (المنوال) في هذه الشركة

الفئات (الاعمار)	60–	63–	66–	69–	72–< 75
التكرار (عدد الموظفين) Fi	5	18	42	27	8

نلاحظ ان اكبر تكرار هو (42) وعليه فان فئة المنوال هي الفئة (66–69) وعليه فان $w = 3 , d_2 = 42 - 27 = 15 , d_1 = 42 - 18 = 24 , L_1 = 66$

$$M_0 = L_1 + \left(\frac{d_1}{d_1+d_2}\right) w$$

$$= 66 + \left(\frac{24}{24+15}\right) * 3 = 67.8$$

س/ من جدول التوزيع التكراري الاتي :

الفئات Classes	2–4	5–7	8–10	11–13	14–16	17–19	20–22
التكرار Fi	6	9	12	20	14	11	8

جد (١) الوسيط (٢) المنوال

س/ من الجدول الاتي :

الفئات Classes	100–	120–	140–	160–	180–	200–	220–< 240
التكرار Fi	3	7	14	20	18	12	6

جد (١) الوسيط (٢) المنوال

٢ – ٤ .مقاييس التشتت Measures of Dispersion

يقصد بالتشتت (او الاختلاف) بأنه التباعد او التقارب الموجود بين قيم المشاهدات التابعة لمتغير ما ، ومقاييس التشتت هي مقاييس لمدى تشتت المشاهدات عن وسطها , هذا وكلما كان مقياس التشتت كبيرا دل ذلك على عدم التجانس بين القيم ، ويكون مقياس التشتت صغيرا عندما تكون الاختلافات بين قيم المشاهدات قليلة ذكرنا في الفصل السابق مقاييس النزعة المركزية والتي تعطينا فكرة عن مكان تمركز المشاهدات بينما مقاييس التشتت تعطينا فكرة عن مدى تجانس او تباين هذه القيم حول مركزها اي درجة انتشارها ان لمقاييس التشتت اهميتها في وصف التوزيعات ومقارنتها مع بعضها حيث ان مقاييس التوسط وحدها لا تكفي لهذا الغرض فقد يتساوى الوسط الحسابي لمجموعتين من القيم مثلا بينما يختلف مدى انتشار قيم المجموعة الاولى عن انتشار قيم المجموعة الثانية كما يتضح من مقارنة المجموعتين التاليتين :

المجموعة الاولى : 22 , 21 , 19 , 18 , 23 , 20 , 17

المجموعة الثانية : 13 , 20 , 45 , 5 , 7 , 15 , 35

فالوسط الحسابي لكل من المجموعتين هو (20) ولكن المجموعة الاولى اكثر تجانسا .

ولمقاييس التشتت اهميتها في تطبيق نظرية العينات والاستنتاج الاحصائي واختبار الفرضيات .

وهناك عدة مقاييس للتشتت اهمها :

١ – التباين والانحراف المعياري (القياسي) Variance and Standard Devietion

أ. التباين والانحراف المعياري لبيانات غير مبوبة

اذا كان لدينا عينة حجمها (n) من المشاهدات $x_i = x_1, x_2, \dots, x_n$ فأن التباين (ويرمز له بـ S^2 يكون

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n} x_1^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{2}}{n-1}$$

والانحراف المعياري (S) لعينة ما هو

الجند التربيعي لتباين تلك العينة اي : $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$

$$= \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}}$$

مثال/ البيانات التالية تبين كمية محصول القطن في خمس مزارع احسب الانحراف القياسي (المعياري)

$x_i = 9,8,6,5,7$

أ) الطريقة الاولى

نجد التباين $S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = \frac{\sum_{i=1}^{S}(x_i-\bar{x})^2}{4}$

نحسب الوسط الحسابي $\bar{x}$

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

$= \frac{9+8+6+5+7}{5} = \frac{35}{5} = 7$

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
9	2	4
8	1	1
6	−1	1
5	−2	4
7	0	0
		$\sum_{i=1}^{5}(x_i - \bar{x})^2 = 10$

التباين ← $S^2 = \frac{\sum_{i=1}^{5}(x_i-\bar{x})^2}{4} = 2.5$

الانحراف المعياري ← $S = \sqrt{2.5} = 1.58$

ب) الطريقة الثانية

$$S^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$$

$$\sum_{i=1}^{5} x_i^2 = 9^2 + 8^2 + 6^2 + 5^2 + 7^2$$

$$= 81 + 64 + 36 + 25 + 49 = 225$$

$$\left(\sum_{i=1}^{5} x_i\right)^2 = (9 + 8 + 6 + 5 + 7)^2 = (35)^2 = 1225$$

$$S^2 = \frac{255 - \frac{1225}{5}}{4} = \frac{255 - 245}{4} = \frac{10}{4} = 2.5$$

$S = \sqrt{2.5} = 1.58$ الانحراف المعياري

٢- التباين والانحراف المعياري لبيانات مبوبة

اذا كان لدينا $x_1, x_2, \dots, x_n$ تمثل مراكز فئات في جدول توزيع تكراري وان تكراراتها هي $f_1, f_2, \dots, f_n$ فان التباين هو :

$$S^2 = \frac{\sum_{i=1}^{n} fi(x_i - \bar{x})^2}{\sum_{i=1}^{n} fi - 1}$$

$$= \frac{\sum_{i=1}^{n} fi\, x_i^2 - \frac{(\sum_{i=1}^{n} fi\, x_i)^2}{\sum_{i=1}^{n} fi}}{\sum_{i=1}^{n} fi - 1}$$

والانحراف المعياري هو الجذر التربيعي للتباين

مثال/ من جدول التوزيع التكراري الاتي جد التباين والانحراف المعياري

الفئات	5−9	10−14	15−19	20−24	25−29
التكرار	2	5	10	6	2

أ) الطريقة الاولى

الفئات	fi	x_i	$fi\,x_i$	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$fi(x_i-\bar{x})^2$
5−9	2	7	14	−10.2	104.04	2*104.04=208.08
10−14	5	12	60	−5.2	27.04	5*27.04=135.5
15−19	10	17	170	−0.2	0.04	0.4
20−24	6	22	132	4.8	23.04	138.24
25−29	2	27	54	9.8	96.04	192.08
	$\sum_{i=1}^{5} fi$ =25		$\sum_{i=1}^{5} fix_i$ =430			$\sum_{i=1}^{5} fi(x_i-\bar{x})^2$ =674

$$\bar{x}=\frac{\sum_{i=1}^{5} fix_i}{\sum_{i=1}^{5} fi}=\frac{430}{25}=17.2$$

التباين $S^2=\frac{\sum_{i=1}^{5} fi\,(x_i-\bar{x})^2}{\sum_{i=1}^{5} fi-1}=\frac{674}{24}=28.08$

الانحراف المعياري $S=\sqrt{28.08}=5.3$

ب) الطريقة الثانية

الفئات	fi	x_i	f_ix_i	x_i^2	$f_ix_i^2$
5−9	2	7	14	49	2*49=98
10−14	5	12	60	144	5*144=720
15−19	10	17	170	289	10*289=2890
20−24	6	22	132	484	6*484=2904
25−29	2	27	54	729	2*729=1458
	$\sum_{i=1}^{5} fi$ =25		$\sum_{i=1}^{5} fix_i$ =430		$\sum_{i=1}^{5} fix_i^2$ = 8070

$$S^2 = \frac{\sum_{i=1}^{5} fix_i^2 - \frac{\left(\sum_{i=1}^{5} fix_i\right)^2}{\sum fi}}{\sum_{i=1}^{5} fi - 1} = \frac{8070 - \frac{(430)^2}{25}}{24}$$

$$= \frac{8070-7396}{24} = \frac{674}{24} = 28.08 \quad \text{التباين}$$

$S = 5.3$ الانحراف المعياري

٣- معامل الاختلاف Coefficient of Variation

نفرض ان $\bar{x}$ يمثل الوسط الحسابي لمجموعة قيم وان S يمثل الانحراف المعياري لها عندئذ يعرف معامل الاختلاف على النحو التالي :

$$C.V = \frac{S}{\bar{x}} * 100$$

يعتبر معامل الاختلاف افضل انواع معاملات التشتت لكونه يعتمد على افضل مقياس نزعة مركزية (وهو الوسط الحسابي) وافضل مقياس تشتت (وهو الانحراف المعياري) وعند اجراء مقارنة بين قيم مجموعتين تتم مقارنة معامل اختلاف الاولى مع معامل اختلاف الثانية وعندئذ يقال عن المجموعة بانها اكثر تجانس اذا كان معامل اختلافها اقل من الاخرى .

مثال/ كان متوسط درجات طلبة الصف الاول احصاء في امتحان الرياضيات (69) درجة وبانحراف معياري قدره (19.3) في حين كان متوسط درجاتهم في امتحان الاحصاء (75) درجة بانحراف معياري قدره (25.5) في اي الامتحانين كان مستوى اداء الطلبة اكثر تقاربا .

نجد معامل الاختلاف لكل امتحان

$$C.V(math.) = \frac{19.3}{69} * 100 = 28\%$$

$$C.V(stat.) = \frac{25.2}{75} * 100 = 34\%$$

وحيث ان معامل الاختلاف في امتحان الرياضيات اقل من معامل اختلاف امتحان الاحصاء فان مستوى اداء الطلبة في امتحان الرياضيات كان اكثر تقارب .

س/ اجريت تجربة لدراسة طول النبات (سم) وكمية المحصول (كغم) لـ (150) نبات من الذرة فكانت النتائج كالاتي :

	الطول		كمية المحصول
	200	$\bar{x}$	800
قارن بين تشتت الصفتين	16	S	36

س/ جدول التوزيع التكراري الاتي يمثل الاجور الاسبوعية لعمال احد المصانع

الفئات	16–21	22–27	28–33	34–39	40–45
Fi	15	16	25	20	10

المطلوب ايجاد (١) التباين والانحراف المعياري (٢) معامل الاختلاف

س/ لكل من المجموعات التالية اوجد :

١) التباين
٢) الانحراف القياسي
٣) معامل الاختلاف

a) $x_i = 2,5,9,11,13$
b) $x_i = 4,10,2,8,4,14,10,12,18$
c) $x_i = 16,2,22,8,6,20,24,14$

٤-الدرجة المعيارية (القياسية) Standard Scores

نحتاج في كثير من الاحيان الى مقارنة مفردتين من مجموعتين مختلفتين وفي هذه الحالة يجب تحويل وحدات كل مفردة الى وحدات قياسية حتما تكون المقارنة ذات معنى وذلك باستخدام الوسط الحسابي والانحراف المعياري لكل مجموعة .

تسمى القيمة Z_i درجة قياسية (معيارية) اذا كانت تساوي

$$Z_i = \frac{x_i - \bar{x}}{s}$$

ان الدرجات المعيارية خالية من الوحدات المستخدمة في القياس واذا حولنا جميع قيم المجموعة (اي مجموعة) الى درجات قياسية فان الوسط الحسابي لهذه الدرجات القياسية يساوي صفرا وان تباينها يساوي واحد اي ان $Z_i \sim N(0,1)$ اي ان Z_i يتوزع طبيعيا بوسط حسابي لها = صفر وتباين = 1 .

مثال/ كانت درجات خمسة طلاب في مادتي الرياضيات والاحصاء هي :

تسلسل الطالب	1	2	3	4	5
درجة الرياضيات x	50	52	69	72	81
درجة الاحصاء y	62	77	82	85	86

المطلوب حساب الدرجات المعيارية لكل موضوع

ان الوسط الحسابي $\bar{x}$ والانحراف المعياري s الى x هو

$$\bar{x} = 64.8 \quad , S_x = 11.96$$

وان الوسط الحسابي $\bar{y}$ والانحراف المعياري S_y الى Y هو

$$\bar{Y} = 78.4 \quad , S_y = 8.78$$

$$Z_{x^1} = \frac{x_1 - \bar{x}}{S_x} = \frac{50 - 64.8}{11.96} = -1.24$$

وبنفس الطريقة نحصل على

$$Z_{x_2} = -1.07 \quad Z_{x_3} = 0.35 \quad Z_{x_4} = 0.60 \quad Z_{x_5} = 1.36$$

وكذلك فان

$$Z_{y_1} = \frac{62-78.4}{8.78} = -1.87$$

وبنفس الطريقة نحصل على

$$Z_{y_2} = -0.16 \quad Z_{y_3} = 0.41 \quad Z_{y_4} = 0.75 \quad Z_{y_5} = 0.87$$

لاحظ ان الطالب الخامس درجته في الاحصاء (86) افضل منها في الرياضيات (81) الا انه من حيث الاهمية فان درجته في الرياضيات هي في الحقيقة افضل منها في الاحصاء حيث ان الدرجة المعيارية في الرياضيات كانت اعلى من الدرجة المعيارية في الاحصاء من خلال هذا المثال يمكن التحقق بأن $S_Z^2 = 1 , \bar{Z} = 0$.

س/ اذا علمت ان الوسط الحسابي والانحراف المعياري لدرجات (50) طالب في امتحان مادتي الفيزياء والرياضيات كانت كما يلي :

المادة	الفيزياء	الرياضيات
الوسط الحسابي	75	62
الانحراف المعياري	8	6

فاذا كانت درجة احد الطلبة في الفيزياء هي (78) وفي الرياضيات (68) ففي اي من هذين الامتحانين كان مستوى الطالب افضل .

س/ اذا كانت المشاهدات (القيم) التالية تمثل انتاج معملي لمادة معينة احسب معامل الاختلاف

$x = 17,20,23,18,22,19,21$

$Y = 15,35,13,20,5,7,45$