$$\sum_{I=1}^{n} X_i \qquad i = (1 - n)$$

القواعد الخاصة بالمجموع :

١ـ يرمز لمجموع عدد المشاهدات (قيم المتغيرx) ابتداءً من المشاهدة الأولى وحتى الأخيرة بالرمز

$$\sum_{I=1}^{n} X_i = x_1 + x_2 + x_3 + \ldots\ldots + x_n$$

Ex. If X = 1, 2, 3, 4, 5 find

$$\sum_{I=1}^{n} xi = x_1 + x_2 + x_3 + x_4 + x_5$$

$$= 1 + 2 + 3 + 4 + 5 = 15$$

٢ـ وهناك مجموع جزئي مثل $\sum_{I=3}^{5} x_i$ (أي مجموع المشاهدات الثالثة والرابعة والخامسة) .

٣ـ ويرمز لمجموع مربعات عدد من المشاهدات (قيم المتغير X) بالرمز :

$$\sum_{I=}^{n} x^2_i = x_1^2 + x_2^2 + x_3^2 + \ldots\ldots + x_n^2$$

ففي المثال السابق :

$$\sum_{I=1}^{n} xi^2 = (1)^2 + (2)^2 + (3)^2 + (4)^2 + (5)^2$$

$$= 1 + 4 + 9 + 16 + 25$$

$$= 55$$

٤ـ يرمز لمربع مجموع عدد من المشاهدات (قيم المتغير X) بالرمز :

$$\left(\sum_{I=1}^{n} x_i\right)^2 = (x_1 + x_2 + x_3 + \ldots + x_n)^2$$

ففي المثال السابق فإن :

$$\left(\sum_{I=1}^{5} x_i\right)^2 = (1+2+3+4+5)^2$$
$$= (15)^2$$
$$= 225$$

ملاحظة : نستنتج من 4 , 3 بأن :

$$\sum_{I=1}^{n} x_i^2 \neq \left(\sum_{I}^{5} x_i\right)^2$$

٥- ويرمز لحاصل ضرب مجموعي ثيم المتغيرين x , y بالرمز :

$$\boxed{\sum_{I=1}^{n} xi \sum_{I=1}^{n} yi = (x_1 + x_2 + x_3 + \ldots\ldots + x_n)(y_1 + y_2 + y_3 + \ldots\ldots + y_n)}$$

مثال : إذا كانت قيم كل من x و y هي كالآتي

X = 2, 3, 5, 1

Y = 4, 2, 3, 5

أوجد كل مما يأتي :

1) $\sum_{I=1}^{3} x_i = x_1 + x_2 + x_3$

$= 2 + 3 + 5 = 10$

2) $\sum_{I=1}^{n} X_i \sum_{I=1}^{n} y_i = (x_1 + x_2 + x_3 + \ldots + x_n)(y_1 + y_2 + y_3 + \ldots + y_n)$

$= (2 + 3 + 5 + 1)(4 + 2 + 3 + 5)$

$= (11)(14)$

$= 154$

٦-يرمز لمجموع حاصل ضرب متغيرين x , y بالرمز :

$$\sum_{I=1}^{n} X_i\ y_i = =(X_1 Y_1 + X_2 Y_2 + X_3 Y_3 + \ldots\ldots + X_n Y_n$$

مثال : إذا كانت قيم كل من x و y هي كالآتي

X = 2, 3, 5, 1

Y = 4, 2, 3, 5

أوجد كل مما يأتي :

1) $\sum_{I=1}^{3} x_i = x_1 + x_2 + x_3$

$= 2 + 3 + 5 = 10$

2) $\sum_{I=1}^{n} X_i\ y_i = (X_1 Y_1 + X_2 Y_2 + X_3 Y_3 + \ldots\ldots + X_n Y_n)$

$= 2(4) + 3(2) + 5(3) + 1(5)$

$= 8 + 6 + 15 + 5$

$= 34$

ملاحظة نستنتج من 5 , 6 أن

$$\sum_{I=1}^{n} X_i \sum_{I=1}^{n} y_i \neq \sum_{I=1}^{n} X_i\ y_i$$

٧-يرمز لمجموع مقلوب قيم X بالرمز :

$$\sum_{I=1}^{n} \frac{1}{X_i} = \frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{X_n}$$

مثال : إذا كانت قيم كل من x و y هي كالآتي

X = 2, 3, 1

أوجد كل مما يأتي :

$$1)\quad \sum_{i=1}^{3} \frac{1}{x_i} = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3}$$

$$= \frac{1}{2} + \frac{1}{3} + \frac{1}{1} = \frac{11}{6}$$

٨-يرمز لمقلوب مجموع قيم المتغير X بالرمز :

$$\boxed{\frac{1}{\sum_{I=1}^{n} x_i} = \frac{1}{x_1 + x_2 + \dots + x_n}}$$

مثال : إذا كانت قيم كل من x و y هي كالآتي

X = 2, 3, 1 , 4

أوجد كل مما يأتي :

$$1)\quad \frac{1}{\sum_{I=1}^{n} x_i} = \frac{1}{x_1 + x_2 + x_3 + x_4}$$

$$= \frac{1}{2 + 3 + 1 + 4}$$

$$= \frac{1}{10}$$

بعض القواعد الأساسية في عملية الجمع :

١- إذا كان (c) أي عدد ثابت فإن :

$$\boxed{\sum_{I=1}^{n} c = nc}$$

البرهان :

$$\sum_{I=1}^{n} c = c_1 + c_2 + c_3 + \dots + c_n$$

$$= nc$$

مثال : أوجد كل مما يأتي :

$$\sum_{I=1}^{4} 7 = 7 + 7 + 7 + 7$$
$$= 4\,(7) = 28$$

٢- إذا كان (c)أي عدد ثابت فإن :

$$\boxed{\sum_{I=1}^{n} cx_i = c \sum_{I=1}^{n} x_i}$$

البرهان :

$$\sum_{I=1}^{n} cx_i = cx_1 + cx_2 + cx_3 + \ldots + cx_n$$
$$= c\,(x_1 + x_2 + x_3 + \ldots + x_n)$$
$$= c \sum_{I=1}^{n} x_i$$

مثال : إذا كانت $x = 1, 2, 3, 4, 5$ أوجد ناتج ما يأتي :

$$1)\ \sum_{I=1}^{n} 4\,x_i = 4\,(1 + 2 + 3 + 4 + 5)$$
$$= 4\,(15) = 30$$

$$2)\ \sum_{I=1}^{n} 2\,x_i^2 = 2\,(x_1^2 + x_2^2 + \ldots + x_5^2 +$$
$$= 2\,(1 + 4 + 9 + 16 + 25)$$
$$= 2\,(55)$$
$$= 110$$

ملاحظة :

*** يرمز لمجموع حاصل قسمة المتغيرين x , y بالرمز :**

$$\boxed{\sum_{I=1}^{n} \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \ldots\ldots + \frac{x_n}{y_n}}$$

* ويرمز لحاصل قسمة مجموعي المتغيرين x , y بالرمز

$$\frac{\sum_{I=1}^{n} X_i}{\sum_{I=1}^{n} y_i} = \frac{x_1 + x_2 + \ldots + x_n}{y_1 + y_2 + \ldots + y_n}$$

H. W. :

If X = 3, 6, 9, 12, 15 Y = 2, 4, 6, 8, 10 Find the following :

1) $\sum_{I=1}^{3} X_i$

2) $\sum_{I=1}^{n} y_i^2$

3) $\sum_{I=1}^{3} X_i \; y_i$

4) $\left(\sum_{I=1}^{3} y_i \right)^2$

5) $\sum_{I=1}^{n} \frac{1}{X_i \; y_i}$

6) $\sum_{I=1}^{n} 2\, x_i\, y_i^2$

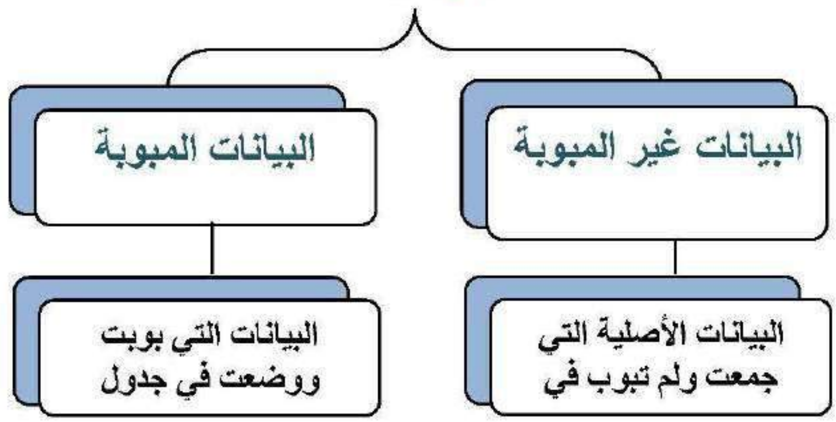
الفصــــل الثـــــاني

طـــــرق عـــــرض البيانـــات

البيانات Data مجموعة القيم التي يتم جمعها من مفردات المجتمع أو العينة لخاصية معينة (متغير).

وإن البيانات الخام لا معنى لها ولا يمكن قياسها إلا إذا خضعت للمعالجات الاحصائية ، فالبيانات قبل عرضها **تسمى بيانات غير مبوبة** (Ungrouped) وبعد عرضها تسمى **بيانات مبوبة** (Grouped) .

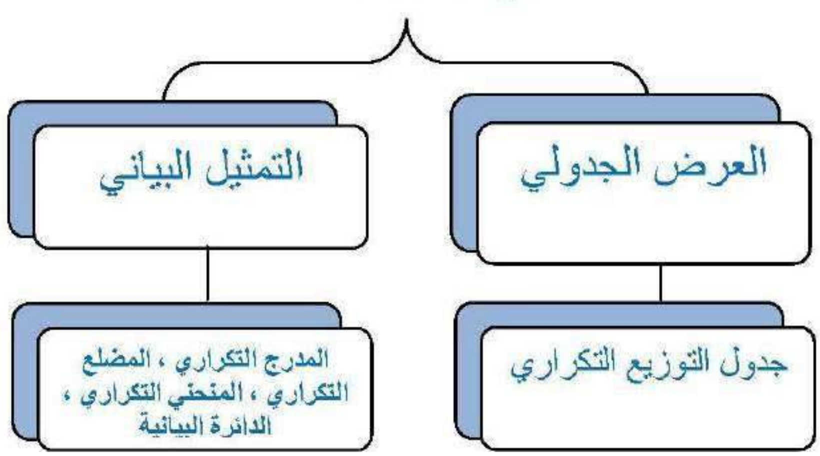
أنواع البيانات



١- **البيانات غير المبوبة (Ungrouped Date)** وهي البيانات الأصلية التي جمعت ولم تبوب في جدول .

٢- **البيانات المبوبة (Grouped Date)** وهي البيانات التي بوبت ووضعت في جدول توزيع تكراري .

طرق عرض البيانات



تقسم طرق عرض البيانات الى قسمين رئيسين هما :

أولاً :العرض الجدولي (Tabular Presentation)

ويطلق عليه أيضاً **الجدول الاحصائي** : وهو عبارة عن جدول منظم ، يسمى (جدولاً بسيطاً) إذا تألف من بعد واحد ، و(جدولاً مركباً) إذا تألف من بعدين .

الجدول البسيط : وهو الجدول الذي توزع فيه البيانات بحسب صفة واحدة ، ويتألف عادة من عمودين ، **الأول** يتضمن تقسيمات الصفة الى فئات أو مجموعات ، و**الثاني** يبين عدد المفردات التابعة لكل فئة أو مجموعة .

مثال : اعرض البيانات الآتية في جدول تكراري :

1, 2, 2, 2, 3, 3, 4, 5, 5, 5, 5, 5, 6, 7, 8, 8, 8, 9, 9, 9

الحل :

X	1	2	3	4	5	6	7	8	9	$\sum x$
F_i	1	3	2	1	4	1	1	2	3	18

الجدول المركب : وهو الجدول الذي توزع فيه البيانات حسب صفتين أو ظاهرتين أو أكثر في نفس الوقت.

مثال : الجدول أدناه يبين توزيع مجموعة من الطلبة بحسب الوزن والطول :

المجموع	71 – 80	61 – 70	51 60	الوزن (كغم) / الطول
30	4	6	20	121 – 140
52	10	40	2	141 – 160
18١	10	6	2	161 – 180
100	24	52	24	المجموع

ويتمثل العرض الجدولي بجدول التوزيع التكراري .

جدول التوزيع التكراري (Frequency Distribution Table)

وهو جدول بسيط يتكون من عمودين **الأول يسمى عمود الفئات (Classes)** ، وتقسم فيه قيم المتغير الى أقسام ومجموعات ، ويرمز له بالرمز (C) ، **والثاني يسمى عمود التكرارات (Frequency)** ويتضمن عدد المرات التي تكررت فيها الفئة ، ويرمز له بالرمز (f_i) .

بعض المفاهيم ذات الصلة بجدول التوزيع التكراري :

١- **التوزيع التكراري :** هو عبارة عن تلخيص وترتيب لبيانات المتغير العشوائي التي سبق وأن جمعت وصنفت مقسمة إلى عدد من المجاميع يسمى كل منها بـ (الفئة Class) . وهذه الفئات قد تكون مرتبة تصاعدياً أو تنازلياً حسب طبيعة البيانات . ويسمى توزيع عدد قيم X حسب الفئات بـ (التوزيع التكراري) .

٢- **المدى الكلي (Total Range):** يعرف المدى الكلي بأنه الفرق ما بين أكبر قيمة وأصغر قيمة في المجموعة . ويرمز له بالرمز (T.R) .

$$\boxed{T.R. = x_L - x_s}$$

٣- **الفئات class :** هي عدد المجاميع التي يتألف منها التوزيع التكراري ، وكل مجموعة تتحدد بحدين ، الأول يسمى **الحد الأدنى للفئة (Lower Class Limit)** ويرمز له بالرمز L.L ، والثاني يسمى **الحد الأعلى للفئة (Upper Class Limit) ويرمز له بالرمز u.L .**

مثل : الفئة (16 – 10) حدها الأدنى 10 وحدها الأعلى 16

٤- **الحدود الحقيقية للفئات :**

لكل فئة حدان حقيقيان حد أدنى حقيقي وحد أعلى حقيقي

الحد الأدنى الحقيقي للفئة = الحد الأدنى – 0.5

مثل : ذا كان الحد الأدنى لفئة هو (10) فإن الحد الأدنى الحقيقي لها هو ؟

الحد الأدنى الحقيقي = الحد الأدنى – 0.5 = 9.5

الحد الأعلى الحقيقي للفئة = الحد الأعلى + 0.5

مثل : ذا كان الحد الأعلى لفئة هو (16) فإن الحد الأعلى الحقيقي لها هو ؟

الحد الأعلى الحقيقي = الحد الأعلى + 0.5 = 16.5

٥- **عدد فئات التوزيع m** . ويمثل عدد المجاميع التي يتألف منها التوزيع التكراري . فإذا رمزنا لعدد الفئات بالرمز m فإن :

$$m = 2.5 \sqrt[4]{n}$$

حيث n تمثل عدد المفردات (قيم المتغير x) .

ويجب أن لا يقل عدد الفئات عن خمسة فئات و لا يزيد عن خمسة عشر فئة .

٦- **تكرار الفئة (Classes frequency) :** هو عدد القيم التي تقع في مدى تلك الفئة ، ويرمز له بالرمز (f_i)

إذ إن f_1 يمثل تكرار الفئة الأولى ، f_2 يمثل تكرار الفئة الثانية ، ... و هكذا .

وإن مجموع التكرارات يجب أن يكون دائماً مساوياً للعدد الكلي لقيم المفردات .

مثلاً : تكرار الفئة الثالثة = 5، فهذا يعني أن هناك 5 قيم من قيم المتغير واقعة في المدى الفئة . .

٧- **مركز الفئة (Center of Classes) :** ويمثل منتصف المدى بين حدي للفئة ، ويرمز له بالرمز x . أي إن :

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى} + \text{الحد الأعلى}}{2}$$

$$X = \frac{L.L + U.L}{2}$$

مثال : إذا كانت لدينا الفئة (16 – 10) أوجد مركز الفئة :

الحل :

$$X = \frac{L.L + U.L}{2}$$

$$= \frac{10 + 16}{2} = 13$$

٨- **طول الفئة (Length of Classes) :** ويسمى أحياناً بالمدى الفئوي ، ويمثل مقدار المدى بين حدي للفئة ، ويرمز له بالرمز (L) .

وإن طول الفئة يتناسب عكسياً مع عدد الفئات ، فكلما زاد طول الفئة قل عدد الفئات والعكس صحيح .

ويجب أن تكون أطوال الفئات متساوية لتسهيل العمليات الحسابية ،ويجب أن تكون أعداد صحيحة موجبة دائماً .

طول الفئة = الحد الأعلى – الحد الأدنى + ١

$$\boxed{L = U.L - L.L + 1}$$

مثال : إذا كان لدينا الفئة (16 – 10) ، أوجد طول الفئة :

الحل :طول الفئة = الحد الأعلى – الحد الأدنى + ١

$$L = 16 - 10 + 1$$
$$= 7$$

كما يمكن استخراجه من :

أ) إذا أمكن استخراج المدى الكلي وعدد الفئات فإن طول الفئة يستخرج من القانون :

$$\boxed{L = \frac{T.R.}{m}}$$

حيث T.R : هو المدى الكلي

m : عدد الفئات

ب) أو من خلال : الفرق بين الحدين الأعلى أو (الأدنى) لفئتين متتاليتين إذا كانت هناك فئتين متتاليتين معلومتين :

$$\boxed{L = (L.L.)_2 - (L.L.)_1}$$

$$\boxed{L = (u.L.)_2 - (u.L.)_1}$$

ج) أو من (الفرق بين مركزي فئتين متتاليتين) :

$$\boxed{L = X_2 - X_1}$$

حيث : X_2 يمثل مركز الفئة الثانية و X_1 يمثل مركز الفئة الأولى.

خطوات إنشاء جدول التوزيع التكراري (تبويب البيانات) :

لنأخذ المثال الافتراضي الآتي :

لتكن $X_1, X_2, X_3, \ldots, X_n$ بيانات المتغير العشوائي X من عينة عشوائية ما عدد مفرداتها n مفردة . ونرغب في تلخيص هذه البيانات في توزيع تكراري عدد فئاته هو m.

ولنفرض أن X_s هي أصغر قيمة ، X_L هي أكبر قيمة في مجموعة بيانات المتغير x ، عندئذ نتبع الخطوات الآتية في إنشاء جدول التوزيع التكراري :

١- تحديد أصغر قيمة وأعلى قيمة من قيم المتغير العشوائي X .

٢- استخراج **المدى الكلي T.R** .

٣- استخراج عدد فئات التوزيع m .

٤- استخراج طول الفئة L .

٥- تحديد الحد الأدنى والأعلى للفئات ، وكتابة حدود الفئات .

٦- استخراج عدد التكرارات لكل فئة f_i .

<u>ملاحظة مهمة</u>: يفضل أن يمتاز التوزيع التكراري بما يأتي :

١- أن تكون الفئات متساوية في الطول .

٢- أن يكون التوزيع التكراري توزيع مغلق .

٣- أن تبدأ فئات التوزيع وتنتهي (حدود الفئات) بأعداد صحيحة ، وذلك لتسهيل اجراء العمليات الحسابية خصوصاً إذا اجريت بدون استخدام الحاسبات .

٤- أن لا يقل عدد فئات التوزيع عن 5 فئات ولا يزيد عن 15 فئة .

٥- أن تكون حدود الفئات محددة بشكل واضح بحيث أن كل قيمة من قيم المتغير تبوب ضمن فئة واحدة فقط من فئات التوزيع .

<u>مثال :</u> البيانات الآتية تمثل درجات 80 طالباً في مادة الاحصاء . المطلوب انشاء جدول تكراري لهذه النباتات <u>أو</u> (اعرض البيانات الآتية في جدول توزيع تكراري) .

80	84	80	87	97	81	74	48	79	80
56	92	70	71	78	82	93	91	70	90
83	93	65	51	75	68	72	73	74	81
93	68	86	43	74	73	83	90	35	86
61	80	91	75	67	72	90	71	76	92
80	95	97	70	74	81	88	91	97	72
89	67	60	82	83	63	60	77	71	59
65	75	79	88	66	70	88	76	63	63

الحل :

1) $x_L = 97 \qquad x_s = 35$

2) $T.R. = x_L - x_s$

$= 97 - 35 = 62$

3) $m = 2.5 \sqrt[4]{n}$

$= 2.5 \sqrt[4]{80}$

$= 7.4 \qquad = 7$

4) $L = \dfrac{T.R.}{m} = \dfrac{62}{7} = 8.8 \simeq 9$

	C	Fi
1	35 -43	2
2	44 – 52	2
3	53 – 61	5
4	62 – 70	14
5	71 – 79	20
6	80 – 88	22
7	89 – 97	15
		80

*ملاحظة:- اذا كان هناك فئة موجودة في الجدول يجب معرفة طول الفئة ثم أكمال الجدول.
مثال : نظم البيانات الاتية:- (24,5,15,16,10,11,19,15,21,15) في جدول تكرارات فئته الاولى(5-9)

C	fi
5 -9	1
10-14	2
15-19	5
20-24	2

مثال : إذا كان عدد مفردات ظاهرة ما هو (150) مفردة وإن أقل قيمة فيها هي (10) وأعلى قيمة هي (90) . أكمل الجدول الآتي :

	C	X	الحدود الحقيقية للفئات
1			
2			
3			
4			
5			
6			
7			
8			
9			

الحل :

1) $x_L = 90 \qquad x_s = 10$

2)T. R. = $x_L - x_s$

$= 90 - 10 += 80$

3) $m = 2.5 \sqrt[4]{n}$

$= 2.5 \sqrt[4]{150}$

$= 8.75 \quad \simeq 9$

4) $L = \frac{T.R.}{m} = \frac{80}{9} \quad \simeq 9$

5) $L = u.L. - L.L. + 1$

$9 = u.L. - 10 + 1$

$u.L. = 18$ ....

6) $X_1 = \dfrac{u.L. + L.L.}{2} = \dfrac{18 + 10}{2} = 14$

$X_2 = X_1 + L = 14 + 9 = 23$

.... وهكذا .

الحد الأدنى الحقيقي للفئة الأولى = الحد الأدنى للفئة – 0.5

$= 10 - 0.5 = 9.5$

الحد الأعلى الحقيقي للفئة الأولى = الحد الأعلى للفئة + 0.5

$= 18 + 0.5 = 18.5$

وهكذا .

	C	X	الحدود الحقيقية للفئات
1	10 – 18	14	9.5 – 18.5
2	19 – 27	23	18.5 – 27.5
3	28 – 36	32	27.5 – 36.5
4	37 – 45	41	36.5 – 45.5
5	46 – 54	50	45.5 – 54.5
6	55 – 63	59	54.5 – 63.5
7	64 – 72	68	63.5 – 72.5
8	73 – 81	77	72.5 – 81.5
9	82 – 90	86	81.5 – 90.5

تمرين1 : أكمل الجدول الآتي :

	C	X	الحدود الحقيقية للفئات
1	10 -	12	
2			
3			

4			
5			
6			

تمرين2 : أكمل الجدول الآتي :

	C	X	الحدود الحقيقية للفئات
1			
2			
3			
4			
5			
6	90 – 99		

تمرين 3 : البيانات الآتية تمثل درجات 30 طالب في مادة الاحصاء :

١- أنشئ جدول التوزيع التكراري لدرجات الطلبة .

٢- استخرج الحدود الحقيقية للفئات ، مراكز الفئات .

13, 15, 25, 67, 88, 33, 90, 98, 78, 75, 44, 54, 31, 40. 89, 55, 76, 43, 38, 51,
34, 72, 69, 37, 33, 42, 88, 78, 79,29

التوزيع التكراري المتجمع (Cumulativ frequency distribution) :

وهو التوزيع الذي يبين كمية التكرار المتجمع عن قيمة معينة من قيم المتغير العشوائي.

وهو على نوعين :

أ- التوزيع **التكراري المتجمع الصاعد** ($F_i\uparrow$) :

وهو التوزيع الذي يبين تراكم التكرارات ابتداءً من الفئة الأولى في التوزيع وانتهاءً بالفئة الأخيرة منه . ويتم حساب التكرارات المتجمعة على أساس الحدود العليا للفئات .

ب- التوزيع **التكراري المتجمع النازل** ($F_i\downarrow$) :

وهو التوزيع الذي يبين تناقص التكرارات ابتداءً من الفئة الأولى في التوزيع وانتهاءً بالفئة الأخيرة منه . ويتم حساب التكرارات المتجمعة على أساس الحدود الدنيا للفئات .

4			
5			
6			

تمرين2 : أكمل الجدول الآتي :

	C	X	الحدود الحقيقية للفئات
1			
2			
3			
4			
5			
6	90 – 99		

تمرين 3 : البيانات الآتية تمثل درجات 30 طالب في مادة الاحصاء :

١- أنشئ جدول التوزيع التكراري لدرجات الطلبة .

٢- استخرج الحدود الحقيقية للفئات ، مراكز الفئات .

13, 15, 25, 67, 88, 33, 90, 98, 78, 75, 44, 54, 31, 40. 89, 55, 76, 43, 38, 51, 34, 72, 69, 37, 33, 42, 88, 78, 79,29

التوزيع التكراري المتجمع (Cumulativ frequency distribution) :

وهو التوزيع الذي يبين كمية التكرار المتجمع عن قيمة معينة من قيم المتغير العشوائي.

وهو على نوعين :

أ- التوزيع **التكراري المتجمع الصاعد ($F_i\uparrow$)** :

وهو التوزيع الذي يبين تراكم التكرارات ابتداءً من الفئة الأولى في التوزيع وانتهاءً بالفئة الأخيرة منه . ويتم حساب التكرارات المتجمعة على أساس الحدود العليا للفئات .

ب- التوزيع **التكراري المتجمع النازل ($F_i\downarrow$)** :

وهو التوزيع الذي يبين تناقص التكرارات ابتداءً من الفئة الأولى في التوزيع وانتهاءً بالفئة الأخيرة منه . ويتم حساب التكرارات المتجمعة على أساس الحدود الدنيا للفئات .

مثال :الآتي جدول توزيع تكراري يمثل توزيع 60 عائلة فلاحية حسب ملكيتها من عدد أشجار البرتقال . المطلوب تكوين التوزيع التكراري المتجمع الصاعد والنازل :

	C	f_i	$\uparrow F_i$	$\downarrow F_i$
1	60 – 74	4	4	60
2	75 – 89	5	9	56
3	90 – 104	10	19	51
4	105 – 119	12	31	41
5	120 – 134	16	47	29
6	135 – 149	7	54	13
7	150 – 164	6	60	6
		60		

ثانياً : التمثيل البياني لجدول التوزيع التكراري

١- المدرج التكراري :

هو عبارة عن مستطيلات رأسية تمتد قواعدها على المحور الأفقي لتمثل أطوال الفئات بينما يمثل ارتفاعها تكرارات الفئات .

خطوات رسم المدرج التكراري :

١- رسم المحورين الأفقي والعمودي .

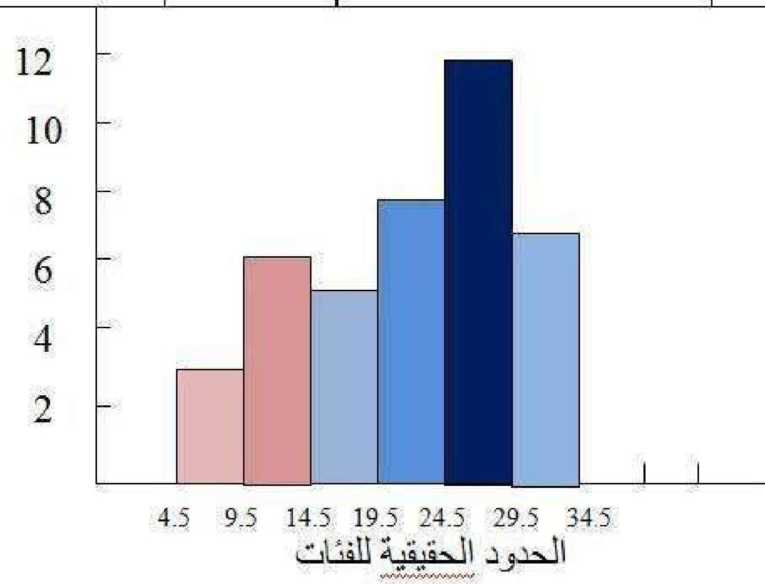
٢- تدريج المحور الأفقي الى أقسام متساوية بحيث يشمل جميع الحدود الحقيقية للفئات ، ويقسم المحور العمودي الى اقسام متساوية بحيث يشمل على اكثر التكرارات .

٣- يرسم على كل فئة مستطيلاً رأسياً تمثل قاعدته طول تلك الفئة وارتفاعه يمثل تكرارات تلك الفئة .

مثال :ارسم المدرج التكراري لجدول التوزيع التكراري الآتي :

	C	F_i	الحدود الحقيقية للفئات	X
1	5 – 9	3	4.5 – 9.5	7
2	10 – 14	6	9.5 – 14.5	12

3	15 – 19	5	14.5 – 19.5	17
4	20 – 24	8	19.5 – 24.5	22
5	25 – 29	12	24.5 – 29.5	27
6	30 – 34	7	29.5 – 34.5	32
$\sum f_i$		41		



٢- المضلع التكراري :

هو عبارة عن خطوط مستقيمة متكسرة تصل بين نقاط كل منها واقعة فوق مركز فئة على ارتفاع يمثل تكرار تلك الفئة .

ملاحظة : عادة يغلق المضلع بأن نصل بداية المضلع بالمحور الأفقي بمركز فئة (خيالية) واقعة الى يسار أول فئة تكرارها صفراً ، ونصل نهاية المضلع بالمحور الأفقي بمركز فئة(خيالية) واقعة الى يمين اخر فئة تكرارها صفراً أيضاً .

خطوات رسم المضلع التكراري :

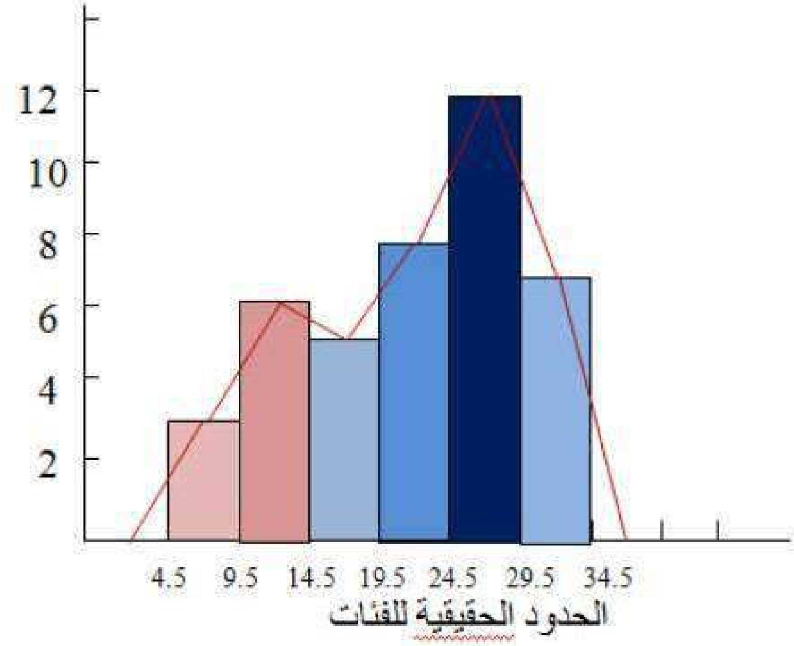
١- رسم المحورين الأفقي والعمودي .

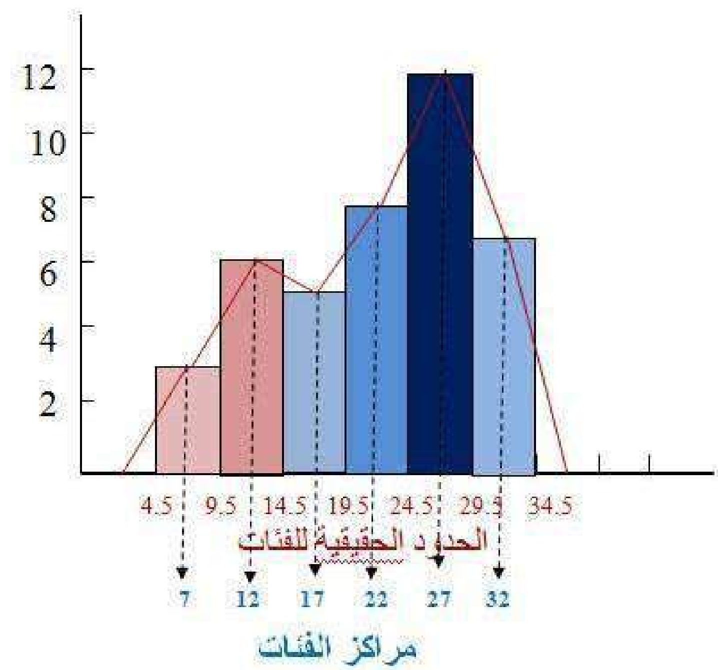
٢- تدريج المحور الأفقي الى أقسام متساوية بحيث يشمل جميع مراكز الفئات ، ويقسم المحور العمودي الى اقسام متساوية بحيث يشمل على اكثر التكرارات .

٣- وضع نقطة أمام مركز كل فئة ارتفاعها يعادل تكرار تلك الفئة .

٤- توصيل تلك النقاط بخطوط مستقيمة .

مثال : ارسم المضلع التكراري للمثال السابق :





٣- المنحني التكراري :

وهو عبارة عن منحني يمر بمعظم النقاط الواقعة على مراكز الفئات والتي ارتفاعها يمثل تكرارات تلك الفئات

ملاحظة : يفضل غلق المنحني التكراري بأن نصل بدايته بالحد الأدنى للفئة الأولى ونهايته بالحد الأعلى للفئة الأخيرة . وتكون مساحة المنحني (مكافئة) وليست مساوية للمضلع التكراري .

٤ـ الدائرة البيانية :

تعد هذه الطريقة أفضل الطرق لتمثيل البيانات ذات الصفة المشتركة ونستطيع بواسطتها أن نقارن الأجزاء بعضها ببعض ثم الجزء (القطاع الدائري) بالكل (الدائرة) .

خطوات رسم الدائرة البيانية :

١- استخراج زاوية القطع = ($\frac{\text{الجزء}}{\text{الكل}}$ × 360)

٢- -نرسم دائرة معينة ونرسم عليها نصف القطر.

٣- نرسم الزاوية المركزية التي ضلعها الابتدائي نصف القطر والممثلة بالقطاع.

مثال/مجموعة من الفاكهة وزعت على طلاب القسم الداخلي وكانت كالآتي :

نوع الفاكهة	تفاح	موز	برتقال	رمان	المجموع
العدد	180	540	90	270	1080

المطلوب تمثيل هذه البيانات بالقطاع الدائري.

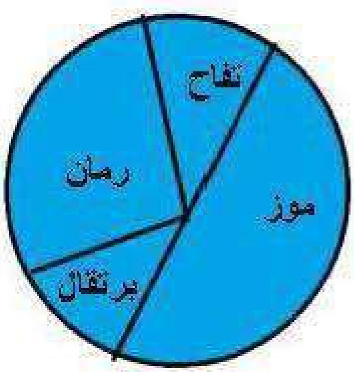
الحل :

زاوية قطاع التفاح $=\frac{180}{1080}\times 360=60^\circ$

زاوية قطاع الموز $=\frac{540}{1080}\times 360=180^\circ$

زاوية قطاع البرتقال $=\frac{90}{1080}\times 360=30^\circ$

زاوية قطاع الرمان $=\frac{270}{1080}\times 360=90^\circ$



تمارين :

١- الآتي توزيع تكراري لدرجات (40) طالباً في مادة الإحصاء الترب

93	82	78	71	70	92	56	81
74	73	72	68	75	51	65	93
35	90	83	73	74	43	86	68
92	76	71	90	72	67	75	91
80	61	72	97	91	88	81	74

اعرض البيانات أعلاه مستخدماً كل من العرض الجدولي والعرض البياني :

٢- أكمل جدول التوزيع التكراري الآتي ، ثم اعرض البيانات في الجدول باستخدام التمثيل البياني :

	C	f_i	X
1	1 -	2	4
2		4	
3		7	
4		12	
5		10	
6		15	
7		8	
8		6	
9		4	

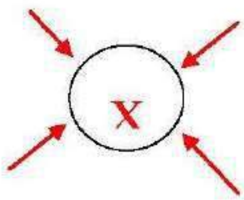
الفصل الثالث

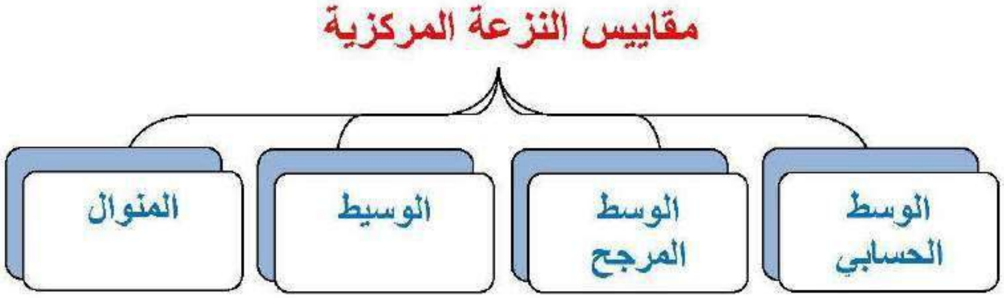
المقاييس الاحصائية

مقاييس النزعة المركزيةMeasures of Central Tendency

إن مصطلح النزعة المركزية يعني التمركز والتكثف نحو رقم معين .

تعرف مقاييس النزعة المركزية بأنها تلك المقاييس التي تبحث عن قيمة تتمركز حولها أغلبية البيانات ، وإن هذه القيمة المتوسطة هي رقم واحد يعبر عن جميع بيانات تلك المجموعة .





١ – الوسط الحسابي (المتوسط) : (The Arithmetic Mean)

وهو القيمة الناتجة عن قسمة مجموع القيم على عددها . ويرمز له بالرمز $\overline{X}$.

طرق إحتساب الوسط الحسابي :

١- في حالة البيانات غير المبوبة :

إذا كان لدينا x_n , …… , x_3 , x_2 , x_1 بيانات المتغير العشوائي x عددها n من المفردات ،فإن الوسط الحسابي لها:

$$\overline{X} = \frac{\sum x_i}{n}$$

حيث : x : القيم للمتغير

n : عدد القيم .

مثال : البيانات الآتية تمثل درجات 15 طالب في مادة الإحصاء (80, 81, 55, 32, 43, 95, 90, 36, 49, 58, 69, 75, 65, 72, 60) أوجد الوسط الحسابي لدرجات الطلاب.

الحل :

$$\bar{X} = \frac{\sum x_i}{n}$$

$$= \frac{60 + 72 + 65 + 75 + 69 + 58 + 49 + 36 + 80 + 81 + 55 + 32 + 43 + 95 + 90}{15}$$

$$= \frac{987}{15} = \mathbf{65.8}$$

أي أن معدل درجات الطلبة هو 65.8 .

٢- في حالة البيانات المبوبة :

إذا كانت $x_1 , x_2 , x_3 , \ldots\ldots , x_n$ تمثل مراكز فئات في توزيع تكراري عدد فئاته m فئة ، و $f_1, f_2, \ldots , f_n$ هي تكرارات تلك الفئات فإن الوسط الحسابي يستخرج من الصيغة الآتية :

$$\boxed{\bar{X} = \frac{\sum (X)(f_i)}{\sum f_i}}$$

حيث : X تمثل مراكز الفئات

f_i تمثل التكرارات

مثال : أوجد الوسط الحسابي للبيانات الآتية :

	C	f_i	X	Xf_i
1	2 – 6	2	4	8
2	7 – 11	8	9	72
3	12 – 16	6	14	84
4	17 – 21	1	19	19
5	22 – 26	5	24	110
$\sum$		22		203

الحل :

$$\bar{X} = \frac{\sum (X f_i)}{\sum f_i} = \frac{203}{22} = \mathbf{9.22}$$

خصائص الوسط الحسابي :

١- مجموع انحرافات قيم المتغير العشوائي **X** عن وسطها الحسابي الذي أحتسب منها يساوي صفر .

للبيانات غير المبوبة : $\sum(x_i - \bar{x}) = 0$

البرهان :

$$\sum(x_i - \bar{x}) = \sum x_i - \sum \bar{x}$$
$$= \sum x_i - n.\bar{x}$$
$$= \sum x_i - n\frac{\sum x_i}{n}$$
$$= \sum x_i - \sum x_i = 0$$

للبيانات المبوبة : $\sum f_i(x_i - \bar{x}) = 0$

البرهان :

$$\sum f_i(x_i - \bar{x}) = \sum f_i x_i - \bar{x}\sum f_i$$
$$= \sum f_i x_i - \frac{\sum f_i x_i}{\sum f_i}\sum f_i$$
$$= \sum f_i x_i - \sum f_i x_i = 0$$

٢- عند إضافة عدد ثابت **k** إلى كل قيمة من قيم المفردات (المشاهدات) فإن قيمة الوسط الحسابي تزداد بمقدار العدد الثابت . أي إن :

$$X_i = y_i + k \longrightarrow \bar{x} = \bar{y} + k$$

البرهان :

$$X_i = y_i + k$$
$$\sum x_i = \sum y_i + n.k$$

نقسم طرفي المعادلة على n

$$\frac{\sum X_i}{n} = \frac{\sum y_i}{n} + \frac{n.k}{n}$$
$$\bar{x} = \bar{y} + k$$

٣- عند ضرب عدد ثابت k إلى كل قيمة من قيم المفردات (المشاهدات) فإن قيمة الوسط الحسابي تكون مساوية لحاصل ضرب العدد الثابت في الوسط الحسابي .

$$X_i = y_i . k \longrightarrow \bar{x} = \bar{y} . k$$

البرهان :

$$X_i = y_i . k$$

$$\sum x_i = \sum y_i . n.k$$

$$\frac{\sum X_i}{n} = \frac{\sum y_i}{n} . \frac{n.k}{n}$$ نقسم طرفي المعادلة علىn

$$\bar{x} = \bar{y} . k$$

٤- اذا ضربت كل مشاهدة من المشاهدات بقدار ثابت a واضيف لها مقدار ثابتb فان الوسط الحسابي الجديد = الوسط الحسابي الاصلي مضروب في a ومضاف له b.

$$x_i = ay_i + b \longrightarrow \bar{x} = a\bar{y} + \bar{b}$$

البرهان :

$$\sum x_i = \sum (ay_i + b)$$

$$\sum X_i = a\sum y_i + \sum b$$

$$\sum x_i = a\sum y_i + n.b$$

$$\frac{\sum X_i}{n} = a\frac{\sum X_i}{n} + \frac{n.k}{n}$$ نقسم طرفي المعادلة علىn

$$\bar{x} = a\bar{y} + b$$

٥- الوسط الحسابي لمجموع قيم متغيرين = مجموع الوسطين الحسابيين للمتغيرين .

$$z_i = x_i + y_i \longrightarrow \bar{z} = \bar{x} + \bar{y}$$

$$z_i = x_i + y_i$$

$$\sum z_i = \sum (x_i + y_i)$$

$$\sum z_i = \sum x_i + \sum y_i$$

نقسم طرفي المعادلة على n

$$\frac{\sum z_i}{n} = \frac{\sum X_i}{n} + \frac{\sum y_i}{n}$$

$$\bar{z} = \bar{x} + \bar{y}$$

٦- **لا يمكن ايجاده في التوزيعات التكرارية والجداول المفتوحة (غير المنتهية) .**

٧- **الوسط الحسابي يتأثر بالقيم المتطرفة .**

٨- **مجموع مربعات انحرافات قيم المتغير العشوائي X عن وسطها الحسابي الذي أحتسب منها يكون أقل ما يمكن .**

٩- **يتأثر الوسط الحسابي بالقيم المتطرفة (صغيرة أو كبيرة) التي تقع على جانبي التوزيع .**

H. W.

١) الآتي توزيع تكراري لدرجات 92 طالب في مادة الاحصاء التربوي ، المطلوب ايجاد متوسط درجات الطلبة في مادة الاحصاء :

	C	F_i
1	21 – 30	4
2	31 – 40	7
3	41 – 50	10
4	51 – 60	15
5	61 – 70	18
6	71 – 80	12
7	81 – 90	6
8	91 – 100	18

٢) **الآتي توزيع تكراري لدرجات (25) طالب في مادة الاحصاء التربوي ، استخرج معدل درجات الطلبة مستخدماً**

أ) حالة البيانات غير المبوبة ب) البيانات المبوبة :

80, 84, 48, 77, 65, 71, 20, 55, 67, 65, 88, 35, 24, 31, 44, 90, 66, 45, 78, 79, 74, 70, 44, 56, 50

٣) أ) أكمل جدول التوزيع التكراري أدناه

ب) ارسم المدرج والمضلع التكراري

ج) استخرج الوسط الحسابي

	C	F_i	X
1	10 -	4	14.5
2		7	
3		10	
4		25	
5		18	
6		12	
7		6	
8		18	

٤) جد قيمة z التي تجعل قيمة الوسط الحسابي في المجموعة z, 37, 31, 28, 25, 20 مساوٍ إلى 30 .

٥) إذا علمت أن الوسط الحسابي للتوزيع التكراري الآتي هو 39 ، جد قيمة f.

	C	f_i
1	10 – 19	2
2		4
3		6
4		8
5	50 - 59	F

٦) صنف صف إل عدد من الطلاب والطالبات ، بحيث كان متوسط عمر الذكور فيه 20 سنة ، ومتوسط عمر الاناث 19.5 سنة ، أوجد متوسط عمر الذكور والاناث معاً .

٢- الوسط الحسابي المرجح (s الموزون) :(Weighted Mean)

يستخدم الوسط الحسابي المرجح ($\overline{X}_w$) عندما تكون بعض المفردات أكثر أهمية من الأحرى مما يستوجب أخذ هذه الأهمية بنظر الاعتبار عند استخراج الوسط الحسابي .

ومن الأمثلة عليه (استخراج معدل الطالب الجامعي في فصل دراسي)، إذ أن الأمر يتطلب احتساب معدل درجات الطالب موزونة بعدد الساعات الاسبوعية لكل مادة من المواد التي تدخل في حساب المعدل .

فإذا كانت $x_1 , x_2 , x_3 , \ldots\ldots , x_n$ بيانات المتغير العشوائي x قوامها (عددها) n من المفردات ، وكانت $w_1, w_2, \ldots , w_n$ تمثل أوزان هذه البيانات فإن الوسط الحسابي المرجح يمكن استخراجه كما يأتي :

$$\overline{X}_w = \frac{\sum w_i x_i}{\sum w_i}$$

حيث : $\overline{X}_w$ هو الوسط الحسابي المرجح

$\sum w_i x_i$ هو مجموع حاصل ضرب كل مفردة بوزنها

$\sum w_i$ هو مجموع أوزان المفردات

مثال : طالب جامعي في السنة الثالثة كانت درجاته في نهاية الفصل كالاتي:

اسم المادة	قياس وتقويم	علم النفس	اللغة العربية	E	ارشاد تربوي	صحة
الدرجة	72	80	90	65	65	70
عدد الساعات	3	3	2	2	2	2

فما هو معدله في الفصل الدراسي ؟

الحل :

$$\overline{X}_w = \frac{\sum w_i x_i}{\sum w_i}$$

$$= \frac{70.2+65.2+65.2+90.2+80.3+72.3}{14}$$

$$= \frac{1036}{14}$$

$$= 74$$

ملاحظة : إن الوسط الحسابي الاعتيادي هو حالة خاصة من الوسط الحسابي المرجح

٣ – الوسيط ا: (The Median)

هي القيمة التي تفصل القيم الى مجموعتين متساويتين أحدهما أكبر من القيمة والأخرى أصغر من القيمة . ويرمز له بالرمز (Me) .

أكبر ← → أصغر

X x x (x) x x x

المجموعة العليا الوسيط المجموعة الدنيا

أكبر ← → أصغر

X x x (**x x**) x x x

المجموعة العليا الوسيط المجموعة الدنيا

طرق حسابه :

أ – في حالة البيانات غير المبوبة :

إذا كان لدينا x_n , ... , x_3 , x_2 , x_1 بيانات المتغير العشوائي x عددها n من المفردات ، ورتبت تصاعدياً أو تنازلياً فإن الوسيط يحتسب من :

١- إذا كانت n عدداً فردياً فإن الوسيط هو القيمة التي يكون ترتيبها $\frac{n+1}{2}$

٢- إذا كانت n عدداً زوجياً فإن الوسيط هو متوسط للقيمتين التي يكون ترتيبهما

$$\frac{n}{2} , \frac{n}{2}+1$$

خطوات استخراجه فى حالة البيانات غير المبوبة :

١- ترتب البيانات تصاعدياً أو تنازلياً .

٢- تحديد قيمة n هل هي عدد فردي أم زوجي .

٣- تحديد ترتيب الوسيط .

٤- القيمة التي تقابل ترتيب الوسيط تكون هي الوسيط إذا كان n فردي ، ومتوسط القيمتين إذا كان n زوجي .

مثال : أوجد الوسيط لدرجات خمسة طلاب في امتحان مادة الإحصاء إذا علمت أن الدرجات هي :

84, 87, 76, 82, 80

الحل :

١- ترتب البيانات ترتيباً تصاعدياً : 76 , 80, 82, 84, 87

٢- تحديد قيمة n : $n = 5$ (عدد فردي)

٣- تحديد ترتيب الوسيط : الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2} = \frac{5+1}{2} = 3$

٤- استخراج قيمة الوسيط : $Me = X_3 = 82$

مثال ٢ : أوجد الوسيط للقيم الآتية : 84, 87, 76, 82, 80, 88

الحل : ترتب البيانات ترتيباً تصاعدياً 76 , 80, 82, 84, 87, 88

$n = 6$ (عدد زوجي)

الوسيط هو القيمة التي ترتيبها $\frac{n}{2} , \frac{n}{2}+1 = \frac{6}{2} , \frac{6}{2}+1$

$= X_3 , X_4 = 82 , 84$

$Me = \frac{82+84}{2} = 83 \qquad = 83$

ب – في حالة البيانات المبوبة :

إذا كانت $x_1 , x_2 , x_3 , \ldots\ldots , x_n$ تمثل مراكز فئات في توزيع تكراري عدد فئاته m فئة ، وكانت $f_1, f_2, \ldots , f_n$ تمثل التكرارات المقابلة لتلك الفئات ، وافرض أن $F_1, F_2 , \ldots , F_n$ تمثل التكرارات المتجمعة الصاعدة المقابلة للحدود العليا لفئات التوزيع ، ولتكن $\frac{\sum f_i}{2}$ تمثل ترتيب الوسيط في هذا التوزيع ، فإذا كان $F_1 < \frac{\sum f_i}{2} < F_2$

عندئذٍ يقال أن قيمة الوسيط هي الفئة التي تسلسلها هو K ، وسوف يتم حساب قيمة الوسيط وفق الصيغة الآتية :

$$Me = L.L. + \frac{\frac{\sum f_i}{2} - F_1}{f_2} \, (L.)$$

حيث :

L L. : الحد الأدنى الحقيقي للفئة الوسيطية.

$\frac{\sum f_i}{2}$: مجموع التكرارات مقسوم على 2 (ترتيب الوسيط)

$F_1\uparrow$: التكرار المتجمع الصاعد للفئة قبل الوسيطية .

f_2 : تكرار الفئة الوسيطية .

L. : طول الفئة .

خطوات إيجاد الوسيط :

١- عمل جدول توزيع تكراري متجمع صاعد .

٢- نستخرج ترتيب الوسيط = $\frac{\sum f_i}{2}$

٣- تحديد الفئة الوسيطية : وهي الفئة التي يكون تكرارها المتجمع الصاعد يساوي ترتيب الوسيط أو أكبر منه .

٤- تطبيق القانون .

مثال : أوجد الوسيط لجدول التوزيع التكراري الآتي جبرياً وبيانياً :

	C	F_i
1	20 – 29	4
2	30 – 39	6
3	40 – 49	2
4	50 – 59	10
5	60 – 69	12
6	70 – 79	8
7	80 – 89	3

الحل :

١) ايجاد التكرار المتجمع الصاعد:

	C	f_i	$F_i\uparrow$
1	20 – 29	4	4
2	30 – 39	6	10
3	40 – 49	2	12
4	50 – 59	10	22
5	60 – 69	12	34
6	70 – 79	8	42
7	80 – 89	3	45

٢) نستخرج ترتيب الوسيط $\frac{\sum f_i}{2} = \frac{45}{2} = 22.5$

٣) نحدد الفئة الوسيطية وهي (69 – 60)

٤) تطبيق القانون

$$Me = L.\,L. + \frac{\frac{\sum f_i}{2} - F_1\uparrow}{f_2}\,(L.)$$

$$= 59.5 + \frac{22.5 - 22}{12}\,(10)$$

$$= 59.5 + 0.41 = 59.91$$

لايجاد الوسيط بيانياً بالرسم

١) نرسم المحورين الافقي والعمودي .

٢) نعين على المحور الافقي الحدود العليا للفئات ، وعلى المحور العمودي التكرارات المتجمعة الصاعدة .

٣) نعين نقاط أعلى كل حد أعلى للفئة ارتفاعها يساوي التكرار المتجمع الصاعد لتلك الفئة

٤) نصل بين هذه النقاط .

خصائص الوسيط :

١- لا يتأثر اطلاقاً بالقيم المتطرفة .

٢- يمكن تعيينه هندسياً .

٣- يمكن حسابه في حالة التوزيعات التكرارية المفتوحة س طرف واحد أو طرفين .

H. W.

١) استخرج الوسط الحسابي المرجح للتوزيعات الآتية :

X_i	10	18	15	11	19	22	20
W_i	1	4	3	2	5	7	6

X_i	1	0	2	2	3	1	4
W_i	2	4	3	5	3	2	2

٢) في جدول التوزيع التكراري الآتي :

أ) رسم المدرج والمضلع التكراري لهذا التوزيع .

ب) أوجد الوسط الحسابي .

ج) أوجد الوسيط جبرياً وبيانياً .

	C	f_i
1	11 – 15	3
2	16 – 20	8
3	21 – 25	15
4	26 – 30	25
5	31 – 35	20
6	36 – 40	17
7	41 – 45	6
8	46 – 50	10

٣) الآتي توزيع تكراري لكميات الكهرباء المستهلكة (كيلو واط) من قبل 40 دار سكنية خلال شهر واحد

الاستهلاك (كيلو واط)	800 -	900 -	1000 -	1100 -	1200 -	1300 فأكثر
عدد الدور	7	8	15	5	3	2

جد الوسط والوسيط للكمية المستهلكة من الكهرباء ثم تحقق من ذلك بالرسم .

٤) أ) أكمل جدول التوزيع التكراري أدناه ، وارسم المدرج والمضلع التكراري ، واستخرج الوسيط جبرياً وبيانياً .

	C	F_i	X
1	10 -	4	14.5
2		7	
3		10	
4		25	
5		18	
6		12	
7		6	
8		18	

٥) جد قيمة z التي تجعل قيمة الوسط الحسابي في المجموعة 20, 25, 28, 31, 37, z مساوٍ إلى 30 ، ثم استخرج الوسيط .

٤ – المنوال : (The Mode)

يعرف المنوال بأنه القيمة الأكثر تكراراً وشيوعاً بين مجموعة من القيم ، ويرمز له بالرمز (Mo) .

طرق حساب المنوال :

١ – في حالة البيانات غير المبوبة :

لتكن x_n , ... , x_3 , x_2 , x_1 بيانات المتغير العشوائي x في توزيع تكراري قوامها (عددها) n من المفردات ، ولنفرض أن x_j قيمة من قيم هذه المجموعة لوحظ أنها تكررت أكثر من غيرها ، عندئذٍ وبحسب تعريف المنوال فإن x_j تمثل المنوال لهذه المجموعة .

أ- إذا لم تتكرر أية قيمة فلا يوجد منوال :

مثال : أوجد المنوال للقيم الآتية : 15, 12, 16, 22, 17 ,18, 19

الحل : لا يوجد منوال لهذه القيم لعدم وجود قيمة مكررة فيها .

ب- إذا تكرر أحدى القيم أكثر من غيرها فإن هناك منوال واحد .

مثال : أوجد المنوال للقيم الآتية : 15, 12, 16, 22, 15 ,18, 19

الحل : المنوال (Mo) = 15 لأنها القيمة الأكثر تكراراً .

ج- إذا كان لقيمتين من قيم التوزيع نفس العدد من التكرارات فإن هناك منوالين .

مثال : أوجد المنوال للقيم الآتية : 15, 12, 16, 22, 15 ,18, 19,19

الحل : يوجد منوالان هما 19 , 15 لأن لهما نفس التكرارات .

٢ – في حالة البيانات المبوبة :

إذا كانت x_n , ...,x_2 , x_1 تمثل مراكز فئات في توزيع تكراري عدد فئاته m فئة ، و ... ,f_1, f_2 , f_n هي تكرارات تلك الفئات **فإن المنوال يستخرج بإتباع الخطوات الآتية :**

١- تحديد الفئة المنوالية : : وهي الفئة التي تقابل أعلى تكرارٍ بين الفئات .

٢- تحديد الفرق ما بين تكرار الفئة المنوالية والفئة السابقة لها ، وليكن f_1.

٣- تحديد الفرق ما بين تكرار الفئة المنوالية والفئة اللاحقة لها ، وليكن f_2.

٤- تطبيق العلاقة الآتية لإيجاد المنوال :

$$Mo = L.L. + \frac{f_1}{f_1 + f_2} (L.)$$

حيث L.L. الحد الأدنى الحقيقي للفئة المنوالية .

L. طول الفئة

مثال : أوجد المنوال جبرياً وبيانياً لجدول التوزيع التكراري الآتي :

	C	f_i
1	20 – 29	2
2	30 – 39	4
3	40 – 49	7
4	50 – 59	10
5	60 – 69	8
6	70 – 79	5
7	80 – 89	3

الحل :

1) 50 – 59 الفئة المنوالية

2) $f_1 = 10 - 7 = 3$

3) $f_2 = 10 - 8 = 2$

4) $Mo = L.L. + \frac{f_1}{f_1 + f_2} (L.)$

$$= 49.5 + \frac{3}{3 + 2} (10)$$

$$= 55.5$$

ولإيجاد المنوال بيانياً (بالرسم) نتبع الخطوات الآتية :

١) نرسم المدرج التكراري .

٢) نحدد الفئة المنوالية : وهي الفئة التي تقابل أكبر تكرار

٣) نحدد الفئة السابقة والفئة اللاحقة الفئة المنوالية .

٤) نرسم خط مستقيم يصل بداية عمود المنوال (العمود الأعلى تكراراً) ببداية العمود الذي بعده ، وخط مستقيم آخر يصل نهاية عمود المنوال بنهاية العمود السابق له .

٥) من نقطة تقاطع المستقيمين نرسم مستقيم عمودي على المحور الافقي ، نقطة تقاطعهما تمثل المنوال .

خصائص المنوال :

١- لا يتأثر بالقيم المتطرفة (صغيرة أو كبيرة) التي تقع على جانبي التوزيع .

٢- يمكن ايجاده بسهولة لأنه القيمة الأكثر تكراراً بين القيم .

٣- يمكن ايجاده في التوزيعات التكرارية المفتوحة (غير المنتهية) والجداول المفتوحة .

العلاقة بين مقاييس النزعة المركزية:

تحدد العلاقة بالقانون الاتي:

$$\boxed{\bar{X} - Mo = 3(\bar{X} - Me)}$$

مثال : في احد التوزيعات التكرارية إذا كانت قيمة $Mo=17$ و $\bar{X}=12$, جد Me

الحل :

$$\bar{X} - Mo = 3(\bar{X} - Me)$$

$$12 - 17 = 3(12 - Me)$$

$$-5 = 36 - 3Me$$

$$3Me = 36 + 5$$

$$3Me = 41$$

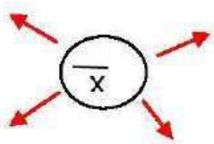
$$Me = \frac{41}{3}$$

الفصل الرابع

مقاييس التشتت

وردت كلمة تشتت في الأدبيات الاحصائية بمعنى مبعثر(Scattered) أي انتشار قيم مجموعة من البيانات .

ويعرف التشتت بأنه الدرجة التي تتجه بها البيانات الكمية للتباعد والانتشار عن بعضها البعض ، أو حول قيمة متوسطة وغالباً ما تكون أحد مقاييس النزعة المركزية (الوسط ،والوسيط) .



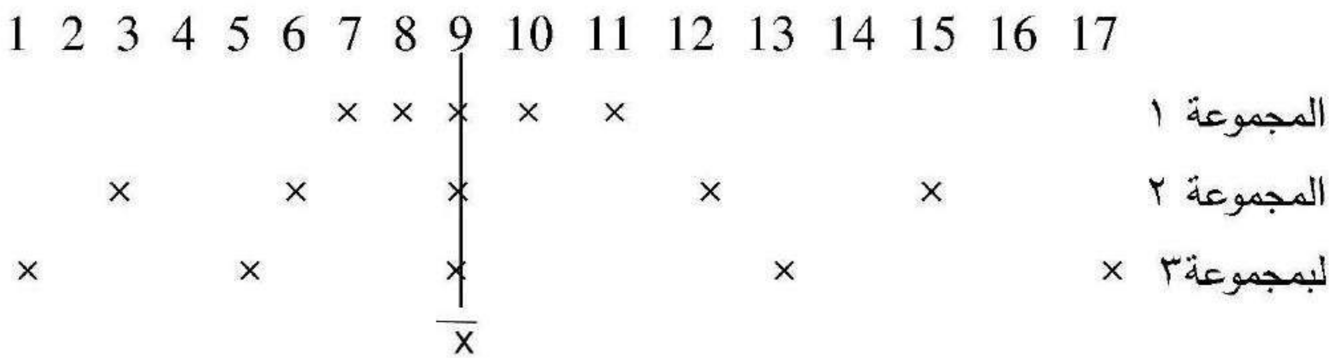
وإن الهدف من دراسة التشتت هو تكوين فكرة عن مدى تجانس قيم مجموعة من المفردات ، وهذا يعني أن دراسة التشتت أمر مفيد في اجراء المقارنة بين قيم مجموعة أو أكثر من البيانات عن ظاهرة معينة .

<u>فعلى سبيل المثال :</u> إن الوسط الحسابي للمجاميع الآتية هو (9)

المجموعة ١ : 7, 8, 9, 10, 11 المجموعة ٢ : 3, 6, 9, 12, 15

المجموعة ٣ : 1, 5, 9, 13, 17

نلاحظ أن المجموعة الأولى أكثر تجانساً (أقل انتشاراً) من والمجوعتين الثانية والثالثة كذلك فإن المجموعة الثانية أكثر تجانساً من المجموعة الثالثة ، كما موضح في الشكل :



نلاحظ من المثال إنه أمكن اجراء المقارنة بين المجموعات الثلاث بعد أن كانت المقارنة بينهما على أساس الوسط الحسابي عديمة الفائدة .

مثال : كانت درجات الحرارة ثلاثة من المحافظات الشمالية هي كالآتي :

المحافظة الأولى : $-2, -1, 0, 1, 2, 3, 4$

المحافظة الثانية : $-3, -2, -1, 0, 2, 4, 7$

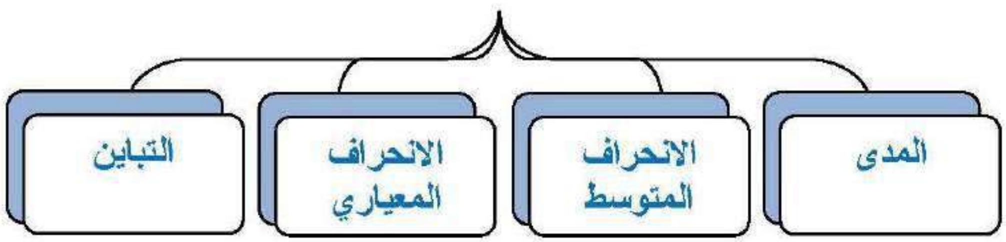
المحافظة الثالثة : $-1, 0, 0, 1, 2, 2, 3$

في أية محافظة كان الجو أكثر استقراراً ؟

مقاييس التشتت

تعرف مقاييس التشتت بأنها المقاييس التي تبحث في مقدار الاختلافات بين البيانات وهي على الأنواع الاتية :

أنواع مقاييس التشتت



١-المدى (Range):

هو الفرق ما بين اكبر قيمة واصغر قيمة في المجموعة ، ويرمز له بالرمز R .

فإذا كانت $x_1, x_2, x_3, \dots, x_n$ تمثل بيانات مجموعة من المفردات قوامها (عددها) n وإن x_l أكبر قيمة فيها ، وإن x_s أصغر قيمة فيها عندئذٍ فإن المدى في هذه المجموعة هو:

$$\boxed{R = x_l - x_s}$$

مثال : أوجد المدى للبيانات الآتية : **3, 4, 5, 7, 8, 9, 10, 12**

الحل : الندى هو الفرق ما بين أكبر قيمة وأصغر قيمة

$$R = x_l - x_s$$

$$R = 12 - 3 = 9$$

٢-الانحراف المتوسط :

ويعرف بأنه مجموع الانحرافات المطلقة لقيم متغير عشوائي عن نقطة اختيارية مثل A مقسوماً على عدد هذه القيم . وغالباً ما نختار النقطة A لأن تكون أحد مقاييس النزعة المركزية الثلاثة (الوسط الحسابي ، الوسيط ، المنوال) . ويرمز له بالرمز M.D .

طرق حساب الانحراف المتوسط :

١-في حالة البيانات غير المبوبة :

لتكن $x_1, x_2, x_3, \ldots , x_n$ تمثل بيانات مجموعة من المفردات قوامها (عددها) n، وكانت A ثابت اختياري ، عندئذٍ يمكن حساب الانحراف المتوسط لهذه المجموعة وفق القانون الآتي :

$$M.D = \frac{\sum |x_i - A|}{n}$$

حيث A : أي مقياس من مقاييس النزعة المركزية .

خطوات استخراج الانحراف المتوسط :

١-تحديد قيمة مقياس النزعة المركزية (الوسط، الوسيط، المنوال).

٢-ايجاد الفرق بين كل قيمة من قيم المتغير (البيانات) وقيمة A .

٣-ايجاد القيمة المطلقة للفرق ، ثم ايجاد المجموع وقسمته على n .

ففي المثال السابق : أوجد الانحراف المتوسط للبيانات الآتية باستخدام (الوسط الحسابي ، الوسيط ، المنوال) : **3, 4, 5, 7, 8, 9, 10, 12**

الحل :

أ) باستخدام الوسط الحسابي

$A = \overline{x}$

$$\overline{x} = \frac{\sum x_i}{}$$

$$\overline{x} = \frac{54}{8} = 6.75$$

x_i	$x_i - \overline{x}$	$\lvert x_i - \overline{x} \rvert$	
3	-3.75	3.75	
4	-2.75	2.75	
5	-1.75	1.75	
7	0.25	025	

8	1.25	1.25	
9	2.25	2.25	
10	3.25	3.25	
12	5.25	5.25	
$\sum$		45.25	

$$M.D = \frac{\sum |x_i - \overline{X}|}{n}$$

$$= \frac{45.25}{8} = 5.65$$

مثال ٢ : أوجد الانحراف المتوسط للبيانات الآتية باستخدام (الوسط الحسابي ، الوسيط ، المنوال) :

2, 3, 4, 5, 6, 10, 11, 13, 12, 14

الحل :

ب) باستخدام الوسط الحسابي

$$\text{if } A = \overline{X}$$

$$\overline{X} = \frac{\sum x_i}{}$$

$$\overline{X} = \frac{80}{10} = 8$$

x_i	$x_i - \overline{X}$	$\mid x_i - \overline{X} \mid$
2	-6	6
3	-5	5
4	-4	4
5	-3	3
6	-2	2
10	2	2
11	3	3
12	4	4
13	5	5
14	6	6
$\sum$		40

$$M.D = \frac{\sum |x_i - \overline{X}|}{n}$$

$$= \frac{40}{10} = 4$$

٢) باستخدام الوسيط :

if A = Me

ترتيب الوسيط

$$\frac{n}{2}, \frac{n}{2}+1 \quad = 5, 6$$

$$Me = \frac{x_5 + x_6}{2}$$

$$= \frac{6 + 10}{2} = 8$$

x_i	$x_i - Me$	$\lvert x_i - Me \rvert$
2	-6	6
3	-5	5
4	-4	4
5	-3	3
6	-2	2
10	2	2
11	3	3
12	4	4
13	5	5
14	6	6
Σ		40

$$M.D = \frac{\sum \lvert x_i - Me \rvert}{n}$$

$$= \frac{40}{10} = 4$$

٣) باستخدام المنوال:

if A = Mo

لا يوجد منوال وعليه الانحراف المتوسط في هذه الحالة هو

$$M.D = \frac{\sum \lvert x_i - Mo \rvert}{n} = \frac{80}{10} = 8$$

٢-في حالة البيانات المبوبة :

إذا كانت x_n , ... , x_2 , x_1 تمثل مراكز فئات في توزيع تكراري عدد فئاته m فئة ، و $f_1, f_2,$ ... , f_n هي تكرارات تلك الفئات، ولنفرض أن A قيمة اختيارية ، عندئذٍ يمكن حساب الانحراف المتوسط باتباع الخطوات الآتية :

١) استخراج قيمة A وهي أحد مقاييس النزعة المركزية

٢) استخراج $f_i x_i$ وهو حاصل ضرب التكرار في مركز الفئة .

٣) ايجاد الفرق مابين $f_i x_i$ و A.

٤) أخذ القيمة المطلقة لها ، وايجاد المجموع لها وقسمتها على مجموع التكرارات .

٥) تطبيق القانون الآتي :

$$\boxed{M.D = \frac{\sum |f_i x_i - A|}{\sum f_i}}$$

وغالباً يتم اختيار A لتكون أحد مقاييس النزعة المركزية .

مثال : الآتي توزيع تكراري لدرجات مجموعة من الطلبة في مادة الاحصاء التربوي،

أوجد الانحراف المتوسط مستخدماً (الوسط ، الوسيط) :

	C	f_i
1	1 – 5	6
2	6 – 10	2
3	11 – 15	10
4	16 – 20	12
5	21 – 25	8
6	26 – 30	3

الحل :

١) في حالة الوسط : if $A = \overline{X}$

$$\overline{X} = \frac{\sum f_i x_i}{\sum f_i} = 15$$

	C	f_i	X	f_i X	f_i X - A	| f_i X – A |
1	1 – 5	5	3	15	0	0
2	6 – 10	3	8	24	9	9
3	11 – 15	11	13	143	128	128
4	16 – 20	8	18	144	129	129

5	21 – 25	5	23	115	100	100
6	26 – 30	3	28	84	69	69
		35		525		435

$$M.D = \frac{\sum |f_i x_i - A|}{\sum f_i}$$

$$= \frac{435}{35} = 12.43$$

٢) في حالة الوسيط :   if A = Me

$$Me = L.L. + \frac{\frac{\sum f_i}{2} - F_1\uparrow}{f_2} (L.)$$

$$= 10.5 + \frac{17.5 - 8}{11} (5) = 14.82$$

	C	f_i	X	$F_1\uparrow$	$f_i X$	$f_i X - A$	$\lvert f_i X - A \rvert$
1	1 – 5	5	3	5	15	0.18	0.18
2	6 – 10	3	8	8	24	9.18	9.18
3	11 – 15	11	13	19	143	128.18	128.18
4	16 – 20	8	18	27	144	129.18	129.18
5	21 – 25	5	23	32	115	100.18	100.18
6	26 – 30	3	28	35	84	69.18	69.18
		35			525		436.08

$$M.D = \frac{\sum |f_i x_i - Me|}{\sum f_i}$$

$$= \frac{436.08}{35} = 12.4594$$

٣- الانحراف المعياري(القياسي) :

ويسمى في بعض الأحيان بالانحراف القياسي ، ويعد أفضل مقاييس التشتت على الاطلاق .

ويعرف بأنه الجذر التربيعي الموجب لمتوسط مجموع مربعات انحرافات قيم المتغير العشوائي عن وسطها الحسابي ، ويرمز له بالرمز s في حالة العينة وبالرمز 6 في حالة المجتمع .

٤- التباين :

ويعرف بأنه متوسط مجموع مربعات انحرافات قيم المتغير العشوائي عن وسطها الحسابي ، ويرمز له بالرمز S^2 في حالة العينة وبالرمز 6^2 في حالة المجتمع .

أي إنه مربع الانحراف المعياري .

طرق حساب الانحراف المعياري والتباين :

١) الانحراف المعياري والتباين للبيانات غير المبوبة:

لحساب الانحراف المعياري والتباين للعينة نستخدم القوانين الآتية :

الطريقة المطولة — الانحراف المعياري

$$S=\sqrt{=\frac{\sum(xi-\bar{x})2}{n-1}}$$

التباين

$$S^2=\frac{\sum(xi-x)2}{n-1}$$

الطريقة المختصرة — الانحراف المعياري

$$S=\sqrt{\frac{\sum(xi^{\,2}-\frac{(\sum xi)^2}{n}}{n-1}}$$

التباين

$$S^2=\frac{\sum(xi^{\,2}-\frac{(\sum xi)^2}{n}}{n-1}$$

اما لحساب الانحراف المعياري والتباين للمجتمع فنستخدم القوانين الآتية :

الطريقة المطولة — الانحراف المعياري

$$6=\sqrt{\frac{\sum(xi-x)2}{n}}$$

التباين

$$\sigma^2=\frac{\sum(xi-x)2}{n}$$

الطريقة المختصرة الانحراف المعياري $6= \sqrt{\frac{\sum(xi^2 - \frac{(\sum xi)^2}{n}}{n}}$

التباين $\sigma^2 = \frac{\sum(xi^2 - \frac{(\sum xi)^2}{n}}{n}$

ملاحظة : يتم استخدام الانحراف المعياري للمجتمع في أغلب الاحيان .

مثال: البيانات الاتية 9,8,6,5,7 تبين كمية محصول القطن في خمس مزارع احسب الانحراف المعياري والتباين لها.

الحل :

X_i	$(xi - x)$	$(xi - x)^2$
9	2	4
8	1	1
6	−1	1
5	−2	4
7	0	0
		10

$\bar{X} = \frac{35}{5} = 7$

$S = \sqrt{\frac{\sum(xi-x)2}{n-1}}$

$= \sqrt{\frac{10}{5-1}}$

$= \sqrt{\frac{10}{4}}$

$= \sqrt{2.5} \quad = 1.58$

$S^2 = 2.5$

٢) الانحراف المعياري والتباين للبيانات المبوبة:

لحساب الانحراف المعياري والتباين للعينة نستخدم القوانين الآتية :

الطريقة المطولة الانحراف المعياري $S=\sqrt{\frac{\sum fi(xi-\bar{x})2}{\sum fi-1}}$

التباين $S^2=\frac{\sum fi(xi-x)2}{\sum fi-1}$

الطريقة المختصرة الانحراف المعياري $S=\sqrt{=\frac{\sum fixi^{\,2}-\frac{(\sum fixi)^2}{\sum fi}}{\sum fi-1}}$

التباين $S^2=\frac{\sum fixi^{\,2}-\frac{(\sum fixi)^2}{\sum fi}}{\sum fi-1}$

اما لحساب الانحراف المعياري والتباين للمجتمع فنستخدم القوانين الآتية :

الطريقة المطولة الانحراف المعياري $6=\sqrt{\frac{\sum fi(xi-x)2}{\sum fi}}$

التباين $6^2=\frac{\sum fi(xi-x)2}{\sum fi-1}$

الطريقة المختصرة الانحراف المعياري $\sigma=\sqrt{\frac{\sum fixi^{\,2}-\frac{(\sum fixi)^2}{\sum fi}}{\sum fi}}$

التباين $6^2=\frac{\sum fixi^{\,2}-\frac{(\sum fixi)^2}{\sum fi}}{\sum fi}$

مثال: احسب الانحراف المعياري لجدول التوزيع التكراري الاتي:

C	Fi	Xi	Fixi	$(xi-\bar{x})$	$(xi-\bar{x})^2$	$Fi(xi-\bar{x})^2$
50-55	5	52.5	262,5	-18,7	349.69	1748.45
60-65	14	62.5	875	-8.7	75,69	1059.66
70-75	22	72.5	1595	1,3	1.69	37.18
80-85	9	82.5	742.5	11.3	127.69	1149.51

90-95	4	92.5	370	21.3	453.69	1814.76
	54		3845			5809.56

$$\overline{X}=\frac{\sum fixi}{\sum fi}$$

$$=\frac{3845}{54} = 71.2$$

$$\sigma=\sqrt{\frac{\sum fi(xi-x)2}{\sum fi}}$$

$$=\sqrt{\frac{5809.56}{54}}$$

$$\sigma^2=\frac{5809.56}{54}$$

$$=107.58$$

الفصل الخامس

مقاييس الارتباط

يعرف الارتباط بأنه درجة أو قيمة العلاقة بين التي تربط متغيرين أو أكثر مع بعضهما البعض ، وهي قيمة حقيقية خالية من وحدات قياس المتغيرات المرتبطة بالعلاقة .

وهناك عدة أنواع من الارتباط يختص كل منها بنوع معين من العلاقات ، وستقتصر دراستنا على معاملي ارتباط بيرسون للمتغيرات الكمية ، ومعامل ارتباط الرتب لسبيرمان الخاص بالمتغيرات الوصفية .

خواص معامل الارتباط :

١) **تتراوح قيمة معامل الارتباط بين $(-1,1)$ أي أن : $-1 \le r \le 1$.**

٢) **إذا كان $r=1$ فهذا يعني وجود <u>علاقة موجبة تامة طردية</u> بين المتغيرين ، أي ان زيادة في قيم احد المتغيرين يصحبه زيادة في قيم المتغير الأخر. (ترابط طردي)**

٣) **إذا كان $r=-1$ فهذا يعني وجود <u>علاقة سالبة تامة عكسية</u> بين المتغيرين ، أي ان زيادة في قيم احد المتغيرين يصحبه نقصان في قيم المتغير الأخر. (ترابط عكسي)**

٤) **إذا كانت $r=0$ فهذا يعني <u>عدم وجود ترابط خطي بينهما</u> ، أي لا توجد علاقة بينهما. (لا يوجد ترابط)**

انواع معاملات الارتباط

- معامل ارتباط بيرسون
- معامل ارتباط سبيرمان

١) معامل الارتباط الخطي البسيط (معامل ارتباط بيرسون) :

ويعرف بأنه القيمة العددية للعلاقة بين متغيرين كميين ، ويعد العالم الانجليزي كارل بيرسون أول من وضع صيغة لهذا المعامل . **ونرمز له بالرمز (r أو $r_{x,y}$ أو $r_{1,2}$).**

حساب قيمة معامل ارتباط بيرسون :

لتكن (x_i , y_i); $i = 1, 2, 3, ..., n$ تمثل أزواج القيم التي يتم الحصول عليها من عينة من البيانات عددها n من البيانات ، عندئذٍ نستخدم القانون اللآتي لاستخراجه :

$$r = \frac{\sum xiyi - \frac{\sum xi \sum yi}{n}}{\sqrt{(\sum xi^2 - \frac{(\sum xi)^2}{n})(\sum yi^2 - \frac{(\sum yi)^2}{n}}}$$

مثال: احسب معامل ارتباط بيرسون للبيانات الاتية والتي تمثل الدرجات الفصلية 7,6, 5.8,10 والدرجات النهائية 8,7,6,8,6 لخمسة طلاب.

X_i	y_i	Y_i^2	X_i^2	x_iy_i
7	8	49	64	56
6	7	36	49	42
5	6	25	36	30
8	8	64	64	64
10	6	100	36	60
249	274	252	35	36

$$r = \frac{\sum xiyi - \frac{\sum xi \sum yi}{n}}{\sqrt{(\sum xi^2 - \frac{(\sum xi)^2}{n})(\sum yi^2 - \frac{(\sum yi)^2}{n}}}$$

$$= \frac{252 - \frac{36\times35}{5}}{\sqrt{(274 - \frac{(36\times36)}{})(249 - \frac{(35\times35)}{5}}}$$

$$= \frac{252-252}{\sqrt{(274-259.2)(249-245)}}$$

$$= \frac{0}{\sqrt{14.8\times4}} = 0$$ لايوجد ارتباط

٢) معامل ارتباط سبيرمان للرتب :

ويعرف بأنه القيمة العددية للعلاقة بين متغيرين وصفيين. ونرمز له بالرمز (r_s).

ويقيس معامل ارتباط الرتب التغير الاقتراني القائم بين ترتيب الافراد بالنسبة لصفة معينة ، ففي بعض الاحيان يمكن وصف مركز الفرد في جماعته عن طريق ترتيبه بينهم في سمة معينة.

حساب قيمة معامل ارتباط سبيرمان للرتب :

لنفرض أن y ، x متغيرين من النوع الوصفي ، ولنفرض أن البيانات التي جمعت عن المتغيرين x,y عددها n هي بهيئة صفات غير قابلة للقياس الكمي ، فلو فرضنا أن (x_i,y_i); $i = 1, 2, 3, ..., n$ ممكنة الترتيب تصاعدياً أو تنازلياً وفق معيار معين يمتاز به كل متغير (مثلاً درجات مجموعة من الطلبة يمكن ترتيبها تصاعدياً على أساس معيار الأقل إلى الأعلى درجة أي ضعيف ، مقبول ، متوسط ، جيد ، جيد جداً ، ممتاز و بالعكس أي الترتيب التنازلي .عندئذٍ واستناداً إلى هذا التعريف إذا تم تخصيص سلسلة الأعداد الطبيعية ..., 1, 2, 3 للتعبير عن صفات الترتيب لكل متغير بحيث أن كل صفة يخصص لها أحد أعداد هذه السلسلة (مثلاً : ضعيف يخصص لها 1 ، ومتوسط يخصص لها 3 ، وهكذا) في حالة عدم تكرار قيمة أي صفة من صفات المتغير ، عندها يستخدم القانون الآتي لإيجاد معامل الارتباط بين المتغيرين x , y :

$$R_s = 1 - \frac{6\sum(Rx-Ry)^2}{n(n^2-1)}$$

<u>مثال 1</u> : كانت تقديرات درجات ستة طلبة في مادتي الاحصاء والرياضيات هي كالآتي :

الاحصاء : جيد ، متوسط ، ضعيف ، مقبول ، جيد جداً ، ممتاز

الرياضيات : متوسط ، جيد ، مقبول ، ضعيف ، ممتاز ، جيد جداً

أوجد معامل الارتباط بين تقديرات درجات الطلاب في امتحاني الاحصاء والرياضيات (هل توجد علاقة ارتباطية بين تقديرات درجات الطلاب في امتحاني الاحصاء والرياضيات؟ وما نوعها ؟)

الحل :

ترتب التقديرات وفق ترتيب تصاعدي أو تنازلي ، وليكن الترتيب تصاعدي وتخصص لها سلسلة الأعداد الطبيعية .

التقديرات : ضعيف ، مقبول ، متوسط ، جيد ، جيد جداً ، ممتاز

1 2 3 4 5 6

ثم يعوض عن كل تقدير بما يقابله من اعداد طبيعية

الاحصاء x	الرياضيات y	R_x	R_y	$d_i = R_x - R_y$	$d_i^2 = (R_x - R_y)^2$
جيد	متوسط	4	3	1	1
متوسط	جيد	3	4	-1	1
ضعيف	مقبول	1	2	-1	1
مقبول	ضعيف	2	1	1	1
جيد جداً	ممتاز	5	6	-1	1
ممتاز	جيد جداً	6	5	1	1
					6

$$R_s = 1 - \frac{6\sum(Rx - Ry)^2}{n(n^2-1)}$$

$$= 1 - \frac{\mathbf{6(6)}}{\mathbf{6(35)}} = 0.8$$

علاقة موجبة طردية جيدة جداً بين درجات الطلاب

مثال 2 : أوجد معامل ارتباط سبيرمان للرتب بين قيم المتغيرين في الجدول أدناه :

N	x	Y
1	6	9
2	8	10
3	10	7
4	7	6
5	9	8

الحل :

نرتب قيم المتغيرين تصاعدياً ونعبر عنها بسلسلة الأعداد الطبيعية ... 1, 2, 3,

القيم : 6 , 7, ,8 , 9 , 10

1 2 3 4 5

N	x	Y	Rx	Ry	(Rx –Ry)	$(Rx - Ry)^2$
1	6	9	1	4	−3	9
2	8	10	3	5	−2	4
3	10	7	5	2	3	9
4	7	6	2	1	1	1
5	9	8	4	3	1	1
						24

$$R_s = 1 - \frac{6\sum(Rx - Ry)^2}{n(n^2-1)}$$

$$= 1 - \frac{6\times24}{5(5^2-1)}$$

$$= 1 - \frac{144}{5\times24}$$

$$= 1 - \frac{144}{120}$$

$= -0.20$ علاقة موجبة طردية ضعيفة

التصميم العشوائي الكامل
(Completely Randomized Design –CRD)

يعد التصميم العشوائي الكامل واحد من أكثر التصاميم أستعمالا في مجال الانتاج الحيواني والنباتي، كما أنه سهل التطبيق فضلا الى ذلك فان من أهم ميزاته هو أمكانية تطبيقه مهما ان عدد المعاملات في التجربة وذلك عدد المكررات في كل معاملة ويمكن تطبيقه حتى في حالة عدم تساوي المكررات بأختلاف المعاملات ، الا أن من أهم محددات هذا التصميم هي عدم أمكانية تطبيقه الا أذا كانت الوحدات التجريبية على درجة عالية من التجانس.

أولا: التصميم العشوائي الكامل (CRD) في حالة تساوي عدد المكررات (مع تسجيل مشاهدة واحدة).

الانموذج الرياضي للتصميم : (Mathematical Model).

$$Y_{ij} = \mu + T_i + e_{ij}$$

أذ أن :

Y_{ij}: قيمة المشاهدة j العائدة للمعاملة i.

μ : المتوسط العام للصفة المدروسة.

T_i: تأثير المعاملة i .

e_{ij}: الخطأ العشوائي الذي يتوزع توزيعا طبيعيا بمتوسط يساوي صفر وتباين قدره $\sigma^2 e$.

جدول تحليل التباين للتصميم :(**ANOVA Table**).

S.O.V. مصادر الاختلاف	d.f. درجات الحرية	S.S. مجموع المربعات	M.S. متوسط المربعات	F. Value قيمة f المحسوبة
Treat. المعاملة	t-1	$SSt = \frac{\sum Yi.^2}{R} - CF$	$MSt = \frac{SSt}{t-1}$	$F = \frac{MSt}{MSe}$
Experimental Error. الخطأ التجريبي	t(r-1)	$SSe = SST - SSt$	$MSe = \frac{SSe}{t(r-1)}$	
Total الكلي	tr-1	$SST = \sum Yij^2 - CF$	-------------	

علما أن :

t: عدد المعاملات في التجربة

r: عدد المشاهدات أو المكررات في كل معاملة

وأن CF يمثل معامل التصحيح ويساوي مربع مجموع القيم مقسوما الى عددها والعدد ناتج من ضرب عدد المعاملات (t) في عدد المكررات لكل معاملة (r).

أي أن :

$$CF = \frac{(Y..)^2}{tr}$$

<u>مثال:</u> أجريت تجربة شملت ثلاث سلالات (معاملات) من الابقار ، لدراسة تأثير السلالة في نسبة الدهن في الحليب وضمت كل معاملات أربعة أبقار أخذت عينة حليب (أنموذج) واحدة من كل منها لقياس نسبة الدهن وكانت كالاتي:

المعاملة (السلالة)	نسبة الدهن (Yij)	المجموع (Yi.)
فريزيان	3 , 3 , 4 , 2	12
براون سويس	4 , 5 , 3 , 4	16
جرسي	4 , 3 , 3 , 3	13
		Y.. = 41 المجموع الكلي

الحل :

يتم حساب معامل التصحيح أولا:

$$CF = \frac{(Y..)^2}{tr} = \frac{(41)^2}{3 \times 4} = 140.8$$

ثم مجموع مربعات المعاملات (SSt):

$$SSt = \frac{\sum Yi.^2}{r} - CF = \frac{(12)^2 + (16)^2 + (13)^2}{4} - 140.8$$

$$SSt = 2.166$$

يتم حساب مجموع المربعات الكلية (SST):

$$SST = \sum Yij^2 - CF$$
$$SST = 3^2 + 3^2 + 4^2 + \text{-----} + 3^2 - 140.8$$
$$SST = 6.92$$

يتم حساب مجموع مربعات الخطأ (SSe):

$$SSe = SST - SSt$$
$$SSe = 6.92 - 2.166$$
$$SSe = 4.75$$

ومن النتائج السابقة يمكن حساب متوسط مربعات كل من المعاملات والخطأ وكما يلي:

متوسط مربعات المعاملات (MSt):

$$MSt = \frac{SSt}{t-1} = \frac{2.166}{3-1} = \frac{2.166}{2} = 1.08$$

متوسط مربعات الخطأ (MSe):

$$MSe = \frac{SSe}{t(r-1)} = \frac{4.75}{3(4-1)} = \frac{4.75}{9} = 0.53$$

ومن خلال متوسط مربعات المعاملة والخطأ يمكن حساب قيمة F وكما يلي:

$$F = \frac{MSt}{MSe} = \frac{1.08}{0.53} = 2.05$$

ومن ثم يتم تكوين جدول تحليل التباين لتحليل البيانات:

جدول تحليل التباين للتصميم :(ANOVA Table).

S.O.V. مصادر الاختلاف	d.f. درجات الحرية	S.S. مجموع المربعات	M.S. متوسط المربعات	F. Value قيمة f المحسوبة
Treat. المعاملة	t-1= 3-1 = 2	SSt = 2.166	MSt = 1.08	
Experimental Error. الخطأ التجريبي	t(r-1) = 3(4-1) = 9	SSe = 4.75	MSe = 0.53	$F = \frac{1.08}{0.53}$
Total الكلي	tr-1 = 3X4- 1 = 12	SST = 6.92	-------------	F = 2.05

تقارن قيمة F المحسوبة (Calculated) وهي (2.05) مع قيمة F الجدولية (Tabulated) من جداول F (منشورة في نهاية كتب تصميم وتحليل التجارب) وفق درجات حرية المعاملة (2) ودرجات حرية الخطأ (9) فأذا كانت المحسوبة أعلى من الجدولية فأن تأثير المعاملة (السلالة) معنويا في الصفة المدروسة، وأذا كانت قيمة F المحسوبة أقل من الجدولية فأن تأثير المعاملة في نسبة الدهن غير معنوي (Non-significant) : ففي المثال السابق التأثير غير معنوي.

ويتم أختبار قيمة F على مستوى أحتمالية 0.05 أي (P<0.05) وأشارتها *

أو على مستوى أحتمالية 0.01 أي (P<0.01) وأشارتها **

وأن * تعني معنوي و ** عالي المعنوية.

سؤال1 واجب: أكمل جدول تحليل التباين الاتي :

S.O.V. مصادر الاختلاف	d.f. درجات الحرية	S.S. مجموع المربعات	M.S. متوسط المربعات	F. Value قيمة f المحسوبة
Treat. المعاملة	3	60	-----------	-----------
Experimental Error. الخطأ التجريبي	--------	----------	15	
Total الكلي	19	-----------		

سؤال2 واجب: أكتب جدول تحليل التباين بالرموز للانموذج الرياضي الاتي:

$Yij = \mu + Ti + eij$

سؤال3 واجب: ما هي ميزات ومحددات تطبيق التصميم العشوائي الكامل (CRD).

ملاحظة: بالامكان أستخراج معامل أختلاف التجربة (CV) وفق القانون الاتي : (من قسمة جذر متوسط مربعات الخطأ MSe (يؤخذ من جدول تحليل التباين) على المتوسط العام للصفة (X) في 100.

$$CV\% = \frac{MSe}{X} \times 100$$

ثانيا: التصميم العشوائي الكامل (CRD) في حالة عدم تساوي المكررات (مع تسجيل مشاهدة واحدة).

الانموذج الرياضي للتصميم : (Mathematical Model).(كما في حالة تساوي المكررات أنف الذكر). أي

$$\mathbf{Yij = \mu + Ti + eij}$$

وتفسير رموزه كما في الانموذج السابق.

جدول تحليل التباين للتصميم :(ANOVA Table).

S.O.V. مصادر الاختلاف	d.f. درجات الحرية	S.S. مجموع المربعات	M.S. متوسط المربعات	F. Value قيمة f المحسوبة
Treat. المعاملة	t-1	$SSt = \frac{\sum Yi.^2}{Ri} - CF$	$MSt = \frac{SSt}{t-1}$	$F = \frac{MSt}{MSe}$
Experimental Error. الخطأ التجريبي	$\sum ri-t$	SSe = SST –SSt	$MSe = \frac{SSe}{\sum ri-t}$	
Total الكلي	$\sum ri-1$	$SST = \sum Yij^2 - CF$	-------------	

علما أن : معامل الاختلاف يحسب كما يلي في حالة عدم تساوي المكررات.

$$CF = \frac{(Y..)^2}{\sum ri}$$

حيث $\sum ri$: هو عدد المشاهدات (المكررات) في التجربة.

<u>مثال:</u> في تجربة شملت أربع معاملات أستعمل فيها فيتامين (E) لدراسة تأثير نسبة الفيتامين (0 , 5 , 10 , 15 %) في العليقة على معدل الزيادة الوزنية في الدجاج المحلي وتم الحصول على البيانات التالية:

المعاملات	الزيادة الوزنية (Yij)	المجموع .Yi	عدد المشاهدات (المكررات) ri
1	10 , 15 , 20 , 22	67	4
2	10 , 12 , 13 , --	35	3
3	7 , 7 , 8 , 10	32	4
4	14 , 12 , --- , ---	26	2
		المجموع الكلي ..Y = 160	$\sum ri = 13$

الحل:

يتم حساب معامل التصحيح أولا:

$$CF = \frac{(Y..)^2}{\sum ri} = \frac{(160)^2}{13} = 1969.6$$

ثم مجموع مربعات المعاملات (SSt): ((مهم جدا)).

$$SSt = \frac{\sum Yi.^2}{ri} - CF = \frac{(67)^2}{4} + \frac{(35)^2}{3} + \frac{(32)^2}{4} + \frac{(26)^2}{2} - 1969.6$$

$$SSt = 153.81$$

يتم حساب مجموع المربعات الكلية (SST):

$$SST = \sum Yij^2 - CF$$
$$SST = 10^2 + \text{----------------------} + 12^2 - 1969.6$$
$$SST = 261.23$$

يتم حساب مجموع مربعات الخطأ (SSe):

$$SSe = SST - SSt$$
$$SSe = 261.23 - 153.81$$
$$SSe = 107.41$$

ومن النتائج السابقة يمكن حساب متوسط مربعات كل من المعاملات والخطأ وكما يلي:

متوسط مربعات المعاملات (MSt):

$$MSt = \frac{SSt}{t-1} = \frac{153.81}{4-1} = \frac{153.81}{3} = 51.27$$

متوسط مربعات الخطأ (MSe):

$$MSe = \frac{SSe}{\sum ri-t} = \frac{107.41}{13-4} = \frac{107.41}{9} = 11.93$$

ومن خلال متوسط مربعات المعاملة والخطأ يمكن حساب قيمة F وكما يلي:

$$F = \frac{MSt}{MSe} = \frac{51.23}{11.93} = 4.30$$

ومن ثم يتم تكوين جدول تحليل التباين لتحليل البيانات

جدول تحليل التباين للتصميم :(**ANOVA Table**).

S.O.V. مصـــــادر الاختلاف	d.f. درجات الحرية	S.S. مجموع المربعات	M.S. متوسط المربعات	F. Value قيمـــة f المحسوبة
Treat. المعاملة	$t-1= 4-1 = 3$	$SSt = 153.81$	$MSt = 51.27$	$F = \frac{51.27}{11.9}$ $F = 4.30$
Experimental Error. الخطـــأ التجريبي	$\sum ri-t = 13-4 = 9$	$SSe = 107.41$	$MSe = 11.93$	
Total الكلي	$\sum ri-1 = 13-1 = 12$	$SST = 261.23$	-------------	

تقارن قيمة F المحسوبة (Calculated) وهي (4.30) مع قيمة F الجدولية (Tabulated) من جداول F وفق درجات حرية المعاملة (3) ودرجات حرية الخطأ (9) ، نجد أن F المحسوبة أعلى من الجدولية لذلك فان تأثير المعاملة (الفيتامين) على معدل الزيادة الوزنية فأذا كانت المحسوبة أعلى من الجدولية فأن تأثير المعاملة معنويا في الصفة المدروسة، وأذا كانت قيمة F المحسوبة أقل من الجدولية فأن تأثير المعاملة في نسبة الدهن غير معنوي (Non-significant) : ففي المثال السابق التأثير غير معمعنويا على مستوى ($P<0.05$).

سؤال واجب: من البيانات الموضحة في الجدول الاتي (لديك ثلاث معاملات عدد مكرراتها غير متساوية)، أوجد جدول تحليل التباين لغاية قيمة F

المعاملات	الزيادة الوزنية (Yij)	المجموع .Yi	عدد المشاهدات (المكررات) ri
1	5 , 9 , 10 , 11	39	4
2	6 , 4 , -- , ---	10	2
3	9 , 10 , 3 , 4	26	4
		المجموع الكلي ..Y = 75	$\sum ri = 10$

الاختبارات المقترحة بعد أجراء التجربة.

1- أختبار أقل فرق معنوي.

(Least Significant Difference – LSD):

يستعمل لمقارنة الفروق المعنوية بين أي متوسطين في التجربة.

خطوات تطبيق الاختبار :

أ- حساب الانحراف القياسي بين متوسط أي معاملتين في التجربة ، مما يلي

$$\text{الانحراف القياسي بين متوسط اي معاملتين} = \frac{2MSe}{r}$$

علما ان 2 ثابت كوننا نقارن بين متوسط كل معاملتين.

MSe : متوسط مربعات الخطأ (يتم الحصول عليه من جدول تحليل التباين).

r: عدد المشاهدات (المكررات) في كل معاملة.

ب- نستخرج قيمة t من جداول t (منشورة في نهاية أي كتاب لتصميم وتحليل التجارب).
على درجات حرية الخطأ فقط ومستوى معنوية 5 % أو 1 %.

ج- نستخرج قيمة LSD من حاصل ضرب الخطوتين السابقتين، أي وفق القانون الاتي:

$$LSD = \frac{2MSe}{r} \times t$$

د- نأخذ الفرق بين متوسطين أي معاملتين في التجربة ونقارنه مع قيمة LSD ، فأذا كان الفرق بين المتوسطين أعلى من الـ LSD فهو معنوي ونلاحظ مستوى المعنوية.

<u>مثال:</u> أجرية تجربة لدراسة تأثير خمسة أنواع من العلائق في معدل الزيادة الوزنية لدى العجول وقد شملت كل معاملة خمس عجول (البيانات موضحة في الجدول الاتي).

المعاملة	معدل الزيادة الوزنية (Yij)	المجموع (.Yi)	المتوسط
1	6 , 8 , 7 , 3 , 10	36	7.2
2	9 , 8 , 11 , 11 10	49	9.8
3	7 , 5 , 5, 9 , 4	30	6.0
4	5 , 3, 4 , 6 , 6	24	4.8
5	8 , 6 , 9 , 9 , 11	43	8.6

		المجموع الكلي 182 :(Y..)	

وبعد أجراء التحليل الاعتيادي للتجربة لغاية الحصول على جدول تحليل التباين (كما في الامثل السابقة) يكون جدول تحليل التباين كالاتي.

S.O.V. مصـــادر الاختلاف	d.f. درجات الحرية	S.S. مجموع المربعات	M.S. متوسط المربعات	F. Value قيمة f المحسوبة
Treat. المعاملة	4	79.44	19.86	6.90 **
Experimental Error. الخطأ التجريبي	20	57.60	2.88	
Total الكلي	24	137.04	----	

أجراء الاختبار (LSD) : بما أن

$$MSe = 2.88 \quad , \quad r = 5$$

وقيمة t من جداول t تساوي 2.08 لذلك :

$$LSD = \frac{2MSe}{r} \times t$$

$$LSD = \frac{2 \times 2.88}{5} \times 2.08 = 2.239$$

الان نأخذ الفرق بين متوسط كل معاملتين زنقارنه مع قيمة LSD ، فأذا كان الفرق بين المتوسطين معنوي نضع علهما حروف مختلفة ، واذا كان الفرق غير معنوي نضع عليها حروف موجبة.

مثلا الفرق بين متوسط المعاملتين t2 و t3 يكون

$$9.8 - 6.0 = 3.8$$

وبما أن 3.8 أكبر من 2.239 لذلك فأن الفرق معنوين بين متوسطي المعاملة الثانية والثالثة: وتوضع بالصيغة

متوسط المعاملة الثانية = a 9.8

متوسط المعاملة الثالثة = b 6.0

وكذلك بما أن الفرق بين متوسط المعاملة الاولى (7.2) ومتوسط المعاملة الثالثة (6.0) يساوي (1.2) أذن الفرق غير معنوي ، توضع بالصيغة :

متوسط المعاملة الاولى = a 7.2

متوسط المعاملة الثالثة = a 6.0

سؤال واجب: أذكر خطوات أجراء أختبار أقل فرق معنوي (LSD).

سؤال واجب: اجريت تجربة وفق تصميم عشوائي كامل (CRD) مع عدم تساوي المشاهدات أو المكر رات (ثلاث معاملات بمكررات مختلفة). المطلوب:

1- كتابة الانموذج الرياضي للتجربة مفسراً رموزه.

2- أيجاد جدول تحليل التباين للبيانات

والبيانات كما في الجدول الاتي:

المعاملات	المشاهدات yij	Yi. مجاميع المعاملات	Ri عدد المكررات
T1	5 , 6 , 3 , 4	18	4
T2	4 , 2 , -- , --	6	2
T3	9 , 6 , 7 , 8	30	4
		Y.. = 54	$\sum ri = 10$

2- أختبار دنكن (Duncan) متعدد الحدود.

وجد هذا الاختبار عام 1955 من قبل الباحث Duncan ويتميز عن باقي الاختبارات بأنه يأخذ الفروق المعنوية بين المتوسطات مهما كان عددها مرة واحدة .

خطوات أجراء الاختبار :

- يتم أستخراج الانحراف القياسي لاي مشاهدة في التجربة وفق الاتي.

جذر الاتي.

$$Syi = \frac{MSe}{r}$$

- أستخراج قيم SSR من جداول دنكن (موجودة في نهاية كتاب تصميم وتحليل التجارب) وحسب عدد المتوسطات الداخلة في المقارنة.
- أستخراج قيم LSR من المعادلة الاتية (حاصل ضرب الخطوتين السابقتين).

$$LSR = \frac{MSe}{r} X\ SSR$$

- يتم ترتيب المتوسطات وقيم LSR تنازليا وبشكل عمودي وكذلك ترتيب المتوسطات تصاعديا وبشكل أفقي وفي كلا الحالتين يترك أخر متوسط. بعد ذلك نأخذ الفرق بين كل متوسطين ونقارنه بقيمة LSR المقابلة لهما ، فأذا كانت قيمة الفرق بين المتوسطين أعلى من قيمة LSR أذن الفرق بين المتوسطين معنوي ، في حين أذا كان الفرق أقل من الـ LSR فهو غير معنوي. وتوضع حروف على المتوسطات كما تم توضيح ذلك أنفا في أختبار LSD.

مثال: تطبيق أختبار دنكن على نفس المثال السابق الذي طبق عليه أختبار LSD.

$$Syi = \frac{MSe}{r} = \frac{2.88}{5} = 0.759$$

	عدد المتوسطات الداخلة في المقارنة			
	2	3	4	5
SSR	2.95	3.09	3.19	3.25
$\frac{MSe}{r}$	0.759			
LSR	2.33	2.35	2.42	2.47

قيم LSR في الجدول ناتجة من ضرب قي SSR في 0.759 .

ولغرض أجراء المقارنة نكون الجدول الاتي:

متوسط المعاملات تنازليا	قيم LSR تنازليا	T4 4.8	T3 6.0	T1 7.2	T5 8.6
T2 : 9.8	2.47	5.0*	3.8*	2.6*	1.2NS
T3 : 8.6	2.42	3.8*	2.6*	1.4	
T1 : 7.2	2.35	2.4*	1.2NS		
T3 : 6.0	2.33	1.2NS			

فمثلا الفرق بين متوسط المعاملة الثانية (9.3) والمعاملة الرابعة (4.8) هو (5.0) كما موضح في الجدول وهذه القيمة أعلى من قيمة LSR المقابلة لها (2.47) لذلك الفرق بين متوسطي المعاملتين الثانية والرابعة معنوي لذا وضعت الاشارة * ولهذا يعطى المتوسط الاعلى a والاقل b .