

ثانيا: - مقياس التشتت او الاختلاف of dispersion or variation Measure :

تتصف الظواهر الحياتية بشكل عام بتفاوت قيمتها من مفردة الى اخرى داخل نفس المجموعة ومن مجموعها لأخرى. أما درجة التفاوت او التشتت بين القيم فأذا تعتمد على طبيعة المتغير والمفردات المشمولة بالدراسة. ومقاييس التشتت هي مقاييس لمدى تشتت قيم المشاهدات عن وسطها.

فعلى سبيل المثال لو اخذنا المجموعتين التاليتين من الارقام

المجموعة الاولى / 6 6 6 6 6 6 6 6 6 كغم

المجموعة الثانية / 2 3 4 5 6 4 4 20 كغم

فإن المتوسط الحسابي للمجموعتين = 6

لذا فان المتوسط الحسابي بمفرده لا يمكنه ان يشخص او يميز بين المجموعات وعليه فنحن بحاجة الى مقياس اخر ولكي نميز المجموعات تمييز قاطع يجب ان نأخذ مقياسين احدهما مقياس نزعه مركزية والاخر مقياس تشتت وهنالك مقاييس للتشتت اهمها: -

● **مقاييس التشتت المطلق:** - اي التي تكون وحداتها نفس وحدات القيم الاصلية واهمها:

● **المدى** the range

● **التباين** variance

● **الانحراف القياسي (المعياري)** standard deviation

● **مقاييس التشتت النسبي:** - وهذه تكون خالية وحدات القياس واهمها:

● **معامل الاختلاف** coefficient of variation

1- **المدى the range:** هو الفرق بين اعلى قيمة وادنى قيمة في مجموعته من القيم ويرمز له R وقانونه

$$\text{هو: } Y_{\min} - Y_{\max} = R$$

مثال: -

$$(a) y_i = 12, 6, 7, 3, 15, 10, 18, 5$$

$$(b) y_i = 9, 3, 8, 8, 9, 8, 9, 18$$

الحل: نلاحظ ان المدى بعد تطبيق القانون في كلا المجموعتين متساو ولكننا نلاحظ حقيقة ان الاختلاف في

المجموعة الاولى (a) أكبر منه في المجموعة (b)

$$a) R = 18 - 3 = 15$$

$$b)R=18-3=15)$$

لأن قيمة المجموعة (b) تتألف معظمها من 8 و 9 . لذلك فإن المدى يكون احيانا مظلا لأنه يعتمد فقط على القيمتين الطرفيتين ((بعد ترتيب القيم تصاعديا او تنازليا)) اللتين كثيرا ما تكونان شاذتين لكن هذا القياس يتميز بسهولة حسابه واعطائه فكره سريعة ومبسطة عن درجه تشتت قيم المجموعة.

ومن عيوبه انه يهمل جميع قيم المجموعة عدا القيمتين العليا والدنيا، ولهذا السبب فإن هذا المقياس محدود الاستخدام مقارنة بالمقاييس الاخرى.

مثال: -

من خلال استخراج المدى كقياس تشتت نلاحظ انه في الحالة الاولى لا يوجد تشتت بينما في الحالة الثانية يوجد تشتت عالي.

المدى = هو أكبر فرق يمكن الحصول عليه بين البيانات

ملاحظة: من الصعب حساب المدى الحقيقي من جدول توزيع تكراري لعدم معرفة لقيمتي الطرفين

(a)

6	6	6	6	6	6	6	6
---	---	---	---	---	---	---	---

2	3	4	5	6	4	4	20
---	---	---	---	---	---	---	----

(b)

$$a)R= 6 - 6 = 0$$

$$B)R = 20 - 2 = 18$$

2- التباين Variance

لكي يمكننا التغلب على مشكلة الاشارات عند جمع الانحرافات والتي تؤدي دائما لأن يكون مجموع الانحرافات لأي عينة عن وسطها الحسابي يساوي صفرا. نقوم بتربيع قيم الانحرافات وبذلك تصبح جميعها موجبة اي نحصل على مجموع مربعات الانحرافات sum of squares والتي يرمز لها (ss) وعلى ذلك فإن :-

$$ss = \sum (y_i - \bar{Y})^2$$

التباين: - هو مقياس لمعرفة درجة التباين بين القيم اي اختلاف القيم عن بعضها.

والتباين يشتق على اساس تحديد الفروق بين كل قيمة والمتوسط الحسابي لكل مجموعة.

ولكي نأخذ في الاعتبار حجم العينة حتى يمكن مقارنة العينات المختلفة الاحجام فأنتنا نقسم مجموع مربعات الانحرافات على درجات الحرية (n-1) وبذلك نحصل على ما يسمى بالتباين (s^2) ويمكن التعبير عن ذلك رياضيا على النحو التالي:

$$s^2 = \frac{\sum(y_i - \bar{Y})^2}{n-1}$$

تباين العينة

ويمكن اعادة كتابة المعادلة اعلاه بشكل اخر مطابق لها جبريا اضافة الى انه يسهل استخدام الآلات الحاسبة في هذا المجال وهذه الصيغة مبينة في المعادلة:

$$s^2 = \frac{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}{n-1} \dots \dots \dots \text{(هي الطريقة المختصرة)}$$

نقسم القيمة كلها على (n-1) بدلا من n لأنه يعطي قيمة اكبر لأن (n-1) اصغر من n و (n-1) يعطي درجة حدية للقيم اكبر .

* ويلاحظ ان القانون السابق هو لحساب تباين العينة اما اذا كانت قيم المشاهدات تمثل المجتمع كله فان التباين يرمز له في هذه الحالة $\sigma^2 = \frac{\sum(y_i - \mu)^2}{N}$ وتلفظ sigma square

حيث ان : M = الوسط الحسابي للمجتمع ----- N = عدد مفردات المجتمع

او:

$$\sigma^2 = \frac{\sum y_i^2 - \frac{(\sum y_i)^2}{N}}{N}$$

يلاحظ بأن التباين يقيس التشتت بوحدات مربعة مثل او وهكذا حسب طبيعة البيانات وغيرها والحل لهذه الحالة ولكي نرجع (وحدات القياس) الى أصلها فأنتنا نأخذ الجذر التربيعي للتباين لنحصل على قيمة (S) اي ان وهو ما يسمى بالانحراف القياسي او المعياري والذي يكون مقاسا بالوحدات الاصلية في الحالات السابقة بال (سم) و (كغم) او الدينار او العامل وهكذا.

الانحراف القياسي (المعياري) standard deviation :

$$\sigma x = \sqrt{\sigma x^2} = \sqrt{\frac{\sum(y_i - \mu)^2}{N}}$$

عبارة عن الجذر التربيعي للتباين

الحالة الاولى الانحراف المعياري للمجتمع (σx)

$$Sx = \sqrt{Sx^2}$$

الحالة الثانية الانحراف المعياري

العينة (Sx)

$$Sx = \sqrt{\frac{\sum(y_i - \bar{Y})^2}{n-1}} = \sqrt{\frac{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}{n-1}}$$

مثال: جد قيمة التباين للقيم التالية التي تمثل وزن اللحم الصافي ل(6) رؤوس من الغنم تمثل عينة :-

32 , 37 , 38 , 40 , 34 , 35

$$\bar{Y}_i = \frac{\sum y_i}{n} = \frac{32 + 37 + \dots + 35}{6} = 36 \text{ كغم}$$

$$S_y^2 = \frac{42}{6-1} = \frac{42}{5} = 8.4 \text{ كغم}^2$$

$(y_i - \bar{Y})^2$	الفرق $y_i - \bar{Y}$	الوزن y_i
16	$32 - 36 = -4$	32
1	$= +1$	37
4	$= +2$	38
16	$= +4$	40
4	$= -2$	34
1	$= -1$	35
$\sum (y_i - \bar{Y})^2 = 42$	0	

اما الانحراف المعياري لهذا المثال هو : $S_y = \sqrt{S_y^2} = \sqrt{8.4} = 2.89 \text{ كغم}$

(أ) بيانات غير مبوبة :-
مثال: البيانات التالية تمثل كمية المحصول للقطعة (كغم) لمحصول القطن في خمسة مزارع احسب الانحراف المعياري القياسي لها: - $Y_i = 9, 8, 6, 5, 7$
الحل: - 1- طريقة حساب الوسط الحسابي او طريقة التعريف وتسمى احيانا بالطريقة المطولة.

y_i	$Y_i - \bar{Y}$	$(y_i - \bar{Y})^2$
9	+2	4
8	+1	1
6	-1	1
5	-2	4
7	0	0
$\sum y_i = 35$ $\bar{Y} = 7$	0	10

$$\bar{Y} = \frac{\sum y_i}{n} = \frac{9 + \dots + 7}{5} = \frac{35}{5} = 7$$

$$\bar{Y} = 7$$

$$S = \sqrt{\frac{\sum (y_i - \bar{Y})^2}{n-1}} = \sqrt{\frac{10}{4}}$$

$$S = \sqrt{2.5} = 1.58 \text{ kg} \setminus m^2$$

(1) طريقة تربيع البيانات (وتسمى بالطريقة المختصرة)

Ss=sum of squares

S= الانحراف القياسي

$$S = \sqrt{\frac{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}{n-1}} = \sqrt{\frac{255 - \frac{(35)^2}{5}}{5-1}}$$

$$S = \sqrt{\frac{255 - \frac{1225}{5}}{4}} = \sqrt{\frac{255 - 245}{4}}$$

$$S = \sqrt{\frac{10}{4}} = \sqrt{2.5} = 1.58 \text{ Kg/m}^2$$

y_i	y_i^2
9	81
8	64
6	36
5	25
7	49
$\sum y_i = 35$	$\sum y_i^2 = 255$

*اما التباين لهذه القيم فهو مربع الانحراف القياسي اي نرفع الجذر التربيعي

$$S^2 = \frac{10}{4} = 2.5 \text{ (kg/m}^2)^2$$

(ب)البيانات المبوبة :

اذا كانت Y_1, Y_2, \dots, Y_n تمثل مراكز الفئات في جدول التوزيع التكراري وان تكراراتها هي f_1, f_2, \dots, f_n على التوالي فان الانحراف القياسي لها هو:

$$S = \sqrt{\frac{\sum f_i(y_i - \bar{Y})^2}{\sum f_i - 1}} = \sqrt{\frac{\sum f_i y_i^2 - \frac{(\sum f_i y_i)^2}{\sum f_i}}{\sum f_i - 1}}$$

مثال: - احسب الانحراف القياسي والتباين لجدول التوزيع التكراري التالي:

1- طريقة حساب الوسط الحسابي او طريقة التعريف (وتسمى احيانا الطريقة المطولة)

الفئات	F_i	Y_i	$f_i y_i$	$y_i - \bar{Y}$	$(y_i - \bar{Y})^2$	$f_i(y_i - \bar{Y})^2$
60-62	5	61	305	- 6.45	41.6025	208.0125
63-65	18	64	1152	- 3.45	11.9025	214.245
66-68	42	67	2814	- 0.45	0.2025	8.505
69-71	27	70	1890	+ 2.55	6.5025	175.5675
72-74	8	73	584	+ 5.55	30.8025	246.42
	100		$\sum f_i y_i = 6745$			$\sum f_i(y_i - \bar{Y})^2 = 852.75$

$$\bar{Y} = \frac{\sum f_i y_i}{\sum f_i} = \frac{6745}{100} = 67.45$$

$$\bar{Y} = 67.45$$

مجموع المربعات ss هو $ss = \sum f_i (y_i - \bar{Y})^2 = 852.7500$

$$s^2 = \frac{ss}{\sum f_i - 1} = \frac{852.75}{99} = 8.6 \quad \text{اما التباين فهو}$$

$$s = \sqrt{s^2} = \sqrt{8.6} = 2.93 \quad \text{اما الانحراف القياسي فهو}$$

(2) طريقة التربيع (وتسمى بالطريقة المختصرة)

الفئات	F_i	Y_i	$F_i y_i$	y_i^2	$f_i y_i^2$
60-62	5	61	305	3721	18605
63-65	18	64	1152	4096	73728
66-68	42	67	2814	4489	188538
69-71	27	70	1890	4900	132300
72-74	8	73	584	5329	42632
	100		6745		$\sum f_i y_i^2$

$$ss = \sum f_i y_i^2 - \frac{(\sum f_i y_i)^2}{\sum f_i}$$

$$s^2 = \frac{ss}{\sum f_i - 1} = \frac{\sum f_i y_i^2 - \frac{(\sum f_i y_i)^2}{\sum f_i}}{\sum f_i - 1} = \frac{455803 - \frac{(6745)^2}{100}}{100 - 1}$$

$$s^2 = \frac{455803 - 454950.25}{99} = \frac{852.75}{99} = 8.6$$

$$s = \sqrt{s^2} = \sqrt{8.6} = 2.9 \quad \text{الانحراف القياسي}$$

مقاييس التشتت النسبي

ان مقاييس التشتت النسبي لها اهميتها عند مقارنة تشتت مجموعتين او أكثر تختلف في وحدات القياس لقيمتها. لان مقاييس التشتت النسبي تكون خالية من وحدات القياس واهم مقاييس التشتت النسبي هي:

معامل الاختلاف Coefficient of variation:

اذا كان S و \bar{Y} هما الانحراف القياسي والوسط الحسابي لمجموعه من القيم على التوالي فان معامل الاختلاف يرمز له **C.V.**

$$C.V. = \frac{S}{\bar{Y}} \times 100 \quad \text{ويحسب وفق المعادلة التالية :-}$$

مثال: - نتائج الامتحان النهائي لدرسي الاحصاء والكيمياء للصف الاول كانتا كالاتي:

ففي اي الموضوعين كان لتشتت الدرجات أكثر؟

$$c.v. = \frac{S}{\bar{Y}} \times 100$$

الاحصاء	الكيمياء	
78	73	
8	7.6	
الاحصاء $= \frac{8}{78} \times 100 = 10.25\%$		
الكيمياء $= \frac{7.6}{73} \times 100 = 10.41\%$		

اي ان التشتت لدرجات الكيمياء كان أكثر.

لاحظ بان لو قارنا التشتت بمقياس الانحراف القياسي مكان التشتت في الاحصاء اكبر منه في الكيمياء.

مثال: - اجريت تجربة لدراسة طول (سم) وكمية المحصول (كغم) ل (150) نبات من الذرة فكانت النتائج كالاتي:
- قارن بين تشتت الصفتين

كمية المحصول	الطول	
800	200	الوسط الحسابي \bar{Y}
36	16	الانحراف القياسي s

$$\text{صفة الطول } c.v. = \frac{S}{\bar{Y}} = \frac{16}{200} \times 100 = 8\%$$

$$\text{صفة كمية المحصول } C.V. = \frac{36}{800} \times 100 = 4.5\%$$

التشتت كان أكبر في صفة الطول من صفة كمية المحصول.