مقاييس التشتت Measures of Variation

مقدمة-:

سبق ان تعرفنا على خصائص الظاهرة وهي مقاييس النزعة المركزية مثل(المتوسطات) إلا ان هذه المقاييس ليست كافية احيانا لوصف الظاهرة ومقارنتها مع ظواهر أخرى , حيث ان بعض الظواهر قد تتشابها في اوساطها ولكن قد تختلف في تباعد او تقارب بياناتها عن اوساطها مثال-:

نفترض لدينا مجموعتين لمادة الاحصاء ودرجات الطلاب كالتالي-:

المجموعة الاولى: (76 , 64 , 71 , 70 , 69)

المجموعة الثانية : (60 , 100 , 80 , 40 , 70)

ولو حسبنا الوسط الحسابي للمجموعتين فان:

$$\text{الوسط الحسابي للمجموعة الاولى} = \frac{69 + 70 + 71 + 64 + 76}{5} = 70$$

$$\text{الوسط الحسابي للمجموعة الثانية} = \frac{70 + 40 + 80 + 100 + 60}{5} = 70$$

الوسط الحسابي للمجموعة الاولى= 70

الوسط الحسابي للمجموعة الثانية= 70

ونلاحظ ان الوسطين متساويان = 70 درجة ولو اكتفينا بمقارنة الوسطين للظاهرة نستنتج ان مستوى الطلاب واحد او ان المجموعتين متجانسة ، ولكن هذا يخالف واقع البيانات ، لان درجات طلاب المجموعة الاولى تظهر متقاربة من بعضها البعض وتتركز حول وسطها الحسابي ، بينما درجات الطلاب في المجموعة الثانية تظهر متباعدة اكثر من بعضها ومنتشرة في مدى اوسع .ولذلك نحن بحاجة الى مقاييس اخرى غير مقاييس النزعة المركزية تقيس مدى تقارب او تباعد مفردات الظواهر بعضها عن بعض ، ولمعرفة مدى تجانس الظواهر في خصائصها , هذه المقاييس تسمى مقاييس التشتت.

نستنتج مما سبق ان الهدف من دراسة التشتت هو تكوين فكرة عن مدى تجانس قيم مجموعة من المفردات وهذا يعني ان دراسة التشتت امر مفيد في اجراء المقارنة بين قيم مجموعتين او اكثر من البيانات عن ظاهرة معينة.

لقد ورد مصطلح (التشتت) في الادبيات الاحصائية بكلمة (التبعثر) اي انتشار قيم مجموعة من المفردات.

تعريف مقاييس التشتت :- هي المقاييس التي تقيس مدى تباعد القيم او تقاربها والتي تستعمل كمؤشر احصائي لتحديد درجة التقارب او التشتت .

مثال: كانت درجات الحرارة في ثلاثة محافظات هي كما يلي:

المحافظة1 : 4 , 3 , 2 ,0 ,1- ,2-

المحافظة2 : 7 , 4 ,2 ,0, 1- ,2- ,3-

المحافظة3: 3 , 2 , 2 ,1 ,0 ,0 ,1-

في اي محافظة كان الجو اكثر استقراراً.

الحل/

من خلال استقراء البيانات لهذه المجموعات نلاحظ ان متوسط درجة الحرارة في كل محافظة هو درجة مئوية واحدة وهذا غير نافع في تحديد المطلوب.إلا انه يلاحظ في المحافظة الثالثة كان الجو اكثر استقراراً من بقية المحافظات وذلك بسبب تجانس قيم درجات الحرارة فيها اكثر من تجانسها في بقية المحافظات , اي ان تشتت القيم فيها كان اقل من بقية المحافظات.

ومقاييس التشتت على نوعين رئيسيين هي :

أ- مقاييس التشتت المطلق : وهي تبين درجة تجانس قيم مجموعة من البيانات بشكل مطلق وتكون مقاسة بنفس وحدات قياس المتغير العشوائي(وحدات الطول ,الوزن,زمن,...الخ)

ب- مقاييس التشتت النسبي : وهي تبين درجة تجانس قيم مجموعة من البيانات بشكل نسبي وتكون خالية من وحدات قياس المتغير العشوائي.

أ- مقاييس التشتت المطلق :

1- المدى Range :- يسمى احيانا بمجال التغير وهو من ابسط مقاييس التشتت المطلق ويعرف بأنه الفرق بين اعلى قيمة واقل قيمة من مجموعة البيانات غير المبوبة.

ويمكن حساب المدى للبيانات كالتالي :

أ- في حالة البيانات الغير مبوبة: المدى R = اعلى قيمة – اقل قيمة

مثال // اوجد المدى للمجموعتان التاليتان والتي تمثل درجات طلبة قسم الادارة في مادة الاحصاء ؟

المجموعة A= 70 , 60 , 50 , 40 , 30

المجموعة B = 52 , 51 , 50 , 49 , 48

R.A = 70 – 30 = 40

R.B = 52 – 48 = 4

المجموعة A تشتتها اكبر من المجموعة B اي المجموعة Bاكثر تجانس من A.

ب- في حالة البيانات المبوبة :

المدى R = الحد الاعلى لأكبر الفئات – الحد الأدنى لأصغر الفئات

مثال : اوجد المدى لجدول التوزيع التكراري التالي:

الفئات	8-16	16-24	24-32	32-40	40-48	المجموع
التكرارات	20	25	8	6	3	62

الحل

$$R = 48 - 8 = 40$$

لا يستخدم المدى كثيرا لأنه يستند الى قيمتين الاولى والأخيرة ويهمل باقي القيم وهذا يعني انه مقياس حساس جدا لأي خطأ قد يحصل في قياس احدى هاتين القيمتين او كليهما كما لا يمكن حسابه في حالة الجداول التكرارية المفتوحة.

مزايا المدى:

1-أبسط وأسهل طريقة لحساب التشتت

2-مقياس سريع لمدى التشتت المفردات أو حينما يكون للمفردات المتطرفة أهمية خاصة.

عيوب المدى:

1-ليس للمدى أهمية كبيرة في البحوث العلمية نظراً لأنه لا يأخذ في الاعتبار تشتت كل المفردات في حسابه.

2-مقياس تقريبي غير دقيق

3-يتأثر تأثراً كبيراً بالقيم المتطرفة

4-يصعب تقدير قيمته من الجداول التكرارية المفتوحة.

2- الانحراف المتوسط Mean Deviation:

يعرف الانحراف المتوسط بأنه معدل مجموع انحرافات القيم المطلقة عن متوسطها.

أ- الانحراف المتوسط لبيانات غير مبوبة :

$$M.D. = \frac{\sum |yi - \bar{y}|}{n}$$

مثال/ اوجد الانحراف المتوسط للبيانات التالية : yi = 11 , 12 ,13 , 12 , 13 , 11

$$\bar{y} = \frac{11 + 12 + 13 + 12 + 13 + 11}{6} = \frac{72}{6} = 12$$

$\|yi - \bar{y}\|$	yi
1	11

12	0
13	1
12	0
13	1
11	1
المجموع	4

الحل:-

$$\boldsymbol{M.D.} = \frac{\sum |\boldsymbol{yi} - \bar{\boldsymbol{y}}|}{\boldsymbol{n}}$$

$$\boldsymbol{M.D.} = \frac{\mathbf{4}}{\mathbf{6}} = \mathbf{0.67}$$

ب – الانحراف المتوسط لبيانات مبوبة :

$$\boldsymbol{M.D.} = \frac{\sum \boldsymbol{fi}\,|\boldsymbol{yi} - \bar{\boldsymbol{y}}|}{\sum \boldsymbol{fi}}$$

مثال / اوجد الانحراف المتوسط للبيانات التالية التي تمثل توزيع طلبة كلية الصيدلة حسب الوزن .

الفئات	التكرار *fi*	مركز الفئات *yi*	yif_i	$\lvert yi - \bar{y} \rvert$	$fi\lvert yi - \bar{y} \rvert$
60 – 62	5	61	305	*6.54*	32.7
63- 65	15	64	960	*3.54*	53.1
66 – 68	45	67	3015	*0.54*	24.3

69 – 71	**27**	**70**	**1890**	***2.46***	**66.42**
74 - 72	**8**	**73**	**584**	***5.46***	**43.68**
المجموغ	***100***		**6754**		**220.2**

$$M.D. = \frac{\sum fi \left| yi - \bar{y} \right|}{\sum fi}$$

$\bar{y} = \frac{\sum fiyi}{\sum fi} = \frac{6754}{100} = 67.54$

$$M.D. = \frac{\sum fi \left| yi - \bar{y} \right|}{\sum fi} = \frac{220.2}{100} = 2.202$$

<u>ميزاته:</u>

1- حسابه يعتمد على كافة البيانات المتاحة.

2- مقياس سهل الفهم والحساب.

3- خضوعه للعمليات الجبرية.

<u>عيوبه:</u>

1- اهمال الاشارات السالبة للفروق عند عملية حسابه

2- تتأثر قيمته في حالة وجود القيم المتطرفة والشاذة

-3 لا يمكن حساب قيمته من الجداول التكرارية المفتوحة .

4- لا يمكن حسابه في حالة البيانات الوصفية.

5-يتأثر وعلى نحو كبير بأخطاء المعاينة.