المحور الخامس: مقاييس التشتت

The fifth axis: Measures of dispersion

محتوى المحاضرة الثامنة

أولا: مقاييس التشتت المطلق(measures of Absolute Dispersion)

I. **المدى العام(Range)**

II. **المدى الربيعي(Interquartile Range)**

III. **الانحراف المتوسط (Mean Deviation)**

IV. **التباين و الانحراف المعياري(Variance and Standard Deviation)**

ثانيا : مقاييس التشتت النسبي(measures of Relative Dispersion)

معامل الإختلاف(Coefficient of Variation)

تمهيد:

التشتت هو مقدار الاختلاف في مفردات ظاهرة ما، وهو يعبر عن مدى بعد البيانات عن وسطها الحسابي، فإذا كانت البيانات بعيدة عن وسطها الحسابي سنجد التشتت كبير وبالتالي بيانات الظاهرة المدروسة غير متجانسة، والعكس صحيح. ولقياس ظاهرة التشتت تستعمل عدة مقاييس مطلقة ونسبية، أبرزها: المدى العام، المدى الربيعي، الإنحراف المتوسط، الإنحراف المعياري، ومعامل الإختلاف.

أولا: مقاييس التشتت المطلق(measures of Absolute Dispersion)

I. المدى العام(Range):

وهو الفرق بين أكبر قيمة وأقل قيمة أو الفرق بين الحد الاعلى للفئة الاخيرة والحد الادنى للفئة الأولى، ويرمز له بالرمز R، حيث:

$$R = X_{max} - X_{min}$$

يتميز هذا المقياس بإعطاء نظرة سريعة عن درجة تشتت البيانات من خلال التركيز على القيمتين الطرفيتين فقط، فكلما زاد مقداره دل ذلك على تشتت أكبر و كلما قل دل على ميل البيانات إلى التجانس، و هو شديد التأثر بالقيم المتطرفة.

II. المدى الربيعي(Interquartile Range):

وهو الفرق بين الربيعي الأول والربيعي الثالث، ويرمز له بالرمز IQ_R، حيث:

$$IQ_R = Q_3 - Q_1$$

يركز هذا المقياس على مجال طوله (% 50) من البيانات المتمركزة في الوسط وهي خاصية مفيدة في حالة التوزيعات شديدة الإلتواء.

مثال1:

تمثل البيانات التالية المبيعات الشهرية بالدينار لمؤسسة ما:

114950; 158000; 230500; 387000; 389950; 479000; 488800; 529000; 575000; 639000; 659000; 1095000; 2500000

1- احسب المدى الربيعي ؟

الحل:

1- حساب المدى الربيعي:

رتبة $Q_1 = \frac{N+1}{4} = \frac{13+1}{4} = 3,5 \Rightarrow Q_1 = \frac{X_3+X_4}{2} \Rightarrow Q_1 = \frac{230500+387000}{2} = 308750$

رتبة $Q_3 = \frac{3(N+1)}{4} = \frac{3(14)}{4} = 10,5 \Rightarrow Q_3 = \frac{X_{10}+X_{11}}{2} \Rightarrow Q_3 = \frac{639000+659000}{2} = 649000$

$IQ_R = Q_3 - Q_1 = 649000 - 308750 \Rightarrow IQ_R = 340250$

III. **الانحراف المتوسط (Mean Deviation):**

يعرف بأنه متوسط الإنحرافات المطلقة للقيم عن وسطها الحسابي ويرمز له بالرمز D_x، ويحسب كما يلي:

أ- **في حالة بيانات غير مبوبة:** $D_x = \frac{\sum|(x_i - \bar{x})|}{n}$

ب- **في حالة بيانات مبوبة:**

- **بيانات متقطعة:** $D_x = \frac{\sum f_i|(x_i - \bar{x})|}{N}$

- **بيانات مستمرة:** $D_x = \frac{\sum f_i|(m_i - \bar{x})|}{N}$

IV. **التباين و الانحراف المعياري(Variance and Standard Deviation):**

1- **التباين:**

هو متوسط مجموع مربع انحرافات القيم عن وسطها الحسابي، ويدل على مدى تشتت قيم المتغير الاحصائي حول احد مقاييس النزعة المركزية(الوسط الحسابي)، ويرمز له بالرمز V_x، ويحسب كما يلي:

أ- **في حالة بيانات غير مبوبة:** $V_x = \frac{\sum(x_i - \bar{x})^2}{n}$

ب- **في حالة بيانات مبوبة:**

- **بيانات متقطعة:**

$$V_x = \frac{\sum f_i(X_i - \bar{x})^2}{N} \quad \text{أو} \quad V_x = \frac{\sum f_i(X_i)^2}{N} - \bar{X}^2$$

- **بيانات مستمرة:**

$$V_x = \frac{\sum f_i(m_i - \bar{x})^2}{N} \quad \text{أو} \quad V_x = \frac{\sum f_i(m_i)^2}{N} - \bar{X}^2$$

2- **الإنحراف المعياري(Standard Deviation):**

وهو أكثر مقاييس التشتت استخداما، و هو الجذر التربيعي الموجب للتباين، ويرمز له بالرمز S_X، ويستخدم للمقارنة بين متغيرين من نفس الطبيعة، ويحسب كما يلي:

$$S_X = \sqrt{V_x}$$

مثال2:

يمثل الجدول التالي توزيع 40 طالب حسب عدد الغيابات:

x_i (عدد الغيابات)	0	1	2	3	4	5	$\sum$
عدد التلاميذ	9	10	7	8	4	2	40

1- أحسب المتوسط الحسابي؟

2- أحسب المدى الربيعي؟

3- أحسب الإنحراف المتوسط؟

4- أحسب الإنحراف المعياري؟

الحل:

1- **حساب المتوسط الحسابي:**

x_i (عدد الغيابات)	f_i (عدد الطلاب)	f_iX_i	$cf_i\uparrow$	$(X_i - \bar{x})$	$f_i\|(x_i - \bar{x})\|$	$(X_i - \bar{x})^2$	$f_i(X_i - \bar{x})^2$
0	9	0	9	0-1.85= -1.85	16.65	3.4225	30.8025
1	10	10	19	1-1.85= -0.85	8.5	0.7225	7.725
2	7	14	26	0.15	1.05	0.0225	0.1575
3	8	24	34	1.15	9.2	1.3225	10.58
4	4	16	38	2.15	8.6	4.6225	18.49
5	2	10	40	3.15	6.3	9.9225	49.6125
$\sum$	40	74	/	/	50.3	/	117.3675

$$\bar{X} = \frac{\sum_{i=1}^{6} f_i X_i}{\sum_{i=1}^{6} f_i} = \frac{74}{40} = 1.85 \approx 2$$

متوسط عدد غيابات الطلبة يساوي2

2- **حساب المدى الربيعي:** $IQ_R = Q_3 - Q_1$

- **تحديد رتبة الربيعي الأول:**

$$\text{رتبة } Q_1 = \frac{N}{4} = \frac{40}{4} = 10$$

نبحث عن هذه القيمة في عمود $cf_i\uparrow$ ، وإذا لم نجدها نأخذ القيمة الأكبر مباشرة وهي 19، ثم نختار القيمة التي تقابلها في عمود x_i ، ومنه $Q_1 = 1$

- تحديد رتبة الربيعي الثالث:

$$\text{رتبة } Q_3 = \frac{3N}{4} = \frac{3(40)}{4} = 30$$

نبحث عن هذه القيمة في عمود $cf_i \uparrow$ ، وإذا لم نجدها نأخذ القيمة التي تليها مباشرة وهي 34، ثم نختار القيمة التي تقابلها في عمود x_i ، ومنه $Q_3 = 3$

_ حساب المدى الربيعي:

$$IQ_R = Q_3 - Q_1 = 3 - 1 \Rightarrow IQ_R = 2$$

أي أن 50% من البيانات منتشرة في مجال قدره 2

3- حساب الإنحراف المتوسط:

$$D_x = \frac{\sum f_i |(x_i - \bar{x})|}{N} = \frac{50.3}{40} \Rightarrow D_x = 1.25$$

أي أن متوسط بعد البيانات عن الوسط الحسابي هو 1.25

4- حساب الإنحراف المعياري:

$$S_X = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{N}} = \sqrt{\frac{117.3675}{40}} \Rightarrow S_X = 1.71$$

II- مقاييس التشتت النسبي(measures of Relative Dispersion):

معامل الإختلاف(Coefficient of Variation):

يستخدم معامل الاختلاف للمقارنة بين توزيعين أو أكثر إذا ما اختلفت وحدات القياس(مثلا أرباح المؤسسة A بالدينار، و أرباح المؤسسة B بالأورو)، أو اختلفت متوسطاتها، وذلك لأنه معامل مجرد من الوحدة، ويعبر عنه بشكل عام بالنسبة المئوية. ويرمز له بالرمز C_v، ويحسب كما يلي:

$$C_v(\%) = \frac{\delta_x}{\bar{x}} . 100$$

تشتت معتبر $Cv\% < 30$

تشتت ضعيف $30 \leq Cv\% < 50$

تشتت كبير $50 \leq Cv\% < 80$

تشتت كبير جدا $80 \leq Cv\% \leq 100$

مثال3:

يقارن مستثمر بين أرباح سهمين حيث:

	متوسط الربح السنوي	الإنحراف المعياري لبيانات الربح
السهم A	30	2
السهمB	19	1.5

المطلوب: أي السهمين يختار هذا المستثمر؟

الحل:

من خلال المعطيات نلاحظ أن الإنحراف المعياري للسهم (A) أكبر من نظيره للسهم(B) و لكن نظرا لعدم تساوي الوسط الحسابي لا تصلح هذه المقارنة، لذلك نحسب معامل الإختلاف لكل خيار:

$$C_{v_A} = \frac{\delta_x}{\bar{x}} . 100 = \frac{2}{30} . 100 = 6.66\ \%$$

$$C_{v_B} = \frac{\delta_x}{\bar{x}} . 100 = \frac{1.5}{19} . 100 = 7.89\ \%$$

نلاحظ أن معامل الإختلاف يكشف عكس النتيجة الأولى، أي أن أرباح السهم(B) أكثر تشتتا، وعليه سيختار المستثمر السهم A(الذي أرباحه أقل تشتتا)