(مقاييس التشتت)

نقصد بالتشتت التباعد بين القيم او التفاوت بينها وكلما كان التباعد كثيراً كان مقدار التشتت كبيراً.

وسيشمل التشتت على ما يلي :-

1. المدى
2. الانحراف الربيعي
3. التباين
4. الانحراف المعياري

1. المدى (Range)

عبارة عن الفرق بين اعلى مشاهدة واقل مشاهدة في حالة البيانات غير المبوبة ويرمز له بالرمز R

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

اما في حالة البيانات المبوبة فان المدى :-

$$R = \text{الحد الاعلى الحقيقي للفئة الاخيرة} - \text{الحد الادنى الحقيقي للفئة الاولى}$$

ملاحظة :- لا نعتمد دائماً على المدى في حساب التشتت لانه يكون احياناً مضللاً بسبب اعتماده على اعلى مشاهدة واقل مشاهدة .

② الانحراف الربيعي (Quartile of Deviation)

- يحسب الانحراف الربيعي من القانون الآتي :-
(في حالة البيانات غير المبوبة)

$$Q = \frac{Q_3 - Q_1}{2}$$

اذ ان :-

Q_1 : يمثل الربيع الاول (الادنى)

Q_3 : يمثل الربيع الثالث (الاعلى)

الربيع الاول Q_1 : هو المشاهدة التي يسبقها في الترتيب ربع المشاهدة ويليها ثلاثة ارباع المشاهدة على فرض ان هذه البيانات مرتبة ترتيباً تصاعدياً.

الربيع الثالث Q_3 : هو المشاهدة التي يسبقها في الترتيب ثلاثة ارباع المشاهدة ويليها ربع المشاهدة وعلى فرض ان هذه البيانات مرتبة تصاعدياً.

ولاجل ايجاد الانحراف الربيعي نتبع الخطوات التالية :-

1) نرتب البيانات تصاعدياً.
2) نجد الوسيط للبيانات.
3) نجد الوسيط للقسم الاول من هذه البيانات وهذا القسم من البيانات يسبق الوسيط فيكون هذا الوسيط هو Q_1.
4) نجد الوسيط للقسم الثاني من البيانات وهذا القسم من البيانات يلي الوسيط فيكون هذا الوسيط هو Q_3.
5) نطبق القانون.

مثال:- جد الانحراف الربيعي للبيانات الآتية:-

9, 12, 10, 8, 2, 4, 7, 5

الحل

نرتب البيانات تصاعديا

2, 4, 5, 7, 8, 9, 10, 12

$$\tilde{X} = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2} = \frac{7+8}{2} = 12.5 \quad (\text{الوسيط})$$

$$Q_1 = \frac{4+5}{2} = 4.5 \quad , \quad Q_3 = \frac{9+10}{2} = 9.5$$

$$Q = \frac{9.5 - 4.5}{2} = 2.5 \quad (\text{الانحراف الربيعي})$$

اما في حالة البيانات المبوبة:-

$$Q_1 = L_1 + \left(\frac{\frac{1}{4}\sum f_i - F_1}{f_1} \right) * C$$

$$Q_3 = L_3 + \left(\frac{\frac{3}{4}\sum f_i - F_3}{f_3} \right) * C$$

إذ ان:-

F_1 : المتكرار المتجمع الصاعد للفئة التي تسبق فئة الربيع الاول مباشرة.

L_1 : الحد الادنى الحقيقي لفئة الربيع الاول.

وفئة الربيع الاول هي تلك الفئة التي تكرارها التجميعي الصاعد اكبر او يساوي $\frac{\sum f_i}{4}$

f_1 : تمثل تكرار فئة الربيع الاول .

c : طول الفئة .

F_3 : التكرار المتجمع الصاعد للفئة التي تسبق فئة الربيع الثالث مباشرة .

L_3 : الحد الادنى الحقيقي لفئة الربيع الثالث .

وفئة الربيع الثالث هي تلك الفئة التي تكرارها التجميعي الصاعد الجبر او يساوي $\frac{3}{4}\sum f_i$.

f_3 : تكرار فئة الربيع الثالث .

وبعد ان نجد Q_1 و Q_3 نطبق القانون $Q = \frac{Q_3 - Q_1}{2}$

س: احسب الانحراف الربيعي لجدول التوزيع التكراري الآتي :-

الفئات	f_i
50 – 59	8
60 – 69	10
70 – 79	16
80 – 89	15
90 – 99	10
100 – 109	5
110 – 119	3
120 – 129	3

المحاضرة ٦

② التباين (variance)

ⓐ للبيانات غير المبوبة

التباين:- عبارة عن متوسط مربعات انحرافات القيم عن وسطها الحسابي اي ان اذا كانت المشاهدات $x_1, x_2, \ldots, x_n$ عددها n ووسطها الحسابي $\bar{x}$ فإن التباين يحسب من الصيغة الآتية:-

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad , n < 60$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n} \quad , n \geq 60$$

وبالامكان اعادة كتابة المعادلة اعلاه بالصيغة الآتية:

$$S^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} \quad , n < 60$$

$$S^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n} \quad , n \geq 60$$

مثال:- اوجد التباين للبيانات الآتية:-

9, 8, 6, 5, 7

الحل

$$\bar{x} = \frac{\sum x_i}{n} = \frac{9+8+6+5+7}{5} = 7$$

x_i	$x_i - \bar{x}$	$(x_i-\bar{x})^2$
9	$9-7=2$	4
8	$8-7=1$	1
6	$6-7=-1$	1
5	$5-7=-2$	4
7	$7-7=0$	0
	$\Sigma = \text{Zero}$	$\Sigma = 10$

$$\therefore S^2 = \frac{\Sigma(x_i-\bar{x})^2}{n-1} = \frac{10}{5-1} = \frac{10}{4} = 2.5$$

(ب) للبيانات المبوبة (جدول توزيع تكراري)

اذا كانت $y_1, y_2, --, y_n$ تمثل مراكز الفئات مع تكراراتها على التوالي $f_1, f_2, --, f_n$ فإن التباين يكتب بالشكل الآتي.

$$S^2 = \frac{\sum_{i=1}^{n} f_i (y_i-\bar{y})^2}{\sum_{i=1}^{n} f_i - 1} \quad \text{or} \quad S^2 = \frac{\Sigma f_i y_i^2 - \frac{(\Sigma f_i y_i)^2}{\Sigma f_i}}{\Sigma f_i - 1}$$

$$\bar{y} = \frac{\Sigma f_i y_i}{\Sigma f_i}$$

مثال:- جد التباين من جدول التوزيع التكراري

الفئات	التكرارات	y_i	$f_i y_i$	$(y_i-\bar{y})^2$	$f_i(y_i-\bar{y})^2$
60 - 62	5	61	305	$(61-67.45)^2 = 41.60$	208.0125
63 - 65	18	64	1152	11.9025	214.245
66 - 68	42	67	2814	0.2025	8.505
69 - 71	27	70	1890	6.5025	175.7675
72 - 74	8	73	5084	30.8025	246.42
	$\sum_{i=1}^{n} f_i = 100$		$\Sigma = 6745$		$\Sigma = 852.75$

$$\bar{y} = \frac{\Sigma f_i y_i}{\Sigma f_i} = \frac{6745}{100} = 67.45 \qquad , S^2 = \frac{852.75}{99} = 8.6$$

(22)

٤) الانحراف المعياري (Standard Devation)

هو القيمة الموجبة لجذر التباين اي ان :

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

الانحراف المعياري.

مثال: جد الانحراف المعياري (القياسي) للبيانات الاتية:

2, 8, 3, 7, 6, 4

الحل

$$S^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} \quad , \quad \bar{x} = \frac{\sum x_i}{n} = \frac{30}{6} = 5$$

$$S^2 = \frac{178 - 6(5)^2}{5} = 5.6 \quad (\text{التباين})$$

$$S = \sqrt{S^2} = \sqrt{5.6} = 2.37 \quad (\text{الانحراف المعياري})$$

ملاحظة: هنالك مقاييس مشتركة بين مقاييس النزعة المركزية ومقاييس التشتت :-

١- معامل الاختلاف (C.V.)

٢- الدرجة المعيارية

٣- الالتواء

① معامل الاختلاف C.V.

$$C.V. = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times 100\,\%$$

مثال :- اذا كانت قيمة الانحراف القياسي (المعياري) والوسط الحسابي لمادتي الاحصاء والرياضيات كالآتي :-

	الوسط الحسابي	الانحراف المعياري
الاحصاء	78	8
الرياضيات	73	7.6

احسب اي من المادتين لها اعلى تشتت .

الحل

$$C.V.\ (\text{للإحصاء}) = \frac{8}{78} \times 100\,\% = 10.25\,\%$$

$$C.V.\ (\text{للرياضيات}) = \frac{7.6}{73} \times 100\,\% = 10.41\,\%$$

$$C.V.\ (\text{للرياضيات}) > C.V.\ (\text{للإحصاء})$$

∴ تشتت مادة الرياضيات اكبر من تشتت مادة الإحصاء .

٣- الدرجة المعيارية

تعرف الدرجة المعيارية بالقانون الاتي :-

$$Z = \frac{X - \bar{X}}{S}$$

وتستخدم هذه الدرجة المعيارية لمقارنة موضوع واحد لشخصين في مجموعتين او موضوعين لشخص واحد. والذي تكون درجته المعيارية الاكبر يكون هو الافضل.

مثال:- حصل طالب على درجة (84) في مادة الرياضيات حيث كان متوسط الدرجات $(\bar{X}_m = 76)$ و الانحراف القياسي $(S_m = 10)$ وحصل على درجة (90) في مادة الفيزياء حيث كان متوسط الدرجات $(\bar{X}_n = 82)$ والانحراف القياسي $(S_n = 16)$ في اي موضوع كانت درجة استيعابه افضل.

الحل

$$Z_m = \frac{X_m - \bar{X}_m}{S_m} = \frac{84 - 76}{10} = 0.8$$

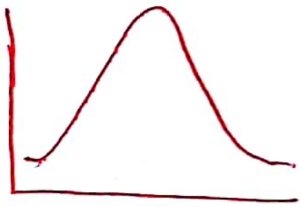
$$Z_n = \frac{X_n - \bar{X}_n}{S_n} = \frac{90 - 82}{10} = 0.5$$

∴ في مادة الرياضيات كانت درجة استيعابه الكبر.

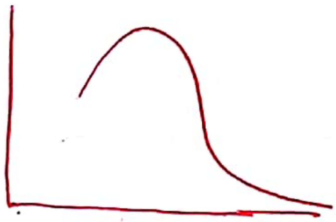
٢- الالتواء (Skewness)

هو عبارة عن درجة القرب او البعد عن التماثل للتوزيع. فإذا كان المنحنى التكراري للتوزيع له ذيل اكبر الى اليمين يسمى التوزيع بأنه ملتوي الى اليمين او موجب الالتواء.

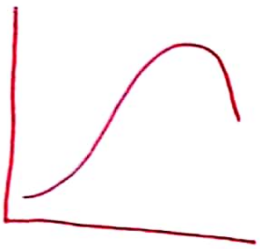
اما اذا كان العكس صحيح فيقال بأن التوزيع ملتوي الى اليسار او سالب الالتواء.



متماثل
لا يوجد التواء
(الالتواء = صفر)



الالتواء موجب
نحو اليمين



الالتواء سالب
نحو اليسار

هنالك عدة مقاييس لحساب معامل الالتواء منها :-

$$S_k = \frac{\bar{X} - \bar{X}_0}{S}$$

اذ ان:

$\bar{X}$: الوسط الحسابي.

$\bar{X}_0$: المنوال

S : الانحراف القياسي (المعياري)

المحاضرة ٨

الارتباط والانحدار (Correlation & Regression)

تعريف :- اذا كان لكل قيمة من قيم المتغير X . يوجد قيمة مقابلة لمتغير آخر y فإن الزوج المرتب (X, y) متغير ذا بعدين . فمثلاً اذا كان لدينا بيانات على طلبة الجامعة تمثل علاقة الطالب في الشهادة ومعدل تخرجه .

* يمكن دراسة هذه البيانات للاجابة على السؤالين :-
ⓐ هل هنالك علاقة بين المتغيرين .
ⓑ اذا كان هنالك علاقة بين المتغيرين فكيف يعبر عنها بمعادلة .

* الارتباط الخطي :-

اذا كان لدينا n من المشاهدات $X_1, \dots, X_n$ و $y_1, y_2, \dots, y_n$ وكانت قيمتي الوسط الحسابي لهما هي $\bar{X}$ و $\bar{y}$ على التوالي فإن معامل الارتباط يرمز له بالرمز r هو

$$r_{x,y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

or

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sqrt{(\sum x_i^2 - n \bar{x}^2)(\sum y_i^2 - n \bar{y}^2)}}$$

* اهم خواص معامل الارتباط

① $-1 \le r_{x,y} \le 1$

② اذا كان $r=1$ فإن الارتباط موجب (عالي جداً) (طردي)

$r=-1$ " " سالب (عالي جداً) (عكسي)

$r=0$ لا يوجد ارتباط (X مستقل عن Y)

مثال:- البيانات التالية تمثل كمية المعروضة من سلعة معينة وسعر الوحدة الواحدة منها ، احسب معامل الارتباط بين الكمية المعروضة والسعر .

X_i	Y_i	$X_i-\bar{X}$	$Y_i-\bar{Y}$	$(X_i-\bar{X})^2$	$(Y_i-\bar{Y})^2$	$(X_i-\bar{X})(Y_i-\bar{Y})$
2	3	-2	-4	4	16	8
2	5	-2	-2	4	4	4
5	7	1	0	1	0	0
4	8	0	1	0	1	0
5	9	1	2	1	4	2
6	11	2	4	4	16	8
3	6	-1	-1	1	1	1
5	8	1	1	1	1	1
4	6	0	-1	0	1	0
$\Sigma X_i=36$	$\Sigma Y_i=63$			$\Sigma=16$	$\Sigma=44$	$\Sigma=24$

$$\bar{X}=\frac{\Sigma X_i}{n}=\frac{36}{9}=4$$

$$\bar{Y}=\frac{\Sigma Y_i}{n}=\frac{63}{9}=7$$

$$, r_{x,y}=\frac{24}{\sqrt{(16)(44)}}=0.91$$

هنالك علاقة قوية بين X و Y .
(علاقة طردية)

الانحدار (Regression)

هو اسلوب يتم فيه التنبؤ بمتوسط متغير عشوائي او عدة متغيرات عشوائية اعتماداً على قيم وقياسات متغيرات عشوائية اخرى، له عدة انواع مثل:-

1. الانحدار الخطي البسيط:- يعني ان المتغير التابع (الاستجابة) Y يعتمد على متغير عشوائي X واحد، وتكون العلاقة بينهم علاقة خطية.
2. الانحدار الخطي المتعدد:- يعني ان المتغير Y يعتمد على اكثر من متغير عشوائي مستقل (التوضيحي) X.
3. الانحدار غير الخطي:- اذا كانت العلاقة بين Y والمتغير المستقل X غير خطية كأن تكون من الدرجة الثانية او علاقة اسية او دوال مثلثية.
4. الانحدار اللوجستي.
5. انحدار بواسون.
6. الانحدار اللامعلمي.
7. الانحدار شبه المعلمي.

⋮

سيتم في الدراسة استعمال الانحدار الخطي البسيط.

الانحدار الخطي البسيط:-

بعد تمثيل الازواج المرتبة بالمستوى الاحداثي نحصل على شكل الانتشار للبيانات فإذا كان الشكل خط مستقيم نقوم بتقدير خط انحدار Y على X بواسطة العلاقة:-

$$\hat{Y} = a + b X$$

حيث ان $\hat{Y}$ تمثل القيمة المتنبأ بها عندما تكون X معلومة.

وان a تمثل المعلمة المقدرة بطريقة المربعات الصغرى لـ ثابت الانحدار.

b تمثل المعلمة المقدرة بطريقة المربعات الصغرى لـ ميل خط الانحدار.

$$b = \frac{n\sum_{i=1}^{n} X_i Y_i - \left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}$$

$$a = \bar{Y} - b\bar{X} \quad , \quad \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \quad , \quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

<u>مثال:</u> اذا كان لدينا الجدول الاتي:

X	Y	XY	X^2
10	6	60	100
13	8	104	169
15	9	135	225
14	8	112	196
9	7	63	81
7	6	42	49
6	5	30	36
6	6	36	36
5	5	25	25
5	5	25	25
$\Sigma = 90$	$\Sigma = 65$	$\Sigma = 632$	$\Sigma = 942$

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

$$b = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{n\Sigma X^2 - (\Sigma X)^2}$$

$$= \frac{(10)(632) - (90)(65)}{(10)(942) - (90)^2}$$

$$= 0.36 = b$$

$$a = \frac{\sum_{i=1}^{n} y_i - b\sum_{i=1}^{n} x_i}{n} = \frac{65 - (0.36)(90)}{10}$$

$$= 3.26 = a$$

$$\therefore \hat{Y} = 3.26 + 0.36\,X \Longleftarrow$$ تمثل معادلة خط الانحدار البسيط

اذا علمت ان قيمة ($X = 16$) اوجد القيمة المتنبأ بها لـ Y.

$$\hat{Y} = 3.26 + 0.36\,(16)$$
$$= 9.02$$

واجب بيتي

مثال: لديك الجدول الاتي:-

X	1	2	3	4	5	6	7	8
Y	13	12	10	7	8	6	5	2

المطلوب:-

1) إيجاد معادلة خط الانحدار البسيط

2) تنبأ قيمة Y عندما $X = 3$ و $X = 15$

المحاضرة

مبادئ العد .

مضروب (مفكوك العدد)

اذا كان n عدد صحيح غير سالب (n عدد طبيعي)، فإن

مضروب العدد (مفكوك العدد) يرمز له بالرمز $n!$ او

$\underline{|n}$ ويحسب كالآتي :

$$\underline{|n} = n! = n\cdot(n-1)(n-2)\text{ -----}\cdot(1)$$

ex:- $\underline{|6}$

$$\underline{|6} = 6! = 6\cdot5\cdot4\cdot3\cdot2\cdot1 = ?$$

$$\underline{|10} = 10! = 10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1 = ?$$

هنالك صيغ على التعويض $n!$ حسب الحاجة في الحل :

$$n! = n(n-1)! \quad \xRightarrow{ex} 6! = 6\cdot5!$$

$$n! = n(n-1)(n-2)! \quad \xRightarrow{ex} 6! = 6\cdot5\cdot4!$$

$$n! = \frac{(n+1)!}{(n+1)} \quad \xRightarrow{ex} 6! = \frac{7!}{7} = \frac{\cancel{7}\cdot6!}{\cancel{7}} = 6!$$

ex:- $0! = \underline{|0} = 1$, Prove that. برهن ذلك

من خلال العلاقة $n! = \frac{(n+1)!}{n+1}$ فإن

$$0! = \frac{1!}{1} = 1$$

مثال :- جد قيمة $\frac{\lfloor 8}{\lfloor 6}$ او $\frac{8!}{6!}$

الحل

$$\frac{8!}{6!} = \frac{8\cdot 7\cdot \cancel{6!}}{\cancel{6!}} = 56$$

مثال :- اذا كان $\frac{n!}{(n-2)!} = 6$ فما هي قيمة الـ n ؟

الحل

$$\frac{n!}{(n-2)!} = \frac{n(n-1)\cancel{(n-2)!}}{\cancel{(n-2)!}} = 6$$

$$\Rightarrow n(n-1) = 6 \Rightarrow n^2 - n = 6$$

$$\Rightarrow n^2 - n - 6 = 0$$

$$\Rightarrow (n+2)(n-3) = 0$$

اما $n = -2$ تهمل لانها بالسالب (كون المفكوك يتعامل مع الاعداد الموجبة فقط)

او $n = 3$

للتحقق من ذلك

$\because n = 3$

$$\frac{3!}{(3-2)!} \Rightarrow \frac{3!}{1!} = \frac{3\cdot 2\cdot 1}{1} = 6$$

التباديل (Permutation)

هو عدد طرق الاختيار المرتب التي يمكن تكوينها من عدة اشياء بأخذ كلها او بعضها ويرمز له بالرمز $P(n,r)$ او P_r^n اي تبديل r عنصر من n من العناصر ، وهو مبدأ من مبادئ العدد يستخدم لكل المسائل التي يكون فيها الترتيب مهم والتكرار غير مسموح ، وهنالك حالات يكون التكرار مسموح .

ليكن $r,n \in N$ وان $r \leq n$ فان تباديل r مأخوذة من n تكتب بالشكل الاتي :-

$$P_r^n = P(n,r) = \frac{n!}{(n-r)!}$$

مثال :- ما هو عدد الكلمات المختلفة التي يمكن تكوينها من كلمة ربح

الحل

$$P_3^3 = P(3,3) = \frac{3!}{0!} = \frac{3.2.1}{1} = 6$$

والكلمات هي (رحب ، برح ، بحر ، حرب ، حبر ، ربح) .

مثال :- ما عدد طرق توزيع 5 اشخاص على 5 وظائف مختلفة بحيث لكل واحد منهم وظيفة واحدة .

$$P_5^5 = \frac{5!}{(5-5)!} = \frac{5!}{0!} = \frac{5.4.3.2.1}{1!} = 120$$

عدد الطرق الممكنة للتوزيع

مثال :- جد قيمة n اذا كان $P_2^n = 42$.

الحل

$$P_2^n = \frac{n!}{(n-2)!} = \frac{n(n-1)\cancel{(n-2)!}}{\cancel{(n-2)!}}$$

$$\therefore P_2^n = n(n-1) = 42$$

$$n^2 - n = 42 \Rightarrow n^2 - n - 42 = 0$$

$$(n-7)(n+6) = 0$$

اما $n = 7$

او $n = -6$ (تهمل كونها سالب)

للتحقق

$$P_2^7 = \frac{7!}{(7-2)!} = \frac{7!}{5!} = \frac{7 \cdot 6 \cdot \cancel{5!}}{\cancel{5!}} = \underline{\underline{42}}$$

واجب بيتي

H.W. :- ماهو عدد الكلمات المختلفة المكونة من حرفين يمكن تشكيلها من هذه الكلمة ABCD .