

احتمالات واحصاء

المحاضرة ٥

الفرقة الثالثة

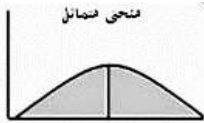
شعب عامه

مقاييس التشتت ٢

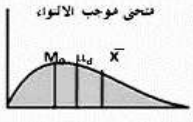
الالتواء

تعريف الالتواء:

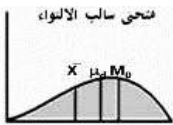
هو انحراف منحنى التوزيع التكراري عن التماثل وقد يكون الالتواء موجباً [اي التواء الي اليمين] او سالباً [إلتواء الي اليسار] .



الوسط = الوسيط = المنوال



الوسط < الوسيط < المنوال



الوسط > الوسيط > المنوال

ففي حاله التوزيعات المتماثله فإن :

الوسط الحسابي = الوسيط = المنوال كما بالشكل ١

وإذا كان التوزيع ملتويًا جهة اليمين فإن :

الوسط الحسابي < الوسيط < المنوال كما بالشكل ٢

وإذا كان التوزيع ملتويًا جهة اليسار فإن :

الوسط الحسابي > الوسيط > المنوال كما بالشكل ٣

وعليه يمكن حساب معامل الإلتواء والذي يرمز له بالرمز α من خلال العلاقات التالية:

$$\alpha = \frac{\bar{x} - m}{s} \dots\dots(1)$$

$$\alpha = \frac{3(\bar{x} - \mu)}{s} \dots\dots(2)$$

حيث :

\bar{x} : الوسط الحسابي

m : المنوال

S : الانحراف المعياري

μ : الوسيط

ويمكن ايضاً قياس الالتواء بدراسه المواقع النسبيه للربيع الاول والوسيط والربيع الثالث

للتوزيع التكراري فاذا كان التوزيع متماثلا فان الفرق بين الوسيط والربيع الاول يساوي

الفرق بين الربع الثالث و الوسيط واذا كان التوزيع ملتويًا جهة اليمين فان الربع الاول

يكون اقرب الي الوسيط مقارنة بالربع الثالث

واذا كان التوزيع ملتويًا جهة اليسار فان الربع الثالث يكون اقرب الي الوسيط مقارنة

بالربع الأول ويكون مقياس الإلتواء في هذه الحالة هو :

$$\alpha = \frac{(Q_3 - \mu) - (\mu - Q_1)}{(Q_3 - \mu) + (\mu - Q_1)}$$

مثال:

احسب معامل الإلتواء للبيانات التالية :

$$\bar{x} = 70 , \mu = 69.4 , m = 68.5 , Q_1 = 50 , Q_3 = 80 , S = 6$$

الحل:

$$\therefore \alpha = \frac{\bar{x} - m}{s} \Rightarrow \alpha = \frac{70 - 68.5}{6} = 0.25$$

$$\therefore \alpha = \frac{3(\bar{x} - m)}{s} \Rightarrow \alpha = \frac{3(70 - 69.4)}{6} = 0.30$$

$$\alpha = \frac{(Q_3 - \mu) - (\mu - Q_1)}{(Q_3 - \mu) + (\mu - Q_1)} = \frac{(80 - 69.4) - (69.4 - 60)}{(80 - 69.4) + (69.4 - 60)} = 0.06$$

ومن هذه النتائج يتضح لنا أن هناك إلتواءً موجباً اي إلتواء جهة اليمين وباستخدام طريقة العزوم نجد أن :

$$\alpha = \frac{m_3}{\sigma^3}$$

$$m_3 = \begin{cases} \frac{1}{n} \sum (x_i - \bar{x})^3 \\ \frac{1}{\sum f_i} \sum f_i (x_i^* - \bar{x})^3 \end{cases}$$

مقياس التفلطح

مقدمة:

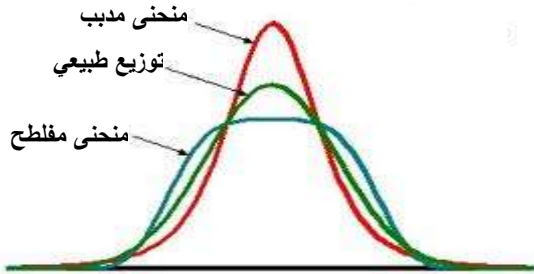
هو مقياس يقيس تدبب [او الإستواء] لمنحني الكثافة عند المتوسط وهو ايضاً مقياس يقيس درجه علو أي منحني توزيع تكراري أو إنخفاضه بالنسبه للمنحني الطبيعي وهو منحني متمائل حول محور رأسي يمر بالمتوسط ويعرف التفلطح والذي يرمز له بالرمز β بالعلاقة الرياضية التاليه :

$$\beta = \frac{m_4}{s^4}$$

m_4 : العزم الرابع حول المتوسط

حيث أن :

$$m_4 = \begin{cases} \frac{1}{n} \sum (x_i - \bar{x})^4 \\ \frac{1}{\sum f_i} \sum f_i (x_i^* - \bar{x})^4 \end{cases}$$



ملاحظه:

إذا كانت :

$\beta = 0$ سميت القمه معتدله

$\beta > 0$ سميت القمه مدبيه

$\beta < 0$ سميت القمه مفطحه

مثال:

احسب مقياس الإلتواء والتفلطح للبيانات التاليه :

x_i	5-	10-	15-	20-	25-	30-
f_i	8	9	14	12	11	10

الحل:

x_i	f_i	x_i^*	$f_i x_i^*$	$(x_i^* - \bar{x})$	$f_i (x_i^* - \bar{x})^2$	$f_i (x_i^* - \bar{x})^3$	$f_i (x_i^* - \bar{x})^4$
5-9	8	7	56	-13	1352	-1757.76	228488
10-14	9	12	108	-8	576	-4608	36864
15-19	14	17	238	-3	126	-378	1134
20-24	12	22	264	2	48	96	192
25-29	11	27	297	7	539	3773	26411
30-34	10	32	320	12	1440	17280	207360
	64		1283		4081	-1413	500449

$$\bar{x} = \frac{1}{\sum f_i} \sum f_i x_i^* = \frac{1283}{64} = 20.1 \approx 20$$

$$S^2 = \frac{1}{\sum f_i} \sum f_i (x_i^* - \bar{x})^2 = \frac{4081}{64} = 63.76$$

$$S = \sqrt{S^2} = 7.98$$

$$m_3 = \frac{1}{\sum f_i} \sum f_i (x_i^* - \bar{x})^3 = \frac{-1413}{64} = -22.1$$

∴ مقياس الالتواء هو :

$$\alpha = \frac{m_3}{S^3} = \frac{-22.1}{531.44} = -0.042$$

وهذا يدل علي أن المنحني ملتوي نحو اليسار قليلا . ولحساب مقياس التفلطح نجد أن مربع التباين هو :

$$S^2 = 63.76 \Rightarrow S^4 = 4065.34$$

وكذلك يمكن حساب العزم الرابع كالتالي :

$$m_4 = \frac{1}{\sum f_i} \sum f_i (x_i^* - \bar{x})^4 = \frac{500449}{64} = 7819.516$$

وعليه فان مقياس التفلطح هو :

$$\beta = \frac{m_4}{S^4} = \frac{7819.516}{4065.34} = 1.92$$

وهي تمثل قمة المنحني للتوزيع التكراري وحيث أن $\beta > 0$ تكون القمة مدببه

الارتباط والانحدار

الانحدار:

في هذا الموضوع سوف تقوم بتعريف أهميه العلاقة بين المتغيرات ونوع انتشار البيانات بين متغيرين اثنين وتمثل أيضا البحث في علاقة رياضية بين متغيرين تمكن من إجراء تنبؤ لقيم مستقبلية للمتغيرات وهو ما يعرف بمعادله الانحدار .

العلاقة بين المتغيرات والتبؤ:

بعد إجراء عدة تجارب عمليه تم إيجاد وعلاقه رياضية تربط بين المتغيرات وهذه العلاقة الرياضيه تقرب البيانات التجريبيه التي تستخدم لاستخلاص معلومات عن هذه المتغيرات تكون ذات فائده في المستقبل .

العلاقة بين متغيرين:

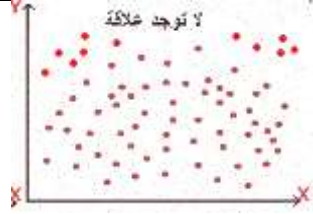
إذا أجرينا تجربه معينه كان فيها المتغير المستقل هو X والمتغير التابع هو Y وقمنا بتجميع البيانات لهذه المتغيره من التجارب التي نجريها لنحصل علي القراءات :
 $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$ فإذا رسمنا هذه النقاط التجريبيه في المستوي xy فإننا نحصل علي شكل الانتشار .

أشكال الانتشار:

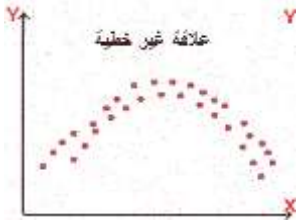
أشكال الإنتشار تأخذ صوراً مختلفه وذلك حسب طبيعه العلاقة بين المتغيرين (x, y) محل الدراسه كما يتضح من الرسوم التاليه:



تكون النقاط منتشرة حول خط مستقيم
تريد قيمة مع زياده قيم X وهذه علاقة
خطية طردية بين X, Y



هذا الشكل النقاط منتشرة بدون ترابط
حول إتجاه محدد وهذا يدل يدل علي
عدم وجود علاقة بين X, Y



النقاط منتشرة حول منحنى وهذه علاقته
غير خطية بين X, Y



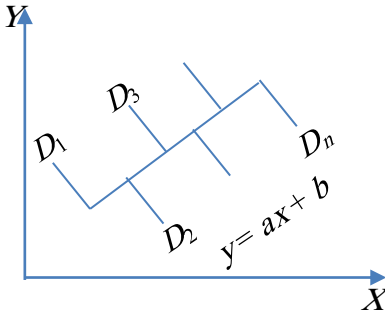
النقاط منتشرة حول خط مستقيم وفيه
تتقص قيم Y مع زيادة قيم X وهذا
معناه وجود علاقة عكسيه بين X, Y
فإذا كان لدينا أزواج من القراءات :

$(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$
مستقيم . وتسمى هذه القراءات (النقاط) بشكل الإنتشار وتقرب هذه النقاط الي خط
مستقيم بطريقة المربعات الصغري .

خطوط المربعات الصغرى

الحالة الاولى:

إيجاد افضل خط مستقيم علي الصورة $Y = aX + b$ [معادله خط إنحدار y علي x]



بتقريب مجموعه النقاط السابقه يمكن الحصول

علي معادلة خط إنحدار y علي x حيث a, b

ثوابت يمكن إيجاد قيمتها بإستخدام طريقة

المربعات الصغرى وبذلك يتحدد الخط المستقيم

تماماً رياضياً .

نفرض أن D_1, D_2, \dots, D_n هي الانحرافات

الرأسيه لمجموعة البيانات عن الخط المستقيم $y = ax + b$

وتعين بالطريقه التاليه :

$$D_1 = Y_1 - (aX_1 + b)$$

$$D_2 = Y_2 - (aX_2 + b)$$

⋮

$$D_n = Y_n - (aX_n + b)$$

ونفرض أن مجموع مربعات هذه الانحرافات هي :

$$s^2 = D_1^2 + D_2^2 + \dots + D_n^2$$

$$s^2 = [Y_1 - (aX_1 + b)]^2 + \dots + [Y_n - (aX_n + b)]^2$$

نلاحظ أن S^2 تعتمد علي a, b فقط أي أنها داله في a, b ويكون S^2 أقل ما يكن

عندما يتحقق الشرطان :

$$\frac{\partial s^2}{\partial a} = 0 \quad , \quad \frac{\partial s^2}{\partial b} = 0$$

بتفاضل S^2 جذئياً بالنسبة ل a, b والمساواه بالصفر نحصل علي المعادلات الاعتدالية التاليه :

$$\sum y = a\sum x + nb \quad (1)$$

$$\sum xy = a\sum x^2 + b\sum x \quad (2)$$

بحل المعادلتين (1) , (2) في a, b نحصل علي أفضل خط مستقيم في الصورة $y = ax + b$ حيث n عدد البيانات .

الحالة الثاني:

إيجاد أفضل خط مستقيم علي الصورة : $x = ay + b$ [معادله خط إنحدار x علي y] باستخدام نفس الطريقه السابقه في حاله الاولي يمكننا الحصول علي المعادلات الإعتداليه التاليه:

$$\sum x = a\sum y + nb \quad (1)$$

$$\sum xy = a\sum y^2 + b\sum y \quad (2)$$

بحل المعادلتين (1) , (2) وتعيين قيمتي a, b يمكننا الحصول علي معادله خط إنحدار x علي y في الصوره $x = ay + b$

مثال:

إذا كان لدينا مجموعه البيانات التاليه :

x	2	4	1	3	5	6
y	2	5	6	3	4	2

فاحسب :

١- معادله خط انحدار y علي x ثم احسب قيمه y عندما $x = 3.5$

٢- معادله خط انحدار x علي y ثم احسب قيمه y عندما $x = 2.3$

الحل:

أ- باستخدام المعادلات الاعتدالية:

$$\sum y = a\sum x + nb \quad , \quad \sum xy = a\sum x^2 + b\sum x$$

x	y	xy	x ²	y ²
2	2	4	4	4
4	5	20	16	25
1	6	6	1	36
3	3	9	9	9
5	4	20	25	16
6	2	12	36	4
21	22	71	91	94

$$\therefore 22 = 21a + 6b$$

$$71 = 91a + 21b$$

بحل هاتين المعادلتين نحصل علي:

$$a = -0.34 \quad , \quad b = 4.9$$

$$\therefore y = -0.34x + 4.9$$

اذن قيمة y عندما x = 3.5 هي:

$$\therefore y|_{x=3.5} = -0.34(3.5) + 4.9 = 3.71$$

ب- لإيجاد أفضل خط مستقيم علي الصورة :

$$x = ay + b$$

نستخدم المعادلات الاعتدالية:

$$\sum x = a\sum y + nb \quad , \quad \sum xy = a\sum y^2 + b\sum y$$

$$\therefore 21 = 22a + 6b$$

$$71 = 94a + 22b$$

بحل هاتين المعادلتين نحصل علي:

$$a = -0.45 , b = 5.15$$

$$\therefore x = -0.45y + 5.15$$

اذن قيمة x عندما $y = 2.3$ هي:

$$\therefore x|_{y=2.3} = -0.45(2.3) + 5.15 = 4.11$$

الإرتباط:

عند دراسة العلاقة بين متغيرين x, y فإن شكل الإنتشار يمكن أن يوضح طبيعه هذه العلاقة . فالعلاقة بين X, Y قد تكون قوية جداً إذا ما وقعت نقاط نقاط شكل الإنتشار كلها علي منحنى أو خط مستقيم وتكون أقل قوة إذا لم تتغير نقاط شكل الإنتشار تماماً بمنحنى أو خط مستقيم . سوف نقتصر في دراستنا هنا علي ما يسمى بالإرتباط الخطي ويظهر في حالة إمكانية تقريب نقاط شكل الإنتشار بخط مستقيم .

عند دراسة الإرتباط في هذه الحالة تنشأ الحالات التالية :

- ١- إرتباط طردي موجب أي أن قيمة Y تزداد بزيادة قيمة X
- ٢- إرتباط عكسي سالب أي أن قيمة Y تقل بزيادة قيمة X
- ٣- لا يوجد إرتباط أي أن قيمة Y لا تتأثر بتغير قيمة X

معامل الإرتباط بين متغيرين (X, Y)

هو مقياس لدرجة العلاقة بين (X, Y) ويرمز له بالرمز r ويحقق معامل الإرتباط العلاقة التالية :

$$r = \begin{cases} 1 & \text{طردي تام} \\ 0 & \text{لا يوجد ارتباط} \\ -1 & \text{عكسي تام} \end{cases}$$

حساب معامل الارتباط

يمكن حساب معامل الارتباط بإحدى الطريقتين .

١- طريقه بيرسون :

بفرض ان لدينا مجموعة من البيانات :

فإن معامل الارتباط يعطي بالعلاقة الرياضية التاليه:

$$r = \frac{\sum xy - n \cdot \bar{x} \cdot \bar{y}}{n \cdot s_x \cdot s_y} \quad (1)$$

حيث أن :

\bar{x} : هو الوسط الحسابي لقيم x

s_x : الانحراف المعياري لقيم x

\bar{y} : هو الوسط الحسابي لقيم y

s_y : هو الانحراف المعياري لقيم y

n : عدد البيانات

المعادلة (1) يمكن كتابتها في الصورة التالية :

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (2)$$

مثال:

إحسب معامل الارتباط للمثال السابق .

الحل:

باستخدام المعادله رقم (٢) نحصل علي :

$$\begin{aligned}
 r &= \frac{n\sum xy - \sum x\sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \\
 &= \frac{6(71) - (21)(22)}{\sqrt{[6(91) - (21)^2][6(94) - (22)^2]}} \\
 &= \frac{426 - 462}{\sqrt{(546 - 441)(564 - 484)}} \\
 &= \frac{-36}{\sqrt{(105)(80)}} \\
 &= -0.39
 \end{aligned}$$

٢- طريقه سبيرمان :

تتلخص هذه الطريقة في تعيين رتب لمختلف القيم علي النحو التالي

أ- بالنسبه لقيم المتغير x تعطي أكبر قيمه الرتبه ١ والقيمه التي تليها الرتبه ٢ وهكذا .

ب- تعيين الرتب لقيم y بنفس الطريقه السابقه .

ت- لا تعطي رتباً مختلفه لقيم مكرره في متغير ما بل إذا وجدت قيمتان [أو أكثر]

متساويتان فانهما يأخذان رتبه واحده هي الوسط الحسابي للرتبتين (او الرتب)

ث- نحسب فرق الرتب D للقيم المناظره .

ج- نحسب معامل الارتباط من خلال العلاقة :

$$r = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

حيث n عدد البيانات .

مثال:

إحسب معامل الارتباط بين المتغيرين (x, y) باستخدام طريقة سيرمان للبيانات التالية :

x	2	4	6	8	10	12
y	1	3	5	7	9	11

الحل:

$$\therefore r = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

x	y	رتبه x	رتبه y	D	D^2
2	1	6	6	0	0
4	3	5	5	0	0
6	5	4	4	0	0
8	7	3	3	0	0
10	9	2	2	0	0
12	11	1	1	0	0

$$\therefore r = 1 - \frac{6(0)}{6(36-1)} = 1$$

طردي تام

ملحوظة:

- أ- إذا أردنا الحصول علي معامل إرتباط بين x, y يكون طردي تام يجب أن تكون بيانات x, y مرتبه ترتيبياً تصاعدياً أو تنازلياً .
- ب- إذا أردنا الحصول علي معامل إرتباط بين x, y يكون عكسي تام يجب أن تكون بيانات أحد المتغيرين مرتبة ترتيبياً تصاعدياً والمتغير الآخر بياناته مرتبه ترتيبياً تنازلياً .

تمارين

١- الجدول التالي يمثل عدد الطلاب الذين يمارسون بعض أنواع الانشطة الرياضية:

النشاط	كرة قدم	هوكي	سباحه	كره سله	كرة طائرة	تنس	كرة يد
بنين	40	25	50	45	35	20	25
بنات	25	10	35	25	30	15	20

المطلوب تمثيل هذه البيانات باستخدام الاعمدة المزوجة وكذلك باستخدام القطاع الدائري

٢- البيانات التالية تمثل درجات 40 طالب في احد الاختبارات

41	42	44	46	24	33	56	45	22	53
44	66	63	67	52	38	36	32	23	44
32	29	28	27	55	65	43	45	52	37
54	44	24	41	39	63	43	32	31	43

أ- احسب متوسط درجات الطلاب - المنوال - الوسيط - معامل الالتواء وبين نوعه
ب- ضع هذه البيانات في جدول تكراري ذي 7 فئات متساوية الطول ثم احسب

الوسط الحسابي - نصف المدى الربيعي - P_{85}

٣- احسب الوسط الهندسي والتوافقي والوسيط ونصف المدى الربيعي ومعامل

الاختلاف للبيانات 13,11,15,17,9,7,14,12,10 وإذا عدلت القيم السابقة

حسب المعادلة $y = 0.9x + 10$: فاحسب الانحراف المتوسط - التباين -

معامل الاختلاف للمتغير y

٤- احسب Q_3, D_6, P_{75} للبيانات التاليه :

x_i	4-	9-	14-	19-	24-	29-	34-
-------	----	----	-----	-----	-----	-----	-----

الاحتمالات والاحصاء

f_i	10	12	8	7	4	3	8
-------	----	----	---	---	---	---	---

٥- ثلاثة من مدرسي كرة القدم اعطوا متوسط درجات امتحاناتهم 79,74,82 في فصولهم المكونه من 32,25,17 طالبا علي الترتيب أوجد متوسط الدرجات في جميع الفصول

٦- حصل طالب في احد الفصول الدراسيہ علي الدرجات التاليہ:
75,80,65,87,70 وكان عدد الساعات المعتمده هي: 4,3,4,2,3 احسب المعدل الفصلي لهذا الطالب. احسب الوسط الحسابي للدرجات باستخدام طريقة الوسط الفرضي

٧- لدراسه العلاقة بين الدخل (X) بالاف الجنيهات والانفاق (Y) بالاف الجنيهات في احدي المدن اخذت عينه من الاسر فكانت لدينا النتائج التاليہ:

$$\bar{X} = 5, \bar{Y} = 4, \sum XY = 750, \sum X^2 = 1480, \sum Y^2 = 500, \sum Y = 80$$

- أ- اوجد معامل الارتباط بين X, Y وبين نوعه
ب- اوجد خط انحدار X علي Y
ت- قدر قيمة الدخل عندما يكون الانفاق 6 الاف جنيهه
٨- احسب معادله خط انحدار Y علي X وكذلك معامل الارتباط مبينا نوعه للبيانات التاليہ:

X	8	10	7	6	9	11	5
Y	10	12	8	10	6	11	6

٩- احسب معامل الارتباط بين تقديرات مادتي الاحصاء والجبر لسته طلاب

موزعة تقديراتهم علي النحو التالي:

6	5	4	3	2	1	الطالب
ضعيف جدا	مقبول	جيد جدا	ضعيف	جيد	ممتاز	الاحصاء
مقبول	ضعيف	ممتاز	ضعيف جدا	جيد	جيد جدا	الجبر

