احصاء حيوي

المحاضرة ٥

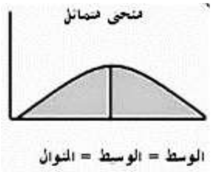
المستوي الثانـي ساعات معتمدة

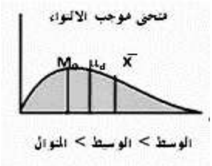
علوم البحار البيولوجيه

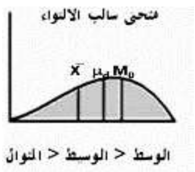
الالتواء

تعريف الالتواء:

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







ففي حاله التوزيعات المتماثله فإن :

الوسط الحسابي = الوسيط = المنوال كما بالشكل ١

واذا كان التوزيع ملتويا جهه اليمين فإن:

الوسط الحسابي > الوسيط > المنوال كما بالشكل ٢

واذا كان التوزيع ملتويا جهه اليسار فإن:

الوسط الحسابي < الوسيط < المنوال كما بالشكل ٣

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

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

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

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 الربيع الثالث و الوسيط واذا كان التوزيع ملتوياً جهه اليمين فان الربيع الاول يكون اقرب الي الوسيط مقارنه بالربيع الثالث

وإذا كان التوزيع ملتوياً جهه اليسار فان الربيع الثالث يكون اقرب الي الوسيط مقارنه بالربيع الأول ويكون مقياس الإلتواء في هذه الحاله هو :

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

مثال:

احسب معامل الإلتواء للبيانات التاليه :

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

الحل:

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

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

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

ومن هذه النتائج يتضح لنا أن هناك إلتواءً موجباً اي إلتواء جهه اليمين وبإستخدام طريقه 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:

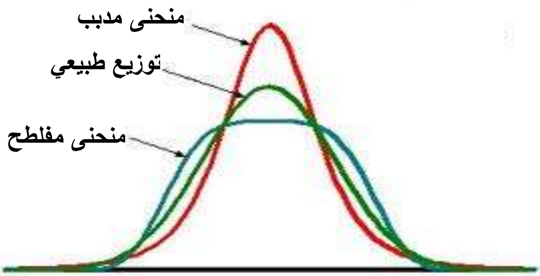
هو مقياس يقيس تدبب [او الإستواء] لمنحني الكثافه عند المتوسط وهو ايضاً مقياس يقيس درجه علو أي منحني توزيع تكراري أو إنخفاضه بالنسبه للمنحني الطبيعي وهو منحني متماثل حول محور رأسي يمر بالمتوسط ويعرف التفلطح والذي يرمز له بالرمز β بالعلاقة الرياضية التاليه :

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

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

$\beta \succ 0$ سميت القمه مدببه

$\beta \prec 0$ سميت القمه مفلطحه

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

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

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

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

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

$\therefore$ 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 $\beta \succ 0$ تكون القمة مدببه

الارتباط والانحدار

الانحدار:

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

العلاقه بين المتغيرات والتبؤ:

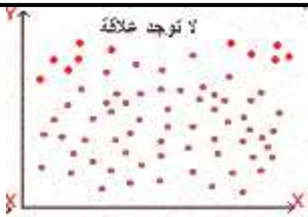
بعد إجراء عده تجارب عمليه تم إيجاد وعلاقه رياضية تربط بين المتغيرات وهذه العلاقه الرياضيه تقرب البيانات التجريبيه التي تستخدم لاستخلاص معلومات عن هذه المتغيرات تكون ذات فائده في المستقبل .

العلاقه 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 :
$(x_1, y_1),(x_2, y_2),(x_3, y_3),...,(x_n, y_n)$ فإذا رسمنا هذه النقاط التجريبيه في المستوي xy فإننا نحصل علي شكل الانتشار.

أشكال الانتشار:

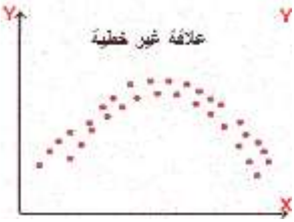
أشكال الإنتشار تأخذ صوراً مختلفه وذلك حسب طبيعه العلاقة بين المتغيرين (x, y) محل الدراسه كما يتضح من الرسوم التالية:



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

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

النقاط منتشره حول خط مستقيم وفية تنقص قيم Y مع زيادة قيم X وهذا معناه وجود علاقة عكسيه بين X , Y



النقاط منتشرة حول منحني وهذه علاقه غير خطية بين X , Y

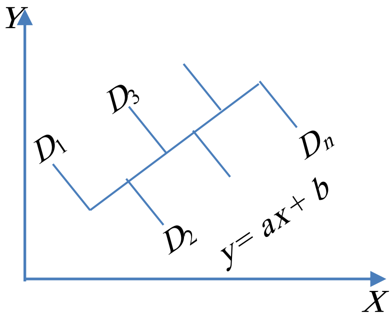
فإذا كان لدينا أزواج من القراءات :

$(x_1, y_1), (x_2, y_2), (x_3, y_3), ..., (x_n, y_n)$ فإنه يمكن تقريب هذه الأزواج بخط مستقيم . وتسمي هذه القراءات (النقاط) بشكل الإنتشار وتقرب هذه النقاط الي خط مستقيم بطريقة المربعات الصغري .

خطوط المربعات الصغري

الحالة الاولي:

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 الصغري وبذلك يتحدد الخط المستقيم تماماً رياضياً .

نفرض أن $D_1, D_2, ..., D_n$ هي الانحرافات الرأسيه لمجموعة البيانات عن الخط المستقيم $y = ax + b$ وتعين بالطريقه التاليه :

$$D_1 = Y_1 - (aX_1 + b)$$
$$D_2 = Y_2 - (aX_2 + b)$$
$$\vdots$$
$$D_n = Y_n - (aX_n + b)$$

ونفرض أن مجموع مربعات هذه الانحرافات هي :

$$s^2 = D_1^{\ 2} + D_2^{\ 2} + ... + D_n^{\ 2}$$
$$s^2 = [Y_1 - (aX_1 + b)]^2 + ... + [Y_n - (aX_n + b)]^2$$

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

$$\frac{\partial s2}{\partial a} = 0 \quad , \qquad \frac{\partial s2}{\partial b} = 0$$

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

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

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

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 الحاله الاولي يمكننا الحصول علي المعادلات الإعتداليه التالية:

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

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

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 $y = 2.3$

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:

$$a = -0.34 \;,\; \text{b} = 4.9$$

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

اذن قيمة y عندما $x = 3.5$ هي:

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

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:

$$a = -0.45 \ , \quad \mathrm{b} = 5.15$$
$$\therefore x = -0.45y + 5.15$$

اذن قيمة x عندما $y = 2.3$ هي:

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

الإرتباط:

عند دراسة العلاقة بين متغيرين x, y فإن شكل الإنتشار يمكن أن يوضح طبيعه هذه العلاقه . فالعلاقة بين X, Y قد تكون قوية جداً إذا ما وقعت نقاط شكل الإنتشار كلها علي منحني أو خط مستقيم وتكون أقل قوة إذا لم تتغير نقاط شكل الإنتشار تماماً بمنحني أو خط مستقيم . سوف نقتصر في دراستنا هنا علي ما يسمي 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 بزياده قيمة X

٢- إرتباط عكسي سالب أي أن قيمة Y تقل بزياده قيمة X

٣- لا يوجد إرتباط أي أن قيمة Y لا تتأثر بتغير قيمة X

معامل الإرتباط بين متغيرين (X, Y)

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

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

حساب معامل الإرتباط

يمكن حساب معامل الإرتباط بإحدي الطريقتين .

١- طريقه بيرسون :

بفرض ان لدينا مجموعة من البيانات :

$(x_1, y_1), (x_2, y_2), (x_3, y_3), ..., (x_n, y_n)$ فإن معامل الإرتباط يعطي بالعلاقة الرياضية التاليه:

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

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 الإنحراف المعياري لقيم y

n : عدد البيانات

المعادلة (1) يمكن كتابتها في الصورة التالية :

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

مثال:

إحسب معامل الارتباط للمثال السابق .

الحل:

بإستخدام المعادله رقم (٢) نحصل علي :

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

$$=\frac{6(71)-(21)(22)}{\sqrt{[6(91)-(21)^2][6(94)-(22)^2]}}$$

$$=\frac{426-462}{\sqrt{(546-441)(564-484)}}$$

$$=\frac{-36}{\sqrt{(105)(80)}}$$

$$=-0.39$$

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- تعين الرتب لقيم y بنفس الطريقه السابقه .

ت- لا تعطي رتباً مختلفه لقيم مكرره في متغير ما بل إذا وجدت قيمتان [أو اكثر] متساويتان فانهما يأخذان رتبة واحده هي الوسط الحسابي للرتبتين (او الرتب)

ث- نحسب فرق الرتب D للقيم المناظرة .

ج- نحسب معامل الإرتباط من خلال العلاقه :

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

حيث n عدد البيانات .

مثال:

إحسب معامل الإرتباط بين المتغيرين (x, y) بإستخدام طريقة سبيرمان للبيانات التالية :

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

الحل:

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

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 ترتيباً تصاعدياً أو تنازلياً .

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

تمارين

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

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 المزدوجة وكذلك باستخدام القطاع الدائري

٢- البيانات التالية تمثل درجات 40 طالب في احد الاختبارات

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

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 المدي الربيعي – P_{85}

٣- احسب الوسط الهندسي والتوافقي والوسيط ونصف المدي الربيعي ومعامل الاختلاف للبيانات 13,11,15,17,9,7,14,12,10 واذا 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 :

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

f_i	10	12	8	7	4	3	8

٥- ثلاثة من مدرسي كره 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 الدراسيه علي الدرجات التاليه: 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 التاليه:

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

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 جنيه

٨- احسب معادله خط انحدار Y علي X وكذلك معامل الارتباط مبينا نوعه للبيانات التالية:

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