محاضرة البث المباشر الثاني

مبادئ الإحصاء

د. فراس حداد

سؤال المحاضرة الثالثة عشر

* احسب التباين والانحراف المعياري والانحراف المتوسط للتوزيع التكراري التالي:

الفئات	التكرار f_i	مركز الفئة x_i	$x_i f_i$	$f_i x_i^2$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$
10-14	12	12	144	1728	10.4	124.8
15-19	9	17	153	2601	5.4	48.6
20-24	8	22	176	3872	0.4	3.2
25-29	5	27	135	3645	4.6	23
30-34	16	32	512	16384	9.6	153.6
Total	50		1120	28230		353.2

$h = 5$, $n = 50$

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{1120}{50} = \boxed{22.4}$$

التباين

$$s^2 = \frac{\left(\sum_{i=1}^{h} f_i x_i^2 - n \bar{x}^2 \right)}{n-1}$$

$$= \frac{28230 - 50(22.4)^2}{50-1}$$

$$= \frac{28230 - 25088}{49} = \boxed{64.122}$$

الانحراف المعياري

$$s = \sqrt{s^2} = \sqrt{64.122} \simeq \boxed{8.008}$$

الانحراف المتوسط

$$M.D = \frac{\sum_{i=1}^{k} |x_i - \bar{x}| \cdot f_i}{n}$$

$$= \frac{353.2}{50} = \boxed{7.064}$$

---

* معامل ارتباط سبيرمان للرتب

يعرف قانون معامل الارتباط للرتب معامل سبيرمان كما يلي :

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}$$

حيث n : عدد الأزواج المرتبة (X, Y) .

d : الفرق بين رتب X ورتب Y .

يستعمل هذا المعامل عندما تكون n عدد الأزواج المرتبة ما بين 25 و 30 او أقل .

مثال: احسب معامل سبيرمان للارتباط بالرتب بين المعدلات التالية لعشرة طلاب في شهادة الدراسة الثانوية والفصل الجامعي الاول:

معدل الطالب في شهادة الثانوية X	(10) 77	(8) 85	(9) 79	(5) 88	(2) 93	(7) 86	(1) 94	(3) 90	(6) 87	(4) 89
معدل الطالب في نهاية الفصل الجامعي Y	61 (10)	72 (7)	65 (9)	71 (8)	80 (3)	74 (6)	82 (1)	81 (2)	76 (5)	78 (4)

الحل: نرتب المعدلات X بحيث نعطي الرتبة 1 لأعلى معدل من بين معدلات X وهكذا للبقية.
ونرتب المعدلات Y بحيث نعطي الرتبة 1 لأعلى معدل بين معدلات Y وهكذا تتابع للبقية

رتب X	رتب Y	الفرق بين الرتب (d)	d^2
10	10	0	0
8	7	1	1
9	9	0	0
5	8	-3	9
2	3	-1	1
7	6	1	1
1	1	0	0
3	2	1	1
6	5	1	1
4	4	0	0
المجموع			14

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

$$= 1 - \frac{6(14)}{10(10^2-1)}$$

$$= 1 - \frac{84}{990}$$

$$= 1 - 0.085 = 0.915$$

نلاحظ في المثال السابق عدم ظهور معدلات متساوية . في حالة وجود بيانات متساوية فيكون تعيين الرتب لهذه البيانات كما يلي :

1) نرتب البيانات كما لو أن ليس فيها بيانات متساوية .

2) نأخذ الوسط الحسابي لرتب كل مجموعة من البيانات المتساوية ونعتبر هذا الوسط الحسابي رتبة كل بيان من هذه المجموعة .

حتى نتعلم كيف نقوم بترتيب البيانات التي تمثل قيم المتغير X والمتغير Y دعونا نأخذ المثال التالي :

مثال : عين الرتب للعلامات التالية :

63, 70, 79, 63, 70, 63, 57, 53, 57, 45, 65
(7) (3) (1) (6) (2) (5) (8) (10) (9) (11) (4)

- نلاحظ ان القيمة 70 مكررة مرتين لذلك نأخذ الوسط الحسابي لرتبتها الاولية فتكون رتبة 70 هي :

رتب 70 هي 2, 3 فنأخذ وسطها الحسابي اي

$$\frac{2+3}{2} = \frac{5}{2} = 2.5$$

فتكون رتبة 70 هو 2.5.

- القيمة 63 مكررة ثلاث مرات ورتبها الاولية هي

5, 6, 7

فيكون وسطهم هو

$$\frac{5+6+7}{3} = \frac{18}{3} = 6$$

∴ رتبة 63 هو 6.

- كذلك القيمة 57 لها الرتب الاولية 8, 9 ووسطهم هو

$$\frac{8+9}{2} = 8$$

فيكون ترتيب البيانات النهائي هو

العلامة	الرتبة
63	6
70	2.5
79	1
63	6
70	2.5
63	6
57	8
53	10
57	8
45	11
65	4

خصائص معامل الارتباط (r)

1) اذا كانت قيمة معامل الارتباط $r = 1$ فإننا نصف الارتباط بين X, Y بانه ارتباط خطي موجب كامل.

2) اذا كانت $r = -1$ كان الارتباط ارتباط خطي سالب كامل.

معنى موجب: اي كلما زادت قيمة المتغير X زادت قيمة المتغير Y.

معنى سالب: اي كلما زادت X نقصت Y. اي العلاقة عكسية.

3) نصف قوة الارتباط عندما $r \neq \pm 1$ كما يلي

الوصف	r
قوي جداً موجب	$0.9 \leq r < 1$
قوي جداً سالب	$-1 < r \leq -0.9$
قوي موجب	$0.5 \leq r < 0.9$
قوي سالب	$-0.9 < r \leq -0.5$
ضعيف موجب	$0 < r < 0.5$
ضعيف سالب	$-0.5 < r < 0$
لا يوجد ارتباط	$r = 0$

معادلة خط الانحدار

اذا كان لدينا عينة من الازواج المرتبة

$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

وصورنا هذه النقاط على المستوى XY نحصل على لوحة الانتشار

ومنها نستدل ان كان يمكن تطبيق خط مستقيم مع شكل الانتشار ام لا .

- اذا فرضنا ان هناك علاقة خطية بين المتغيرين X، Y امكن التعبير عنها بالمعادلة

$$Y = A + BX + \varepsilon$$

حيث ε: الخطأ بالتقدير.

المطلوب هو تقدير A, B
لذلك نفرض ان تقدير A هو a
وتقدير B هو b . نيكون
تقدير y هو

$$\hat{y} = a + bx$$

وهو خط انحدار y على x
الذي حصلنا عليه بتعويض
قيمة a , b .

حيث $$b = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$a = \bar{y} - b\bar{x}$$

حيث $\bar{x}$: الوسط الحسابي لـ

$x_1, \ldots, x_n$

$\bar{y}$: الوسط الحسابي

$y_1, \ldots, y_n$

مثال: اوجد معادلة خط الانحدار

y على x للبيانات في

الجدول التالي

x	y	xy	x^2
4	2	8	16
10	6	60	100
9	8	72	81
12	11	132	144
8	5	40	64
5	4	20	25
48	36	332	430

الحل: معادلة خط الانحدار

$$\hat{y} = a + bx$$

$$b = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

$$a = \bar{y} - b\bar{x}$$

$$\bar{x} = \frac{48}{6} = 8$$

$$\bar{y} = \frac{36}{6} = 6$$

$$\therefore b = \frac{332 - 6(8)(6)}{430 - 6(8)^2}$$

$$= 0.96$$

$$a = 6 - 0.96(8)$$
$$= -1.68$$

∴ معادلة خط الانحدار

$$\hat{Y} = -1.68 + 0.96X$$

* اوجد القيمة التقديرية للمتغير Y عندما X=9

$$\hat{Y} = -1.68 + 0.96(9)$$
$$= 6.96$$

∴

$$e = Y - \hat{Y}$$
$$= 8 - 6.96$$
$$= 1.04$$