



كلية هندسة العمارة

الإحصاء والاحتمالات

Statistics & Probability

المحاضرة
العاشرة

الأستاذ الدكتور محمود محمد ديب طيوب

الانحدار الخطي البسيط

Simple linear Regression

تعريف الانحدار

خط الانحدار ورسم شكل الانتشار Scatter

تحديد معادلة الانحدار

حساب ثوابت معادلة الانحدار بالطريقة الملائمة

استخدام معادلة الانحدار في التنبؤ

تحديد محال الثقة للقيم التقديرية بدلالة المعادلة

تقويم فعالية التكميل

تحليل تباين الانحدار ANOVA

الانحدار Regression

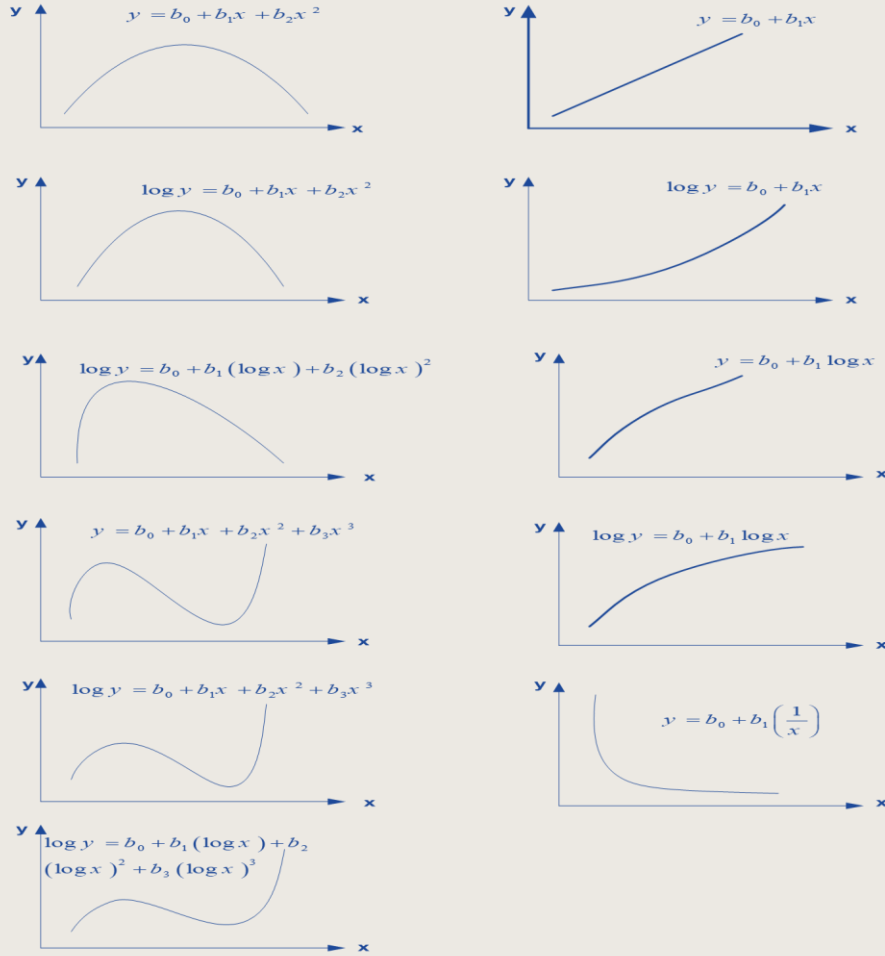
تمهيد

أن معامل الارتباط هو مجرد مقدار يقيس درجة اقتران متغير بمتغير آخر، الانحدار Regression يعد من الموضوعات الإحصائية الهامة التي تتناول مسألة التنبؤ، فالانحدار ظاهرة طبيعية ترجمت إلى مفهوم إحصائي فمن المعروف أن الكثير من الخصائص تتوزع عند الأفراد في مجتمع إحصائي توزيعاً اعتدالياً أو تقترب منه، أي أن نسبة من الأفراد تتجمع حول وسط التوزيع، ...

خط الانحدار:

خط مستقيم تحكمه معادلة إحصائية يمر عبر أو قرب أكبر عدد من النقاط أفضل تمثيل أي خط الملائمة الأفضل Line of best fit أو ما يسمى بخط الانحدار (Regression Line). المعيار المستخدم في تحديد خط الملائمة الأفضل هو انحرافات القيم عن خط معين، فإذا كان مجموع مربعات الانحرافات أقل ما يمكن فإن ذلك الخط هو خط الملائمة المطلوب

وتأخذ العلاقات بشكل عام بين المتغيرات عدة أشكال



يبين أشكال معادلات تمثيل العلاقة بين المتغيرات المدروسة

معادلة خط الانحدار :

يخضع خط الملائمة الأفضل /خط الانحدار/ لمعادلة خط مستقيم تظهر فيها قيم المتغير y والمتغير x وثوابت معادلة الانحدار أي ميل خط الانحدار Slope ونقطة تقاطعه مع المحور y وبما أن قيم y الواقعة على خط الانحدار يقصد بها القيم المتنبأ بها من x فإن معادلة خط الانحدار تأخذ الشكل الآتي:

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

حيث أن:

$\hat{y}$: القيم التقديرية المحسوبة للمتغير التابع التي تقع على خط الانحدار بينما تمثل y القيم الفعلية المشاهدة للمتغير التابع.

b: معامل الانحدار وهو عبارة عن ميل خط الانحدار ويعبر عن معدل التغير في y عندما تتغير قيم x وحدة واحدة.

a: هو معدل قيمة y عندما تكون $x=0$ أي ترتيب نقطة تقاطع مستقيم الانحدار مع محور قيم y (ثابت الانحدار). x : قيم المتغير المستقل.

5-5-6 طرق حساب معادلة خط انحدار y على x باستخدام الدرجات الخام:

نرمز لميل خط الانحدار بـ b والنقطة التي يقطع فيها خط الانحدار محور y بالرمز a

وتكون المعادلة: $\hat{y} = bx + a$

ويحسب الثابت b بالعلاقة الآتية:

$$b = \frac{n \sum x \cdot y - \sum x \cdot \sum y}{n \sum x^2 - (\sum x)^2}$$

أو بالعلاقة الآتية (صيغة كرايمر):

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

أما الثابت a فيحسب بالعلاقة الآتية:

$$a = \frac{\sum y - b \sum x}{n}$$

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

مثال

دراسة العلاقة بين الدخل والانفاق لخمسة أسر /الف وحدة نقدية في الشهر

الفرد	1	2	3	4	5
الدخل x	2	3	7	18	20
الانفاق Y	5	7	6	12	10

المطلوب: أوجد معادلة انحدار y على x.

الحل: نكون الجدول المساعد

1	2	3	4	5	6	7
n	x	y	x ²	y ²	x.y	$\hat{y}$
1	2	5	4	25	10	5.51
2	3	7	9	49	21	5.822
3	7	6	49	36	42	9.20
4	18	12	324	144	216	11.56
5	20	10	400	100	200	11.987
Σ	50	40	786	354	489	40
n=5	Σx	Σy	Σx^2	Σy^2	$\Sigma x.y$	$\Sigma \hat{y}$

ومعادلة انحدار y على x هي: $\hat{y} = bx + a$

ولحساب ثوابت المعادلة نحسب:

$$\bar{x} = \frac{\sum x}{n} = \frac{50}{5} = 10 \quad \text{— المتوسط الحسابي لـ } x$$

$$\bar{y} = \frac{\sum y}{n} = \frac{40}{5} = 8 \quad \text{— المتوسط الحسابي لـ } y$$

حساب الثابت b:

$$b = \frac{n \sum x.y - \sum x . \sum y}{n \sum x^2 - (\sum x)^2}$$

$$b = \frac{5 \times 489 - 50 \times 40}{5 \times 786 - (50)^2} = \frac{2445 - 2000}{3930 - 2500}$$

$$b = \frac{445}{1430} = 0.311$$

ويمكن حساب الثابت b بالعلاقة الآتية:

$$b = \frac{\sum x.y - n \bar{x} \bar{y}}{\sum x^2 - n \bar{x}^2} = \frac{489 - 5 \times 10 \times 8}{786 - 5 \times 10^2}$$

$$b = \frac{89}{286} = 0.311$$

أما الثابت فيحسب بالعلاقة الآتية:

$$a = \frac{\sum y - b \sum x}{n} = \frac{40 - 0.311 \times 50}{5} = 4.89$$

أو تحسب بالعلاقة التالية:

$$a = \bar{y} - b\bar{x} = 8 - 0.311 \times 10 = 4.89$$

وبالتالي تصبح المعادلة على النحو الآتي:

$$\hat{y} = 0.311x + 4.89$$

وهي معادلة انحدار y على x .

ويمكن استخدام هذه المعادلة للتنبؤ بقيم y بمعلومية قيم x العمود رقم (7) جدول:

$$\hat{y}_1 = 0.311 \times 2 + 4.89 = 5.51$$

$$M \quad M \quad M$$

$$y_5 = 0.311 \times 20 + 4.89 = 11.987$$

حساب معادلة انحدار x على y باستخدام الدرجات الخام:

لقد أوجدنا سابقاً معادلة انحدار y على x وقد حدّدنا هذا الخط بحيث يجعل مجموع مربعات المسافات المبينة بالشكل الانتشاري الموازي لمحور y من النقاط إلى خط الانحدار نهاية صغرى. والمشكلة المقدّمة كانت التنبؤ بأقل قدر ممكن من الخطأ بدرجات الاتزان الانفعالي بمعلومية نسب أو معامل الذكاء. أما إذا أردنا التنبؤ بنسب الذكاء بمعلومية درجات الاتزان الانفعالي في هذه الحالة علينا حساب معادلة انحدار x على y وهذا الخط يجب أن يجعل مجموع مربعات المسافات الموازية لمحور x من النقاط إلى خط الانحدار نهاية صغرى.

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

وتستخدم معادلة الانحدار الآتية:

حيث:

$\hat{x}$: قيم x المقدرة بمعلومية قيم y .

b' : معامل الانحدار أي الميل.

y : قيم المتغير y (وهو هنا متغير مستقل).

a' : ترتيب نقطة تقاطع مستقيم الانحدار مع محور x .

وتحسب ثوابت المعادلة بالعلاقة الآتية:

$$b' = \frac{n \sum x.y - \sum x . \sum y}{n \sum x^2 - (\sum x)^2}$$

أو بالعلاقة:

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

أما الثابت a' فيحسب بالعلاقة:

$$a' = \frac{\sum x - b' \sum y}{n}$$

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

أو بالعلاقة:

مثال

بالعودة إلى معطيات المثال السابق نجد أن:

$$\begin{array}{llll} \Sigma y^2 = 354 & \Sigma x = 50 & \Sigma y = 40 & n=5 \\ \bar{y} = 8 & \bar{x} = 10 & \Sigma x \cdot y = 489 & \Sigma x^2 = 786 \end{array}$$

وبتطبيق العلاقات السابقة نجد أن:

$$\begin{aligned} b' &= \frac{5 \times 489 - 40 \times 50}{5 \times 354 - (40)^2} \\ b' &= \frac{445}{170} = 2.62 \\ b' &= \frac{489 - 10 \times 8 \times 5}{354 - 5 \times 8^2} = \frac{89}{34} = 2.62 \\ a' &= \frac{50 - 2.62 \times 40}{5} = -10.96 \\ a' &= 10 - 2.62 \times 8 = -10.96 \end{aligned}$$

وتصبح المعادلة على النحو الآتي:

$$\hat{x} = 2.62y - 10.96$$

وهي معادلة انحدار x على y .

ويمكن استخدام هذه المعادلة للتنبؤ بقيم x بمعلومية قيم y فمثلاً:

$$\hat{x}_1 = 2.62 \times 5 - 10.96 = 2.147$$

$$M \quad \quad \quad M \quad \quad \quad M$$

$$\hat{x}_5 = 2.62 \times 10 - 10.96 = 15.23$$

ويتضح مما سبق أن خط انحدار y على x يختلف عن خط انحدار x على y وكل منهما يعبر عن علاقة تقريبية بين المتغيرين. ولكنهما ينطبقان ويصبحان خطاً واحداً إذا كان معامل الارتباط بين المتغيرين تاماً ($r = 1$) أما إذا لم يكن الارتباط تاماً فإنه يمكن أن نضرب المعادلتين السابقتين

ونحسب معامل الارتباط كما يلي:

$$r = \sqrt{b' \times b}$$

$$r = \sqrt{2.62 \times 0.311} = \sqrt{0.81482} = 0.903$$

ويمكن حساب معامل الارتباط بين المتغيرين بالعلاقة الآتية:

$$\begin{aligned} r &= \frac{n \sum x . y - \sum x . \sum y}{\sqrt{\left[n \sum x^2 - (\sum x)^2 \right] \left[n \sum y^2 - (\sum y)^2 \right]}} \\ &= \frac{5 \times 489 - 50 \times 40}{\sqrt{\left[5 \times 786 - (50)^2 \right] \left[5 \times 354 - (40)^2 \right]}} \\ &= \frac{2445 - 2000}{\sqrt{1430 \times 170}} = \frac{445}{\sqrt{243100}} = \frac{445}{439.06} = 0.903 \end{aligned}$$

إيجاد معادلتى خطى الانحدار باستخدام طريقة الانحرافات:

يمكن إيجاد معادلتى خطى الانحدار بالعلاقتين التاليتين:

$$b = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

في معادلة انحدار y على x

أما الثابت a فيحسب كما شاهدنا سابقاً أي $a = \bar{y} - b\bar{x}$ أما في معادلة انحدار x على y فتحسب

بالعلاقة الآتية:

$$b' = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\Sigma(y - \bar{y})^2}$$

أما الثابت a' فيحسب بالعلاقة الآتية:

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

- حساب معادلة انحدار y على x باستخدام معامل الارتباط لمعلوم والانحراف المعياري لكل من المتغيرين:

- الانحراف المعياري للمتغير x يساوي:

$$\sigma_x = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$$

- الانحراف المعياري للمتغير y يساوي:

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - \bar{y}^2}$$

ويحسب معامل الارتباط بأي من العلاقات المعروفة والمستخدمه لحسابه ومنه:

$$b = r \frac{\sigma_y}{\sigma_x}$$

أما الثابت a فيحسب بالعلاقة: $a = \bar{y} - b\bar{x}$

وذلك في معادلة انحدار y على x أي $\hat{y} = bx + a$

وبالتعويض عن قيمة كل من b و a في معادلة انحدار y على x وهي:

نجد أن $\hat{y} = bx + a$

$$\hat{y} = \left(r \frac{\sigma_y}{\sigma_x} \times x \right) + \left(\bar{y} - \frac{\sigma_y}{\sigma_x} \times \bar{x} \right)$$

$$\hat{y} = \bar{y} + r \frac{\sigma_y}{\sigma_x} (x - \bar{x})$$

لتوضيح كيفية تطبيق معادلة خط الانحدار باستخدام معامل الارتباط من المعطيات الآتية:

اختبار آخر العام	اختبار نصف العام
$\bar{y} = 80$	$\bar{x} = 78$
$\sigma_y = 8$	$\sigma_x = 5$

إذا علمت أن قيمة **معامل الارتباط بين الاختبارين $r = 0.60$** وللحصول على الدرجة المتنبأ بها يجب أن نحصل على معادلة خط انحدار y على x لأننا نود التنبؤ بدرجة الطالب في اختبار آخر السنة y بمعلومية درجته في اختبار نصف العام x وفق العلاقة الآتية:

$$\hat{y} = \bar{y} + r \frac{\sigma_y}{\sigma_x} (x - \bar{x})$$

مع العلم أن درجة الطالب في اختبار نصف العام في إحد المقررات كانت $x=68$

$$\hat{y} = 80 + 0.60 \frac{8}{5} (68 - 78) = 70.40$$

إذا حصل الطالب على الدرجة 85 في اختبار نصف العام. ما الدرجة المتنبأ بها في اختبار نهاية العام وذلك وفق معطيات المثال السابق.

$$\hat{y} = 80 + 0.60 \frac{8}{5} (85 - 78) = 86.72$$

* لإيجاد معادلة انحدار x على y باستخدام معامل الارتباط يمكن استخدام المعادلة الآتية:

$$\hat{x} = \bar{x} + r \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

وبتطبيق معادلتني خط الانحدار على معطيات مثال نسب الذكاء والانتزان الانفعالي نجد أن:

$$\hat{y} = \bar{y} + r \frac{\sigma_x}{\sigma_y} (x - \bar{x})$$

إذا علمت أن:

$$\bar{y} = 8$$

$$r = 0.903$$

$$\bar{x} = 10$$

$$\sigma_x = \sqrt{\frac{786}{5} - 10^2} = 7.56$$

$$\sigma_y = 2.61$$

وبتطبيق المعادلة نجد أن:

$$\begin{aligned}\hat{y} &= 8 + 0.902 \frac{2.61}{7.56} (x - 10) \\ &= 8 + 0.311(x - 10)\end{aligned}$$

$$\hat{y} = 8 + 0.311x - 3.11$$

$$\hat{y} = 0.311x + 4.89$$

وهي معادلة انحدار y على x فإذا كانت $x=2$ فإن قيمة $\hat{y}$ المقدرة تساوي $\hat{y} = 0.311 \times 2 + 4.89 = 5.51$

مثال

حصلت مجموعة من الطلاب على درجات في اختبار مادة الرياضيات (x) واختبار في الاحصاء (y) وكانت المعطيات الآتية:

$$\begin{aligned} \bar{x} &= 10 & \Sigma y &= 84 & n &= 7 & \Sigma x &= 70 \\ \sigma_y &= 4.1 & \sigma_x &= 4.81 & \bar{y} &= 12 \end{aligned}$$

$r = 0.90$ معامل الارتباط بين المتغيرين

المطلوب: حساب انحدار y على x وذلك إذا كانت قيمة $x = 15$.

الحل:

باستخدام معادلة الانحدار y على x نجد ما يلي:

$$\hat{y} = r \frac{\sigma_y}{\sigma_x} (x - \bar{x}) + \bar{y}$$

$$\hat{y} = 0.90 \frac{4.1}{4.81} (15 - 10) + 12$$

$$\hat{y} = (0.90 \times 0.85 \times 5) + 12$$

$$\hat{y} = 3.82 + 12 = 15.82$$

ويمكن أيضاً استنتاج قيمة x من قيمة y من المعادلة الآتية:

$$\hat{x} = r \frac{\sigma_x}{\sigma_y} (y - \bar{y}) + \bar{x}$$

$$\hat{x} = 0.90 \frac{4.81}{4.1} (y - 12) + 10$$

$$\hat{x} = 0.90 \times 1.17 (y - 12) + 10$$

$$\hat{x} = 0.90 \times (1.17y - 14.04) + 10$$

$$\hat{x} = 1.053y - 12.63 + 10$$

$$\hat{x} = 1.053y - 2.63$$

إذا افترضنا أن قيمة $y = 15.82$ فنجد أن قيمة $\hat{x}$ تساوي:

$$\hat{x} = 1.053y \times 15.82 - 2.63 = 14.028$$

الخطأ المعياري للتقدير:

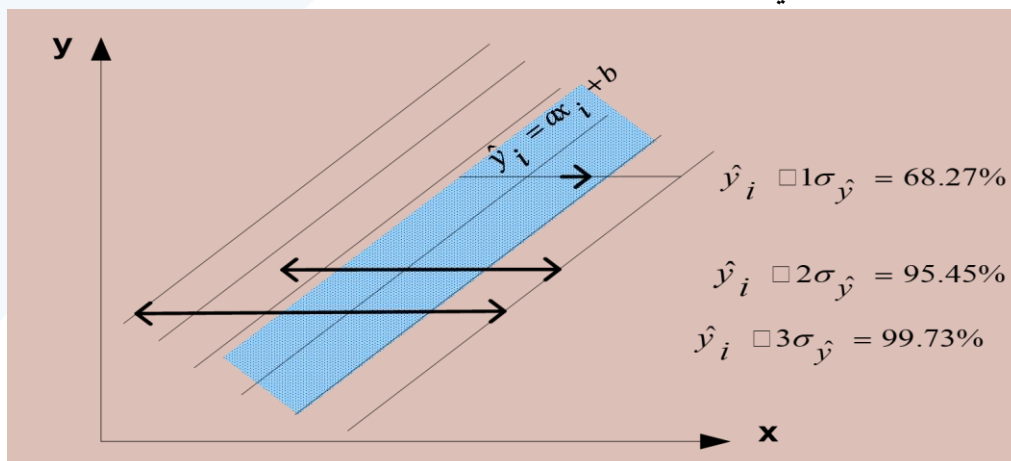
هو عبارة عن مقياس مدى تشتت القيم الفعلية حول خط الانحدار مقياس لدرجة الدقة في حساب القيم المقدرة للمتغير التابع. يستخدم الخطأ المعياري للتقدير كمقياس لدرجة الدقة في تقدير قيم المتغير التابع وحيث أن التشتت حول الوسط الحسابي يقاس بالجذر التربيعي لمتوسط مربعات الانحرافات عن الوسط الحسابي. فإنه يمكننا حساب التشتت حول خط الانحدار باستخدام نفس الأسلوب أي بمتوسط مربعات انحرافات القيم الفعلية عن الخط. أي أنه يمكن حساب الخطأ المعياري لمعادلة انحدار y على x بالصيغة الرياضية التالية:

$$\sigma_{\hat{y}} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}$$

حيث أن: $\sigma_{\hat{y}}$ الخطأ المعياري للتقدير؛ y_i القيم الفعلية؛ $\hat{y}_i$ القيم المقدرة باستخدام معادلة الانحدار؛ n عدد القيم. ونلاحظ أن للخطأ المعياري نفس خصائص الانحراف المعياري ويفسر كما يفسر تماماً الانحراف المعياري. فلو رسمنا على جانبي مستقيم الانحدار خطين موازيين على بعد واحدات من الخطأ المعياري نلاحظ أنه:

- على بعد واحد خطأ معياري بين هذين الخطين لوقع 68.27% من نقاط الانتشار.
- على بعد 2 خطأ معياري بين هذين الخطين لوقع 95.45% من نقاط الانتشار.
- على بعد 3 خطأ معياري بين هذين الخطين لوقع 99.73% عن نقاط الانتشار.

وذلك وفق الشكل التالي :



***الطريقة المختصرة لحساب الخطأ المعياري للتقدير:**

إن حساب الخطأ المعياري بالطريقة المباشرة (أسلوب الفروقات) طويل ويحتاج إلى عمليات حسابية معقدة ويمكن اختصار ذلك باستخدام معادلة تباين الخطأ المعياري للتقدير وفق العلاقة الآتية:

$$\sigma_{\hat{y}}^2 = \frac{\sum y_i^2 - a \sum y_i - b \sum x_i y_i}{n - 2}$$

وبأخذ الجذر التربيعي نحصل على الخطأ المعياري للتقدير

لنعود إلى معطيات المثال (الذكاء × الاتزان الانفعالي) نجد أن الخطأ المعياري للتنبؤ يساوي:

$$\begin{aligned}\sigma_{\hat{y}} &= \sqrt{\frac{\sum y^2 - a \sum y - b \sum x \cdot y}{n - 2}} \\ \sigma_{\hat{y}} &= \sqrt{\frac{354 - 4.89 \times 40 - 0.311 \times 4.89}{5 - 2}} \\ &= \sqrt{\frac{6.321}{3}} = 1.4516\end{aligned}$$

ويستخدم الخطأ المعياري للتقدير في تحديد مجال الثقة للقيم المقدرة باحتمال معين.

مثال

بالعودة إلى معطيات المثال السابق أوجد قيمة $\hat{y}$ المقدرة إذا كانت $x = 25$

$$\hat{y}_{25} = 0.311 \times 25 + 4.89 = 12.67$$

أوجد مجال الثقة للقيمة المقدرة باحتمال 95% إذا علمت أن القيمة الجدولية 1.96:

$$P \left(\hat{y} - Z_{1-\frac{\alpha}{2}} \sigma_{\hat{y}} \leq y \leq \hat{y} + Z_{1-\frac{\alpha}{2}} \sigma_{\hat{y}} \right) = 95\%$$

$$P (12.67 - 1.96 \times 1.4516 \leq y \leq 12.67 + 1.96 \times 1.4516) = 0.95$$

$$P (12.67 \pm 2.845) = 0.95$$

$$P (9.825 ; 15.515) = 0.95$$

ويمكن حساب الخطأ المعياري للتنبؤ بدرجات اختبار فهم المقروء y بمعلومية درجات اختبار القبول

$$\sigma_{\hat{y}} = \sigma_y \sqrt{1-r^2} \quad \text{بالعلاقة الآتية:}$$

مثال

لتكن لدينا المعطيات عن المتغير x و y :

اختبار القبول (x)	اختبار فهم المقروء (y)
$\bar{x} = 32.24$	$\bar{y} = 27.12$
$\sigma_x = 11.82$	$\sigma_y = 9.45$

ومعامل الارتباط بين المتغيرين $r = 0.78$.

أوجد الخطأ المعياري للتنبؤ بقيم y بمعلومية قيم x يكون:

$$\sigma_{\hat{y}} = \sigma_y \sqrt{1 - r^2} = 9.45 \sqrt{1 - 0.78^2}$$

$$\sigma_{\hat{y}} = 9.45 \times 0.6258 = 5.91$$

هذا وتجدر الملاحظة إلى أن استخدام الخطأ المعياري في التنبؤ يتطلب أن تحقق بعض الفروض في البيانات وهي:

- 1- أن تكون العينة ممثلة للمجموعة التي ستطبق عليها معادلة الانحدار.
- 2- أن تكون أخطاء التنبؤ موزعة توزيعاً اعتدالياً.
- 3- أن تكون أخطاء التنبؤ موزعة توزيعاً اعتدالياً على جميع نقاط خط الانحدار /تجانس التباين/.

-تصحيح الخطأ المعياري في التنبؤ:

ربما يكون من الأفضل تصحيح تقدير الخطأ المعياري في التنبؤ إذا استخدم الباحث عينة قليلة العدد ($n < 50$) قبل أن يعمم على هذا التقدير على المجتمع الأصلي الذي سميت منه العينة ويمكن إجراء التصحيح بالعلاقة الآتية:

الخطأ المعياري في التنبؤ بقيم y بمعلومية قيم x بعد تصحيحه:

$$\sigma_{\hat{y}c} = \sigma_{\hat{y}} \times \sqrt{\frac{n}{n - 2}}$$

حيث أن:

$\sigma_{\hat{y}}$: الخطأ المعياري قبل التصحيح.

n : عدد أفراد العينة.

كما يمكن إجراء هذا التصحيح عند حساب الخطأ المعياري في التنبؤ بقيم y بمعلومية قيم x بالعلاقة الآتية أيضاً:

$$\sigma_{\hat{y}_c} = \sigma_y \times \sqrt{(1-r^2) \left(\frac{n}{n-2} \right)}$$

حيث أن:

σ_y : الانحراف المعياري للمتغير y .

r : معامل الارتباط.

n : عدد أفراد العينة.

تقويم جودة التمثيل والتنبؤ بواسطتها:

إن حصولنا على معادلة لتمثيل العلاقة بين المتغيرين x و y لا يعني نهاية المطاف، وإنه يجب معرفة مدى فعالية هذه المعادلة في تمثيل العلاقة المذكورة ولذلك لابد من إيجاد مقياس أو معيار لقياس مدى فعالية التمثيل، ومن البديهي أن يكون أحد عناصر ذلك القياس مدى تطابق قيم y الفعلية مع قيم $\hat{y}$ النظرية ولبيان ذلك نقوم بحساب عدة تباينات أهمها:

1- التباين غير المفسر بالعلاقة الآتية:

$$S_{y\hat{y}}^2 = \frac{1}{n} * \sum (y - \hat{y})^2$$

- فإذا كانت يعني $S_{y\hat{y}}^2$ أن هناك تطابق تام بين قيم y الفعلية وقيم $\hat{y}$ النظري أي: $y = \hat{y}$ وهذا يعني أن تباين التمثيل $S_{y\hat{y}}^2$ يكون معدوماً وبالتالي تكون العلاقة بين المتغيرين x و y علاقة تابعة وكذلك جميع تغيرات y تفسر بواسطة المتغير x وحده.

- فإذا كانت قيمة $S_{y\hat{y}}^2$ معدومة فهذا يعني أن قيم y الفعلية لا تتطابق مع قيم $\hat{y}$ النظرية وكلما كانت قيمة $S_{y\hat{y}}^2$ كبيرة يعني ذلك إما أن الارتباط ضعيف أو أن المعادلة غير ملائمة وبالتالي عدم التطابق ينتج عن عوامل أخرى غير مفسرة ولذلك سمي التباين غير المفسر $\text{variance non expliquée}$.

-التباين المفسر ويعرف بالعلاقة الآتية:

$$S^2_{\hat{y}\bar{y}} = \frac{1}{n} \sum_{i=1}^n (\hat{y} - \bar{y})^2$$

وهو عبارة عن تباين القيم النظرية لـ $\hat{y}$ عن المتوسط الحسابي العام للمتغير التابع y .

- فإذا كانت قيمة $S^2_{\hat{y}\bar{y}} = 0$ أي أن جميع قيم $\hat{y} = \bar{y}$ أي أن جميع قيم $\hat{y}$ تساوي قيمة ثابتة $\bar{y}$ وهذا يعني أن y غير مرتبطة بـ x .

- فإذا كانت قيمة $S^2_{\hat{y}\bar{y}}$ غير معدومة وكلما كانت كبيرة كان الارتباط بين x و y ارتباط متين وقوي ودال ومفسراً بواسطة المتغير x .

وهذا يعني أن المعادلة المستخدمة لدراسة العلاقة بين المتحولين مناسبة وفعالة ولذلك سمي بالتباين المفسر $\text{variance expliquée}$.

3-التباين الكلي:

هو عبارة عن مجموع التباين غير المفسر والتباين المفسر أي:

$$\sigma_y^2 = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2 + \frac{1}{n} \sum_{i=1}^n (\hat{y} - \bar{y})^2$$

ويجب أن يساوي الى تباين القيم الفعلية للمتغير التابع y أي:

$$\frac{1}{n} \sum_{i=1}^n (y - \bar{y})^2 = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2 + \frac{1}{n} \sum_{i=1}^n (\hat{y} - \bar{y})^2$$

- حساب معامل التحديد:

هو عبارة عن نسبة التباين المفسر على التباين الكلي:

$$D = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{y} - \bar{y})^2}{\frac{1}{n} \sum_{i=1}^n (y - \bar{y})^2} = \frac{S_{\hat{y}\bar{y}}^2}{\sigma_y^2}$$

وقيمته تتراوح بين $0 \leq D \leq 1$ وكلما كانت قيمته قريبة من الواحد الصحيح كان التمثيل بواسطة المعادلة تمثيلاً فعالاً.

ونظراً لحاجتنا إلى تباين التمثيل $S_{y\hat{y}}^2$ فإننا نعوض قيمة $S_{y\hat{y}}^2$ من المعادلة (6-45) فإننا نجد أن:

$$D = \frac{\sigma_y^2 - S_{y\hat{y}}^2}{\sigma_y^2} = 1 - \frac{S_{y\hat{y}}^2}{\sigma_y^2}$$

وهي العلاقة الأكثر استخداماً لحساب معامل التحديد D.

وتجدر الإشارة الى أنه إذا كانت العلاقة خطية فإن معامل التحديد D يرتبط بمعامل الارتباط $D = r_{yx}^2$.

حساب خطأ التمثيل (الخطأ المعياري):

عند تمثيل سلسلة ارتباطية بواسطة معادلة معينة فإن مقدار الخطأ الناتج عن ذلك يسمى بخطأ التمثيل ويحسب بأخذ الجذر التربيعي للتباين غير المفسر وفق العلاقة الآتية:

$$S_{y\hat{y}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2}$$

مثال

يبين الجدول الآتي قيم المتغيرين x و y :

n	1	2	3	4	5
x	2	3	7	18	20
y	5	7	6	12	10

المطلوب:

1- حساب معادلة انحدار y على x وفق المعادلة: $\hat{y} = bx + a$.

2- تقويم فعالية التمثيل.

3- حساب معامل التحديد D .

4- التنبؤ بقيمة $\hat{y}$ عندما $x=28$.

5- حدد مجال الثقة للقيمة المتنبأ بها عند مستوى دلالة 5%, اذا علمت أن القيمة الجدولية $Z_{1-\frac{\alpha}{2}} = 1.96$

الحل:

ننشئ الجدول المساعد:

n	x	y	x^2	y^2	x.y	القيم النظرية $\hat{y}$	$(y - \hat{y})^2$	$(\hat{y} - \bar{y})^2$
1	2	5	4	25	10	5.511	$-5.511)^2=0.2611$ (5	$-8)^2=6.196$ (5.511
2	3	7	9	49	21	5.821	1.390041	4.748041
3	7	6	49	36	42	7.07	1.1449	0.8649
4	18	12	324	144	216	10.49	2.2801	6.2001
5	20	10	400	100	200	11.11	1.2321	9.6721
Σ	50	40	786	354	489	40	6.308241	27.681141

حساب ثوابت المعادلة $\hat{y} = bx + a$

$$b = \frac{\sum x \cdot y - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} = \frac{498 - 5 \cdot 50 \cdot 40}{786 - 5 \cdot 10^2} = 0.311$$

$$a = \bar{y} - b\bar{x} = 8 - 0.311 \cdot 10 = 4.89$$

وبذلك تصبح المعادلة: $\hat{y} = 0.311x + 4.89$ وهي معادلة انحدار y على x .

- حساب القيم النظرية لـ $\hat{y}$ بدلالة المعادلة حيث نعوض x بقيمتها على النحو الآتي:

$$\hat{y}_1 = 0.311 \cdot 2 + 4.89 = 5.5104$$

وهكذا بالنسبة لباقي قيم x وترتب في الجدول.

- حساب التباين غير المفسر:

$$S^2_{y\hat{y}} = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2 = \frac{1}{5} (6.308241)^2 = 1.26165$$

- حساب التباين المفسر:

$$S^2_{\bar{y}\hat{y}} = \frac{1}{n} \sum_{i=1}^n (\hat{y} - \bar{y})^2 = \frac{1}{5} (27.681141)^2 = 5.53623$$

- حساب التباين الكلي:

$$\sigma_y^2 = S^2_{y\hat{y}} + S^2_{\bar{y}\hat{y}} = 1.26165 + 5.53623 = 6.79788$$

- حساب معامل التحديد D:

$$D = \frac{S^2_{\bar{y}\hat{y}}}{\sigma_y^2} = \frac{5.53623}{6.79788} = 0.8144$$

ويمكن حساب معامل التحديد بواسطة معامل الارتباط، إذا علمت أن قيمة معامل الارتباط بين المتغيرين

x و y يساوي $r_{yx} = 0.902$ ومنه يكون معامل التحديد يساوي :

$$D = r^2 = 0.902^2 = 0.814$$

-حساب خطأ التمثيل ويساوي الجذر التربيعي للتباين غير المفسر:

$$S_{y\hat{y}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2} = \sqrt{1.26165} = 1.12323$$

-التنبؤ بقيمة $\hat{y}$ عندما تكون $x=28$ من معادلة التمثيل نجد أن:

$$\hat{y}_{28} = 0.311 * 28 + 4.89 = 13.598$$

حساب مجال الثقة للقيمة المتنبأ بها عند مستوى دلالة 5% نجد أن:

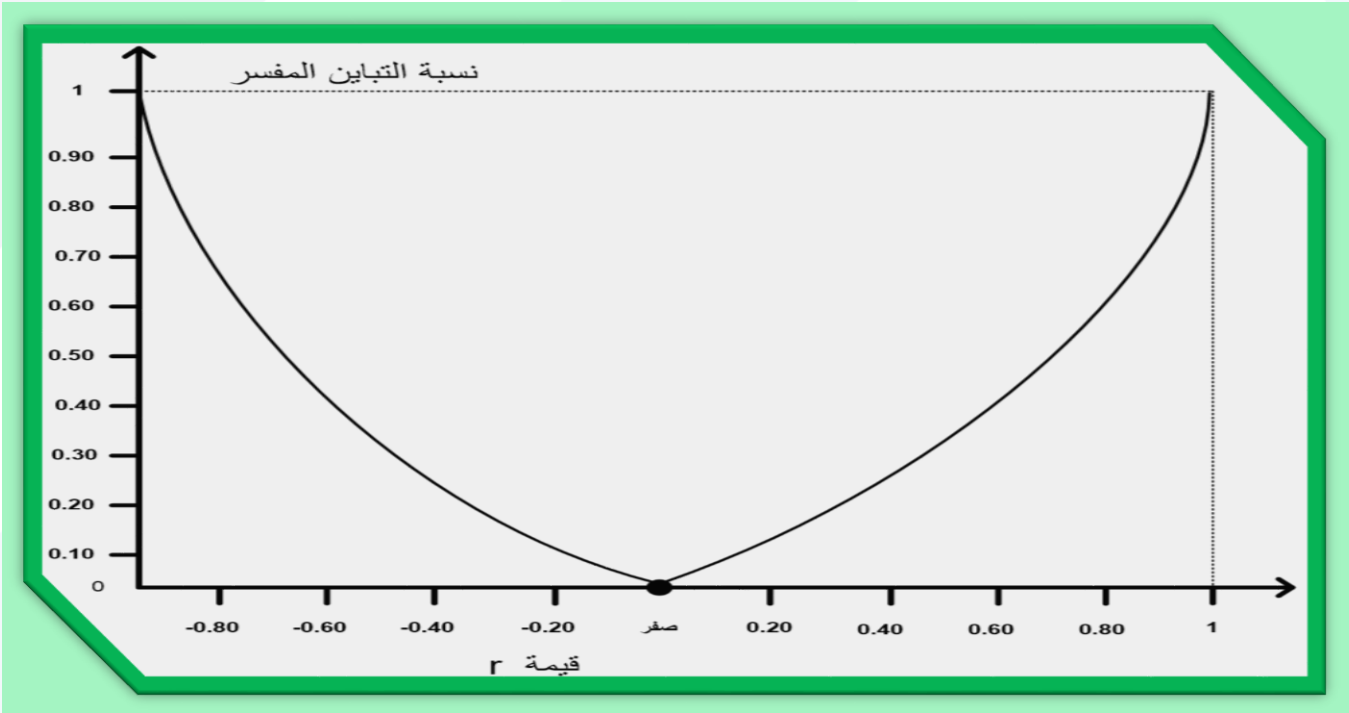
$$p \left(\hat{y}_{28} - Z_{1-\frac{\alpha}{2}} S_{y\hat{y}} \leq \hat{y} \leq \hat{y}_{28} + Z_{1-\frac{\alpha}{2}} S_{y\hat{y}} \right) = 0.05$$

$$p (13.598 - 1.96 * 1.12323 \leq \hat{y} \leq 13.598 + 1.96 * 1.12323) = 0.05$$

$$p (13.598 \text{ m } 2.2015308) = 0.05$$

$$p (11.39647 \text{ و } 15.7995308) = 0.05$$

وهو مجال الثقة للقيمة المتنبأ بها لـ $\hat{y}$ بدلالة قيمة x .



شكل(4-6): نسبة تباين أحد المتغيرين الذي يمكن تفسيره بمعلومية تباين الآخر عندما تأخذ (r) معامل الارتباط قيماً مختلفة

ونظراً لأن r^2 تمثل نسبة التباين الذي يمكن تفسيره فإن $1-r^2$ تمثل نسبة التباين الذي لا نستطيع تفسيره بمعلومية الارتباط بين x و y لذلك يسمى المقدار $1-r^2$ بمعامل الاغتراب ونرمز له بالرمز k^2 أي أن k^2 تمثل نسبة التباين في المتغير y الذي يلزم تفسيره بمعلومية متغيرات أخرى تختلف عن المتغير x . ويمكن تلخيص العلاقة بين k^2 , r^2 كالتالي:

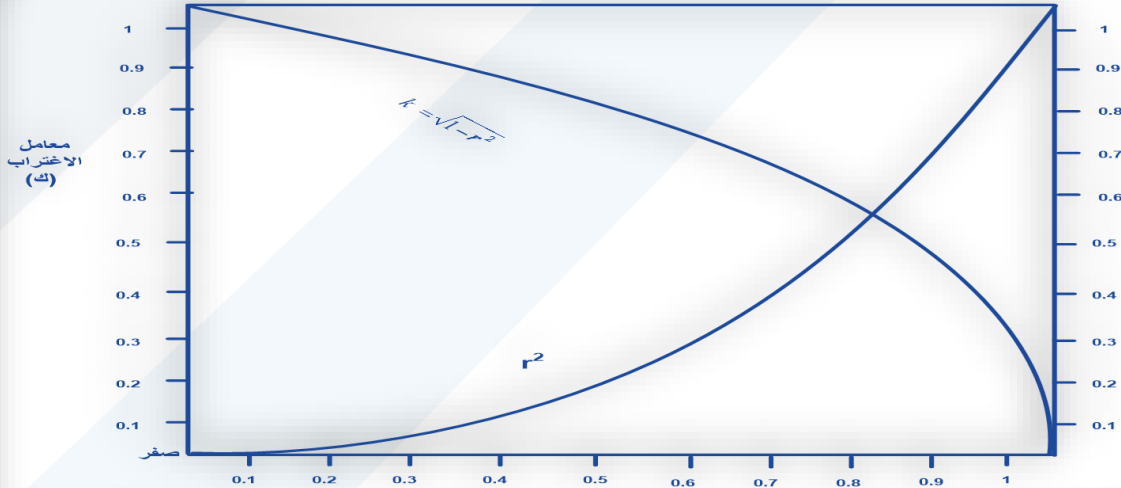
$$k^2 = 1 - r^2$$

$$k^2 + r^2 = 1$$

ويتضح من العلاقة $k^2 + r^2 = 1$ أن مجموع مربعي كل من k و r يساوي الواحد الصحيح فإذا كانت $r=0.50$ فإن k لا تساوي 0.50 إنما

$$\begin{aligned} k &= \sqrt{1-r^2} = \sqrt{1-0.50^2} \\ &= \sqrt{1-0.25} = \sqrt{0.75} = 0.86603 \end{aligned}$$

وإذا كانت $r=0.7071$ فإن $k=0.7071$ أيضاً وهنا فقط تكون $r^2 + k^2 = 0.50 + 0.50 = 1$ أي عندما تكون $r=0.7071$ فإنه يتساوى وجود علاقة مع عدم وجودها. ويمكن تمثيل العلاقة بين (r) , (k) , $(1-k)$ بالشكل التالي:



شكل (5-6): العلاقة بين معامل الارتباط r ومعامل التحديد r^2 ومعامل الاغتراب k^2

والحقيقة تدل العلاقة $r^2 + k^2 = 1$ على معادلة دائرة مركزها نقطة الأصل، ونصف قطرها الوحدة وقد اقتصر الشكل على تمثيل القيم الموجبة فقط لكل (r) , (k) .

تقدير تباين الانحدار . / تحليل تباين في النموذج الخطي البسيط: ANOVA

ذكرنا سابقاً أن معادلة الانحدار من العينة هي: $\hat{y} = bx + a$

هي تقدير لمعادلة انحدار y على x في المجتمع ذي البعدين.

وباستعمال معادلة الانحدار من العينة نستطيع تقدير قيم y الفعلية بالقيم التقديرية $\hat{y}$ ولكي نحكم على جودة التقدير نحتاج إلى معرفة مدى دقة هذا التقدير، وبما أن قيمة a ، b حسبت من عينة من الأزواج المرتبة حجمها n فإن هذه القيم تخضع لتغير المعاينة وبالتالي عند استعمال $\hat{y}$ كتقدير لقيمة y الفعلية لا بد من معرفة درجة الدقة المنوطة بهذا التقدير وبالتالي لا بد من معرفة حسن المطابقة أي حسن مطابقة خط الانحدار على النقاط في شكل الانتشار وللإجابة على هذا السؤال لا بد من التعرف على:

1- مجموعة المربعات الكلية (SST) Sum of Squares:

وهي عبارة عن مجموع مربعات انحرافات قيم y الفعلية عن وسطها الحسابي $\bar{y}$ ويرمز لها بـ SST:

$$SSt = \sum (y_i - \bar{y})^2$$

وبتقسيم هذا المجموع على عدد درجات الحرية $n-1$ يعطي تقديراً للتباين الكلي لقيم y الفعلية عن وسطها الحسابي.

الانحرافات الكلية:

$$SSt = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum y_i^2 - n \bar{y}^2$$

الانحرافات الكلية

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = b^2 \left[\sum x^2 - n \bar{x}^2 \right]$$

$$= b \left[\sum x.y - n \bar{x} \bar{y} \right]$$

الانحرافات بسبب الانحدار

وتحسب b بالعلاقة التالية:

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

أو بالعلاقة التالية:

$$\sum y_i = b \sum x + an$$

$$\sum x.y = b \sum x^2 + a \sum x$$

$$SSE = SST - SSR$$

البواقي

مثال

لنعود إلى معطيات مثال نسب الذكاء والاتزان الانفعالي:

$$\sum y^2 = 354 \quad \sum y = 40 \quad \sum x^2 = 786 \quad \sum x = 50 \quad n=5$$

$$a=4.81 \quad b=0.311 \quad \bar{y} = 8 \quad \bar{x} = 10 \quad \sum x.y = 489$$

$\hat{y} = 0.311x + 4.81$ - حساب مجموع مربعات الانحرافات الكلية:

$$SS_t = \sum_{i=1}^n (y - \bar{y})^2 = \sum y^2 - n\bar{y}^2$$

$$SS_t = 354 - 5 \times 8^2 = 34$$

- مجموع مربعات الانحرافات بسبب الانحدار:

$$SS_r = \sum_{i=1}^n (\hat{y} - \bar{y})^2 = b^2 [\sum x^2 - n\bar{x}^2]$$

$$= b [\sum x.y - n\bar{x}\bar{y}]$$

$$SS_r = 0.311^2 [786 - 5 \cdot 10^2] = 27.67$$

$$SS_r = 0.311 [489 - 5 \times 10 \times 8] = 27.67$$

- مجموع مربعات انحرافات الأخطاء:

$$SS_e = \sum_{i=1}^n (y - \hat{y})^2 = SS_t - SS_r$$

$$SS_e = 34 - 27.67 = 6.321$$

وترتب النتائج في جدول تحليل ANOVA كما يلي:

مصدر التباين S.O.V	درجات الحرية d.f	مجموع المربعات SS	متوسط مربعات الانحرافات MS	$\hat{F}$
بسبب الانحدار SSR	a-1 2-1=1	27.67	$M_{sx} = \frac{27.67}{1} = 27.67$	$\hat{F} = \frac{M_{sr}}{M_{se}}$ $\hat{F} = \frac{27.67}{2.107} = 13.13$
بسبب الأخطاء SSE	n-2 5-2=3	6.321	$M_{sx} = \frac{6.321}{3} = 2.107$	
SST الكلي	N-1 5-1=4	34		

القرار: إن القيمة الجدولية لفisher عند مستوى دلالة 5% درجات حرية (1.3) $F_{0.01\alpha}(1.3) = 10.12$ بما أن قيمة F المحسوبة أكبر من قيمة F_{α} نرفض الفرضية الابتدائية H_0 وبالتالي فإن الانحدار أو الارتباط بين المتغيرين ارتباط حقيقي ودال إحصائياً بمعنى كلما ارتفع معامل الذكاء لدى الفرد ازداد اتزاناً.

نهاية محاضرة
الانحدار

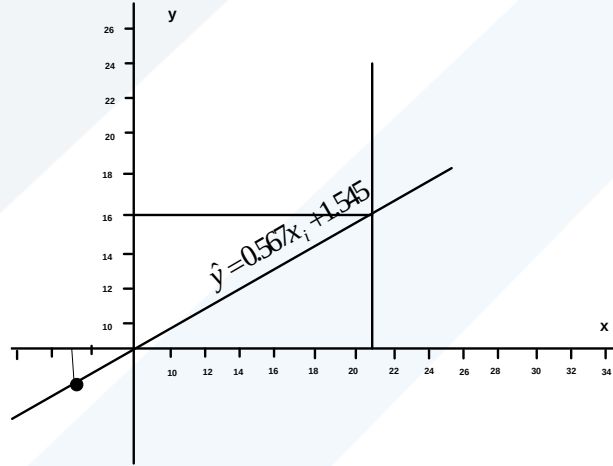
تمارين محلولة في الانحدار والارتباط

يبين الجدول الآتي دخل (x) وانفاق (y) عشر أسر:

الأسر	1	2	3	4	5	6	7	8	9	10	Σ
الدخل x	36	20	18	14	15	30	17	22	16	32	
الإنفاق y	20	14	12	10	10	24	14	18	12	26	

والمطلوب:

- 1- تحديد شكل الانتشار.
- 2- حساب معامل الارتباط الخطي لبيرسون.
- 3- حساب معادلة انحدار y على x.
- 4- حساب معادلة انحدار x على y.



شكل (6-7) يبين مستقيم انحدار y على x (تقريبي)

$$\hat{y} = 0.567x_i + 1.545$$

$$x = 0 \Rightarrow \hat{y} = 1.5452 \approx 1.60$$

$$\hat{y} = 0 \Rightarrow x = \frac{1.60}{0.567} = -2.43$$

الحل: ننشئ الجدول المساعد:

i	x	y	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	x^2	y^2	$x.y$
1	36	20	14+	196	4+	16	56+	1296	400	720
2	20	14	2-	4	2-	4	4+	400	196	280
3	18	18	4-	16	4-	16	16+	324	144	216
4	14	10	8-	64	6-	36	48+	196	100	140
5	15	10	7-	49	6-	36	42+	225	100	150
6	30	24	8+	64	8+	64	64+	900	576	720
7	17	14	5-	25	2-	4	10+	289	196	238
8	22	18	0	0	2+	4	0	484	324	346
9	16	12	6-	36	4-	16	24+	256	144	192
10	32	26	10+	100	10	100	100	1024	676	832
Σ	220	160		554		296	364	5394	2856	3884

1- حساب الوسط الحسابي للمتغير x:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{220}{10} = 22$$

مؤشر المتغير x:

$$\sigma_x^2 = 55.4 \quad n=10$$

$$\sum x_i = 220 \quad \sum x_i^2 = 5394$$

$$\bar{x} = 22 \quad \sigma_x = 7.44$$

2- حساب الوسط الحسابي للمتغير y:

$$\bar{y} = \frac{\sum y_i}{n} = \frac{160}{10} = 16$$

مؤشرات المتغير y:

$$\sigma_y^2 = 29.599 \approx 29.6 \quad n=10$$

$$\sum y_i = 160 \quad \sum y_i^2 = 2856$$

$$\bar{y} = 16 \quad \sigma_y = 5.44$$

1- حساب معامل الارتباط:

$$R_{yx} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n \cdot \sigma_x \cdot \sigma_y} = \frac{364}{10 \times 7.44 \times 5.44} = \frac{364}{404.86} = 0.898$$

2- حساب معامل الارتباط بالصيغة التالية:

$$R_{yx} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} = \frac{364}{\sqrt{554} \sqrt{296}} = \frac{364}{404.949} = 0.898$$

3- الصيغة التالية:

$$R_{yx} = \frac{\frac{\sum x_i y_i}{n} - \bar{x} \cdot \bar{y}}{\sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} \sqrt{\frac{\sum y_i^2}{n} - \left(\frac{\sum y_i}{n}\right)^2}}$$

$$= \frac{\frac{3884}{10} - 22.16}{\sqrt{\frac{5394}{10} - \left(\frac{220}{10}\right)^2} \sqrt{\frac{2856}{10} - \left(\frac{160}{10}\right)^2}} = \frac{388.4 - 352}{7.44 \times 5.44} = \frac{36.4}{40.4736} = 0.898$$

4- المعادلة التالية:

$$R_{yx} = \frac{n \sum x_i y_i - \sum x_i \cdot \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

$$= \frac{10 \times 3884 - 220 \times 160}{\sqrt{10 \times 5394 - (220)^2} \sqrt{10 \times 2856 - (160)^2}} = \frac{3640}{74.43 \times 54.4}$$

$$= \frac{3640}{4049.4} = 0.898$$

- اختيار دلالة معامل الارتباط:

. الفرضية الابتدائية: $H_0 : R_{yx} = 0$

. الفرضية البديلة: $H_1 : R_{yx} \neq 0$

$$t = \frac{R_{yx}}{\sqrt{\frac{1-R^2}{n-2}}} = \frac{0.898}{\sqrt{\frac{1-0.898^2}{10-2}}} = \frac{0.898}{0.156} = 5.76$$

من جدول توزيع ستودينت نجد أن القيمة الجدولية $t_{0.05} = 2.306$ بما أن قيمة t الفعلية أكبر من الجدولية نرفض الفرضية الابتدائية ونرى بأنه يوجد ارتباط معنوي بين المتغيرين المدروسين.

دراسة الانحدار:

نفترض أن العلاقة بين المتغيرين خطية ونوع خط مستقيم أي معادلة التمثيل:

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

حساب ثوابت معادلة التمثيل انحدار y على x.

- طريقة المعادلات الطبيعية:

$$\sum y_i = b \sum x_i + na \quad (1)$$

$$\sum y_i x_i = b \sum x_i^2 + a \sum x_i \quad (2)$$

وبالتطبيق نجد أن:

$$160 = b \times 220 + 10a$$

$$3884 = 5394b + 220a$$

وبضرب المعادلة الأولى بـ 22 فنحصل على التالي:

$$\begin{aligned} 3520 &= 4840b + 22a \\ 3884 &= 5394b + 220a \\ -364 &= -554 + 0 \end{aligned}$$

ومنه نجد أن فيه a تساوي :

$$b = \frac{-364}{-554} = 0.657$$

وبتعويض a بقيمتها في المعادلة الأولى نجد أ،:

$$\begin{aligned} 160 &= 220 \times 0.657 + 10a \\ 160 &= 144.548 + 10a \\ 15.452 &= 10a \\ a &= \frac{15.452}{10} = 1.5452 \Rightarrow 1.546 \end{aligned}$$

ومنه تصبح المعادلة:

$$\begin{aligned} \hat{y}_i &= bx + a \\ \hat{y}_i &= 0.657x_i + 1.5452 \end{aligned}$$

يمكن حساب قيمة الثابت a : $a = \bar{y} - b\bar{x} = 16 - 0.657 \times 22 = 1.546$

- يمكن حساب قيمة ثوابت المعادلة بطرق عدة منها:

1- الطريقة المختصرة:

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{364}{554} = 0.657$$

أما حساب الثابت b' في حالة حساب معادلة انحدار x على y .

$$b' = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2} = \frac{364}{296} = 1.229$$

2- حساب الثابت b بطريقة كرايمر:

$$\begin{aligned} b &= \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} \\ &= \frac{3884 - 10 \times 22 \times 16}{5394 - 10 \times 22^2} = \frac{364}{554} = 0.657 \end{aligned}$$

أما حساب الثابت b' في معادلة انحدار x على y .

$$\begin{aligned} b' &= \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum y_i^2 - n(\bar{y})^2} \\ &= \frac{364}{2856 - 10 \times (16)^2} \\ &= \frac{364}{296} = 1.229 \end{aligned}$$

$$a' = \bar{x} - b'\bar{y} = 22 - 1.229 \times 16 = 2.336$$

- أما إذا كان معلوماً لديك الانحراف المعياري لكلا المتغيرين ومعامل الارتباط فيمكن حساب قيمة الثابت a وفق التالي:

$$b = R_{yx} \frac{\sigma_y}{\sigma_x} = 0.898 \frac{5.44}{7.44} = 0.657$$

وإذا كان الثابت a معلوماً والانحرافان المعياريان لكلا المتغيرين يمكن حساب معامل الارتباط مباشرة

$$R_{yx} = b \frac{\sigma_x}{\sigma_y} = 0.657 \frac{7.44}{5.44} = 0.898$$

ومهما كانت طريقة حساب الثابت a فإن قيمة الثابت b فيحسب مباشرة من العلاقة التالية: $a = \bar{y} = b\bar{x}$
* تقويم فعالية التمثيل:

i	y _i	$\hat{y}_i$	$(y_i - \hat{y}_i)$	$(y_i - \hat{y}_i)^2$	$(\hat{y}_i - \bar{y})$	$(\hat{y}_i - \bar{y})^2$
1	20	25.198	5.198-	27.02	9.198	84.603
2	14	14.686	0.686-	0.4706	1.314-	1.7266
3	12	13.372	1.372-	1.8823	2.628-	6.9064
4	10	10.744	0.744-	0.554	5.256-	27.626
5	10	11.401	1.401-	1.963	4.599-	21.1508
6	24	21.256	2.744+	7.53	5.256	27.626
7	14	12.715	1.285	1.651	3.285-	10.7912
8	18	16	2	4	0	0
9	12	12.058	0.058-	0.00336	3.942-	15.54
10	26	22.57	3.43	11.7649	6.57	43.165
Σ	160	160		65.83916		239.135

$$\hat{y}_i = 0.657 \times 36 + 1.546$$

- تقويم فعالية التمثيل:

1- حساب التباين غير المفسر:

$$S_{y\hat{Y}_i}^2 = \frac{1}{n} \sum (y_i - \hat{y}_i)^2 = \frac{56.84}{10} = 5.684$$

2- حساب التباين المفسر:

$$S_{\hat{y}_i}^2 = \frac{1}{n} \sum (\hat{y}_i - \bar{y})^2 = \frac{239.135}{10} = 23.9135$$

3- حساب التباين الكلي:

$$\sigma^2 = S_{y\hat{y}_i}^2 + S_{\hat{Y}_i}^2 = 5.684 + 23.99135 = 29.598$$

- حساب الخطأ المعياري للتقدير:

$$\sigma_{\hat{y}} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}} = \sqrt{5.684} = 2.384$$

- الطريقة المختصرة لحساب الخطأ المعياري للتقدير:

$$\begin{aligned} \sigma_{\hat{y}}^2 &= \frac{\sum y_i^2 - a \sum y_i - b \sum x_i y_i}{n} \\ &= \frac{2856 - 1.546 \times 160 - 0.657 \times 3884}{10} = \frac{56.852}{10} \\ &= 5.685 \Rightarrow \sigma_{\hat{y}} = \sqrt{5.685} = 2.384 \end{aligned}$$

استخدام معادلة التمثيل بالتنبؤ:

ما هو مقدار الانفاق المتوقع إذا كان الدخل يساوي 40/.

$$\hat{y}_{40} = 0.657 \times 40 + 1.546 = 27.826$$

- تحديد مجال الثقة للانفاق المتوقع:

$$P \left(\hat{y}_i - z_{1-\frac{\alpha}{2}} \sigma_{\hat{y}_i} \right) = 0.95$$

$$P (27.826 - 1.96 \times 2.384 \leq \hat{y}_i \leq 27.826 + 1.96 \times 2.384)$$

$$P (27.826 \text{ m } 4.673) = 0.95$$

$$P (23.153 , 32.498) = 0.95$$

- حساب معامل التحديد:

$$D^2 = \frac{\text{الجزء الذي يفسر}}{\text{الجزء الذي يفسر + الجزء غير المفسر}} = \frac{23.9135}{29.598} = 0.808$$

$$D^2 = R_{yx}^2 = 0.898^2 = 0.808$$

الجزء غير المفسر / البواقي / $R_s = 1 - d^2 = 1 - 0.808 = 0.193$

أي أن هناك % 19.3 من تغيرات المتحول y لم يفسر بدلالة تغيرات المتحول x .

* تقدير تباين الانحدار ANOVA:

1- الطريقة المختصرة لحساب التباينات:

$$SST = \sum y^2 - n \bar{y}^2 = 2856 - 10 \times 16^2 = 296$$

2- الانحرافات بسبب الانحدار:

$$SSR = b^2 \left[\sum x^2 - n \bar{x}^2 \right]$$

$$= 0.657^2 \left[5394 - 10 \times 22^2 \right] = 239.14$$

$$SSE = SST - SSR = 296 - 239.14 = 56.86$$

S.O.V	df	SS	MS	F
بسبب الانحدار SSR	a-1 2-1=1	239.14	239.14	$\hat{F} = \frac{M_{SE}}{M_{SE}}$ $= \frac{239.14}{7.1075} = 33.65$
الانحرافات عن الانحدار SSE الخطأ	n-2 10-2=8	56.86	$S_{yx} = \frac{56.86}{8} = 7.1075$	
SST التباين الكلي	n-1 10-1=9	296		

$$F_{\alpha 0.05}(1.9) = 5.12$$

$$\hat{F} > F_{\alpha}$$

نرفض الفرضية الابتدائية H_1 وبالتالي فإن الانحدار والارتباط بين المتغيرين دال إحصائياً.

