مسألة: إذا توافرت لديك المعلومات الآتية من إحدى المزارع في منطقتك والتي تبيّن متوسط إنتاج الحيوان السنوي من الحليب في المزرعة مقدراً بالكيلوغرام ومتوسط كمية الأعلاف المقدمة له بالكيلوغرام.

المطلوب:

أ. حساب كل من معامل الارتباط بيرسون، ومعامل التحديد لمتوسط كميات الإنتاج والأعلاف، وتفسير النتائج.

ب. تشكيل معادلة الانحدار التي تعيّن العلاقة بين كمية إنتاج الحيوان وكمية الأعلاف المقدمة له.

ج. حساب إنتاج الحليب المتوقع عندما تكون كمية الأعلاف المقدمة: 300، 1200، 1500، 2000 كغ.

د. حساب كمية الأعلاف اللازمة للحصول على 650، 800، 1100، 1250 كغ من الحليب.

هـ. اختبار معنوية معاملات الارتباط والانحدار باستخدام الاختبارات الآتية:

- ✓ اختبار معنوية المعامل r باستخدام اختبار T عند مستويي المعنوية 5% و 1%.
- ✓ اختبار معنوية المعاملات a و b باستخدام اختبار الخطأ المعياري أو القياسي.
- ✓ اختبار معنوية المعاملات a و b باستخدام اختبار T عند مستويي المعنوية 5% و 1%.

xi	yi	$X_i-\bar{X}$ x_i	$Y_i-\bar{Y}$ y_i	$(X_i-\bar{X})*(Y_i-\bar{Y})$ x_i*y_i	$(X_i-\bar{X})^2$ ${x_i}^2$	$(Y_i-\bar{Y})^2$ ${y_i}^2$
400	100	-400	-250	100000	160000	62500
650	250	-150	-100	15000	22500	10000
750	350	-50	0	0	2500	0
800	400	0	50	0	0	2500
900	500	100	150	15000	10000	22500
1000	450	200	100	20000	40000	10000
1100	400	300	50	15000	90000	2500
5600	2450	0	0	165000	325000	110000
800	350					

- **حساب معامل الارتباط:**

$$r=\frac{\sum_{i=1}^{n}(X_i-\bar{X})*(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2*\sum_{i=1}^{n}(Y_i-\bar{Y})^2}}$$

$$r = \frac{165000}{\sqrt{325000 * 110000}} = 0.8726617108 \approx 0.87$$

نستنتج أنّه لدينا علاقة طردية بين إنتاج الحيوان وكمية الأعلاف المقدمة له ولدينا ارتباط موجب قوي جداً، حيث أنّ زيادة كمية الأعلاف المقدمة للحيوان تؤثّر بقوة على إنتاجه.

- **حساب معامل التحديد:**

$$\mathrm{B} = r^2 = 0.8726617108^2 = 0.7615384615 \approx 0.76$$

نستنتج من ذلك أنّ:

76% من أسباب زيادة إنتاج الحيوان تعود للزيادة في كمية الأعلاف.

أمّا 24% من أسباب زيادة إنتاج الحيوان تعود إلى أسباب أخرى غير مدروسة.

- **تشكيل معادلة الانحدار:**

نقوم بتشكيل معادلة الانحدار التي تعيّن العلاقة بين كمية إنتاج الحيوان من الحليب وكمية الأعلاف، كما يلي:

$$b = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) * (Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$b = \frac{165000}{325000} = 0.5076923077 \approx 0.51$$

بعد أن قمنا بحساب قيمة b نقوم بحساب قيمة a ، كما يلي:

$$a = \bar{Y} - b\bar{X}$$

$$a = 350 - 800 * 0.5076923077 = -56.15384615 \approx -56$$

بعد استنتاج (a و b) نشكّل معادلة الانحدار:

$$Y = -56 + 0.51X$$

b: موجب، وهذا يعني أنّ العامل المستقل X يؤثّر إيجاباً على العامل Y أي أنّ زيادة كمية الأعلاف تؤدي إلى زيادة إنتاج الحيوان.

كلما زادت كمية الأعلاف بمقدار وحدة كاملة يزداد إنتاج الحيوان بمقدار وحدة كاملة والتي هي b=0.51.

- **حساب إنتاج الحليب المتوقع:**

$$y = -56 + 0.51x$$

عندما تكون كمية الأعلاف 300 kg:

$$y = -56 + 0.51 * 300 = 97\ kg$$

عندما تكون كمية الأعلاف 1200 kg:

$$y = -56 + 0.51 * 1200 = 556\ kg$$

عندما تكون كمية الأعلاف 1500 kg:

$$y = -56 + 0.51 * 1500 = 709\ kg$$

عندما تكون كمية الأعلاف 2000 kg:

$$y = -56 + 0.51 * 2000 = 964\ kg$$

- **حساب كمية الأعلاف اللازمة:**

$$y = -56 + 0.51x$$

$$x = \frac{y + 56}{0.51}$$

للحصول على 650 كغ من الحليب.

$$x = \frac{650 + 56}{0.51} = 1384.313725 \approx 1384\ kg$$

للحصول على 800 كغ من الحليب.

$$x = \frac{800 + 56}{0.51} = 1678.431373 \approx 1678\ kg$$

للحصول على 1100 كغ من الحليب.

$$x = \frac{1100 + 56}{0.51} = 2266.666667 \approx 2267\ kg$$

للحصول على 1250 كغ من الحليب.

$$x = \frac{1250 + 56}{0.51} = 2560.784314 \approx 2561\ kg$$

- **اختبار معنوية معاملات الارتباط والانحدار:**

<u>اختبار معنوية المعامل r باستخدام اختبار T عند مستويي المعنوية 5% و 1%:</u>

للقيام بهذا الاختبار نقارن T المحسوبة مع T الجدولية عند درجات حرية $n-k$ ودرجة الخطأ أو مستويي المعنوية 1% و 5%.

فإذا كانت T المحسوبة > T الجدولية يعني هذا أن الارتباط بين X و Y معنوي ومؤكّد إحصائياً أي غير ناتج عن الصدفة، ولا يمكن إهماله.

أمّا إذا كانت T المحسوبة < T الجدولية يعني أن الارتباط بين X و Y غير معنوي وغير مؤكّد إحصائياً أي ناتج عن الصدفة، ويمكن إهماله.

$$\mathrm{T} = r\sqrt{\frac{n-k}{1-r^2}}$$

T: T المحسوبة.

r: معامل الارتباط. r^2: معامل التحديد.

$n-k$: درجات الحرية حيث:

n: تساوي عدد قيم x أو y.

k: تمثل عدد المعلمات المقدرة في معادلة الانحدار (a, b).

إذا أردنا اختبار معنوية معامل الارتباط في مثالنا نحسب قيمة T الجدولية عند درجة الخطأ أو مستوى المعنوية 5% والتي تساوي 2.571 وعند 1% والتي تساوي 4.032 ثم نحسب قيمة T المحسوبة كما يلي:

$$\mathrm{T} = r\sqrt{\frac{n-k}{1-r^2}}$$

$$T = 0.8726617108\sqrt{\frac{7-2}{1-0.7615384615}} = 3.995965707 \approx 3.996$$

بعد مقارنة قيمة T المحسوبة مع T الجدولية نلاحظ أن T المحسوبة > T الجدولية عند درجة الخطأ أو مستوى المعنوية 5% ممَّا يعني أنّ الارتباط بين X و Y معنوي ومؤكّد إحصائياً أي غير ناتج عن الصدفة، وذلك عند مستوى المعنوية 5%. وأن T المحسوبة < T الجدولية عند درجة الخطأ أو مستوى المعنوية 1% ممَّا يعني أنّ الارتباط بين X و Y غير معنوي وغير مؤكّد إحصائياً أي ناتج عن الصدفة، وذلك عند مستوى المعنوية 1%.

اختبار معنوية المعاملات a و b باستخدام اختبار الخطأ المعياري أو القياسي:

X_i	Y_i	x_i	y_i	$x_i * y_i$	${x_i}^2$	${y_i}^2$	$\hat{Y}$	e_i	${e_i}^2$	${X_i}^2$
400	100	-400	-250	100000	160000	62500	148	-48	2304	160000
650	250	-150	-100	15000	22500	10000	275.5	-25.5	650.25	422500
750	350	-50	0	0	2500	0	326.5	23.5	552.25	562500
800	400	0	50	0	0	2500	352	48	2304	640000
900	500	100	150	15000	10000	22500	403	97	9409	810000
1000	450	200	100	20000	40000	10000	454	-4	16	1000000
1100	400	300	50	15000	90000	2500	505	-105	11025	1210000
5600	2450	0	0	165000	325000	110000				4805000

لإجراء هذا الاختبار نقوم أولاً بحساب تباين الحد العشوائي وذلك باستخدام القانون:

$$S_e{}^2 = \frac{\sum {e_i}^2}{n-k}$$

e_i: الخطأ ويساوي الفرق بين Y_i و $\hat{Y}$ أي أن $e_i = Y_i - \hat{Y}$.

k: تمثل عدد المعلمات المقدرة في معادلة الانحدار (a, b).

n: حجم العينة أي عدد قيم X أو Y.

نقوم بحساب قيم $\hat{Y}$ من خلال تعويض قيم X بمعادلة الانحدار التي قمنا سابقاً بتشكيلها والتي تعيّن العلاقة بين كمية إنتاج الحيوان وكمية الأعلاف المقدمة له.

$$Y = -56 + 0.51X$$

ثم نحسب الخطأ e_i من خلال طرح قيم $\hat{Y}$ المقدرة من قيم Y الحقيقية المقابلة لها باستخدام المعادلة:

$$e_i = Y_i - \hat{Y}$$

ثم نقوم بتربيع قيم الخطأ $\boldsymbol{e_i}$ للحصول على قيم $\boldsymbol{e_i}^{\boldsymbol{2}}$.

$$S_e{}^2 = \frac{\sum e_i{}^2}{n - k}$$

$$S_e{}^2 = \frac{26260.5}{7 - 2} = \frac{26260.5}{5} = 5252.1$$

ثم نقوم بحساب مربعات قيم X_i التي سنستخدمها في الخطوة الآتية.

في الخطوة الثانية تكون بحساب قيم الانحراف المعياري لكل من قيم a و b المقدرة من خلال الصيغ:

$$S_a = \sqrt{\frac{S_e{}^2 \sum X_i{}^2}{n \sum x_i{}^2}}$$

$$S_a = \sqrt{\frac{5252.1 * 4805000}{7 * 325000}} \approx 105$$

$$S_b = \sqrt{\frac{S_e{}^2}{\sum x_i{}^2}}$$

$$S_b = \sqrt{\frac{105.3228224}{325000}} = 0.127$$

توضيح: بعد حساب الانحراف المعياري للمعلمتين نجد أمامنا حالتين:

إمّا أن تكون قيم الانحراف المعياري مساوية للصفر وبالتالي فإن القيمة المقدرة للمعلمة من خلال العينة لا تنحرف عن وسطها الحسابي وتمثّل المجتمع تمثيلاً تاماً، ولا يوجد أي نوع من الخطأ. وهي حالة نظرية نادرة الحدوث.

أو تكون قيم الانحرافات المعيارية للمعلمتين أكبر من الصفر (كما في هذه المسألة). وهنا كلما ازدادت قيم هذه الانحرافات كلما أصبحت المعلمات المقدرة أقل تمثيلاً للمجتمع.

ولنختبر معنوية المعلمات المقدرة من خلال اختبار الخطأ العشوائي نستخدم ما يسمى بفرض العدم والفرض البديل:

Null Hypothesis H_0: a = 0 , b = 0

Alternative Hypothesis H_1: a ≠0 , b ≠0

هنا لتحديد أي الفرضيتين سنقبل، نقوم بمقارنة قيمة الانحراف المعياري لكل معلمة بنصف القيمة المقدرة للمعلمة نفسها. حيث نواجه حالتين:

بالنسبة للمعلمة b: إذا كان $S_b < \frac{b}{2}$ هنا نرفض فرض العدم ونقبل الفرض البديل القائل بأن قيمة المعلمة تختلف معنوياً عن الصفر. ونفس الكلام ينطبق على المعلمة a.

بالنسبة للمعلمة b: إذا كان $S_b > \frac{b}{2}$ أي أن الخطأ كبير نسبياً وهنا نقبل فرض العدم القائل بأن المعلمة المقدرة لا تختلف معنوياً عن الصفر. ونفس الكلام ينطبق على المعلمة a.

نقوم بحساب نصف قيمة b المقدرة، ثم نقارنها مع $S_b = 0.127$ كالآتي:

$$\frac{b}{2} = \frac{0.51}{2} = 0.255$$

نلاحظ أن $S_b < \frac{b}{2}$ وبالتالي نرفض فرض العدم ونقبل الفرض البديل القائل بأن قيمة المعلمة المقدرة b تختلف معنوياً عن الصفر.

نقوم بحساب نصف قيمة a المقدرة، ثم نقارنها مع $S_a = 105$ كالآتي:

$$\frac{a}{2} = \frac{-56}{2} = -28$$

نلاحظ أن $S_a > \frac{a}{2}$ وبالتالي فإن الخطأ كبير نسبياً وهنا نقبل فرض العدم القائل بأن المعلمة المقدرة a لا تختلف معنوياً عن الصفر.

<u>اختبار معنوية المعاملات a و b باستخدام اختبار T عند مستويي المعنوية 5% و 1%:</u>

يستخدم اختبار T عندما: يكون تباين المجتمع مجهولاً وحجم العينة صغير n<30 كما في مثالنا.

عند إجراء الاختبار يجب معرفة أن المتوسط الحسابي للتوزيع الطبيعي 0 والانحراف المعياري له 1.

لإجراء الاختبار نتّبع الخطوات التالية:

أولاً نقوم تحديد T المحسوبة كالآتي:

$$\boldsymbol{T_a} = \frac{\boldsymbol{\hat{a} - a}}{S_a}$$

$$T_b = \frac{\hat{b} - b}{S_b}$$

تعطى قيم S_a و S_b من خلال العلاقتين السابقتين في اختبار الخطأ المعياري. وتقدر القيمتان $\hat{a}$ و $\hat{b}$ من خلال نموذج الانحدار. وتبقى قيم a و b مجهولة بالنسبة لنا وهي تعبر عن قيم المعلمات الحقيقية للمجتمع ككل. لذا يجب أن نفترض قيم خاصة بمعلمات المجتمع. وبالطبع هذه القيم تحتمل الصواب والخطأ لذلك يتعين اختبارها. ولقد جرى العرف على استخدام القيمة المفترضة صفر لمعلمات المجتمع، وهي تعني في المعلمة الانحدارية أنه لا توجد علاقة بين المتغيرين X و Y وذلك لاختبار فرض العدم والفرض البديل فتصبح العلاقتان السابقتان:

$$\boldsymbol{T_a = \frac{\hat{a}}{S_a}}$$

$$\boldsymbol{T_b = \frac{\hat{b}}{S_b}}$$

ثمّ نقوم بتحديد T الجدولية. تحدد من جدول توزيع T عند درجات حرية n-k ومستوى معنوية محدد (0.05 أو 0.025 أو 0.01).

إن توزيع T توزيع طبيعي متماثل وسطه الحسابي T=0 وتباينه أو انحرافه المعياري $\frac{n-1}{n-3}$ وهو يقترب من الواحد كلما كبر حجم العينة أي أنه يقترب من توزيع Z ويخلف توزيع T عن توزيع Z في أن الأول مصمم على أساس درجات حرية، بعكس الثاني.

بما أننا ندرس مدى تأثير كمية الأعلاف على زيادة إنتاج الحليب (وليس على النقصان) لذلك نستخدم اختبار الطرف الواحد حيث أن b حتماً موجبة.

لاختبار المعنوية للمعلمات المقدرة نستخدم <u>اختبار الطرف الواحد:</u>

والذي يتحدد بفرض العدم:

Null Hypothesis $H_0: b_i = b_0$

والفرض البديل:

Alternative Hypothesis $H_1: b_i < b_0$ أو $b_i > b_0$

نقوم بحساب قيمة T المحسوبة كما في الصيغة الآتية وبهذه الحالة تكون $b_0 = 0$:

$$\boldsymbol{T_b = \frac{\hat{b}}{S_b}}$$

بعد ذلك نبحث عن قيمة T الجدولية في الجداول الإحصائية عند مستوى معنوية 5% و 1% ودرجات حرية n-k فإذا كانت:

|T المحسوبة | < T الجدولية: نقبل فرض العدم ونرفض الفرض البديل وتكون المعلمة المقدرة غير معنوية إحصائياً.

|T المحسوبة | > T الجدولية: نقبل الفرض البديل وتكون المعلمة المقدرة معنوية إحصائياً.

ونفس الكلام ينطبق على المعلمة a.

في الجداول الإحصائية عند مستوى معنوية 5% ودرجات حرية n-k=7-2=5 تكون قيمة T الجدولية = 2.571

أمّا عند مستوى معنوية 1% ودرجات حرية n-k=7-2=5 تكون قيمة T الجدولية = 4.032

$$T_a = \frac{\hat{a}}{S_a} = \frac{-56}{105} \approx -0.53$$

|T المحسوبة | < T الجدولية: نقبل فرض العدم ونرفض الفرض البديل وتكون المعلمة المقدرة a غير معنوية إحصائياً عند مستويي المعنوية 1% و 5%.

$$T_b = \frac{\hat{b}}{S_b} = \frac{0.51}{0.127} \approx 4$$

|T المحسوبة | > T الجدولية: نقبل الفرض البديل وتكون المعلمة المقدرة b معنوية إحصائياً عند مستوى المعنوية 5%.

|T المحسوبة | < T الجدولية: نقبل فرض العدم ونرفض الفرض البديل وتكون المعلمة المقدرة b غير معنوية إحصائياً عند مستوى المعنوية 1%.

- عند تطبيق **اختبار التباين ANOVA** نجد أن:

	df	SS	MS	F
Regression	1	83769.231	83769.231	15.968*
Residual	5	26230.769	5246.154	
Total	6	110000		

بعد مقارنة قيمة F المحسوبة مع قيمة F الجدولية عند مستوى المعنوية 5% (F=9.5) ومستوى المعنوية 1% (F=16.5) نستنتج أن معادلة الانحدار معنوية.