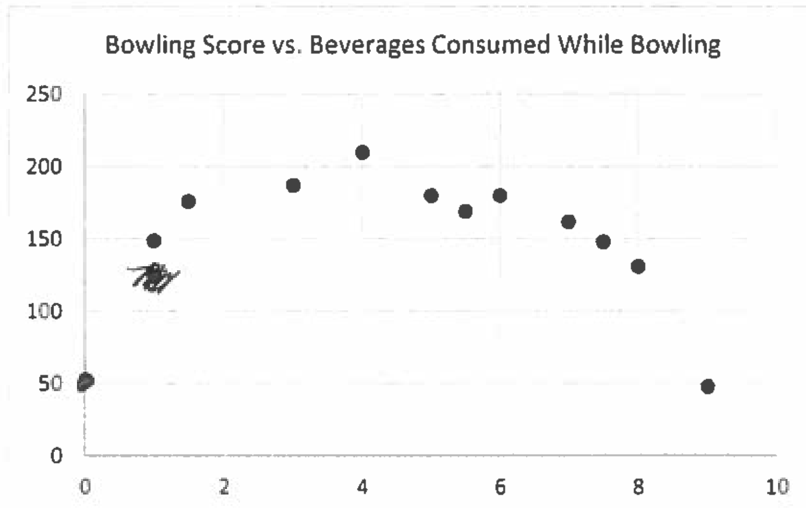
ISE 2211 • Exam VI • Tax Day 2019!   NAME SOLUTION

There appears to be a relationship between bowling score (y) vs. the number of tasty beverages consumed while bowling (x). The following data was collected over the course of twelve games.

	Beverages (x)	Bowling Score (y)
1	7.5	148
2	1	149
3	6	180
4	8	131
5	4	210
6	7	162
7	0	52
8	9	48
9	1.5	176
10	5.5	169
11	3	187
12	5	180



Determine least-squares estimates for slope (β_1) and intercept (β_0) of the simple linear regression model for bowling score.

Formulae:

$$\hat{\beta}_1 = \frac{\sum y_i x_i - \frac{(\sum y_i)(\sum x_i)}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} = \frac{S_{XY}}{S_{XX}} \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$\sum x_i = 57.5$, $\sum y_i = 1792$ (+1)

$\sum y_i^2 = 295924$, $\sum x_i^2 = 369.75$ (+1)

$\bar{x} = 4.792$, $\bar{y} = 149.3$ (+1)

$\sum y_i x_i = 8447.5$ (+1) (+1)

$S_{xy} = 8447.5 - \frac{1792 \cdot 57.5}{12}$

$S_{xy} = -139.2$ (+1)

$S_{xx} = 369.75 - \frac{57.5^2}{12}$

$S_{xx} = 94.23$ (+1)

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-139.2}{94.23} = -1.477$

$\hat{\beta}_0 = 149.3 - (-1.477)4.792 = 156.4$

Write an equation for the estimated regression line ($\hat{y}$) with your actual numbers for $\hat{\beta}_0$ and $\hat{\beta}_1$.

$$\hat{y} = 156.4 - 1.477x$$

Write a 95% confidence interval on the mean bowling score at x = 4 beverages.

$$t_{\alpha/2,\, n-2} = t_{.025,10} = 2.228 \quad +1$$

$$SS_T = \sum y_i^2 - n\bar{y}^2$$

$$= 295924 - 12 \cdot 149.3^2 = 28438 \quad +1$$

$$SS_E = SS_T - \hat{\beta}_1 S_{xy}$$

$$= 28438 - (-1.477)(-139.2)$$

$$= 28233 \quad +0$$

$$\hat{\sigma}^2 = \frac{SS_E}{n-2} = \frac{28233}{10} = 2823 \quad +1$$

$$\hat{\mu}_{Y|4} = 156.4 - 1.477 \cdot 4 = 150.5 \quad +1$$

$$\mu_{Y|4} \qquad 150.5 \pm 2.228\sqrt{2823\left[\frac{1}{12} + \frac{(4-4.792)^2}{94.23}\right]}$$

$$115 < \mu_{Y|4} < 186 \quad +1$$

Write a 95% prediction interval on the 13th bowling score at $x = 4$ beverages.

$$\hat{y}_0 = \mu_{Y|4} = 150.5 \quad +1$$

$$Y_0 : \quad \hat{y}_0 \pm t_{\alpha/2, n-2} \sqrt{\hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}$$

$$150.5 \pm 2.228 \sqrt{2823 \left[1 + \frac{1}{12} + \frac{(4 - 4.792)^2}{94.23}\right]}$$

$$26.9 < Y_0 < 274 \quad +1$$

Write a 95% confidence interval on the value of intercept.

$$\beta_0 : \quad \hat{\beta}_0 \pm t_{\alpha/2, n-2} \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]} \quad +1$$

$$156.4 \pm 2.228 \sqrt{2823 \left[\frac{1}{12} + \frac{4.792^2}{94.23}\right]}$$

$$88.7 < \beta_0 < 224 \quad +1$$

Write a 95% confidence interval on the value of slope and use it to test the following hypotheses that the slope is zero. If you fail to reject the null hypothesis, what does it clearly mean from the scatter plot of the data?

H_0: $\hat{\beta}_1 = 0$

H_1: $\hat{\beta}_1 \neq 0$

$$\beta_1 : \quad \hat{\beta}_1 \pm t_{\alpha/2, n-2} \sqrt{\hat{\sigma}^2 / S_{xx}}$$

$$-1.477 \pm 2.228 \sqrt{2823/14.23}$$

$$-13.67 < \beta_1 < 10.72 \quad +1$$

contains zero; fail to reject Ho +1

not linear! +1

Test the following hypotheses ~~Write a 95% confidence interval~~ on the correlation coefficient ρ, if y and x may both be considered random variables. (Ignore the fact that $n \ngtr 30$.) @ $\alpha = 0.05$

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

$$R^2 = 1 - \frac{SS_E}{SS_T}$$

$$= 1 - \frac{28233}{28438}$$

$$R^2 = 0.007209 \quad +1$$

$$t_0 = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} = \frac{\sqrt{.007209 \cdot 10}}{\sqrt{1 - 0.007209}}$$

$$t_0 = 0.2695 \quad +1$$

$$t_{critical} = t_{\alpha/2, n-2} = 2.228$$

$$t_0 \ngtr t_{\alpha/2, n-2}$$ fail to reject Ho +1