

Name:

STAT 516 – Fall 2026 – Version B
Exam 1

1. A simple linear regression was fit (based on $n = 8$ observations) in order to predict sugar content (Y, in grams) of fruit based on the number of days (X) after the fruit had been picked. Look at the partial R output given below:

```
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  12.0419     2.5527   4.717  0.00327 **
days        -0.8216     0.5105  -1.609  0.15867
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.803 on 6 degrees of freedom
Multiple R-squared:  0.3015,    Adjusted R-squared:  0.1851
F-statistic:  2.59 on 1 and 6 DF,  p-value: 0.1587
```

(a) State the estimated regression equation, using the correct notation.

(b) Use the estimated regression equation to predict the sugar content of a fruit that was picked 5 days ago.

$$\hat{Y} = 12.0419 - 0.8216(5) = 7.934 \text{ grams}$$

(c) Find a 90% prediction interval for the sugar content of a fruit that was picked 5 days ago.

Note that $\bar{X} = 4.25$ for this data set. [Hint: Recall that the standard error of $\hat{\beta}_1$ is $\sqrt{\frac{MSE}{S_{xx}}}$, and that what R calls “Residual standard error” is merely $\sqrt{MSE}$.]

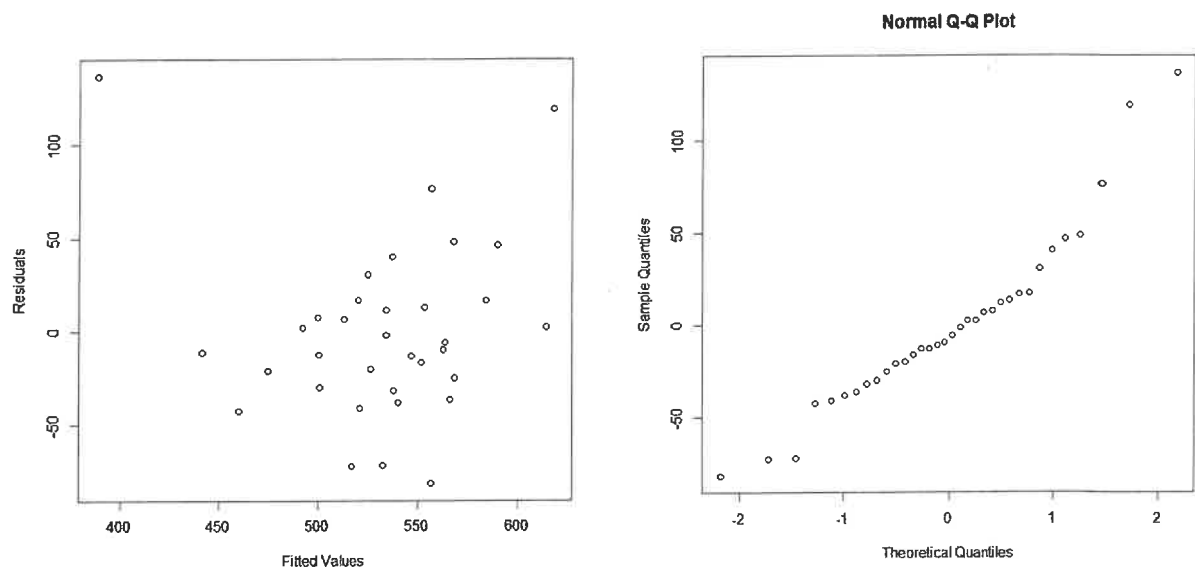
(See Version A for MSE, S_{xx} , and t -value)

90% PI:

$$7.934 \pm 1.943 \sqrt{14.4628 \left[1 + \frac{1}{8} + \frac{(5 - 4.25)^2}{55.496} \right]}$$

$$7.934 \pm 1.943 (4.05182)$$

$$(0.061, 15.807)$$



3. Choose the best answer: If for two variables X and Y , the correlation coefficient $r > 0$, then as X increases, Y tends to:

- (A) decrease (B) remain constant (C) increase

4. For a data set of pro basketball players, a regression was performed with the dependent variable being annual salary (Y) in millions of dollars and the independent variable being points per game (X). A 90% prediction interval for the Y -value for a new player with $X = 7.3$ was found to be (2.18, 5.56).

(a) What is the best interpretation of this prediction interval?

- (A) With probability 0.90, the salary for a new player who scores 7.3 points per game is between 2.18 and 5.56 million dollars.
 (B) With 90% confidence, the change in expected salary for a one-point increase in scoring is between 2.18 and 5.56 million dollars.
 (C) With 90% confidence, the mean salary for players who score 7.3 points per game is between 2.18 and 5.56 million dollars.
 (D) About 90% of the pro basketball players in the sample have a salary of between 2.18 and 5.56 million dollars.

(b) If you had to predict *with a single number* the salary of a player who scores 7.3 points per game based on this regression analysis, what would the answer be?

- (A) \$ 2.18 million (B) \$5.56 million (C) \$7.3 million (D) \$3.87 million

(c) In this data set, what is a possible value for the 90% confidence interval for the mean salary of all players scoring 7.3 points per game? (Choose one)

- (A) (1.37, 6.37) (B) (3.41, 4.33) (C) (2.18, 5.56) (D) (3.91, 4.83)