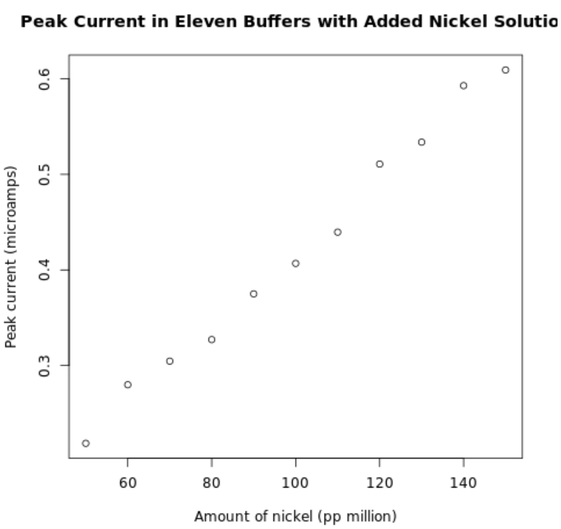
Assignment Assignment-10 due 12/08/2022 at 11:59pm PST

Problem 1. **(2 points)**

An engineer measures the peak current (in microamps) when a solution containing an amount of nickel (in parts per 10^6) is added to a buffer. The experiment was repeated for eleven different values of nickel solutions. A scatterplot showing the data is given below:



Suppose a regression line was added to the plot above. If an additional measurement had been taken with the nickel of 72 parts per 10^6 and a peak current of 0.25 microamps, adding this observation would:

- A. decrease the intercept, decrease the slope.
- B. increase the intercept, decrease the slope.
- C. decrease the intercept, increase the slope.
- D. increase the intercept, increase the slope.
- E. not affect the regression line.

Solution:

Answer(s) submitted:

- C

(correct)

Correct Answers:

- C

Problem 2. **(4 points)**

For each problem, select the best response.

(a) A study found a correlation of r = -0.61 between the gender of a worker and his or her income. You may correctly conclude

- A. this is incorrect because r makes no sense here.
- B. women earn more than men on average.
- C. women earn less than men on average.
- D. an arithmetic mistake was made. Correlation must be positive.
- E. None of the above.

(b) For a biology project, you measure the weight in grams and the tail length in millimeters of a group of mice. The correlation is r = 0.8. If you had measured tail length in centimeters instead of millimeters, what would be the correlation? (There are 10 millimeters in a centimeter.)

- A. (0.8)(10) = 8
- B. 0.8
- C. 0.8/10 = 0.08
- D. None of the above.

Answer(s) submitted:

- A
- B

(correct)

Correct Answers:

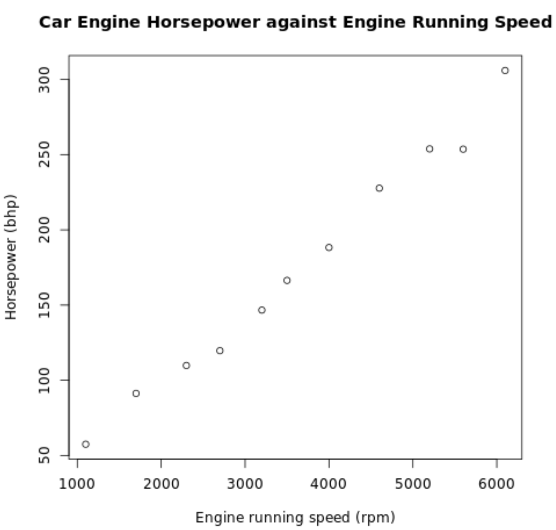
- A
- B

Problem 3. **(14 points)**

The horsepower (Y, in bhp) of a motor car engine was measured at a chosen set of values of running speed (X, in rpm). The data are given below (the first row is the running speed in rpm and the second row is the horsepower in bhp):

	1100	1700	2300	2700	3200	3500	4000	4600	5200	5600	6100
p)	57.51	91.29	109.84	119.69	146.65	166.4	188.26	227.67	253.8	253.61	305.89

The mean and sum of squares of the rpm are 3636.3636 rpm and 171940000.0000 rpm ^{2} respectively; the mean of the horsepower values is 174.6009 bhp and the sum of the products of the two variables is 8252156.0000 rpm bhp. A scatterplot displaying the data is shown below:



Please provide answers to the following to 3 decimals places where appropriate:

Part a)

Compute the regression line for these data, and provide your estimates of the slope and intercept parameters. **To calculate the slope and intercept please keep intermediate results to 6 decimal places (when you calculate the intercept do not round the value of the slope).**

Slope: ___

Intercept: ___

Note: For the parts below, **use the slope and intercept values in Part a, corrected to 3 decimal places** to calculate answers.

Part b)

Based on the regression model, what level of horsepower would you expect the engine to produce if running at 2400 rpm?

Answer: ___

Part c)

Assuming the model you have fitted, if increase the running speed by 100 rpm, what would you expect the change in horsepower to be?

Answer: ___

Part d)

The standard error of the estimate of the slope coefficient was found to be 0.001769. Provide a 95% confidence interval for the true underlying slope.

Confidence interval: (___, ___)

Part e)

Without extending beyond the existing range of speed values or changing the number of observations, we would expect that increasing the variance of the rpm speeds at which the horsepower levels were found would make the confidence interval in (d)

- A. either wider or narrower depending on the values chosen.
- B. unchanged.
- C. narrower.
- D. wider.

Part f)

If testing the null hypothesis that horsepower does not depend linearly on rpm, what would be your test statistic? (For this part, you are to calculate the test statistic by hand using appropriate values from the answers you provided in part (a) accurate to 3 decimal places, and values given to you in part (d).)

Answer: ___

Part g)

Assuming the test is at the 1% significance level, what would you conclude from the above hypothesis test?

- A. Since the observed test statistic does not fall in either the upper or lower 1/2 percentiles of the t distribution with 9 degrees of freedom, we cannot reject the null hypothesis that the horsepower does not depend linearly on rpm.
- B. Since the observed test statistic falls in either the upper or lower 1/2 percentiles of the t distribution with 9 degrees of freedom, we can reject the null hypothesis that the horsepower does not depend linearly on rpm.
- C. Since the observed test statistic does not fall in either the upper or lower 1 percentiles of the t distribution with 9 degrees of freedom, we cannot reject the null hypothesis that the horsepower does not depend linearly on rpm.
- D. Since the observed test statistic falls in either the upper or lower 1/2 percentiles of the t distribution with 9 degrees of freedom, we cannot reject the null hypothesis that the horsepower does not depend linearly on rpm.
- E. Since the observed test statistic does not fall in either the upper or lower 1/2 percentiles of the t distribution with 9 degrees of freedom, we can reject the null hypothesis that the horsepower does not depend linearly on rpm.

Solution:

Part a)
The slope term is
$\hat{\beta}_1 = \frac{n\sum_i x_i y_i - \sum_i x_i \sum_i y_i}{n\sum_i x_i^2 - (\sum_i x_i)^2} = 0.048$
and the intercept is
$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 0.492.$

Part b)
The fitted value at rpm 2400 is given by $\hat{\beta}_0 + \hat{\beta}_1(2400) = 115.692.$

Part c)
The predicted increase will be $100\hat{\beta}_1 = 4.800.$

Part d)
We will use $SE(\hat{\beta}_1)$, the estimated standard deviation of $\hat{\beta}_1$, given here as 0.001769. The relevent percentile on the t distribution is $t_{11-2}(0.975) = 2.262$ and the confidence interval is therefore
$\hat{\beta}_1 \pm (2.262)(0.001769) = (0.044, 0.052).$

Part e) The answer is "narrower" since the variance of $\hat{\beta}_1$ is inversely proportional to the variance of the rpm values chosen.

Part f) The test statistic is
$\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = 27.134.$

Part g) If the test statistic falls above $t_{11-2}(0.995) = 3.250$ (or, much less likely, below $t_{11-2}(0.005) = -3.250$), the answer is
"Since the observed test statistic falls in either the upper or lower 1/2 percentiles of the t distribution with 9 degrees of freedom, we can reject the null hypothesis that the horsepower does not depend linearly on rpm;"
otherwise, far less likely, the response is
"Since the observed test statistic does not fall in either the upper or lower 1/2 percentiles of the t distribution with 9 degrees of freedom, we cannot reject the null hypothesis that the horsepower does not depend linearly on rpm."

Answer(s) submitted:

- `0.04780`
- `0.492357`
- `115.212357`
- `4.78`
- `0.044`
- `0.052`
- `C`
- `27.134`
- `B`

(correct)
Correct Answers:

- `0.048`
- `0.492`
- `115.692`
- `4.8`
- `0.044`
- `0.052`
- `C`
- `27.134`
- `B`