Questions & Answers

I. Exercise 11.135: Simple Linear Regression.

Read Exercise 11.135 on page 608 in the text. This is a simple linear regression problem. The explanatory (independent) variable X is the Contraceptive Prevalence (% of married women who use contraceptives). The response (dependent) variable Y is the Rate of Fertility (average number of children per woman). There are sample data from 27 developing countries. Do the data provide sufficient evidence at the 5% significance level to conclude that the regression equation using Contraceptive Prevalence as the explanatory variable is useful for predicting the Fertility Rate?

The Least Squares (LS) regression equation (use three decimals for the y-int and slope) is

________________________________________________

The estimated standard error of the regression model (rounded to 1 decimal), that describes the average prediction error in absolute value while predicting the Fertility Rate, is given by \( s = \) _______ children per woman.

The coefficient of determination is \( R^2 = \) _______ meaning that _______ of the total variance of the Fertility Rate is explained by ___________________________ using a linear equation.

The hypotheses (in symbols and words) for testing the utility of the regression line are:

Ho: _______________________________________________________________________

Ha: _______________________________________________________________________.
The “t” test statistic for the significance of the slope is equal to __________ with an
associated p-value = __________.

Hence, we ______________________ Ho at $\alpha = 0.05$. The data provide __________
evidence to conclude that the regression equation is useful for predicting the Fertility Rate.

A 95% confidence interval for the slope of the (population) regression line is given by
LL = _______________ and UL = _______________

The predicted number of children per woman (rounded to one decimal place) for a country
with 60% of married women using contraceptives is __________. Show your calculation
below.
II. Example 11.7: Simple Linear Regression.

Read Example 11.7 on page 596 in the text. This is a simple linear regression problem.

The Explanatory variable X is ____________________________ and the Response variable Y is ____________________________.

There are sample data from \( n = 15 \) fires. Do the data provide sufficient evidence at the 5% significance level to conclude that the regression equation is useful for predicting the response \( Y \)? Follow the steps below.

The LS regression equation (round the y-int and slope to 2 decimal places) is:

________________________________________________________________________

The estimated standard error of the regression model (rounded to 2 decimals) is

\( s = \) _________. Hence 95% of all prediction errors while predicting the fire damage are expected to fall between _________ dollars and _________ dollars.

The coefficient of determination is equal to __________ meaning that

________________________________________________________________________

________________________________________________________________________

The hypotheses (in symbols and words) for testing the utility of the regression equation are:

\( H_0: \) _________________________________________________________________

\( H_a: \) _________________________________________________________________
The value of the “F” test statistic for the utility of the regression equation is _______.

with a p-value = _______. Therefore, we ______________________ Ho at α = 0.05.

The data provide ___________________ evidence at α = 0.05 to conclude that the
regression equation with __________________________as explanatory variable is useful for predicting ________________________________.

The predicted fire damage (in dollars) for a burning house located 2.5 miles away from the nearest fire station is _________________. Show your calculation below.

Note: Highlight the results in the SPSS output used for this section
III. Exercise 13.55: Testing multinomial probabilities with a Chi-Square test for categorical data.

Read Exercise 13.55 on page 754 in the text. This is a problem about multinomial probabilities (one-way table). We are testing hypotheses about four probabilities simultaneously. The test statistic has a chi-square probability distribution with degrees of freedom $df = k-1$ where $k$ represents the number of category probabilities to be tested. The test statistic involves the observed and expected frequencies for each category. We are comparing frequencies on four different Hearing Loss statuses. Do the data provide sufficient evidence at the 5% significance level to indicate a difference between the probabilities of the four Hearing Loss statuses?

The hypotheses (numerically and in words) for this problem are:

$H_0$: ____________________________ 

______________________________

$H_a$: ____________________________ 

______________________________

The expected frequencies (under $H_0$) for the four Hearing categories are:

Status 1 = _____, Status 2 = _____, Status 3 = _____, Status 4 = _____,

Since $df = _____$ the rejection region at $\alpha = 0.05$ is $RR = \{ \_______________ \}$. The value of the test statistic is equal to ___________ and hence we ____________________ the null hypothesis. The data provide ______________ evidence to conclude that

______________________________

______________________________

The p-value of the test is equal to ___________ indicating ____________________
IV. Exercise 13.66: Chi-Square Test for Independence/Dependence of Two Categorical Variables

Read Exercise 13.66 on page 757 in the text. This is a test of hypotheses about the independence/dependence of two categorical variables. The test statistic has a Chi-Square probability distribution with degrees of freedom $df = (\text{rows} - 1)(\text{cols} - 1)$. The test statistic is based on the difference between the observed and expected frequency for each cell of the contingency (two-way) table. Do the data provide sufficient evidence at the 5% significance level to conclude that the Row Variable and Column Variable are dependent?

The row variable is _______________________ and has _____ levels. The column variable is __________ ______________ and has _____ levels. The number of No Stops trials is _____. The number of trials with exactly 2 kills is equal to _____.

The hypotheses (in words) using the context of this problem are:

Ho: ________________________________________________________________

_____________________________________________________________________

Ha: ________________________________________________________________

_____________________________________________________________________.

Since $df = _____$ the rejection region is given by RR = { ________________ }.

The test statistic is equal to ___________ with a p-value _____________.

Hence, we __________________________ the null hypothesis at $\alpha = .05$. The data provide __________________ evidence at $\alpha = .05$ to conclude that

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

Note: Highlight the results in the SPSS output used for this section