Elementary Statistics

Test 4

May 04, 2009

Name:……..

18+19+18+16+16+2+2+1+1+2+2+1+2=100

1. National average (9 month) salary of a starting faculty member in a regional university is $42,000 in 2005. A researcher wants to find out whether the inflation adjusted salary is more than $42,000 for starting faculty members in 2008. She collects data from 22 faculty members and found out that the sample mean is $43,250 and standard deviation is $5,500 for the inflation adjusted salary. Test whether the inflation adjusted average salary has increased in three years at 0.05 level of significance. Find the p-value too.

   $H_0$:

   $H_a$:

   LOS:

   Formula and calculation:

   Graph:

   Decision:

2. The average hourly wage last year for members of a certain hospital custodians was $6.32 with a standard deviation of $0.54. This year a sample of 50 custodians had an average hourly wage of $6.51. Test the administration's claim that the average hourly wage has increased at 0.05 level of significance. Assume that the hourly wages are normally distributed. Also find the p-value.

$H_0$:

$H_a$:

LOS:

Formula and calculation:

Graph:

Decision:

Decision in terms of the problem:

p-value:

3. Dr. Jay's advises usually take two career paths. Some go to graduate school for PhD programs in statistics and others go to become actuaries and take professional exams while working. He wants to check whether the actuaries earn more than the PhDs, 7 years after graduating from college. His data is summarized below.

| Group | Sample size | Mean salary in thousands of US dollars (Adjusted for 12 month) | Standard deviation in thousands of US dollars |
|---|---|---|---|
| 1.Actuaries | 35 | \$80 | 5.8 |
| 2.PhDs | 40 | \$77 | 6.2 |

Test an appropriate hypotheses at 0.05 level of significance.

$H_0$:

$H_a$:

LOS:

Formula and calculation:

Graph:

Decision:

4. It has been claimed that 30% of all families moving away from California move to Arizona. In a random sample of the records of several large van lines, it is found that the belongings of 104 of 400 families moving away from California were shipped out to Arizona. Test whether the true proportion of families moving away from California move to Arizona is less that 30% at 0.01 level of significance.

$H_0$:

$H_a$:

LOS:

Formula and calculation:

Graph:

Decision:

p-value:

5.

a. Find the regression line.

b. Find the correlation coefficient.

c. Test $H_0 : a = 0$ against $H_a : a \neq 0$. Use $\alpha$ =0.01. (intercept)

p-value= Decision:

d. Test $H_0 : b = 0$ against $H_a : b \neq 0$. Use $\alpha$ =0.01. (Slope)

p-value= Decision:

e. Find a **95%** confidence interval for the intercept.

f. What is the predicted value of $Y$ when $X$ =10?

SUMMARY OUTPUT

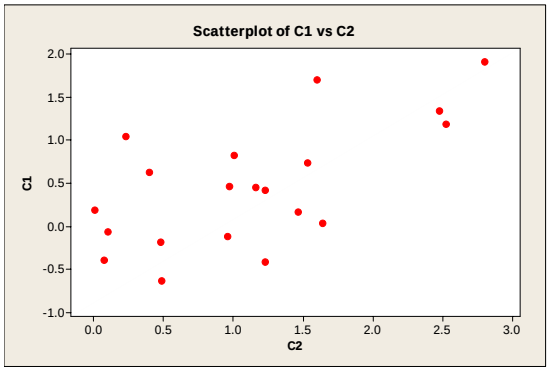
| *Regression Statistics* | |
|---|---|
| Multiple R | 0.767558 |
| R Square | 0.589146 |
| Adjusted R Square | 0.563467 |
| Standard Error | 2.115461 |
| Observations | 18 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 1 | 102.675 | 102.675 | 22.94324 | 0.0002 |
| Residual | 16 | 71.60278 | 4.475174 | | |
| Total | 17 | 174.2778 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
|---|---|---|---|---|---|---|
| Intercept | 2.986111 | 1.086715 | 2.747834 | 0.014298 | 0.682379 | 5.289843 |
| X Variable 1 | 0.925 | 0.193114 | 4.789911 | 0.0002 | 0.515616 | 1.334384 |

6. Guess the correlation coefficient between C1 and C2.



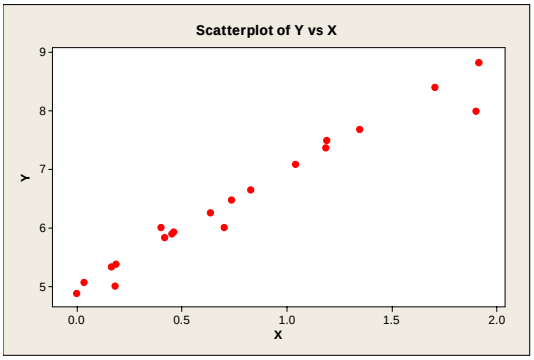

(a) -0.78

(b) -0.04

(c) 0.64

(d) 0.93

7. Use your pencil to fit a line to the data and guess the followings.




a. $Y$ -intercept =

b. Slope =

c. What is the value of $Y$ when $X$ =1.0?

8. What is the highest value the correlation coefficient can take?

9. If the slope is positive then the correlation coefficient will be _____.

10. Reject the null hypothesis when the p-value is _____ than $\alpha$ .

11. What is the only way to reduce the probability of the two types of errors simultaneously?

12. Once you fit a regression line to a set of data, can you use the line to predict Y values for X values far away from X values in data? Yes / No