1 (12 Points)

Suppose that the population correlation coefficient for two variables is positive, but that if we control for a confounding variable, we find that the correlation coefficient is negative.

This is an example of:

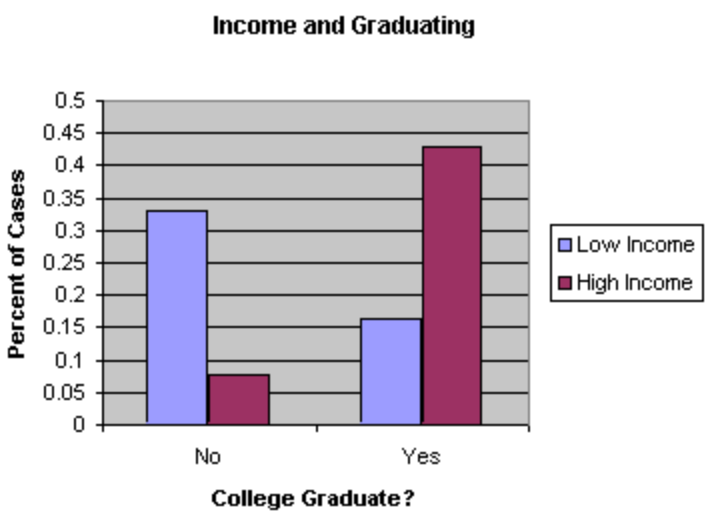
- [1] The Hawthorne effect.
- [2] Simpson's paradox.
- [3] A skewed distribution.
- [4] The Bourne identity.
- [0] no answer or skip this item

Submit Answer

2 (12 Points)

The figure describes data on the income level and college graduation status of a sample of 79 people.

Which of the following statements can NOT be concluded from the graph?



- [1] There were fewer people graduating from college than not graduating from college.
- [2] The risk of low income was higher for people not graduating from college than for people graduating from college.
- [3] People were more likely to have high income if they had graduated from college than if they did not graduate from college.
- [4] The odds of having high income were better for college graduates than for those who did not graduate college.
- [0] no answer or skip this item

Submit Answer

3 (12 Points)

Suppose the number of oak trees distributed in a forest of Northwestern Pennsylvania is a Poisson process with an average of 70 trees per acre.

What is the expected number of trees if the forest covers 80,000 acres?

- [2] 56,000
- [3] 5,600,000
- [4] about 1,142
- [0] no answer or skip this item

Submit Answer

4 (12 Points)

The age at which babies, carried to full term, begin walking have a distribution that can be approximated by a normal curve with a mean equal to 10.38 months and a standard deviation equal to 1.40 months. Use a normal curve calculator of your choice to answer this question.

What percentage of babies carried to full term begin walking in the range from 8.28 months to 12.48 months?

- [1] about 93.3%
- [2] about 86.7%
- [3] about 50.0%
- [4] about 6.7%
- [0] no answer or skip this item

Submit Answer

5 (12 Points)

Choose the option that is best among the four options given for describing the two given situations. Situation I: The distribution of weights of the males in a college class. Situation II: Income of the 30 wealthiest corporations in the United States.

- [1] Both situations could be exponential.
- [2] Situation I could be exponential after shifting, but not Situation II.
- [3] Situation II could be exponential after shifting, but not Situation I.
- [4] Neither could be exponential, even after shifting.
- [0] no answer or skip this item

Submit Answer

6 (12 Points)

A consumer group is interested in comparing the gas consumption (MPG) of four different octane levels of gasoline. Six measurements were taken for each octane level. The summary statistics are given in the following table. An analysis of variance was performed on the data. The F-statistic for testing whether the mean MPG measurements are equal for all four octane levels is 0.67. The table summarizes this information.

Octane Level	Number of measurements	Mean MPG	Variance
87	6	18.60	16.04
89	6	20.42	8.318
91	6	21.43	10.641
93	6	21.40	28.228

What are the degrees of freedom associated with this analysis of variance?

- [1] 4 and 6 df
- [2] 3 and 20 df
- [3] 3 and 24 df
- [4] 4 and 24 df
- [5] 2 and 22 df
- [0] no answer or skip this item

Submit Answer

7 (12 Points)

Boxes of a certain type of cake mix are supposed to contain 24 ounces of cake mix. The machine that fills the boxes puts on average m ounces in the boxes. A random sample of 64 such boxes was obtained and the weights of the cake mix inside each was found. The sample mean was 23.926 ounces, and the sample standard deviation was s = 0.45 ounces. Any deviation from the null is of interest.

What is the standard error of the sample mean?

- [1] about 0.007
- [2] about 0.056
- [3] about 0.450
- [4] about 2.991
- [0] no answer or skip this item

Submit Answer

8 (12 Points)

A larger sample size increases power because:

- [1] Variability in the sampling distributions has increased.
- [2] Variability in the sampling distributions has decreased.
- [3] The real mean gets further away from the value specified in the null hypothesis.
- [4] Hypothesis tests based on extremely large samples always have extremely high power.
- [0] no answer or skip this item

Submit Answer

9 (12 Points)

What can happen if you use a two-sample t-test for independent samples to analyze paired data?

- [1] The two-sample procedure could lead to the wrong conclusion by producing a wider confidence interval than it should and failing to detect a difference in the two means, when it should.
- [2] The two-sample procedure could lead to the wrong conclusion by producing a narrower confidence interval than it should, but will still detect a difference in the two means when it should.
- [3] The two-sample procedure will still lead to the correct conclusion, producing the same identical confidence interval, and failing to detect a difference in the two means when it should.
- [4] The two-sample procedure will still lead to the correct conclusion producing the same identical confidence interval, and the same identical conclusion with regard to the hypotheses.
- [0] no answer or skip this item

Submit Answer

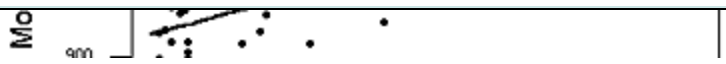
10 (12 Points)

For 64 students, the least squares line for X = Socialhours (number of hours per week spent socializing) versus Y = GPA is GPA = 3.33 - 0.01 * Socialhours. The sum of squares of the residuals is 18.311.

What is the sample standard deviation of the residuals?

- [1] about 0.0100
- [2] about 0.2861
- [3] about 0.2953
- [4] about 0.5349
- [5] about 0.5435
- [0] no answer or skip this item

Submit Answer



- [1] Pittsburgh, PA.
- [2] Birmingham, AL.
- [3] Chicago, IL.
- [4] San Jose, CA.
- [5] New Orleans, La.
- [0] no answer or skip this item

Submit Answer

12 (12 Points)

A company's salary system includes the following three variables:

- *Income*, the increase in total personal income as a result of the percentage of the population holding a bachelor's degree in thousands of dollars,
- *Rate*, the percentage of high school students enrolling in college during the year following high school graduation,
- *Debt*, the average loan amount for students in dollars

The table below contains least squares estimates of the slopes for some linear models. Answer the following question based on those least squares estimates.

Model	Estimates
Income = a + b * Rate + E	**a: 62.83, b: +0.281**
Income = a + b * Debt+ E	**a: 49.98, b: +0.406**
Income = a + b * Rate + c * Debt + E	**a: 29.95, b: +0.273, c: +0.390**

Fixing the rate, an increase of one dollar in the average loan amount is associated with an increase of _____ in income.

- [1] 0.273 thousand dollars
- [2] 0.281 thousand dollars
- [3] 0.390 thousand dollars
- [4] 0.406 thousand dollars
- [0] no answer or skip this item

Submit Answer

13 (12 Points)

Which factor cannot be controlled by investigators in determining the sample size necessary to make accurate estimates and conduct powerful statistical tests?

- [1] The specific types of statistics and/or hypothesis tests employed by the researcher.
- [2] The natural variation of the response variable.
- [3] The researcher's goals for accuracy and power.
- [4] The general design structure.
- [0] no answer or skip this item

Submit Answer

14 (12 Points)

Which statement best describes the concept of ANOVA?

- [1] ANOVA is defined as the variation among responses from experimental uses.
- [2] It is a technique for splitting the variation among responses into its different components.
- [3] It is a technique for computing the total of the squared distances between the responses and the corresponding group means.
- [4] It is a technique for computing the variation among responses from experimental units at the same level of factors.
- [0] no answer or skip this item

Submit Answer

15 (12 Points)

Students in a technology class performed a study examining the pollution level of automobiles. Three categories of exhaust filters were: standard, new1 and new2; each of the three filters was randomly placed on each of twelve vehicles. The pollution levels of the automobiles in each of the three categories were recorded. The following ANOVA table was produced:

Analysis of Variance Results					
Source	**df**	**SS**	**MS**	**F-ratio**	**P**
Factor (Exhaust)	2	3501	1751	2.19	0.1278
Error	3	26373	799		
Total	35	29874			

How is the F-ratio computed?

- [1] The F-ratio is computed as the ratio of Error Mean Square to the Factor (Exhaust) Mean Square.
- [2] The F-ratio is computed as the ratio of Factor (Exhaust) Sum of Squares to the Error Sum of Squares.
- [3] The F-ratio is computed as the ratio of Factor (Exhaust) Mean Square to the Error Mean Square.
- [4] The F-ratio is computed as the ratio of Error Sum of Squares to the Factor (Exhaust) Sum

of Squares.

- [0] no answer or skip this item

Submit Answer

16 (12 Points)

A study is conducted to compare the mean volatility (risk) among four mutual stock funds, FastGrowth, TechFunds, ClassicCompanies and IncomeGrowth. The mean volatility of the mutual funds computed from the samples are 18.3, 18.2, 13.47 and 14.9 respectively for the four mutual stock funds.

What is the research question of interest here?

- [1] Are the differences in sample mean volatilities supported by the data?
- [2] Are the differences in the population mean volatilities supported by the data, or could the observed differences in sample mean volatilities have occurred by chance?
- [3] Are the sample mean volatilities equal?
- [4] Are there differences in the ratios of volatilities by four different mutual stock funds?
- [0] no answer or skip this item

Submit Answer

17 (12 Points)

Which of the following statements is false?

- [1] The experimentwise error rate increases as the number of population means increases.
- [2] The experimentwise error rate decreases as the comparisonwise error rate decreases.
- [3] The experimentwise error rate increases as the number of population means decreases.
- [4] The number of possible treatment pairs increases as the number of population means increases.
- [0] no answer or skip this item

Submit Answer

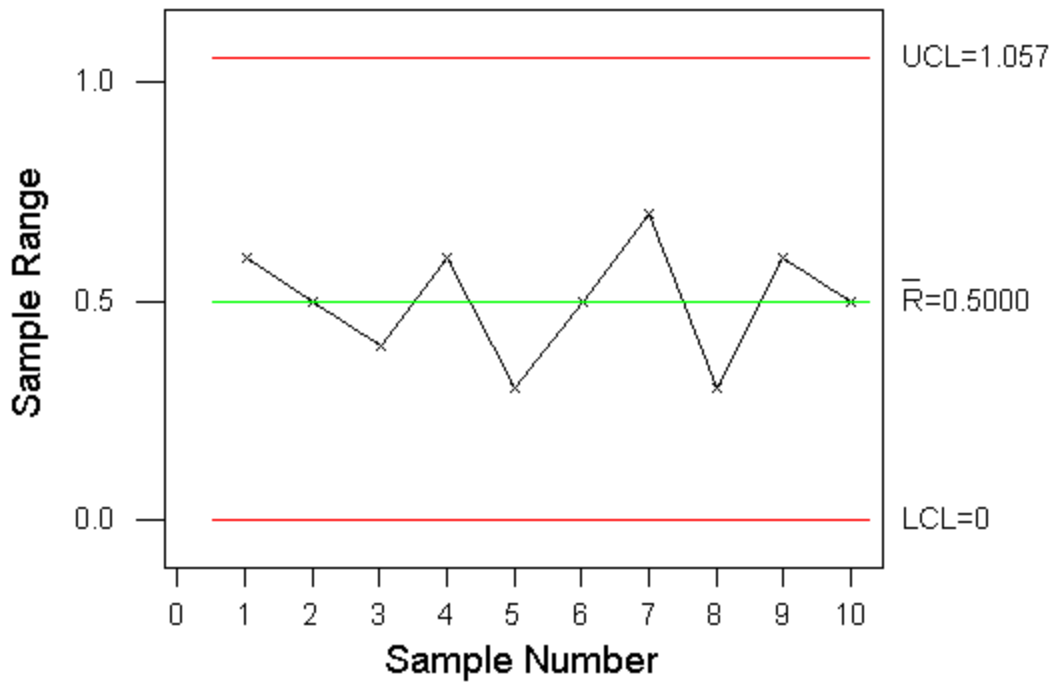
18 (12 Points)

A process that is said to be in-control is a process that has:

- [1] Only special cause variation present.
- [2] Only common cause variation present.
- [3] Both special and common cause variation present.
- [4] Neither special nor common cause variation present.
- [0] no answer or skip this item

Submit Answer

19 (12 Points)



This is a control chart for the sample ranges obtained in a study of the thickness of metal wires. Ten samples of five wires each were obtained, the thickness measured, the sample ranges calculated, and the accompanying range chart prepared. Measurements are in microns.

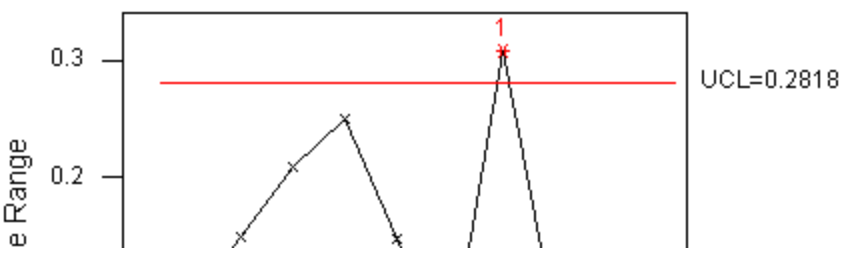
Is the variability of thickness of metal wires in control?

- [1] Test 2 does not apply, since the process centering (as measured by the mean) is not in control.
- [2] If no processing changes are made, and if no outside sources of variability enter the process, we have reason to believe that an estimate will not be valid for the average thickness of metal wires made in the near future.
- [3] A trend like those described in Test 2 is present, thus an estimate will not be valid for the center of the thickness of metal wires made in the near future.
- [4] If no processing changes are made, and if no outside sources of variability enter the process, we have reason to believe that an estimate will be valid for the average thickness of metal wires made in the near future.
- [0] no answer or skip this item

Submit Answer

20 (12 Points)

A control chart for the sample ranges obtained in a study of the thickness of silicon oxide on wafers used in integrated circuits follows. Ten samples of four wafers each were



obtained, the thickness of the silicon oxide measured, the sample ranges calculated, and the accompanying chart prepared. Measurements are in microns.

Which test applies to this range chart?

- [1] Test 1 does apply; we can not use the sample data to obtain a valid estimate of the process variability until the ranges are within the control limits.
- [2] Test 2 does apply; we can not use the sample data to obtain a valid estimate of the process variability until the ranges are within the control limits.
- [3] Test 3 does apply; we can not use the sample data to obtain a valid estimate of the process variability until the ranges are within the control limits.
- [4] Test 4 does apply; we can not use the sample data to obtain a valid estimate of the process variability until the ranges are within the control limits.
- [0] no answer or skip this item

Submit Answer

21 (12 Points)

Many economic time series, such as unemployment rates, gross national product, and number of housing starts, are only produced quarterly and would have fluctuations due to (i) _____ effects.

On the other hand, time series such as annual precipitation in a fixed location may exhibit strong (ii) _____ behavior. This may be linked to other phenomena, such as tracks of ocean currents, which have cyclical, but irregular patterns. In this case, it would not make sense to view these cyclical components as being of a fixed period, as required for a seasonal component.

Indicate the missing two words.

Submit Answer

22 (12 Points)

A horticulturist is interested in comparing mean yields of crops as a result of using 4 different fertilizers. He divides an entire field into 20 equal plots. He then randomly assigns each of 4 fertilizers to 5 different plots. After harvesting the crop from each plot, he records the yield from each plot and prepares an ANOVA table. The computations showed:

ANOVA for Yield				
Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	F-ratio
Fertilizer				

Error			0.80	
Total		17.30		

Find the degrees of freedom for (i) Fertilizers, (ii) Error, and (iii) Total.

Submit Answer

23 (12 Points)

The number of server crashes per week can be modeled as a Poisson distribution. Suppose that your manufacturer says that there will be on average only 1 server crash every three years for a model you are purchasing.

(i) What is the rate per month and (ii) what is the chance that you have one or more server crashes in the first month?

Submit Answer

24 (12 Points)

The police department in a large city gives an entrance exam to all applicants for positions in the department. The distribution of scores is approximated by a normal curve with a mean of 70 and a standard deviation of 8. Suppose that applicants are considered for a job only if the score on the entrance exam is above the 80th percentile.

Based on the normal curve model, what entrance score is required for job consideration?

Submit Answer

25 (12 Points)

A professional card dealer dealt 250 hands of five cards each. (Each hand was dealt after shuffling a complete deck.) The average number of Aces in the 250 hands was 0.512. The population mean is 5/13 = 0.385 if the shuffling is fair. The sample standard deviation of the 250 hands was s = 1.12.

(i) Find the standard error of the sample mean and (ii) find the z-statistic.

Submit Answer

26 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Number_customers_4Hour_shift.dat"**. This is one out of 5 questions that will work with this data set.

Conduct a one way ANOVA on all 4 columns of the data set (i.e., for Jenny, Paul, Lee, and Laura), assuming the data represent the number of customers on randomly selected 4 hour shifts for these 4 employees. Report the mean number of customers for these 4 employees.

Submit Answer

27 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Number_customers_4Hour_shift.dat"**. This is one out of 5 questions that will work with this data set.

Based on your one way ANOVA table calculated above, state the null and alternative hypotheses that relate to this one way ANOVA.

Submit Answer

28 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Number_customers_4Hour_shift.dat"**. This is one out of 5 questions that will work with this data set.

Based on your one way ANOVA table calculated above, how do we calculate MS Error? First indicate a formula using terms such as SS Error, df Treatment, etc. and then fill in the numerical values.

Submit Answer

29 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Number_customers_4Hour_shift.dat"**. This is one out of 5 questions that will work with this data set.

Based on your one way ANOVA table calculated above, report the p-value, state whether we reject (or do not reject) the null hypothesis, and draw a conclusion.

Submit Answer

30 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Number_customers_4Hour_shift.dat"**. This is one out of 5 questions that will work with this data set.

For an upcoming promotion, a supervisor wants to determine whether Laura on average serves more customers per 4 hour shift than Lee. Conduct a two sample T-test (with pooled variances). Report the df, t-statistic, and p-value, and draw a conclusion.

Submit Answer

31 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Socio-econ_indicators_States96.dat"**. This is one out of 5 questions that will work with this data set.

Conduct a multiple linear regression analysis with Poverty, Crime, College, and Metro as four explanatory variables and Income as the response variable. Save the residuals.

Report the F-statistic and p-value from the analysis of variance (ANOVA) table for the multiple regression model and interpret the p-value.

Submit Answer

32 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Socio-econ_indicators_States96.dat"**. This is one out of 5 questions that will work with this data set.

Based on your multiple regression model calculated above, report and interpret the p-values associated with Poverty, Crime, College, and Metro.

Submit Answer

33 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Socio-econ_indicators_States96.dat"**. This is one out of 5 questions that will work with this data set.

Based on your multiple regression model calculated above, draw a scatterplot of the Residuals (y) vs. Income (x). Describe the overall appearance of this residual plot, a possible trend, and identify the most striking outlier (i.e., say which state this is).

Submit Answer

34 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Socio-econ_indicators_States96.dat"**. This is one out of 5 questions that will work with this data set.

Based on your scatterplot of the Residuals (y) vs. Income (x) and the most striking outlier you identified previously, can you think of any possible statistical AND economic explanations for this outlier?

Submit Answer

35 (15 Points)

Use **WebStat**. Load from "Data > Sample Data" the data set **"Socio-econ_indicators_States96.dat"**. This is one out of 5 questions that will work with this data set.

Based on your multiple regression model calculated above, recalculate the model for those variable(s) that had significant slopes (i.e., that had p-values < 0.05 associated with them). Indicate your final model in the form Income = Intercept + Estimate1 * Xvar1 + ... + Estimaten + Xvarn and fill in the numeric values for Intercept, Estimate1 through Estimaten, and indicate what Xvar1 through Xvarn stand for. Should this be our final model or can we remove any further variable(s)?

Submit Answer

When you are done answering all questions above to your satisfaction, press the button below to complete your test.

Mark Test Completed