PSY 423 Exam 1

S. Budoff

2023-09-20

Instructions

This is a two part exam.

Part 1 will test you theoretical knowledge in terms of your ability to correctly identify statistical language we have gone over, and your ability to visualize some of these ideas. Part 1 is closed note-book and should be completed during the class period. For questions that ask you to doodle, I am not looking or accuracy in the sense of a computer rendering, however, if there is a key feature you are trying to convey and are worried I will not see it please write an explanation on the side. Please answer your questions on this paper exam and turn it in when you are ready. Each question is worth 2 points.

Part 2 will be take home and you can use any resources you would like, except for your friends/family/acquaintances/housemates/co-workers/etc (arch-enemies are ok but I will need concrete proof that you hate each other, with p-values and examples). This portion of the exam will be easiest to complete using JAMOVI, but you are welcome to use another statistical language or chatGPT plus R. If you choose to use a Large Language Model, please take screen shots of your full chat log as I would like to see you "showing your work" with these tools. Each question will be worth 1 point and I will scale the resultant score to be 50 points.

IMPORTANT NOTE ON GRADING, please do not be stressed about your grade, I will offer unpenalized retakes during office hours if you are unhappy with how you do today. Take a deep breath and good luck :)

Question 1:

Which of the following measures of central tendency is most appropriate to use when dealing with a dataset that includes extreme outliers?

A. Mean

B. Median

C. Mode

D. Variance

Question 2:

What is the primary goal of a measure of central tendency in statistics?

A. To describe the spread of data points.

B. To identify extreme outliers.

C. To summarize the most typical or representative value in a dataset.

D. To calculate the variability within the data.

Question 3:

Which of the following statements best describes the Central Limit Theorem (CLT)?

A. The CLT states that the mean of a sample is equal to the mean of the population.

B. The CLT states that as the sample size increases, the sampling distribution of the sample mean approaches a normal distribution, regardless of the shape of the population distribution.

C. The CLT only applies to populations with a normal distribution.

D. The CLT applies to the standard deviation of a population.

Question 4:

You collect 50 random samples from a non-normally distributed population, and for each sample, you calculate the sample mean. According to the Central Limit Theorem, what can you expect regarding the distribution of these sample means?

A. The sample means will follow a normal distribution.

B. The sample means will follow the same distribution as the population.

C. The sample means will become increasingly skewed.

D. The sample means will not follow any specific distribution.

Question 5:

In a survey, you collect data on the annual incomes of 100 individuals in a city. The data appears to be a right skewed distribution with a long tail.

Doodle this distribution's curve:

Question 6

Regarding Question 5, to estimate a typical income in the city, which measure of central tendency should you use, and why?

A. Mean, because it is robust to outliers.

B. Median, because it is not affected by extreme values.

C. Mode, because it identifies the most frequently occurring income.

D. Range, because it accounts for variability.

Question 7:

Which of the following measures of variability gives extreme outliers a vote but not a veto?

A. Interquartile Range (IQR)

B. Standard Deviation

C. Median

D. Range

Question 8:

In a dataset with high variability, which of the following is true?

A. The values are closely clustered around the mean.

B. The range is narrow.

C. The standard deviation is small.

D. The values are spread out over a wide range.

Question 9:

Which measure of variability accounts for the sample size?

A. Standard Error of the Mean

B. Variance

C. Interquartile Range (IQR)

D. Standard Deviation

Question 10:

What does a p-value represent in hypothesis testing?

A. The probability that the null hypothesis is true.

B. The probability of making a Type I error.

C. The probability of observing the data, assuming the null hypothesis is true.

D. The probability of observing the data, assuming the alternative hypothesis is true.

Question 11:

In hypothesis testing, a smaller p-value indicates:

A. Stronger evidence in favor of the null hypothesis.

B. Stronger evidence in favor of the alternative hypothesis.

C. You are very unwilling to make a mistake.

D. A less reliable statistical test.

Question 12:

If the p-value in a hypothesis test is 0.03, what is typically considered the significance level (α) in most research settings?

A. $\alpha = 0.01$

B. $\alpha = 0.05$

C. $\alpha = 0.10$

D. $\alpha = 0.20$

Question 13:

You perform a t-test and obtain a p-value of 0.07 with a significance level (α) set at 0.05. What is the most appropriate course of action?

A. Reject the null hypothesis.

B. Fail to reject the null hypothesis.

C. Conduct a larger sample size.

D. Lower the alpha level.

Question 14:

Which of the following statements about p-values is true?

A. A p-value always falls between 0 and 1.

B. A p-value represents the strength of an effect.

C. A p-value can directly provide the effect size.

D. A p-value is not affected by the choice of the significance level.

Scenario for questions 15-20:

A researcher is conducting a study to compare the mean exam scores of two different groups of students, Group A and Group B. Group A received a new teaching method, while Group B received the traditional teaching method. Group A hada mean score of 85, and group B had a mean score of 87. The researcher collects exam scores from both groups and wants to determine if there is a statistically significant difference between the mean exam scores of the two groups.

Question 15:

What is the correct null hypothesis (H_0) for this study?

A. There is a statistically significant difference between the mean exam scores of Group A and Group B.

B. The mean exam score of Group A is equal to the mean exam score of Group B.

C. The new teaching method is better than the traditional teaching method.

D. The mean exam scores of Group A and Group B are not equal.

Question 16:

What is the correct alternative hypothesis (H_a) for this study?

A. There is a statistically significant difference between the mean exam scores of Group A and Group B.

B. The mean exam score of Group A is equal to the mean exam score of Group B.

C. The new teaching method is better than the traditional teaching method.

D. The mean exam scores of Group A and Group B are not equal.

Question 17:

Which statistical test should the researcher use to compare the mean exam scores of Group A and Group B?

A. Chi-squared test

B. Independent samples t-test

C. Paired samples t-test

D. Single sample t-test

Question 18:

The researcher conducted the statistical test and obtained a p-value of 0.02 with a chosen α level of 0.05. How should the researcher interpret this p-value?

A. The research rejects the null hypothesis

B. The researcher fails to reject the null hypothesis

C. There is a very strong effect size.

D. The p-value is not relevant for this test.

Question 19:

After calculating Cohen's d, the researcher obtains a value of 0.75. How should the researcher interpret Cohen's d in the context of this study?

A. The mean exam scores of the two groups are practically identical.

B. The mean exam scores of the two groups are moderately different.

C. The mean exam scores of the two groups are very different.

D. Cohen's d does not provide information about the mean exam scores.

Question 20:

Please doodle the distributions for these exam scores and label on your x-axis the means (μ_A and μ_B) for each condition.

Question 21

What do you expect to happen to your 95% confidence interval as the sample size decreases

A. The confidence interval becomes narrower.

B. The confidence interval becomes wider.

C. The confidence interval remains the same.

D. The confidence interval becomes more accurate.

Question 22

When should you use the t-test instead of the z-test?

A. When the sample size is small, and the population standard deviation is known.

B. When the sample size is large, and the population standard deviation is known.

C. When the sample size is small, and the population standard deviation is unknown.

D. When the sample size is large, and the population standard deviation is unknown.

Last Push

For the following quesitons you are given formal math equations, please circle the part or parts of the equation that have to do with the given plain english idea.

Question 23:

In the z distribution standardization formula which part(s) of the equation shifts the data?

$$z = \frac{X - \mu}{\sigma}$$

Question 24:

In the z distribution standardization formula which part(s) of the equation stretches/squishes the data?

$$z = \frac{X - \mu}{\sigma}$$

Question 25:

In both versions of the t-test formula which part(s) of the equation accounts for the the sample size?

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

MC Answer key - Sam Time 9:02

1) B
2) C
3) B
4) A
5) DOODLE
6) B
7) B
8) D
9) D
10) A
11) C
12) B
13) B
14) A
15) B
16) D
17) B
18) A
19) B
20) DOODLE
21) B
22) C