

University of Toronto (UofT) PSY201 Statistics I Midterm 1 Practice Exam (Sample)

Study Guide



Everything you need from our exam experts!

Copyright © 2025 by Examzify - A Kaluba Technologies Inc. product.

ALL RIGHTS RESERVED.

No part of this book may be reproduced or transferred in any form or by any means, graphic, electronic, or mechanical, including photocopying, recording, web distribution, taping, or by any information storage retrieval system, without the written permission of the author.

Notice: Examzify makes every reasonable effort to obtain from reliable sources accurate, complete, and timely information about this product.

SAMPLE

Questions

SAMPLE

- 1. In statistics, what is the primary use of categorical variables?**
 - A. To quantify measurements**
 - B. To describe a phenomenon qualitatively**
 - C. To provide numerical data for analysis**
 - D. To indicate a specific value in a data set**
- 2. What type of data is affected most significantly by extreme values?**
 - A. Mode**
 - B. Median**
 - C. Mean**
 - D. Weighted mean**
- 3. If the highest score in a test is 95 and the lowest is 10, what is the range?**
 - A. 85**
 - B. 75**
 - C. 90**
 - D. 80**
- 4. What is the relationship between standard deviation and variance?**
 - A. Standard deviation is the average of the variance**
 - B. Variance is the square root of standard deviation**
 - C. Standard deviation is the square root of the variance**
 - D. Variance is always larger than standard deviation**
- 5. How is variance defined in statistical terms?**
 - A. A measure of the average distance from the mean**
 - B. A measure of how much the values in a data set differ from the mean**
 - C. The difference between the highest and lowest values**
 - D. A count of the frequency of each value**

- 6. What does the significance level (alpha) represent in hypothesis testing?**
- A. The probability of making a Type I error**
 - B. The threshold for determining statistical significance**
 - C. The range of values for the p-value**
 - D. The variance within a dataset**
- 7. What is the definition of a variable in research?**
- A. A constant characteristic for every individual**
 - B. A characteristic or condition that varies among individuals**
 - C. A measurement that cannot change**
 - D. A type of data that is always fixed**
- 8. When analyzing data, what is a limitation of using the range as a measure of variability?**
- A. It is difficult to calculate**
 - B. It does not measure the 'average' score**
 - C. It does not account for all values in the distribution**
 - D. It is difficult to interpret**
- 9. What is the definition of range in statistics?**
- A. The average of all scores in a distribution**
 - B. The middle score in a distribution**
 - C. The difference between the highest and lowest scores in a distribution**
 - D. The total number of scores in a distribution**
- 10. What aspect does a research design focus on?**
- A. The methodology used to gather data**
 - B. The ethical considerations involved in the study**
 - C. The theoretical framework supporting the research**
 - D. The personal bias of the researcher**

Answers

SAMPLE

1. B
2. C
3. A
4. C
5. B
6. B
7. B
8. C
9. C
10. A

SAMPLE

Explanations

SAMPLE

1. In statistics, what is the primary use of categorical variables?

- A. To quantify measurements**
- B. To describe a phenomenon qualitatively**
- C. To provide numerical data for analysis**
- D. To indicate a specific value in a data set**

The primary use of categorical variables is to describe a phenomenon qualitatively. Categorical variables are used to categorize or label distinct groups or characteristics without implying any numerical value or order. For example, they can represent categories such as gender, hair color, or type of pet, which are all classifications that provide descriptive information rather than numerical data. This qualitative nature allows researchers to analyze patterns, trends, and relationships based on classifications, making it pivotal in fields such as social sciences where understanding group differences and similarities is essential. In contrast to quantitative variables, which can be measured and expressed numerically, categorical variables serve to provide insights into the nature or attributes of the data being studied.

2. What type of data is affected most significantly by extreme values?

- A. Mode**
- B. Median**
- C. Mean**
- D. Weighted mean**

The mean is the measure of central tendency most significantly affected by extreme values, also known as outliers. This is because the mean is calculated by summing all values in the data set and then dividing by the number of observations. When an extreme value, whether very high or very low, is included in this calculation, it can dramatically change the sum and, consequently, the resulting mean. For instance, if a set of test scores is composed primarily of scores in the 70s and 80s, a single score of 0 or 100 will pull the mean down or push it up, respectively, creating a misrepresentation of the typical score. This sensitivity to extreme values means that the mean can provide a skewed view of the data's central tendency, particularly in datasets with outliers. In contrast, the median, which is the middle value when all observations are sorted, is not affected by extreme values since it depends solely on the rank order of values, not their specific magnitudes. Similarly, the mode, being the most frequently occurring value, is not influenced by extreme values either. The weighted mean, depending on its construction, can be comparatively less influenced than the simple mean but can still reflect the impact of outliers if extreme

3. If the highest score in a test is 95 and the lowest is 10, what is the range?

A. 85

B. 75

C. 90

D. 80

To find the range of a set of scores, you subtract the lowest score from the highest score. In this case, the highest score is 95 and the lowest score is 10. The calculation would be as follows: $\text{Range} = \text{Highest Score} - \text{Lowest Score}$ $\text{Range} = 95 - 10$ $\text{Range} = 85$ This calculation indicates that the spread of scores from the lowest to the highest is 85, which is why the answer is 85. This concept is essential in statistics as it provides a quick sense of the distribution of the data points, reflecting how much variation exists within a set of scores.

4. What is the relationship between standard deviation and variance?

A. Standard deviation is the average of the variance

B. Variance is the square root of standard deviation

C. Standard deviation is the square root of the variance

D. Variance is always larger than standard deviation

The relationship between standard deviation and variance is fundamentally rooted in how they are calculated. Standard deviation is defined as the square root of the variance. Variance itself is a measure of how much data points in a dataset spread out from the mean and is calculated by taking the average of the squared differences from the mean. Since variance is based on squared values, it can often yield larger numbers, which can make interpretation challenging when assessing variability in the original units of the data. This is where the standard deviation comes into play. Taking the square root of variance transforms it back to the original units of measurement, making it more understandable and directly comparable to the data itself. This mathematical relationship shows that while variance represents the overall spread of the dataset through squaring, standard deviation provides a more intuitive sense of that spread by reverting to the original measurement scale. Thus, understanding that standard deviation is the square root of variance is key in statistical analysis and interpretation of data.

5. How is variance defined in statistical terms?

- A. A measure of the average distance from the mean
- B. A measure of how much the values in a data set differ from the mean**
- C. The difference between the highest and lowest values
- D. A count of the frequency of each value

Variance is defined as a measure of how much the values in a data set differ from the mean. It quantifies the degree of spread in a set of data points. Specifically, variance assesses the average of the squared differences from the mean, providing insight into the distribution of data points around the mean value. A higher variance indicates greater dispersion among the values, while a lower variance suggests that the values are closer to the mean. The other choices highlight different concepts in statistics. The first option refers to the concept of standard deviation instead of variance, which indeed measures the average distance from the mean but is not the definition of variance. The third choice talks about the range, which simply measures the distance between the highest and lowest values, rather than the overall spread of the data set. The fourth option relates to frequency distribution, which counts how often each value occurs in the data but does not provide information about variability or dispersion.

6. What does the significance level (alpha) represent in hypothesis testing?

- A. The probability of making a Type I error
- B. The threshold for determining statistical significance**
- C. The range of values for the p-value
- D. The variance within a dataset

The significance level, commonly denoted as alpha, plays a crucial role in hypothesis testing by representing the probability threshold at which researchers decide whether to reject the null hypothesis. It specifically indicates the level of risk the researcher is willing to take in making a Type I error, which occurs when the null hypothesis is true, but is incorrectly rejected. Although the option recognizing it as a threshold for determining statistical significance is closely related to the concept, the most direct representation of alpha is its role as the probability of making a Type I error. By setting alpha, researchers define a cutoff for p-values after conducting a statistical test. If the obtained p-value is less than or equal to alpha, the result is deemed statistically significant, leading to the rejection of the null hypothesis. Furthermore, understanding alpha as a threshold helps contextualize the relationship between hypothesis testing and statistical significance, but it primarily serves to quantify the risk of incorrectly concluding that an effect exists when there is none. The other choices either address related but distinct aspects or do not capture the primary function of alpha in hypothesis testing.

7. What is the definition of a variable in research?

- A. A constant characteristic for every individual**
- B. A characteristic or condition that varies among individuals**
- C. A measurement that cannot change**
- D. A type of data that is always fixed**

A variable in research is defined as a characteristic or condition that varies among individuals. This fundamental concept is essential in statistics and research methodology, as it allows researchers to differentiate between the subjects being studied and to analyze the relationships between different characteristics. Variables can take on different values or categories, which is crucial for conducting experiments, surveys, and observational studies. In the context of data analysis, recognizing that a variable can change or take on different outcomes among individuals enables researchers to explore trends, make comparisons, and draw conclusions based on the varying data collected. This variability is what makes quantitative analysis possible, as it allows for statistical testing and the evaluation of patterns within the data.

8. When analyzing data, what is a limitation of using the range as a measure of variability?

- A. It is difficult to calculate**
- B. It does not measure the 'average' score**
- C. It does not account for all values in the distribution**
- D. It is difficult to interpret**

The range is defined as the difference between the highest and lowest values in a data set. While it provides a quick measure of variability, a significant limitation is that it does not account for all values in the distribution. Specifically, the range is solely dependent on the two extreme values (the maximum and minimum), which can be significantly affected by outliers. For instance, if a single observation in a dataset is much higher or lower than the others, it can create a misleading representation of the overall variability. Thus, while the range may suggest a larger spread, it does not provide a full picture of how data points are distributed within that range. In contrast, measures like the interquartile range or standard deviation consider all values in the distribution and provide a more comprehensive understanding of variability. This is why option C highlights a crucial limitation of the range, making it an insufficient measure of variability on its own.

9. What is the definition of range in statistics?

- A. The average of all scores in a distribution**
- B. The middle score in a distribution**
- C. The difference between the highest and lowest scores in a distribution**
- D. The total number of scores in a distribution**

The definition of range in statistics is indeed the difference between the highest and lowest scores in a distribution. This is a fundamental measure of variability that indicates how spread out the values in a dataset are. To compute the range, one simply subtracts the smallest value (the minimum) from the largest value (the maximum). This measure provides a quick sense of the extent of the data, highlighting how much variation exists between extreme values. A larger range suggests more variability in the dataset, while a smaller range indicates that the values are more closely clustered together. The other choices do not accurately reflect the concept of range. The average of all scores pertains to the mean, the middle score relates to the median, and the total number of scores represents the count rather than a measure of spread. Understanding the range is essential for interpreting the variability within data, making option C the correct choice.

10. What aspect does a research design focus on?

- A. The methodology used to gather data**
- B. The ethical considerations involved in the study**
- C. The theoretical framework supporting the research**
- D. The personal bias of the researcher**

A research design primarily focuses on the methodology used to gather data. It encompasses the overall plan for a study, including how data will be collected, measured, and analyzed. This design serves as a blueprint for the research process, ensuring that the data collection methods will adequately address the research questions and hypotheses. The methodology includes various components, such as selecting the type of research (e.g., experimental, survey, observational), defining the population of interest, choosing sampling methods, and determining the tools and instruments for data collection. A well-structured research design is crucial for the validity and reliability of the results, as it directly impacts how effectively the research can draw conclusions about the studied phenomena. Other aspects mentioned, such as ethical considerations, theoretical frameworks, and potential biases of the researcher, are important to the research process but are not the primary focus of research design. Ethical considerations relate to the integrity and safety of participants, theoretical frameworks provide the context and rationale for the study, and researcher biases involve the personal beliefs or perspectives that might influence the analysis or interpretation of data. However, these elements do not define the structure and approach of how data is actually gathered, which is the core of research design.