

# BIOSTAT702 Midterm 1 – Practice Questions

Fall 2025

## Wooclap Questions from Class

Match the population to its description.

1. Study Sample
  2. Target Population
  3. Analysis Set
- 
- a) All adults over age 50 at risk for cardiovascular disease
  - b) 1,740 participants with full lab and survey data at 1-year follow up
  - c) 2,000 patients recruited from 5 hospitals in North Carolina

In which of the following cases should you include p-values in your Table 1?

- a) When you are conducting an observational study and want to compare groups before adjustment.
- b) When you want to test whether baseline characteristics differ significantly between treatment groups in a randomized trial.
- c) When you want to highlight imbalances in covariates that may need adjustment in your analysis.
- d) When your sample size is large and you expect small differences to be statistically significant.
- e) None of the above — p-values should not be included in Table 1.

What is the primary difference between standard deviation of a set of data points and standard error?

- a) Standard deviation describes variability in individual data points, while standard error describes variability in sample statistics like the mean.
- b) Standard deviation is only used in hypothesis testing, while standard error is used to describe variability in raw data.
- c) Standard deviation measures how far a sample mean is from the population mean, while standard error measures how spread out the data are.
- d) There is no difference; the terms are interchangeable.

Which of the following is not a commonly cited flaw of null hypothesis significance testing (NHST) and the use of p-values?

- a) P-values do not provide the probability that the null hypothesis is true.

- b) The conventional threshold of  $p < 0.05$  is arbitrary and can lead to inconsistent conclusions.
- c) P-values are sensitive to sample size, which can lead to misleading conclusions.
- d) NHST encourages researchers to focus on estimation and uncertainty rather than binary decisions.

Sort the following hypothesis tests from LEAST to MOST number of assumptions required.

- a) Wilcoxon Signed Rank Test
- b) T-Test
- c) Sign Test

Suppose we carry out a one-sample proportion test on a sample of students, where 32% are above 6 feet tall. Testing against a null hypothesis of 30%, we obtain a p-value of 0.04, but the 95% confidence interval for the proportion is (29.5%, 34.5%), which includes the null value. Why might this apparent discrepancy occur?

- a) The confidence interval is centered around the null value, so it should match the p-value exactly.
- b) The confidence interval is incorrect because it should exclude the null value if the p-value is below 0.05.
- c) The test statistic uses the standard error based on the null value, while the confidence interval uses the standard error based on the sample estimate.
- d) The discrepancy is due to rounding errors in the confidence interval.

## Other Practice Questions

Choose from [decreases / stays the same / increases] to fill in the following blanks regarding confidence intervals for a mean.

- a) Holding everything else fixed, as the standard error decreases, the width of the confidence interval  
\_\_\_\_\_.
- b) Holding everything else fixed, as the Type I error probability increases, the width of the confidence interval  
\_\_\_\_\_.

If a random variable  $Y$  has population variance  $\sigma^2$ , then the distribution of the mean of a random sample will have variance  $\sigma^2/n$  even if  $n = 2$ .

- a) True
- b) False

Confidence Intervals for a population parameter will always contain the sample estimate of the parameter.

- a) True
- b) False

## Some Topics / Concepts to Know

- Bias, Internal/External Validity
- P-values
- Sampling Distributions / Standard Error
- Confidence Intervals
- Differences between Hypothesis Tests for the "center" or "typical value" of a distribution (i.e., mean/median)
- Basics of Linear Regression
- Linear Regression sums of squares