

Instructions

Calculators (simple & graphing) will be allowed. You may bring two pages of hand-written, double sided notes. You will receive a copy of the z-table, t-table and chi-square table. You will have 3 hours to complete the exam.

Length

The final will have less questions per topic than the quizzes but additional questions on the new content after quiz 4. You will have to work fluently to complete it. If you typically took the full class to complete the quizzes, you may require the full 3 hours to complete the final exam.

- - - - -

Study Design

- From a description of an experiment
 - Identify the IV/DV
 - Explain potential confounds
- Explain the importance of the following to a scientific experiment
 - Manipulation
 - Control
 - Random Assignment
 - Counterbalancing
- Explain the difference between causation/correlation

Frequency

- Understand the difference between frequency, relative frequency and cumulative frequency
- Understand not just how mean/median/mode are calculated, but why they have different values depending on the shape of the distribution
- Identify skew and multiple modes in a histogram/distribution

Mean & Standard Deviation

- Understand the parts of the mean and standard deviation formulas
- Know when to use the population versus sample version of standard deviation
- Have an intuitive understanding of variability and how standard deviation is a measure of variability
- Understand the difference between the phrases “variability”, “variance” and “standard deviation”

Normal Distribution

Describe key characteristics of the normal distribution

Determine if a given distribution is (reasonably) normal or not

Understand what it means that the normal distribution is a fundamental assumption in our statistical tests

Distribution of Sampling Means

Know how the distribution of sampling means is created

Understand why it is important to hypothesis testing

Understand how the mean and standard deviation of the DSM compare to the mean and standard deviation of the original population

Understand how and why the DSM changes shape with larger sample sizes

Understand what are the entities (things contributing a value) in the population, in a sample and the DSM

Hypothesis Testing

Know the 5 steps of the hypothesis test as used in class

Understand what it means that a hypothesis test is making an inference about data (it is an inferential statistic) versus simply describing data (descriptive statistics)

Be able to describe how a hypothesis test:

- defines what happens by chance
- identifies events that rarely happen by chance
- makes a decision about what is going on based on how rare an event is

Understand the difference between alpha and p (students frequently confuse this)

Describe the difference between the null and alternative hypothesis

Know whether the statistical test is testing the null or alternative hypothesis

Know the 4 different ways we can label a test decision (clue: statsig, p, retain/reject, words of original problem)

Explain why a 1-tailed test is more sensitive than a 2-tailed test, but also why we do not just use a 1-tailed test all the time

Test selection

Identify the appropriate test based on type of data and available information

For each test:

- Select and calculate the correct standard error
- Select and calculate the test statistic
- Calculate degrees of freedom
- Know whether 1vs2 tailed applies, and how to choose
- Find an appropriate critical value

Interpret the meaning of a statistically-significant-outcome and a not-significant-outcome for each test

Power

Understand 4 factors that affect power

Understand what it means that we must assume an effect in order to calculate power

Understand how power relates to a percentage of samples in the alternative distribution

Understand how a scientist would use the result of a power calculation to make decisions about her study/analysis

Confidence Interval

Understand what it means that a hypothesis test is "testing" against a hypothesis and a confidence interval is only estimating something

Understand how confidence level affects the size of the confidence interval

Understand why we can assume the population mean is within the confidence interval (how the confidence interval works)

Outcomes & Errors

Be able to derive the four possible outcomes from memory

Understand what it means that "we can never know reality / if there is really an effect"

Understand the two types of error

Understand how to determine the probabilities of each outcome

2-sample t-tests

Explain why the 2-sample dependent test is more sensitive, but also why it is not always used instead of the independent

Explain why we use weighted averages and pooled variances, and why we can not just take the average of the two sample averages/variances

Correlation

Be able to estimate the r value based on a scatter plot or description of the relationship

Be able to describe the relationship between two variables based on the r value

Identify factors which affect r value (e.g. outliers can make or ruin a strong correlation)

Know typical r values, directions of relationships, strength of relationships

Understand what it means that r squared is "the percent of variance explained"

Know the difference between the two different meanings of "correlation":

- correlation as a mathematic description of a relationship between variables

- correlation as opposed to causation

Explain what the SE of r represents or is used for

Hypothesis test of r

Understand why there is a hypothesis test for correlation

Know how to determine degrees of freedom

Be able to explain what a significant versus non-significant hypothesis test means

Regression

Know what a regression line is used for

Explain the advantages of the "least squares" regression line versus other possible regression lines

Know what the standard error of the estimate is measuring

Chi-square

Differentiate between expected/observed and proportion/frequency

Be able to determine what expected proportions to test against in the null hypothesis based on the problem/question wording.

Understand what degrees of freedom are based on and how this is different from past tests

Chi-square distribution

Explain what it means that there is a distribution of chi-square values

Identify the properties of the chi-square distribution

Know what the chi-square table is used for and how to lookup values

Test for independence

Be able to determine expected proportions for each cell

Understand what a significant versus non-significant test means

Graphing

Identify the type of data being displayed in a graph and what statistical test would be appropriate

Identify how variability in the data is being displayed (if at all)

Estimate (roughly) whether a given statistical test would be significant based on the graph data

Published Results

From an excerpt from a journal, be able to identify:

- the test used
- the number of subjects
- the test statistic
- whether the test was significant

For a given dataset and test result, be able to identify a correct description