Hypothesis Testing

LECTURE 9

Objectives

- Define terms
- Develop good hypothesis sets
- Explain statistical errors (Type I and Type II), statistical power, and effect size.
- Choose between one-tailed and two-tailed tests.
- Use the ten-step procedure to setup problems.

Introduction to hypothesis testing

- A major goal of statistical analysis is to describe a population by examining a sample from that population.
  - Draw conclusions about the parameters of the population.
  - Descriptive statistics

Introduction to hypothesis testing

- Rather than use samples and derived statistics as approximations of the sample population, we can use sample statistics to support or reject hypotheses about the true nature of the population parameters.
  - Inferential statistics

Hypothesis sets

- Hypotheses need to be:
  - Exclusive (Only one of the hypotheses can be true).
  - Exhaustive (All possible outcomes are included).

Hypothesis sets

- Start by stating a null hypothesis.
- A null hypothesis states that there is no difference.
  - Remember earlier that the we stated that a null hypothesis stated that there was no relationship between to the independent variable and the dependent variable -- changing the independent variable will make no difference in the dependent variable.

Hypothesis sets

- Null example: The population mean is not significantly different from 0.
  - H_o:μ=0
- Null example: The population mean is not significantly different from 3.5.
  - H_o: μ =3.5

Hypothesis sets

- If it is concluded that the null hypothesis (H_o) is false then an alternate hypothesis (H_A) is assumed to be true.
  - H_o: μ =0 H_A: $\mu \neq 0$
  - H_o: μ =3.5 H_A: $\mu \neq 3.5$

Hypothesis sets

- You need to state a null hypothesis and alternate hypothesis for each statistical test performed and before the data is collected.

Truncated hypothesis sets

- You usually know something about the system you are studying and can exclude some outcomes as being impossible.
  - H_o: μ =0 H_A: μ >0
  - H_o: μ =3.5 H_A : μ <3.5

Statistical errors in hypothesis testing

- How small a probability will be required to reject the null hypothesis?
  - The probability used as the criterion for rejection is called the significance level and is denoted by α.
    - A probability of 5% or less is commonly used as the criterion for rejecting H_o
  - The value of the test statistic corresponding to α is termed the critical value of the test statistic.

Statistical errors in hypothesis testing

- The significance level needs to be set when the null hypothesis and alternate hypotheses are set (prior to data collection).

Statistical errors in hypothesis testing

	H_o is true	H_o is false
Reject H_o	Type I error	No error
Retain H_o	No error	Type II error

Statistical errors in hypothesis testing

- Type I error
  - Committed with a frequency of α.
- Type II error
  - Committed with a frequency of β. β is generally not known.
  - But β increases as α decreases.
- β and α can both be decreased by increasing n (sample size).

Statistical errors in hypothesis testing

- The power of a statistical test is the ability of the test to reject a false null hypothesis (no error) and is calculated by (1-β).

Effect Size

- Effect size is how much a change in the independent variable causes in the dependent variable.
  - This influences the power of the test.
    - A small effect size, decreases the power of a test.
  - To determine the appropriate sample size, a researcher needs to decide on the minimum power they want and on the minimum effect size they think is meaningful.
    - More on this later.

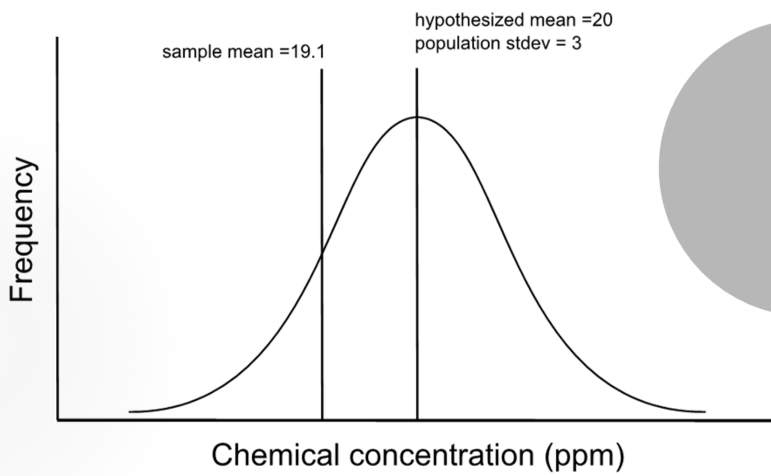
Example

- Suppose a municipality wants to test the hypothesis that particular chemical in the water supply averages 20 ppm.
- The hypothesis set would be:
  - H_o: μ =20 ppm
  - H_A: $\mu \neq$ 20 ppm

Example

- Since we can not test all of the water in the city, we must test a sample of the water.
  - n = 100
  - x = 19.1
  - σ = 3

Example



Example

$$z \cong \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z \cong \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{19.1 - 20}{\frac{3}{\sqrt{100}}} = \frac{-0.9}{\frac{3}{10}} = \frac{-0.9}{0.3} = -3$$

Example

- Z= -3.00, p=0.4987 from Table A
- Since we are interested in the proportion more extreme than this:
  - p = 0.500-0.4987 = 0.0013
- So the probability that we get a sample mean of 19.1 ppm when the actual mean is 20 is p=0.0013.

Example

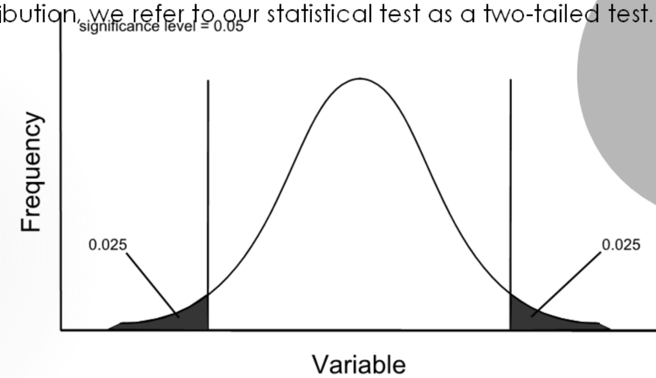
- Since $p<0.05$, we would reject the null hypothesis H_o: $\mu = 20$ ppm
- We would retain the alternate hypothesis H_A: $\mu \neq 20$ ppm

One- or two-tailed

- The probability corresponding to the various significance levels can either be split between the two tails of the distribution or placed all in one tail.

One- or two-tailed

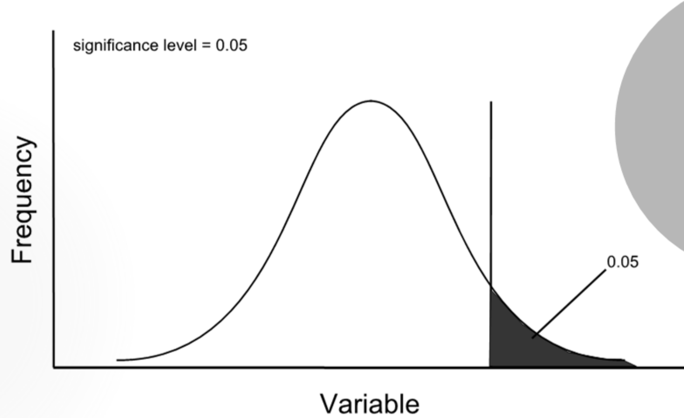
- If the probability of an unlikely outcome is divided between the two tails of the distribution, we refer to our statistical test as a two-tailed test.



One- or two-tailed

- If, however, we are dealing with a truncated hypotheses set in which we assume that any departure from Ho must be in a particular direction, then we should choose to put the entire probability of an unlikely outcome in one tail. We call this type of test a one-tailed test.

One- or two-tailed



Ten step procedure

- Technical paper 14