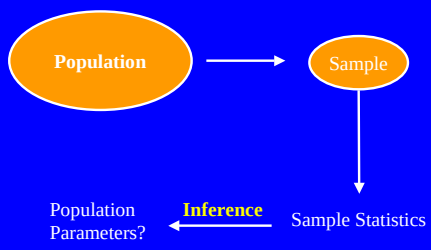
Hypothesis Testing: Basic Concepts

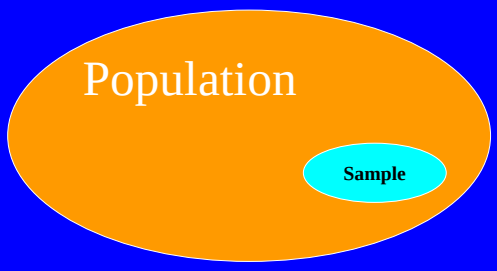
Psychology 320

Gary S. Katz, Ph.D.

Inferential Statistics



Samples from Populations



Sampling from Populations

Non-Random Samples	**Almost-Random Samples**
Systematic Sample	Stratified Random Sample
Convenience Sample	Cluster Sample

Truly Random Samples

Simple Random Sample

Sampling Distributions

- In reality, we take only ***one*** sample of a specific size (*N*) from a population and calculate a statistic of interest.
- Based upon this single statistic from a single sample, we want to know:

"How likely is it that I could get a sample statistic of this value from a population if the corresponding population parameter was ___"

Sampling Distributions

- BUT, in order to answer that question, we need to know what the entire range of values this statistic *could* be.
- How can we find this out?
- Draw *all possible* samples of size *N* from the population and calculate a sample statistic on *each of these samples*.

Sampling Distributions

- A distribution of *all possible* statistics calculated from *all possible* samples of size *N* drawn from a population is called a **Sampling Distribution.**
- Three things we want to know about *any* distribution?
  - Central Tendency
  - Dispersion
  - Shape

Sampling Distributions

- Central Tendency
  - "typical value"
  - usually estimates the population parameter
  - e.g., $\bar{x} \rightarrow \mu$
- Dispersion - Standard Error
  - "variability"
  - **the SD of a sampling distribution is called a Standard Error**
- Shape - depends upon the statistic and assumptions

Why Bother with Sampling Distributions?

- Sampling distributions are the heart of statistical inference using **hypothesis testing.**
- Based upon
  - a hypothesis
  - a sampling procedure
  - a measurement procedure
  - a statistic
  - a sampling distribution
- We make "guesses" about populations.

Steps in Hypothesis Testing

- State the research hypothesis (H_1) and the null hypothesis (H_0).
- Decide how unlikely our sample results have to be in order to reject H_0.
- Decide how to test the hypothesis.
- Decide upon the rejection region of the sampling distribution.
- Calculations!
- Decide whether to reject or retain H_0.

H_0 and H_1

- H_0 & H_1
  - Always stated in terms of population parameters, never sample statistics.
  - Together, define the entire range of possibilities.

H_0: $\mu = 0$

H_1: $\mu \neq 0$

The Null Hypothesis

- H_0
  - the null hypothesis
  - always specifies **a single value** (has the equals sign in it somewhere)
  - says,
    - "Nothing happened."
    - "There's no difference."
    - "Chance is a good explanation for your results."
    - "Your sample statistic doesn't differ from your hypothesized population parameter."

The Null Hypothesis

- H_0
  - In general, is something we **don't** want to find support for.
  - Although, there are some exceptions!
    - H_0 : There is no difference between men and women in performance of military tasks.
    - H_1: There **is** a difference between men and women in performance of military tasks.

The Research Hypothesis

- H_1
  - Also known as the "alternative hypothesis"
  - always specifies **a range of values** (has an inequality in it somewhere) says,
    - "Something happened."
    - "There **is** a difference."
    - "Chance is not a good explanation for your results."
    - "Your sample statistic differs 'significantly' from your hypothesized population parameter."

The Research Hypothesis

- H_1
  - In general, is something we **want** to find support for.
  - Although there are some exceptions!
    - H_0 : There is no difference between men and women in performance of military tasks.
    - H_1: There **is** a difference between men and women in performance of military tasks.

Null & Research Hypotheses

H_0: $\mu = 7.5$lbs

H_1: $\mu \neq 7.5$lbs

Suppose you draw a sample of 100 pregnant women and get the following result:

$$\bar{x} = 6.5lbs$$

Which (H_0 or H_1) describes the universe?

Sampling Error

H_0: $\mu = 7.5$ $\quad \bar{x} = 6.5$

H_1: $\mu \neq 7.5$ $\quad N = 100$

Which (H_0 or H_1) describes the universe?

If we were to draw a sample of 100 people, would you be guaranteed a sample mean of 6.5?

***Sampling Error*: Variability of sample statistics from sample to sample.**

The Big Question

- Is the variation you found due to
  - Sampling Error

  OR
  - Some systematic (real) variability?

The Unsettling Answer(s)

- **YOU** decide on the odds for casting your decision in favor of H_0 or H_1.
- Even after you decide, you're still never certain if you made the right decision.
- All your choices are based upon probability distributions (sampling distributions).
- ***SO, "never say prove in my class."***

Steps in Hypothesis Testing

- State the research hypothesis (H_1) and the null hypothesis (H_0).
- Decide how unlikely our sample results have to be in order to reject H_0.
- Decide how to test the hypothesis.
- Decide upon the rejection region of the sampling distribution.
- Calculations!
- Decide whether to reject or retain H_0.

Rejection Level

- Also known as "alpha" : α
- Also known as the "significance level."
- The probability with which we are willing to reject H_0 when it is in fact correct.
- Determines the "Rejection Region"
  - the set of outcomes of an experiment that will lead to the rejection of H_0

Rejection Level

- Completely up to the discretion of the researcher.
- Typical values:
  .05 (5%) .01 (1%)
- So, scientists regularly reject the null hypothesis if the odds of finding their sample outcome given a specified population parameter are less than 1 in 20.

Steps in Hypothesis Testing

- State the research hypothesis (H_1) and the null hypothesis (H_0).
- Decide how unlikely our sample results have to be in order to reject H_0.
- Decide how to test the hypothesis.
- Decide upon the rejection region of the sampling distribution.
- Calculations!
- Decide whether to reject or retain H_0.

But "where is alpha?"

- Can be located in three different places:
  - Upper tail
  - Lower tail
  - Both tails
- Upper Tail & Lower Tail are "one tailed tests"
- Both Tails are "two tailed tests."

Upper tailed tests

- All of alpha is in the rejection region in the upper tail of the sampling distribution.
- Used to test hypotheses of the form:

H_0: $\mu \leq 100$

H_1: $\mu > 100$

Lower tailed tests

- All of alpha is in the rejection region in the lower tail of the sampling distribution.
- Used to test hypotheses of the form:

H_0: $\mu \geq 100$

H_1: $\mu < 100$

Two Tailed Tests

- Alpha is split into two rejection regions in both tails of the sampling distribution.
- Used to test hypotheses of the form:

H_0: $\mu = 100$

H_1: $\mu \neq 100$

Errors of Inference & Power

- Type I Error
  - The error of rejecting H_0 when it is in fact true.
  - Equal to alpha (α)
- Type II Error
  - The error of retaining (not rejecting) H_0 when it is in fact false.
  - Equal to beta (β)
- Power
  - The probability of correctly rejecting a false H_0
  - Equal to 1- β

Probabilities, Power & Errors

As α gets smaller, β gets larger!

	True State of the Universe	
Decision	**H_0 True**	**H_0 False**
Reject H_0	Type I error $p = \alpha$	**Correct Decision** $p = 1 - \beta$ **Power**
Retain H_0	**Correct Decision** $p = 1 - \alpha$	Type II error $p = \beta$