Hypothesis Testing

Lisa Kuramoto
Centre for Clinical Epidemiology and Evaluation

Overview

Outline

- Review
- Hypothesis testing
- Examples

Learning objectives

- to able to state a good hypothesis
- to be aware of steps for hypothesis testing
- to be able to make decisions based on hypothesis tests

Review

Terminology

- population

  group of individuals or objects that we would like to study

- sample

  subset of a population

Review

Terminology

- population parameter

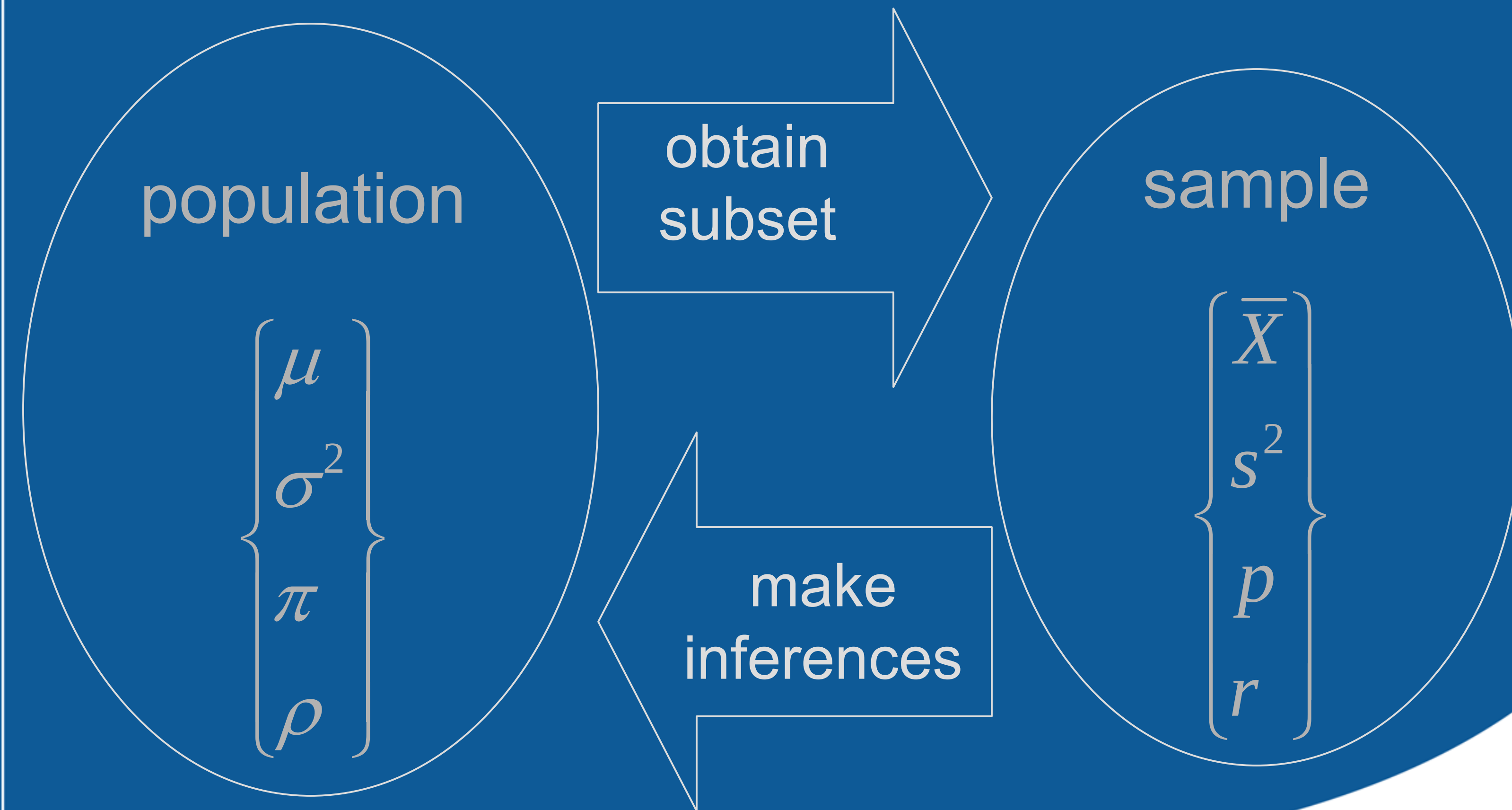
  an unknown quantity that describes a population
- sample statistic

  an estimate of the population parameter
- statistical inference

  process of drawing conclusions about a population based on observations in a sample

Review

Framework for statistical inference



Vancouver CoastalHealth Research Institute
Healthier lives through discovery

Introduction

Motivation of hypothesis testing

- to help make decisions
- to determine if chance alone explains our observations

Hypothesis testing

Hypothesis testing

- process of making decisions based on the likelihood of the data

Main components

- develop hypothesis
- test hypothesis
- make decision

Develop the hypothesis

Research question

- define study population
- specify question of interest

Statistical question

- translate research question into statistical question
- identify population parameter
- state hypothesis

UBC

Vancouver CoastalHealth Research Institute
Healthier lives through discovery

Develop the hypothesis

State hypothesis

- null hypothesis, H_O

  typically states that there is no relationship between the response and explanatory variable(s)

- alternative hypothesis, H_A

  typically states that there is a relationship between the response and explanatory variable(s)

UBC
Vancouver Coastal Health Research Institute
Healthier lives through discovery

Develop the hypothesis

Components to think about

- population(s)
- population parameter
- directional or non-directional

Develop the hypothesis

Set significance level, α

- cutoff point that determines what decision will be made concerning the null hypothesis
- α is between 0 and 1 (eg. 0.10, 0.05, 0.01)
- the probability with which the researcher is willing to reject H_0 when it is actually true

Develop the hypothesis

Expressions for the significance level, α

- alpha level (α) was fixed at 0.05
- H_0 will be rejected if $p < 0.05$
- statistical significance was set at the 5% level
- differences between groups were examined at the 0.05 significance level

Develop the hypothesis

Check your understanding

- state a “good” hypothesis

Test hypothesis

Data collection

- collect and summarize sample data
- researcher's decision to reject H_0 is made after observing sample data
- evidence to reject H_0 is based on sample data that is converted into a numeric value (ie. test statistic or p-value)

Test hypothesis

Test statistics

- quantify evidence to reject or not reject H_0
- have some distribution based on theory

Test hypothesis

Test statistics

- most test statistics, t_{obs}, have the following form:

$$t_{obs} = \frac{\text{sample statistic - population parameter under } H_0}{\text{standard error of sample statistic}}$$

- measure the difference between the value from the data and the hypothesized value, and scales that value by the variation in such sample statistics based on the data (i.e. the standard error)

Vancouver Coastal Health Research Institute
Healthier lives through discovery

Test hypothesis

p-value

- quantify evidence to reject or not reject H_0
- probability of observed data assuming the null hypothesis is true
- probability of observing a test statistic that is as extreme or more extreme than currently observed assuming the null hypothesis is true

UBC

Vancouver Coastal Health Research Institute
Healthier lives through discovery

Test hypothesis

Check your understanding

- name a test statistic

Make decision

critical value

- value of the statistic marking the boundary between the acceptance region and rejection region

significance level, α

- predetermined probability with which the researcher is willing to reject H_0 when it is actually true

Make decision

Decision criterion

- compare the test statistic with the critical value

  → test statistics greater (in absolute value) than the critical value indicate reject H_0

- compare the sample-based p-value with the significance level, α

  → p-values less than α indicate reject H_0

Make decision

Expressions to reject H_0

- H_0 was rejected
- statistically significant
- p-value is less than 0.05
- there is significant evidence to reject H_0

Make decision

Expressions to fail to reject H_0

- H_0 was not rejected
- not statistically significant
- p-value is greater than 0.05
- no significant evidence to reject H_0

Vancouver Coastal Health Research Institute
Healthier lives through discovery

Make decision

Interpret the decision

- translate decision into context of research
- does statistical significance imply clinical significance?
- does statistical non-significance imply clinical non-significance?

Make decision

Check your understanding

- how are decisions are made?

Example: scenario

Suppose that you would like to compare the effect of a newly developed drug (drug A) and a current drug (drug B) on systolic blood pressure (sbp).

How should we proceed?

Example: develop hypothesis

What is the hypothesis?

H_0: the mean sbp among patients on drug A is the same as the mean sbp among patients on drug B

ie. $\mu_A = \mu_B$ or $\mu_A - \mu_B = 0$

H_A: the mean sbp among patients on drug A is different from the mean sbp of patients on drug B

ie. $\mu_A \neq \mu_B$ or $\mu_A - \mu_B \neq 0$

Vancouver Coastal Health Research Institute
Healthier lives through discovery

Example: develop hypothesis

significance level, α

Q: How often would you be comfortable with rejecting H_O when there is no difference?

A: 1 out of 20 times

→ $\alpha = 0.05$

Example: test hypothesis

What does the data show?

- we took a random sample of 35 patients on drug A and 35 patients on drug B.
- we observed an average sbp of 107 mmHg and 125 mmHg among drug A and drug B patients, respectively.
- we observed an sd of 19 and 20 mmHg, respectively.

How will we test our hypothesis?

independent two-sample t-test
with two-sided alternative

UBC

Vancouver CoastalHealth Research Institute
Healthier lives through discovery

Example: test hypothesis

What is the test statistic?

$$t_{obs} = \frac{\bar{x}_B - \bar{x}_A}{s_p\sqrt{\frac{1}{n_B} + \frac{1}{n_A}}} = \frac{125 - 107}{19.5\sqrt{\frac{1}{35} + \frac{1}{35}}} \approx 3.9$$

$$s_p = \sqrt{\frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}} = \sqrt{\frac{(35-1)(19^2) + (35-1)(20^2)}{35+35-2}}$$

≈ 19.5 mmHg

Example: make decision

Decision criterion

Find critical value:

$t_{\alpha/2,\, n_1+n_2-2} = 2.0$

Compare test statistic with critical value:

$t_{obs} = 3.9 > 2.0 = t_{\alpha/2,\, n_1+n_2-2}$

Make decision:

→ reject H_0

Example: make decision

Decision criterion

Find p-value:

$p\text{-value} = 0.0001$

Compare p-value with α:

$p\text{-value} = 0.0001 < 0.05 = \alpha$

Make decision:

⇒ reject H_0

Example: make decision

Decision

There is evidence to support that there is a significant difference in mean sbp between patients on drug A and patients on drug B (p=0.0001).

Example: confidence interval

95% CI for difference between mean sbp:

$$(x_B - x_A) \pm t_{\alpha/2, n_A+n_B-2} s_p \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$$

$$\Rightarrow (125 - 107) \pm (2.0)(4.7)$$

$$\Rightarrow 18 \pm 9.3$$

$\Rightarrow (8.7, 27.3)$ mmHg

Vancouver Coastal Health Research Institute
Healthier lives through discovery

Example: confidence interval

Interpretation

The mean sbp for patients on drug B is on average 18 (95% CI 9-27) mmHg higher than the mean sbp for patients on drug A.

Relation to confidence intervals

Hypothesis testing and confidence intervals

- hypothesis testing and confidence intervals are related
- H_0 is not rejected at the α significance level if the 100(1- α)% CI captures the parameter value specified under H_0
- H_0 is rejected at the α significance level if the 100(1- α)% CI does not capture the parameter value specified under H_0

Vancouver Coastal Health Research Institute
Healthier lives through discovery

Fallacies

- statistical significance implies clinical significance
- small p-values imply large effects
- not rejecting H_0 implies that it is true
- p-value is the probability that H_0 is true

Examples from literature

Check your understanding

- find some examples of hypotheses
- find some examples of how decisions were made
- find some examples of decisions

UBC

Vancouver CoastalHealth Research Institute
Healthier lives through discovery

Errors in hypothesis testing

- type I error, α

  chance of mistakenly rejecting H_O

- type II error, β

  chance of mistakenly accepting H_O

Errors in hypothesis testing

Truth table

	State of Reality	
Decision	H_O true	H_O false
Reject H_O	type I error α	correct decision $1-\beta$
Accept H_O	correct decision $1-\alpha$	type II error β

UBC

Vancouver CoastalHealth Research Institute
Healthier lives through discovery

Follow-up

Thoughts

- how do we control the chance of error?
- how do we know if we were able to detect a clinically significant result?

→ evaluate power and sample size before beginning study

References

- Norman, GR and Streiner, DL. (2000). Biostatistics: The Bare Essentials 2E. Hamilton: BC Decker.
- Song, M et al. (2005). "A Randomized, Controlled Trial to Improve Advance Care Planning among Patients undergoing Cardiac Surgery." Medical Care. 43:1049-1053.
- van Bemmel, T et al. (2005). "In an observational study elderly patients had an increased risk of falling due to home hazards." Journal of Clinical Epidemiology. 58:63-67.

UBC

Vancouver Coastal Health Research Institute
Healthier lives through discovery