University of Saskatchewan

Biostatistics Review
T-test, ANOVA, Chi-square

Public Health and Preventative Medicine Residents
Daniel Fuller
(Slides from Dr. Pahwa)

Objectives

- Describe how ANOVA, Chi-square, T-test are used to explain differences in health and health behaviors.
- Explore a health issue effectively by appropriately differentiating when to select each of these
- Explore a health issue effectively by critically appraising the use of each of these

Introduction

- Choosing an appropriate statistical technique depends on

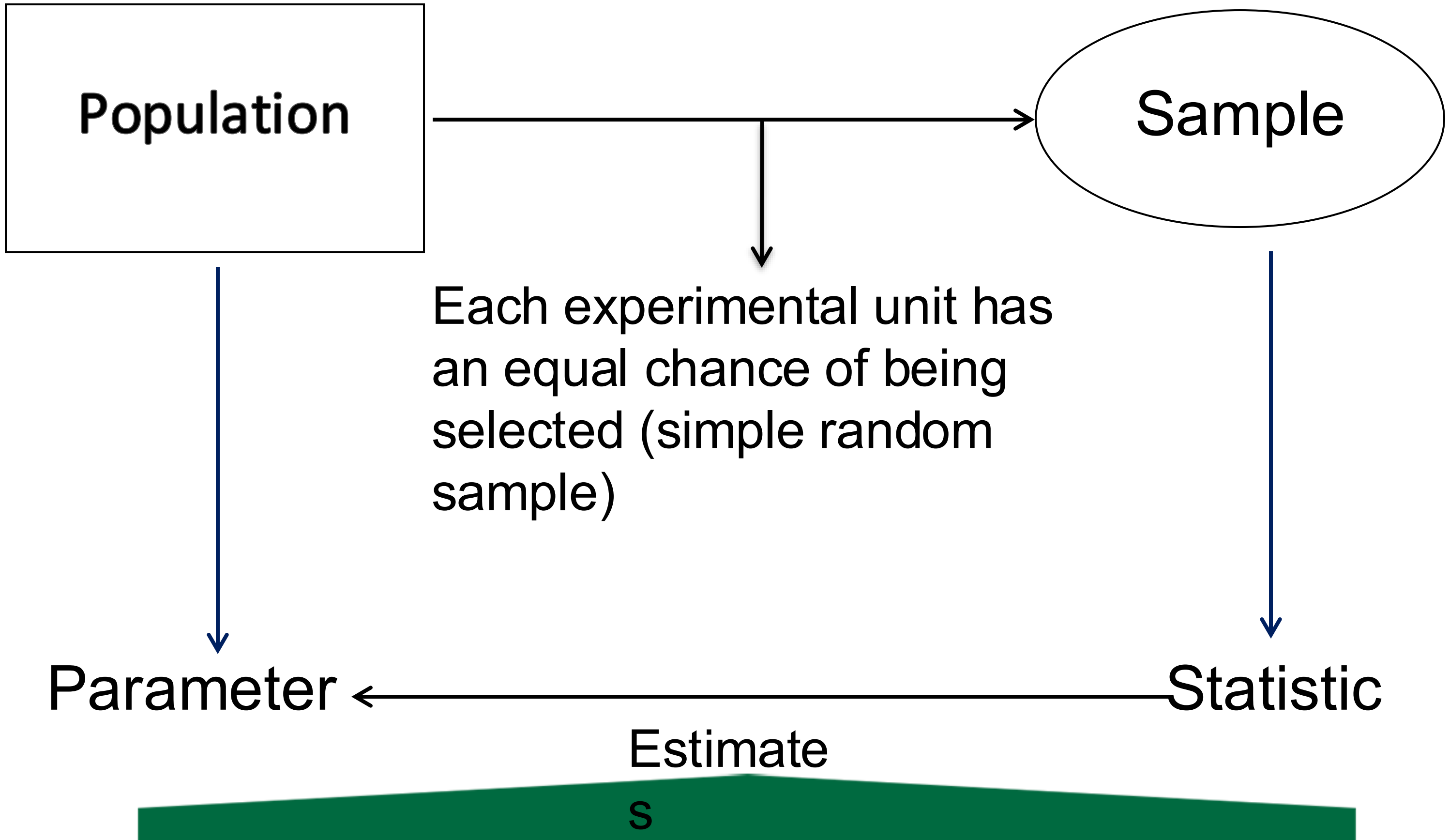
Research question(s)

Type(s) of variable(s) under investigation

UNIVERSITY OF SASKATCHEWAN

Population and Sample

The basic idea of statistics is simple: you want to extrapolate from the data you have collected to make general conclusions.



Populations and Samples

- Population Parameter (N)
  a) Characteristic of the population
     - Age = 45years
     - BP =115/75

- Sample Statistic (n)
  a) Estimates the population parameter
     - Age = 44years
     - BP =120/80

  b) May or may not be a good estimate

Populations and Samples

- Very difficult to measure everyone in an entire population
  a) **Census** = Measuring the entire populations
- Need to sample people from the population
  a) Two goals with samples
    - Big enough to represent the population
    - Not biased toward a certain group or variable
- **This is main reason we calculate confidence intervals, have *p* values, and use statistical tests. We want to understand uncertainty between our samples and the true population.**

Types of Data

- Data can be broadly categorized as **categorical** or **continuous**
  a) **Categorical data**
     - Limited number of unique (“discrete”) possible values
     - Differences between categories can’t be described with consistent measures
  b) **Continuous data**
     - Unlimited possible values on an unbroken scale
     - Measurable/quantifiable differences between values

Types of Data

- **Categorical**
  a) Limited number of possible values
  b) Difference between categories not quantifiable
- **Types of Categorial data**
  a) Nominal: Categories with *no ranking or order*
    - Men/Women, City A vs City B, blood types, marital status

  b) Ordinal: *Ranked/ordered* categories
    - A/B/C academic grades, cancer stages, self-rated health (excellent, very good, good…)

Types of Data

- **Continuous**
  a) Unbroken scale, Unlimited possible values
- **Types of continuous data**
  a) Interval
  - There *is* a consistent difference between values on a scale, but the ratio between values is not meaningful (not clear definition of “zero”)
    - Temperature in degrees C: 20℃ is not ‘twice as much heat’ as 10℃, and “0℃” does not mean “no heat”

  b) Ratio
  - Ratio of values has specific meaning; true zero point on scale
    - 20kg is twice as much mass as 10kg; “0kg” really does mean “no mass”

UNIVERSITY OF SASKATCHEWAN

T-test

UNIVERSITY OF SASKATCHEWAN

Introduction

- Compare two groups
  - Male vs Female (blood pressure)
  - Young vs old (mental health)
- Assess relationship/association among a set of variables.
  a) Is there an association between smoking and lung cancer?
  b) Is there an association between exposure to pesticides and non-Hodgkin's lymphoma?

Types of t-tests

- One sample t-test
  a) A **one-sample Student's *t*-test** is a location test of whether the mean of a population has a value specified in a null hypothesis.Independent samples t-test
- Two-sample *t*-tests
  a) A **two-sample** location test of the null hypothesis such that the means of two populations are equal.

Types of t-tests

- Paired t-test
  a) Paired samples *t*-tests typically consist of a sample of matched pairs of similar units, or one group of units that has been tested twice (a "repeated measures" *t*-test).
  b) Example: Participants are examined prior to a treatment, say for high blood pressure, and the same subjects are examined again after treatment with a blood-pressure-lowering medication.
- Independent samples t-test
  a) The independent samples *t*-test is used when two separate sets of independent and identically distributed samples are obtained, and one variable from each of the two populations is compared.

Purpose of t test

1. to estimate and test hypotheses about population variances

and

2. to estimate and test hypotheses about population means.

Example: Paired sample t-test

- **Paired Samples**
- 11 patients was administered a treatment on different occasions. Each patient first received the placebo. One month later they received a new drug. The average blood pressures when taking the placebo were 120/80. The average blood pressures when taking the treatment were 121/79.
- The t-test asks the following question
  a) Is there any evidence of a difference in mean systolic blood pressure during the use of these two drugs?
  b) Do the t-test with your eyes… any difference?

Example from: Colton T. Statistics in medicine. Boston: Little, Brown and Company; 1974. Chapter 4, Inference on means; p.99-150. As

Example: Independent samples

- A psychologist measured reaction time (in seconds) of 18 individuals who were not given any stimulus and 16 individuals who were given an alcoholic stimulus. The data are given below.
- Do the data indicate that the reaction time is increased significantly by alcoholic stimulus?

- Group (Nothing): mean = 1.96s
- Group (Alcohol): mean = 3.90s

Reatime	Group (Nothing)	Reatime	Group (Alcohol)
3.00	1	5.00	2
2.00	1	4.00	2
1.00	1	4.25	2
4.00	1	4.00	2
1.50	1	3.75	2
2.00	1	3.00	2
1.00	1	4.50	2
2.25	1	2.75	2
2.50	1	2.50	2
2.25	1	3.25	2
1.75	1	3.80	2
1.50	1	4.20	2
2.00	1	4.60	2
1.75	1	4.50	2
1.00	1	4.45	2
1.50	1	4.00	2
2.00	1		
2.25	1		

Fundamental Assumptions

Population assumptions:

1. All populations are normally distributed
2. All populations have the same variance

Sampling assumptions :

1. The samples are independent of each other
2. Each sample is obtained by using random sampling
3. The outcome will be most interpretable if the data are measured on an interval or ratio scale/continuous

Example: Independent samples

- Group (Nothing):
  - mean = 1.96s
  - variance = 0.55
- Group (Alcohol):
  - mean = 3.90s
  - variance = 0.50

- Are the variances different? Use your eyes?
- Can also do statistical test
  - Levene's Test
  - F-test for equality of variance

ANOVA

ANalysis Of VAriance (ANOVA)

- What if you wanted to test the difference between means but you have more than two groups? t-test only work for 2 groups.
- ANOVA is developed as a generalization of the t-test so we can test more than two groups (time can be a group here…).

- Outcome variable = continuous
- Main predictor variable = nominal variable or ordinal with more than two levels

ANalysis Of VAriance (ANOVA)

Example:

Predictor = Mental distress score among cancer patients – divided into four groups: (i) stage 1, (ii) stage II, (iii) stage III, and (iv) stage IV.

Outcome = Mental distress score - continuous

UNIVERSITY OF SASKATCHEWAN

Purpose of ANOVA

1. to estimate and test hypotheses about population variances

and

2. to estimate and test hypotheses about population means.

ANOVA

1. For comparing multiple populations means
2. Makes only one comparison (Tests one null hypothesis)
3. To determine if any of the population means differ from the rest

Tells you there is a difference between the groups somewhere but not specifically which groups are different.

Types of ANOVA

- One-way ANOVA
  a) One-way ANOVA is used to test for differences among two or more independent groups means
     - Different levels of effect of some medicine on groups of patients.
  b) Groups should be independent, and there is an order in the groups (such as mild, moderate and severe disease), oIndependent samples t-test

UNIVERSITY OF SASKATCHEWAN

Types of ANOVA

- Repeated measures ANOVA
  a) Repeated measures ANOVA is used when the same subjects are used for each factor (e.g., in a longitudinal study).
- Factorial ANOVA
  a) Used when there is more than one factor.

University of Saskatchewan

Example: One-way ANOVA

- A study was conducted to compare the effectiveness of three comprehensive therapeutic programs for severe anxiety. Three methods were employed:
  - Treatment 1:
    - Medication only.
    - Mean anxiety score = 80.
  - Treatment 2:
    - Cognitive Behavioural Therapy.
    - Mean anxiety score = 82.
  - Treatment 3:
    - Medication + Cognitive Behavioural Therapy.
    - Mean anxiety score = 100.

Example: One-way ANOVA

- Treatment 1:
  - Medication only.
  - Mean anxiety score = 80.
- Treatment 2:
  - Cognitive Behavioural Therapy.
  - Mean anxiety score = 82.
- Treatment 3:
  - Medication + Cognitive Behavioural Therapy.
  - Mean anxiety score = 100.

- ANOVA will tell you if there is a difference between these scores.
- ANOVA will NOT tell you which specifically, which scores are different from each other. Need posthoc tests for that…

Post hoc tests

- Multiple comparison/post hoc tests can be used to answer the question which population means are different?
- All post-hoc methods are essentially based on the t-test but include appropriate corrections for the fact that we are comparing more than one pair of means.
  - Tukey
  - Bonferroni
  - Scheffé
  - Bunch more… all old statistician dudes names…

https://en.wikipedia.org/wiki/Post_hoc_analysis

Fundamental Assumptions

Population assumptions:
1. All populations are normally distributed.
2. All populations have the same variance.

Sampling assumptions :
1. The samples are independent of each other.
2. Each sample is obtained by using independent random sampling.
3. The outcome will be most interpretable if the data are measured on an interval or ratio scale/continuous.

ANOVA is just linear regression

- ANOVA is a special case of linear regression which in turn is a special case of the general linear model.
- All consider the observations to be the sum of a model (fit) and a residual (error) to be minimized.
- Any modern analysis basically just do a linear regression instead of ANOVA.
  - Much better because you can include covariates and the linear model has been further generalized so you can relax many assumptions of ANOVA (e.g. equal variance, independence)

Chi-Squared Test (χ^2)

χ^2

When dependent and independent variables are categorical:

- How do you investigate if there is an association between two categorical variables?
    a) Chi-square test
- How do you measure the strength of this association?
    a) Estimate an odds ratio

- This is similar to ANOVA in that we only get “is there a difference” not ”where” or ”how strong” is that difference.

		Disease		
		Diseased	Not Diseased	
Exposure	Exposed	a	b	$a+b$
	Not exposed	c	d	$c+d$
		$a+c$	$b+d$	

Chi-squares test – to measure the association

Examples: **Exposure to**: smoking (personal and second hand); grain dust, pesticides
Disease: lung cancer, chronic cough; asthma etc.

Crosstabs

		Disease		
		Diseased	Not Diseased	
Exposure	Exposed	35	65	100
	Not exposed	15	85	100
		50	150	200

UNIVERSITY OF SASKATCHEWAN

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
expos * illwell	200	100.0%	0	.0%	200	100.0%

70% of diseased were exposed to the risk factor

expos * illwell Crosstabulation

			illwell		Total
			.00 diseased	1.00 no diseased	
expos	.00 exposed to a risk factor	Count	35	65	100
		% within illwell	70.0%	43.3%	50.0%
	1.00 not exposed to a risk factor	Count	15	85	100
		% within illwell	30.0%	56.7%	50.0%
Total		Count	50	150	200
		% within illwell	100.0%	100.0%	100.0%

UNIVERSITY OF SASKATCHEWAN

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	10.667[b]	1	.001		
Continuity Correction[a]	9.627	1	.002		
Likelihood Ratio	10.903	1	.001		
Fisher's Exact Test				.002	.001
Linear-by-Linear Association	10.613	1	.001		
N of Valid Cases	200				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 25.00.

Statistical Tests

Exposure measure (independent variable)	**Outcome measure (dependent variable)**	**Example**	**Test**
Categorical, two groups only	Continuous	Comparing blood pressure between men and women	T-test
Categorical, two or more groups	Continuous	Comparing blood pressure between four med school classes	Analysis of Variance (ANOVA)
Categorical	Categorical	Comparing likelihood of dementia (Y/N) by sex/gender	Chi-square test (χ_2)
Any (multiple)	Categorical	Comparing likelihood of dementia (Y/N) by sex/gender *and* age	Logistic regression
Any (multiple)	Continuous	Comparing blood pressure by sex/gender and age	Linear regression

Modern analysis

- Very rare to see chi-square tests in modern analyses. More common to see logistic regression which estimates both the strength of the association (odds ratio) and the statistical significance.

University of Saskatchewan

Examples from the literature

- Cohort profile: the Saskatchewan Rural Health Study—adult component | BMC Research Notes (springer.com)
- Racial discrimination and depression among on-reserve First Nations people in rural Saskatchewan | SpringerLink
- Gender Differences in the Association of Individual and Contextual Exposures with Lung Function in a Rural Canadian Population | SpringerLink
- Household Income and Psychological Distress: Exploring Women's Paid and Unpaid Work as Mediators - PMC (nih.gov)