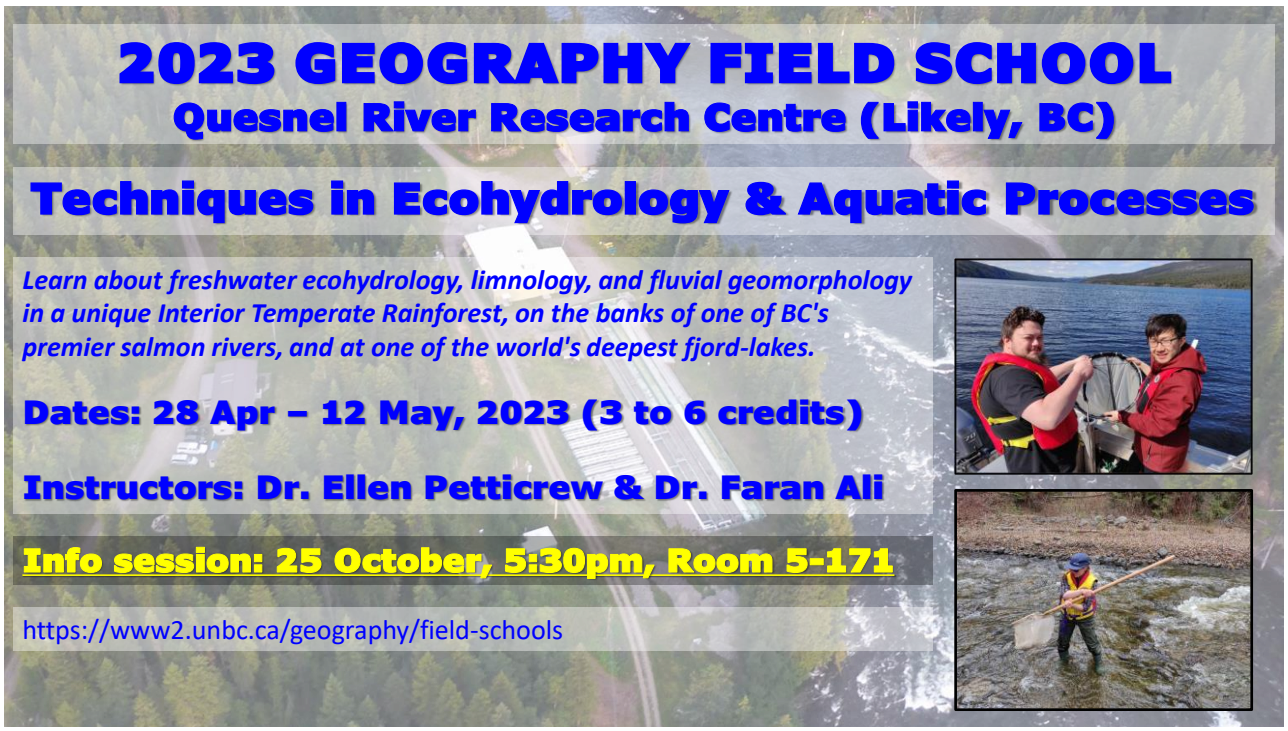
GEOG 204

LECTURE 10

1



2

Hypothesis Testing

- A hypothesis is an informed explanation, prediction, or supposition about something
  - Informed because there is evidence (but limited)
  - It has to be testable to so that one can conclude if its true or not
  - It has to be falsifiable (a possibility exists that it can be proven false)

- Bad hypothesis examples
  - Water is better than fire: not testable, criterion/variable is undefined
  - One can live to be 300 years old: not false falsifiable because no one gets to live that long

3

Hypothesis Testing

- A procedure that leads to a conclusive statement regarding a specified hypothesis

- The general goal is to make an ***inference*** about the magnitude of one or more population parameters based on sample data

4

Hypothesis Testing

- Two complimentary hypotheses

H_0 = Null Hypothesis	There is **no significant** difference between two parameters
H_A = Alternate Hypothesis	There is **a significant** difference between two parameters

H_0: *Sample Statistic = Population Parameter*
H_A: *Sample Statistic ≠ Population Parameter*

5

5

Hypothesis Testing

- The goal is to calculate probability that the null hypothesis is true
- If this probability is acceptably low
  - (meaning that it is likely not true)
  - then the null hypothesis is rejected in favour of the alternate hypothesis.
  - The sample results can be said to be significantly different

6

6

Hypothesis Testing

- Steps in hypothesis testing
  1. State the null and alternate hypotheses
  2. Select appropriate statistical test
  3. Select level of significance
  4. Delineate regions of rejection and nonrejection of null hypothesis
  5. Calculate test statistic
  6. Make decision regarding null and alternate hypotheses

7

7

State the null and alternate hypotheses

- Consider the formulation of a hypotheses concerning the mean of the population

$$H_0: \mu = \mu_H$$

$H_A: \mu \neq \mu_H$ (nondirectional)
$H_A: \mu < \mu_H$ (directional)
$H_A: \mu > \mu_H$ (directional)

8

8

Errors that can occur

- Type I Error
  - A decision is made **to reject** a null hypothesis as false when in fact it is true
- Type II Error
  - A decision is made ***to not reject*** the null hypothesis when it is actually false

9

9

Select the appropriate statistical test

$$Z = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu}{s/\sqrt{n}} \; (if \; n \geq 30)$$

$$t = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}} = \frac{\bar{X}-\mu}{s/\sqrt{n-1}} \; (if \; n < 30)$$

where

$Z, t = test\ statistic$

$\bar{X} = sample\ mean$

$\mu = population\ mean$

$\sigma_{\bar{X}} = standard\ error\ of\ the\ mean$

$\sigma = population\ standard\ deviation$

$n = sample\ size$

10

10

Select the level of significance

- Place a probability on the likelihood of a sampling error
  - Usually a low significance level ($\alpha = 0.05$ *or* $\alpha = 0.01$)
  - As such the conclusion is stated in terms of the level of significance
- If a null hypothesis is rejected at a 0.05 level of significance
  - There is a 5% chance that a Type I error has occurred
  - 5% likelihood that a null hypothesis has been improperly rejected

11

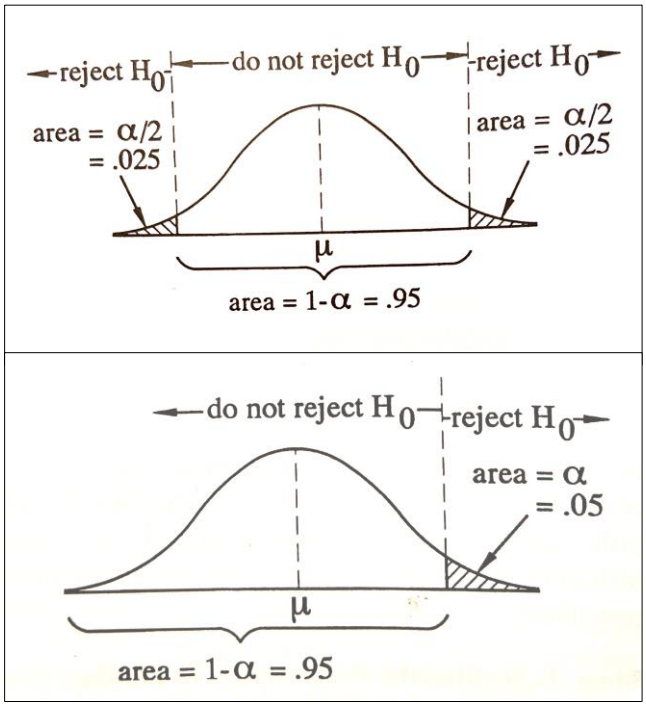
11

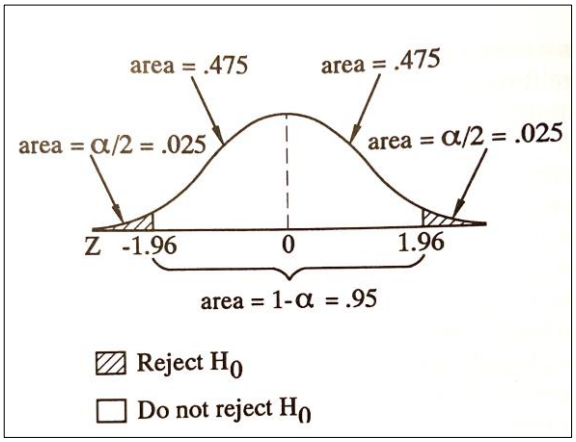
Delineate regions of rejection and nonrejection of null hypothesis

- The level of significance is used to set regions of **rejection** and **nonrejection**
- Determine the critical Z-values that delimite the boundaries separating the rejection and nonrejection regions

12

12





If
Z < -1.96 or if **Z > 1.96** reject H_0

Conversely if
-1.96 ≤ Z ≤1.96 do not reject H_0

13

13

Calculate the Test Statistic

- For our purposes here let's suppose the following contrived example
- A population has a mean = 65 and standard deviation = 25
  - Two samples are drawn from the population
    - Tsample has a size of 15 with a mean = 52.35 and standard deviation = 30.52
    - Zsample has a size of 75 with a mean = 57.87 and standard deviation = 24.93

$Z = (57.87 - 65)/(24.93/(75)^{0.5}) = -7.13/2.88 = -2.476$
$t = (52.35 - 65)/(30.52/(15-1)^{0.5}) = -12.65/8.156 = -1.55$

14

14

Decide on the hypotheses

- From the last slide, the Z < -1.96. Therefore, we reject the null hypothesis (no significant difference between the sample mean and the population mean)

15

Example

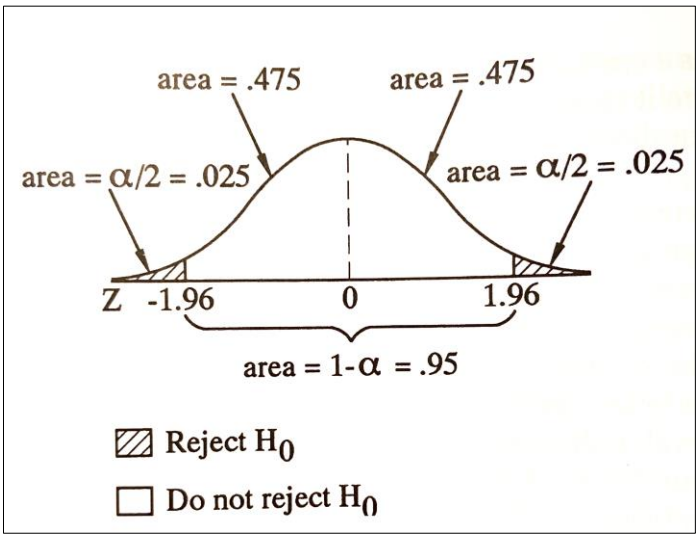
- A population has a mean = 65
  - Two samples are drawn from the population
    - Zsample has a size of 250 with a mean = 60.871 and standard deviation = 24.96
  - State the hypotheses
  - Decide on the level of significance
  - Calculate the Z statistic

16

The P-Value or Probability Value

- The area within the rejection region(s) represents the p-value.
- This p-value, is calculated in in four steps
  1. The test statistic for Z is calculated
  2. The probability or relative area under the normal curve is determined for that Z-value.
  3. The rejection area is determined by subtracting the probability (in step 2) from 0.5000.
  4. This area is doubled if a nondirectional (two-tailed) alternate hypothesis is used.

17



You can find the Z-Table
in most statistics books even online

18