STA 291
Lecture 27

- ***Final exam 6:00-8:00PM Thursday May 6***
- ***Room: will be in the Classroom Building***

- Makeup final exam: Friday May 7

- 10:00am – 12:00 noon

Last online Homework

- **Last Online homework assignment will be posted** this week

Example: two sample test

- ***Comparing Two Populations***
  - ***Two Independent Samples***

    (not paired)

Chap.13 Comparison of Two population means with Independent Samples

- Two ***Independent*** Samples (not paired)
  - Different subjects in the different samples
  - Two subpopulations (e.g., male/female)
  - The two samples constitute independent samples from two subpopulations, sample size can be different and often are different

Example

- Weight gain (grams) of female rats between 28 and 84 days after birth. 12 were fed with high protein diet, 7 were fed with low protein diet.

- High protein: 134, 146, 104, 119, 124, 161, 107, 83,113, 129, 97, 123.
- Low protein: 70, 118, 101, 85, 107, 132, 94.

- Two samples. un-equal sample size. Parameters: 2 mu's (the population mean values)

$$H_0 : m_1 = m_2$$

(two sided) $H_A : m_1 \neq m_2$

- Compute t_obs = 1.891436
- P-value = 0.0757 (use two sided formula)

- This calculation of t_obs can be done by calculator,
- But more often by a software.

- Besides, an extensive t-table is not always available on paper.

- Sometime we just report a P-value.

- So, in this example if we decided to use alpha=0.05 the conclusion would be “not reject Ho”, since 0.0757 is NOT less than alpha

- Sample size is small. Here 12 and 7.
- Have to use t-table, (the substitute of Z-table would result in large errors)
- Usually done by software. We are not required to work with software in sta291 exam, but we should be able to workout everything else given the computer output, or P-value.

Confidence Interval for the Difference of Two Means: Example

- In the 1982 General Social Survey, 350 subjects reported the time spent every day watching television. The sample mean was 2.7 hours, with standard deviation 2.1
- In the 1994 General Social Survey, 395 subjects reported a mean time spent watching television of 3.5 hours, with standard deviation 2.5
- *Is it plausible that the mean was the same in both years?*

- both mu's unknow (for year 1982 and 1994) since we never was sure about the year 1982. (no census was done)
- Two sided alternative. We did not see something like "was average TV time *increased*". The default one is to use two-sided alternative

$$H_A : m_1 \neq m_2$$

- TV programs are getting better, but other competing form (internet, computers etc) getting the people away from TV

Significance Test for the Difference of Two Means

- Let mu1 be the mean in 1982, and mu2 be the mean in 1994

$H_0 : m_1 = m_2$ which is equivalent to $H_0 : m_2 - m_1 = 0$,

$$t_{obs} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\frac{2.7-3.5}{\sqrt{\frac{2.1^2}{350}+\frac{2.5^2}{395}}}=-4.745$$

- P-value = 2P(Z > 4.745) = 0.000002085

or less than 2x0.000?= 0.000? by our Z table

- Highly significant!

- Strictly speaking I should look up the t-table….P-value = 0.00000250
- Did not change our conclusion.

- In general when sample size(s) > 100, normal table and t-table are very similar

- Will have one (or two) long question that is not multiple choice.
- This is where you can earn partial credit even if the final answer is wrong

- Get prepared by reviewing
  - **Formula sheets**
  - **Lecture notes**
  - **The first two exams**
  - **Online homework questions**
  - **Material from lab sessions**
  - **Textbook**
  - **Old exams**

- Testing hypothesis (we covered 4 cases)

Null and alternative hypothesis

P-value, significance

Type I and type II errors

- Computation of the test statistic:

either z or t (follow formula sheet)

And table

Multiple choice Q

- If a test turns out to be significant at alpha-level 0.01. (what exactly this mean for the p-value?)
- Will the same test also be significant at 0.05 level?

- P-value is ***NOT*** the probability that the H0 is true.

- A small p-value mean that we saw something happened that is ***hard to explain*** by H0 (a small probability event)

- A large p-value do not automatically means H0 is true. (2 possibilities: either H0 is true or there is too few data/info)

Correspondence Between Confidence Intervals and Tests

Only apply to 2-sided alternative hypothesis setup.

But works for either proportion “p” or mean “mu”. In fact this correspondence is valid in much wider context.

- Confidence intervals:

Interpretation, coverage probability, confidence level

Student t-confidence interval

Attendance Survey Question

- Please write down your name and section number
- Today's Questions

- DO you like to use software in sta291?

Exam II curve: conversion formula

- If your original score is 83 or above, then converted score is $= 90+(x-83)10/17$
- If your original score is 71 → 82, then converted score is $= 80+(x-71)9/11$
- If your original score is 59 → 70, then converted score is $= 70+(x-59)9/11$
- If your original score is 48 → 58, then converted score is $60+(x-48)9/10$
- If your original score is 1 → 47, then converted score is $x\ 59/47$