Chapter 11: Inference For Distributions

11.1 Inference for the mean of a population - Day 1

<u>OBJ</u>: You will find the p-value + conf. int. without σ.

<u>Last Chapter</u>		<u>This Chapter</u>
Pop $\geq 10n$, SRS, Normal	$\longrightarrow$	SRS, Normal
σ was Known	$\longrightarrow$	σ is unknown, $\therefore$ use s (Sample St. Dev)
$P\left(Z > \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right)$	$\longrightarrow$	$P\left(T > \frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)$
$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$	$\longrightarrow$	$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$

- $s/\sqrt{n}$ is called the std. error of the sample mean $\bar{x}$ (SEM)
- There is a different t-Dist. for each sample size!
- Does <u>NOT</u> have a NORMAL Distribution!

Find t^* for a 90% conf. int. w/ $n=20$ — Ans: $t^* = 1.729$

Find t^* for a 90% conf. int. w/ $n=45$ — Ans: $t^* = 1.684$ (use $n=40$)
- See Ex 11.2 Step 3 if needed

Facts About t-Distributions - Bottom of pg 618

Look @ Table C - shows as K inc., t critical values approach the normal values.

11.1-11.4 (Part a's)

11.8

11.9

<u>HW</u>: 11.7, 11.10, 11.11

Tomorrow: (pgs 628-635)

11.1 Inference for The Mean of a Population - Day 2

OBJ: You will find a p-value + a conf. int. of a matched pairs design.

Matched Pairs Design - Subjects are matched in pairs and each ~~treatment~~ is given to one subject in each pair. The experimenter can toss a coin to assign two treatments to the two subjects in each pair.

Matched Pairs t Procedures - Apply the one-sample t-procedures to the observed differences.

Look @ Computer output on pg. 631

Show t-test and t-interval on TI-84

11.13

HW: 11.12 Tomorrow (pg 635-640)

11.1 - Day 3

Robust Procedures - A conf. int. or sig. test is called Robust if the conf. int. or p-value does not Δ very much when the assumptions of the procedures are violated. (Look @ ℙ Before as well)

Example 11.6 AND ℙ Above - Did Yesterday!! (Similar Prob.)

Using the t-Procedures (And ℙ Above) - pg 636

- SRS, Pop. Dist ≈ NORMAL (SRS more Important) * Part I
- If $n < 15$: DATA ≈ Normal (If not normal or has outliers, - Don't use)
- If $n \geq 15$: Can use except in the presence of outliers or strong skewness.
- Large Samples ($n \geq 30$): Can use even for strong skewness.

11.20

HW: 11.21, 11.31

11.2 Comparing Two Means

<u>OBJ</u>: You will compare two populations or two treatments.

<u>Two-Sample Problems</u>

- The goal of inference is to compare the responses to two treatments or to compare the characteristics of two populations.
- We have a seperate sample from each treatment or each population.

Bottom of pg 648

Example 11.9

11-37

11.38

<u>Comparing Two Population Means</u> - pg 649

- Examine two-sample data graphically: stemPlots (Small n), Histograms/Box Plots (Larger Samples)
- A comparison of the mean responses in the two populations is common when both pop. dist. are symmetric and especially when they are at least ≈ Normal.
- <u>Conditions</u>: Two SRS's from two distinct pop.
  These samples are independent.
  Both pop. are normally distributed.

- Notation to Describe Two pop.

Population	Variable	Mean	StD. Dev
1	x_1	μ_1	σ_1
2	x_2	μ_2	σ_2

- We want to compare the two pop. means, either by giving a conf. int. for their difference $\mu_1 - \mu_2$ or by Testing the hypothesis of no Difference $H_0: \mu_1 = \mu_2$

- Notation to Describe Two Samples

Population	Sample Size	Sample Mean	Sample St. Dev
1	n_1	$\bar{x}_1$	s_1
2	n_2	$\bar{x}_2$	s_2

⟹

- To do inference about the difference $\mu_1-\mu_2$ between the means of the two populations, we start from the difference $\bar{x}_1-\bar{x}_2$ between the means of the two samples. (Just like we used $\bar{x}$ to estimate μ)

Example 11.10

<u>The Sampling Distribution $\bar{x}_1-\bar{x}_2$:</u>

- The mean of $\bar{x}_1-\bar{x}_2$ is $\mu_1-\mu_2$ (Just like the mean of $\bar{x}$ is μ - P.I.N. Histogram)
- The variance of the difference is the <u>sum</u> of the variances of $\bar{x}_1-\bar{x}_2$ which is: $\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ (Note: Variances add, std dev. do NOT - see pg 420)
- If the two pop. dist. are both normal, then the dist. of $\bar{x}_1-\bar{x}_2$ is also normal.

$\bar{x}-\bar{x}$ is NORMAL ∴ <u>Two-Sample Z statistic</u>

$$Z=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \Bigg\} \quad \boxed{Z=\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}}$$

$\mu_1-\mu_2$? (Wait for it....)

Since we don't know σ_1 or σ_2 all the time in problems, we estimate. The result is the <u>Standard Error</u> (or estimated Std. Dev)

$$SE=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

Don't know σ? ∴ <u>Two Sample t-statistic.</u>

$$t=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

This statistic does <u>not</u> have a t distribution. Why? We replaced <u>two</u> Std. Dev's. A t dist. occurs when you only replace <u>one</u> Std. Dev.!

<u>BUT</u>, Use it Anyway (It is all we have): Choose an option:
** Use critical values from t dist. w/ degrees of freedom:
1) From <u>all</u> data → very accurate (df = decimal) - Formula on pg. 659.
2) From the <u>smaller</u> of n_1-1 and n_2-1 ⇒ Always conservative

TI-84 Does option 1 if by hand do option 2.
Always List df = $\boxed{?}$ in your problems.

⇒)

Confidence Interval For $\mu_1 - \mu_2$:

* $\boxed{(\bar{X}_1 - \bar{X}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ Has conf. Level at least C w/ $t(K)$ Dist. w/ K the smaller of $n_1 - 1$ and $n_2 - 1$.

To Test the Hypothesis $H_0: \mu_1 = \mu_2$ compute the two-sample t-statistic:

* $\boxed{t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}}$

- where did $\mu_1 - \mu_2$ go? Well, if $\mu_1 = \mu_2$ then $\mu_1 - \mu_2 = 0$!
- Use p-values or critical values for the $t(K)$ Distribution

* These two-sample t procedures always err on the safe side, Reporting Higher p-values and Lower confidence than are actually true.

* As the sample sizes inc., Prob. values based on t w/ df equal to the smaller of $n_1 - 1$ and $n_2 - 1$ become more accurate.

Example 11.11

Example 11.12 + P After

Day 1

11.40

HW: 11.39, 11.41, 11.42c

Day 2

Next Day: class/HW: 11.50, 11.64, 11.67, 11.72