1 Inference for Quantitative Data: Means

1.1 Constructing a One Sample t-Interval

Review:

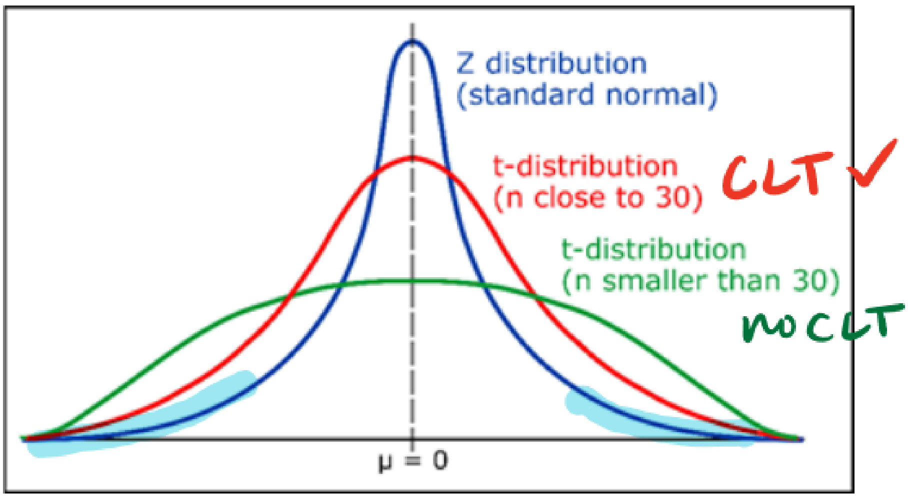
- Shape
  - Normal population indicates a normal sampling distribution
  - $n < 30$ then data needs to be stated as approximately normal.
  - $n \geq 30$ satisfies Central Limit Theorem which guarantees approximately normal
- Center
  - As long as you are taking a random sample or using random assignment, $\mu_{\overline{x}} = \mu$
- When sampling, as long as 10% condition is met, $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$

As long as the above conditions are met, the sampling distribution of $\overline{x} : \overline{x}\ N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

When we do not know the population standard deviation (which we usually don't), we must estimate it from our sample using the sample standard deviation, s. However, when we do, the test statistic (z) that we previously used changes. The new test statistic is called the t-statistic and has a new distribution associated with it.

The new t-distribution is not exactly like the standard normal curve, but it is very close:

- It is still centered at 0.
- It is bell shaped.
- Its sprea dis slightly greater than the normal distribution.
- It has more area in the tails.



The special thing about t-distributions is the fact that it is a family of distributions. There is a unique density curve for each sample size (and the dependence on sample size is taken care of by degrees of freedom: $n-1$).

The tails on a t-distribution are "fatter" than a standard normal distribution. This is true because the smaller our sample size, the more variation we have, hence the fatter tails. The larger our sample size gets, the closer the t-distribution will move towards the standard normal distribution.

Constructing a Confidence Interval:

- Define the Parameter
  - $\mu =$ true mean of population parameter in context
- Check the Assumptions and Conditions
  - Random: The sample should be a random sample of the population or random assignment in an experiment
  - Independence (10% Condition): The sample size, n, must be no larger than 10% of the population.
  - Normality: There are multiple ways to verify this condition.
    - * Stated in Problem: It may be stated in the problem that the sampling distribution is approximately Normal.
    - * Central Limit Theorem: When n is large, the sampling distribution of the sample means is approximately normal.
    - * Visual Representation: You are a given a graphical representation (histogram or boxplot) that depicts a shape that is approximately normal. You may also be given data that you have to graph yourself (include a sketch). You are looking for no strong skew or outliers.

    Note: If you get a sample size less than 30, only use T-procedures if there are no outliers and no strong skew.
  - Name the Inference Method: One Sample t-Interval
  - Calculate the Interval: point estimate $\pm$ margin of error

$$\overline{x} \pm t^*_{n-1}\left(\frac{s}{\sqrt{n}}\right)$$

    To calculate t^*_{n-1}:

    2nd-VARS-4:invT()

    - * Enter in the desired area.
      - · 90% Confidence Interval: 0.95
      - · 95% Confidence Interval: 0.975
      - · 99% Confidence Interval: 0.995
    - * Enter the degrees of freedom
      - · Calculation: $n - 1$ where n is the sample size

Write your conclusion in Context:

- We are ______% confident that the interval from ______to ______captures the true mean of population parameter in context.

Example

The Tribal Urban District Assessment (TUDA) is a government sponsored study of student achievement in large urban school districts. TUDA gives a reading test scored from 0 to 500. A score of 243 is "basic" reading level and a score of 281 is "proficient." Scores for a random sample of 1,470 eighth graders in a district has a mean score of 240 with a standard deviation of 42.17.

(a) Construct and interpret a 99% confidence interval for the mean reading test score for all of this district's eight graders.

$\mu =$ true mean reading test score of all the district eight graders.

- Random: Random sample of 1470 district eighth graders.
- Independence: $n = 1470 \leq 0.10$(all of this district's eighth graders)
- Normal: $n = 1470 \geq 30$. CLT applies, so sampling dist. is approx. normal.

One Sample t-Interval

$\overline{x} \pm t^* \left(\frac{s}{\sqrt{n}}\right)$, with $t^* = 2.5792$ gives interval $(237.1632, 242.8368)$.

We are 99% confident the interval from 237.1632 to 242.8368 points captures the true mean reading test score of all the eight graders in this district.

Using a calculator we can do the following:

Calculator Steps	
STAT – Tests – 8: TInterval…	
Inpt: Data (Actual Data)	Stats (Summary Statistics)
List: L_1 (Enter Data) Freq: 1 (Don't Change) C-Level: Confidence Level	$\overline{x}$: sample mean s_x : sample SD n : sample size C-Level: Confidence Level

(b) Based on your interval, is there convincing evidence that the mean reading test score for all the eighth graders in this district is less than the basic reading level? Justify your answer using your confidence interval.

Since our CI is entirely below 243 points there is convincing evidence that the mean reading test score is below basic reading level for all the eighth graders in this district.

Example

A teacher's car records the fuel efficiency (mpg) and resets every time she fills up her gas tank. She randomly selected 20 samples of fuel efficiency from her car's computer.

26.8	24.7	26.6	27.2	28.4	27.2	27.0	26.4	24.6	26.8
26.2	26.0	24.2	25.8	25.9	26.8	26.6	27.3	25.4	24.9

Given $\mu =$ true mean fuel efficiency for this teacher's car, verify the conditions for inference are met.

- Random: Randomly selected 20 fuel efficiency samples.
- Independence: $n = 20 \leq 0.10$(all car fill ups)
- Normal:

No strong skew or outliers so sampling distribution is approximately normal.

To graph a boxplot with a calculator:

- STAT-Edit-1:Edit. Enter Data into L1
- 2nd - Y= - Turn Plot1 On
- Change Type to Boxplot w/ Outliers
- ZOOM-9:ZoomStat
- TRACE - Identify the five number summary
- Sketch on your paper - looking for no strong skew or outliers

Determing Sample Size:

$$ME = t^*_{n-1}\left(\frac{s}{\sqrt{n}}\right)$$

Example

The health teachers at a high school want to estimate the body mass index (BMI) of students at their school. The BMI of US high school students follows a normal distribution with a standard deviation of 8.1%. How large of a sample is needed to estimate the mean BMI of this high school's students within 3% with 98% confidence?

Use invT to get $t^* = 2.4258$.

Plugging this into the above formula and solving for n gives $n = 43$ people.

1.2 Constructing a One Sample t-Test

- Define the parameter: $\mu =$ true mean of population parameter in context
- State the Hypotheses
  - Null Hypothesis: $H_0 : \mu = \mu_0$
  - Alternative Hypothesis: $H_A : \mu < \mu_0$, $H_A : \mu > \mu_0$, $H_A : \mu \neq \mu_0$
- Check the Assumptions and Conditions
  - Randomness: The sample should be a random sample of the population or random assignment in an experiment.

- 10% Condition: The sample size, n, must be no larger than 10% of the population.
- Approx. Normal Condition: There are multiple ways to verify this condition. (These ways are the same as the one sample t-interval)

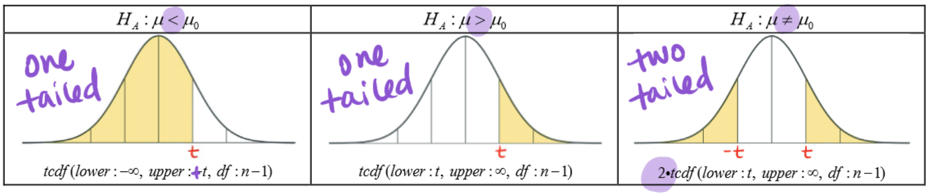
- Name the Inference Method: One Sample t-Test
- Calculate the Test Statistic

The test statistic is test statistic $= \frac{\text{statistic-parameter}}{\text{standard deviation of statistic}}$

One Sample t-Test: $t = \frac{\overline{x}-\mu_0}{\left(\frac{s}{\sqrt{n}}\right)}$.

Where $\overline{x}$ is sample mean, μ_0 is null mean, s is sample standard deviation, and n is sample size.

- Obtain the p-Value



- Make a Decision
  - If the p-value is less than α, we reject the null hypothesis.
  - If the p-value is greater than α, we fail to reject the null hypothesis.

State the conclusion in context:

- Since our p-value of ______is (less/greater) than ______, we (reject/fail to reject) the null hypothesis. There (is/is not) convincing evidence that alternative hypothesis.

Example

At the Hawaii Pineapple Company, managers are interested in the sizes of the pineapples grown in the company's fields. Last year, the mean weight of the pineapples harvested from one large field was 32 ounces. A new irrigation system was installed in this field after the growing season. Managers wonder whether this change will affect the mean weight of future pineapples grown in the field. To find out, they select and weigh a random sample of 50 pineapples from this year's crop. Their sample has a mean of 31.935 ounces and a standard deviation of 2.394 ounces. Does this data give convincing evidence that the mean weight of pineapples produced in the field has changed this year?

$\mu =$ true mean weight (ounces) of pineapples produced in this field

$H_0 : \mu = 32$, $H_A : \mu \neq 32$.

- Random: Random sample of 50 pineapples from this field
- Independence: $n = 50 \leq 0.10$(all pineapples in this field)
- $n = 50 \geq 30$, CLT applies to sampling dist. is approx. Normal

One Sample t-Test

$t = -0.1920$ (from the formulas above). $P(t < -1.92) = 0.4243$. Since this is two tailed, $p = 0.8486$.

Since the p-value of 0.8486 is greater than $\alpha = 0.05$, we fail to reject the null. There is not convincing evidence that the mean weight of pineapples produced in the field has changed this year.

Example

A study was conducted on whether time perception, an indication of a person's ability to concentrate, is impaired during caffeine withdrawl. Twenty randomly selected coffee drinkers abstained from caffeine for 24 hours. They were asked to estimate how much time had passed during a 45-second time period. The data is displayed below.

70	66	73	74	60	56	40	53	68	58
57	51	71	48	57	56	71	65	68	54

Is there convincing evidence at the $\alpha = 0.05$ significance level that caffeine abstinence had a negative impact on time perception (causing elapsed time to be overestimated?)

$\mu =$ true average estimation of time elapsed (seconds)

$H_0 : \mu = 45$, $H_A : \mu > 45$

- Random: 20 randomly selected coffee drinkers
- Independence: $n = 20 \leq 0.10$(all coffee drinkers)
- Normal:

No strong skew or outliers, sampling dist. is approx. normal.

One Sample t-Test

Calculator Steps	
STAT – Tests – 2: T-Test…	
Inpt: Data (Actual Data)	Stats (Summary Statistics)
μ_0 : null mean List: L_1 (enter data) Freq: 1 (don't change) μ : alternative hypothesis	μ_0 : null mean $\bar{x}$: sample mean s_x : sample SD n : sample size μ : alternative hypothesis

From this we get $t = 7.5888$ and $p = 0.0000002$.

Since the p-value of 0.0000002 is less than $\alpha = 0.05$, we reject the null. There is convincing evidence that caffeine abstinence had a negative impact on time perception.

1.3 Inference for Paired Data

Comparative studies are more convincing than single-sample investigations. For that reason, one-sample inference is less common that comparative inference. Study designs that involve making two observations on the same individual or one observation on each of two similar individuals, result in paired data.

When paired data result from measuring the same quantitative variable twice, we can make comparisons by analyzing the differences in each pair. If the conditions for inference are met, we can use one-sample t-procedures to perform inference about the mean difference: μ_D. These methods are called matched pairs procedures.

Example

A researcher studied a random sample of identical twins who had been separated and adopted at birth. In each case, one twin (Twin A) was adopted by a high-income family and the other (Twin B) by a low-income family. Both twins were given an IQ test as adults. Here are their scores.

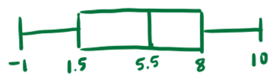
Pair	1	2	3	4	5	6	7	8	9	10	11	12
Twin A's IQ (High Income)	128	104	108	100	116	105	100	100	103	124	114	102
Twin B's IQ (Low Income)	120	99	99	94	111	97	99	94	104	114	113	100
Difference (High – Low)	8	5	9	6	5	8	1	6	-1	10	1	2

Construct and interpret a 95% confidence interval for the true mean difference in IQ scores among twins raised in high-income and low-income households.

$\mu_D =$ true mean difference in IQ scores between twins raised in high-income and low-income households.

- Random sample of 12 sets of identical twins
- $n = 12 \leq 0.10$(all sets of identical twins)

Normal:



One Sample t-Interval for Matched Pairs.

You can calculate using $\overline{x}_D \pm t^*_{n-1}\left(\frac{s_D}{\sqrt{n}}\right)$ or STAT-Tests-8:TInterval:

The interval is $(2.7495, 7.2505)$.

We are 95% confident that the interval from 2.7495 to 7.2505 captures the true difference (High-Low) between the IQ scores of identical twins raised on high income and low income households.

Example

Researchers designed an experiment to study the effects of caffeine withdrawl. They recruited 11 volunteers who were diagnosed as being caffeine dependent to serve as subjects. Each subject was barred from coffee, sodas, and other subtances with caffeine during the duration of the experiment. During one two-day period, subjects took capsules containing their normal caffeine intake. During another two-days period, they took placebo capsules. The order in which the subjects took caffeine and the placebo is randomized. At the end of each two-day period, a test for depression was given to all 11 subjects. Researchers wanted to know whether being deprived of caffeine would lead to an increase in depression. The table displays data on the subjects' depression test scores. Higher scores show more symptoms of depression.

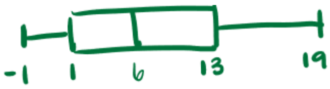
Subject	1	2	3	4	5	6	7	8	9	10	11
Depression Score (Placebo)	16	23	5	7	14	24	6	3	15	12	0
Depression Score (Caffeine)	5	5	4	3	8	5	0	0	2	11	1
Difference (Placebo - Caffeine)	11	18	1	4	6	19	6	3	13	1	-1

Does the data provide convincing evidence at the $\alpha = 0.01$ significance level that caffeine withdrawl increases depression score, on average, for subjects like the ones in this experiment?

μ_D: true mean difference in depression test scores betewen normal caffeine intake and placebo capsules.

$H_0 : \mu_D = 0$, $H_A : \mu_D > 0$

- Order of treatments is randomized
- Assume volunteers are independent of one another



No strong skew or outliers so sampling dist. is approx. normal.

One Sample t-Test for matched pairs.

You can calculate using $t = \frac{\overline{x}_D - 0}{\left(\frac{s_D}{\sqrt{n}}\right)}$ and 2nd-VARS-6:tcdf or STAT-Tests-2:T-Test

The t value is 3.5304 and $p = 0.0027$

Since the p-value of 0.0027 is less than $\alpha = 0.01$, we reject the null. There is convincing evidence that caffeine withdrawl increases depression score, on average, for subjects like those in this experiment.

1.4 Inference for Comparing Two Sample Means

Constructing a Two Sample T-Interval

- Define the Parameter:
  - μ_1 = true mean of population parameter in context for Sample 1
  - μ_2 = true mean of population parameter in context for Sample 2
- Check the Assumptions and Conditions: This is the same as for a one sample t-interval, just they have to apply for both populations.
- Name the Inference Method: Two Sample t-Interval for $\mu_1 - \mu_2$
- Calculate the Interval

  $$(\overline{x}_1 - \overline{x}_2) \pm t^*_{n-1} \left(\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}} \right)$$

  Calculating the t^*_{n-1} value is similar to earlier, the degrees of freedom is $n - 1$ for the smaller sample size (also known as the conservative df)
- Write your conclusion in context.

We are ______% confident the interval from ______to ______units captures the true mean difference Pop 1-Pop 2 between Context of Question OR

We are ______% confident that the true mean of Pop 1 in Context is between ______and ______units (higher/lower) than Pop 2

Example

College financial aid offices expect students to use summer earnings to help pay for college. But how large are these earnings? The University of Texas studied this question by asking random samples of 675 male and 621 female students with summer jobs how much they earned. Their data is summarized below.

Group	n	$\bar{x}$	$s_{\bar{x}}$
Males	675	$1884.52	$1368.37
Females	621	$1360.39	$1037.46

Construct and interpret a 90% confidence interval for the true mean difference between summer earnings of male or female students at the University of Texas.

μ_M = mean summer earnings of UT males, μ_F = mean summer earnings of UT females.

- Random sample of 675 male and 621 female UT students.
- $n_m = 675 \leq 0.10$(all male UT students), $n_F = 621 \leq 0.10$(all female UT students)
- $n_m = 675 > 30$, $n_F = 621 \geq 30$, CLT applies, sampling dist. is approx. normal

Two Sample t-Interval for $\mu_M - \mu_F$

Using the formula above, we find that $t^* = 1.6473$, and the interval to be $(413.54, 634.72)$.

We are 90% confident the interval from 413.62 to 634.64 dollars captures the true mean difference in summer earnings of male and female students at UT.

To use a calculator:

Calculator Steps	
STAT – Tests – 0: 2-SampTInt...	
Inpt: Data (Actual Data)	Stats (Summary Statistics)
List1: L1 (Sample 1 Data)	$\bar{x}1$: Sample 1 Mean
List2: L2 (Sample 2 Data)	$Sx1$: Sample 1 SD
Freq1: 1 (Don't Change)	$n1$: Sample 1 Size
Freq2: 1 (Don't Change)	$\bar{x}2$: Sample 2 Mean
C-Level: Confidence Level	$Sx2$: Sample 2 SD
Pooled: No (Don't Change)	$n2$: Sample 2 Size
	C-Level: Confidence Level
	Pooled: No (Don't Change)

Constructing a two sample t-Test

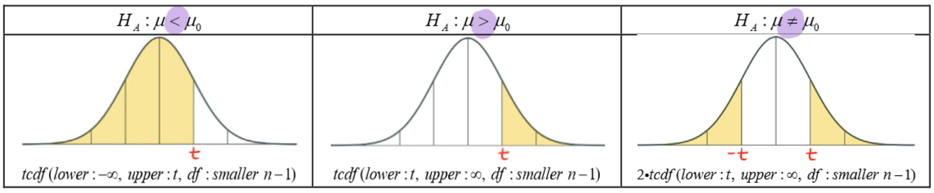
- Define the Parameter - Same as the two sample t-Interval
- State the Hypotheses:
  - Null Hypothesis: $H_0 : \mu_1 = \mu_2$
  - Alternative Hypothesis: $H_A : \mu_1 < \mu_2$, $H_A : \mu_1 > \mu_2$, $H_A : \mu_1 \neq \mu_2$
- Check the Assumptions and Conditions: Same as two sample t-Interval

- Calculate the Test Statistic

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - 0}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}}$$

  where $\overline{x}$ is sample mean, s is sample standard deviation, and n is sample size.

- Obtain the P-Value



- Make a Decision: You know how to do this by now
- State your conclusion in Context

  Since our p-value of ______is (less/greater) than ______, we (reject/fail to rejet) the null hypothesis. There (is/is not) convincing evidence that alternative hypothesis.

Example

Many people take gingko supplements advertised to improve memory. Are these over-the-counter supplements effective? In a study, elderly adults were random assigned to the treatment group or control group. The 104 participants who were assigned to the treatment group took 40 mg of gingko 3 times a day for 6 weeks. The 115 participants assigned to the control group took a placebo pill 3 times a day for 6 weeks. At the end of the 6 weeks, a memory test was administered. Higher scores indicate better memory function. Summary values are given in the following table.

Treatment	n	$\overline{x}$	$s_{\overline{x}}$
Gingko	104	5.7	0.6
Placebo	115	5.5	0.5

Based on these results, is there significant evidence that taking 40 mg of gingko 3 times a day is effective in increasing performance on a memory test?

μ_G = mean memory score for Gingko treatment group.

μ_P = mean memory score for placebo treatmnet group.

$H_0 : \mu_G = \mu_p$, $H_A : \mu_G > \mu_p$

- Randomly assigned elders to treatment or control group
- Assume independence among elderly memory test scores
- $n_G = 104 \geq 30$, $n_p = 115 \geq 30$. CLT applies, sampling dist. is approx. normal.

Two Sample t-Test

Using the formula above to calculate t, we get $t = 2.6642$, and $p = 0.0045$.

Since the p-value of 0.0045 is less than $\alpha = 0.05$, we reject the null. There is convincing evidence that taking Gingko is effective in increasing memory score.

This is how to do it in your calculator:

Calculator Steps	
STAT – Tests – 4: 2-SampTTest...	
Inpt: Data (Actual Data)	Stats (Summary Statistics)
List1: L1 (Sample 1 Data)	$\overline{x}1$: Sample 1 Mean
List2: L2 (Sample 2 Data)	$Sx1$: Sample 1 SD
Freq1: 1 (Don't Change)	$n1$: Sample 1 Size
Freq2: 1 (Don't Change)	$\overline{x}2$: Sample 2 Mean
$\mu1$: Alternative Hypothesis	$Sx2$: Sample 2 SD
Pooled: No (Don't Change)	$n2$: Sample 2 Size
	$\mu1$: Alternative Hypothesis
	Pooled: No (Don't Change)