Chapter 7
Section 7.1: Inference for the Mean of a Population

Now…let's look at a similar situation
- Take an SRS of size n
- Normal Population : $N(\mu, \sigma)$.
- **Both** μ *and* σ are unknown parameters.

Unlike what we used in Chapter 6, when we knew the population standard deviation σ

- Confidence interval for the population mean μ: $\bar{x} \pm z * \dfrac{\sigma_X}{\sqrt{n}}$

- Hypothesis test statistic for the population mean μ: $z_0 = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

- Used the standard normal distribution.

In this instance, we have to use an estimate of σ, the sample standard deviation *s.*

Using *s* instead of σ means that we are no longer able to use the standard normal distribution. Instead, we will have to use the **student's *t* distribution**.

The student's *t* distribution is completely determined by the number of degrees of freedom. When looking at the distribution of $\overline{X}$, we will use the *t* distribution with n-1 degrees of freedom, the *t(n-1)* distribution.

Using the t distribution

- Suppose that an SRS of size *n* is drawn from a $N(\mu, \sigma)$ population.

- There is a different *t* distribution for each sample size, so *t(k)* stands for the *t* distribution with *k* degrees of freedom.

- Degrees of freedom = *k = n – 1* = sample size – 1

- As k increases, the *t* distribution looks more like the normal distribution (because as *n* increases, $s \rightarrow \sigma$). *t(k)* distributions are symmetric about 0 and are bell shaped, they are just a bit wider than the normal distribution.

- The *t* table shows upper tails only, so
    - if t^* is negative, *P(t < t*) = P(t > |t*|)*.
    - if you have a 2-sided test, multiply the *P(t > |t*|)* by 2 to get the area in both tails.
    - The normal table showed lower tails only, so the *t*-table is backwards.

Table entry for p and C is the critical value t^* with probability p lying to its right and probability C lying between $-t^*$ and t^*.
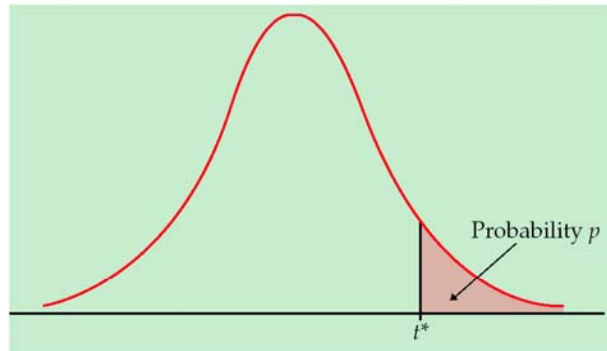
Probability p

TABLE D t distribution critical values

df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
					Upper tail probability p							
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
					Confidence level C							

2

The One-Sample *t* Confidence Interval:

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$ where *t** is the value for the *t*(*n*-1) density curve with area C between –*t** and *t**.

Finding t* on the table:

Start at the bottom line to get the right column for your confidence level, and then work up to the correct row for your degrees of freedom.

What happens if your degrees of freedom isn't on the table, for example df = 79? Always round DOWN to the next lowest degrees of freedom to be conservative.

Example 1

a) Find *t** for an 80% confidence interval if the sample size is 20.

b) Find *t** for an 98% confidence interval if the sample size is 35.

c) Find a 95% confidence interval for the population mean if the sample mean is 42, the sample standard deviation is 1.9, and the sample size is 50.

The One-Sample *t* test:

- State the Null and Alternative hypothesis.

- Find the test statistic:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- Calculate the *p*-value. In terms of a random variable *T* having the *t*(*n*-1) distribution, the *P*-value for a test of H_0 against

$$H_a : \mu > \mu_0 \text{ is } P(T \geq t)$$

$$H_a : \mu < \mu_0 \text{ is } P(T \leq t)$$

$$H_a : \mu \neq \mu_0 \text{ is } 2 P(T \geq |t|)$$

These *P*-values are exact if the population distribution is normal and are approximately correct for large *n* in other cases.

- Compare the *P*-value to the α level.

    If *P*-value $\leq$ α, then reject H_0 (significant results)
    If *P*-value > α, then fail to reject H_0 (non-significant results)

- State conclusions in terms of the problem. "Reject / Do not reject the null hypothesis. There is / is not enough evidence that the population average _____ is _____."

Robustness of the *t* Procedures

A confidence interval or statistical test is called robust if the confidence level or *P*-value does not change very much when the assumptions of the procedure are violated. The *t* procedures are robust against non-normality of the population when there are no outliers, especially when the distribution is roughly symmetric and unimodal.

When is it appropriate to use the t procedures?

- Unless a small sample is used, the assumption that the data comes from a SRS is more important than the assumption that the population distribution is normal.

- n<15: Use *t* procedures only if the data are close to normal with no outliers.

- n>15: The *t* procedure can be used except in the presence of outliers or strong skewness.

- n is large (n ≥ 40): The *t* procedure can be used even for clearly skewed distributions.

You can check to see if data is normally distributed using a histogram or normal quantile plot.

Example 2

a. An agricultural expert performs a study to measure yield on a tomato field. Studying 10 plots of land, she finds the mean yield is 34 bushels with a sample standard deviation of 12.75. Find a 95% confidence interval for the unknown mean yield of tomatoes.

b. Conduct a hypothesis test with $\alpha = 0.05$ to determine if the mean yield of tomatoes is less than 42 bushels. State your conclusion in terms of the story.

c. Draw a picture of the t curve with the number and symbol for the mean you used in your null hypothesis (μ_0), the sample mean ($\bar{x}$), the standard error ($\hat{\sigma}_{\bar{x}} = \frac{s}{\sqrt{n}}$), $t = 0$, and the test statistic (t_0). Shade the appropriate part of the curve which shows the P-value.
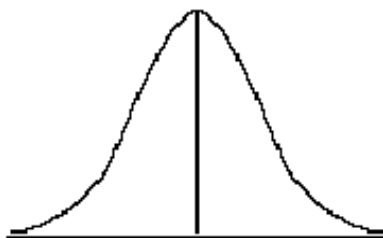
Example 3 (Exercise 7.37)

How accurate are radon detectors of a type sold to homeowners? To answer this question, university researchers placed 12 detectors in a chamber that exposed them to 105 picocuries per liter of radon. The detector readings were as follows:

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	119.3	104.8	101.7	96.6

a) Is there convincing evidence that the mean reading of all detectors of this type differs from the true value of 105? Use $\alpha = 0.10$ for the test. Carry out a test in detail and write a brief conclusion. (SPSS tells us the mean and standard deviation of the sample data are 104.13 and 9.40, respectively.)
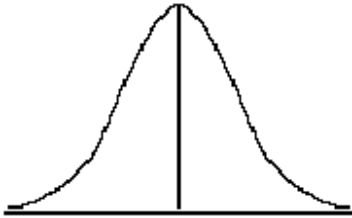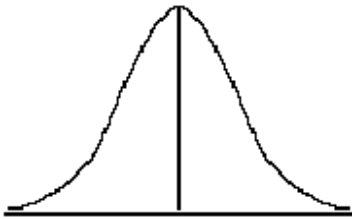
b) Find a 90% confidence interval for the population mean.

Now re-do the above example using SPSS completely.

To do just a confidence interval: enter data, then Analyze→Descriptive Statistics→ Explore. Click on "Statistics" and change the CI to 90%. Then hit "OK."

If you need to do a hypothesis test and a CI, go to Analyze→Compare Means→ One-sample T test. Change the "test value" to 105 (since that is our H_0), change "options" to 90%, and hit "OK." (This will give you the output below.)

One-Sample Test

|	Test Value = 105					
					90% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
radon detector readings	-.319	11	.755	-.8667	-5.739	4.005

$H_0 : \mu = 105$ $H_a : \mu \neq 105$	
$H_0 : \mu = 105$ $H_a : \mu < 105$	
$H_0 : \mu = 105$ $H_a : \mu > 105$	

You must choose your hypotheses BEFORE you examine the data. When in doubt, do a two-sided test.

<u>Normal quantile plots</u>

In SPSS, go to Graphs→ Q-Q. Move your variable into "variable" column and hit "OK."

Normal Q-Q Plot of Radon Detector Reading
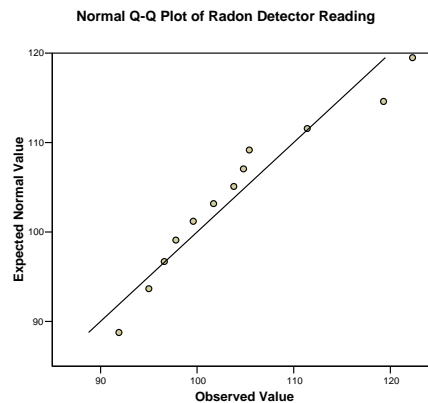
Look to see how closely the data points (dots) follow the diagonal line. The line will always be a 45-degree line. Only the data points will change. The closer they follow the line, the more normally distributed the data is.

What happens if the t procedure is not appropriate? What if you have outliers or skewness with a smaller sample size ($n < 40$)?

- **Outliers**: Investigate the cause of the outlier(s).

    o Was the data recorded correctly? Is there any reason why that data might be invalid (an equipment malfunction, a person lying in their response, etc.)? If there is a good reason why that point could be disregarded, try taking it out and compare the new confidence interval or hypothesis test results to the old ones.

    o If you don't have a valid reason for disregarding the outlier, you have to the outlier in and not use the t procedures.

- **Skewness**:

    o If the skewness is not too extreme, the t procedures are still appropriate if the sample size is bigger than 15.

If the skewness is extreme or if the sample size is less than 15, you can use nonparametric procedures. One type of nonparametric test is similar to the t procedures except it uses the median instead of the mean. Another possibility would be to transform the data, possibly using logarithms. A statistician should be consulted if you have data which doesn't fit the t procedures requirements. *We won't cover nonparametric procedures or transformations for non-normal data in this course, but your book has supplementary chapters (14 and 15) on these topics online if you need them later in your own research. They are also discussed on pages 465-470 of your book.*

What do you do when you have 2 lists of data instead of 1?

First decide whether you have 1 sample with 2 measurements OR 2 independent samples with one measurement each.

1. **Matched Pairs** (covered in 7.1)

- One group of individuals with 2 different measurements on them
- Same individuals, different measurements
- Examples: pre- and post-tests, before and after measurements
- Based on the difference obtained between the 2 measurements

1) Find the difference = post test – pre test (or before - after, etc.), in the individual measurements.
2) Find the sample mean $\bar{d}$ and sample standard deviation s of these differences.
3) Use the *t* distribution because the standard deviation is estimated from the data.

- Confidence interval: $$\bar{d} \pm t^* \frac{s}{\sqrt{n}}$$

- Hypothesis testing:
  H_0: $\mu_{diff} = 0$

  t test statistic: $t_0 = \dfrac{\bar{d} - 0}{s/\sqrt{n}}$

  Conclusion: "Reject / Do Not Reject H_0. There is / There is not enough evidence that the population mean difference in _____ is _____."

Example 4 (Exercise 7.31)

Researchers are interested in whether Vitamin C is lost when wheat soy blend (CSB) is cooked as gruel. Samples of gruel were collected, and the vitamin C content was measured (in mg per 100 grams of gruel) before and after cooking. Here are the results:

Sample	1	2	3	4	5	Mean	St. Dev.
Before	73	79	86	88	78	80.8	6.14
After	20	27	29	36	17	25.8	7.53
Before-After	53	53	57	52	61	55.0	3.94

    a)   Set up an appropriate hypothesis test and carry it out for these data. State your conclusions in a sentence. Use $\alpha=.10$.

    b)   Find a 90% confidence interval for the mean vitamin C content loss.

Example 5

In an effort to determine whether sensitivity training for nurses would improve the quality of nursing provided at an area hospital, eight different nurses were selected and their nursing skills were given a score from 1-10. After this initial screening, a training program was administered, then the same nurses were rated again. Below is a table of their pre- and post-training scores. Conduct a test to determine whether training improved the quality of nursing provided.

Individuals	Pre-training score	Post-training score
1	2.56	4.54
2	3.22	5.33
3	3.45	4.32
4	5.55	7.45
5	5.63	7.00
6	7.89	9.80
7	7.66	5.33
8	6.20	6.80

Enter the pre and post training scores to SPSS. Then Analyze→Compare Means→Paired-Samples T-test. Then input both variable names and hit the arrow key. If you need to change the confidence interval, go to "Options." SPSS will always do the left column of data – the right column of data for the order of the difference. If this bothers you, just be careful how you enter the data into the program

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Post-training score	6.3212	8	1.82086	.64377
	Pre-training score	5.2700	8	2.01808	.71350

Data entered as written above with pre-training in left column and post-training in right column:

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	pretraining - posttraining	-1.05125	1.47417	.52120	-2.28369	.18119	-2.017	7	.084

Data entered backwards from how it is written above with post-training in left column and pre-training in right column:

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Post-training score - Pre-training score	1.05125	1.47417	.52120	-.18119	2.28369	2.017	7	.084

What's different? What's the same? Which one matches the way that you defined μ_{diff}?

a. What are your hypotheses?

b. What is the test statistic?

c. What is the P-value?

d. What is your conclusion in terms of the story if α=.05?

e. What is the 95% confidence interval of the difference in nursing scores?

2-Sample Comparison of Means (covered in 7.2)

- A group of individuals is divided into 2 different experimental groups
- **Each group has different individuals who may receive different treatments**
- Responses from each sample are independent of each other.
- Examples: treatment vs. control groups, male vs. female, 2 groups of different women

- Hypothesis Test

    H_0: $\mu_A = \mu_B$ (same as H_0: $\mu_A - \mu_B = 0$)

    H_a: $\mu_A > \mu_B$ or H_a: $\mu_A < \mu_B$ or H_a: $\mu_A \neq \mu_B$ (pick one)

    2-Sample *t* Test Statistic is used for hypothesis testing when the standard deviations are ESTIMATED from the data (these are approximately *t* distributions, but not exact)

    $$t_0 = \frac{(\overline{x}_A - \overline{x}_B)}{\sqrt{\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}}} \sim t \text{ distribution with df} = \min(n_A - 1, n_B - 1)$$

    Conclusion: "Reject / Do Not Reject H_0. There is / There is not enough evidence that the difference between the population means for _____ and _____ is _____."

- **Confidence Interval** for $\mu_A - \mu_B$:

    $$(\overline{x}_A - \overline{x}_B) \pm t^* \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}} \text{ where } t^* \sim t \text{ distribution with df} = \min\left(n_A - 1, n_B - 1\right)$$

- ***Equal sample sizes are recommended, but not required.

Assumptions for Comparing Two Means

1. Two independent random samples from two distinct populations or two treatment groups from randomized comparative experiments are compared. The same variable is measured on both samples. The assumption of independence says that one sample has no influence on the other.
2. Both populations are normally distributed.
3. The means μ_1 and μ_2 and standard deviations σ_1 and σ_2 of both populations are unknown.

Robustness of the Two Sample t Procedures

The two-sample *t* procedures are more robust than the one-sample *t* methods, particularly when the distributions are not symmetric. They are robust in the following circumstances:

- If two samples are of equal size and the two populations that the samples come from have similar distributions, then the *t* distribution is accurate for a variety of distributions, even when the sample sizes are as small as $n_1 = n_2 = 5$.
- When the two population distributions are different, larger samples are needed.
  - If $n_1 + n_2 < 15$: Use two-sample *t* procedures if the data are close to normal. If the data are clearly non-normal or if outliers are present, do not use *t*.
  - If $15 \leq n_1 + n_2 < 40$: The *t* procedures can be used except in the presence of outliers or strong skewness.
  - If $n_1 + n_2 \geq 40$: The *t* procedures can be used even for clearly skewed distributions

Example 6

A group of 15 college seniors are selected to participate in a manual dexterity skill test against a group of 20 industrial workers. Skills are assessed by scores obtained on a test taken by both groups. Conduct a 5% alpha hypothesis test to determine whether the industrial workers had better manual dexterity skills than the students. Descriptive statistics are listed below. Also construct a 95% confidence interval for this problem.

group	n	$\overline{x}$	s	df
students	15	35.12	4.31	
workers	20	37.32	3.83	

Example 7 *(Exercise 7.84)*

The SSHA is a psychological test designed to measure the motivation, study habits, and attitudes towards learning of college students. These factors, along with ability, are important in explaining success in school. A selective private college gives the SSHA to an SRS of both male and female first-year students. The data for the women are as follows:

154	109	137	115	152	140	154	178	101
103	126	126	137	165	165	129	200	148

Here are the scores for the men:

108	140	114	91	180	115	126	92	169	146
109	132	75	88	113	151	70	115	187	104

     a) Test whether the mean SSHA score for men is different than the mean score for women. State your hypotheses, carry out the test using SPSS, obtain a *P*-value, and give your conclusions. Use a 10% significance level.

*When you enter your data into SPSS, have 2 variables: gender (type: string) and score (numeric). In the gender column, state whether a score is from a man or a woman, and in the score column, state all 38 scores. Analyze→Compare Means→Independent-Samples T Test. Move **score** into "Test Variable(s)" box. Move **gender** into "Grouping Variable" box, and then click "Define Groups" and state which "woman" and "man" as group 1 and group 2, hit "Continue". We will need a 90% confidence interval in part c, so go to "Options" to change it.*

Group Statistics

	gender	N	Mean	Std. Deviation	Std. Error Mean
score	woman	18	141.06	26.436	6.231
	man	20	121.25	32.852	7.346

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									90% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
score	Equal variances assumed	.862	.359	2.032	36	.050	19.806	9.745	3.353	36.258
	Equal variances not assumed			2.056	35.587	.047	19.806	9.633	3.538	36.073

What do we do with this "Equal variances assumed" and "Equal variances not assumed"? Always go with the bottom row, "Equal variances not assumed." This is the more conservative approach.

  b.) Most studies have found that the mean SSHA score for men is lower than the mean score in a comparable group of women. Test this supposition here.

  c.) Give a 90% confidence interval for the difference in means of SSHA scores of male and female first-year students at this college.