MATH 3307
Lesson 29

Confidence Interval for a Population Mean

Recall that formula for a confidence interval is *statistic* $\pm$ *margin of error*. When we are making an inference about a population mean, the statistic will be our sample mean, $\bar{x}$.
The critical value we use to find the margin of error for our calculation will be based on whether the population or sample standard deviation is known. When the population standard deviation is known, we use the formula $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$ and when it is unknown, we will need to find the sample standard deviation, s, and use the formula $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$ where t^* is the t-critical value based on $n-1$ degrees of freedom.

If you know population standard deviation (sigma), then use a z-test
If you do not know population standard deviation, then use a t-test.

Confidence Intervals

General formula: *statistic* ± *margin of error*

One-sample z-test: $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$

Two-proportion z-test: $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

One-sample t-test: $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$

One-proportion z-test: $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

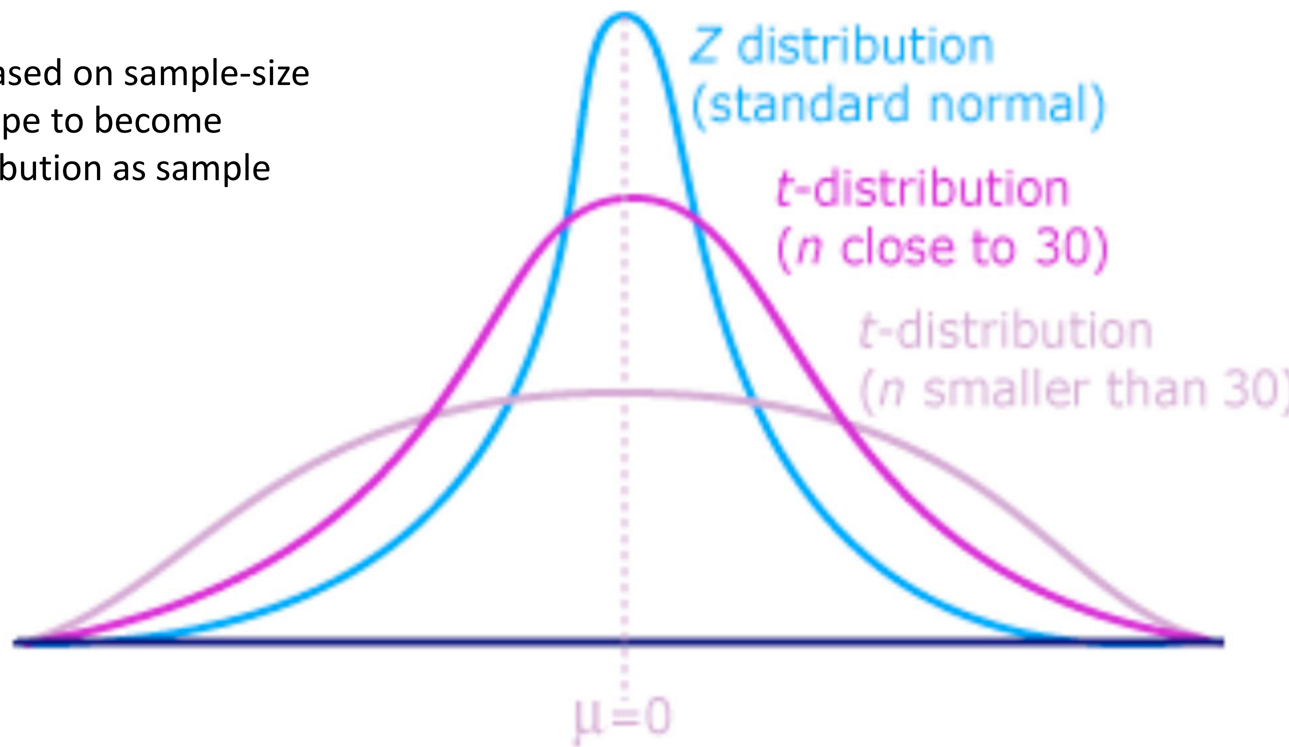
Two-sample z-test: $(\bar{x}_1 - \bar{x}_2) \pm z^* \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Two-sample t-test: $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

So, what is *t**?

t-distribution vs. standard normal distribution:

t-distribution is based on sample-size (it will change shape to become closer to a z-distribution as sample size increases).



How do we find critical values for a *t*-distribution?

Look at the online text book, under appendices.

Degrees of freedom are on the left and the top end (1-confidence level/2) are given at the top.

	Upper tail probability *p*											
df	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
1	1.000	1.376	1.963	3.078	6.314	12.706	15.895	31.821	63.657	127.321	318.309	636.619
2	0.8165	1.0607	1.3862	1.8856	2.9200	4.3027	4.8487	6.9646	9.9248	14.0890	22.3271	31.5991
3	0.7649	0.9785	1.2498	1.6377	2.3534	3.1824	3.4819	4.5407	5.8409	7.4533	10.2145	12.9240
4	0.7407	0.9410	1.1896	1.5332	2.1318	2.7764	2.9985	3.7469	4.6041	5.5976	7.1732	8.6103
5	0.7267	0.9195	1.1558	1.4759	2.0150	2.5706	2.7565	3.3649	4.0321	4.7733	5.8934	6.8688
6	0.7176	0.9057	1.1342	1.4398	1.9432	2.4469	2.6122	3.1427	3.7074	4.3168	5.2076	5.9588
7	0.7111	0.8960	1.1192	1.4149	1.8946	2.3646	2.5168	2.9980	3.4995	4.0293	4.7853	5.4079
8	0.7064	0.8889	1.1081	1.3968	1.8595	2.3060	2.4490	2.8965	3.3554	3.8325	4.5008	5.0413
9	0.7027	0.8834	1.0997	1.3830	1.8331	2.2622	2.3984	2.8214	3.2498	3.6897	4.2968	4.7809
10	0.6998	0.8791	1.0931	1.3722	1.8125	2.2281	2.3593	2.7638	3.1693	3.5814	4.1437	4.5869
11	0.6974	0.8755	1.0877	1.3634	1.7959	2.2010	2.3281	2.7181	3.1058	3.4966	4.0247	4.4370
12	0.6955	0.8726	1.0832	1.3562	1.7823	2.1788	2.3027	2.6810	3.0545	3.4284	3.9296	4.3178
13	0.6938	0.8702	1.0795	1.3502	1.7709	2.1604	2.2816	2.6503	3.0123	3.3725	3.8520	4.2208
14	0.6924	0.8681	1.0763	1.3450	1.7613	2.1448	2.2638	2.6245	2.9768	3.3257	3.7874	4.1405
15	0.6912	0.8662	1.0735	1.3406	1.7531	2.1314	2.2485	2.6025	2.9467	3.2860	3.7328	4.0728
16	0.6901	0.8647	1.0711	1.3368	1.7459	2.1199	2.2354	2.5835	2.9208	3.2520	3.6862	4.0150

Assumptions:

The assumptions for a population mean are:

1. The sample must be an SRS from the population of interest.
2. The data must come from a normally distributed population. If this is not the case or if we are unsure whether the population is normally distributed, the sampling distribution of $\overline{x}$ must be normally distributed. (Recall from section 4.4 that we can assume that the sampling distribution of $\overline{x}$ is normal for values of n greater than 30.)

For the second assumption:
Either: the population is stated as being normally distributed.
or
the sample size is greater than 30.

Changes is the Confidence Interval:

The following will cause an increase in the width of the confidence interval:

- An increase in the Margin of Error
- An increase in the Standard Deviation (σ or s).
- An increase in the Confidence Level.
- A decrease in the Sample Size.

Note: Changes in the sample mean ($\bar{x}$) will affect the center of the interval but will have *no effect* on the interval width.

Examples:

Look for the first mention of the sample. If standard deviation is given after that, it is sample standard deviation (use t-test). If standard deviation is given before that, then it is population standard deviation (use a z-test).

1. Suppose your class is investigating the weights of Snickers 1-ounce fun-size candy bars to see if customers are getting full value for their money. Assume that the weights are normally distributed with standard deviation $\sigma = .005$ ounces. Several candy bars are randomly selected and weighed with sensitive balances borrowed from the physics lab.

z - test

The weights are: .95 1.02 .98 .97 1.05 1.01 .98 1.00

We want to determine a 90% confidence interval for the true mean, μ.

```
> assign("x",c(.95,1.02,.98,.97,1.05,1.01,.98,1.0))
> mean(x)
[1] 0.995
> qnorm(1.90/2)
[1] 1.644854
> mean(x)-qnorm(1.90/2)*.005/sqrt(8)
[1] 0.9920923
> mean(x)+qnorm(1.90/2)*.005/sqrt(8)
[1] 0.9979077
```

$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$

a. What is the sample mean?

b. Determine z^*.

c. Determine the 90% confidence interval.

d. Write a sentence that explains the significance of the confidence interval.

[0.992, 0.997]

We are 90% confident that the population mean is between 0.992 and 0.998 oz.

If we repeat this process several times, 90% of intervals will contain the population mean

Popper 25

You select a sample of 50 people with a mean height of 72 inches from a population that has a standard deviation of 3 inches.

1. What would the margin of error be for a 95% confidence interval?

a. 0.698 b. 0.832 c. 16.748 d. 1.644

```
qnorm(1.95/2)*3/sqrt(50)
1] 0.8315423
```

2. What is the width of confidence interval?

a. 1.396 b. 0.349 c. 1.66 d. 0.415

```
2*0.8315423
1] 1.663085
```

3. What is the confidence interval?

a. [71.17, 72.83] b. [69.43, 75.11] c. [72.41, 75.21] d. [47.31, 53.65]

```
72-qnorm(1.95/2)*3/sqrt(50)
1] 71.16846
72+qnorm(1.95/2)*3/sqrt(50)
1] 72.83154
```

Popper 25 Continued:

4. Give an interpretation of this interval:

a. A randomly selected person from the population has a 95% chance being within the interval.
b. There is a 95% chance that the population mean will fall within the interval.
c. If numerous similar samples are selected, 95% of them will contain the population mean.
d. All of these choices
e. 2 of the choices a, b or c.

5. What will cause the width of the confidence interval to increase?

a. decrease n
b. decrease confidence level
c. decrease standard deviation
d. increase the mean