Confidence Intervals for Small Sample Means

Dr Tom Ilvento
Department of Food and Resource Economics

University of Delaware

Overview

- We discussed a confidence interval for proportions and means
- For the confidence interval we needed
  - An estimator – either the sample proportion or the mean
  - A sample from which I obtain my estimate
  - Knowledge of the sampling distribution of the estimator
  - A level or confidence (Confidence Coefficient) and a level of error (alpha)
  - **As long as the sample size is LARGE!**

$$p_s \pm Z_{\alpha/2}\sqrt{\frac{p_s(1-p_s)}{n}}$$

$$\bar{x} \pm Z_{\alpha/2}\frac{s}{\sqrt{n}}$$

Confidence Interval for a Mean

- If σ is known, we use σ in the formula and use a z-value for the confidence interval.
- However, we rarely know the standard deviation of the population
- So I use my sample estimate, give as s.
- But I have a concern with this approach with the mean. I have two estimates
  - The estimate of the mean
  - The estimate of the standard deviation (s), which is used to estimate the standard error

Issues of Normality and Sample Size when using s

- **Is it a problem to use the sample estimate of s when σ is not known?**
  - If the population is distributed normally, we have less of a problem – the sampling distribution will be normal
  - And if the population is not distributed normally, as long as the sample size is large, the Central Limit Theorem says as the sample size becomes larger, the sampling distribution will start to approach normality
- **But what do we do if σ is unknown, and the sample size is small?**

Relax and have a beer!

- W.S. Gossett worked for Guinness Brewery in Ireland around 1900
- In quality control tests he noticed the problem of using the z-distribution with small samples
- His solution was the t-distribution
- Based on the variable being distributed approximately normal

The t-distribution

- Similar to the standard normal distribution
- The **t-distribution** varies with **n** (sample size) via **degrees of freedom**
  - **df = n-1**
- As n gets larger, the t-distribution approximates the z distribution
- **Here's a Table comparing values of z and t for a 95% C.I.**

Sample Size	Z-Value	t-value
10	too small	2.262
20	too small	2.093
30	1.960	2.045
50	1.960	2.010
100	1.960	1.984
500	1.960	1.965
1000	1.960	1.962

The t-table

- Organized with degrees of freedom as rows
- Probabilities in the right tail (α) are the columns
  - Our table has α in two tails as well
- We substitute the t-value from the table for a z-value in the C.I.
  - Example, 95% C.I., 15 d.f.
- **In the case of a small sample, n < 30, the Central Limit Theorem doesn't hold.**
- **In order to do a C.I., a big assumption with a small sample is that the population is distributed approximately normal**

Critical Values of t-Distribution

The table shows the critical t-values for a given alpha level (one-tailed or two-tailed) and degrees of freedom. The degrees of freedom are the rows.

d.f.	Area in Two Tails 0.2000	0.1000	0.0500	0.0200	0.0100	0.0020	0.0010
v	Area in One Tail 0.1000	0.0500	0.0250	0.0100	0.0050	0.0010	0.0005
1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
31	1.309	1.696	2.040	2.453	2.744	3.375	3.633
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
33	1.308	1.692	2.035	2.445	2.733	3.356	3.611
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
35	1.306	1.690	2.030	2.438	2.724	3.340	3.591
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
37	1.305	1.687	2.026	2.431	2.715	3.326	3.574
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
39	1.304	1.685	2.023	2.426	2.708	3.313	3.558
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
45	1.301	1.679	2.014	2.412	2.690	3.281	3.520
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	1.291	1.662	1.987	2.368	2.632	3.183	3.402
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
150	1.287	1.655	1.976	2.351	2.609	3.145	3.357
Infinity	1.282	1.645	1.960	2.326	2.576	3.090	3.291

The t-table

- The Confidence Interval based on a t-value is given below
- The t-value is interpreted like the z-value
- NOTE: the t-value represents the corresponding value at α/2, which is in the right tail of the curve
- So a t-value for 30 degrees of freedom at the .025 level is 2.042
- This corresponds to a z-value of 1.96
- And is used for a 95% C.I.

$$\bar{x} \pm t_{\alpha/2,n-1\text{d.f.}} \left(\frac{s}{\sqrt{n}} \right)$$

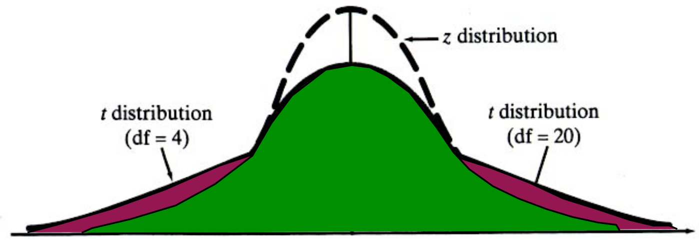
Critical Values of t-Distribution

The table shows the critical t-values for a given alpha level (one-tailed or two-tailed) and degrees of freedom. The degrees of freedom are the rows.

d.f.	Area in Two Tails 0.2000	0.1000	0.0500	0.0200	0.0100	0.0020	0.0010
v	Area in One Tail 0.1000	0.0500	0.0250	0.0100	0.0050	0.0010	0.0005
1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
31	1.309	1.696	2.040	2.453	2.744	3.375	3.633
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
33	1.308	1.692	2.035	2.445	2.733	3.356	3.611
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
35	1.306	1.690	2.030	2.438	2.724	3.340	3.591
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
37	1.305	1.687	2.026	2.431	2.715	3.326	3.574
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
39	1.304	1.685	2.023	2.426	2.708	3.313	3.558
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
45	1.301	1.679	2.014	2.412	2.690	3.281	3.520
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	1.291	1.662	1.987	2.368	2.632	3.183	3.402
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
150	1.287	1.655	1.976	2.351	2.609	3.145	3.357
Infinity	1.282	1.645	1.960	2.326	2.576	3.090	3.291

The t-distribution

- There is a different t-distribution for each d.f.
- The following shows t-distributions compared to the standard normal
  - t for d.f. 4
  - t for d.f. 20
- Compared to the standard normal, the t-distribtion is flatter and has fatter tails
- **As the d.f. gets larger, the t-value approaches z**



Basic Steps for C.I. for Small Sample Mean

- Set a probability that an interval estimator encloses the population parameter **p = .95**
- Set an alpha level as 1-p **(1-p) = .05**
- Divide the alpha by 2 **=.025**
- Calculate the degrees of freedom as n-1 **if n=24, d.f. = 23**
- Locate the ½ probability value for your degrees of freedom in the t-Table
- **This only applies if our variable is approximately normal!**

Critical Values of t-Distribution

The table shows the critical t-values for a given alpha level (one-tailed or two-tailed) and degrees of freedom. The degrees of freedom are the rows.

	Area in Two Tails						
d.f.	0.2000	0.1000	0.0500	0.0200	0.0100	0.0020	0.0010
	Area in One Tail						
v	0.1000	0.0500	0.0250	0.0100	0.0050	0.0010	0.0005
1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
31	1.309	1.696	2.040	2.453	2.744	3.375	3.633
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
33	1.308	1.692	2.035	2.445	2.733	3.356	3.611
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
35	1.306	1.690	2.030	2.438	2.724	3.340	3.591
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
37	1.305	1.687	2.026	2.431	2.715	3.326	3.574
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
39	1.304	1.685	2.023	2.426	2.708	3.313	3.558
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
45	1.301	1.679	2.014	2.412	2.690	3.281	3.520
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	1.291	1.662	1.987	2.368	2.632	3.183	3.402
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
150	1.287	1.655	1.976	2.351	2.609	3.145	3.357
Infinity	**1.282**	**1.645**	**1.960**	**2.326**	**2.576**	**3.090**	**3.291**

C.I. for a Mean

- Use the Population parameter **σ** if it **is known**, and a **z-value**

$$\bar{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

- Use the sample estimate s, and a **t-value**, **if σ is not known**

$$\bar{x} \pm t_{n-1d.f.}\left(\frac{s}{\sqrt{n}}\right)$$

- **To be safe, software packages present all confidence intervals and hypotheses tests using a t-value rather than a z-value**

95% C.I. for Well Water Example using a t-value

- Let's say there are 2,500 households in the area
- I could try to test them all, but at $50 a test it would cost $125,000 and weeks of work
- So, I decide to take 50 well water samples, and test for the presence of nitrogen
  - n = 50 **d.f. = 49**
  - Mean = 7 mg/l
  - s = 3.003 mg/l
  - S.E. = 3.003/(50)$^{.5}$ = .425

95% C.I. for Well Water Example using a t-value

$$\bar{x} \pm t_{n-1d.f.}\left(\frac{s}{\sqrt{n}}\right)$$

- To solve for this 95% B.O.E.
  - $t_{.05/2,\ 49\ d.f.} = 2.010$
  - S.E. = .425
  - Mean = 7.000
- 7.000 ± 2.010(.425)
- **7.000 ± .854**
- **6.146 to 7.854**

Nitrate+Nitrite	
Mean	**7.000**
Standard Error	**0.425**
Median	7.050
Mode	7.100
Standard Deviation	**3.003**
Sample Variance	9.018
Kurtosis	-0.723
Skewness	0.101
Range	11.600
Minimum	1.600
Maximum	13.200
Sum	350.000
Count	50
Confidence Level(95.0%)	**0.853**

Excel gives you the B.O.E. based on the Confidence Coefficient

Problem 1

- Spinifex pigeons in Western Australia rely entirely on seeds for food
- Our research will examine stomach contents of 16 pigeons
- Recorded the weight in grams of dry seed of each pigeon - **assumed to be approximately normal**
- Sample Statistics
  - n=16
  - Mean = 1.373
  - s = 1.034
- **Construct a 99% C.I.**

Spinifex 99% C.I. f

- Calculate the degrees of freedom as n-1 **if n=16, d.f. = 15**
- t-value for 99% C.I.
  - α = .01
  - α/2 = .005 in each tail
  - $t_{.005\text{ with }15\ d.f.} = 2.947$
- 1.373 ± 2.947(.2585)
- 1.373 **± .762**
- **.611 to 2.135**

$$\bar{x} \pm t_{\alpha/2,n-1\text{d.f.}}\left(\frac{s}{\sqrt{n}}\right)$$

Critical Values of t-Distribution

The table shows the critical t-values for a given alpha level (one-tailed or two-tailed) and degrees of freedom. The degrees of freedom are the rows.

	Area in Two Tails						
d.f.	**0.2000**	**0.1000**	**0.0500**	**0.0200**	**0.0100**	**0.0020**	**0.0010**
	Area in One Tail						
v	**0.1000**	**0.0500**	**0.0250**	**0.0100**	**0.0050**	**0.0010**	**0.0005**
1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
31	1.309	1.696	2.040	2.453	2.744	3.375	3.633
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
33	1.308	1.692	2.035	2.445	2.733	3.356	3.611
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
35	1.306	1.690	2.030	2.438	2.724	3.340	3.591
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
37	1.305	1.687	2.026	2.431	2.715	3.326	3.574
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
39	1.304	1.685	2.023	2.426	2.708	3.313	3.558
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
45	1.301	1.679	2.014	2.412	2.690	3.281	3.520
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	1.291	1.662	1.987	2.368	2.632	3.183	3.402
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
150	1.287	1.655	1.976	2.351	2.609	3.145	3.357
Infinity	**1.282**	**1.645**	**1.960**	**2.326**	**2.576**	**3.090**	**3.291**

Confidence Interval Problem

- A furniture company wants to test a random sample of sofas to determine how long the cushions last
- They simulate people sitting on the sofas by dropping a heavy object on the cushions until they wear out – they count the number of drops it takes
- This test involves 9 sofas
  - Mean = 12,648.889
  - s = 1,898.673
- Assume it follows a normal distribution. Generate a 95% Confidence Interval for this problem

Answer

- n = 9, small sample, approximately normal, so use t-distribution with 8 d.f.
  - **$t_{.05/2,\ 8\ d.f.} = 2.306$**
- Standard error = 1,898.673/(9)$^{.5}$
  - **SE= 632.891**
- 95% C.I.
  - 12,648.889 ± 2.306(632.891)
  - 12,648.889 ± **1,459.447**
  - **11,189.442 to 14,108.336**

Sofa Confidence Interval using Excel

- I entered the data into a column in Excel
- I then used the following sequence
  - Tools...Data Analysis...Descriptive Statistics
- I then follow the options, including:
  - Identify the Input Range, marking a label is in the first row
  - Output range
  - Descriptive statistics
  - A 95% Confidence Interval

Drops	
Mean	12648.889
Standard Error	632.891
Median	12742
Mode	#N/A
Standard Deviation	1898.673
Sample Variance	3604958.111
Kurtosis	-0.676
Skewness	-0.372
Range	5886
Minimum	9459
Maximum	15345
Sum	113840
Count	9
Confidence Level(95.0%)	1459.450

A few more points on small sample confidence intervals

- If we cannot assume a normal distribution
  - The probability associated with our interval is not (1 -α)
  - **We really shouldn't construct a C.I.**
  - Or we should get more data - larger sample size!
- If σ is known, we can use the z instead of the t, but we still need to have an approximately normal distribution for small samples
- **It is ok, even preferred, to use the t-distribution for larger samples**

Summary

- Confidence Intervals provide an **interval estimate** of a Population Parameter
- Requires knowledge of the sampling distribution of the estimator
  - We treat our estimate from a sample as one of many possible estimates from many possible samples
  - Our confidence is in the process
- Figure a C.I. Probability level as (1-α)
  - (1 -α) is referred to as the **Confidence Coefficient**
  - Where **α** is the probability of being wrong in our assertion
  - and **α/2** represents the probability in either tail of the sampling distribution

Summary

- For proportions, you can use a z-score provided the sample size is large enough (Binomial approximation)

$$p \pm z_{\alpha/2}\sqrt{\frac{p\,q}{n}}$$

- For the mean
  - If σ is known, use a z-value for the C.I. similar to proportions

$$\overline{x} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

  - If σ is unknown, use the t-table with n-1 degrees of freedom

$$\overline{x} \pm t_{\alpha/2,n-1 d.f.}\frac{s}{\sqrt{n}}$$

- In this class, I will allow you to use a z-value for hand problems when n > 30.
- If the **sample size is small** (<30), and the distribution is approximately normal, you **must use the t-table** with n-1 degrees of freedom