

Confidence Intervals:

Based on a sample, calculate the interval of values in which the population parameter (mean, proportion, standard deviation) may lie with a degree of certainty. This degree of certainty (level of confidence) is given as a percentage or the equivalent decimal.

The most used confidence intervals are 80%, 90%, 95%, 98% and 99%.

The population parameter lies in between the sample point estimate plus or minus an error:

$$\text{Sample point estimate} \pm \text{Error}$$

The three confidence intervals in STA2023:

1. Confidence Interval for the population mean, population standard deviation σ , known.

$$\bar{x} \pm \text{Error}$$

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

2. Confidence Interval for the population mean, population standard deviation σ , unknown. Use the sample standard deviation (for a normal distributed variable), use the *t distribution*.

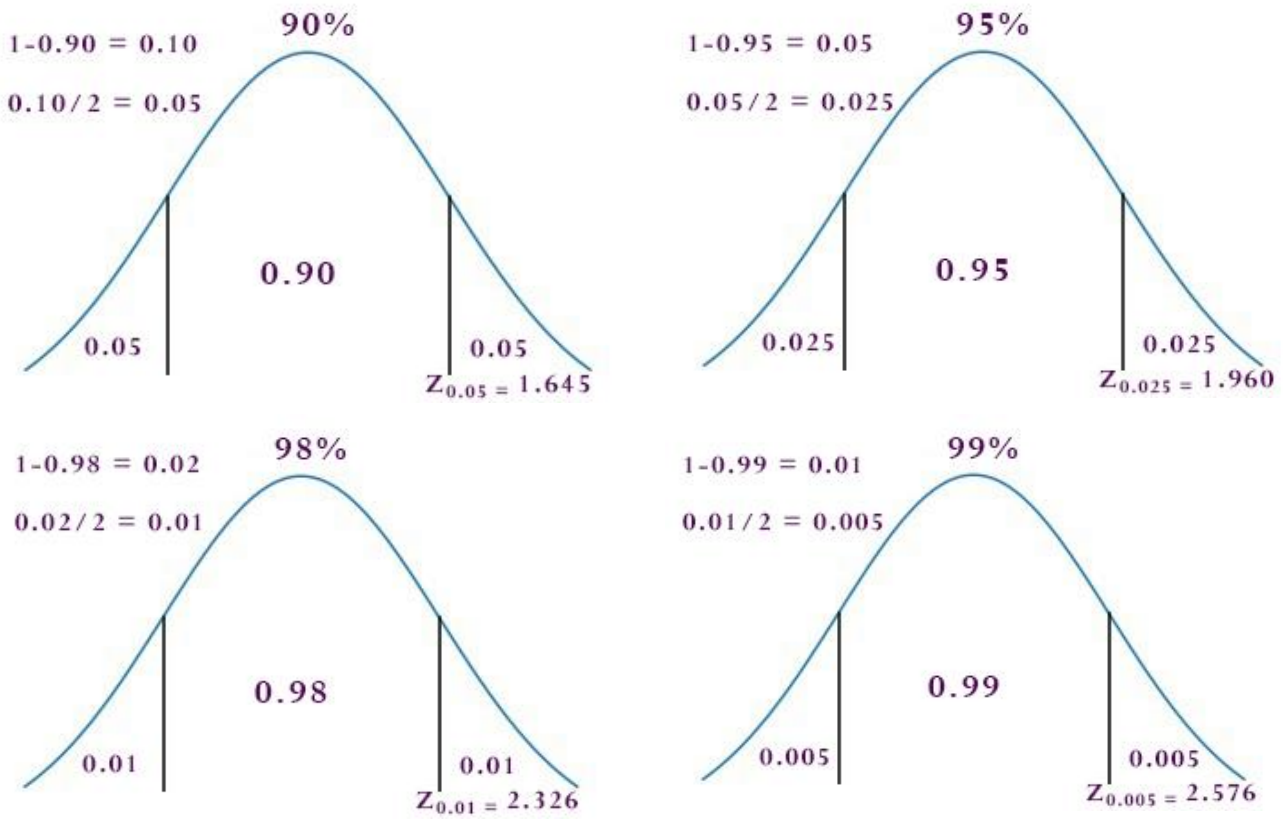
$$\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

3. Confidence Interval for the population proportion:

$$\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} \quad \text{where} \quad \hat{q} = 1 - \hat{p}$$

<p>Sample size for means:</p> $n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2$	<p>Sample size por proportions:</p> $n = \hat{p} \cdot \hat{q} \left(\frac{z_{\alpha/2}}{E} \right)^2$ <p>If \hat{p} unknown, use 0.5</p>
--	---

Critical values, $Z_{\alpha/2}$



In short:

90%, $Z_{\alpha/2} = 1.645$ 95%, $Z_{\alpha/2} = 1.960$ 98%, $Z_{\alpha/2} = 2.326$ 99%, $Z_{\alpha/2} = 2.576$

TABLE A-3		t Distribution: Critical t Values				
		Area in One Tail				
		0.005	0.01	0.025	0.05	0.10
Degrees of Freedom	99%CI	98%CI	Area in Two Tails			80%CI
	0.01	0.02	0.05	0.10	0.20	
1	63.657	31.821	12.706	6.314	3.078	
2	9.925	6.965	4.303	2.920	1.886	
3	5.841	4.541	3.182	2.353	1.638	
4	4.604	3.747	2.776	2.132	1.533	
5	4.032	3.365	2.571	2.015	1.476	
6	3.707	3.143	2.447	1.943	1.440	
7	3.499	2.998	2.365	1.895	1.415	
8	3.355	2.896	2.306	1.860	1.397	
9	3.250	2.821	2.262	1.833	1.383	
10	3.169	2.764	2.228	1.812	1.372	
11	3.106	2.718	2.201	1.796	1.363	
12	3.055	2.681	2.179	1.782	1.356	

See complete table at the end of this document.

Degrees of freedom, $df = n - 1$ where n is the sample size.

Examples:

Question 1

The lengths, in inches, of adult corn snakes are normally distributed with a population standard deviation of 8 inches and an unknown population mean. A random sample of 25 snakes is taken and results in a sample mean of 58 inches. Find a 99% confidence interval estimate for the population mean.

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \qquad 58 \pm 2.576 \cdot \frac{8}{\sqrt{25}} \qquad 58 \pm 4.1216$$

by subtracting and adding $E = 4.1216$, yields the 99% Conf Interval: (53.88, 62.12)

Interpretation:

We can estimate with 99% confidence that the true population mean length of adult corn snakes is between 53.88 and 62.12 inches.

Question 2

Suppose the scores of a standardized test are normally distributed. If the population standard deviation is 3 points, what minimum sample size is needed to be 90% confident that the sample mean is within 2 points of the true population mean?

Sample size,

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2 \quad \text{Therefore, } n = \left(\frac{1.645 \cdot 3}{2} \right)^2 = 6.08 \text{ rounded (always) up} = 7$$

Question 3

The weekly salaries of sociologists in the United States are normally distributed and have a known population standard deviation of 425 dollars and an unknown population mean. A random sample of 22 sociologists is taken and gives a sample mean of 1520 dollars.

Find the margin of error for the confidence interval for the population mean with a 98% confidence level.

$$\text{ME, margin of error given by } z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad \text{Therefore, } \text{ME} = 2.326 \cdot \frac{425}{\sqrt{22}} = 210.76$$

Question 4

The commute times for workers in a city are normally distributed with an unknown population mean and standard deviation.

If a random sample of 20 workers is taken and results in a sample mean of 21 minutes and sample standard deviation of 6 minutes, find a 95% confidence interval estimate for the population mean using the Student's t-distribution.

df	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
...
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831

Find the margin of error, for a 95% confidence interval estimate for the population mean using the Student's t-distribution.

$$\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \quad \text{Margin of error, ME given by} \quad t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$\text{Therefore ME} = 2.093 \cdot \frac{6}{\sqrt{20}} = 2.8081 = 2.81$$

$$\text{Interval, } \bar{x} \pm \text{Error} \quad \text{or} \quad 21 \pm 2.81$$

by subtracting and adding E, yields the 95% Confidence interval: (18.19, 23.81)

Question 5

Emma wants to estimate the percentage of people who use public transportation in a city. She surveys 140 individuals and finds that 62 use public transportation. Find the confidence interval for the population proportion with a 99% confidence level.

Answer: This a proportion interval. Find \hat{p} first:

$$\hat{p} = \frac{x}{n} = \frac{62}{140} = 0.4429 \quad \text{and} \quad \hat{q} = 1 - \hat{p} \quad \text{Therefore,} \quad \hat{q} = 1 - 0.4429 = 0.5571$$

$$\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} \quad \text{substituting values, becomes:}$$

$$0.4429 \pm 2.576 \cdot \sqrt{\frac{0.4429 \times 0.5571}{140}}$$

$$0.4429 \pm 0.1081$$

$$(0.335, \quad 0.551)$$

Question 6

Suppose a clothing store wants to determine the current percentage of customers who are over the age of forty.

How many customers should the company survey in order to be 90% confident that the estimated (sample) proportion is within 4 percentage points of the true population proportion of customers who are over the age of forty?

Answer: Sample size for proportions:

$n = \hat{p} \cdot \hat{q} \left(\frac{z_{\alpha/2}}{E} \right)^2$ since \hat{p} is unknown (sometimes researchers use a previous study result, that is the case in which \hat{p} is known; otherwise, like in this case, use 0.5; the $z_{\alpha/2} = 1.645$ for a 90% CI, 4% as decimal, $E = 0.04$; substituting values:

$$n = 0.5 \cdot 0.5 \left(\frac{1.645}{0.04} \right)^2 = 422.81 = 423$$

TABLE A-3		<i>t</i> Distribution: Critical <i>t</i> Values				
		Area in One Tail				
		0.005	0.01	0.025	0.05	0.10
Degrees of Freedom	99%CI	98%CI	95%CI	90%CI	80%CI	
	0.01	0.02	Area in Two Tails 0.05	0.10	0.20	
1	63.657	31.821	12.706	6.314	3.078	
2	9.925	6.965	4.303	2.920	1.886	
3	5.841	4.541	3.182	2.353	1.638	
4	4.604	3.747	2.776	2.132	1.533	
5	4.032	3.365	2.571	2.015	1.476	
6	3.707	3.143	2.447	1.943	1.440	
7	3.499	2.998	2.365	1.895	1.415	
8	3.355	2.896	2.306	1.860	1.397	
9	3.250	2.821	2.262	1.833	1.383	
10	3.169	2.764	2.228	1.812	1.372	
11	3.106	2.718	2.201	1.796	1.363	
12	3.055	2.681	2.179	1.782	1.356	
13	3.012	2.650	2.160	1.771	1.350	
14	2.977	2.624	2.145	1.761	1.345	
15	2.947	2.602	2.131	1.753	1.341	
16	2.921	2.583	2.120	1.746	1.337	
17	2.898	2.567	2.110	1.740	1.333	
18	2.878	2.552	2.101	1.734	1.330	
19	2.861	2.539	2.093	1.729	1.328	
20	2.845	2.528	2.086	1.725	1.325	
21	2.831	2.518	2.080	1.721	1.323	
22	2.819	2.508	2.074	1.717	1.321	
23	2.807	2.500	2.069	1.714	1.319	
24	2.797	2.492	2.064	1.711	1.318	
25	2.787	2.485	2.060	1.708	1.316	
26	2.779	2.479	2.056	1.706	1.315	
27	2.771	2.473	2.052	1.703	1.314	
28	2.763	2.467	2.048	1.701	1.313	
29	2.756	2.462	2.045	1.699	1.311	
30	2.750	2.457	2.042	1.697	1.310	
31	2.744	2.453	2.040	1.696	1.309	
32	2.738	2.449	2.037	1.694	1.309	
33	2.733	2.445	2.035	1.692	1.308	
34	2.728	2.441	2.032	1.691	1.307	
35	2.724	2.438	2.030	1.690	1.306	
36	2.719	2.434	2.028	1.688	1.306	
37	2.715	2.431	2.026	1.687	1.305	
38	2.712	2.429	2.024	1.686	1.304	
39	2.708	2.426	2.023	1.685	1.304	
40	2.704	2.423	2.021	1.684	1.303	
45	2.690	2.412	2.014	1.679	1.301	
50	2.678	2.403	2.009	1.676	1.299	
60	2.660	2.390	2.000	1.671	1.296	
70	2.648	2.381	1.994	1.667	1.294	
80	2.639	2.374	1.990	1.664	1.292	
90	2.632	2.368	1.987	1.662	1.291	
100	2.626	2.364	1.984	1.660	1.290	
200	2.601	2.345	1.972	1.653	1.286	
300	2.592	2.339	1.968	1.650	1.284	
400	2.588	2.336	1.966	1.649	1.284	
500	2.586	2.334	1.965	1.648	1.283	
1000	2.581	2.330	1.962	1.646	1.282	
2000	2.578	2.328	1.961	1.646	1.282	
Large	2.576	2.326	1.960	1.645	1.282	

large n
t becomes z

99% 98% 95% 90% 80%