Chapter 8: Confidence Interval

Dr. Siddiqua Mazhar

Mid Michigan College

Nov 4, 2021

Introduction

- Inferential Statistics
- Results on Sample $\rightarrow$ Results on Population

- Results on Sample $\rightarrow$ Results on Population
- Our results will be Averages or Proportions

Confidence Intervals for Population Mean

Estimating Population Mean from a Sample Mean

- Random Sample
- Population Standard Deviation, σ, is known
- $n \geq 30$ Or, Population is normally distributed

Point Estimate

$\overline{X}$ is the point estimate for μ

Example The sample of 30 students in Mid Michigan College has mean height is 5 ft 2 inches, so we estimate the average height of a student in Mid Michigan 5 ft 2 inches.

Computing bounds of Margin of Error

The maximum difference between $\overline{X}$ and μ or magnitude of $\overline{X} - \mu$ is called margin of error.
Distribution that is approximately normal with mean

$$\mu_{\overline{X}} = \mu \quad \textit{and} \quad \textit{standard error}, \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

$$E = Z_{\frac{\alpha}{2}}.\frac{\sigma}{\sqrt{n}}$$

(1)

$$P(-z_{\frac{\alpha}{2}} < z < z_{\frac{\alpha}{2}}) = C.L.$$

(2)
Using the formula

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Substituting this expression for z score into Equation (1) gives

$$P(-z_{\frac{\alpha}{2}} < \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < z_{\frac{\alpha}{2}}) = C.L.$$

Multiplying all parts of the above inequality by $\sigma/\sqrt{n}$ gives us

$$P(-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}) = C.L.$$

We conclude that

$$P(-E < \overline{X} - \mu < E) = c$$

For μ we can rewrite it in the following

$$\overline{X} - E < \mu < \overline{X} + E$$

Confidence Interval

Confidence Interval is a range of numbers where we expect the population mean to be.

Confidence Level

Confidence Level is a level of how confident we are that the confidence interval actually contains the population parameter.

$$1 - \alpha$$

where α is the complement of the confidence level.

Most common:

- .90 or 90% $\rightarrow \alpha = .10$
- .95 or 95% $\rightarrow \alpha = .05$
- .99 or 99% $\rightarrow \alpha = .01$

Critical value:

A z-score is the number such that area under the standard normal curve between $-z_{\alpha/2}$ and $z_{\alpha/2}$ equals the confidence level. Alternatively, a z-score that seperates the likely region from the unlikely region.

Confidence Level	$\propto$ Chance of being wrong	$\frac{\propto}{2}$ Divide alpha by 2	$z_{\propto/2}$ $\left\lvert invNorm(\frac{\propto}{2}, 0, 1)\right\rvert$
99%	.01	.005	2.58
95%	.05	.025	1.96
92%	.08	.04	1.75
90%	.10	.05	1.64
84%	.16	.08	1.41
80%	.20	.10	1.28
50%	.50	.25	.67

Steps:

- Check requirements
- Use Confidence Level to find $z_{\frac{\alpha}{2}}$ (Critical Value)
- Find marginal Error $E = z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}}$
- Find Confidence Interval: $\overline{X} - E < \mu < \overline{X} + E$

Confidence Interval for Ppopulation Mean (Z-Interval)

On Calculator Press the STAT key, select Tests

Choose option 7: ZInterval

Inpt:Data Stats (Select Stats)

$\sigma =$ populationn standard deviation

$\overline{x} =$ Sample mean

n= Sample Size

C-Level: Confidence Level
Press Calculate

Example 1

Random sample of 40 students. The average resting heart-rate for the sample was 76.3 bpm. Assume the population standard deviation is 12.5 bpm. Construct a 99% confidence interval for the average resting heart-rate of the population.

Example 2

A study of 35 golfers showed that their average score on a particular course was 92. The population standard deviation we know to be 5.

- What is the point estimate in this scenario?
- Find and interpret the 95% confidence interval of the mean score of all golfers.

Understanding Results

- Saying the AVERAGE of all individuals will be between the numbers.
- Not saying that all individuals fall between the numbers.

Example 3

Assume we have a sample of 50 U.S. voters. The sample has an average IQ level of 98. We know that the population standard deviation in IQ tests is about 15.
Find and interpret the 95% confience interval of the mean IQ level for all voters.

Further Questions

- If someone said that the average IQ of voters is below 85, would that be believable?
- If someone said that they know a voter with an IQ below 85, would that be believable?

Estimating μ, σ not known, (T - Interval)

If you don't know σ, you can't use a Z-score.

Instead, we use a T-score.

- Random sample
- $n \geq 30$ or sample is from a population that is normally distributed.

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

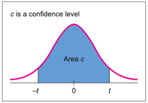
$$T = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}}$$

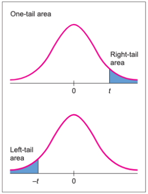
Critical Values and Degrees of Freedom

Critical values are given by $T_{\alpha/2}$

Degrees of Freedom is $n-1$

Critical Values and Degrees of Freedom





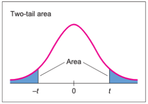


TABLE 6 Critical Values for Student's t Distribution

one-tail area	0.250	0.125	0.100	0.075	0.050	0.025	0.010	0.005	0.0005
two-tail area	0.500	0.250	0.200	0.150	0.100	0.050	0.020	0.010	0.0010
d.f. \ c	0.500	0.750	0.800	0.850	0.900	0.950	0.980	0.990	0.999
1	1.000	2.414	3.078	4.165	6.314	12.706	31.821	63.657	636.619
2	0.816	1.604	1.886	2.282	2.920	4.303	6.965	9.925	31.599
3	0.765	1.423	1.638	1.924	2.353	3.182	4.541	5.841	12.924
4	0.741	1.344	1.533	1.778	2.132	2.776	3.747	4.604	8.610
5	0.727	1.301	1.476	1.699	2.015	2.571	3.365	4.032	6.869
6	0.718	1.273	1.440	1.650	1.943	2.447	3.143	3.707	5.959
7	0.711	1.254	1.415	1.617	1.895	2.365	2.998	3.499	5.408
8	0.706	1.240	1.397	1.592	1.860	2.306	2.896	3.355	5.041
9	0.703	1.230	1.383	1.574	1.833	2.262	2.821	3.250	4.781
10	0.700	1.221	1.372	1.559	1.812	2.228	2.764	3.169	4.587
11	0.697	1.214	1.363	1.548	1.796	2.201	2.718	3.106	4.437
12	0.695	1.209	1.356	1.538	1.782	2.179	2.681	3.055	4.318
13	0.694	1.204	1.350	1.530	1.771	2.160	2.650	3.012	4.221
14	0.692	1.200	1.345	1.523	1.761	2.145	2.624	2.977	4.140
15	0.691	1.197	1.341	1.517	1.753	2.131	2.602	2.947	4.073
16	0.690	1.194	1.337	1.512	1.746	2.120	2.583	2.921	4.015
17	0.689	1.191	1.333	1.508	1.740	2.110	2.567	2.898	3.965
18	0.688	1.189	1.330	1.504	1.734	2.101	2.552	2.878	3.922
19	0.688	1.187	1.328	1.500	1.729	2.093	2.539	2.861	3.883
20	0.687	1.185	1.325	1.497	1.725	2.086	2.528	2.845	3.850
21	0.686	1.183	1.323	1.494	1.721	2.080	2.518	2.831	3.819
22	0.686	1.182	1.321	1.492	1.717	2.074	2.508	2.819	3.792
23	0.685	1.180	1.319	1.489	1.714	2.069	2.500	2.807	3.768
24	0.685	1.179	1.318	1.487	1.711	2.064	2.492	2.797	3.745
25	0.684	1.198	1.316	1.485	1.708	2.060	2.485	2.787	3.725
26	0.684	1.177	1.315	1.483	1.706	2.056	2.479	2.779	3.707
27	0.684	1.176	1.314	1.482	1.703	2.052	2.473	2.771	3.690
28	0.683	1.175	1.313	1.480	1.701	2.048	2.467	2.763	3.674
29	0.683	1.174	1.311	1.479	1.699	2.045	2.462	2.756	3.659
30	0.683	1.173	1.310	1.477	1.697	2.042	2.457	2.750	3.646
35	0.682	1.170	1.306	1.472	1.690	2.030	2.438	2.724	3.591
40	0.681	1.167	1.303	1.468	1.684	2.021	2.423	2.704	3.551
45	0.680	1.165	1.301	1.465	1.679	2.014	2.412	2.690	3.520
50	0.679	1.164	1.299	1.462	1.676	2.009	2.403	2.678	3.496
60	0.679	1.162	1.296	1.458	1.671	2.000	2.390	2.660	3.460
70	0.678	1.160	1.294	1.456	1.667	1.994	2.381	2.648	3.435
80	0.678	1.159	1.292	1.453	1.664	1.990	2.374	2.639	3.416
100	0.677	1.157	1.290	1.451	1.660	1.984	2.364	2.626	3.390
500	0.675	1.152	1.283	1.442	1.648	1.965	2.334	2.586	3.310
1000	0.675	1.151	1.282	1.441	1.646	1.962	2.330	2.581	3.300
∞	0.674	1.150	1.282	1.440	1.645	1.960	2.326	2.576	3.291

For degrees of freedom d.f. not in the table, use the closest d.f. that is *smaller*.

Example

Sample 23 from a normal distribution population. Find critical value ($T_{\frac{\alpha}{2}}$) for a confidence level 95%

Margin of Error: E

$$E = T_{\frac{\alpha}{2}} . \frac{s}{\sqrt{n}}$$

Confidence Intetvals

$$\overline{X} - E < \mu < \overline{X} + E$$

Using Calculator (T Interval)

On Calculator Press the STAT key, select Tests

Choose option 8: TInterval

Inpt:Data Stats (Select Stats)

$\overline{x}$= Sample mean

S_x= sample standard deviation

n= Sample Size

C-Level: Confidence Level
Press Calculate

Example 1

Construct a 95% confidence interval for the average age of people denied a promotion. In a random sample of 23 people, the average age was 47.0 with a standard deviation of 7.2. Assume this sample comes from a population that is normally distributed.

Example 2

According to an internet post, a sample of 15 college havyweight wrestlers has an average weight of 255.88 lbs with a sample standard deviation of 19.29 lbs. Find the 90% confidence interval of the true mean weight of all college heavyweight wrestlers.

Example 2

- If a coach claimed that on average the weight of heavyweight wrestlers was 280 pounds, would the claim be believable?
- If a coach claimed that this wrestler was 280 pounds, would that be believable?

Example 3

The number of deaths by all rifles per year from 2012 to 2016 is given below. Treating these data as a random sample, construct a 98% confidence interval of the average number of deaths by rifles per year.

Rifle Deaths	298	285	258	258	374

Finding Sample Size

We said the as you get bigger samples, you are more likely to get samples that have averages close to the population average.

Sample Size

- Want to find sample size that gives us a certain range.
- Use Z interval for means
- Use Proportion interval for proportions

Sample Size

- Error for Means

$$E = Z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}}$$

- Error for proportions

$$E = Z_{\frac{\alpha}{2}} . \frac{\sqrt{\hat{p}\hat{q}}}{\sqrt{n}}$$

Example 1

You wish to find the average height of all people from Mexico. You want to make sure you are accurate to within 0.5 inches of the real mean height with a 95% confidence interval. You know that the standard deviation is 3 in. How large should the sample be?

Example 2

A researcher is interested in estimating the average monthly salary of sports reporters in a large city. He wants to be 90% confident that his estimate is correct. If the standard deviation is \$1100, how large a sample is needed to get the desired information and to be accurate to within \$150.

Example 3

A medical researcher wishes to determine the percentage of females who take vitamins. He wishes to be 99% confident that the estimate is within 2 percentage points of the true proportion. A recent study of 180 females showed that 25% took vitamins. How large should my sample size be?