8.1 Confidence Intervals for the Mean (σ Known or $n \geq 30$)

The average age of college students is on the rise.

An actuary of a liability insurance company knows that a change in the average age of the population of a college campus changes the likelihood of the occurrence of certain age related illnesses and accidents on that campus. Thus, the actuary would like to be fairly certain of the current mean age of a particular campus population in order to establish liability insurance costs. The insurance company wants to know with 95% certainty an estimate of the mean student age with a margin of error no greater than ±1 year.

The actuary collected the following data set of the ages of **30** students on campus:

17	18	18	26	33	41	60	19	18	55
48	25	30	24	17	21	42	18	35	36
16	17	26	20	19	54	71	40	30	20

The mean of this data set is $\bar{X}$ = **30.47**. The actuary also knows that the standard deviation σ for the campus population is 2 years.

Can the actuary be at least 95% certain that μ is within ±1 year of 30.47?

By determining the 95% **confidence interval for the mean** μ, that is, an interval of values that has a probability of 95% of containing the mean μ, using the sample mean 30.47 and standard deviation 2, we will be able to answer this question.
First we will introduce some vocabulary.

First we will introduce some new vocabulary.

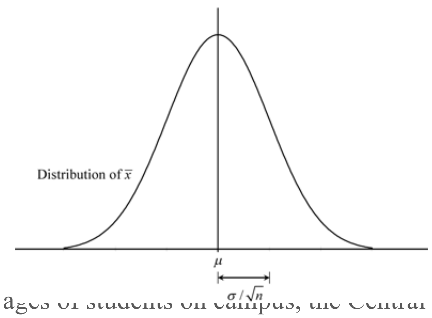
A *parameter* is a value used to represent a certain population characteristic.
In our example, the actuary wants to know the parameter, μ, of the campus population.

A *point estimate* is the value that is obtained from a sample of data and used to estimate the value of a parameter.
In our example, $\bar{X} = 30.47$ is the point estimate, obtained from the data of 30 selected students, and is used to estimate the parameter, μ, of the entire campus population.

A *confidence interval* is a specific interval of values within which the parameter lies determined by using data obtained from a sample given a specific confidence level.
We will determine an interval of values, hopefully no greater than $\bar{X} \pm 1$, within which the parameter, μ, lies by using the point estimate of $\bar{X} = 30.47$ and σ = 2, with 95% confidence.

Recall:
Central Limit Theorem:

As the sample size n increases without limit, the shape of the distribution of the sample means taken from a population will approach a normal distribution with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$.

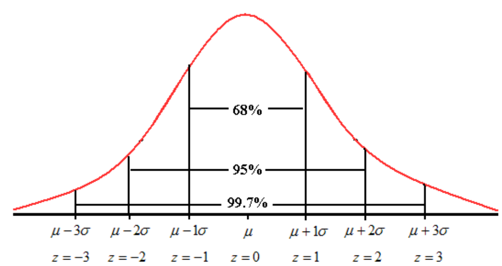


In our example, if many sample means of size 30 were taken of the ages of students on campus, the Central Limit Theorem assures that the distribution of these means would approach a normal distribution with mean μ and standard deviation $\frac{2}{\sqrt{30}}$.

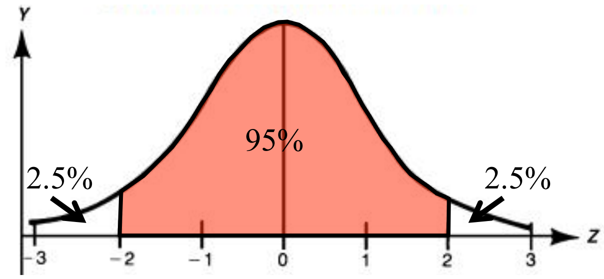
Empirical Rule:

In a normal distribution with mean μ, approximately

- **68%** of the data will fall within 1 standard deviation of μ.
- **95%** of the data will fall within 2 standard deviations of μ.
- **99.7%** of the data will fall within 3 standard deviations of μ.



In our example, since the distribution of the sample means approaches a normal distribution, then it is with 95% certainty that the actuary's sample mean $\bar{X}$ = 30.47 will fall within 2 standard deviations of the population mean μ.



That is

$$P(\mu - 2\left(\frac{\sigma}{\sqrt{n}}\right) < \bar{X} < \mu + 2\left(\frac{\sigma}{\sqrt{n}}\right)) = 0.95$$

From this we have

$$P(\bar{X} - 2\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{X} + 2\left(\frac{\sigma}{\sqrt{n}}\right)) = 0.95$$

Substituting 30.47 for $\bar{X}$, 30 for n, and 2 for σ, we have

$$P(30.47 - 2\left(\frac{2}{\sqrt{30}}\right) < \mu < 30.47 + 2\left(\frac{2}{\sqrt{30}}\right)) = 0.95$$

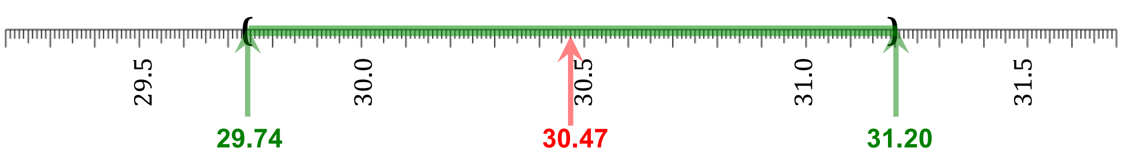
$$P(30.47 - 0.73 < \mu < 30.47 + 0.73) = 0.95$$

$$P(29.74 < \mu < 31.20) = 0.95$$

This gives us that the 95% confidence interval for the mean is **(29.74, 31.20)**.

Notice that

$$29.47 = 30.47 - 1 < 29.74 < \mu < 31.20 < 30.47 + 1 = 31.47$$



Thus, the actuary can report with 95% certainty that μ is within ±1 year of 30.47.

A 95% confidence level implies that the population mean will fall outside the confidence interval in only 5% of all possible samples of size 30. We assign the greek letter α to this 5%. We can find confidence levels other than those associated with the empirical rule by finding z-scores associated with particular α values. Notice that for

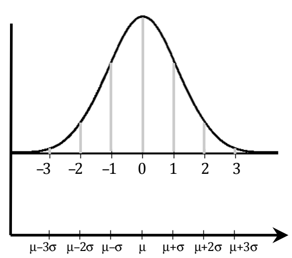
$$\alpha = 0.05,$$

$$1 - \alpha = 0.95,$$

the desired confidence level.

The actuary was also curious about the 90% confidence interval for the mean. She wanted to compare this with the 95% confidence interval. $\overline{X} = 30.47,\ \sigma = 2,\ n = 30$

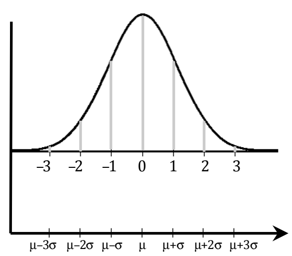
Here are the steps she followed:

<table>
<tr><th>Step 1</th><th>Step 2</th></tr>
<tr><td>$\frac{\sigma}{\sqrt{n}} =$ CL =</td><td>Calculate the error bound for the mean:
$z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$</td></tr>
<tr><td rowspan="2">Shade approximately 90% of the area under the curve centered about the mean and label the x-axis with the error bounds.
</td><th>Step 3</th></tr>
<tr><td>Find the confidence interval such that
$P(\overline{X} - z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)) = \text{CL}$

The 90% confidence interval for the mean is

(,   )</td></tr>
</table>

Can the actuary achieve a 99.7% confidence level within 1 year of accuracy? Repeat the steps from above. $\overline{X} = 30.47,\ \sigma = 2,\ n = 30$

Step 1	Step 2
$\dfrac{\sigma}{\sqrt{n}} =$ CL =	Calculate the error bound for the mean: $z_{\frac{\alpha}{2}}\left(\dfrac{\sigma}{\sqrt{n}}\right)$
Shade approximately 99.7% of the area under the curve centered about the mean and label the *x*-axis with the error bounds. 	**Step 3** Find the confidence interval such that $P(\overline{X} - z_{\frac{\alpha}{2}}\left(\dfrac{\sigma}{\sqrt{n}}\right) < \mu < \overline{X} + z_{\frac{\alpha}{2}}\left(\dfrac{\sigma}{\sqrt{n}}\right)) = \text{CL}$ The 99.7% confidence interval for the mean is (,)

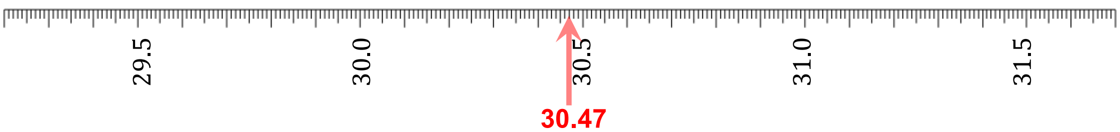
Compare the results from the 3 calculations of confidence intervals

90% confidence interval

$$P\left(30.47 - 1.65\left(\frac{2}{\sqrt{30}}\right) < \mu < 30.47 + 1.65\left(\frac{2}{\sqrt{30}}\right)\right) = 0.90$$

$$P(30.47 - 0.60 < \mu < 30.47 + 0.60) = 0.90$$

$$P(29.87 < \mu < 31.07) = 0.90$$

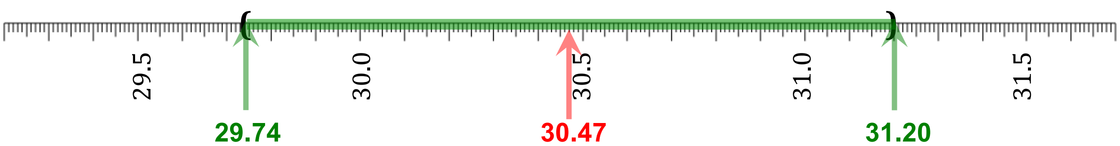


95% confidence interval

$$P\left(30.47 - 2\left(\frac{2}{\sqrt{30}}\right) < \mu < 30.47 + 2\left(\frac{2}{\sqrt{30}}\right)\right) = 0.95$$

$$P(30.47 - 0.73 < \mu < 30.47 + 0.73) = 0.95$$

$$P(29.74 < \mu < 31.20) = 0.95$$

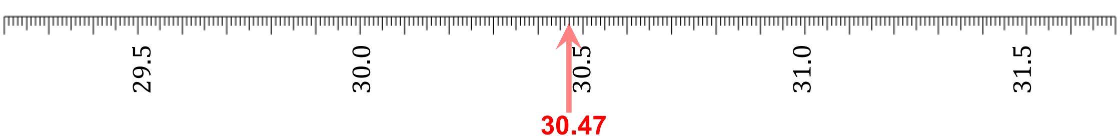


99.7% confidence interval

$$P\left(30.47 - 2.97\left(\frac{2}{\sqrt{30}}\right) < \mu < 30.47 + 2.97\left(\frac{2}{\sqrt{30}}\right)\right) = 0.997$$

$$P(30.47 - 1.08 < \mu < 30.47 + 1.08) = 0.997$$

$$P(29.39 < \mu < 31.55) = 0.997$$



Notice that as $z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$ get larger, the interval gets larger, that is, the margin of error about the sample mean becomes larger. We call $z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$ the error bound of the mean.

We can specify the maximum error of the estimate. In the example, the president specified the maximum allowable error by asking for the average age of college students within 1 year of accuracy. Let's find the sample size required to calculate a 99.7% confidence interval with a margin of error no greater than 1.

That is, we want

$$P(\,30.47 - 1 < \mu < 30.47 + 1\,) = 0.997$$

And since

$$P(\,30.47 - z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < 30.47 + z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)\,) = 0.997$$

then, we want

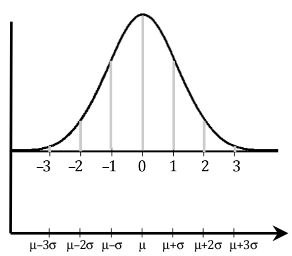
$$1 = z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$$

For a 99.7% confidence interval, we can use the values 2.97 for $z_{\frac{\alpha}{2}}$ and 2 for σ and solve the resulting equation for *n*.

$$\begin{aligned}
1 &= 2.97\left(\frac{2}{\sqrt{n}}\right) \\
\sqrt{n} &= 2.97\left(\frac{2}{\sqrt{n}}\right)\sqrt{n} \\
\sqrt{n} &= 2.97\ (2) \\
\sqrt{n} &= 5.94 \\
\left(\sqrt{n}\right)^2 &= (5.94)^2 \\
n &= 35.28
\end{aligned}$$

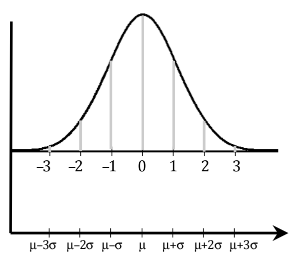
Additional Practice

1. Fuel Efficiency of Cars and Trucks Since 1975 the average fuel efficiency of U. S. cars and light trucks (SUVs) has increased from 13.5 to 24.1 mpg, an increase of over 75%! A random sample of 40 cars from a large community got a mean mileage of 27.2 mpg per vehicle. The population standard deviation is 4.7 mpg per vehicle. Estimate the true mean gas mileage with 95% confidence.

<table>
<tr><th>Step 1</th><th>Step 2</th></tr>
<tr><td>$n =$ CL $=$
$\frac{\sigma}{\sqrt{n}} =$ $\alpha =$</td><td>Calculate the error bound for the mean:
$z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$</td></tr>
<tr><td>Shade approximately 95% of the area under the curve centered about the mean and label the x-axis with the error bounds.
</td><td>Step 3
Find the confidence interval such that
$P(\bar{X} - z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{X} + z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)) = \text{CL}$
The 95% confidence interval for the mean is
(,)</td></tr>
</table>

2. Day Care Tuition A random sample of 50 day care centers provided an average yearly tuition of \$3987 with a population standard deviation of \$630. Find the 90% confidence level interval of the true mean.

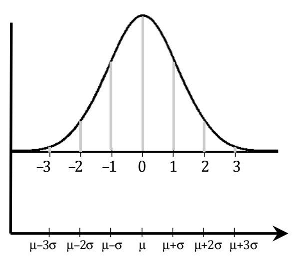
(a)

<table>
<tr><th>Step 1</th><th>Step 2</th></tr>
<tr><td>$\frac{\sigma}{\sqrt{n}}$ = CL =</td><td>Calculate the error bound for the mean:
$z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$</td></tr>
<tr><td rowspan="2">Shade approximately 90% of the area under the curve centered about the mean and label the x-axis with the error bounds.

</td><th>Step 3</th></tr>
<tr><td>Find the confidence interval such that
$P(\bar{X} - z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{X} + z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)) = \text{CL}$

The 90% confidence level interval for the mean is
(,)</td></tr>
</table>

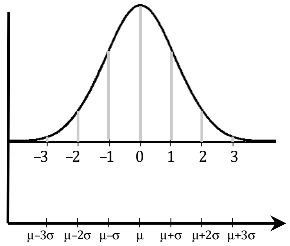
(b) If a day care center were starting up and wanted to keep tuition low, what would be a reasonable amount to charge?

3. Length of Growing Season The growing seasons for a random sample of 35 U. S. cities were recorded, yielding a sample mean of 190.7 days and the population standard deviation of 54.2 days. Estimate for all U. S. cities the true mean of the growing season with 95% confidence.

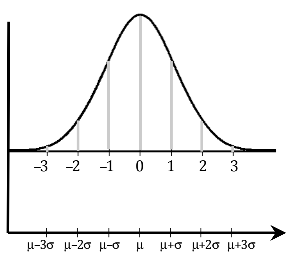
<table>
<tr><th>Step 1</th><th>Step 2</th></tr>
<tr><td>$\frac{\sigma}{\sqrt{n}} =$ CL =</td><td>Calculate the error bound for the mean:
$z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$</td></tr>
<tr><td>Shade approximately 95% of the area under the curve centered about the mean and label the x-axis with the error bounds.

</td><td>Step 3
Find the confidence interval such that
$P(\bar{X} - z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{X} + z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)) = \text{CL}$

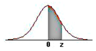
The 95% confidence level interval for the mean is
(,)</td></tr>
</table>

4. Monthly Gasoline Expenditures How large a sample is needed to estimate the population mean monthly gasoline expenditure within \$10 with 95% confidence? The population standard deviation is \$59.50.

Step 1	Step 2	Step 3
Shade approximately 95% of the area under the curve centered about the mean.	Use the z-table or a statistics calculator to find the corresponding z-values	Use $E = z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$ to find the minimum sample size needed.
		

5. Birth Weights of Infants A health care professional wishes to estimate the birth weights of infants. How large a sample must be obtained if she desires to be 90% confident that the true mean is within 2 ounces of the sample mean?
Assume $\sigma = 8$ ounces.

Step 1	Step 2	Step 3
Shade approximately 90% of the area under the curve centered about the mean.	Use the z-table or a statistics calculator to find the corresponding z-values	Use $E = z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$ to find the minimum sample size needed.
		



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0.1	0.03983	0.04380	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
0.2	0.07926	0.08317	0.08706	0.09095	0.09483	0.09871	0.10257	0.10642	0.11026	0.11409
0.3	0.11791	0.12172	0.12552	0.12930	0.13307	0.13683	0.14058	0.14431	0.14803	0.15173
0.4	0.15542	0.15910	0.16276	0.16640	0.17003	0.17364	0.17724	0.18082	0.18439	0.18793
0.5	0.19146	0.19497	0.19847	0.20194	0.20540	0.20884	0.21226	0.21566	0.21904	0.22240
0.6	0.22575	0.22907	0.23237	0.23565	0.23891	0.24215	0.24537	0.24857	0.25175	0.25490
0.7	0.25804	0.26115	0.26424	0.26730	0.27035	0.27337	0.27637	0.27935	0.28230	0.28524
0.8	0.28814	0.29103	0.29389	0.29673	0.29955	0.30234	0.30511	0.30785	0.31057	0.31327
0.9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398	0.33646	0.33891
1.0	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214
1.1	0.36433	0.36650	0.36864	0.37076	0.37286	0.37493	0.37698	0.37900	0.38100	0.38298
1.2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	0.39973	0.40147
1.3	0.40320	0.40490	0.40658	0.40824	0.40988	0.41149	0.41308	0.41466	0.41621	0.41774
1.4	0.41924	0.42073	0.42220	0.42364	0.42507	0.42647	0.42785	0.42922	0.43056	0.43189
1.5	0.43319	0.43448	0.43574	0.43699	0.43822	0.43943	0.44062	0.44179	0.44295	0.44408
1.6	0.44520	0.44630	0.44738	0.44845	0.44950	0.45053	0.45154	0.45254	0.45352	0.45449
1.7	0.45543	0.45637	0.45728	0.45818	0.45907	0.45994	0.46080	0.46164	0.46246	0.46327
1.8	0.46407	0.46485	0.46562	0.46638	0.46712	0.46784	0.46856	0.46926	0.46995	0.47062
1.9	0.47128	0.47193	0.47257	0.47320	0.47381	0.47441	0.47500	0.47558	0.47615	0.47670
2.0	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.48030	0.48077	0.48124	0.48169
2.1	0.48214	0.48257	0.48300	0.48341	0.48382	0.48422	0.48461	0.48500	0.48537	0.48574
2.2	0.48610	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.48840	0.48870	0.48899
2.3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2.6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736
2.8	0.49744	0.49752	0.49760	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807
2.9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861
3.0	0.49865	0.49869	0.49874	0.49878	0.49882	0.49886	0.49889	0.49893	0.49896	0.49900