Topic 17 - Confidence Intervals

Statistics for Managers

June 3, 1999

Statistical inference concerning parameters of a (normal) population.

- **Estimation**
- **Hypothesis Testing**
- **Estimation of the Mean μ of a Normal Population N(μ,σ)**
- **Point Estimate of μ**

A statistic for estimating the mean of a normal population is the sample mean:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

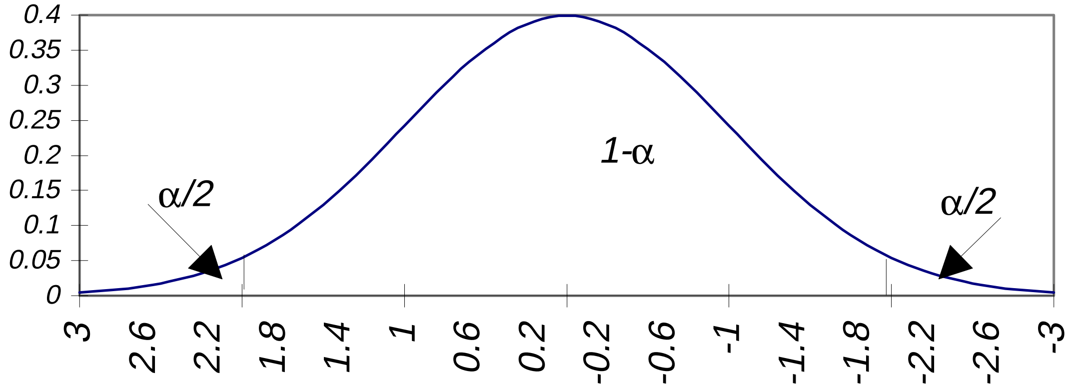
The sampling distribution of the mean is normal:

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Central Limit Theorem: Even if the population is not normal, the sampling distribution is normal provided the sample size <u>*n* is larger than 30</u>:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$$



Scale of *z*:

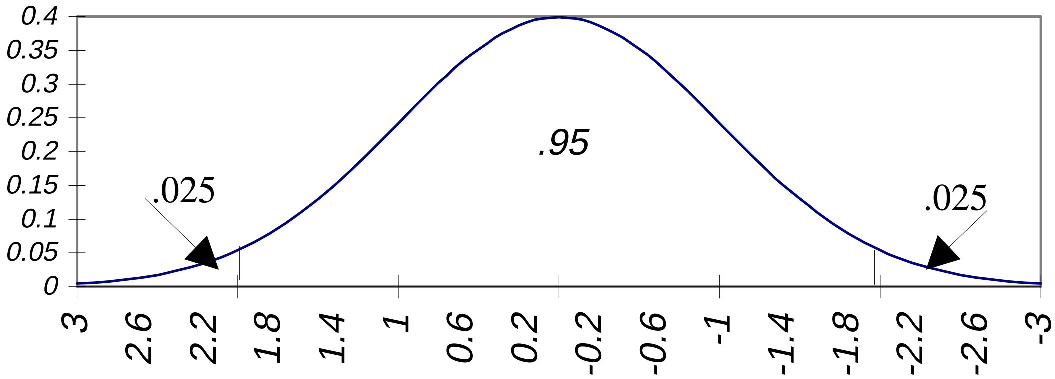
$-z_{\alpha/2}$ $\qquad$ 0 $\qquad$ $+z_{\alpha/2}$

Scale of $\overline{X}$:

$\mu - z_{\alpha/2}\sigma\big/\sqrt{n}$ $\qquad$ μ $\qquad$ $\mu + z_{\alpha/2}\sigma\big/\sqrt{n}$

Example

In what interval will the mean of a sample of 25 observations from a N(2,4) population lie 95% of the time in repeated sampling?



Scale of *z*:

$-z_{.025} = -1.96$ $\quad$ 0 $\quad$ $+z_{.025} = 1.96$

Scale of $\overline{X}$:

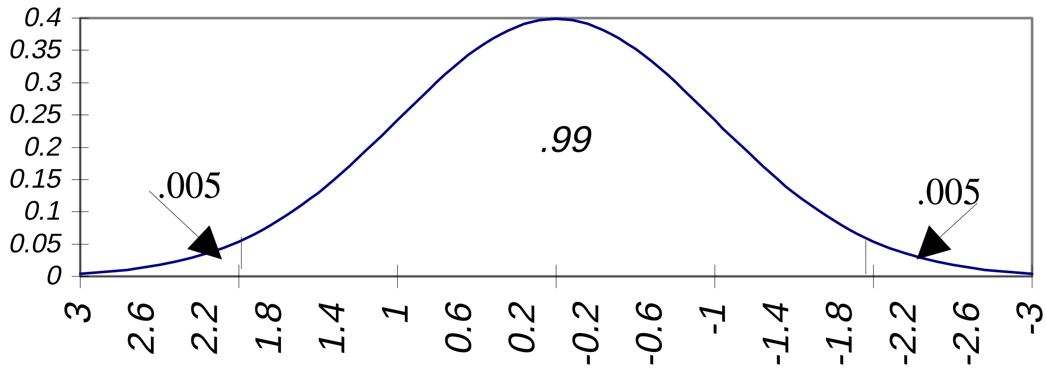
$2 - 1.96 \cdot 4 / \sqrt{25} = .4320$ $\quad$ **2** $\quad$ $2 + 1.96 \cdot 4 / \sqrt{25} = 3.5680$

Answer:

The sample mean will fall in the interval (.432, 3.568) in 95% of the time during repeated sampling.

Example:

In how many units of the unknown mean with the sample mean lie 99% of the time if the sample size is 9 and the population is N(μ,1.5)



Scale of *z*:

$-z_{.005} = -2.576$ 0 $z_{.005} = 2.576$

Scale of $\overline{X}$:

$\mu - 2.576 \cdot 1.5/\sqrt{9}$ $\mu + 2.576 \cdot 1.5/\sqrt{9}$

$= \mu - 1.288$ μ $= \mu + 1.288$

Answer:

The mean of a sample of size 9 will fall within 1.288 units of the true mean in 99% of all random samples when sampling from N(μ,1.5).

Confidence Intervals with Known Population SD

The 100(1-α)% confidence interval for the mean of a N(μ,σ) population using a sample of size *n*:

$$\overline{X} - z_{\alpha/2}\sigma\big/\sqrt{n} \le \mu \le \overline{X} + z_{\alpha/2}\sigma\big/\sqrt{n}$$

The formula is valid even when the population is not normal, if the <u>sample size is greater than 30</u>, and the <u>SD is known</u>. (Central Limit Theorem).

Example:

Find a 90% confidence interval for the population mean in sample of size 25 from N(μ,3) using the sample mean equal to 11.

Answer:

90% = 100(1-α)% implies α = .10, so α/2 = .05 and $z_{.05}$ = 1.645

$$11 - 1.645 \cdot 3\big/\sqrt{25} \le \mu \le 11 + 1.645 \cdot 3\big/\sqrt{25}$$
$$10.013 \le \mu \le 11.987$$

Example:

Find a 95% confidence interval for the population mean in sample of size 100 from a population with SD equal to 2 using the sample mean equal to 8.5.

Answer:

95% = 100(1-α)% implies α = .05, so α/2 = .025 and $z_{.025}$ = 1.96

$$8.5 - 1.96 \cdot 2\big/\sqrt{100} \le \mu \le 8.5 + 1.96 \cdot 2\big/\sqrt{100}$$
$$8.108 \le \mu \le 8.892$$

Confidence Intervals with Unknown Population SD

In most situations, one does not know the standard deviation of the population.

In this case, we estimate the population standard deviation with the sample standard deviation.

<u>In the sample size is 30 or larger</u>, the above formulations apply equally well by just substituting in for the unknown population SD.

However, if the sample sizes are less than 30, we have to use a slightly different approach.

Student *t*:

The Students t-statistic for estimating the mean of a normal population is *t*:

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

Where *s* is the sample standard deviation based on size *n*:

$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)}{n-1}} \; .$$

The sampling distribution of the *t* statistic is the *t* distribution:

$$t = \frac{\overline{X} - \mu}{s / \sqrt{n}} \sim t(n-1) \; .$$

Where *t*(*n*-1) is the Student t distribution with *n*-1 degrees of freedom.

t Table

	α/2			
df	*0.050*	*0.025*	*0.100*	*0.005*
1	6.314	12.706	3.078	63.656
2	2.920	4.303	1.886	9.925
3	2.353	3.182	1.638	5.841
4	2.132	2.776	1.533	4.604
5	2.015	2.571	1.476	4.032
6	1.943	2.447	1.440	3.707
7	1.895	2.365	1.415	3.499
8	1.860	2.306	1.397	3.355
9	1.833	2.262	1.383	3.250
10	1.812	2.228	1.372	3.169
11	1.796	2.201	1.363	3.106
12	1.782	2.179	1.356	3.055
13	1.771	2.160	1.350	3.012
14	1.761	2.145	1.345	2.977
15	1.753	2.131	1.341	2.947
16	1.746	2.120	1.337	2.921
17	1.740	2.110	1.333	2.898
18	1.734	2.101	1.330	2.878
19	1.729	2.093	1.328	2.861
20	1.725	2.086	1.325	2.845
21	1.721	2.080	1.323	2.831
22	1.717	2.074	1.321	2.819
23	1.714	2.069	1.319	2.807
24	1.711	2.064	1.318	2.797
25	1.708	2.060	1.316	2.787
26	1.706	2.056	1.315	2.779
27	1.703	2.052	1.314	2.771
28	1.701	2.048	1.313	2.763
29	1.699	2.045	1.311	2.756
inf.	1.645	1.960	1.282	2.576

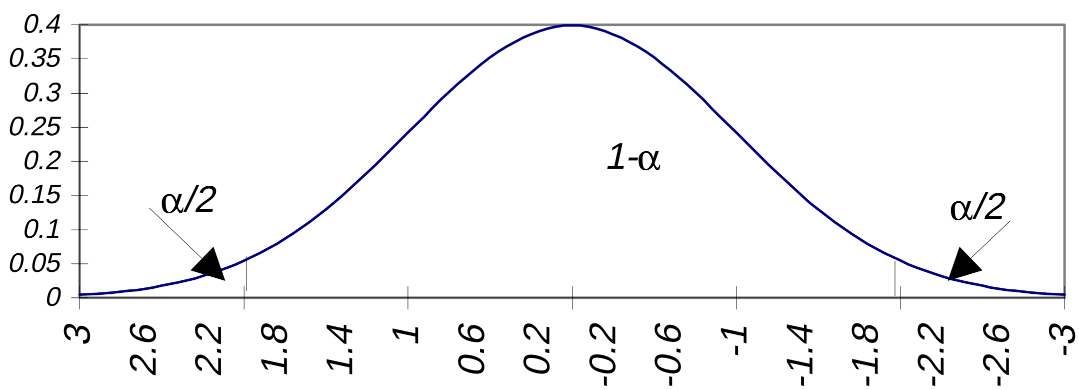
Confidence Interval Formula, Unknown SD:

The 100(1-α)% confidence interval for the mean of a N(μ,σ) population using a sample of size *n*:

$$\bar{X} - t_{\alpha/2} s\big/\sqrt{n} \le \mu \le \bar{X} + t_{\alpha/2} s\big/\sqrt{n}$$

The formula is valid even when the population is not normal, if the <u>sample size is greater than 30</u>. (Central Limit Theorem).

- Table 1



Scale of *t*:

$-t_{\alpha/2}$	0	$+t_{\alpha/2}$

Scale of $\bar{X}$:

$\mu - t_{\alpha/2} s\big/\sqrt{n}$	μ	$\mu + t_{\alpha/2} s\big/\sqrt{n}$

Example:

Find a 90% confidence interval for the population mean in sample of size 25 from N(μ,σ) using the sample mean equal to 11 and SD equal 3.01.

Answer:

90% = 100(1-α)% implies α = .10, so α/2 = .05, *n* = 25 $t_{.05}$ = 1.711 with 24 degrees of freedom

$$11 - 1.711 \cdot 3.01 \big/ \sqrt{25} \le \mu \le 11 - 1.711 \cdot 3.01 \big/ \sqrt{25}$$
$$9.970 \le \mu \le 12.030$$