CONFIDENCE LEVELS MADE SIMPLE

PROCEDURE

(1) Let c be the confidence level percentage written as a decimal.

For example: If the confidence level is 95%, then $c = 0.95$.

(2) Let z_0 be an instance of the standard normal random variable Z such that $P(Z \le z_0) = \dfrac{1+c}{2}$.

For example: If $c = 0.95$, then $P(Z \le z_0) = \dfrac{1+0.95}{2} = \dfrac{1.95}{2} = 0.9750$.

Then use *Integrating.xls* and *Solver* to compute z_0. In this example, $z_0 = 1.96$.

(3) Let X be a random variable of which we have a sample of size n. We can compute the mean, $\bar{x}$, of our sample and the standard deviation, s, of our sample.

(4) The c confidence interval is $\left(\bar{x} - \dfrac{z_0 \cdot s}{\sqrt{n}},\ \bar{x} + \dfrac{z_0 \cdot s}{\sqrt{n}}\right)$.

This means that $P\left(\bar{x} - \dfrac{z_0 \cdot s}{\sqrt{n}} \le \mu_X \le \bar{x} + \dfrac{z_0 \cdot s}{\sqrt{n}}\right) = c$ where μ_X is the true mean of X.

EXAMPLE

We have a sample of size 7 from a random variable X. The mean of our sample is 50 and the standard deviation is 10. Find the 90% confidence interval for μ_X.

Solution:

$c = 0.90$

$P(Z \le z_0) = \dfrac{1+0.90}{2} = \dfrac{1.90}{2} = 0.9500$. Using *Integrating.xls* and *Solver* we find that $z_0 = 1.64$.

Definition	Computation		Plot Interval		Integration Interval		$\int_a^b f(x)\,dx$
Formula for $f(x)$	x	$f(x)$	A	B	a	b	
0.39894	-5	1.5E-06	-5	5	-5	1.64486	0.9500

The confidence interval is $\left(50 - \dfrac{(1.64)(10)}{\sqrt{7}},\ 50 + \dfrac{(1.64)(10)}{\sqrt{7}}\right) = (43.80,\ 56.20)$.

This means that the probability of the true mean of X lying between 43.80 and 56.20 is 90%. So if we take a huge number of such samples, μ_X (the true mean of X) will be within the computed confidence interval close to 90% of the time. That is, the true mean of X will be in approximately 90% of the huge number of samples that we take. [SEE SLIDE # 128.]

DERIVATION OF THE
CONFIDENCE LEVEL

Find an interval about the sample mean that has a 95% probability of containing the true mean.

Compute a number "a" such that

$$P(\bar{x} - a \le \mu_x \le \bar{x} + a) = 0.95$$
$$P(-a \le \mu_x - \bar{x} \le a) = 0.95$$
$$P(a \ge \bar{x} - \mu_x \ge -a) = 0.95$$
$$P(-a \le \bar{x} - \mu_x \le a) = 0.95$$

$$P\left(\frac{-a}{\frac{\sigma_x}{\sqrt{n}}} \le \frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}} \le \frac{a}{\frac{\sigma_x}{\sqrt{n}}}\right) = 0.95$$

$$P\left(\frac{-a\sqrt{n}}{\sigma_x} \le Z \le \frac{a\sqrt{n}}{\sigma_x}\right) = 0.95$$

$$P\left(Z < \frac{-a\sqrt{n}}{\sigma_x} \quad or \quad Z > \frac{a\sqrt{n}}{\sigma_x}\right) = 1 - 0.95$$

$$P\left(Z < \frac{-a\sqrt{n}}{\sigma_x}\right) + P\left(Z > \frac{a\sqrt{n}}{\sigma_x}\right) = 0.05$$

$$2 \cdot P\left(Z > \frac{a\sqrt{n}}{\sigma_x}\right) = 0.05$$

$$P\left(Z > \frac{a\sqrt{n}}{\sigma_x}\right) = 0.025$$

$$P\left(Z \le \frac{a\sqrt{n}}{\sigma_x}\right) = 0.9750$$

$$Let \quad z_0 = \frac{a\sqrt{n}}{\sigma_x}. \quad Now \ find \ z_0 \ so \ that \ P(Z \le z_0) = 0.9750 .$$

Using *integrating.xls*, we integrate the standard normal probability density function with z_0 as the upper limit of integration and with 0.9750 as the solution. This gives $z_0 = 1.96$.

The administrator makes his sample of $n = 50$ and finds that the mean of the sample is $88,989 and that the standard deviation of the sample is $22,358. Since his sample size is fairly large, his sample standard deviation is a good estimate for the true sample standard deviation (σ_x).

So $1.96 = \dfrac{a\sqrt{n}}{\sigma_x}$. *Solving for "a" gives* $a = \dfrac{1.96 \cdot \sigma_x}{\sqrt{n}}$.

Substituting in $P\left(\bar{x} - a \le \mu_x \le \bar{x} + a\right) = 0.95$ *gives*

$$P\left(\bar{x} - \frac{1.96\ \sigma_x}{\sqrt{n}} \le \mu_x \le \bar{x} + \frac{1.96\ \sigma_x}{\sqrt{n}}\right) = 0.95.$$

Substituting the numbers above gives

$$P\left(88989 - \frac{1.96(22358)}{\sqrt{50}} \le \mu_x \le 88989 + \frac{1.96(22358)}{\sqrt{50}}\right) = 0.95$$

$$P\left(82791 \le \mu_x \le 95187\right) = 0.95$$

Hence, the 95% confidence level is the interval ($82791, $95187).