BUS 172
Descriptive Statistics

Lecture 17,18, & 19

Example: Suppose the population proportion of managers who participated in the training program is P = 0.60. With a simple random sample of size 30, the personnel director wants to know the probability of obtaining a value of sample proportion that is within 0.05 of the population proportion; That is, what is the probability of obtaining a sample with a sample proportion between 0.55 and 0.65?

***n* = 30**

***p* = P() = 0.60**

***q* = 0.40**

***np* = 18 *nq* = 12**

$$= P\left(\frac{0.55-0.60}{\sqrt{\frac{0.60(0.40)}{30}}} < z < \frac{0.65-0.60}{\sqrt{\frac{0.60(0.40)}{30}}}\right) = P(-0.56 < z < 0.56)$$

$$= 0.7123 - 0.2877 = 0.4246$$

OK to use the normal approximation

If npq>9 then $\hat{p}\hat{q}$ can replace P and q to give better approximation.

Interval Estimation

✓A point estimator is a sample statistic used to estimate a population parameter. The sample mean $\bar{x}$ is a point estimator of the population mean μ and the sample proportion $\hat{p}$ is a point estimator of the population proportion P.

✓Because a point estimator cannot be expected to provide the exact value of the population parameter, an interval estimate is often computed by adding and subtracting a value, called the margin of error, to the point estimate.

✓The general form of an interval estimate is as follows:

Point estimate ± Margin of error

✓The purpose of an interval estimate is to provide information about how close the point estimate, provided by the sample, is to the value of the population parameter.

Interval Estimation

✓Using the standard normal probability table, we find that 95% of the values of any normally distributed random variable are within 1.96 standard deviations of the mean.

✓Thus, when the sampling distribution of $\bar{x}$ is normally distributed, 95% of the values must be within $\pm 1.96\sigma_{\bar{x}}$ of the mean μ.

✓Thus, when the sampling distribution of $\hat{p}$ is normally distributed, 95% of the values must be within $\pm 1.96\sigma_{\hat{p}}$ of the mean P.

Point estimator of population mean $\mu : \bar{x}$

Margin of error $(n \geq 30) : \pm 1.96 \dfrac{s}{\sqrt{n}}$

•For a binomial population,

Point estimator of population proportion $p : \hat{p} = x/n$

Margin of error $(n \geq 30) : \pm 1.96 \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

Example: A homeowner randomly samples 64 homes similar to her own and finds that the average selling price is \$252,000 with a standard deviation of \$15,000. Estimate the average selling price for all similar homes in the city.

Point estimator of $\mu : \bar{x} = 252{,}000$

Margin of error: $\pm 1.96 \dfrac{s}{\sqrt{n}} = \pm 1.96 \dfrac{15{,}000}{\sqrt{64}} = \pm 3675$

95% confidence interval:

$$\bar{x} \pm 1.96 \sigma_{\bar{x}} = \bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}} = 252000 \pm 1.96 \frac{15000}{\sqrt{64}} = 252000 \pm 3675$$

Example: A quality control technician wants to estimate the proportion of soda cans that are underfilled. He randomly samples 200 cans of soda and finds 10 underfilled cans.

$n = 200$ $p =$ proportion of underfilled cans

Point estimator of p: $\hat{p} = x/n = 10/200 = .05$

Margin of error: $\pm 1.96\sqrt{\dfrac{\hat{p}\hat{q}}{n}} = \pm 1.96\sqrt{\dfrac{(.05)(.95)}{200}} = \pm .03$

95% confidence interval:

$$\hat{p} \pm 1.96\sigma_{\hat{p}} = \hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.05 \pm 1.96\sqrt{\frac{(0.05)(0.95)}{200}} = 0.05 \pm 0.03$$

Confidence interval for population mean:

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

Confidence interval for Population Proportion:

$$\hat{p} \pm z_{\alpha/2}\sigma_{\hat{p}} = \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Where,

$1-\alpha$ = confidence coefficient

α = level of significance

$Z_{\alpha/2}$= Z value providing an area of $\alpha/2$ in the negative tail of the standard normal prob. distribution

95% confidence interval= confidence coefficient, $1-\alpha$ =0.95

i.e. α = 0.05. Hence

$$|z_{\alpha/2}| = |z_{0.05/2}| = |z_{0.025}| = |-1.96| = 1.96$$

Example: A random sample of n = 50 males showed a mean average daily intake of dairy products equal to 756 grams with a standard deviation of 35 grams. Find a 95% confidence interval for the population average μ.

Solution:95% confidence interval= confidence coefficient, 1-α =0.95 i.e. α = 0.05. Hence

$$|z_{\alpha/2}| = |z_{0.05/2}| = |z_{0.025}| = |-1.96| = 1.96$$

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

$$\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} \Rightarrow \quad 756 \pm 1.96\frac{35}{\sqrt{50}} \quad \Rightarrow \quad 756 \pm 9.70$$

$$\text{or } 746.30 < \mu < 765.70 \text{ grams.}$$

Example: A random sample of n = 50 males showed a mean average daily intake of dairy products equal to 756 grams with a standard deviation of 35 grams. Find a 99% confidence interval for the population average μ.

Solution:99% confidence interval= confidence coefficient, 1-α =0.99 i.e. α = 0.01. Hence

$$|z_{\alpha/2}| = |z_{0.01/2}| = |z_{0.005}| = |-2.575| = 2.575$$

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

$$\bar{x} \pm 2.575\frac{\sigma}{\sqrt{n}} \Rightarrow \quad 756 \pm 2.575\frac{35}{\sqrt{50}} \Rightarrow 756 \pm 12.74$$

$$\text{or } 746.26 < \mu < 768.74 \text{ grams.}$$

Example: Of a random sample of n = 150 college students, 104 of the students said that they had played on a soccer team during their K-12 years. Estimate the proportion of college students who played soccer in their youth with a 98% confidence interval.

Solution:98% confidence interval= confidence coefficient, 1-α =0.98 i.e. α = 0.02. Hence

$$|z_{\alpha/2}| = |z_{0.02/2}| = |z_{0.01}| = |-2.33| = 2.33$$

$$\hat{p} \pm z_{\alpha/2}\sigma_{\hat{p}} = \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.69 \pm 2.33\sqrt{\frac{(0.69)(0.31)}{150}} \Rightarrow 0.69 \pm 0.0879$$

$$\text{or } \; 0.602 < P < 0.778 \qquad 60.2\% < P < 77.8\%$$

Confidence interval for population mean(σ unknown)

$$\bar{x} \pm t_{\alpha/2 \atop (n-1)D.F} \sigma_{\bar{x}} = \bar{x} \pm t_{\alpha/2 \atop (n-1)D.F} \frac{s}{\sqrt{n}}$$

Where,

1-α = confidence coefficient, α = level of significance

$t_{\alpha/2}$= t value providing an area of α/2 in t prob. distribution at (n-1) degrees of freedom. As the number of degrees of freedom increases, the difference between the t distribution and the standard normal (z) distribution becomes smaller and smaller.

95% confidence interval= confidence coefficient, 1-α =0.95

i.e. α = 0.05. Hence

$$t_{\alpha/2 \atop (n-1)D.F} = t_{0.05/2 \atop (23)D.F} = t_{0.025 \atop (23)D.F} = 2.069$$

Degrees of freedom refer to the number of independent pieces of information that go into the computation of $\Sigma(x_i - \bar{x})^2$. The n pieces of information involved in computing $\Sigma(x_i - \bar{x})^2$ are as follows: $x_1 - \bar{x}, x_2 - \bar{x}, \ldots, x_n - \bar{x}$. In Section 3.2 we indicated that $\Sigma(x_i - \bar{x}) = 0$ for any data set. Thus, only $n - 1$ of the $x_i - \bar{x}$ values are independent; that is, if we know $n - 1$ of the values, the remaining value can be determined exactly by using the condition that the sum of the $x_i - \bar{x}$ values must be 0. Thus, $n - 1$ is the number of degrees of freedom associated with $\Sigma(x_i - \bar{x})^2$ and hence the number of degrees of freedom for the t distribution in expression (8.2).

Example: For a random sample of seven automobiles radar indicated the following speeds, in miles per hour:

79 73 68 77 86 71 69

S= 6.4, $\bar{x}$ =74.71

Assuming a normal probability distribution for population, find the margin of error and interval at 95% confidence interval for the mean speed of all automobiles.

Solution:95% confidence interval= confidence coefficient, 1-α =0.95 i.e. α = 0.05. Hence

$$t_{\alpha/2 \, (n-1)D.F} = t_{0.05/2 \, (7-1)D.F} = t_{0.025 \, (8)D.F} = 2.306$$

$$\bar{x} \pm t_{\alpha/2 \, (n-1)D.F} \, \sigma_{\bar{x}} = \bar{x} \pm t_{\alpha/2 \, (n-1)D.F} \frac{s}{\sqrt{n}}$$

$$= 74.71 \pm 2.306 \frac{6.4}{\sqrt{7}} = 74.71 \pm 5.58$$

or $69.13 < \mu < 80.29$ miles per hour.