Stat 576
10-3-17

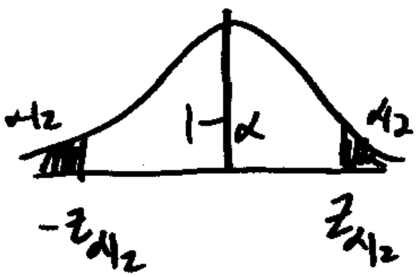
Confidence Intervals

①

SRS WR:

By the Central Limit Theorem,

$$\frac{\bar{y}-\bar{Y}}{\sqrt{V(\bar{y})}} \xrightarrow{\mathcal{D}} N(0,1) \quad \text{as } n\to\infty$$



$$P\Bigg(\underbrace{-z_{\alpha/2} < \frac{\bar{y}-\bar{Y}}{\sqrt{V(\bar{y})}} < z_{\alpha/2}}\Bigg) \approx 1-\alpha$$

$$\underline{\bar{y} - z_{\alpha/2}\sqrt{V(\bar{y})} < \bar{Y} < \bar{y} + z_{\alpha/2}\sqrt{V(\bar{y})}}$$

this is a $(1-\alpha)$100% C.I.

②

$$\underset{\substack{\uparrow \\ \text{point estimator}}}{\bar{y}} \pm \underbrace{z_{\alpha/2}\sqrt{V(\bar{y})}}_{\substack{\uparrow \\ \text{margin of error}}} = \bar{y} \pm z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$$

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

Usually, σ^2 must be estimated by s^2,

resulting in $\boxed{\bar{y} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}}$ for large n

SRSWOR: ③

Hájek (1960) proved an extension to the C.L.T.

For large $n, N, N-n$,

$$\frac{\bar{y}-\bar{Y}}{\sqrt{V(\bar{y})}} \approx N(0,1)$$

Result $\bar{y} \pm z_{\alpha/2}\sqrt{\frac{S^2}{n}\,\text{fpc}}$ is a $(1-\alpha)100\%$ C.I. for $\bar{Y}$

Also s^2 can estimate S^2, so

$$\boxed{\bar{y} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}\sqrt{1-\frac{n}{N}}}$$

Suppose that we specify a margin of error, E. Find the necessary sample size. ④

WR: Set $z_{\alpha/2}\frac{s}{\sqrt{n}} = E$ & solve for n

$$n = \left(\frac{z_{\alpha/2}\, s}{E}\right)^2$$

WOR: Set $z_{\alpha/2}\frac{s}{\sqrt{n}}\sqrt{1-\frac{n}{N}} = E$

$$z^2\frac{s^2}{n}\left(1-\frac{n}{N}\right) = E^2$$

$$\frac{z^2s^2}{n} - \frac{z^2s^2}{N} = E^2$$ (5)

$$z^2s^2 = n\left(\frac{z^2s^2}{N} + E^2\right)$$

$$n = \frac{z^2s^2}{\frac{z^2s^2}{N} + E^2} = \frac{\frac{z^2s^2}{E^2}}{1 + \frac{1}{N}\frac{z^2s^2}{E^2}}$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}},$$ where n_0 is the WR solution

(6)

Example: In a sample of 100 items selected from a population of 1000 items, we find $\bar{y} = 63$ and $s = 10$.

Find a 95% C.I. for $\bar{Y}$

WR: $63 \pm 1.96 \frac{10}{\sqrt{100}} = 63 \pm 1.96$

WoR: $63 \pm 1.96 \frac{10}{\sqrt{100}} \sqrt{1 - \frac{100}{1000}} = 63 \pm 1.86$

Suppose the desired $\mathcal{E}$ is 1

WR: $n = \frac{z^2 S^2}{\mathcal{E}^2} = \frac{1.96^2\, 10^2}{1^2} = 384.16 \rightarrow \underline{385}$

WOR: $n = \frac{n_0}{1 + \frac{n_0}{N}} = \frac{384.16}{1 + \frac{384.16}{1000}} = 277.6 \rightarrow \underline{278}$

<u>Proportions</u> Assume $Y_j = \begin{cases} 1 \\ 0 \end{cases}$ for $j = 1, \dots, N$

Then $\bar{Y} = \frac{1}{N} \sum_{j=1}^{N} Y_j = \frac{\#1\text{s}}{N} = P$ = population proportion

We know that $\bar{y}$ is an unbiased estimator of $\bar{Y}$

$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i = \frac{\#1\text{s}}{n} = \hat{p}$ = sample proportion

$$\sigma^2 = \frac{1}{N} \sum_{j=1}^{N} Y_j^2 - (\bar{Y})^2$$

$$= \frac{1}{N} \sum_{j=1}^{N} Y_j - P^2 = P - P^2 = P(1-P) = PQ$$

$$S^2 = \frac{N}{N-1} \sigma^2 = \frac{N}{N-1} PQ$$

9

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right]$$

$$= \frac{1}{n-1}\left[\underbrace{\sum_{i=1}^{n} y_i}_{\substack{n\bar{y} \\ n\hat{p}}} - n\hat{p}^2\right] = \frac{1}{n-1} n(\hat{p} - \hat{p}^2)$$

$$= \frac{n}{n-1}\hat{p}(1-\hat{p})$$

$$= \frac{n}{n-1}\hat{p}\,\hat{q}$$

C.I. for p, SRSWR

$$\hat{p} \pm z_{\alpha/2} \frac{\sqrt{\frac{n}{n-1}\hat{p}\hat{q}}}{\sqrt{n}}$$

10

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n-1}}$$

C.I. for p, SRSWOR

$$\hat{p} \pm z_{\alpha/2} \frac{\sqrt{\frac{n}{n-1}\hat{p}\hat{q}}}{\sqrt{n}}\sqrt{1-\frac{n}{N}}$$

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n-1}}\sqrt{1-\frac{n}{N}}$$

(11)

Sample size determination

WR: Set $z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n-1}} = E$

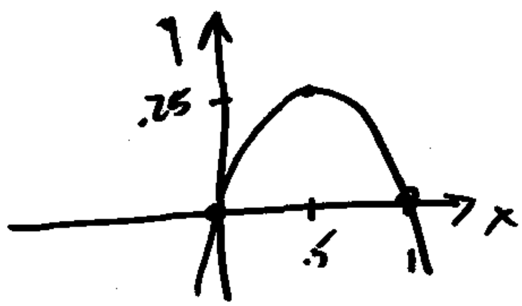
$$n = 1 + \frac{z^2\hat{p}\hat{q}}{E^2} = n_0$$

WOR: Set $z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n-1}}\sqrt{1-\dfrac{n}{N}} = E$

$$n = \frac{n_0}{1+\dfrac{n_0-1}{N}}$$

(12)

$$y = x(1-x) = x - x^2$$



If a pilot study is not feasible, use $\hat{p}\hat{q} = .25$ to overestimate n