Test One

STA304 H1 S/STA1003 HS, FORMULA SHEET, 2014W

INTRODUCTION AND SIMPLE RANDOM SAMPLING

Population mean: $\mu_y = \frac{1}{N}\sum_{i=1}^{N} y(e_i) = \frac{1}{N}\sum_{i=1}^{N} y_i$

Population variance: $\sigma_y^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \mu_y)^2$

$= \frac{1}{N}\sum_{i=1}^{N} y_i^2 - \mu_y^2,\ \tilde{\sigma}_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu_y)^2$

Population total: $\tau_y = \sum_{1}^{N} y(e_i) = \sum_{1}^{N} y_i, \tau_y = N\mu_y,\ \mu_y = \frac{1}{N}\tau_y$

Population proportion: $p = \frac{1}{N}\sum_{1}^{N} y(e_i) = \frac{M}{N}$

$= \frac{\#\{\text{elements with the property}\}}{N} = \mu_y$

Population ratio: $R = \frac{\mu_y}{\mu_x} = \frac{N\mu_y}{N\mu_x} = \frac{\tau_y}{\tau_x} = R_{y|x}$

Sample mean: $\bar{y} = \frac{1}{n}\sum_{1}^{n} y_i$

Sample variance: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n-1}(\sum_{i=1}^{n} y_i^2 - n\bar{y}^2)$

Error of estimation: $|\hat{\theta} - \theta|$ Error bound: $B_\theta = 2 \times \hat{\sigma}(\hat{\theta})$

95% Confidence interval:

$\hat{\theta} \pm B_\theta = \hat{\theta} \pm 2\hat{\sigma}(\hat{\theta}) = [\hat{\theta} - 2\hat{\sigma}(\hat{\theta}), \hat{\theta} + 2\hat{\sigma}(\hat{\theta})]$

Ch. 4. Simple Random Sampling (SRS)

$\binom{N}{n}$ # of all possible samples of size n

Estimators of population mean and total:

$\hat{\mu} = \bar{y} = \frac{1}{n}\sum_{1}^{n} y_i, \hat{\tau} = N\hat{\mu} = N\bar{y} = \frac{N}{n}\sum_{1}^{n} y_i$

$$\hat{\sigma}^2 = \begin{cases} S^2 & \text{, unbiased in SRS with replacement} \\ \tilde{S}^2 = \frac{N-1}{N} S^2 & \text{, unbiased in SRS without replacement} \end{cases}$$

Theoretical var. and estimated var. of sample mean

$Var(\hat{\mu}) = Var(\bar{y}) = \frac{N-n}{N-1}\frac{\sigma_y^2}{n},\ \hat{V}ar(\bar{y}) = \frac{N-n}{N}\frac{S^2}{n}$

Theoretical variance of the estimator for total

$Var(\hat{\tau}) = N^2 Var(\bar{y}) = N^2 \frac{N-n}{N-1}\frac{\sigma_y^2}{n}$

Estimated variance of the estimator for total

$\hat{V}ar(\hat{\tau}) = N^2 \hat{V}ar(\bar{y}) = N^2 \frac{N-n}{N}\frac{S^2}{n} = N(N-n)\frac{S^2}{n}$

Simple random sampling (cont.)

Error bound: $B_\mu = 2\hat{\sigma}_{\hat{\mu}} = 2 \times SD(\hat{\mu}) = 2 \times SD(\bar{y})$

$B_\tau = 2\hat{\sigma}_{\hat{\tau}} = 2 \times \hat{S}D(\hat{\tau}) = 2 \times \hat{S}D(N\bar{y}) = 2N \times \hat{S}D(\bar{y}) = N \times B_\mu$

Confidence interval:

For μ: $\hat{\mu} \pm B_\mu = \bar{y} \pm B_\mu = [\hat{\mu} - B_\mu, \hat{\mu} + B_\mu]$

For τ: $\hat{\tau} \pm B_\tau = N\bar{y} \pm B_\tau = N(\bar{y} \pm B_\mu) = [N(\bar{y} - B_\mu), N(\bar{y} + B_\mu)]$

Selecting the sample size for given cost

$C = C(n) = c_0 + c_1 \times n,\ n = \frac{C - c_0}{c_1} = \frac{C'}{c_1}, C' = C - c_0$

Selecting the sample size for given error bound

$n = \frac{N\tilde{\sigma}^2}{ND + \tilde{\sigma}^2} = \frac{N\sigma^2}{(N-1)D + \sigma^2}, D = D_\mu = \left(\frac{B_\mu}{2}\right)^2, D = D_\tau = \left(\frac{B_\tau}{2N}\right)^2$

Estimating proportion

$\hat{p} = \bar{y} = \frac{1}{n}\sum_{1}^{n} y_i = \frac{a}{n},\ \hat{M} = \hat{\tau} = N\hat{p} = N\frac{a}{n}$

Variance of proportion, theoretica and estimated

$Var(\hat{p}) = \frac{N-n}{N-1}\frac{p(1-p)}{n}, \hat{V}ar(\hat{p}) = \frac{N-n}{N}\frac{\hat{p}(1-\hat{p})}{n-1} = \frac{N-n}{N}\frac{\hat{p}\hat{q}}{n-1}$

Error bound and CI for proportion

$B_p = 2\hat{S}D(\hat{p}) = 2\sqrt{\hat{V}ar(\hat{p})},\ \hat{p} \pm B_p$

Sample size for estimation of p

$n = \frac{N\sigma_y^2}{(N-1)D + \sigma_y^2} = \frac{Npq}{(N-1)D + pq},\ D = D_p = \left(\frac{B_p}{2}\right)^2$

Upper bound on sample size

$n_p \le n_{max} = \frac{N \times 0.25}{(N-1)D + 0.25}$