

SRS = simple random sampling

Stat 576

4-1-25

Population consists of N items

①

Collect a sample of n items

WR = with replacement: same item can be chosen more than once

WOR = without replacement: no item can be chosen more than once

Sampling WOR, there are $\binom{N}{n}$ different possible samples, each of which must have the same probability.

$$\binom{N}{n} = \frac{N!}{n! (N-n)!}$$

②

(RATSTATS)

Notation:

Population: $Y_1, Y_2, \dots, Y_N$

population mean: $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$

population variance: $\sigma^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2$

Sample: $y_1, y_2, \dots, y_n$

sample variance:

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

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

SRSWR

Use $\bar{y}$ to estimate $\bar{Y}$

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

$$= \frac{1}{n} \sum_{i=1}^n E[y_i]$$

$$= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^N y_j \cdot \underbrace{\frac{1}{N}}_{\bar{Y}}$$

$$= \frac{1}{n} n \bar{Y} = \bar{Y} \quad \bar{y} \text{ is an unbiased estimator of } \bar{Y}$$

(3)

$$E[aX + bY]$$

$$= aE[X] + bE[Y]$$

$$V[aX + bY]$$

$$= a^2 V[X] + b^2 V[Y]$$

$$+ 2ab \operatorname{Cov}(X, Y)$$

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

$$= \frac{1}{n^2} \left[\sum_{i=1}^n V[y_i] + 2 \sum_{i < j} \operatorname{Cov}(y_i, y_j) \right]$$

because of independence

$$= \frac{1}{n^2} \sum_{i=1}^n V[y_i]$$

$$= \frac{1}{n^2} \sum_{i=1}^n \left[\underbrace{\sum_{j=1}^N (y_j - \bar{Y})^2 \cdot \frac{1}{N}}_{\sigma^2} \right] = \frac{1}{n^2} \cdot n \sigma^2 = \frac{\sigma^2}{n}$$

(4)

How do we estimate σ^2 ?

(5)

Consider s^2 as a candidate

$$\begin{aligned} E[s^2] &= E\left[\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2\right] \\ &= \frac{1}{n-1} E\left[\sum (y_i^2 - 2y_i\bar{y} + \bar{y}^2)\right] \\ &= \frac{1}{n-1} E\left[\sum y_i^2 - 2\bar{y} \underbrace{\sum y_i}_{n\bar{y}} + n\bar{y}^2\right] \\ &= \frac{1}{n-1} E\left[\sum y_i^2 - n\bar{y}^2\right] \end{aligned}$$

$$\begin{aligned} &= \frac{1}{n-1} \left[\sum_{i=1}^n E[y_i^2] - n E[\bar{y}^2] \right] \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n (V[y_i] + (E[y_i])^2) - n(V[\bar{y}] + (E[\bar{y}])^2) \right] \\ &= \frac{1}{n-1} \left[n(\sigma^2 + \bar{y}^2) - n\left(\frac{\sigma^2}{n} + \bar{y}^2\right) \right] \\ &= \frac{1}{n-1} [(n-1)\sigma^2] = \sigma^2 \end{aligned}$$

Recall (6)

$V[X] = E[X^2] - (E[X])^2$

so $E[X^2] = V[X] + (E[X])^2$

⑦

Results for SRSWR

$$E[\bar{y}] = \bar{Y}$$

$$V[\bar{y}] = \frac{\sigma^2}{n}$$

$$E[s^2] = \sigma^2$$

$$\text{So } \hat{V}[\bar{y}] = \frac{s^2}{n}$$