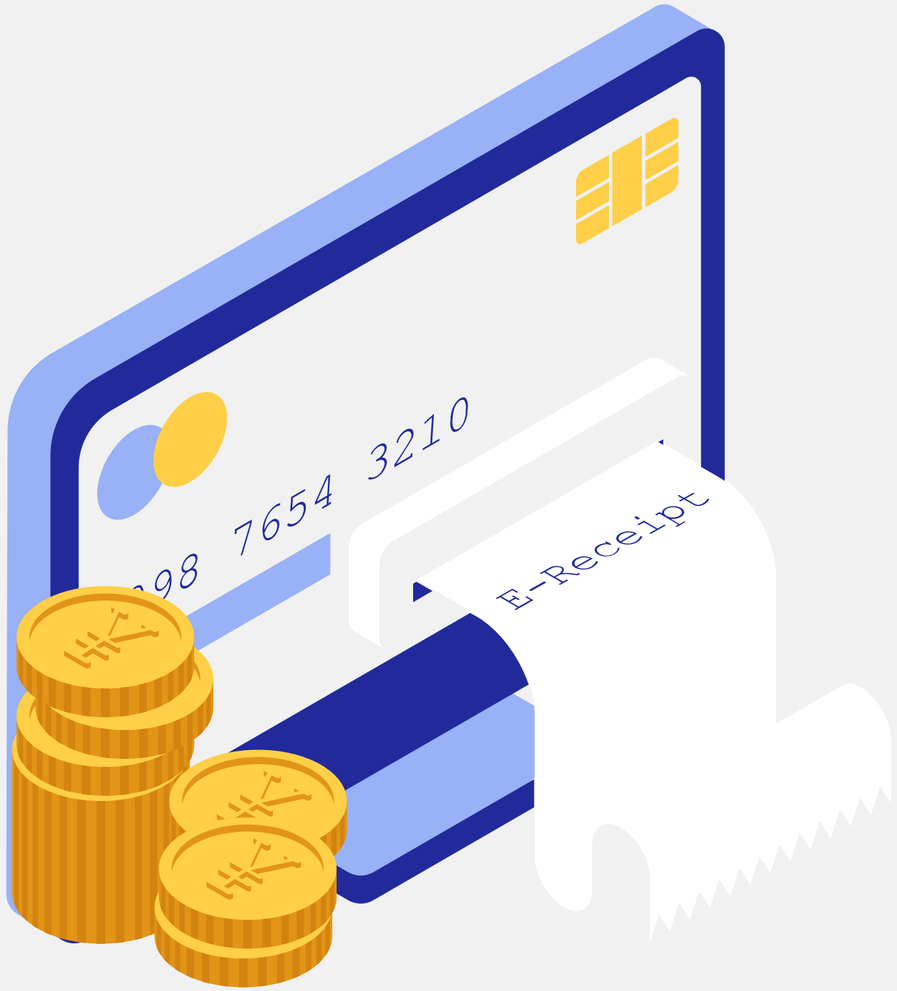
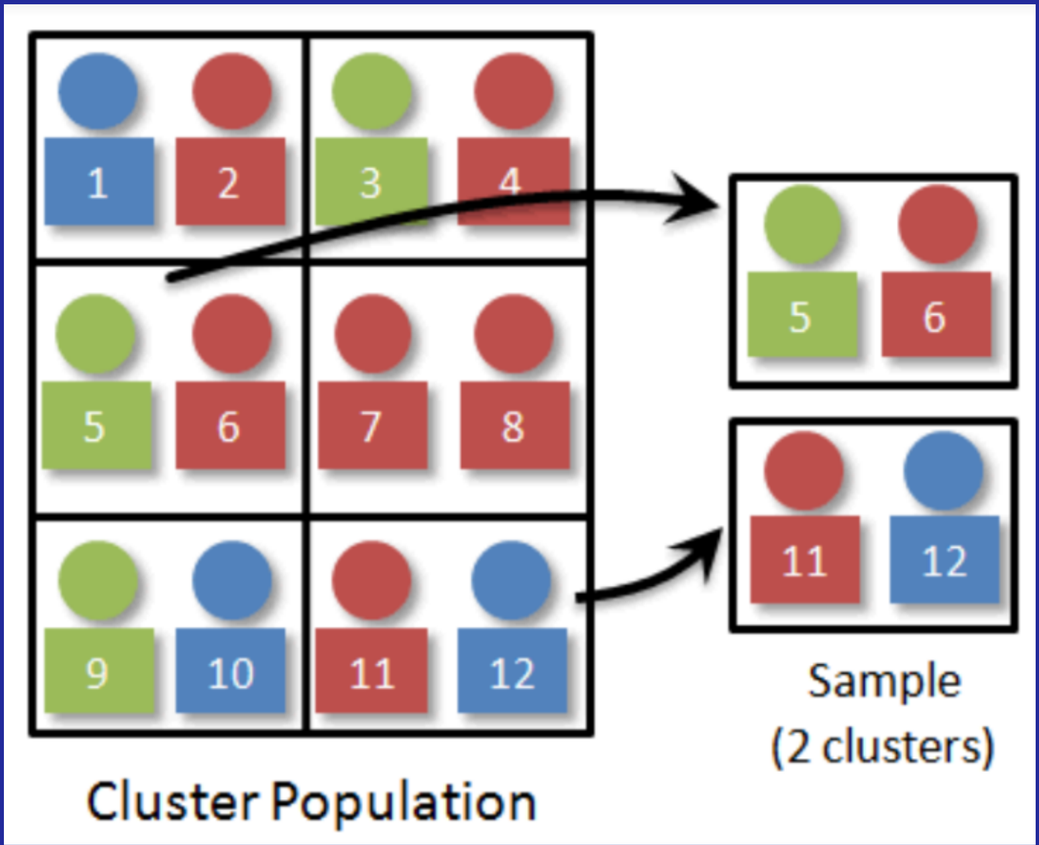


STA1221 MPD P2 - PRAKTIKUM 6

Cluster Random Sampling

AZKA AL AZKIYA

What is Cluster Random Sampling



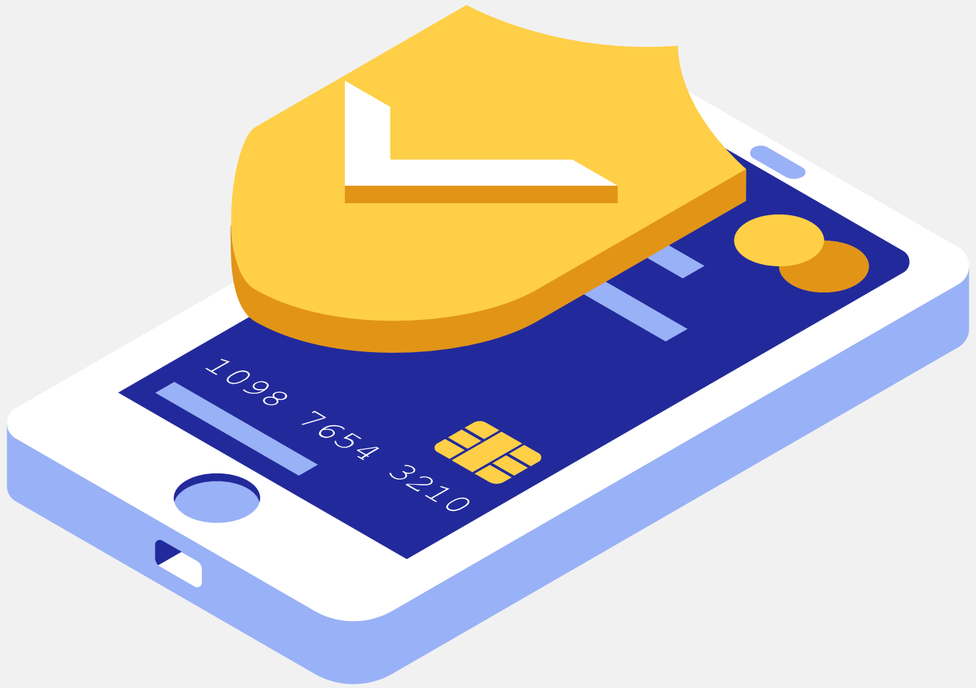
Cluster Sample (Contoh Gerombol) adalah suatu contoh berpeluang yang satuan penarikan contohnya berupa gerombol (kumpulan elemen). Dalam PCG semua elemen dalam gerombol yang terpilih sebagai contoh diamati semua

Ilustrasi

Populasi: seluruh penduduk yang terdaftar sebagai pemilih dalam suatu Pemilu

Elemen: setiap penduduk yang terdaftar sebagai pemilih

Gerombol: TPS (sebagai sampling unit)



Kapan pakai Cluster Random Sampling

Jika kerangka penarikan contoh elemen tidak tersedia, atau untuk mendapatkannya perlu biaya yang besar

Jika biaya untuk memperoleh amatan meningkat dengan semakin jauhnya jarak antar elemen populasi

Stratified vs Cluster

Keragaman di dalam lapisan homogen, antar lapisan heterogen

Keragaman di dalam gerombol heterogen, antar gerombol homogen

Cara menentukan Cluster

01 mendefinisikan gerombol

02 Pertimbangannya: (1) Kedekatan geografis antar elemen dalam gerombol, (2) Ukuran gerombol yang mudah ditangani

Mana yang lebih baik? Mengambil banyak gerombol yang berukuran kecil atau sedikit gerombol yang berukuran besar?

Tergantung keragaman & sudut pandang

Notasi

N = the number of clusters in the population

n = the number of clusters selected in a simple random sample

m_i = the number of elements in cluster i, $i = 1, \ldots, N$

$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} m_i$ = the average cluster size for the sample

$M = \sum_{i=1}^{N} m_i$ = the number of elements in the population

$\overline{M} = M/N$ = the average cluster size for the population

y_i = the total of all observations in the ith cluster

Pendugaan Rataan

Ratio estimator of the population mean μ:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i} \tag{8.1}$$

Estimated variance of $\bar{y}$:

$$\hat{V}(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{s_r^2}{n\bar{M}^2} \tag{8.2}$$

where

$$s_r^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y}m_i)^2}{n - 1} \tag{8.3}$$

Here $\overline{M}$ can be estimated by $\overline{m}$ if M is unknown.

Pendugaan Total

Estimator of the population total τ:

$$M\bar{y} = M\frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i} \tag{8.4}$$

Estimated variance of $M\bar{y}$:

$$\hat{V}(M\bar{y}) = M^2\hat{V}(\bar{y}) = N^2\left(1 - \frac{n}{N}\right)\frac{s_r^2}{n} \tag{8.5}$$

Pendugaan Total (Tidak tergantung M)

Estimator of the population total τ, which does not depend on M:

$$N\bar{y}_t = \frac{N}{n}\sum_{i=1}^{n} y_i \qquad (8.7)$$

Estimated variance of $N\bar{y}_t$:

$$\hat{V}(N\bar{y}_t) = N^2\hat{V}(\bar{y}_t) = N^2\left(1 - \frac{n}{N}\right)\frac{s_t^2}{n} \qquad (8.8)$$

where

$$s_t^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y}_t)^2}{n-1} \qquad (8.9)$$

Ukuran gerombol sama, perbandingan dengan SRS

$$\bar{\bar{y}}_c = \frac{1}{m}\left[\frac{1}{n}\sum_{i=1}^{n} y_i\right] = \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij}$$

$$\hat{V}(\bar{\bar{y}}_c) = \left(1 - \frac{n}{N}\right)\left(\frac{1}{nm^2}\right)\left(\frac{1}{n-1}\right)\sum_{i=1}^{n}(y_i - \bar{y}_t)^2$$

$$\hat{V}(\bar{\bar{y}}_c) = \left(1 - \frac{n}{N}\right)\frac{1}{nm}\text{MSB}$$

ANOVA

Sumber keragaman	df (db)	SS (JK)	MS (KT)
Antar gerombol	n-1	SSB	MSB=SSB/(n-1)
Dalam gerombol	n(m-1)	SSW	MSW=SSW/(n(m-1))
Total	nm-1	SST	

Note: p = n dan r = m

Efisiensi relatif dengan SRS

$$\widehat{\mathrm{RE}(\bar{\bar{y}}_c/\bar{y})} = \frac{\hat{s}^2}{\mathrm{MSB}}$$

$$\hat{s}^2 = \frac{N(m-1)\mathrm{MSW} + (N-1)\mathrm{MSB}}{Nm-1}$$

$$\approx \frac{1}{m}[(m-1)\mathrm{MSW} + \mathrm{MSB}]$$

Ukuran Contoh

Approximate sample size required to estimate μ, with a bound B on the error of estimation:

$$n = \frac{N\sigma_r^2}{ND + \sigma_r^2} \qquad \textbf{(8.12)}$$

where σ_r^2 is estimated by s_r^2 and $D = (B^2\overline{M}^2)/4$.

Approximate sample size required to estimate τ, using $N\bar{y}_t$ with a bound B on the error of estimation:

$$n = \frac{N\sigma_t^2}{ND + \sigma_t^2} \qquad \textbf{(8.15)}$$

where σ_t^2 is estimated by s_t^2 and $D = B^2/4N^2$.

Penduga Proporsi

Estimator of the population proportion p:

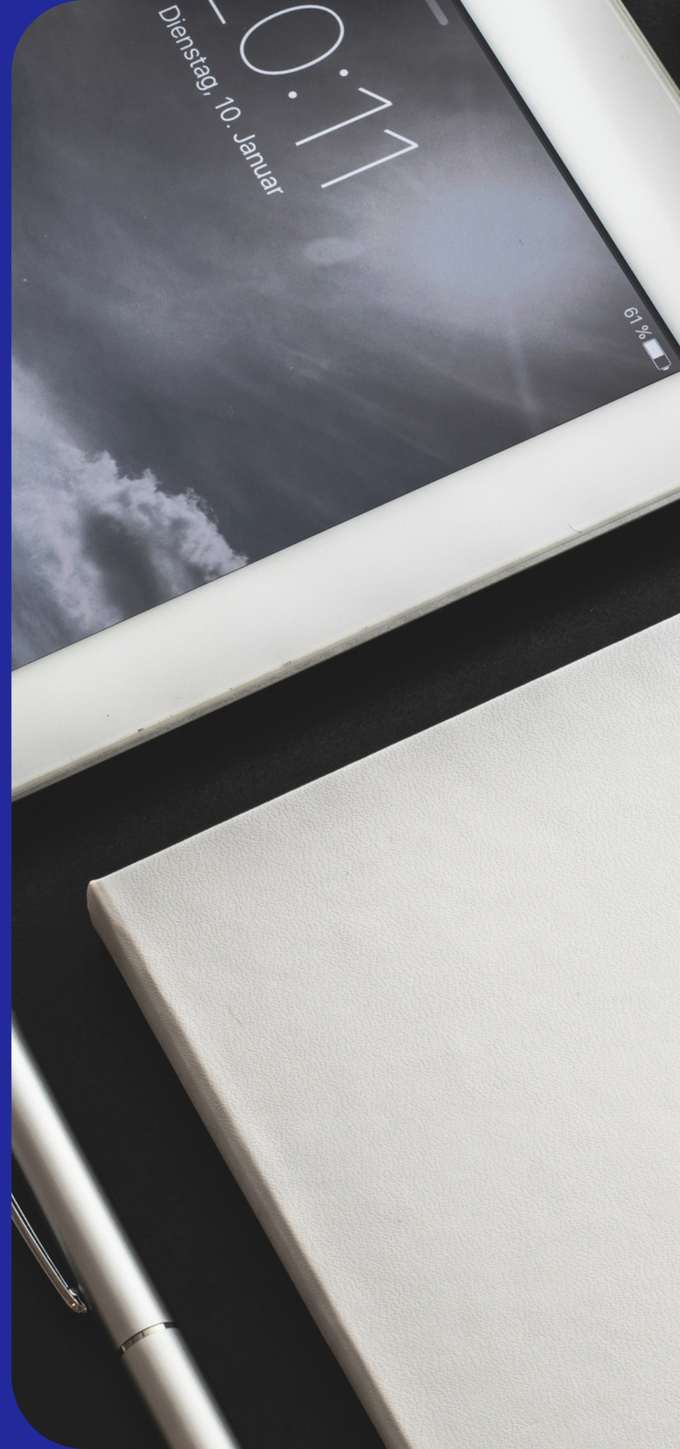
$$\hat{p} = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} m_i} \tag{8.16}$$

Estimated variance of $\hat{p}$:

$$\hat{V}(\hat{p}) = \left(1 - \frac{n}{N}\right)\frac{s_p^2}{n\overline{M}^2} \tag{8.17}$$

where

$$s_p^2 = \frac{\sum_{i=1}^{n}(a_i - \hat{p}m_i)^2}{n - 1} \tag{8.18}$$



Ukuran Contoh Penduga Proporsi

$$n = \frac{N\sigma_p^2}{ND + \sigma_p^2}$$

where $D = B^2\overline{M}^2/4$, and σ_p^2 is estimated by

$$s_p^2 = \frac{\sum_{i=1}^{n}(a_i - \hat{p}m_i)^2}{n - 1}$$