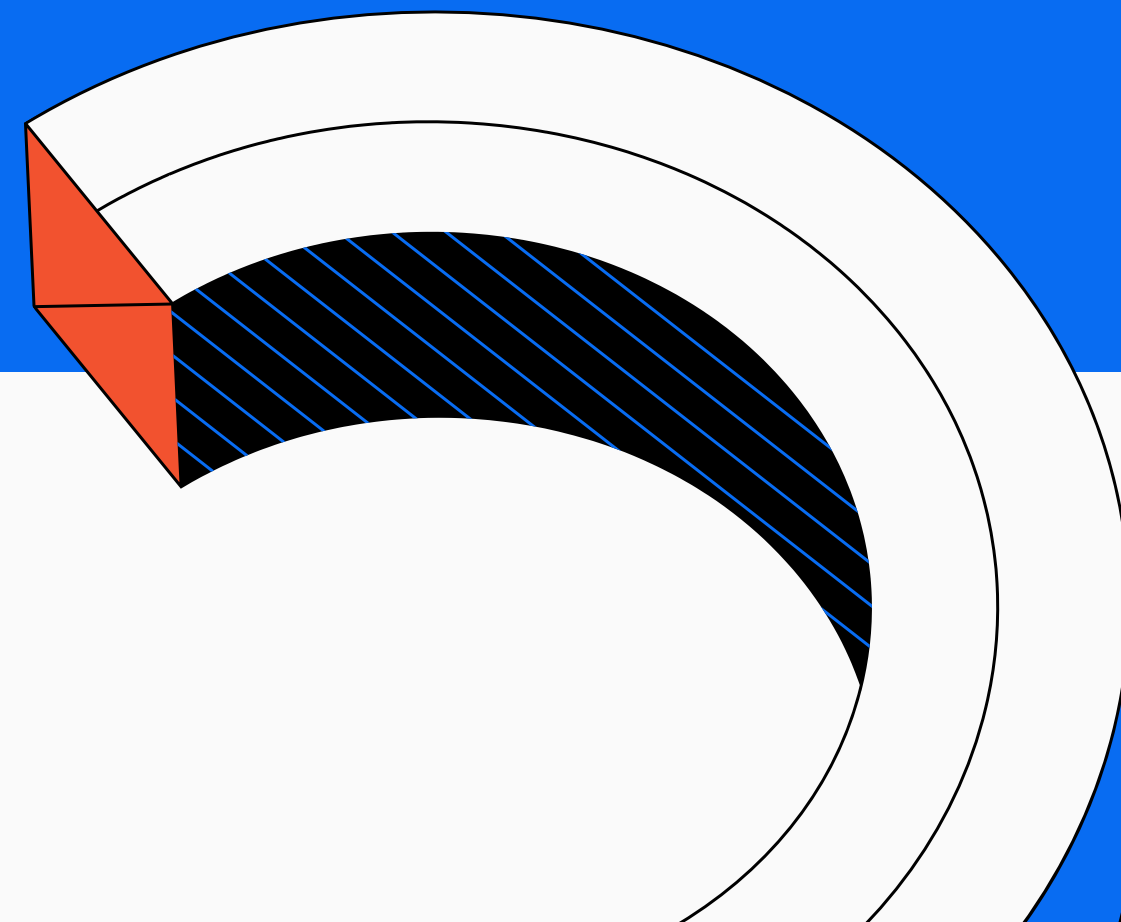
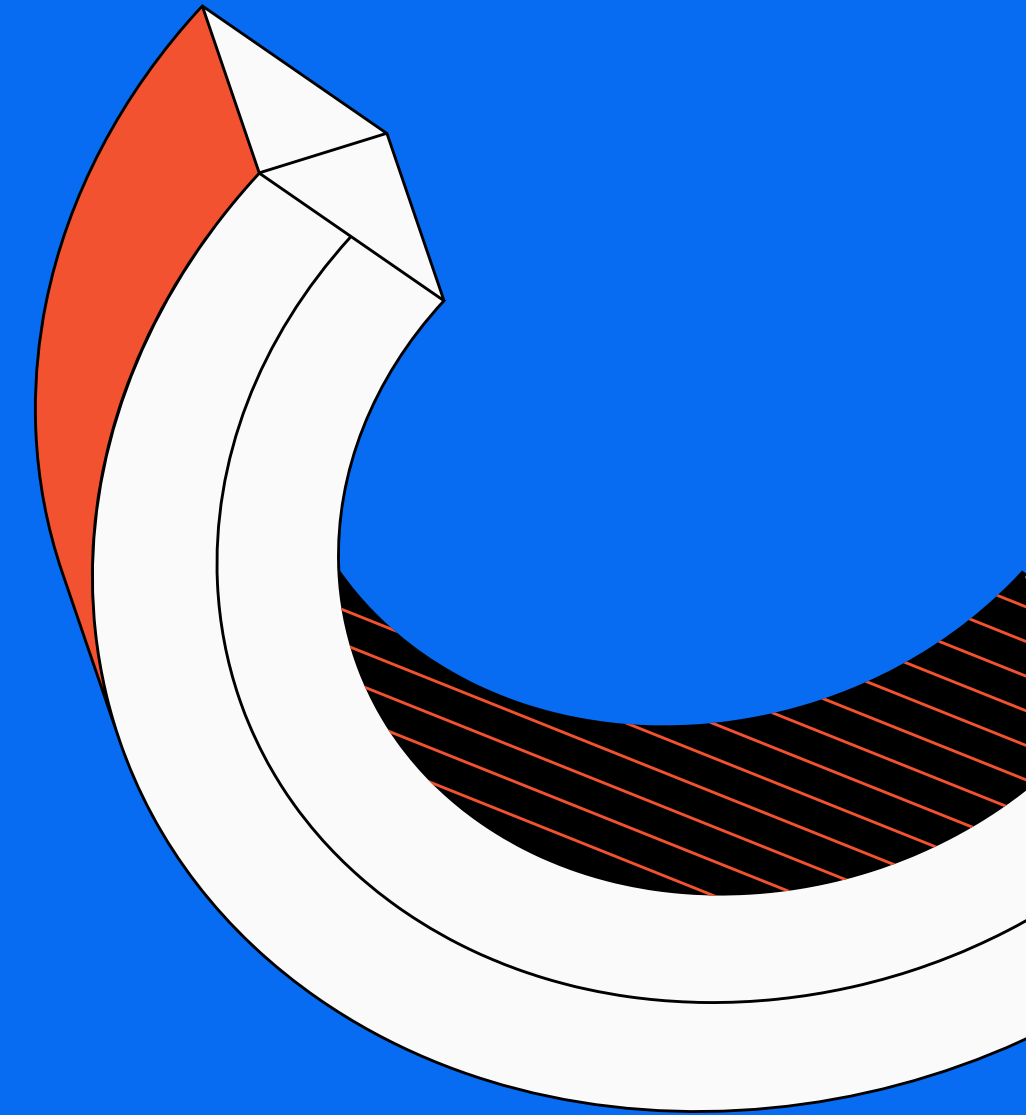


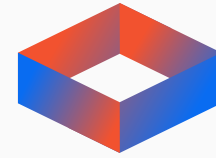
Praktikum STA1221 - Metode Pengumpulan Data

Paralel 2 Pertemuan 4



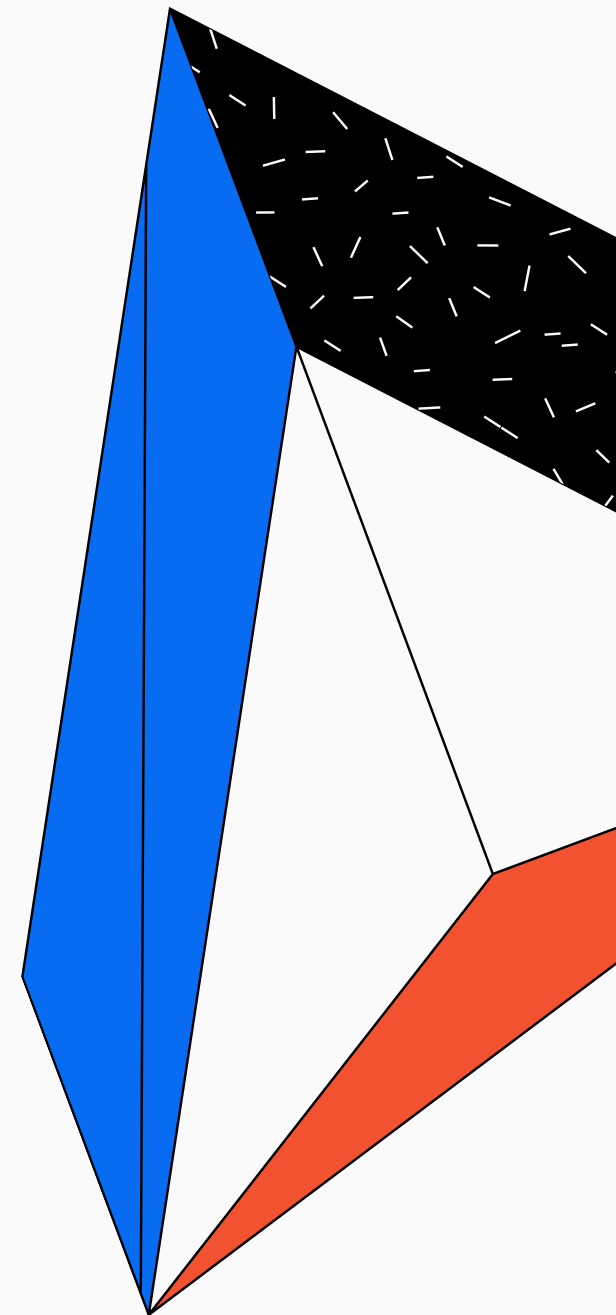
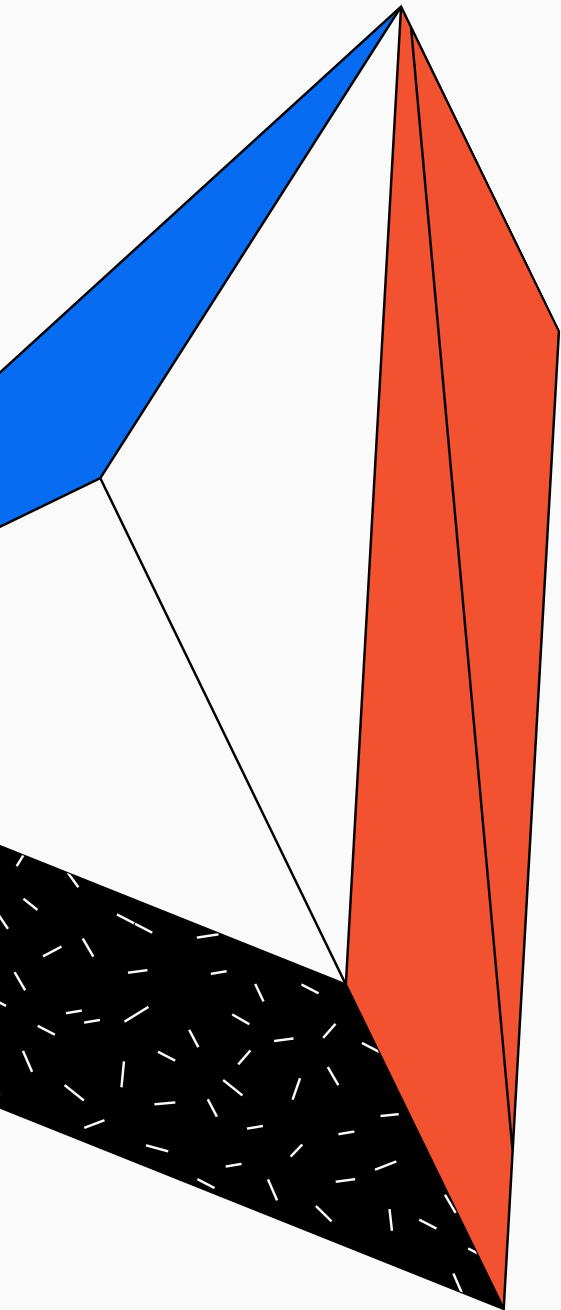
Azka AI Azkiya





Stratified Random Sampling

Contoh acak berlapis didapatkan dengan cara membagi populasi menjadi beberapa kelompok yang **tidak saling tumpang tindih**, dan kemudian **mengambil secara acak (simple random sampling)** dari setiap kelompok-kelompok itu. Kelompok tersebut dinamakan **LAPISAN** atau **STRATA**.



Karakteristik



Karakteristik lapisan atau strata adalah keragaman antar anggota kelompok _____ dan keragaman antar kelompok _____

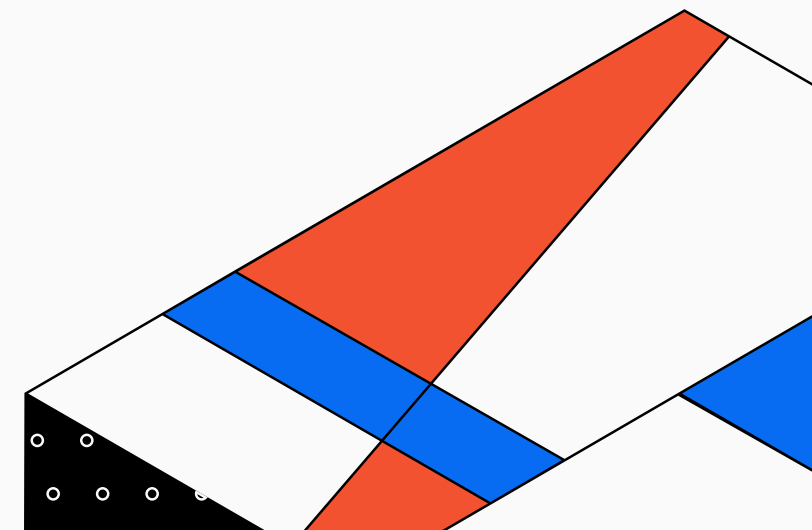
Ada yang bisa kasih contoh strata?

Stratified (PCAB) vs Simple (PCAS)

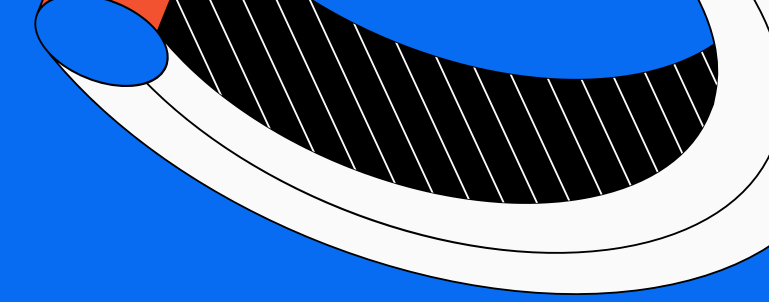
- **BOE lebih kecil***
- **Mengurangi biaya survei**
- **dimungkinkan untuk melakukan pendugaan parameter di setiap sub-secara terpisah,**

- **BOE lebih besar***
- **Biaya survei lebih mahal**
- **Tidak dimungkinkan untuk melakukan pendugaan parameter di setiap sub-secara terpisah,**

***Syarat dan Ketentuan berlaku**



Pendugaan Rataan Populasi



Estimator of the population mean μ :

$$\bar{y}_{st} = \frac{1}{N} [N_1 \bar{y}_1 + N_2 \bar{y}_2 + \cdots + N_L \bar{y}_L] = \frac{1}{N} \sum_{i=1}^L N_i \bar{y}_i \quad (5.1)$$

Estimated variance of \bar{y}_{st} :

$$\begin{aligned} \hat{V}(\bar{y}_{st}) &= \frac{1}{N^2} [N_1^2 \hat{V}(\bar{y}_1) + N_2^2 \hat{V}(\bar{y}_2) + \cdots + N_L^2 \hat{V}(\bar{y}_L)] \\ &= \frac{1}{N^2} \left[N_1^2 \left(1 - \frac{n_1}{N_1} \right) \left(\frac{s_1^2}{n_1} \right) + \cdots + N_L^2 \left(1 - \frac{n_L}{N_L} \right) \left(\frac{s_L^2}{n_L} \right) \right] \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \left(1 - \frac{n_i}{N_i} \right) \left(\frac{s_i^2}{n_i} \right) \end{aligned} \quad (5.2)$$

Selang Kepercayaan & BOE

Selang kepercayaan $(1-\alpha)100\%$ bagi μ :

$$\bar{y}_{st} \pm t_{\frac{\alpha}{2}} \sqrt{\hat{V}(\bar{y}_{st})}$$

$$BOE = 2 \sqrt{\hat{V}(\bar{y}_{st})}$$

Kenapa BOE rumusnya pake 2?

Pendugaan Total Populasi

Estimator of the population total τ :

$$N\bar{y}_{\text{st}} = N_1\bar{y}_1 + N_2\bar{y}_2 + \cdots + N_L\bar{y}_L = \sum_{i=1}^L N_i\bar{y}_i \quad (5.3)$$

Estimated variance of $N\bar{y}_{\text{st}}$:

$$\hat{V}(N\bar{y}_{\text{st}}) = N^2\hat{V}(\bar{y}_{\text{st}}) = \sum_{i=1}^L N_i^2 \left(1 - \frac{n_i}{N_i}\right) \left(\frac{s_i^2}{n_i}\right) \quad (5.4)$$

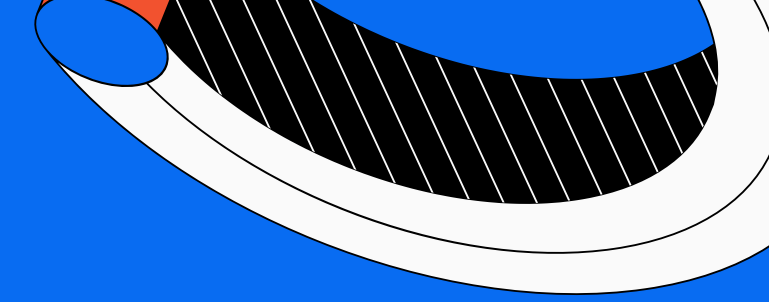
Selang Kepercayaan & BOE

Selang kepercayaan $(1-\alpha)100\%$ bagi τ :

$$\hat{\tau}_{st} \pm t_{\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{\tau}_{st})}$$

$$BOE = 2 \sqrt{\hat{V}(\hat{\tau}_{st})}$$

Penentuan ukuran contoh



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

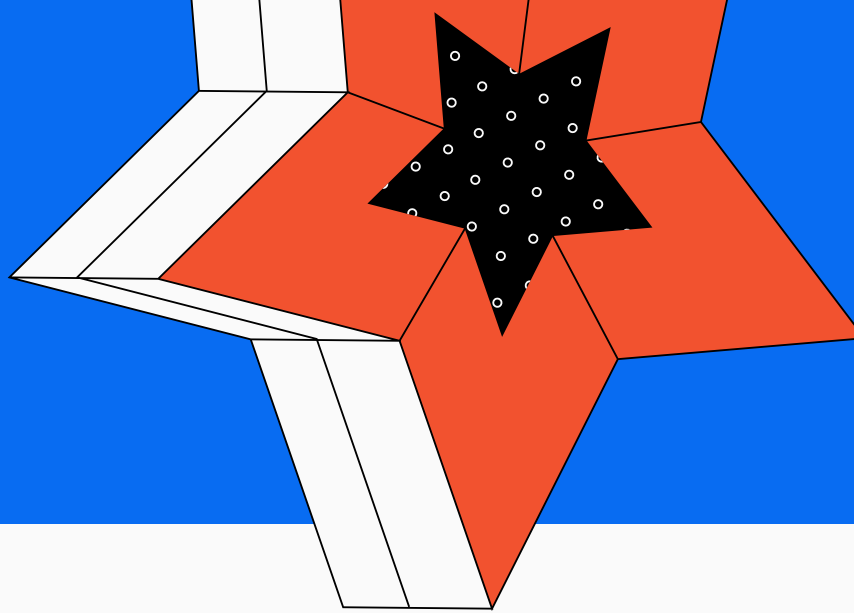
$$n = \frac{\sum_{i=1}^L N_i^2 \sigma_i^2 / a_i}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2} \quad (5.6)$$

where a_i is the fraction of observations allocated to stratum i , σ_i^2 is the population variance for stratum i , and

$$D = \frac{B^2}{4} \quad \text{when estimating } \mu$$

$$D = \frac{B^2}{4N^2} \quad \text{when estimating } \tau$$

Alokasi



Skema alokasi terbaik dipengaruhi oleh 3 faktor:



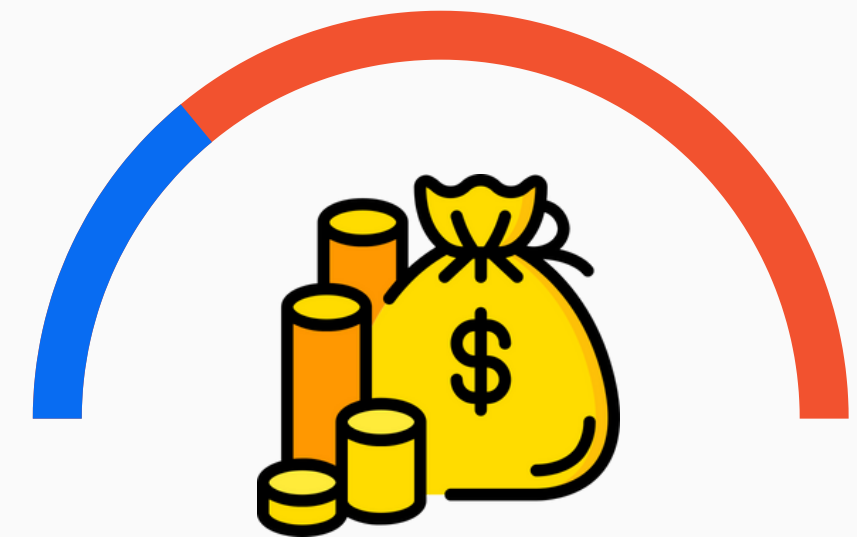
Faktor 1

**Jumlah elemen
dalam tiap stratum**



Faktor 2

**Keragaman observasi
dalam setiap stratum**



Faktor 3

**Biaya untuk
mendapatkan
observasi di tiap
stratum**

Alokasi Optimum

Penentuan ukuran contoh jika diketahui biaya

$$n = \frac{\left(\sum_{k=1}^L N_k \sigma_k / \sqrt{c_k} \right) \left(\sum_{i=1}^L N_i \sigma_i \sqrt{c_i} \right)}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2}$$

Alokasi Optimum

Approximate allocation that minimizes cost for a fixed value of $V(\bar{y}_{st})$ or minimizes $V(\bar{y}_{st})$ for a fixed cost:

$$n_i = n \left(\frac{N_i \sigma_i / \sqrt{c_i}}{N_1 \sigma_1 / \sqrt{c_1} + N_2 \sigma_2 / \sqrt{c_2} + \cdots + N_L \sigma_L / \sqrt{c_L}} \right) \quad (5.7)$$
$$= n \left(\frac{N_i \sigma_i / \sqrt{c_i}}{\sum_{k=1}^L N_k \sigma_k / \sqrt{c_k}} \right)$$

where N_i denotes the size of the i th stratum, σ_i^2 denotes the population variance for the i th stratum, and c_i denotes the cost of obtaining a single observation from the i th stratum. Note that n_i is directly proportional to N_i and σ_i and inversely proportional to $\sqrt{c_i}$.

c_i = biaya untuk memperoleh satu unit contoh dari lapisan ke- i

Alokasi Neyman & Proporsional

Alokasi Neyman(biaya sama)

$$n_i = n \left(\frac{N_i \sigma_i}{\sum_{k=1}^L N_k \sigma_k} \right)$$

$$n = \frac{\left(\sum_{k=1}^L N_k \sigma_k \right)^2}{N^2 D + \sum_{i=1}^L N_i \sigma_i^2}$$

Alokasi Proporsional

$$n_i = n \left(\frac{N_i}{N} \right)$$

$$n = \frac{\sum_{i=1}^L N_i \sigma_i^2}{ND + \frac{1}{N} \sum_{i=1}^L N_i \sigma_i^2}$$

Pendugaan Proporsi Populasi

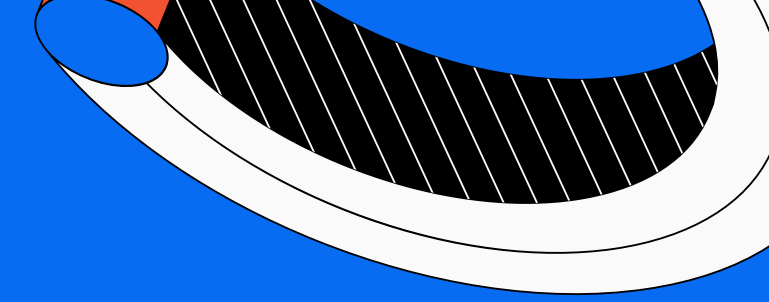
Estimator of the population proportion p :

$$\hat{p}_{\text{st}} = \frac{1}{N^2} (N_1 \hat{p}_1 + N_2 \hat{p}_2 + \cdots + N_L \hat{p}_L) = \frac{1}{N} \sum_{i=1}^L N_i \hat{p}_i \quad (5.13)$$

Estimated variance of \hat{p}_{st} :

$$\begin{aligned} \hat{V}(\hat{p}_{\text{st}}) &= \frac{1}{N^2} [N_1^2 \hat{V}(\hat{p}_1) + N_2^2 \hat{V}(\hat{p}_2) + \cdots + N_L^2 \hat{V}(\hat{p}_L)] \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \hat{V}(\hat{p}_i) \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \left(1 - \frac{n_i}{N_i}\right) \left(\frac{\hat{p}_i \hat{q}_i}{n_i - 1}\right) \end{aligned} \quad (5.14)$$

Ukuran Contoh untuk Pendugaan Proporsi Populasi



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

$$n = \frac{\sum_{i=1}^L N_i^2 p_i q_i / a_i}{N^2 D + \sum_{i=1}^L N_i p_i q_i} \quad (5.15)$$

where a_i is the fraction of observations allocated to stratum i , p_i is the population proportion for stratum i , and $D = B^2/4$.

Latihan 1

Seorang peneliti ingin mengetahui rata – rata konsumsi gas elpiji per rumah tangga per bulan (kg) di suatu kawasan tertentu. Berdasarkan penghasilan, karakteristik rumah tangga di kawasan tersebut terbagi menjadi 3 kelompok, yaitu kelompok A, B, dan C. dimana penghasilan rumah tangga kelompok A > kelompok B > kelompok C. penarikan sampel yang digunakan adalah penarikan contoh acak berlapis. Informasi awal yang diperoleh dari penelitian sebelumnya sebagai berikut :

| Kelompok A | Kelompok B | Kelompok C |
|------------------|-------------------|------------------|
| $N1 = 100$ | $N2 = 200$ | $N3 = 350$ |
| $\sigma_1^2 = 9$ | $\sigma_2^2 = 16$ | $\sigma_3^2 = 4$ |

a. Berdasarkan alokasi neyman, tentukan ukuran sampel masing – masing kelompok jika jumlah sampel total ada 90!

Jika statistik yang diperoleh dari sampel sebagai berikut :

| Kelompok A | Kelompok B | Kelompok C |
|--------------------|--------------------|--------------------|
| $n1 = 15$ | $n2 = 40$ | $n3 = 35$ |
| $\bar{y}_1 = 30.2$ | $\bar{y}_2 = 14.5$ | $\bar{y}_3 = 10.4$ |
| $S1 = 2$ | $S2 = 5$ | $S3 = 3$ |

b. Hitung \bar{y}_{st} dan $V(\bar{y}_{st})$!

Latihan 2

Distributor makanan grosir di kota besar ingin tahu apakah ia harus meluncurkan produk baru ke stoknya. Untuk membantu mengambil keputusan, ia berencana menambahkan produk baru ini ke sampel restoran yang ia handle untuk memperkirakan rata – rata penjualan bulanan. Dia hanya melayani empat daerah besar di kota. Oleh karena itu, untuk kenyamanan administrasi, ia memutuskan untuk menggunakan sampel acak berstrata dengan setiap daerah sebagai strata. Ada 24 restoran di strata 1, 36 di strata 2, 30 di strata 3 dan 30 di strata 4. Dengan demikian $N_1 = 24$, $N_2 = 36$, $N_3 = 30$, $N_4 = 30$ dan $N = 120$. Distributor memiliki waktu dan uang yang cukup untuk mendapatkan sampel sebesar $n = 20$ restoran. Karena dia tidak memiliki informasi awal, dan karena biaya pengambilan sampel adalah sama di setiap strata, ia memutuskan untuk menggunakan alokasi proporsional. Produk baru diperkenalkan di 4 restoran yang dipilih secara acak dari daerah 1, 6 restoran dari daerah 2, dan 5 masing – masing dari daerah 3 dan 4. Data penjualan setelah sebulan ditunjukkan pada table di bawah ini. Perkirakan rata – rata penjualan untuk bulan itu beserta BoEnya!

Latihan 2

| Stratum 1 | Stratum 2 | Stratum 3 | Stratum 4 |
|-------------------------------------|--------------------------------------|-------------------------------------|---------------------------------------|
| 94 | 91 | 108 | 92 |
| 90 | 99 | 96 | 110 |
| 102 | 93 | 100 | 94 |
| 110 | 105 | 93 | 91 |
| | 111 | 93 | 113 |
| | 101 | | |
| $\hat{y}_1 = 99$ $s_1^2 = 78.67$ | $\hat{y}_2 = 100$ $s_2^2 = 55.60$ | $\hat{y}_3 = 98$ $s_3^2 = 39.50$ | $\hat{y}_4 = 100$ $s_4^2 = 112.50$ |

Latihan 3

A chain department store tertarik untuk memperkirakan proporsi akun yang menunggak. The chain tersebut terdiri dari empat toko. Untuk mengurangi biaya pengambilan sampel, stratified random sampling digunakan dengan masing – masing toko sebagai strata. Karena tidak ada informasi tentang proporsi populasi yang sebelumnya, alokasi proporsional digunakan. Dari table dibawah ini, dugalah proporsi akun yang menunggak tersebut beserta BoEnya!

| | Stratum I | Stratum II | Stratum III | Stratum IV |
|---|--------------|---------------|----------------|---------------|
| Number of accounts receivable | $N_1 = 65$ | $N_2 = 42$ | $N_3 = 93$ | $N_4 = 25$ |
| Sample size | $n_1 = 14$ | $n_2 = 9$ | $n_3 = 21$ | $n_4 = 6$ |
| Sample number of delinquent accounts | 4 | 2 | 8 | 1 |

THANK
YOU

