

METODE SURVEI
PERTEMUAN 9

METODE SAMPLING: SYSTEMATIC SAMPLING

Dosen Pengampu: Atik Maratis Suhartini, S.E, M.Si



Disusun Oleh:
2KS2 / Kelompok 3

Diah Ayu Nur Rahmadani	(222313048)
Harry Pentaleon Rumahorbo	(222313112)
Muhammad Daffa Dekananda	(222313233)
Satria Imka Dwi Putra	(222313372)
Seto Haidar Yudhistira	(222313375)
Zahira Priyan Husna	(222313439)

PROGRAM STUDI D-IV KOMPUTASI STATISTIK
POLITEKNIK STATISTIKA STIS
TAHUN AJARAN 2024/2025

A. Definisi

Systematic sampling adalah metode di mana unit pertama dipilih menggunakan angka acak, dan unit-unit yang tersisa dipilih dengan pola yang telah ditentukan sebelumnya. Dalam systematic sampling, peneliti memilih anggota populasi untuk sampel mereka pada interval reguler (atau k) yang ditentukan sebelumnya.

Metode ini merupakan kombinasi antara random sampling dan convenience sampling, di mana randomisasi hanya dilakukan sekali di awal untuk menentukan titik mulai, kemudian seleksi dilakukan secara sistematis dengan interval yang tetap.

Cara Kerja Systematic Sampling:

1. Menyusun Kerangka Sampel: Populasi harus tersusun dalam bentuk daftar yang lengkap dan terurut
2. Menghitung Interval Sampling (k): $k = N/n$, di mana N adalah ukuran populasi dan n adalah ukuran sampel yang diinginkan
3. Pemilihan Titik Awal: Pilih secara acak satu angka antara 1 hingga k
4. Seleksi Sistematis: Mulai dari titik awal, pilih setiap unit ke- k hingga mencapai ukuran sampel yang diinginkan

Contoh Sederhana:

Misalkan ada 1000 mahasiswa ($N=1000$) dan kita ingin sampel 100 mahasiswa ($n=100$):

- Interval $k = 1000/100 = 10$
- Pilih acak angka antara 1-10, misalnya angka 3
- Sampel yang terpilih: mahasiswa nomor 3, 13, 23, 33, 43, ..., 993

B. Alasan/ Syarat Sampling

Systematic sampling digunakan karena beberapa alasan berikut:

- Kemudahan dan Efisiensi: Hanya memerlukan satu angka acak untuk memilih unit pertama, sehingga lebih cepat dan murah dibandingkan Simple Random Sampling (SRS) yang memerlukan pengacakan berulang.
- Representasi Populasi: Jika populasi diurutkan berdasarkan karakteristik tertentu (misalnya, luas lahan atau jumlah ternak), systematic sampling dapat menghasilkan sampel yang lebih representatif dibandingkan SRS.
- Kesesuaian dengan Struktur Populasi: Cocok untuk populasi dengan susunan acak atau terurut dengan tren linear, tetapi kurang efektif untuk populasi dengan variasi periodik.

Syarat penggunaannya meliputi:

- Populasi harus memiliki kerangka sampel yang jelas dengan nomor urut. Interval k sebaiknya bilangan bulat atau dibulatkan ke bilangan bulat terdekat jika N tidak habis dibagi n.
- Pengurutan populasi harus dipertimbangkan untuk menghindari bias, terutama pada populasi dengan pola periodik.

C. Kelebihan dan Kekurangan

1. Kelebihan

- **Pelaksanaan Mudah:** Prosesnya lebih cepat dan sederhana karena hanya memerlukan satu angka acak dan interval tetap.
- **Representatif:** Dengan pengurutan yang tepat, sampel cenderung lebih merata dan representatif, terutama pada populasi terurut.
- **Presisi Tinggi:** Pada populasi dengan tren linear, varians systematic sampling sering kali lebih kecil dibandingkan SRS

2. Kelemahan

- **Kesulitan Estimasi Varians:** Penduga varians yang tidak bias sulit diperoleh dari sampel sistematis tunggal.
- **Risiko Bias:** Pengurutan yang tidak tepat atau populasi dengan variasi periodik dapat menghasilkan sampel yang tidak representatif.
- **Ketergantungan pada Kerangka Sampel:** Kerangka sampel yang buruk dapat meningkatkan varians dan mengurangi efisiensi.

D. Estimasi

- **Rata-rata Sampel (\bar{y}_{sy})**

$$\bar{y}_{sy} = \frac{1}{n} \sum_{j=1}^n y_j$$

Keterangan:

\hat{y}_{sy} : Rata-rata sampel, yaitu estimasi rata-rata populasi berdasarkan sampel.

n : Ukuran sampel, jumlah elemen yang diambil dalam sampel.

y_j : Elemen dalam sampel ke-j, yaitu nilai yang diamati dalam unit sampel ke-j.

Rata-rata sampel ini merupakan penduga tak bias untuk rata-rata populasi jika $N = nk$ (Linear Systematic Sampling) atau selalu tak bias pada Circular Systematic Sampling.

- **Total Populasi (\hat{y}_{sy})**

$$\hat{y}_{sy} = N \cdot \bar{y}_{sy}$$

Keterangan:

\bar{y}_{sy} : Estimasi rata-rata sampel, yang digunakan untuk menghitung total populasi.

N : Ukuran total populasi, yaitu jumlah seluruh elemen dalam populasi.

Estimasi total populasi diperoleh dengan mengalikan rata-rata sampel dengan jumlah unit populasi N .

- **Varians Rata- rata Sampel ($V(\bar{y}_{sy})$)**

$$V(\bar{y}_{sy}) = \frac{S^2}{n} \frac{N-1}{N} [1 + (n-1)\rho]$$

Keterangan:

$V(\bar{y}_{sy})$: Varians dari rata-rata sampel, yang mengukur ketidakpastian dalam estimasi rata-rata.

S^2 : Varians populasi, mengukur sebaran data dalam populasi.

n : Ukuran sampel, jumlah elemen yang diambil dalam sampel.

N : Ukuran populasi, jumlah elemen dalam populasi.

ρ : Koefisien korelasi intra-kelompok, yang mengukur ketergantungan antar elemen dalam kelompok.

Di mana S^2 adalah varians populasi:

$$S^2 = \frac{1}{N-1} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{Y})^2$$

Keterangan:

S^2 : Varians populasi, menunjukkan seberapa besar nilai-nilai dalam populasi tersebar dari rata-rata populasi.

N : Ukuran total populasi.

k : Jumlah kelompok dalam sampel atau data.

y_{ij} : Nilai observasi untuk elemen sampel ke-i dan ke-j.

\bar{Y} : Rata-rata populasi atau kelompok.

dan ρ adalah koefisien korelasi intraklas:

$$\rho = \frac{2}{kn(n-1)} \sum_{i=1}^k \sum_{j < j'} (y_{ij} - \bar{Y})(y_{ij'} - \bar{Y}) \cdot \frac{N}{(N-1)S^2}$$

Keterangan:

ρ : Koefisien korelasi intra-kelompok, yang mengukur tingkat ketergantungan antara elemen-elemen dalam kelompok.

k : Jumlah kelompok.

n : Ukuran sampel.

y_{ij} : Nilai observasi untuk elemen sampel ke-i dan ke-j.

\bar{Y} : Rata-rata populasi atau kelompok.

- **Interval Kepercayaan 95%**

$$\bar{y}_{sy} \pm z_{\alpha/2} \cdot \sqrt{V(\bar{y}_{sy})}$$

Keterangan:

\hat{y}_{sy} : Estimasi rata-rata sampel.

$z_{\alpha/2}$: Nilai kritis dari distribusi normal standar pada tingkat kepercayaan tertentu (misalnya 1.96 untuk kepercayaan 95)

$V(\hat{y}_{sy})$: Varians rata-rata sampel, yang mengukur ketidakpastian dalam estimasi rata-rata.

Untuk total populasi:

$$\hat{y}_{sy} \pm z_{\alpha/2} \cdot \sqrt{N^2 \cdot V(\bar{y}_{sy})}$$

Keterangan:

\hat{y}_{sy} : Estimasi rata-rata populasi dari sampel sistematis.

y_{sy} : Rata-rata sampel yang dihitung.

$z_{\alpha/2}$: Nilai z untuk tingkat kepercayaan pada distribusi normal.

$V(\hat{y}_{sy})$: Varians estimasi rata-rata dari sampel sistematis.

dengan $Z_{\frac{\alpha}{2}} = 1.96$ untuk tingkat kepercayaan 95%.

- **Varians dengan Paired Selection Model (PSM) (untuk n genap)**

$$v(\bar{y}_{sy}) = \frac{1-f}{n^2} \sum_{i=1}^{n/2} (y_{2i} - y_{2i-1})^2, \quad f = \frac{n}{N}$$

Keterangan:

n : Jumlah sampel yang dipilih.

N : Ukuran total populasi.

f : Fraksi sampel terhadap populasi, $f = \frac{n}{N}$.

- **Varians dengan Successive Difference Model (SDM)**

$$v(\bar{y}_{sy}) = \frac{1-f}{n(n-1)} \sum_{i=1}^{n-1} (y_{i+1} - y_i)^2$$

Keterangan:

Σ : Simbol penjumlahan.

y_{2i}, y_{2i-1} : Elemen dalam data yang dipilih dengan interval tertentu.

$(y_{2i} - y_{2i-1})^2$: Kuadrat dari selisih dua elemen berturut-turut dalam data.

$\frac{1}{n(n-1)}$: Faktor untuk menghitung varians dalam sampel.

y_i, y_{i+1} : Elemen-elemen berturut-turut dalam data yang dihitung variansnya.

$(y_i - y_{i+1})^2$: Kuadrat dari selisih antara elemen berturut-turut.

E. Contoh Soal

Misalkan ada $N=9$ populasi, yaitu 1,2,3,4,5,6,7,8,9. Kemudian diambil sampel sebanyak 3 dengan sistematis. Hitunglah rata-rata dan variansnya!

Jawab:

	Sampel 1		Sampel 2		Sampel3	
	y_{1j}	y_{1j}^2	y_{2j}	y_{2j}^2	y_{3j}	y_{3j}^2
	1	1	2	4	3	9
	4	16	5	25	6	36
	7	49	8	64	9	81
Total	12	66	15	93	18	126
rata-rata i	4		5		6	
rata-rata total	5					

Penghitungan:

Cara I

$$V(\bar{y}_{sy}) = \frac{1}{k} \sum_i^k (\bar{y}_i - \bar{Y})^2 = \frac{1}{3} [(4-5)^2 + (5-5)^2 + (6-5)^2] = \frac{2}{3}$$

Cara II

$$S^2 = \frac{\sum_i^k \sum_j^n (y_{ij} - \bar{Y})^2}{N-1} = \frac{1}{9-1} [(1-5)^2 + \dots + (9-5)^2] = \frac{60}{8}$$

$$\text{Varians within} = \frac{\sum_i^k \sum_j^n (y_{ij} - \bar{y}_i)^2}{N} = \frac{1}{9} [(1-4)^2 + \dots + (9-6)^2] = \frac{54}{9}$$

$$V(\bar{Y}_{sy}) = \frac{N-1}{N} S^2 - \frac{k(n-1)}{N} S_{wsy}^2 = \frac{9-1}{9} \frac{60}{8} - \frac{54}{9} = \frac{2}{3}$$

Cara III

$$\begin{aligned} \sum_{j < u} (y_{ij} - \bar{Y})(y_{iu} - \bar{Y}) &= (y_{11} - \bar{Y})(y_{12} - \bar{Y}) + (y_{11} - \bar{Y})(y_{13} - \bar{Y}) + (y_{12} - \bar{Y})(y_{13} - \bar{Y}) \\ &= (1-5)(4-5) + (1-5)(7-5) + (4-5)(7-7) = -6 \end{aligned}$$

Dengan cara yang sama untuk $i = 2$ di dapat -9 dan $i = 3$ di dapat -6 serta $S^2 = 60/8$

$$\begin{aligned} \rho &= \frac{2}{n-1} \frac{1}{N-1} \frac{1}{S^2} \sum_i \sum_{j < u} (y_{ij} - \bar{Y})(y_{iu} - \bar{Y}) = \frac{2}{3-1} \frac{1}{9-1} \frac{1}{60/8} (-6-9-6) = -\frac{21}{60} \\ V(\bar{Y}_{sy}) &= \frac{S^2}{n} \frac{N-1}{N} [1 + (n-1)\rho] = \frac{60/8}{3} \frac{9-1}{9} \left[1 + (3-1) \left(-\frac{21}{60} \right) \right] = \frac{2}{3} \end{aligned}$$

Interpretasi:

Rata-rata sampel sistematis dari populasi $\{1, 2, \dots, 9\}$ yang diambil sebanyak 3 elemen adalah 5, dan varians rata-ratanya sebesar $2/3$. Ketiga cara yang digunakan—baik pendekatan langsung dari data sampel, pendekatan varians populasi dan within group, maupun pendekatan korelasi (ρ)—semuanya memberikan hasil varians yang sama, yaitu $2/3$. Hal ini menegaskan bahwa metode systematic sampling dapat dianalisis melalui berbagai pendekatan statistik yang saling konsisten, dan pentingnya mempertimbangkan struktur keteraturan data (seperti korelasi antar elemen) dalam menghitung varians estimasi.