

TEKNIK SAMPLING DAN DATA WRANGLING

Simple Random Sampling (SRS)

Simple Random Sample

Simple Random Sampling With Replacement (SRSWR)

Setelah satu elemen dipilih, elemen itu dikembalikan lagi ke populasi (satu elemen bisa terpilih lebih dari sekali)

$$\frac{(N + n - 1)!}{n!(N - 1)!}$$

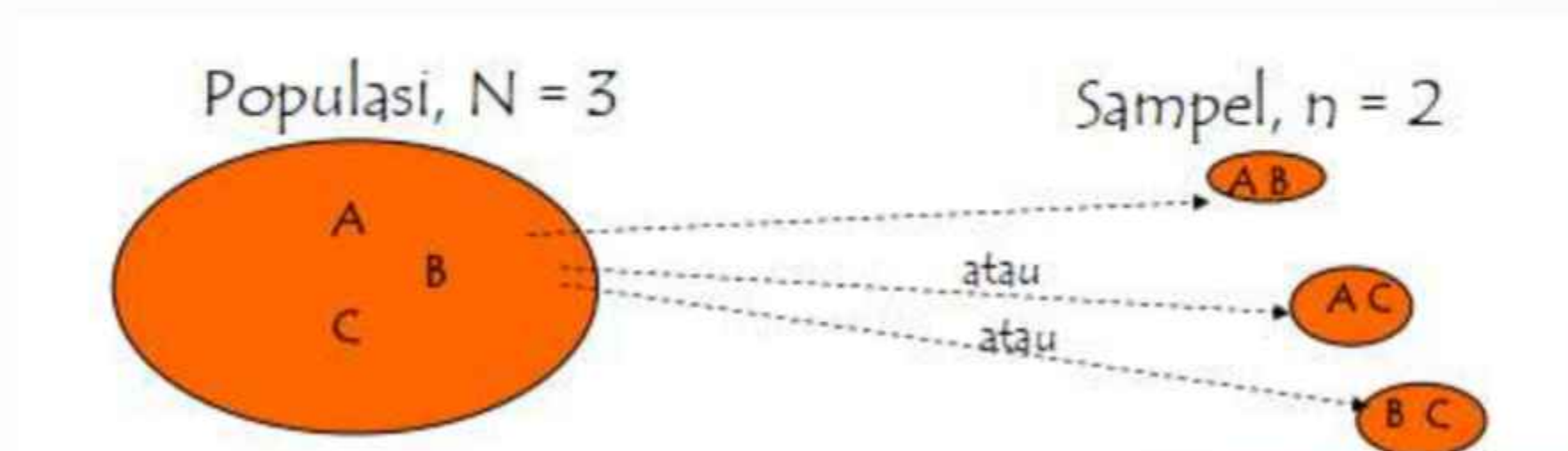
Simple Random Sampling Without Replacement (SRSWOR)

Setelah satu elemen dipilih, elemen itu tidak dikembalikan (satu elemen hanya bisa terpilih sekali)

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

Konsep SRS

Populasi berukuran $N = 3$ diambil sampel berukuran $n = 2$, maka banyak sampel yang mungkin adalah 3 yaitu AB, AC, dan BC yang masing-masing mempunyai peluang/kesempatan yang sama besar untuk terpilih menjadi sampel, dengan peluang sebesar $1/3$.



Keuntungan & Kelemahan SRS

Keuntungan:

- cara pengambilan sampel dan teknik estimasi parameteranya sederhana.

Kelemahan:

- hanya cocok untuk populasi yang relatif homogen
- hanya cocok untuk cakupan survei yang tidak terlalu luas, karena membutuhkan kerangka sampel sampai elemen
- biaya tinggi untuk populasi yang besar.

Simple Random Sample

INDEPENDENT CHOICE OF DIGIT

- Tentukan baris, kolom, dan halaman Tabel Angka Random (TAR) yang digunakan untuk memulai penelusuran angka random.
- Jika jumlah populasi sebanyak N unit dan jumlah digits dari N adalah sebanyak r digit, maka telusuri r digit angka dari baris dan kolom permulaan.
- Jika angka random (AR) $\leq N$, maka unit yang nomor urutnya sama dengan AR tersebut terpilih sebagai sampel.
- Jika angka random (AR) $= 0$, maka unit ke- N (terakhir) terpilih sebagai sampel.
- Jika angka random (AR) $> N$, maka lanjutkan penelusuran ke angka random di baris selanjutnya pada kolom yang sama.
- Lakukan pengambilan AR sampai jumlah sampel terpenuhi.
- Jika sudah sampai pada kolom terakhir dan belum mendapatkan angka random sebanyak sampel, lanjutkan ke kolom berikutnya pada baris pertama.

Simple Random Sample

INDEPENDENT CHOICE OF DIGIT

Misalkan kita ingin mengambil sampel SRS $n=6$ dari populasi $N=60$. Pembacaan TAR dimulai dari halaman 1, baris 1, kolom 1.

$N=60$, $r=2$ (jumlah digit populasi).

Sampel terpilih:

- SRS WR: 57, 57, 26, 48, 22, 19
- SRS WOR: 57, 26, 48, 22, 19, 46

Baris	Kolom (1-5)
1	88 347
2	57 140
3	74 686
4	68 013
5	57 477
6	89 127
7	26 519
8	48 045
9	22 531
10	84 887
11	72 047
12	19 645
13	46 884
14	92289

SRSWR

Baris	Kolom (1-5)
1	88 347
2	57 140
3	74 686
4	68 013
5	57 477
6	89 127
7	26 519
8	48 045
9	22 531
10	84 887
11	72 047
12	19 645
13	46 884
14	92289

SRSWOR

Simple Random Sample

REMAINDER APPROACH

- Dari N unit populasi dan jumlah digits dari N adalah sebanyak r digit, maka tentukan nilai N' yaitu kelipatan terbesar dari N dengan jumlah digit yang sama. N' adalah batas atas dari angka random yang akan dipilih. Misal: $N=32$, $r=2$, $N'=96$
- Jika $AR \leq N$, maka unit yang nomor urutnya sama dengan AR tersebut terpilih sebagai sampel.
- Jika $AR = 0$, maka unit ke- N (terakhir) terpilih sebagai sampel.
- Jika $N < AR \leq N'$, maka lakukan operasi pembagian: $AR / N = k$ (sisa s)
- Unit dengan nomor urut = s terpilih sebagai sampel.
- Jika $s = 0$, unit ke- N (terakhir) terpilih sebagai sampel.
- Jika $AR > N'$, maka lanjutkan penelusuran ke angka random di baris selanjutnya pada kolom yang sama.

Simple Random Sample

REMAINDER APPROACH

Misalkan kita ingin mengambil sampel SRS WOR $n=3$ dari populasi $N=36$ dengan remainder approach. Pembacaan TAR dimulai dari halaman 1, baris 1, kolom 2.

$N=36$, $r=2$, $N' = 72$

Angka random:

- 83 → tolak, karena lebih dari N'
- 71 → $71/36 = 1$, sisa 35 (unit ke-35 terpilih sampel)
- 46 → $46/36 = 1$, sisa 10 (unit ke-10 terpilih sampel)
- 80 → tolak, karena lebih dari N'
- 74 → tolak, karena lebih dari N'
- 91 → tolak, karena lebih dari N'
- 65 → $65/36 = 1$, sisa 29 (unit ke-29 terpilih sampel)

Sampel terpilih: 35, 10, 29

Baris	Kolom (1-5)
1	8 8 347
2	5 7 140
3	7 4 686
4	6 8 013
5	5 7 477
6	8 9 127
7	2 6 519
8	48045
9	22531
10	84887
11	72047
12	19645
13	46884
14	92289

Simple Random Sample

QUOTIENT APPROACH

- Dari N unit populasi dan jumlah digits dari N adalah sebanyak r digit, maka tentukan nilai N' yaitu kelipatan terbesar dari N dengan jumlah digit yang sama. Misal: $N=32$, $r=2$, $N'=96$
- Hitung nilai $q = N'/N$
- Angka random (AR) yang diambil adalah mulai dari 0 sampai $(N' - 1)$
- Hitung $t = AR/q$ (pembulatan ke bawah)
- Sampel terpilih = unit dengan nomor urut $(t - 1)$
- Lakukan pengambilan AR sampai jumlah sampel terpenuhi
- Jika sudah sampai kolom terakhir dan belum mendapatkan AR sebanyak sampel, lanjutkan ke kolom berikutnya baris pertama

Simple Random Sample

QUOTIENT APPROACH

Misalkan kita ingin mengambil sampel SRS WOR $n=3$ dari populasi $N=36$ dengan quotient approach.

Pembacaan TAR dimulai dari halaman 1, baris 1, kolom 2.

$$N=36, r=2, N' = 72, (N' - 1) = 71, q = 72/36 = 2$$

Angka random:

- 83 → tolak, karena lebih dari $N' - 1$
- 71 → $t = 71/q = 71/2 = 35$ (unit ke-34 terpilih sampel)
- 46 → $t = 46/q = 46/2 = 23$ (unit ke-22 terpilih sampel)
- 80 → tolak, karena lebih dari $N' - 1$
- 74 → tolak, karena lebih dari $N' - 1$
- 91 → tolak, karena lebih dari $N' - 1$
- 65 → $t = 65/q = 65/2 = 32$ (unit ke-31 terpilih sampel)

Sampel terpilih: 34, 22, 31

Baris	Kolom (1-5)
1	8 8 347
2	5 7 1 40
3	7 4 6 86
4	6 8 0 13
5	5 7 4 77
6	8 9 1 27
7	2 6 5 19
8	48045
9	22531
10	84887
11	72047
12	19645
13	46884
14	92289

Estimasi Rata-Rata

Nilai yang diestimasi	SRS	
	WR	WOR
Rata-rata	$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$	
Varians rata-rata	$v(\bar{y}) = \frac{s^2}{n}$	$v(\bar{y}) = \frac{N-n}{N} \cdot \frac{s^2}{n}$
Standar error	$se(\bar{y}) = \sqrt{v(\bar{y})}$	
Relative standar error (RSE)	$rse(\bar{y}) = \frac{se(\bar{y})}{\bar{y}} \times 100\%$	
$(1 - \alpha)\%$ Confidence Interval	$\bar{y} - Z_{\alpha/2} \cdot se(\bar{y}) < \bar{Y} < \bar{y} + Z_{\alpha/2} \cdot se(\bar{y})$	

Estimasi Total

Nilai yang diestimasi	SRS	
	WR	WOR
Total	$\hat{Y} = \frac{N}{n} \sum_{i=1}^n y_i = N\bar{y}$	
Varians total	$v(\hat{Y}) = N^2 \cdot \frac{s^2}{n}$ $= N^2 \cdot v(\bar{y})$	$v(\hat{Y}) = N^2 \cdot \frac{N-n}{N} \cdot \frac{s^2}{n}$ $= N^2 \cdot v(\bar{y})$
Standar error	$se(\hat{Y}) = \sqrt{v(\hat{Y})}$	
Relative standar error (RSE)	$rse(\hat{Y}) = \frac{se(\hat{Y})}{\hat{Y}} \times 100\%$	
$(1 - \alpha)\%$ Confidence Interval	$\hat{Y} - Z_{\alpha/2} \cdot se(\hat{Y}) < Y < \hat{Y} + Z_{\alpha/2} \cdot se(\hat{Y})$	

Estimasi Proporsi

Nilai yang diestimasi	SRS	
	WR	WOR
Rata-rata	$p = \frac{1}{n} \sum_{i=1}^n y_i$	
Varians rata-rata	$v(p) = \frac{pq}{n-1}$	$v(p) = \frac{N-n}{N} \cdot \frac{pq}{n-1}$
Standar error	$se(p) = \sqrt{v(p)}$	
Relative standar error (RSE)	$rse(p) = \frac{se(p)}{p} \times 100\%$	
$(1 - \alpha)\%$ Confidence Interval	$p - Z_{\alpha/2} \cdot se(p) < P < p + Z_{\alpha/2} \cdot se(p)$	

Estimasi Proporsi

Nilai yang diestimasi	SRS	
	WR	WOR
Total	$\hat{A} = Np$	
Varians total	$v(\hat{A}) = N^2 \cdot \frac{pq}{n-1}$	$v(\hat{A}) = N^2 \cdot \frac{N-n}{N} \cdot \frac{pq}{n-1}$
Standar error	$se(\hat{A}) = \sqrt{v(\hat{A})}$	
Relative standar error (RSE)	$rse(\hat{A}) = \frac{se(\hat{A})}{\hat{A}} \times 100\%$	
$(1 - \alpha)\%$ Confidence Interval	$\hat{A} - Z_{\alpha/2} \cdot se(\hat{A}) < A < \hat{A} + Z_{\alpha/2} \cdot se(\hat{A})$	

Thank You