II. LANDASAN TEORI

Definisi 2.1 Sampling

Sampling adalah proses pengambilan atau memilih n buah elemen dari populasi yang berukuran N (Lohr, 1999).

Dalam melakukan sampling, terdapat teori dasar yang disebut teori sampling. Teori sampling mencoba mengembangkan metode/rancangan pemilihan sampel, sehingga dengan biaya sekecil mungkin dapat menghasilkan pendugaan parameter yang mendekati parameter populasinya. Teori sampling bertujuan untuk membuat sampling menjadi lebih efisien. Pengertian efisien dalam teori dasar sampling adalah rancangan sampling yang menghasilkan dugaan yang paling mendekati parameter populasi, membutuhkan biaya pengumpulan data yang sekecil-kecilnya (Cochran, 1991).

Rancangan sampling yang efisien adalah rancangan sampling yang dapat menghemat waktu, tenaga dan biaya tanpa mengurangi keakuratan data, dan informasi yang diperoleh benar-benar menggambarkan karakteristik populasi dengan baik.

Eriyanto (2007) mengemukakan bahwa pemakaian sampel akan berguna jika dapat digunakan sebagai alat pendugaan (inferensia). Nilai populasi disebut sebagai parameter, sementara nilai sampel disebut statistik.

Definisi 2.2 Teknik Sampling

Teknik sampling (teknik penarikan sampel) merupakan upaya penelitian untuk mendapatkan sampel yang representatif atau mewakili, yang dapat menggambarkan populasinya.

Yang termasuk dalam teknik sampling antara lain:

a. Sampling Acak Sederhana atau Simple Random Sampling

Sampling acak sederhana merupakan bentuk paling sederhana dari pengambilan sampel. Sampel acak sederhana dari n ukuran sampel diambil ketika setiap kemungkinan irisan (subset) dari n unit dalam populasi memiliki kesempatan yang sama untuk terpilih sebagai sampel. Sampel acak sederhana dapat digunakan apabila dalam satu populasi bersifat homogen (memiliki karakteristik populasi sama) .

Definisi 2.3 Sampling Acak Sederhana

Sampling acak sederhana adalah sampling acak, dimana setiap elemen memiliki peluang yang sama untuk dipilih dari populasi. Sampling acak sederhana dilakukan apabila:

1. Elemen populasi yang bersangkutan homogen (memiliki karakteristik populasi sama).
2. Hanya diketahui identitas-identitas dari satuan sampling (elemen) dalam populasi, sedangkan keterangan lain mengenai populasi, seperti tingkat keragaman, dan pembagian ke dalam golongan-golongan tidak diketahui.

(Hasan, 2001)

Dalam Cochran (1991) definisi sampling acak sederhana adalah sebuah metode untuk memilih n unit dari N sehingga setiap elemen dari ${}_NC_n$ sampel yang berbeda memepunyai kesempatan yang sama untuk dipilih.

Penduga rata-rata pada sampling acak sederhana adalah:

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

Penduga varian untuk $\hat{\mu}$ adalah:

$$V(\bar{y}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

Dimana:

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

b. Sampling Acak Berlapis atau *Stratified Random Sampling*

Definisi 2.4 Sampling Acak Berlapis

Sampling acak berlapis adalah bentuk sampling acak yang elemen populasinya dibagi kedalam kelompok-kelompok homogen yang disebut strata. Sampling acak berlapis dilakukan apabila:

1. Elemen-elemen populasinya heterogen (karakteristik populasinya tidak sama).
2. Ada kriteria yang digunakan sebagai dasar untuk menstratifikasikan elemen populasi ke dalam stratum-sratum.
3. Dapat diketahui dengan tepat jumlah unit/satuan samplingnya dari setiap stratum dalam populasi.

(Hasan, 2001)

Cochran (1991) mengemukakan bahwa dalam sampling acak berlapis, unit populasi (N) dibagi ke dalam sub populasi, masing-masing N_1, N_2,…., N_H. Sub populasi ini tidak boleh tumpang tindih, dan bila seluruh sub populasi dijumlahkan, maka diperoleh:

$N_1 + N_2 + \ldots + N_H = N$

Sub populasi disebut lapisan (strata). Untuk memperoleh keuntungan yang maksimal dari pelapisan (*stratification*), nilai N_H harus diketahui. Bila lapisan telah ditentukan, sebuah sampel diambil dari masing-masing lapisan secara acak untuk lapisan berbeda. Ukuran sampel dalam lapisan dinotasikan dengan:

$n_1 + n_2 + \ldots + n_H = n$

Alasan yang mendasar untuk menggunakan sampling acak berlapis dibandingkan sampling acak sederhana adalah:

a) Memaksimalkan informasi yang diperoleh. Stratifikasi memungkinkan untuk menghasilkan batas kesalahan sampling (sampling error) yang lebih kecil dari pada sampling acak sederhana dengan ukuran sampel yang sama. Jika hasil pengukuran di dalam suatu strata relative homogen.

b) Biaya observasi pada suatu survei bias dikurangi dengan melakukan stratifikasi pada elemen-elemen populasi ke dalam kelompok yang tepat.

c) Memungkinkan diterapkannya metode dan prosedur yang berbeda untuk setiap strata yang diambil.

Penduga rata-rata pada sampling acak berlapis adalah:

$$\bar{y} = \frac{1}{N} \sum_{i=1}^{L} N_i \bar{y}_i$$

Penduga varian untuk $\hat{\mu}$ adalah:

$$V(\bar{y}) = \frac{1}{N^2} \sum_{i=1}^{L} N_i^2 \left(\frac{N_i - n_i}{N_i}\right) \left(\frac{S_i^2}{n_i}\right)$$

$$S_i^2 = \frac{\sum_{i=1}^{n} (Y_{ij} - \bar{Y})^2}{n-1}$$

c. Sampling Kelompok atau *Cluster Sampling*

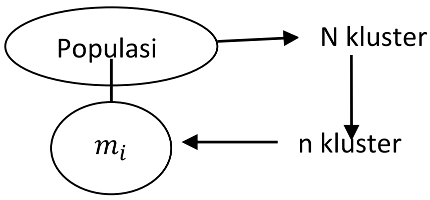
Definisi 2.5 Sampling Kelompok

Sampling kelompok adalah pengambilan sampel dari beberapa unit sampling yang merupakan kelompok dari elemen (Scheaffer, Mendenhall dan Ott., 1996). Sampling kelompok digunakan apabila populasi geografis elemen-elemen populasi berjauhan, keterbatasan biaya dan selain itu juga karena tidak tersedianya *sampling frame* secara lengkap, atau terlalu mahal untuk memperoleh *sampling frame*.

2.1 *One-Stage Cluster Sampling*

One-Stage Cluster Sampling membagi populasi menjadi kelompok atau kluster. Beberapa kluster kemudian dipilih secara acak sebagai wakil dari populasi, kemudian sampel diambil dari dalam kluster yang terpilih secara acak yang akan dijadikan sebagai sampel penelitian. *One-stage cluster sampling* dapat digunakan apabila posisi geografis elemen-elemen populasi berjauhan, tidak tersedianya *sampling frame* secara lengkap, dan untuk efisiensi biaya, mengingat besarnya biaya untuk melakukan survai.

Tahapan pada *one-stage cluster sampling* adalah sebagai berikut



Penduga bagi rata-rata populasi, yaitu:

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

Dengan $y_i = \sum_{i=1}^{m_i} y_{ij}$

Pada *one-stage cluster*, pemilihan kluster dilakukan secara acak sehingga penduga variansnya adalah sebagai berikut:

$$\hat{V}(\bar{y}_{one}) = \left(\frac{N-n}{Nn\bar{M}^2}\right) S^2$$

dengan:

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

Bukti:

Akan dibuktikan melalui penduga rata-rata pada persamaan 1:

$$\hat{V}(\bar{y}_{one}) = \mathrm{V}\left(\frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i}\right)$$

$$= \frac{1}{\left(\sum_{i=1}^{n} m_i\right)^2} V\left(\sum_{i=1}^{n} y_i\right) \qquad (2)$$

Diketahui bahwa $\bar{m} = \frac{\sum_{i=1}^{n} m_i}{n}$, sehingga $\sum_{i=1}^{n} m_i = n\bar{m}$. Untuk kasus *one-stage cluster sampling*, $m_i = M_i$, sehingga $\bar{M} = \frac{\sum_{i=1}^{n} m_i}{n}$ atau $\sum_{i=1}^{n} m_i = n\bar{M}$. Dengan demikian, persamaan (2) menjadi:

$$\hat{V}(\bar{y}_{one}) = \frac{1}{n^2\bar{M}^2} V\left(\sum_{i=1}^{n} y_i\right)$$

$$= \frac{1}{n^2\bar{M}^2} V(n\,\bar{y}^*) \text{ : dengan } \bar{y}^* = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$= \frac{1}{n^2\bar{M}^2}\; n^2 V(\bar{y}^*)$$

$$= \frac{1}{\bar{M}^2}\; V(\bar{y}^*) \qquad (3)$$

Perhatikan bahwa $\bar{y}^* = \frac{\sum_{i=1}^{n} y_i}{n}$, dengan y_i adalah $\sum_{j=1}^{m_i} y_{ij}$ atau merupakan total pengamatan pada kluster terpilih ke-i. Dengan demikian, penduga total bagi kluster ke-i adalah $\bar{y}m_i$. Mengikuti konsep penduga varians pada kasus *simple random sampling*, maka diperoleh:

$$V(\bar{y}^*) = \left(\frac{N-n}{N}\right)\frac{S^2}{n} \qquad (4)$$

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

Dengan mensubtitusikan persamaan (4) ke dalam persamaan (3), maka diperoleh

$V(\bar{y}^*) = \frac{1}{\bar{M}^2}\left(\frac{N-n}{N}\right)\frac{S^2}{n}$ atau

$$= \frac{N-n}{Nn\bar{M}^2}\, S^2 \qquad \text{(terbukti).}$$

Catatan:

N= jumlah kluster

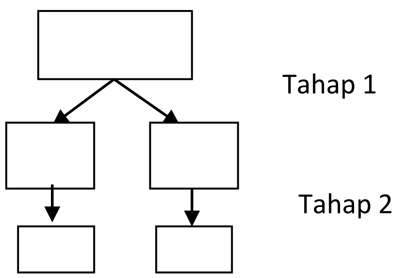
n= jumlah kluster terpilih

$m_i = M_i$= jumlah elemen/unit sampel kluster terpilih ke-i

2.2 *Two-stage Cluster sampling*

Metode *two-stage kluster sampling* merupakan pengembangan dari metode kluster sampling dimana pengambilan sampel dilakukan secara dua tahap, yaitu tahap pertama, memilih beberapa kluster dalam populasi secara acak dan tahap kedua memilih beberapa unit sampel dari tiap kluster terpilih secara acak (Scheafer et.al., 1990).

Tahapan pada *two-stage cluster sampling* dapat diilustrasikan sebagai berikut:



Notasi:

N = Jumlah kluster dalam populasi

n = Jumlah kluster terpilih

M_i = Jumlam elemen/unit sampling dari kluster ke-i

m_i = Jumlam elemen/unit sampling yang dipilih dari kluster terpilih ke-i

M = $\sum_{i=1}^{N} M_i$ = jumlah elemen/unit sampling dalam populasi

$\bar{M} = \frac{M}{N}$ = rata-rata jumlah elemen/unit sampling masing-masing kluster

Penduga bagi rata-rata populasi yaitu

$$\bar{y}_{two} = \frac{1}{\bar{M}} \frac{\sum_{i=1}^{n} M_i \, \bar{y}_i}{n}$$

Bukti:

Akan dibuktikan melalui pendekatan total populasi. Diketahui bahwa total populasi = $N\bar{y}$ dan M merupakan total jumlah elemen dalam populasi, maka:

$$\bar{y}_{two} = \frac{N \, \bar{y}_{total}}{M}$$

$$= \frac{N}{M} \frac{\sum_{i=i}^{n} y_i}{n} \qquad (5)$$

Dengan y_i merupakan total pengamatan dari kluster terpilih ke-i:

$$y_i = \sum_{j=1}^{m_i} y_{ij}$$

$$= M_i \bar{y}_i$$

Dengan demikian persamaan (5) menjadi:

$$\bar{y}_{two} = \frac{1}{\bar{M}} \frac{\sum_{i=1}^{n} M_i \, \bar{y}_i}{n} \qquad \text{(terbukti)}$$

Penduga varians bagi penduga rata-rata populasi adalah:

$$\hat{V}(\bar{y}_{two}) = \left(\frac{N-n}{N}\right) \left(\frac{1}{n\bar{M}^2}\right) S_b^2 + \frac{1}{nN} \sum_{i=1}^{n} \left(\frac{\bar{M}-m}{\bar{M}}\right) \frac{S_2^2}{m}$$

Bukti:

Menurut scheaffer, et.al (1996) penduga ragam bagi penduga rata-rata populasi untuk kasus *two-stage cluster sampling* dapat diuraikan sebagai berikut:

$V(\bar{y}_{two}) = V_1[E_2(\bar{y}_{two})\ ] + E_1[V_2(\bar{y}_{two})]$

Pertama menguraikan:

$$V_1[E_2(\bar{y}_{two})\ ] = V_1\left[\frac{1}{n}\sum_{i=1}^{n}\bar{y}_i\right]$$

$$= V_1[\bar{y}]$$

$$= \frac{N-n}{N}\frac{S_1^2}{n}$$

dengan $S_1^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\bar{y}_i - \bar{y})^2$

menurut scheaffer, et.al (1996) dengan menguraikan,

$$\frac{1}{\bar{M}^2}\ S_b^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\bar{y}_i - \mu)^2$$

maka diperoleh $\left(\frac{N-n}{N}\right)\left(\frac{1}{n\bar{M}^2}\right) S_b^2$.

Kemudian yang kedua menguraikan:

$$E_1[V_2(\bar{y}_{two})] = E_1\left(V_2\left(\frac{1}{n}\sum_{i=1}^{n}\bar{y}_i\right)\right)$$

$$= E_1\left[\frac{1}{n^2}\left(\sum_{i=1}^{n}\bar{y}_i\right)\right]$$

$$= E_1\left(\frac{1}{n^2}\left(V_2(\bar{y}_1) + V_2(\bar{y}_2) + \cdots + V_2(\bar{y}_n)\right)\right)$$

$$= E_1\left(\frac{1}{n^2}\ n\,V(\bar{y}_i)\right)$$

$$= E_1\left(\frac{1}{n}\ V(\bar{y}_i)\right)$$

$$= \frac{1}{n}E_1[V(\bar{y}_i)]$$

$$= \frac{1}{n}\frac{1}{N}\sum_{i=1}^{n} V(\bar{y}_i)$$

$$= \frac{1}{n}\frac{1}{N}\sum_{i=1}^{n}\frac{\bar{M}-m}{\bar{M}}\,\frac{S_2^2}{m}$$

dengan $S_2^2 = \frac{\sum_{i=1}^{n}\left(y_{ij}-\bar{y}_i\right)^2}{m-1}$

Dengan demikian diperoleh penduga varians bagi penduga rata-rata populasinya adalah:

$$\hat{V}(\bar{y}_{two}) = \left(\frac{N-n}{N}\right)\left(\frac{1}{n\bar{M}^2}\right)S_b^2 + \frac{1}{nN}\sum_{i=1}^{n}\left(\frac{\bar{M}-m}{\bar{M}}\right)\frac{S_2^2}{m}$$