

# PENAKSIRAN RATAAN DAN VARIANS POPULASI PADA SAMPEL ACAK TERSTRATIFIKA DENGAN *AUXILIARY VARIABLE*

Raupong, M. Saleh AF, Hasruni Satya Taruma

## Abstrak

Penaksiran rata-rata dan variansi suatu populasi dilakukan melalui sampel yang diharapkan hasilnya dapat mewakili karakteristik populasi tersebut. Teknik pengambilan sampel yang baik akan memberikan hasil penelitian yang baik pula. Kebanyakan peneliti menggunakan teknik pengambilan sampel acak tetapi hasil penaksiran dengan teknik ini masih menghasilkan variansi yang besar dan tidak sesuai untuk jumlah sampel yang besar sehingga dibutuhkan teknik pengambilan sampel lain, salah satunya adalah pengambilan sampel acak terstratifikasi. Teknik ini masih lebih baik dibandingkan dengan sampel acak sederhana karena menghasilkan variansi yang lebih minimum dan sesuai untuk jumlah sampel yang besar. Tetapi dengan hanya mengandalkan teknik pengambilan sampel, masih tidak menjamin hasil penaksiran akan lebih baik sehingga dalam penelitian ini ditambahkan beberapa teknik lain seperti penambahan *auxiliary variable* dan alokasi optimum dalam penentuan jumlah sampel. Adanya penggabungan antara teknik pengambilan sampel yang distratifikasi, penambahan *auxiliary variable*, serta alokasi optimum memperlihatkan hasil penaksiran yang lebih baik dalam menggambarkan karakteristik populasi melalui sampel.

**Kata kunci:** Sampel acak sederhana, sampel acak terstratifikasi, *auxiliary variable*, alokasi optimum.

## 1. Pendahuluan

Dalam kehidupan kita sehari-hari tidak terlepas dari statistika. Pada dasarnya statistika berhubungan dengan data, fakta, dan informasi. Salah satu kegiatan yang utama dalam statistika adalah pengumpulan data. Pelaksanaan pengumpulan data selalu berhadapan dengan objek yang diteliti atau yang diselidiki. Meskipun penelitian hanya mengambil sebagian dari seluruh objek yang diteliti, tetapi hasilnya dapat mewakili atau mencakup seluruh objek yang diteliti.

Pengambilan sampel penelitian ini diterapkan suatu cara atau teknik-teknik tertentu, sehingga sampel tersebut sedapat mungkin mewakili populasinya. Teknik ini biasa disebut "teknik sampling". Dalam penelitian survei, teknik sampling ini sangat penting dan perlu diperhatikan dengan seksama sebab teknik pengambilan sampel yang baik akan mempengaruhi validitas hasil penelitian tersebut. Salah satu teknik pengambilan sampel yang biasa orang pakai adalah pengambilan sampel acak terstratifikasi (*stratified random sampling*). Teknik ini digunakan apabila peneliti berpendapat bahwa ada perbedaan ciri, atau karakteristik antara strata-strata yang ada, dan perbedaan strata tersebut mempengaruhi variabel. Apabila anggota populasi terbagi atas tingkat-tingkat atau strata maka pengambilan sampel tidak boleh mengabaikan adanya strata dan setiap strata harus diwakili sebagai sampel.

Tujuan utama kita mengambil sampel dari suatu populasi adalah memperoleh informasi mengenai parameter populasi atau singkatnya untuk mengetahui parameter populasi itu. Dalam statistika inferensia, jika parameter populasi tidak diketahui, maka untuk mengetahuinya

dilakukan penaksiran dari statistik suatu sampel dan diharapkan hasilnya mendekati apa yang sebenarnya terjadi pada populasi.<sup>1</sup>

Dalam melakukan penaksiran, tentunya terdapat bias pada hasil akhir dan diharapkan bias tersebut memiliki hasil yang sangat minimum karena setiap peneliti pastinya mempunyai harapan agar hasil penelitiannya menghasilkan penaksiran yang baik.

Salah satu cara dalam meminimumkan bias penaksiran adalah dengan menambahkan suatu variabel dan variabel tersebut biasa disebut dengan *auxiliary variable*. *Auxiliary variable* adalah variabel tambahan yang biasanya para peneliti ingin menggunakannya dalam metode estimasi untuk memperoleh penaksir yang paling baik.

Tujuan dari penulisan ini adalah menentukan penaksiran rata-rata dan variansi pada sampel acak terstratifikasi, membandingkan hasil penaksiran rata-rata dan variansi sederhana dengan hasil penaksiran dengan *auxiliary variable* dan alokasi optimum pada sampel acak terstratifikasi serta Menentukan hasil penaksiran variansi menggunakan *Mean Square Error*.

## 2. Tinjauan Pustaka

### 2.1 Populasi dan Sampel

Populasi adalah keseluruhan subjek penelitian. Apabila seseorang ingin meneliti semua elemen yang ada dalam wilayah penelitian, maka penelitiannya merupakan penelitian populasi atau disebut studi populasi, atau juga studi sensus. Penelitian populasi dilakukan apabila peneliti ingin melihat semua atribut yang terdapat di dalam populasi. Oleh karena semua subjeknya meliputi semua yang terdapat dalam populasi maka juga disebut sensus.

Sampel adalah anggota dari populasi. Sampel dapat dipilih berdasarkan random, strata atau wilayah tetapi didasarkan atas adanya tujuan tertentu.

### 2.2 Beberapa Metode Sampling

Pengambilan sampel harus dilakukan sedemikian rupa sehingga diperoleh sampel yang benar-benar dapat menggambarkan keadaan populasi yang sebenarnya atau *representative*.

Dari beberapa cara pengambilan sampel yang telah dipaparkan tetapi hanya dua yang akan dibahas pada penulisan kali ini yaitu Pengambilan Sampel Acak Sederhana (*Simple Random Sampling*) dan Pengambilan Sampel Terstratifikasi (*Stratified Sampling*).

#### 2.2.1 Pengambilan Sampel Acak Sederhana (*Simple Random Sampling*)

Pengambilan sampel acak sederhana adalah sebuah metode untuk memilih  $n$  unit dari  $N$  sehingga setiap elemen dari  ${}_N C_n$  sampel yang berbeda mempunyai kesempatan yang sama untuk dipilih. Dalam praktek, penarikan sampel acak sederhana dipilih unit per unit. Unit-unit dalam populasi diberi nomor dari 1 sampai  $N$ . Serangkaian bilangan acak 1 dan  $N$  kemudian dipilih, dengan cara menggunakan sebuah tabel bilangan acak atau dengan cara menggunakan sebuah program komputer yang menghasilkan tabel bilangan acak. Pada setiap penarikan, proses yang digunakan harus memberikan kesempatan terpilih yang sama untuk setiap bilangan dalam populasi. Unit-unit yang terpilih ini sebanyak  $n$  merupakan sampel.

---

<sup>1,2,3</sup> Jurusan Matematika FMIPA Universitas Hasanuddin Makassar, Jl. Perintis Kemerdekaan Km.10 Makassar

### 2.2.2 Pengambilan Sampel Terstratifikasi (*Stratified Sampling*)

#### ➤ Uraian

Dalam penarikan sampel berlapis populasi  $N$  unitnya pertama-tama dibagi ke dalam subpopulasi, masing-masing  $N_1, N_2, \dots, N_L$  unit. Subpopulasi ini tidak boleh tumpang-tindih dan bila seluruh subpopulasi dijumlahkan, maka diperoleh

$$N_1 + N_2 + \dots + N_L = N$$

Subpopulasi 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, pengambilan dilakukan secara acak untuk setiap lapisan. Ukuran sampel lapisannya dinotasikan dengan  $n_1, n_2, \dots, n_L$  dan  $x_{hi}$  adalah nilai dalam sampel ke- $h$  pada pengamatan ke- $i$ .

Jika sebuah sampel acak sederhana diambil dari setiap lapisan, seluruh prosedur digambarkan sebagai *penarikan sampel acak terstratifikasi (berlapis)*.

#### ➤ Notasi

Akhiran  $h$  menunjukkan lapisan dan  $i$  unit di dalam lapisan tersebut. Notasinya merupakan perluasan dari yang digunakan sebelumnya. Seluruh simbol-simbol berikut menunjukkan pada lapisan ke- $h$ .

$N_h$	Jumlah unit lapisan yang terpilih
$n_h$	Jumlah unit dalam sampel pada lapisan terpilih
$y_{hi}$	Nilai yang diperoleh untuk unit ke- $i$
$W_h = \frac{N_h}{N}$	Bobot lapisan
$f_h = \frac{n_h}{N_h}$	Fraksi penarikan sampel di dalam lapisan
$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h}$	Rata-rata populasi (rata-rata sebenarnya)
$\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$	Rata-rata sampel
$S_h^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2}{N_h - 1}$	Varians pada lapisan ke- $h$ yang terpilih

#### ➤ Sifat-sifat Penaksiran

Penaksiran yang digunakan dalam penarikan sampel berlapis adalah  $\bar{y}_{st}$  di mana

$$\bar{y}_{st} = \frac{\sum_{h=1}^L N_h \bar{y}_h}{N} = \sum_{h=1}^L W_h \bar{y}_h \quad (1)$$

dengan  $N = N_1 + N_2 + \dots + N_L$

Rata-rata sampel,  $\bar{y}$  dapat ditulis sebagai berikut

$$\bar{y} = \frac{\sum_{h=1}^L n_h \bar{y}_h}{n} \quad (2)$$

Pembuktian telah memperlihatkan bahwa  $\bar{y}_{st}$  adalah penaksiran tak bias dari rata-rata populasi  $\bar{Y}$  maka dapat dituliskan

$$E(\bar{y}_{st}) = \bar{Y} = \sum_{h=1}^L W_h \bar{Y}_h \quad (3)$$

Sehingga variansi dari penaksiran  $\bar{y}_{st}$  untuk penarikan sampel acak terstratifikasi adalah

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h} = \sum_{h=1}^L W_h^2 \frac{S_h^2}{n_h} (1 - f_h) \quad (4)$$

[1]

### 2.3 Variabel Tambahan (*Auxiliary Variabel*)

Selain variabel penelitian  $y_i$ , satu atau lebih variabel tambahan  $x_i$  dapat berhubungan dengan unit  $i$  dari populasi. Misalnya, jika variabel penelitian adalah volume pohon, diameter payudara atau "bola mata" penaksiran volume dapat berfungsi sebagai variabel tambahan.

Informasi tambahan dapat digunakan baik dalam desain sampling atau estimasi.

Pada tahap estimasi, hubungan antara  $y_i$  dan  $x_i$  kadang-kadang dapat dimanfaatkan untuk menghasilkan penaksiran yang lebih tepat daripada yang dapat diperoleh dari data  $y$  saja. Rasio dan penduga regresi adalah contoh dari penggunaan variabel tambahan dalam penaksiran [6].

### 2.4 Mean Square Error

Untuk membandingkan sebuah penaksir bias dengan penaksir tidak bias, atau dua penaksir dengan jumlah bias yang berbeda, maka suatu kriteria yang berguna adalah dengan menghitung rata-rata kuadrat galat (*mean square error*) pada penaksir yang diukur dari nilai populasi yang ditaksirkan. Secara umum,

$$\begin{aligned} MSE(\bar{y}) &= E(\bar{y} - \bar{Y})^2 = E(\bar{y}^2 - 2\bar{y}\bar{Y} + \bar{Y}^2) \\ &= E(\bar{y}^2) + E(\bar{Y}^2) - 2E(\bar{y}\bar{Y}) \\ &= var(\bar{y}) + (bias(\bar{y}))^2 \end{aligned}$$

### 2.5 Alokasi Optimum (*Optimum Allocation*)

Dalam penarikan sampel acak terstratifikasi, nilai dari ukuran sampel  $n_h$  dalam masing-masing strata dipilih oleh pengambil sampel, alasannya mungkin memilih untuk meminimumkan  $V(\bar{y}_{st})$  dengan biaya tertentu. Biaya dalam setiap stratum  $c_h$  adalah proporsional dengan ukuran sampel [1].

Alokasi optimum strata ini menghasilkan alokasi penaksiran dengan varians terkecil yang mungkin untuk total biaya survei. Hal-hal yang perlu diketahui sebelum menentukan alokasi optimum untuk  $n_h$  adalah tiap sampel telah terbagi ke dalam masing-masing strata, rata-rata biaya per unit  $c_h$ , standar deviasi  $S_h$  dan bobot  $W_h$ .

Maka alokasi yang memberikan  $V(\bar{y}_{st})$  minimum adalah

$$n_h = n \frac{N_h S_h / \sqrt{c_h}}{\sum_{h=1}^L N_h S_h / \sqrt{c_h}} \quad (5)$$

Alokasi optimum memberikan nilai  $n_h$  baru yang kemudian akan digunakan untuk lebih meminimumkan  $V(\bar{y}_{st})$ .

### 2.6 Penaksir Regresi untuk Sampel Acak Sederhana

Misalkan terdapat populasi terbatas berukuran  $N$  dimana  $s$  sampel berukuran  $n$  diambil secara acak sederhana tanpa penggantian. Kita misalkan bahwa  $y_i$  dan  $x_i$  masing-masing adalah variabel penelitian dan variabel tambahan yang diperoleh untuk setiap unit dalam sampel dimana  $i = 1, 2, \dots, N$  dan rata-rata populasi  $\bar{X}$  dari  $x_i$  diketahui maka penaksiran regresi linear  $\bar{Y}$  dari rata-rata populasi  $y_i$ , adalah

$$\bar{y}_{lr} = \bar{y} + z(\bar{X} - \bar{x}) \quad (6)$$

dimana notasi  $lr$  menyatakan *regresi linear* dan  $z$  adalah suatu koefisien penaksiran dari perubahan dalam  $y$  bila  $x$  meningkat.

Perhitungan  $z$  menggunakan metode penaksiran kuadrat terkecil dari  $Z$  (*koefisien regresi populasi*), maka

$$z = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (7)$$

**Teorema 7.3** dalam [1] Bila  $z$  adalah penaksiran kuadrat terkecil dari  $Z$  dan,

$$\bar{y}_{lr} = \bar{y} + z(\bar{X} - \bar{x}) \quad (8)$$

kemudian dalam sampel acak sederhana berukuran  $n$ , dengan  $n$  besar,

$$V(\bar{y}_{lr}) = \frac{(1-f)}{n} S_y^2 (1 - \rho^2) \quad (9)$$

dimana  $\rho = S_{yx}/S_y S_x$  adalah korelasi populasi antara  $y$  dan  $x$ , maka didapatkan

$$E(\bar{y}_{lr} - \bar{Y})^2 = V(\bar{y}_{lr}) = \frac{(1-f)}{n} S_y^2 (1 - \rho^2) \quad (10)$$

[1]

Bukti bahwa  $z$  dapat meminimumkan variansi adalah karena  $z_0$  adalah koefisien regresi linear dari  $y$  pada  $x$  dalam populasi terbatas.  $B$  tidak tergantung pada sifat dari setiap sampel yang diambil dan secara teoritis dapat ditentukan lebih dulu.

## 2.7 Penaksir Regresi untuk Sampel Acak Terstratifikasi

Terdapat dua jenis penaksiran regresi dalam pengambilan sampel acak terstratifikasi. Penaksiran pertama yaitu penaksiran regresi terpisah  $\bar{y}_{lrs}$  (huruf  $s$  berarti terpisah) dihitung untuk setiap rata-rata lapisan, yaitu

$$\bar{y}_{lrh} = \bar{y}_h + z_h(\bar{X}_h - \bar{x}_h) \quad (11)$$

Kemudian, dengan  $W_h = N_h/N$ ,

$$\bar{y}_{lrs} = \sum_h W_h \bar{y}_{lrh} \quad (12)$$

Penaksiran ini mempertimbangkan bahwa koefisien regresi  $Z_h$  bervariasi dari lapisan ke lapisan. Varians minimum dari  $\bar{y}_{lrs}$  didapatkan bila  $z_h = Z_h$ , koefisien regresi sebenarnya dalam lapisan  $h$ , sehingga dapat dituliskan sebagai berikut

$$V_{min}(\bar{y}_{lrs}) = \sum_h \frac{W_h^2(1-f_h)}{n_h} \left( S_{yh}^2 - \frac{S_{y_xh}^2}{S_{xh}^2} \right) \quad (13)$$

Penaksiran regresi yang kedua adalah  $\bar{y}_{lrc}$  (huruf  $c$  berarti gabungan). Penaksiran ini menganggap  $Z_h$  sama untuk seluruh lapisan.

$$\bar{y}_{lrc} = \bar{y}_{st} + z(\bar{X} - \bar{x}_{st}) \quad (14)$$

Sedangkan penaksiran dari  $V(\bar{y}_{lrc})$  adalah sebagai berikut

$$v(\bar{y}_{lrc}) = \sum_h \frac{W_h^2(1-f_h)}{n_h(n_h-1)} \sum_h [(y_{hi} - \bar{y}_h) - z_c(x_{hi} - \bar{x}_h)]^2 \quad (15)$$

[1]

## 3. Hasil dan Pembahasan

### 4.1 Penaksiran Parameter Rataan $\bar{y}$ dan Variansi Populasi $V(\bar{y}_{lr})$ pada Sampel Acak Sederhana

#### 4.1.1 Penaksiran rata-rata $\bar{y}$

Rataan sampel  $\bar{y}$  adalah penaksiran tak bias dari rata-rata populasi  $\bar{Y}$ .

*Bukti:*

$$E(\bar{y}) = \frac{(N-1)!}{(n-1)!(N-n)!} \frac{n!(N-n)!}{nN!} (y_1 + y_2 + \dots + y_N) \\ = \bar{Y}$$

Sama halnya dengan  $\bar{x} = \bar{X}$

#### 4.1.2 Penaksiran $z$

Penaksiran  $z$  sebagai koefisien regresi pada pers. regresi  $y_i = \bar{y} + z(x_i - \bar{x}) + \varepsilon_i$  ditaksir dengan metode kuadrat terkecil.

*Bukti:*

$$\varepsilon = y_i - \bar{y} - z(x_i - \bar{x})$$

$$l = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \bar{y} - z(x_i - \bar{x}))^2$$

$$\frac{dl}{dz} = 2 \sum_{i=1}^n (y_i - \bar{y} - z(x_i - \bar{x}))(-x_i + \bar{x})$$

$$z = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

#### 4.1.3 Penaksiran variansi populasi $V(\bar{y}_{tr})$

Sebelum membuktikan Penaksiran variansi  $V(\bar{y}_{tr})$  terlebih dahulu persamaan varians dari rataan  $\bar{y}$  untuk sampel acak sederhana yang akan dibuktikan yaitu

$$V(\bar{y}) = E(\bar{y} - \bar{Y})^2 = \frac{S^2}{n} (1 - f), \text{ di mana } S^2 \text{ adalah variansi populasi}$$

*Bukti:*

$$\begin{aligned} n(\bar{y} - \bar{Y}) &= (y_1 - \bar{Y}) + (y_2 - \bar{Y}) + \dots + (y_n - \bar{Y}) & (11.1) \\ E[(y_1 - \bar{Y})^2 + (y_2 - \bar{Y})^2 + \dots + (y_n - \bar{Y})^2] \\ &= \frac{n}{N} [(y_1 - \bar{Y})^2 + (y_2 - \bar{Y})^2 + \dots + (y_N - \bar{Y})^2] \\ E[(y_1 - \bar{Y})(y_2 - \bar{Y}) + (y_1 - \bar{Y})(y_3 - \bar{Y}) + \dots + (y_{n-1} - \bar{Y})(y_n - \bar{Y})] \\ &= \frac{n(n-1)}{N(N-1)} [(y_1 - \bar{Y})(y_2 - \bar{Y}) + (y_1 - \bar{Y})(y_3 - \bar{Y}) + \dots \\ &\quad + (y_{N-1} - \bar{Y})(y_N - \bar{Y})] \end{aligned}$$

Kuadratkan pers.(11.1) kemudian dikalikan dengan  $\frac{n}{N}$  diperoleh

$$\begin{aligned} &= \frac{n}{N} \left\{ (y_1 - \bar{Y})^2 + (y_2 - \bar{Y})^2 + \dots + (y_N - \bar{Y})^2 \right. \\ &\quad \left. + 2 \frac{(n-1)}{(N-1)} [(y_1 - \bar{Y})(y_2 - \bar{Y}) + (y_1 - \bar{Y})(y_3 - \bar{Y}) + \dots \right. \\ &\quad \left. + (y_{N-1} - \bar{Y})(y_N - \bar{Y})] \right\} \\ &= \frac{n}{N} \left\{ (y_1 - \bar{Y})^2 + \dots + (y_n - \bar{Y})^2 + 2 \frac{(n-1)}{(N-1)} (y_1 - \bar{Y})(y_2 - \bar{Y}) + \dots \right. \\ &\quad \left. + (y_{N-1} - \bar{Y})(y_N - \bar{Y}) \right\} \end{aligned}$$

Karena  $\sum_{i=1}^N y_i = N\bar{Y}$ , maka

$$\begin{aligned} &= \frac{n}{N} \left\{ \left( \frac{(N-1) - (n-1)}{(N-1)} \right) \left( \sum_{i=1}^N (y_i - \bar{Y})^2 \right) + \frac{(n-1)}{(N-1)} [N\bar{Y} - N\bar{Y}]^2 \right\} \\ n^2(\bar{y} - \bar{Y})^2 &= \frac{n}{N} \left\{ \left( \frac{N-n}{N-1} \right) \sum_{i=1}^N (y_i - \bar{Y})^2 \right\} & (11.2) \end{aligned}$$

Setelah dibagi dengan  $n^2$

$$\begin{aligned} E(\bar{y} - \bar{Y})^2 &= \frac{1}{Nn} \left( \frac{N-n}{N-1} \right) \sum_{i=1}^N (y_i - \bar{Y})^2 & (11.3) \\ &= \frac{S^2}{n} \frac{N-n}{N} \end{aligned}$$

Karena  $f = n/N$ , maka

$$V(\bar{y}) = \frac{S^2}{n} (1 - f)$$

Pers.(11) merupakan pengembangan dari  $V(\bar{y}) = \frac{S^2}{n} (1 - f)$  dengan penambahan  $z_0$  sebagai konstanta yang ditentukan dari sampel, berikut adalah pembuktiannya

$$\begin{aligned} V(\bar{y}_{lr}) &= \frac{(1-f) \sum_{i=1}^N [(y_i - \bar{Y}) - z_0(x_i - \bar{X})]^2}{n(N-1)} \\ &= \frac{(1-f)}{n} (S_y^2 - 2z_0 S_{yx} + z_0^2 S_x^2) \end{aligned}$$

Nilai  $z_0$  yang meminimumkan  $V(\bar{y}_{lr})$  adalah

$$\begin{aligned} z_0 = z &= \frac{S_{yx}}{S_x^2} = \frac{\sum_{i=1}^N (y_i - \bar{Y})(x_i - \bar{X})}{(x_i - \bar{X})^2} \\ &= \frac{(1-f)}{n} S_y^2 \left( 1 - \frac{S_{yx}^2}{S_y^2 S_x^2} \right) \end{aligned}$$

Dimana  $\frac{S_{yx}}{S_y S_x} = \rho$ , maka

$$V(\bar{y}_{lr}) = \frac{(1-f)}{n} S_y^2 (1 - \rho^2)$$

## 4.2 Penaksiran Parameter Rataan $\bar{y}_{lrs}$ dan Variansi $V(\bar{y}_{lrs})$ Populasi pada Sampel Acak Terstratifikasi (Penaksiran Terpisah)

### 4.2.1 Penaksiran rata-rata populasi $\bar{y}_{lrs}$

Penaksiran ini sama halnya dengan penaksiran sampel acak sederhana karena penaksiran dilakukan per strata dan sampelnya dipilih secara acak kemudian hasil penaksiran rata-rata masing-masing strata digabungkan untuk menghasilkan

$$\bar{y}_{lrs} = \sum_h W_h \bar{y}_{lrh} = \bar{Y}$$

Bukti:

$$\begin{aligned} E(\bar{y}_{lrs}) &= E\left(\sum_{h=1}^L W_h \bar{y}_h\right) = \sum_{h=1}^L W_h (E(\bar{y}_h)) \\ &= \frac{1}{N} \sum_{h=1}^L Y_h = \bar{Y} \end{aligned}$$

### 4.2.2 Penaksiran variansi populasi $V(\bar{y}_{lrs})$

Penaksiran ini berdasarkan pers. (11.3) dengan penggantian langsung variabel-variabelnya, berikut adalah pembuktiannya.

$$\begin{aligned} E(\bar{y}_{lrs} - \bar{Y})^2 &= \frac{N_h - n_h}{N_h n_h} W_h^2 \frac{\sum_{i=1}^N (\bar{y}_{lrh} - \bar{Y})^2}{N-1} \\ &= \frac{N_h - n_h}{N_h} W_h^2 \frac{S_{yh}^2}{n_h} \end{aligned}$$

di mana  $f_h = n_h/N_h$ , sehingga

$$= \frac{(1-f_h) W_h^2}{n_h} S_{yh}^2$$

Persamaan di atas mengalami pengembangan dengan penambahan  $z_h$ . Nilai  $z_h$  yang meminimumkan variansi adalah  $z_h = \frac{S_{y_x h}}{S_{x_h}^2} = \frac{\sum_i (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h)}{\sum_i (x_{hi} - \bar{x}_h)^2}$ , sehingga

$$\begin{aligned}
V(\bar{y}_{lrs}) &= \frac{(1-f_h)W_h^2 \sum_{i=1}^{N_h} [(y_{hi} - \bar{y}_h) - z_h(x_{hi} - \bar{x}_h)]^2}{n_h(N_h - 1)} \\
&= \frac{(1-f_h)W_h^2}{n_h} \left( S_{yh}^2 - \frac{S_{yxh}^2}{S_{xh}^2} \right)
\end{aligned}$$

### 4.3 Penaksiran Parameter Rataan $\bar{y}_{lrc}$ dan Variansi $V(\bar{y}_{lrc})$ Populasi pada Sampel Acak Terstratifikasi (Penaksiran Gabungan)

#### 4.3.1 Penaksiran rata-rata $\bar{y}_{lrc}$

$$\bar{y}_{lrc} = \bar{y}_{st} + z(\bar{X} - \bar{x}_{st})$$

$$E(\bar{y}_{st}) = E\left(\sum_{h=1}^L W_h \bar{y}_h\right)$$

$$\begin{aligned}
&= \sum_{h=1}^L \frac{N_h}{Nn_h} n_h \bar{y}_h = \frac{1}{N} \sum_{h=1}^L Y_h \\
&= \bar{Y}
\end{aligned}$$

#### 4.3.2 Penaksiran variansi $V(\bar{y}_{lrc})$

Penaksiran variansi gabungan ini sama dengan penaksiran variansi terpisah yaitu  $E(\bar{y}_{lrc} - \bar{Y})^2 = \frac{(1-f_h)W_h^2}{n_h} S_{yh}^2$  tetapi yang membedakan adalah koefisien regresi  $z_c$ . Di mana  $z_c$  didapatkan dari penaksiran kuadrat terkecil dari persamaan regresi

$\bar{y}_{lrc} = \bar{y}_{st} + z(\bar{X} - \bar{x}_{st})$  dan didapatkan

$$\begin{aligned}
Z_c &= \frac{\sum_h \frac{W_h^2(1-f_h)S_{yxh}}{n_h}}{\sum_h \frac{W_h^2(1-f_h)S_{xh}^2}{n_h}} \\
&= \frac{\sum_h \frac{W_h^2(1-f_h)}{n_h(n_h-1)} \sum_i (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h)}{\sum_h \frac{W_h^2(1-f_h)}{n_h(n_h-1)} \sum_i (x_{hi} - \bar{x}_h)^2}
\end{aligned}$$

Kemudian substitusi nilai  $z_c$  ke persamaan variansi  $E(\bar{y}_{lrc} - \bar{Y})^2 = \frac{(1-f_h)W_h^2}{n_h} S_{yh}^2$  sehingga didapatkan persamaan

$$V(\bar{y}_{lrc}) = \frac{(1-f_h)W_h^2}{n_h(n_h-1)} \sum_{i=1}^{N_h} [(y_{hi} - \bar{y}_h) - z_c(x_{hi} - \bar{x}_h)]^2$$

### 4.4 Pembuktian persamaan MSE

$$\begin{aligned}
E(\bar{y} - \bar{Y})^2 &= E(\bar{y}^2 - 2\bar{y}\bar{Y} + \bar{Y}^2) \\
&= E(\bar{y}^2) - 2E(\bar{y})(\bar{Y}) + E(\bar{Y}^2)
\end{aligned}$$

Diketahui

$$V(\bar{y}) = E(\bar{y}^2) - (E(\bar{y}))^2$$

$$E(\bar{y}^2) = V(\bar{y}) + (E(\bar{y}))^2$$

sehingga

$$= \text{Var}(\bar{y}) + \text{bias}^2(\bar{y})$$

## 3.2. Aplikasi pada Data



Data yang digunakan dalam penelitian ini yakni data primer mengenai mahasiswa FMIPA UNHAS khususnya angkatan 2011 di mana grade SNMPTN jurusan sebagai pembeda strata.

Jumlah sampel tiap jurusan pada sampel acak terstratifikasi adalah

- Jur.Fisika = 49 orang
- Jur. Matematika = 47 orang
- Jur.Biologi = 42 orang
- Jur.Kimia = 48 orang

IPK mahasiswa sebagai variabel penelitian  $Y$  dan penghasilan orang tua per bulan sebagai *auxiliary variable*  $X$ .

Berdasarkan pada hasil penaksiran rata-rata dan variansi minimum yang telah dijelaskan sebelumnya, maka didapatkan hasil sebagai berikut:

**Tabel. 1 Perbandingan jumlah  $n_h$  secara persentase dan alokasi optimum**

Jurusan	$N_h$	$n_h$ secara persentase	$n_h$ secara alokasi optimum
Fisika	111	58	49
Matematika	110	57	47
Biologi	97	47	42
Kimia	68	24	48

**Tabel 2. Perbandingan Hasil Taksiran Rataan dan Variansi**

Metode	Taksiran	Rataan	Variansi
Acak sederhana		3,1786	0,0003
Terstratifikasi	Terpisah	3,1749	0,0002
	Gabungan	3,1774	0,0016
Mean Square Error	Terpisah	-	0,0002
	Gabungan	-	0,0016
Alokasi optimum		-	0,0014

## 5. Kesimpulan dan Saran

### 5.1. Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan beserta penjelasan yang dipaparkan, maka kesimpulan yang dapat ditarik adalah sebagai berikut:

1. Besarnya penghasilan orang tua mahasiswa tidak mempengaruhi tinggi atau rendahnya nilai IPK seorang mahasiswa.
2. Penaksiran variansi menggunakan teknik sampel acak terstratifikasi lebih baik dibandingkan dengan teknik sampel acak sederhana karena teknik tersebut membagi populasi ke dalam strata-strata yang variansinya lebih kecil sehingga membuat keragaman dalam populasi juga menjadi kecil.
3. Penambahan *auxiliary variable* ternyata lebih meminimumkan variansi populasi dibandingkan dengan hanya menggunakan variabel penelitian saja karena *auxiliary*

*variable* ini memberikan informasi lebih mengenai variabel penelitian sehingga variabel ini banyak dimanfaatkan untuk menghasilkan penaksiran yang lebih tepat.

4. Penaksiran variansi pada sampel acak terstratifikasi secara terpisah lebih menghasilkan penaksiran yang lebih baik dibandingkan dengan penaksiran gabungan karena dalam kasus ini keragaman dalam strata lebih kecil sedangkan antar strata keragamannya besar.
5. Penambahan alokasi optimum lebih meminimumkan hasil penaksiran variansi galat karena alokasi optimum memberikan jumlah sampel yang lebih memperlihatkan karakteristik dari populasi.

### 5.1 Saran

Pada penelitian ini membahas mengenai penambahan *auxiliary variable* dalam meminimumkan hasil penaksiran rata-rata dan variansi pada variabel penelitian dan berikut adalah saran yang dapat diberikan oleh penulis:

1. Penelitian ini hanya melibatkan satu *auxiliary variable* yang hasilnya cukup efisien dalam menghasilkan variansi minimum tetapi diharapkan pada penelitian selanjutnya dapat melibatkan dua atau lebih *auxiliary variable* yang berkorelasi dengan variabel penelitian agar hasil penaksiran yang dihasilkan juga semakin baik.
2. Penelitian ini menggunakan penaksir regresi sehingga penelitian selanjutnya dapat memakai penaksir rasio untuk metode penaksirannya.

## DAFTAR PUSTAKA

- [1] Cochran, W. G. 1991. *Sampling Techniques Third Edition (Terjemahan)*. Jakarta: UI-Press.
- [2] Kadilar, C. dan Cingi, H., 2003. "Ratio Estimators in Stratified Random Sampling". *Biometrical Journal* 45 (2003) 2,218-225. Turkey.
- [3] Koyuncu, Nursel dan Kadilar, Cem. 2010. *On Improvement in Estimating Population Mean in Stratified Random Sampling*. Taylor & Francis: Turkey.
- [4] Singh Rajesh, dkk. 2008. "Ratio Estimators in Simple Random Sampling Using Information on Auxiliary Attribute". *Pak.j.stat.oper.res.Vol.IV No.1 2008* pp47-53. India.
- [5] Sisodia, B.V.S. and Dwivedi, V. K. 1981. "A Modified Ratio Estimator Using Coefficient of Variation of Auxiliary Variable". *Journal of Indian Society Agricultural Statistics* 33,1318.
- [6] Tompson, Steven K. 1992. *Sampling*. New York: John Wiley & Sons, Inc.
- [7] Upadhyaya, L.N. and Singh, H.P., 1999: "Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean". *Biometrical Journal* 41, 5, 627636.