

**“PENAKSIRAN RATAAN DAN VARIANSI POPULASI
PADA SAMPEL ACAK TERSTRATIFIKASI DENGAN
AUXILIARY VARIABLE”**

Skripsi

Disusun oleh :

HASRUNI SATYA TARUMA

H121 09 001



JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

2013

**“PENAKSIRAN RATAAN DAN VARIANSI POPULASI
PADA SAMPEL ACAK TERSTRATIFIKASI DENGAN
AUXILIARY VARIABLE”**

S K R I P S I

*Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada
Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam*

Universitas Hasanuddin

Makassar

Oleh:

HASRUNI SATYA TARUMA

H 121 09 001

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

2013

P E R N Y A T A A N

Saya yang bertanda tangan di bawah ini menyatakan dengan sesungguhnya-benar bahwa skripsi yang saya buat dengan judul:

**“PENAKSIRAN RATAAN DAN VARIANSI POPULASI
PADA SAMPEL ACAK TERSTRATIFIKASI DENGAN
AUXILIARY VARIABLE”**

adalah benar hasil kerja saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun.

Makassar, 23 Mei 2013

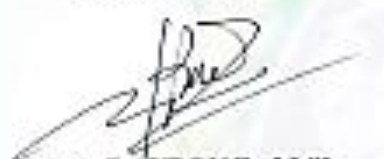


HASRUNI SATYA TARUMA
NIM : H 121 09 001

**“PENAKSIRAN RATAAN DAN VARIANSI POPULASI
PADA SAMPEL ACAK TERSTRATIFIKASI DENGAN
AUXILIARY VARIABLE”**

Disetujui Oleh :

Pembimbing Utama


Drs. RAUPONG, M.Si
NIP. 19621015 198810 1 001

Pembimbing Pertama


Drs. M. SALEH A.F., M. Si
NIP. 19540804 197802 1 001

Pada tanggal : 23 Mei 2013

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN**

2013

Pada hari ini, Senin tanggal 23 Mei 2013, Panitia Ujian Skripsi menerima dengan baik skripsi yang berjudul :

**“PENAKSIRAN RATAAN DAN VARIANSI POPULASI
PADA SAMPEL ACAK TERSTRATIFIKASI DENGAN
AUXILIARY VARIABLE”**

yang diajukan untuk memenuhi salah satu syarat guna memperoleh gelar Sarjana Sains pada Program Studi Statistika Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Makassar, 23 Mei 2013

PANITIA UJIAN SKRIPSI

Tanda Tangan

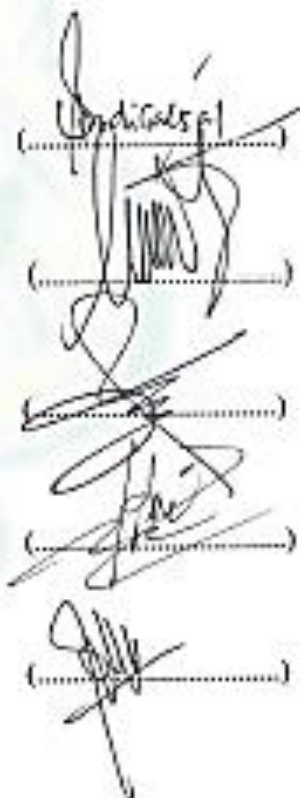
1. Ketua : Andi Galsan Mahie, S.Si, M.Si

2. Sekretaris : Muh. Nur, S.Si, M.Si

3. Anggota : Andi Kresna Juyi, S.Si, M.Si

4. Anggota : Drs. Raupong, M.Si

5. Anggota : Drs. M. Saleh AF., M. Si



KATA PENGANTAR

Alhamdulillah rabbil alamin. Puji syukur senantiasa penulis panjatkan kehadirat **Allah SWT** atas segala limpahan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi ini, sebagai salah satu prasyarat untuk menyelesaikan studi di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin. Tak lupa pula salam dan salawat terlantunkan kepada Nabi besar **Muhammad SAW** (sebaik-baik panutan dan idola bagi umat manusia), sebagai pembawa risalah dan kebenaran sejati yang senantiasa menjadi pedoman dan panutan dalam kehidupan kita.

Skripsi yang berjudul **“Penaksiran Rataan dan Variansi Populasi pada Sampel Acak Terstratifikasi dengan *Auxiliary Variable*”** disusun sebagai salah satu syarat mendapatkan gelar Sarjana Sains. Skripsi ini dapat terselesaikan dengan baik atas dukungan dari berbagai pihak baik moril maupun materil, untuk itu pada kesempatan ini penulis menyampaikan terima kasih dan penghargaan yang tidak terhingga kepada almarhum **Ayahanda Lataribba dan Ibunda Hasmawiah** sebagai orang tua yang telah memberikan segalanya, yang kuhormati serta iringan doa-doanya demi keberhasilan pendidikan penulis. Juga kepada kakak-kakakku tercinta **Maslinda Taruma dan keluarga, Masrudi Taruma dan keluarga, Masdalifah Taruma dan keluarga, Maswita Taruma dan keluarga**, serta kakak termuda **Mas ‘Ali Wahyudi Taruma**. Tak lupa juga buat semua keluarga yang telah memberikan dukungan dan semangat kepada penulis.

Demikian pula penulis menyampaikan rasa terima kasih dan penghargaan yang setinggi-tingginya kepada:

1. Bapak **Drs. Raupong, M.Si** dan **Drs. M. Saleh AF., M.Si** selaku pembimbing utama dan pembimbing pertama yang penuh kesabaran, kesungguhan, dan kebaikan hatinya telah banyak memberikan petunjuk serta bimbingan sehingga kesulitan penulis dapat teratasi dan akhirnya dapat menyelesaikan skripsi ini dengan baik.
2. **Dr. Hasmawati, M. Si.** selaku Ketua Jurusan Matematika FMIPA Unhas, dan Bapak **Dr. Nurdin, S. Si, M. Si.** selaku Sekretaris Jurusan Matematika FMIPA dan para Dosen Jurusan Matematika yang telah memberikan bekal ilmunya selama perkuliahan serta para staf yang telah memberikan bantuan dan dorongan selama penulis menjalani perkuliahan.
3. Bapak **Andi Kresna Jaya, S.Si, M. Si.** selaku penasehat akademik yang telah memberikan perhatiannya selama perkuliahan.
4. Bapak **Andi Galsan Mahie, S.Si, M.Si**, Bapak **Muh. Nur, S.Si, M.Si**, dan Bapak **Andi Kresna Jaya, S.Si, M.Si**, selaku penguji yang selalu bersedia meluangkan waktunya dan memberikan masukan-masukan dalam perbaikan skripsi ini.
5. Para pegawai administrasi baik di tingkat jurusan maupun di tingkat fakultas. Terutama **Pak Sutamin**, dan **Pak Nasir**, yang telah memudahkan penulis dalam menyelesaikan segala administrasi selama menempuh perkuliahan di Jurusan Matematika.

6. Teman-teman “Statistika09”: **Ifa, Ipin, Kiki, Yuli, Yuni, Try, Iman, Isna, Hesty, Chimank, Jejen, Whay, Mimi, Evi, Risma, Ayu, Ida, Anda, Fitri, Ira, Hera, Irzan, Tenri, Naser, Fitrah, Jumi, Juned, Fahrur, Ume, Endy, Mirsam, Vinni, dan Niki.** Terima kasih atas persahabatan dan dukungannya selama ini.
7. Kanda-kanda senior dan adik-adik junior warga Himatika FMIPA Unhas, terima kasih atas bantuan, dorongan, dan kebersamaannya.
8. Kepada teman-teman seperjuangan dari tanah Maspul, **Nurfaidah Hasnur, Fifik Astuti Hasanuddin, S.Si, Lilia Afriana dan Siti Huspanti Husain.**
9. Kepala Badan Pusat Statistik (BPS) Kabupaten Gowa dan seluruh staf BPS Kabupaten Gowa atas sambutannya pada saat pelaksanaan Kerja Praktek.
10. Keluarga KKN, Bapak Kades Ujungpero, *alm* Abdul Rahman Ozi beserta keluarga dan warga sekitar, teman-teman posko KKN Kabupaten Wajo, Desa Ujungpero. Terima kasih atas kebersamaan yang terjalin selama pelaksanaan Kuliah Kerja Nyata (KKN).
11. **Sdr. Ahmad Taufiq,** terima kasih atas dorongan semangat, nasehat, berbagai pelajaran hidup untuk berjuang dan bersabar dan terima kasih karena ingin direpotkan oleh urusan penulis mulai dari pertama bertemu hingga sekarang *and thank's for this advice “everything will be fine”*.
12. Semua pihak yang telah banyak berpartisipasi, baik secara langsung maupun tidak langsung, dalam pembuatan tugas akhir ini yang tak sempat penulis sebutkan satu per satu. Semoga segenap bantuan dan partisipasinya bernilai ibadah dan mendapat pahala yang setimpal di sisi Allah SWT.

Penulis mengharapkan kritik dan saran yang membangun dalam upaya peningkatan kualitas skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi penulis khususnya dan semua pihak pada umumnya. Akhirnya semoga tulisan ini dapat memberikan manfaat kepada semua pihak yang membutuhkannya. Amin.

Makassar, 23 Mei 2013

Penulis

PENAKSIRAN RATAAN DAN VARIANSI POPULASI PADA SAMPEL ACAK TERSTRATIFIKASI DENGAN *AUXILIARY VARIABLE*

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 terstratifikasi, 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.

ESTIMATION MEAN AND VARIANCE POPULATION IN *STRATIFIED RANDOM SAMPLING WITH AUXILIARY VARIABLE*

ABSTRACT

Estimation of mean and variance population through by sample and the expected results can represent the characteristics of the population. Best sampling technique will give best results too. Most researchers use random sampling techniques but assessment results with this technique still produces a large variance and is not suitable for a large number of samples that required other sampling techniques, one of another is stratified random sampling. This technique is still better than a simple random sample because it produces the minimum variance and more suitable for large number of samples. But with only rely on sampling techniques, still no guarantee the results will be better so the assessment in this study added several other techniques such as the addition of auxiliary variables and the determination of optimum allocation of sample size. The combined between the stratified sampling technique, the addition of auxiliary variables, and the optimum allocation assessment showed better results in describing characteristics of the population through the sample.

Keywords : *Simple random sampling, stratified random sampling, auxiliary variable, optimum allocation.*

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR KEOTENTIKAN	ii
KATA PENGANTAR	vi
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI	xii
DAFTAR TABEL.....	xv
DAFTAR SINGKATAN DAN DAFTAR SIMBOL	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan	4
1.5 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
2.1 Populasi dan Sampel	5
2.2 Beberapa Metode Sampling	6
2.2.1 Pengambilan Sampel Acak Sederhana.....	6
2.2.2 Pengambilan Sampel Terstratifikasi.....	7
2.3 <i>Auxiliary Variable</i>	12
2.4 <i>Mean Square Error</i>	12
2.5 Alokasi Optimum	13
2.6 Penaksir Regresi untuk Sampel Acak Sederhana	14
2.7 Penaksir Regresi untuk Sampel Acak Terstratifikasi	15

BAB III METODE PENELITIAN	18
3.1 Sumber Data	18
3.2 Populasi dan Sampel	18
3.3 Identifikasi Variabel	19
3.4 Metode Analisis	19
3.5 Diagram Alur Kerja	21
BAB IV HASIL DAN PEMBAHASAN	22
4.1 Penaksiran Rataan $\bar{y}$ dan Variansi Populasi $V(\bar{y}_{lr})$ pada Sampel Acak Sederhana	22
4.1.1 Penaksiran rataan $\bar{y}$	22
4.1.2 Penaksiran z	22
4.1.3 Penaksiran variansi populasi $V(\bar{y}_{lr})$	23
4.2 Penaksiran Rataan $\bar{y}_{lrs}$ dan Variansi $V(\bar{y}_{lrs})$ Populasi pada Sampel Acak Terstratifikasi (Penaksiran Terpisah)	27
4.2.1 Penaksiran rataan populasi $\bar{y}_{lrs}$	27
4.2.2 Penaksiran variansi populasi $V(\bar{y}_{lrs})$	28
4.3 Penaksiran Rataan $\bar{y}_{lrc}$ dan Variansi $V(\bar{y}_{lrc})$ Populasi pada Sampel Acak Terstratifikasi (Penaksiran Gabungan)	29
4.3.1 Penaksiran rataan $\bar{y}_{lrc}$	29
4.3.2 Penaksiran variansi $V(\bar{y}_{lrc})$	29
4.4 Pembuktian Persamaan <i>Mean Square Error</i>	30
4.5 Aplikasi Data Penaksiran Rataan dan Variansi	31
BAB V PENUTUP	41
5.1 Kesimpulan	41
5.2 Saran	42

DAFTAR PUSTAKA	xviii
----------------------	-------

LAMPIRAN	
----------	--

DAFTAR TABEL

Tabel 1. Perbandingan jumlah n_h secara persentase dan alokasi optimum	36
Tabel 2. Perbandingan Hasil Taksiran Rataan dan Variansi	38
Tabel 3. Data IPK (y) dan Penghasilan Orang Tua dalam Jutaan Rupiah (x) secara Persentase	xix
Tabel 4. Data IPK (y) dan Penghasilan Orang Tua dalam Jutaan Rupiah (x) secara Alokasi Optimum.	xx

DAFTAR SINGKATAN DAN DAFTAR SIMBOL

Simbol	Keterangan
n	Jumlah sampel
N	Jumlah populasi
h	Strata
i	Unit
L	Jumlah strata
y	Sampel variable penelitian
x	Sampel <i>Auxiliary variable</i>
z	Koefisien regresi
Y	Populasi variable penelitian
X	Populasi <i>auxiliary variable</i>
ρ	Korelasi
S	Standar deviasi populasi
S^2	Variansi populasi
y_i	Nilai pada pengamatan ke- i pada variable penelitian
x_i	Nilai pada pengamatan ke- i pada <i>auxiliary variable</i>
$\bar{y}$	Rataan sampel variable penelitian
$\bar{x}$	Rataan sampel <i>auxiliary variable</i>
$\bar{Y}$	Rataan populasi variable penelitian
$\bar{X}$	Rataan populasi <i>auxiliary variable</i>
n_h	Ukuran sampel yang terpilih
N_h	Ukuran populasi yang terpilih
c_h	Biaya pengambilan sampel pada strata ke- h
z_h	Koefisien regresi pada strata ke- h
z_c	Koefisien regresi gabungan dari seluruh strata
$\bar{Y}_h$	Rataan populasi untuk variable penelitian
$\bar{y}_h$	Rataan sampel pada strata ke- h untuk variable penelitian
$\bar{X}_h$	Rataan populasi untuk <i>auxiliary variable</i>
$\bar{x}_h$	Rataan sampel pada strata ke- h untuk <i>auxiliary variable</i>
W_h	Bobot lapisan
f_h	Fraksi penarikan sampel di dalam lapisan
S_h	Standar deviasi pada strata ke- h
S_y	Standar deviasi variable penelitian
S_x	Standar deviasi <i>auxiliary variable</i>
S_h^2	Variansi pada lapisan ke- h
S_y^2	Variansi untuk variable penelitian
S_x^2	Variansi untuk <i>auxiliary variable</i>
S_{yx}	Kovariansi antara x dan y

st	Menandakan pengerjaan pada sampel acak terstratifikasi
$\bar{y}_{st}$	Rataan dari keseluruhan sampel terstratifikasi untuk variable penelitian
$\bar{x}_{st}$	Rataan dari keseluruhan sampel terstratifikasi untuk <i>auxiliary variable</i>
y_{hi}	Nilai pada strata ke- h pada pengamatan ke- i untuk variable penelitian
x_{hi}	Nilai pada strata ke- h pada pengamatan ke- i untuk <i>auxiliary variable</i>
lr	Menandakan linear
$\bar{y}_{lr}$	Rataan yang menggabungkan antara variabel penelitian dan <i>auxiliary variable</i> untuk sampel acak sederhana
S_{yh}^2	Variansi variabel penelitian pada strata ke- h
S_{xh}^2	Variansi <i>auxiliary variable</i> pada strata ke- h
S_{yjh}^2	Kovariansi antara x dan y pada strata ke- h
$\bar{y}_{lrs}$	Rataan yang menggabungkan antara variabel penelitian dan <i>auxiliary variable</i> untuk penaksiran terpisah pada sampel acak terstratifikasi
$\bar{y}_{lrc}$	Rataan yang menggabungkan antara variabel penelitian dan <i>auxiliary variable</i> untuk penaksiran gabungan pada sampel acak terstratifikasi
MSE	<i>Mean Square Error</i>

BAB I

PENDAHULUAN

1.1. Latar Belakang

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. Objek tersebut dapat berupa manusia, hewan, tumbuhan, benda mati, dsb. Dalam melakukan penelitian, kadang-kadang penelitian dilakukan terhadap seluruh objek (*sensus*), tetapi sering juga peneliti hanya mengambil sebagian saja dari seluruh objek tersebut (*survey*). Meskipun penelitian hanya mengambil sebagian dari seluruh objek yang diteliti, tetapi hasilnya dapat mewakili atau mencakup seluruh objek yang diteliti.

Keseluruhan objek penelitian atau objek yang diteliti tersebut adalah populasi penelitian atau universe. Sedangkan sebagian yang diambil dari keseluruhan objek yang diteliti dan dianggap mewakili seluruh populasi ini disebut “sampel penelitian”. Dalam mengambil sampel penelitian ini digunakan 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.

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 efisien. Untuk mengetahui bias tersebut maka digunakan penaksir *mean square error*, di mana penaksir ini selain memperlihatkan seberapa besar bias yang terjadi juga memperlihatkan variansi penaksiran.

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 efisien. Di beberapa kasus, di samping rata-rata variabel tambahan, berbagai parameter yang terkait dengan variabel tambahan seperti standar deviasi, koefisien variasi, kemiringan, koefisien korelasi kurtosis, dll juga dapat diketahui. Untuk kasus ini, banyak penulis seperti Upadhyaya dan Singh (1999) serta Sisodia dan Dwivedi (1981) mengembangkan berbagai penaksir untuk meningkatkan rasio penaksir dalam sampel acak sederhana. Kadilar dan Cingi (2003) mengadaptasi penaksir yang didapatkan oleh Upadhyaya dan Singh (1999) untuk sampel acak terstratifikasi.

Oleh karena alasan-alasan tersebutlah maka penulis menulis skripsi dengan judul ***“Penaksiran Rataan dan Variansi Populasi pada Sampel Acak Terstratifikasi dengan Auxiliary Variable”***.

1.2. Rumusan Masalah

Rumusan masalah yang dapat diangkat oleh penulis adalah

1. Bagaimana menentukan penaksiran rata-rata dan variansi pada sampel acak terstratifikasi?
2. Bagaimana menentukan penaksiran rata-rata dan variansi dengan penambahan *auxiliary variable* dan alokasi optimum pada sampel acak terstratifikasi?
3. Bagaimana menentukan penaksiran variansi menggunakan *Mean Square Error* pada sampel acak terstratifikasi?

1.3. Batasan Masalah

Penulisan skripsi kali ini penulis membatasi pada pengestimasian rata-rata dan variansi populasi pada sampel acak terstratifikasi dengan adanya variabel tambahan (*auxiliary variable*).

1.4. Tujuan

Tujuan dari penulisan skripsi ini adalah

1. Menentukan penaksiran rata-rata dan variansi pada sampel acak terstratifikasi.
2. Membandingkan hasil penaksiran rata-rata dan variansi sederhana dengan hasil penaksiran dengan *auxiliary variable* dan alokasi optimum pada sampel acak terstratifikasi.
3. Menentukan hasil penaksiran variansi menggunakan *Mean Square Error*.

1.5. Manfaat

Digunakan sebagai tambahan informasi dan wawasan tentang pengestimasian rata-rata dan variansi populasi pada sampel acak terstratifikasi dengan *auxiliary variable* serta mampu mengaplikasikannya dalam kehidupan sehari-hari dan diharapkan dapat menjadi tambahan kepustakaan sebagai sarana dalam pengembangan ilmu pengetahuan khususnya di jurusan matematika dalam kajian statistik.

BAB II

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. Walaupun banyak cara yang dapat dilakukan tetapi ada syarat-syarat yang harus dipenuhi yaitu:

- a. Pengambilan sampel harus didasarkan atas ciri-ciri atau sifat tertentu yang merupakan ciri pokok populasi.
- b. Subjek yang diambil sebagai sampel benar-benar merupakan subjek yang paling banyak mengandung ciri-ciri yang terdapat pada populasi atau merupakan key subjek.
- c. Penentuan karakteristik populasi harus dilakukan dengan cermat di dalam studi pendahuluan.

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*. Beberapa cara pengambilan sampel penelitian adalah sebagai berikut *Simple Random Sampling*, *Stratified Sampling*, *Cluster Sampling*, *Area Probability Sample*, *Propotional Sampling*, dll.

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 ${}_NC_n$ 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.

Dalam menggunakan tabel bilangan acak untuk memilih sebuah sampel acak sederhana, tahap pertama adalah memberi nomor unit-unit dalam populasi dari 1 sampai N . Jika angka pertama dari N adalah sebuah nomor antara 5 dan 9, cara berikutnya adalah memilih yang sesuai. Misalkan $N = 528$, dan kita ingin $n = 10$. Pilih tiga kolom dari tabel bilangan acak, katakan kolom 25 sampai 27. Pada tiga kolom tersebut pilihlah 10 bilangan yang berbeda antara 001 sampai 528. Angka yang terpilih adalah 36, 509, 364, 417, 348, 127, 149, 186, 290, dan 162. Untuk dua kolom terakhir, kita loncat pada kolom 30 sampai 32. Dalam pemilihan ulang disarankan untuk memulai dengan titik mulai berbeda.

Kerugian dari metode ini ialah bahwa bilangan tiga angka 000 dan 529 sampai 999 tidak digunakan, walaupun loncatan bilangan tidak memakan waktu yang lama. Bila angka pertama dari N kurang dari 5, metode ini masih disenangi jika n kecil dan sebuah tabel angka acak yang besar tersedia. (Cochran: 1991)

2.2.2 Pengambilan Sampel Terstratifikasi (*Stratified Sampling*)

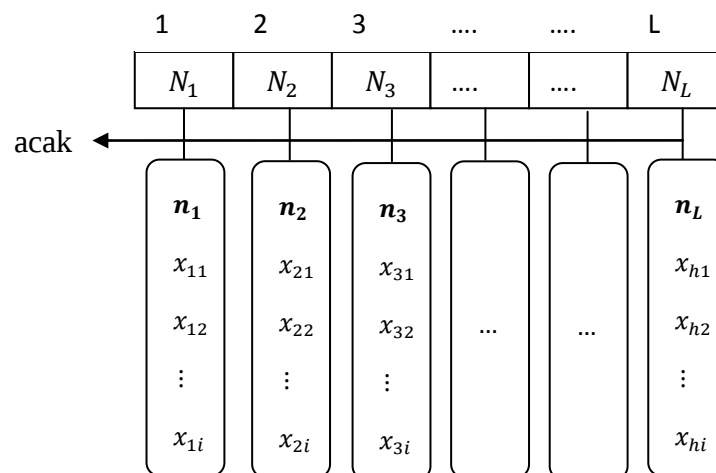
➤ Uraian

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 . Berikut adalah konstruksi pengambilan sampel acak terstratifikasi



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

Pelapisan adalah sebuah teknik biasa. Ada beberapa alasan prinsip untuk penggunaannya, prinsip yang dimaksud antara lain:

1. Jika data diketahui penelitian yang diinginkan untuk subpopulasi tertentu dari populasi, ada baiknya memperlakukan setiap subpopulasi sebagai suatu populasi tertentu.
2. Administrasi yang baik dapat memakai kegunaan pelapisan; sebagai contoh agen survei dapat menggunakan kantor-kantor cabang yang masing-masing dapat mengawasi survey sebagai bagian dari populasi.
3. Masalah penarikan sampel dapat berbeda dalam bagian populasi yang berbeda. Dengan populasi manusia, orang-orang yang hidup dalam adat kebiasaan (misalkan hotel, rumah sakit, penjara, dll) seringkali ditempatkan pada lapisan yang berbeda dengan orang-orang yang tinggal di rumah-rumah biasa, karena pendekatan yang berbeda untuk penarikan sampelnya adalah sesuai untuk dua keadaan tersebut. Dalam penarikan sampel perusahaan kita dapat memperoleh sebuah daftar dari perusahaan-perusahaan besar, yang ditempatkan pada lapisan yang terpisah. Beberapa jenis daerah penarikan sampel dipergunakan untuk perusahaan-perusahaan kecil.
4. Pelapisan dapat menghasilkan suatu manfaat dalam penelitian perkiraan dari karakteristik seluruh populasi. Hal ini memungkinkan untuk membagi sebuah populasi yang heterogen menjadi subpopulasi-subpopulasi, dengan setiap subpopulasi menjadi homogen. Subpopulasi ini dinamakan *lapisan*, dengan

implikasi pembagiannya ke dalam lapisan. Jika tiap-tiap lapisan homogen, maka pengukuran varians antar lapisan menjadi kecil dan perkiraan yang diteliti dari setiap rata-rata lapisan dapat diperoleh dari sebuah sampel yang kecil dalam lapisan tersebut. Perkiraan ini kemudian dapat dikombinasikan dengan suatu perkiraan yang teliti untuk seluruh populasi.

Teori dari penarikan sampel berhubungan dengan sifat-sifat perkiraan dari sampel berlapis dan dengan pemilihan yang terbaik dari ukuran sampel n_h untuk memperoleh ketelitian yang maksimum.

➤ 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

$$S_h^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2}{N_h - 1} \quad \text{Varians pada lapisan ke-}h \text{ yang terpilih}$$

Perhatikan bahwa pembagi untuk varians adalah $N_h - 1$. (Cochran: 1991)

➤ 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 $N = N_1 + N_2 + \dots + N_L$

Penaksiran $\bar{y}_{st}$ adalah tidak sama dengan rata-rata sampel biasa.

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

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

(Cochran: 1991)

2.3 Variabel Tambahan (*Auxiliary Variable*)

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. Jika variabel penelitian adalah jumlah hewan dalam komplotan, variabel tambahan dapat mencakup area plot, jenis vegetasi, atau ketinggian rata-rata plot. Dalam banyak survei populasi manusia, nilai dari variabel penelitian dari sensus sebelumnya dapat berfungsi sebagai variabel tambahan.

Informasi tambahan dapat digunakan baik dalam desain sampling atau estimasi. Stratifikasi berdasarkan jenis vegetasi atau elevasi merupakan penggunaan informasi tambahan dalam desain. Sampling dengan pengembalian dengan pemilihan proporsional probabilitas dengan ukuran- misalnya, ukuran plot atau ukuran pohon adalah penggunaan lain dari informasi tambahan dalam desain.

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. (Thompson: 1992)

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 diperkirakan.

Secara umum,

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

Penghitungan *MSE* sebagai satu kriteria ketelitian sebuah penaksir berhubungan dengan dua penaksiran yang memiliki *MSE* yang sama. Ini tidak sepenuhnya benar karena distribusi frekuensi galat $(\bar{y} - \bar{Y})$ dari ukuran yang berbeda tidak akan sama untuk dua penaksiran tersebut, jika keduanya masing-masing memiliki nilai bias yang berbeda. (Cochran: 1991)

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. (Cochran: 1991)

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

- a. Tiap sampel telah terbagi ke dalam masing-masing strata.

- b. Rata-rata biaya per unit c_h untuk mengumpulkan data dalam setiap strata diketahui.
- c. Standar deviasi S_h diketahui untuk setiap strata.
- d. Bobot W_h diketahui untuk setiap strata.
- e. Ukuran sampel n diketahui.

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. Dalam penaksiran regresi terdapat dua cara dalam pemberian z yaitu z telah ditentukan terlebih dulu dan z dihitung dari sampel. (Cochran: 1991)

Penaksiran yang diterapkan kali ini adalah z dihitung dari sebuah penaksiran sampel efektif. Perhitungannya 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 Cochran (1991) 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

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

(Cochran: 1991)

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

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

Penaksiran regresi yang kedua adalah $\bar{y}_{lrc}$ (huruf c berarti gabungan).

Penaksiran ini menganggap Z_h sama untuk seluruh lapisan. Untuk menghitung $\bar{y}_{lrc}$, pertama-tama kita tentukan

$$\bar{y}_{st} = \sum_h W_h \bar{y}_h$$

$$\hat{x}_{st} = \sum_{h=1}^L N_h \bar{x}_h$$

Maka

$$\bar{X} = \hat{x}_{st} = \frac{1}{N} \sum_{h=1}^L N_h \bar{x}_h$$

$\bar{x}_{st}$ merupakan hasil perhitungan langsung dari sampel, kemudian

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

Pers.(16) dapat menunjukkan bahwa $\bar{y}_{lrc}$ adalah juga suatu penaksiran tak bias dari $\bar{Y}$. Karena $\bar{y}_{lrc}$ adalah penaksiran biasa dari suatu sampel berlapis untuk $y_{hi} + b(\bar{X} - x_{hi})$, kita dapat menggunakan persamaan (4) untuk menghasilkan

$$V(\bar{y}_{lrc}) = \sum_h \frac{W_h^2(1-f_h)}{n_h} (S_{yh}^2 - 2zS_{yhx} + z^2S_{xh}^2) \quad (15)$$

Nilai z yang meminimumkan variansi ini adalah

$$Z_c = \sum_h \frac{w_h^2(1-f_h)S_{yxh}}{n_h} \bigg/ \sum_h \frac{w_h^2(1-f_h)S_{xh}^2}{n_h} \quad (16)$$

Nilai Z_c adalah suatu rata-rata tertimbang dari koefisien regresi lapisan

$$Z_h = S_{yxh}/S_{xh}^2.$$

Dengan penaksiran gabungan, variansi dapat diminimumkan bila $z = Z_c$ sehingga untuk penaksiran sampel dipilih koefisien regresi sampel sebagai berikut

$$z_c = \sum_h \frac{w_h^2(1-f_h)}{n_h(n_h-1)} \sum_i (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h) \bigg/ \sum_h \frac{w_h^2(1-f_h)}{n_h(n_h-1)} \sum_i (x_{hi} - \bar{x}_h)^2 \quad (17)$$

Jika kesalahan pengambilan sampel dari z_c dapat diabaikan maka

$$V(\bar{y}_{lrc}) = \sum_h \frac{w_h^2(1-f_h)}{n_h} (S_{yh}^2 - 2Z_c S_{yxh} + Z_c^2 S_{xh}^2) \quad (18)$$

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

$$v(\bar{y}_{lrc}) = \sum_h \frac{w_h^2(1-f_h)}{n_h(n_h-1)} \sum_h [(y_{hi} - \bar{y}_h) - z_c(x_{hi} - \bar{x}_h)]^2 \quad (19)$$

(Cochran: 1991)