TEKNIK PENARIKAN SAMPEL ACAK BERLAPIS DAN APLIKASINYA PADA PENDUGAAN PARAMETER

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Pendidikan
Program Studi Pendidikan Matematika

Oleh :

Dahlia Adiati

NIM : 931414007
NIRM : 930052010501120007

PROGRAM STUDI PENDIDIKAN MATEMATIKA
JURUSAN PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN
UNIVERSITAS SANATA DHARMA
YOGYAKARTA
1998

SKRIPSI

TEKNIK PENARIKAN SAMPEL ACAK BERLAPIS DAN APLIKASINYA PADA PENDUGAAN PARAMETER

Oleh :

Dahlia Adiati
NIM : 93 1414 007
NIRM : 930052010501120007

Telah disetujui oleh :

Pembimbing I

Johanes EP

Drs.Y.Eka Priyatma, M.Sc tanggal : 18 - 12 - 1998

Pembimbing II

Drs. Ig. Dwiatmoko

Ir.Ig. Aris Dwiatmoko, M.Sc tanggal : 18 - 12 - 1998

SKRIPSI

TEKNIK PENARIKAN SAMPEL ACAK BERLAPIS DAN APLIKASINYA PADA PENDUGAAN PARAMETER

Yang dipersiapkan dan disusun oleh

Dahlia Adiati
NIM : 93 1414 007
NIRM : 930052010501120007

Telah dipertahankan didepan Panitia Penguji
pada tanggal : 3 Desember 1998
dan dinyatakan telah memenuhi syarat

SUSUNAN PANITIA

	Nama lengkap	Tanda tangan
Ketua	: Drs. F. Kartika Budi, M.Pd	. . . .
Sekretaris	: Dr. St. Suwarsono	. . . .
Anggota	: Ir. Ig. Aris Dwiatmoko, M.Sc	. . . .
	Drs. Y. Eka Priyatma, M.Sc	. . . .
	Dr. Y. Marpaung	. . . .

Yogyakarta, 1998
Fakultas Keguruan dan Ilmu Pendidikan
Universitas Sanata Dharma
Dekan FKIP

Dr. Paulus Suparno, S.J., MST

WAKTU TUHAN

Didalam hidup ini semua ada waktunya,
Ada waktunya kita menabur, ada waktu menuai
Mungkin doamu bagai tak terjawab
Namun yakinlah tetap Tuhan tak akan terlambat
Juga tak akan lebih cepat
Semuanya Dia jadikan indah tepat pada waktunya
Tuhan tak akan terlambat, juga tak akan tinggalkanmu
Ajarlah kami setia selalu menanti waktumu Tuhan

Karya kecil ini kupersembahkan untuk :

Papi dan mami tercinta,

Arief Gunawan, Maria, dan Budi yang terkasih

KATA PENGANTAR

Syukur dan puji bagi Tuhan Yang Maha Esa atas rahmat dan kasih-Nya sehingga skripsi yang berjudul Teknik Penarikan Sampel Acak Berlapis dapat terselesaikan.

Penyusunan skripsi ini dimaksudkan untuk memenuhi salah satu syarat memperoleh gelar Sarjana Pendidikan Program Studi Pendidikan Matematika di Jurusan Pendidikan Matematika Dan Ilmu Pengetahuan Alam Fakultas Keguruan dan Ilmu Pendidikan Universiatas Sanata Dharma Yogyakarta.

Pada kesempatan ini, penyusun ingin mengucapkan terima kasih kepada :

- Bapak Drs. Y. Eka Priyatma, M.Sc selaku pembimbing I yang telah membimbing dan memberikan masukan yang berharga.
- Bapak Ir. Ig. Aris Dwiatmoko, M.Sc selaku pembimbing II yang dengan teliti, sabar dan penuh pengertian membimbing dalam penyusunan skripsi ini.
- Bapak dan Ibu dosen yang telah membimbing dan mendidik penulis selama belajar di Universitas Sanata Dharma.
- Papi, mami, kakak-kakak, dan adik yang telah memberikan bantuan moril maupun spirituil selama penulis kuliah.
- Semua teman- teman dan semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu penulis selama penulisan skripsi hingga selesai.

Akhirnya penulis menyadari, bahwa masih banyak kekurangan dan kelemahan dalam penulisan skripsi ini karenanya segala masukan dan saran yang membangun akan diteima dengan senang hati.

Penyusun

DAFTAR ISI

ABSTRAK

Salah satu cara yang dilakukan manusia dalam memecahkan masalahnya adalah dengan mengadakan penelitian. Penelitian yang dilakukan menyangkut sampel yang merupakan bagian dari populasi. Tulisan ini akan membahas suatu teknik penarikan sampel, yang disebut teknik penarikan sampel acak berlapis.

Dalam teknik penarikan sampel acak berlapis populasi dibagi dalam beberapa lapisan. Tiap lapisan harus homogen dan mempunyai sifat saling asing, yang berarti anggota elemen pada lapisan yang satu bukan merupakan anggota elemen pada lapisan yang lain. Salah satu syarat dalam teknik penarikan sampel acak berlapis adalah harus ada kerangka, yang merupakan daftar unit penarikan sampel. Berdasarkan lapisan-lapisan tersebut, tiap lapisan diambil sampel secara acak sederhana. Suatu kuantitas yang dihitung dari sampel yang sudah terambil (disebut statistik) digunakan untuk menduga parameter populasi yang tidak diketahui (total populasi, rata-rata populasi, proporsi populasi). Dengan adanya pelapisan, tiap lapisan yang homogen mempunyai variansi pendugaan yang minimum, sehingga penduga akan mendekati nilai parameter yang sebenarnya dalam populasi.

Langkah yang harus dilakukan sebelum menerapkan teknik penarikan sampel acak berlapis adalah adanya data pendahuluan yang dapat dipakai sebagai dasar pelapisan. Berdasarkan data tersebut peneliti dapat memperkirakan keadaan populasi, dan menerapkan salah satu metode yaitu metode alokasi sebanding, metode alokasi Neyman, metode alokasi optimum untuk menentukan ukuran sampel.

Penarikan Sampel Acak Berlapis 2 tahap dilakukan untuk mengambil nilai dalam sub lapisan yang lebih kecil sehingga dapat meningkatkan efisiensi statistik sampel. Populasi dibagi dalam beberapa lapisan, kemudian tiap lapisan diambil sampel acak sederhana yang merupakan unit utama. Dari tiap unit utama diambil sampel acak sederhana yang merupakan unit sekunder.

ABSTRACT

One way of solving human problem is by research. The research covers a representative sample of the population. This thesis intended to explain about sampling techniques was namely stratified random sampling.

In stratified random sampling, population is divided into some strata. Every stratum must be homogenous and independent each other. It means that the element in a stratum is not part of the other stratum. One of the requirement in stratified random sampling is the list of sampling units, called frame. Based on those strata, the researcher can take a simple random sample from every stratum. A quantity which calculated from sample is called statistics. This statistics is used to estimate the unknown population parameter (total of population, average of population, proportion of population). In related with stratification, every stratum which is homogenous has their minimum variance of estimation so that the estimation will nearly reach the true value.

The step which the researcher should do before apply stratified random sampling is having a prior data which can be used as a basic of stratification. Based on those data, the researcher could predict the population condition and apply one of the method that is proportional allocation method, Neyman allocation method, and optimum allocation method to determine the sample size.

Two stages stratified random sampling is used to take observations from smaller stratum so that it increases the efficiency of statistics. The first is to select a sample of units, often called the primary unit, and the second is to select a sample of elements from each chosen primary unit, often called secondary unit.

BAB I

PENDAHULUAN

Penelitian merupakan salah satu usaha yang dilakukan manusia untuk memecahkan suatu permasalahan. Penelitian yang dilakukan oleh peneliti mempunyai tujuan untuk memecahkan masalah yang timbul dan sekaligus mengembangkan lebih lanjut ilmu pengetahuan dibalik masalah yang sedang diteliti tersebut.

Penelitian yang dilakukan oleh peneliti melibatkan obyek-obyek yang akan menjadi sasaran penelitian. Obyek penelitian tersebut dapat berhingga jumlahnya, dapat juga tak hingga jumlahnya. Peneliti tidak mungkin meneliti seluruh objek penelitian jika jumlahnya banyak sekali, karena keterbatasan dana, waktu, tenaga. Salah satu cara yang dilakukan oleh peneliti adalah mengambil sebagian obyek penelitian dari seluruh obyek penelitian dan selanjutnya dilakukan observasi atas masalah yang diselidiki. Hasil penelitian dari pengambilan sebagian obyek penelitian tersebut dibuat generalisasi dengan prosedur tertentu, dan akhirnya diharapkan dapat menjawab tujuan penelitian

Sering kali peneliti dihadapkan dengan pertanyaan bagaimana cara mengambil sebagian elemen penelitian yang dapat mewakili elemen penelitian keseluruhan. Selain itu juga bagaimana cara menduga nilai rata-rata populasi, nilai total populasi, dan nilai proporsi populasi. Berdasarkan masalah tersebut, maka skripsi ini bertujuan untuk membantu menjawab dan membahas tentang penarikan sampel yang sesuai dengan tujuan penelitian.

Pada kesempatan ini, penulis akan menyajikan pokok bahasan mengenai suatu bagian penting dari penelitian yaitu teori penarikan sampel, dan secara khusus akan dibahas tentang teknik penarikan sampel acak berlapis. Teori penarikan sampel tersebut merupakan suatu studi tentang hubungan antara populasi dan sampel yang diambil dari populasi tersebut. Teori ini kiranya sangat penting karena dapat digunakan untuk menduga ciri populasi yang tidak diketahui (seperti rata-rata dan variansi populasi), yang disebut parameter populasi, dan berdasarkan apa yang diketahui dari kuantitas sampel yang berasal dari populasi yang bersangkutan (seperti nilai rata-rata dan variansi sampel), yang disebut statistik.

Untuk mempelajari teknik penarikan Sampel Acak Berlapis diperlukan pengetahuan dasar probabilitas. Pengertian sampel, populasi, parameter, statistik dan konsep-konsep dasar probabilitas seperti probabilitas bersyarat, variabel random, nilai harapan, kesalahan penarikan sampel, dan selang kepercayaan akan dibahas pada bab II.

Pada bab III pertama-tama akan dibahas penarikan Sampel Acak Berlapis, dan definisi sampel yang repersentatif. Sub bab ketiga akan membahas cara pemilihan elemen anggota sampel yang meliputi cara undian dan dengan menggunakan tabel angka acak, dan dilanjutkan pembahasan tentang kelemahan penarikan Sampel Acak Sederhana. Sub bab yang kelima membahas penarikan Sampel Acak Berlapis dan dilanjutkan pembahasan mengenai kelebihan penarikan Sampel Acak Berlapis. Sub bab yang ketujuh menjelaskan notasi-notasi yang digunakan dalam penarikan Sampel Acak Berlapis. Sub bab yang kedelapan dan kesembilan masing-masing membahas penentuan ukuran Sampel Acak Berlapis untuk pendugaan nilai rata-rata dan nilai proporsi dimana masing-masing menggunakan 3 metode yaitu metode alokasi sebanding, metode alokasi Neyman, dan metode alokasi optimum. Sub bab kesepuluh membahas pembentukan lapisan dan sub bab terakhir membahas jumlah lapisan.

Bab IV membahas pendugaan parameter dalam penarikan Sampel Acak Berlapis. Tiga sub bab yang pertama akan membahas pendugaan nilai total populasi, nilai rata-rata populasi, proporsi populasi. Sub bab keempat membahas ketelitian relatif pada penarikan Sampel Acak Sederhana dan penrikan Sampel Acak Berlapis. Sub bab kelima membahas sub penarikan Sampel Acak Berlapis yang meliputi pembahasan sub penrikan Sampel Acak Berlapis dengan unit-unit berukuran sama dan unit-unit berukuran tidak sama.

Beberapa materi prasyarat yang diperlukan dalam pembahasan tulisan ini adalah teori himpunan, persamaan diferensial, dan teori probabilitas.

Sebagai pembatasan, beberapa teorema tidak akan dibuktikan, misalnya teorema limit pusat, teorema nilai harapan, dan beberapa definisi tidak tulis seperti definisi kelas kejadian, dan definisi distribusi marjinal. Penulis mengaplikasikannya secara langsung dalam pembahasan-pembahasan yang menggunakan teorema tersebut. Seluruh bab dalam tulisan ini dibatasi pada pengambilan elemen dalam lapisan yang mempunyai probabilitas yang sama untuk terpilih dalam sampel, dan pada pembahasan sub penarikan sampel hanya dibatasi sampai dua tahap.

Metode penulisan yang digunakan dalam tulisan ini adalah metode studi pustaka.

Pada akhir tulisan ini akan diberikan beberapa kesimpulan yang berkaitan dengan pembahasan teori penarikan Sampel Acak Berlapis pada bab-bab sebelumnya.

BAB II

LANDASAN TEORI

Pada bab ini penulis akan membahas pengertian populasi dan sampel, langkah-langkah dalam penarikan sampel. Selain itu akan dibahas pula definisi-definisi dan teorema-teorema yang perlu diketahui sebagai dasar untuk membahas penarikan sampel acak berlapis.

2.1. Pendahuluan

Salah satu usaha yang dilakukan manusia untuk memecahkan masalah di berbagai bidang kehidupan adalah dengan mengadakan penelitian. Pada dasarnya penelitian merupakan suatu bentuk penyelidikan yang bersifat kritis untuk memperoleh keterangan atas sesuatu persoalan tertentu pada suatu daerah atau wilayah. Karena adanya berbagai keterbatasan (biaya, waktu, dan tenaga) penelitian dilakukan dengan hanya mengambil sebagian objek dari sasaran penelitian dan menganalisanya untuk kemudian digeneralisasi. Misalkan, seorang peneliti ingin memperoleh keterangan mengenai pendapatan rumah tangga di daerah A, maka peneliti mengadakan penelitian terhadap rumah tangga di daerah A tersebut. Dalam hal ini peneliti tidak perlu mengamati seluruh rumah tangga yang ada di daerah A, tetapi peneliti hanya mengambil sebagian saja rumah tangga untuk dijadikan sampel penelitiannya. Sejauh mana ketepatan dari kesimpulan tentang sifat populasi berdasarkan metode penelitian akan sangat tergantung pada teknik penarikan sampel. Dengan demikian dalam penelitian, teknik penarikan sampel memegang peranan penting, karena hanya sampel yang benar-benar mewakili populasi yang akan menghasilkan keterangan yang mampu menggambarkan sifat-sifat populasi tersebut.

Beberapa keuntungan yang didapat dengan melakukan pengambilan sampel :

- Biaya berkurang . Data yang didapat berasal dari sejumlah bagian kecil populasi, sehingga biaya yang dikeluarkan akan lebih murah daripada biaya untuk meneliti seluruh populasi.
- Tingkat ketelitian lebih besar. Sebuah sampel mungkin memberikan hasil yang lebih teliti dari pada pencacahan lengkap, jika dipakai tenaga yang sedikit, berkualitas baik dan pengawasan terhadap pekerjaan diperketat, maka hasil dapat diproses dengan baik.
- Kecepatan pengumpulan data lebih besar. Data dikumpulkan dan diringkas lebih cepat dengan sebuah sampel daripada dengan sebuah populasi.

2.2. Pengantar Probabilitas

Probabilitas merupakan konsep dasar yang diperlukan dalam teknik penarikan sampel. Penarikan sampel yang dilakukan berhubungan dengan probabilitas terjadinya suatu kejadian. Bagaimana cara menentukan besar peluang terjadinya suatu kejadian dalam ruang sampel digunakan konsep probabilitas.

Beberapa pengertian, teorema, contoh probabilitas yang berhubungan dengan teknik penarikan sampel akan dibahas .

DEFINISI 2.2.1. Percobaan adalah suatu proses yang secara teoritis dapat diulang tak hingga kali dengan kondisi tidak berubah dan mempunyai hasil yang terdefinisi.

CONTOH 2.2.1. Percobaan melempar sekeping mata uang, percobaan mengukur jumlah curah hujan selama bulan Desember di Kaliurang.

DEFINISI 2.2.2. **Ruang Sampel (S)** adalah himpunan yang unsur-unsurnya menyatakan semua kemungkinan hasil suatu percobaan.

Setiap unsur dari ruang sampel disebut titik sampel.

DEFINISI 2.2.2.1. **Ruang Sampel Diskret** adalah ruang sampel yang berhingga atau tak berhingga terbilang

Sedangkan ruang sampel yang tidak diskret disebut **ruang sampel kontinu.**

CONTOH 2.2.2.1. Sebuah percobaan bertujuan mengamati sisi apa yang muncul dari percobaan melemparkan dua mata uang bersama-sama satu kali. Ruang sampelnya adalah : S = {MM,MB,BM,BB}, M adalah sisi muka, dan B adalah sisi belakang.

Sedangkan MM merupakan salah satu titik sampel dari S.

CONTOH 2.2.2.2. Percobaan mengamati daya hidup (dalam satuan waktu) lampu akan menghasilkan ruang sampel kontinu.

DEFINISI 2.2.3. **Kejadian** adalah himpunan bagian dari ruang sampel S

CONTOH 2.2.3. Dari contoh 2.7.2.1. A = {BB,MM} adalah kejadian munculnya sisi sama.

DEFINISI 2.2.4. Probabilitas Klasik

Jika suatu percobaan dapat menghasilkan N titik sampel yang berbeda dan masing-masing berkemungkinan sama untuk terjadi, dan jika tepat ada sebanyak n dari titik-titik sampel tersebut merupakan unsur dari kejadian A, maka probabilitas kejadian A adalah P(A) = n/N.

DEFINISI 2.2.5. Fungsi Probabilitas

Fungsi probabilitas P(.) adalah suatu fungsi dengan domain himpunan kelas kejadian (yang dilambangkan dengan U) dan kodomain interval [0,1] yang memenuhi aksioma-aksioma berikut :

- Bernilai tak negatif $P(A) \geq 0, \forall A \in U$
- $P(S) = 1$
- Jika $A_1, A_2, \ldots$ adalah kejadian-kejadian yang saling asing dalam U (yaitu $A_i \cap A_j = \varnothing$ untuk $i \neq j$, $i,j = 1,2,\ldots$) dan jika

  $A_1 \cup A_2 \cup \ldots = \bigcup_{i=1}^{\infty} A_i \in U$, maka $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

P(A) dibaca " probabilitas kejadian A" atau "probabilitas bahwa kejadian A terjadi ".

Sifat-sifat P(.)

Untuk setiap teorema berikut, andaikan bahwa S dan U diberikan dan P(.) adalah fungsi probabilitas yang mempunyai domain U.

TEOREMA 2.2.1. $P(\varnothing) = 0$

Bukti :

$A \cap \varnothing = \varnothing$ dan $A \cup \varnothing = A$

A dan $\varnothing$ adalah dua kejadian yang saling asing

Maka $P(A \cup \varnothing) = P(A) + P(\varnothing)$

$P(A) \quad = P(A) + P(\varnothing)$

$P(\varnothing) \quad = P(A) - P(A) = 0$

TEOREMA 2.2.2. Jika $A \subset B$ maka $P(A) \le P(B)$

Bukti :

Karena $A \subset B$ maka $B = A \cup (A^c \cap B)$ ... (1)

dan $A \cap (A^c \cap B) = \varnothing$ ...(2)

Dari (1) $P(B) = P(A \cup (A^c \cap B))$. Karena (2), maka $P(B) = \underset{\ge 0}{P(A)} + \underset{\ge 0}{P(A^c \cap B)}$

Karena $P(A^c \cap B) \ge 0$ maka $P(A) \le P(B)$

TEOREMA 2.2.3.

Jika A dan B adalah sembarang kejadian, maka

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Bukti :

Pendekatan yang digunakan adalah mengekspresikan kejadian $A \cup B$ dan A sebagai gabungan dari kejadian yang saling lepas.

$A \cup B = (A \cap B^c) \cup B$

$A = (A \cap B) \cup (A \cap B^c)$

jadi $P(A \cup B) = P(A \cap B^c) + P(B)$

karena $P(A) = P(A \cap B) + P(A \cap B^c)$

maka $P(A \cup B) = P(A \cap B^c) + P(B)$

$= [P(A) - P(A \cap B)] + P(B)$

$= P(A) + P(B) - P(A \cap B)$

CONTOH 2.2.5 probabilitas

Dari contoh 2.2.2. diketahui S = {MM,MB,BM,BB}. Probabilitas kejadian munculnya sisi sama adalah $P(A) = P\{BB,MM\} = {}^2/_4 = ½$

2.3. Probabilitas Bersyarat

Probabilitas terjadinya suatu kejadian A jika diketahui kejadian B terjadi disebut probabilitas bersyarat dan dinyatakan dengan lambang P(A/B), dibaca probabilitas A dengan syarat B terjadi.

DEFINISI 2.3.1. Probabilitas Bersyarat

Probabilitas bersyarat kejadian A, jika diketahui kejadian B terjadi didefinisikan sebagai :

$$P(A/B) = \quad = \frac{P(A \cap B)}{P(B)} \text{ , dimana } P(B) > 0$$

DEFINISI 2.3.2. Kejadian Saling Bebas

Kejadian A dan B disebut 2 buah kejadian yang bebas bila dan hanya bila $P(A \cap B) = P(A).P(B)$

TEOREMA 2.3.1.

$$P(A \cap B) = P(A).P(B/A)$$

$$= P(B).P(A/B)$$

<u>Bukti</u> : akibat langsung dari definisi probabilitas bersyarat.

CONTOH 2.3.1. Probabilitas bersyarat

Petugas kantor pos yang bertugas mengantarkan surat terdiri dari 12 orang laki-laki dan 8 wanita. Dari 12 orang laki-laki tersebut terdapat 4 orang gendut dan 8 orang kurus. Diantara 8 wanita terdapat 5 orang gendut dan 3 orang kurus. Jika setiap orang mempunyai peluang sama untuk bertugas mengantarkan surat ke daerah tertentu dan misalkan bahwa :

L = laki-laki

W = wanita

G = gendut

K = kurus

maka : $P(L) = {}^{12}/_{20}$ $\quad P(W) = {}^{8}/_{20}$ $\quad P(G) = {}^{9}/_{20}$ $\quad P(K) = {}^{11}/_{20}$

Tetapi jika orang yang datang mengantarkan surat adalah gendut, maka peluang bahwa orang tersebut laki-laki adalah:

$P(L/G) = P(L \cap G) / P(G) = ({}^{4}/_{20}) / ({}^{9}/_{20}) = {}^{4}/_{9}$

2.4. Variabel Random

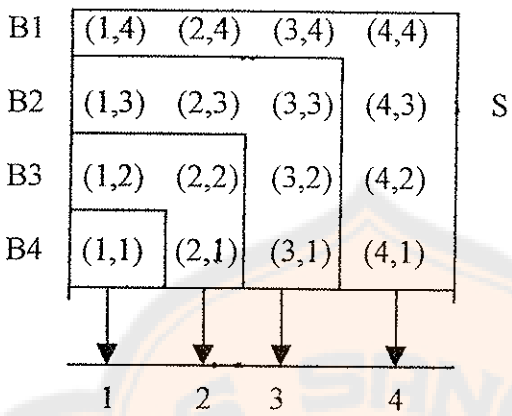
Peluang timbulnya suatu kejadian dalam ruang sampel dideskripsikan dalam model matematika yang diekspresikan dalam bentuk nilai-nilai numeris dari hasil percobaan . Hal tersebut menimbulkan gagasan untuk mendefinisikan sebuah fungsi yang dikenal dengan variabel random, yang memetakan setiap hasil dalam suatu percobaan dengan bilangan real.

DEFINISI 2.4.1. Variabel Random

Variabel random, misalnya X adalah suatu fungsi yang didefinisikan pada ruang sampel S, yang memetakan setiap elemen $e \in S$ ke bilangan real, $X(e)=x$

CONTOH 2.4. Variabel Random

Sebuah tetrahedral diberi nomor masing-masing sisinya dengan bilangan tetapi semua kemungkinan nilainya dapat didaftarkan dan dapat didefinisikan sebuah variabel random X. Secara khusus, jika $e = (i,j)$, $i,j \in \{1,2,3,4\}$ maka $X(e) = \max(i,j)$. Ruang sampel S dan X digambarkan dalam diagram berikut :



Dari diagram tersebut terlihat bahwa 16 titik sampel dalam S dipetakan ke bilangan real 1,2,3,4 oleh fungsi (variabel random) X

$X(1,4) = 4, \quad X(2,3) = 3$

DEFINISI 2.4.2. Variabel Random Diskrit

Suatu variabel random disebut variabel random diskrit bila daerah hasilnya merupakan suatu himpunan diskrit.

CONTOH 2.4.2. Variabel random diskrit

Banyak barang yang cacat dalam sampel k barang atau banyak korban meninggal di suatu jalan raya pertahun.

DEFINISI 2.4.3. Variabel Random Kontinu

Suatu variabel random disebut variabel random kontinu bila daerah hasilnya merupakan suatu himpunan non diskrit (kontinu).

CONTOH 2.4.3 Variabel random kontinu

Pada pengukuran tinggi badan siswa SD kelas 5, jika X adalah variabel random yang menyatakan tinggi badan siswa SD kelas 5 maka X adalah variabel random kontinu karena X menjalani nilai-nilai dalam suatu interval.

DEFINISI 2.4.4. Fungsi Probabilitas Diskret

Probabilitas variabel random diskrit didefinisikan sebagai berikut : $P(X = x) = P(\{e \in S \mid X(e) = x\})$

$P(X = x)$ biasa ditulis dengan $f(x_i)$

Jadi f adalah fungsi probabilitas diskrit jika memenuhi sifat berikut :

1. $f(xi) \geq 0$
2. $\sum_{\text{semua } x_i} f(x_i) = 1$

CONTOH 2.4.4. Fungsi probabilitas binomial didefinisikan sebagai berikut

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0,1,2,\ldots, n$$

$P(X=k) > 0$ untuk setiap k dan melalui teorema binomial,

$$1 = [p+(1-p)]^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}.$$

Untuk X = k dalam perluasan binomial $[p+(1-p)]^n$ memberikan probabilitas munculnya nilai dari k. Distribusi yang didefinisikan di atas dinamakan distribusi binomial dan X dinamakan variabel random binomial.

DEFINISI 2.4.5. Fungsi Probabilitas Kontinu

Jika X variabel random kontinu maka suatu kejadian akan berkaitan dengan suatu interval. Probabilitas variabel random X terletak antara a dan b atau $P(a \leq X \leq b)$ dapat diperoleh dengan mengandaikan ada fungsi f(x) sedemikian hingga luas dibawah kurva fungsi ini pada interval [a,b] sama dengan $P(a \leq X \leq b)$

Jadi $P(a \leq X \leq b) = \int_a^b f(x)dx$ Fungsi probabilitas variabel random kontinu juga dikenal dengan nama fungsi densitas.

f(x) merupakan fungsi densitas kontinu jika memenuhi sifat berikut :

1. $f(x) \geq 0$ untuk semua $x \in \Re$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

CONTOH 2.4.5.

Distribusi normal. Variabel random X dikatakan berdistribusi normal bila fungsi probabilitasnya berbentuk:

$$f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi.\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \text{untuk } -\infty < X < \infty \quad ,$$

dimana $\pi = 3{,}14159$ dan $e = 2{,}171828$, $\sigma \geq 0$, μ real.

Distribusi peluang kontinu yang paling penting dalam bidang statistika adalah seperti lonceng yang dapat digunakan dalam banyak sekali kumpulan data yang terjadi di dalam praktek-praktek penelitian. Persamaan matematik bagi distribusi peluang variabel random normal tersebut bergantung pada dua parameter mean dan simpangan baku. Oleh karena itu dilambangkan nilai-nilai fungsi densitas normal bagi X adalah N(X, μ , σ). Pada tulisan ini distribusi normal berperan dalam pembahasan selang kepercayaan, kesalahan sampling.

DEFINISI 2.4.6. Variabel Random Bebas Stokastik

Variabel random X_i disebut bebas stokastik bila fungsi densitas bersamanya

$$f_{X_{1,2,\ldots,k}}(x_1, x_2, \ldots, x_k) = \begin{cases} \prod_{i=1}^{n} p(x_i), \text{untuk } X_i \text{ diskrit} \\ \prod_{i=1}^{n} f(x_i), \text{untuk } X_i \text{ kontinu} \end{cases}$$

Konsep variabel random bebas stokastik sangat penting untuk memberikan dasar konsep sampel acak yang akan dibahas dalam bab Sampel Acak Berlapis.

DEFINISI 2.4.7. Sampel Acak

Jika diketahui variabel random $X_1, X_2, \ldots, X_n$ mempunyai suatu fungsi densitas $f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)$ dapat difaktorkan menjadi $f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n)$. $f(.)$ adalah fungsi densitas untuk masing-masing X_i, maka $X_1, X_2, \ldots, X_n$ disebut sampel acak berukuran n dari suatu populasi dengan fungsi densitas $f(.)$.

2.5. Nilai Harapan

Konsep nilai harapan memegang peranan penting dalam statistika. Contoh yang paling mudah adalah rata-rata dan variansi suatu variabel random. Keduanya adalah parameter-parameter yang hampir selalu muncul dalam teknik-teknik analisis statistika elementer atau lanjut. Yang dimaksud dengan nilai harapan dinyatakan dengan definisi berikut.

DEFINISI 2.5.1. Nilai Harapan, E(X)

$$E(X) = \begin{cases} \sum_{\forall x_i} x_i \cdot p(x_i), & \text{jika X diskrit dengan fungsi probabilitas } p(x) \\ \int_{-\infty}^{\infty} x.f(x)dx, & \text{jika X kontinu dengan fungsi densitas } f(x) \end{cases}$$

TEOREMA 2.5. Nilai Harapan, E(X)

Misalkanlah X suatu variabel random dengan distribusi peluang f(x).

Nilai harapan fungsi g(x) adalah

$$E[g(X)] = \sum_{\forall x} g(x)f(x) \quad \text{bila X diskrit}$$

$$= \int_{-\infty}^{\infty} g(x)f(x)dx \quad \text{bila X kontinu}$$

TEOREMA SIFAT-SIFAT NILAI HARAPAN 2.5.

2.5.1. Bila a dan b konstanta, maka E (aX + b) = aE(X) + b

Bukti : Menurut teorema nilai harapan,

$$E(aX + b) = \int_{-\infty}^{\infty} (ax+b)f(x)dx = a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx$$

Integral pertama di sebelah kanan adalah E(x) dan integral kedua sama dengan 1. Jadi E (aX + b) = a E(X) + b.

2.5.2. E (X+Y) = E(X) + E(Y)

Bukti : Misalkan g(x,y) adalah fungsi dari variabel random X dan Y.

$$E(g(x,y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\, f(x,y)\, dx\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\,f(x,y)dx\,dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\,f(x,y)dx\,dy$$

$$= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x,y)\,dy\,dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_{-\infty}^{\infty} x\,f(x)\,dx + \int_{-\infty}^{\infty} y\,f(y)\,dy$$

$$= E(X) + E(Y)$$

2.5.3. Misalkan X dan Y dua variabel random independen, maka

$$E(XY) = E(X)\,E(Y)$$

Bukti : $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)\,dx\,dy$

Karena X dan Y independen, maka dapat ditulis $f(x,y) = g(x)\,h(y)$ dengan $g(x)$ dan $h(y)$ menyatakan masing-masing distribusi marginal X dan Y.

Jadi, $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,g(x)\,h(y)\,dx\,dy$

$$= \int_{-\infty}^{\infty} x\,g(x)\left[\int_{-\infty}^{\infty} y\,h(y)\,dy\right]dx = \int_{-\infty}^{\infty} x\,g(x)\,E(y)\,dx$$

$$= E(y)\int_{-\infty}^{\infty} x\,g(x)\,dx = E(Y)\,E(X)$$

CONTOH 2.5.1.1.

Carilah nilai harapan banyaknya kimiawan dalam panitia 3 orang yang dipilih secara acak dari 4 kimiawan dan 3 biolog.

Jawab :

Misalkan X menyatakan banyaknya kimiawan dalam panitia. Distribusi peluang X adalah $f(x) = \dfrac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}, x = 0,1,2,3$

Beberapa perhitungan menghasilkan $f(0) = {}^{1}/_{35}$, $f(1) = {}^{12}/_{35}$ $f(2) = {}^{18}/_{35}$, $f(3) = {}^{4}/_{35}$

Jadi, $E(X) = (0)\,({}^{1}/_{35}) + (1)\,({}^{12}/_{35}) + (2)\,({}^{18}/_{35}) + (3)\,({}^{4}/_{35}) = {}^{12}/_{7} = 1{,}7$

Jadi bila suatu panitia beranggotakan 3 orang dipilih secara acak berulang-ulang dari 4 kimiawan dan 3 biolog, maka rata-ratanya akan beranggota 1,7 kimiawan.

CONTOH 2.5.1.2.

Diberikan variabel random kontinu X dengan fungsi densitas

$$f(x) = \begin{cases} \frac{1}{2}, 0 \le x \le 2 \\ 0, \text{selainnya} \end{cases}$$

maka nilai harapannya $E(X) = 0\int_{-\infty}^{0} x.dx + \frac{1}{2}\int_{0}^{2} x\,dx + 0\int_{2}^{\infty} x\,dx = \frac{1}{4}x^2\Big]_0^2 = 1$

2.6. Variansi dan Kovariansi

Salah satu nilai harapan yang penting adalah variansi yang merupakan nilai harapan fungsi $g(x) = (x-\mu)^2$, dimana μ=E(x)

DEFINISI 2.6.1. Variansi variabel random X

$$\begin{aligned} \text{Var}(X) &= E[(X-\mu)^2] \\ &= E[X^2] - \mu^2 \end{aligned}$$

Akar pangkat dua dari Var (X) adalah standar deviasi dari X dan diberi notasi σ_X.

Kegunaan dari variansi adalah untuk mengukur keragaman data.

TEOREMA 2.6.

Bila a adalah kontanta, maka Var(aX) = a^2 Var X

Bukti : $\text{Var}(aX) = E(aX-a\mu)^2$

$$= a^2 E(X-\mu)^2$$

$$= a^2 \text{ Var } X$$

DEFINISI 2.6.2. Kovariansi dari X dan Y

$$\mathrm{Kov}(X,Y) = E[(X-\overline{X})(Y-\overline{Y})]$$

$$= E(X.Y) - E(\overline{X}).E(\overline{Y})$$

2.7. Populasi dan Sampel

Dalam suatu penelitian, berdasarkan ukuran populasinya dibedakan menjadi 2 yaitu populasi terhingga dan tak terhingga. Misalkan seorang peneliti ingin mengetahui mutu SMU di Yogyakarta, melalui NEM siswa SMU di Yogyakarta. Dalam hal ini peneliti akan memperoleh data yang terhingga banyaknya. Pada pihak lain, jika kita menentukan daya tahan hidup sebuah lampu setiap kali mencatat waktu pada saat lampu tersebut mati, kita akan mendapatkan himpunan nilai yang tak hingga banyaknya.

Keseluruhan pengamatan yang menjadi perhatian peneliti baik terhingga maupun tak hingga, menyusun apa yang disebut *populasi* . Populasi berkaitan dengan sembarang pengamatan yang menarik perhatian peneliti, apakah itu sekelompok orang, binatang, bilangan atau benda apa saja. Berikut akan didefinisikan pengertian populasi.

DEFINISI 2.7.1. Populasi adalah

- himpunan (yang lengkap atau sempurna) dari semua *unit observasi yang mungkin*, di mana unit observasi adalah individu atau kelompok yang dapat memberikan keterangan tentang apa yang ingin diamati atau dipelajari oleh seorang peneliti (I Gusti Ngurah Agung, 1992:12).

- himpunan berhingga atau tak hingga dari individu- individu. Populasi tersebut analog dengan istilah semesta pembicaraan dalam teori himpunan. Secara praktis, populasi sinonim dengan suatu kumpulan yang anggotanya tidak selalu berupa organisme hidup (Kendall,S.M.G, W.r. Buckland. 1982).

Kedua definisi tersebut pada dasarnya mempunyai inti yang sama dan saling melengkapi.

Populasi merupakan suatu kumpulan, yang pendefinisiannya tergantung pada minat dan kepentingan masing-masing peneliti. Peneliti yang satu mungkin ingin membuat pengamatan tentang karakteristik mahasiswa–mahasiswa di semua perguruan tinggi di Indonesia. Peneliti yang lain mungkin ingin meneliti karakteristik mahasiswa-mahasiswa di suatu perguruan tinggi tertentu saja. Kedua peneliti tersebut sama-sama menganggap populasi sebagai kumpulan mahasiswa yang akan menjadi objek penelitian.

Dalam beberapa konteks tertentu, peneliti juga menggunakan istilah populasi untuk kumpulan pengukuran yang dilakukan terhadap sekumpulan orang, tempat atau benda. Sebagai contoh, peneliti berminat menyelidiki usia semua mahasiswa di suatu perguruan tinggi, yang dipandang sebagai populasi adalah kumpulan usia mahasiswa tersebut. Populasi dapat bersifat nyata (real), dapat juga hipotetik. Contoh populasi nyata adalah segenap mahasiswa yang sekarang terdaftar di suatu universitas. Sedangkan contoh populasi hipotetik digambarkan dalam contoh berikut. Seorang peneliti sedang merancang sebuah eksperimen untuk mengevaluasi keefektifan tiga jenis obat penenang, dan merencanakan akan memilih sejumlah kelinci percobaan secara acak yang masing-masing akan diberi salah satu dari ketiga jenis obat tersebut. Peneliti dapat menganggap masing-masing dari ketiga kelompok tersebut sebagai sampel dari suatu populasi yang beranggotakan sejumlah besar kelinci percobaan yang dapat diberi obat tersebut. Populasi semacam ini secara riil belum ada, namun bersifat hipotetik (masih mungkin ada).

CONTOH 2.7.1.

Andaikan seorang peneliti ingin mengetahui prestasi belajar siswa di SMU Yogyakarta, melalui rata-rata NEM SMU di Yogyakarta, populasinya adalah semua NEM SMU di Yogyakarta.

Inferensia Statistika mencakup semua metode yang berhubungan dengan analisis sebagian data untuk kemudian sampai pada peramalan atau penarikan kesimpulan mengenai keseluruhan himpunan data induknya meskipun peneliti tidak mungkin atau tidak praktis untuk mengamati keseluruhan individu untuk menyusun populasi, karena keterbatasan dana, waktu, dan tenaga. Peneliti ingin memperoleh kesimpulan mengenai populasi. Oleh karena itu, peneliti dapat memilih sebagian anggota populasi untuk membantu peneliti menarik kesimpulan mengenai populasi tersebut. Berikut akan didefinisikan pengertian sampel.

DEFINISI 2.7.2. Sampel adalah himpunan bagian dari populasi

CONTOH 2.7.2 NEM dari 5 SMU di Yogyakarta.

2.8. Populasi Sasaran dan Populasi yang Disampelkan

Kata populasi digunakan untuk menyatakan kumpulan dari mana sampel dipilih. Definisi populasi mungkin tidak menimbulkan masalah jika penarikan sampel dilakukan dari kelompok bola lampu listrik yang digunakan untuk memperkirakan rata-rata lamanya sebuah bola lampu menyala. Akan tetapi, dalam pemilihan sampel dari sebuah populasi areal pertanian akan muncul kasus-kasus batas. Kasus tersebut harus ditangani secara praktis, yaitu pelaksana penelitian harus dapat memutuskan di lapangan tanpa banyak keraguan, apakah unit yang meragukan tersebut merupakan bagian dari populasi tersebut atau tidak. Dengan kata lain, populasi yang menjadi sasaran penelitian harus terdefinisi secara jelas.

Contoh berikut akan dipakai untuk memperjelas pengertian populasi sasaran. Misalkan menteri P&K ingin mengetahui mutu SMU di P. Jawa. Karena SMU di P. Jawa banyak, maka peneliti membagi daerah menjadi 3 lapisan, yaitu SMU yang berada di kota besar, SMU yang berada di kota sedang, SMU yang berada di kota kecil, dengan alasan peneliti mengasumsikan bahwa 3 lapisan tersebut menggambarkan mutu SMU yang berbeda-beda. Kota besar mempunyai sekolah yang lebih banyak, didukung tenaga pengajar yang profesional dan fasilitas atau sarana belajar mengajar yang lebih menunjang, kota sedang didukung tenaga pengajar dan sarana belajar mengajar yang cukup memadai. Demikian juga sebaliknya dengan kota kecil mempunyai jumlah sekolah yang lebih sedikit dan sarana belajar mengajar yang kurang memadai. Penjelasan 3 lapisan tersebut berhubungan dengan mutu sekolah yang bersangkutan. Tiap lapisan dibagi lagi menjadi 2 sub lapisan, yaitu SMU swasta dan SMU negeri. Secara umum, peneliti masih menganggap adanya perbedaan mutu SMU swasta dan SMU negri karena SMU negri dari segi fasilitas dan pengadaan guru mendapat subsidi dari pemerintah sedangkan sekolah swasta harus mengadakan secara swadaya. Hal ini dapat mempengaruhi perbedaan kemampuan siswa negri dan swasta, meskipun beberapa SMU swasta memperlihatkan mutu yang lebih baik dari SMU negri. Langkah berikutnya, tiap lapisan diambil sampel secara acak (misalnya tiap sub lapisan diambil 5 sekolah atau tergantung jumlah SMU di tiap lapisan), untuk diadakan penelitian. Berikut didefinisikan populasi sasaran yang dijelaskan diatas.

DEFINISI 2.8.1. Populasi Sasaran adalah keseluruhan elemen yang akan dibicarakan dimana informasi akan diambil.

Populasi sasaran harus well defined (terdefinisi dengan baik), dapat berupa populasi yang nyata atau hipotetis. Obyek penelitian adalah menemukan suatu populasi sasaran tertentu, biasanya dapat berupa populasi secara keseluruhan atau peneliti bisa melakukan pembatasan yang lebih terinci tentang populasi yang diperhatikannya.

CONTOH 2.8.1. Berdasarkan penelitian di atas, yang menjadi populasi sasaran adalah seluruh SMU di P. Jawa.

DEFINISI 2.8.2. Populasi yang Disampelkan.

Jika $X_1,X_2,X_3,\ldots,X_n$ suatu variabel random dari suatu populasi dengan fungsi densitas f (.), maka populasi ini disebut populasi yang disampelkan.

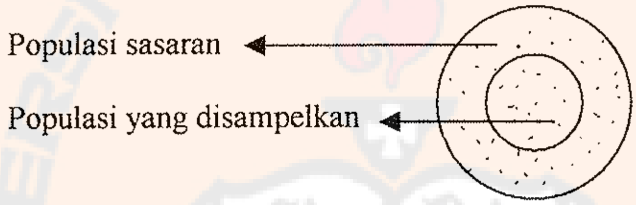
CONTOH 2.8.2.

Dari kasus diatas, seluruh SMU di tiga lapisan merupakan populasi yang disampelkan. Keseluruh SMU di tiga lapisan tersebut merupakan populasi yang dipakai untuk mengambil sampel dalam melakukan penelitian. Dalam praktek sering diasumsikan nilai-nilai berasal dari fungsi densitas normal.

2.9. Hubungan Populasi Sasaran dan Populasi yang Disampelkan

Hubungan antara populasi sasaran dan populasi yang disampelkan dapat dijelaskan sebagai berikut. Populasi yang disampelkan merupakan himpunan bagian dari populasi sasaran. Semua elemen dalam populasi yang disampelkan adalah elemen dalam populasi sasaran, tetapi elemen dalam populasi sasaran belum tentu menjadi elemen dalam populasi yang disampelkan.

Hubungan tersebut dapat digambarkan dalam diagram berikut:



CONTOH 2.9. Lihat contoh 2.8.1. dan contoh 2.8.2. SMU di tiga lapisan merupakan himpunan bagian dari SMU di P. Jawa.

2.10. Kerangka (frame) dan Teknik Pembentukan Kerangka

Peneliti membutuhkan populasi yang jelas, dan elemen populasi juga harus diketahui oleh peneliti. Sebelum sampel diambil dari populasi, populasi dibagi dalam bagian-bagian yang disebut unit penarikan sampel. Unit tersebut harus mencakup seluruh populasi dan tidak boleh tumpang tindih, dalam arti bahwa setiap sampel dalam populasi hanya menjadi anggota satu dan hanya satu unit. Kadang-kadang unit yang cocok sudah jelas, seperti dalam sebuah populasi bola lampu, unitnya hanya satu bola lampu. Kadang-kadang ada pilihan terhadap unit. Dalam memilih sampel penduduk sebuah kota, unitnya dapat terdiri dari seorang anggota keluarga. Dalam penarikan sampel sebuah tanaman pertanian, unitnya dapat berupa sebuah lapangan, daerah pertanian yang bentuk dan ukurannya memenuhi aturan yang ditentukan. Para penarik sampel harus bersikap kritis terhadap daftar-daftar yang telah dikumpulkan sesuai dengan tujuannya. Pada kenyataannya sering daftar tersebut tidak sesuai dengan tujuan, daftar-daftar tidak lengkap, atau berisi sejumlah duplikasi yang tidak diketahui. Dalam praktek sering dijumpai kesulitan memperoleh kerangka yang baik. Oleh karena itu, diperlukan pendefinisian yang jelas tentang unit-unit penarikan sampel yang identitasnya akan dicantumkan dalam kerangka

DEFINISI 2.10. Kerangka adalah daftar unit penarikan sampel

CONTOH 2.10.

1. Berdasarkan kasus 2.8.1. kerangkanya dapat berupa semua nama SMU di P. Jawa.
2. Pengamatan sebuah pasar, kerangkanya berupa daftar semua nama pedagang di pasar tersebut.

2.10.1. Tipe-tipe Pembentukan Kerangka

Terdapat 2 tipe kerangka :

1. Tipe Daftar

Populasi yang diketahui dibagi dalam beberapa unit penarikan sampel secara terpisah. Tiap unit tidak boleh tumpang tindih. Format umum dari tipe ini adalah:

Unit sampling no 1

Unit sampling no 2

....

unit sampling no M

CONTOH tipe daftar :

Diketahui populasi nomor seri majalah yang terbit tiap hari untuk tahun 1990-1997. Dilakukan pembagian nomor seri tiap tahun.

Untuk tahun 1990 = 90001,90002,...,90365

Untuk tahun 1992 = 91001.91002,...,91365

Untuk tahun 1997 = 97001,97002,...,97365

Jadi ada 7 unit penarikan sampel yang berupa seluruh nomor seri majalah yang terbit tiap hari untuk tiap tahun.

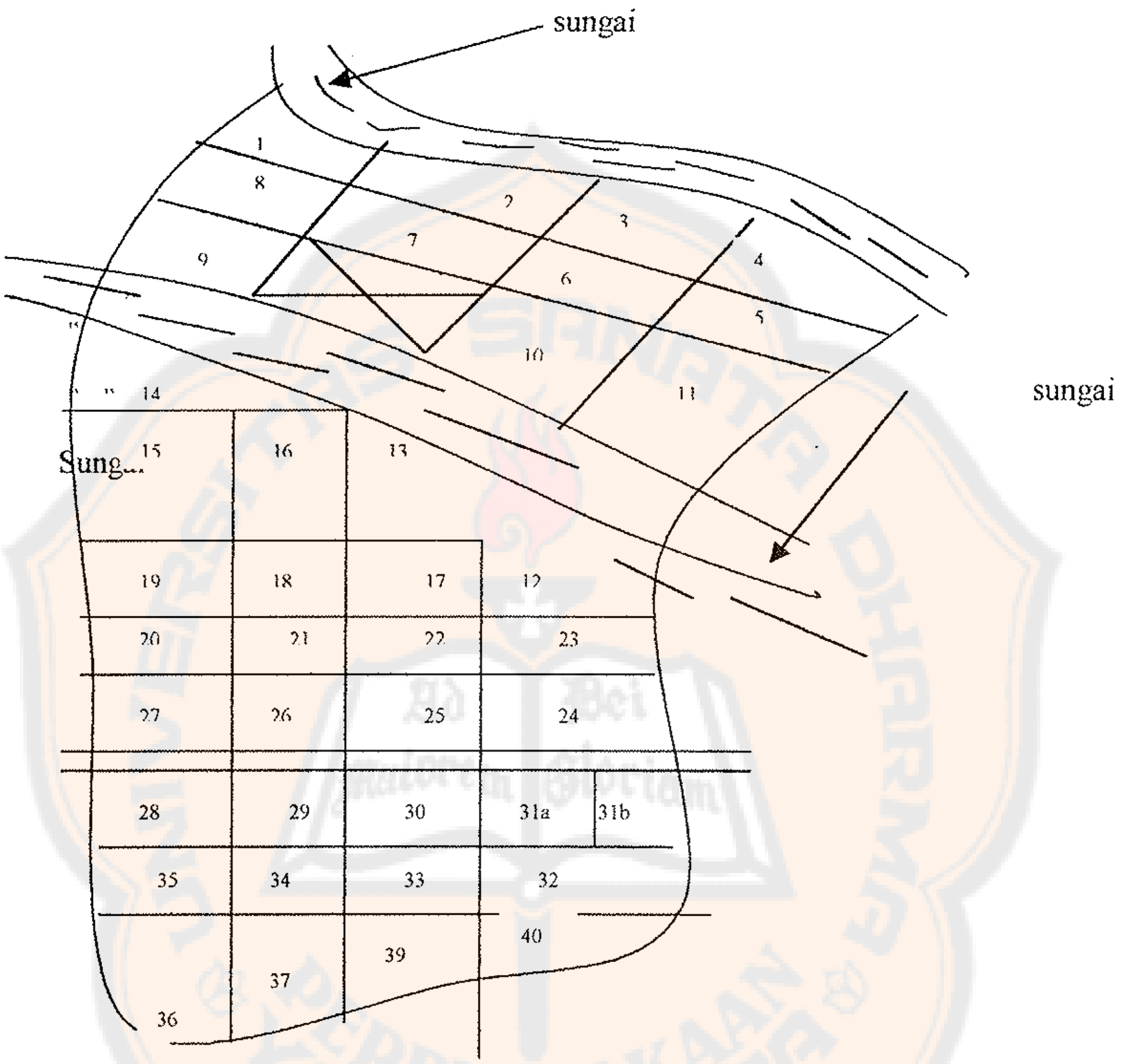
Keseluruhan populasi adalah {90001,90002,...,97365}, seluruh nomor seri majalah yang terbit tiap hari dari tahun 1990-1997.

2. Tipe Pemetaan

Format umum dari tipe pemetaan adalah suatu peta dari beberapa daerah (misalnya suatu negara, provinsi, dll) yang dibagi kedalam beberapa daerah-daerah kecil.

CONTOH tipe pemetaan :

Penelitian diadakan di suatu kota besar yang terdiri dari rumah-rumah (wilayah) jalan-jalan, dan sungai. Kota besar tersebut dibagi dalam daerah-daerah yang kecil, misalnya kelurahan (pada gambar ada pada daerah yang bernomor), yang akan menjadi unit penarikan sampel.


sungai
sungai
Sung...
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31a
31b
32
33
34
35
36
37
39
40

2.11. Langkah-langkah dalam penarikan sampel

Ada beberapa langkah yang perlu diperhatikan dalam penarikan sampel :

1. Merumuskan tujuan dari survei.
2. Mendefinisikan populasi penelitian secara jelas yang merupakan agregat dari mana sampel dipilih.
3. Mengumpulkan data yang relevan berdasarkan tujuan survei.
4. Memilih instrumen pengukuran dan metode pengumpulan data yang lengkap.
5. Menentukan kerangka penarikan sampel.
6. Memilih sampel dari kerangka penarikan sampel yang telah ditetapkan, dengan menentukan ukuran sampel yang didasarkan pada pertimbangan tingkat ketelitian yang diinginkan oleh peneliti, kendala dana, waktu, dan tenaga yang ada.
7. Pengorganisasian pekerjaan lapangan.

2.12. Parameter dan Statistik

Notasi yang digunakan statistikawan dalam mengolah data statistik sepenuhnya tergantung pada apakah data tersebut merupakan populasi atau suatu sampel yang diambil dari suatu populasi. Misalkan sekelompok data berikut yang berupa banyaknya kesalahan ketik pada tiap halaman yang dilakukan oleh seorang sekretaris ketika mengetik sebuah dokumen setebal 10 halaman: 1,0,1,2,3,1,1,4,0, dan 2. Misalkan dapat dikatakan bahwa banyaknya kesalahan terbesar adalah 4, atau menyatakan nilai tengah hitung (rata-rata) 10 hitungan itu adalah 1,5. Bilangan 4 dan 1,5 merupakan deskripsi bagi populasi tersebut. Nilai-nilai tersebut disebut parameter. Sudah menjadi kebiasaan untuk melambangkan parameter dengan huruf Yunani. Secara tradisi rata-rata populasi dilambangkan dengan μ. Jadi untuk populasi kesalahan ketik, $\mu = 1{,}5$.

DEFINISI 2.12.1. Parameter adalah suatu nilai berdasarkan data yang diobservasi dari suatu populasi secara keseluruhan.

Dari definisi tersebut parameter ditafsirkan sebagai karakteristik dari populasi. Dalam teori probabilitas populasi dicirikan dengan distribusi probabilitas dari variabel random yang merupakan fungsi dari parameter. Oleh karenanya distribusi probabilitas variabel random (P) sering diindekskan dengan suatu parameter θ, θ (tidak selalu nilai real) merupakan anggota dari ruang parameter (Ω), di mana Ω merupakan himpunan dari semua nilai yang mungkin dari parameter θ. P adalah satu anggota dari beberapa keluarga distribusi ($\wp = \{ P_{\theta}, \theta \in \Omega \}$.

Contoh parameter antara lain adalah rata-rata populasi μ, variansi populasi σ^2. Parameter biasanya tidak diketahui, dan dengan statistiklah nilai parameter tersebut ditaksir atau diduga. Sebagai contoh rata-rata sampel $\overline{X}$ untuk menduga rata-rata populasi μ yang tidak diketahui dari populasi sampelnya yang diambil.

CONTOH 2.12.1. Rata-rata banyaknya kesalahan ketik yang dilakukan oleh sekretaris adalah 1-2 per halaman.

Misalkan bahwa data tersebut merupakan sebuah sampel 10 halaman yang diambil dari sebuah naskah yang jauh lebih tebal. Jelaslah bahwa sekarang populasinya tersusun atas data yang jauh lebih besar, dan hanya dimiliki informasi sebagian yang diberikan oleh sampel. Dengan demikian 4 dan 1,5 menjadi ukuran deskripsi sampel, dan tidak lagi merupakan parameter populasi. Suatu nilai yang dihitung dari sampel disebut statistik.

DEFINISI 2.12.2. **Statistik** adalah suatu fungsi dari variabel random yang diobservasi dalam suatu sampel.

Statistik digunakan untuk membuat kesimpulan (atau pendugaan) tentang parameter populasi yang tidak diketahui. Statistik yang sering dijumpai adalah rata-rata sampel $\bar{x}$, variansi sampel s^2, yang masing-masing dapat dipakai untuk menduga μ dan σ^2.

CONTOH 2.12.2. Dari sampel acak kesalahan ketik diperoleh $\bar{x} = 1,5$.

2.13. Distribusi Sampling Statisitik

Penelitian yang dilakukan biasanya menghasilkan statistik untuk menduga parameter populasi. Suatu statistik dihitung dari suatu sampel yang diambil dari suatu populasi, dan berdasarkan statistik tersebut dibuat pernyataan mengenai nilai parameter populasi. Misalkan petugas perusahaan ingin mengatur mesin minuman agar dapat mengeluarkan minuman ringan sebanyak 240 ml setiap kali. Petugas perusahaan tersebut melakukan percobaan pada 40 gelas dan mendapatkan $\bar{x} = 236$ ml, dan berdasarkan nilai tersebut diputuskan bahwa mesin tersebut secara rata-rata mengeluarkan minuman sebanyak $\mu = 240$ ml setiap kali. Ke-40 gelas tersebut dapat dianggap sebagai suatu sampel dari suatu populasi tak hingga minuman yang dikeluarkan oleh mesin tersebut.

Karena banyak sekali sampel acak yang mungkin dapat ditarik dari suatu populasi yang sama , maka dapat dikatakan bahwa setiap statistik apakah , S^2 akan bervariasi dari sampel yang satu ke sampel yang lain. Jadi suatu staitistik sesungguhnya merupakan suatu variabel random yang nilainya bergantung pada sampel yang diamati. Karena suatu statistik merupakan variabel random maka statistik mempunyai sebaran (distribusi).

DEFINISI 2.13. Distribusi Sampling Statistik

Distribusi Sampling Statistik adalah distribusi probabilitas suatu statistik.

CONTOH 2.13.

Distribusi peluang untuk $\overline{X}$ adalah distribusi sampling untuk rata-rata. Wajar bila menyebut simpangan baku distribusi sampling statistik sebagai simpangan baku statistik tersebut. Jadi simpangan baku rata-rata adalah simpangan baku distribusi sampling statistik untuk $\overline{x}$

TEOREMA 2.13.

a. Andaikan sampel acak berukuran n diambil dari suatu populasi dengan rata-rata μ maka nilai harapan dari $\overline{x}$ akan sama dengan μ ($\mu_{\overline{x}} = \mu$)

Bukti :

$$
\begin{aligned}
E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) &= E\left(\frac{X_1}{n} + \frac{X_2}{n} + \ldots + \frac{X_n}{n}\right) \\
&= E\left(\frac{X_1}{n}\right) + E\left(\frac{X_2}{n}\right) + \ldots + E\left(\frac{X_n}{n}\right) \\
&= \frac{1}{n}(E(X_1) + E(X_2) + \ldots E(X_n)) \\
&= \frac{1}{n}(\mu + \mu + \ldots + \mu) \\
&= \frac{1}{n}.n\mu = \mu
\end{aligned}
$$

b. Andaikan sampel acak berukuran n diambil dari suatu populasi dengan deviasi standar (simpangan baku) σ maka simpangan baku dari $\overline{x}$ akan mendekati (sama) simpangan baku populasi dibagi dengan akar kuadrat ukuran sampel

($\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$)

Bukti :

$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ sama saja dengan $\sigma_{\bar{x}}^2 = \dfrac{\sigma^2}{n} = \mathrm{Var}(\bar{x})$

$$
\begin{aligned}
\mathrm{Var}(\bar{x}) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \\
&= \mathrm{Var}\left(\frac{1}{n}(x_1 + x_2 + \ldots + x_n)\right) \\
&= \frac{1}{n^2}\mathrm{Var}(x_1 + x_2 + \ldots + x_n) \\
&= \frac{1}{n^2} n\,\mathrm{Var}(\sum_{i=1}^{n} x_i) = \frac{\mathrm{Var}(x_i)}{n} = \frac{\sigma^2}{n}
\end{aligned}
$$

Kalau populasi terbatas, tetapi pemilihan sampel dilakukan tanpa pengembalian maka $\mathrm{Var}(\bar{x}) = \sigma_{\bar{x}}^{2} = \dfrac{\sigma^2}{n} \cdot \dfrac{N-n}{N-1}$,

Faktor $\sqrt{\dfrac{N-n}{N-1}}$ disebut faktor koreksi populasi terbatas. Untuk N yang relatif besar dibandingkan dengan ukuran sampel n, faktor koreksi tersebut akan <u>mendekati satu</u>, sehingga nilai $\sigma_{\bar{x}}^{2}$ akan menghampiri σ^2/n.

TEOREMA 2.13.1. Limit Pusat

Bila sampel acak berukuran n ditarik dari suatu populasi dengan rata-rata μ dan variansi σ^2, maka rata-rata sampel $\bar{x}$ akan menyebar menghampiri distribusi normal. Berdasarkan teorema 2.13. a dan 2.1.3.b maka $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ merupakan sebuah nilai untuk variabel random normal standar Z.

Hampiran normal dalam dalil 2.13. digunakan bila $n \geq 30$, bagaimanapun bentuk popoulasinya. Bila $n < 30$, hampiran tersebut baik digunakan bila populasi asalnya tidak terlalu berbeda dari populasi normal. Bila populasi asalnya normal, maka $\bar{x}$ akan menyebar normal, sekecil apapun ukuran sampelnya.

CONTOH 2.13.

Ada populasi sebanyak 6 elemen, yaitu 6 orang karyawan suatu perusahaan kecil (N = 6)

X = upah harian (ribuan)

$X_1 = 1, X_2 = 2, X_3 = 3, X_4 = 4, X_5 = 5, X_6 = 6$

Hitung μ dan σ^2

Kemudian ambil sampel acak, sebanyak 2 karyawan (n = 2). Hitung seluruh kemungkinan sampel dan kemudian hitung $E(\bar{x})$ dan $\sigma^2_{\bar{x}}$ (pengambilan sampel dilakukan dengan pengembalian dan tanpa pengembalian).

Jawab: $\mu = \frac{1}{6}\sum X_i = \frac{21}{6} = 3{,}5$

$$\sigma^2 = \frac{1}{6}\sum (X_i - 3{,}5)^2 = \frac{17{,}5}{6}$$

a. Dengan pengembalian

Ada 36 sampel berelemen dua yang dapat diambil dari populasi , yaitu

$\{1,1\} : \bar{x}_1 = 1$	$\{1,2\} : \bar{x}_2 = 1{,}5$	$\{1,3\} : \bar{x}_3 = 2$	$\{1,4\} : \bar{x}_4 = 2{,}5$
$\{1,5\} : \bar{x}_5 = 3$	$\{1,6\} : \bar{x}_6 = 3{,}5$	$\{2,1\} : \bar{x}_7 = 1{,}5$	$\{2,2\} : \bar{x}_8 = 2$
$\{2,3\} : \bar{x}_9 = 2{,}5$	$\{2,4\} : \bar{x}_{10} = 3$	$\{2,5\} : \bar{x}_{11} = 3{,}5$	$\{2,6\} : \bar{x}_{12} = 4$

$\{3,1\} : \bar{x}_{13} = 2$	$\{3,2,\} : \bar{x}_{14} = 2,5$	$\{3,3\} : \bar{x}_{15} = 3$	$\{3,4\} : \bar{x}_{16} = 3,5$
$\{3,5\} : \bar{x}_{17} = 4$	$\{3,6\} : \bar{x}_{18} = 4,5$	$\{4,1\} : \bar{x}_{19} = 2,5$	$\{4,2\} : \bar{x}_{20} = 3$
$\{4,3\} : \bar{x}_{21} = 3,5$	$\{4,4\} : \bar{x}_{22} = 4$	$\{4,5\} : \bar{x}_{23} = 4,5$	$\{4,6\} : \bar{x}_{24} = 5$
$\{5,1\} : \bar{x}_{25} = 3$	$\{5,2\} : \bar{x}_{26} = 3,5$	$\{5,3\} : \bar{x}_{27} = 4$	$\{5,4\} : \bar{x}_{28} = 4,5$
$\{5,5\} : \bar{x}_{29} = 5$	$\{5,6\} : \bar{x}_{30} = 5,5$	$\{6,1\} : \bar{x}_{31} = 3,5$	$\{6,2\} : \bar{x}_{32} = 4$
$\{6,3\} : \bar{x}_{33} = 4,5$	$\{6,4\} : \bar{x}_{34} = 5$	$\{6,5\} : \bar{x}_{35} = 5,5$	$\{6,6\} : \bar{x}_{36} = 6$

$$E(\bar{x}) = \mu_{\bar{x}} = \frac{1}{36}\sum_{i=1}^{36}\bar{x}_i = \frac{126}{36} = 3,5$$

$$\sigma_{\bar{x}} = \sqrt{\frac{1}{36}\sum_{i=1}^{36}(\bar{x}_i - 3,5)^2} = \sqrt{\frac{52,5}{36}} = 1,21$$

Dari hasil (a) terlihat bahwa $E(\bar{x}) = \mu = 3,5$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{17,5}{6.2}} = 1,21$$

b. Tanpa Pengembalian

Ada 30 sampel berelemen dua yang dapat diambil dari populasi, yaitu :

$\{1,2\} : \bar{x}_1 = 1,5$	$\{1,3\} : \bar{x}_2 = 2$	$\{1,4\} : \bar{x}_3 = 2,5$	$\{1,5\} : \bar{x}_4 = 3$
$\{1,6\} : \bar{x}_5 = 3,5$	$\{2,1\} : \bar{x}_6 = 1,5$	$\{2,3\} : \bar{x}_7 = 2,5$	$\{2,4\} : \bar{x}_8 = 3$
$\{2,5\} : \bar{x}_9 = 3,5$	$\{2,6\} : \bar{x}_{10} = 4$	$\{3,1\} : \bar{x}_{11} = 2$	$\{3,2\} : \bar{x}_{12} = 2,5$
$\{3,4\} : \bar{x}_{13} = 3,5$	$\{3,5\} : \bar{x}_{14} = 4$	$\{3,6\} : \bar{x}_{15} = 4,5$	$\{4,1\} : \bar{x}_{16} = 2,5$
$\{4,2\} : \bar{x}_{17} = 3$	$\{4,3\} : \bar{x}_{18} = 3,5$	$\{4,5\} : \bar{x}_{19} = 4,5$	$\{4,6\} : \bar{x}_{20} = 5$
$\{5,1\} : \bar{x}_{21} = 3$	$\{5,2\} : \bar{x}_{22} = 3,5$	$\{5,3\} : \bar{x}_{23} = 4$	$\{5,4\} : \bar{x}_{24} = 4,5$
$\{5,6\} : \bar{x}_{25} = 5,5$	$\{6,1\} : \bar{x}_{26} = 3,5$	$\{6,2\} : \bar{x}_{27} = 4$	$\{6,3\} : \bar{x}_{28} = 4,5$
$\{6,4\} : \bar{x}_{29} = 5$	$\{6,5\} : \bar{x}_{30} = 5,5$		

$$E(\bar{x}) = \mu_{\bar{x}} = \frac{1}{30}105 = 3,5$$

$$\sigma_{\bar{x}}^2 = \frac{1}{30}\sum_{i=1}^{30}(\bar{x}_i - 3,5)^2 = \frac{35}{30} = 1,17$$

Disini terlihat bahwa $\mu_{\bar{x}} = \mu = 3,5$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\frac{N-n}{N-1} = \frac{17,5}{6.2}\frac{6-2}{6-1} = \frac{17,5}{12}\frac{4}{5} = 1,17$$

2.14. Penduga Parameter

Statistika banyak berhubungan dengan penarikan kesimpulan mengenai parameter populasi. Penarikan kesimpulan ini bersifat tidak pasti , karena hanya didasarkan pada hasil yang berasal dari sampel. Tingkat kebenaran penarikan kesimpulan diukur dengan seberapa besar peluang kesimpulan tersebut benar. Andaikan seorang guru ingin mengetahui berapa rata-rata lama belajar / perhari siswa SMU DIY . Peneliti dapat mengambil sampel sebesar 300 orang secara acak yang representatif dari seluruh SMU di DIY. Berdasarkan hasil perhitungan dari data sampel didapat rata-rata lama belajarnya adalah 2 jam/hari. Dari hasil tersebut dapat diduga bahwa rata-rata lama belajar siswa SMU DIY tiap hari adalah 2 jam. Berdasarkan penjelasan tentang statistik maka $\bar{Y} = 2$ jam adalah suatu nilai dugaan bagi μ. Contoh lain penduga bagi parameter adalah s^2 yang merupakan penduga bagi σ^2.

Pendugaan parameter merupakan usaha penentuan nilai parameter yang sedang diselidiki. Untuk melakukan pendugaan nilai suatu parameter dapat ditempuh dengan 2 cara. Cara pertama merupakan penentuan nilai tunggal yang mendekati nilai parameter itu dengan sebaik-baiknya (pendugaan titik). Cara kedua dapat merupakan penentuan suatu selang nilai dengan peluang yang besar mencakup nilai parameter yang diselidiki (pendugaan selang).

Secara umum dalam pembahasan selanjutnya, aplikasi teori penarikan sampel acak berlapis difokuskan pada pendugaan parameter θ, dimana θ adalah rata-rata populasi (μ), total populasi (τ), proporsi populasi (p). Distribusi sampling statistik yang bersesuaian dengan parameter-parameter tersebut menggunakan distribusi normal. Dengan demikian hal yang berkaitan dengan distribusi sampling akan selalu berkaitan dengand istribusi normal standar.

DEFINISI 2.15.1.1. Selang Kepercayaan

Selang kepercayaan (1 - α) bagi parameter populasi θ adalah suatu interval nilai $[\hat{\theta} - z_{\alpha/2}\sigma_{\hat{\theta}}, \hat{\theta} + z_{\alpha/2}\sigma_{\hat{\theta}}]$ sedemikian hingga $\theta \in [\hat{\theta} - z_{\alpha/2}\sigma_{\hat{\theta}}, \hat{\theta} + z_{\alpha/2}\sigma_{\hat{\theta}}]$ dan

$$P(\hat{\theta} - z\alpha/_2\, \sigma_{\hat{\theta}} < \theta < \hat{\theta} + z\alpha/_2\, \sigma_{\hat{\theta}}) = 1 - \alpha$$

dimana $\hat{\theta}$ adalah nilai yang dihitung dari sampel sebagai penduga tak bias.

Pernyataan taraf kepercayaan 95% (misalnya) mempunyai implikasi bahwa jika rencana penarikan sampel berukuran sama dengan teknik yang sama dilakukan berulang kali, misalnya 100 kali penarikan sampel, kemudian dari setiap sampel dibuat pernyataan tentang pendugaan selang, maka sekitar 95 kali dari selang nilai $(\hat{\theta} - z\alpha/_2\, \sigma_{\hat{\theta}}, \hat{\theta} + z\alpha/_2\, \sigma_{\hat{\theta}})$ mencakup parameter populasi akan benar, dan hanya sekitar 5 kali bahwa pernyataan diatas akan salah.

Data sampel yang diperoleh melalui penarikan sampel menghasilkan nilai statistik yang dapat dipakai sebagai penduga parameter. Nilai statistik tidak bisa tepat sama dengan nilai parameter populasi, tetapi dapat ditentukan sejauh mana ketepatan pendugaan tersebut. Langkah-langkah pendugaan parameter dapat digambarkan dalam diagram berikut :

Penarikan sampel $\rightarrow$ data dari penarikan sampel $\rightarrow$ statistik $\rightarrow$ penduga parameter.

Para peneliti, administrator dalam bidang pendidikan, atau pemerintahan semuanya berkepentingan dalam masalah pendugaan. Apakah itu berupa pendugaan keefektifan suatu obat baru, banyaknya siswa yang masuk perguruan tinggi dalam dekade mendatang, suatu kesimpulan statistik mengenai parameter populasi harus dibuat. Prosedur pendugaan nilai parameter populasi yang belum diketahui harus dibuat dari informasi yang dikandung oleh data sampel yang didasarkan pada distribusi sampling statistik yang telah dibahas pada pokok bahasan sebelumnya. Distribusi sampling statistik tersebut memungkinkan peneliti untuk mengaitkan suatu taraf kepercayaan tertentu dengan setiap kesimpulan statistik yang dibuat, sebagai suatu ukuran seberapa jauh peneliti menaruh kepercayaan pada ketepatan statistik dalam menduga parameter populasinya.

Selanjutnya akan didefinisikan penduga dan ciri penduga yang baik.

DEFINISI 2.14.1.2. Penduga

Suatu statistik T $(X_1,X_2,X_3,\ldots,X_n)$ dikatakan sebagai penduga titik dari $\underset{\sim}{\theta}$, $\underset{\sim}{\theta} \in \Theta$, jika T adalah suatu fungsi korespondensi satu-satu dari R_n ke Θ, yaitu

$$T : R_n \xrightarrow{\text{onto}} \Theta,$$

Dimana $\underset{\sim}{\theta}$ adalah vektor dari parameter -parameter yang tidak diketahui pada distribusi tertentu dan Θ adalah ruang parameter atau himpunan dari semua nilai $\underset{\sim}{\theta}$ yang mungkin. T disebut penduga parameter.

2.14.2. Ciri Penduga Yang Baik

Besarnya parameter yang tidak diketahui menyebabkan diperlukannya metode pendugaan parameter yang menjamin bahwa hasil pendugaan harus sedekat mungkin dengan parameter yang diduga. Dengan kata lain, penduga yang dihasilkan haruslah dengan "baik" mendekati nilai parameter sebenarnya.

Suatu parameter yang tidak diketahui dimungkinkan mempunyai lebih dari satu penduga titik. Jika dihadapkan pada dua pilihan penduga tertentu, maka diperlukan ciri yang menjadi acuan untuk menentukan penduga mana yang lebih baik. Adapun ciri-ciri penduga yang baik adalah :

1. Tidak bias
2. Konsisten
3. Mempunyai variansi minimum

DEFINISI 2.14.2.1. Penduga Tak Bias

Suatu statistik $\hat{\theta}$ disebut sebagai penduga tak bias dari parameter θ, jika

$E(\hat{\theta}) = \theta$

Penduga tak bias menyatakan bahwa bila dilakukan pengambilan sampel berukuran n secara berulang-ulang dan setiap kali dihitung penduganya maka yang diharapkan adalah rata-rata penduga tersebut sama dengan parameter yang diduga. Sebaliknya $\hat{\theta}$ merupakan penduga yang berbias, jika $E(\hat{\theta}) \neq \theta$ dan besarnya bias tersebut adalah $E(\hat{\theta}) - \theta$.

DEFINISI 2.14.2.2. Penduga Konsisten

$\hat{\theta}$ merupakan penduga konsisten apabila nilai dugaan dari $\hat{\theta}$ mendekati θ bila samplenya diperbesar sampai tak hingga (∞)

DEFINISI 2.14.2.3.

$\hat{\theta}$ merupakan penduga terbaik atau penduga yang mempunyai variansi minimum, jika memenuhi syarat-syarat berikut :

a) $E(\hat{\theta}) = \theta$

b) $\sigma^2_{\hat{\theta}}$ minimum, artinya apabila dibandingkan dengan penduga θ lainnya, maka θ mempunyai variansi terkecil.

CONTOH 2.14.

Misalkan $X_1, X_2, \ldots X_n$ merupakan sampel acak dari suatu variabel random dengan rata-rata μ dan variansinya σ^2. Telah diketahui bahwa $\bar{x}$ adalah penduga tak bias untuk μ dan variansinya $\frac{\sigma^2}{n}$. Berdasarkan teorema Chebychev diketahui

bahwa $P(|\theta - \hat{\theta}| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} E[(\theta - \hat{\theta})^2]$

$= \frac{1}{\varepsilon^2} MSE(\theta)$, MSE merupakan rata-rata simpangan kuadrat.

Jadi $\lim\limits_{n \to \infty} P(|\mu - \bar{x}| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \frac{\sigma^2}{n}$, maka MSE $(\bar{x}) = \frac{\sigma^2}{n} \to 0$ untuk $n \to \infty$.

$$\lim_{n \to \infty} P(|\mu - \bar{x}| \geq \varepsilon) = 0$$

$\bar{x}$ merupakan penduga konsisten untuk μ.

Dua macam penduga dapat dibandingkan atas dasar efisiensi relatif, misalnya: θ_1 dan θ_2 masing-masing merupakan penduga bagi θ, maka efisiensi relatif dari θ_2 terhadap θ_1 adalah : $R(\hat{\theta}_2, \hat{\theta}_1) = \dfrac{E(\hat{\theta}_1 - \theta)^2}{E(\hat{\theta}_2 - \theta)^2}$

Jika R>1 maka sacara relatif θ_2 lebih efisien daripada θ_1; jika R=1 maka kedua penduga θ_1 dan θ_2 mempunyai tingkat efisiensi yang sama, sedangkan jika R<1 maka secara relatif θ_1 lebih efisien daripada θ_2.

2.15. Kesalahan Sampling

Penarikan sampel bermaksud untuk melakukan pendugaan parameter populasi berdasarkan nilai statistik sampel. Timbul masalah dalam pendugaan tersebut yaitu berapa besar kesalahan yang dibuat sebagai akibat menarik kesimpulan tentang sifat populasi berdasarkan sampel yang dipelajari. Persoalan tersebut dalam statistika disebut kesalahan sampling. Setelah mengetahui besar kesalahan sampling, peneliti dapat membuat suatu selang kepercayaan dengan tingkat kepercayaan tertentu untuk meyakini suatu pernyataan diterima atau tidak.

DEFINISI 2.15.1. Kesalahan Sampling

Jika $\hat{\theta}$ merupakan penduga tak bias bagi parameter θ, maka kesalahan sampling didefinisikan sebagai jarak antara $\hat{\theta}$ dan θ atau penyimpangan mutlak dari $\hat{\theta}$ dengan θ, yang dinotasikan sebagai KS = $|\hat{\theta} - \theta|$. KS dan galat (G) pada dasarnya mempunyai definisi yang sama. Galat merupakan penyimpangan pendugaan yang telah ditetapkan oleh peneliti sebelum melakukan penelitian, sedangkan KS merupakan kesalahan pendugaan berdasarkan data yang didapat dari penelitian , dan diharapkan KS lebih kecil dari galat pendugaan. Bila (1- α) = tingkat keyakinan, maka perhatikan kurva normal berikut :

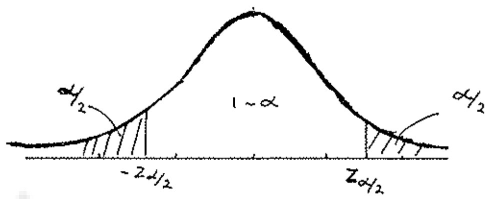
$$P(-z_{\alpha/2} \leq z \leq z_{\alpha/2}) = 1 - \alpha$$

$$-z_{\alpha/2} \leq \frac{\hat{\theta} - \theta}{\sigma\hat{\theta}} \leq z_{\alpha/2}$$

$-z_{\alpha/2}\ \sigma_{\hat\theta} < \hat\theta - \theta < z_{\alpha/2}\ \sigma_{\hat\theta}$

$-z_{\alpha/2}\ \sigma_{\hat\theta} < KS < z_{\alpha/2}\ \sigma_{\hat\theta}$

$|KS| < z_{\alpha/2}\ \sigma_{\hat\theta}$



Jadi besarnya kesalahan sampling sebesar $\pm\, z_{\alpha/2}\, \sigma_\theta$

Nilai $z_{\alpha/2}$ diperoleh dari tabel normal. Sedangkan $\sigma_{\hat\theta}$ sering disebut sebagai galat pendugaan $\hat\theta$

CONTOH 2.15. Kesalahan sampling

Misalnya ada 1000 perusahaan di Jateng. Dari 1000 perusahaan diambil 625 perusahaan secara acak. Rata-rata pendugaan modal perusahaan ($\overline{X}$) yang dihitung berdasarkan penelitian sampel Rp 150 juta dengan simpangan baku sebesar Rp 30 juta. Untuk pendugaan μ digunakan tingkat kepercayaan sebesar 95%. Berikut akan ditentukan besar kesalahan sampling dan selang kepercayaannya

Jawab : $\overline{X} = 150$ $\quad \hat\sigma = 30$ $\quad N = 1000$ $\quad n = 625$

$(1-\alpha) = 0{,}95$ $\qquad \alpha = 1 - 0{,}95 = 0{,}05$ $\quad z_{\alpha/2} = z_{0{,}025} = 1{,}96$

$KS = \pm\, z_{\alpha/2}\, \sigma_X = \pm\, 1{,}96\ (30/\sqrt{625}) = \pm\, 1{,}96\ (1{,}2) = \pm\, 2{,}352$

$P(\hat\theta - z_{\alpha/2}\,\sigma_{\hat\theta} < \theta < \hat\theta + z_{\alpha/2}\,\sigma_{\hat\theta}) = 1 - \alpha$

$P(\overline{X} - z_{0{,}025}\,\sigma_x \le \mu \le \overline{X} + z_{0{,}025}\,\sigma_x) = 1 - 0{,}05$

$P(150 - 2{,}352 \le \mu \le 150 + 2{,}352) = 0{,}95$

$P(\ 147{,}648 \le \mu \le 152{,}352\ ) = 0{,}95$

Dari hasil diatas dapat dikatakan selang nilai (147,65 ; 152,35) mencakup parameter μ dengan tingkat kepercayaan sebesar 95%.

BAB III

PENARIKAN SAMPEL ACAK BERLAPIS

Dalam bab ini akan dibahas Sampel Acak Sederhana, alasan perlunya stratifikasi, arti Penarikan Sampel Acak Berlapis dan rumus-rumus yang berhubungan dengan Penarikan Sampel Acak Berlapis.

3.1. Pengambilan Sampel Secara Acak

Contoh berikut akan dipakai untuk menjelaskan pokok bahasan pengambilan sampel secara acak. Seorang ibu ingin mengetahui apakah sepanci gulai sudah cukup bumbunya, atau belum. Ibu tersebut hanya mengambil satu sendok untuk dicicipi, dengan terlebih dahulu mengaduk gulai tersebut. Dalam contoh tersebut "populasi" yang diselidiki adalah sepanci gulai sedangkan "sampel" adalah gulai satu sendok yang terambil. Pengadukan gulai sebelum sampel diambil dimaksudkan agar bumbu gulai tersebar secara merata. Dengan demikian satu sendok gulai yang terambil akan mewakili dengan baik seluruh isi panci.

Apabila suatu benda dipilih secara sembarang dari sekumpulan benda yang tersedia dan kesempatan atau peluang setiap benda dari kumpulan benda tersebut untuk terpilih sama besar, maka dapat dikatakan bahwa benda tersebut telah terpilih secara acak. Sampel yang terambil perlu dipilih secara acak karena untuk mengatasi berbagai penyimpangan yang bisa terjadi. Berdasarkan contoh di atas jika gulai tidak di aduk terlebih dahulu maka ibu tersebut mungkin hanya mengambil bagian atas dari gulai dan mencicipinya. Bumbu gulai mungkin saja akan mengendap didasar panci dan tidak terambil oleh ibu, sehingga ibu tersebut akan merasa hambar, dan demikian pula sebaliknya.

DEFINISI 3.1.1. Sampel Acak Sederhana (SAS)

Suatu sampel dikatakan telah terpilih secara acak (sederhana) dari populasi apabila setiap anggota populasi mempunyai kesempatan yang sama besar untuk dijadikan anggota sampel. Karena anggota sampel tersebut dipilih secara acak, maka sampel tersebut juga disebut sampel acak.

Pada umumnya, persoalan yang dihadapi adalah bagaimana cara mengambil sampel dari suatu populasi agar diperoleh keterangan yang sebaik-baiknya sesuai dengan biaya dan waktu yang tersedia. Salah satu cara mengambil sampel adalah pengambilan sampel acak Berikut akan didefinisikan penarikan sampel acak.

DEFINISI 3.1.2. **Penarikan Sampel Acak** adalah metode pengambilan sampel dari suatu populasi berukuran N sedemikian rupa sehingga semua sampel yang mungkin terambil memiliki probabilitas yang sama untuk terpilih.

Suatu sampel yang ditarik secara acak tidaklah mengandung bias, dalam arti bahwa tidak satu sampelpun mempunyai peluang lebih besar untuk terpilih dibandingkan dengan sembarang sampel lainnya. Akan tetapi jika sampel tersebut tidak terambil secara acak, ada faktor yang tidak diketahui yang mungkin membuat peneliti condong memilih suatu sampel yang mengandung bias.

CONTOH 3.1.2. Penarikan Sampel Acak

1. Andaikan populasi yang hendak diteliti adalah semua siswa kelas 4 di sekolah X. Andaikan terdapat 200 siswa, jika seorang anak diambil secara acak dari

populasi tersebut maka peluang terambilnya anak tersebut adalah $^1/_{200}$. Jika siswa yang terpilih pertama tidak dikembalikan, peluang siswa lain untuk terambil adalah $^1/_{199}$, prosedur tersebut disebut penarikan sampel tanpa pengembalian. Jika anak pada pengambilan sampel yang pertama dikembalikan kedalam populasi, maka peluang siswa lain untuk terambil adalah $^1/_{200}$, prosedurnya penarikan sampel tersebut disebut penarikan sampel dengan pengembalian.

2. Misalkan ada suatu populasi berukuran N = 5, yang terdiri dari unit-unit A,B,C,D,E. Jika peneliti menginginkan sampel berukuran n = 2 (pengambilan sampel dilakukan tanpa pengembalian) , maka dari populasi tersebut terdapat sampel sebanyak 20 sampel, yaitu :
   AB,AC,AD,AE,BA,BC,BD,BE,CA,CB,CD,CE,DA,DB,DC,DE,EA,EB,EC, ED. Setiap sampel akan mempunyai peluang yang sama untuk terpilih sebagai sampel dalam penelitian, yaitu $^1/_{20}$. Misalkan pengambilan sampel dilakukan dengan pengembalian maka akan terdapat 25 sampel, yaitu: AA,AB,AC,AD,AE,BA,BB,BC,BD,BE,CA,CB,CC,CD,CE,DA,DB,DC,DD, DE,EA,EB,EC,ED,EE. Dapat dibuat ketentuan umum bahwa dari suatu populasi berukuran N yang ditarik suatu sampel berukuran n (pengambilan sampel dilakukan dengan dan tanpa pengembalian), maka nilai peluang bagi sampel terpilih adalah 1 dibagi semua kemungkinan sampel yang bisa terambil.

3.2. Sampel yang Representatif

Sebelum diuraikan macam dan prosedur penarikan sampel akan dibahas dulu hakekat kedudukan sampel baik terhadap populasi maupun terhadap kegiatan penelitian secara keseluruhan. Andaikan peneliti dapat melakukan observasi pada semua subjek dalam penelitian yang dilakukan, sampel tidak diperlukan lagi. Namun karena keterbatasan waktu, dana, dan kemampuan maka seorang peneliti tidak mungkin meneliti satu persatu elemen dalam populasi tersebut. Satu masalah yang muncul dalam penelitian adalah apakah data/sampel yang terambil tersebut betul-betul mewakili populasi dari mana sampel diambil. Dalam penelitian, peneliti mengambil sampel acak dengan harapan dan mengasumsikan sampel tersebut representatif, bahwa karakteristik populasi tersebut akan terdapat pula dalam sampel yang terambil. Ada beberapa hal yang mempengaruhi representativitas sampel, yaitu :

(1) homogenitas populasi

(2) jumlah (besar) sampel yang dipilih

(3) banyaknya karakteristik subjek yang akan dipelajari dan

(4) teknik penarikan sampel yang memadai.

Makin homogen distribusi atau keadaan karakter subjek dalam suatu populasi maka makin mudah dicapai representativitas sampel, karena dimanapun sampel tersebut diambil akan mendapatkan hasil yang sama. Sebagai contoh misalnya, distribusi eritrosit dalam darah sedemikian homogen, sehingga dari tiap tetes darah yang diambil dari bagian tubuh manapun akan diperoleh sifat yang sama. Sebaliknya tempat tinggal penduduk kaya dan miskin di suatu tempat tidak berdistribusi secara merata, maka pemilihan sampel pada tiap bagian daerah tidak akan menggambarkan distribusi kaya-miskin yang sama.

Makin banyak subjek yang dijadikan sampel makin tinggi tingkat representativitasnya. Hal tersebut karena makin banyak subjek yang dipilih, berarti makin besar proporsi sampel terhadap populasi, sehingga makin 'dekat' karakteristik subjek sampel dengan karakteristik populasi.

Makin banyak karakteristik subjek yang dipelajari, yang secara praktis berarti makin banyak variabel yang akan diteliti, mengakibatkan keadaan populasi makin kurang homogen sebab masing-masing variabel mempunyai distribusinya sendiri dalam subjek populasi. Dengan demikian, banyaknya karakteristik akan menurunkan tingkat representativitas sampel. Untuk mengatasi hal tersebut maka dibutuhkan jumlah sampel yang lebih besar dan teknik pemilihan sampel yang memadai. Teknik pemilihan sampel yang memadai adalah teknik pemilihan subjek-subjek penelitian yang sesuai dengan keadaan populasi. Berikut akan didefinisikan sampel yang representatif.

DEFINISI 3.2. **Sampel yang Representatif** adalah sampel yang memiliki karakteristik populasi yang relevan dengan tujuan penelitian yang bersangkutan.

CONTOH 3.2.

Misalkan seorang peneliti ingin mengetahui total pengeluaran penduduk DIY untuk bulan Januari 1998. Karakteristik populasi yang relevan dengan tujuan penelitian adalah masyarakat tingkat bawah, masyarakat tingkat sedang, masyarakat tingkat atas. Jadi dalam pengambilan sampel acak, peneliti berharap bahwa sampel yang terambil memuat atau mengandung karakteristik populasi tersebut.

3.3. Cara Pemilihan Elemen Anggota Sampel pada SAS

Terdapat cara-cara memilih elemen anggota sampel agar sampel yang dihasilkan bersifat acak, sebagai berikut :

3.3.1. *Pemilihan dengan cara undian*

Setiap kali seorang peneliti akan melakukan penelitian, peneliti harus menentukan sampel dengan jelas. Misalkan seorang peneliti akan menilai khasiat dua jenis zat penenang. Karena percobaan kepada manusia terkait dengan 'etika' dan konsekuensi-konsekuensi yang mungkin timbul (adanya efek samping) maka percobaan dilakukan terlebih dahulu pada unit sampel 'bukan manusia'. Peneliti memutuskan untuk mencobakan kedua macam zat penenang tersebut kepada 15 ikan mujair. Setiap ikan dari lima belas ekor yang tersedia harus memiliki kesempatan yang sama untuk dapat ditempatkan dalam akuarium yang mana saja dari ketiga akuarium yang tersedia.

Langkah-langkah yang dilakukan oleh peneliti adalah setiap ikan diberi kalung dan dinomori dari satu sampai lima belas. Setelah itu dengan cara mengundi, dipilih lima nomor pertama dan lima nomor kedua. Sisanya langsung menjadi kelompok lima ekor lainnya. Setelah tiga kelompok tersebut ditentukan, maka menggunakan lima gulung kertas yang masing-masing memuat angka 1 hingga 5, diundilah kelompok mana yang menerima salah satu dari tiga perlakuan yang sudah ditentukan oleh peneliti. Demikian seterusnya sampai tiga kelompok tesebut dikenai perlakuan yang berbeda-beda.

Contoh diatas merupakan penjelasan dari cara pengambilan sampel dengan undian. Langkah-langkah pengambilan sampel dengan cara undian adalah sebagai berikut:

a) Mendaftarkan semua unit yang ada dan diberi nomor urut 1,2,...,N. N adalah ukuran populasi .Membentuk kerangka penarikan sampel.
b) Membuat gulungan kertas yang berukuran sama sebanyak N buah dengan masing-masing bertuliskan nomor unit tersebut.
c) Memasukkan N buah gulungan kertas tersebut kedalam kotak dan kemudian dikocok.
d) Menarik atau mengambil gulungan kertas sebanyak ukuran sampel n yang telah ditentukan. Dengan cara menutup mata orang yang mengambil gulungan kertas tersebut.

Berdasarkan cara tersebut, maka akan terpilih n buah unit dari populasi berukuran N untuk membentuk sampel acak sederhana berukuran n.

3.3.2. *Cara Pengambilan Sampel Menggunakan Tabel Angka Acak*

Tabel angka acak merupakan gambaran tentang sampel acak. Tabel angka acak tersebut memuat angka-angka yang dihasilkan oleh mesin, sedemikian rupa sehingga tidak tampak tatanan atau urutan atau sistem dalam angka-angka yang terpapar tersebut. Berikut ini akan dijelaskan bagaimana mengambil sampel secara acak dengan menggunakan tabel angka acak.

Cara ini digunakan untuk populasi yang besar. Misalkan dari suatu populasi berukuran N=1000 akan ditarik sampel berukuran n=30. Langkah-langkahnya sebagai berikut :

a. Membentuk kerangka penarikan sampel dengan jalan mendaftar semua unit yang ada dalam populasi. Unit-unit tersebut diberi nomor urut :001,002,...,999,000. Penentuan jumlah digit disesuaikan dengan ukuran populasi, sebagai berikut :

- Jika N di antara 1-10, menggunakan 1 digit.

- Jika N di antara 11-100, menggunakan 2 digit

- Jika N di antara 101-1000, menggunakan 3 digit.

-dan seterusnya.

b. Tentukan secara sembarang dengan menggunakan ujung pensil atau lainnya pada tabel angka acak. Pembacaan dapat dilakukan secara vertikal (kebawah) atau horisontal (ke samping) secara berurutan dengan mengambil angka acak sesuai jumlah digit yang telah ditentukan.

c. Misalkan penunjukan pertama kali adalah pada angka yang terdapat dalam baris 30 dan kolom 35. Tabel angka acak yang diberikan berukuran 10000, terdiri dari 100 baris dan 100 kolom. Misalkan pembacaan dilakukan secara vertikal (ke bawah), maka 30 angka acak berukuran 3 digit yang diperoleh dapat dilihat dalam tabel pada lampiran 1

Dalam kasus di atas, urutan pembacaan dilakukan secara setiap baris pada posisi kolom yang sama, dimulai dari baris 30,31,...,59. Jika ukuran sampel n besar, maka pembacaan dapat dilanjutkan terus hingga baris 99 pada posisi kolom yang sama, kemudian dilanjutkan terus, dimulai lagi dari baris 00 dengan posisi kolom yang berubah ke samping kanan bertambah sesuai jumlah digit yang ditentukan, misalnya kalau bertambah tiga digit maka menjadi kolom 38-40, dan seterusnya.

Dalam pembacaan angka acak, apabila ada angka acak yang muncul lebih dari satu kali, maka angka acak yang muncul kemudian diabaikan serta pembacaan dilanjutkan terus sampai terpilih n buah angka acak berukuran jumlah digit tertentu.

3.4. Kelemahan Penarikan Sampel Acak Sederhana (SAS)

Suatu sampel berukuran n yang ditarik secara acak dari suatu populasi berukuran N dinamakan sampel acak sederhana. Sampel acak sederhana adalah suatu contoh pengambilan sampel yang termudah dari suatu populasi. Penarikannya didasari anggapan bahwa semua anggota populasi misalkan para pengendara sepeda motor, diperkirakan mempunyai sikap yang sama terhadap pertanyaan penelitian, misalnya persepsi terhadap kegunaan helm pengaman. Kalaupun ada perbedaan sikap, maka perbedaan sikap tersebut hanya disebabkan karena kebetulan saja, dan bukan karena adanya sikap yang sangat berbeda.

Permasalahan menjadi berbeda kalau sejak semula diperkirakan bahwa sebagian pengendara sepeda motor didalam populasi tersebut berbeda pandangan mengenai sikap mereka terhadap penggunaan helm. Diperkirakan misalnya bahwa laki-laki lebih mudah menerima peraturan tersebut, sedangkan bagi kaum wanita yang di dalam suasana pekerjaannya harus tampak rapi, helm tersebut akan merusak penampilan rambutnya. Maka tanggapan pengendara sepeda motor mungkin sekali juga akan bergantung pada jenis kelaminnya. Oleh karena itu, kalau seorang peneliti akan menarik suatu sampel acak, peneliti sebaiknya menariknya tidak langsung dari seluruh populasi tetapi peneliti terlebih dahulu membagi populasi menjadi beberapa lapisan. Dalam contoh di atas populasi dibagi dalam dua lapisan yaitu laki-laki dan wanita.

Secara umum, jika peneliti tidak membagi menjadi beberapa lapisan maka akan diperoleh variansi yang besar, karena adanya variasi. Timbul pertanyaan bagaimana variasi data yang ada dalam populasi tersebut, dan berkaitan dengan pertanyaan berapa besar variansi populasi. Peneliti memang tidak dapat menentukan berapa besar variansi populasi karena parameter tidak diketahui, namun pendugaan tentang sejauh mana variasi sifat populasi seyogyanya dapat diantisipasi oleh peneliti dengan memperhatikan beberapa fenomena yang ada. Fenomena tersebut mungkin dapat dilihat melalui pola kehidupan sehari-hari yang ada dalam masyarakat populasi tersebut, mungkin dapat dilihat melalui bangunan rumah tempat tinggal yang ada atau mungkin dapat diperkirakan dari jenis pekerjaan yang ditekuni sehari-hari, dan sebagainya. Berdasarkan fenomena yang ada serta diketahui berkaitan erat dengan sifat populasi yang akan dipelajari, peneliti dapat memperkirakan variasi sifat populasi tersebut sehingga dapat menentukan teknik penarikan sampel tertentu sesuai dengan keadaaan yang dihadapi. Suatu cara yang langsung dapat memperkecil variansi adalah dengan memperbesar ukuran sampel. Tetapi memperbesar ukuran sampel biasanya akan memperbesar biaya. Maka salah satu pemecahannya adalah dengan membagi populasi menjadi beberapa lapisan.

Dari pembahasan diatas, kelemahan dari penarikan sampel acak sederhana adalah kurangnya keakuratan data yang terkumpul. Bila populasi heterogen, peneliti tidak dapat menduga besar variansi populasi dengan baik, sehingga diperlukan penduga variansi dengan membagi populasi kedalam subpopulasi (lapisan) yang relatif homogen berdasarkan kriteria tertentu.

Untuk melihat kelemahan penarikan SAS dapat ditempuh dengan melihat pengaruhnya pada pendugaan selang kepercayaan, yang akan dibahas pada pokok bahasan ketelitian relatif penarikan SAS dan penarikan SAB (4.4.). Bila populasi tidak homogen maka variansi besar dan kesalahan sampling besar, sehingga selang kepercayaan menjadi lebar yang berarti akurasi pendugaan parameter kecil.

Ada usaha lain yang dapat dilakukan agar hasil pendugaan mendekati parameter populasi, yaitu dengan memperbanyak data (ukuran sampel diperbesar). Jika n besar maka kesalahan sampling akan semakin kecil sehingga selang kepercayaan sempit dan hasil akan semakin mendekati parameter populasi. Akan tetapi usaha ini akan mempermahal ongkos penelitian.

3.5. Penarikan Sampel Acak Berlapis (SAB)

Karena adanya kelemahan pada penarikan SAS maka diperlukan penarikan sampel yang lain yang dapat mengatasi kelemahan tersebut. Dalam penarikan SAB, populasi yang ukurannya besar dan heterogen dibagi dalam beberapa lapisan yang merupakan subpopulasi yang homogen. Subpopulasi yang homogen akan memperkecil besarnya variansi populasi, sehingga peneliti dapat menentukan nilai dugaan parameter populasi, dan nilai dugaan lainnya yang berkaitan dengan populasi dengan lebih baik. Dengan demikian penentuan lapisan merupakan langkah pertama yang harus dikerjakan dalam menggunakan teknik penarikan

SAB. Penentuan lapisan adalah merupakan suatu hal yang menentukan berhasil tidaknya peneliti memperkecil variansi nilai rata-rata sampel yang dengan demikian akan meningkatkan keandalan pendugaan.

DEFINISI 3.5. Penarikan Sampel Acak Berlapis

Penarikan Sampel Acak Berlapis adalah suatu teknik penarikan sampel melalui pemisahan unit-unit populasi ke dalam kelompok-kelompok yang saling asing dan relatif homogen, yang disebut strata atau lapisan-lapisan. Dari setiap lapisan dipilih sampel acak sederhana. Rancangan tersebut dilakukan pada populasi yang heterogenitasnya diwarnai oleh adanya beberapa kelompok atau lapisan subjek, dengan batas yang jelas antar kelompok tersebut.

3.6. Kelebihan Penarikan Sampel Acak Berlapis

Dalam penarikan sampel acak berlapis, data menjadi lebih homogen, karena sudah dibagi dalam beberapa lapisan. Berdasarkan data yang homogen tersebut, peneliti dapat menentukan pendugaan parameter untuk tiap lapisan dengan lebih akurat. Ada beberapa kelebihan penarikan Sampel Acak Berlapis.

1. Data akan menjadi lebih homogen pada setiap lapisan dibandingkan dengan data pada populasi secara keseluruhan.
2. Biaya untuk pelaksanaan SAB lebih murah daripada SAS karena alasan administrasi.
3. Jika SAB digunakan, maka pendugaan parameter populasi secara terpisah dapat diperoleh untuk setiap lapisan tanpa perlu melakukan penarikan sampel tambahan.

Akibat berkurangnya variasi dalam setiap lapisan, pendugaan berdasarkan SAB mempunyai variansi yang lebih kecil dibandingkan dengan pendugaan yang sama berdasarkan SAS, pada ukuran sampel (n) yang sama.

Setiap lapisan, biasanya dalam penelitian sosial ekonomi, akan memiliki elemen yang lebih sedikit dan mencakup daerah geografis yang lebih kecil dibandingkan populasi secara keseluruhan. Oleh karena itu akan lebih mudah untuk memilih sampel dan mengumpulkan data dalam lapisan yang lebih kecil. Peneliti juga dapat membagi tim peneliti untuk dapat bekerja dalam setiap lapisan dan dengan demikian pelaksanaan penelitian akan menjadi lebih cepat selesai. Hal tersebut berhubungan dengan biaya penelitian yang dikeluarkan dan pengorganisasian data penelitian.

Ada tiga persyaratan yang harus dipenuhi dalam melakukan penarikan SAB. Syarat tersebut adalah :

1. Harus ada kriteria yang akan dipergunakan sebagai dasar untuk membagi populasi tersebut ke dalam lapisan-lapisan. Dengan kata lain harus ditentukan secara jelas dan tegas kriteria yang dipergunakan sebagai dasar pelapisan. Yang dapat dijadikan kriteria untuk pembagian populasi adalah variabel yang akan dipelajari atau jika hal tersebut tidak memungkinkan maka dapat menggunakan variabel-variabel lain yang menurut pengetahuan peneliti mempunyai korelasi yang kuat dengan variabel yang akan dipelajari tersebut. Sebagai contoh dikemukakan seorang peneliti yang ingin mempelajari tingkat produksi harian dikawasan industri tertentu sangat tidak seragam karena dikawasan tersebut terdiri dari industri-industri berskala kecil, sedang, dan

besar. Peneliti dapat memutuskan variabel terbaik yang dijadikan sebagai dasar pelapisan adalah tingkat produksi harian itu sendiri, tetapi seperti diketahui bahwa sifat dari variabel tersebut tidak mungkin diketahui sebelum melakukan penelitian. Sebagai pengganti variabel tingkat produksi harian, peneliti dapat menggunakan variabel jumlah tenaga kerja yang ada di setiap perusahaan yang juga mencerminkan skala dari industri tersebut. Dalam hal tersebut peneliti berkeyakinan bahwa variabel jumlah tenaga kerja berkorelasi erat dengan variabel tingkat produksi harian, dengan demikian variabel jumlah tenaga kerja dapat dijadikan sebagai dasar untuk pelapisan. Dalam hal ini, peneliti dapat mendefinisikan skala industri berdasarkan jumlah tenaga kerja yang ada di industri tersebut. Dengan demikian peneliti dapat membagi populasi industri ke dalam industri berskala kecil, sedang, dan besar. Untuk kasus diatas, variabel yang dijadikan sebagai dasar pelapisan adalah jumlah tenaga kerja yang ada di industri tersebut yang mampu mencerminkan skala industri. Berdasarkan pengelompokan industri-industri ke dalam industri berskala kecil, sedang, dan besar. Berdasarkan tingkat penggunaan tenaga kerja, maka diharapkan variasi produksi harian dalam setiap lapisan akan menjadi lebih homogen.

2. Harus ada data pendahuluan dari populasi mengenai kriteria yang dipergunakan sebagai dasar pelapisan. Untuk kasus di atas harus ada data yang lengkap dari populasi tentang jumlah tenaga kerja yang ada dalam setiap industri dan dengan demikian setiap industri dapat dikelompokan secara tegas apakah termasuk industri berskala kecil, sedang, atau besar.

3. Harus diketahui dengan tepat jumlah unit-unit dari setiap lapisan dalam populasi tersebut. Dalam hal ini harus diketahui secara tepat jumlah perusahaan industri berskala kecil, sedang, dan besar. Dengan demikian untuk kasus di atas, harus diketahui ukuran dari setiap lapisan, yaitu : N_1,N_2, dan N_3; dimana $N_1+N_2+N_3= N$. Jadi jumlahan dari ukuran setiap lapisan dalam populasi harus sama dengan ukuran populasi tersebut. Peneliti kemudian menetapkan kerangka penarikan sampel. Pada kasus di atas yang dijadikan unit penarikan sampel adalah perusahaan industri.

Berdasarkan uraian di atas, maka secara singkat dapat dibuat langkah-langkah umum prosedur penarikan SAB sebagi berikut :

a. Bagilah populasi kedalam lapisan-lapisan berdasarkan kriteria yang dijadikan sebagai dasar pelapisan dengan menciptakan kondisi sedemikian rupa agar variasi variabel yang dipelajari dalam lapisan menjadi relatif homogen dan akan berlaku bahwa variansi dalam lapisan menjadi lebih kecil daripada variansi antar lapisan.

b. Daftarkan secara jelas unit-unit dalam setiap lapisan kemudian diberi nomor urut. Sebagai contoh untuk lapisan ke satu diberi nomor urut dari 001,002,…,N_1; demikian pula untuk lapisan kedua diberi nomor urut dari 001,002,…N_2, dan seterusnya. Nomor urut diberikan dengan jumlah digit yang sesuai dengan ukuran dari setiap lapisan.

c. Lakukan penarikan SAS dengan menggunakan undian atau tabel angka acak untuk setiap unit dalam lapisan. Unit-unit yang terpilih secara acak sederhana dari setiap lapisan akan membentuk SAB. Dengan demikian terlihat bahwa teknik penarikan SAB adalah menerapkan teknik penarikan SAS dalam setiap lapisan.

d. Unit-unit yang terpilih secara acak sederhana dalam setiap lapisan akan menjadi anggota sampel yang seterusnya dipelajari sifat-sifat yang sesuai dengan maksud penelitian tersebut.

CONTOH 3.5.1.

Seorang guru olah raga ingin mengetahui rata-rata waktu lari 100 m, murid kelas 3 SMU di DIY. Guru tersebut mengetahui bahwa kemampuan tiap siswa berlainan, maka guru OR membagi mereka dalam 2 lapisan yang relatif homogen yaitu laki-laki dan perempuan. Misalkan diketahui ukuran stratum 1 dan 2 adalah N_1 = 1100 siswa laki-laki kelas 3 SMU DIY, N_2 = 900 siswa perempuan kelas 3 SMU DIY. Jadi populasi berukuran $N=N_1+N_2$ = 2000 siswa. Menggunakan SAB ditentukan ukuran sampel tiap lapisan, n_1 = 500 siswa laki-laki kelas 3 dari 10 sekolah di DIY, n_2 = 350 siswa perempuan kelas 3 dari 10 sekolah di DIY. Kecepatan rata-rata lari siswa dari sampel didapat (dalam satuan menit) .

Lapisan 1	1,5 2 2,5 1 1,75 2,33 2,66 3 2,17 2 2,5
Lapisan 2	2 3 3,5 3,17 2,75 3,75 4 3,42 3,17 3,5

Dari hasil tersebut diperoleh $\overline{x}_1$ = 2,042 dan $\overline{x}_2$ =3,225 yang dipakai untuk menduga rata-rata masing-masing lapisan

Jadi untuk waktu lari 100 m, murid kelas 3 SMU DIY untuk lapisan 1 (siswa laki-laki) adalah 2,042 dan untuk lapisan 2 (siswa perempuan) adalah 3,225.

CONTOH 3.5.2.

Peneliti ingin mengetahui kegiatan apakah yang dilakukan anak SD pada waktu liburan sekolah di kota Jakarta.. Peneliti membagi populasi menjadi 3 lapisan yaitu anak dari masyarakat kelas bawah, anak dari masyarakat kelas menengah dan anak dari masyarakat kelas atas. Langkah yang dilakukan adalah wawancara di perumahan kumuh, perumahan biasa, perumahan elite, dengan mengambil sampel 10 keluarga untuk tiap daerah tersebut. Peneliti mengasumsikan bahwa perumahan kumuh mencerminkan keadaan masyarakat kelas bawah, perumahan biasa mencerminkan keadaan masyarakat kelas menengah, dan perumahan elite mencerminkan masyarakat kelas atas. Di dapat hasil sebagai berikut :

	Masyarakat bawah	Masyarakat menengah	Masyarakat atas
berlibur	3	8	15
bekerja	11	7	3
kursus	2	9	4
dirumah saja	6	7	3

Dari hasil di atas dapat disimpulkan bahwa kegiatan yang dilakukan anak dari masyarakat menengah kebawah pada hari libur adalah bekerja, anak dari penduduk menengah relatif merata dalam memilih kegiatan, dan anak dari penduduk menengah ke atas sebagian besar pergi berlibur.

3.7. Notasi dalam Penarikan Sampel Acak Berlapis

Sebelum membahas lebih jauh tentang penentuan ukuran sampel serta pendugaan nilai parameter populasi dalam SAB, maka perlu dikemukakan beberapa notasi yang berkaitan dengan penggunaan rumus-rumus.

Huruf h menunjukkan lapisan ke-h sedangkan huruf i menunjukkan elemen unit ke-i di dalam lapisan tersebut

1) N_h = total banyaknya unit dalam lapisan ke-h
2) n_h = banyaknya unit dalam sampel yang diambil secara acak dari lapisan ke-h (ukuran sampel yang ditarik dari lapisan ke-h)
3) y_{hi} = nilai yang diperoleh untuk unit ke-i dalam lapisan ke-h
4) $\mu_h = \frac{\sum_{i=1}^{N_h} y_{h_i}}{N_h}$ = nilai rata-rata yang sesungguhnya (parameter rata-rata) dari lapisan ke-h
5) $\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{h_i}}{n_h}$ = nilai rata-rata sampel dari lapisan ke-h sebagai penduga bagi μ_h
6) $W_h = \frac{N_h}{N}$ = pembobot bagi lapisan ke-h yang merupakan rasio ukuran lapisan ke-h terhadap ukuran populasi
7) $f_h = \frac{n_h}{N_h}$ = fraksi penarikan sampel dalam lapisan ke-h
8) σ_h^2 = variansi yang sesungguhnya dari lapisan ke-h
9) S_h^2 = variansi sampel dari lapisan ke-h sebagai penduga bagi parameter σ_h^2,

   dimana $S_h^2 = \frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2}{n_h - 1}$
10) c_h = ongkos penarikan setiap unit dalam lapisan ke-h

11) p_h = dugaan nilai proporsi bagi lapisan ke-h, yang merupakan penduga bagi parameter nilai proporsi dari lapisan ke-h, P_h

Disamping notasi untuk lapisan ke-h, maka berikut ini adalah notasi yang dipergunakan untuk populasi secara keseluruhan.

1) μ = rata-rata populasi
2) σ^2 = variansi populasi
3) N = ukuran populasi, $N = N_1+N_2+N_3+\ldots+N_L$
4) P = proporsi populasi
5) n = ukuran sampel acak berlapis , $n=n1+n2+n3+\ldots n_L$
6) $\bar{y}_{st}$ = rata-rata sampel sebagai penduga bagi nilai rata-rata populasi
7) p_{st} = proporsi sampel sebagai penduga bagi nilai proporsi bagi nilai proporsi populasi

3.8. Penentuan Ukuran Sampel Acak Berlapis Untuk Pendugaan Nilai Rata-rata

Dalam pokok bahasan ini akan dibahas tentang penentuan ukuran SAB untuk pendugaan nilai rata-rata. Sebelumnya pendugaan nilai rata-rata sudah dibahas dalam distribusi sampling statistik dan dalam selang kepercayaan (pada contoh 2.15.). Selang kepercayaan bagi nilai rata-rata tidak lain adalah penentuan selang nilai yang berbentuk $\overline{X} - KS < \mu < \overline{X} + KS$ yang dengan peluang besar memuat nilai rata-rata yang sesungguhnya.

Di dalam penarikan Sampel Acak Berlapis ada 2 hal yang perlu diperhatikan dalam menentukan ukuran sampel acak (n) pada masalah pendugaan tersebut, yaitu :

1. Menentukan ukuran sampel keseluruhan yang akan diambil, yaitu sebesar n.
2. Mengalokasikan sampel berukuran n ke dalam tiap lapisan di mana untuk tiap lapisan ke-h (h=1,2,3,...,L) terdiri dari n_h unit, sedemikian sehingga berlaku :

$$n = n_1 + n_2 + \ldots + n_L = \sum_{n=1}^{L} n_h$$

Alokasi sampel berukuran n ke dalam setiap lapisan dapat menggunakan beberapa metode alokasi yang akan di bahas kemudian dengan memperhatikan 3 faktor berikut :

- Total banyaknya unit-unit dalam setiap lapisan
- Variansi nilai-nilai pengamatan dalam setiap lapisan
- Ongkos untuk memperoleh satu nilai pengamatan dari setiap lapisan, yang tidak lain merupakan ongkos per unit penarikan sampel dari setiap lapisan

Banyaknya unit dalam setiap lapisan mempengaruhi kuantitas informasi dalam sampel. Semakin besar proporsi sampel terhadap populasi, semakin besar pula banyaknya informasi yang diperoleh. Sebagai contoh, suatu sampel berukuran n = 20 dari populasi berukuran N = 200 unit akan mengandung lebih banyak informasi daripada sampel berukuran n = 20 dari populasi berukuran N = 2000 unit. Dengan demikian suatu lapisan yang berukuran besar akan memperoleh alokasi sampel yang lebih besar.

Variasi nilai-nilai pengamatan dalam setiap lapisan perlu mendapat perhatian peneliti, sebab apabila nilai-nilai pengamatan dalam lapisan itu kurang homogen maka diperlukan sampel yang lebih besar dari lapisan itu agar memperoleh suatu pendugaan yang baik bagi parameter populasi.

Jika ongkos per unit penarikan sampel berbeda-beda untuk setiap lapisan maka perlu diambil sampel yang lebih kecil ukurannya bagi lapisan yang memiliki ongkos penarikan sampel yang tinggi agar total ongkos penarikan sampel menjadi minimum.

Dari ketiga faktor tersebut, besar variansi, ongkos, atau total banyaknya unit-unit penarikan sampel akan berbanding lurus dengan ukuran sampel acak berlapis. Dengan adanya perbedaan ketiga faktor tersebut akan berpengaruh langsung pada ukuran sampel acak berlapis.

Dalam teori pendugaan parameter, galat pendugaan akan berbanding terbalik dengan ukuran sampel. Bila n kecil maka galat pendugaan menjadi besar sehingga selang kepercayaan semakin lebar. Semakin besar tingkat kepercayaan yang dinginkan peneliti maka diperlukan n yang semakin besar, karena dengan n semakin besar peneliti semakin percaya bahwa hasil pendugaan akan mendekati parameter populasi. Jadi n berbanding lurus dengan tingkat kepercayaan.

Ada tiga metode pengalokasian sampel berukuran n kedalam setiap lapisan yaitu metode alokasi sebanding, metode alokasi Neyman, dan metode alokasi Optimum.

3.8.1. *Metode Alokasi Sebanding (Proportional Allocation Method)*

Gagasan dasar metode alokasi sebanding adalah pengalokasian ukuran sampel yang proporsional terhadap ukuran lapisan. Dengan demikian ukuran sampel yang ditarik, n_h, berbanding secara proporsional dengan ukuran lapisan, N_h. Metode alokasi sebanding dipergunakan apabila total banyaknya unit-unit penarikan sampel (N_h) berbeda-beda untuk tiap lapisan sedangkan variansi lapisan (σ_h^2) dan ongkos per unit penarikan sampel dari setiap lapisan (c_h) relatif sama.

Dalam metode alokasi sebanding ukuran sampel yang ditarik, n_h, berbanding secara proporsional dengan ukuran lapisan, N_h . Pengalokasian sampel untuk tiap lapisan adalah $n_h = \frac{N_h}{N}.n$ ... 3.a.1.

Sudah dibahas pada pokok pahasan 2.14.1. tentang selang kepercayaan bahwa $|\theta - \hat{\theta}| = \pm z_{\alpha/2}\sigma_{\hat{\theta}}$, misalkan $|\theta - \hat{\theta}| = G$, maka $G^2 = Z^2\sigma_{\hat{\theta}}^2$

$$G^2 = \frac{Z^2\sigma_{N\bar{x}}^2}{N^2}, \text{di mana } \sigma_{N\bar{x}}^2 = \sum_{h=1}^{L}\frac{N_h^2\sigma_h^2}{n_h}\left(1-\frac{n_h}{N_h}\right) = \sum_{h=1}^{L}\frac{N_h^2\sigma_h^2}{n_h} - \sum_{h=1}^{L}N_h\sigma_h^2 \quad (*)$$

Berdasarkan (3.a.1.) diperoleh (*) $\sigma_{N\bar{x}}^2 = \left(\frac{N}{n} - 1\right)\sum N_h\sigma_h^2$

Maka $N^2G^2 = Z^2\left(\frac{N}{n} - 1\right)\sum N_h\sigma_h^2$

$$N^2G^2 + Z^2\sum N_h\sigma_h^2 = \frac{NZ^2\sum N_h\sigma_h^2}{n}$$

$$n = \frac{NZ^2\sum_{h=1}^{L}N_h\sigma_h^2}{N^2G^2 + Z^2\sum_{h=1}^{L}N\sigma_h^2} \qquad \text{................(3 a.2)}$$

Z = nilai tabel kurva normal yang besarnya tergantung pada tingkat kepercayaan pendugaan yang diinginkan peneliti.

Tingkat kepercayaan	80%	90%	95%	99%	100%
Nilai Z	1,290	1,645	1,960	2,575	3,000

CONTOH 3.8.1. metode alokasi sebanding

Misalkan suatu survei bermaksud mengetahui banyak anak dalam suatu keluarga di Mrican pada tahun 1997, dimana keluarga didefinisikan sebagai unit yang beranggotakan ayah, ibu dan anak. Peneliti menganggap bahwa variasi ukuran keluarga di Mrican sangat tinggi karena keadaan ibu yang mempengaruhi jumlah anak dalam keluarga. Oleh karena itu peneliti membagi populasi ke dalam lapisan berdasarkan beberapa karakteristik yaitu ibu yang mengikuti program KB dengan alat (lapisan 1), dengan metode Siblings (lapisan 2), tanpa perhitungan (lapisan 3), dan ibu yang mempunyai penyakit tertentu (yang menyebabkan susah punya anak) (lapisan 4), dan lain-lain (lapisan 5). Andaikan peneliti menginginkan galat pendugaan nilai rata-rata $G\,\overline{y}_{st} = \left|\overline{y}_{st} - \mu\right|$, untuk selanjutnya di singkat dengan G, tidak lebih besar dari 0,5, artinya pendugaan nilai rata-rata sampel acak berlapis tidak boleh menyimpang lebih besar daripada 0,5. Tingkat kepercayaan yang diinginkan peneliti dalam melakukan pendugaan berdasarkan sampel adalah 90 %. Berikut ini adalah data hipotetis, jumlah elemen dalam lapisan dan variansi dalam lapisan.

Lapisan (h)	N_h	σ_h^2
1	448	6
2	131	10
3	81	5
4	108	7
5	100	5
	N = 868	

1. Galat pendugaan, G = 0.5

2. Taraf kepercayaan 90 %, berarti $Z_{tabel} = 1.645$.

Jadi ukuran sampel acak berlapis, n, adalah :

$$n = \frac{N . Z^2 \sum_{h=1}^{L} N_h \sigma_h^2}{N^2 . G^2 + Z^2 . \sum_{h=1}^{L} N_h \sigma_h^2}$$

$$n = \frac{(868).(1.645)^2 (5659)}{(868)^2 . (0.5)^2 + (1.645)^2 (5659)} = 65.26$$

$$= 65.26$$

$$= 66 \text{ (dibulatkan)}$$

Dengan demikian berdasarkan metode alokasi sebanding, ukuran sampel acak berlapis,n, yang harus diambil sekurang-kurangnya 66 rumah tangga. Alokasi sampel tiap lapisan ditentukan sebagai berikut:

$$n_h = \frac{N_h}{N} . n; h = 1,2,3,4,5.$$

$$n_1 = \frac{N_1}{N} n = \frac{448}{868} . (66) = 34.06 = 34 \text{(dibulatkan)}$$

$$n_2 = \frac{N_2}{N} n = \frac{131}{868} . (66) = 9.96 = 10 \text{(dibulatkan)}$$

$$n_3 = \frac{N_3}{N} n = \frac{100}{868} . (66) = 7{,}60 = 8 \text{(dibulatkan)}$$

$$n_4 = \frac{N_4}{N} n = \frac{81}{868} . (66) = 6{,}16 = 6 \text{(dibulatkan)}$$

$$n_5 = \frac{N_5}{N} n = \frac{108}{868} . (66) = 8{,}21 = 8 \text{(dibulatkan)}$$

Jadi : $n_1 + n_2 + n_3 + n_4 + n_5 = n = 66$

3.8.2. *Metode Alokasi Neyman (Neyman Allocation Method)*

Metode alokasi Neyman dipergunakan apabila variansi setiap lapisan berbeda-beda besarnya sedangkan ongkos per unit penarikan sampel dari setiap stratum (ch) dianggap relatif sama besarnya. Pengalokasian sampel tiap lapisan

adalah $n_h = \dfrac{n_h \sigma_h}{\sum_{h=1}^{L} n_h \sigma_h} . n$ ... (3.b.1)

Berdasarkan (3.b.1.) diperoleh (*) $\sigma^2_{N\bar{x}} = \dfrac{1}{n}\left(\Sigma N_h \sigma_h\right)^2 - \Sigma N_h \sigma_h^2$

Maka $N^2 G^2 = Z^2\left(\dfrac{1}{n}\left(\Sigma N_h \sigma_h\right)^2 - \Sigma N_h \sigma_h^2\right)$

$$N^2 G^2 + Z^2 \Sigma N_h \sigma_h^2 = \frac{Z^2\left(\Sigma N_h \sigma_h\right)^2}{n}$$

$$n = \frac{Z^2\left(\sum_{h=1}^{L} N_h \sigma_h^2\right)^2}{N^2 G^2 + Z^2 \sum_{h=1}^{L} N_h \sigma_h^2} \quad \text{...(3b.2)}$$

Dengan $\sum_{h=1}^{L} n_h = n$

CONTOH 3.8.2

Contoh soal seperti pada contoh 3.8.1.

$$n = \frac{Z^2 \left(\sum_{h=1}^{L} N_h S_h\right)^2}{N^2 . G^2 + Z^2 . \sum_{h=1}^{L} N_h . S_h^2}$$

$$n = \frac{(1.645)^2 . (2202.3)^2}{(868)^2 . (0.5)^2 + (1.645)^2 (5659)} = 64.44$$

$n = 65$(dibulatkan)

Dengan demikian berdasarkan metode alokasi Neyman ditentukan ukuran sampel acak berlapis sekurang-kurangnya 65 rumah tangga.

Alokasi sampel untuk tiap lapisan ditentukan sebagai berikut ;

$$n_1 = \frac{1097,6}{2202,3}(65) = 32,39 \quad = 32 \text{ rumah tangga}$$

$$n_2 = \frac{414,2}{2202,3}(65) = 12,22 \quad = 12 \text{ rumah tangga}$$

$$n_3 = \frac{223,6}{2202,3}(65) = 6,60 \quad = 7 \text{ rumah tangga}$$

$$n_4 = \frac{181,1}{2202,3}(65) = 5,35 \quad = 5 \text{ rumah tangga}$$

$$n_5 = \frac{285,8}{2202,3}(65) = 8,44 \quad = 9 \text{ rumah tangga}$$

3.8.3. *Metode Alokasi Optimum*

Metode alokasi optimum dipergunakan apabila variansi tiap lapisan serta ongkos tiap unit penarikan sampel dari setiap lapisan berbeda-beda besarnya dari lapisan yang satu dengan lapisan yang lain. Pengalokasian sampel untuk tiap lapisan adalah $n_h = \dfrac{N_h \sigma_h / \sqrt{c_h}}{\sum_{h=1}^{L} N_h \sigma_h / \sqrt{c_h}}(n)$ ...(3.c.1.)

Berdasarkan (3.c.1.) diperoleh (*) $\sigma^2_{N\bar{x}} = \dfrac{\sum N_h \sigma_h \sqrt{c_h} \sum N_h \sigma_h / \sqrt{c_h}}{n} - \sum N_h \sigma_h^2$

Maka $N^2 G^2 = Z^2 \left(\dfrac{\sum N_h \sigma_h \sqrt{c_h} \sum N_h \sigma_h / \sqrt{c_h}}{n} - \sum N_h \sigma_h^2 \right)$

$$N^2 G^2 + Z^2 \sum N_h \sigma_h^2 = \frac{Z^2 \left(\sum N_h \sigma_h \sqrt{c_h} \sum N_h \sigma_h / \sqrt{c_h} \right)}{n}$$

$$n = \frac{Z^2 \left(\sum_{h=1}^{L} N_h \sigma_h \sqrt{c_h} \right) \left(\sum N_h \sigma_h / \sqrt{c_h} \right)}{N^2 G^2 + Z^2 \sum_{h=1}^{L} N_h \sigma_h^2} \quad \text{....(3.c.2.)}$$

CONTOH 3.8.3

Contoh seperti pada contoh 3.8.1 dan andaikan biaya penarikan sampel (dalam satuan biaya) tiap lapisan pada contoh 3.8.1. sebagai berikut :

Ongkos (c_h)	2 2 3 3 1

$$n = \frac{Z^2\left(\sum_{h=1}^{L} N_h \sigma_h \sqrt{c_h}\right)\left(\sum_{h=1}^{L} N_h \sigma_h / \sqrt{c_h}\right)}{N^2 G^2 + Z^2 \sum_{h=1}^{L} N_h \sigma_h^2}$$

$$n = \frac{(1{,}645)^2 (3170{,}0).(1562{,}3)}{(868)^2.(0{,}5)^2 + (1{,}645)^2 (5659)}$$

$n = 65{,}80 = 66$ rumah tangga

Berdasarkan metode alokasi optimum, maka ukuran sampel acak berlapis, n , yang perlu diambil adalah 66 rumah tangga, dengan alokasi sampel tiap lapisan adalah sebagai berikut :

$$n_h = \frac{N_h \sigma_h / c_h}{\sum_{h=1}^{L} N_h \sigma_h / \sqrt{c_h}} .n;\ h = 1,2,3,4,5$$

$$n_1 = \frac{776{,}2}{1562{,}3}(66) = 32{,}79 \quad = 33 \text{ rumah tangga}$$

$$n_2 = \frac{292{,}9}{1562{,}3}(66) = 12{,}37 \quad = 12 \text{ rumah tangga}$$

$$n_3 = \frac{223{,}6}{1562{,}3}(66) = 9{,}45 \quad = 10 \text{ rumah tangga}$$

$$n_4 = \frac{104{,}6}{1562{,}3}(66) = 4{,}41 = 4 \text{ rumah tangga}$$

$$n_5 = \frac{165{,}0}{1562{,}3}(66) = 6{,}97 \quad = 7 \text{ rumah tangga}$$

$n_1+n_2+n_3+n_4+n_5 = n = 66$ rumah tangga.

Berdasarkan 3 metode alokasi penentuan ukuran sampel di atas maka dapat diambil kesimpulan sebagai berikut:

❖ Metode alokasi sebanding digunakan apabila total banyaknya unit-unit penarikan sampel (N_h) berbeda-beda untuk tiap lapisan, dan variansi lapisan σ_h^2, serta ongkos per unit penarikan sampel (c_h)dari tiap lapisan relatif sama.

- Metode alokasi Neyman digunakan apabila N_h dan σ_h^2 berbeda-beda untuk tiap lapisan sedangkan ongkos (c_h) per lapisan dianggap relatif sama besarnya.
- Metode alokasi optimum digunakan apabila N_h, σ_h^2,c_h, untuk tiap lapisan besarnya berbeda-beda.

3.9. Penentuan Ukuran Sampel Acak Berlapis untuk Pendugaan Proporsi

Dalam banyak hal, populasi yang diteliti dapat digolongkan ke dalam dua kategori, yaitu kategori yang memiliki suatu sifat tertentu dan kategori yang tidak memilki sifat tertentu tersebut. Misalnya ingin diketahui proporsi atau persentase petani yang menanam padi varietas unggul, proporsi ibu-ibu yang mengikuti program KB, proporsi rumah tangga yang memiliki kendaraan pribadi, respons masyarakat terhadap suatu proyek pembangunan di daerahnya, dan sebagainya. Dalam hal ini populasi dikategorikan menjadi petani penanam padi varietas unggul dan petani bukan penanam, ibu-ibu peserta KB dan bukan, dll.

Untuk pembahasan selanjutnya, parameter proporsi populasi akan dinotasikan dengan P, sedangkan penduga bagi P adalah nilai proporsi sampel yang dinotasikan dengan p.

Menurut teori penarikan sampel, selang kepercayaan bagi nilai proporsi tidak lain adalah penentuan selang nilai yang berbentuk $\hat{p} - KS \leq P \leq \hat{p} + KS$ yang dengan peluang yang besar memuat nilai proporsi yang sesungguhnya. Berikut ini akan dibahas tentang penentuan ukuran SAB untuk pendugaan nilai proporsi, dalam tiga metode alokasi.

3.9.1. *Metode Alokasi Sebanding (Proportional Alocation Method)*

Metode alokasi sebanding dipergunakan apabila ukuran lapisan (N_h) berbeda-beda untuk tiap lapisan, sedangkan proporsi, P_h, dan ongkos per unit penarikan sampel, c_h, dari setiap lapisan diperkirakan relatif sama besarnya. Pengalokasian sampel tiap lapisan adalah $n_h = \frac{N_h}{N}.n$ ...(4.a.1.)

Telah dibahas pada pokok bahasan 2.14.1. tentang selang kepercayaan bahwa

$\left|\theta - \hat{\theta}\right| = \pm z_{\alpha/2}\sigma_{\hat{\theta}}$. Misalkan $\left|\theta - \hat{\theta}\right| = G$, maka $G^2 = Z^2\sigma_{\hat{\theta}}^2$,

$G^2 = \frac{Z^2\sigma_{\hat{p}}^2}{N^2}$, di mana dalam pendugaan proporsi diketahui bahwa

$$\sigma_{\hat{p}}^2 = \sum \frac{N_h^2 P_h(1-P_h)}{n_h}\left(1 - \frac{n_h}{N_h}\right) = \sum \frac{N_h^2 P_h(1-P_h)}{n_h} - \sum N_h P_h(1-P_h) \quad \text{...(\#)}$$

Berdasarkan (4.a.1.) diperoleh (#) $\sigma_{\hat{p}}^2 = N\sum \frac{N_h P_h(1-P_h)}{n} - \sum N_h P_h(1-P_h)$

Maka $N^2G^2 = Z^2\left(\frac{N\sum N_h P_h(1-P_h)}{n} - \sum N_h P_h(1-P_h)\right)$

$$N^2G^2 + Z^2\sum N_h P_h(1-P_h) = \frac{NZ^2\sum N_h P_h(1-P_h)}{n}$$

$$n = \frac{NZ^2\sum N_h P_h(1-P_h)}{N^2G^2 + Z^2\sum N_h P_h(1-P_h)} \quad \text{....(4.a.2.)}$$

Besarnya P_h dalam penentuan ukuran sampel ditentukan berdasarkan penelitian pendahuluan sebelum dilakukan penelitian yang sebenarnya.

CONTOH 3.9.1

Contoh penentuan ukuran sampel untuk pendugaan nilai proporsi. Misalkan suatu survei bermaksud mengetahui proporsi pemilikan TV dalam berbagai ukuran di DIY. Populasi dibagi dalam beberapa lapisan yaitu pemilik TV berukuran ≤14", 16"-18", 20"-22", 24"-26", ≥27". Galat pendugaannya 5%, dan tingkat kepercayaannya 90%. Hasil survey yang diperoleh sebagai berikut:

Lapisan (h)	N_h	P_h (perkiraan)
1	448	0.7
2	131	0.7
3	81	0.7
4	108	0.7
5	100	0.7
Total	N = 868	

Dapat dihitung besar ukuran sampel menggunakan metode alokasi sebanding.

$$n = \frac{N Z^2 \sum_{h=1}^{L} N_h P_h (1 - P_h)}{N^2 G^2 + Z^2 \sum_{h=1}^{L} N_h P_h (1 - P_h)}$$

$$n = \frac{(868).(1{,}645)^2 (182{,}3)}{(868)^2 0{,}05^2 + (1{,}645)^2 (182{,}3)}$$

$$= 180{,}15 = 181 \text{ rumah tangga}$$

Dengan demikian berdasarkan metode alokasi sebanding ukuran sampel acak berlapis, n, yang harus diambil adalah 181 rumah tangga, dengan alokasi sampel untuk tiap lapisan, n_h, ditentukan sebagai berikut :

$$n_h = \frac{N_h}{N} n; h = 1,2,3,4,5$$

$$n_1 = \frac{448}{868}(181) = 93.42 = 93 \text{rumahtangga}$$

$$n_2 = \frac{131}{868}(181) = 27.32 = 27 \text{rumahtangga}$$

$$n_3 = \frac{81}{868}(181) = 16.89 = 17 \text{rumahtangga}$$

$$n_4 = \frac{108}{868}(181) = 22.52 = 23 \text{rumahtangga}$$

$$n_5 = \frac{100}{868}(181) = 20.85 = 21 \text{rumahtangga}$$

3.9.2. *Metode Alokasi Neyman (Neyman Allocation Method)*

Metode alokasi Neyman dipergunakan apabila populasi yang dihadapi diperkirakan memiliki proporsi lapisan, P_h, yang berbeda-beda besarnya untuk tiap lapisan, sedangkan ongkos tiap unit penarikan sampel, c_h, relatif sama besarnya untuk tiap lapisan. Rumus yang digunakan untuk menghitung besar sampel yang diambil dari populasi dan besar sampel untuk tiap lapisan adalah sebagai berikut :

Pengalokasian ukuran sampel untuk tiap lapisan adalah :

$$n_h = \frac{N_h\sqrt{P_h(1-P_h)}}{\sum_{h=1}^{L} N_h\sqrt{P_h(1-P_h)}}.n \qquad .....(4b.1)$$

Berdasarkan (4.b.1.) diperoleh (#)

$$\sigma_{\hat{p}}^2 = \frac{1}{n}\left(\sum N_h\sqrt{P_h(1-P_h)}\right)^2 - \sum N_h P_h(1-P_h)$$

Maka $N^2G^2 = Z^2\left(\frac{1}{n}\left(\sum N_h\sqrt{P_h(1-P_h)}\right)^2 - \sum N_h P_h(1-P_h)\right)$

$$N^2G^2 + Z^2\sum N_h P_h(1-P_h) = \frac{Z^2\left(\sum N_h\sqrt{P_h(1-P_h)}\right)^2}{n}$$

$$n = \frac{Z^2\left(\sum_{h=1}^{L} N_h\sqrt{P_h(1-P_h)}\right)^2}{N^2G^2 + Z^2\sum_{h=1}^{L} N_h P_h(1-P_h)} \qquad ....(4b.2.)$$

Dengan $\sum n_h = n$

CONTOH 3.9.2

Seperti pada contoh 3.9.1

Lapisan (h)	Nh	Ph (perkiraan)	Nh Ph (1-Ph)	$N_h\sqrt{P_h(1-P_h)}$
1	448	0,8	71,68	179.2
2	131	0.7	27,51	60.03
3	81	0.65	18,43	38.63
4	108	0.5	27	54
5	100	0.7	21	45,83

$$n = \frac{Z^2(\sum_{h=1}^{L} N_h\sqrt{P_h(1-P_h)})^2}{N^2G^2 + Z^2\sum_{h=1}^{L} N_h P_h(1-P_h)}$$

$$= \frac{(1{,}645)^2(377{,}69)^2}{(868)^2(0{,}05)^2 + (1{,}645)^2(165{,}62)} = 165{,}55 \approx 166$$

$$n_1 = \frac{179{,}2}{377{,}69}.166 = 78{,}76 = 79$$

$$n_2 = \frac{60{,}03}{377{,}69}.166 = 26{,}38 = 26$$

$$n_3 = \frac{38{,}63}{377{,}69}.166 = 16{,}98 = 17$$

$$n_4 = \frac{54}{377{,}69}.166 = 23{,}73 = 24$$

$$n_5 = \frac{45{,}83}{377{,}69}.166 = 20{,}14 = 20$$

$n_1 + n_2 + n_3 + n_4 + n_5 = n = 166$

3.9.3. Metode Alokasi Optimum (Optimum Allocation Method)

Metode alokasi optimum dipergunakan apabila populasi yang dihadapi memiliki proporsi lapisan, P_h, serta ongkos tiap unit penarikan sampel berbeda-beda untuk tiap lapisan, maka metode ini akan sangat sesuai untuk menentukan ukuran sampel acak berlapis serta ukuran sampel untuk tiap lapisan. Pengalokasian ukuran sampel untuk tiap lapisan adalah :

$$n_h = \frac{N_h\sqrt{P_h(1-P_h)}/\sqrt{c_h}}{\sum_{h=1}^{L} N_h\sqrt{P_h(1-P_h)}/\sqrt{c_h}}.n \quad ..(4.c.1.)$$

Berdasarkan (4.c.1.) diperoleh (#)

$$\sigma_{\hat{p}}^2 = \frac{\sum N_h\sqrt{P_h(1-P_h)}\sqrt{c_h}\sum N_h\sqrt{P_h(1-P_h)}/\sqrt{c_h}}{n} - \sum N_h P_h(1-P_h)$$

maka

$$N^2G^2 = Z^2\left(\frac{\sum N_h\sqrt{P_h(1-P_h)}\sqrt{c_h}\sum N_h\sqrt{P_h(1-P_h)}/\sqrt{c_h}}{n} - \sum N_hP_h(1-P_h)\right)$$

$$N^2G^2 + Z^2\sum N_hP_h(1-P_h) = \frac{Z^2\left(\sum N_h\sqrt{P_h(1-P_h)}\sqrt{c_h}\sum N_h\sqrt{P_h(1-P_h)}/\sqrt{c_h}\right)}{n}$$

$$n = \frac{Z^2(\sum_{h=1}^{L} N_h\sqrt{P_h(1-P_h)}\sqrt{c_h}).(\sum_{h=1}^{L} N_h\sqrt{P_h(1-P_h)}/\sqrt{c_h})}{N^2G^2 + Z^2\sum_{h=1}^{L} N_hP_h(1-P_h)} \qquad \text{........(4c.2.)}$$

CONTOH 3.9.3

Soal seperti pada contoh 3.9.1

Lapisan (h)	N_h	P_h (perkiraan)	C_h	$\sqrt{C_h}$	$N_h P_h$ $(1-P_h)$	$N_h\sqrt{P_h(1-P_h)}\sqrt{C_h}$	$\frac{n_h\sqrt{P_h(1-P_h)}}{\sqrt{C_h}}$
1	448	0,8	2	1,414	71,68	116,117	126,733
2	131	0,7	2	1,414	27,51	84,885	42,455
3	81	0,65	3	1,732	18,428	66,915	22,306
4	108	0,5	3	1,732	27	93,528	31,178
5	100	0,7	1	1	21	45,826	45,826
Total	868				165,618	407,271	268,498

$$n = \frac{Z^2(\sum N_h\sqrt{P_h(1-P_h)}\sqrt{C_h})(\sum N_h\sqrt{P_h(1-P_h)}/\sqrt{C_h}}{N^2G^2 + Z^2\sum N_hP_h(1-P_h)}$$

$$= \frac{(1{,}645)^2(407{,}271)(268{,}498)}{(868)^2(0{,}05)^2 + (1{,}645)^2(165{,}618)}$$

$$= \frac{295907{,}75}{2049{,}56 + 448{,}166} = 118{,}47 \approx 119$$

Berdasarkan metode alokasi optimum , n, yang diambil 119 dengan alokasi setiap lapisan sebagai berikut :

$$n_h = \frac{N_h\sqrt{P_h(1-P_h)}/\sqrt{C_h}}{\Sigma N_h\sqrt{P_h(1-P_h)}/\sqrt{C_h}}.n$$

$$n_1 = \frac{126{,}733}{268{,}498}.118 = 55{,}697 \approx 56$$

$$n_2 = \frac{42{,}455}{268{,}498}.118 = 18{,}658 \approx 19$$

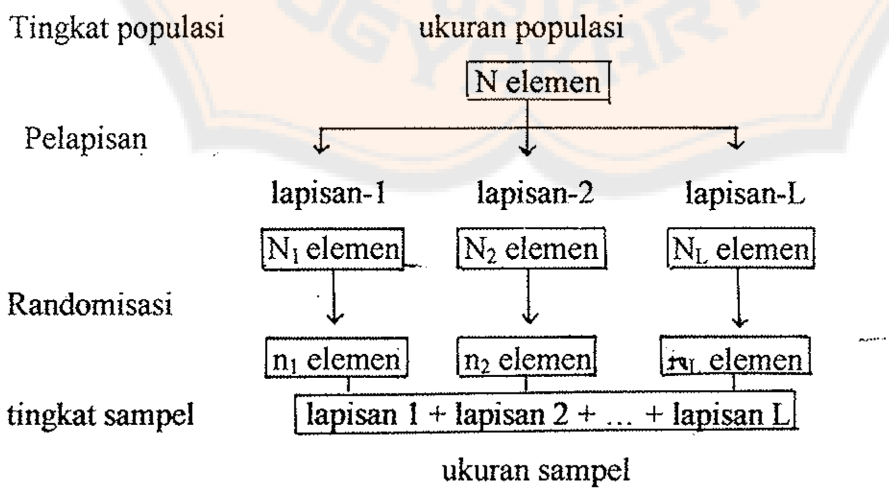
$$n_3 = \frac{22{,}306}{268{,}498}.118 = 9{,}803 \approx 10$$

$$n_4 = \frac{31{,}178}{268{,}498}.118 = 13{,}702 \approx 14$$

$$n_5 = \frac{45{,}826}{268{,}498}.118 = 20{,}139 \approx 20$$

Berdasarkan pembahasan tentang alokasi sampel untuk pendugaan rata-rata dan proporsi yang telah dijelaskan di atas pada hakekatnya ekuivalen. Tetapi terdapat perbedaan di antara keduanya, yaitu hanyalah pada rumus variansinya yang disubstitusikan pada masing-masing rumus. Pada pendugaan nilai variansi rata-rata $\hat{V}(\bar{y}) = \frac{S}{\sqrt{n}}$, sedangkan variansi pendugaan nilai proporsi $\hat{V}(\hat{p}) = \hat{p}(1-\hat{p})$

Secara umum, langkah-langkah yang perlu ditempuh dalam penarikan SAB dapat digambarkan dalam diagram berikut :



Penentuan metode alokasi sampel mana yang akan dipakai untuk pendugaan nilai rata-rata populasi dan total populasi dapat digambarkan dengan diagram berikut :

Populasi dibagi dalam L lapisan, dan tiap lapisan memuat N_h elemen, dengan ongkos C_h dan besarnya variansi σ_h^2

N_h berbeda, C_h, S_h^2 sama	N_h, σ_h^2 berbeda, C_h sama	N_h, C_h, σ_h^2 berbeda
Alokasi Sebanding	Alokasi Neyman	Alokasi Optimum
$n = \dfrac{N Z^2 \sum_{h=1}^{L} N_h \sigma_h^2}{N^2 G^2 + Z^2 \sum_{h=1}^{L} N_h \sigma_h^2}$	$n = \dfrac{Z^2 \left(\sum_{h=1}^{L} N_h \sigma_h \right)^2}{N^2 G^2 + Z^2 \sum_{h=1}^{L} N_h \sigma_h^2}$	$n = \dfrac{Z^2 \left(\sum_{h=1}^{L} N_h \sigma_h \sqrt{C_h} \right) \left(\sum_{h=1}^{L} N_h \sigma_h / \sqrt{C_h} \right)}{N^2 G^2 + Z^2 \sum_{h=1}^{L} N_h \sigma_h^2}$
$n_h = \dfrac{N_h}{N} n$	$n_h = \dfrac{N_h \sigma_h}{\sum_{h=1}^{L} N_h \sigma_h} n$	$n_h = \dfrac{N_h \sigma_h / \sqrt{C_h}}{\sum_{h=1}^{L} N_h \sigma_h / \sqrt{C_h}}$

Penentuan ukuran sampel untuk pendugaan nilai proporsi populasi sama dengan rumus penentuan ukuran sampel untuk pendugaan rata-rata populasi dan pendugaan total populasi, hanya saja dalam rumus proporsi, variansi (σ_h^2) disubstitusi dengan P_h (1-P_h).

3.10. Pembentukan Lapisan

Salah satu krietria pembentukan lapisan adalah adanya batas antar lapisan yang jelas, tiap lapisan relatif homogen dan tiap lapisan harus saling asing. Pada sub bab ini akan dibahas bagaimana menentukan batas antar lapisan. Seperti sudah dijelaskan pada sub bab sebelumnya bahwa pembentukan lapisan harus sedemikian rupa sehingga karakteristik dalam lapisan relatif homogen. Sebagai contoh, telah dijelaskan pada pembahasan terdahulu tentang kelemahan SAS (3.4). Pandangan yang berbeda terjadi antara pengguna helm pria dan wanita. Jika dalam populasi tidak ada lapisan yang jelas maka akan mempengaruhi kesimpulan akhir, karena tidak adanya karakteristik yang jelas dalam populasi tersebut. Jika unit-unit penarikan sampel dipilih secara acak dari setiap lapisan, maka variansi lapisan tersebut relatif kecil karena sifat lapisan yang relatif homogen. Hal tersebut diperoleh dengan cara mengalokasikan unit-unit yang karakteristiknya homogen kedalam lapisan yang sama. Hal ini dapat dicapai dengan baik bila diketahui disribusi variabel random yang dipelajari (Y). Lapisan tersebut dapat dibentuk dengan memotong distribusi tersebut pada titik-titik yang dipandang sesuai sebagai batas lapisan.

Berikut ini akan dibahas bagaimana menentukan batas-batas lapisan bila distribusi peluang variabel random Y diketahui. Andaikan distribusi Y kontinu dengan fungsi densitas f(y), a < y < b. Jika dibentuk L lapisan, dan domain Y dibatasi oleh titik-titik $y_1<y_2<y_3<\ldots<y_{L-1}$. Frequensi relatif pada lapisan h disimbolkan dengan W_h, dan variansi lapisan ke-h disimbolkan σ_h^2 , maka didefinisikan sebagai berikut :

$$\left.\begin{aligned} W_h &= \int_{y_{h-1}}^{y_h} f(t)dt && \text{adalah frekuensi lapisan} \\ W_h\,\mu_h &= \int_{y_{h-1}}^{y_h} t.f(t).dt, && \text{rata}-\text{rata lapisan ke}-\text{h (nilai harapan)} \\ W_h\,\sigma_h^2 &= \int_{y_{h-1}}^{y_h} t^2.f(t).dt - W_h.\mu_h^2, && \text{Variansi lapisan ke}-\text{h} \end{aligned}\right\} \ldots 1$$

Rata-rata populasi $\mu = \sum_{h=1}^{L} W_h\,\mu_h$, dan penduganya diperoleh dari suatu sampel acak

berlapis adalah $\bar{y} = \sum W_h\,\bar{y}_h$ dengan variansi $V\ (\bar{y}) = \sum W_h^2 \frac{\sigma_h^2}{n_h}$. ...2 Terlihat bahwa W_h

σ_h^2 merupakan variansi pada lapisan ke-h, jadi bisa dikatakan bahwa $V(\bar{y})$

merupakan suatu fungsi dari batas-batas lapisan $y_1, y_2, y_3, \ldots, y_{L-1}$. Sedangkan rata-rata populasi tidak lain adalah rata-rata terbobot dari lapisan, dengan W_h sebagai pembobotnya. Masalah yang timbul selanjutnya adalah bagaimana menentukan batas-batas lapisan (y_i) sehingga V ($\bar{y}$) menjadi minimum untuk alokasi n_h yang diberikan. Berikut ini akan dibahas penentuan batas-batas lapisan berdasarkan jenis alokasi sampel dalam penentuan ukuran sampel pada SAB.

3.10.1. Penentuan Batas Lapisan pada Alokasi Proporsional

Jika $n_h = n.W_h$, maka $V(\bar{y}) = \frac{1}{n} \cdot \sum W_h \sigma_h^2$. Untuk menentukan nilai-nilai terbaik dari y_h, menurut Des Raj (1968), $\sum W_h \sigma_h^2$ harus diminimumkan dengan menurunkan terhadap y_h, dan menyamakan dengan nol.

$$\begin{aligned}
\text{Misalkan } A &= \sum W_h \sigma_h^2 \\
&= \sum \left[\int_{y_{h-1}}^{y_h} t^2 f(t)dt - W_h \mu_h^2 \right] \\
&= \int_{y_{h-1}}^{y_h} t^2.f(t)dt - W_h \mu_h^2 + \int_{y_h}^{y_{h+1}} t^2.f(t)dt - W_{h+1} \mu_{h+1}^2 \\
&= -\frac{(W_h \mu_h)^2}{W_h} - \frac{(W_{h+1} \mu_{h+1})^2}{W_{h+1}}
\end{aligned}$$

$$\frac{\partial W_h}{\partial y_h} = \frac{\partial \left(\int_{y_{h+1}}^{y_h} f(t)dt \right)}{\partial y_h} = f(y_h)$$

$$\frac{\partial W_{h+1}}{W_h} = \frac{\partial \left(\int_{y_h}^{y_{h+1}} f(t)dt \right)}{\partial y_h} = \frac{\partial \left(\int_{y_{h+1}}^{y_h} -f(t)dt \right)}{\partial y_h} = -f(y_h)$$

$$\frac{\partial (W_h \mu_h)^2}{\partial y_h} = 2 W_h \mu_h \frac{\partial (W_h \mu_h)}{\partial y_h}$$

$$= 2 W_h \mu_h \frac{\partial \left(\int_{y_{h+1}}^{y_h} t.f(t)dt \right)}{\partial y_h}$$

$$= 2 W_h \mu_h y_h f(y_h)$$

$$\frac{\partial (W_{h+1} \mu_{h+1})^2}{\partial y_h} = 2 W_{h+1} \mu_{h+1} \frac{\partial (W_{h+1} \mu_{h+1})}{\partial y_h}$$

$$= 2 W_{h+1} \mu_{h+1} \frac{\partial \left(\int_{y_h}^{y_{h+1}} tf(t)dt \right)}{\partial y_h}$$

$$= 2 W_{h+1} \mu_{h+1} (-y_h f(y_h))$$

$$= -2 W_{h+1} \mu_{h+1} y_h f(y_h)$$

Karena $\partial A/\partial y_h = 0$, maka

$$\frac{\partial \left[-\frac{(W_h \mu_h)^2}{W_h} - \frac{(W_{h+1} \mu_{h+1})^2}{W_{h+1}} \right]}{\partial y_h} = -\frac{\partial \left(\frac{(W_h \mu_h)^2}{W_h} \right)}{\partial y_h} - \frac{\partial \left(\frac{(W_{h+1} \mu_{h+1})^2}{W_{h+1}} \right)}{\partial y_h}$$

$$= -\left[\frac{2 W_h \mu_h y_h f(y_h) W_h - f(y_h)(\mu_h W_h)^2}{W_h^2} \right] -$$

$$\left[\frac{-2 W_{h+1} \mu_{h+1} y_h f(y_h) W_{h+1} + f(y_h)(W_{h+1} \mu_{h+1})^2}{W_{h+1}^2} \right]$$

$$= -(2\mu_h y_h f(y_h) - f(y_h)\mu_h^2) - (-2\mu_{h+1} y_h f(y_h) + f(y_h)\mu_{h+1}^2)$$

$$(-2\mu_h + 2\mu_{h+1})y_h = -\mu_h^2 + \mu_{h+1}^2$$

$$2y_h(-\mu_h + \mu_{h+1}) = (\mu_h + \mu_{h+1})(-\mu_h + \mu_{h+1})$$

$$2y_h = \mu_h + \mu_{h+1}$$

$$y_h = \frac{\mu_h + \mu_{h+1}}{2}$$

Hal tersebut menunjukkan bahwa nilai terbaik y_h adalah nilai rata-rata dari rata-rata 2 lapisan yang dibatasi olehnya.

CONTOH 3.10.1. Alokasi Proporsional

Misalkan penelitian yang dilakukan bertujuan mengetahui rata-rata pendapatan penduduk desa Tidak Tahu. Peneliti akan membagi populasi berdasarkan tingkat konsumsi (pengeluaran) penduduk, karena penduduk yang tingkat konsumsinya tinggidiasumsikan penduduk tersebut berpendapatan tinggi juga. Jadi tingkat konsumsi penduduk berhubungan erat dengan pendapatan penduduk tersebut. Pnenliti membagi populasi dalam 3 lapisan yaitu masyarakat kalangan bawah, masyarakat kalangan tengah, masyarakat kalangan atas. Data yang didapat sebagai berikut : (data dalam ribuan)

Kalangan bawah	Kalangan tengah	Kalangan atas
100,110,145,105,	202,250,210,225,	325,450,510,470,
190,175,155,160,	275,265,190,305,	390,410,425,435,
200,160,125,140,	260,285,290,230,	375,350,430,440,
180,175,200	235,250,270	360,380,375

$N_1 = 15$ $\quad$ $N_2 = 15$ $\quad$ $N_3 = 15$

$\mu_1 = 154{,}667$ $\quad$ $\mu_2 = 249{,}333$ $\quad$ $\mu_3 = 408{,}333$

Terlihat bahwa pengeluaran terletak pada interval [100,510]

Akan dicari batas-batas lapisan (yi)

$$y_1 = \frac{\mu_1 + \mu_2}{2} = \frac{154,667 + 249,333}{2} = 201,999 \approx 202$$

$$y_2 = \frac{\mu_2 + \mu_3}{2} = \frac{249,333 + 408,333}{2} = 328,833 \approx 329$$

Kalangan bawah terletak pada interval 100-202

Kalangan tengah terlatak pada interval 203-329

Kalangan atas terletak pada interval 330-510

3.10.2. Batas Antar Lapisan pada Metode Alokasi Optimum

Akan ditentukan batas-batas antar lapisan berdasarkan bila penentuan ukuran sampel menggunakan metode alokasi optimum.

Didefinisikan $n_h = \dfrac{n W_h \sigma_h}{\sum W_h \sigma_h}$, $V(\bar{y}) = \dfrac{1}{n}(W_h \sigma_h)^2$

Karena y_h merupakan penyelesaian yang akan dicari dengan turunan, maka y_h akan muncul dalam kuantitas tersebut hanya melalui $W_h\sigma_h$ dan $W_{h+1}\sigma_{h+1}$.

Misalkan $C = W_h\sigma_h + W_{h+1}\sigma_{h+1}$ untuk variansi dalam y_h

Dari (1) diperoleh :

$$\frac{\partial\left(W_h\sigma_h^2\right)}{\partial y_h} = \frac{\partial \int_{y_{h+1}}^{y_h} t^2.f(t)dt}{\partial y_h} - \frac{\partial W_h\mu_h^2}{\partial y_h} = y_h^2.f(y_h) - \frac{\partial \frac{\left(W_h\mu_h\right)^2}{W_h}}{\partial y_h}$$

$$= y_h^2.f(y_h) - \frac{\partial \frac{\left(\int_{y_{h-1}}^{y_h} t.f(t)dt\right)^2}{\int_{y_{h-1}}^{y_h} f(t)dt}}{\partial y_h}$$

$$= y_h^2 f(y_h) - \frac{2\left(\int_{y_{h-1}}^{y_h} t.f(t)dt\right)\left(y_h f(y_h)\right)\left(\int_{y_{h-1}}^{y_h} f(t)dt\right) - f(y_h)\left(\int_{y_{h-1}}^{y_h} t.f(t)dt\right)^2}{\left(\int_{y_{h-1}}^{y_h} f(t)dt\right)^2}$$

$$= y_h^2 f(y_h) - \left(\frac{2W_h\mu_h y_h f(y_h) W_h - f(y_h)\left(W_h\mu_h\right)^2}{W_h^2}\right)$$

$$= y_h^2 f(y_h) - 2\mu_h y_h f(y_h) + f(y_h)\mu_h^2$$

$$= f(y_h)\left(y_h^2 - 2\mu_h y_h + \mu_h^2\right)$$

$$\frac{\partial\left(W_h\sigma_h^2\right)}{\partial y_h} = f(y_h)\left(y_h-\mu_h\right)^2 \qquad \ldots\ 3$$

Dari (3) diperoleh :

$$\frac{\partial\left(W_h\sigma_h^2\right)}{\partial y_h} = f(y_h)\left(y_h-\mu_h\right)^2$$

$$\sigma_h^2\frac{\partial W_h}{\partial y_h} + W_h 2\sigma_h\frac{\partial \sigma_h}{\partial y_h} = f(y_h)\left(y_h-\mu_h\right)^2$$

$$W_h\frac{\partial \sigma_h}{\partial y_h} = \frac{f(y_h)\left(y_h-\mu_h\right)^2 - \sigma_h^2 f(y_h)}{2\sigma_h}$$

$$= \left(2\sigma_h\right)^{-1} f(y_h)\left[\left(y_h-\mu_h\right)^2 - \sigma_h^2\right]$$

maka $\frac{\partial(W_h\sigma_h)}{\partial y_h} = \sigma_h \frac{\partial W_h}{\partial y_h} + W_h \frac{\partial \sigma_h}{\partial y_h}$

$$= \sigma_h f(y_h) + (2\sigma_h)^{-1} f(y_h)\left[(y_h-\mu_h)^2 - \sigma_h^2\right]$$

$$= \sigma_h f(y_h) + (2\sigma_h)^{-1} f(y_h)(y_h-\mu_h)^2 - \frac{f(y_h)\sigma_h}{2}$$

$$= \frac{\sigma_h f(y_h)}{2} + (2\sigma_h)^{-1} f(y_h)(y_h-\mu_h)^2$$

$$= (2\sigma_h)^{-1} f(y_h)\left[(y_h-\mu_h)^2 + \sigma_h^2\right]$$

Demikian juga untuk $\frac{\partial(W_{h+1}\sigma_{h+1})}{\partial y_h} = -(2\sigma_{h+1})^{-1} f(y_h)\left[(y_h-\mu_{h+1})^2 + \sigma_{h+1}^2\right]$

Dari $\frac{\partial C}{\partial y_h} = 0$ maka

$$\frac{\partial(W_h\sigma_h + W_{h+1}\sigma_{h+1})}{\partial y_h} = 0$$

$$\frac{\partial W_h\sigma_h}{\partial y_h} + \frac{\partial W_{h+1}\sigma_{h+1}}{\partial y_h} = 0$$

$$(2\sigma_h)^{-1} f(y_h)\left[(y_h-\mu_h)^2 + \sigma_h^2\right] - (2\sigma_{h+1})^{-1} f(y_h)\left[(y_h-\mu_{h+1})^2 + \sigma_{h+1}^2\right] = 0$$

$$\frac{(y_h-\mu_h)^2 + \sigma_h^2}{\sigma_h} - \frac{(y_h-\mu_{h+1})^2 + \sigma_{h+1}^2}{\sigma_{h+1}} = 0;\ \ h = 1,2,3,\ldots,L-1$$

Persamaan diatas sulit dipecahkan, karena μ_h dan S_h tergantung pada y_h, dengan alasan ini diperlukan pendekatan untuk meminimumkan $V(\bar{y})$. Sebuah metode pendugaan yang baik, dikemukakan oleh Dalenius dan Hodges (1959) (dalam Cochran, 1991), disajikan untuk meminimumkan $\Sigma\, W_h\sigma_h$

Misalkan $Z(y) = \int_{y_0}^{y} \sqrt{f(t)}dt$

Menurutnya jika lapisannya banyak sekali dan ukurannya kecil, f(y) kira-kira akan mendekati konstan (seragam) di dalam sebuah lapisan tertentu.

Andaikan $W_h = \int_{y_{h-1}}^{y_h} f(t)dt = f_h(y_h - y_{h-1})$ 4

$$\sigma_h = \frac{1}{\sqrt{12}}\left(y_h - y_{h-1}\right) \qquad \ldots 5$$

$$Z_h - Z_{h-1} = \int_{y_{h-1}}^{y_h} \sqrt{f(t)}dt = \sqrt{f_h}\left(y_h - y_{h-1}\right)$$

Dimana f_h adalah nilai "konstan" dari f(y) dalam lapisan h. Dengan menggunakan (4) dan (5) maka didapatkan

$$\sqrt{12}\sum_{h=1}^{L} W_h \sigma_h = \sum_{h=1}^{L} f_h \left(y_h - y_{h-1}\right)^2 = \sum_{h=1}^{L}\left(Z_h - Z_{h-1}\right)^2$$

Karena (Z_L-Z_0) tetap, maka penjumlahan pada ruas kanan menjadi minimum dengan membuat (Z_h-Z_{h-1}) konstan. Untuk f(y) tertentu, aturannya adalah dengan membentuk kumulatif f(y) dan pilih y_h sehingga menghasilkan interval yang sama pada skala $\sqrt{f(y)}$ kumulatif.

CONTOH 3. 10.2. Alokasi Optimum

Contoh diambil dari buku Cochran 1991:hal 147

Penghitungan Batasan Lapisan dengan Aturan Kumulatif $\sqrt{f(y)}$

Pinjaman Industri / Jumlah pinjaman %	f(y)	kum $\sqrt{f(y)}$	Pinjaman Industri / Jumlah pinjaman %	f(y)	kum $\sqrt{f(y)}$
0-5	3464	58,9	50-55	126	340,3
5-10	2516	109,1	55-60	107	350,6
10-15	2157	155,5	60-65	82	359,7
15-20	1581	195,3	65-70	50	366,8
20-25	1142	229,1	70-75	39	373,0
25-30	746	256,4	75-80	25	378,0
30-35	512	279,0	80-85	16	352,0
35-40	376	289,4	85-90	9	386,4
40-45	265	314,7	90-95	2	387,8
45-50	207	329,1	95-100	3	389,5

Data tersebut menggambarkan distribusi frekensi dari presentase pinjaman bank yang ditujukan untuk pinjaman industri dalam sebuah populasi dari 13,435 Bank AS (Mc Evoy,1956). Distribusi tersebut miring (skew) dengan modusnya pad nilai akhir yang rendah. Dalam kolom kumulatif $\sqrt{f(y)}$, 58,9 = $\sqrt{3464}$; 109,1 = $\sqrt{3464}+\sqrt{2516}$ dan seterusnya. Misalkan peneliti menginginkan 5 lapisan, karena jumlah kumulatif $\sqrt{f(y)}$ adalah 389,5, titik pembaginya berada pada skala 77,9 diperoleh dari 389,5:5; dan 155,8 diperoleh dari 77,9 + 77,9; 233,7; dan 311,6. Jadi tiap lapisan mempunyai interval kumulatif ± 77,9. Titik terdekat yang tersedia adalah sebagai berikut:

	Lapisan				
	1	2	3	4	5
Batasan	0-5%	5-15%	15-25%	25-45%	45-100%
Interval kum $\sqrt{f(y)}$	58,9	96,6	73,6	85,6	74,8

3.11. Jumlah Lapisan

Ada satu pertanyaan yang perlu untuk dijawab, yaitu berapa jumlah lapisan yang diperlukan dalam pengambilan sampel. Pada bab II telah dijelaskan bahwa pembentukan lapisan dapat menghasilkan variansi yang lebih kecil daripada tanpa pelapisan. Karena dengan pelapisan sampel, yang diambil relatif homogen dan mempunyai karakteristik yang lebih khusus. Sebenarnya, pelapisan dapat dibuat pada suatu nilai sehingga hanya satu unit dipilih dari setiap lapisan. Dengan demikian jumlah lapisan dapat dibuat sebanyak jumlah unit yang dipilih.

Misalkan lapisan dibentuk dari nilai distribusi seragam (misalkan distribusi y) dalam interval (a, a+d). Menurut distribusi seragam, σ_y^2 sebelum pelapisan adalah $d^2/12$, sehingga dengan sampel acak sederhana berukuran n, $V(y) = d^2/12$.

Jika L lapisan yang berukuran sama dibuat, maka variansi dalam setiap lapisan adalah $\sigma_{yh}{}^2 = d^2/12L^2$. Oleh karena itu, untuk sampel acak berlapis, dengan

$W_h = \frac{1}{L}$ dan $n_h = \frac{n}{L}$, variansi pendugaan rata – rata

$$V(\bar{y}_{st}) = \frac{1}{n}\left(\sum_{h=1}^{L} W_h \sigma_{y_h}\right)^2 = \frac{1}{n}\left(\sum_{h=1}^{L} \frac{1}{L}\frac{d}{\sqrt{12.L}}\right)^2 = \frac{d^2}{12nL^4} = \frac{V(\bar{y})}{L^2}$$

Maka distribusi seragam, variansi dari $\bar{y}_{st}$ berbanding terbalik dengan kuadarat jumlah lapisan. Dengan demikian, $L \approx \frac{1}{\sqrt{V(\bar{y}_{st})}}$. Pada populasi sama, semakin kecil variansi pendugaan, yang berarti semakin tinggi ketelitian yang diharapkan, maka jumlah lapisan yang diperlukanpun semakin banyak.

BAB IV

PENDUGAAN PARAMETER DALAM PENARIKAN SAMPEL ACAK BERLAPIS

Pada bab terdahulu telah dibahas prosedur penarikan Sampel Acak Berlapis (SAB) dan penentuan ukuran sampel untuk pendugaan total populasi dan proporsi populasi. Bab IV ini merupakan kelanjutan pembahasan bab terdahulu yaitu membahas pendugaan parameter total populasi, rata-rata populasi, dan proporsi populasi. Selain itu akan dibahas juga penarikan SAB dua tahap mengenai pendugaan untuk nilai total populasi, nilai rata-rata populasi, dan nilai proporsi populasi.

4.1. Pendugaan Nilai Total Populasi

Andaikan peneliti ingin mengetahui nilai total variabel tertentu dalam suatu populasi berdasarkan tujuan penelitian, (seperti menduga jumlah penduduk Yogya yang transmigrasi ke P. Sulawesi, menduga total penggunaan waktu per minggu oleh rumah tangga di daerah tertentu dalam hal menonton TV). Untuk menjawab masalah tersebut, maka dilakukan pendugaan total populasi yang dilambangkan dengan τ. Total populasi diperoleh dengan mengalikan ukuran populasi (N) dengan rata-rata populasi, μ, jadi $\tau = N \cdot \mu$.

DEFINISI 4.1.

Penduga bagi total populasi, τ, dengan metode pelapisan didefinisikan sebagai berikut: $\hat{\tau}_{st} = \sum_{h=1}^{L} N_h \overline{y}_h$

dimana $\hat{\tau}_{st}$ adalah penduga total populasi dengan metode pelapisan.

Jumlah populasi disimbolkan dengan τ, dimana populasi terbagi menjadi L lapisan, dan total tiap lapisan disombolkan dengan τ_h. Jadi $\tau = \sum_{h=1}^{L} \tau_h$ (jumlah total populasi merupakan jumlah total untuk tiap lapisan). Karena peneliti tidak tahu pasti total populasi, maka total populasi di duga oleh $\hat{\tau}_{st}$, dan τ_h diduga oleh $\hat{\tau}_h = N_h \bar{y}_h$ sehingga $\sum \hat{\tau}_h = \hat{\tau}_{st}$. Rumus-rumus yang berhubungan dengan pendugaan nilai total populasi diberikan dalam teorema berikut:

TEOREMA 4.1.1.

Andaikan suatu populasi yang terdiri dari N elemen dibagi ke dalam L lapisan, dan lapisan ke-h terdiri dari N_h elemen dengan suatu total nilai τ_h untuk variabel y. Pemilihan sampel dalam setiap lapisan independen dengan pemilihan dari lapisan lain. Misalkan $\hat{\tau}_h$ adalah penduga tak bias dari τ_h, yang didasarkan pada ukuran sampel (n_h) yang diambil dari lapisan. Jika $\hat{V}(\hat{\tau}_h)$ adalah penduga tak bias dari $V(\hat{\tau}_h)$, maka

$$\hat{\tau}_{st} = \sum \hat{\tau}_h \qquad V(\hat{\tau}_{st}) = \sum V(\hat{\tau}_h) \triangleq \sum \hat{V}(\hat{\tau}_h)$$

<u>Bukti :</u>

<u>Diketahui</u> :

Populasi berukuran N dibagi dalam L lapisan, dan tiap lapisan berukuran N_h dan total tiap lapisan adalah τ_h .

$\hat{\tau}_h$ adalah penduga tak bias dari τ_h maka E ($\hat{\tau}_h$) = τ_h

$\hat{V}(\hat{\tau}_h)$ adalah penduga tak bias dari $V(\hat{\tau}_h)$ maka E ($\hat{V}(\hat{\tau}_h)$)= $V(\hat{\tau}_h)$

Akan dibuktikan :

1. $\hat{\tau} = \sum \hat{\tau}_h$

2. $V(\hat{\tau}) = \sum V(\hat{\tau})_h \triangleq \hat{V}(\hat{\tau}_h)$

Bukti :

1. Karena $E(\hat{\tau}_h) = \tau_h$ maka menurut sifat nilai harapan variabel yang independen $E(\sum E(\hat{\tau}_h)) = \sum E(\hat{\tau}_h) = \sum \tau_h = \tau$ (Berdasarkan pokok bahasan Nilai Harapan pada bab 2). Karena τ_h tidak diketahui maka nilai total variabel y diduga oleh $\hat{\tau}_h$.

   Jadi $\Sigma\ \hat{\tau}_h = \hat{\tau}$ penduga tak bias dari τ.

2. Dengan argumen yang sama, menurut sifat variansi variabel random yang independen maka $V(\hat{\tau}) = V((\sum \hat{\tau}_h) = \sum V(\hat{\tau}_h)$

Teorema tersebut menyatakan bahwa dugaan total populasi merupakan jumlah dari dugaan setiap total lapisan. Begitu juga dalam pendugaan variansi-variansi. Jadi tidak ada prinsip-prinsip baru yang dilibatkan dalam analisis data yang dihasilkan , yaitu bahwa masalah pendugaan dapat diselesaikan dalam lapisan.

Telah dibahas dalam bab 3 tentang variansi dugaan dalam pengambilan SAS karena dalam lapisan sampel diambil secara acak sederhana maka akan dibuktikan teorema berikut :

TEOREMA 4.1.2.

Jika $\hat{\tau}_{st} = N\bar{y}_{st}$ adalah penduga tak bias dari total populasi τ, maka

$$V(\hat{\tau}_{st}) = \sum_{h=1}^{L} N_h(N_h - n_h)\frac{\sigma_h^2}{n_h}$$

Bukti : $V(\hat{\tau}_{st}) = V(N\,\bar{y}_{st})$

$= V(\Sigma N_h \bar{y}_h)$

$= \Sigma V(N_h \bar{y}_h)$

$= \Sigma N_h^2 V(\bar{y}_h)$

Berdasarkan variasi penduga y pada SAS maka $V(\bar{y}) = \frac{\sigma^2}{n} \frac{N-n}{N}$ sehingga

$$V(\bar{y}_{st}) = \Sigma N_h^2 \frac{\sigma_h^2}{n_h} \frac{N_h - n_h}{N_h}$$

$$= \Sigma N_h (N_h - n_h) \frac{\sigma_h^2}{n_h}$$

Berdasarkan teorema 4.1.2. maka akar pangkat dua dari $\hat{V}(\hat{\tau}_{st})$ merupakan galat baku dugaan dari total SAB, $s(\hat{\tau}_{st}) = s(N\bar{y}_{st})$. Andaikan peneliti menggunakan tingkat kesalahan sebesar α, dimana $0<\alpha<1$, maka dapat ditentukan batas kesalahan sampling nilai total dapat dinyatakan dalam persamaan sebagai berikut: $KS(N\bar{y}_{st}) = KS(\hat{\tau}_{st}) = z_{\alpha/2}\, s(\hat{\tau}_{st})$

Jika diketahui bentuk pernyataan peluang $(1-\alpha)$, maka KS nilai total dapat dinyatakan dalam persamaan sebgai berikut : $P\{G(\hat{\tau}_{st}) < KS(\hat{\tau}_{st})\} = 1 - \alpha$

Dengan selang kepercayaan $(1 - \alpha)$ 100% untuk nilai total populasi, τ, didapat persamaan sebagai berikut: $P\{\hat{\tau}_{st} - KS(\hat{\tau}_{st}) < \tau < \hat{\tau}_{st} + KS(\hat{\tau}_{st})\} = 1 - \alpha$.

Untuk menentukan kesalahan sampling tersebut diatas digunakan tabel kurva normal,

CONTOH 4.1. Pendugaan Total Populasi

Dengan menggunakan contoh 3.8.1. akan dihitung penduga total populasi, yaitu penduga total banyak anak di Mrican pada tahun 1997. Populasi dibagi dalam 5 lapisan. Lapisan 1 adalah wanita yang mengikuti program KB dengan menggunakan alat, lapisan 2 adalah wanita yang mengikuti program KB dengan metode siblings, lapisan 3 adalah tanpa perhitungan, lapisan 4 adalah wanita yang menderita sakit tertentu sehingga susah mempunyai anak, serta lapisan 5 adalah lain-lain. Data diambil pada 31 Desembar 1997 (akhir tahun). Dari contoh 3.8.1. telah diketahui ukuran sampel tiap lapisan, yaitu $n_1 = 34$, $n_2 = 10$, $n_3 = 8$, $n_4 = 6$, $n_5 = 8$

Andaikan setelah dilakukan survey diperoleh data banyaknya anak sebagai berikut:

Lapisan 1	Lapisan 2	Lapisan 3	Lapisan 4	Lapisan 5
5,4,1,2,3,2,2,3,1,1,5,2, 1,2,3,3,2,2,4,5,5,1,2,3, 1,2,1,2,3,4,1,1,2,2	6,6,4,5,1,2,2, 3,4,5	4,2,3,5,4,6,7, 8	1,1,2,1,3,1	2,4,5,1,5,7,3,4,

Maka rata-rata dan variansi sampel banyak anak setiap lapisan adalah :

$\bar{y}_1 = 2{,}44$	$\bar{y}_2 = 3{,}8$	$\bar{y}_3 = 4{,}87$	$\bar{y}_4 = 1{,}5$	$\bar{y}_5 = 3{,}88$
$S_1^2 = 1{,}69$	$S_2^2 = 3{,}07$	$S_3^2 = 4{,}13$	$S_4^2 = 0{,}7$	$S_5^2 = 3.55$

Jadi total banyak anak di Mrican adalah:

$$\hat{\tau}_{st} = \sum N_h \bar{y}_h = (448.2{,}44 + 131.3{,}8 + 100.4{,}87 + 81.1{,}5 + 108.3{,}88)$$

$$= 2618{,}46$$

Variasi dugaan dari $\hat{\tau}_{st}$ adalah:

$$\hat{V}(\hat{\tau}_{st}) = \sum_{h=1}^{L} N_h^2 \frac{S_h^2}{n_h}\left(\frac{N_h - n_h}{N_h}\right)$$

$$= \left(448^2 \frac{(1{,}69)}{34}\left(\frac{448-34}{448}\right) + 131^2 \frac{(3{,}07)}{10}\left(\frac{131-10}{131}\right) + \right.$$

$$\left. 100^2 \frac{(4{,}13)}{8}\left(\frac{100-8}{100}\right) + 81^2 \frac{(0{,}7)}{6}\left(\frac{81-6}{81}\right) + 108^2 \frac{(3{,}55)}{8}\left(\frac{108-8}{108}\right)\right)$$

$$= 24336{,}057$$

$$s(\hat{\tau}_{st}) = \sqrt{24336{,}057} = 156$$

Dengan menggunakan tingkat kepercayaan (1-α)100% = 90%, maka dapat ditentukan batas kesalahan penarikan sampel nilai total, KS $(\hat{\tau}_{st})$, sebagai berikut :

$KS(\hat{\tau}_{st}) = z_{\alpha/2}s(\hat{\tau}_{st}) = 1{,}645.156 = 256{,}62$

Dengan diketahuinya batas kesalahan sampling nilai total populasi, maka dapat dibuat pendugaan selang bagi nilai total populasi, dengan taraf kepercayaan 90% sebagai berikut:

$$P\{\hat{\tau}_{st} - KS(\hat{\tau}_{st}) < \tau < \hat{\tau}_{st} + KS(\hat{\tau}_{st})\} = 1 - \alpha$$

$$P\{2618{,}46 - 256{,}62 < \tau < 2618{,}46 + 256{,}62\} = 0{,}90$$

$$P\{2361{,}84 < \tau < 2875{,}08\} = 0{,}90$$

Berdasarkan hasil hitungan di atas maka dapat ditarik kesimpulan:

1. Total banyak anak di Mrican pada tahun 1997 diduga 2619 anak.
2. Dengan menggunakan taraf kepercayaan 90% peneliti yakin bahwa pendugaan total yang dibuat tidak akan menyimpang lebih besar dari 257 anak.
3. Dengan menggunakan taraf kepercayaan 90% yakin bahwa pada selang nilai [2361 , 2876] akan memuat total anak di Mrican pada tahun 1997.

Pada pokok bahasan tersebut peneliti tidak hanya dapat mengetahui nilai dugaan total populasi tetapi dapat juga mengetahui dugaan total banyak anak tiap lapisan. Misalkan peneliti ingin mengetahui total banyak anak pada lapisan 1.

Pada contoh 4.1 telah diketahui $N_1 = 448$, $n_1 = 34$, $\bar{y}_1 = 2{,}44$, $S_1^2 = 1{,}69$, dan karena sampel pada lapisan 1 diambil secara acak maka perhitungan nilai total tersebut menggunakan perhitungan pada penarikan SAS.

$$\hat{\tau}_1 = N_1\bar{y}_1 = 448.2{,}44 = 1093{,}12 \text{, dan}$$

$$\hat{V}(\hat{\tau}_1) = \hat{V}(N_1\bar{y}_1)$$

$$= N_1^2 . \frac{S_1^2}{n_1}\left(\frac{N_1 - n_1}{N_1}\right)$$

$$= (448)^2 . \frac{1{,}69}{34}\left(\frac{448 - 34}{448}\right) = 9219{,}049$$

$$s(\hat{\tau}_1) \quad = 96{,}02$$

Dengan menggunakan taraf kepercayaan 90% maka batas kesalahan sampling total lapisan ke-1 adalah sebagai berikut:

$$KS(\hat{\tau}_1) = z_{\alpha/2}\, s(\hat{\tau}_1) = 1{,}645 \,.\, 96{,}02 = 157{,}946$$

Pendugaan selang untuk nilai total lapisan ke-1, dengan taraf kepercayaan 90% adalah sebagai berikut:

$$P\,\{(\hat{\tau}_1) - KS(\hat{\tau}_1) < \tau_1 < \hat{\tau}_1 + KS(\hat{\tau}_1)\} = 1 - \alpha$$

$$P\,\{1093{,}12 - 157{,}946 < \tau_1 < 1093{,}12 + 157{,}946\} = 0{,}90$$

$$P\,\{935{,}174 < \tau_1 < 1251{,}066\} = 0{,}90$$

Berdasarkan hasil di atas dapat ditarik kesimpulan untuk lapisan ke-1 adalah sebagai berikut:

1. Total anak dalam lapisan 1 diduga 1094 anak.
2. Dengan menggunakan taraf kepercayaan 90% yakin bahwa penduga total lapisan 1 yang diperoleh tidak akan menyimpang lebih besar dari pada 158 anak.
3. Dengan menggunakan taraf kepercayaan 90% diyakini bahwa selang nilai [935 , 1252] akan memuat total jumlah anak yang sebenarnya dalam lapisan 1.

4.2. Pendugaan Nilai Rata-rata Populasi dengan SAB

Pada pokok bahasan ini akan dibahas bagaimana menduga rata-rata populasi dengan menggunakan SAB. Misalkan ada suatu populasi berukuran N dibagi menjadi L lapisan dengan masing-masing lapisan berukuran N_1, N_2 ..... N_L. Kemudian dari setiap lapisan ditarik sampel berukuran n_1, n_2,.....n_L sedemikian sehingga $n_1 + n_2 + ... + n_L = n$ yang merupakan ukuran sampel acak berlapis. Untuk tiap n_i dapat dihitung nilai rata-rata sampel dan variansi sampel untuk setiap lapisan.

Berdasarkan hasil hitungan tersebut maka akan diperoleh nilai-nilai $\bar{y}_1, \bar{y}_2, \ldots\ldots, \bar{y}_L$ yang merupakan nilai rata-rata sampel untuk lapisan 1, 2, ..., L, dan nilai variansi sampel S_1^2, S_2^2, ..., S_L^2 dari lapisan 1, 2, ..., L. Secara singkat dapat dikatakan bahwa berdasarkan pengamatan yang dilakukan atas dasar SAB, akan diperoleh nilai rata-rata sampel $\bar{y}_h$ untuk lapisan ke-h (h = 1, 2, ..., L), serta diperoleh juga nilai variansi sampel S_h^2 untuk lapisan ke-h (h = 1, 2,....,L). Berdasarkan nilai-nilai yang diperoleh dari SAB maka akan diduga nilai rata-rata populasi μ. Pendugaan untuk nilai rata-rata populasi merupakan akibat dari pendugaan untuk nilai total populasi.

AKIBAT 4.1.

Pendugaan untuk nilai rata-rata populasi, μ, menggunakan nilai rata-rata SAB, $\bar{y}_{st}$, diduga dalam rumus berikut:

$$\hat{\mu} = \bar{y}_{st} = \frac{1}{N}\left(N_1\bar{y}_1 + N_2\bar{y}_2 + \ldots\ldots + N_L\bar{y}_L\right)$$

$$= \frac{1}{N}\sum_{h=1}^{L} N_h\bar{y}_h \quad = \sum W_h\bar{y}_h$$

Variansi dugaan dari rata-rata sampel acak berlapis diberikan dalam teorema berikut:

TEOREMA 4.2.1.

Variansi dugaan dari nilai rata-rata SAB, $V(\bar{y}_{st})$ adalah:

$$V(\bar{y}_{st}) = \sum_{h=1}^{L} W_h^2 V(\bar{y}_h)$$

<u>Bukti</u>: Diketahui akibat 4.1. bahwa

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

Jadi $V(\bar{y}_{st}) = \sum_{h=1}^{L} W_h^2 V(\bar{y}_h)$

Dari penarikan SAS sudah didefinisikan bahwa:

$$V(\bar{y}_h) = \frac{\sigma_h^{\ 2}}{n_h} \frac{N_h - n_h}{N_h}$$

Maka $V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^{L} N_h^2 V(\bar{y}_h)$

$$= \frac{1}{N^2} \sum_{h=1}^{L} N_h^2 \frac{\sigma_h^2}{n_h} \frac{N_h - n_h}{N_h}$$

$$= \sum W_h^2 \frac{\sigma_h^2}{n_h} (1 - f_h)$$

Berdasarkan teorema 4.2.1. dapat dihitung simpangan baku dugaan dari nilai rata-rata SAB, $s(\bar{y}_{st})$ sebagai akar pangkat dua dari $\hat{V}(\bar{y}_{st})$. Dari simpangan baku tersebut dapat diperkirakan besarnya kesalahan sampling rata-rata, G $(\bar{y}_{st}) = |\bar{y}_{st} - \mu|$, yang dapat dinyatakan dalam bentuk pernyataan peluang (1-α), seperti dalam persamaan berikut: $P\{G(\bar{y}_{st}) < z_{\alpha/2} S(\bar{y}_{st})\} = 1 - \alpha$

Dengan demikian batas kesalahan pendugaan nilai rata-rata populasi berdasarkan nilai rata-rata sampel acak berlapis, $KS(\bar{y}_{st})$, pada taraf kepercayaan (1-α) 100% adalah: $KS(\bar{y}_{st}) = z_{\alpha/2}\, s(\bar{y}_{st})$.

Selang kepercayaan (1-α) 100% untuk nilai rata-rata populasi, μ, diberikan dalam persamaan berikut: $P\{\bar{y}_{st} - KS(\bar{y}_{st}) < \mu < \bar{y}_{st} + KS(\bar{y}_{st})\} = 1-\alpha$

CONTOH 4.2. Pendugaan nilai rata-rata populasi.

Contoh seperti pada contoh 4.1. Peneliti ingin mengetahui rata-rata banyaknya anak dalam suatu keluarga di Mrican pada tahun 1997. Diketahui data dan hasil hitungan pada contoh 4.1. Rata-rata sampel dengan metode pelapisan adalah

$$\bar{y}_{st} = \frac{1}{N}\sum_{h=1}^{4} N_h \bar{y}_h$$

$$= \frac{1}{868}\left((448.2{,}44)+(131.3{,}8)+(100.4{,}87)+(81.1{,}5)+(108.3{,}88)\right)=3{,}02$$

Variansi dugaan untuk nilai rata-rata $\hat{V}(\bar{y}_{st})$ adalah:

$$\hat{V}(\bar{y}_{st}) = \frac{1}{N^2}\sum N_h^2 \frac{S_h^2}{n_h}\left(\frac{N_h - n_h}{N_h}\right)$$

$$= \frac{1}{(868)^2}\left\{(448)^2\frac{1{,}69}{34}\left(\frac{448-34}{448}\right)+(131)^2\frac{3{,}07}{10}\left(\frac{131-10}{131}\right)+\right.$$

$$\left.(100)^2\frac{4{,}13}{8}\left(\frac{100-8}{100}\right)+(81)^2\frac{0{,}7}{6}\left(\frac{81-6}{81}\right)+(108)^2\frac{3{,}55}{8}\left(\frac{108-8}{108}\right)\right\}$$

$$= 0{,}033$$

$$s(\bar{y}_{st}) = \sqrt{\hat{V}(\bar{y}_{st})} = 0{,}182$$

Dengan menggunakan taraf kepercayaan 90% maka batas kesalahan pendugaan nilai rata-rata populasi, $KS(\bar{y}_{st})$ adalah:

$$KS(\bar{y}_{st}) = z_{\alpha/2} s(\bar{y}_{st}) = 1{,}645 \,.\, 0{,}182 = 0{,}29$$

Dengan diketahuinya batas kesalahan sampling nilai rata-rata populasi $KS(\bar{y}_{st})$, maka dapat dibuat pendugaan selang bagi rata-rata populasi dengan taraf kepercayaan 90%, sebagai berikut:

$$P(\bar{y}_{st} - KS(\bar{y}_{st}) < \mu < \bar{y}_{st} + KS(\bar{y}_{st})) = 1\text{-}\alpha$$

$$P(3{,}02\text{-}0{,}29 < \mu < 3{,}02 + 0{,}29) = 0{,}90$$

$$P(2{,}73 < \mu < 3{,}31) = 0{,}90$$

Berdasarkan berbagai perhitungan di atas, dapat ditarik kesimpulan tentang banyaknya anak dalam suatu keluarga di Mrican, sebagai berikut:

1. Rata-rata banyak anak dalam suatu keluarga di Mrican pada tahun 1997 adalah 2 sampi dengan 4 anak.
2. Dengan menggunakan taraf kepercayaan 90% diyakini bahwa pendugaan rata-rata populasi tidak akan menyimpang lebih besar dari pada 1 anak.
3. Dengan menggunakan taraf kepercayaan 90% diyakini bahwa selang nilai [2 , 4] akan mencakup rata-rata banyaknya anak dalam suatu keluarga di Mrican pada tahun 1997.

Seringkali dalam penelitian yang menggunakan teknik penarikan SAB disamping ingin menarik kesimpulan tentang sifat populasi secara keseluruhan, juga ingin menarik kesimpulan tentang lapisan tertentu. Misalkan peneliti ingin mengetahui lebih jelas tentang lapisan 2.

Berdasarkan perhitungan awal, didapat hasil pendugaan rata-rata dan variansi pada lapisan 2 sebagai berikut: $\bar{y}_2 = 3{,}8$, $S_2^2 = 3{,}07$.

Sudah dijelaskan pada awal pembahasan tentang teknik penarikan SAB, bahwa penarikan sampel untuk tiap lapisan dilakukan secara acak sederhana. Jadi setiap sampel yang ditarik dari lapisan dapat dianggap sebagai sampel acak sederhana.

Berikut ini akan dibahas sejauh mana keandalan nilai dugaan tersebut, melalui penentuan $\hat{V}(\bar{y}_2)$.

$$\hat{V}(\bar{y}_2) = \frac{S_2^2}{n_2}\left(\frac{N_2 - n_2}{N_2}\right) = \frac{3{,}07}{10}.\left(\frac{131-10}{131}\right) = 0{,}28$$

$$s(\bar{y}_2) = \sqrt{0{,}28} = 0{,}53$$

Dengan menggunakan taraf kepercayaan 90%, maka batas kesalahan sampling nilai rata-rata lapisan ke-2, yaitu:

$$KS(\bar{y}_2) = z_{\alpha/2} s(\bar{y}_2) = (1{,}645)\,(0{,}53) = 0{,}88$$

Penduga selang bagi nilai rata-rata lapisan ke-2, dengan taraf kepercayaan 90%, adalah sebagai berikut:

$$P(\bar{y}_2 - KS(\bar{y}_2) < \mu_2 < \bar{y}_2 + KS(\bar{y}_2)) = 0{,}90$$

$$P(3{,}8 - 0{,}88 < \mu_2 < 3{,}8 + 0{,}88) = 0{,}90$$

$$P(2{,}92 < \mu_2 < 4{,}68) = 0{,}90$$

Berdasarkan perhitungan tersebut dapat ditarik kesimpulan tentang jumlah anak untuk lapisan ke-2, sebagai berikut:

1. Rata-rata banyaknya anak dalam keluarga pada lapisan ke-2 adalah 3-4 anak.

2. Dengan menggunakan taraf kepercayaan 90% diyakini bahwa dugaan terhadap nilai rata-rata lapisan yang sesungguhnya mempunyai penyimpangan tidak lebih dari 1 anak.
3. Dengan menggunakan taraf kepercayaan 90% diyakini bahwa selang nilai [2,5] anak akan mencakup rata-rata jumlah anak yang sesungguhnya dalam keluarga pada lapisan ke-2.

Keandalan pendugaan akan semakin tinggi apabila batas kesalahan sampling relatif kecil dibandingkan nilai penduga bagi parameter tersebut, atau selang kepercayaan bagi nilai parameter tidak terlalu lebar.

Dengan dimungkinkannya menentukan pendugaan pada tiap lapisan maka berdasarkan teknik penarikan SAB, tampak bahwa peneliti memperoleh informasi yang lebih banyak dibandingkan teknik penarikan SAS. Teknik penarikan SAB dapat menghasilkan kesimpulan tentang populasi secara keseluruhan, maupun kesimpulan tentang sifat setiap lapisan serta dapat memeriksa sejauh mana keandalan pendugaan yang dibuat baik menyangkut populasi maupun menyangkut setiap lapisan dalam populasi tersebut.

4.3. Pendugaan Proporsi Populasi

Konsep proporsi sering dijumpai dalam kehidupan sehari-hari, misalnya berapa proporsi masyarakat yang mengikuti program KB, proporsi penganggur, dll. Pada pokok bahasan ini, akan dibahas pendugaan porporsi atau prosentase populasi yang memiliki karakteristik tertentu. Jadi akan dibahas proporsi populasi P, variansi dugaan dari P_{st} adalah $\hat{V}(P_{st})$, kesalahan sampling nilai proporsi $KS(P_{st})$, serta selang kepercayaan $(1-\alpha)$ 100% untuk parameter proporsi populasi.

Misalkan suatu populasi berukuran N, dibagi dalam L lapisan dengan masing-masing lapisan berukuran $N_1, N_2, \ldots, N_L$. Kemudian dari setiap lapisan diambil SAS berturut-turut sebanyak $n_1, n_2, \ldots, n_L$. Dari sampel tersebut dihitung nilai proporsi sampel $p_1, p_2, \ldots, p_L$. Berdasarkan SAB berukuran n, $n = n_1 + n_2 + \ldots + n_L$, dapat dilakukan pendugaan nilai proporsi populasi P berdasarkan nilai proporsi SAB, P_{st}.

DEFINISI 4.3. Penduga Proporsi Populasi

Penduga untuk proporsi populasi P dengan SAB didefinisikan dalam rumus sebagai berikut: $p_{st} = \frac{1}{N}(N_1 p_1 + N_2 p_2 + \ldots\ldots + N_L p_L) = \frac{1}{N}\sum_{h=1}^{L} N_h p_h$

TEOREMA 4.3. Variansi dugaan dari p_{st}, yaitu

$\hat{V}(p_{st}) = \frac{1}{N^2}\sum N_h^2 \frac{p_h(1-p_h)}{n_h - 1}\left(\frac{N_h - n_h}{N_h}\right)$ dengan simpangan baku dugaan nilai proporsi sampel $s(p_{st})$ adalah $s(p_{st}) = \sqrt{\hat{V}(p_{st})}$.

Bukti : Bukti ini adalah kasus khusus dari teorema umum untuk varians dan pendugaan rata-rata. Dari teorema 4.2.1. $V(\bar{y}st) = \frac{1}{N^2}\sum N_h (N_h - n_h)\frac{\sigma_h^2}{n_h}$

Dari variansi proporsi sampel didapat $\sigma_h^2 = \frac{N_h}{N_h - 1} P_h(1-P_h)$

$$\text{Jadi } V(\hat{p}st) = \frac{1}{N^2}\sum N_h (N_h - n_h)\frac{N_h}{N_h - 1}\frac{P_h(1-P_h)}{n_h}$$

$$= \sum W_h^2 \frac{P_h(1-P_h)}{n_h}(1-f_h)$$

Peneliti dapat menentukan batas kesalahan sampling, dengan menggunakan tingkat kesalahan sebesar α. Batas kesalahan sampling ditentukan sebagai berikut:

$$KS(p_{st}) = z_{\alpha/2}\, s(p_{st})$$

Selang kepercayaan (1-α) 100% untuk parameter proporsi populasi P, digunakan dalam persamaan berikut: $P\{p_{st} - KS(p_{st}) < P < p_{st} + KS(p_{st})\} = 1-\alpha$

CONTOH 4.3. Pendugaan Proporsi Populasi

Seorang peneliti ingin mengetahui proporsi keluarga di DIY yang menyekolahkan anaknya sampai ke Perguruan Tinggi (PT) pada tahun 1996. Berdasarkan skala kemampuan ekonomi keluarga yang bermacam-macam, maka populasi dibagi dalam 3 lapisan yaitu keluarga kaya, keluarga sedang, keluarga miskin. Konsep garis kemiskinan yang digunakan mengacu pada kebutuhan minimum 2100 kalori perkapita per hari ditambah dengan kebutuhan minimum non makanan yang merupakan kebutuhan dasar seseorang. Besarnya nilai pengeluaran (dalam rupiah) untuk memenuhi kebutuhan dasar minimum makanan dan non makanan tersebut disebut garis kemiskinan. Penduduk yang tidak mampu memenuhi kebutuhan dasar minimum dikategorikan sebagai penduduk miskin. Pada tahun 1996, diketahui batas garis kemiskinan adalah Rp 35.841,00 per kapita per bulan. Berdasarkan data Susenas 1996, secara umum pengeluaran rata-rata per kapita sebulan masyarakat DIY sebesar Rp 75.609,00. Sedangkan golongan dalam keluarga dapat dilihat sebagai berikut :

- Golongan miskin : bila pengeluaran per kapita sebulan < Rp 39.999,00
- Golongan sedang : bila pengeluaran per kapita sebulan Rp 40.000,00-Rp149.999,00
- Golongan kaya : pengeluaran per kapita sebulan > Rp 150.000,00.

Andaikan diketahui di DIY terdapat 900 keluarga, dengan perincian: keluarga kaya sebanyak N_1 = 150 keluarga, keluarga sedang sebanyak N_2 = 550 keluarga, dan keluarga miskin sebanyak N_3 = 200 keluarga. Dengan menggunakan teknik penarikan SAB, ditentukan ukuran sampel untuk tiap lapisan. Metode yang digunakan adalah Metode Alokasi Neyman. Peneliti menggunakan tingkat kepercayaan 90% berarti $Z_{tabel} = 1{,}645$ dan kesalahan sampling tidak lebih besar dari 0,05 berarti G = 0,05.

Jadi $$n = \frac{Z^2\left(\sum_{h=1}^{L} N_h \sqrt{P_h(1-P_h)}^2\right)}{N^2G^2 + Z^2\sum_{h=1}^{L} N_h P_h(1-P_h)}$$

$$n = \frac{(1{,}645)^2\left(150\sqrt{0{,}8.0{,}2} + 550\sqrt{0{,}6.0{,}4} + 200\sqrt{0{,}65.0{,}35}\right)^2}{900^2.0{,}05^2 + (1{,}645)^2(150.0{,}8.0{,}2 + 550.0{,}6.0{,}4 + 200.0{,}65.0{,}35)}$$

$$= \frac{488398{,}705}{205 + 545{,}259} = 190{,}019 \approx 191 \text{ (pembulatan)}$$

Sampel yang harus diambil adalah 191 keluarga, dengan alokasi sampel untuk setiap lapisan n_h adalah sebagai berikut:

$$n_h = \frac{N_h\sqrt{P_h(1-P_h)}}{\sum_{h=1}^{L} N_h\sqrt{P_h(1-P_h)}} n; h = 1,2,3$$

$$n_1 = \frac{150\sqrt{0{,}6.0{,}4}}{424{,}8379}.191 = 121{,}14 \approx 27$$

$$n_2 = \frac{550\sqrt{0{,}6.0{,}4}}{424{,}8379}.191 = 121{,}14 \approx 121$$

$$n_3 = \frac{200\sqrt{0{,}65.0{,}35}}{424{,}8379}191 = 42{,}89 \approx 43$$

Berdasarkan hasil hitungan di atas dan penelitian yang dilakukan oleh peneliti, didapat data sebagai berikut:

Lap (h)	N_h	P_h	n_h	Banyak keluarga yang menyekolahkan	p_h
1	150	0,80	27	21	0,80

2	550	0,60	121	73	0,60
3	200	0,35	43	28	0,65

Berdasarkan data dalam tabel di atas maka dapat dilakukan pendugaan proporsi populasi berdasarkan proporsi SAB, sebagai berikut:

$$p_{st} = \frac{1}{N}\sum_{h=1}^{L} N_h p_h$$

$$= \frac{1}{900}(150.0{,}8 + 550.0{,}6 + 200.0{,}65) \qquad = 0{,}64$$

Dan variansi dugaan dari proporsi SAB; $\hat{V}(p_{st})$ adalah:

$$\hat{V}(p_{st}) = \frac{1}{N^2}\sum_{h=1}^{L} N_h^2 \frac{p_h(1-p_h)}{n_h - 1}\left(\frac{N_h - n_h}{N_h}\right)$$

$$= \frac{1}{900^2}\left\{\frac{(150)^2(0{,}16)(0{,}82)}{26} + \frac{(550)^2(0{,}24)(0{,}78)}{120} + \frac{(200)^2(0{,}23)(0{,}79)}{42}\right\}$$

$$= 0{,}0009$$

$$s(p_{st}) = \sqrt{0{,}0009} = 0{,}03$$

Dengan menggunakan taraf kepercayaan 90% dapat ditentukan batas kesalahan sampling proporsi, KS(p_{st}), sebagai berikut:

$$KS(p_{st}) = z_{\alpha/2} s(p_{st}) = 1{,}645 \;.\; 0{,}03 \;=\; 0{,}05$$

Selang kepercayaan 90% untuk nilai parameter proporsi P, dapat dinyatakan sesuai dengan persamaan berikut:

$$P(p_{st} - KS(p_{st}) < P < p_{st} + KS(p_{st})) = 0{,}90$$

$$P(0{,}64 - 0{,}05 < P < 0{,}64 + 0{,}05) = 0{,}90$$

$$P(0{,}59 < P < 0{,}69) = 0{,}90$$

Berdasarkan hasil di atas dapat ditarik kesimpulan sebagai berikut:

1. Proporsi keluarga di DIY yang menyekolahkan anaknya sampai ke Perguruan Tinggi adalah 0,64.
2. Dengan menggunakan taraf kepercayaan 90% yakin bahwa dugaan nilai proporsi tidak akan menyimpang lebih besar daripada 0,05.
3. Dengan menggunakan taraf kepercayaan 90% yakin bahwa selang nilai (0,59, 0,69) akan mencakup nilai proporsi yang sebenarnya dari keluarga di DIY yang menyekolahkan anaknya sampai ke Perguruan Tinggi.

Di samping perhitungan di atas, dapat juga dihitung proporsi tiap lapisan. Misalkan ingin mengetahui proporsi pada lapisan ke-3. Dari tabel diketahui bahwa $N_3 = 200$, $n_3 = 20$, $p_3 = 0,65$. Karena dalam teknik penarikan SAB, sampel yang diambil dari setiap lapisan berdasarkan teknik penarikan SAS, maka sampel yang terpilih dari setiap lapisan merupakan SAS. Jadi setiap pendugaan parameter lapisan tertentu menggunakan rumus dalam teknik penarikan SAS. Berdasarkan data yang ada diketahui bahwa nilai sampel untuk lapisan ke-3 adalah $p_3 = 0,65$, jadi proporsi keluarga pada lapisan ke-3 yang menyekolahkan anaknya sampai PT adalah 65%. Selanjutnya akan dilihat sejauh mana keandalan pendugaan tersebut.

Variansi dugaan dari nilai proporsi sampel untuk lapisan ke-3 adalah:

$$\hat{V}(p_3) = \frac{p_3(1-p_3)}{n_3 - 1}\left(\frac{N_3 - n_3}{N_3}\right)$$

$$= \frac{0,65(1-0,65)}{43-1}\left(\frac{200-43}{200}\right) = 0,004$$

$$s(p_3) = \sqrt{0,004} = 0,065$$

Batas kesalahan sampling nilai proporsi, KS(P_3), adalah:

$$KS(p_3) = z_{\alpha/2} s(p_3) \text{ (taraf kepercayaan yang digunakan 90\%)}$$

$$= 1{,}645 \; . \; 0{,}065 \; = 0{,}107$$

Dengan demikian selang kepercayaan 90% bagi nilai parameter proporsi lapisan ke-3, P_3 adalah:

$$P\{p_3 - KS(p_3) < P_3 < p_3 + KS(p_3)\} = 1\text{-}\alpha$$

$$P\{0{,}65 - 0{,}107 < P_3 < 0{,}65 + 0{,}107\} = 0{,}90$$

$$P\{0{,}543 < P_3 < 0{,}757\} = 0{,}90$$

Dari perhitungan di atas, dapat ditarik kesimpulan sebagai berikut:

1. Proporsi keluarga miskin yang menyekolahkan anaknya sampai PT adalah 65%.
2. Dengan menggunakan taraf kepercayaan 90% yakin bahwa pendugaan yang dilakukan tidak akan menyimpang lebih besar dari 10,7%.
3. Dengan menggunakan taraf kepercayaan 90% yakin bahwa selang nilai (0,543, 0,757) akan mencakup nilai parameter proporsi keluarga miskin sesungguhnya yang menyekolahkan anaknya sampai PT.

Selain masalah-masalah di atas, peneliti dapat juga mengetahui berapa banyak keluarga di daerah DIY yang menyekolahkan anaknya sampai PT. Jadi akan dihitung total banyaknya keluarga dalam populasi yang menyekolahkan anak sampai PT, pada lapisan tertentu.

Misalkan banyaknya keluarga yang menyekolahkan anaknya sampai PT dilambangkan dengan A. Jadi penduga untuk total populasi A adalah jumlah total populasi dikalikan dengan penduga proporsi yang memiliki karakteristik A:

$$\hat{A}_{st} = Np_{st}$$

$$= (900) \,.\, 0,64 = 576$$

Variansi dugaan dari $\hat{A}_{st}$, $(\hat{V}(\hat{A}_{st}))$ adalah:

$$\hat{V}(\hat{A}_{st}) = \hat{V}(Np_{st}) = N^2\hat{V}(p_{st}) = \sum_{h=1}^{L} N_h^2 \frac{P_h(1-P_h)}{n_h - 1}\left(\frac{N_h - n_h}{N_h}\right)$$

$$\hat{V}(\hat{A}_{st}) = 756,618$$

dan $\quad s(\hat{A}_{st}) = \sqrt{756,618} = 27,507$

Dengan menggunakan taraf kepercayaan (1-α) 100% maka batas kesalahan sampling nilai total, $KS(\hat{A}_{st})$ adalah:

$$KS = z_{\alpha/2} . S(\hat{A}_{st}) = 1,645 \,.\, 27,507 = 45,25$$

Jadi $G(\hat{A}_{st}) = |\hat{A}_{st} - A|$ dapat dinyatakan dalam persamaan berikut:

$$P\{G(\hat{A}_{st}) < KS(\hat{A}_{st})\} = 1 - \alpha$$

$$P\{G(\hat{A}_{st}) < 45,25\} = 0,90$$

Selang kepercayaan (1-α) 100% untuk nilai parameter total populasi A, ditulis dalam persamaan berikut:

$$P\{\hat{A}_{st} - KS(\hat{A}_{st}) < A < \hat{A}_{st} + KS(\hat{A}_{st})\} = 1 - \alpha$$

$$P\{576 - 45,25 < A < 576 + 45,25\} = 0,90$$

$$P\{530,75 < A < 621,25\} = 0,90$$

$$P\{530 < A < 622\} = 0,90$$

Dari perhitungan di atas dapat diambil kesimpulan sebagai berikut:

1. Total banyaknya keluarga di DIY yang menyekolahkan anaknya sampai ke PT sebanyak 576 keluarga.
2. Dengan menggunakan taraf kepercayaan 90% diyakini bahwa pendugaan yang dilakukan tidak akan menyimpang lebih dari 46 keluarga.
3. Dengan menggunakan taraf kepercayaan 90% diyakini bahwa selang nilai [530, 622] akan mancakup total banyaknya keluarga (nilai yang sesungguhnya) di DIY yang menyekolahkan anaknya ke-PT.

Berdasarkan perhitungan dan kesimpulan di atas, dapat juga dihitung pendugaan total banyaknya keluarga dalam tiap lapisan yang menyekolahkan anaknya sampai ke PT. Misalkan ingin mengadakan pendugaan total banyaknya keluarga dalam lapisan ke-3 yang menyekolahkan anaknya sampai ke PT. Telah diketahui bahwa $N_3 = 200$, $n_3 = 20$, $p_3 = 0{,}65$, dan $\hat{V}(p_3) = 0{,}004$; $S(P_3) = 0{,}065$

Dari data tersebut maka dapat dilakukan pendugaan:

$$\hat{A} = N_3 p_3 = 200 \,.\, 0{,}65 = 130$$

Variansi dugaan dari $\hat{A}_3$ adalah:

$$\hat{V}(\hat{A}_3) = \hat{V}(N_3 p_3) = N_3^2 \hat{V}(p_3)$$

$$= (200)^2 \,(0{,}004) = 160$$

$$s(\hat{A}_3) = \sqrt{160} = 12{,}65$$

Dengan menggunakan taraf kepercayaan 90% maka dapat ditentukan batas galat pendugaan nilai total, sebagai berikut:

$$KS(\hat{A}_3) = z_{\alpha/2} . s(\hat{A}_3) = 1{,}645 \,.\, 12{,}65 = 20{,}81$$

Jadi selang kepercayaan 90% untuk nilai parameter A_3 adalah:

$$P\{\hat{A}_3 - KS(\hat{A}_3) < A_3 < \hat{A}_3 KS(\hat{A}_3)\} = 1-\alpha$$

$$P\{130 - 20{,}81 < A_3 < 130 + 20{,}81\} = 0{,}90$$

$$P\{109{,}19 < A_3 < 150{,}81\} = 0{,}90$$

Dari hasil hitungan di atas dapat diambil kesimpulan sebagai berikut:

1. Total banyaknya keluarga miskin yang menyekolahkan anaknya sampai ke PT adalah 130 keluarga.
2. Dengan menggunakan taraf kepercayaan 90% yakin bahwa pendugaan yang dilakukan tidak akan menyimpang lebih besar dari 21 keluarga.
3. Dengan menggunakan taraf kepercayaan 90% yakin bahwa selang nilai (109; 151) akan mencakup nilai sebenarnya dari total banyaknya keluarga miskin yang menyekolahkan anaknya sampai ke PT.

4.4.Ketelitian Relatif pada Penarikan SAB dan Penarikan SAS

Dalam pokok bahasan ini akan dibahas perbandingan ketelitian pendugaan parameter antara penarikan SAS dan penarikan SAB dengan alokasi proporsional dan alokasi optimum. Perbandingan tersebut dapat menunjukkan bagaimana manfaat pelapisan dapat diperoleh. Variansi dari pendugaan rata-rata untuk berbagai metode alokasi sampel berturut-turut dinotasikan dengan Vran, Vprop, Vopt.

Dimana Vran = variansi dari pendugaan rata-rata pada SAS, Vprop = variansi pendugarata-rata pada alokasi proporsional, Vopt = Variansi penduga rata-rata pada alokasi optimum.

TEOREMA 4.4.

Jika $^1/_{N_h}$ diabaikan dan n tetap dengan alokasi optimum, maka

Vopt < Vprop <Vran

Bukti : Telah diketahui bahwa

Variansi pada SAS

$$V_{ran} = (1-f)\frac{\sigma^2}{n} \quad \ldots \ (1)$$

Dari teorema 4.3.

$$V_{prop} = \frac{(1-f)}{n}\sum W_h \sigma_h^2$$
$$= \frac{\sum W_h \sigma_h^2}{n} - \frac{\sum W_h \sigma_h^2}{N} \quad \ldots \text{(teorema 4.2.1.) (2)}$$

Dalam penarikan SAB $V(\bar{y}_{st})$ diminimumkan untuk total ukuran sampel n yang tetap jika $n_h = n\frac{W_h \sigma_h}{\sum W_h \sigma_h} = n\frac{N_h \sigma_h}{\sum N_h \sigma_h}$. Suatu rumus untuk variansi minimum dengan n tetap diperoleh dengan menggantikan nilai n_h ke dalam rumus (2) untuk $V(\bar{y}_{st})$ maka diperoleh $V_{opt} = \frac{(\sum W_h \sigma_h)^2}{n} - \frac{\sum W_h \sigma_h^2}{N} \quad \ldots (3)$

Dalam persamman aljabar baku untuk analisis variansi (Cochran,1991) populasi berlapis, diperoleh :

$$(N-1)\sigma^2 = \sum_h \sum_i (y_{hi} - \mu)^2 \quad (4)$$
$$= \sum_h \sum_i [(\mu_h - \mu) + (y_{hi} - \mu_h)]^2$$
$$= \sum_h \sum_i [(\mu_h - \mu)^2 + 2(\mu_h - \mu)(y_{hi} - \mu_h) + (y_{hi} - \mu_h)^2]$$
$$= \sum_h \sum_i (\mu_h - \mu)^2 + 2\sum_h \sum_i (\mu_h - \mu)(y_{hi} - \mu_h) + \sum_h \sum_i (y_{hi} - \mu_h)^2$$

Suku yang ditengah sama dengan nol, karena

$$\sum_i \left(y_{hi} - \mu_h\right) = \sum y_{hi} - N\mu_h = \sum y_{hi} - N\frac{\sum y_{hi}}{N} = 0$$

Suku yang pertama tidak mengandung indeks i, jadi dapat ditulis

$$\sum_h \sum_i (\mu_h - \mu)^2 = N\sum_h \left(\mu_h - \mu\right)^2$$

sehingga

$$(N-1)\sigma^2 = \sum_h N_h \left(\mu_h - \mu\right)^2 + \sum_h \sum_i \left(y_{hi} - \mu_h\right)^2$$

$$(N-1)\sigma^2 = \sum_h N_h \left(\mu_h - \mu\right)^2 + \sum_h (N_h - 1)\sigma_h^2 \quad \ldots(5)$$

Jika $^1/_{N_h}$ dan $^1/_N$ diabaikan, maka (5) diperoleh

$$\sigma^2 = \sum W_h \sigma_h^2 + \sum W_h \left(\mu_h - \mu\right)^2 \quad \ldots(6)$$

*) Akan dibuktikan Vran > Vprop

<u>Bukti</u> :

Dari (1)

$$\begin{aligned} Vran &= (1-f)\frac{\sigma^2}{n} \\ &= (1-f)\left(\frac{\sum W_h \sigma_h^2 + \sum W_h \left(\mu_h - \mu\right)^2}{n}\right) \\ &= \frac{(1-f)}{n}\sum W_h \sigma_h^2 + \frac{(1-f)}{n}\sum W_h \left(\mu_h - \mu\right)^2 \\ &= Vprop + \frac{(1-f)}{n}\sum W_h \left(\mu_h - \mu\right)^2 \quad \ldots(7) \end{aligned}$$

terbukti bahwa Vran > Vprop

**) Akan dibuktikan Vprop > Vopt

<u>Bukti :</u>

Dari (2) dan (3) perbedaannya adalah

$$Vprop - Vopt = \frac{1}{n}\left[\sum W_h \sigma_h^2 - \left(\sum W_h \sigma_h\right)^2\right]$$

$$= \frac{1}{n}\left[\sum W_h \sigma_h^2 - 2\sum W_h \sigma_h \sum W_h \sigma_h + \left(\sum W_h \sigma_h\right)^2\right]$$

$$= \frac{1}{n}\left[\sum W_h \left(\sigma_h - \bar{\sigma}\right)^2\right]$$

dan andaikan $\bar{\sigma} = \sum W_h \sigma_h$ adalah rata-rata tertimbang dari σ_h

$$Vprop = \frac{1}{n}\left[\sum W_h \left(\sigma_h - \bar{\sigma}\right)^2\right] + Vopt \ldots (8)$$

terbukti bahwa Vprop > Vopt karena suku pertama ruas kanan selalu positif.

***) Akan dibuktikan Vran > Vopt

Bukti :

Dari (7) dan (8) diperoleh :

$$Vran = Vopt + \frac{1}{n}\sum W_h \left(\sigma_h - \bar{\sigma}\right)^2 + \frac{(1-f)}{n}\sum W_h \left(\mu_h - \mu\right)^2$$

terbukti bahwa Vran > Vopt

Jadi kesimpulannya dari *, **, *** terbukti Vopt < Vprop < Vran.

Jika $^1/_{N_h}$ tidak diabaikan, pengganti untuk σ^2 dari (6) menjadi

$$\sigma^2 = \frac{\sum_h (N_h - 1)\sigma_h^2 + \sum N_h \left(\mu_h - \mu\right)^2}{N - 1} \qquad \ldots (9)$$

Dari (9) $\sum W_h \left(\mu_h - \mu\right)^2 = \sigma^2 - \sum W_h \sigma_h^2$

Berdasarkan persamaan (9) maka persamaan (7) menjadi

$$Vran = Vprop + \frac{(1-f)}{n(N-1)}\left[\sum N_h \left(\mu_h - \mu\right)^2 - \frac{1}{N}\sum (N - N_h)\sigma_h^2\right] \quad \ldots (12)$$

Hal tersebut berarti bahwa jika $\sum N_h(\mu_h - \mu)^2 < \frac{1}{N}\sum(N - N_h)\sigma_h^2$

Maka pelapisan proporsional memberikan variansi yang lebih tinggi dari pada penarikan SAS.

CONTOH 4.4

Contoh seperti pada contoh 4.2.. Diketahui data dan hasil penelitian sebagai berikut :

$$N = 868, \quad n = \frac{NZ^2\sigma^2}{NG^2 + Z^2\sigma^2} = \frac{868(1,645)^2 27}{868(1)^2 + (1,645)^2 27} = 67,39 \approx 68 \; ;$$

$Z = 1,645, \; \sigma^2 = 27$

Data-data tersebut adalah (dalam satuan anak):

6,6,1,2,2,6,7,3,4,5,5,2,1,7,7,2,3,4,1,1,5,1,2,3,3,3,6,7,6,2,1,3,6,8,2,1,1,3,4,4,5,2,1,5,4,5 ,3,2,3,4,5,2,3,4,8,8,9,3,6,2,5,6,6,3,2,1,1,2,3,3,4,5,1,2,2,3,3,4,4,5,5,6,7,3,2,4,4,5,6,6,7, 5,4,3,2,1,1,2,1,2,3,4,4,5,1,1,2,2,3,3,1,3,4,5,6,1,1,1,2,4,7,1,2,3,3,4,5.

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{256}{68} = 3,76$$

$$S^2 = \frac{\sum_{i=1}^{n}(y_i - y)^2}{n-1} = \frac{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}{n(n-1)} = \frac{68.1276 - 65536}{68.67} = 4,66$$

$$\hat{V}(\bar{y}) = \frac{s^2}{n} = 0,068$$

$$s_{\bar{y}} = \sqrt{\hat{V}(\bar{y})} = 0,26$$

$$P(\bar{y} - t_{\alpha/2;n-1}s_{\bar{y}} < \mu < \bar{y} + t_{\alpha/2;n-1}s_{\bar{y}}) = 1 - \alpha$$

$$P(3,76 - 1,64.0,26 < \mu < 3,76 + 1,64.0,26) = 0,90$$

$$P(3,33 < \mu < 4,186) = 0,90$$

Berdasarkan contoh 4.2. dan contoh 4.4. tentang pendugaan banyaknya anak dalam satu keluarga dapat dilihat adanya perbedaan hasil diantara keduanya. Taraf kepercayaan yang digunakan dalam pendugaan $\hat{V}(\bar{y}_{st})$ dalam SAB dan pendugaan $\hat{V}(\bar{y})$ dalam SAS sama yaitu 90% tetapi nilai dugaan $\hat{V}(\bar{y})$ dalam SAS lebih besar dari nilai dugaan $\hat{V}(\bar{y}_{st})$ dalam SAB. Dalam SAS $\hat{V}(\bar{y})$ = 0,068 dan selang kepercayaan untuk μ adalah P $(3,33 < \mu < 4,186) = 0,90$. Dalam SAB, didapat $\hat{V}(\bar{y}_{st}) = 0,033$ dan selang kepercayaan untuk μ adalah $P(2,73 < \mu < 3,31) = 0,90$. Selang ini lebih sempit dari pada selang yang diperoleh dengan SAS. Hal ini berarti pendugaan pada SAB lebih baik daripada pada SAS. Dalam arti bahwa $\hat{V}(\bar{y}_{st})$ dalam SAB menyebabkan keandalan pendugaan yang lebih tinggi untuk menduga $V(\bar{y}_{st})$ dibandingkan keandalan pendugaan $\hat{V}(\bar{y}_{st})$ dalam SAS untuk menduga $V(\bar{y}_{st})$, dengan kata lain ketelitian relatif $R(\bar{y}_{st}, \bar{y}) = \dfrac{V(\bar{y}_{st})}{V(\bar{y})} < 1$. Jadi jika dihubungkan dengan pokok bahsan 4.4. maka jelas bahwa Vopt < Vran.

4.5. Sub Penarikan SAB

Dasar pemikiran yang melatar belakangi sehingga dirancang penarikan sampel dua tahap atau penarikan sampel berlapis dua tahap adalah bahwa pada umumnya suatu lapisan terdiri atas banyak sekali elemen sehingga sulit untuk melakukan pengukuran pada setiap elemen tersebut, atau lapisan tersebut terdiri atas elemen-elemen yang pada dasarnya memiliki karakteristik yang sama sehingga pengukuran tidak perlu dilakukan terhadap setiap elemen dalam lapisan tetapi cukup pada beberapa elemen dalam lapisan tersebut. Alasan tersebut berkaitan dengan prinsip efisiensi dalam penarikan sampel, artinya agar suatu teknik penarikan sampel tertentu mampu memberikan informasi lebih banyak pada tingkat biaya tertentu, sehingga apabila suatu teknik penarikan sampel mampu mengurangi ongkos tanpa mengurangi informasi yang diperoleh, maka teknik penarikan sampel tersebut yang paling baik untuk dipergunakan. Pembahasan pada pokok bahasan ini lebih memfokuskan pada penarikan sampel acak sederhana, semata-mata karena alasan bahwa biasanya dalam praktek survei teknik tersebut yang paling banyak diterapkan. Misalkan bahwa setiap unit dalam populasi dibagi dalam beberapa lapisan yang disebut lapisan utama. Didalam penarikan SAB tiap lapisan diambil sampel secara acak. Sampel yang sudah terpilih tersebut dibagi lagi dalam beberapa sub lapisan. Dari setiap sub lapisan tersebut diambil sampel 2 tahap atau lebih. Teknik penarikan sampel tersebut disebut penarikan sampel dua tahap. Karena sampelnya diambil melalui dua tahap maka sampel yang kedua disebut sub penarikan sampel. Tahap pertama memilih sampel dari unit-unit utama dan tahap kedua memilih sampel dari unit-unit tahap kedua atau sub unit dari setiap unit utama yang terpilih. Teknik penarikan sampel lebih dari 2 tahap hanyalah merupakan pengembangan dari teknik penarikan sampel 2 tahap. Pembahasan dibatasi sampai dengan penarikan sampel 2 tahap.

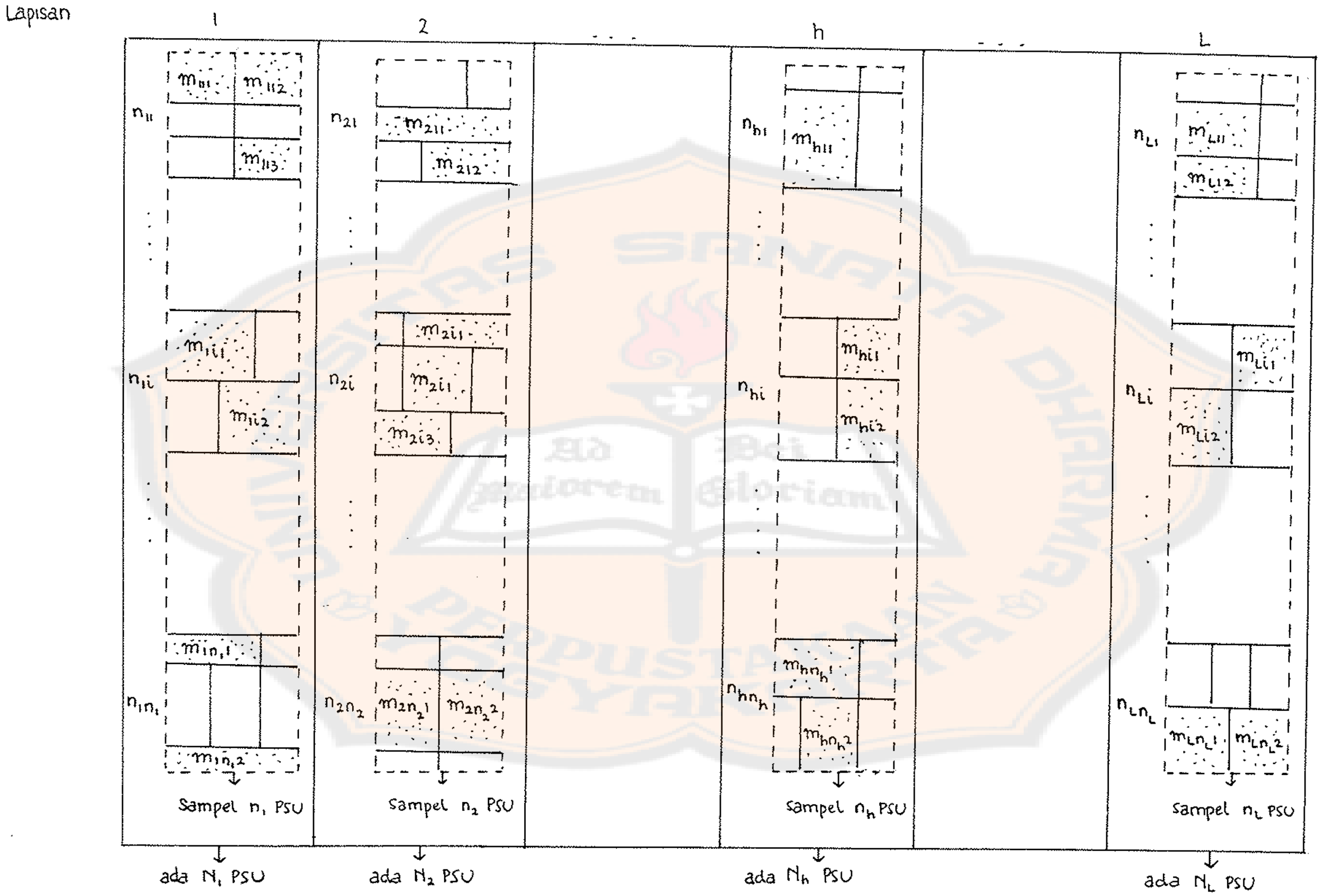
Keuntungan utama dari penarikan sampel dua tahap adalah bahwa cara tersebut lebih fleksibel dari pada penarikan sampel satu tahap. Artinya peneliti mempunyai kesempatan untuk mengambil nilai dalam sub lapisan yang lebih kecil sehingga bisa meningkatkan efisiensi statistik sampel. Sub Unit yang terbentuk mempunyai sifat yang lebih homogen. Jadi variansinya kecil dan berakibat nilai atau kesimpulan yang didapat akan semakin mendekati nilai yang sesungguhnya.

4.5.1. Prosedur Penarikan Sampel

Prosedur penarikan sampel dua tahap adalah seagai berikut:

1. Tentukan populasi dengan jelas
2. Populasi dibagi dalam L lapisan
3. Tiap lapisan mempunyai N_h unit utama atau Primari Unit Sampling (PSU) (h =1,2,…,L)
4. Dari N_h unit utama (PSU) dalam masing-masing lapisan tersebut diambil sampel acak sederhana berukuran n_h (h =1,2,…,L)
5. Dari setiap unit utama yang terpilih, masing-masing mempunyai M_{hi} elemen (i =1,2,…n_h). Dari M_{hi} elemen tersebut diambil sampel acak berukuran m_{hi} (i =1,2,…n_h) yang merupakan unit sekunder atau Secundary Sampling Unit (SSU).

Dalam penarikan sampel dua tahap, peneliti memerlukan 2 kerangka penarikan sampel yaitu tahap pertama adalah kerangka penarikan sampel unit utama (misalnya daftar semua perguruan tinggi di wilayah tertentu), sedangkan tahap kedua adalah kerangka penarikan sampel unit sekunder (misalnya daftar nama-nama mahasiswa dari setiap perguruan tinggi yang terpilih pada pemilihan sampel tahap pertama). Prosedur tersebut dapat digambarkan dalam diagram berikut :


Lapisan
1
2
h
L
n_{11}
m_{111}
m_{112}
m_{113}
n_{1i}
m_{1i1}
m_{1i2}
n_{1n_1}
m_{1n_11}
m_{1n_12}
n_{21}
m_{211}
m_{212}
n_{2i}
m_{2i1}
m_{2i2}
m_{2i3}
n_{2n_2}
m_{2n_21}
m_{2n_22}
n_{h1}
m_{h11}
n_{hi}
m_{hi1}
m_{hi2}
n_{hn_h}
m_{hn_h1}
m_{hn_h2}
n_{L1}
m_{L11}
m_{L12}
n_{Li}
m_{Li1}
m_{Li2}
n_{Ln_L}
m_{Ln_L1}
m_{Ln_L2}
Sampel n_1 PSU
Sampel n_2 PSU
Sampel n_h PSU
Sampel n_L PSU
ada N_1 PSU
ada N_2 PSU
ada N_h PSU
ada N_L PSU

Pada sub penarikan SAB terdapat 2 situasi, yaitu sub penarikan sampel dengan unit berukuran sama dan sub penarikan sampel dengan unit berukuran tidak sama. Akan dijelaskan satu persatu mengenai pendugaan pada kedua situasi tersebut.

4.5.2. Sub Penarikan SAB dengan unit-unit berukuran sama

Prosedur penarikan sampel pada pembahasan ini sama seperti prosedur penarikan sampel yang dijelaskan pada 4.5.1. Hanya saja jumlah elemen dalam sampel PSU berukuran sama, begitu juga jumlah elemen dalam sampel SSU berukuran sama. Selanjutnya akan dijelaskan variansi dari pendugaan rata-rata penarikan sampel dua tahap.

4.5.2.1. Penduga Rata-rata Populasi dan Variansinya Pada Penarikan Sampel duaTahap

Berikut ini akan dibahas bagaimana menentukan penduga rata-rata populasi dan variansinya. Pada penarikan sampel dua tahap, rencana penarikan sampelnya pertama memberikan sebuah metode pemilihan n unit. Kemudian untuk setiap unit terpilih, diberikan metode untuk memilih sejumlah tertentu subunit-subunit. Dalam mencari rata-rata dan variansi pendugaan, rata-ratanya harus meliputi seluruh sampel yang diturunkan dengan proses dua tahap tersebut. Untuk sebuah pendugaan $E(\hat{\theta}) = E_1[E_2(\hat{\theta})]$, dimana E menyatakan nilai harapan atau rata-rata seluruh sampel, E_2 menyatakan rata-rata seluruh pemilihan yang mungkin pada tahap kedua dari sekumpulan n unit tetap, dan E_1 menyatakan rata-rata seluruh pemilihan pada tahap pertama.

$$V(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = E_1E_2(\hat{\theta} - \theta)^2$$

Tetapi, $E_2(\hat{\theta}-\theta)^2 = E_2(\hat{\theta}^2) - 2\theta E_2(\hat{\theta}) + \hat{\theta}^2$

$$= [E_2(\hat{\theta})]^2 + V_2(\hat{\theta}) - 2\theta E_2(\hat{\theta}) + \theta^2$$

Rata-rata sekarang seluruh pemilihan tahap pertama. Karena $E_1E_2(\hat{\theta}) = \theta$

$$V(\hat{\theta}) = E1[E2(\hat{\theta})]^2 - \theta^2 + E_1[V_2(\hat{\theta})]$$

$$= V_1[E_2(\hat{\theta})] + E_1[V_2(\hat{\theta})]$$

Sebelum mendefinisikan penduga rata-rata populasi dan variansinya akan didefinisikan dulu rata-rata sampel untuk tiap unit dan sub unit.

y_{hij} = elemen ke-j dalam unit utama ke-i pada lapisan ke-h

m_{hi} = banyaknya elemen sampel dalam unit utama ke-i pada lapisan ke-h

$\bar{y}_{hi} = \dfrac{\sum_{j=1}^{m_{hi}} y_{hij}}{m_{hi}}$ = rata-rata sampel untuk unit utama i lapisan ke-h sebagai penduga dari rata-rata unit ke-i lapisan ke-h (μ_{hi})

n_h = banyaknya unit utama yang terambil secara acak pada lapisan ke-h

$\bar{\bar{y}}_{st}$ = rata-rata sampel untuk unit tahap kedua sebagai penduga rata-rata populasi keseluruhan μ

$\bar{\bar{y}}_h = \dfrac{\sum_{i=1}^{n_h} \bar{y}_{hi}}{n_h}$ = rata-rata sampel untuk lapisan ke-h sebagai penduga rata-rata lapisan ke-h (μ_h)

DEFINISI 4.5.2.1. Penduga rata-rata

Andaikan suatu populasi terdiri dari L lapisan. Lapisan ke-h terdiri atas N_h unit-unit utama (PSU) masing-masing dengan M_h unit tahap kedua (SSU), ukuran sampel yang bersesuaian adalah n_h dan m_h.

Penduga rata-rata populasi per unit tahap kedua adalah:

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

$W_n = \frac{N_h M_h}{\sum N_h M_h}$ adalah ukuran relatif dari lapisan pada unit tahap kedua.

$\bar{\bar{y}}_h$ = rata-rata sampel dalam lapisannya.

TEOREMA 4.5.2.1.1. Variansi penduga rata-rata

Bila n_h unit utama dan m_h elemen dari masing-masing unit yang telah diambil dipilih dengan penarikan SAS, maka variansi dalam masing-masing lapisan adalah:

$$V\left(\bar{\bar{y}}_{st}\right) = \sum_h W_h^2 \left(\frac{1-f_{1h}}{N_h} \sigma_{1h}^2 + \frac{1-f_{2h}}{n_h m_h} \sigma_{2h}^2 \right)$$

$$f_{1h} = \frac{n_h}{N_h}, \quad f_{2h} = \frac{m_h}{M_h}$$

$$\sigma_{1h}^2 = \frac{\sum_{i=1}^{N_h} (\mu_{hi} - \mu_h)^2}{N_h - 1}$$ = variansi pada lapisan ke-h(variansi padaPSU)

$$\sigma_{2h}^2 = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{M_{hi}} (y_{hij} - \mu_{hi})^2}{N_h (N_h - 1)}$$ variansi pada sub lapisan dalam lapisan ke-h

(variansi pada SSU)

Bukti:

Akan ditunjukkan terlebih dahulu bahwa $\overline{\overline{y}}_h$ adalah penduga tak bias dari μ_h

Dengan penarikan sampel acak sederhana pada kedua tahap.

$$E(\overline{\overline{y}}_h) = E_1[E_2(\overline{\overline{y}}_h)] = E_1\left(\sum_{i=1}^{n_h} \frac{\mu_{hi}}{n_h}\right) = \left(\sum_{i=1}^{N_h} \frac{\mu_{hi}}{N_h}\right) = \mu_h ,$$

Untuk $V(\overline{\overline{y}}_h)$, digunakan rumus sebagai berikut :

$V(\overline{\overline{y}}_h) = V_1[E_2(\overline{\overline{y}}_h)] + E_1[V_2(\overline{\overline{y}}_h)]$,.....4.5.2.1.1.a

Dimana $V_2(\overline{\overline{y}}_h)$ = variansi seluruh sub sampel pilihan yang mungkin untuk sekumpulan unit-unit tertentu.

Karena $E_2(\overline{\overline{y}}_h) = \sum_{i=1}^{n_h} \mu_{hi} \Big/ n_h$, suku pertama pada ruas kanan adalah variansi dari rata-rata per sub unit untuk sebuah sampel acak sederhana satu tahap berukuran n_h unit. Oleh karena itu, dari variansi random

$V_1[E_2(\overline{\overline{y}}_h)] = \left(\frac{N_h - n_h}{N_h}\right)\frac{\sigma_{1h}^2}{n_h}$,....4.5.2.1.1.b

Selanjutnya, dengan $\overline{\overline{y}}_h = \sum_{i=1}^{n_h} \overline{y}_{hi} / n_h$ dan penarikan SAS digunakan pada tahap kedua,

$$V_2(\overline{\overline{y}}_h) = \frac{M_h - m_h}{M_h n_h^2} \frac{\sum_{i=1}^{n_h} \sigma_{2hi}^2}{m_h}$$

dimana $\sigma_{2hi}^2 = \sum_{j=1}^{M_h} (y_{hij} - \mu_{hi})^2 \Big/ (M_h - 1)$ adalah variansi diantara sub unit utama ke-i.

Bila dirata-ratakan seluruh sampel tahap pertama, $\sum_{i=1}^{n_h} \sigma_{2hi}^2 \Big/ N_h = \sigma_{2h}^2$

Oleh karena itu, $E_1[V_2(\bar{\bar{y}}_h)] = \left(\frac{M_h - m_h}{M_h}\right)\frac{\sigma_{2h}^2}{m_h n_h}$,.... 4.5.2.1.1.c

Dari 4.5.2.1.1.a didapat dengan penjumlahan 4.5.2.1.1.b dan 4.5.2.1.1.c

$$V(\bar{\bar{y}}_h) = \frac{(N_h - n_h)}{N_h}\frac{\sigma_{1h}^2}{n_h} + \left(\frac{M_h - m_h}{M_h}\right)\frac{\sigma_{2h}^2}{m_h n_h}$$

Telah didefinisikan bahwa $V(\bar{\bar{y}}_{st}) = \sum_{h=1}^{L} W_h^2 V(\bar{\bar{y}}_h)$

$$\text{Jadi } V(\bar{\bar{y}}_{st}) = \sum_{h=1}^{L} W_h^2\left(\frac{N_h - n_h}{N_h}\frac{\sigma_{1h}^2}{n_h} + \frac{M_h - m_h}{M_h}\frac{\sigma_{2h}^2}{m_h n_h}\right)$$

$$= \sum_{h=1}^{L} W_h^2\left(\frac{1 - f_{1h}}{n_h}\sigma_{1h}^2 + \frac{1 - f_{2h}}{n_h m_h}\sigma_{2h}^2\right)$$

dimana $f_{1h} = \frac{n_h}{N_h}$ dan $f_{2h} = \frac{m_h}{M_h}$ merupakan fraksi penarikan sampel tahap1 dan tahap2.

TEOREMA 4.5.2.1.2.

Penduga yang tak bias dari $V\left(\bar{\bar{y}}_{st}\right)$ adalah:

$$v\left(\bar{\bar{y}}_{st}\right) = \sum_h W_h^{\ 2}\left[\frac{1 - f_{1h}}{n_h}S_{1h}^{\ 2} + \frac{f_{1h}(1 - f_{2h})}{n_h m_h}S_{2h}^{\ 2}\right]$$

$$S_{1h}^2 = \frac{\sum_{i=1}^{n_h}\left(\bar{y}_{hi} - \bar{\bar{y}}_h\right)^2}{n_h - 1}, \quad S_{2h}^2 = \frac{\sum_{i=1}^{n_h}\sum_{j=1}^{m_h}\left(y_{hij} - \bar{y}_{hi}\right)^2}{n_h(m_h - 1)}$$

Bukti:

Berdasarkan rumus analisis variansi sampel

$$(n_h - 1)S_{1h}^2 = \sum_{i=1}^{n_h}\left(y_{hi} - \bar{\bar{y}}_h\right)^2 = \sum_{i=1}^{n_h}\bar{y}_i^2 - n\,\bar{\bar{y}}_h^2$$

Oleh karena itu

$$(n_h-1)E_2(S_{1h}^2)=\sum_{i=1}^{n_h}\mu_{hi}^2+\sum_{i=1}^{n_h}\frac{(1-f_{2h})}{m_h}\sigma_{2hi}^2-n\mu_{n_h}^2-\frac{1}{n_h}\sum_{i=1}^{n_h}\frac{(1-f_{2h})}{m_h}\sigma_{2hi}^2$$

dimana $\mu_{n_h}=\sum_{i=1}^{n_h}\mu_{hi}/n_h$. Suku terakhir pada ruas kanan terpenuhi karena sub penarikan sampel adalah bebas dalam unit-unit yang berbeda. Maka,

$$(n_h-1)E_2(S_{1h}^2)=\sum_{i=1}^{n_h}\left(\mu_{hi}-\mu_{n_h}\right)^2+\frac{(n_h-1)(1-f_{2h})}{n_h m_h}\sum_{i=1}^{n_h}\sigma_{2hi}^2$$

Kalikan dengan $(1\text{-}f_{1h})/n_h(n_h\text{-}1)$ dan rata-ratakan seluruh penarikan SAS tahap pertama, $E\frac{(1-f_{1h})}{n_h}S_{1h}^2=\frac{(1-f_{1h})}{n_h}\sigma_{1h}^2+\frac{(1-f_{1h})(1-f_{2h})}{m_h n_h}\sigma_{2h}^2$

Dengan membandingkan hasil (teorema 4.5.2.1.1.) untuk $V(\bar{\bar{y}}_{st})$, perhatikan bahwa σ_{2h}^2 terlalu kecil dengan $f_{1h}(1\text{-}f_{2h})\sigma_{2h}^2/m_h n_h$. Karena $E_1E_2(s_{2h}^2)=\sigma_{2h}^2$ maka penduga tak bias dari $V(\bar{\bar{y}}_{st})$ adalah $v(\bar{\bar{y}}_{st})=\sum_{h=1}^{L}W_h^2\left[\frac{1-f_{1h}}{n_h}s_{1h}^2+\frac{f_{1h}(1-f_{2h})}{n_h m_h}s_{2h}^2\right]$

CONTOH 4.5.2.

Contoh ini diambil dari penelitian Wayan Ardhana, dkk (dalam Majalah Analisis Pendidikan, 13 April 1985). Penelitian tersebut bertujuan mengetahui adakah pengaruh usia masuk sekolah terhadap kemajuan belajar anak-anak di SD. Kemajuan belajar diukur dengan menggunakan nilai rata-rata lima bidang studi Akademik (PMP, Bahasa Indonesia, Matematika, IPA,IPS) yang diperoleh dari catur wulan III dikelas VI, sedangkan kecerdasan di ukur dengan menggunakan tes "matriks progresif" yang dikembangkan oleh Raven. Skor yang diperoleh dari tes tersebut kemudian ditransformasikan ke dalam nilai IQ setelah mempertimbangkan faktor usia murid. Populasi sasarannya adalah SD Negeri di Jawa Timur. Pengambilan sampel dilakukan secara bertahap, karena luas daerah populasi yang sangat luas dengan keadaan lingkungan yang berbeda. Peneliti hanya memusatkan daerah penelitian pada daerah bekas karesidenan, dan bukan wilayah kabupaten, semata-mata didasarkan pada pertimbangan praktis. Disamping itu, wilayah bekas karesidenan dianggap telah mencerminkan keadaan geografis di Jatim, yang mewakili daerah surplus, juga yang mewakili daerah minus. Pada tingkat pertama daerah populasi yang disampelkan dibagi dalam 7 wilayah, yaitu (1) Madiun,(2) Kediri, (3) Bojonegoro, (4) Madura, (5) Surabaya, (6) Malang,(7) Besuki.

Pada penelitian ini diperoleh data rata-rata IQ dan rata-rata kemampuan belajar untuk anak usia dibawah 7 tahun dan anak usia diatas 7 tahun. Disamping itu analisis datanya menggunakan uji hipotesis, dan pada skripsi ini penulis tidak membahas uji hipotesis. Berdasarkan pokok bahasan 4.5. ini, maka penulis hanya menitikberatkan pada pengambilan sampel yang telah dilakukan oleh Wayan Ardhana, dkk. Sedangkan data yang digunakan untuk contoh 4.5.2. adalah rata-rata IQ dan rata-rata kemampuan belajar (K-bel) untuk usia dibawah 7 tahun.

Didapat data-data sebagai berikut (dari tahap 1)

Wilayah h	banyaknya SD N_h	ukuran sampel n_h
1	25	4
2	21	4
3	17	4
4	18	4
5	28	4
6	20	4
7	15	4

n_h merupakan banyaknya sampel SD yang diambil secara acak dari tiap lapisan N_h, n_h juga sekaligus sebagai banyaknya sub lapisan (lapisan pada tahap 2).

Didapat data-data sebagai berikut: (dari tahap 2)

h	n_h	M_{hi}	m_{hi}	Rata-rata IQ	K-Bel
1	SD Klegen I	200	40	98,11	72,09
	SD Winongo I	240	40	88,36	90,00
	SD Babadan I	180	40	67,41	71,15
	SD Mejaran I	200	40	96,57	85,07
		$\Sigma M_{hi} = 820$		$\bar{\bar{y}}_h = 87{,}61$	79,58
2	SD Banjaran IV	250	40	69,01	81,15
	SD Ngadirejo I	150	40	82,79	99,23
	SD Keras I	160	40	77,57	71,58
	SD Purwodadi	180	40	80,35	80,00
		$\Sigma M_{hi} = 740$		$\bar{\bar{y}}_h = 77{,}43$	82,99
3	SD Bojonegoro I	230	40	92,94	63,38
	SD Kadipaten I	210	40	93,63	71,13
	SD Kaliditu I	200	40	95,54	75,61
	SD Sumberejo I	210	40	93,97	74,52
		$\Sigma M_{hi} = 850$		$\bar{\bar{y}}_h = 94{,}02$	71,16
4	SD Barurambat Barat Kota V	190	40	88,33	67,58
	SD Bugih II	200	40	77,29	59,83
	SD Pakong I	180	40	82,43	60,76
	SD Bandaran	175	40	70,51	65,51
		$\Sigma M_{hi} = 745$		$\bar{\bar{y}}_h = 79{,}64$	63,42
5	SD Kusumabangsa	280	40	103,00	68,60
	SD Tembokdukuh	275	40	82,75	63,80

	SD Cerme Lor I	250	40	98,75	75,17
	SD Cerme Kidul I	260	40	102,51	73,63
		$\Sigma M_{hi} = 1065$		$\bar{\bar{y}}_h = 96,75$	70,3
6	SD Bareng IV	240	40	110,55	72,40
	SD Sukorejo I	250	40	105,45	66,30
	SD Lawang V	230	40	111,00	70,76
	SD Ngaglik I	255	40	98,91	69,14
		$\Sigma M_{hi} = 975$		$\bar{\bar{y}}_h = 106,48$	69,65
7	SD Dabasah V	160	40	104,88	70,00
	SD Bundungan I	170	40	83,25	66,55
	SD Pakisan I	200	40	103,96	71,51
	SD Sukosari I	190	40	95,78	72,26
		$\Sigma M_{hi} = 720$		$\bar{\bar{y}}_h = 96,97$	70,07

M_{hi} merupakan banyaknya murid pada SD ke-i dalam daerah/wilayah h

m_{hi} merupakan sampel yang diambil dari M_{hi}

M_h merupakan banyaknya murid pada wilayah-h

$$\bar{\bar{y}}_{st1} = \frac{\sum N_h M_h \bar{\bar{y}}_h}{\sum N_h M_h}$$

$$= \frac{(25\cdot 820\cdot 87{,}61)+(21\cdot 740\cdot 77{,}43)+(17\cdot 850\cdot 94{,}02)+(18\cdot 745\cdot 79{,}64)}{(25\cdot 820)+(21\cdot 740)+(17\cdot 850)+(15\cdot 720)}$$

$$\frac{+(28\cdot 1065\cdot 96{,}75)+(20\cdot 975\cdot 106{,}48)+(15\cdot 720\cdot 71{,}16)}{+(28\cdot 820)+(21\cdot 41)+(17\cdot 850)}$$

$$= \frac{11434549{,}6}{124020} = 92{,}199$$

$$\bar{\bar{y}}_{st2} = \frac{(20\cdot 820\cdot 79{,}58)+(21\cdot 740\cdot 82{,}99)+(17\cdot 850\cdot 71{,}16)+(18\cdot 745\cdot 63{,}42)+}{(25\cdot 850)+(21\cdot 740)+(17\cdot 850)+(18\cdot 745)+}$$

$$\frac{(28\cdot 1065\cdot 70{,}3)+(20\cdot 975\cdot 69{,}65)+(15\cdot 720\cdot 70{,}07)}{(28\cdot 1065)+(20\cdot 975)+(15\cdot 720)}$$

$$= \frac{9011055{,}8}{124020} = 72{,}66$$

Jadi penduga rata-rata IQ anak SD di Jawa Timur ($\bar{\bar{y}}_{st\,1}$) adalah 92,199 dan penduga rata-rata kemampuan belajar anak SD di Jawa Timur ($\bar{\bar{y}}_{st\,2}$) adalah 72,66. Berdasarkan contoh diatas, maka ada 7 PSU dengan $N_1 = 25$, $N_2 = 21$, $N_3 = 17$, $N_4 = 18$, $N_5 = 28$, $N_6 = 20$, $N_7 = 15$, dan tiap PSU diambil sampel $n_h = 4$. N_h merupakan SSU. Jadi tiap PSU ada 4 SSU. Karena tidak diketahui data secara individual maka tidak dapat dihitung v ($\bar{\bar{y}}_{st}$).

4.5.2.2. Alokasi Optimum dengan Penarikan Sampel Acak Berlapis

Dalam survei, biaya yang dikeluarkan di lapangan berperan besar dalam menentukan jumlah unit penarikan sampel yang optimum. Pada pokok bahasan ini akan dibahas bagaimana menentukan unit yang optimum untuk memimumkan V (variansi) dengan C(biaya) yang tetap.

Total biaya penelitian dengan n tergantung pada dua komponen yaitu ongkos per unit dari unit tahap 1 dan tahap 2 dalam lapisan ke-h. Total biaya penelitian dapat didefinisikan sebagai berikut:

$$C = \sum_h C_{1h} n_h + \sum_h C_{2h} n_h m_h$$

Komponen biaya pertama C_{1h} adalah proporsional terhadap jumlah unit/lapisan utama dalam sampel, yang kedua $C_2 n_h m_h$ proporsional terhadap jumlah unit/lapisan dari tahap kedua. Untuk meminimumkan $V(\overline{\overline{y}}_{st})$ dengan kendala C, maka diperoleh persamaan berikut:

$$V(\overline{\overline{y}}_{st}) = \lambda(C - C_{1h}n_h - C_{2h}n_h m_h) \qquad \text{(e)}$$

λ adalah sebuah pengganda Lagrange, merupakan sebuah fungsi dari variabel n_h dan $(n_h m_h)$. Karena m_h masuk ke V dan C hanya dalam kombinasi $n_h m_h$, andaikan $k = n_h m_h$. Dengan demikian, persamaan (e) dapat ditulis sebagai berikut:

$$\left(\frac{1}{n_h} - \frac{1}{N_h}\right)\sigma_{1h}^2 + \left(\frac{1}{k} - \frac{1}{MN}\right)\sigma_{2h}^2 - \lambda(C - C_{1h}n_h - C_{2h}k) = 0$$

Bila diturunkan terhadap n_h dan k, menghasilkan:

$$\frac{\sigma_{1h}^2}{n_h^2} = -\lambda C_{1h} \Rightarrow \lambda = -\frac{\sigma_{1h}^2}{n_h^2 C_{1h}}$$

$$\frac{\sigma_{2h}^2}{k^2} = -\lambda C_{2h} \Rightarrow \lambda = -\frac{\sigma_{2h}^2}{k^2 C_{2h}}$$

Oleh karena itu diperoleh hubungan:

$$\frac{k}{n_k} = m_{opt} = \frac{\sigma_{2h}}{\sigma_{2h}}\sqrt{\frac{C_{1h}}{C_{2h}}}$$

Dengan mengingat bahwa $\sigma_{u_h} = \sqrt{\dfrac{\sigma_{1h}^2 - \sigma_{2h}^2}{m_h}}$

Penduga dari m_{opt} adalah:

$$\hat{m}_{opt} = \frac{\sigma_{2h}\sqrt{m_h}}{\sqrt{\sigma_{1h}^2 - \sigma_{2h}^2}}\sqrt{\frac{C_{1h}}{C_{2h}}}$$

4.5.3. Sub Penarikan Sampel Acak Berlapis dengan Unit-unit Berukuran Tidak Sama

Pada pokok bahasan 4.5.2. telah dibahas sub penarikan sampel SAB dengan unit-unit berukuran sama. Pokok bahasan ini membahas sub penarikan sampel acak berlapis dengan unit-unit berukuran tidak sama dan bagaimana menduga rata-rata populasi. Sebelum mendefinisikan penduga dari μ akan didefinisikan lebih dahulu banyaknya elemen dalam sub unit ke-i dalam lapisan ke-h.

DEFINISI 4.5.3.1.

$m_{hi} = k_h N_h M_{hi}$

m_{hi} = banyak elemen dalam unit ke-i dalam lapisan ke-h.

$$k_h = \frac{\text{nilai harapan dari jumlah elemen dalam sampel}}{\text{jumlah elemen dalam populasi}}$$

DEFINISI 4.5.3.2.

Penduga dari μ adalah $\bar{\bar{y}}_{st} = \frac{1}{M}\sum_h \frac{N_h}{n_h}\sum_i M_{hi}\,\bar{y}_{hi}$

CONTOH 4.5.3.

Contoh penelitian ini diambil dari penelitian I Gusti Putu Suweta (dalam majalah Analisis Pendidikan, 25 Mei 1984). Peneliti ingin mengukur rata-rata berat badan sapi di Pulau Bali. Populasi dibagi dalam 8 lapisan, yaitu 8 wilayah berdasarkan *derajat keasaman* (pH) *media lingkungan*, *tingkat kebasahan* media lingkungan dan *suhu media lingkungan*. Kedelapan wilayah tersebut adalah:

1. Wilayah-1, pH media asam, air kurang, suhu sejuk merupakan wilayah ekosistem, tenggelam dengan pola tanam palawija.

2. Wilayah-2, ph media asam, air sedang, suhu sejuk merupakan wilayah ekosistem sawah dengan pola tanam diversifikasi palawija.
3. Wilayah-3, pH media netral, air kurang, suhu sejuk, merupakan wilayah ekosistem tegalan dengan pola tanam palawija.
4. Wilayah-4, pH media netral, air kurang, suhu panas, merupakan wilayah ekosistem tegalan dengan pola tanam palawija.
5. Wilayah-5, pH media netral, air sedang, suhu sejuk merupakan wilayah ekosistem sawah dengan pola tanam diversifikasi palawija - padi.
6. Wilayah-6, pH media netral, air sedang, suhu panas merupakan wilayah ekosistem sawah dengan pola tanam diversifikasi padi palawija.
7. Wilayah-7, pH media netral, air banyak, suhu panas merupakan wilayah ekosistem sawah dengan pola tanam monokultur padi.
8. Wilayah 8, pH media alkalis, air kurang, suhu panas merupakan wilayah ekosistem tegalan dengan pola tanam palawija.

Diketahui data sebagai berikut :

Diketahui dari kerangka sampel jumlah sapi keseluruhan di P. Bali M = 4590

h	Banyaknya peternak sapi N_h	Banyaknya sampel n_h
1	10	4
2	15	5
3	14	6
4	18	6
5	19	6
6	19	6
7	15	5
8	13	4
	$\Sigma = 123$	$\Sigma = 41$

Lap ke-h	n_h	Banyaknya Sapi (M_{hi})	m_{hi}	Rata-rata berat	$M_{hi}.\bar{y}_{hi}$
1	A_{11}	25	5	143,4	3585
	B_{12}	35	7	145	5075
	C_{13}	40	8	141	5640
	D_{14}	41	8	139	5699
		$\Sigma M_{hi} = 141$		$\bar{\bar{y}}_h = 142,1$	
2	A_{21}	31	6	139,5	4324,5
	B_{22}	34	7	149	5066
	C_{23}	41	8	150,5	6170,5
	D_{24}	42	8	146,3	6144,6
	E_{25}	36	7	147	5292
		$\Sigma M_{hi} = 184$		$\bar{\bar{y}}_h = 146,46$	
3	A_{31}	37	8	132,5	4902,5
	B_{32}	35	7	136	4760
	C_{33}	28	7	137,5	3850
	D_{34}	29	5	130,8	3793,2
	E_{35}	32	5	139,5	4462
		$\Sigma M_{hi} = 161$		$\bar{\bar{y}}_h = 135,26$	
4	A_{41}	47	9	111,9	5259,3
	B_{42}	45	9	134	6030
	C_{43}	43	8	127	5461
	D_{44}	38	8	120	4560
	E_{45}	39	8	115	4485
	F_{46}	40	8	118	4720
		$\Sigma M_{hi} = 252$		$\bar{\bar{y}}_h = 120,98$	
5	A_{51}	40	8	151	6040
	B_{52}	37	7	151	5587
	C_{53}	38	7	153	5814
	D_{54}	30	6	147	4410
	E_{55}	29	6	145	4205
	F_{56}	33	6	151	4983
		$\Sigma M_{hi} = 207$		$\bar{\bar{y}}_h = 149,67$	
6	A_{61}	34	6	133	4522
	B_{62}	38	7	133,5	5073
	C_{63}	37	7	125	4625
	D_{64}	39	8	129	5031
	E_{65}	41	8	140	5740
	F_{66}	40	8	128	5120
		$\Sigma M_{hi} = 229$		$\bar{\bar{y}}_h = 131,42$	

7	A_{71}	41	8	134	5494
	B_{72}	40	8	137	5480
	C_{73}	43	8	135	5805
	D_{74}	39	8	130	5070
	E_{75}	45	9	131,5	5917,5
		$\Sigma M_{hi} = 208$		$\bar{\bar{y}}_h = 133,5$	
8	A_{81}	38	8	129,5	4921
	B_{82}	35	7	131,5	4602,5
	C_{83}	38	8	137	5206
	D_{84}	39	8	135	5265
		$\Sigma M_{hi} = 150$		$\bar{\bar{y}}_h = 133,25$	

Penduga dari μ adalah

$$\bar{\bar{y}}_{st} = \frac{1}{M}\sum_h \frac{N_h}{n_h}\sum_i M_{hi}\,\bar{y}_{hi}$$

$$= \frac{1}{4590}\left(\begin{array}{l}\frac{10}{4}(19999) + 3(26997,6) + \frac{14}{5}(21767,7(+3930515,3 + \frac{19}{6}(30111) + \\ 3(27766,5) + \frac{13}{4}(625409,05)\end{array}\right)$$

$$= 136,25$$

Jadi penduga rata-rata berat sapi di P. Bali adalah 136,25 kg.

BAB V

KESIMPULAN

Berdasarkan uraian pada bab-bab terdahulu, secara umum dapat disimpulkan :

Teknik penarikan sampel acak berlapis digunakan untuk mengatasi kelemahan yang ada dalam teknik penarikan sampel acak sederhana. Populasi yang sudah ditentukan dengan jelas oleh peneliti dibagi dalam beberapa lapisan, sehingga jumlah semua elemen dari tiap lapisan merupakan jumlah seluruh elemen populasi. Pelapisan dilakukan berdasarkan kriteria yang jelas sesuai tujuan penelitian, sehingga tiap lapisan bersifat homogen dan saling asing. Dari tiap lapisan diambil sampel secara acak sederhana dan dilakukan penghitungan. Hasil hitungan tersebut (statistik) digunakan untuk menduga parameter populasi.

Salah satu masalah yang timbul ketika akan dilakukan pengambilan sampel adalah berapa banyak sampel yang harus diambil dalam tiap lapisan. Langkah pertama yang perlu dilakukan sebelum mengadakan penelitian adalah harus ada data pendahuluan yang dapat di pakai sebagai dasar pelapisan. Ada tiga metode yang dapat digunakan untuk menentukan ukuran sampel, yaitu metode alokasi sebanding, metode alokasi Neyman, metode aloksi optimum. Metode alokasi sebanding digunakan apabila total banyaknya unit-unit penarikan sampel (N_h) berbeda-beda untuk tiap lapisan, dan variansi lapisan σ_h^2, serta ongkos per unit penarikan sampel (c_h) dari tiap lapisan relatif sama. Metode alokasi Neyman digunakan apabila N_h dan σ_h^2 berbeda-beda untuk tiap lapisan sedangkan ongkos

(c_h) per lapisan dianggap relatif sama besarnya. Sedangkan metode alokasi optimum digunakan apabila N_h, S_h^2, c_h, untuk tiap lapisan besarnya berbeda-beda.

Statistik yang dihasilkan dari sampel digunakan untuk menduga total populasi, menduga rata-rata populasi, dan menduga nilai proporsi populasi. Dengan adanya pelapisan, tiap lapisan yang homogen mempunyai variansi pendugaan yang minimum, sehingga penduga akan mendekati nilai parameter yang sebenarnya dalam populasi.

Salah satu alasan mengapa peneliti menggunakan teknik penarikan sampel acak berlapis adalah untuk mengatasi kelemahan dari penarikan sampel acak sederhana. Dalam penarikan sampel acak sederhana, data yang terkumpul kurang akurat, karena data dalam populasi sangat heterogen sehingga menghasilkan variansi yang besar. Bila populasi heterogen, peneliti tidak dapat menduga besar variansi populasi dengan baik, sehingga diperlukan penduga variansi dengan membagi populasi ke dalam sub populasi (lapisan) yang relatif homogen berdasarkan kriteria tertentu. Peneliti dalam melakukan penarikan sampel acak berlapis memperoleh hasil yang maksimal yang berarti peneliti dapat melakukan pendugaan paramter untuk tiap lapisan.

Dalam teknik penarikan sampel acak berlapis sering dijumpai bahwa dalam suatu lapisan masih terdapat lapisan-lapisan (sub lapisan) berdasarkan kriteria tertentu. Penarikan sampel acak berlapis dua tahap digunakan oleh peneliti untuk mengambil nilai dalam sub lapisan yang lebih kecil sehingga bisa meningkatkan efisiensi statistik sampel. Populasi telah terdefinisi dengan jelas dibagi dalam L lapisan. Tiap lapisan diambil sampel acak sederhana yang merupakan unit-unit.

utama (PSU), dan tiap PSU diambil sampel acak sederhana yang merupakan unit-unit sekunder (SSU). Peneliti dapat memperoleh hasil yang maksimal karena perhitungan dugaan parameter dilakukan dalam 2 tahap. Peneliti melakukan pendugaan parameter dalam tiap SSU yang selanjutnya dilakukan pendugaan parameter dalam tiap PSU, dan akhirnya berdasarkan hasil tersebut dapat dilakukan pendugaan parameter untuk populasi secara keseluruhan.

DAFTAR PUSTAKA

Agung, I.G. N, 1992, *Metode Penelitian Sosial Pengertian dan Pemakaian Praktis.* PT Gramedia Pustaka Utama.

Cochran, W.G. 1991. *Teknik Penarikan Sampel* . Jakarta : Penerbit Universitas Indonesia.

Cochran, W.G. 1977. *Sampling Techniques.* Canada : John Wiley & Sons,Inc.

Dalenius,T. 1985. *Elements of Survey Sampling.* Notes Prepared for the svedisk Agency for Research Cooperation with Developing Countries.

Gaspersz, V. 1991. *Teknik Penarikan Contoh untuk Penelitian Survei.* Bandung : Penerbit Tarsito.

Kendall, S.M.G, W.r. Buckland.1982. *A Dictionary of Statistical terms,* John Wiley and Sons,Inc.

Kruskal, W.H, J. M. Tanur. 1978. - *International Encyclopedia of Statistics.* New York : A Division of Macmillan Publishing Co,Inc.

Mendenhall,W. – *Elementary Survey Sampling.* Belmont California : Wadsworth Publishing Conny, Inc.

Nasution, A.H. 1992. *Panduan Berpikir dan Meneliti Secara Ilmiah Bagi Remaja.* Jakarta : PT Gramedia Widiasarana Indonesia.

Popham,W.J. 1973. *Educational Statistics Use and Interpretation.* Los Angeles : University of California.

Raj,D. 1968. *Sampling Theory.* America : McGraw-Hill,Inc.

Virola, R.A. *Monograph on Inference.*

Walpole, R.E. 1992. Pengantar Statistika. Jakarta : PT. Gramedia Pustaka Utama

__________. 1996. *Laporan Perekonomian DIY.* Yogyakarta : Biro Pusat Statistik.

LAMPIRAN

Penarikan Sampel Menggunakan Angka Acak

Urutan pembacaan angka acak	Angka acak dalam susunan tiga digit	Terletak dalam baris dan kolom (baris; kolom)
1	118	(30; 35 - 37)
2	908	(31; 35 - 37)
3	036	(32; 35 - 37)
4	130	(33; 35 - 37)
5	326	(34; 35 - 37)
6	053	(35; 35 - 37)
7	161	(36; 35 - 37)
8	830	(37; 35 - 37)
9	279	(38; 35 - 37)
10	831	(39; 35 - 37)
11	735	(40; 35 - 37)
12	222	(41; 35 - 37)
13	317	(42; 35 - 37)
14	807	(43; 35 - 37)
15	038	(44; 35 - 37)
16	135	(45; 35 - 37)
17	613	(46; 35 - 37)
18	205	(47; 35 - 37)
19	650	(48; 35 - 37)
20	833	(49; 35 - 37)
21	607	(50; 35 - 37)
22	655	(51; 35 - 37)
23	918	(52; 35 - 37)
24	778	(53; 35 - 37)
25	917	(54; 35 - 37)
26	434	(55; 35 - 37)
27	476	(56; 35 - 37)
28	709	(57; 35 - 37)
29	896	(58; 35 - 37)
30	567	(59; 35 - 37)

TABEL ANGKA ACAK

	00-04	05-09	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
00	88758	66605	33843	43623	62774	25517	09560	41880	85126	60755
01	35661	42832	16240	77410	20686	26656	59698	86241	13152	49187
02	26335	03771	46115	88133	40721	06787	95962	60841	91788	86386
03	60826	74718	56527	29508	91975	13695	25215	72237	06337	73439
04	95044	99896	13763	31764	93970	60987	14692	71039	34165	21297
05	83746	47694	06143	42741	38338	97694	69300	99864	19641	15083
06	27998	42562	63402	10056	81668	48744	08400	83124	19896	18805
07	82685	32323	74625	14510	65927	28017	80568	14756	54937	76379
08	18386	13862	10988	04197	18770	72757	71418	81133	69503	44037
09	21717	13141	22707	68165	58440	19167	08421	23872	03036	34208
10	18446	83052	31842	08634	11867	86070	08464	20565	74390	36541
11	66027	75177	47398	66423	70160	16232	67343	36205	50036	59411
12	51420	96779	54309	67456	78967	79638	68869	49062	02196	55109
13	27045	62626	73159	91149	96509	44204	92237	29969	49315	11804
14	13094	17725	14103	00067	68843	63565	93578	24756	10814	15185
15	92382	62518	17752	53163	63852	44840	02592	88572	03107	90169
16	16215	50609	49326	77232	90155	69955	93892	70445	00906	57002
17	09342	14528	64727	71403	84156	34063	35613	35670	10549	07468
18	38148	79001	03509	79424	39625	73315	18811	86230	99682	82896
19	23689	19997	72382	15247	80205	58090	43804	94548	82695	22799
20	25407	37726	73099	51057	68733	75766	77991	72641	95386	70138
21	25349	69456	19693	85568	93676	18661	69016	10332	83137	88257
22	02322	77491	56095	03055	37738	18216	81781	32245	84081	18436
23	15072	33261	99219	43307	39239	79712	94753	41450	30944	53912
24	27002	31036	85276	74547	84809	36252	09373	69171	15606	77209
25	66181	83316	40386	54316	29505	86032	34563	93204	72973	90760
26	09779	01822	45537	13128	51128	82703	75350	25179	86104	40638
27	10791	07706	87481	26197	24857	27805	42710	63471	08804	23455
28	74833	35767	31312	76611	67389	04591	39687	13596	88730	66850
29	17583	24036	83701	28570	63561	00098	60784	76098	84217	34997
30	45601	46977	39325	09286	41133	34031	94867	11849	75171	57682
31	60683	33112	65995	54203	18070	65437	13624	99896	80915	71987
32	29956	81169	18877	15296	94366	16317	34239	03643	66081	12242
33	91713	84235	75296	69875	82414	05197	66596	13083	46278	73498
34	85704	86589	82837	57822	95963	83021	90752	32661	64751	83903
35	17921	26111	35373	86494	48266	01888	65735	05315	79328	13367
36	13929	71341	80188	89827	48277	07229	71953	16128	65074	28782
37	03248	18280	21067	01311	61805	80201	47889	83052	31029	06023
38	50583	17972	12690	00452	93766	16414	01212	27964	02766	28786
39	10636	46975	09449	45986	34672	46916	63881	93117	53947	95218
40	43896	41278	42205	10425	66560	59967	90139	73563	29875	79033
41	76714	80963	74907	16890	15492	27489	06067	22287	19760	13056
42	22393	45719	02683	62428	45177	57562	49243	31748	64278	05731
43	70942	92042	22776	47761	13503	16037	30875	80754	47491	96012
44	92011	60326	86346	26738	01983	04186	41388	03848	78354	14964
45	66156	00126	45685	67607	70796	04889	98128	13599	93710	23974
46	96292	41348	20898	02227	76512	53185	03057	61375	10760	26889
47	19680	07146	53951	10935	23333	76233	13706	20502	60405	09745
48	67347	51112	24536	60151	05498	61678	87569	65066	17790	55413
49	95888	59255	06898	99137	50871	81265	42223	83303	40694	61953

	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99
00	70896	44520	64720	49898	78088	76740	47460	83150	78905	59870
01	56809	42909	25853	47624	29486	14196	75841	00393	42390	24847
02	66109	84775	07515	49949	61482	91836	48126	80778	21302	24975
03	18071	36263	14053	52526	44347	04923	68100	57805	19521	15345
04	98732	15120	91754	12657	74675	78500	01247	49719	47635	55514
05	36075	83967	22268	77971	31169	68584	21336	72541	66959	39708
06	04110	45061	78062	18911	27855	09419	56459	00695	70323	04538
07	75658	58509	24479	10202	13150	95946	55087	38398	18718	95561
08	87403	19142	27208	35149	34889	27003	14181	44813	17784	41036
09	00005	52142	65021	64438	69610	12154	98422	65320	79996	01935
10	43674	47103	48614	70823	78252	82403	93424	05236	54588	27757
11	68597	68874	35567	98463	99671	05634	81533	47406	17228	44455
12	91874	70208	06308	40719	02772	69589	79936	07514	44950	35190
13	73854	19470	53014	29375	62256	77488	74388	53949	49607	19816
14	65926	34117	55344	68155	38099	56009	03513	05926	35584	42328
15	40005	35246	49440	40295	44390	83043	26090	80201	02934	49260
16	46686	29890	14821	69783	34733	11803	64845	32065	14527	38702
17	02717	61518	39583	72863	50707	96115	07416	05041	36756	61065
18	17048	22261	35573	28944	96889	51823	57268	03866	27658	91950
19	75304	53248	42151	93928	17343	88322	28683	11252	10355	65175
20	97844	62947	62230	30500	92816	85232	27222	91701	11057	83257
21	07611	71163	82212	20653	21499	51496	40715	78952	33029	64207
22	47744	04603	44522	62783	39347	72310	41460	31052	40814	94297
23	54293	43576	88116	67415	34908	15238	40561	73940	56050	31078
24	67556	93979	73363	00300	11217	74405	18937	79080	60834	48307
25	86581	73041	95809	73986	49408	53316	90841	73808	53421	82315
26	28020	86282	83365	76600	11261	74354	20968	60770	12141	09539
27	42578	52471	37840	30872	75074	79027	57813	62831	54715	26693
28	47290	15997	86163	10571	81911	92124	92971	80860	41012	58666
29	24856	63911	13221	77028	06573	33667	30732	47280	12926	67276
30	16352	24036	60799	76281	83402	44709	78930	82969	84468	36910
31	89060	79852	97354	28324	39638	86936	06702	74304	39873	19496
32	07637	30412	04921	26471	09605	07355	20466	49793	40539	21077
33	37711	47786	37468	31963	16908	50283	80864	08252	72655	58926
34	82994	53232	58202	73318	62471	49650	15888	73370	98748	69181
35	31722	67288	12110	04776	15168	68852	92347	90789	66961	01162
36	93619	78050	19364	38037	25706	90879	05215	00260	14426	88207
37	65557	24495	04713	23688	26623	41356	47049	60676	72236	01214
38	88001	91382	05129	36041	10257	55558	89979	58061	28957	10701
39	96648	70303	18191	62404	26558	92804	15415	02865	52449	78509
40	04118	51573	59356	02426	35010	37104	98316	44602	96478	08433
41	19317	27753	39431	26996	04465	69695	61374	06317	42225	62025
42	37182	91221	17307	68507	85725	81898	22588	22241	80337	89033
43	82990	03607	29560	60413	59743	75000	03806	13741	79671	25416
44	97294	21991	11217	98087	79124	52275	31088	32085	23089	21498
45	86771	69504	13345	42544	59616	07867	78717	82840	74669	21515
46	26046	55559	12200	95106	56496	76662	44880	89157	84209	01332
47	39689	05999	92290	79024	70271	93352	90272	94195	26042	54477
48	83265	89573	01437	43786	52986	49041	17952	35035	88985	84671
49	15128	35791	11296	45319	06330	82027	90808	54354	43091	30387

TABEL ANGKA ACAK (LANJUTAN)

	00-04	05-09	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49
50	54441	64681	93190	00993	62130	44484	46293	60717	50239	76319
51	08573	52937	81274	95106	89117	65849	41356	65549	78787	50442
52	81067	68052	14270	19718	88499	63303	13533	91882	51136	60828
53	39737	58891	75278	98046	52284	40164	72442	77824	72900	14886
54	34958	76090	08827	61623	31114	66952	83645	91786	29633	78294
55	61117	72424	92626	71952	69709	81259	58472	43409	84454	88648
56	99167	14149	57474	32268	85424	90378	34682	47606	89295	02420
57	13130	13064	36185	48133	35319	05720	76317	70953	50823	06793
58	65563	11831	82402	46929	91446	72037	17205	89600	59084	55718
59	28737	49502	06060	52100	43704	50839	22538	56768	83467	19313
60	50353	74022	59767	49927	45882	74099	18758	57510	58560	07050
61	65208	96466	29917	22862	69972	35178	32911	08172	06277	62795
62	21323	38118	26696	81741	25131	20087	67452	19670	35898	50636
63	67875	29831	59330	46570	69768	30671	01031	95995	68417	68665
64	82641	26260	86551	39081	70512	37899	38851	40568	51204	24056
65	91899	39643	59089	12526	37730	68848	71399	28513	69018	10289
66	12950	31118	93125	69756	31036	55097	97211	92480	49745	42461
67	00329	27127	95171	97217	05034	26676	49629	13591	50525	13185
68	63986	16698	82881	01524	39919	32381	67488	05223	89537	59490
69	55725	75605	57912	20977	35722	51931	89565	77579	93085	06467
70	21751	36877	56357	78809	40748	69727	56652	12162	40528	75269
71	43820	80926	26795	57553	28319	25376	51795	26123	51102	89853
72	66669	02880	02987	33615	54206	29013	75872	88678	17726	60640
73	49511	66725	19779	50416	42800	71733	82052	28504	15593	51799
74	71003	87596	61296	95019	21568	86131	66096	65403	47166	78638
75	52715	01593	69481	93111	38016	13000	01293	66830	03914	75357
76	21998	31729	89963	11573	49442	69467	40265	56066	36024	25705
77	58970	96827	18377	31564	23555	86338	79250	43166	96929	97732
78	67592	59149	42554	42719	13553	48560	81167	10747	92552	19867
79	10298	18429	09357	96436	11237	88039	81020	00128	75731	37779
80	88120	28811	42628	81617	59024	52032	31251	72017	43875	48320
81	07627	88424	23381	29680	14027	75905	27037	22113	77873	78711
82	37917	93581	04979	21041	95252	62450	05937	61670	44894	47262
83	14783	95119	68464	08726	74818	91700	05961	23554	74649	50540
84	05378	32640	64562	15303	13168	23189	88198	63617	58566	56047
85	19640	96709	22047	07825	40583	99500	39969	96593	32254	37158
86	20514	11081	51131	56469	33947	77703	35679	45774	06776	67062
87	96763	56219	81213	62416	84451	14696	38195	70435	45948	67690
88	49139	61075	31558	59740	52759	55323	95226	01385	20158	54054
89	16294	50548	71317	32168	86071	47314	65393	56367	46910	51269
90	31391	94301	79273	32613	05862	36211	93960	00671	67631	23952
91	98032	87203	03227	66021	99666	98368	39222	36056	81992	20121
92	40700	31826	94774	11366	81391	33602	69608	84119	93204	26825
93	68692	66849	29366	77540	14978	06508	10824	65416	23629	63029
94	19047	10784	19607	20296	31804	72984	60060	50353	23260	58909
95	82867	69266	50733	62630	00956	61500	89913	30049	82321	62367
96	26528	28928	52600	72997	80943	04084	86662	90025	14360	64867
97	51166	00607	49962	30724	81707	14548	25844	47336	57492	02207
98	97245	15440	55182	15368	85136	98869	33712	95152	50973	98658
99	54998	88830	95639	45104	72676	28220	82576	57381	34438	24565

	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99
50	58649	85086	16502	97541	76611	91229	34987	86718	87208	05426
51	97306	52449	55596	66739	36525	97563	29169	31235	79276	10831
52	09942	79344	78160	11015	55777	22047	57615	15717	86239	36378
53	83842	28631	74893	47911	92170	38181	30416	54860	44120	73031
54	73778	30395	20163	76111	13712	33449	99224	18206	51418	70006
55	88381	56550	47467	59663	61117	39716	32927	06168	06217	45477
56	31044	21404	15968	21357	30772	81482	38807	67231	84283	63552
57	00909	63837	9132°	81106	11740	50193	86806	21931	18054	49601
58	69882	37028	41732	37425	"9832	03320	20690	32653	90145	03029
59	26059	78324	22501	73825	16927	31545	15695	74216	98372	28547
60	38573	98078	38982	33078	93524	45606	53463	20391	81637	37269
61	70624	00063	81455	16924	12848	23801	55481	78978	26795	10553
62	49806	23976	05610	29804	38988	25024	76951	02341	63219	75864
63	05161	67523	48316	13613	08511	35231	38312	14969	67279	50502
64	76582	62153	53801	51219	30124	32599	49099	83959	68408	20147
65	16660	80176	75062	75588	24381	27874	20618	11428	32265	07692
66	60166	42424	97470	88451	81270	80070	72959	26220	59939	31127
67	28953	03272	31460	41691	57736	72052	22762	96323	27616	53123
68	47536	86439	95210	96386	38701	15484	07426	70675	06888	81203
69	73457	26657	36983	72410	30244	97711	25652	09373	66218	64077
70	11190	66193	66287	09116	48140	37669	02932	50799	17255	06181
71	57062	78964	44455	14036	36098	40773	11688	33150	07459	36127
72	99624	67254	67302	18991	97687	54099	94884	42283	63258	50651
73	97521	83669	85968	16135	30133	51312	17831	75016	89278	68953
74	40273	04838	13661	64757	17461	78085	60094	27010	80945	66139
75	57260	06176	49963	29760	69546	61336	39129	41985	18572	96128
76	03451	47098	63495	71227	79304	29753	99131	18419	71791	81515
77	62331	20492	15393	84270	24396	32962	21632	92965	38670	44923
78	32290	51079	06512	38806	93327	80086	19088	59887	98416	24918
79	28014	80428	92853	31333	32648	16734	43418	90124	15086	48444
80	18950	16091	29543	65817	07002	73115	91115	20271	50250	25061
81	17403	69503	01866	15019	07263	13039	83844	80143	39048	62654
82	27909	50489	65613	21843	71746	65868	16208	46781	93402	12323
83	87076	53174	12165	84495	47947	60706	64034	31655	65169	93070
84	89044	45974	14524	46906	26052	51851	84197	61694	57429	63395
85	98048	64400	24705	75711	36232	57624	41424	77366	52790	84705
86	09345	12956	49770	80311	32319	48238	16952	92088	51222	82865
87	07086	77628	76195	47584	62411	40397	71857	54823	26536	56792
88	93128	25657	46872	11206	06831	87944	97914	64670	45760	34353
89	85137	70964	29947	27795	25547	37682	96105	26848	09389	64326
90	32796	39024	13814	98546	46585	84108	74603	94812	73968	68766
91	62496	26371	89880	52078	47781	95260	83464	65942	91764	53727
92	62707	81825	40987	97656	89714	52177	23779	07482	91678	40128
93	05500	28982	86124	19554	80818	94935	61924	31828	79369	23507
94	79176	31445	59498	85132	24582	26024	24002	63718	79164	43556
95	10653	29954	97568	91541	33139	84525	72271	02546	61818	14381
96	30524	06495	00886	40666	68574	49574	19705	16429	90981	08103
97	69050	22019	74066	14500	14506	06423	38332	34191	82663	85323
98	27908	78802	63146	07674	98871	63831	72449	42705	26513	19683
99	64520	16618	47409	19574	78136	46047	01277	79146	95759	36781