TAKSIRAN MEAN DAN TOTAL PADA

TWO STAGE ADAPTIVE CLUSTER SAMPLING

MAYRAMADAN MADYA PUTRA

0305010351

UNIVERSITAS INDONESIA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

DEPARTEMEN MATEMATIKA

DEPOK

2009

TAKSIRAN MEAN DAN TOTAL PADA

TWO STAGE ADAPTIVE CLUSTER SAMPLING

Skripsi diajukan sebagai salah satu syarat

untuk memperoleh gelar Sarjana Sains

Oleh:

MAYRAMADAN MADYA PUTRA

0305010351

DEPOK

2009

SKRIPSI : TAKSIRAN MEAN DAN TOTAL PADA *TWO STAGE ADAPTIVE CLUSTER SAMPLING*

NAMA : MAYRAMADAN MADYA PUTRA

NPM : 0305010351

SKRIPSI INI TELAH DIPERIKSA DAN DISETUJUI

DEPOK, JULI 2009

Dra. RIANTI SETIADI, M.Si	FEVI NOVKANIZA, S.Si., M.Si
PEMBIMBING I	PEMBIMBING II

Tanggal Lulus Ujian Sidang Sarjana: 9 Juli 2009

Penguji I : Dra. Rianti Setiadi, M.Si

Penguji II : Dr. Dian Lestari

Penguji III : Dr. Kiki Ariyanti S.

KATA PENGANTAR

Segala Puji dan Syukur hanya kepada Allah SWT, yang telah memberikan segala nikmat dan karunianya kepada penulis sehingga tugas akhir ini dapat diselesaikan. Tugas akhir ini dapat selesai juga karena bantuan doa, moril, bimbingan, dan dorongan dari banyak pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada :

1. Keluarga tercinta, ibu dan bapak penulis, kak Norman, kak Aco, kak Yuli, Susan, Messa, kak Meli dan tante Mis yang telah memberikan semangat kepada penulis untuk terus berjuang menyelesaikan tugas akhir ini.
2. Pembimbing tugas akhir penulis, Dra. Rianti Setiadi, M.Si dan Fevi Novkaniza, S. Si., M.Si, yang dengan sabar membimbing, memberi saran, dan bantuan selama proses pembuatan hingga tersusunnya tugas akhir ini.
3. Pembimbing akademik, Dra. Siti Nurrohmah, M.Si yang telah membimbing dan memberikan saran selama penulis menimba ilmu di Matematika UI.
4. Para dosen Departemen Matematika FMIPA-UI yang telah memberikan ilmu yang berguna kepada penulis, terutama kepada Dr. Yudi Satria MT, Dr. Dian Lestari, Dra. Saskya Mary, M. Si, Dra. Suarsih Utama, Rahmi Rusin S.Si, MSc.Tech, dan Mila Novita S.Si., M.Si yang telah memberikan saran, nasehat atau semangat kepada penulis selama pembuatan tugas akhir.
5. Seluruh karyawan Departemen Matematika FMIPA-UI, terutama mba Santi, pak Saliman, mas Irwan, dan pak Anshori yang telah membantu selama proses registrasi seminar.

6. Teman-teman seperjuangan yang mengambil skripsi, Ida , Vani, Amri, Khuriyanti, Shinta, Ratih, Rizky, Rifky, Riesa, Syarah, Maul, Uun, Iif.
7. Vani, Ida, Shinta, Fika, Wicha, Dia, Yanu, Desti, Anggie yang telah memberikan semangat yang luar biasa kepada penulis.
8. Semua teman-teman angkatan 2005, Fika, Wicha, Dia, Ratna, Melati, Raisa, Nisma, Othe, Miranti, Rani, Desti, Anggie, Jessie, Akmal, Anggie, Puji, Shally, Gyo, Pute, Aini, Rif'ah, Rara, Yanu, Ranti, TH, Fery, Maria, Andre, Karlina, QQ, Aya, Merry, Yuni, Fia, Dian, Mia, Hamdan, Asep, Trian, Ridwan, Aris, Hairu, Udin, dll.
9. Semua teman-teman angkatan 2003, 2004, 2006, 2007, dan 2008 terutama Dicky 03, Gele 03, Ajat 04, Bong 04, Ias 04, Rimbun 04, Nadya 04, Avidati 04, Lee 06, Rita 06, Syafirah 06, Yuri 06, ArRizqiyatul 06, Alberta 06, Winda 07, Farah 07, Shafira 07, Amanda 07, Hikmah 07, Widya 07, Nedia 07, Adit 07, Dhanar 07, Syahrul 07, Azhari 07, Zulfalah 07, yang telah membantu dan memberikan semangat.

Penulis menyadari bahwa penulisan tugas akhir ini kurang sempurna. Oleh karena itu penulis ingin memohon maaf bila masih terdapat kesalahan, karena penulis hanyalah manusia biasa yang tak luput dari kesalahan, dan perlu diingat bahwa kesempurnaan hanya milik Allah SWT. Semoga tugas akhir ini bermanfaat bagi orang yang membacanya.

Penulis

2009

ABSTRAK

Metode *two stage adaptive cluster sampling* (2S-ACS) sangat baik digunakan untuk mengambil sampel dimana elemen yang akan diteliti sangat jarang atau berkelompok. Pada 2S-ACS, pengambilan sampel diawali dengan membagi wilayah penelitian menjadi unit-unit primer. Masing-masing unit primer dibagi menjadi unit-unit sampling. Pada tahap pertama, dipilih beberapa unit primer secara SRS. Pada tahap kedua, dari masing-masing unit primer yang terpilih pada tahap pertama, diambil beberapa unit sampling sebagai sampel awal. Kemudian, dilakukan proses penambahan sampel pada masing-masing unit sampling yang terpilih pada sampel awal. Ada dua skema yang dapat digunakan untuk menambahkan sampel, yaitu skema *overlapping* dan skema *nonoverlapping*. Pada skema *overlapping*, proses penambahan sampel diperbolehkan melewati batas unit primer, sedangkan pada skema *nonoverlapping* tidak diperbolehkan melewati batas unit. Pada masing-masing skema akan digunakan taksiran Horvitz-Thompson dan taksiran Hansen-Hurwitz untuk menaksir mean dan total populasi. Taksiran yang diperoleh adalah taksiran yang tak bias. Pada tugas akhir ini akan diberikan contoh penerapan *two stage adaptive cluster sampling* dengan menggunakan skema *overlapping* dan skema *nonoverlapping.*

Kata kunci : taksiran Horvitz-Thompson; taksiran Hansen-Hurwitz; *two stage adaptive cluster sampling*; unit primer; unit sampling.

ix+106 hal.;lamp.;gamb.;tab.;
Bibliografi : 10 (1967-2002)

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengambilan sampel merupakan hal yang penting dalam melakukan suatu penelitian. Pengambilan sampel dilakukan untuk memperoleh data yang akan digunakan dalam penelitian. Ada berbagai macam metode yang dapat digunakan dalam pengambilan sampel. Pemilihan metode pengambilan sampel juga harus diperhatikan karena akan mempengaruhi hasil penelitian. Jika metode pengambilan sampel yang digunakan tidak sesuai, maka hasil yang diperoleh dari penelitian tersebut dapat menyesatkan. Oleh karena itu, dalam pengambilan sampel perlu diketahui bagaimana keadaan dari populasi.

Populasi merupakan kumpulan dari elemen-elemen yang merupakan objek penelitian dimana pengukuran akan diambil. Elemen-elemen tersebut seringkali dapat dikelompokkan ke dalam kelompok-kelompok yang merupakan partisi dari populasi terkait. Kelompok-kelompok ini disebut sebagai unit sampling. Pengambilan sampel dilakukan terhadap unit-unit sampling sedangkan pengukuran dilakukan terhadap elemen-elemen yang terdapat pada unit sampling yang terpilih sebagai sampel.

Adakalanya unit sampling dan elemen merupakan hal yang sama. Sebagai contoh, misalkan akan diteliti rata-rata pertambahan berat ayam setelah diberi vitamin tertentu. Kemudian, ayam-ayam yang akan diteliti berada di dalam beberapa kandang dan pengambilan sampel dilakukan pada kandang ayam. Kemudian, pengukuran dilakukan pada ayam-ayam yang berada di dalam kandang yang terpilih menjadi sampel. Dalam penelitian ini unit samplingnya adalah kandang sedangkan elemennya adalah ayam. Andaikan ayam-ayam yang akan diteliti tidak berada di dalam kandang dan pengambilan sampel dilakukan secara langsung pada ayam, maka unit sampling dan elemen dalam penelitian ini adalah sama yaitu ayam.

Dalam penelitian, terutama penelitian lapangan populasi merupakan daerah yang mencakup seluruh wilayah penelitian. Wilayah penelitian dapat dibagi menjadi beberapa daerah penelitian/subwilayah yang merupakan partisi dari wilayah penelitian. Sampel dipilih dari daerah penelitian/subwilayah dan pengukuran dilakukan terhadap objek penelitian yang berada di dalam daerah penelitian/ subwilayah yang terpilih sebagai sampel. Dalam hal ini daerah penelitian/subwilayah merupakan unit sampling dan objek penelitian merupakan elemen.

Dalam penelitian sering terjadi suatu kasus di mana elemen yang akan diteliti sangat jarang atau saling berkelompok. Dalam kasus ini, ketika pengambilan sampel dari unit sampling dilakukan dengan metode konvensional, ada kemungkinan bahwa elemen yang akan diteliti tidak ditemukan dalam unit sampling yang terpilih sebagai sampel, sehingga data yang cukup tidak akan diperoleh. Ada juga kemungkinan lain bahwa elemen yang diperoleh dari unit sampling yang terpilih sebagai sampel hanya sedikit, tetapi elemen-elemen yang akan diteliti dapat ditemukan juga pada unit-unit sampling yang bertetanggaan dengan unit sampling yang terpilih sebagai sampel. Seandainya elemen-elemen yang terdapat pada unit-unit sampling yang bertetanggaan dengan unit sampling yang terpilih menjadi sampel, juga dapat dimasukkan menjadi elemen sampel, maka data yang diperoleh akan lebih memadai.

Thompson (1990, 1992) memperkenalkan suatu metode pengambilan sampel yang akan memasukkan elemen-elemen yang ditemukan pada unit-unit sampling yang bertetanggaan dengan unit sampling yang terpilih sebagai sampel untuk menjadi elemen sampel. Metode ini dikenal dengan *adaptive cluster sampling*.

Cara pengambilan sampel pada *adaptive cluster sampling* diawali dengan membagi wilayah penelitian menjadi subwilayah-subwilayah yang disebut unit-unit sampling. Kemudian, ambil sampel awal (*initial sampel*) dari unit-unit sampling dengan teknik sampling yang sudah diketahui, misalnya *simple random sampling (SRS)*. Jika dalam sampel awal tersebut terdapat elemen yang akan diteliti, maka unit-unit disekitarnya dijadikan sampel dan diteliti apakah terdapat elemen dan memenuhi kondisi tertentu yang telah ditentukan oleh peneliti. Apabila terdapat elemen yang akan diteliti dan memenuhi kondisi yang diinginkan oleh peneliti, maka unit-unit disekitarnya juga akan dijadikan sampel dan diteliti apakah memenuhi kondisi tersebut.

Proses ini terus dilakukan hingga tidak ada lagi unit-unit disekitar unit sampling yang terpilih menjadi sampel dan memenuhi kondisi yang diinginkan oleh peneliti. Pada *adaptive cluster sampling* dapat diperoleh taksiran parameter (dalam tugas akhir ini dibatasi untuk mean dan total) yang tak bias dengan taksiran variansi yang tak bias pula.

Metode *adaptive cluster sampling* memiliki beberapa kelemahan, yaitu ada kemungkinan bahwa elemen pada sampel akhir yang terbentuk akan sangat banyak. Kelemahan berikutnya adalah memerlukan banyak usaha/tenaga dalam meneliti unit-unit sampling yang menjadi sampel awal karena ada kemungkinan unit-unit sampling tersebut terletak berjauhan. Hal ini menjadi suatu kesulitan jika daerah yang menjadi objek penelitian sangat luas.

Dalam tugas akhir ini akan diperkenalkan suatu metode yang dapat digunakan untuk mengatasi permasalahan dalam *adaptive cluster sampling,* yaitu elemen pada sampel akhir sangat banyak, dan atau sampel awal terletak berjauhan. Metode ini dikenal dengan *two stage adaptive cluster sampling.* Berbeda dengan *adaptive cluster sampling*, pada *two stage adaptive cluster sampling,* daerah yang menjadi objek penelitian dibagi dahulu menjadi beberapa unit primer yang diperkirakan homogen. Kemudian, masing-masing unit primer dibagi menjadi unit-unit sampling. Selanjutnya, ambil beberapa unit primer secara SRS dan dari unit primer yang terpilih, pilih unit sampling sebagai sampel awal (*initial sample*) secara SRS. Setelah sampel awal terpilih, proses selanjutnya sama seperti *adaptive* *cluster sampling.* Permasalahan yang perlu diselesaikan pada *two stage adaptive cluster sampling* adalah bagaimana mendapatkan taksiran mean populasi yang tak bias dengan taksiran variansi yang tak bias pula. Hal ini yang akan diselesaikan dalam tugas akhir ini.

1.2 Permasalahan

Permasalahan dalam tugas akhir ini adalah bagaimana mendapatkan taksiran tak bias untuk mean dan total dengan taksiran variansi yang tak bias, pada *two stage adaptive cluster sampling.*

1.3 Tujuan Penulisan

Tujuan penulisan tugas akhir ini adalah :

- Menjelaskan tentang metode *two stage adaptive cluster sampling.*
- Mencari taksiran tak bias untuk mean dan total dengan taksiran variansi yang tak bias, pada *two stage adaptive cluster sampling.*

1.4 Pembatasan Masalah

Batasan masalah dalam skripsi ini antara lain :

- Pemilihan unit primer dan sampel awal *(initial sampel)* hanya dilakukan dengan s*imple random sampling* (SRS) tanpa pengembalian.
- Banyaknya unit sampling pada masing-masing unit primer adalah sama.
- Banyaknya unit sampling yang terpilih sebagai sampel awal pada masing-masing unit primer adalah sama.

1.5 Sistematika Penulisan

Penulisan tugas akhir ini dibagi menjadi lima bab, yaitu

BAB I : Pendahuluan

Pada bab ini dibahas latar belakang, permasalahan, tujuan, pembatasan masalah, dan sistematika penulisan tugas akhir ini.

BAB II : Landasan Teori

Pada bab ini dibahas mengenai landasan teori tugas akhir ini, yaitu *simple random sampling*, *two stage sampling*, *unequal probability sampling*, dan *adaptive cluster sampling*.

BAB III : *Two stage adaptive cluster sampling*

Pada bab ini akan dijelaskan mengenai skema yang digunakan pada *two stage adaptive cluster sampling*. Pada masing-masing skema akan dijelaskan taksiran Horvitz-Thompson dan taksiran Hansen-Hurwitz untuk menentukan taksiran mean dan total populasi.

BAB IV : Contoh Penerapan

Pada bab ini diberikan suatu ilustrasi / contoh penerapan metode *two stage adaptive cluster sampling* dengan menggunakan skema *overlapping* dan skema *nonoverlapping.*

BAB V : Penutup

Bab ini terdiri dari kesimpulan dan saran.

BAB II

LANDASAN TEORI

Pada bab ini akan dibahas mengenai teori-teori yang mendasari topik pada tugas akhir ini, yaitu *Simple Random Sampling (SRS)*, *Two Stage Sampling, Unequal Probability Sampling,* dan *Adaptive cluster sampling.*

2.1 Simple Random Sampling

Simple Random Sampling (SRS) adalah metode pengambilan sampel dimana setiap kombinasi sampel yang mungkin, mempunyai probabilitas yang sama untuk terpilih menjadi sampel. Sampel yang dipilih secara SRS disebut *Simple Random Sample*. SRS merupakan bentuk dasar dari probability sampling dan secara teoritis menjadi dasar dari bentuk sampling yang lebih rumit. Dalam pengambilan sampel, SRS dapat dilakukan dengan dua cara yaitu dengan pengembalian atau tanpa pengembalian. Dalam tugas akhir ini hanya akan dibahas SRS tanpa pengembalian.

Teorema 2.1:

Dalam SRS tanpa pengembalian, probabilitas suatu unit terpilih menjadi anggota sampel adalah sama yaitu $\frac{n}{N}$, dimana n adalah ukuran sampel sedangkan N adalah ukuran populasi.

Bukti :

Misalkan dipunyai unit pada populasi : $\{u_1, u_2, ..., u_N\}$

Pandang suatu unit pada populasi, yaitu u_1

$p_j = \Pr(u_1$ muncul pada pengambilan ke j);

$j = 1,2,...,n$

$j = 1$ maka $p_1 = \dfrac{1}{N}$

$j = 2$ maka $p_2 = \Pr(u_1$ tidak muncul pada pengambilan pertama dan

u_1 muncul pada pengambilan kedua)

misalkan : F = kejadian u_1 muncul pada pengambilan pertama

F' = kejadian u_1 tidak muncul pada pengambilan pertama

G = kejadian u_1 muncul pada pengambilan kedua

dengan demikian, $\Pr(F) = \dfrac{1}{N}$, $\Pr(F') = \left(1 - \dfrac{1}{N}\right)$, $\Pr(G \mid F') = \left(\dfrac{1}{N-1}\right)$

sehingga :

$$
\begin{aligned}
p_2 &= \Pr(F' \cap G) \\
&= \Pr(F').\Pr(G \mid F') \\
&= \left(1 - \frac{1}{N}\right).\left(\frac{1}{N-1}\right) \\
&= \frac{1}{N}
\end{aligned}
$$

$j = 3$ maka $p_3 = \Pr(u_1$ tidak muncul pada pengambilan pertama,

u_1 tidak muncul pada pengambilan kedua, dan

u_1 muncul pada pengambilan ketiga)

misalkan : $G' =$ kejadian u_1 tidak muncul pada pengambilan kedua

$H =$ kejadian u_1 muncul pada pengambilan ketiga

dengan demikian, $\Pr(G' \mid F') = \left(1 - \dfrac{1}{N-1}\right)$ dan

$\Pr(H \mid F' \cap G') = \left(\dfrac{1}{N-2}\right)$, sehingga :

$$
\begin{aligned}
p_3 &= \Pr(F' \cap G' \cap H) \\
&= \Pr(F' \cap G').\Pr(H \mid F' \cap G') \\
&= \Pr(F').\Pr(G' \mid F').\Pr(H \mid F' \cap G') \\
&= \left(1 - \frac{1}{N}\right).\left(1 - \frac{1}{N-1}\right).\left(\frac{1}{N-2}\right) \\
&= \frac{1}{N}
\end{aligned}
$$

$\vdots$

$$p_j = \frac{1}{N};\ j = 1,2,\ldots,n$$

Misalkan : $\pi_1 = \Pr(u_1$ terpilih dalam sampel)

$D_j =$ kejadian u_1 muncul pada pengambilan ke j

$\pi_1 = \Pr(D_1 \cup D_2 \cup \ldots \cup D_n)$

karena kejadian D_j saling lepas, maka :

$$
\begin{aligned}
\pi_1 &= \Pr(D_1 \cup D_2 \cup \ldots \cup D_n) \\
&= P(D_1) + P(D_2) + \ldots + P(D_n) \\
&= \frac{1}{N} + \frac{1}{N} + \ldots + \frac{1}{N} \\
&= \frac{n}{N}
\end{aligned}
$$

Hal yang sama juga berlaku untuk unit $u_2, u_3, \ldots, u_N$.

Dengan demikian telah terbukti bahwa dalam SRS tanpa pengembalian, probabilitas suatu unit terpilih menjadi anggota sampel adalah sama yaitu $\frac{n}{N}$.

Selanjutnya, akan dibahas mengenai taksiran mean dan total pada SRS, beserta variansi dan taksiran variansi dari taksiran mean dan total.

2.1.1 Taksiran Mean

Misalkan $y_1, y_2, ..., y_n$ menyatakan suatu *simple random sample* dari suatu populasi $u_1, u_2, ..., u_N$ sedemikian sehingga $y_i = u_i z_i$; $i = 1, 2, ..., N$ dimana z_i adalah variabel indikator, yaitu :

$$z_i = \begin{cases} 1, \text{ jika unit ke i terpilih sebagai sampel} \\ 0, \text{ jika unit ke i tidak terpilih sebagai sampel} \end{cases}$$

Mean populasi adalah rata-rata dari semua nilai pengamatan di populasi, yaitu :

$$\mu = \frac{1}{N}\left(u_1 + u_2 + ... + u_N\right) = \frac{1}{N}\sum_{i=1}^{N} u_i$$

taksiran mean pada SRS adalah :

$$\bar{y} = \frac{1}{n}\left(y_1 + y_2 + ... + y_n\right) = \frac{1}{n}\sum_{i=1}^{n} y_i \ .$$

Akan ditunjukkan bahwa $\bar{y}$ adalah taksiran tak bias untuk μ. Telah dibuktikan bahwa dalam SRS tanpa pengembalian, probabilitas suatu unit terpilih sebagai sampel adalah $\frac{n}{N}$, maka :

$$\begin{aligned} E[z_i] &= 0.\Pr(z_i = 0) + 1.\Pr(z_i = 1) \\ &= 1.\Pr(\text{unit ke i terpilih sebagai sampel}) \\ &= \frac{n}{N} \end{aligned} \tag{2.1.1}$$

sehingga

$$\begin{aligned} E[\bar{y}] &= E\left[\frac{1}{n}\sum_{i=1}^{n} y_i\right] \\ &= E\left[\frac{1}{n}\sum_{i=1}^{N} u_i z_i\right] \\ &= \frac{1}{n}\sum_{i=1}^{N} u_i . E[z_i] \\ &= \frac{1}{n}\sum_{i=1}^{N} u_i . \frac{n}{N} \\ &= \frac{1}{N}\sum_{i=1}^{N} u_i \\ &= \mu \end{aligned}$$

Jadi, $\bar{y}$ adalah taksiran tak bias untuk μ.

Untuk mencari variansi dari $\bar{y}$, terlebih dahulu akan dicari variansi dan kovariansi dari variabel indikator z_i, yaitu :

$$\begin{aligned}\operatorname{var}[z_i] &= E[z_i^2] - \left[E[z_i]\right]^2 \\ &= E[z_i] - \left[E[z_i]\right]^2 \\ &= \frac{n}{N} - \left(\frac{n}{N}\right)^2 \\ &= \frac{n}{N}\left(1 - \frac{n}{N}\right) \qquad (2.1.2)\end{aligned}$$

$$\begin{aligned}E[z_i z_j] &= \Pr(z_i = 1, z_j = 1) \\ &= \frac{\binom{N-2}{n-2}}{\binom{N}{n}} \\ &= \frac{n(n-1)}{N(N-1)} \qquad (2.1.3)\end{aligned}$$

$$\begin{aligned}\operatorname{cov}[z_i, z_j] &= E[z_i z_j] - E[z_i].E[z_j] \\ &= \frac{n(n-1)}{N(N-1)} - \frac{n}{N}.\frac{n}{N} \\ &= -\frac{n}{N}.\left(1 - \frac{n}{N}\right).\frac{1}{N-1} \qquad (2.1.4)\end{aligned}$$

Dengan demikian,

$$\begin{aligned}\operatorname{var}[\bar{y}] &= \operatorname{var}\left[\frac{1}{n}\sum_{i=1}^{n} y_i\right] \\ &= \operatorname{var}\left[\frac{1}{n}\sum_{i=1}^{N} u_i z_i\right] \\ &= \frac{1}{n^2}\left[\sum_{i=1}^{N} u_i^2 \operatorname{var}[z_i] + \sum_{i \neq j} u_i u_j \operatorname{cov}[z_i, z_j]\right] \\ &= \frac{1}{n^2}\left[\sum_{i=1}^{N} u_i^2.\frac{n}{N}\left(1 - \frac{n}{N}\right) + \sum_{i \neq j} u_i u_j.\left(-\frac{n}{N}.\left(1 - \frac{n}{N}\right).\frac{1}{N-1}\right)\right] \\ &= \frac{1}{n^2}.\frac{n}{N}\left(1 - \frac{n}{N}\right)\left[\sum_{i=1}^{N} u_i^2 - \sum_{i \neq j} \frac{u_i u_j}{N-1}\right], \text{ karena } \sum_{i \neq j} u_i u_j = \left(\sum_{i=1}^{N} u_i\right)^2 - \sum_{i=1}^{N} u_i^2, \text{ maka}\end{aligned}$$

$$= \frac{1}{nN}\left(1-\frac{n}{N}\right)\left[\sum_{i=1}^{N}u_i^2-\left\{\left(\sum_{i=1}^{N}u_i\right)^2-\sum_{i=1}^{N}u_i^2\right\}\Big/(N-1)\right]$$

$$= \frac{1}{nN}\left(1-\frac{n}{N}\right)\left[\sum_{i=1}^{N}u_i^2\left(1+\frac{1}{N-1}\right)-\frac{1}{N-1}\left(\sum_{i=1}^{N}u_i\right)^2\right]$$

$$= \frac{1}{nN}\left(1-\frac{n}{N}\right).\frac{1}{N-1}\left[N.\sum_{i=1}^{N}u_i^2-\left(\sum_{i=1}^{N}u_i\right)^2\right]$$

$$= \frac{1}{n}\left(\frac{N-n}{N}\right).\frac{1}{N-1}\left[\sum_{i=1}^{N}u_i^2-\frac{1}{N}\left(\sum_{i=1}^{N}u_i\right)^2\right]$$

$$= \frac{1}{n}\left(\frac{N-n}{N}\right).\frac{1}{N-1}\left[\sum_{i=1}^{N}\left(u_i-\mu\right)^2\right]$$

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

Dengan demikian,

$$\mathrm{var}[\bar{y}] = \frac{\sigma^2}{n}\left(\frac{N-n}{N}\right) \qquad (2.1.5)$$

dengan $\sigma^2 = \sum_{i=1}^{N}\frac{\left(u_i-\mu\right)^2}{N-1}$. Taksiran dari variansi tersebut adalah

$$\widehat{\mathrm{var}}[\bar{y}] = \frac{s^2}{n}\left(\frac{N-n}{N}\right) \qquad (2.1.6)$$

dimana $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2$. Selanjutnya akan ditunjukkan bahwa taksiran variansi tersebut adalah taksiran yang tak bias, tetapi akan ditunjukkan dulu bahwa $E[s^2] = \sigma^2$

$$
\begin{aligned}
E[s^2] &= E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2\right] \\
&= \frac{1}{n-1}.E\left[\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2\right] \\
&= \frac{1}{n-1}.E\left[\sum_{i=1}^{n}\left((y_i-\mu)-(\bar{y}-\mu)\right)^2\right] \\
&= \frac{1}{n-1}.E\left[\sum_{i=1}^{n}\left(y_i-\mu\right)^2-n\left(\bar{y}-\mu\right)^2\right] \\
&= \frac{1}{n-1}.\left\{E\left[\sum_{i=1}^{n}\left(y_i-\mu\right)^2\right]-nE\left[\left(\bar{y}-\mu\right)^2\right]\right\} \\
&= \frac{1}{n-1}.\left\{\sum_{i=1}^{n}E\left[\left(y_i-\mu\right)^2\right]-nE\left[\left(\bar{y}-\mu\right)^2\right]\right\} \\
&= \frac{1}{n-1}.\left\{\sum_{i=1}^{n}\text{var}[y_i]-n\,\text{var}[\bar{y}]\right\} \\
&= \frac{1}{n-1}.\left\{n\sigma^2-n.\text{var}[\frac{1}{n}\sum_{i=1}^{n}y_i]\right\} \\
&= \frac{1}{n-1}.\left\{n\sigma^2-\frac{1}{n}.\text{var}[\sum_{i=1}^{n}y_i]\right\} \\
&= \frac{1}{n-1}.\left\{n\sigma^2-\frac{1}{n}.\sum_{i=1}^{n}\text{var}[y_i]\right\} \\
&= \frac{1}{n-1}.\left\{n\sigma^2-\frac{1}{n}.n\sigma^2\right\} \\
&= \frac{1}{n-1}.\left\{(n-1).\sigma^2\right\} \\
&= \sigma^2 \qquad (2.1.7)
\end{aligned}
$$

maka,

$$
\begin{aligned}
E[\widehat{\text{var}[\bar{y}]}] &= E\left[\frac{s^2}{n}\left(\frac{N-n}{N}\right)\right] \\
&= \frac{N-n}{n.N}.E[s^2] \\
&= \frac{\sigma^2}{n}\left(\frac{N-n}{N}\right) \\
&= \text{var}[\bar{y}] \qquad (2.1.8)
\end{aligned}
$$

Dengan demikian, taksiran variansi tersebut adalah taksiran yang tak bias.

2.1.2 Taksiran Total

Total populasi adalah

$$\tau = N\mu = \sum_{i=1}^{N} u_i$$

sedangkan taksirannya adalah

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

lalu, karena

$$E[\hat{\tau}] = E[N\bar{y}] = NE[\bar{y}] = N\mu = \tau\,.$$

maka taksiran tersebut merupakan taksiran yang tak bias untuk total.

Variansi untuk taksiran total adalah

$$\text{var}[\hat{\tau}] = \text{var}[N\bar{y}] = N^2\,\text{var}[\bar{y}] = N^2.\frac{\sigma^2}{n}\left(\frac{N-n}{N}\right) = N(N-n)\frac{\sigma^2}{n}$$

dengan taksirannya adalah

$$\widehat{\text{var}}[\hat{\tau}] = \widehat{\text{var}}[N\bar{y}] = N^2\,\widehat{\text{var}}[\bar{y}] = N^2.\frac{s^2}{n}\left(\frac{N-n}{N}\right) = N(N-n)\frac{s^2}{n}$$

lalu, karena :

$$E[\widehat{\text{var}}[\hat{\tau}]] = E[\widehat{\text{var}}[N\bar{y}]] = N^2E[\widehat{\text{var}}[\bar{y}]] = N^2\,\text{var}[\bar{y}] = \text{var}[N\bar{y}] = \text{var}[\hat{\tau}]$$

maka dimana taksiran variansi tersebut merupakan taksiran yang tak bias untuk variansi dari taksiran total.

2.2 Two Stage Sampling

Dalam pengambilan sampel secara *simple random sampling* (SRS) tidak selamanya efektif, walaupun SRS merupakan metode sampling yang paling sederhana. Misalkan dalam penelitian yang mencakup daerah yang sangat luas, seperti hutan. Jika pengambilan sampel dilakukan secara SRS, maka peneliti akan memerlukan banyak tenaga, waktu dan biaya untuk meneliti unit-unit yang terpilih sebagai sampel karena kemungkinan besar unit-unit tersebut terletak berjauhan. Agar lebih efektif, pengambilan sampel dapat dilakukan dengan dua tahap. Pada tahap pertama pengambilan sampel dilakukan pada unit primer, kemudian pada tahap kedua pengambilan sampel dilakukan pada unit sekunder yang berada di dalam unit primer yang terpilih sebagai sampel. Pengambilan sampel seperti ini disebut *Two Stage Sampling*. Pada tugas akhir ini, pengambilan sampel pada masing-masing tahap hanya dilakukan dengan SRS tanpa pengembalian.

Misalkan M adalah banyaknya unit primer pada populasi dan N_i adalah banyaknya unit sekunder pada unit primer ke i. Misalkan τ_{ij} menyatakan total pengukuran pada unit sekunder ke j dalam unit primer ke i pada populasi, maka total pengukuran pada unit primer ke i adalah $\tau_i = \sum_{j=1}^{N_i} \tau_{ij}$, sedangkan total populasi adalah $\tau = \sum_{i=1}^{M} \sum_{j=1}^{N_i} \tau_{ij} = \sum_{i=1}^{M} \tau_i$. Selanjutnya, mean pada unit primer ke i adalah $\mu_i = \tau_i / N_i$, sedangkan mean populasi adalah $\mu = \tau / N$, dimana $N = \sum_{i=1}^{M} N_i$ yaitu banyaknya unit sekunder pada populasi.

2.2.1 Taksiran Mean Dan Total Populasi

Misalkan m menyatakan banyaknya unit primer yang terpilih sebagai sampel pada tahap awal pengambilan sampel, sebut s_1 sebagai kumpulan unit primer dalam sampel. Kemudian n_i menyatakan banyaknya unit sekunder pada unit primer ke i yang terpilih sebagai sampel pada tahap kedua pengambilan sampel, y_{ij} adalah total pengukuran pada unit sekunder ke j di dalam unit primer ke i, yang terpilih sebagai sampel. Karena n_i unit sekunder dipilih secara SRS pada tahap kedua, maka taksiran mean dan total pada unit primer ke i adalah

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

$$\hat{\tau}_i = N_i \hat{\mu}_i = \frac{N_i}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

Selanjutnya, taksiran mean dan total populasi adalah

$$\hat{\mu} = \frac{M}{Nm} \sum_{i=1}^{m} \hat{\tau}_i$$

$$\hat{\tau} = N.\hat{\mu} = \frac{M}{m} \sum_{i=1}^{m} \hat{\tau}_i$$

Misalkan s_1 adalah himpunan semua unit primer yang terpilih sebagai sampel. Akan ditunjukkan bahwa diberikan himpunan s_1, taksiran total pada unit primer ke *i* adalah taksiran tak bias untuk total pada unit primer ke *i.* Dengan perkataan lain, akan ditunjukkan bahwa $E[\hat{\tau}_i \mid s_1] = \tau_i$ sebagai berikut:

Misalkan z_{ij} adalah variabel indikator, yaitu :

$$z_{ij} = \begin{cases} 1, \text{ jika unit sampling ke j pada unit primer ke i terpilih sebagai sampel} \\ 0, \text{ jika unit sampling ke j pada unit primer ke i tidak terpilih sebagai sampel} \end{cases}$$

dimana $\Pr(z_{ij} = 1 \mid s_1) = \dfrac{n_i}{N_i}$, maka :

$$\begin{aligned} E[\hat{\tau}_i \mid s_1] &= E\left[\frac{N_i}{n_i}\sum_{j=1}^{n_i} y_{ij} \mid s_1\right] \\ &= E\left[\frac{N_i}{n_i}\sum_{j=1}^{N_i} \tau_{ij} z_{ij} \mid s_1\right] \\ &= \frac{N_i}{n_i}\sum_{j=1}^{N_i} \tau_{ij} E[z_{ij} \mid s_1] \\ &= \frac{N_i}{n_i}\sum_{j=1}^{N_i} \tau_{ij} \left\{0.\Pr(z_{ij} = 0 \mid s_1) + 1.\Pr(z_{ij} = 1 \mid s_1)\right\} \\ &= \frac{N_i}{n_i}\sum_{j=1}^{N_i} \tau_{ij} \frac{n_i}{N_i} \\ &= \sum_{j=1}^{N_i} \tau_{ij} \\ &= \tau_i \end{aligned}$$

Kemudian, akan ditunjukkan bahwa $\hat{\mu}$ dan $\hat{\tau}$ adalah taksiran tak bias untuk mean dan total populasi. Misalkan z_i adalah variabel indikator, yaitu :

$$z_i = \begin{cases} 1, \text{ jika unit primer ke i terpilih sebagai sampel} \\ 0, \text{ jika unit primer ke i tidak terpilih sebagai sampel} \end{cases}$$

dengan $\Pr(z_i = 1) = \frac{m}{M}$, maka :

$$
\begin{aligned}
E[\hat{\mu}] &= E[E(\hat{\mu} \mid s_1)] \\
&= E\left[E\left(\frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right)\right] \\
&= E\left[\frac{M}{Nm}\sum_{i=1}^{m}E[\hat{\tau}_i \mid s_1]\right] \\
&= E\left[\frac{M}{Nm}\sum_{i=1}^{m}\tau_i\right] \\
&= E\left[\frac{M}{Nm}\sum_{i=1}^{M}\tau_i z_i\right] \\
&= \frac{M}{Nm}\sum_{i=1}^{M}\tau_i E[z_i] \\
&= \frac{M}{Nm}\sum_{i=1}^{M}\tau_i . \frac{m}{M} \\
&= \frac{1}{N}\sum_{i=1}^{M}\tau_i \\
&= \frac{\tau}{N} \\
&= \mu
\end{aligned}
$$

sedangkan untuk taksiran total

$$
\begin{aligned}
E[\hat{\tau}] &= E[E(\hat{\tau} \mid s_1)] \\
&= E[E(N.\hat{\mu} \mid s_1)] \\
&= N.E[E(\hat{\mu} \mid s_1)] \\
&= N.\mu \\
&= \tau
\end{aligned}
$$

Jadi, $\hat{\mu}$ dan $\hat{\tau}$ adalah taksiran tak bias untuk μ dan τ .

Selanjutnya akan dicari variansi dari $\hat{\mu}$ dan $\hat{\tau}$. Variansi dari $\hat{\mu}$ dapat diperoleh dengan menggunakan dekomposisi sebagai berikut :

$$\text{var}[\hat{\mu}] = \text{var}[E(\hat{\mu} \mid s_1)] + E[\text{var}(\hat{\mu} \mid s_1)] \qquad (2.2.2)$$

Ada dua bagian pada (2.2.2) yang akan dicari

Bagian pertama dari ruas kanan persamaan (2.2.2) adalah

$$
\begin{aligned}
\operatorname{var}[E(\hat{\mu} \mid s)] &= \operatorname{var}\left[E\left(\frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right)\right] \\
&= \operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}E[\hat{\tau}_i \mid s_1]\right] \\
&= \operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}\tau_i\right] \\
&= \frac{M^2}{N^2}\operatorname{var}\left[\frac{1}{m}\sum_{i=1}^{m}\tau_i\right] \\
&= \frac{M^2}{N^2}\operatorname{var}\left[\bar{\tau}\right] \\
&= \frac{M^2}{N^2}.\left(\frac{M-m}{M}\right)\frac{\sigma_u^2}{m} \\
&= \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m}
\end{aligned}
$$

dimana σ_u^2 adalah variansi antar unit primer, yaitu $\sigma_u^2 = \frac{1}{M-1}\sum_{i=1}^{M}\left(\tau_i - \mu_1\right)^2$,

dengan $\mu_1 = \frac{1}{M}\sum_{i=1}^{M}\tau_i$ yaitu mean per unit primer.

Bagian kedua dari ruas kanan persamaan (2.2.2) adalah

$$
\begin{aligned}
E[\mathrm{var}(\hat{\mu} \mid s_1)] &= E\left[\mathrm{var}\left(\frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right)\right] \\
&= E\left[\left(\frac{M}{Nm}\right)^2 \sum_{i=1}^{m}\mathrm{var}[\hat{\tau}_i \mid s_1]\right] \\
&= E\left[\left(\frac{M}{Nm}\right)^2 \sum_{i=1}^{m} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right] \\
&= E\left[\left(\frac{M}{Nm}\right)^2 \sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.z_i\right] \\
&= \left(\frac{M}{Nm}\right)^2 \sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.E[z_i] \\
&= \left(\frac{M}{Nm}\right)^2 \sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.\frac{m}{M} \\
&= \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}
\end{aligned}
$$

dimana σ_i^2 adalah variansi di dalam unit primer ke *i*, yaitu

$\sigma_i^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(\tau_{ij} - \mu_i\right)^2$, dengan $\mu_i = \tau_i / N_i$ yaitu mean di dalam unit primer ke *i*, untuk *i = 1, 2, …, N.*

Jadi, variansi dari $\hat{\mu}$ adalah

$$
\mathrm{var}[\hat{\mu}] = \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i} \qquad (2.2.3)
$$

Taksiran dari (2.2.3) adalah

$$
\widehat{\mathrm{var}}[\hat{\mu}] = \frac{1}{N^2}M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i} \qquad (2.2.4)
$$

dimana $s_u^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(\hat{\tau}_i - \hat{\mu}_1\right)^2$, dengan $\hat{\mu}_1 = \frac{1}{m}\sum_{i=1}^{m}\hat{\tau}_i$ dan

$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(y_{ij} - \hat{\mu}_i\right)^2$, untuk *i= 1, 2,…m.*

Akan ditunjukkan bahwa (2.2.4) adalah taksiran tak bias untuk (2.2.3)

$$
\begin{aligned}
E[\widehat{\text{var}}[\hat{\mu}]] &= E[E(\widehat{\text{var}}[\hat{\mu}] \mid s_1)] \\
&= E\left[E\left[\left(\frac{1}{N^2}M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}N_i(N_i - n_i)\frac{s_i^2}{n_i}\right) \mid s_1\right]\right] \\
&= E\left[E\left(\frac{1}{N^2}M(M-m)\frac{s_u^2}{m} \mid s_1\right) + E\left(\frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}N_i(N_i - n_i)\frac{s_i^2}{n_i} \mid s_1\right)\right] \\
&= E\left[E\left(\frac{1}{N^2}M(M-m)\frac{s_u^2}{m} \mid s_1\right)\right] + E\left[E\left(\frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}N_i(N_i - n_i)\frac{s_i^2}{n_i} \mid s_1\right)\right] \\
&= \underbrace{\frac{1}{N^2}\frac{M(M-m)}{(m-1)m}\underbrace{E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i - \hat{\mu}_1\right)^2 \mid s_1\right]\right]}_{(*)}}_{\text{(I)}} + \frac{1}{N^2}\underbrace{E\left[\frac{M}{m}\sum_{i=1}^{m}N_i(N_i - n_i)\frac{E[s_i^2 \mid s_1]}{n_i}\right]}_{\text{(II)}}
\end{aligned}
$$

Ada dua bagian yang akan dicari, pada bagian pertama (I) akan dicari dahulu bentuk (*) yaitu

$$
\begin{aligned}
E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i - \hat{\mu}_1\right)^2 \mid s_1\right]\right] &= E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i^2 - 2\hat{\tau}_i\hat{\mu}_1 + \left(\hat{\mu}_1\right)^2\right) \mid s_1\right]\right] \\
&= E\left[E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i^2 - 2\hat{\mu}_1\sum_{i=1}^{m}\hat{\tau}_i + m\left(\hat{\mu}_1\right)^2\right) \mid s_1\right]\right] \\
&= E\left[E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i^2 - 2\frac{1}{m}\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 + \frac{1}{m}\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2\right) \mid s_1\right]\right] \\
&= E\left[E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i^2 - \frac{1}{m}\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2\right) \mid s_1\right]\right]
\end{aligned}
$$

$$= E\left[E\left(\sum_{i=1}^{m}\hat{\tau}_i^2 \mid s_1\right) - \frac{1}{m}E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 \mid s_1\right]\right]$$

$$= \underbrace{E\left[E\left(\sum_{i=1}^{m}\hat{\tau}_i^2 \mid s_1\right)\right]}_{(a)} - \underbrace{E\left[\frac{1}{m}E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 \mid s_1\right]\right]}_{(b)}$$

bentuk (a) dan (b) dapat dicari sebagai berikut :

(a) $$E\left[E\left(\sum_{i=1}^{m}\hat{\tau}_i^2 \mid s_1\right)\right] = E\left[\sum_{i=1}^{m}E\left(\hat{\tau}_i^2 \mid s_1\right)\right]$$

$$= E\left[\sum_{i=1}^{m}\left(\left\{E[\hat{\tau}_i \mid s_1]\right\}^2 + \text{var}[\hat{\tau}_i \mid s_1]\right)\right]$$

$$= E\left[\sum_{i=1}^{m}\left(\tau_i^2 + N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right)\right]$$

$$= E\left[\sum_{i=1}^{m}\tau_i^2\right] + E\left[\sum_{i=1}^{m}N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right]$$

$$= E\left[\sum_{i=1}^{M}\tau_i^2 z_i\right] + E\left[\sum_{i=1}^{M}N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.z_i\right]$$

$$= \frac{m}{M}\sum_{i=1}^{M}\tau_i^2 + \frac{m}{M}\sum_{i=1}^{M}N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

(b) $$\left[\frac{1}{m}E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 \mid s_1\right]\right] = E\left[\frac{1}{m}\left(\left\{E\left[\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right]\right\}^2 + \text{var}\left[\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right]\right)\right]$$

$$= E\left[\frac{1}{m}\left(\left\{\sum_{i=1}^{m}E\left(\hat{\tau}_i \mid s_1\right)\right\}^2 + \sum_{i=1}^{m}\text{var}[\hat{\tau}_i \mid s_1]\right)\right]$$

$$= E\left[\frac{1}{m}\left(\left\{\sum_{i=1}^{m}\tau_i\right\}^2 + \sum_{i=1}^{m}N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right)\right]$$

$$= E\left(m\bar{\tau}^2 + \frac{1}{m}\sum_{i=1}^{m}N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right)$$

$$= mE[\bar{\tau}^2] + \frac{1}{m}E\left[\sum_{i=1}^{m}N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right]$$

$$= m\left(\left\{E[\bar{\tau}]\right\}^2 + \text{var}[\bar{\tau}]\right) + \frac{1}{m}E\left[\sum_{i=1}^{M}N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.z_i\right]$$

$$= m\left(\left\{E[\frac{1}{m}\sum_{i=1}^{M}\tau_i z_i]\right\}^2 + \left(\frac{M-m}{M}\right)\frac{\sigma_u^2}{m}\right) + \frac{1}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.E\left[z_i\right]$$

$$= m\left(\left\{\frac{1}{M}\sum_{i=1}^{M}\tau_i\right\}^2 + \left(\frac{M-m}{M}\right)\frac{\sigma_u^2}{m}\right) + \frac{1}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= m\mu_1^2 + \left(\frac{M-m}{M}\right)\sigma_u^2 + \frac{1}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

sehingga (*) menjadi

$$E\left[E\left[\sum_{i=1}^{m}(\hat{\tau}_i - \hat{\mu}_1)^2 \mid s_1\right]\right] = E\left[\sum_{i=1}^{m} E_2[\hat{\tau}_i^2 \mid s_1]\right] - E\left[\frac{1}{m}E_2\left[\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 \mid s_1\right]\right]$$

$$= \left\{\frac{m}{M}\sum_{i=1}^{M}\tau_i^2 + \frac{m}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right\}$$

$$- \left\{m\mu_1^2 + \left(\frac{M-m}{M}\right)\sigma_u^2 + \frac{1}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right\}$$

$$= \frac{m}{M}\sum_{i=1}^{M}\tau_i^2 - m\mu_1^2 - \left(\frac{M-m}{M}\right)\sigma_u^2 + \frac{m}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$- \frac{1}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= \frac{m}{M}\left(\sum_{i=1}^{M}\tau_i^2 - M\mu_1^2\right) - \left(1 - \frac{m}{M}\right)\sigma_u^2$$

$$+ \frac{(m-1)}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= \frac{m}{M}.\frac{M-1}{M-1}.\sum_{i=1}^{M}(\tau_i - \mu_1)^2 - \sigma_u^2 + \frac{m}{M}\sigma_u^2$$

$$+ \frac{(m-1)}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= \frac{m}{M}.(M-1)\sigma_u^2 - \sigma_u^2 + \frac{m}{M}\sigma_u^2 + \frac{(m-1)}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= m\sigma_u^2 - \frac{m}{M}\sigma_u^2 - \sigma_u^2 + \frac{m}{M}\sigma_u^2 + \frac{(m-1)}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= m\sigma_u^2 - \sigma_u^2 + \frac{(m-1)}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= (m-1)\sigma_u^2 + \frac{(m-1)}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= (m-1)\left[\sigma_u^2 + \frac{1}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right]$$

maka bentuk (I) adalah

$$\frac{1}{N^2}\frac{M(M-m)}{(m-1)m}E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i - \hat{\mu}_1\right)^2 | s_1\right]\right]$$

$$= \frac{1}{N^2}\frac{M(M-m)}{(m-1)m}.(m-1)\left[\sigma_u^2 + \frac{1}{M}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right]$$

$$= \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{\left(M-m\right)}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

Selanjutnya akan dicari bentuk kedua (II) yaitu :

$$\frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{E[s_i^2 | s_1]}{n_i}\right] = \frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}\right]$$

$$= \frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.z_i\right]$$

$$= \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.E[z_i]$$

$$= \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}.\frac{m}{M}$$

$$= \frac{1}{N^2}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

maka,

$$E[\widehat{\text{var}}[\hat{\mu}]] = \left(\frac{1}{N^2} M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{(M-m)}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i} \right)$$

$$+ \left(\frac{1}{N^2}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i} \right)$$

$$= \frac{1}{N^2} M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\left(\frac{M}{m} - 1\right)\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i} + \frac{1}{N^2}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= \frac{1}{N^2} M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$= \text{var}[\hat{\mu}]$$

Jadi, karena $E[\widehat{\text{var}}[\hat{\mu}]] = \text{var}[\hat{\mu}]$ maka $\widehat{\text{var}}[\hat{\mu}]$ merupakan taksiran tak bias untuk $\text{var}[\hat{\mu}]$.

Selanjutnya, taksiran variansi untuk taksiran total adalah sebagai berikut

$$\widehat{\text{var}}[\hat{\tau}] = \widehat{\text{var}}[N.\hat{\mu}]$$

$$= N^2.\widehat{\text{var}}[\hat{\mu}]$$

$$= M(M-m)\frac{s_u^2}{m} + \frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i}$$

Kemudian, karena

$$E[\widehat{\text{var}}[\hat{\tau}]] = E[\widehat{\text{var}}[N.\hat{\mu}]] = N^2 E[\widehat{\text{var}}[\hat{\mu}]] = N^2.\text{var}[\hat{\mu}] = \text{var}[N.\hat{\mu}] = \text{var}[\hat{\tau}]$$

maka, taksiran variansi tersebut merupakan taksiran tak bias untuk variansi dari taksiran total.

2.3 *Unequal Probability Sampling*

Jika dalam SRS probabilitas terpilihnya suatu unit adalah sama, tetapi adakalanya unit-unit dalam populasi memiliki probabilitas yang berbeda untuk terpilih sebagai sampel. Untuk kasus dimana unit-unit memiliki peluang yang berbeda untuk terpilih sebagai sampel, ada dua jenis taksiran yang dapat digunakan untuk menaksir mean dan total populasi, yaitu taksiran Horvitz-Thompson dan taksiran Hansen-Hurwitz.

2.3.1 Taksiran Horvitz-Thompson

Taksiran ini dapat digunakan ketika pengambilan sampel dilakukan dengan atau tanpa pengembalian. Misalkan dipunyai unit pada populasi $\{u_1, u_2, ..., u_N\}$ berukuran *N*, kemudian diambil suatu sampel $\{y_1, y_2, ..., y_n\}$ berukuran *n* dimana probabilitas terpilihnya unit ke i sebagai sampel tidak sama, sebut π_i. Taksiran Horvitz-Thompson untuk mean adalah

$$\hat{\mu}_{HT} = \frac{1}{N}\sum_{i=1}^{n}\frac{y_i}{\pi_i}$$

Akan ditunjukkan bahwa taksiran ini adalah taksiran yang tak bias untuk mean. Sebelumnya, didefinisikan suatu variabel indikator z_i yang bernilai 1 jika unit ke *i* terpilih menjadi sampel dan bernilai 0 jika unit ke *i* tidak terpilih menjadi sampel. Jadi,

$$E[z_i] = 0.\Pr(z_i = 0) + 1.\Pr(z_i = 1)$$
$$= 1.\Pr(\text{unit ke } i \text{ terpilih sebagai sampel})$$
$$= \pi_i$$

sehingga

$$E[\hat{\mu}_{HT}] = E\left[\frac{1}{N}\sum_{i=1}^{n}\frac{y_i}{\pi_i}\right]$$
$$= E\left[\frac{1}{N}\sum_{i=1}^{N}\frac{u_i z_i}{\pi_i}\right]$$
$$= \frac{1}{N}\sum_{i=1}^{N}\frac{u_i}{\pi_i}E[z_i]$$
$$= \frac{1}{N}\sum_{i=1}^{N}\frac{u_i}{\pi_i}.\pi_i$$
$$= \frac{1}{N}\sum_{i=1}^{N}u_i$$
$$= \mu \qquad (2.3.1)$$

maka, taksiran tersebut merupakan taksiran yang tak bias untuk mean.

Untuk mencari variansi dari taksiran Horvitz-Thompson, terlebih dahulu akan dicari variansi dan kovariansi dari variabel indikator z_i , yaitu :

$$E[z_i^{\,2}] = 0^2.\Pr(z_i = 0) + 1^2.\Pr(z_i = 1)$$
$$= 1.\Pr(\text{unit ke i terpilih menjadi sampel})$$
$$= \pi_i$$

$$\text{var}[z_i] = E[z_i^{\,2}] - [E[z_i]]^2$$
$$= \pi_i - \pi_i^{\,2}$$
$$= \pi_i(1 - \pi_i)$$

$$\text{cov}[z_i, z_j] = E[z_i z_j] - E[z_i].E[z_j]$$
$$= \pi_{ij} - \pi_i \pi_j$$

Dimana π_{ij} adalah probabilitas unit ke i dan ke j terpilih menjadi sampel.

Variansi dan taksiran variansi dari taksiran mean Horvitz-Thompson adalah

$$\begin{aligned}
\mathrm{var}[\hat{\mu}_{HT}] &= \mathrm{var}\left[\frac{1}{N}\sum_{i=1}^{n}\frac{y_i}{\pi_i}\right] \\
&= \mathrm{var}\left[\frac{1}{N}\sum_{i=1}^{N}\frac{u_i z_i}{\pi_i}\right] \\
&= \frac{1}{N^2}\mathrm{var}\left[\sum_{i=1}^{N}\frac{u_i z_i}{\pi_i}\right] \\
&= \frac{1}{N^2}\left\{\sum_{i=1}^{N}\left(\frac{u_i}{\pi_i}\right)^2 \mathrm{var}[z_i] + \sum_{i=1}^{N}\sum_{j\neq i}\frac{u_i u_j}{\pi_i \pi_j}\mathrm{cov}[z_i, z_j]\right\} \\
&= \frac{1}{N^2}\left\{\sum_{i=1}^{N}\left(\frac{u_i}{\pi_i}\right)^2 .\pi_i(1-\pi_i) + \sum_{i=1}^{N}\sum_{j\neq i}\frac{u_i u_j}{\pi_i \pi_j}.\left(\pi_{ij} - \pi_i \pi_j\right)\right\} \\
&= \frac{1}{N^2}\left\{\sum_{i=1}^{N} u_i^2 \frac{(1-\pi_i)}{\pi_i} + \sum_{i=1}^{N}\sum_{j\neq i} u_i u_j\left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}\right)\right\} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}\right) u_i u_j
\end{aligned}$$

$$\widehat{\mathrm{var}}[\hat{\mu}_{HT}] = \frac{1}{N^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j \pi_{ij}}\right) y_i y_j$$

Akan dibuktikan bahwa taksiran variansi tersebut adalah taksiran yang tak bias untuk variansi dari taksiran mean:

$$E[\widehat{\text{var}}[\hat{\mu}_{HT}]] = E\left[\frac{1}{N^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j\pi_{ij}}\right)y_iy_j\right]$$
$$= E\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j\pi_{ij}}\right)u_iu_jz_iz_j\right]$$
$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j\pi_{ij}}\right)u_iu_j.E[z_iz_j]$$
$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j\pi_{ij}}\right)u_iu_j.\pi_{ij}$$
$$= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}\right)u_iu_j$$
$$= \text{var}[\hat{\mu}_{HT}] \qquad (2.3.2)$$

Karena $E[\widehat{\text{var}}[\hat{\mu}_{HT}]] = \text{var}[\hat{\mu}_{HT}]$, maka taksiran variansi tersebut merupakan taksiran yang tak bias.

Selanjutnya, taksiran Horvitz-Thompson untuk total adalah

$$\hat{\tau} = N.\hat{\mu}_{HT} = \sum_{i=1}^{n}\frac{y_i}{\pi_i} = \sum_{i=1}^{N}\frac{u_iz_i}{\pi_i}.$$

Berdasarkan (2.3.1), diperoleh

$$E[\hat{\tau}_{HT}] = E[N.\hat{\mu}_{HT}] = N.E[\hat{\mu}_{HT}] = N.\mu = \tau$$

sehingga taksiran Horvitz-Thompson untuk total tersebut merupakan taksiran tak bias untuk total.

Kemudian, variansi dari taksiran total dapat diperoleh sebagai berikut

$$\text{var}[\hat{\tau}_{HT}] = \text{var}[N\hat{\mu}_{HT}] = N^2\,\text{var}[\hat{\mu}_{HT}] = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j}\right)u_iu_j$$

dan taksiran variansi

$$\widehat{\text{var}}[\hat{\tau}_{HT}] = \widehat{\text{var}}[N\hat{\mu}_{HT}] = N^2\widehat{\text{var}}[\hat{\mu}_{HT}] = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j\pi_{ij}}\right)y_i y_j \ .$$

Berdasarkan (2.3.2), dapat diperoleh

$$E[\widehat{\text{var}}[\hat{\tau}_{HT}]] = E[\widehat{\text{var}}[N.\hat{\mu}_{HT}]] = N^2E[\widehat{\text{var}}[\hat{\mu}_{HT}]] = N^2\,\text{var}[\hat{\mu}_{HT}] = \text{var}[N.\hat{\mu}_{HT}] = \text{var}[\hat{\tau}_{HT}]$$

sehingga $\widehat{\text{var}}[\hat{\tau}_{HT}]$ adalah taksiran tak bias untuk $\text{var}[\hat{\tau}_{HT}]$.

2.3.2 Taksiran Hansen-Hurwitz

Taksiran ini biasa digunakan ketika pengambilan sampel dilakukan dengan pengembalian. Misalkan dari *N* unit pada populasi $\{u_1, u_2, ..., u_N\}$ diambil suatu sampel berukuran *n* $\{y_1, y_2, ..., y_n\}$ dengan pengembalian. Misalkan p_i adalah probabilitas terpilihnya unit ke *i*. Taksiran Hansen-Hurwitz untuk mean adalah

$$\hat{\mu}_{HH} = \frac{1}{nN}\sum_{i=1}^{n}\frac{y_i}{p_i}$$

Misalkan f_i menyatakan berapa kali unit *i* terpilih sebagai sampel, dengan demikian f_i berdistribusi binomial $(f_i \sim b(n, p_i))$, sehingga taksiran Hansen-Hurwitz dapat pula dinyatakan sebagai :

$$\hat{\mu}_{HH} = \frac{1}{nN}\sum_{i=1}^{n}\frac{y_i}{p_i} = \frac{1}{nN}\sum_{i=1}^{N}\frac{u_i f_i}{p_i}$$

Karena $f_i \sim b(n, p_i)$, maka $E[f_i] = np_i$, sehingga taksiran Hansen-Hurwitz dapat pula dinyatakan dengan

$$\begin{aligned}\hat{\mu}_{HH} &= \frac{1}{nN}\sum_{i=1}^{n}\frac{y_i}{p_i}\\ &= \frac{1}{nN}\sum_{i=1}^{N}\frac{u_i f_i}{p_i}\\ &= \frac{1}{N}\sum_{i=1}^{N}u_i\,\frac{f_i}{E[f_i]}\end{aligned}$$

Akan dibuktikan bahwa taksiran tersebut adalah taksiran yang tak bias untuk mean, yaitu :

$$\begin{aligned}E[\hat{\mu}_{HH}] &= E\left[\frac{1}{nN}\sum_{i=1}^{n}\frac{y_i}{p_i}\right]\\ &= E\left[\frac{1}{N}\sum_{i=1}^{N}u_i\,\frac{f_i}{E[f_i]}\right]\\ &= \frac{1}{N}\sum_{i=1}^{N}u_i\,\frac{E[f_i]}{E[f_i]}\\ &= \frac{1}{N}\sum_{i=1}^{N}u_i\\ &= \mu\end{aligned} \tag{2.3.3}$$

Jadi, $\hat{\mu}_{HH}$ merupakan taksiran yang tak bias untuk mean.

Untuk mencari taksiran variansi dari $\hat{\mu}_{HH}$ yang tak bias, didefinisikan variabel random T dengan nilai $\frac{u_i}{Np_i}$ dan $\Pr\left(T = \frac{u_i}{Np_i}\right) = p_i,\ (i = 1,2,\ldots,N)$,

maka

$$\mu_T = E[T] = \sum \frac{u_i}{Np_i}.\Pr\left(T = \frac{u_i}{Np_i}\right)$$
$$= \sum_{i=1}^{N} \frac{u_i}{Np_i}.p_i$$
$$= \mu$$

$$\sigma_T^2 = \mathrm{var}[T] = E[(T - \mu_T)^2]$$
$$= \sum \left(\frac{u_i}{Np_i} - \mu_T\right)^2.\Pr\left(T = \frac{u_i}{Np_i}\right)$$
$$= \sum_{i=1}^{N} \left(\frac{u_i}{Np_i} - \mu\right)^2.p_i$$

Misalkan *t* adalah sampel random dengan distribusi dari variabel random *T*, dengan nilai $\frac{y_i}{Np_i}$; $i = 1,2,...,n$, sehingga

$$\hat{\mu}_{HH} = \frac{1}{nN}\sum_{i=1}^{n}\frac{y_i}{p_i} = \frac{1}{n}\sum_{i=1}^{n} t_i = \bar{t}$$

dengan

$$E[\hat{\mu}_{HH}] = E[\bar{t}\,] = \mu_T = \mu$$

Variansi dan taksiran variansi dari $\hat{\mu}_{HH}$ adalah

$$\mathrm{var}[\hat{\mu}_{HH}] = \mathrm{var}\left[\frac{1}{n}\sum_{i=1}^{n} t_i\right] = \frac{1}{n^2}\mathrm{var}\left[\sum_{i=1}^{n} t_i\right]$$
$$= \frac{1}{n^2}.n.\mathrm{var}[t_i]$$
$$= \frac{\sigma_T^2}{n}$$
$$= \frac{1}{n}\sum_{i=1}^{N}\left(\frac{u_i}{Np_i} - \mu\right)^2.p_i$$

$$
\begin{aligned}
\widehat{\text{var}}[\hat{\mu}_{HH}] &= \frac{s_T^2}{n} \\
&= \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(t_i - \bar{t}\right)^2 \\
&= \frac{1}{n(n-1)}\sum_{i=1}^{n}\left(\frac{y_i}{Np_i} - \hat{\mu}_{HH}\right)^2
\end{aligned}
$$

Selanjutnya akan dibuktikan bahwa $\widehat{\text{var}}[\hat{\mu}_{HH}]$ adalah taksiran yang tak bias untuk $\text{var}[\hat{\mu}_{HH}]$

$$
\begin{aligned}
E[\widehat{\text{var}}[\hat{\mu}_{HH}]] &= E\left[\frac{s^2}{n}\right] \\
&= \frac{1}{n}.E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(t_i - \bar{t}\right)^2\right] \\
&= \frac{1}{n(n-1)}.E\left[\sum_{i=1}^{n}\left(t_i - \bar{t}\right)^2\right] \\
&= \frac{1}{n(n-1)}.E\left[\sum_{i=1}^{n}\left(\left(t_i - \mu\right) - \left(\bar{t} - \mu\right)\right)^2\right] \\
&= \frac{1}{n(n-1)}.E\left[\sum_{i=1}^{n}\left(t_i - \mu\right) - n\left(\bar{t} - \mu\right)^2\right] \\
&= \frac{1}{n(n-1)}.\left\{E\left[\sum_{i=1}^{n}\left(t_i - \mu\right)^2\right] - nE\left[\left(\bar{t} - \mu\right)^2\right]\right\} \\
&= \frac{1}{n(n-1)}.\left\{\sum_{i=1}^{n}E\left[\left(t_i - \mu\right)^2\right] - nE\left[\left(\bar{t} - \mu\right)^2\right]\right\} \\
&= \frac{1}{n(n-1)}.\left\{\sum_{i=1}^{n}\text{var}[t_i] - n\,\text{var}[\bar{t}]\right\} \\
&= \frac{1}{n(n-1)}.\left\{n\sigma_T^2 - n.\text{var}[\frac{1}{n}\sum_{i=1}^{n}t_i]\right\} \\
&= \frac{1}{n(n-1)}.\left\{n\sigma_T^2 - \frac{1}{n}.\text{var}[\sum_{i=1}^{n}t_i]\right\} \\
&= \frac{1}{n(n-1)}.\left\{n\sigma_T^2 - \frac{1}{n}.\sum_{i=1}^{n}\text{var}[t_i]\right\}
\end{aligned}
$$

$$= \frac{1}{n(n-1)} . \left\{ n\sigma_T^2 - \frac{1}{n} . n\sigma_T^2 \right\}$$
$$= \frac{1}{n(n-1)} . \left\{ (n-1) . \sigma_T^2 \right\}$$
$$= \frac{\sigma_T^2}{n}$$
$$= \text{var}[\hat{\mu}_{HH}] \qquad (2.3.4)$$

Jadi, $\widehat{\text{var}}[\hat{\mu}_{HH}]$ adalah taksiran yang tak bias untuk $\text{var}[\hat{\mu}_{HH}]$.

Taksiran Hansen-Hurwitz untuk taksiran total adalah

$$\hat{\tau}_{HH} = N.\hat{\mu}_{HH} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{p_i} = \frac{1}{n}\sum_{i=1}^{N}\frac{u_i n_i}{p_i}$$

Berdasarkan (2.3.3), dapat diperoleh

$$E[\hat{\tau}_{HH}] = E[N.\hat{\mu}_{HH}] = N.E[\hat{\mu}_{HH}] = N.\mu = \tau$$

sehingga taksiran Hansen-Hurwitz untuk total adalah taksiaran yang tak bias.

Selanjutnya, variansi dari taksiran total tersebut adalah

$$\text{var}[\hat{\tau}_{HH}] = \text{var}[N\hat{\mu}_{HH}] = N^2\,\text{var}[\hat{\mu}_{HH}] = \frac{N^2}{n}\sum_{i=1}^{N}\left(\frac{u_i}{Np_i} - \mu\right)^2 .p_i$$

dan taksiran variansi

$$\widehat{\text{var}}[\hat{\tau}_{HH}] = \widehat{\text{var}}[N\hat{\mu}_{HH}] = N^2\widehat{\text{var}}[\hat{\mu}_{HH}] = \frac{N^2}{n(n-1)}\sum_{i=1}^{n}\left(\frac{y_i}{Np_i} - \hat{\mu}_{HH}\right)^2 .p_i$$

Kemudian, berdasarkan (2.3.4), dapat diperoleh

$$E[\widehat{\text{var}}[\hat{\tau}_{HH}]] = E[\widehat{\text{var}}[N.\hat{\mu}_{HH}]] = N^2E[\widehat{\text{var}}[\hat{\mu}_{HH}]] = N^2\,\text{var}[\hat{\mu}_{HH}] = \text{var}[N.\hat{\mu}_{HH}] = \text{var}[\hat{\tau}_{HH}]$$

sehingga $\widehat{\text{var}}[\hat{\tau}_{HH}]$ merupakan taksiran tak bias untuk $\text{var}[\hat{\tau}_{HH}]$.

2.4 *Adaptive Cluster Sampling*

Metode *adaptive cluster sampling* digunakan untuk mengambil sampel dimana elemen yang akan diteliti sangat jarang dan atau berkelompok. Metode ini ada dua jenis, tetapi pada tugas akhir ini hanya dibahas metode *adaptive cluster sampling* dengan sampel awal dipilih secara SRS tanpa pengembalian. Metode ini banyak digunakan dalam penelitian mengenai spesies langka, penyakit langka, atau mendeteksi wilayah dengan kandungan barang tambang terbanyak.

2.4.1 Keadaan Populasi

Populasi terdiri dari unit-unit sampling yang saling lepas. Pada metode *adaptive cluster sampling*, setiap unit-unit sampling (misalkan unit *i*), mempunyai tetangga yaitu unit-unit yang berada di sebelah kiri, kanan, depan, dan belakang unit *i*. Selain itu, peneliti akan menetapkan suatu syarat atau kondisi C untuk unit sampling. Dari unit-unit sampling yang ada pada populasi, diantaranya ada yang memenuhi kondisi C, ada pula yang tidak memenuhi kondisi C.

Misalkan unit *i* adalah suatu unit sampling pada populasi yang memiliki tetangga unit *p*, *q*, *r*, dan *s*. Jika unit *i* tidak memenuhi kondisi C, maka unit *i* disebut sebagai *network* berukuran satu. Jika unit *i* memenuhi kondisi C, tetapi tidak ada satu pun unit tetangganya yang memenuhi kodisi C, maka unit *i* juga disebut *network* berukuran satu. Jika unit *i* memenuhi kondisi C dan ada unit tetangganya yang memenuhi kondisi C (misalkan unit *p* dan *r*), maka unit *p* dan *r* digabung dengan unit *i* membentuk sebuah *network* berukuran tiga. *Network* ini dilambangkan dengan A_i, yaitu *network* yang dibentuk dari unit *i.* Sementara unit *q* dan *s* yang tidak memenuhi kondisi C tidak digabung dengan unit *i* dan masing-masing disebut sebagai *network* berukuran satu. Lalu, jika tetangga dari unit *p* dan *r* ada yang memenuhi kondisi C, maka unit-unit tersebut juga digabung dengan *network* A_i, sehingga ukuran *network* A_i bertambah. Proses ini terus terjadi hingga tidak ada unit-unit yang memenuhi kondisi C.

Berdasarkan penjelasan tersebut, dapat dikatakan bahwa populasi terdiri dari *network-network* berukuran tertentu yang saling lepas.

2.4.2 Cara Pengambilan Sampel

Pengambilan sampel pada *adaptive cluster sampling* diawali dengan membagi wilayah penelitian menjadi *N* unit-unit sampling. Dari *N* unit-unit sampling pada populasi, dipilih suatu sampel awal berukuran n_1 unit secara SRS. Misalkan unit *i* adalah salah satu anggota dari sampel awal. Unit *i* tersebut diteliti apakah memenuhi kondisi C atau tidak. Jika unit *i* memenuhi kondisi C, maka unit-unit sampling yang bertetanggaan dengan unit *i* diperiksa. Jika unit sampling yang merupakan tetangga dari unit *i* juga memenuhi kondisi C, maka unit tersebut dijadikan sampel dan unit sampling yang bertetanggaan dengannya juga diperiksa. Jika unit yang bertetanggaan dengan tetangga unit *i* juga memenuhi kondisi C, maka unit tersebut juga dijadikan sampel dan seterusnya. Proses ini dinamakan proses penambahan sampel dan berhenti ketika unit-unit sampling yang bertetanggaan dengan sampel tidak memenuhi kondisi C. Dari proses penambahan sampel, unit *i* membentuk suatu kumpulan unit-unit sampling yang merupakan suatu *network* berukuran tertentu yang dinotasikan dengan A_i. *Network* A_i yang dibentuk oleh unit *i*, juga merupakan *network* dalam populasi. Oleh karena itu, jika unit *i* terpilih dalam sampel awal maka *network* A_i dapat dianggap sebagai *network* yang terpilih dalam sampel. Hal yang sama dengan unit *i* juga dilakukan terhadap unit sampling lain yang terpilih pada sampel awal.

Jika unit sampling yang terpilih pada sampel awal tidak memenuhi kondisi C, maka proses penambahan sampel tidak dilakukan dan unit tersebut merupakan *network* berukuran satu. Kemudian, jika suatu sampel awal memenuhi kondisi C, tetapi unit-unit sampling tetangganya tidak memenuhi kondisi C, maka sampel awal tersebut merupakan *network* berukuran satu pula.

2.4.3 Penaksiran Mean Dan Total Populasi

Metode *adaptive cluster sampling* merupakan *unequal probability sampling*. Hal ini disebabkan karena semakin besar ukuran suatu *network*, maka probabilitas *network* tersebut terpilih sebagai sampel semakin besar. Taksiran untuk mean dan total yang dapat digunakan dalam kasus *unequal probability sampling* seperti ini adalah taksiran Horvitz-Thompson dan taksiran Hansen-Hurwitz.

2.4.3.1 Penaksiran Mean Dan Total Populasi Dengan Taksiran Horvitz-Thompson

Misalkan *N* adalah banyaknya unit sampling pada populasi, n_1 adalah banyaknya sampel awal, *K* banyaknya *network* pada populasi, κ adalah banyaknya *network* pada sampel, τ_k^* adalah total pengukuran pada *network* ke k dalam populasi, y_k^* adalah total pengukuran pada *network* ke k dalam sampel, Υ_i adalah total pengukuran pada unit ke *i* dalam populasi, y_i adalah total pengukuran pada unit ke *i* dalam sampel, dan π_k adalah probabilitas *network* ke k terpilih sebagai sampel. Taksiran Horvitz-Thompson untuk mean tersebut didefinisikan sebagai berikut :

$$\hat{\mu} = \frac{1}{N}\sum_{k=1}^{\kappa}\frac{y_k^*}{\pi_k} \tag{2.4.1}$$

Bentuk π_k dapat dicari sebagai berikut :

Misalkan x_k adalah banyaknya unit sampling pada *network* ke *k* dan p_k adalah probabilitas *network* ke k tidak terpilih sebagai sampel. Pada *network* ke k ada sebanyak x_k unit sampling, sedangkan $N - x_k$ unit tidak berada pada *network* k. Sehingga banyaknya cara memilih n_1 unit dari $N - x_k$ unit adalah $\binom{N - x_k}{n_1}$, sedangkan banyaknya semua kemungkinan sampel adalah $\binom{N}{n_1}$, maka probabilitas *network* ke k tidak terpilih sebagai sampel adalah $p_k = \binom{N - x_k}{n_1} \Big/ \binom{N}{n_1}$. Jadi, probabilitas *network* ke k terpilih sebagai sampel adalah

$$\pi_k = 1 - p_k = 1 - \left[\binom{N - x_k}{n_1} \Big/ \binom{N}{n_1} \right] \tag{2.4.2}$$

Selanjutnya akan ditunjukkan bahwa $\hat{\mu}$ pada (2.4.1) adalah taksiran tak bias untuk mean. Sebelumnya, didefinisikan z_k adalah variabel indikator, yaitu

$$z_k = \begin{cases} 1, \text{ jika } \textit{network} \text{ ke k terpilih sebagai sampel} \\ 0, \text{ jika } \textit{network} \text{ ke k tidak terpilih sebagai sampel} \end{cases}$$

dimana $\Pr(z_k = 1) = \pi_k$ dan

$$\begin{aligned} E[z_k] &= 0.\Pr(z_k = 0) + 1.\Pr(z_k = 1) \\ &= 1.\Pr(\text{network k terpilih menjadi sampel}) \\ &= \pi_k \end{aligned}$$

maka

$$
\begin{aligned}
E[\hat{\mu}] &= E\left[\frac{1}{N}\sum_{k=1}^{K}\frac{y_k^*}{\pi_k}\right] \\
&= E\left[\frac{1}{N}\sum_{k=1}^{K}\frac{\tau_k^* z_k}{\pi_k}\right] \\
&= \frac{1}{N}\sum_{k=1}^{K}\frac{\tau_k^* E[z_k]}{\pi_k} \\
&= \frac{1}{N}\sum_{k=1}^{K}\frac{\tau_k^* \pi_k}{\pi_k} \\
&= \frac{1}{N}\sum_{k=1}^{K}\tau_k^* \\
&= \frac{1}{N}\sum_{i=1}^{N}\Upsilon_i \\
&= \mu \qquad (2.4.3)
\end{aligned}
$$

sehingga $\hat{\mu}$ merupakan taksiran yang tak bias untuk mean.

Untuk mencari variansi dari taksiran mean pada (2.4.1), terlebih dahulu akan dicari nilai dari $\text{var}[z_k]$ dan $\text{cov}[z_k, z_{k'}]$, yaitu :

$$
\begin{aligned}
\text{var}[z_k] &= E[z_k{}^2] - [E[z_k]]^2 \\
&= \left[0^2.\Pr(z_k = 0) + 1^2.\Pr(z_k = 1)\right] - [E[z_k]]^2 \\
&= \pi_k - \pi_k{}^2 \\
&= \pi_k(1-\pi_k) \qquad (2.4.4)
\end{aligned}
$$

$$
\begin{aligned}
E[z_k z_{k'}] &= \sum_{z_k=0}^{1}\sum_{z_{k'}=0}^{1} z_k z_{k'}.\Pr(z_k, z_{k'}) \\
&= 1.1.\Pr(z_k = 1, z_{k'} = 1) \\
&= \Pr(\text{network ke k dan network ke k' terpilih sebagai sampel}) \\
&= \pi_{kk'}
\end{aligned}
$$

$$
\begin{aligned}
\text{cov}[z_k, z_{k'}] &= E[z_k z_{k'}] - E[z_k].E[z_{k'}] \\
&= \pi_{kk'} - \pi_k \pi_{k'} \qquad (2.4.5)
\end{aligned}
$$

Sehingga variansi dari (2.4.1) dapat ditentukan sebagai berikut

$$
\begin{aligned}
\text{var}[\hat{\mu}] &= \text{var}\left[\frac{1}{N}\sum_{k=1}^{\kappa}\frac{y_k^*}{\pi_k}\right] \\
&= \text{var}\left[\frac{1}{N}\sum_{k=1}^{K}\frac{\tau_k^* z_k}{\pi_k}\right] \\
&= \frac{1}{N^2}\text{var}\left[\sum_{k=1}^{K}\frac{\tau_k^* z_k}{\pi_k}\right] \\
&= \frac{1}{N^2}\left\{\sum_{k=1}^{K}\left(\frac{\tau_k^*}{\pi_k}\right)^2 \text{var}[z_k] + \sum_{k=1}^{K}\sum_{k'\neq k}\frac{\tau_k^*\tau_{k'}^*}{\pi_k\pi_{k'}}\text{cov}[z_k, z_{k'}]\right\} \\
&= \frac{1}{N^2}\left\{\sum_{k=1}^{K}\left(\frac{\tau_k^*}{\pi_k}\right)^2 .\pi_k(1-\pi_k) + \sum_{k=1}^{K}\sum_{k'\neq k}\frac{\tau_k^*\tau_{k'}^*}{\pi_k\pi_{k'}}.\left(\pi_{kk'} - \pi_k\pi_{k'}\right)\right\} \\
&= \frac{1}{N^2}\left\{\sum_{k=1}^{K}\tau_k^{*2}\frac{(1-\pi_k)}{\pi_k} + \sum_{k=1}^{K}\sum_{k'\neq k}\tau_k^*\tau_{k'}^*\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}}\right)\right\} \\
&= \frac{1}{N^2}\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}}\right)\tau_k^*\tau_{k'}^* \qquad (2.4.6)
\end{aligned}
$$

dengan taksirannya adalah

$$
\widehat{\text{var}}[\hat{\mu}] = \frac{1}{N^2}\sum_{k=1}^{\kappa}\sum_{k'=1}^{\kappa}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}\pi_{kk'}}\right)y_k^* y_{k'}^* \qquad (2.4.7)
$$

Bentuk $\pi_{kk'}$ dapat dicari sebagai berikut :

Misalkan :

$p_{kk'} = P$(network k dan network k' tidak terpilih sebagai sampel), maka

$$
\begin{aligned}
p_{kk'} &= P(z_k \neq 1 \cap z_{k'} \neq 1) \\
&= \binom{N - x_k - x_{k'}}{n_1} \Big/ \binom{N}{n_1}
\end{aligned}
$$

Kemudian, komplemen dari $p_{kk'}$ adalah

$$
\begin{aligned}
p_{kk'}^c &= P\left[(z_k \neq 1 \cap z_{k'} \neq 1)^c\right] \\
&= P(z_k = 1 \cup z_{k'} = 1)
\end{aligned}
$$

dengan demikian, $p^c_{kk'}$ adalah probabilitas *network* k atau *network* k' terpilih sebagai sampel, maka

$$\begin{aligned} p^c_{kk'} &= 1 - p_{kk'} \\ &= 1 - P(z_k \neq 1 \cap z_{k'} \neq 1) \\ &= 1 - \binom{N - x_k - x_{k'}}{n_1} \Big/ \binom{N}{n_1} \end{aligned}$$

Jadi, probabilitas *network* k dan *network* k' terpilih sebagai sampel adalah

$$\begin{aligned} \pi_{kk'} &= \pi_k + \pi_{k'} - p^c_{kk'} \\ &= \left\{ 1 - \left[\binom{N - x_k}{n_1} \Big/ \binom{N}{n_1} \right] \right\} + \left\{ 1 - \left[\binom{N - x_{k'}}{n_1} \Big/ \binom{N}{n_1} \right] \right\} - \left\{ 1 - \binom{N - x_k - x_{k'}}{n_1} \Big/ \binom{N}{n_1} \right\} \\ &= 1 - \left[\binom{N - x_k}{n_1} + \binom{N - x_{k'}}{n_1} - \binom{N - x_k - x_{k'}}{n_1} \right] \Big/ \binom{N}{n_1} \end{aligned} \qquad (2.4.8)$$

Selanjutnya akan ditunjukkan bahwa $\widehat{\text{var}}[\hat{\mu}]$ adalah taksiran tak bias untuk $\text{var}[\hat{\mu}]$:

$$\begin{aligned} E[\widehat{\text{var}}[\hat{\mu}]] &= E\left[\frac{1}{N^2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'} \pi_{kk'}} \right) y^*_k y^*_{k'} \right] \\ &= E\left[\frac{1}{N^2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'} \pi_{kk'}} \right) \tau^*_k \tau^*_{k'} \, z_k z_{k'} \right] \\ &= \frac{1}{N^2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'} \pi_{kk'}} \right) \tau^*_k \tau^*_{k'} . E[z_k z_{k'}] \\ &= \frac{1}{N^2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'} \pi_{kk'}} \right) \tau^*_k \tau^*_{k'} . \pi_{kk'} \\ &= \frac{1}{N^2} \sum_{k=1}^{K} \sum_{k'=1}^{K} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'}} \right) \tau^*_k \tau^*_{k'} \\ &= \text{var}[\hat{\mu}] \end{aligned} \qquad (2.4.9)$$

maka taksiran variansi tersebut merupakan taksiran yang tak bias.

Selanjutnya, taksiran Horvitz-Thompson untuk total adalah

$$\hat{\tau} = N.\hat{\mu} = \sum_{k=1}^{\kappa} \frac{y_k^*}{\pi_k} .$$

Akan ditunjukkan bahwa $\hat{\tau}$ adalah taksiran tak bias untuk total. Berdasarkan (2.4.3) diperoleh

$$E[\hat{\tau}] = E[N.\hat{\mu}] = N.E[\hat{\mu}] = N\mu = \tau . \tag{2.4.10}$$

sehingga $\hat{\tau}$ adalah taksiran tak bias untuk total.

Variansi dari taksiran total adalah

$$\text{var}[\hat{\tau}] = \text{var}[N\hat{\mu}] = N^2 \text{var}[\hat{\mu}] = \sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'}}\right)\tau_k^* \tau_{k'}^*$$

dan taksiran variansi

$$\widehat{\text{var}}[\hat{\tau}] = \widehat{\text{var}}[N\hat{\mu}] = N^2 \widehat{\text{var}}[\hat{\mu}] = \sum_{k=1}^{\kappa}\sum_{k'=1}^{\kappa}\left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'} \pi_{kk'}}\right) y_k^* y_{k'}^* .$$

Akan ditunjukkan bahwa taksiran variansi tersebut adalah taksiran yang tak bias. Berdasarkan (2.4.9), diperoleh

$$\begin{aligned} E[\widehat{\text{var}}[\hat{\tau}]] &= E[\widehat{\text{var}}[N\hat{\mu}]] \\ &= E[N^2 \widehat{\text{var}}[\hat{\mu}]] \\ &= N^2 E[\widehat{\text{var}}[\hat{\mu}]] \\ &= N^2 \text{var}[\hat{\mu}] \\ &= \text{var}[N\hat{\mu}] \\ &= \text{var}[\hat{\tau}] \end{aligned} \tag{2.4.11}$$

sehingga terbukti bahwa $\widehat{\text{var}}[\hat{\tau}]$ adalah taksiran tak bias untuk $\text{var}[\hat{\tau}]$.

2.4.3.2 Penaksiran Mean Dan Total Populasi Dengan Taksiran Hansen-Hurwitz

Misalkan *N* adalah banyaknya unit sampling pada populasi, y_i adalah total pengukuran pada unit ke *i,* dan m_i adalah banyaknya unit sampling yang ada di *network* A_i. Taksiran Hansen-Hurwitz untuk mean didefinisikan sebagai berikut :

$$\tilde{\mu} = \frac{1}{N}\sum_{i=1}^{n}\frac{y_i}{E[f_i]}$$

atau dapat ditulis sebagai :

$$\tilde{\mu} = \frac{1}{N}\sum_{i=1}^{N}\frac{\Upsilon_i f_i}{E[f_i]} \tag{2.4.12}$$

Kemudian, misalkan f_i adalah banyaknya sampel awal yang ada di *network* A_i (*network* yang dibentuk oleh unit ke *i*). Dengan demikian, f_i berdistribusi hypergeometrik (Hyper[*N*, m_i, n_1]) dengan $E[f_i] = n_1 m_i / N$, dan persamaan (2.4.12) dapat ditulis sebagai berikut :

$$\tilde{\mu} = \frac{1}{n_1}\sum_{i=1}^{N}\frac{\Upsilon_i f_i}{m_i} \tag{2.4.13}$$

Taksiran pada persamaan (2.4.13) dinyatakan sebagai penjumlahan dari *N* unit sampling. Untuk mempermudah dalam perhitungan, persamaan

(2.4.13) dapat dinyatakan sebagai penjumlahan dari n_1 *network* yang terpilih sebagai sampel, yaitu ;

$$\begin{aligned}\tilde{\mu} &= \frac{1}{n_1}\sum_{i=1}^{n_1}\frac{1}{m_i}\sum_{j\in A_i} y_j \\ &= \frac{1}{n_1}\sum_{i=1}^{n_1} w_i \\ &= \bar{w} \qquad (2.4.14)\end{aligned}$$

dimana w_i adalah rata-rata pengukuran di A_i yang terdiri dari m_i unit. Akan dibuktikan bahwa $\tilde{\mu}$ adalah taksiran tak bias untuk μ, yaitu :

$$\begin{aligned}E[\tilde{\mu}] &= E\left[\frac{1}{n_1}\sum_{i=1}^{n_1} w_i\right] \\ &= E\left[\frac{1}{n_1}\sum_{i=1}^{n_1}\frac{1}{m_i}\sum_{j\in A_i} y_j\right] \\ &= E\left[\frac{1}{n_1}\sum_{i=1}^{N}\frac{\Upsilon_i f_i}{m_i}\right] \\ &= \frac{1}{n_1}\sum_{i=1}^{N}\frac{\Upsilon_i}{m_i}.E[f_i] \\ &= \frac{1}{n_1}\sum_{i=1}^{N}\frac{\Upsilon_i}{m_i}.\frac{n_1 m_i}{N} \\ &= \frac{1}{N}\sum_{i=1}^{N}\Upsilon_i \\ &= \mu \qquad (2.4.15)\end{aligned}$$

karena $E[\tilde{\mu}] = \mu$, maka $\tilde{\mu}$ adalah taksiran tak bias untuk μ.

Dalam pemilihan sampel awal berukuran n_1 digunakan SRS tanpa pengembalian. Berdasarkan konsep SRS dan persamaan (2.1.5) dan (2.1.6) variansi dan taksiran variansi dari $\tilde{\mu}$ dapat diperoleh sebagai berikut :

(2.4.15)

$$\text{var}[\tilde{\mu}] = \frac{N - n_1}{Nn_1(N-1)} \sum_{i=1}^{N} \left(w_i - \mu\right)^2 \qquad (2.4.16)$$

$$\widehat{\text{var}}[\tilde{\mu}] = \frac{N - n_1}{Nn_1(n_1-1)} \sum_{i=1}^{n_1} \left(w_i - \tilde{\mu}\right)^2 \qquad (2.4.17)$$

Dengan memisalkan $\sigma^2 = \frac{1}{(N-1)} \sum_{i=1}^{N} \left(w_i - \mu\right)^2$ dan $s^2 = \frac{1}{(n_1-1)} \sum_{i=1}^{n_1} \left(w_i - \tilde{\mu}\right)^2$,

maka cara yang sama seperti (2.1.7) dan (2.1.8) dapat digunakan untuk menunjukkan bahwa

$$E[\widehat{\text{var}}[\tilde{\mu}]] = \text{var}[\tilde{\mu}] \qquad (2.4.18)$$

sehingga, $\widehat{\text{var}}[\tilde{\mu}]$ adalah taksiran tak bias untuk $\text{var}[\tilde{\mu}]$.

Selanjutnya, taksiran Hansen-Hurwitz untuk taksiran total adalah

$$\tilde{\tau} = N.\tilde{\mu} = \frac{N}{n_1} \sum_{i=1}^{n_1} \frac{1}{m_i} \sum_{j \in A_i} y_j = \frac{N}{n_1} \sum_{i=1}^{n_1} w_i = N.\bar{w}$$

Akan ditunjukkan bahwa $\tilde{\tau}$ adalah taksiran tak bias untuk total. Berdasarkan (2.4.15), diperoleh

$$E[\tilde{\tau}] = E[N.\tilde{\mu}] = N.E[\tilde{\mu}] = N\mu = \tau$$

Sehingga $\hat{\tau}$ adalah taksiran tak bias untuk total.

Berdasarkan (2.4.16) dan (2.4.17), variansi dan taksiran variansi dari taksiran total dapat diperoleh sebagai berikut

$$\text{var}[\tilde{\tau}] = \text{var}[N\tilde{\mu}] = N^2\, \text{var}[\tilde{\mu}] = \frac{N\left(N - n_1\right)}{n_1(N-1)} \sum_{i=1}^{N} \left(w_i - \mu\right)^2$$

$$\widehat{\text{var}}[\tilde{\tau}] = \widehat{\text{var}}[N\tilde{\mu}] = N^2\, \widehat{\text{var}}[\tilde{\mu}] = \frac{N\left(N - n_1\right)}{n_1(n_1-1)} \sum_{i=1}^{n_1} \left(w_i - \tilde{\mu}\right)^2 .$$

Akan ditunjukkan bahwa $\widehat{\text{var}}[\tilde{\tau}]$ adalah taksiran tak bias untuk $\text{var}[\tilde{\tau}]$.

Berdasarkan (2.4.18), maka

$$\begin{aligned}
E[\widehat{\text{var}}[\tilde{\tau}]] &= E[\widehat{\text{var}}[N\tilde{\mu}]] \\
&= E[N^2\widehat{\text{var}}[\tilde{\mu}]] \\
&= N^2 E[\widehat{\text{var}}[\tilde{\mu}]] \\
&= N^2 \text{var}[\tilde{\mu}] \\
&= \text{var}[N\tilde{\mu}] \\
&= \text{var}[\tilde{\tau}]
\end{aligned}$$

sehingga $\widehat{\text{var}}[\tilde{\tau}]$ adalah taksiran tak bias untuk $\text{var}[\tilde{\tau}]$.

BAB III

TWO STAGE ADAPTIVE CLUSTER SAMPLING

(2S-ACS)

Pada bagian sebelumnya telah dibahas mengenai metode *adaptive cluster sampling.* Metode ini sangat berguna digunakan apabila elemen yang akan diteliti bersifat sangat jarang dan atau berkelompok. Akan tetapi, metode *adaptive cluster sampling* memiliki beberapa kelemahan, yaitu ada kemungkinan bahwa elemen pada sampel akhir yang terbentuk akan sangat banyak. Kelemahan berikutnya adalah memerlukan banyak usaha/tenaga dalam meneliti unit-unit sampling yang menjadi sampel awal karena ada kemungkinan unit-unit sampling tersebut terletak berjauhan. Hal ini menjadi suatu kesulitan jika daerah yang menjadi objek penelitian sangat luas. Oleh karena itu, pada bagian ini akan dijelaskan mengenai suatu metode yang dapat digunakan untuk mengatasi kelemahan-kelemahan tersebut, yaitu metode *two stage adaptive cluster sampling* (2S-ACS).

Berbeda dengan *adaptive cluster sampling*, pada *two stage adaptive cluster sampling,* daerah yang menjadi objek penelitian dibagi dahulu menjadi unit-unit primer yang diperkirakan homogen. Kemudian, masing-masing unit primer dibagi menjadi unit-unit sampling. Selanjutnya, ambil beberapa unit primer secara SRS dan dari unit primer yang terpilih, pilih unit sampling sebagai sampel awal (*initial sample*) secara SRS. Masing-masing sampel awal diteliti apakah terdapat elemen dan memenuhi kondisi C. Jika dalam sampel awal terdapat elemen yang akan diteliti dan memenuhi kondisi C, maka unit-unit sampling yang bertetanggaan dengan sampel awal diperiksa. Jika pada unit sampling yang bertetanggaan dengan sampel awal terdapat elemen dan memenuhi kondisi C, maka unit-unit tersebut dijadikan sampel. Lalu, unit yang bertetanggan dengan unit-unit yang memenuhi kondisi C juga diperiksa, jika memenuhi kondisi C maka unit tersebut dijadikan sampel pula. Proses ini dinamakan proses penambahan sampel.

Dalam penambahan sampel, pada 2S-ACS ada dua skema yang dapat digunakan, yaitu skema *overlapping* dan skema *nonoverlapping*. Pada skema *overlapping*, proses penambahan sampel diperbolehkan melewati batas unit primer, dan proses penambahan sampel berhenti ketika tidak ada lagi unit-unit sampling yang bertetanggaan dengan sampel yang memenuhi kondisi C. Pada skema *nonoverlapping*, proses penambahan sampel tidak diperbolehkan melewati batas unit primer. Dengan demikian, pada skema *nonoverlapping*, proses penambahan sampel berhenti ketika tidak ada lagi unit-unit sampling yang bertetanggaan dengan sampel yang memenuhi kondisi C atau ketika penambahan sampel telah mencapai batas unit primer.

Sama halnya dengan *adaptive cluster sampling*, sampel yang diperoleh pada 2S-ACS merupakan *network-network* dengan ukuran tertentu. Jika suatu unit sampling yang terpilih pada sampel awal tidak memenuhi kondisi C, maka unit tersebut merupakan *network* berukuran satu. Kemudian, jika suatu unit sampling pada sampel awal memenuhi kondisi C, tetapi unit-unit sampling tetangganya tidak memenuhi kondisi C, maka sampel awal tersebut merupakan *network* berukuran satu pula.

Misalkan *M* adalah banyaknya unit primer pada populasi dan N_i adalah banyaknya unit sampling pada unit primer ke *i*. Misalkan τ_{ij} menyatakan total pengukuran pada unit sampling ke *j* dalam unit primer ke *i*, maka total pengukuran pada unit primer ke *i* adalah $\tau_i = \sum_{j=1}^{N_i} \tau_{ij}$, sedangkan total pengukuran pada populasi adalah $\tau = \sum_{i=1}^{M} \sum_{j=1}^{N_i} \tau_{ij} = \sum_{i=1}^{M} \tau_i$. Selanjutnya, mean / rata-rata pengukuran pada unit primer ke *i* adalah $\mu_i = \tau_i / N_i$, sedangkan mean pengukuran pada populasi adalah $\mu = \tau / N$, dimana $N = \sum_{i=1}^{M} N_i$ yaitu banyaknya unit sampling pada populasi. Misalkan pula *m* adalah banyaknya unit primer yang terpilih sebagai sampel pada tahap pertama, n_i adalah banyaknya unit sampling pada unit primer ke *i* (*i = 1, 2, …, m*) yang terpilih sebagai sampel awal pada tahap kedua, dan n_0 adalah total sampel awal dimana $n_0 = \sum_{i=1}^{m} n_i$.

Dalam hal penaksiran mean dan total, metode *two stage adaptive cluster sampling* merupakan *unequal probability sampling*. Jadi, taksiran Horvitz-Thompson dan taksiran Hansen-Hurwitz akan digunakan untuk menaksir mean dan total populasi pada masing-masing skema (skema *overlapping* dan skema *nonoverlapping*).

3.1 Skema *Overlapping*

Pada skema *overlapping,* proses penambahan sampel diperbolehkan melewati batas unit primer. Jadi, pada skema ini penambahan sampel berhenti ketika tidak ada lagi unit sampling yang bertetanggaan dengan sampel yang memenuhi kondisi C. Dengan perkataan lain, penambahan sampel pada 2S-ACS dengan skema *overlapping* sama dengan penambahan sampel pada *adaptive cluster sampling,* walaupun cara pengambilan sampel kedua metode adalah berbeda. Oleh karena itu, bentuk taksiran Horvitz-Thompson untuk mean dan total pada skema *overlapping* sama dengan bentuk taksiran pada *adaptive cluster sampling.*

3.1.1 Penaksiran Mean Dan Total Populasi Dengan Taksiran Horvitz-Thompson

Misalkan *N* adalah banyaknya unit sampling pada populasi, *K* adalah banyaknya *network* pada populasi, κ adalah banyaknya *network* pada sampel, τ_k^* adalah total pengukuran pada *network* ke *k* dalam populasi, y_k^* adalah total pengukuran pada *network* ke *k* dalam sampel, Υ_i adalah total pengukuran pada unit ke *i* dalam populasi, y_i adalah total pengukuran pada unit ke *i* dalam sampel, dan π_k adalah probabilitas *network* ke *k* terpilih sebagai sampel, maka taksiran Horvitz-Thompson untuk mean adalah

$$\hat{\mu} = \frac{1}{N}\sum_{k=1}^{\kappa}\frac{y_k^*}{\pi_k} \quad (3.1.1)$$

Sama halnya dengan (2.4.3), maka taksiran mean pada (3.1.1) adalah taksiran tak bias untuk mean. Lalu, variansi dan taksiran variansi dari taksiran mean sama seperti (2.4.6) dan (2.4.7), yaitu

$$\text{var}[\hat{\mu}] = \frac{1}{N^2}\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}}\right)\tau_k^*\tau_{k'}^* \quad (3.1.2)$$

$$\widehat{\text{var}}[\hat{\mu}] = \frac{1}{N^2}\sum_{k=1}^{\kappa}\sum_{k'=1}^{\kappa}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}\pi_{kk'}}\right)y_k^*y_{k'}^* \quad (3.1.3)$$

dimana π_{kj} adalah probabilitas *network* ke *k* dan *network* ke *k'* terpilih sebagai sampel. Berdasarkan (2.4.9), maka taksiran pada (3.1.3) adalah taksiran tak bias untuk (3.1.2). Selanjutnya, taksiran Horviz-Thompson untuk total adalah

$$\hat{\tau} = N\hat{\mu} = \sum_{k=1}^{\kappa}\frac{y_k^*}{\pi_k}$$

sama halnya dengan (2.4.10), maka $\hat{\tau}$ adalah taksiran tak bias untuk total. Kemudian, variansi dan taksiran variansi dari taksiran total adalah

$$\text{var}[\hat{\tau}] = \sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}}\right)\tau_k^*\tau_{k'}^*$$

$$\widehat{\text{var}}[\hat{\tau}] = \sum_{k=1}^{\kappa}\sum_{k'=1}^{\kappa}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}\pi_{kk'}}\right)y_k^*y_{k'}^*$$

dan langkah seperti (2.4.11) dapat digunakan untuk menunjukkan bahwa $\widehat{\text{var}}[\hat{\tau}]$ adalah taksiran tak bias untuk $\text{var}[\hat{\tau}]$.

Meskipun bentuk taksiran Horvitz-Thompson pada *two stage adaptive cluster sampling* dengan skema *overlapping* dan *adaptive cluster sampling* adalah sama, namun bentuk π_k dan $\pi_{kk'}$ pada kedua metode berbeda. Bentuk π_k dan $\pi_{kk'}$ akan dijelaskan sebagai berikut :

Misalkan x_{ik} adalah banyaknya unit sampling pada *network* ke k di dalam unit primer *i*, B_k adalah himpunan unit primer yang beririsan dengan *network* ke k, g_k adalah banyaknya anggota himpunan B_k , C_{ik} adalah kejadian bahwa paling sedikit satu unit sampling dari *network* ke k yang ada pada unit primer *i*, terpilih sebagai sampel awal, C_k adalah kejadian bahwa paling sedikit satu unit sampling pada *network* ke k terpilih sebagai sampel awal, maka :

$$C_k = \bigcup_{i \in B_k} C_{ik}$$

sehingga

$$\begin{aligned}\pi_k &= \Pr(C_k) = \Pr(\bigcup_{i \in B_k} C_{ik}) \\ &= \sum_{i \in B_k} \Pr(C_{ik}) - \sum_{i} \sum_{i' < i} \Pr(C_{ik} \cap C_{i'k}) + \ldots + (-1)^{g_k + 1} \Pr(\bigcap_{i \in B_k} C_{ik})\end{aligned}$$

Akan dicari terlebih dahulu bentuk dari $\Pr(\bigcap_{i \in B_k} C_{ik})$. Misalkan s_1 adalah himpunan dari unit primer yang terpilih sebagai sampel pada tahap pertama. Jika, diberikan s_1 kejadian C_{ik} saling bebas, maka

$$\begin{aligned}\Pr(\bigcap_{i \in B_k} C_{ik}) &= \sum_{S_1} \Pr\left(\bigcap_{i \in B_k} C_{ik} \mid S_1\right) . \Pr(S_1) \\ &= \Pr(S_1) . \sum_{S_1 \ni B_k} \prod_{i \in B_k} \Pr\left(C_{ik} \mid S_1\right)\end{aligned}$$

Karena $\Pr(S_1) = 1\Big/\binom{M}{m}$ dan ada sebanyak $\binom{M-g_k}{m-g_k}$ cara bagi unit primer

$i \in B_k$ untuk berada dalam sampel, maka

$$\Pr(\bigcap_{i \in B_k} C_{ik}) = 1\Big/\binom{M}{m} . \sum_{S_1 \ni B_k} \prod_{i \in B_k} \Pr\left(C_{ik} \mid S_1\right)$$

$$= \frac{\binom{M-g_k}{m-g_k}}{\binom{M}{m}} \prod_{i \in B_k} \Pr\left(C_{ik} \mid S_1\right)$$

Sementara $\Pr\left(C_{ik} \mid S_1\right)$ dapat diartikan pula sebagai probabilitas *network* ke k pada unit primer *i* terpilih sebagai sampel, maka untuk mencari bentuk $\Pr\left(C_{ik} \mid S_1\right)$ dapat dilakukan cara yang sama seperti (2.4.2), sehingga

$$\Pr(\bigcap_{i \in B_k} C_{ik}) = \frac{\binom{M-g_k}{m-g_k}}{\binom{M}{m}} \prod_{i \in B_k} \left[1 - \frac{\binom{N_i - x_{ik}}{n_i}}{\binom{N_i}{n_i}} \right]$$

Dengan cara yang sama seperti diatas, $\Pr(C_{ik} \cap C_{i'k})$ dapat ditentukan, yaitu :

Ada sebanyak $\binom{M-2}{m-2}$ cara bagi unit primer *i* dan *i'* untuk berada dalam sampel, sehingga

$$\begin{aligned}\Pr(C_{ik} \cap C_{i'k}) &= \sum_{S_1} \Pr\left(C_{ik} \cap C_{i'k} \mid S_1\right).\Pr(S_1) \\ &= \Pr(S_1).\sum_{S_1 \ni i,i'} \Pr\left(C_{ik} \cap C_{i'k} \mid S_1\right) \\ &= 1\Big/\binom{M}{m}.\sum_{S_1 \ni i,i'} \Pr\left(C_{ik} \mid S_1\right).\Pr\left(C_{i'k} \mid S_1\right) \\ &= \binom{M-2}{m-2}\Big/\binom{M}{m}.\Pr\left(C_{ik} \mid S_1\right).\Pr\left(C_{i'k} \mid S_1\right) \\ &= \frac{m(m-1)}{M(M-1)}.\left[1-\binom{N_i - x_{ik}}{n_i}\Big/\binom{N_i}{n_i}\right]\left[1-\binom{N_{i'} - x_{i'k}}{n_{i'}}\Big/\binom{N_{i'}}{n_{i'}}\right]\end{aligned}$$

Begitupun untuk mencari bentuk $\Pr(C_{ik})$. Karena ada sebanyak $\binom{M-1}{m-1}$ cara bagi unit primer *i* untuk berada dalam sampel, maka

$$\begin{aligned}\Pr(C_{ik}) &= \sum_{S_1} \Pr\left(C_{ik} \mid S_1\right).\Pr(S_1) \\ &= \Pr(S_1).\sum_{S_1 \ni i} \Pr\left(C_{ik} \mid S_1\right) \\ &= 1\Big/\binom{M}{m}.\sum_{S_1 \ni i} \Pr\left(C_{ik} \mid S_1\right) \\ &= \binom{M-1}{m-1}\Big/\binom{M}{m}.\Pr\left(C_{ik} \mid S_1\right) \\ &= \frac{m}{M}.\left[1-\binom{N_i - x_{ik}}{n_i}\Big/\binom{N_i}{n_i}\right]\end{aligned}$$

Dengan demikian,

$$\pi_k = \sum_{i \in B_k} \Pr(C_{ik}) - \sum_{i} \sum_{i'<i} \Pr(C_{ik} \cap C_{i'k}) + \ldots + (-1)^{g_k+1} \Pr(\bigcap_{i \in B_k} C_{ik})$$

$$= \sum_{i \in B_k} \frac{m}{M} \left[1 - \frac{\binom{N_i - x_{ik}}{n_i}}{\binom{N_i}{n_i}} \right] - \sum_{i} \sum_{i'<i} \frac{m(m-1)}{M(M-1)} \left[1 - \frac{\binom{N_i - x_{ik}}{n_i}}{\binom{N_i}{n_i}} \right] \left[1 - \frac{\binom{N_{i'} - x_{i'k}}{n_{i'}}}{\binom{N_{i'}}{n_{i'}}} \right] + \ldots$$

$$+ (-1)^{g_k+1} \frac{m(m-1)\ldots(m-g_k+1)}{M(M-1)\ldots(M-g_k+1)} \prod_{i \in B_k} \left[1 - \frac{\binom{N_i - x_{ik}}{n_i}}{\binom{N_i}{n_i}} \right] \qquad (3.1.4)$$

Selanjutnya akan dicari bentuk $\pi_{kk'}$ yaitu probabilitas *network* ke k dan ke k' terpilih sebagai sampel. Misalkan $C_{kk'}$ adalah kejadian bahwa paling tidak satu unit sampling di dalam *network* ke k dan ke k' yang ada di dalam unit primer *i*, terpilih sebagai sampel awal. Karena $C_k = \bigcup_{i \in B_k} C_{ik}$, maka

$$C_{kk'} = \left[\bigcup_{i \in B_k} C_{ik} \right] \cap \left[\bigcup_{i \in B_{k'}} C_{ik'} \right]$$

Sehingga

$$\pi_{kk'} = \Pr\left(\left[\bigcup_{i \in B_k} C_{ik} \right] \cap \left[\bigcup_{i \in B_{k'}} C_{ik'} \right] \right)$$

$$= \Pr\left(\left[\bigcup_{i \in B_k} C_{ik} \right] \right) + \Pr\left(\left[\bigcup_{i \in B_{k'}} C_{ik'} \right] \right) - \Pr\left(\left[\bigcup_{i \in B_k} C_{ik} \right] \cup \left[\bigcup_{i \in B_{k'}} C_{ik'} \right] \right)$$

$$= \pi_k + \pi_{k'} - \Pr\left(\left[\bigcup_{i \in B_k} C_{ik} \right] \cup \left[\bigcup_{i \in B_{k'}} C_{ik'} \right] \right)$$

Sementara $\left[\bigcup_{i\in B_k} C_{ik}\right]\cup\left[\bigcup_{i\in B_{k'}} C_{ik'}\right]=\bigcup_{i\in B_{kk'}} C_{ikk'}$, yaitu kejadian bahwa paling tidak satu unit sampling dari *network* ke k atau ke k' yang ada di dalam unit primer i, terpilih sebagai sampel awal. Untuk mencari $\Pr\left(\bigcup_{i\in B_{kk'}} C_{ikk'}\right)$, dapat digunakan cara yang sama ketika mencari $\Pr(\bigcup_{i\in B_k} C_{ik})$, yaitu

$$\Pr\left(\bigcup_{i\in B_{kk'}} C_{ikk'}\right)=\sum_{i\in B_{kk'}}\Pr(C_{ikk'})-\sum_{i}\sum_{i'<i}\Pr(C_{ikk'}\cap C_{i'kk'})+\ldots+(-1)^{g_{kk'}+1}\Pr(\bigcap_{i\in B_{kk'}} C_{ikk'})$$

dimana $B_{kk'}$ adalah himpunan unit primer yang beririsan dengan *network* ke k atau ke k', dan $g_{kk'}$ adalah banyaknya anggota himpunan $B_{kk'}$. Oleh karena itu, bentuk $\Pr\left(\bigcup_{i\in B_{kk'}} C_{ikk'}\right)$, sama seperti bentuk (3.1.4), dengan x_{ik} diubah menjadi $x_{ik}+x_{ik'}$, sehingga diperoleh

$$\begin{aligned}\Pr\left(\bigcup_{i\in B_{kk'}} C_{ikk'}\right)&=\sum_{i\in B_{kk'}}\frac{m}{M}\left[1-\frac{\binom{N_i-(x_{ik}+x_{ik'})}{n_i}}{\binom{N_i}{n_i}}\right]\\&-\sum_{i}\sum_{i'<i}\frac{m(m-1)}{M(M-1)}\left[1-\frac{\binom{N_i-(x_{ik}+x_{ik'})}{n_i}}{\binom{N_i}{n_i}}\right]\left[1-\frac{\binom{N_{i'}-(x_{ik}+x_{ik'})}{n_{i'}}}{\binom{N_{i'}}{n_{i'}}}\right]+\ldots\\&+(-1)^{g_{kk'}+1}\frac{m\left(m-1\right)\ldots(m-g_{kk'}+1)}{M\left(M-1\right)\ldots(M-g_{kk'}+1)}\prod_{i\in B_{kk'}}\left[1-\frac{\binom{N_i-(x_{ik}+x_{ik'})}{n_i}}{\binom{N_i}{n_i}}\right]\end{aligned}$$

Akhirnya diperoleh bentuk $\pi_{kk'}$ sebagai berikut :

$$\pi_{kk'} = \pi_k + \pi_{k'} - \left\{ \sum_{i \in B_{kk'}} \frac{m}{M} \left[1 - \frac{\binom{N_i - (x_{ik} + x_{ik'})}{n_i}}{\binom{N_i}{n_i}} \right] - \right.$$

$$\sum_i \sum_{i' < i} \frac{m(m-1)}{M(M-1)} \left[1 - \frac{\binom{N_i - (x_{ik} + x_{ik'})}{n_i}}{\binom{N_i}{n_i}} \right] \left[1 - \frac{\binom{N_{i'} - (x_{ik} + x_{ik'})}{n_{i'}}}{\binom{N_{i'}}{n_{i'}}} \right] + \ldots$$

$$\left. + (-1)^{g_{kk'}+1} \frac{m(m-1)\ldots(m - g_{kk'} + 1)}{M(M-1)\ldots(M - g_{kk'} + 1)} \prod_{i \in B_{kk'}} \left[1 - \frac{\binom{N_i - (x_{ik} + x_{ik'})}{n_i}}{\binom{N_i}{n_i}} \right] \right\} \quad (3.1.5)$$

Meskipun bentuk π_k dan $\pi_{kk'}$ sangat rumit, tetapi g_k dan $g_{kk'}$ biasanya kecil sehingga hanya beberapa suku yang dibutuhkan dalam perhitungan.

3.1.2 Penaksiran Mean Dan Total Populasi Dengan Taksiran Hansen-Hurwitz

Misalkan y_{ij} adalah total pengukuran pada unit sampling *j* di dalam unit primer *i*, p_i adalah probabilitas unit primer ke *i* terpilih sebagai sampel dan $p_{ii'}$ adalah probabilitas unit primer ke *i* dan *i'* terpilih sebagai sampel. Didefinisikan A_{ij} yaitu *network* yang dibangun oleh unit sampling ke *j* dalam unit primer *i,* dan A_{ijl} yaitu bagian dari *network* A_{ij} yang ada di dalam unit primer *l.* Misalkan f_{ijl} adalah banyaknya unit sampling pada sampel awal yang ada di *network* A_{ijl} dan a_{ijl} adalah banyaknya unit sampling pada *network* A_{ijl}. Dengan demikian, banyaknya sampel awal pada *network* A_{ij} adalah $f_{ij.} = \sum_{l=1}^{M} f_{ijl}$. Taksiran Hansen-Hurwitz untuk mean adalah

$$\tilde{\mu} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{y_{ij}}{E[f_{ij.}]} \tag{3.1.6}$$

Bentuk $E[f_{ij.}]$ dapat dicari sebagai berikut :

Misalkan z_l adalah variabel indikator, yaitu

$$z_l \begin{cases} 1, \text{ jika unit primer ke l terpilih sebagai sampel} \\ 0, \text{ jika unit primer ke l tidak terpilih sebagai sampel} \end{cases}$$

dimana $\Pr(z_l = 1) = p_l$. Jadi, apabila unit primer ke *l* tidak terpilih sebagai sampel maka $f_{ijl} = 0$, dan apabila unit primer ke *l* terpilih sebagai sampel maka $f_{ijl} \sim Hyper(N_l, a_{ijl}, n_l)$ dengan $E[f_{ijl} \mid z_l = 1] = n_l a_{ijl} / N_l$, sehingga

$$\begin{aligned}
E[f_{ij.}] &= \sum_{l=1}^{M} E[f_{ijl}] \\
&= \sum_{l=1}^{M} E[E[f_{ijl} \mid z_l]] \\
&= \sum_{l=1}^{M} \left\{ E[f_{ijl} \mid z_l = 0].\Pr[z_l = 0] + E[f_{ijl} \mid z_l = 1].\Pr[z_l = 1] \right\} \\
&= \sum_{l=1}^{M} E[f_{ijl} \mid z_l = 1].\Pr[z_l = 1] \\
&= \sum_{l=1}^{M} \frac{n_l a_{ijl}}{N_l}.p_l \\
&= \sum_{l=1}^{M} n_l a_{ijl} p_l / N_l
\end{aligned}$$

Nilai $E[f_{ij.}]$ adalah sama untuk setiap unit sampling pada *network* A_{ij}, maka taksiran mean pada (3.1.6) dapat pula ditulis sebagai

$$\begin{aligned}\tilde{\mu} &= \frac{1}{N}\sum_{i=1}^{m}\left(\sum_{j=1}^{n_i}\frac{1}{E[f_{ij.}]}\sum_{(i',j')\in A_{ij}} y_{i'j'}\right)\\ &= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{Y_{ij}}{E[f_{ij.}]}\end{aligned} \qquad (3.1.7)$$

dimana Y_{ij} adalah total pengukuran pada *network* A_{ij}.

Diasumsikan bahwa banyaknya unit sampling pada unit primer *i* adalah sama, dan banyaknya unit sampling yang terpilih sebagai sampel awal pada masing-masing unit primer juga sama, maka nilai p_i dan n_i/N_i untuk setiap unit primer adalah sama. Oleh karena itu, nilai ekspektasi dari $f_{ij.}$ menjadi $E[f_{ij.}] = \frac{n_i p_i}{N_i}\sum_{l=1}^{M} a_{ijl}$, sehingga taksiran mean pada (3.1.7) dapat ditulis sebagai berikut

$$\begin{aligned}\tilde{\mu} &= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{Y_{ij}}{E[f_{ij.}]}\\ &= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{Y_{ij}}{\sum_{l=1}^{M} a_{ijl}}\frac{N_i}{p_i n_i}\\ &= \frac{1}{N}\sum_{i=1}^{m}\frac{N_i}{p_i n_i}\sum_{j=1}^{n_i}\frac{Y_{ij}}{\sum_{l=1}^{M} a_{ijl}}\end{aligned}$$

misalkan $w_{ij} = Y_{ij}\Big/\sum_{l=1}^{M} a_{ijl}$ adalah mean / rata-rata pengukuran dari *network* A_{ij} dan $\bar{w}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}$ adalah taksiran mean pada unit primer ke *i*, maka

$$\tilde{\mu} = \frac{1}{N}\sum_{i=1}^{m}\frac{N_i}{p_i n_i}\sum_{j=1}^{n_i} w_{ij}$$
$$= \frac{1}{N}\sum_{i=1}^{m}\frac{N_i \bar{w}_i}{p_i}$$

Lalu, karena pemilihan unit primer dilakukan secara SRS tanpa pengembalian, maka $p_i = m/M$, sehingga bentuk taksiran mean menjadi

$$\tilde{\mu} = \frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} N_i \bar{w}_i \tag{3.1.8}$$

Kemudian, karena $\bar{w}_i$ adalah taksiran mean pada unit primer ke *i*, maka $N_i \bar{w}_i$ adalah taksiran total pada unit primer ke *i.* Jadi, dengan memisalkan $\hat{\tau}_i = N_i \bar{w}_i$, maka (3.1.8) dapat ditulis sebagai berikut

$$\tilde{\mu} = \frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} \hat{\tau}_i \tag{3.1.9}$$

Taksiran mean pada (3.1.9) sama seperti bentuk taksiran mean pada persamaan (2.2.1) yang ada pada pembahasan *Two Stage Sampling*.

Misalkan $W_i = \sum_{j=1}^{N_i} w_{ij}$ adalah total pada unit primer ke *i*, $\bar{W}_i = W_i / N_i$ adalah mean pada unit primer ke *i,* $\sigma_i^2 = \sum_{j=1}^{N_i} \frac{(w_{ij} - \bar{W}_i)^2}{N_i - 1}$ adalah variansi di dalam unit primer ke *i*, $\bar{W}. = \sum_{i=1}^{M} W_i \Big/ M$ adalah rata-rata per unit primer, dan $\sigma_u^2 = \sum_{i=1}^{M} \frac{(W_i - \bar{W}.)^2}{M - 1}$ adalah variansi antar unit primer.

Misalkan s_1 adalah himpunan semua unit primer yang terpilih sebagai sampel. Akan ditunjukkan bahwa diberikan himpunan s_1, taksiran mean dan total pada unit primer ke *i* adalah taksiran tak bias untuk mean dan total pada unit primer ke *i.* Dengan perkataan lain, akan ditunjukkan $E[\bar{w}_i \mid s_1] = \bar{W}_i$ dan $E[\hat{\tau}_i \mid s_1] = W_i$ sebagai berikut :

Misalkan z_{ij} adalah variabel indikator, yaitu :

$$z_{ij} = \begin{cases} 1, \text{ jika unit sampling ke j pada unit primer ke i terpilih sebagai sampel} \\ 0, \text{ jika unit sampling ke j pada unit primer ke i tidak terpilih sebagai sampel} \end{cases}$$

dimana $\Pr(z_{ij} = 1 \mid s_1) = \dfrac{n_i}{N_i}$, maka

$$\begin{aligned}
E[\bar{w}_i \mid s_1] &= E\left[\frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij} \mid s_1\right] \\
&= E\left[\frac{1}{n_i}\sum_{j=1}^{N_i} w_{ij} z_{ij} \mid s_1\right] \\
&= \frac{1}{n_i}\sum_{j=1}^{N_i} w_{ij} E[z_{ij} \mid s_1] \\
&= \frac{1}{n_i}\sum_{j=1}^{N_i} w_{ij}.\left\{0.\Pr(z_{ij} = 0 \mid s_1) + 1.\Pr(z_{ij} = 1 \mid s_1)\right\} \\
&= \frac{1}{n_i}\sum_{j=1}^{N_i} w_{ij}.\left\{1.\Pr(z_{ij} = 1 \mid s_1)\right\} \\
&= \frac{1}{n_i}\sum_{j=1}^{N_i} w_{ij}.\frac{n_i}{N_i} \\
&= \frac{1}{N_i}\sum_{j=1}^{N_i} w_{ij} \\
&= W_i / N_i \\
&= \bar{W}_i
\end{aligned}$$

$E[\hat{\tau}_i \mid s_1] = E[N_i \bar{w}_i \mid s_1] = N_i E[\bar{w}_i \mid s_1] = N_i \bar{W}_i = W_i$ Kemudian, akan ditunjukkan bahwa $\tilde{\mu}$ adalah taksiran tak bias untuk mean populasi.

Misalkan z_i adalah variabel indikator, yaitu :

$$z_i = \begin{cases} 1, \text{ jika unit primer ke i terpilih sebagai sampel} \\ 0, \text{ jika unit primer ke i tidak terpilih sebagai sampel} \end{cases}$$

dimana $\Pr(z_i = 1) = \dfrac{n}{N}$, sehingga

$$\begin{aligned}
E[\tilde{\mu}] &= E[E[\tilde{\mu} \mid s_1]] \\
&= E\left[E\left[\frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} \hat{\tau}_i \mid s_1 \right]\right] \\
&= E\left[\frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} E[\hat{\tau}_i \mid s_1] \right] \\
&= E\left[\frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} W_i \right] \\
&= E\left[\frac{1}{N}\frac{M}{m}\sum_{i=1}^{M} W_i z_i \right] \\
&= \frac{1}{N}\frac{M}{m}\sum_{i=1}^{M} W_i E[z_i] \\
&= \frac{1}{N}\frac{M}{m}\sum_{i=1}^{M} W_i \left\{ 0.\Pr(z_i = 0) + 1.\Pr(z_i = 1) \right\} \\
&= \frac{1}{N}\frac{M}{m}\sum_{i=1}^{M} W_i \left\{ 1.\Pr(z_i = 1) \right\} \\
&= \frac{1}{N}\frac{M}{m}\sum_{i=1}^{M} W_i \frac{m}{M} \\
&= \frac{1}{N}\sum_{i=1}^{M} W_i \\
&= \frac{\tau}{N} \\
&= \mu
\end{aligned} \tag{3.1.10}$$

Jadi, $\tilde{\mu}$ adalah taksiran tak bias untuk mean populasi

Selanjutnya, berdasarkan persamaan (3.1.9), variansi dari $\tilde{\mu}$ dapat dicari dengan menggunakan dekomposisi sebagai berikut

$$\text{var}[\tilde{\mu}] = \text{var}[E(\tilde{\mu} \mid s_1)] + E[\text{var}(\tilde{\mu} \mid s_1)]$$

bentuk pertama dari dekomposisi variansi dapat dicari sebagai berikut :

$$
\begin{aligned}
\text{var}[E(\tilde{\mu} \mid s_1)] &= \text{var}\left[E\left(\frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right)\right] \\
&= \text{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}E[\hat{\tau}_i \mid s_1]\right] \\
&= \text{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}W_i\right] \\
&= \frac{M^2}{N^2}\text{var}\left[\frac{1}{m}\sum_{i=1}^{m}W_i\right] \\
&= \frac{M^2}{N^2}\text{var}\left[\bar{W}\right] \\
&= \frac{M^2}{N^2}.\left(\frac{M-m}{M}\right)\frac{\sigma_u^2}{m} \\
&= \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m}
\end{aligned}
$$

sedangkan bentuk kedua dari dekomposisi variansi dapat dicari sebagai berikut :

$$
\begin{aligned}
E[\text{var}(\tilde{\mu} \mid s_1)] &= E\left[\text{var}\left(\frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right)\right] \\
&= E\left[\left(\frac{M}{Nm}\right)^2\sum_{i=1}^{m}\text{var}[\hat{\tau}_i \mid s_1]\right] \\
&= E\left[\left(\frac{M}{Nm}\right)^2\sum_{i=1}^{m}N_i(N_i-n_i)\frac{\sigma_i^2}{n_i}\right] \\
&= E\left[\left(\frac{M}{Nm}\right)^2\sum_{i=1}^{M}N_i(N_i-n_i)\frac{\sigma_i^2}{n_i}.z_i\right]
\end{aligned}
$$

$$= \left(\frac{M}{Nm}\right)^2 \sum_{i=1}^{M} N_i (N_i - n_i) \frac{\sigma_i^2}{n_i} . E[z_i]$$

$$= \left(\frac{M}{Nm}\right)^2 \sum_{i=1}^{M} N_i (N_i - n_i) \frac{\sigma_i^2}{n_i} . \frac{m}{M}$$

$$= \frac{1}{N^2} \frac{M}{m} \sum_{i=1}^{M} N_i (N_i - n_i) \frac{\sigma_i^2}{n_i}$$

Jadi, variansi dari $\tilde{\mu}$ adalah

$$\text{var}[\tilde{\mu}] = \frac{1}{N^2} M(M-m) \frac{\sigma_u^2}{m} + \frac{1}{N^2} \frac{M}{m} \sum_{i=1}^{M} N_i (N_i - n_i) \frac{\sigma_i^2}{n_i} \qquad (3.1.11)$$

karena bentuk (3.1.11) sama seperti (2.2.3) pada yang ada pada pembahasan *two stage sampling*, maka bentuk taksirannya juga sama seperti (2.2.4), yaitu :

$$\widehat{\text{var}}[\tilde{\mu}] = \frac{1}{N^2} M(M-m) \frac{s_u^2}{m} + \frac{1}{N^2} \frac{M}{m} \sum_{i=1}^{m} N_i (N_i - n_i) \frac{s_i^2}{n_i} \qquad (3.1.12)$$

dimana $s_u^2 = \frac{1}{m-1} \sum_{i=1}^{m} \left(N_i \bar{w}_i - \sum_{i=1}^{m} N_i \bar{w}_i \Big/ m \right)^2$, dan $s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(w_{ij} - \bar{w}_i \right)^2$,

untuk *i= 1, 2,…,m.* Dengan cara yang sama seperti pada pembahasan *two stage sampling*, dapat dibuktikan bahwa taksiran variansi pada (3.1.12) adalah taksiran takbias untuk (3.1.11) atau $E[\widehat{\text{var}}[\tilde{\mu}]] = \text{var}[\tilde{\mu}]$.

Selanjutnya, taksiran Hansen-Hurwitz untuk total adalah

$$\tilde{\tau} = N.\tilde{\mu}$$

$$= \frac{M}{m} \sum_{i=1}^{m} N_i \bar{w}_i$$

$$= \frac{M}{m} \sum_{i=1}^{m} \hat{\tau}_i$$

sedangkan variansi dan taksiran variansi dari taksiran total tersebut adalah

$$\begin{aligned}\text{var}[\tilde{\tau}] &= \text{var}[N.\tilde{\mu}] \\ &= M(M-m)\frac{\sigma_u^2}{m} + \frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i} \\ \widehat{\text{var}}[\tilde{\tau}] &= \widehat{\text{var}}[N.\tilde{\mu}] \\ &= M(M-m)\frac{s_u^2}{m} + \frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i}\end{aligned}$$

dengan mnggunakan cara yang sama seperti pada taksiran mean, dapat dibuktikan bahwa taksiran total dan taksiran variansi tersebut adalah taksiran tak bias untuk total dan variansi.

3.2 Skema *Nonoverlapping*

Pada skema *nonoverlapping*, proses penambahan sampel tidak diperbolehkan melewati batas unit primer. Dengan demikian, proses penambahan sampel berhenti ketika tidak ada lagi unit sampling yang bertetanggaan dengan sampel yang memenuhi kondisi C dan juga ketika penambahan sampel telah mencapai batas unit primer.

3.2.1 Penaksiran Mean Dan Total Populasi Dengan Taksiran Horvitz-Thompson

Misalkan *N* adalah banyaknya unit sampling pada populasi, K_i adalah banyaknya *network* pada unit primer *i*, κ_i adalah banyaknya *network* pada unit primer *i* yang terpilih sebagai sampel, τ^*_{ik} adalah total pengukuran pada *network k* di dalam unit primer *i* pada populasi, y^*_{ik} adalah total pengukuran pada *network k* di dalam unit primer *i* pada sampel, π_{ik} adalah probabilitas *network k* pada unit primer *i* terpilih sebagai sampel, dan $\hat{\tau}_i = \sum_{k=1}^{\kappa_i} y^*_{ik} \big/ \pi_{ik}$ adalah taksiran total pada unit primer *i*, maka taksiran Horvitz-Thompson untuk mean adalah

$$\begin{aligned} \hat{\mu}_1 &= \frac{M}{Nm} \sum_{i=1}^{m} \sum_{k=1}^{\kappa_i} \frac{y^*_{ik}}{\pi_{ik}} \\ &= \frac{M}{Nm} \sum_{i=1}^{m} \hat{\tau}_i \end{aligned} \tag{3.2.1}$$

Dengan menggunakan cara yang sama seperti pada (2.4.2), akan diperoleh bentuk π_{ik} sebagai berikut

$$\pi_{ik} = 1 - p_{ik} = 1 - \left[\binom{N_i - x_{ik}}{n_i} \Big/ \binom{N_i}{n_i} \right] \tag{3.2.2}$$

Kemudian, akan ditunjukkan bahwa $\hat{\tau}_i$ adalah taksiran tak bias untuk total pada unit primer *i*. Misalkan s_1 adalah himpunan unit primer dalam sampel, z_{ik} adalah variabel indikator, yaitu

$$z_{ik} = \begin{cases} 1, \text{ jika network j pada unit primer i terpilih sebagai sampel} \\ 0, \text{ jika network j pada unit primer i tidak terpilih sebagai sampel} \end{cases}$$

dimana $\Pr(z_{ik} = 1 \mid s_1) = \pi_{ik}$, maka

$$
\begin{aligned}
E[\hat{\tau}_i \mid s_1] &= E\left[\sum_{k=1}^{K_i} \frac{y_{ik}^*}{\pi_{ik}} \mid s_1\right] \\
&= E\left[\sum_{k=1}^{K_i} \frac{\tau_{ik}^* z_{ik}}{\pi_{ik}} \mid s_1\right] \\
&= \sum_{k=1}^{K_i} \frac{\tau_{ik}^*}{\pi_{ik}}.E[z_{ik} \mid s_1] \\
&= \sum_{k=1}^{K_i} \frac{\tau_{ik}^*}{\pi_{ik}}.\left\{0.\Pr(z_{ik} = 0 \mid s_1) + 1.\Pr(z_{ik} = 1 \mid s_1)\right\} \\
&= \sum_{k=1}^{K_i} \frac{\tau_{ik}^*}{\pi_{ik}}.\left\{1.\Pr(z_{ik} = 1 \mid s_1)\right\} \\
&= \sum_{k=1}^{K_i} \frac{\tau_{ik}^*}{\pi_{ik}}.\pi_{ik} \\
&= \sum_{k=1}^{K_i} \tau_{ik}^* \\
&= \tau_i
\end{aligned}
$$

Jadi, $\hat{\tau}_i$ adalah taksiran tak bias untuk τ_i (total pada unit primer *i*).

Kemudian, akan akan ditunukkan bahwa $\hat{\mu}_1$ adalah taksiran tak bias untuk mean populasi. Misalkan z_i adalah variabel indikator, yaitu

$$
z_i = \begin{cases} 1, & \text{jika unit primer ke i terpilih sebagai sampel} \\ 0, & \text{jika unit primer ke i tidak terpilih sebagai sampel} \end{cases}
$$

dimana $\Pr(z_i = 1) = \dfrac{m}{M}$, maka

$$
\begin{aligned}
E[\hat{\mu}_1] &= E[E[\hat{\mu}_1 \mid s_1]] \\
&= E\left[E\left[\frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right]\right] \\
&= E\left[\frac{M}{Nm}\sum_{i=1}^{m}E[\hat{\tau}_i \mid s_1]\right] \\
&= E\left[\frac{M}{Nm}\sum_{i=1}^{m}\tau_i\right] \\
&= E\left[\frac{M}{Nm}\sum_{i=1}^{M}\tau_i z_i\right] \\
&= \frac{M}{Nm}\sum_{i=1}^{M}\tau_i E[z_i] \\
&= \frac{M}{Nm}\sum_{i=1}^{M}\tau_i . \frac{m}{M} \\
&= \frac{1}{N}\sum_{i=1}^{M}\tau_i \\
&= \mu
\end{aligned}
$$

Jadi, $\hat{\mu}_1$ merupakan taksiran tak bias untuk mean populasi.

Untuk mencari variansi dari $\hat{\mu}_1$, sebelumnya akan dijelaskan dahulu mengenai variansi dari $\hat{\tau}_i = \sum_{k=1}^{\kappa_i} y^*_{ik} / \pi_{ik}$, atau variansi dari *network* pada unit primer *i*, yaitu :

Berdasarkan (2.4.4) dan (2.4.5), diperoleh $\mathrm{var}[z_{ik} \mid s_1] = \pi_{ik}(1-\pi_{ik})$ dan $\mathrm{cov}[(z_{ik}, z_{ik'}) \mid s_1] = \pi_{ikk'} - \pi_{ik}\pi_{ik'}$, sehingga

$$\begin{aligned}
\operatorname{var}[\hat{\tau}_i \mid s_1] &= \operatorname{var}\left[\sum_{k=1}^{\kappa_i} \frac{y_{ik}^*}{\pi_{ik}} \mid s_1\right] \\
&= \operatorname{var}\left[\sum_{k=1}^{K_i} \frac{\tau_{ik}^* z_{ik}}{\pi_{ik}} \mid s_1\right] \\
&= \sum_{k=1}^{K_i}\left(\frac{\tau_{ik}^*}{\pi_{ik}}\right)^2 \operatorname{var}[z_{ik} \mid s_1] + \sum_{k=1}^{K_i}\sum_{k'\neq k} \frac{\tau_{ik}^*\tau_{ik'}^*}{\pi_{ik}\pi_{ik'}} \operatorname{cov}[(z_{ik}, z_{ik'}) \mid s_1] \\
&= \sum_{k=1}^{K_i}\left(\frac{\tau_{ik}^*}{\pi_{ik}}\right)^2 \pi_{ik}(1-\pi_{ik}) + \sum_{k=1}^{K_i}\sum_{k'\neq k} \tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right) \\
&= \sum_{k=1}^{K_i}(\tau_{ik}^*)^2\left(\frac{1-\pi_{ik}}{\pi_{ik}}\right) + \sum_{k=1}^{K_i}\sum_{k'\neq k} \tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right) \\
&= \sum_{k=1}^{K_i}\sum_{k'=1}^{K_i} \tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)
\end{aligned}$$

Dengan demikian, variansi dari $\hat{\mu}_1$ dapat ditentukan sebagai berikut

$$\begin{aligned}
\operatorname{var}[\hat{\mu}_1] &= \operatorname{var}[E(\hat{\mu}_1 \mid s_1)] + E[\operatorname{var}(\hat{\mu}_1 \mid s_1)] \\
&= \operatorname{var}\left[E\left[\frac{M}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{\kappa_i}\frac{y_{ik}^*}{\pi_{ik}} \mid s_1\right]\right] + E\left[\operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{\kappa_i}\frac{y_{ik}^*}{\pi_{ik}} \mid s_1\right]\right] \\
&= \operatorname{var}\left[E\left[\frac{M}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{K_i}\frac{\tau_{ik}^* z_{ik}}{\pi_{ik}} \mid s_1\right]\right] + E\left[\operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right]\right] \\
&= \operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{K_i}\frac{\tau_{ik}^*}{\pi_{ik}}E[z_{ik} \mid s_1]\right] + E\left[\left(\frac{M}{Nm}\right)^2\sum_{i=1}^{m}\operatorname{var}\left[\hat{\tau}_i \mid s_1\right]\right] \\
&= \operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{K_i}\frac{\tau_{ik}^*}{\pi_{ik}}.\pi_{ik}\right] + E\left[\left(\frac{M}{Nm}\right)^2\sum_{i=1}^{m}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^*\tau_{ik'}^*\right)\right] \\
&= \operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{K_i}\tau_{ik}^*\right] + E\left[\left(\frac{M}{Nm}\right)^2\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^*\tau_{ik'}^*\right).z_i\right] \\
&= \operatorname{var}\left[\frac{M}{Nm}\sum_{i=1}^{m}\tau_i\right] + \left(\frac{M}{Nm}\right)^2\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^*\tau_{ik'}^*\right).E[z_i] \\
&= \left(\frac{M}{N}\right)^2\operatorname{var}\left[\frac{1}{m}\sum_{i=1}^{m}\tau_i\right] + \left(\frac{M}{Nm}\right)^2\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^*\tau_{ik'}^*\right).\frac{m}{M}
\end{aligned}$$

$$= \left(\frac{M}{N}\right)^2 \left(\frac{M-m}{M}\right)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right)$$

$$= \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right)$$

Dengan demikian

$$\text{var}[\hat{\mu}_1] = \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right)$$

dimana $\sigma_u^2 = \sum_{i=1}^{M}\frac{(\tau_i - \bar{\tau})^2}{M-1}$ adalah variansi antar unit primer dan $\bar{\tau} = \frac{1}{M}\sum_{i=1}^{M}\tau_i$

adalah rata-rata antar unit primer. Dengan memisalkan

$V_i = \left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right)$, variansi tersebut dapat ditulis sebagai berikut

$$\text{var}[\hat{\mu}_1] = \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}V_i$$

Taksiran dari variansi tersebut adalah

$$\widehat{\text{var}}[\hat{\mu}_1] = \frac{1}{N^2}M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}\hat{V}_i \qquad (3.2.3)$$

dimana $s_u^2 = \sum_{i=1}^{m}\frac{(\hat{\tau}_i - \sum_{i=1}^{m}\hat{\tau}_i/m)^2}{m-1}$ dan $\hat{V}_i = \left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}\pi_{ik}\pi_{ik'}}\right)y_{ik}^{*}y_{ik'}^{*}\right)$.

Selanjutnya, dengan menggunakan cara yang sama seperti pada (2.4.8), akan diperoleh bentuk $\pi_{ikk'}$ sebagai berikut

$$\pi_{ikk'} = 1 - \left[\binom{N_i - x_{ik}}{n_i} + \binom{N_i - x_{ik'}}{n_i} - \binom{N_i - x_{ik} - x_{ik'}}{n_i}\right] \Big/ \binom{N_i}{n_i} \qquad (3.2.4)$$

Selanjutnya akan ditunjukkan bahwa $\widehat{\text{var}}[\hat{\mu}_1]$ adalah taksiran tak bias untuk $\text{var}[\hat{\mu}_1]$, yaitu :

$$E[\widehat{\text{var}}[\hat{\mu}_1]] = E[E(\widehat{\text{var}}[\hat{\mu}_1] \mid s_1)]$$

$$= E\left[E\left[\left(\frac{1}{N^2}M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}\pi_{ik}\pi_{ik'}}\right)y_{ik}^{*}y_{ik'}^{*}\right)\right) \mid s_1\right]\right]$$

$$= E\left[E\left(\frac{1}{N^2}M(M-m)\frac{s_u^2}{m} \mid s_1\right) + E\left(\frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}\pi_{ik}\pi_{ik'}}\right)y_{ik}^{*}y_{ik'}^{*}\right) \mid s_1\right)\right]$$

$$= E\left[E\left(\frac{1}{N^2}M(M-m)\frac{s_u^2}{m} \mid s_1\right)\right] + E\left[E\left(\frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}\pi_{ik}\pi_{ik'}}\right)y_{ik}^{*}y_{ik'}^{*}\right) \mid s_1\right)\right]$$

$$= \underbrace{\frac{1}{N^2}\frac{M(M-m)}{(m-1)m}\underbrace{E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i - \sum_{i=1}^{m}\hat{\tau}_i/m\right)^2 \mid s_1\right]\right]}_{(*)}}_{\text{(I)}}$$

$$+\frac{1}{N^2}\underbrace{E\left[\frac{M}{m}\sum_{i=1}^{m}E\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}\pi_{ik}\pi_{ik'}}\right)y_{ik}^{*}y_{ik'}^{*} \mid s_1\right)\right]}_{\text{(II)}}$$

Ada dua bagian yang akan dicari, pada bagian pertama (I) akan dicari dahulu bentuk (*) yaitu

$$E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i - \sum_{i=1}^{m}\hat{\tau}_i/m\right)^2 \mid s_1\right]\right] = E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i^2 - 2\hat{\tau}_i\left(\sum_{i=1}^{m}\hat{\tau}_i/m\right) + \left(\sum_{i=1}^{m}\hat{\tau}_i/m\right)^2\right) \mid s_1\right]\right]$$

$$= E\left[E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i^2 - 2\left(\frac{1}{m}\sum_{i=1}^{m}\hat{\tau}_i\right)\sum_{i=1}^{m}\hat{\tau}_i + m\left(\frac{1}{m}\sum_{i=1}^{m}\hat{\tau}_i\right)^2\right) \mid s_1\right]\right]$$

$$= E\left[E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i^2 - 2\frac{1}{m}\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 + \frac{1}{m}\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2\right) \mid s_1\right]\right]$$

$$= E\left[E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i^2 - \frac{1}{m}\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2\right) \mid s_1\right]\right]$$

$$= E\left[E_2\left(\sum_{i=1}^{m}\hat{\tau}_i^2 \mid s_1\right) - \frac{1}{m}E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 \mid s_1\right]\right]$$

$$= \underbrace{E\left[E\left(\sum_{i=1}^{m}\hat{\tau}_i^2 \mid s_1\right)\right]}_{(a)} - \underbrace{E\left[\frac{1}{m}E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 \mid s_1\right]\right]}_{(b)}$$

bentuk (a) dan (b) dapat dicari sebagai berikut :

(a)

$$E\left[E\left(\sum_{i=1}^{m}\hat{\tau}_i^2 \mid s_1\right)\right] = E\left[\sum_{i=1}^{m}E\left(\hat{\tau}_i^2 \mid s_1\right)\right]$$

$$= E\left[\sum_{i=1}^{m}\left(\left\{E[\hat{\tau}_i \mid s_1]\right\}^2 + \text{var}[\hat{\tau}_i \mid s_1]\right)\right]$$

$$= E\left[\sum_{i=1}^{m}\left(\tau_i^2 + \sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right)\right]$$

$$= E\left[\sum_{i=1}^{m}\tau_i^2\right] + E\left[\sum_{i=1}^{m}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right]$$

$$= E\left[\sum_{i=1}^{M}\tau_i^2 z_i\right] + E\left[\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}.z_i\right]$$

$$= \frac{m}{M}\sum_{i=1}^{M}\tau_i^2 + \frac{m}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

(b)

$$E\left[\frac{1}{m}E\left[(\textstyle\sum_{i=1}^{m}\hat{\tau}_i)^2 \mid s_1\right]\right] = E\left[\frac{1}{m}\left(\left\{E\left[\textstyle\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right]\right\}^2 + \text{var}\left[\textstyle\sum_{i=1}^{m}\hat{\tau}_i \mid s_1\right]\right)\right]$$

$$= E\left[\frac{1}{m}\left(\left\{\textstyle\sum_{i=1}^{m}E\left(\hat{\tau}_i \mid s_1\right)\right\}^2 + \textstyle\sum_{i=1}^{m}\text{var}[\hat{\tau}_i \mid s_1]\right)\right]$$

$$= E\left[\frac{1}{m}\left(\left\{\textstyle\sum_{i=1}^{m}\tau_i\right\}^2 + \textstyle\sum_{i=1}^{m}\left\{\displaystyle\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right)\right]$$

$$= E\left(m\hat{\tau}^2 + \frac{1}{m}\textstyle\sum_{i=1}^{m}\left\{\displaystyle\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right)$$

$$= mE[\hat{\tau}^2] + \frac{1}{m}E\left[\textstyle\sum_{i=1}^{m}\left\{\displaystyle\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right]$$

$$= m\left(\left\{E[\hat{\tau}]\right\}^2 + \text{var}[\hat{\tau}]\right) + \frac{1}{m}E\left[\textstyle\sum_{i=1}^{M}\left\{\displaystyle\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^*\tau_{ik'}^*\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}.z_i\right]$$

$$= m\left(\left\{E[\frac{1}{m}\sum_{i=1}^{M}\tau_i z_i]\right\}^2 + \left(\frac{M-m}{M}\right)\frac{\sigma_u^2}{m}\right.$$

$$\left. + \frac{1}{m}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}.E\left[z_i\right]\right.$$

$$= m\left(\left\{\frac{1}{M}\sum_{i=1}^{M}\tau_i\right\}^2 + \left(\frac{M-m}{M}\right)\frac{\sigma_u^2}{m}\right)$$

$$+ \frac{1}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

$$= m\left(\bar{\tau}\right)^2 + \left(\frac{M-m}{M}\right)\sigma_u^2 + \frac{1}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

sehingga (*) menjadi

$$E\left[E\left[\sum_{i=1}^{m}\left(\hat{\tau}_i - \sum_{i=1}^{m}\hat{\tau}_i/m\right)^2 \mid s_1\right]\right]$$

$$= E\left[\sum_{i=1}^{m}E[\hat{\tau}_i^2 \mid s_1]\right] - E\left[\frac{1}{m}E\left[\left(\sum_{i=1}^{m}\hat{\tau}_i\right)^2 \mid s_1\right]\right]$$

$$= \left\{\frac{m}{M}\sum_{i=1}^{M}\tau_i^2 + \frac{m}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right\}$$

$$-\left\{m\left(\bar{\tau}\right)^2 + \left(\frac{M-m}{M}\right)\sigma_u^2 + \frac{1}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right\}$$

$$= \frac{m}{M}\sum_{i=1}^{M}\tau_i^2 - m\left(\bar{\tau}\right)^2 - \left(\frac{M-m}{M}\right)\sigma_u^2 + \frac{m}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

$$-\frac{1}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{j=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

$$= \frac{m}{M}\left(\sum_{i=1}^{M}\tau_i^2 - M\left(\bar{\tau}\right)^2\right) - \left(1-\frac{m}{M}\right)\sigma_u^2$$

$$+ \frac{(m-1)}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

$$= \frac{m}{M} \cdot \frac{M-1}{M-1} \cdot \sum_{i=1}^{M} (\tau_i - \overline{\tau})^2 - \sigma_u^2 + \frac{m}{M} \sigma_u^2$$

$$+ \frac{(m-1)}{M} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\}$$

$$= \frac{m}{M} . (M-1)\sigma_u^2 - \sigma_u^2 + \frac{m}{M} \sigma_u^2 + \frac{(m-1)}{M} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\}$$

$$= m\sigma_u^2 - \frac{m}{M} \sigma_u^2 - \sigma_u^2 + \frac{m}{M} \sigma_u^2 + \frac{(m-1)}{M} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\}$$

$$= m\sigma_u^2 - \sigma_u^2 + \frac{(m-1)}{M} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\}$$

$$= (m-1)\sigma_u^2 + \frac{(m-1)}{M} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\}$$

$$= (m-1) \left[\sigma_u^2 + \frac{1}{M} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\} \right]$$

maka bentuk (I) adalah

$$\frac{1}{N^2} \frac{M(M-m)}{(m-1)m} E\left[E\left[\sum_{i=1}^{m} \left(\hat{\tau}_i - \sum\nolimits_{i=1}^{m} \hat{\tau}_i / m \right)^2 \mid s_1 \right] \right]$$

$$= \frac{1}{N^2} \frac{M(M-m)}{(m-1)m} . (m-1) \left[\sigma_u^2 + \frac{1}{M} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\} \right]$$

$$= \frac{1}{N^2} M(M-m) \frac{\sigma_u^2}{m} + \frac{1}{N^2} \frac{(M-m)}{m} \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K_i} \sum_{k'=1}^{K_i} \tau_{ik}^* \tau_{ik'}^* \left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}} \right) \right\}$$

Selanjutnya akan dicari bentuk kedua (II) yaitu :

$$\frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{m}E\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ikk'}.\pi_{ik}\pi_{ik'}}\right)y_{ik}^{*}y_{ik'}^{*} \mid s_1\right)\right]$$

$$=\frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{m}E\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ikk'}.\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}.z_{ik}z_{ik'} \mid s_1\right)\right]$$

$$=\frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ikk'}.\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}.E[z_{ik}z_{ik'} \mid s_1]\right)\right]$$

$$=\frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ikk'}.\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}.\pi_{ikk'}\right)\right]$$

$$=\frac{1}{N^2}E\left[\frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right).z_i\right]$$

$$=\frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right).E[z_i]$$

$$=\frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right).\frac{m}{M}$$

$$=\frac{1}{N^2}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^{*}\tau_{ik'}^{*}\right)$$

maka,

$$E[\widehat{\text{var}}[\hat{\mu}_1]]=\left(\frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m}+\frac{1}{N^2}\frac{(M-m)}{m}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right)$$

$$+\left(\frac{1}{N^2}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}\right)$$

$$=\frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m}+\frac{1}{N^2}\left(\frac{M}{m}-1\right)\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

$$+\frac{1}{N^2}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

$$=\frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m}+\frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}\left\{\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\tau_{ik}^{*}\tau_{ik'}^{*}\left(\frac{\pi_{ikk'}-\pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\right\}$$

$$=\text{var}[\hat{\mu}_1]$$

Jadi, karena $E[\widehat{\text{var}}[\hat{\mu}_1]] = \text{var}[\hat{\mu}_1]$ maka $\widehat{\text{var}}[\hat{\mu}_1]$ merupakan taksiran tak bias untuk $\text{var}[\hat{\mu}_1]$.

Selanjutnya, taksiran Horvitz-Thompson untuk total adalah

$$\begin{aligned}\hat{\tau}_1 &= \frac{M}{m}\sum_{i=1}^{m}\sum_{k=1}^{\kappa_i}\frac{y_{ik}^*}{\pi_{ik}} \\ &= \frac{M}{m}\sum_{i=1}^{m}\hat{\tau}_i\end{aligned}$$

sedangkan variansi dan taksiran variansu untuk taksiran total tersebut adalah

$$\begin{aligned}\text{var}[\hat{\tau}_1] &= \text{var}[N.\hat{\mu}_1] \\ &= M(M-m)\frac{\sigma_u^2}{m} + \frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^*\tau_{ik'}^*\right) \\ \widehat{\text{var}}[\hat{\tau}_1] &= \widehat{\text{var}}[N.\hat{\mu}_1] \\ &= M(M-m)\frac{s_u^2}{m} + \frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}.\pi_{ik}\pi_{ik'}}\right)y_{ik}^*y_{ik'}^*\right)\end{aligned}$$

Dengan menggunakan cara yang sama seperti pada taksiran mean, dapat ditunjukkan bahwa taksiran total dan taksiran variansi tersebut adalah taksiran tak bias untuk total dan variansi dari taksiran total.

3.2.2 Penaksiran Mean Dan Total Populasi Dengan Taksiran Hansen-Hurwitz

Misalkan y_{ij} adalah total pengukuran pada unit sampling *j* di dalam unit primer *i*, a_{ij} adalah banyaknya unit sampling pada *network* A_{ij}, f_{ij} adalah banyaknya sampel awal pada *network* A_{ij}, p_i adalah probabilitas unit primer ke *i* terpilih sebagai sampel, dan p_{ij} adalah probabilitas unit primer ke *i* dan *j* terpilih sebagai sampel. Taksiran Hansen-Hurwitz untuk mean didefinisikan sebagai berikut

$$\tilde{\mu}_1 = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{y_{ij}}{E[f_{ij}]} \tag{3.2.5}$$

Bentuk $E[f_{ij}]$ dapat dicari sebagai berikut :

Misalkan z_i adalah variabel indikator, yaitu

$$z_i \begin{cases} 1, \text{ jika unit primer ke i terpilih sebagai sampel} \\ 0, \text{ jika unit primer ke i tidak terpilih sebagai sampel} \end{cases}$$

dimana $\Pr(z_i = 1) = p_i$. Jadi, apabila unit primer ke *i* tidak terpilih sebagai sampel maka $f_{ij} = 0$, dan apabila unit primer ke *i* terpilih sebagai sampel maka $f_{ij} \sim Hyper(N_i, a_{ij}, n_i)$ dengan $E[f_{ij} \mid z_i = 1] = \frac{n_i a_{ij}}{N_i}$ sehingga

$$\begin{aligned} E[f_{ij}] &= E[E[f_{ij} \mid z_i]] \\ &= E[f_{ij} \mid z_i = 0].\Pr[z_i = 0] + E[f_{ij} \mid z_i = 1].\Pr[z_i = 1] \\ &= E[f_{ij} \mid z_i = 1].\Pr[z_i = 1] \\ &= \frac{n_i a_{ij}}{N_i}.p_i \\ &= n_i a_{ij} p_i / N_i \end{aligned}$$

Nilai $E[f_{ij}]$ adalah sama untuk setiap unit pada *network* A_{ij}, maka taksiran mean pada (3.2.5) dapat pula ditulis sebagai

$$\tilde{\mu}_1 = \frac{1}{N}\sum_{i=1}^{m}\left(\sum_{j=1}^{n_i}\frac{1}{E[f_{ij}]}\sum_{(i',j')\in A_{ij}} y_{i'j'}\right)$$

$$= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{Y_{ij}}{E[f_{ij}]} \tag{3.2.6}$$

dimana Y_{ij} adalah total pengukuran pada *network* A_{ij} .

Misalkan $w_{ij}^{\#} = Y_{ij}/a_{ij}$ adalah mean dari *network* A_{ij} , dan

$\bar{w}_i^{\#} = \frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}^{\#}$ adalah taksiran mean pada unit primer ke *i*. Lalu, karena

$E[f_{ij}] = n_i a_{ij} p_i / N_i$ maka taksiran mean pada (3.2.6) dapat ditulis sebagai

$$\begin{aligned}\tilde{\mu}_1 &= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{Y_{ij}}{E[f_{ij}]}\\ &= \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{Y_{ij}}{a_{ij}}\frac{N_i}{n_i p_i}\\ &= \frac{1}{N}\sum_{i=1}^{m}\frac{N_i}{p_i n_i}\sum_{j=1}^{n_i}\frac{Y_{ij}}{a_{ij}}\\ &= \frac{1}{N}\sum_{i=1}^{m}\frac{N_i}{p_i}\frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}^{\#}\\ &= \frac{1}{N}\sum_{i=1}^{m}\frac{N_i}{p_i}\bar{w}_i^{\#}\end{aligned}$$

karena pemilihan unit primer dilakukan secara SRS tanpa pengembalian, maka $p_i = m/M$, sehingga

$$\tilde{\mu}_1 = \frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} N_i \bar{w}_i^{\#} \tag{3.2.7}$$

Cara yang sama seperti (3.1.10) dapat digunakan untuk menunjukkan bahwa $\tilde{\mu}_1$ adalah taksiran tak bias untuk mean populasi.

Pada dasarnya, bentuk modifikasi taksiran Hansen-Hurwitz pada skema *overlapping* dan skema *nonoverlapping* adalah sama, tetapi pada skema *nonoverlapping*, *network* dibatasi oleh batas unit primer. Hal ini terlihat pada taksiran mean pada (3.1.8) dan (3.2.7), akan tetapi pada (3.1.8) simbol yang digunakan adalah $\bar{w}_i$, sedangkan pada (3.2.7) digunakan simbol $\bar{w}_i^{\#}$. Oleh karena bentuk taksiran mean pada (3.2.7) sama seperti (3.1.8), maka bentuk variansi dan taksiran variansi dari (3.2.7) adalah sama seperti bentuk variansi dan taksiran variansi pada (3.1.11) dan (3.1.12), yaitu

$$\text{var}[\tilde{\mu}_1] = \frac{1}{N^2} M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i} \qquad (3.2.8)$$

$$\widehat{\text{var}}[\tilde{\mu}_1] = \frac{1}{N^2} M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i} \qquad (3.2.9)$$

dimana $s_u^2 = \dfrac{1}{m-1}\sum_{i=1}^{m}\left(N_i\bar{w}_i^{\#} - \sum_{i=1}^{m} N_i\bar{w}_i^{\#} \big/ m\right)^2$, dan $s_i^2 = \dfrac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(w_{ij}^{\#} - \bar{w}_i^{\#}\right)^2$,

untuk *i= 1, 2,…,m.*

Untuk membuktikan bahwa taksiran variansi pada (3.2.9) adalah taksiran tak bias untuk (3.2.8), cara yang sama seperti pada skema *overlapping* dapat digunakan.

Selanjutnya, taksiran Hansen-Hurwitz untuk total adalah

$$\begin{aligned}\tilde{\tau}_1 &= N.\tilde{\mu}_1 \\ &= \frac{M}{m}\sum_{i=1}^{m} N_i\bar{w}_i^{\#}\end{aligned}$$

sedangkan variansi dan taksiran variansi dari taksiran total tersebut adalah

$$\begin{aligned}\text{var}[\tilde{\tau}_1] &= \text{var}[N.\tilde{\mu}_1] \\ &= M(M-m)\frac{\sigma_u^2}{m} + \frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i} \\ \widehat{\text{var}}[\tilde{\tau}_1] &= \widehat{\text{var}}[N.\tilde{\mu}_1] \\ &= M(M-m)\frac{s_u^2}{m} + \frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i}\end{aligned}$$

Dengan cara yang sama seperti pada taksiran mean, taksiran total dan taksiran variansi tersebut adalah taksiran tak bias untuk total dan variansi dari taksiran total.

BAB IV

CONTOH PENERAPAN METODE

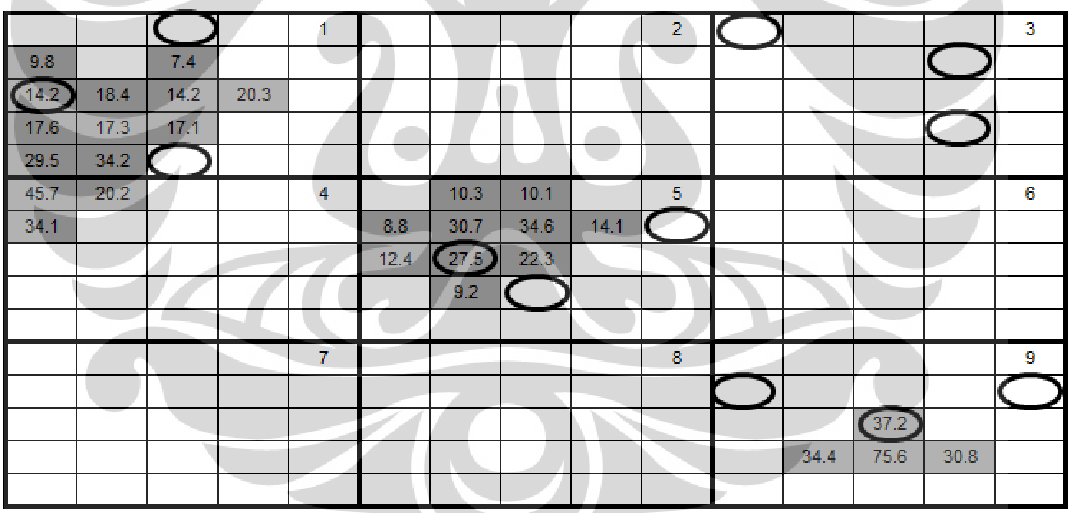
TWO STAGE ADAPTIVE CLUSTER SAMPLING

Pada pembahasan ini, akan dijelaskan suatu contoh penggunaan metode *two stage adaptive cluster sampling* untuk menaksir rata-rata berat jamur (dalam kg) per 100 meter persegi dan total berat jamur disuatu hutan. Jamur tersebut hidup berkelompok dan cukup sulit ditemukan. Misalkan hutan tersebut mempunyai luas 22.500 meter persegi dan dibagi menjadi sembilan unit primer, dimana masing-masing unit primer dibagi menjadi unit sampling sebanyak 25 unit, dengan masing-masing unit sampling seluas 100 meter persegi.

Misalkan pada tahap pertama, dipilih empat unit primer sebagai sampel. Pada masing-masing unit primer, dipilih tiga unit sampling sebagai sampel awal. Dengan demikian, dalam contoh ini diperoleh $M = 9$, $N_i = 25$, $N = 225$, $m = 4$, dan $n_i = 3$. Dalam contoh ini, akan digunakan kondisi $C = \{y_j \mid y_j > 0,\ j = 1,2,\ldots,N\}$. Selanjutnya setiap unit sampling yang terpilih sebagai sampel awal tersebut diperiksa apakah memenuhi kondisi C atau tidak. Kedua skema (skema *overlapping* dan skema *nonoverlapping*) akan diterapkan dalam contoh ini dan juga kedua jenis taksiran (Horvitz-Thompson dan Hansen-Hurwitz) akan digunakan untuk menaksir mean dan total populasi.

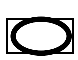
4.1 Skema *Overlapping*

Dengan menggunakan skema *overlapping*, sampel yang diperoleh dapat dilihat pada Gambar (4.1). Pada Gambar (4.1), unit primernya adalah kotak persegi yang dicetak tebal dan masing-masing unit primer diberi nomor pada pojok kanan unit primer. Sementara kota-kotak persegi sebanyak 225 menunjukkan unit sampling dengan angka-angka pada unit sampling menunjukkan total pengukuran pada unit sampling tersebut.



Gambar 4.1. Sampel Dengan Skema *Overlapping*

Keterangan :

 : unit sampling yang terpilih sebagai sampel awal

Berdasarkan Gambar 4.1, terlihat bahwa sampel yang diperoleh sebanyak 12 buah *network*. Dari 12 *network* pada sampel, ada sebuah *network* berukuran 14, satu *network* berukuran 10, satu *network* berukuran 4, dan sembilan *network* berukuran satu. Masing-masing total pengukurannya adalah $y_1^* = 300$, $y_2^* = 180$, $y_3^* = 178$, dan $y_4^* = y_5^* = ... = y_{12}^* = 0$. Kemudian, akan dicari taksiran Horvitz-Thompson dan Hansen-Hurwitz untuk mean dan total.

4.1.1 Taksiran Mean Dan Total Dengan Taksiran Horvitz-Thompson

Telah disebutkan bahwa x_{ik} adalah banyaknya unit sampling pada *network* ke k di dalam unit primer *i*, B_k himpunan unit primer yang beririsan dengan *network* ke k, dan g_k adalah banyaknya unit primer yang beririsan dengan *network* ke k. Dengan demikian,

$$\begin{aligned} &x_{11} = 11,\ x_{41} = 3,\ x_{52} = 10,\ x_{93} = 4 \\ &B_1 = \{1,4\},\ B_2 = \{5\},\ B_3 = \{9\} \\ &g_1 = 2,\ g_2 = 1,\ g_3 = 1 \end{aligned} \tag{4.1.1}$$

Selanjutnya akan dicari π_1, π_2, dan π_3, sebagai berikut :

Telah disebutkan bahwa C_{ik} adalah kejadian bahwa salah satu unit sampling dari *network* ke k yang ada pada unit primer *i*, terpilih sebagai sampel awal. Kemudian, karena keadaan (4.1.1), maka berdasarkan bentuk (3.1.4), nilai π_1, π_2, dan π_3 dapat diperoleh sebagai berikut :

$$
\begin{aligned}
\pi_1 &= \Pr(C_{11} \cup C_{41}) \\
&= \Pr(C_{11}) + \Pr(C_{41}) - \Pr(C_{11} \cap C_{41}) \\
&= \frac{4}{9}.\left[1 - \binom{25-11}{3} \Big/ \binom{25}{3}\right] + \frac{4}{9}.\left[1 - \binom{25-3}{3} \Big/ \binom{25}{3}\right] \\
&\quad - \frac{4.3}{9.8}.\left[1 - \binom{25-11}{3} \Big/ \binom{25}{3}\right].\left[1 - \binom{25-3}{3} \Big/ \binom{25}{3}\right] \\
&= 0.47461 \\
\pi_2 &= \Pr(C_{52}) \\
&= \frac{4}{9}.\left[1 - \binom{25-10}{3} \Big/ \binom{25}{3}\right] \\
&= 0.35652 \\
\pi_3 &= \Pr(C_{92}) \\
&= \frac{4}{9}.\left[1 - \binom{25-4}{3} \Big/ \binom{25}{3}\right] \\
&= 0.18744
\end{aligned}
$$

Dengan demikian, taksiran Horvitz-Thompson untuk mean dan total dapat diperoleh yaitu

$$
\begin{aligned}
\hat{\mu} &= \frac{1}{225} \sum_{k=1}^{3} \frac{y_k^*}{\pi_k} \\
&= \frac{1}{225}\left(\frac{300}{0,47461} + \frac{180}{0,35652} + \frac{178}{0,18744}\right) \\
&= 9,27385 \\
\hat{\tau} &= N.\hat{\mu} = 2086,62
\end{aligned}
$$

Jadi, rata-rata berat jamur adalah 9,27385 kg per 100 meter persegi, sedangkan total berat jamur pada hutan tersebut adalah 2086,62 kg.

Berikutnya akan dicari taksiran variansi dari taksiran mean tersebut. Sebelumnya, akan dicari dahulu nilai π_{12}, π_{13}, dan π_{23} sebagai berikut :

Telah disebutkan bahwa $B_{kk'}$ adalah himpunan unit primer yang beririsan dengan *network* ke k atau *network* ke k', $g_{kk'}$ adalah banyaknya unit primer yang beririsan dengan *network* ke k atau *network* ke k', dan $x_{ikk'} = x_{ik} + x_{ik'}$ adalah banyaknya unit sampling pada unit primer *i* yang beririsan dengan *network* ke k atau *network* ke k'. Dengan demikian,

$$
\begin{aligned}
&B_{12} = \{1,4,5\}, B_{13} = \{1,4,9\}, B_{23} = \{5,9\} \\
&g_{12} = 3, g_{13} = 3, g_{23} = 2 \\
&x_{112} = 11, x_{412} = 3, x_{512} = 10 \\
&x_{113} = 11, x_{413} = 3, x_{913} = 4 \\
&x_{523} = 10, x_{923} = 4
\end{aligned}
\tag{4.1.2}
$$

Jadi, karena kondisi (4.1.2), maka berdasarkan bentuk (3.1.5), nilai π_{12}, π_{13}, dan π_{23} dapat diperoleh sebagai berikut

$$
\begin{aligned}
\pi_{12} = \pi_1 + \pi_2 - &\\
&\left\{ \sum_{i \in B_{12}} \frac{m}{M} \left[1 - \binom{N_i - x_{i12}}{n_i} \bigg/ \binom{N_i}{n_i} \right] \right. \\
&- \sum_i \sum_{i'<i} \frac{m(m-1)}{M(M-1)} \left[1 - \binom{N_i - x_{i12}}{n_i} \bigg/ \binom{N_i}{n_i} \right] \left[1 - \binom{N_i - x_{i'12}}{n_i} \bigg/ \binom{N_i}{n_i} \right] \\
&+ \frac{m(m-1)(m-2)}{M(M-1)(M-2)} \left[1 - \binom{N_1 - x_{112}}{n_1} \bigg/ \binom{N_1}{n_1} \right] \left[1 - \binom{N_4 - x_{412}}{n_4} \bigg/ \binom{N_4}{n_4} \right] \\
&\left. \cdot \left[1 - \binom{N_5 - x_{512}}{n_5} \bigg/ \binom{N_5}{n_5} \right] \right\}
\end{aligned}
$$

$$
\begin{aligned}
\pi_{12} &= 0.47461 + 0.35652 - \\
&\quad \{\tfrac{4}{9}(0.84174 + 0.33043 + 0.802174) - (0.04636 + 0.11254 + 0.044178) + 0.010625\} \\
&= 0.14609
\end{aligned}
$$

$$\pi_{13} = \pi_1 + \pi_3 -$$
$$\left\{ \sum_{i \in B_{13}} \frac{m}{M} \left[1 - \binom{N_i - x_{i13}}{n_i} \Big/ \binom{N_i}{n_i} \right] \right.$$
$$- \sum_{i} \sum_{i' < i} \frac{m(m-1)}{M(M-1)} \left[1 - \binom{N_i - x_{i13}}{n_i} \Big/ \binom{N_i}{n_i} \right] \left[1 - \binom{N_i - x_{i'13}}{n_i} \Big/ \binom{N_i}{n_i} \right]$$
$$+ \frac{m(m-1)(m-2)}{M(M-1)(M-2)} \left[1 - \binom{N_1 - x_{113}}{n_1} \Big/ \binom{N_1}{n_1} \right] \left[1 - \binom{N_4 - x_{413}}{n_4} \Big/ \binom{N_4}{n_4} \right]$$
$$\left. . \left[1 - \binom{N_9 - x_{913}}{n_9} \Big/ \binom{N_9}{n_9} \right] \right\}$$

$$\pi_{13} = 0.47461 + 0.18744 -$$
$$= \{0.37411 + 0.14686 + 0.18744 - (0.04636 + 0.05917 + 0.02323) + 0.00559\}$$
$$= 0.07681$$

$$\pi_{23} = \pi_2 + \pi_3 -$$
$$\left\{ \frac{m}{M} \left[1 - \binom{N_5 - x_{523}}{n_5} \Big/ \binom{N_5}{n_5} \right] + \frac{m}{M} \left[1 - \binom{N_9 - x_{923}}{n_9} \Big/ \binom{N_9}{n_9} \right] \right.$$
$$\left. - \frac{m(m-1)}{M(M-1)} \left[1 - \binom{N_5 - x_{523}}{n_5} \Big/ \binom{N_5}{n_5} \right] \left[1 - \binom{N_9 - x_{923}}{n_9} \Big/ \binom{N_9}{n_9} \right] \right\}$$
$$\pi_{23} = 0.31304 + 0.18744 - \{(0.31304 + 0.18744) - 0.04951\}$$
$$= 0.056385$$

sehingga taksiran variansi dari taksiran mean dan taksiran total adalah

$$\widehat{\text{var}}[\hat{\mu}] = \frac{1}{225^2} \sum_{k=1}^{12} \sum_{k'=1}^{12} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'} \pi_{kk'}} \right) y_k^* y_{k'}^*$$
$$= \frac{1}{225^2} \sum_{k=1}^{3} \sum_{k'=1}^{3} \left(\frac{\pi_{kk'} - \pi_k \pi_{k'}}{\pi_k \pi_{k'} \pi_{kk'}} \right) y_k^* y_{k'}^*$$
$$= \frac{1}{225^2} \{ 209919{,}0131 + 164023{,}7954 + 732778{,}79$$
$$- 101005{,}1609 - 189983{,}4187 - 177574{,}8051 \}$$
$$= 12{,}60559$$
$$\widehat{\text{var}}[\hat{\tau}] = N^2 . \widehat{\text{var}}[\hat{\mu}] = 638158{,}2138$$

4.1.2 Taksiran Mean Dan Total Dengan Taksiran Hansen-Hurwitz

Telah dijelaskan bahwa Y_{ij} adalah total pengukuran pada *network* A_{ij}.

Berdasarkan Gambar 4.1, diperoleh bahwa

$$\begin{array}{llll} Y_{11} = 300, & Y_{31} = 0, & Y_{51} = 180, & Y_{91} = 178, \\ Y_{12} = 0, & Y_{32} = 0, & Y_{52} = 0, & Y_{92} = 0, \\ Y_{13} = 0, & Y_{33} = 0, & Y_{53} = 0, & Y_{93} = 0, \end{array}$$

Untuk mencari taksiran mean dan total dengan taksiran Hansen-Hurwitz, terlebih dahulu akan ditentukan $w_{ij} = Y_{ij} \Big/ \sum_{l=1}^{M} a_{ijl}$ dan $\bar{w}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}$ sebagai berikut :

$$\begin{array}{llll} w_{11} = 21{,}42857 & w_{31} = 0 & w_{51} = 18 & w_{91} = 44{,}5 \\ w_{12} = 0 & w_{32} = 0 & w_{52} = 0 & w_{92} = 0 \\ w_{13} = 0 & w_{33} = 0 & w_{53} = 0 & w_{93} = 0 \\ \bar{w}_1 = 7{,}14286 & \bar{w}_3 = 0 & \bar{w}_5 = 6 & \bar{w}_9 = 14{,}83 \end{array}$$

Dengan demikian, taksiran Hansen-Hurwitz untuk mean dan total dapat diperoleh sebagai berikut

$$\begin{aligned} \tilde{\mu} &= \frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} N_i \bar{w}_i \\ &= \frac{9.25}{225.4}(7{,}14286 + 0 + 6 + 14.83) \\ &= 6{,}99405 \\ \tilde{\tau} &= N.\tilde{\mu} = 1573{,}661 \end{aligned}$$

Jadi, rata-rata berat jamur adalah 6,99405 kg per 100 meter persegi, sedangkan total berat jamur pada hutan tersebut adalah 1573,661 kg.

Selanjutnya akan dicari taksiran variansi dari taksiran mean dan total tersebut. Sebelumnya akan dicari dahulu $s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(w_{ij} - \bar{w}_i\right)^2$ i = 1, 3, 5, 9 dan $s_u^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(N_i\bar{w}_i - \sum_{i=1}^{m} N_i\bar{w}_i / m\right)^2$ sebagai berikut :

$$
\begin{aligned}
s_1^2 &= \frac{1}{2}\left\{(21,42857 - 7,1428571)^2 + (0 - 7,1428571)^2 + (0 - 7,1428571)^2\right\} \\
&= 153,06122 \\
s_3^2 &= 0 \\
s_5^2 &= \frac{1}{2}\left\{(18-6)^2 + (0-6)^2 + (0-6)^2\right\} \\
&= 108 \\
s_9^2 &= \frac{1}{2}\left\{(44,5 - 14,83)^2 + (0 - 14,83)^2 + (0 - 14,83)^2\right\} \\
&= 660,083
\end{aligned}
$$

$$
\frac{1}{m}\sum_{i=1}^{m} N_i\bar{w}_i = \frac{25}{4}(7,14286 + 0 + 6 + 14,83) = 174,8512
$$

$$
\begin{aligned}
s_u^2 &= \frac{1}{m-1}\sum_{i=1}^{m}\left(N_i\bar{w}_i - \sum_{i=1}^{m} N_i\bar{w}_i / m\right)^2 \\
&= \frac{1}{3}\Big\{\left[(25)(7,14286) - 174,8512\right]^2 + \left[(25)(0) - 174,8512\right]^2 \\
&\qquad + \left[(25)(6) - 174,8512\right]^2 + \left[(25)(14,83) - 174,8512\right]^2\Big\} \\
&= 23204,4537
\end{aligned}
$$

Jadi, taksiran variansi dari taksiran mean dan taksiran total dapat dicari dengan menggunakan bentuk (3.1.12), yaitu

$$\widetilde{\text{var}}[\tilde{\mu}] = \frac{1}{N^2} M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i}$$

$$= \frac{1}{225^2}\left\{9.5.\frac{23204,4537}{4} + \frac{9.25.22}{4.3}\left(153,06122 + 0 + 108 + 660,083\right)\right\}$$

$$= 12,66168$$

$$\widetilde{\text{var}}[\tilde{\tau}] = N^2.\widetilde{\text{var}}[\tilde{\mu}] = 641022,2337$$

Dengan menggunakan skema *overlapping*, hasil taksiran yang diperoleh adalah sebagai berikut :

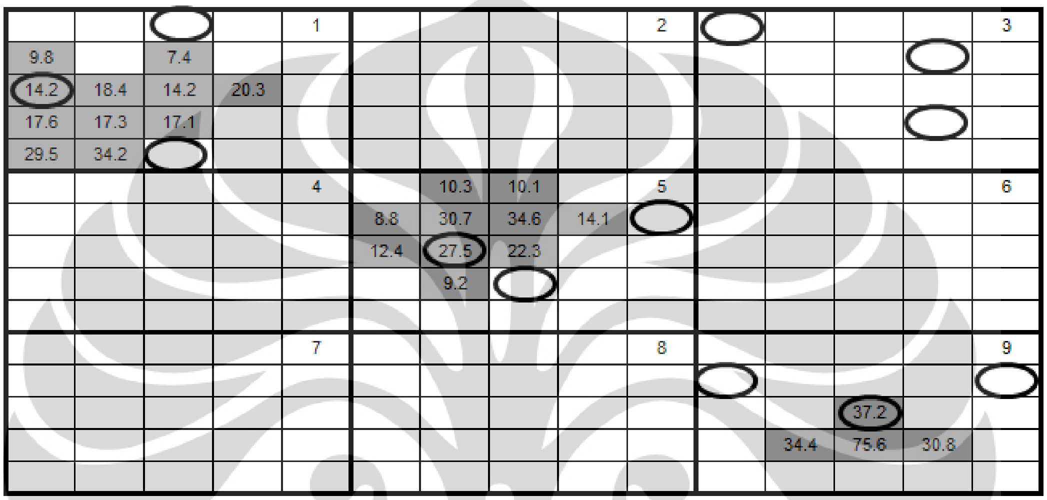
Tabel 4.1. Hasil Taksiran Pada Skema *Overlapping*

Taksiran Horvitz-Thompson	**Taksiran Hansen-Hurwitz**
$\hat{\mu} = 9,27385$	$\tilde{\mu} = 6.99405$
$\widehat{\text{var}}[\hat{\mu}] = 12,60559$	$\widetilde{\text{var}}[\tilde{\mu}] = 12,66217$
$\hat{\tau} = 2086,616$	$\tilde{\tau} = 1573,661$
$\widehat{\text{var}}[\hat{\tau}] = 638158,2138$	$\widetilde{\text{var}}[\tilde{\tau}] = 641022,2337$

Dari contoh terlihat bahwa untuk skema *overlapping*, ternyata taksiran Horvitz-Thompson sedikit lebih efisien dibandingkan taksiran Hansen-Hurwitz.

4.2 Skema *Nonoverlapping*

Apabila skema *nonoverlapping* digunakan dalam penelitian ini, maka sampel yang diperoleh adalah sebagai berikut



Gambar 4.2. Sampel Dengan Skema *Nonoverlapping*

Jumlah *network* yang diperoleh pada skema ini adalah sama dengan jumlah *network* yang diperoleh pada skema *overlapping*, namun pada skema ini ukuran *network* pertama adalah 11 sedangkan pada skema *overlapping* ukuran *network* pertamanya adalah 14. Misalkan y_{ik}^{*} adalah total pengukuran dari *network* ke k yang ada pada unit primer *i*. Kemudian, total pengukuran pada masing-masing *network* adalah $y_{11}^{*} = 200$, $y_{52}^{*} = 180$, $y_{93}^{*} = 178$, sedangkan total pengukuran dari sembilan *network* lainnya adalah nol.

4.2.1 Taksiran Mean Dan Total Dengan Taksiran Horvitz-Thompson

Untuk menentukan taksiran mean dan total dengan taksiran Horvitz_Thompson, terlebih dahulu akan dicari nilai dari π_{11}, π_{52}, dan π_{93} dengan menggunakan persamaan (3.2.2), yaitu

$$\pi_{11} = 1 - \binom{25-11}{3} \Big/ \binom{25}{3} = 0{,}84174$$

$$\pi_{52} = 1 - \binom{25-8}{3} \Big/ \binom{25}{3} = 0{,}80217$$

$$\pi_{93} = 1 - \binom{25-4}{3} \Big/ \binom{25}{3} = 0{,}42174$$

Dengan demikian, taksiran Horvitz-Thompson untuk mean dan total adalah

$$\begin{aligned}
\hat{\mu}_1 &= \frac{M}{Nm} \sum_{i=1}^{m} \sum_{k=1}^{\kappa_i} \frac{y_{ik}^*}{\pi_{ik}} \\
&= \frac{9}{225.4} \left\{ \frac{200}{0{,}84174} + \frac{180}{0{,}80217} + \frac{178}{0{,}42174} \right\} \\
&= 8{,}84055
\end{aligned}$$

$$\hat{\tau}_1 = N.\hat{\mu}_1 = 1989{,}125$$

Jadi, rata-rata berat jamur adalah 8,84 kg per 100 meter persegi, sedangkan total berat jamur pada hutan tersebut adalah 1989,125 kg.

Selanjutnya untuk mencari taksiran variansi dari taksiran mean dan taksiran total yang ada pada persamaan (3.2.3), terlebih dahulu akan dicari nilai $\hat{V}_1$, $\hat{V}_3$, $\hat{V}_5$, $\hat{V}_9$ dan $s_u^2 = \sum_{i=1}^{m} \frac{(\hat{\tau}_i - \sum_{i=1}^{m} \hat{\tau}_i / m)^2}{m-1}$ sebagai berikut :

$$\hat{V}_1 = \sum_{k=1}^{3}\sum_{k'=1}^{3}\left(\frac{\pi_{1kk'} - \pi_{1k}\pi_{1k'}}{\pi_{1kk'}\pi_{1k}\pi_{1k'}}\right)y_{1k}^{*}y_{1k'}^{*}$$
$$= \left(\frac{\pi_{111} - \pi_{11}\pi_{11}}{\pi_{111}\pi_{11}\pi_{11}}\right)\left(y_{11}^{*}\right)^2$$
$$= \left(\frac{1-\pi_{11}}{\pi_{11}^2}\right)\left(y_{11}^{*}\right)^2$$
$$= 8934,66976$$

$$\hat{V}_3 = 0$$

$$\hat{V}_5 = \sum_{k=1}^{3}\sum_{k'=1}^{3}\left(\frac{\pi_{5kk'} - \pi_{5k}\pi_{5k'}}{\pi_{5kk'}\pi_{5k}\pi_{5k'}}\right)y_{5k}^{*}y_{5k'}^{*}$$
$$= \left(\frac{\pi_{522} - \pi_{52}\pi_{52}}{\pi_{522}\pi_{52}\pi_{52}}\right)\left(y_{52}^{*}\right)^2$$
$$= \left(\frac{1-\pi_{52}}{\pi_{52}^2}\right)\left(y_{52}^{*}\right)^2$$
$$= 9960,73766$$

$$\hat{V}_9 = \sum_{k=1}^{3}\sum_{k'=1}^{3}\left(\frac{\pi_{9kk'} - \pi_{9k}\pi_{9k'}}{\pi_{9kk'}\pi_{9k}\pi_{9k'}}\right)y_{9k}^{*}y_{9k'}^{*}$$
$$= \left(\frac{\pi_{922} - \pi_{92}\pi_{92}}{\pi_{522}\pi_{52}\pi_{52}}\right)\left(y_{92}^{*}\right)^2$$
$$= \left(\frac{1-\pi_{92}}{\pi_{92}^2}\right)\left(y_{92}^{*}\right)^2$$
$$= 103009,2$$

$$\sum_{i=1}^{m}\hat{\tau}_i \Big/ m = \frac{1}{4}(237,60331 + 0 + 224,39024 + 422,06186)$$
$$= 221,0139$$

$$s_u^2 = \sum_{i=1}^{m}\frac{(\hat{\tau}_i - \sum_{i=1}^{m}\hat{\tau}_i / m)^2}{m-1}$$
$$= \frac{1}{3}\Big\{(237,60331 - 221,0139)^2 + (0 - 221,0139)^2 + (224,39024 - 221,0139)^2$$
$$+ (422,06186 - 221,0139)^2\Big\}$$
$$= 29851,34$$

Dengan demikian, taksiran variansi dari taksiran mean dan total dapat diperoleh sebagai berikut

$$\widehat{\text{var}}[\hat{\mu}_1] = \frac{1}{N^2} M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}\widehat{V}_i$$
$$= \frac{1}{225^2}\left\{9.5.\frac{29851,34}{4} + \frac{9}{4}.\left(8934,66976 + 0 + 9960,73766 + 103009,2\right)\right\}$$
$$= 12,05161$$
$$\widehat{\text{var}}[\hat{\tau}_1] = N^2.\widehat{\text{var}}[\hat{\mu}_1] = 610112,9881$$

4.2.2 Taksiran Mean Dan Total Dengan Taksiran Hansen-Hurwitz

Dengan menggunakan skema *nonoverlapping*, diperoleh total pengukuran *network* sebagai berikut:

$$\begin{array}{llll} Y_{11} = 200, & Y_{31} = 0, & Y_{51} = 180, & Y_{91} = 178, \\ Y_{12} = 0, & Y_{32} = 0, & Y_{52} = 0, & Y_{92} = 0, \\ Y_{13} = 0, & Y_{33} = 0, & Y_{53} = 0, & Y_{93} = 0, \end{array}$$

Kemudian akan dicari $w_{ij}^{\#} = Y_{ij}/a_{ij}$ dan $\bar{w}_i^{\#} = \frac{1}{n_i}\sum_{j=1}^{n_i} w_{ij}^{\#}$, sebagai berikut :

$$\begin{array}{llll} w_{11}^{\#} = 18,18182 & w_{31}^{\#} = 0 & w_{51}^{\#} = 18 & w_{91}^{\#} = 44,5 \\ w_{12}^{\#} = 0 & w_{32}^{\#} = 0 & w_{52}^{\#} = 0 & w_{92}^{\#} = 0 \\ w_{13}^{\#} = 0 & w_{33}^{\#} = 0 & w_{53}^{\#} = 0 & w_{93}^{\#} = 0 \\ \bar{w}_1^{\#} = 6,06061 & \bar{w}_3^{\#} = 0 & \bar{w}_5^{\#} = 6 & \bar{w}_9^{\#} = 14,83 \end{array}$$

Maka taksiran Hansen-Hurwitz untuk mean dan total dapat dicari sebagai berikut

$$\tilde{\mu}_1 = \frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} N_i \bar{w}_i^{\#}$$
$$= \frac{9.25}{225.4}.\left(6,06061 + 0 + 6 + 14,83\right)$$
$$= 6,72348$$
$$\tilde{\tau}_1 = N.\tilde{\mu}_1 = 1512,784$$

Jadi, rata-rata berat jamur adalah 6,72348 kg per 100 meter persegi, sedangkan total berat jamur pada hutan tersebut adalah 1512,784 kg.

Selanjutnya akan dicari taksiran variansi dari taksiran mean dan total tersebut. Sebelumnya akan dicari dahulu variansi antar unit sampling pada unit primer ke 1, 3, 5, dan 9 dengan menggunakan bentuk

$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(w_{ij}^{\#} - \bar{w}_i^{\#}\right)^2$, i = 1, 3, 5, 9, dan $s_u^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(N_i \bar{w}_{ij}^{\#} - \sum_{i=1}^{m} N_i \bar{w}_{ij}^{\#} \big/ m\right)^2$

sebagai berikut :

$$s_1^2 = \frac{1}{2}\left\{(18,18182 - 6,06061)^2 + (0 - 6,06061)^2 + (0 - 6,06061)^2\right\}$$
$$= 110,19284$$
$$s_3^2 = 0$$

$$s_5^2 = \frac{1}{2}\left\{(18 - 6)^2 + (0 - 6)^2 + (0 - 6)^2\right\}$$
$$= 108$$
$$s_9^2 = \frac{1}{2}\left\{(44,5 - 14,83)^2 + (0 - 14,83)^2 + (0 - 14,83)^2\right\}$$
$$= 660,083$$

$$\frac{1}{m}\sum_{i=1}^{m} N_i \bar{w}_i^{\#} = \frac{25}{4}(6,06061 + 0 + 6 + 14,83) = 168,0871$$

$$s_u^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(N_i\bar{w}_{ij}^{\#} - \sum\nolimits_{i=1}^{m} N_i\bar{w}_{ij}^{\#} \big/ m\right)^2$$
$$= \frac{1}{3}\Big\{\left[(25)(6,06061) - 168,0871\right]^2 + \left[(25)(0) - 168,0871\right]^2$$
$$+\left[(25)(6) - 168,0871\right]^2 + \left[(25)(14,83) - 168,0871\right]^2\Big\}$$
$$= 23320,3603$$

Jadi, taksiran variansi dari taksiran mean dan taksiran total dapt ditentukan dengan menggunakan bentuk (3.1.12), yaitu

$$\widetilde{\text{var}}[\tilde{\mu}_1] = \frac{1}{N^2}M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}N_i(N_i - n_i)\frac{s_i^2}{n_i}$$
$$= \frac{1}{225^2}\left\{9.5.\frac{23320,3603}{4} + \frac{9.25.22}{4.3}\left(110,19284 + 0 + 108 + 660,083\right)\right\}$$
$$= 12,33863$$
$$\widetilde{\text{var}}[\tilde{\tau}_1] = N^2.\widetilde{\text{var}}[\tilde{\mu}_1] = 624642,9741$$

Dengan menggunakan skema *nonoverlapping*, hasil taksiran yang diperoleh sebagai berikut:

Tabel 4.2. Hasil Taksiran Pada Skema *Nonoverlapping*

Taksiran Horvitz-Thompson	Taksiran Hansen-Hurwitz
$\hat{\mu}_1 = 8,84055$	$\tilde{\mu}_1 = 6,72348$
$\widehat{\text{var}}[\hat{\mu}_1] = 12,05161$	$\widetilde{\text{var}}[\tilde{\mu}_1] = 12,33863$
$\hat{\tau}_1 = 1989,125$	$\tilde{\tau}_1 = 1512,784$
$\widehat{\text{var}}[\hat{\tau}_1] = 610112,9881$	$\widetilde{\text{var}}[\tilde{\tau}_1] = 624642,9741$

Dengan skema *nonoverlapping*, taksiran Horvitz-Thompson lebih baik dibandingkan taksiran Hansen-Hurwitz.

Berdasarkan contoh tersebut dapat diambil kesimpulan bahwa untuk kedua skema, taksiran Horvitz-Thompson lebih efisien daripada taksiran Hansen-Hurwitz. Kemudian, taksiran Horvitz-Thompson dan taksiran Hansen-Hurwitz pada skema *Nonverlapping* menghasilkan taksiran variansi yang lebih kecil daripada skema *Overlapping,* sehingga pada contoh ini skema *Nonverlapping* lebih efisien daripada skema *Overlapping.*

Walaupun demikian, hal ini tidak selalu terjadi, karena secara umum antara kedua jenis taksiran tersebut tidak ada yang lebih efisien secara seragam. Begitupun mengenai skema yang digunakan, skema mana yang lebih baik tidak dapat dibandingkan, karena pemilihan skema ditentukan oleh keadaan lapangan. Akan tetapi, berdasarkan berbagai macam penelitian yang telah dilakukan dan juga berdasarkan literatur-literatur yang telah ada (Thompson and Seber, 1996; Roesch, 1993; M. Salehi, 2003), menyarankan bahwa taksiran Horvitz-Thompson dengan skema *Nonoverlapping* cenderung lebih efisien untuk berbagai macam situasi.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penjelasan pada BAB III, dapat diperoleh kesimpulan sebagai berikut :

1. Pada metode *two stage adaptive cluster sampling*, ada dua skema yang dapat digunakan yaitu skema *overlapping* dan skema *nonoverlapping*. Pada skema *overlapping*, proses penambahan sampel diperbolehkan melewati batas unit primer, sedangkan skema *nonoverlapping*, proses penambahan sampel tidak diperbolehkan melewati batas unit primer
2. Berikut adalah taksiran tak bias untuk mean dan total populasi, dengan variansi dan taksiran variansinya yang tak bias pula :
   - Skema *Overlapping*, dengan taksiran Horvitz-Thompson

$$\hat{\mu} = \frac{1}{N}\sum_{k=1}^{\kappa}\frac{y_k^*}{\pi_k} \qquad \hat{\tau} = \sum_{k=1}^{\kappa}\frac{y_k^*}{\pi_k}$$

$$\text{var}[\hat{\mu}] = \frac{1}{N^2}\sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}}\right)\tau_k^*\tau_{k'}^* \qquad \text{var}[\hat{\tau}] = \sum_{k=1}^{K}\sum_{k'=1}^{K}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}}\right)\tau_k^*\tau_{k'}^*$$

$$\widehat{\text{var}}[\hat{\mu}] = \frac{1}{N^2}\sum_{k=1}^{\kappa}\sum_{k'=1}^{\kappa}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}\pi_{kk'}}\right)y_k^*y_{k'}^* \qquad \widehat{\text{var}}[\hat{\tau}] = \sum_{k=1}^{\kappa}\sum_{k'=1}^{\kappa}\left(\frac{\pi_{kk'} - \pi_k\pi_{k'}}{\pi_k\pi_{k'}\pi_{kk'}}\right)y_k^*y_{k'}^*$$

- Skema *Overlapping*, dengan taksiran Hansen-Hurwitz :

$$\tilde{\mu} = \frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} N_i \bar{w}_i \qquad \tilde{\tau} = \frac{M}{m}\sum_{i=1}^{m}\hat{\tau}_i$$

$$\text{var}[\tilde{\mu}] = \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$\widetilde{\text{var}}[\tilde{\mu}] = \frac{1}{N^2}M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i}$$

$$\text{var}[\tilde{\tau}] = M(M-m)\frac{\sigma_u^2}{m} + \frac{M}{m}\sum_{i=1}^{M} N_i(N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$\widetilde{\text{var}}[\tilde{\tau}] = M(M-m)\frac{s_u^2}{m} + \frac{M}{m}\sum_{i=1}^{m} N_i(N_i - n_i)\frac{s_i^2}{n_i}$$

- Skema *Nonoverlapping*, dengan taksiran Horvitz-Thompson :

$$\hat{\mu}_1 = \frac{M}{Nm}\sum_{i=1}^{m}\sum_{k=1}^{\kappa_i}\frac{y_{ik}^*}{\pi_{ik}} = \frac{M}{Nm}\sum_{i=1}^{m}\hat{\tau}_i \qquad \hat{\tau}_1 = \frac{M}{m}\sum_{i=1}^{m}\sum_{k=1}^{\kappa_i}\frac{y_{ik}^*}{\pi_{ik}} = \frac{M}{m}\sum_{i=1}^{m}\hat{\tau}_i$$

$$\text{var}[\hat{\mu}_1] = \frac{1}{N^2}M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^*\tau_{ik'}^*\right)$$

$$\widehat{\text{var}}[\hat{\mu}_1] = \frac{1}{N^2}M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}\pi_{ik}\pi_{ik'}}\right)y_{ik}^*y_{ik'}^*\right)$$

$$\text{var}[\hat{\tau}_1] = M(M-m)\frac{\sigma_u^2}{m} + \frac{M}{m}\sum_{i=1}^{M}\left(\sum_{k=1}^{K_i}\sum_{k'=1}^{K_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ik}\pi_{ik'}}\right)\tau_{ik}^*\tau_{ik'}^*\right)$$

$$\widehat{\text{var}}[\hat{\tau}_1] = M(M-m)\frac{s_u^2}{m} + \frac{M}{m}\sum_{i=1}^{m}\left(\sum_{k=1}^{\kappa_i}\sum_{k'=1}^{\kappa_i}\left(\frac{\pi_{ikk'} - \pi_{ik}\pi_{ik'}}{\pi_{ikk'}\pi_{ik}\pi_{ik'}}\right)y_{ik}^*y_{ik'}^*\right)$$

- Skema *Nonoverlapping*, dengan taksiran Hansen-Hurwitz :

$$\tilde{\mu}_1 = \frac{1}{N}\frac{M}{m}\sum_{i=1}^{m} N_i \bar{w}_i^{\#} \qquad \tilde{\tau}_1 = \frac{M}{m}\sum_{i=1}^{m} N_i \bar{w}_i^{\#}$$

$$\text{var}[\tilde{\mu}_1] = \frac{1}{N^2} M(M-m)\frac{\sigma_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{M} N_i (N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$\widetilde{\text{var}}[\tilde{\mu}_1] = \frac{1}{N^2} M(M-m)\frac{s_u^2}{m} + \frac{1}{N^2}\frac{M}{m}\sum_{i=1}^{m} N_i (N_i - n_i)\frac{s_i^2}{n_i}$$

$$\text{var}[\tilde{\tau}_1] = M(M-m)\frac{\sigma_u^2}{m} + \frac{M}{m}\sum_{i=1}^{M} N_i (N_i - n_i)\frac{\sigma_i^2}{n_i}$$

$$\widetilde{\text{var}}[\tilde{\tau}_1] = M(M-m)\frac{s_u^2}{m} + \frac{M}{m}\sum_{i=1}^{m} N_i (N_i - n_i)\frac{s_i^2}{n_i}$$

3. Dalam penelitian dimana elemen yang diteliti bersifat sangat jarang atau saling berkelompok, 2S-ACS lebih baik untuk digunakan dibandingkan dengan metode konvensional dan *adaptive cluster sampling*, karena :
   - Data yang diperoleh pada 2S-ACS akan lebih memadai dibandingkan metode konvensional, karena pada 2S-ACS dilakukan proses penambahan sampel dari unit-unit tetangga sampel awal yang memenuhi suatu kondisi tertentu.
   - 2S-ACS lebih efektif dalam hal meneliti unit sampling yang terpilih sebagai sampel awal, dibandingkan *adaptive cluster sampling*. Hal ini disebabkan karena pada 2S-ACS, pemilihan sampel awal hanya

dilakukan pada unit sampling yang berada di dalam unit primer yang terpilih pada tahap pertama, sehingga sampel awal tidak terlalu menyebar dan peneliti lebih mudah untuk meneliti sampel awal tersebut. Dengan demikian, meskipun wilayah penelitian sangat luas, tetapi usaha/tenaga yang dibutuhkan akan lebih sedikit dibandingkan metode *adaptive cluster sampling*.

5.2 Saran

1. Pembahasan pada tugas akhir ini dapat dilanjutkan dengan membahas taksiran Rao-Blackwell untuk mean dan total pada metode *Adaptive cluster sampling* dan *Two stage adaptive cluster sampling*. Dengan menggunakan taksiran Rao-Blackwell, akan dihasilkan variansi yang lebih kecil dibandingkan dengan taksiran Horvitz-Thompson dan taksiran Hansen-Hurwitz.

2. Ada baiknya dipelajari mengenai perbandingan antara taksiran Hansen-Hurwitz dan taksiran Horvitz-Thompson untuk *adaptive cluster sampling.*

DAFTAR PUSTAKA

Agustianto, Ikrar Aulia. (2001). *Menaksir Parameter Populasi dengan Metode Adaptive Cluster Sampling*. Departemen Matematika FMIPA-UI, Depok.

Cochran, W. G. (1997). *Sampling Techniques* (Ed. Ke-3). New York: Wiley.

Hogg, R. V. dan Craig, A. T. (1995). *Introduction to Mathematical Statistics* (Ed. Ke-5). New Jersey: Prentice Hall Inc.

Mendenhall, W., Ott, L., dan Scheaffer, R. L. (1996). *Elementary Survey Sampling* (Ed. Ke-5). Boston: Wadsworth.

Raj, Des. (1968). *Sampling Theory*. New York: McGraw-Hill.

Salehi, M. M. dan Seber, G. F. A. (1997). Two Stage Adaptive Cluster Sampling. *Biometrics*. Vol. 53, 959-970.

Thompson, S. K. (1990). Adaptive Cluster Sampling. *Journal of The American Statistical Association.* Vol. 85, 1050-1059.

Thompson, S. K. (2002). *Sampling* (Ed. Ke-2). New York: Wiley.

Thompson, S. K. and Seber, G. A. F. (1996). *Adaptive Sampling*. New York: Wiley.

Yamane, Taro. (1967). *Elementary Sampling Theory.* New York: Prentice Hall.

LAMPIRAN 1

Double Ekspektasi

Jika X_1 dan X_2 adalah variabel random (bertipe diskrit/kontinu), maka dapat ditunjukkan bahwa :

$$E[X_2] = E[E[X_2 \mid X_1]]$$

Bukti :

Misalkan X_1 dan X_2 adalah variabel random bertipe diskrit yang memiliki pdf bersama $f(x_1, x_2)$, pdf marginal dari X_1 adalah $f_1(x_1) = \sum_{x_1} f(x_1, x_2)$, dan pdf marginal dari X_2 adalah $f_2(x_2) = \sum_{x_2} f(x_1, x_2)$, maka

$$\begin{aligned}
E[X_2] &= \sum_{x_1} \sum_{x_2} x_2 f(x_1, x_2) \\
&= \sum_{x_1} \left[\sum_{x_2} x_2 \frac{f(x_1, x_2)}{f_1(x_1)} \right] f_1(x_1) \\
&= \sum_{x_1} \left[\sum_{x_2} x_2 f(x_1 \mid x_2) \right] f_1(x_1) \\
&= \sum_{x_1} E[X_2 \mid x_1] f_1(x_1) \\
&= E[E[X_2 \mid X_1]]
\end{aligned}$$

Jadi, terbukti bahwa $E[X_2] = E[E[X_2 \mid X_1]]$. Hal yang sama dapat dilakukan untuk variabel random jenis kontinu.

LAMPIRAN 2

Dekomposisi Variansi

Jika X_1 dan X_2 adalah variabel random (bertipe diskrit/kontinu), maka dapat ditunjukkan bahwa :

$$\text{var}[X_2] = E\left(\text{var}[X_2 \mid X_1]\right) + \text{var}\left(E[X_2 \mid X_1]\right)$$

Bukti :

$$\text{var}[X_2] = E[X_2^2] - \{E[X_2]\}^2$$

Berdasarkan pembuktian pada lampiran 1, diperoleh :

$$\begin{aligned}
\text{var}[X_2] &= E[E[X_2^2 \mid X_1]] - \left\{E[E[X_2 \mid X_1]]\right\}^2 \\
&= E[E[X_2^2 \mid X_1]] + E\left[\left(E[X_2 \mid X_1]\right)^2\right] - E\left[\left(E[X_2 \mid X_1]\right)^2\right] - \left\{E[E[X_2 \mid X_1]]\right\}^2 \\
&= \left\{E[E[X_2^2 \mid X_1]] - E\left[\left(E[X_2 \mid X_1]\right)^2\right]\right\} + \left\{E\left[\left(E[X_2 \mid X_1]\right)^2\right] - \left\{E[E[X_2 \mid X_1]]\right\}^2\right\} \\
&= E\left\{E[X_2^2 \mid X_1] - \left(E[X_2 \mid X_1]\right)^2\right\} + \left\{E\left[\left(E[X_2 \mid X_1]\right)^2\right] - \left\{E[E[X_2 \mid X_1]]\right\}^2\right\} \\
&= E\left(\text{var}[X_2 \mid X_1]\right) + \text{var}\left(E[X_2 \mid X_1]\right)
\end{aligned}$$

Jadi, terbukti bahwa jika X_1 dan X_2 adalah variabel random, maka

$\text{var}[X_2] = E\left(\text{var}[X_2 \mid X_1]\right) + \text{var}\left(E[X_2 \mid X_1]\right)$.