UNIVERSITAS INDONESIA

TAKSIRAN REGRESI UNTUK *MEAN* POPULASI PADA *STRATIFIED RANDOM SAMPLING*

SKRIPSI

SPINA ELECTRA SUSILA

0304010579

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

DEPARTEMEN MATEMATIKA

DEPOK

JUNI 2010

UNIVERSITAS INDONESIA

TAKSIRAN REGRESI UNTUK *MEAN* POPULASI PADA *STRATIFIED RANDOM SAMPLING*

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana

SPINA ELECTRA SUSILA

0304010579

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

DEPARTEMEN MATEMATIKA

DEPOK

JUNI 2010

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri,

dan semua sumber baik yang dikutip maupun dirujuk

telah saya nyatakan dengan benar.

Nama	**: Spina Electra Susila**
NPM	**: 0304010579**
Tanda Tangan	**:**
Tanggal	**: Juni 2010**

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama	: Spina Electra Susila
NPM	: 0304010579
Program Studi	: Matematika
Judul Skripsi	: Taksiran Regresi untuk *Mean* Populasi pada *Stratified Random Sampling*

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia.

DEWAN PENGUJI

Pembimbing	: Dra. Rianti Setiadi M.S.	(..)
Pembimbing	: Fevi Novkaniza S.Si., M.Si.	(..)
Penguji	: Dra. Siti Aminah M.Kom	(..)
Penguji	: Dr. Dian Lestari DEA	(..)

Ditetapkan di : Depok

Tanggal : 2 Juli 2010

KATA PENGANTAR

Assalamualaikum Wr.Wb.

Puji dan syukur kehadirat Allah SWT atas segala anugerah dan karuniaNya yang terus diberikan kepada penulis sehingga penulis dapat menyelesaikan tugas akhir ini sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dari Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia. Shalawat dan salam kepada suri tauladan seluruh umat manusia, Nabi Muhammad SAW, serta para pengikutnya yang senantiasa istiqomah hingga akhir zaman.

Penulis menyadari bahwa masih terdapat kekurangan dalam tugas akhir ini. Penulis mengharapkan saran dan kritik yang membangun dari para pembaca untuk menyempurnakan tugas akhir ini.

Dengan penuh rasa syukur, penulis ingin mengucapkan terima kasih kepada semua pihak yang telah memberikan bantuan dalam pembuatan tugas akhir ini, terutama kepada :

1. Ibunda tercinta, Hj. Sri Rumiyati dan ayahanda tercinta H. Susila Zaenal Muhammad, MT yang telah mendidik dan menyayangiku dengan sepenuh hati dan kesabaran, mengajarkanku tentang hidup, selalu mendoakanku, dan memberikan segalanya untuk kemajuan anak – anaknya. Terima kasih atas segala kasih sayang dan cintanya selama ini.
2. Ibu Dra. Rianti Setiadi, M.S dan bu Fevi Novkaniza S.Si.,M.Si sebagai pembimbing skripsi yang telah memotivasi penulis, menyediakan waktu dan perhatiannya serta memberikan banyak ilmu selama penulis menyelesaikan tugas akhir ini.
3. Ibu Dra. Denny Riama Silaban, M.Kom, Ibu Yekti Widyaningsih, dan ibu Siti Aminah, M.Kom selaku Pembimbing Akademik angkatan 2004.
4. Adik penulis, Muhammad Faisal Susila yang selalu menyayangi dan menemani hari – hari penulis. Adikku semangat kuliahnya walaupun di Yogyakarta.

5. Seluruh dosen – dosen terutama bu Sarini, bu Nur, dan bu Saskya yang telah menjadi penguji kolokium skripsi saya dan karyawan Departemen Matematika FMIPA UI yang tidak dapat disebutkan namanya satu per satu.
6. Novi Andra S.Si yang selalu menyayangi, menemani, dan menjadi inspirasi serta tempat mengeluarkan keluh kesah penulis baik saat sedih maupun gembira. Terima kasih atas semua yang telah kita jalani bersama – sama.
7. Sahabat – sahabat penulis di kampus, yaitu Leli Zuita yang selalu mau mendengarkan curahan hati penulis dan Febrini Cesarina yang selalu memberikan tumpangan kendaraan saat kuliah. Terima kasih atas bantuannya.
8. Teman – teman seperjuangan skripsi khususnya satu bimbingan bu Rianti atau mbak Fevi atau keduanya.
9. Keluarga BK (Bembi, Tyas, Dewi, Nana,TB, Yanthie), khususnya Bembi. Terima kasih karena telah menjadi PO Logika+ di tahun 2005 sehingga dapat berkenalan dengan seseorang yang sangat penulis sayangi dan berarti di matematika angkatan 2003 hingga sekarang.
10. Teman – teman penulis, matematika angkatan 2004. Terima kasih atas kebersamaan yang sangat berarti bagi penulis semasa kuliah, khususnya melalui acara Daroepat.
11. Teman – teman yang selalu berada di labkom skripsi gedung A. Terima kasih atas hiburannya.
12. Teman – teman angkatan 2001,2002,2003, 2005, 2006, dan 2007.
13. Keluarga besar penulis : kakek, nenek, mbah buyut, pa'de, bu'de, om, bule', kakak – kakak dan adik – adik sepupu,saudara – saudara penulis di Magelang, Yogyakarta, Klaten, Jambi, Palembang dan Jakarta. Terima kasih atas doanya.
14. Keluarga besar SMPN 3 Depok (Bento) dan SMAN 3 Depok (Smuntie) yang selalu penulis rindukan.
15. SMPN 6 Rumbai, Riau. Terima kasih walaupun hanya 1 (satu) cawu penulis dapat menikmati indahnya bersekolah disana.
16. TK Suci Cijantung, SDN Mekarjaya XII Depok dan SDN 005 Rumbai.
17. Serta seluruh pihak yang tidak dapat disebutkan namanya satu per satu yang telah sangat membantu penulis selama ini.

Penulis mohon maaf atas segala kesalahan dan kekurangan dalam penulisan tugas akhir ini. Akhir kata, penulis berharap bahwa semoga tugas akhir ini dapat bermanfaat.

Wassalam.

Depok, 07 Juni 2010

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan dibawah ini.

Nama : Spina Electra Susila
NPM : 0304010579
Program Studi : Sarjana (S1)
Departemen : Matematika
Fakultas : Matematika dan Ilmu Pengetahuan Alam
Jenis karya : Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia **Hak Bebas Royalti Noneksklusif (** ***Non – exclusive Royalty – Free Right*** **)** atas karya ilmiah saya yang berjudul :
Taksiran Regresi untuk *Mean* Populasi pada *Stratified Random Sampling*
beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian penyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok
Pada tanggal : 7 Juni 2010
Yang menyatakan

(Spina Electra Susila)

ABSTRAK

Nama : Spina Electra Susila
Program Studi : S1 Matematika
Judul : Taksiran Regresi untuk Mean Populasi pada Stratified Random Sampling

Tulisan ini membahas suatu metode penaksiran mean dengan suatu metode yang disebut taksiran regresi pada *stratified random sampling*. Taksiran regresi ini digunakan untuk menaksir mean suatu variabel dengan melibatkan variabel lain yang berkorelasi kuat dan mempunyai hubungan linier dengan variabel terkait berdasarkan *simple random sample*. Taksiran regresi ini terbagi menjadi 2 (dua) macam, yaitu taksiran regresi terpisah dan taksiran regresi gabungan. Selain itu, dalam tulisan ini akan diberikan contoh penggunaannya dengan menggunakan data fiktif.

Kata kunci : taksiran regresi, *stratified random sampling*, *simple random sample*
xiii + 92 halaman; 18 tabel; 4 lampiran
Daftar Pustaka : 6 (1968 – 2009)

ABSTRACT

Name : Spina Electra Susila
Study Program : S1 Mathematics
Title : Regression Estimation for Population's Mean in Stratified Random Sampling

This paper discusses a method of estimating the mean by a method called regression estimates in stratified random sampling. Estimated regression was used to estimate the mean of a variable with the involvement of other variables that correlate strongly and has a linear relationship with related variables based on simple random sample. The estimated regression is divided into 2 (two) types, namely the estimated regression estimated separately and the combined regression. Moreover, in this paper will be given examples of its use by using fictitious data.

Keywords : regression estimation, stratified random sampling, simple random sample
xiii + 92 pages; 18 tables; 4 attachments
Bibliography : 6 (1968 – 2009)

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Dalam suatu penelitian, biasanya ingin diukur suatu karakteristik dari populasi yang diamati yang sering disebut parameter. Hal ini sulit untuk dilakukan karena keterbatasan waktu dan dana yang tersedia. Untuk mengatasi hal ini dilakukan pengambilan sampel dari populasi tersebut. Sampel yang diambil haruslah yang dapat mewakili populasi sehingga taksiran dari parameter yang diteliti mendekati nilai parameter yang sebenarnya. Untuk mendapatkan sampel yang mewakili populasi dibutuhkan suatu tehnik pengambilan sampel yang tepat dengan keadaan populasi. Secara garis besar tehnik pengambilan sampel dapat digolongkan menjadi 2 (dua) macam, yaitu *Probability Sampling* yang merupakan tehnik pengambilan sampel dimana peluang terpilihnya setiap anggota sampel dapat ditentukan dan *Non Probability Sampling* yang merupakan tehnik pengambilan sampel dimana peluang terpilihnya setiap anggota sampel tidak dapat ditentukan.

Simple Random Sampling (SRS) dan *Stratified Random Sampling* adalah suatu tehnik pengambilan sampel yang merupakan bagian dari Probability Sampling. SRS baik digunakan jika keadaan populasi homogen. Tetapi apabila keadaan populasinya heterogen, SRS kurang baik digunakan. Jika kondisi populasinya heterogen, digunakan Stratified Random Sampling. Pada prinsipnya dalam Stratified Random Sampling, elemen – elemen populasi yang heterogen dikelompokkan ke dalam kelompok – kelompok yang homogen yang tidak *overlapping* yang disebut *strata*. Kemudian dari setiap *stratum* dipilih elemen – elemen secara SRS.

Biasanya taksiran parameter diperoleh berdasarkan data suatu variabel terkait yang misalnya mempunyai nilai $y_1, y_2, ..., y_n$. Tetapi terkadang pada kenyataannya terdapat variabel – variabel lain yang berkorelasi kuat dengan variabel terkait, sebut sebagai *X*. Informasi dari variabel tambahan dapat digunakan dalam penaksiran parameter. Penggunaan dari variabel tambahan dapat menghasilkan taksiran yang lebih teliti dibandingkan taksiran yang didapat hanya dengan menggunakan data dari variabel yang diamati saja,*Y*.

Taksiran parameter yang diperoleh dengan melibatkan variabel tambahan X ini adalah taksiran regresi atau taksiran rasio. Apabila terdapat hubungan linier antara variabel X dengan variabel Y dapat dituliskan sebagai $Y = a + bX$, maka informasi kelinieran tersebut dapat diperhitungkan dalam penaksiran parameter sehingga diperoleh taksiran yang lebih teliti. Metode penaksiran parameter yang melibatkan hubungan linier antara Y dan X seperti yang telah dijelaskan diatas adalah dengan taksiran regresi. Taksiran regresi dapat digunakan untuk menaksir parameter mean populasi berdasarkan data di *Simple Random Sample* ataupun *Stratified Random Sample*. Taksiran regresi pada Stratified Random Sample terbagi menjadi 2 (dua), yaitu taksiran regresi terpisah dan taksiran regresi gabungan

Masalah yang akan diangkat dalam tugas akhir ini adalah bagaimana taksiran regresi diterapkan untuk mencari taksiran mean populasi berdasarkan Stratified Random Sample dengan menggunakan taksiran regresi baik terpisah maupun gabungan dan memeriksa apakah taksiran yang didapat merupakan taksiran yang bias atau taksiran yang tidak bias, khususnya untuk sampel yang berukuran besar.

1.2 PERUMUSAN MASALAH

Berdasarkan uraian diatas, maka rumusan masalah dalam tugas akhir ini adalah :

1. Bagaimana mencari taksiran regresi untuk mean populasi berdasarkan Stratified Random Sample.
2. Bagaimana sifat dari taksiran regresi yang didapat beserta taksiran variansinya..

1.3 TUJUAN PENULISAN

Pada tugas akhir ini, tujuan penulisannya adalah :

1. Membahas taksiran regresi untuk mean populasi berdasarkan Stratified Random Sample.
2. Memeriksa sifat taksiran regresi yang didapat beserta taksiran variansinya.

1.4 PEMBATASAN MASALAH

Pada tugas akhir ini, pembatasan masalahnya adalah sebagai berikut :

1. Pengambilan sampel pada Stratified Random Sampling dilakukan dengan menggunakan alokasi proporsional.
2. Setiap *stratum* dipilih elemen – elemen secara Simple Random Sampling (SRS) tanpa pengambilan.
3. Sampel yang diambil berukuran besar.
4. Variabel tambahan yang terlibat hanya satu variabel, yaitu X.

1.5 SISTEMATIKA PENULISAN

Penulisan tugas akhir yang merupakan hasil studi pustaka ini, dibagi menjadi lima bab, yaitu :

Bab I membahas mengenai latar belakang, perumusan masalah, tujuan penulisan, pembatasan masalah dan sistematika penulisan.

Bab II membahas teori – teori dasar yang akan digunakan dalam taksiran regresi untuk mean populasi pada Stratified Random Sampling. Diantaranya adalah teori mengenai taksiran mean populasi dari Simple Random Sampling (SRS), taksiran untuk mean populasi pada Stratified Random Sampling, dan taksiran regresi untuk mean dan populasi pada SRS.

Bab III membahas taksiran regresi untuk mean populasi pada Stratified Random Sampling.

Bab IV membahas penggunaan dari taksiran regresi untuk mean populasi pada Stratified Random Sampling.

Bab V berisi kesimpulan dan saran untuk tugas akhir ini.

BAB II

LANDASAN TEORI

Pada bab ini akan dijelaskan teori – teori yang akan digunakan dalam taksiran regresi untuk mean populasi pada stratified random sampling. Diantaranya adalah teori mengenai *Simple Random Sampling* (SRS), *Stratified Random Sampling,* dan taksiran regresi pada Simple Random Sampling.

2.1 SIMPLE RANDOM SAMPLING

2.1.1 Pendahuluan

Simple Random Sampling (SRS) adalah teknik pengambilan sampel dimana setiap sampel yang berukuran *n* diambil dari populasi yang berukuran *N* dengan cara sedemikian sehingga setiap sampel yang mungkin mempunyai probabilitas yang sama untuk terpilih menjadi sampel. Teknik ini baik digunakan untuk keadaan populasi yang homogen.

Ada 2 (dua) cara pengambilan sampel dalam Simple Random Sampling (SRS), yaitu :

1. Simple Random Sampling tanpa pengembalian (SRSWOR)
   Elemen yang telah terpilih menjadi sampel tidak diikutsertakan kembali ke dalam populasi sehingga setiap elemen yang terpilih menjadi sampel memiliki peluang yang sama untuk terpilih.
2. Simple Random Sampling dengan pengembalian (SRSWR)
   Elemen yang telah terpilih menjadi sampel diikutsertakan kembali ke dalam populasi sehingga elemen yang terpilih menjadi sampel memiliki kemungkinan untuk terpilih menjadi sampel lebih dari satu kali.

Karena unit yang sama tidak memberikan tambahan informasi, maka yang sering digunakan adalah SRSWOR.

Pada SRS, setiap unit memiliki probabilitas yang sama untuk terpilih menjadi anggota sampel. Hal ini dapat dibuktikan dalam teorema berikut.

Teorema 2.1.1.1

Dalam Simple Random Sampling tanpa pengembalian, probabilitas suatu unit terpilih menjadi anggota sampel adalah $\frac{n}{N}$.

Bukti :

Misalkan diketahui nilai – nilai dari populasi, yaitu $\{u_1, u_2, ..., u_N\}$.

Kemudian didefinisikan

$p_j = \Pr(u_1 \text{ muncul pada pengambilan ke-}j)$; $j = 1, 2, ..., n$

Misalkan

A_j = Kejadian u_1 muncul pada pengambilan ke – j dengan $j = 1, 2, ..., n$

$A_j\,'$ = Kejadian u_1 tidak muncul pada pengambilan ke – j dengan $j = 1, 2, ..., n$

Untuk $j = 1$, maka :

$$p_1 = \Pr(u_1 \text{ muncul pada pengambilan ke-1})$$

$$p_1 = \frac{1}{N}$$

Untuk $j = 2$, maka :

$$p_2 = \Pr\begin{pmatrix} u_1 \text{ tidak muncul pada pengambilan ke-1} \\ \text{dan } u_1 \text{ muncul pada pengambilan ke-2} \end{pmatrix}$$

$$p_2 = \Pr(A_1' \cap A_2)$$

$$p_2 = \Pr(A_1')\Pr(A_2 | A_1')$$

$$p_2 = \left(1 - \frac{1}{N}\right)\left(\frac{1}{N-1}\right)$$

$$p_2 = \left(\frac{N-1}{N}\right)\left(\frac{1}{N-1}\right)$$

$$p_2 = \frac{1}{N}$$

Untuk $j = 3$, maka :

$$p_3 = \Pr\left(u_1 \text{ muncul pada pengambilan ke-3}\right)$$

$$p_3 = \Pr\begin{pmatrix} u_1 \text{ tidak muncul pada pengambilan ke-1,} \\ u_1 \text{ tidak muncul pada pengambilan ke-2,} \\ \text{dan } u_1 \text{ muncul pada pengambilan ke-3} \end{pmatrix}$$

$$p_3 = \Pr\left(A_1' \cap A_2' \cap A_3\right)$$

$$p_3 = \Pr\left(A_1' \cap A_2'\right)\Pr\left(A_3 \middle| A_1' \cap A_2'\right)$$

$$p_3 = \left(1 - \frac{1}{N}\right)\left(1 - \frac{1}{N-1}\right)\left(\frac{1}{N-2}\right)$$

$$p_3 = \left(\frac{N-1}{N}\right)\left(\frac{N-2}{N-1}\right)\left(\frac{1}{N-2}\right)$$

$$p_3 = \frac{1}{N}$$

Dengan cara yang sama, akan diperoleh :

$$p_j = \frac{1}{N} \text{ dimana } j = 1,2,\ldots,n \qquad \ldots(2.1.1.1)$$

Hal tersebut dapat juga dilakukan untuk setiap u_i ; $i = 1,2,\ldots,N$

Selanjutnya, misalkan :

$$\pi_1 = \Pr\left(u_1 \text{ terpilih dalam sampel}\right)$$

$$A_j = \text{Kejadian } u_1 \text{ muncul pada pengambilan ke-}j$$

Karena kejadian A_j merupakan kejadian yang saling lepas maka

$\pi_1 = \Pr\left(A_1 \cup A_2 \cup \ldots \cup A_n\right)$, sehingga

$$\pi_1 = \Pr\left(A_1 \cup A_2 \cup \ldots \cup A_n\right)$$

$$\pi_1 = \Pr\left(A_1\right) + \Pr\left(A_2\right) + \ldots \Pr\left(A_n\right)$$

$$\pi_1 = \frac{1}{N} + \frac{1}{N} + \ldots + \frac{1}{N}$$

$$\pi_1 = \frac{n}{N} \qquad \ldots(2.1.1.2)$$

Hal yang sama juga berlaku untuk setiap u_i, $i = 2,3,\ldots,N$

Dengan demikian terbukti bahwa dalam Simple Random Sampling (SRS) tanpa pengembalian, probabilitas suatu elemen terpilih menjadi anggota sampel adalah sama yaitu $\frac{n}{N}$.

(Terbukti)

2.1.2 Taksiran untuk Mean Populasi dan Variansinya

Misalkan diketahui nilai – nilai dari populasi, yaitu $\{u_1, u_2, ..., u_N\}$.

Didefinisikan mean populasi, μ_Y yaitu sebagai berikut :

$$\mu_Y = E(u_i)$$

$$\mu_Y = \sum_{i=1}^{N} u_i p(u_i)$$

$p(u_i)$ adalah probabilitas u_i terpilih dalam sampel.

$$\mu_Y = \sum_{i=1}^{N} u_i \frac{1}{N}$$

$$\mu_Y = \frac{1}{N}\sum_{i=1}^{N} u_i$$

Misalkan pula $S = \{y_1, y_2, ..., y_n\}$ adalah *Simple Random Sample* yang diambil dari populasi $\{u_1, u_2, ..., u_N\}$.

Definisikan taksiran mean populasi, $\bar{y}$ yaitu :

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

Akan dibuktikan bahwa :

1. $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ adalah taksiran yang tidak bias untuk mean populasi, μ_Y.

2. $Var(\bar{y}) = \frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$ dimana $\sigma_Y^2 = \frac{1}{N}\sum_{i=1}^{N}(u_i - \mu)^2$.

3. $\widehat{Var}(\bar{y}) = \frac{s_Y^2}{n}\left(\frac{N-n}{N}\right)$ dengan $s_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ adalah taksiran yang tidak bias untuk $Var(\bar{y})$.

Pembuktian :

1. Untuk membuktikan bahwa $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ adalah taksiran yang tidak bias untuk mean populasi, μ_Y yaitu dengan menunjukkan bahwa $E(\bar{y}) = \mu_Y$.

Bukti :

Misalkan didefinisikan sebuah variabel random *Z*, dimana :

$$z_i = 1 \; ; \; u_i \in S \; ; \; i = 1,2,...,N$$

$$z_i = 0 \; ; \; u_i \notin S \; ; \; i = 1,2,...,N \qquad \text{…(2.1.2.1)}$$

Akan dicari $E(Z_i)$, yaitu :

$$E(Z_i) = \sum_{z_i=0}^{1} z_i \Pr(Z_i = z_i)$$

$$E(Z_i) = 0.\Pr(Z_i = 0) + 1.\Pr(Z_i = 1)$$

$$E(Z_i) = \Pr(Z_i = 1)$$

Dari definisi (2.1.2.1) diatas diperoleh :

$$E(Z_i) = \Pr(Z_i = 1)$$

$$E(Z_i) = \Pr(u_i \in S)$$

Berdasarkan teorema 2.1.1.1, yaitu $\Pr(u_i \in S) = \frac{n}{N}$, diperoleh :

$$E(Z_i) = \frac{n}{N} \qquad \text{…(2.1.2.2)}$$

Kemudian perhatikan terlebih dahulu bentuk:

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

Maka $E(\bar{y})$ menjadi :

$$E(\bar{y}) = E\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)$$

$$E(\bar{y}) = E\left(\frac{1}{n}\sum_{i=1}^{N} u_i z_i\right)$$

$$E(\bar{y}) = \frac{1}{n}E\left(\sum_{i=1}^{N} u_i z_i\right)$$

$$E(\bar{y}) = \frac{1}{n}\left(\sum_{i=1}^{N} u_i E(z_i)\right)$$

$$E(\bar{y}) = \frac{1}{n}\left(\sum_{i=1}^{N} u_i \frac{n}{N}\right)$$

$$E(\bar{y}) = \frac{1}{N}\sum_{i=1}^{N} u_i$$

$$E(\bar{y}) = \mu_Y$$

Karena telah diperoleh $E\left(\bar{y}\right)=\mu_Y$, maka terbukti bahwa $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$ adalah taksiran yang tidak bias untuk μ_Y.

(Terbukti)

2. Selanjutnya akan dibuktikan bahwa $Var\left(\bar{y}\right)=\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$, dimana

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

Bukti :

$$Var\left(\bar{y}\right)=Var\left(\frac{1}{n}\sum_{i=1}^{n}y_i\right)$$

$$Var\left(\bar{y}\right)=Var\left(\frac{1}{n}\sum_{i=1}^{N}u_iz_i\right)$$

$$Var\left(\bar{y}\right)=\frac{1}{n^2}Var\left(\sum_{i=1}^{N}u_iz_i\right)$$

$$Var\left(\bar{y}\right)=\frac{1}{n^2}\left\{E\left(\sum_{i=1}^{N}u_iz_i\right)^2-\left(E\left(\sum_{i=1}^{N}u_iz_i\right)\right)^2\right\}$$

$$Var\left(\bar{y}\right)=\frac{1}{n^2}\left\{E\left(\sum_{i=1}^{N}\left(u_iz_i\right)^2+\sum_{i=1}^{N}\sum_{j\neq i}u_iu_jz_iz_j\right)-\left(\sum_{i=1}^{N}E\left(u_iz_i\right)\right)^2\right\}$$

$$\begin{aligned}Var\left(\bar{y}\right)=\frac{1}{n^2}\Bigg\{&E\left(\sum_{i=1}^{N}\left(u_iz_i\right)^2\right)+E\left(\sum_{i=i}^{N}\sum_{j\neq i}u_iu_jz_iz_j\right)\\&-\left(\sum_{i=1}^{N}\left(E\left(u_iz_i\right)\right)^2+\sum_{i=1}^{N}\sum_{j\neq i}E\left(u_iz_i\right)E\left(u_jz_j\right)\right)\Bigg\}\end{aligned}$$

$$\begin{aligned}Var\left(\bar{y}\right)=\frac{1}{n^2}\Bigg\{&\left(\sum_{i=1}^{N}E\left(u_iz_i\right)^2\right)-\sum_{i=1}^{N}\left(E\left(u_iz_i\right)\right)^2+\left(\sum_{i=i}^{N}\sum_{j\neq i}E\left(u_iu_jz_iz_j\right)\right)\\&-\left(\sum_{i=1}^{N}\sum_{j\neq i}E\left(u_iz_i\right)E\left(u_jz_j\right)\right)\Bigg\}\end{aligned}$$

$$Var\left(\bar{y}\right)=\frac{1}{n^2}\left\{\sum_{i=1}^{N}\left(E\left(u_iz_i\right)^2-\left(E\left(u_iz_i\right)\right)^2\right)+\sum_{i=1}^{N}\sum_{j\neq i}\left(E\left(u_iu_jz_iz_j\right)-E\left(u_iz_i\right)E\left(u_jz_j\right)\right)\right\}$$

$$Var\left(\bar{y}\right)=\frac{1}{n^2}\left\{\sum_{i=1}^{N}\left(u_i^2E\left(z_i\right)^2-u_i^2\left(E\left(z_i\right)\right)^2\right)+\sum_{i=1}^{N}\sum_{j\neq i}\left(u_iu_jE\left(z_iz_j\right)-u_iu_jE\left(z_i\right)E\left(z_j\right)\right)\right\}$$

$$Var\left(\bar{y}\right)=\frac{1}{n^2}\left\{\sum_{i=1}^{N}u_i^2\left(E\left(z_i\right)^2-\left(E\left(z_i\right)\right)^2\right)+\sum_{i=1}^{N}\sum_{j\neq i}u_iu_j\left(E\left(z_iz_j\right)-E\left(z_i\right)E\left(z_j\right)\right)\right\}$$

$$Var(\bar{y}) = \frac{1}{n^2}\left\{\sum_{i=1}^{N} u_i^2 Var(z_i) + \sum_{i=1}^{N}\sum_{j\neq i} u_i u_j Cov(z_i, z_j)\right\} \qquad ...(2.1.2.3)$$

Terlebih dahulu akan dicari $Var(z_i)$ dan $Cov(z_i, z_j)$.

$$Var(z_i) = E(z_i^2) - (E(z_i))^2$$

$$Var(z_i) = \sum_{z_i=0}^{1} z_i^2 \Pr(Z_i = z_i) - (E(z_i))^2$$

$$Var(z_i) = 0^2.\Pr(Z_i = 0) + 1^2.\Pr(Z_i = 1) - (E(z_i))^2$$

$$Var(z_i) = \Pr(Z_i = 1) - (E(z_i))^2$$

$$Var(z_i) = E(z_i) - (E(z_i))^2 \qquad \ldots(2.1.2.4)$$

Substitusikan persamaan (2.1.2.2) ke dalam persamaan (2.1.2.4) menjadi :

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

$$Var(z_i) = \frac{n}{N}\left(1 - \frac{n}{N}\right) \qquad \ldots(2.1.2.5)$$

Kemudian akan dicari nilai dari $Cov(z_i, z_j)$, yaitu :

$$Cov(z_i, z_j) = E(z_i z_j) - E(z_i)E(z_j)$$

$$Cov(z_i, z_j) = \sum_{z_i, z_j = 0}^{1} z_i z_j \Pr(Z_i = z_i, Z_j = z_j) - E(z_i)E(z_j)$$

$$Cov(z_i, z_j) = 0.0.\Pr(Z_i = 0, Z_j = 0) + 0.1.\Pr(Z_i = 0, Z_j = 1) + 1.0.\Pr(Z_i = 1, Z_j = 0)$$
$$+ 1.1.\Pr(Z_i = 1, Z_j = 1) - E(z_i)E(z_j)$$

$$Cov(z_i, z_j) = 1.1.\Pr(Z_i = 1, Z_j = 1) - E(z_i)E(z_j)$$

$$Cov(z_i, z_j) = \Pr(Z_i = 1, Z_j = 1) - E(z_i)E(z_j) \qquad \ldots(2.1.2.6)$$

Terlebih dahulu akan dicari $E(z_i)E(z_j)$.

Berdasarkan definisi (2.1.2.1) sebelumnya diperoleh :

$$E(z_i z_j) = \Pr(u_i \in S, u_j \in S)$$

$$E(z_i z_j) = \Pr(u_i \text{ dan } u_j \text{ masuk dalam sampel})$$

$$E\left(z_i z_j\right) = \frac{\binom{N-2}{n-2}}{\binom{N}{n}}$$

$$E\left(z_i z_j\right) = \frac{\dfrac{(N-2)!}{(N-n)!(n-2)!}}{\dfrac{N!}{(N-n)!n!}}$$

$$E\left(z_i z_j\right) = \frac{(N-2)!}{(N-n)!(n-2)!}.\frac{(N-n)!n!}{N!}$$

$$E\left(z_i z_j\right) = \frac{(N-2)!n!}{(n-2)!N!}$$

$$E\left(z_i z_j\right) = \frac{(N-2)!}{(n-2)!}.\frac{n(n-1)(n-2)!}{N(N-1)(N-2)!}$$

$$E\left(z_i z_j\right) = \frac{n(n-1)}{N(N-1)} \qquad \ldots(2.1.2.7)$$

Substitusikan persamaan (2.1.2.7) ke dalam persamaan (2.1.2.6) sehingga diperoleh nilai dari $Cov\left(z_i, z_j\right)$, yaitu :

$$Cov\left(z_i, z_j\right) = E\left(z_i z_j\right) - E\left(z_i\right)E\left(z_j\right)$$

$$Cov\left(z_i, z_j\right) = \frac{n(n-1)}{N(N-1)} - E\left(z_i\right)E\left(z_j\right) \qquad \ldots(2.1.2.8)$$

Substitusikan (2.1.2.2) ke dalam persamaan (2.1.2.8) sehingga diperoleh :

$$Cov\left(z_i, z_j\right) = \frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2$$

$$Cov\left(z_i, z_j\right) = \frac{n}{N}\left(\frac{n-1}{N-1} - \frac{n}{N}\right)$$

$$Cov\left(z_i, z_j\right) = \frac{n}{N}\left(\frac{N(n-1) - n(N-1)}{N(N-1)}\right)$$

$$Cov\left(z_i, z_j\right) = \frac{n}{N}\left(\frac{Nn - N - Nn + n}{N(N-1)}\right)$$

$$Cov(z_i, z_j) = \frac{n}{N}\left(\frac{n-N}{N(N-1)}\right)$$

$$Cov(z_i, z_j) = \frac{n}{N}\frac{1}{N-1}\left(\frac{n-N}{N}\right)$$

$$Cov(z_i, z_j) = \frac{n}{N}\frac{1}{N-1}\left(\frac{n}{N}-1\right) \qquad \ldots(2.1.2.9)$$

untuk $i = 1,2,\ldots,N$ dan $j = 1,2,\ldots,N$ dimana $i \neq j$

Substitusikan persamaan (2.1.2.5) dan (2.1.2.9) ke dalam persamaan (2.1.2.3) sehingga diperoleh :

$$Var(\bar{y}) = \frac{1}{n^2}\left\{\sum_{i=1}^{N} u_i^2 \frac{n}{N}\left(1-\frac{n}{N}\right) + \sum_{i=1}^{N}\sum_{j\neq i} u_i u_j \frac{n}{N}\frac{1}{N-1}\left(\frac{n}{N}-1\right)\right\}$$

$$Var(\bar{y}) = \frac{1}{n^2}\left\{\sum_{i=1}^{N} u_i^2 \frac{n}{N}\left(1-\frac{n}{N}\right) + \sum_{i=1}^{N}\sum_{j\neq i} u_i u_j \frac{n}{N}\left(-\frac{1}{N-1}\right)\left(1-\frac{n}{N}\right)\right\}$$

$$Var(\bar{y}) = \frac{1}{n^2}\frac{n}{N}\left(1-\frac{n}{N}\right)\left\{\sum_{i=1}^{N} u_i^2 + \left(-\frac{1}{N-1}\right)\sum_{i=1}^{N}\sum_{j\neq i} u_i u_j\right\}$$

$$Var(\bar{y}) = \frac{1}{n}\frac{1}{N}\left(1-\frac{n}{N}\right)\left\{\sum_{i=1}^{N} u_i^2 - \frac{1}{N-1}\sum_{i=1}^{N}\sum_{j\neq i} u_i u_j\right\}$$

$$Var(\bar{y}) = \frac{1}{n}\frac{1}{N}\left(\frac{N-n}{N}\right)\left\{\sum_{i=1}^{N} u_i^2 - \frac{1}{N-1}\sum_{i=1}^{N}\sum_{j\neq i} u_i u_j\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left\{\left(\frac{1}{N}\right)^2 \sum_{i=1}^{N} u_i^2 - \frac{1}{N-1}\left(\left(\frac{1}{N}\right)^2\left(\left(\sum_{i=1}^{N} u_i\right)^2 - \sum_{i=1}^{N} u_i^2\right)\right)\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left\{\left(\frac{1}{N} + \frac{1}{N-1}\left(\frac{1}{N}\right)\right)\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \frac{1}{N-1}\left(\frac{1}{N}\right)^2\left(\sum_{i=1}^{N} u_i\right)^2\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left\{\left(\frac{N}{N-1}\frac{1}{N}\right)\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \frac{1}{N-1}\left(\frac{1}{N}\right)^2\left(\sum_{i=1}^{N} u_i\right)^2\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left\{\left(\frac{1}{N-1}\right)\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \frac{1}{N-1}\left(\frac{1}{N}\right)^2\left(\sum_{i=1}^{N} u_i\right)^2\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left(\frac{1}{N-1}\right)\left\{\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \left(\frac{1}{N}\right)^2\left(\sum_{i=1}^{N} u_i\right)^2\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left(\frac{1}{N-1}\right)\left\{\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} u_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} u_i\right)\right\} \qquad \ldots(2.1.2.10)$$

Karena $\mu_Y = \frac{1}{N}\sum_{i=1}^{N} u_i$, maka persamaan (2.1.2.10) menjadi :

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left(\frac{1}{N-1}\right)\left\{\frac{1}{N}\sum_{i=1}^{N} u_i^2 - \frac{N}{N}\mu^2\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left(\frac{1}{N-1}\right)\left\{\frac{1}{N}\left(\sum_{i=1}^{N} u_i^2 - N\mu^2\right)\right\}$$

$$Var(\bar{y}) = \frac{1}{n}(N-n)\left(\frac{1}{N-1}\right)\left\{\frac{1}{N}\sum_{i=1}^{N} (u_i - \mu)^2\right\}$$

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

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

Dengan demikian diperoleh $Var(\bar{y}) = \frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$ dimana $\sigma_Y^2 = \frac{1}{N}\sum_{i=1}^{N}(u_i - \mu)^2$.

(Terbukti)

3. Berikutnya akan dibuktikan bahwa $\widehat{Var}(\bar{y}) = \frac{s_Y^2}{n}\left(\frac{N-n}{N}\right)$ dengan $s_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ adalah taksiran yang tidak bias untuk $Var(\bar{y})$.

Bukti :

Untuk membuktikan $\widehat{Var}(\bar{y})$ adalah taksiran yang tidak bias untuk $Var(\bar{y})$, yaitu dengan menunjukkan bahwa $E\left(\widehat{Var}(\bar{y})\right) = Var(\bar{y})$.

$$E\left(\widehat{Var}(\bar{y})\right) = E\left(\frac{s_Y^2}{n}\left(\frac{N-n}{N}\right)\right)$$

$$E\left(\widehat{Var}(\bar{y})\right) = \left(\frac{N-n}{N}\right)\frac{1}{n}E\left(s_Y^2\right) \qquad \ldots(2.1.2.11)$$

Terlebih dahulu akan dicari $E\left(s_Y^2\right)$, yaitu :

$$E\left(s_Y^2\right) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}E\left(\sum_{i=1}^{n}(y_i - \bar{y})^2\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}E\left(\sum_{i=1}^{n}\left((y_i-\mu)-(\bar{y}-\mu)\right)^2\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}E\left(\sum_{i=1}^{n}(y_i-\mu)^2 - n(\bar{y}-\mu)^2\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(E\left(\sum_{i=1}^{n}(y_i-\mu)^2\right) - nE(\bar{y}-\mu)^2\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}E(y_i-\mu)^2 - nVar(\bar{y})\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}\left(\sum_{j=1}^{N}\left(y_{ij}-\mu\right)^2 P\left(y_j\right)\right) - n\left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}\left(\left(\sum_{j=1}^{N}y_{ij}^2 - N\mu^2\right)\frac{1}{N}\right) - n\left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}\frac{1}{N}\left(\sum_{j=1}^{N}\left(u_{ij}z_{ij}\right)^2 - N\mu^2\right) - n\left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}\frac{1}{N}\left(\sum_{j=1}^{N}u_{ij}^2 - N\mu^2\right) - n\left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}\frac{1}{N}\left(\sum_{j=1}^{N}\left(u_{ij}-\mu\right)^2\right) - n\left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)\right)$$

$$E\left(s_Y^2\right) = \frac{1}{n-1}\left(\frac{n}{N}\sum_{j=1}^{N}\left(u_j-\mu\right)^2 - n\left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)\right)$$

$$E\left(s_Y^2\right) = \frac{n}{n-1}\left(\frac{1}{N}\sum_{j=1}^{N}\left(u_j-\mu\right)^2 - \left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)\right)$$

$$E\left(s_Y^2\right) = \frac{n}{n-1}\left(\sigma_Y^2 - \frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)$$

$$E\left(s_Y^2\right) = \frac{\sigma_Y^2}{n-1}\left(n - \left(\frac{N-n}{N-1}\right)\right)$$

$$E\left(s_Y^2\right) = \frac{\sigma_Y^2}{n-1}\left(\frac{N(n-1)}{N-1}\right)$$

$$E\left(s_Y^2\right) = \frac{N}{N-1}\sigma_Y^2 \qquad \ldots(2.1.2.12)$$

Kemudian substitusikan persamaan (2.1.2.12) ke dalam persamaan (2.1.2.11) menjadi :

$$E\left(\widehat{Var}\left(\bar{y}\right)\right)=\left(\frac{N-n}{N}\right)\left(\frac{1}{n}\right)\left(\frac{N}{N-1}\right)\sigma_Y^2$$

$$E\left(\widehat{Var}\left(\bar{y}\right)\right)=\left(\frac{N-n}{N-1}\right)\frac{\sigma_Y^2}{n}$$

$$E\left(\widehat{Var}\left(\bar{y}\right)\right)=Var\left(\bar{y}\right)$$

Karena telah diperoleh $E\left(\widehat{Var}\left(\bar{y}\right)\right)=Var\left(\bar{y}\right)$, maka terbukti bahwa $\widehat{Var}\left(\bar{y}\right)=\frac{s_Y^2}{n}\left(\frac{N-n}{N}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\bar{y}\right)$.

(Terbukti)

2.1.3 Taksiran Total Populasi dan Variansinya

Definisikan total populasi, τ_Y yaitu :

$$\tau_Y = N\mu_Y$$

Akan dibuktikan bahwa :

1. $\widehat{\tau_Y} = N\bar{y}$ adalah taksiran yang tidak bias untuk τ_Y.
2. $Var\left(\widehat{\tau_Y}\right)=N^2\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$
3. $\widehat{Var}\left(\widehat{\tau_Y}\right)=N^2\frac{s_Y^2}{n}\left(\frac{N-n}{N}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\widehat{\tau_Y}\right)$.

Pembuktian :

1. Pertama akan dibuktikan bahwa $\widehat{\tau_Y} = N\bar{y}$ adalah taksiran yang tidak bias untuk τ_Y.

Bukti :

Untuk membuktikan bahwa $\widehat{\tau_Y} = N\bar{y}$ adalah taksiran yang tidak bias untuk τ_Y, yaitu dengan menunjukkan bahwa $E\left(\widehat{\tau_Y}\right)=\tau_Y$.

$E\left(\widehat{\tau_Y}\right) = E\left(N\bar{y}\right)$

$E\left(\widehat{\tau_Y}\right) = NE\left(\bar{y}\right)$

Karena $\bar{y}$ adalah taksiran yang tidak bias untuk μ_Y, maka :

$E\left(\widehat{\tau_Y}\right) = N\mu_Y$

$E\left(\widehat{\tau_Y}\right) = \tau_Y$

Karena telah diperoleh $E\left(\widehat{\tau_Y}\right) = \tau_Y$, maka terbukti bahwa $\widehat{\tau_Y} = N\bar{y}$ adalah taksiran yang tidak bias untuk τ_Y.

(Terbukti)

2. Selanjutnya akan dibuktikan $Var\left(\widehat{\tau_Y}\right) = N^2 \frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$

Bukti :

$Var\left(\widehat{\tau_Y}\right) = Var\left(N\bar{y}\right)$

$Var\left(\widehat{\tau_Y}\right) = N^2 Var\left(\bar{y}\right)$

Karena dalam simple random sample $Var\left(\bar{y}\right) = \frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$, maka :

$Var\left(\widehat{\tau_Y}\right) = N^2\left(\frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)\right)$

$Var\left(\widehat{\tau_Y}\right) = N^2 \frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$

Dengan demikian diperoleh $Var\left(\widehat{\tau_Y}\right) = N^2 \frac{\sigma_Y^2}{n}\left(\frac{N-n}{N-1}\right)$.

(Terbukti)

3. Berikutnya akan dibuktikan $\widehat{Var}\left(\widehat{\tau_Y}\right) = N^2 \frac{s_Y^2}{n}\left(\frac{N-n}{N}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\widehat{\tau_Y}\right)$

Bukti :

Untuk membuktikan $\widehat{Var}\left(\widehat{\tau_Y}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\widehat{\tau_Y}\right)$ yaitu dengan menunjukkan bahwa $E\left(\widehat{Var}\left(\widehat{\tau_Y}\right)\right) = Var\left(\widehat{\tau_Y}\right)$.

$$E\left(\widehat{Var}\left(\widehat{\tau_Y}\right)\right) = E\left(N^2 \frac{s_Y^2}{n}\left(\frac{N-n}{N}\right)\right)$$

$$E\left(\widehat{Var}\left(\widehat{\tau_Y}\right)\right) = N^2 \frac{1}{n}\left(\frac{N-n}{N}\right)E\left(s_Y^2\right)$$

Karena $E\left(s_Y^2\right) = \frac{N}{N-1}\sigma_Y^2$, maka :

$$E\left(\widehat{Var}\left(\widehat{\tau_Y}\right)\right) = N^2 \frac{1}{n}\left(\frac{N-n}{N}\right)\frac{N}{N-1}\sigma_Y^2$$

$$E\left(\widehat{Var}\left(\widehat{\tau_Y}\right)\right) = Var\left(\widehat{\tau_Y}\right)$$

Karena telah diperoleh $E\left(\widehat{Var}\left(\widehat{\tau_Y}\right)\right) = Var\left(\widehat{\tau_Y}\right)$, maka terbukti bahwa $\widehat{Var}\left(\widehat{\tau_Y}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\widehat{\tau_Y}\right)$.

(Terbukti)

2.2 STRATIFIED RANDOM SAMPLING

2.2.1 Pendahuluan

Stratified Random Sampling adalah teknik pengambilan sampel dengan mengelompokkan elemen – elemen populasi yang heterogen ke dalam kelompok – kelompok yang homogen yang tidak *overlapping* yang disebut strata, kemudian dari setiap stratum dipilih elemen – elemen secara SRS.

Misalkan diketahui populasi yang berukuran *N* kemudian dibagi menjadi subpopulasi – subpopulasi sebanyak L yang tidak overlapping dengan masing – masing berukuran $N_1, N_2, ..., N_L$ dan $N_1 + N_2 + ... + N_L = N$. Subpopulasi ini disebut strata. Kemudian dari setiap stratum diambil sampel yang masing – masing berukuran $n_1, n_2, ..., n_L$ dengan menggunakan Simple Random Sampling (SRS).

Sebelum masuk ke sub bab mengenai taksiran untuk mean dan variansinya, maka berdasarkan uraian diatas akan terlebih dahulu diperkenalkan beberapa notasi dalam Stratified Random Sampling.

Adapun notasi – notasinya adalah sebagai berikut :

N_i : Ukuran populasi stratum ke-i

n_i : Ukuran sampel yang diambil dari stratum ke-i secara SRS

μ_{Y_i} : Mean populasi pada stratum ke-*i*

$\bar{y}_i$: Mean sampel pada stratum ke-i yang diambil secara SRS

τ_{Y_i} : Total populasi pada stratum ke-i yang diambil secara SRS

2.2.2 Taksiran untuk Mean dan Variansinya

Karena sampel ukuran n_i diambil secara SRS untuk setiap stratum ke-*i*, maka :

$$\bar{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}$$

adalah taksiran yang tidak bias untuk μ_{Y_i}.

Definisikan taksiran mean populasi, $\bar{y}_{st}$ yang diperoleh dari sampel, yaitu :

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

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

Akan dibuktikan bahwa :

1. $\bar{y}_{st} = \frac{1}{N}\sum_{i=1}^{L} N_i\bar{y}_i$ adalah taksiran yang tidak bias untuk μ_Y.

2. $Var\left(\bar{y}_{st}\right) = \sum_{i=1}^{L}\left(\frac{N_i - n_i}{N_i - 1}\right)\left(\frac{N_i}{N}\right)^2 \frac{\sigma_{Y_i}^2}{n_i}$.

3. $\widehat{Var}\left(\bar{y}_{st}\right) = \sum_{i=1}^{L}\left(1 - \frac{n_i}{N_i}\right)\left(\frac{N_i}{N}\right)^2 \frac{s_{Y_i}^2}{n_i}$ adalah taksiran yang tidak bias untuk $Var\left(\bar{y}_{st}\right)$.

Pembuktian :

1. Untuk membuktikan bahwa $\bar{y}_{st} = \frac{1}{N}\sum_{i=1}^{L} N_i \bar{y}_i$ adalah taksiran yang tidak bias untuk mean populasi, μ_Y, akan ditunjukkan bahwa $E(\bar{y}_{st}) = \mu_Y$.

Bukti :

$$E(\bar{y}_{st}) = E\left(\frac{1}{N}\sum_{i=1}^{L} N_i \bar{y}_i\right)$$

$$E(\bar{y}_{st}) = \frac{1}{N} E\left(\sum_{i=1}^{L} N_i \bar{y}_i\right)$$

$$E(\bar{y}_{st}) = \frac{1}{N}\left(\sum_{i=1}^{L} N_i E(\bar{y}_i)\right)$$

$$E(\bar{y}_{st}) = \frac{1}{N}\left(\sum_{i=1}^{L} N_i \mu_{Y_i}\right)$$

$$E(\bar{y}_{st}) = \frac{1}{N}\left(N_1\mu_{Y_1} + N_2\mu_{Y_2} + \ldots + N_L\mu_{Y_L}\right)$$

$$E(\bar{y}_{st}) = \frac{1}{N}\left(\tau_{Y_1} + \tau_{Y_2} + \ldots + \tau_{Y_L}\right) = \frac{1}{N}\sum_{i=1}^{L}\tau_{Y_i}$$

$$E(\bar{y}_{st}) = \frac{1}{N}\tau_Y$$

$$E(\bar{y}_{st}) = \mu_Y$$

Karena telah diperoleh $E(\bar{y}_{st}) = \mu_Y$, maka terbukti bahwa $\bar{y}_{st}$ adalah taksiran yang tidak bias untuk mean populasi, μ_Y .

(Terbukti)

2. Selanjutnya akan dibuktikan bahwa $Var(\bar{y}_{st}) = \sum_{i=1}^{L}\left(\frac{N_i - n_i}{N_i - 1}\right)\left(\frac{N_i}{N}\right)^2 \frac{\sigma_{Y_i}^2}{n_i}$.

Bukti :

$$Var(\bar{y}_{st}) = Var\left(\frac{1}{N}\sum_{i=1}^{L} N_i \bar{y}_i\right)$$

$$Var(\bar{y}_{st}) = \frac{1}{N^2}\left(\sum_{i=1}^{L} N_i^2 Var(\bar{y}_i)\right) \qquad \text{...(2.2.2.1)}$$

Karena sampel yang diambil dari setiap stratum adalah secara SRS, berarti :

$$Var(\bar{y}_i) = \left(\frac{N_i - n_i}{N_i - 1}\right)\frac{\sigma_{Y_i}^2}{n_i} \qquad \text{...(2.2.2.2)}$$

Substitusikan persamaan (2.2.2.2) ke dalam persamaan (2.2.2.1) sehingga diperoleh :

$$Var\left(\bar{y}_{st}\right)=\frac{1}{N^2}\left(\sum_{i=1}^{L}N_i^2\left(\frac{N_i-n_i}{N_i-1}\right)\frac{\sigma_{Y_i}^2}{n_i}\right)$$

$$Var\left(\bar{y}_{st}\right)=\sum_{i=1}^{L}\left(\frac{N_i-n_i}{N_i-1}\right)\left(\frac{N_i}{N}\right)^2\frac{\sigma_{Y_i}^2}{n_i}$$

Dengan demikian diperoleh $Var\left(\bar{y}_{st}\right)=\sum_{i=1}^{L}\left(\frac{N_i-n_i}{N_i-1}\right)\left(\frac{N_i}{N}\right)^2\frac{\sigma_{Y_i}^2}{n_i}$.

(Terbukti)

3. Berikutnya akan dibuktikan bahwa $\widehat{Var}\left(\bar{y}_{st}\right)=\sum_{i=1}^{L}\left(1-\frac{n_i}{N_i}\right)\left(\frac{N_i}{N}\right)^2\frac{s_{Y_i}^2}{n_i}$ adalah taksiran yang tidak bias untuk $Var\left(\bar{y}_{st}\right)$.

Bukti :

Untuk membuktikan bahwa $\widehat{Var}\left(\bar{y}_{st}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\bar{y}_{st}\right)$ akan ditunjukkan bahwa $E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=Var\left(\bar{y}_{st}\right)$.

$$E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=E\left(\sum_{i=1}^{L}\left(1-\frac{n_i}{N_i}\right)\left(\frac{N_i}{N}\right)^2\frac{s_{Y_i}^2}{n_i}\right)$$

$$E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=\sum_{i=1}^{L}E\left(\left(1-\frac{n_i}{N_i}\right)\left(\frac{N_i}{N}\right)^2\frac{s_{Y_i}^2}{n_i}\right)$$

$$E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=\sum_{i=1}^{L}\left(\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{N_i}{N}\right)^2\left(\frac{1}{n_i}\right)E\left(s_{Y_i}^2\right)\right) \quad ...(2.2.2.3)$$

Karena sampel yang diambil dari setiap stratum adalah secara SRS, berarti :

$$E\left(s_{Y_i}^2\right)=\left(\frac{N_i}{N_i-1}\right)\sigma_{Y_i}^2 \quad ...(2.2.2.4)$$

Kemudian substitusikan persamaan (2.2.2.4) ke dalam persamaan (2.2.2.3) sehingga diperoleh :

$$E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=\sum_{i=1}^{L}\left(\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{N_i}{N}\right)^2\left(\frac{1}{n_i}\right)\left(\frac{N_i}{N_i-1}\right)\sigma_{Y_i}^2\right)$$

$$E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=\sum_{i=1}^{L}\left(\left(\frac{N_i-n_i}{N_i-1}\right)\left(\frac{N_i}{N}\right)^2\left(\frac{\sigma_{Y_i}^2}{n_i}\right)\right)$$

$$E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=Var\left(\bar{y}_{st}\right)$$

Karena telah diperoleh $E\left(\widehat{Var}\left(\bar{y}_{st}\right)\right)=Var\left(\bar{y}_{st}\right)$, maka terbuktilah bahwa $\widehat{Var}\left(\bar{y}_{st}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\bar{y}_{st}\right)$.

(Terbukti)

2.2.3 Alokasi Sampel

Dalam Stratified Random Sampling satu hal yang perlu diperhatikan adalah bagaimana alokasi elemen sampel pada setiap stratum. Ada banyak cara untuk mengalokasikan *n* elemen sampel pada *L* strata.

Misalkan n_i ukuran sampel stratum ke-*i*, $i = 1,2,...,L$ dimana $n = \sum_{i=1}^{L} n_i$.

Alokasi sampel yang baik dipengaruhi oleh beberapa faktor, yaitu :

1. Ukuran populasi tiap stratum, N_i .

   Apabila stratum dengan N_i makin besar maka n_i makin besar.

2. Variansi populasi tiap stratum.

   Variansi makin besar maka ukuran sampel n_i makin besar.

3. Biaya pengambilan unit sampel pada stratum ke-*i*.

   Makin mahal biayanya maka ukuran sampel n_i makin kecil.

Alokasi sampel dapat diperoleh dengan 2 (dua) cara, yaitu :

1. Meminimumkan variansi dengan biaya tetap.
2. Meminimumkan biaya dengan variansi tetap.

Misalkan *C* adalah total biaya dan c_0 adalah biaya awal dan c_i adalah biaya pengambilan unit sampel pada stratum ke-*i*.

$$C = c_0 + \sum_{i=1}^{L} c_i n_i$$

2.2.3.1 Alokasi Sampel dengan Meminimumkan Variansi dengan Biaya Konstan

Akan dicari alokasi sampel dengan meminimumkan :

$$Var(\bar{y}_{st}) = \sum_{i=1}^{L} \left(\frac{N_i}{N}\right)^2 \left(1 - \frac{n_i}{N_i}\right) \frac{\sigma_{Y_i}^2}{n_i}$$

$$Var(\bar{y}_{st}) = \sum_{i=1}^{L} \left(\frac{N_i}{N}\right)^2 \frac{\sigma_{Y_i}^2}{n_i} - \sum_{i=1}^{L} \frac{N_i}{N^2} \sigma_{Y_i}^2$$

dengan syarat $\sum_{i=1}^{L} c_i n_i = C - c_0$

Dengan menggunakan metode pengali Lagrange dengan 1 kendala dimana λ adalah pengali Lagrangenya, diperoleh persamaan :

$$Q(n_1, n_2, ..., n_L) = \sum_{i=1}^{L} \left(\frac{N_i}{N}\right)^2 \frac{\sigma_{Y_i}^2}{n_i} - \sum_{i=1}^{L} \frac{N_i}{N^2} \sigma_{Y_i}^2 + \lambda \left(\sum_{i=1}^{L} c_i n_i - (C - c_0) \right)$$

$$\frac{\partial Q}{\partial n_i} = 0 \quad \Rightarrow \quad -\left(\frac{N_i \sigma_{Y_i}}{N n_i}\right)^2 + \lambda c_i = 0$$

$$\lambda c_i = \left(\frac{N_i \sigma_{Y_i}}{N n_i}\right)^2$$

$$\lambda c_i n_i^2 = \left(\frac{N_i \sigma_{Y_i}}{N}\right)^2$$

$$n_i = \frac{1}{\sqrt{\lambda}} \frac{N_i \sigma_{Y_i}}{N \sqrt{c_i}}$$

Dari hasil diatas bahwa dengan meminimumkan variansi dengan biaya konstan maka alokasi sampel pada stratum ke-*i* yaitu $n_i \sim \frac{N_i \sigma_{Y_i}}{N \sqrt{c_i}}$.

2.2.3.2 Alokasi Sampel Dengan Meminimumkan Biaya Dengan Variansi Konstan

Akan dicari alokasi sampel dengan meminimumkan :

$$C = c_0 + \sum_{i=1}^{L} c_i n_i$$

dengan syarat $Var\left(\bar{y}_{st}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\frac{\sigma_{Y_i}^2}{n_i}-\sum_{i=1}^{L}\frac{N_i}{N^2}\sigma_{Y_i}^2$

Dengan menggunakan metode pengali Lagrange dengan 1 kendala dimana γ adalah pengali Lagrangenya, diperoleh persamaan :

$$P\left(n_1,n_2,...,n_L\right)=c_0+\sum_{i=1}^{L}c_in_i+\gamma\left(\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\frac{\sigma_{Y_i}^2}{n_i}-\sum_{i=1}^{L}\frac{N_i}{N^2}\sigma_{Y_i}^2\right)$$

$$\frac{\partial P}{\partial n_i}=0 \quad\Rightarrow\quad c_i-\gamma\left(\frac{N_i\sigma_{Y_i}}{Nn_i}\right)^2=0$$

$$c_i=\gamma\left(\frac{N_i\sigma_{Y_i}}{Nn_i}\right)^2$$

$$c_in_i^2=\gamma\left(\frac{N_i\sigma_{Y_i}}{N}\right)^2$$

$$n_i=\sqrt{\gamma}\,\frac{N_i\sigma_{Y_i}}{N\sqrt{c_i}}$$

Dari hasil diatas bahwa dengan meminimumkan biaya dengan variansi konstan maka alokasi sampel pada stratum ke-*i* yaitu $n_i\sim\frac{N_i\sigma_i}{\sqrt{c_i}}$.

Dengan menggunakan kedua cara diatas ternyata pada setiap stratum ke-*i*, $n_i\sim\frac{N_i\sigma_{Y_i}}{\sqrt{c_i}}$. Hal ini menjelaskan uraian sebelumnya mengenai 3 faktor yang mempengaruhi alokasi sampel yang baik.

Selanjutnya akan dibuktikan bahwa ukuran sampel yang optimal pada setiap stratum jika banyaknya sampel yang dipilih *n* adalah :

$$n_i=n\left(\frac{N_i\sigma_{Y_i}\Big/\sqrt{c_i}}{\sum_{i=1}^{L}N_i\sigma_{Y_i}\Big/\sqrt{c_i}}\right)$$

Bukti :

Dari sebelumnya diketahui bahwa : $n_i\sim\frac{N_i\sigma_{Y_i}}{\sqrt{c_i}}$

Berarti :

$$n_i = k\frac{N_i \sigma_{Y_i}}{\sqrt{c_i}} \qquad ...(2.2.3.1)$$

untuk suatu konstanta *k*.

Karena $C = c_0 + \sum_{i=1}^{L} c_i n_i$, maka :

$$C = c_0 + k\sum_{i=1}^{L} c_i \left(\frac{N_i \sigma_{Y_i}}{\sqrt{c_i}}\right)$$

$$C - c_0 = k\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i}$$

$$k = \frac{C - C_0}{\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i}} \qquad …(2.2.3.2)$$

Kemudian substitusikan persamaan (2.2.3.2) ke dalam (2.2.3.1) menjadi :

$$n_i = \frac{C - C_0}{\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i}} \frac{N_i \sigma_{Y_i}}{\sqrt{c_i}} \qquad ...(2.2.3.3)$$

Substitusikan persamaan (2.2.3.3) ke dalam persamaan $n = \sum_{i=1}^{L} n_i$ menjadi :

$$n = \sum_{i=1}^{L} \left(\frac{C - C_0}{\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i}} \frac{N_i \sigma_{Y_i}}{\sqrt{c_i}} \right)$$

$$n = \frac{C - C_0}{\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i}} \sum_{i=1}^{L} \frac{N_i \sigma_{Y_i}}{\sqrt{c_i}}$$

$$C - C_0 = n\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i} \Bigg/ \sum_{i=1}^{L} \frac{N_i \sigma_{Y_i}}{\sqrt{c_i}} \qquad …(2.2.3.4)$$

Substitusikan persamaan (2.2.3.4) ke dalam (2.2.3.3) menjadi :

$$n_i = \frac{n\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i}}{\sum_{i=1}^{L} \frac{N_i \sigma_{Y_i}}{\sqrt{c_i}}} \frac{1}{\sum_{i=1}^{L} \sqrt{c_i} N_i \sigma_{Y_i}} \frac{N_i \sigma_{Y_i}}{\sqrt{c_i}}$$

$$n_i = n\left(\frac{N_i\sigma_{Y_i}\Big/\sqrt{c_i}}{\sum_{i=1}^{L} N_i\sigma_{Y_i}\Big/\sqrt{c_i}}\right)$$

(Terbukti)

Apabila biaya dianggap sama, yaitu $c_1 = c_2 = \ldots = c_L$, maka $n_i = n\left(n_i S_{Y_i}\Big/\sum_{i=1}^{L} N_i S_{Y_i}\right)$. Metode pemilihan seperti ini disebut **Alokasi Neyman**.

Apabila biaya dan variansinya dianggap sama, yaitu $c_1 = c_2 = \ldots = c_L$ dan $\sigma_{Y_1}^2 = \sigma_{Y_2}^2 = \ldots = \sigma_{Y_L}^2$, maka $n_i = n\left(N_i\Big/\sum_{i=1}^{L} N_i\right)$ atau $n_i = n\left(\frac{N_i}{N}\right)$. Metode pemilihan seperti ini disebut **Alokasi Proporsional**.

2.3 TAKSIRAN REGRESI PADA SIMPLE RANDOM SAMPLING

2.3.1 Pendahuluan

Dalam suatu penelitian mean populasi dari Y sebut μ_Y biasanya ditaksir hanya dengan menggunakan data dari variabel Y. Tetapi jika terdapat variabel tambahan, sebut variabel X yang berkorelasi kuat dan berhubungan secara linier dengan Y maka informasi ini dapat digunakan untuk mendapatkan taksiran μ_Y yang lebih baik.

Salah satu metode untuk mencari taksiran μ_Y dengan memperhitungkan hubungan linier antara variabel Y dengan variabel X adalah dengan **taksiran regresi**. Dalam bagian ini akan dibahas mengenai taksiran regresi untuk mean populasi pada Simple Random Sampling (SRS).

Misalkan dipunyai variabel Y dan X dimana hubungan antara variabel Y dengan variabel X dapat dituliskan dalam model regresi linier sebagai berikut :

$$y = \beta_0 + \beta_1 x + \varepsilon$$

dimana β_0 dan β_1 adalah parameter yang tidak diketahui dan ε adalah variabel errornya.

Parameter β_0 dan β_1 dapat ditaksir menggunakan metode *Least Squares* dengan meminimumkan jumlah kuadratik errornya berdasarkan suatu sampel ukuran *n* yang diambil secara SRS.

Sebut :

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

adalah persamaan jumlah kuadratik error dari model regresi diatas.

Sebut $\widehat{\beta}_0$ adalah taksiran dari β_0 dan $\widehat{\beta}_1$ adalah taksiran dari β_1.

Untuk mendapatkan taksiran Least Squares dari parameter β_0 dan β_1 adalah dengan menyelesaikan solusi dari :

$$\left.\frac{\partial S}{\partial \beta_0}\right|_{\widehat{\beta}_0, \widehat{\beta}_1} = 0 \quad \text{dan} \quad \left.\frac{\partial S}{\partial \beta_1}\right|_{\widehat{\beta}_0, \widehat{\beta}_1} = 0$$

$$\left.\frac{\partial S}{\partial \beta_0}\right|_{\widehat{\beta}_0, \widehat{\beta}_1} = 0$$

$$-2\sum_{i=1}^{n}\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) = 0$$

$$\sum_{i=1}^{n}\left(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i\right) = 0$$

$$\sum_{i=1}^{n} y_i - n\widehat{\beta}_0 - \widehat{\beta}_1 \sum_{i=1}^{n} x_i = 0$$

$$n\widehat{\beta}_0 = \sum_{i=1}^{n} y_i - \widehat{\beta}_1 \sum_{i=1}^{n} x_i$$

$$\widehat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} y_i - \widehat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \overline{x}$$

dimana $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ dan $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$$\left.\frac{\partial S}{\partial \beta_1}\right|_{\hat\beta_0,\hat\beta_1} = 0$$

$$-2\sum_{i=1}^{n}\left(y_i - \hat\beta_0 - \hat\beta_1 x_i\right)x_i = 0$$

$$\sum_{i=1}^{n}\left(y_i - \hat\beta_0 - \hat\beta_1 x_i\right)x_i = 0$$

$$\sum_{i=1}^{n} x_i y_i - \hat\beta_0\sum_{i=1}^{n} x_i - \hat\beta_1\sum_{i=1}^{n} x_i^2 = 0$$

$$\sum_{i=1}^{n} x_i y_i - \left(\bar y - \hat\beta_1\bar x\right)\sum_{i=1}^{n} x_i - \hat\beta_1\sum_{i=1}^{n} x_i^2 = 0$$

$$\sum_{i=1}^{n} x_i y_i - \bar y\sum_{i=1}^{n} x_i + \hat\beta_1\left(\bar x\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i^2\right) = 0$$

$$\hat\beta_1 = \frac{\bar y\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i}{\bar x\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i y_i - \bar y\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \bar x\sum_{i=1}^{n} x_i}$$

$$\hat\beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar x\bar y}{\sum_{i=1}^{n} x_i^2 - n\bar x^2} = \frac{\sum_{i=1}^{n}(y_i - \bar y)(x_i - \bar x)}{\sum_{i=1}^{n}(x_i - \bar x)^2}$$

Misalkan $s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar x)^2$ adalah taksiran variansi dari *X* dan $s_{YX} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar y)(x_i - \bar x)$ adalah taksiran kovariansi dari *Y* dan *X* dari simple random sample berukuran *n*, maka :

$$\hat\beta_1 = \frac{s_{YX}}{s_X^2}$$

Dengan demikian diperoleh taksiran Least Square untuk parameter β_0 dan β_1, yaitu :

$$\hat\beta_0 = \bar y - \hat\beta_1\bar x \text{ dan } \hat\beta_1 = \frac{\sum_{i=1}^{n}(y_i - \bar y)(x_i - \bar x)}{\sum_{i=1}^{n}(x_i - \bar x)^2} = \frac{s_{YX}}{s_X^2}$$

2.3.2 Taksiran Regresi untuk Mean Populasi dan Variansinya

Didefinisikan taksiran regresi untuk μ_Y, yaitu :

$$\widehat{\mu}_{Yr} = \widehat{\beta}_0 + \widehat{\beta}_1 \mu_X \qquad \ldots(2.3.2.1)$$

dengan,

$$\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \overline{x} \qquad \ldots(2.3.2.2)$$

dan

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n}(y_i x_i - y_i \overline{x} - x_i \overline{y} + \overline{yx})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n} y_i x_i - \overline{x}\sum_{i=1}^{n} y_i - \overline{y}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \overline{yx}}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \overline{x}\sum_{i=1}^{n} y_i - n\overline{y}\frac{\sum_{i=1}^{n} x_i}{n} + n\overline{yx}}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n} y_i x_i - \overline{x}\sum_{i=1}^{n} y_i - n\overline{yx} + n\overline{yx}}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n} y_i x_i - \overline{x}\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i (x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad \ldots(2.3.2.3)$$

atau dengan perkataan lain taksiran regresi untuk μ_Y dapat dituliskan sebagai :

$$\widehat{\mu}_{Yr} = \overline{y} + \widehat{\beta}_1 (\mu_X - \overline{x}) \qquad \ldots(2.3.2.4)$$

Untuk selanjutnya akan dibuktikan bahwa :

1. $\widehat{\mu}_{Yr}$ adalah taksiran yang bias untuk μ_Y.
2. Jika *n* besar maka $\widehat{\mu}_{Yr} \approx \mu_Y$.
3. Untuk sampel yang besar, variansi dari $\widehat{\mu}_{Yr}$ adalah $Var(\widehat{\mu}_{Yr}) \approx \frac{N-n}{nN} S_Y^2 (1-\rho^2)$ dimana $S_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu_Y)^2$, $S_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_X)^2$, $S_{YX} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu_Y)(x_i - \mu_X)$ dan

$\rho = \dfrac{S_{YX}}{S_Y S_X}$ adalah berturut – turut adalah variansi populasi dari *Y*, *X*, kovariansi antara *Y* dan *X*, dan koefisien korelasi populasi antara *Y* dan *X*.

4. Misalkan didefinisikan taksiran variansi dari $\hat{\mu}_{Yr}$ adalah :

$$\widehat{Var}\left(\hat{\mu}_{Yr}\right) = \frac{N-n}{nN(n-2)} \sum_{i=1}^{n} \left((y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x}) \right)^2$$

maka dapat dibuktikan bahwa $E\left(\widehat{Var}\left(\hat{\mu}_{Yr}\right)\right) \approx Var\left(\hat{\mu}_{Yr}\right)$ untuk sampel yang besar.

Pembuktian :

1. Untuk membuktikan bahwa $\hat{\mu}_{Yr}$ adalah taksiran yang bias untuk μ_Y adalah dengan menunjukkan bahwa $E\left(\hat{\mu}_{Yr}\right) \neq \mu_Y$ atau $E\left(\hat{\mu}_{Yr} - \mu_Y\right) \neq 0$

Bukti :

$$E\left(\hat{\mu}_{Yr}\right) = E\left(\bar{y} + \hat{\beta}_1 (\mu_X - \bar{x})\right)$$

$$E\left(\hat{\mu}_{Y}\right) = E(\bar{y}) + E\left(\hat{\beta}_1 (\mu_X - \bar{x})\right)$$

Dalam SRS $\bar{y}$ adalah taksiran yang tidak bias untuk μ_Y atau $E(\bar{y}) = \mu_Y$ dan $\bar{x}$ adalah taksiran yang tidak bias untuk μ_X atau $E(\bar{x}) = \mu_X$, maka

$$E\left(\hat{\mu}_{Yr}\right) = \mu_Y + E\left(\hat{\beta}_1 (\mu_X - \bar{x})\right)$$

Kemudian,

$$E\left(\hat{\mu}_{Yr}\right) - \mu_Y = E\left(\hat{\beta}_1 (\mu_X - \bar{x})\right)$$

$$E\left(\hat{\mu}_{Yr} - \mu_Y\right) = E\left(\hat{\beta}_1 (\mu_X - \bar{x})\right)$$

$$E\left(\hat{\mu}_{Yr} - \mu_Y\right) = -E\left(\hat{\beta}_1 (\bar{x} - \mu_X)\right)$$

$$E\left(\hat{\mu}_{Yr} - \mu_Y\right) = -E\left(\hat{\beta}_1 \bar{x} - \hat{\beta}_1 \mu_X\right)$$

$$E\left(\hat{\mu}_{Yr} - \mu_Y\right) = -\left(E\left(\hat{\beta}_1 \bar{x}\right) - E\left(\hat{\beta}_1 \mu_X\right)\right)$$

Jika μ_X diketahui, maka

$$E\left(\widehat{\mu}_{Yr}-\mu_Y\right)=-\left(E\left(\widehat{\beta}_1\bar{x}\right)-\mu_X E\left(\widehat{\beta}_1\right)\right)$$

$$E\left(\widehat{\mu}_{Yr}-\mu_Y\right)=-\left(E\left(\widehat{\beta}_1\bar{x}\right)-E\left(\bar{x}\right)E\left(\widehat{\beta}_1\right)\right)$$

$$E\left(\widehat{\mu}_{Yr}-\mu_Y\right)=-\left(E\left(\widehat{\beta}_1\bar{x}\right)-E\left(\widehat{\beta}_1\right)E\left(\bar{x}\right)\right)$$

Berdasarkan definisi dari kovariansi antara 2 variabel random, maka

$$E\left(\widehat{\mu}_{Yr}-\mu_Y\right)=-Cov\left(\widehat{\beta}_1,\bar{x}\right)$$

Karena $E\left(\widehat{\mu}_{Yr}-\mu_Y\right)=-Cov\left(\widehat{\beta}_1,\bar{x}\right)$ dimana $\widehat{\beta}_1=\dfrac{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)\left(x_i-\bar{x}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$

bergantung dengan $\bar{x}$, maka $-Cov\left(\widehat{\beta}_1,\bar{x}\right)\neq 0$ atau dengan perkataan lain

$$E\left(\widehat{\mu}_{Yr}-\mu_Y\right)=-Cov\left(\widehat{\beta}_1,\bar{x}\right)\neq 0.$$

Dengan demikian terbukti bahwa $\widehat{\mu}_{Yr}$ adalah taksiran yang bias untuk μ_Y dengan bias sebesar $-Cov\left(\widehat{\beta}_1,\bar{x}\right)$.

(Terbukti)

2. Berikutnya akan dibuktikan bahwa untuk *n* yang besar maka $\widehat{\mu}_{Yr}\approx\mu_Y$

Bukti :

Diketahui bahwa taksiran regresi untuk μ_Y adalah $\widehat{\mu}_{Yr}$, maka untuk membuktikan bahwa untuk *n* yang besar , $\widehat{\mu}_{Yr}\approx\mu_Y$ adalah dengan menunjukkan bahwa $\bar{y}\approx\mu_Y$ dan $\widehat{\beta}_1\left(\mu_X-\bar{x}\right)\approx 0$.

Pertama akan dibuktikan terlebih dahulu bahwa untuk *n* yang besar $\bar{y}\approx\mu_Y$ yaitu menunjukkan bahwa $\bar{y}$ konvergen probabilitas ke μ_Y dengan menunjukkan bahwa $\forall\varepsilon>0$, $\lim\limits_{n\to\infty}\Pr\left(\left|\bar{y}-\mu_Y\right|<\varepsilon\right)=1$ atau $\lim\limits_{n\to\infty}\Pr\left(\left|\bar{y}-\mu_Y\right|\geq\varepsilon\right)=0$ berlaku.

Untuk membuktikan ini digunakan ketaksamaan Chebyshev yaitu :

$$\Pr\left(\left|\bar{y}-E\left(\bar{y}\right)\right|\geq\varepsilon\right)=\Pr\left(\left|\bar{y}-E\left(\bar{y}\right)\right|\geq k\sqrt{Var\left(\bar{y}\right)}\right)\leq\frac{1}{k^2}$$

Ambil sembarang $\varepsilon > 0$.

Dalam SRS $E(\bar{y}) = \mu_Y$ dan $Var(\bar{y}) = \frac{N-n}{nN} S_Y^2$, maka

$$\Pr\left(\left|\bar{y} - E(\bar{y})\right| \geq k\sqrt{Var(\bar{y})}\right) = \Pr\left(\left|\bar{y} - \mu_Y\right| \geq kS_Y\sqrt{\frac{N-n}{nN}}\right)$$
$$\leq \frac{1}{k^2}$$

dimana $k = \frac{\varepsilon}{S_Y}\sqrt{\frac{nN}{N-n}}$ atau $\frac{1}{k^2} = \frac{S_Y^2}{\varepsilon^2}\frac{N-n}{nN}$

Dengan demikian,

$$\underset{n\to\infty}{Lim}\Pr\left(\left|\bar{y} - \mu_Y\right| \geq kS_Y\sqrt{\frac{N-n}{nN}}\right) \leq \underset{n\to\infty}{Lim}\frac{S_Y^2}{\varepsilon^2}\frac{N-n}{nN}$$
$$= \frac{S_Y^2}{\varepsilon^2}\underset{n\to\infty}{Lim}\left(\frac{1}{n} - \frac{1}{N}\right)$$
$$= 0$$

Karena $\underset{n\to\infty}{Lim}\Pr\left(\left|\bar{y} - \mu_Y\right| \geq kS_Y\sqrt{\frac{N-n}{nN}}\right) = 0$, maka terbukti bahwa $\bar{y}$ konvergen probabitas ke μ_Y.

Dengan demikian untuk *n* yang besar, $\bar{y} \approx \mu_Y$.

Selanjutnya akan dibuktikan bahwa $\hat{\beta}_1(\mu_X - \bar{x}) \approx 0$ dengan cara menunjukkan bahwa $\mu_X - \bar{x} \approx 0$.

Untuk membuktikan hal diatas adalah dengan menunjukkan bahwa $\bar{x}$ konvergen probabilitas ke μ_X dengan cara menunjukkan bahwa $\forall \varepsilon > 0$, $\underset{n\to\infty}{Lim}\Pr\left(\left|\bar{x} - \mu_X\right| < \varepsilon\right) = 1$ atau $\underset{n\to\infty}{Lim}\Pr\left(\left|\bar{x} - \mu_X\right| \geq \varepsilon\right) = 0$ berlaku.

Untuk membuktikan ini digunakan ketaksamaan Chebyshev yaitu :

$$\Pr\left(\left|\bar{x} - E(\bar{x})\right| \geq \varepsilon\right) = \Pr\left(\left|\bar{x} - E(\bar{x})\right| \geq k\sqrt{Var(\bar{y})}\right) \leq \frac{1}{k^2}$$

Ambil sembarang $\varepsilon > 0$.

Dalam SRS $E(\bar{x}) = \mu_X$ dan $Var(\bar{x}) = \frac{N-n}{nN} S_X^2$, maka

$$\Pr\left(\left|\bar{x}-E\left(\bar{x}\right)\right|\geq k\sqrt{Var\left(\bar{x}\right)}\right)=\Pr\left(\left|\bar{x}-\mu_X\right|\geq kS_X\sqrt{\frac{N-n}{nN}}\right)$$
$$\leq\frac{1}{k^2}$$

dimana $k=\dfrac{\varepsilon}{S_X}\sqrt{\dfrac{nN}{N-n}}$ atau $\dfrac{1}{k^2}=\dfrac{S_X^2}{\varepsilon^2}\dfrac{N-n}{nN}$

Dengan demikian,

$$\lim_{n\to\infty}\Pr\left(\left|\bar{x}-\mu_X\right|\geq kS_X\sqrt{\frac{N-n}{nN}}\right)\leq\lim_{n\to\infty}\frac{S_X^2}{\varepsilon^2}\frac{N-n}{nN}$$
$$=\frac{S_X^2}{\varepsilon^2}\lim_{n\to\infty}\left(\frac{1}{n}-\frac{1}{N}\right)$$
$$=0$$

Karena $\lim_{n\to\infty}\Pr\left(\left|\bar{x}-\mu_X\right|\geq kS_X\sqrt{\dfrac{N-n}{nN}}\right)=0$, maka terbukti bahwa $\bar{x}$ konvergen probabitas ke μ_X.

Dengan demikian untuk *n* yang besar, $\bar{x}\approx\mu_X$ atau $\mu_X-\bar{x}\approx 0$.

Karena $\mu_X-\bar{x}\approx 0$, maka $\hat{\beta}_1\left(\mu_X-\bar{x}\right)\approx 0$.

Karena telah dibuktikan bahwa untuk *n* yang besar $\bar{y}\approx\mu_Y$ dan $\hat{\beta}_1\left(\mu_X-\bar{x}\right)\approx 0$, maka

$$\hat{\mu}_{Yr}=\bar{y}+\hat{\beta}_1\left(\mu_X-\bar{x}\right)$$
$$\hat{\mu}_{Yr}\approx\mu_Y$$

Dengan demikian terbukti bahwa untuk *n* yang besar maka $\hat{\mu}_Y\approx\mu_Y$.

(Terbukti)

3. Selanjutnya akan dibuktikan bahwa untuk sampel yang besar $Var\left(\hat{\mu}_{Yr}\right)\approx\dfrac{N-n}{nN}S_Y^2\left(1-\rho^2\right)$ dimana $S_Y^2=\dfrac{1}{N-1}\sum_{i=1}^{N}\left(y_i-\mu_Y\right)^2$, $S_X^2=\dfrac{1}{N-1}\sum_{i=1}^{N}\left(x_i-\mu_X\right)^2$, $S_{YX}=\dfrac{1}{N-1}\sum_{i=1}^{N}\left(y_i-\mu_Y\right)\left(x_i-\mu_X\right)$

dan $\rho = \dfrac{S_{YX}}{S_Y S_X}$ adalah berturut – turut adalah variansi populasi dari *Y*, *X*, kovariansi antara *Y* dan *X*, dan koefisien korelasi populasi antara *Y* dan *X*.

Bukti :

Sebelum membuktikan hal diatas, pertama definisikan terlebih dahulu

$$y_i = \mu_Y + \beta_1^* \left(x_i - \mu_X \right) + e_i \qquad \text{…(2.3.2.5)}$$

atau

$$e_i = y_i - \mu_Y - \beta_1^* \left(x_i - \mu_X \right) \qquad \text{…(2.3.2.6)}$$

untuk setiap elemen populasi, $i = 1, 2, ..., N$ dengan

$$\beta_1^* = \frac{\sum_{i=1}^{N} \left(y_i - \mu_Y \right)\left(x_i - \mu_X \right)}{\sum_{i=1}^{N} \left(x_i - \mu_X \right)^2} \qquad \text{…(2.3.2.7)}$$

Akan ditunjukkan bahwa $E\left(\widehat{\beta}_1\right) = \beta_1^*$ dimana $\widehat{\beta}_1 = \dfrac{\sum_{i=1}^{n} y_i \left(x_i - \bar{x} \right)}{\sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2}$.

Substitusikan persamaan (2.3.2.5) ke dalam persamaan (2.3.2.3) menjadi :

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(\mu_Y + \beta_1^* \left(x_i - \mu_X \right) + e_i \right)\left(x_i - \bar{x} \right)}{\sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(\mu_Y \left(x_i - \bar{x} \right) + \beta_1^* \left(x_i - \mu_X \right)\left(x_i - \bar{x} \right) + e_i \left(x_i - \bar{x} \right) \right)}{\sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(\mu_Y \left(x_i - \bar{x} \right) + \beta_1^* \left(x_i - \bar{x} + \bar{x} - \mu_X \right)\left(x_i - \bar{x} \right) + e_i \left(x_i - \bar{x} \right) \right)}{\sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n} \left(\mu_Y \left(x_i - \bar{x} \right) + \beta_1^* \left(\left(x_i - \bar{x} \right) + \left(\bar{x} - \mu_X \right) \right)\left(x_i - \bar{x} \right) + e_i \left(x_i - \bar{x} \right) \right)}{\sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}\left(\mu_Y\left(x_i-\bar{x}\right)+\beta_1^*\left(x_i-\bar{x}\right)^2+\beta_1^*\left(\bar{x}-\mu_X\right)\left(x_i-\bar{x}\right)+e_i\left(x_i-\bar{x}\right)\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}\left(\mu_Y\left(x_i-\bar{x}\right)+\beta_1^*\left(x_i-\bar{x}\right)^2\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}+\frac{\beta_1^*\sum_{i=1}^{n}\left(\bar{x}-\mu_X\right)\left(x_i-\bar{x}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}+\frac{\sum_{i=1}^{n}e_i\left(x_i-\bar{x}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$$

$$\hat{\beta}_1 = \frac{\mu_Y\sum_{i=1}^{n}\left(x_i-\bar{x}\right)+\beta_1^*\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2+\beta_1^*\sum_{i=1}^{n}\left(\bar{x}-\mu_X\right)\left(x_i-\bar{x}\right)+\sum_{i=1}^{n}e_i\left(x_i-\bar{x}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$$

$$\hat{\beta}_1 = \beta_1^* + \frac{\sum_{i=1}^{n}e_i\left(x_i-\bar{x}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2} \qquad \ldots(2.3.2.8)$$

Misalkan $c_i = \dfrac{x_i-\bar{x}}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$ untuk $i = 1,2,\ldots,n$, maka

$$E\left(\hat{\beta}_1\right) = E\left(\beta_1^* + \sum_{i=1}^{n} y_i e_i\right)$$

$$E\left(\hat{\beta}_1\right) = E\left(\beta_1^*\right) + E\left(\sum_{i=1}^{n} c_i e_i\right)$$

$$E\left(\hat{\beta}_1\right) = \beta_1^* + \sum_{i=1}^{n} c_i E\left(e_i\right) \qquad \ldots(2.3.2.9)$$

Sebut μ_e adalah mean populasi *e* dimana $\mu_e = \frac{1}{N}\sum_{i=1}^{N} e_i$, maka

$$E\left(e_i\right) = \sum_{i=1}^{N} e_i p\left(e_i\right)$$

$$E\left(e_i\right) = \sum_{i=1}^{N} e_i \frac{1}{N}$$

$$E\left(e_i\right) = \frac{1}{N}\sum_{i=1}^{N} e_i$$

$$E\left(e_i\right) = \mu_e$$

Karena $E\left(e_i\right) = \mu_e$, maka

$E\left(\hat{\beta}_1\right) = \beta_1^* + \sum_{i=1}^{n} c_i \mu_e$

$E\left(\hat{\beta}_1\right) = \beta_1^* + \mu_e \sum_{i=1}^{n} c_i$

Karena $\sum_{i=1}^{n} c_i = \frac{\sum_{i=1}^{n}(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{n\bar{x} - n\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 0$, maka

$E\left(\hat{\beta}_1\right) = \beta_1^*$

Dengan demikian terbukti bahwa $E\left(\hat{\beta}_1\right) = \beta_1^*$.

Selanjutnya misalkan diambil simple random sample $e_i = y_i - \mu_Y - \beta_1^*(x_i - \mu_X)$ ukuran n.

Definisikan :

$\bar{e} = \frac{1}{n}\sum_{i=1}^{n} e_i$

$\bar{e} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \mu_Y - \beta_1^*(x_i - \mu_X)\right)$

$\bar{e} = \bar{y} - \mu_Y - \beta_1^*(\bar{x} - \mu_X)$ …(2.3.2.10)

atau

$\bar{y} = \bar{e} + \mu_Y + \beta_1^*(\bar{x} - \mu_X)$ …(2.3.2.11)

Maka, ekspektasi dari $\bar{e}$ adalah sebagai berikut :

$E(\bar{e}) = E\left(\bar{y} - \mu_Y - \beta_1^*(\bar{x} - \mu_X)\right)$

$E(\bar{e}) = E(\bar{y}) - E(\mu_Y) - \beta_1^*\left(E(\bar{x}) - E(\mu_X)\right)$

$E(\bar{e}) = \mu_Y - \mu_Y - \beta_1^*(\mu_X - \mu_X)$

$E(\bar{e}) = 0$ …(2.3.2.12)

Karena sampel diambil secara SRS maka $\bar{e}$ adalah taksiran yang tidak bias untuk μ_e atau dengan perkataan lain,

$E(\bar{e}) = \mu_e$ …(2.3.2.13)

Substitusikan persamaan (2.3.2.13) ke dalam (2.3.2.12) sehingga diperoleh

$\mu_e = 0$ …(2.3.2.14)

Kemudian substitusikan persamaan (2.3.2.11) ke dalam (2.3.2.2) sehingga diperoleh

$$\widehat{\mu}_{Yr} = \left(\overline{e} + \mu_Y + \beta_1^*\left(\overline{x} - \mu_X\right)\right) + \widehat{\beta}_1\left(\mu_X - \overline{x}\right)$$

$$\widehat{\mu}_{Yr} = \mu_Y - \beta_1^*\left(\mu_X - \overline{x}\right) + \widehat{\beta}_1\left(\mu_X - \overline{x}\right) + \overline{e}$$

$$\widehat{\mu}_{Yr} = \mu_Y + \left(\widehat{\beta}_1 - \beta_1^*\right)\left(\mu_X - \overline{x}\right) + \overline{e} \qquad \ldots(2.3.2.15)$$

Bentuk persamaan (2.3.2.15) inilah yang nantinya akan digunakan untuk mencari $Var\left(\widehat{\mu}_{Yr}\right)$.

Berdasarkan langkah – langkah yang telah dibuktikan sebelumnya, maka selanjutnya akan dibuktikan terlebih dahulu bahwa $\widehat{\beta}_1$ konvergen probabilitas ke β_1^* dengan menunjukkan bahwa $\forall\varepsilon > 0$, $\lim\limits_{n\to\infty} \Pr\left(\left|\widehat{\beta}_1 - \beta_1^*\right| < \varepsilon\right) = 1$ atau $\lim\limits_{n\to\infty} \Pr\left(\left|\widehat{\beta}_1 - \beta_1^*\right| \geq \varepsilon\right) = 0$.

Sebelum membuktikan hal diatas, maka akan dicari terlebih dahulu $Var\left(\widehat{\beta}_1\right)$.

Dengan cara yang sama dari sebelumnya yaitu membuktikan $E\left(\widehat{\beta}_1\right) = \beta_1^*$, maka diperoleh

$$Var\left(\widehat{\beta}_1\right) = Var\left(\frac{\sum_{i=1}^{n} y_i\left(x_i - \overline{x}\right)}{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}\right)$$

$$Var\left(\widehat{\beta}_1\right) = Var\left(\sum_{i=1}^{n} c_i y_i\right)$$

$$Var\left(\widehat{\beta}_1\right) = \sum_{i=1}^{n} c_i^2 Var\left(y_i\right)$$

$$Var\left(\widehat{\beta}_1\right) = \sum_{i=1}^{n} c_i^2\left[\left(E\left(y_i - E\left(y_i\right)\right)^2\right)\right]$$

$$Var\left(\widehat{\beta}_1\right) = \sum_{i=1}^{n} c_i^2\left[\left(E\left(y_i - \mu_Y\right)^2\right)\right]$$

$$Var\left(\widehat{\beta}_1\right) = \sum_{i=1}^{n} c_i^2\left[\left(E\left(y_i^2 - 2\mu_Y y_i + \mu_Y^2\right)\right)\right]$$

$$Var\left(\widehat{\beta}_1\right) = \sum_{i=1}^{n} c_i^2\left[\left(E\left(y_i^2\right) - 2\mu_Y E\left(y_i\right) + E\left(\mu_Y^2\right)\right)\right]$$

$$Var\left(\widehat{\beta}_1\right)=\sum_{i=1}^{n}c_i^2\left[\left(E\left(y_i^2\right)-2\mu_Y\mu_Y+\mu_Y^2\right)\right]$$

$$Var\left(\widehat{\beta}_1\right)=\sum_{i=1}^{n}c_i^2\left[\left(E\left(y_i^2\right)-2\mu_Y^2+\mu_Y^2\right)\right]$$

$$Var\left(\widehat{\beta}_1\right)=\sum_{i=1}^{n}c_i^2\left[\left(E\left(y_i^2\right)-\mu_Y^2\right)\right]$$

$$Var\left(\widehat{\beta}_1\right)=\sum_{i=1}^{n}c_i^2\left[\left(\sum_{i=1}^{N}y_i^2p\left(y_i\right)-\mu_Y^2\right)\right]$$

$$Var\left(\widehat{\beta}_1\right)=\sum_{i=1}^{n}c_i^2\left[\frac{1}{N}\sum_{i=1}^{N}y_i^2-\mu_Y^2\right] \qquad \ldots(2.3.2.16)$$

Misalkan definisikan $\sigma_Y^2=\frac{1}{N}\sum_{i=1}^{N}\left(y_i-\mu_Y\right)^2$ atau $\sigma_Y^2=\frac{1}{N}\sum_{i=1}^{N}y_i^2-\mu_Y^2$, maka persamaan (2.3.2.16) menjadi

$$Var\left(\widehat{\beta}_1\right)=\sum_{i=1}^{n}c_i^2\sigma_Y^2$$

$$Var\left(\widehat{\beta}_1\right)=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{\left(\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)^2}\sigma_Y^2$$

$$Var\left(\widehat{\beta}_1\right)=\frac{\sigma_Y^2}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2} \qquad \ldots(2.3.2.17)$$

Berdasarkan definisi dari $s_X^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2$ atau $\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2=\left(n-1\right)s_X^2$ maka persamaan (2.3.2.17) menjadi

$$Var\left(\widehat{\beta}_1\right)=\frac{\sigma_Y^2}{\left(n-1\right)s_X^2} \qquad \ldots(2.3.2.18)$$

Setelah memperoleh $Var\left(\widehat{\beta}_1\right)$ yaitu $Var\left(\widehat{\beta}_1\right)=\frac{\sigma_Y^2}{\left(n-1\right)s_X^2}$ maka selanjutnya akan dibuktikan $\widehat{\beta}_1$ konvergen probabilitas ke β_1^* dengan cara menunjukkan bahwa $\forall\varepsilon>0$, $\underset{n\to\infty}{Lim}\Pr\left(\left|\widehat{\beta}_1-\beta_1^*\right|<\varepsilon\right)=1$ atau $\underset{n\to\infty}{Lim}\Pr\left(\left|\widehat{\beta}_1-\beta_1^*\right|\geq\varepsilon\right)=0$.

Ambil sembarang $\varepsilon>0$.

Untuk membuktikan ini digunakan ketaksamaan Chebyshev yaitu,

$$\Pr\left(\left|\hat{\beta}_1 - E\left(\hat{\beta}_1\right)\right| \geq \varepsilon\right) = \Pr\left(\left|\hat{\beta}_1 - E\left(\hat{\beta}_1\right)\right| \geq k\sqrt{Var\left(\hat{\beta}_1\right)}\right)$$
$$= \Pr\left(\left|\hat{\beta}_1 - \beta_1^*\right| \geq k\,\frac{\sigma_Y}{s_X\sqrt{n-1}}\right)$$
$$\leq \frac{1}{k^2}$$

dimana $k = \dfrac{\varepsilon s_X \sqrt{n-1}}{\sigma_Y}$ atau $\dfrac{1}{k^2} = \dfrac{\sigma_Y^2}{\varepsilon^2 s_X^2 (n-1)}$.

Dengan demikian,

$$\lim_{n\to\infty} \Pr\left(\left|\hat{\beta}_1 - \beta_1^*\right| \geq \varepsilon\right) \leq \lim_{n\to\infty} \frac{\sigma_Y^2}{\varepsilon^2 s_X^2 (n-1)}$$
$$= \frac{\sigma_Y^2}{\varepsilon^2 s_X^2} \lim_{n\to\infty} \frac{1}{n-1}$$
$$= 0$$

Karena $\lim_{n\to\infty} \Pr\left(\left|\hat{\beta}_1 - \beta_1^*\right| \geq \varepsilon\right) = 0$, maka terbukti bahwa $\hat{\beta}_1$ konvergen probabilitas ke β_1^*.

Dengan demikian dapat disimpulkan bahwa untuk sampel yang besar $\hat{\beta}_1$ akan mendekati β_1^* atau dengan perkataan lain,

$$\hat{\beta}_1 \approx \beta_1^* \qquad \ldots(2.3.2.19)$$

atau

$$\hat{\beta}_1 - \beta_1^* \approx 0 \qquad \ldots(2.3.2.20)$$

Dari sebelumnya telah dibuktikan bahwa $\bar{x}$ konvergen probabilitas ke μ_X atau dengan perkataan lain $\mu_X - \bar{x} \approx 0$.

Langkah berikutnya adalah substitusikan (2.3.2.20) dan $\mu_X - \bar{x} \approx 0$ ke dalam (2.3.2.15) menjadi

$$\hat{\mu}_{Yr} = \mu_Y + \left(\hat{\beta}_1 - \beta_1^*\right)\left(\mu_X - \bar{x}\right) + \bar{e}$$
$$\hat{\mu}_{Yr} \approx \mu_Y + \bar{e}$$

atau

$$\bar{e} \approx \hat{\mu}_Y - \mu_Y \qquad \ldots(2.3.2.21)$$

Kemudian,

$$E\left(\hat{\mu}_{Yr}\right) \approx E\left(\mu_Y + \bar{e}\right)$$

$$E\left(\hat{\mu}_{Yr}\right) \approx E\left(\mu_Y\right) + E\left(\bar{e}\right)$$

Substitusikan persamaan (2.3.2.12) menjadi :

$$E\left(\hat{\mu}_{Yr}\right) \approx E\left(\mu_Y\right)$$

$$E\left(\hat{\mu}_{Yr}\right) \approx \mu_Y \qquad \ldots(2.3.2.22)$$

Hal ini menjelaskan bahwa untuk sampel yang besar, taksiran regresi $\hat{\mu}_{Yr} = \bar{y} + \hat{\beta}_1\left(\mu_X - \bar{x}\right)$ akan menjadi taksiran yang tidak bias untuk μ_Y.

Berdasarkan informasi – informasi yang telah diperoleh sebelumnya, maka langkah berikutnya akan dibuktikan bahwa untuk sampel yang besar,

$$Var\left(\hat{\mu}_{Yr}\right) \approx \frac{N-n}{nN} S_Y^2\left(1-\rho^2\right)$$

Berdasarkan definisi dari variansi,

$$Var\left(\hat{\mu}_{Yr}\right) = E\left(\left(\hat{\mu}_{Yr} - E\left(\hat{\mu}_{Yr}\right)\right)^2\right)$$

$$Var\left(\hat{\mu}_{Yr}\right) \approx E\left(\left(\hat{\mu}_{Yr} - \mu_Y\right)^2\right) \qquad \ldots(2.3.2.23)$$

Substitusikan persamaan (2.3.2.21) ke persamaan (2.3.2.23) menjadi

$$Var\left(\hat{\mu}_{Yr}\right) \approx E\left(\bar{e}^2\right) \qquad \ldots(2.3.2.24)$$

Substitusikan persamaan (2.3.2.12) ke dalam (2.3.2.24) menjadi

$$Var\left(\hat{\mu}_{Yr}\right) \approx E\left(\bar{e}^2\right) - \left(E\left(\bar{e}\right)\right)^2$$

$$Var\left(\hat{\mu}_{Yr}\right) \approx Var\left(\bar{e}\right) \qquad \ldots(2.3.2.25)$$

Karena sampel diambil secara SRS, maka

$$Var\left(\bar{e}\right) = \left(\frac{N-n}{N}\right)\frac{S_e^2}{n} \qquad \ldots(2.3.2.26)$$

dimana $S_e^2 = \dfrac{1}{N-1}\sum_{i=1}^{N}\left(e_i - \mu_e\right)^2$.

Substitusikan persamaan (2.3.2.26) ke dalam (2.3.2.25) menjadi

$$Var\left(\hat{\mu}_{Yr}\right) \approx \left(\frac{N-n}{N}\right)\frac{S_e^2}{n} \qquad \ldots(2.3.2.27)$$

Langkah berikutnya adalah dengan membuktikan bahwa $S_e^2 = S_Y^2\left(1-\rho^2\right)$ dimana ρ^2 adalah koefisien korelasi populasi antara Y dan X.

Berdasarkan definisi,

$$S_e^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(e_i - \mu_e\right)^2 \qquad \ldots(2.3.2.28)$$

Substitusikan persamaan (2.3.2.16) ke dalam (2.3.2.28) sehingga diperoleh

$$S_e^2 = \frac{1}{N-1}\sum_{i=1}^{N} e_i^2 \qquad \ldots(2.3.2.29)$$

Berikutnya akan dicari bentuk dari $\sum_{i=1}^{N} e_i^2$ dengan $e_i = y_i - \mu_Y - \beta_1^*\left(x_i - \mu_X\right)$.

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N}\left(y_i - \mu_Y - \beta_1^*\left(x_i - \mu_X\right)\right)^2$$

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N}\left(\left(y_i - \mu_Y\right) - \beta_1^*\left(x_i - \mu_X\right)\right)^2$$

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N}\left(\left(y_i - \mu_Y\right)^2 - 2\beta_1^*\left(y_i - \mu_Y\right)\left(x_i - \mu_X\right) + \left(\beta_1^*\right)^2\left(x_i - \mu_X\right)^2\right)$$

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N}\left(y_i - \mu_Y\right)^2 - 2\beta_1^*\sum_{i=1}^{N}\left(y_i - \mu_Y\right)\left(x_i - \mu_X\right) + \left(\beta_1^*\right)^2\sum_{i=1}^{N}\left(x_i - \mu_X\right)^2 \qquad \ldots(2.3.2.30)$$

atau

$$\sum_{i=1}^{N} e_i^2 = \left(N-1\right)S_Y^2 - 2\beta_1^*\left(N-1\right)S_{YX} + \left(\beta_1^*\right)^2\left(N-1\right)S_X^2 \qquad \ldots(2.3.2.31)$$

Karena $\beta_1^* = \dfrac{\sum_{i=1}^{N}\left(y_i - \mu_Y\right)\left(x_i - \mu_X\right)}{\sum_{i=1}^{N}\left(x_i - \mu_X\right)^2}$ atau

$\beta_1^*\sum_{i=1}^{N}\left(x_i - \mu_X\right)^2 = \sum_{i=1}^{N}\left(y_i - \mu_Y\right)\left(x_i - \mu_X\right)$, maka persamaan (2.3.2.30) diatas menjadi :

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N}\left(y_i - \mu_Y\right)^2 - 2\left(\beta_1^*\right)^2\sum_{i=1}^{N}\left(x_i - \mu_X\right)^2 + \left(\beta_1^*\right)^2\sum_{i=1}^{N}\left(x_i - \mu_X\right)^2$$

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} (y_i - \mu_Y)^2 - \left(\beta_1^*\right)^2 \sum_{i=1}^{N} (x_i - \mu_X)^2 \qquad \ldots(2.3.2.32)$$

Kemudian substitusikan persamaan (2.3.2.32) ke dalam (2.3.2.29) menjadi

$$S_e^2 = \frac{1}{N-1}\left(\sum_{i=1}^{N} (y_i - \mu_Y)^2 - \left(\beta_1^*\right)^2 \sum_{i=1}^{N} (x_i - \mu_X)^2 \right)$$

$$S_e^2 = \frac{1}{N-1}\sum_{i=1}^{N} (y_i - \mu_Y)^2 - \left(\beta_1^*\right)^2 \frac{1}{N-1}\sum_{i=1}^{N} (x_i - \mu_X)^2$$

$$S_e^2 = S_Y^2 - \left(\beta_1^*\right)^2 S_X^2 \qquad \ldots(2.3.2.33)$$

Persamaan (2.3.2.33) juga memiliki bentuk lain dengan cara mensubstitusikan persamaan (2.3.2.31) ke dalam (2.3.2.29) menjadi

$$S_e^2 = \frac{1}{N-1}\left((N-1)S_Y^2 - 2\beta_1^*(N-1)S_{YX} + \left(\beta_1^*\right)^2 (N-1)S_X^2 \right)$$

$$S_e^2 = S_Y^2 - 2\beta_1^* S_{YX} + \left(\beta_1^*\right)^2 S_X^2 \qquad \ldots(2.3.2.34)$$

Karena koefisien korelasi populasi antara *Y* dan *X* adalah $\rho = \frac{S_{YX}}{S_X S_Y}$ atau

$S_{YX} = \rho S_X S_Y$ dan $\beta_1^* = \frac{S_{YX}}{S_X^2}$ atau $S_{YX} = \beta_1^* S_X^2$ maka,

$$\left(\beta_1^*\right)^2 = \rho^2 \frac{S_X^2 S_Y^2}{\left(S_X^2\right)^2}$$

$$\left(\beta_1^*\right)^2 = \rho^2 \frac{S_Y^2}{S_X^2} \qquad \ldots(2.3.2.35)$$

Substitusikan persamaan (2.3.2.35) ke dalam (2.3.2.33) menjadi

$$S_e^2 = S_Y^2 - \rho^2 \frac{S_Y^2}{S_X^2} S_X^2$$

$$S_e^2 = S_Y^2 - \rho^2 S_Y^2$$

$$S_e^2 = S_Y^2 \left(1 - \rho^2\right) \qquad \ldots(2.3.2.36)$$

Substitusikan persamaan (2.3.2.36) yang telah diperoleh ke dalam (2.3.2.27) menjadi

$$Var\left(\hat{\mu}_{Yr}\right) \approx \left(\frac{N-n}{nN}\right) S_Y^2 \left(1 - \rho^2\right) \qquad \ldots(2.3.2.37)$$

atau dengan mensubstitusikan persamaan (2.3.2.34) ke persamaan (2.3.2.27) diperoleh bentuk lain dari persamaan (2.3.2.37) yaitu

$$Var\left(\hat{\mu}_{Yr}\right) \approx \left(\frac{N-n}{nN}\right) S_Y^2 - 2\beta_1^* S_{YX} + \left(\beta_1^*\right)^2 S_X^2 \qquad \ldots(2.3.2.38)$$

Dengan demikian diperoleh $Var\left(\hat{\mu}_{Yr}\right) \approx \frac{N-n}{nN} S_Y^2\left(1-\rho^2\right)$ untuk sampel yang besar.

(Terbukti)

4. Berikutnya akan dibuktikan bahwa untuk sampel yang besar,

$$\widehat{Var}\left(\hat{\mu}_{Yr}\right) = \frac{N-n}{nN(n-2)} \sum_{i=1}^{n}\left((y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x})\right)^2$$

adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_Y\right)$.

Bukti :

Untuk membuktikan bahwa dalam sampel yang besar $\widehat{Var}\left(\hat{\mu}_{Yr}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yr}\right)$ adalah dengan menunjukkan bahwa

$E\left(\widehat{Var}\left(\hat{\mu}_{Yr}\right)\right) \approx Var\left(\hat{\mu}_{Yr}\right)$.

Didefinisikan bahwa,

$$\widehat{Var}\left(\hat{\mu}_{Yr}\right) = \frac{N-n}{nN(n-2)} \sum_{i=1}^{n}\left((y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x})\right)^2 \qquad \ldots(2.3.2.39)$$

maka,

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yr}\right)\right) = E\left(\frac{N-n}{nN(n-2)} \sum_{i=1}^{n}\left((y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x})\right)^2\right)$$

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yr}\right)\right) = \frac{N-n}{nN} E\left(\frac{1}{n-2} \sum_{i=1}^{n}\left((y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x})\right)^2\right) \qquad \ldots(2.3.2.40)$$

Akan dibuktikan terlebih dahulu bahwa $\frac{1}{n-2}\sum_{i=1}^{n}\left((y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x})\right)^2$ adalah taksiran yang tidak bias untuk S_e^2 dengan menunjukkan bahwa untuk sampel yang besar,

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2\right)\approx S_e^2$$

Dari sebelumnya diketahui bahwa $e_i = y_i - \mu_Y - \beta_1^*(x_i - \mu_X)$ dan $\bar{e} = \bar{y} - \mu_Y - \beta_1^*(\bar{x} - \mu_X)$, maka

$$e_i - \bar{e} = y_i - \mu_Y - \beta_1^*(x_i - \mu_X) - \left(\bar{y} - \mu_Y - \beta_1^*(\bar{x} - \mu_X)\right)$$

$$e_i - \bar{e} = (y_i - \bar{y}) - \beta_1^*\left((x_i - \mu_X) - (\bar{x} - \mu_X)\right)$$

$$e_i - \bar{e} = (y_i - \bar{y}) - \beta_1^*(x_i - \bar{x})$$

$$e_i - \bar{e} = (y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x}) + \hat{\beta}_1(x_i - \bar{x}) - \beta_1^*(x_i - \bar{x})$$

$$e_i - \bar{e} = (y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x}) + \left(\hat{\beta}_1 - \beta_1^*\right)(x_i - \bar{x})$$

atau apabila ditulis dalam bentuk lain yaitu :

$$(y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x}) = (e_i - \bar{e}) - \left(\hat{\beta}_1 - \beta_1^*\right)(x_i - \bar{x}) \qquad \ldots(2.3.2.41)$$

Berdasarkan persamaan (2.3.2.41) maka bentuk $\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2$ menjadi

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2 = \sum_{i=1}^{n}\left((e_i-\bar{e})-\left(\hat{\beta}_1-\beta_1^*\right)(x_i-\bar{x})\right)^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2 = \sum_{i=1}^{n}(e_i-\bar{e})^2 - 2\left(\hat{\beta}_1-\beta_1^*\right)\sum_{i=1}^{n}(e_i-\bar{e})(x_i-\bar{x}) + \left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2 = \sum_{i=1}^{n}(e_i-\bar{e})^2 - 2\left(\hat{\beta}_1-\beta_1^*\right)\sum_{i=1}^{n}(e_ix_i - e_i\bar{x} - \bar{e}x_i + \overline{ex}) + \left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2 = \sum_{i=1}^{n}(e_i-\bar{e})^2 - 2\left(\hat{\beta}_1-\beta_1^*\right)\left(\sum_{i=1}^{n}e_ix_i - \bar{x}\sum_{i=1}^{n}e_i - \bar{e}\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\overline{ex}\right) + \left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-2\left(\widehat{\beta}_1-\beta_1^*\right)\left(\sum_{i=1}^{n}e_ix_i-\bar{x}\sum_{i=1}^{n}e_i-n\bar{e}\frac{\sum_{i=1}^{n}x_i}{n}+n\overline{ex}\right)$$
$$+\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-2\left(\widehat{\beta}_1-\beta_1^*\right)\left(\sum_{i=1}^{n}e_ix_i-\bar{x}\sum_{i=1}^{n}e_i-n\overline{ex}+n\overline{ex}\right)$$
$$+\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-2\left(\widehat{\beta}_1-\beta_1^*\right)\left(\sum_{i=1}^{n}e_ix_i-\bar{x}\sum_{i=1}^{n}e_i\right)$$
$$+\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-2\left(\widehat{\beta}_1-\beta_1^*\right)\left(\sum_{i=1}^{n}e_i(x_i-\bar{x})\right)$$
$$+\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-2\left(\widehat{\beta}_1-\beta_1^*\right)\left(\sum_{i=1}^{n}e_i(x_i-\bar{x})\right)$$
$$+\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad \ldots(2.3.2.42)$$

Selanjutnya akan dicari terlebih dahulu bentuk lain $\sum_{i=1}^{n}e_i(x_i-\bar{x})$ sehingga persamaan diatas dapat disederhanakan.

Dari sebelumnya diketahui bahwa :

$$\widehat{\beta}_1=\beta_1^*+\frac{\sum_{i=1}^{n}e_i(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \qquad \ldots(2.3.2.43)$$

atau

$$\sum_{i=1}^{n}e_i(x_i-\bar{x})=\left(\widehat{\beta}_1-\beta_1^*\right)\sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad \ldots(2.3.2.44)$$

Substitusikan persamaan (2.3.2.44) ke persamaan (2.3.2.42) menjadi

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-2\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2+\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Selanjutnya,

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\bar{e})^2-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\mu_e-\bar{e}+\mu_e)^2-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}\left((e_i-\mu_e)-(\bar{e}-\mu_e)\right)^2-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}\left((e_i-\mu_e)^2-2(\bar{e}-\mu_e)(e_i-\mu_e)+(\bar{e}-\mu_e)^2\right)$$
$$-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\mu_e)^2-2(\bar{e}-\mu_e)\sum_{i=1}^{n}(e_i-\mu_e)+\sum_{i=1}^{n}(\bar{e}-\mu_e)^2$$
$$-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\mu_e)^2-2(\bar{e}-\mu_e)n(\bar{e}-\mu_e)+n(\bar{e}-\mu_e)^2$$
$$-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\mu_e)^2-2n(\bar{e}-\mu_e)^2+n(\bar{e}-\mu_e)^2$$
$$-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2=\sum_{i=1}^{n}(e_i-\mu_e)^2-n(\bar{e}-\mu_e)^2-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Dengan demikian,

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2\right)=E\left(\frac{1}{n-2}\left(\sum_{i=1}^{n}(e_i-\mu_e)^2-n(\bar{e}-\mu_e)^2\right.\right.$$
$$\left.\left.-\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2\right)\right)$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2\right)=\frac{1}{n-2}\left(E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)-nE\left((\bar{e}-\mu_e)^2\right)\right.$$
$$\left.-E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}(x_i-\bar{x})^2\right)\right)$$

Karena sampel diambil secara SRS, maka

$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=E\left(\sum_{i=1}^{n}\left(e_i^2-2\mu_e e_i+\mu_e^2\right)\right)$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=E\left(\sum_{i=1}^{n}e_i^2\right)-2\mu_e E\left(\sum_{i=1}^{n}e_i\right)+E\left(\sum_{i=1}^{n}\mu_e^2\right)$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=\sum^{N}\sum_{i=1}^{n}e_i^2 p(e_i)-2\mu_e\sum^{N}\sum_{i=1}^{n}e_i p(e_i)+n\mu_e^2$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=\frac{1}{N}n\sum_{i=1}^{N}e_i^2-2\mu_e\frac{1}{N}n\sum_{i=1}^{N}e_i+n\mu_e^2$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=\frac{n}{N}\left(\sum_{i=1}^{N}e_i^2-2\mu_e\sum_{i=1}^{N}e_i+N\mu_e^2\right)$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=\frac{n}{N}\left(\sum_{i=1}^{N}e_i^2-2\mu_e\sum_{i=1}^{N}e_i+\sum_{i=1}^{N}\mu_e^2\right)$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=\frac{n}{N}\sum_{i=1}^{N}\left(e_i^2-2\mu_e e_i+\mu_e^2\right)$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=\frac{n}{N}\sum_{i=1}^{N}(e_i-\mu_e)(e_i-\mu_e)$$
$$E\left(\sum_{i=1}^{n}(e_i-\mu_e)^2\right)=\frac{n}{N}\sum_{i=1}^{N}(e_i-\mu_e)^2 \qquad \ldots(2.3.2.45)$$

dan

$$E\left((\bar{e}-\mu_e)^2\right)=E\left(\bar{e}^2-2\mu_e\bar{e}+\mu_e^2\right)$$
$$E\left((\bar{e}-\mu_e)^2\right)=E\left(\bar{e}^2\right)-2\mu_e E(\bar{e})+E\left(\mu_e^2\right)$$
$$E\left((\bar{e}-\mu_e)^2\right)=E\left(\bar{e}^2\right)-2\mu_e^2+\mu_e^2$$
$$E\left((\bar{e}-\mu_e)^2\right)=E\left(\bar{e}^2\right)-\mu_e^2$$
$$E\left((\bar{e}-\mu_e)^2\right)=E\left(\bar{e}^2\right)-\left(E(\bar{e})\right)^2$$
$$E\left((\bar{e}-\mu_e)^2\right)=Var(\bar{e})$$
$$E\left((\bar{e}-\mu_e)^2\right)=\frac{N-n}{N}\frac{S_e^2}{n} \qquad \ldots(2.3.2.46)$$

Substitusikan persamaan (2.3.2.45) dan (2.3.2.46) menjadi

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left((y_i-\bar{y})-\hat{\beta}_1(x_i-\bar{x})\right)^2\right)=\frac{1}{n-2}\left(\frac{n}{N}\sum_{j=1}^{N}\left(e_j-\mu_e\right)^2-n\left(\frac{N-n}{N}\frac{S_e^2}{n}\right)\right.$$
$$\left.-E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)\right)$$

Selanjutnya akan dicari terlebih dahulu $E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)$.

Dari sebelumnya diketahui bahwa $\hat{\beta}_1-\beta_1^*=\dfrac{\sum_{i=1}^{n}e_i\left(x_i-\bar{x}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$, maka

$$E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)=E\left(\left(\frac{\sum_{i=1}^{n}e_i\left(x_i-\bar{x}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)$$

$$E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)=E\left(\frac{\left(\sum_{i=1}^{n}e_i\left(x_i-\bar{x}\right)\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\right)$$

Misalkan $e_i=y_i-\mu_Y-\beta_1^*\left(x_i-\mu_X\right)$ dan $e_j=y_j-\mu_Y-\beta_1^*\left(x_j-\mu_X\right)$ saling bebas, maka $E\left(e_ie_j\right)=E\left(e_i\right)E\left(e_j\right)$.

$$E\left(e_i\right)=E\left(y_i-\mu_Y-\beta_1^*\left(x_i-\mu_X\right)\right)$$
$$E\left(e_i\right)=E\left(y_i\right)-E\left(\mu_Y\right)-\beta_1^*E\left(x_i-\mu_X\right)$$
$$E\left(e_i\right)=\mu_Y-\mu_Y-\beta_1^*\left(\mu_X-\mu_X\right)$$
$$E\left(e_i\right)=0$$

dan analog dengan cara yang sama dapat dibuktikan $E\left(e_j\right)=0$, maka $E\left(e_ie_j\right)=E\left(e_i\right)E\left(e_j\right)=0.0=0$.

$$E\left(\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)=E\left(\frac{\sum_{i=1}^{n}e_i^2\left(x_i-\overline{x}\right)^2+2\sum_{i=1}^{n}\sum_{i<j}e_ie_j\left(x_i-\overline{x}\right)\left(x_j-\overline{x}\right)}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}\right)$$

$$E\left(\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)=E\left(\frac{\sum_{i=1}^{n}e_i^2\left(x_i-\overline{x}\right)^2}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}\right)+2E\left(\frac{\sum_{i=1}^{n}\sum_{i<j}e_ie_j\left(x_i-\overline{x}\right)\left(x_j-\overline{x}\right)}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}\right)$$

Misalkan $\dfrac{\left(x_i-\overline{x}\right)^2}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}=d_i$ untuk $i=1,2,\dots,n$, maka :

$$E\left(\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)=E\left(\sum_{i=1}^{n}d_ie_i^2\right)+2\frac{\sum_{i=1}^{n}\sum_{i<j}E\left(e_ie_j\right)\left(x_i-\overline{x}\right)\left(x_j-\overline{x}\right)}{\left(\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)^2}$$

Karena $E\left(e_ie_j\right)=E\left(e_i\right)E\left(e_j\right)=0.0=0$, maka

$$E\left(\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)=E\left(\sum_{i=1}^{n}d_ie_i^2\right)$$

$$E\left(\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)=\sum_{i=1}^{n}d_iE\left(e_i^2\right) \qquad \dots(2.3.2.47)$$

Karena $\mu_e=0$, maka

$$E\left(e_i^2\right)=\frac{1}{N}\sum_{i=1}^{N}e_i^2$$

$$E\left(e_i^2\right)=\frac{1}{N}\left(\sum_{i=1}^{N}e_i^2-0\right)$$

$$E\left(e_i^2\right)=\frac{1}{N}\left(\sum_{i=1}^{N}\left(e_i-\mu_e\right)^2\right)$$

$$E\left(e_i^2\right)=\sigma_e^2 \qquad \dots(2.3.2.48)$$

Substitusikan persamaan (2.3.2.48) ke dalam (2.3.2.47) menjadi

$$E\left(\left(\widehat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right)=\sum_{i=1}^{n}d_i\sigma_e^2 \qquad \dots(2.3.2.49)$$

Karena $\frac{\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}=d_i$ untuk $i=1,2,\dots,n$, maka persamaan (2.3.2.49) menjadi

$$E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)=\frac{\sum_{i=i}^{n}\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\sigma_e^2$$

$$E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)=\sigma_e^2 \qquad \dots(2.3.2.50)$$

Karena $\sigma_e^2=\frac{1}{N}\sum_{i=1}^{N}\left(e_i-\mu_e\right)^2$ dan $\left(N-1\right)S_e^2=\sum_{i=1}^{N}\left(e_i-\mu_e\right)^2$, maka

$$\sigma_e^2=\frac{N-1}{N}S_e^2 \qquad \dots(2.3.2.51)$$

Substitusikan persamaan (2.3.2.51) ke dalam (2.3.2.50) menjadi

$$E\left(\left(\hat{\beta}_1-\beta_1^*\right)^2\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)=\frac{N-1}{N}S_e^2$$

Dengan demikian diperoleh :

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left(\left(y_i-\bar{y}\right)-\hat{\beta}_1\left(x_i-\bar{x}\right)\right)^2\right)=\frac{1}{n-2}\left(\frac{n}{N}\sum_{j=1}^{N}\left(e_j-\mu_e\right)^2-n\left(\frac{N-n}{N}\frac{S_e^2}{n}\right)-\frac{N-1}{N}S_e^2\right)$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left(\left(y_i-\bar{y}\right)-\hat{\beta}_1\left(x_i-\bar{x}\right)\right)^2\right)=\frac{1}{n-2}\left(n\frac{N-1}{N}S_e^2-\frac{N-n}{N}S_e^2-\frac{N-1}{N}S_e^2\right)$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left(\left(y_i-\bar{y}\right)-\hat{\beta}_1\left(x_i-\bar{x}\right)\right)^2\right)=\frac{1}{n-2}S_e^2\left(n\frac{N-1}{N}-\frac{N-n}{N}-\frac{N-1}{N}\right)$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left(\left(y_i-\bar{y}\right)-\hat{\beta}_1\left(x_i-\bar{x}\right)\right)^2\right)=\frac{1}{n-2}S_e^2\left(\frac{nN-n-N+n-N+1}{N}\right)$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left(\left(y_i-\bar{y}\right)-\hat{\beta}_1\left(x_i-\bar{x}\right)\right)^2\right)=\frac{1}{n-2}S_e^2\left(\frac{nN-2N+1}{N}\right)$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left(\left(y_i-\bar{y}\right)-\hat{\beta}_1\left(x_i-\bar{x}\right)\right)^2\right)=\frac{1}{n-2}S_e^2\left(\left(n-2\right)+\frac{1}{N}\right)$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left(\left(y_i-\bar{y}\right)-\hat{\beta}_1\left(x_i-\bar{x}\right)\right)^2\right)=S_e^2+\frac{1}{n-2}\frac{1}{N}S_e^2 \qquad \dots(2.3.2.52)$$

Karena untuk ukuran sampel yang besar populasi juga harus berukuran besar, maka

$$\frac{1}{n-2}\frac{1}{N}\approx 0 \qquad \ldots(2.3.2.53)$$

Substitusikan persamaan (2.3.2.53) ke dalam (2.3.2.52) menjadi

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2\right)\approx S_e^2$$

Dengan demikian diperoleh,

$$E\left(\widehat{Var}\left(\widehat{\mu}_{Yr}\right)\right)=\frac{N-n}{nN}E\left(\frac{1}{n-2}\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2\right)$$

$$E\left(\widehat{Var}\left(\widehat{\mu}_{Yr}\right)\right)\approx\frac{N-n}{nN}S_e^2$$

Karena dari sebelumnya telah dibuktikan bahwa $S_e^2=S_Y^2\left(1-\rho^2\right)$, maka

$$E\left(\widehat{Var}\left(\widehat{\mu}_{Yr}\right)\right)\approx\frac{N-n}{nN}S_Y^2\left(1-\rho^2\right)$$

$$E\left(\widehat{Var}\left(\widehat{\mu}_{Yr}\right)\right)\approx Var\left(\widehat{\mu}_{Yr}\right)$$

Jadi terbukti bahwa $\widehat{Var}\left(\widehat{\mu}_{Yr}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\widehat{\mu}_{Yr}\right)$ untuk sampel yang besar dengan

$$\widehat{Var}\left(\widehat{\mu}_{Yr}\right)=\frac{N-n}{nN(n-2)}\sum_{i=1}^{n}\left((y_i-\bar{y})-\widehat{\beta}_1(x_i-\bar{x})\right)^2.$$

(Terbukti)

BAB III

TAKSIRAN REGRESI UNTUK MEAN POPULASI

PADA STRATIFIED RANDOM SAMPLING

Telah dijelaskan sebelumnya yaitu pada Bab II bahwa taksiran regresi dapat dipergunakan pada Simple Random Sampling, maka pada Bab III ini akan dijelaskan mengenai taksiran regresi pada Stratified Random Sampling (StRS) yang mencakup penjelasan mengenai taksiran terpisah dan gabungan untuk mean populasi dan variansinya serta memeriksa apakah taksiran yang diperoleh adalah taksiran yang bias atau tidak bias.

3.1. PENDAHULUAN

Seperti yang telah diketahui pada bab sebelumnya bahwa Stratified Random Sampling (StRS) adalah teknik pengambilan sampel dengan mengelompokkan elemen – elemen populasi yang heterogen ke dalam kelompok – kelompok yang homogen yang tidak *overlapping* yang disebut strata, kemudian dari setiap stratum dipilih elemen – elemen secara Simple Random Sampling (SRS).

Misalkan diketahui suatu populasi berukuran *N* yang dibagi menjadi L strata, masing – masing berukuran $N_1, N_2, ..., N_L$ dan $N_1 + N_2 + ... + N_L = N$. Kemudian dari setiap stratum diambil sampel yang masing – masing berukuran $n_1, n_2, ..., n_L$ dengan menggunakan Simple Random Sampling (SRS) dengan $n_1 + n_2 +, ... + n_L = n$. Sebut *Y* adalah variabel yang akan diteliti dan *X* adalah variabel tambahan yang berkorelasi kuat dengan *Y*.

Misalkan :

μ_{Y_i} : Mean populasi dari *Y* untuk stratum ke-*i* dengan $i = 1, 2, ..., L$.

$\hat{\mu}_{Yr_i}$: Taksiran regresi untuk μ_{Y_i} dengan $i = 1, 2, ..., L$.

$\bar{y}_i$: Mean sampel dari *Y* untuk stratum ke-i yang diperoleh secara

SRS dengan $i = 1, 2, \ldots, L$.

$\overline{x}_i$: Mean sampel dari *X* untuk stratum ke-*i* yang diperoleh secara SRS dengan $i = 1, 2, \ldots, L$.

μ_{X_i} : Mean populasi dari *X* untuk stratum ke-*i* dengan $i = 1, 2, \ldots, L$.

N_i : Ukuran populasi stratum ke-i dengan $i = 1, 2, \ldots, L$.

n_i : Ukuran sampel yang diambil dari stratum ke-i secara SRS.

L : Banyaknya strata

3.2. TAKSIRAN REGRESI UNTUK MEAN POPULASI DAN VARIANSINYA

Taksiran regresi pada StRS dapat dicari dengan 2 (dua) cara yang berbeda, yaitu :

1. Taksiran regresi terpisah
2. Taksiran regresi gabungan

3.2.1 TAKSIRAN REGRESI TERPISAH UNTUK MEAN POPULASI DAN VARIANSINYA

Seperti yang telah diketahui sebelumnya bahwa pada StRS pada setiap stratumnya sampel diambil secara SRS . Misalkan didapat taksiran regresi untuk mean populasi pada stratum ke-*i* yaitu :

$$\widehat{\mu}_{Yr_i} = \overline{y}_i + b_i\left(\mu_{X_i} - \overline{x}_i\right) \qquad \text{...(3.2.1.1)}$$

dengan

$$b_i = \frac{\sum_{j=1}^{n_i}\left(y_{ij} - \overline{y}_i\right)\left(x_{ij} - \overline{x}_i\right)}{\sum_{j=1}^{n_i}\left(x_{ij} - \overline{x}_i\right)^2} \qquad \text{...(3.2.1.2)}$$

adalah taksiran koefisien regresi $\widehat{\beta}_{1i} = b_i$.

Definisikan taksiran regresi terpisah untuk mean populasi yaitu :

$$\widehat{\mu}_{Ys} = \sum_{i=1}^{L} \frac{N_i}{N} \widehat{\mu}_{Yr_i}$$ ...(3.2.1.3)

atau,

$$\widehat{\mu}_{Ys} = \sum_{i=1}^{L} \frac{N_i}{N} \left(\overline{y}_i + b_i \left(\mu_{X_i} - \overline{x}_i \right) \right)$$
...(3.2.1.4)

Selanjutnya akan dibuktikan bahwa :

1. $\widehat{\mu}_{Ys}$ adalah taksiran yang bias untuk μ_Y .
2. Jika n_i besar pada setiap stratumnya, maka $\widehat{\mu}_{Ys} \approx \mu_Y$.
3. Untuk sampel n_i yang besar pada setiap stratumnya, variansi dari $\widehat{\mu}_{Ys}$ adalah :

   $$Var\left(\widehat{\mu}_{Ys}\right) \approx \sum_{i=1}^{L} \left(\frac{N_i}{N}\right)^2 \left(\frac{N_i - n_i}{N_i}\right) \frac{1}{n_i} S_{Y_i}^2 \left(1 - \rho_i^2\right)$$ …(3.2.1.5)

   dimana $S_{Y_i}^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} \left(y_{ij} - \mu_{Y_i}\right)^2$, $S_{X_i}^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} \left(x_{ij} - \mu_{X_i}\right)^2$,

   $S_{YX_i} = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} \left(y_{ij} - \mu_{Y_i}\right)\left(x_{ij} - \mu_{X_i}\right)$, dan $\rho_i^2 = \frac{S_{YX_i}}{S_{Y_i} S_{X_i}}$

   untuk $i = 1, 2, \ldots, L$.
4. Misalkan didefinisikan taksiran variansi dari $Var\left(\widehat{\mu}_{Ys}\right)$ adalah :

   $$\widehat{Var}\left(\widehat{\mu}_{Ys}\right) = \sum_{i=1}^{L} \left(\frac{N_i}{N}\right)^2 \left(\frac{N_i - n_i}{N_i}\right) \frac{1}{n_i (n_i - 2)} \left(\sum_{j=1}^{n_i} \left(\left(y_{ij} - \overline{y}_i\right) - b_i \left(x_{ij} - \overline{x}_i\right) \right)^2 \right)$$
   …(3.2.1.6)

   maka dapat dibuktikan bahwa $E\left(\widehat{Var}\left(\widehat{\mu}_{Ys}\right)\right) \approx Var\left(\widehat{\mu}_{Ys}\right)$ untuk n_i yang besar pada setiap stratumnya.

Pembuktian :

1. Untuk membuktikan bahwa $\widehat{\mu}_{Ys}$ adalah taksiran yang bias untuk μ_Y adalah dengan menunjukkan bahwa $E\left(\widehat{\mu}_{Ys}\right) \neq \mu_Y$ atau $E\left(\widehat{\mu}_{Ys} - \mu_Y\right) \neq 0$.

Bukti :

Pada bab II sebelumnya telah dibuktikan bahwa $\widehat{\mu}_{Yr}$ adalah taksiran yang bias untuk μ_Y dengan bias sebesar $-Cov\left(\widehat{\beta}_1, \overline{x}\right)$. Dalam StRS didalam setiap stratum sampel diambil secara SRS, maka $\widehat{\mu}_{Yr_i}$ adalah taksiran yang bias untuk μ_{Y_i} dengan bias sebesar $-Cov\left(\widehat{\beta}_{1i}, \overline{x}_i\right)$ atau $-Cov\left(b_i, \overline{x}_i\right)$.

$$E\left(\widehat{\mu}_{Ys}\right) = E\left(\sum_{i=1}^{L} \frac{N_i}{N} \widehat{\mu}_{Yr_i}\right)$$
$$E\left(\widehat{\mu}_{Ys}\right) = \sum_{i=1}^{L} \frac{N_i}{N} E\left(\widehat{\mu}_{Yr_i}\right)$$
$$E\left(\widehat{\mu}_{Ys}\right) = \sum_{i=1}^{L} \frac{N_i}{N} \left(\mu_{Y_i} - Cov\left(b_i, \overline{x}_i\right)\right)$$
$$E\left(\widehat{\mu}_{Ys}\right) = \mu_Y - \sum_{i=1}^{L} \frac{N_i}{N} Cov\left(b_i, \overline{x}_i\right)$$

Dengan demikian terbukti bahwa $\widehat{\mu}_{Ys}$ adalah taksiran yang bias untuk μ_Y dengan bias sebesar $-\sum_{i=1}^{L} \frac{N_i}{N} Cov\left(b_i, \overline{x}_i\right)$ untuk $i = 1, 2, .., L$.

(Terbukti)

2. Berikutnya akan dibuktikan bahwa untuk sampel n_i yang besar pada setiap stratumnya, $\widehat{\mu}_{Ys} \approx \mu_Y$.

Bukti :

Pada bab II sebelumnya telah dibuktikan bahwa untuk n yang besar $\widehat{\mu}_{Yr} \approx \mu_Y$. Dalam StRS didalam setiap stratum sampel diambil secara SRS, maka $\widehat{\mu}_{Yr_i} \approx \mu_{Y_i}$ untuk sampel n_i yang besar .

$$\widehat{\mu}_{Ys} = \sum_{i=1}^{L} \frac{N_i}{N} \widehat{\mu}_{Yr_i}$$
$$\widehat{\mu}_{Ys} \approx \sum_{i=1}^{L} \frac{N_i}{N} \mu_{Y_i}$$
$$\widehat{\mu}_{Ys} \approx \mu_Y$$

Dengan demikian terbukti bahwa untuk n_i yang besar pada setiap stratumnya, $\widehat{\mu}_{Ys} \approx \mu_Y$.

(Terbukti)

3. Selanjutnya akan dibuktikan bahwa untuk sampel n_i yang besar pada setiap stratumnya, $Var\left(\hat{\mu}_{Ys}\right) \approx \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i}S_{Y_i}^2\left(1-\rho_i^2\right)$ dimana $S_{Y_i}^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(y_{ij} - \mu_{Y_i}\right)^2$, $S_{X_i}^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(x_{ij} - \mu_{X_i}\right)^2$, dan $\rho_i^2 = \frac{S_{YX_i}}{S_{Y_i}S_{X_i}}$ untuk $i = 1, 2, \ldots, L$.

Bukti :

Pada bab II sebelumnya telah dibuktikan bahwa untuk sampel yang besar $Var\left(\hat{\mu}_{Yr}\right) \approx \frac{N-n}{nN}S_Y^2\left(1-\rho^2\right)$. Dalam StRS didalam setiap stratum sampel diambil secara SRS, maka $Var\left(\hat{\mu}_{Yr_i}\right) \approx \frac{N_i - n_i}{N_i}\frac{1}{n_i}S_{Y_i}^2\left(1-\rho_i^2\right)$ untuk sampel n_i yang besar.

$$Var\left(\hat{\mu}_{Ys}\right) = Var\left(\sum_{i=1}^{L}\frac{N_i}{N}\hat{\mu}_{Yr_i}\right)$$

$$Var\left(\hat{\mu}_{Ys}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 Var\left(\hat{\mu}_{Yr_i}\right)$$

Karena untuk setiap stratumnya sampel diambil secara SRS dimana $Var\left(\hat{\mu}_{Yr_i}\right) \approx \frac{N_i - n_i}{N_i}\frac{1}{n_i}S_{Y_i}^2\left(1-\rho_i^2\right)$ untuk n_i yang besar, maka

$$Var\left(\hat{\mu}_{Ys}\right) \approx \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i}S_{Y_i}^2\left(1-\rho_i^2\right)$$

Dengan demikian diperoleh

$Var\left(\hat{\mu}_{Ys}\right) \approx \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i}S_{Y_i}^2\left(1-\rho_i^2\right)$ untuk n_i yang besar pada setiap stratumnya.

(Terbukti)

4. Selanjutnya akan dibuktikan bahwa untuk n_i yang besar pada setiap stratumnya,

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i\left(n_i-2\right)}\left(\sum_{j=1}^{n_i}\left(\left(y_{ij}-\bar{y}_i\right)-b_i\left(x_{ij}-\bar{x}_i\right)\right)^2\right)$$

adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Ys}\right)$.

Bukti :

Pada bab II sebelumnya telah dibuktikan bahwa untuk sampel yang besar $\widehat{Var}\left(\hat{\mu}_{Yr}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yr}\right)$. Karena untuk setiap stratumnya sampel diambil secara SRS dimana $\widehat{Var}\left(\hat{\mu}_{Yr_i}\right)=\frac{N_i-n_i}{n_iN_i\left(n_i-2\right)}\sum_{j=1}^{n_i}\left(\left(y_{ij}-\bar{y}_i\right)-\hat{\beta}_{1i}\left(x_{ij}-\bar{x}_i\right)\right)^2$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yr_i}\right)$ untuk sampel n_i yang besar, maka (3.2.1.6) menjadi

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\widehat{Var}\left(\hat{\mu}_{Yr_i}\right) \qquad \dots(3.2.1.7)$$

Jadi,

$$E\left(\widehat{Var}\left(\hat{\mu}_{Ys}\right)\right)=E\left(\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\widehat{Var}\left(\hat{\mu}_{Yr_i}\right)\right)$$

$$E\left(\widehat{Var}\left(\hat{\mu}_{Ys}\right)\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2E\left(\widehat{Var}\left(\hat{\mu}_{Yr_i}\right)\right)$$

Karena untuk sampel n_i yang besar $\widehat{Var}\left(\hat{\mu}_{Yr_i}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yr_i}\right)$, maka

$$E\left(\widehat{Var}\left(\hat{\mu}_{Ys}\right)\right)\approx\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2Var\left(\hat{\mu}_{Yr_i}\right) \qquad \dots(3.2.1.8)$$

Karena untuk setiap stratumnya sampel diambil secara SRS dimana $Var\left(\hat{\mu}_{Yr_i}\right)\approx\frac{N_i-n_i}{N_i}\frac{1}{n_i}S_{Y_i}^2\left(1-\rho_i^2\right)$ untuk sampel n_i yang besar, maka (3.2.1.8) menjadi

$$Var\left(\hat{\mu}_{Ys}\right)\approx\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2Var\left(\hat{\mu}_{Yr_i}\right) \qquad \dots(3.2.1.9)$$

Substitusikan (3.2.1.8) ke dalam (3.2.1.9) menjadi

$$E\left(\widehat{Var}\left(\widehat{\mu}_{Ys}\right)\right) \approx Var\left(\widehat{\mu}_{Ys}\right)$$

Karena $E\left(\widehat{Var}\left(\widehat{\mu}_{Ys}\right)\right) \approx Var\left(\widehat{\mu}_{Ys}\right)$ maka terbukti bahwa $\widehat{Var}\left(\widehat{\mu}_{Ys}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\widehat{\mu}_{Ys}\right)$ untuk sampel n_i yang besar pada setiap stratumnya dengan

$$\widehat{Var}\left(\widehat{\mu}_{Ys}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i\left(n_i - 2\right)}\left(\sum_{j=1}^{n_i}\left(\left(y_{ij} - \bar{y}_i\right) - b_i\left(x_{ij} - \bar{x}_i\right)\right)^2\right).$$

(Terbukti)

3.2.2 TAKSIRAN REGRESI GABUNGAN UNTUK MEAN POPULASI DAN VARIANSINYA

Seperti yang telah diketahui sebelumnya bahwa pada StRS pada setiap stratumnya sampel diambil secara SRS. Taksiran regresi untuk mean populasi pada stratum ke-*i* yaitu :

$$\widehat{\mu}_{Yr_i} = \bar{y}_i + b_i\left(\mu_{X_i} - \bar{x}_i\right) \quad ...(3.2.2.1)$$

dengan

$$b_i = \frac{\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_i\right)\left(x_{ij} - \bar{x}_i\right)}{\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2} \quad ...(3.2.2.2)$$

adalah taksiran koefisien regresi $\widehat{\beta}_{1i} = b_i$.

Didefinisikan taksiran regresi gabungan untuk mean populasi, μ_Y yaitu sebagai berikut :

$$\widehat{\mu}_{Yc} = \bar{y}_{st} + b_c\left(\mu_x - \bar{x}_{st}\right) \quad …(3.2.2.3)$$

dengan

$$b_c = \frac{\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\left(\frac{1}{n_i\left(n_i - 1\right)}\right)\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_i\right)\left(x_{ij} - \bar{x}_i\right)}{\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\left(\frac{1}{n_i\left(n_i - 1\right)}\right)\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2} \quad …(3.2.2.4)$$

Selanjutnya akan dibuktikan bahwa :

1. $\hat{\mu}_{Yc}$ adalah taksiran yang bias untuk μ_Y.

2. Variansi dari $\hat{\mu}_{Yc}$ adalah :

$$Var\left(\hat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}\left(S_{Y_i}^2-2b_cS_{YX_i}^2+b_c^2S_{X_i}^2\right) \quad \ldots(3.2.2.5)$$

dimana $S_{Y_i}^2=\dfrac{1}{N_i-1}\sum_{j=1}^{N_i}\left(y_{ij}-\mu_{Y_i}\right)^2$, $S_{X_i}^2=\dfrac{1}{N_i-1}\sum_{j=1}^{N_i}\left(x_{ij}-\mu_{X_i}\right)^2$, dan

$$S_{YX_i}=\frac{1}{N_i-1}\sum_{j=1}^{N_i}\left(y_{ij}-\mu_{Y_i}\right)\left(x_{ij}-\mu_{X_i}\right).$$

3. Misalkan didefinisikan taksiran variansi dari $Var\left(\hat{\mu}_{Yc}\right)$ adalah :

$$\widehat{Var}\left(\hat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(s_{Y_i}^2-2b_cs_{YX_i}^2+b_c^2s_{X_i}^2\right) \quad \ldots(3.2.2.6)$$

maka dapat dibuktikan bahwa $\widehat{Var}\left(\hat{\mu}_{Yc}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yc}\right)$.

Pembuktian :

1. Akan dibuktikan bahwa $\hat{\mu}_{Yc}$ adalah taksiran yang tidak bias untuk μ_Y.

Bukti :

Untuk membuktikan bahwa $\hat{\mu}_{Yc}$ adalah taksiran yang tidak bias untuk μ_Y yaitu dengan menunjukkan bahwa $E\left(\hat{\mu}_{Yc}\right)=\mu_Y$ atau $E\left(\hat{\mu}_{Yc}\right)-\mu_Y=0$.

$$E\left(\hat{\mu}_{Yc}\right)=E\left(\bar{y}_{st}+b_c\left(\mu_X-\bar{x}_{st}\right)\right)$$
$$E\left(\hat{\mu}_{Yc}\right)=E\left(\bar{y}_{st}\right)+b_cE\left(\mu_X-\bar{x}_{st}\right)$$
$$E\left(\hat{\mu}_{Yc}\right)=E\left(\bar{y}_{st}\right)+b_c\left(E\left(\mu_X\right)-E\left(\bar{x}_{st}\right)\right)$$

Karena dalam StRS, $\bar{y}_{st}$ adalah taksiran yang tidak bias untuk μ_Y atau $E\left(\bar{y}_{st}\right)=\mu_Y$ dan $\bar{x}_{st}$ adalah taksiran yang tidak bias untuk μ_X atau $E\left(\bar{x}_{st}\right)=\mu_X$, maka

$$E\left(\hat{\mu}_{Yc}\right) = \mu_Y + b_c\left(\mu_X - \mu_X\right)$$

$$E\left(\hat{\mu}_{Yc}\right) = \mu_Y$$

Karena $E\left(\hat{\mu}_{Yc}\right) = \mu_Y$ maka terbukti bahwa $\hat{\mu}_{Yc}$ adalah taksiran yang tidak bias untuk μ_Y.

(Terbukti)

2. Selanjutnya akan dibuktikan bahwa variansi dari $\hat{\mu}_{Yc}$ adalah :

$$Var\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i}\left(S_{Y_i}^2 - 2b_c S_{YX_i}^2 + b_c^2 S_{X_i}^2\right)$$

dimana $S_{Y_i}^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(y_{ij} - \mu_{Y_i}\right)^2$, $S_{X_i}^2 = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(x_{ij} - \mu_{X_i}\right)^2$, dan $S_{YX_i} = \frac{1}{N_i - 1}\sum_{j=1}^{N_i}\left(y_{ij} - \mu_{Y_i}\right)\left(x_{ij} - \mu_{X_i}\right)$.

Bukti :

$$Var\left(\hat{\mu}_{Yc}\right) = Var\left(\sum_{i=1}^{L}\frac{N_i}{N}\hat{\mu}_{Yr_i}^*\right)$$

$$Var\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 Var\left(\hat{\mu}_{Yr_i}^*\right)$$

$$Var\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 E\left(\left(\hat{\mu}_{Yr_i}^* - E\left(\left(\hat{\mu}_{Yr_i}^*\right)\right)\right)^2\right) \quad \ldots(3.2.2.7)$$

Substitusikan (3.2.2.4) ke (3.2.2.10) menjadi :

$$Var\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 E\left(\left(\bar{y}_i + b_c\left(\mu_{X_i} - \bar{x}_i\right) - E\left(\bar{y}_i + b_c\left(\mu_{X_i} - \bar{x}_i\right)\right)\right)^2\right)$$

$$Var\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 E\left(\left(\bar{y}_i + b_c\left(\mu_{X_i} - \bar{x}_i\right) - \left(E\left(\bar{y}_i\right) + b_c\mu_{X_i} - b_c E\left(\bar{x}_i\right)\right)\right)^2\right)$$

Karena untuk setiap stratumnya sampel diambil secara SRS dimana $\bar{y}_i$ dan $\bar{x}_i$ berturut – turut adalah taksiran yang tidak bias untuk μ_{Y_i} dan μ_{X_i}, maka

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 E\left(\left(\overline{y}_i+b_c\left(\mu_{X_i}-\overline{x}_i\right)-\left(\mu_{Y_i}+b_c\mu_{X_i}-b_c\mu_{X_i}\right)\right)^2\right)$$

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 E\left(\left(\overline{y}_i+b_c\left(\mu_{X_i}-\overline{x}_i\right)-\mu_{Y_i}\right)^2\right)$$

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 E\left(\overline{y}_i^2-2b_c\overline{y}_i\left(\mu_{X_i}-\overline{x}_i\right)^2+b_c^2\left(\mu_{X_i}-\overline{x}_i\right)^2-\mu_{Y_i}\right)$$

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2 E\left(\left(\overline{y}_i^2-\mu_{Y_i}\right)^2-2b_c\overline{y}_i\left(\mu_{X_i}-\overline{x}_i\right)+b_c^2\left(\mu_{X_i}-\overline{x}_i\right)^2\right)$$

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(E\left(\left(\overline{y}_i^2-\mu_{Y_i}\right)^2\right)-2b_cE\left(\overline{y}_i\left(\mu_{X_i}-\overline{x}_i\right)\right)+b_c^2E\left(\left(\mu_{X_i}-\overline{x}_i\right)^2\right)\right)$$

Karena $Var\left(\overline{y}_i\right)=E\left(\left(\overline{y}_i^2-\mu_{Y_i}\right)^2\right)$, $Var\left(\overline{x}_i\right)=E\left(\left(\overline{x}_i^2-\mu_{X_i}\right)^2\right)$, dan

$Cov\left(\overline{y}_i,\overline{x}_i\right)=E\left(\overline{y}_i\left(\mu_{X_i}-\overline{x}_i\right)\right)$, maka

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(Var\left(\overline{y}_i\right)-2b_cCov\left(\overline{y}_i,\overline{x}_i\right)+b_c^2Var\left(\overline{x}_i\right)\right)$$

Karena untuk setiap stratumnya sampel diambil secara SRS dimana $Var\left(\overline{y}_i\right)=\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}S_{Y_i}^2$, $Var\left(\overline{x}_i\right)=\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}S_{X_i}^2$, dan

$Cov\left(\overline{y}_i,\overline{x}_i\right)=\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}S_{YX_i}$, maka

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}\left(S_{Y_i}^2-2b_cS_{YX_i}^2+b_c^2S_{X_i}^2\right)$$

Dengan demikian diperoleh variansi dari $\widehat{\mu}_{Yc}$ yaitu :

$$Var\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}\left(S_{Y_i}^2-2b_cS_{YX_i}^2+b_c^2S_{X_i}^2\right).$$

(Terbukti)

3. Berikutnya akan dibuktikan bahwa :

$$\widehat{Var}\left(\widehat{\mu}_{Yc}\right)=\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(s_{Y_i}^2-2b_cs_{YX_i}^2+b_c^2s_{X_i}^2\right)$$

adalah taksiran yang tidak bias untuk $Var\left(\widehat{\mu}_{Yc}\right)$.

Bukti :

Untuk membuktikan bahwa $\widehat{Var}\left(\hat{\mu}_{Yc}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yc}\right)$ adalah dengan menunjukkan bahwa $E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = Var\left(\hat{\mu}_{Yc}\right)$.

Didefinisikan bahwa

$$\widehat{Var}\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(s_{Y_i}^2 - 2b_c s_{YX_i}^2 + b_c^2 s_{X_i}^2\right) \qquad \ldots(3.2.2.8)$$

maka,

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = E\left(\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(s_{Y_i}^2 - 2b_c s_{YX_i}^2 + b_c^2 s_{X_i}^2\right)\right)$$

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}E\left(s_{Y_i}^2 - 2b_c s_{YX_i}^2 + b_c^2 s_{X_i}^2\right)$$

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(E\left(s_{Y_i}^2\right) - 2b_c E\left(s_{YX_i}^2\right) + b_c^2 E\left(s_{X_i}^2\right)\right)$$

Karena untuk setiap stratumnya sampel diambil secara SRS dimana $s_{Y_i}^2, s_{X_i}^2$, dan $s_{YX_i}^2$ berturut – turut adalah taksiran yang tidak bias untuk $S_{Y_i}^2, S_{X_i}^2$, dan $S_{YX_i}^2$, maka

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(E\left(s_{Y_i}^2\right) - 2b_c E\left(s_{YX_i}^2\right) + b_c^2 E\left(s_{X_i}^2\right)\right)$$

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(S_{Y_i}^2 - 2b_c S_{YX_i}^2 + b_c^2 S_{X_i}^2\right)$$

$$E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = Var\left(\hat{\mu}_{Yc}\right)$$

Karena $E\left(\widehat{Var}\left(\hat{\mu}_{Yc}\right)\right) = Var\left(\hat{\mu}_{Yc}\right)$, maka terbukti bahwa $\widehat{Var}\left(\hat{\mu}_{Yc}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yc}\right)$.

(Terbukti)

BAB IV

APLIKASI TAKSIRAN REGRESI UNTUK MEAN POPULASI PADA *STRATIFIED RANDOM SAMPLING*

Pada bab ini akan dibahas contoh penggunaan taksiran mean populasi dengan menggunakan metode taksiran regresi pada *Stratified Random Sampling* (StRS) baik dengan taksiran regresi terpisah maupun taksiran regresi gabungan.

Misalkan suatu pengamatan akan dilakukan untuk mengetahui mean dari variabel *Y* dengan variabel tambahan *X* dimana *Y* dan *X* mempunyai korelasi yang kuat. Populasi dapat dibagi menjadi 3 strata, sebut stratum 1, stratum 2, dan stratum 3.

Misalkan,

N_1 : Ukuran populasi stratum ke-1

N_2 : Ukuran populasi stratum ke-2

N_3 : Ukuran populasi stratum ke-3

μ_{X_1} : Mean populasi dari *X* pada stratum ke-1

μ_{X_2} : Mean populasi dari *X* pada stratum ke-2

μ_{X_3} : Mean populasi dari *X* pada stratum ke-3

$S^2_{X_1}$: Variansi populasi dari *X* pada stratum ke-1

$S^2_{X_2}$: Variansi populasi dari *X* pada stratum ke-2

$S^2_{X_3}$: Variansi populasi dari *X* pada stratum ke-3

Misalkan diketahui :

$N_1 = 210$, $N_2 = 150$, dan $N_3 = 140$

dengan ukuran populasi keseluruhan, yang dinotasikan dengan *N*, adalah :

$$N = \sum_{i=1}^{3} N_i = N_1 + N_2 + N_3 = 210 + 150 + 140 = 500$$

Selain itu diketahui :

$\mu_{X_1} = 111.2011$, $\mu_{X_2} = 54.28348$, dan $\mu_{X_3} = 25.96974$.

dengan $\mu_X = 70.261$

Informasi diatas akan digunakan untuk menaksir mean dari *Y*.

Untuk kepentingan tersebut diambil sampel secara StRS dan dicatat nilai *X* dan *Y* yang terpilih sebagai sampel. Dalam contoh ini akan diambil sampel sebanyak 250 dengan alokasi proporsional, yaitu sebagai berikut :

$$n_i = \frac{N_I}{N} n$$

untuk setiap stratum dengan $i = 1,2,3$ dan $n = 250$.

Maka, ukuran sampel untuk masing – masing stratum adalah :

$$n_1 = \frac{210}{500}(250) = 105$$

$$n_2 = \frac{150}{500}(250) = 75$$

$$n_3 = \frac{140}{500}(250) = 70$$

Misalkan variabel Y_i adalah variabel *Y* pada stratum ke – *i* dan X_i adalah variabel *X* pada stratum ke – *i* mempunyai korelasi cukup kuat. Misalkan hubungan antara Y_i dan X_i dapat dituliskan dalam model regresi linier sebagai berikut :

$$Y_i = \beta_{0i} + \beta_{1i} X_i + \varepsilon_i$$

dimana β_{0i} dan β_{1i} adalah parameter yang tidak diketahui dan ε_i adalah variabel errornya untuk setiap stratum dengan $i = 1,2,3$. Parameter β_{0i} dan β_{1i} dapat ditaksir menggunakan metode *Least Squares*.

Dalam bab ini akan diberikan 2 contoh dimana taksiran β_{1i}, sebut b_i dari ketiga stratum berbeda – beda dan contoh dimana nilai b_i dari ketiga stratum relatif sama.

4.1 CONTOH 1

Pada contoh 1 ini akan dibahas contoh dimana nilai b_i bervariasi untuk ketiga stratum. Adapun data untuk contoh 1 ini ada pada lampiran 3.

Selanjutnya dari data tersebut diperoleh hasil pengolahan data sebagai berikut :

Untuk stratum ke-1,

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.897[a]	.804	.802	17.29840

a. Predictors: (Constant), x1

Tabel 4.1.1. Model Summary untuk stratum ke-1

Dari **Tabel 4.1.1** diatas, koefisien korelasi antara Y_1 dan X_1 cukup kuat, yaitu 0.897, berarti terdapat hubungan linier yang cukup kuat antara Y_1 dan X_1.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	56.458	11.936	
	x1	2.219	.108	.897

a. Dependent Variable: y1

Tabel 4.1.2. Taksiran koefisien model linier regresi untuk stratum ke-1

Dari **Tabel 4.1.2** diatas, $b_{01} \neq 0$ atau dengan perkataan lain garis regresi tidak melalui titik origin.

Descriptive Statistics

	N	Mean	Std. Deviation	Variance
x1	105	1.0948E2	15.71608	246.995
y1	105	2.9938E2	38.89077	1.512E3
Valid N (listwise)	105			

Tabel 4.1.3. Statistik deskriptif untuk stratum ke-1

Dari **Tabel 4.1.3** didapat $\bar{y}_1 = 299.38$, $\bar{x}_1 = 109.48$, $s^2_{Y_1} = 1512$, $s^2_{X_1} = 246.995$, dan $s_{YX_1} = 548.256$.

Untuk stratum ke-2,

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.948[a]	.898	.897	3.86275

a. Predictors: (Constant), x2

Tabel 4.1.4. Model Summary untuk stratum ke-2

Dari **Tabel 4.1.4** diatas, koefisien korelasi antara Y_2 dan X_2 cukup kuat, yaitu 0.948, berarti terdapat hubungan linier yang cukup kuat antara Y_2 dan X_2 .

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	70.582	2.566	
	x2	1.197	.047	.948

a. Dependent Variable: y2

Tabel 4.1.5. Taksiran koefisien model linier regresi untuk stratum ke-2

Dari **Tabel 4.1.5** diatas, $b_{02} \neq 0$ atau dengan perkataan lain garis regresi tidak melalui titik origin.

Descriptive Statistics

	N	Mean	Std. Deviation	Variance
x2	75	53.5580	9.51551	90.545
y2	75	1.3468E2	12.01677	144.403
Valid N (listwise)	75			

Tabel 4.1.6. Statistik deskriptif untuk stratum ke-2

Dari **Tabel 4.1.6** didapat $\bar{y}_2 = 134.68$, $\bar{x}_2 = 53.5580$, $s^2_{Y_2} = 144.403$, $s^2_{X_2} = 90.545$, dan $s_{YX_2} = 108.399$.

Untuk stratum ke-3,

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.905[a]	.818	.816	1.60012

a. Predictors: (Constant), x3

Tabel 4.1.7. Model Summary untuk stratum ke-3

Dari **Tabel 4.1.7** diatas, koefisien korelasi antara Y_3 dan X_3 cukup kuat, yaitu 0.905, berarti terdapat hubungan linier yang cukup kuat antara Y_3 dan X_3 .

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	91.937	1.217	
	x3	.812	.046	.905

a. Dependent Variable: y3

Tabel 4.1.8. Taksiran koefisien model linier regresi untuk stratum ke-2

Dari **Tabel 4.1.8** diatas, $b_{03} \neq 0$ atau dengan perkataan lain garis regresi tidak melalui titik origin.

Descriptive Statistics

	N	Mean	Std. Deviation	Variance
x3	70	25.9138	4.15471	17.262
y3	70	1.1297E2	3.72798	13.898
Valid N (listwise)	70			

Tabel 4.1.9. Statistik deskriptif untuk stratum ke-3

Dari **Tabel 4.1.9** didapat $\bar{y}_3 = 112.97$, $\bar{x}_3 = 25.9138$, $s^2_{Y_3} = 13.898$, $s^2_{X_3} = 17.262$, dan $s_{YX_3} = 14.017$.

Selanjutnya akan dilakukan penaksiran mean populasi dari *Y* dengan menggunakan 2 metode yaitu taksiran regresi terpisah dan gabungan.

4.1.1 TAKSIRAN REGRESI TERPISAH UNTUK MEAN POPULASI DAN VARIANSINYA

Seperti yang telah diketahui pada bab sebelumnya bahwa taksiran regresi terpisah untuk mean populasi adalah :

$$\hat{\mu}_{Ys} = \sum_{i=1}^{L} \frac{N_i}{N}\left(\bar{y}_i + b_i\left(\mu_{X_i} - \bar{x}_i\right)\right)$$

Stratum ke-1 didapat $b_1 = 2.219$ dan diketahui $\mu_{X_1} = 111.2011$.

Stratum ke-2 didapat $b_2 = 1.197$ dan diketahui $\mu_{X_2} = 54.28348$.

Stratum ke-3 didapat $b_3 = 0.812$ dan diketahui $\mu_{X_3} = 25.96974$.

Dengan demikian taksiran regresi terpisah untuk mean populasi adalah :

$$\hat{\mu}_{Ys} = \sum_{i=1}^{3} \frac{N_i}{N}\left(\bar{y}_i + b_i\left(\mu_{X_i} - \bar{x}_i\right)\right)$$

$$\hat{\mu}_{Ys} = \frac{210}{500}\left(299.38 + 2.219\left(111.2011 - 109.48\right)\right)$$

$$+\frac{150}{500}\left(134.68 + 1.197\left(54.28348 - 53.5580\right)\right)$$

$$+\frac{140}{500}\left(112.97 + 0.812\left(25.96974 - 25.9138\right)\right)$$

$$\hat{\mu}_{Ys} = 127.343 + 40.980 + 31.644$$

$$\hat{\mu}_{Ys} = 199.967$$

dengan taksiran variansi populasinya adalah :

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx \sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i\left(n_i - 2\right)}\left(\sum_{j=1}^{n_i}\left(\left(y_{ij} - \bar{y}_i\right) - b_i\left(x_{ij} - \bar{x}_i\right)\right)^2\right)$$

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx \sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i\left(n_i - 2\right)}\left(n_i - 1\right)\left(s_{Y_i}^2 - 2b_i s_{YX_i} + b_i^2 s_{X_i}^2\right)$$

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx \left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\left(\frac{1}{105(103)}\right)(104)\left(1512 - 2(2.219)(548.256) + (2.219)^2\, 246.995\right)$$

$$+\left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\left(\frac{1}{75(73)}\right)(74)\left(144.403 - 2(1.197)(108.399) + (1.197)^2\, 90.545\right)$$

$$+\left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\left(\frac{1}{70(68)}\right)(69)\left(13.898 - 2(0.812)(14.017) + (0.812)^2\, 17.262\right)$$

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx 0.26755$$

Jadi, dengan menggunakan taksiran regresi terpisah diperoleh taksiran mean untuk *Y* adalah 199.967 dengan taksiran variansinya adalah 0.26755.

4.1.2 TAKSIRAN REGRESI GABUNGAN UNTUK MEAN POPULASI DAN VARIANSINYA

Seperti yang telah diketahui pada bab sebelumnya bahwa taksiran regresi gabungan untuk mean populasi adalah :

$$\hat{\mu}_{Yc} = \bar{y}_{st} + b_c\left(\mu_x - \bar{x}_{st}\right)$$

Akan dicari terlebih dahulu nilai dari $\bar{y}_{st}$, $\bar{x}_{st}$, dan b_c .

$$\bar{y}_{st} = \sum_{i=1}^{3} \frac{N_i}{N}\bar{y}_i = \frac{210}{500}(299.38) + \frac{150}{500}(134.68) + \frac{140}{500}(112.97) = 197.78$$

$$\bar{x}_{st} = \sum_{i=1}^{3} \frac{N_i}{N}\bar{x}_i = \frac{210}{500}(109.48) + \frac{150}{500}(53.5580) + \frac{140}{500}(25.9138) = 69.305$$

$$b_c = \frac{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{1}{n_i(n_i-1)}\right)\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)(x_{ij}-\bar{x}_i)}{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{1}{n_i(n_i-1)}\right)\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2} = \frac{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}s_{YX_i}}{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}s^2_{X_i}}$$

$$b_c = \frac{\left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\frac{1}{105}(548.256) + \left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\frac{1}{75}(108.399) + \left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\frac{1}{70}(14.017)}{\left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\frac{1}{105}(246.995) + \left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\frac{1}{75}(90.545) + \left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\frac{1}{70}(17.262)}$$

$$b_c = \frac{0.533}{0.271} = 1.967$$

Diketahui, $\mu_x = 70.261.$

Dengan demikian taksiran regresi gabungan untuk mean populasi adalah :

$$\hat{\mu}_{Yc} = \bar{y}_{st} + b_c(\mu_x - \bar{x}_{st})$$

$$\hat{\mu}_{Yc} = 197.78 + 1.967(70.526 - 69.305)$$

$$\hat{\mu}_{Yc} = 200.181$$

dengan taksiran variansi populasinya adalah :

$$\widehat{Var}(\hat{\mu}_{Yc}) = \sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(s^2_{Y_i} - 2b_c s_{YX_i} + b_c^2 s^2_{X_i}\right)$$

$$\begin{aligned}\widehat{Var}(\hat{\mu}_{Yc}) = {} & \left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\left(\frac{1}{105}\right)\left(1512 - 2(1.967)(548.256) + (1.967)^2\, 246.995\right) \\ & + \left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\left(\frac{1}{75}\right)\left(144.403 - 2(1.967)(108.399) + (1.967)^2\, 90.545\right) \\ & + \left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\left(\frac{1}{70}\right)\left(13.898 - 2(1.967)(14.017) + (1.967)^2\, 17.262\right)\end{aligned}$$

$$\widehat{Var}(\hat{\mu}_{Yc}) = 0.316$$

Jadi, dengan menggunakan taksiran regresi gabungan diperoleh taksiran mean untuk *Y* adalah 200.181 dengan taksiran variansinya adalah 0.316.

Karena pada taksiran regresi terpisah taksiran variansi untuk mean populasinya lebih kecil bila dibandingkan dengan taksiran regresi gabungan, yaitu

$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx 0.26755 < \widehat{Var}\left(\hat{\mu}_{Yc}\right) \approx 0.316$, maka taksiran regresi terpisah lebih sesuai digunakan dalam kasus ini.

4.2 CONTOH 2

Pada contoh 2 ini akan dibahas contoh dimana nilai b_i yang relatif sama untuk ketiga stratum. Adapun data untuk contoh 2 ini ada pada lampiran 4.

Selanjutnya dari data tersebut diperoleh hasil pengolahan data sebagai berikut :

Untuk stratum ke-1,

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.834[a]	.696	.693	16.01404

a. Predictors: (Constant), x1

Tabel 4.2.1. Model Summary untuk stratum ke-1

Dari **Tabel 4.2.1** diatas, koefisien korelasi antara Y_1 dan X_1 cukup kuat, yaitu 0.834, berarti terdapat hubungan linier yang cukup kuat antara Y_1 dan X_1.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta
1	(Constant)	74.964	11.503	
	x1	1.585	.103	.834

a. Dependent Variable: y1

Tabel 4.2.2. Taksiran koefisien model linier regresi untuk stratum ke-1

Dari **Tabel 4.2.2** diatas, $b_{01} \neq 0$ atau dengan perkataan lain garis regresi tidak melalui titik origin.

Descriptive Statistics

	N	Mean	Std. Deviation	Variance
x1	105	1.1040E2	15.21143	231.387
y1	105	2.4989E2	28.89549	834.950
Valid N (listwise)	105			

Tabel 4.2.3. Statistik deskriptif untuk stratum ke-1

Dari **Tabel 4.2.3** didapat $\bar{y}_1 = 249.89$, $\bar{x}_1 = 110.40$, $s^2_{Y_1} = 834.950$, $s^2_{X_1} = 231.387$, dan $s_{YX_1} = 366.577$.

Untuk stratum ke-2,

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.954[a]	.910	.909	4.49775

a. Predictors: (Constant), x2

Tabel 4.2.4. Model Summary untuk stratum ke-2

Dari **Tabel 4.2.4** diatas, koefisien korelasi antara Y_2 dan X_2 cukup kuat, yaitu 0.954, berarti terdapat hubungan linier yang cukup kuat antara Y_2 dan X_2.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	74.320	3.187	
	x2	1.582	.058	.954

a. Dependent Variable: y2

Tabel 4.2.5. Taksiran koefisien model linier regresi untuk stratum ke-2

Dari **Tabel 4.2.5** diatas, $b_{02} \neq 0$ atau dengan perkataan lain garis regresi tidak melalui titik origin.

Descriptive Statistics

	N	Mean	Std. Deviation	Variance
y2	75	1.5968E2	14.88166	221.464
x2	75	53.9558	8.97241	80.504
Valid N (listwise)	75			

Tabel 4.2.6. Statistik deskriptif untuk stratum ke-2

Dari **Tabel 4.2.6** didapat $\bar{y}_2 = 159.68$, $\bar{x}_2 = 53.9558$, $s^2_{Y_2} = 221.464$, $s^2_{X_2} = 80.504$, dan $s_{YX_2} = 127.38$.

Untuk stratum ke-3,

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.842[a]	.709	.705	3.69306

a. Predictors: (Constant), x3

Tabel 4.2.7. Model Summary untuk stratum ke-3

Dari **Tabel 4.2.7** diatas, koefisien korelasi antara Y_3 dan X_3 cukup kuat, yaitu 0.842, berarti terdapat hubungan linier yang cukup kuat antara Y_3 dan X_3.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	161.558	3.127	
	x3	1.583	.123	.842

a. Dependent Variable: y3

Tabel 4.2.8. Taksiran koefisien model linier regresi untuk stratum ke-3

Dari **Tabel 4.2.8** diatas, $b_{03} \neq 0$ atau dengan perkataan lain garis regresi tidak melalui titik origin.

Descriptive Statistics

	N	Mean	Std. Deviation	Variance
x3	70	25.1666	3.61474	13.066
y3	70	2.0140E2	6.79669	46.195
Valid N (listwise)	70			

Tabel 4.2.9. Statistik deskriptif untuk stratum ke-3

Dari **Tabel 4.2.9** didapat $\bar{y}_3 = 201.40$, $\bar{x}_3 = 25.1666$, $s^2_{Y_3} = 46.195$, $s^2_{X_3} = 13.066$, dan $s_{YX_3} = 20.686$.

Selanjutnya akan dilakukan penaksiran mean populasi dari *Y* dengan menggunakan 2 metode yaitu taksiran regresi terpisah dan gabungan.

4.2.1 TAKSIRAN REGRESI TERPISAH UNTUK MEAN POPULASI DAN VARIANSINYA

Dari tabel sebelumnya,

Stratum ke-1 didapat $b_1 = 1.585$ dan diketahui $\mu_{X_1} = 111.2011$

Stratum ke-2 didapat $b_2 = 1.582$ dan diketahui $\mu_{X_2} = 54.28348$

Stratum ke-3 didapat $b_3 = 1.583$ dan diketahui $\mu_{X_3} = 25.96974$

Dengan demikian taksiran regresi terpisah untuk mean populasi adalah :

$$\hat{\mu}_{Ys} = \sum_{i=1}^{3} \frac{N_i}{N}\left(\bar{y}_i + b_i\left(\mu_{X_i} - \bar{x}_i\right)\right)$$

$$\begin{aligned}\hat{\mu}_{Ys} &= \frac{210}{500}\left(249.89 + 1.585\left(111.2011 - 110.40\right)\right)\\ &\quad + \frac{150}{500}\left(201.40 + 1.582\left(54.28348 - 53.9558\right)\right)\\ &\quad + \frac{140}{500}\left(159.68 + 1.583\left(25.96974 - 25.1666\right)\right)\end{aligned}$$

$$\hat{\mu}_{Ys} = 104.82 + 60.575 + 45.066$$

$$\hat{\mu}_{Ys} = 210.461$$

dengan taksiran variansi populasinya adalah :

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx \sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i\left(n_i - 2\right)}\left(\sum_{j=1}^{n_i}\left(\left(y_{ij} - \bar{y}_i\right) - b_i\left(x_{ij} - \bar{x}_i\right)\right)^2\right)$$

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx \sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i\left(n_i - 2\right)}\left(n_i - 1\right)\left(s_{Y_i}^2 - 2b_i s_{YX_i} + b_i^2 s_{X_i}^2\right)$$

$$\begin{aligned}\widehat{Var}\left(\hat{\mu}_{Ys}\right) &\approx \left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\left(\frac{1}{105(103)}\right)(104)\left(834.950 - 2(1.585)(366.577) + (1.585)^2\, 231.387\right)\\ &\quad + \left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\left(\frac{1}{75(73)}\right)(74)\left(221.464 - 2(1.582)(127.38) + (1.582)^2\, 80.504\right)\\ &\quad + \left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\left(\frac{1}{70(68)}\right)(69)\left(46.195 - 2(1.583)(20.686) + (1.583)^2\, 13.066\right)\end{aligned}$$

$$\widehat{Var}\left(\hat{\mu}_{Ys}\right) \approx 0.235$$

Jadi, dengan menggunakan taksiran regresi terpisah diperoleh taksiran mean dari *Y* adalah 210.461 dengan taksiran variansinya adalah 0.235.

4.2.2 TAKSIRAN REGRESI GABUNGAN UNTUK MEAN POPULASI DAN VARIANSINYA

Akan dicari terlebih dahulu nilai dari $\bar{y}_{st}$, $\bar{x}_{st}$, dan b_c .

$$\bar{y}_{st} = \sum_{i=1}^{3} \frac{N_i}{N}\bar{y}_i = \frac{210}{500}(249.89) + \frac{150}{500}(159.68) + \frac{140}{500}(201.40) = 209.2498$$

$$\bar{x}_{st} = \sum_{i=1}^{3} \frac{N_i}{N}\bar{x}_i = \frac{210}{500}(110.40) + \frac{150}{500}(53.9558) + \frac{140}{500}(25.1666) = 69.601$$

$$b_c = \frac{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{1}{n_i(n_i-1)}\right)\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)(x_{ij}-\bar{x}_i)}{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{1}{n_i(n_i-1)}\right)\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2} = \frac{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}s_{YX_i}}{\sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}s_{X_i}^2}$$

$$b_c = \frac{\left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\frac{1}{105}(366.577) + \left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\frac{1}{75}(127.38) + \left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\frac{1}{70}(20.686)}{\left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\frac{1}{105}(231.387) + \left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\frac{1}{75}(80.504) + \left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\frac{1}{70}(13.066)}$$

$$b_c = \frac{\frac{210}{500^2}(366.577) + \frac{150}{500^2}(127.38) + \frac{140}{500^2}(20.686)}{\frac{210}{500^2}(231.387) + \frac{150}{500^2}(80.504) + \frac{140}{500^2}(13.066)} = \frac{0.308 + 0.076 + 0.0116}{0.194 + 0.0483 + 0.00731} = 1.584$$

Diketahui, $\mu_x = 70.261$.

Dengan demikian taksiran regresi gabungan untuk mean populasi adalah :

$$\hat{\mu}_{Yc} = \bar{y}_{st} + b_c(\mu_x - \bar{x}_{st})$$

$$\hat{\mu}_{Yc} = 209.2498 + 1.584(70.261 - 69.601)$$

$$\hat{\mu}_{Yc} = 210.29$$

dengan taksiran variansi populasinya adalah :

$$\widehat{Var}(\hat{\mu}_{Yc}) = \sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(s_{Y_i}^2 - 2b_c s_{YX_i} + b_c^2 s_{X_i}^2\right)$$

$$\widehat{Var}\left(\hat{\mu}_{Yc}\right)=\left(\frac{210}{500}\right)^2\left(\frac{105}{210}\right)\left(\frac{1}{105}\right)\left(834.950-2\left(1.584\right)\left(366.577\right)+\left(1.584\right)^2 231.387\right)$$
$$+\left(\frac{150}{500}\right)^2\left(\frac{75}{150}\right)\left(\frac{1}{75}\right)\left(221.464-2\left(1.584\right)\left(127.38\right)+\left(1.584\right)^2 80.504\right)$$
$$+\left(\frac{140}{500}\right)^2\left(\frac{70}{140}\right)\left(\frac{1}{70}\right)\left(46.195-2\left(1.584\right)\left(20.686\right)+\left(1.584\right)^2 13.066\right)$$
$$\widehat{Var}\left(\hat{\mu}_{Yc}\right)=0.233$$

Jadi, dengan menggunakan taksiran regresi gabungan diperoleh taksiran mean dari *Y* adalah 210.29 dengan taksiran variansinya adalah 0.233.

Karena pada taksiran regresi gabungan taksiran variansi untuk mean populasinya lebih kecil bila dibandingkan dengan taksiran regresi terpisah, yaitu $\widehat{Var}\left(\hat{\mu}_{Yc}\right)\approx 0.233<\widehat{Var}\left(\hat{\mu}_{Ys}\right)\approx 0.235$, maka taksiran regresi gabungan lebih sesuai digunakan dalam kasus ini dibandingkan dengan menggunakan taksiran terpisah.

4.3 KESIMPULAN

Dari kedua contoh sebelumnya dapat dijelaskan bahwa :

1. Untuk contoh 1, b_i untuk setiap stratum adalah :

   Untuk stratum ke – 1, $b_1=2.219$

   Untuk stratum ke – 2, $b_2=1.197$

   Untuk stratum ke – 3, $b_3=0.812$

   Disini terlihat bahwa b_i berbeda untuk ketiga stratum.

   Dari analisa data didapat $\widehat{Var}\left(\hat{\mu}_{Ys}\right)=0.26755<\widehat{Var}\left(\hat{\mu}_{Yc}\right)=0.316$ atau dengan perkataan lain dalam hal ini taksiran regresi terpisah lebih baik dibandingkan taksiran regresi gabungan.

2. Untuk kasus 2, b_i untuk setiap stratum adalah :

   Untuk stratum ke – 1, $b_1=1.585$

Untuk stratum ke – 2, $b_2 = 1.582$

Untuk stratum ke – 3, $b_3 = 1.583$

Disini terlihat b_i relatif sama untuk ketiga stratum.

Dari analisa data didapat $\widehat{Var}\left(\hat{\mu}_{Yc}\right) = 0.233 < \widehat{Var}\left(\hat{\mu}_{Ys}\right) = 0.235$ atau dengan perkataan lain dalam hal ini taksiran regresi gabungan lebih baik dibandingkan taksiran regresi terpisah.

Dari hasil diatas dapat ditunjukkan terlihat bahwa jika b_i bervariasi antar stratum, maka taksiran regresi terpisah lebih baik dipergunakan bila dibandingkan dengan taksiran regresi gabungan. Demikian pula sebaliknya, jika b_i relatif sama antar stratum, maka taksiran regresi gabungan lebih sesuai dipergunakan bila dibandingkan dengan taksiran regresi terpisah.

BAB V

PENUTUP

Pada bab ini akan dibahas mengenai kesimpulan dan saran yang diperoleh hasil tugas akhir ini.

5.1 KESIMPULAN

Adapun kesimpulan yang diperoleh dari tugas akhir ini adalah :

1. Jika terdapat variabel *X* yang berkorelasi kuat dengan variabel *Y*, maka taksiran regresi dapat digunakan untuk mencari taksiran mean populasi dari *Y* berdasarkan *Stratified Random Sample*.
2. Taksiran regresi untuk mean populasi dari *Y* pada Stratified Random Sampling dapat dibedakan menjadi 2 (dua) yaitu :
   a. Taksiran regresi terpisah untuk mean populasi *Y* adalah :

$$\widehat{\mu}_{Ys} = \sum_{i=1}^{L} \frac{N_i}{N}\left(\overline{y}_i + b_i\left(\mu_{X_i} - \overline{x}_i\right)\right)$$

dengan

$$Var\left(\widehat{\mu}_{Ys}\right) \approx \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i}S_{Y_i}^2\left(1-\rho_i^2\right)$$

dan

$$\widehat{Var}\left(\widehat{\mu}_{Ys}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i - n_i}{N_i}\right)\frac{1}{n_i\left(n_i - 2\right)}\left(\sum_{j=1}^{n_i}\left(\left(y_{ij} - \overline{y}_i\right) - b_i\left(x_{ij} - \overline{x}_i\right)\right)^2\right)$$

dimana untuk sampel yang besar setiap stratumnya $\widehat{\mu}_{Ys}$ adalah taksiran yang tidak bias untuk μ_Y dan $\widehat{Var}\left(\widehat{\mu}_{Ys}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\widehat{\mu}_{Ys}\right)$

untuk $i = 1,2,...,L$, $j = 1,2,...,n_i$, $b_i = \dfrac{\sum_{j=1}^{n_i}\left(y_{ij}-\overline{y}_i\right)\left(x_{ij}-\overline{x}_i\right)}{\sum_{j=1}^{n_i}\left(x_{ij}-\overline{x}_i\right)^2}$, dan

$\rho_i^2 = \dfrac{S_{YX_i}}{S_{Y_i}S_{X_i}}$.

b. Taksiran regresi gabungan untuk mean populasi *Y* adalah :

$$\hat{\mu}_{Yc} = \overline{y}_{st} + b_c\left(\mu_x - \overline{x}_{st}\right)$$

dengan

$$Var\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\frac{1}{n_i}\left(S_{Y_i}^2 - 2b_cS_{YX_i}^2 + b_c^2S_{X_i}^2\right)$$

dan

$$\widehat{Var}\left(\hat{\mu}_{Yc}\right) = \sum_{i=1}^{3}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\frac{1}{n_i}\left(s_{Y_i}^2 - 2b_cs_{YX_i} + b_c^2s_{X_i}^2\right)$$ dimana

$\hat{\mu}_{Yc}$ adalah taksiran yang tidak bias untuk μ_Y dan $\widehat{Var}\left(\hat{\mu}_{Yc}\right)$ adalah taksiran yang tidak bias untuk $Var\left(\hat{\mu}_{Yc}\right)$

untuk $i = 1,2,...,L$, $j = 1,2,...,n_i$,

$$B_c = \frac{\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\left(\frac{1}{N_i\left(N_i-1\right)}\right)\sum_{j=1}^{N_i}\left(y_{ij}-\mu_{Y_i}\right)\left(x_{ij}-\mu_{X_i}\right)}{\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(1-\frac{n_i}{N_i}\right)\left(\frac{1}{N_i\left(N_i-1\right)}\right)\sum_{j=1}^{N_i}\left(x_{ij}-\mu_{X_i}\right)^2},$$

dan $b_c = \dfrac{\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{1}{n_i\left(n_i-1\right)}\right)\sum_{j=1}^{n_i}\left(y_{ij}-\overline{y}_i\right)\left(x_{ij}-\overline{x}_i\right)}{\sum_{i=1}^{L}\left(\frac{N_i}{N}\right)^2\left(\frac{N_i-n_i}{N_i}\right)\left(\frac{1}{n_i\left(n_i-1\right)}\right)\sum_{j=1}^{n_i}\left(x_{ij}-\overline{x}_i\right)^2}$.

3. Dalam contoh kasus terlihat bahwa jika nilai b_i bervariasi antar strata, maka taksiran regresi terpisah lebih baik daripada taksiran gabungan.

Sebaliknya jika nilai b_i relatif sama antar strata, maka taksiran regresi gabungan lebih tepat dipergunakan.

5.2 SARAN

Untuk selanjutnya tulisan ini dapat dikembangkan untuk mencari :

1. Taksiran regresi dalam teknik sampling yang lain, misalkan pada *Cluster Sampling*.
2. Taksiran regresi jika terdapat hubungan non linier antara *Y* dan *X*.
3. Taksiran regresi jika terdapat hubungan linier dan atau non linier antara *Y* dan beberapa variabel *X*.

DAFTAR PUSTAKA

Chaudhuri, Arijit dan Horst Stenger. 2005. *Survey Sampling : Theory and Methods*, Second edition. Florida : Chapman & Hall/CRC

Cochran, W.G. 1977. *Sampling Techniques*, Third edition. New York : Wiley

Fuller, Wayne A. 2009. *Sampling Statistics*. New Jersey : Wiley

Kish, Leslie. 1995. *Survey Sampling*. New York : John Wiley & Sons, Inc

Raj, Des. 1968. *Sampling Theory*, T M H Edition. New York : McGraw-Hill, Inc

Rao, Poduri S.R.S. 2000. *Sampling Methodologies : With Applications*. Florida : Chapman & Hall/CRC

Lampiran 1

Ketaksamaan Chebyshev

Misalkan X adalah variabel random dengan mean μ dan variansi σ^2. Maka untuk setiap $k > 0$

$$\Pr\left(\left|X-\mu\right| \geq k\sigma\right) \leq \frac{1}{k^2}$$

atau

$$\Pr\left(\left|X-\mu\right| < k\sigma\right) \geq 1-\frac{1}{k^2}$$

Lampiran 2

Konvergen Probabilitas

Misalkan $\{X_n\}$ adalah barisan variabel random dan X adalah variabel random. $\{X_n\}$ dikatakan konvergen probabilitas ke X jika untuk setiap $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr\left(\left|X_n - X\right| < \varepsilon\right) = 1$$

atau

$$\lim_{n\to\infty} \Pr\left(\left|X_n - X\right| > \varepsilon\right) = 0$$

Lampiran 3

Data Sampel untuk Contoh 1

y_1	x_1	y_2	x_2	y_3	x_3
318.6	111.485	152.75	63.0	116.711	29.648
326.7	115.173	118.115	47.253	111.432	26.001
304.2	104.375	124.306	49.767	110.974	25.705
306.9	116.213	124.597	51.51	110.459	25.798
266.4	94.811	132.974	56.472	112.076	27.984
313.2	118.026	136.347	59.412	108.729	23.904
331.2	118.933	136.964	60.75	111.604	25.305
273.6	100.116	138.585	60.75	110.461	25.707
353.88	119.891	134.264	55.884	111.003	25.323
351.9	124.466	134.734	57.48	115.452	29.299
439.2	154.179	122.566	42.3	110.32	24.356
288.0	109.548	122.754	45.165	109.041	23.047
309.6	99.446	133.1	56.7	109.648	23.342
267.3	94.309	137.889	60.417	111.922	24.703
347.4	138.447	136.833	60.768	105.805	26.149
302.4	124.093	126.708	48.768	110.8	23.28
351.0	141.119	134.102	55.902	109.793	23.458
315.0	111.676	127.488	49.44	110.248	22.98
427.5	161.001	131.34	53.61	110.865	23.408
383.4	135.895	125.076	43.692	113.776	27.688
324.0	125.882	132.881	51.072	115.490	29.96
325.8	127.4	130.414	49.5	109.198	21.175
356.22	128.996	124.952	46.614	107.715	20.218
245.88	97.465	152.198	66.75	114.958	30.391
249.3	92.796	150.312	68.688	114.960	30.532
284.4	100.638	151.725	68.808	117.933	33.567
273.6	98.319	138.681	51.48	114.287	27.29
324.0	119.705	130.825	44.025	111.348	25.37
287.64	103.153	125.656	44.04	113.465	26.972
269.82	90.776	130.022	49.86	111.135	22.876
270.9	91.202	139.851	55.362	117.284	30.395
288.9	101.159	125.8	42.885	110.057	22.331
298.8	112.606	147.396	61.893	118.453	33.12
302.4	123.11	143.011	60.39	115.419	29.912
296.64	121.53	144.90	62.085	115.863	29.358
297.0	107.728	133.088	48.768	113.783	26.437
282.42	106.36	146.364	58.995	115.244	27.473
286.02	102.566	157.395	68.034	116.323	29.478
278.1	100.391	149.118	65.823	117.109	29.175
257.4	87.755	136.671	54.108	109.624	21.745
299.7	114.19	153.562	67.467	115.116	28.744
263.7	91.918	133.545	52.074	112.677	25.306
286.2	99.691	129.098	49.956	108.477	20.361
261.0	96.943	134.573	53.424	110.47	22.868
271.8	100.409	129.929	50.4	109.177	20.002

(lanjutan)

y_1	x_1	y_2	x_2	y_3	x_3
290.16	121.488	116.215	41.127	107.276	18.132
281.7	101.116	117.225	39.414	112.706	23.956
265.5	91.9	122.72	43.353	107.945	19.096
315.54	117.058	120.188	39.606	112.857	24.922
354.6	129.205	123.788	44.523	112.854	23.696
356.4	121.871	121.666	43.269	112.756	23.142
375.3	131.789	122.854	43.353	109.852	20.811
315.0	121.844	139.013	56.976	120.28	32.88
369.0	140.466	124.737	43.935	110.404	21.365
260.82	92.941	124.676	42.225	110.758	21.601
251.82	83.786	159.656	65.16	112.57	26.044
252.0	83.786	140.469	56.157	119.586	33.394
314.1	117.88	121.941	43.629	109.17	22.676
367.2	139.625	163.086	70.2	121.725	34.253
241.02	85.426	131.573	50.16	113.146	24.938
292.5	114.672	127.603	42.972	113.15	24.718
238.5	95.311	144.518	65.769	116.713	28.555
270.9	93.892	159.577	74.025	119.181	30.275
211.5	76.761	142.793	60.576	114.724	26.32
311.4	101.587	156.941	72.603	124.519	39.109
235.8	77.722	145.496	62.559	113.091	24.562
282.6	104.952	146.541	59.94	114.264	25.52
225.0	89.189	128.407	48.384	112.138	22.462
270.54	91.224	158.915	74.724	120.542	31.39
297.0	115.044	117.301	41.127	109.140	18.005
275.4	99.487	130.778	46.62		
276.3	94.295	117.732	39.414		
301.14	99.293	119.974	39.414		
284.4	94.16	129.356	49.446		
284.4	101.832	121.651	44.652		
284.22	102.115				
281.88	102.756				
295.2	102.674				
271.8	105.283				
302.4	105.01				
270.18	93.513				
274.14	115.399				
268.2	94.62				
304.2	114.179				
310.5	113.033				
304.2	110.06				
288.0	107.288				
275.4	102.214				
372.24	121.728				
252.54	103.011				
296.64	104.485				
316.8	123.569				
342.0	131.707				

(lanjutan)

y_1	x_1	y_2	x_2	y_3	x_3
320.4	115.599				
314.82	101.917				
315.0	127.948				
306.0	114.334				
300.6	128.903				
291.24	105.241				
316.26	121.053				
259.2	103.559				
337.32	116.152				
295.2	107.099				
302.4	116.969				
339.84	118.327				

Lampiran 4

Data Sampel untuk Contoh 2

y_1	x_1	y_2	x_2	y_3	x_3
256.54	111.485	157.577	55.884	202.09	27.958
262.255	115.173	158.186	57.48	200.59	26.305
260.35	120.656	142.439	42.3	201.9	26.327
265.43	118.933	164.986	61.632	189.964	20.03
217.678	98.153	142.682	45.165	201.469	28.016
224.79	100.116	149.526	51.09	198.14	25.323
235.331	101.722	156.071	56.7	206.082	29.299
241.3	105.614	143.575	50.085	196.921	24.356
226.06	101.382	147.798	48.768	201.295	27.944
280.035	124.466	155.751	53.976	194.638	23.047
318.135	136.485	148.808	49.44	196.016	22.779
269.113	124.675	153.793	53.61	195.723	23.342
341.63	154.179	145.687	43.692	196.173	22.446
250.19	99.446	144.906	47.679	188.862	26.149
297.18	136.181	155.788	51.072	197.779	23.28
245.364	124.0	161.699	54.927	195.981	23.458
220.345	94.309	152.594	49.5	200.224	26.301
212.725	86.323	149.831	46.716	204.412	27.364
245.11	124.093	180.785	66.75	203.09	27.688
279.4	141.119	178.345	68.688	197.625	23.254
254.0	111.676	173.096	63.342	194.918	21.175
333.375	161.001	163.294	51.48	195.665	22.192
302.26	135.895	150.218	44.835	205.2	30.391
331.724	86.323	153.127	44.025	205.204	30.532
207.391	93.873	146.438	44.04	200.334	25.409
205.232	97.465	152.088	49.86	199.99	25.64
232.41	100.638	164.624	52.908	198.376	22.876
266.065	127.667	168.897	60.39	206.009	27.806
258.953	122.714	172.594	62.085	209.352	30.395
260.35	119.705	173.236	58.995	196.453	22.331
223.52	104.066	176.8	65.823	206.815	29.358
262.89	103.816	161.911	55.989	206.774	28.944
237.998	103.43	150.892	49.956	203.103	26.437
233.68	100.296	173.993	60.543	205.711	27.473
235.585	101.159	151.967	50.4	207.637	29.478
241.3	107.728	152.374	48.765	205.483	28.744
239.014	108.974	156.534	48.0	199.441	25.782
233.553	102.566	135.527	39.414	195.53	21.801
243.205	114.19	140.174	41.127	200.643	24.632
217.805	91.918	147.961	45.18	194.881	20.002
234.95	104.589	144.02	44.523	191.488	18.132
232.41	104.731	155.429	49.923	201.307	23.268
215.9	96.943	153.936	45.96	201.180	23.956
257.175	111.771	190.437	65.16	192.683	19.096
231.14	98.172	165.607	56.157	201.445	23.69

(lanjutan)

y_1	x_1	y_2	x_2	y_3	x_3
223.52	100.409	141.629	43.629	201.27	23.142
250.825	115.88	173.14	58.158	196.087	20.811
230.505	101.116	194.876	70.2	196.87	21.184
254.381	117.058	154.095	50.16	197.072	21.365
243.205	109.943	155.754	52.578	197.704	21.601
262.89	113.66	148.957	42.972	213.462	33.394
281.94	129.205	170.847	65.769	194.869	22.676
283.21	121.871	168.614	60.576	201.965	24.938
265.43	111.77	160.668	54.399	208.333	28.555
295.91	133.468	186.923	72.603	204.783	26.32
296.545	131.789	172.112	62.559	192.891	21.842
254.0	121.844	162.481	54.654	203.962	25.52
288.925	124.573	150.134	46.8	192.4	21.356
215.773	92.941	133.654	41.127	203.826	25.889
209.423	83.786	189.479	74.724	200.386	23.532
209.55	83.786	135.625	41.127	201.766	24.258
253.365	117.88	156.204	52.878	198.31	20.306
299.593	138.046	174.223	58.476	215.167	31.39
285.75	133.744	151.226	49.446	196.036	18.022
222.885	93.892	142.081	42.888	201.976	22.386
198.12	77.722	148.956	46.047	220.392	32.104
222.631	91.224	141.255	44.652	210.122	25.134
241.3	115.044	184.195	69.765	215.039	28.454
239.395	100.195	181.074	67.836	215.839	29.486
226.06	99.487	157.209	56.4	223.069	33.783
234.95	100.269	185.816	67.5		
226.695	94.295	173.938	62.715		
244.221	99.293	193.465	74.178		
232.41	94.16	170.827	59.064		
246.126	101.401	150.777	48.768		
232.41	101.832				
204.597	95.413				
240.03	102.674				
245.11	105.01				
234.95	101.009				
222.377	93.513				
241.3	100.776				
211.963	90.749				
292.1	119.767				
255.27	117.526				
246.38	110.06				
250.698	113.305				
209.931	103.011				
241.046	104.485				
273.05	131.707				
257.81	115.599				
253.873	101.917				
271.145	123.508				
276.86	128.477				

(lanjutan)

y_1	x_1	y_2	x_2	y_3	x_3
264.16	122.738				
237.236	105.241				
254.889	121.053				
264.16	119.057				
290.83	130.058				
232.918	100.433				
269.748	116.152				
240.03	107.099				
271.526	118.327				
232.41	106.274				
239.395	105.177				