PENDUGAAN PARAMETER PADA MODEL REGRESI LINIER SEDERHANA DENGAN VARIABEL X MENGANDUNG KESALAHAN PENGUKURAN

SKRIPSI

Diajukan untuk Memenuhi Persyaratan Penyelesaian Program Sarjana Sains
Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Jember

Oleh :

Adiyanto Juniawan
NIM : 971810101031

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
MEI, 2003

MOTTO :

Karya termasyur harus dibaca oleh siapa saja yang mencintai ketelitian dalam penyelidikan matematis.

(Neils Hendrik Abel)

Saya tidak tahu bagaimana Saya tampak pada dunia, tetapi bagi Saya sendiri Saya nampaknya hanyalah seperti seorang anak laki-laki yang bermain-main di pantai, dan mengalihkan diri sendiri sekarang dan kemudian menemukan koral yang lebih haus atau kerang yang lebih indah daripada yang biasa, sementara samudera besar dari kebenaran semuanya terbentang dihadapan Saya tak terungkapkan.

(Issac Newton)

Berikan Aku Sepuluh Pemuda, Niscaya Akan Aku Goncang Dunia.

(Ir. Soekarno)

Kupersembahkan buat :

- Ayahanda ***Supardi*** dan Ibunda ***Sumiati***
- Kakakku *Mbak Wati* dan Suaminya, *Mbak Ari* dan Suaminya, ***Mas Hariyanto*** dan Istrinya, Adik-adikku Si *Kembar Tutik* dan *Puri* dan Si Ragil *Oon*, serta tidak lupa Keponakanku ***Dana* dan *Dirga***
- Keluarga Besar di *Bangsalsari* dan *Jember*
- Keluarga Besar *HIMATIKA* dan *MIPA*
- Sahabat-sahabat sebangsa dan setanah air yang berpikir merdeka!

DEKLARASI

Skripsi ini berisi hasil kerja / penelitian mulai bulan Maret 2002 sampai dengan bulan November 2002. Bersama ini saya nyatakan bahwa isi skripsi ini adalah hasil pekerjaan saya sendiri keculai jika disebutkan sumbernya dan skripsi ini belum pernah diajukan pada institusi lain.

Jember, Mei 2003

Penulis

(Adiyanto Juniawan)

ABSTRAK

Pendugaan Parameter Pada Model Regresi Linier Sederhana Dengan Variabel X Mengandung Kesalahan Pengukuran. ADIYANTO JUNIAWAN, 971810101031. Skripsi, Mei 2003, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Jember.

Dalam percobaan yang melibatkan lebih dari satu variabel, model linier memegang peranan penting, yaitu melakukan analisis hubungan antar variabel. Model regresi linier yang paling sederhana adalah model regresi linier sederhana yang hanya melibatkan satu variabel respon atau variabel independen (Y) dan satu variabel independen (X). Model regresi linier sederhana, diasumsikan variabel X nilainya bersifat tetap dan diukur dengan mengabaikan kesalahan pengukuran. Penduga parameter dicari dengan menggunakan metode kuadrat terkecil. Di lapangan sering dijumpai variabel X mengandung kesalahan pengukuran. Tujuan dari penelitian ini adalah untuk mempelajari perbandingan mean dan standar deviasi penduga parameter antara model regresi biasa dengan variabel X diukur dengan mengabaikan kesalahan pengukuran dengan model Berkson yang variabel X mengandung kesalahan pengukuran. Penelitian ini dilakukan dengan cara menurunkan model Berkson dengan menggunakan metode kuadrat terkecil untuk mendapatkan penduga parameternya. Dengan bantuan paket statistika S-Plus digunakan untuk mendapatkan mean dan standar deviasi penduga parameter dari kedua model, kemudian dipergunakan untuk membandingkan kedua model tersebut. Dari hasil perbandingan, selanjutnya dilakukan uji beda dua rata-rata dengan uji statistiknya menggunakan distrubusi Z untuk melihat apakah ada perbedaan mean penduga parameter antara kedua model. Dari hasil penelitian, diperoleh bahwa secara statistik tidak ada beda rata-rata antara model regresi biasa dengan model Berkson.

Kata kunci : model regresi, model Berkson, parameter, mean, standar deviasi.

PENGESAHAN

Skripsi ini telah dipertahankan di depan Tim penguji dan diterima oleh Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember pada :

Hari : SELASA

Tanggal : 13 MAY 2003

Tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember

Tim Penguji.

Ketua	Sekretaris
(Dosen Pembimbing Utama)	(Dosen Pembimbing Anggota)
Drs. I Made Tirta, MSc, PhD.	Rita Ratih T, SSi, MSi.
NIP. 131 474 500	NIP. 132 243 343

Anggota,

Yuliani S. Dewi, SSi, MSi.	M. Fatekurohman, SSi, MSi
NIP. 132 258 183	NIP. 132 210 538

Mengesahkan,

Dekan FMIPA Universitas Jember

DEPARTEMEN PENDIDIKAN NASIONAL UNIVERSITAS JEMBER FAKULTAS

Ir. Sumadi, MS

NIP. 130 368 784

KATA PENGANTAR

Dengan mengucapkan rasa puji syukur kepada Tuhan Yang Maha Kuasa, penulis panjatkan atas rahmat dan hidayah-Nya yang telah diberi kekuatan untuk menyelesaikan penulisan skripsi ini.

Penulis menyampaikan rasa terima kasih yang sebesar-besarnya dan penghargaan yang setinggi-tingginya kepada yang terhormat :

1. Ir. Sumadi, MS selaku Dekan Fakultas MIPA Universitas Jember.
2. Drs. Kusno, DEA, PhD selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Jember.
3. Drs. I Made Tirta, MSc, PhD Selaku Dosen Pembimbing Utama yang telah memberikan petunjuk dan bimbingan selama penulisan skripsi ini.
4. Rita Ratih T, SSi, MSi Selaku Dosen Pembimbing Anggota yang telah memberikan petunjuk dan bimbingan selama penulisan skripsi ini.
5. Yuliani S. Dewi, SSi, MSi dan M. Fatekurohman, SSi, MSi yang telah memberikan saran dan masukan serta petunjuk dalam perbaikan penulisan skripsi ini.
6. Ayahanda dan Ibunda serta Kakak dan Adik-adikku yang saya hormati dan cintai, yang telah memberikan dorongan dan dukungan bail materiil maupun spirituil selama ini.
7. Serta rekan-rekan HIMATIKA dan MIPA yang senasib dan seperjuangan yang sudi memberikan saran dan kritiknya.

Akhirnya, penulis berharap skripsi ini dapat memberikan kontribusi terhadap kemajuan ilmu pengetahuan, khususnya bidang ilmu statistika.

Jember, Mei 2003

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

I. PENDAHULUAN

1.1 Latar Belakang

Dalam percobaan yang melibatkan lebih dari satu variabel, model linier memegang peranan penting, yaitu melakukan analisis hubungan antar variabel (Ratih, 2000). Model linier yang paling sederhana yang sudah dikenal adalah model regresi linier sederhana yang hanya melibatkan satu variabel respon atau variabel dependen dan satu variabel independen. Untuk selanjutnya, model linier ini disebut model regresi linier sederhana atau model linier klasik.

Persamaan umum untuk model regresi linier sederhana adalah

$$Y = \beta_0 + \beta_1 X + \varepsilon \qquad \text{.... (1.1)}$$

dengan:

Y adalah variabel respon;

β_0 dan β_1 adalah parameter yang tidak diketahui;

X adalah variabel independen yang diasumsikan nilainya bersifat tetap dan diukur dengan mengabaikan kesalahan pengukuran;

ε adalah variabel random kesalahan (*error*) dengan asumsi $E(\varepsilon) = 0$ dan $var(\varepsilon) = \sigma^2$.

Dari model regresi (1.1), variabel X diasumsikan nilainya bersifat tetap dan diukur dengan mengabaikan kesalahan pengukuran. Sebagai contoh, dalam bidang pertanian, misalkan bahwa hasil produksi padi dihubungkan dengan tingkatan pemberian dosis pemupukan. Analisis regresi dapat digunakan untuk membuat sebuah model yang menggambarkan hasil sebagai sebuah fungsi tingkat pemupukan. Model tersebut dapat digunakan untuk meramal pada sebuah tingkat pemupukan tertentu. Ini dapat juga digunakan untuk tujuan optimalisasi atau tujuan proses kontrol.

Dalam banyak bidang penelitian, seringkali dijumpai bahwa variabel X mengandung kesalahan pengukuran, misalkan kesalahan pengukuran pada saat penimbangan berat dosis pupuk seperti contoh di atas. Sebagai ilustrasi, misalkan ingin diketahui apakah produksi padi dipengaruhi oleh pemberian dosis pemupukan Nitrogen (N) sampai suatu tingkatan pemupukan tertentu? Misalkan pemberian dosis pemupukannya adalah 1 kg, 2 kg, 3 kg, dan seterusnya. Karena kesalahan pengukuran mungkin terjadi pada saat penimbangan dosis pupuk Nitrogen tadi, maka berat yang tercatat bisa jadi lebih dari atau kurang dari berat dosis pupuk yang ditetapkan. Sehingga yang akan diterima oleh tanaman atau padi tersebut, akan bias jadi lebih dari atau kurang dari yang tercatat.

Dalam kesalahan pengukuran ini, variabel X disini juga merupakan variabel yang diamati. Sehingga dalam hal ini, variabel X sesungguhnya bukan lagi bersifat tetap (*fixed*) tetapi sudah bersifat acak (random). Jadi, apakah dengan adanya kesalahan pengukuran ini, model regresi (1.1) masih dapat digunakan untuk mengetahui atau meramalkan pengaruh pemberian dosis pemupukan Nitrogen sampai tingkatan tertentu terhadap produksi padi tersebut? Menurut Hodges dan Moore (1972), dinyatakan bahwa pendugaan parameter untuk model regresi linier yang variabel independennya mengandung kesalahan pengukuran menghasilkan bias dan variabel independennya berkorelasi dengan variabel random kesalahan (Weisberg, 1985). Oleh karena itu, asumsi untuk model regresi (1.1) perlu terlebih dahulu diubah untuk disesuaikan dengan data. Kemudian data ini digunakan untuk melakukan pendugaan parameter yang tidak diketahui. Penduga parameter dicari dengan menggunakan metode kuadrat terkecil.

Model regresi (1.1) perlu dikembangkan, sehingga diharapkan pendugaan parameter menjadi takbias dan variabel independen tidak berkorelasi dengan variabel random kesalahannya. Berkson [1950] menunjukkan langkah penting dalam menyelesaikan kasus dengan variabel X mengandung kesalahan pengukuran (Montgonery dan Peck, 1991). Pendekatan yang dilakukan oleh Berkson adalah dengan mengasumsikan bahwa variabel yang diamati X_i^* nilainya bersifat tetap (*fixed*) atau sebagai nilai target, sedangkan nilai sesungguhnya, yaitu X_i sebagai variabel random. Sebagai contoh, misalkan dalam penimbangan berat dosis pupuk dalam produksi padi seperti contoh di atas tadi. Dalam mencatat berat dosis pupuk tersebut, mungkin pencatat kurang teliti atau kurang akurat dalam melakukan pencatatan berat dosis pupuk seperti yang ditunjukkan oleh timbangan.

Sehingga kesalahan pengukuran mungkin terjadi disini. Dalam hal ini, berat dosis pupuk, yaitu 1 kg, 2 kg, 3 kg, dan seterusnya ditetapkan sebagai nilai target atau diasumsikan nilainya bersifat tetap, sedangkan nilai sesungguhnya diasumsikan bersifat random. Jika variabel yang diamati X_i^* diasumsikan nilainya bersifat tetap dan variabel sesungguhnya X diasumsikan nilainya bersifat random, maka model Berkson dapat dicari penduga parameternya sebagaimana model regresi(1.1). Sehingga dengan penggunaan asumsi tersebut, pendugaan parameter pada model regresi (1.1) dengan variabel independennya mengandung kesalahan pengukuran dapat diselesaikan.

Tulisan ini hanya dibatasi hanya pada masalah perbandingan mean dan standar deviasi pendugaan parameter antara model regresi biasa dengan model Berkson dengan variabel independennya mengandung kesalahan pengukuran.

1.2 Permasalahan

Permasalahan yang dihadapi sebagai berikut ini.

1. Bagaimana perbandingan mean dan standar deviasi penduga parameter antara model regresi biasa dengan model Berkson?
2. Apakah secara matematis dan statistik ada perbedaaan mean parameter antara model regresi biasa dengan model Berkson?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini sebagai berikut.

1. Untuk mengetahui perbandingan mean dan standar deviasi penduga parameter antara model regresi biasa dengan model Berkson.
2. Untuk mengetahui secara matematis dan statistik perbedaan mean penduga parameter antara model regresi biasa dengan model Berkson.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini sebagai berikut.

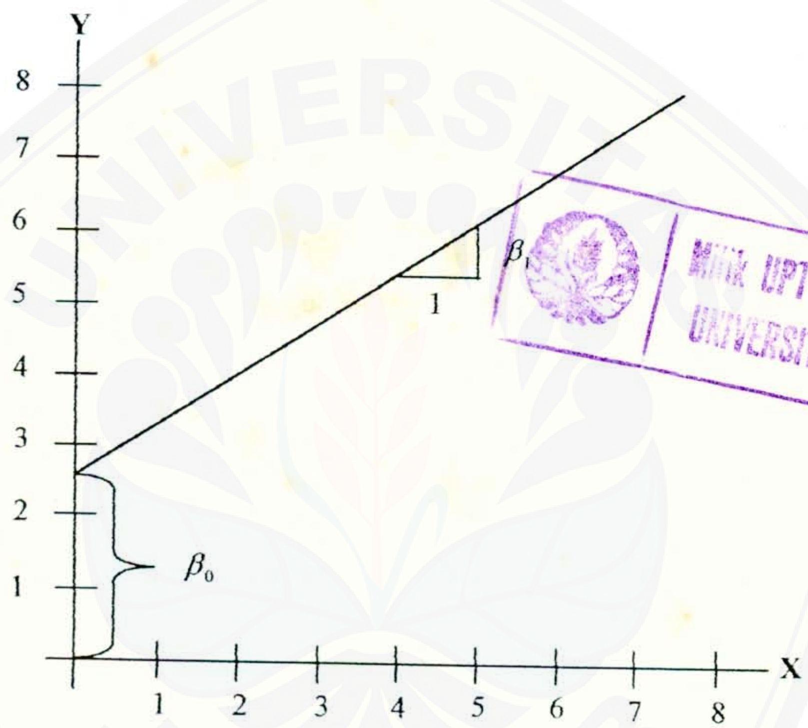
1. Dapat memberi suatu alternatif penyelesaian secara langsung terhadap masalah-masalah dalam model regresi linier.

2. Mengembangkan model regresi linier biasa dengan variabel independennya mengandung kesalahan pengukuran.

II. TINJAUAN PUSTAKA

2.1 Model Regresi Linier Sederhana

Model regresi linier sederhana, menurut Montgomery dan Peck (1991), adalah model regresi yang hanya didasarkan pada satu variabel independen tunggal X dan sebuah variabel respon/dependen Y yang membentuk sebuah garis lurus, seperti seperti diillustrasikan pada Gambar 2.1 di bawah ini.



Gambar 2.1 Illustrasi Konstanta dan Gradien

Dari model regresi (1.1) sebelumnya, model regresi dapat dinyatakan seperti di bawah ini. Misalkan untuk respon ke-i sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad \text{untuk } i = 1,2,\ldots,n \qquad \text{.... (2.1)}$$

Model (2.1) diasumsikan bahwa antara variabel random kesalahan tidak berkorelasi, artinya antara nilai kesalahan satu tidak bergantung pada nilai kesalahan lainnya (Montgomery dan Peck, 1991). Di dalam model regresi (2.1), variabel X diasumsikan nilainya bersifat tetap (*fixed*) dan diukur dengan mengabaikan kesalahan pengukuran (Weisberg, 1985), sehingga variabel independen tidak berkorelasi dengan variabel random kesalahannya. Secara formal untuk respon ke-i dengan $i = 1,2,...,n$, model regresi (2.1) mempunyai sifat khusus seperti berikut ini.

1. Model: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, untuk $i = 1,2,...,n$, atau untuk keseluruhan respon dapat dituliskan dalam bentuk matriks $\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$.
2. Asumsi: $\varepsilon_i \sim NID(0,\sigma^2)$ dan ε_i independen dengan ε_j untuk setiap $i \neq j$.

Konsekuensinya dari asumsi (1) dan (2), variabel respon Y_i juga berdistribusi normal dengan mean $\beta_0 + \beta_1 X_i$ dan varians konstan σ^2, yaitu:

$$Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2) \quad .... (2.2)$$

(Tirta, 2000).

Dari model regresi (2.1), maka dapat diperjelas lagi sebagai berikut.

1. Mean Y_i adalah fungsi linier dari X_i

$$E(Y_i) = \beta_0 + \beta_1 X_i = \mu_i, \text{ untuk } i = 1,...,n \quad .... (2.3a)$$

2. Varians Y_i tidak bergantung pada X_i, yaitu:

$$var(Y_i) = \sigma^2 \quad .... (2.3b)$$

Untuk model regresi (2.1), syarat yang diminta oleh model regresi (2.1) rumusannya dapat dipertegas lagi, yaitu:

1. Hubungan antara Y dan X adalah linier;
2. Respon ke-i dan j adalah saling independen, artinya tidak ada korelasi antara satu respon dengan respon lainnya;
3. Variabel independen dan variabel random kesalahan adalah saling independen (bebas), karena variabel independen diasumsikan nilainya bersifat tetap dan diukur dengan mengabaikan kesalahan pengukuran.

Dari uraian di atas, pendugaan parameter model regresi (2.1) dapat dicari dengan menggunakan metode kuadrat terkecil dengan asumsi yang dikenakan pada model regresi (2.1).

2.2 Metode Pendugaan Parameter

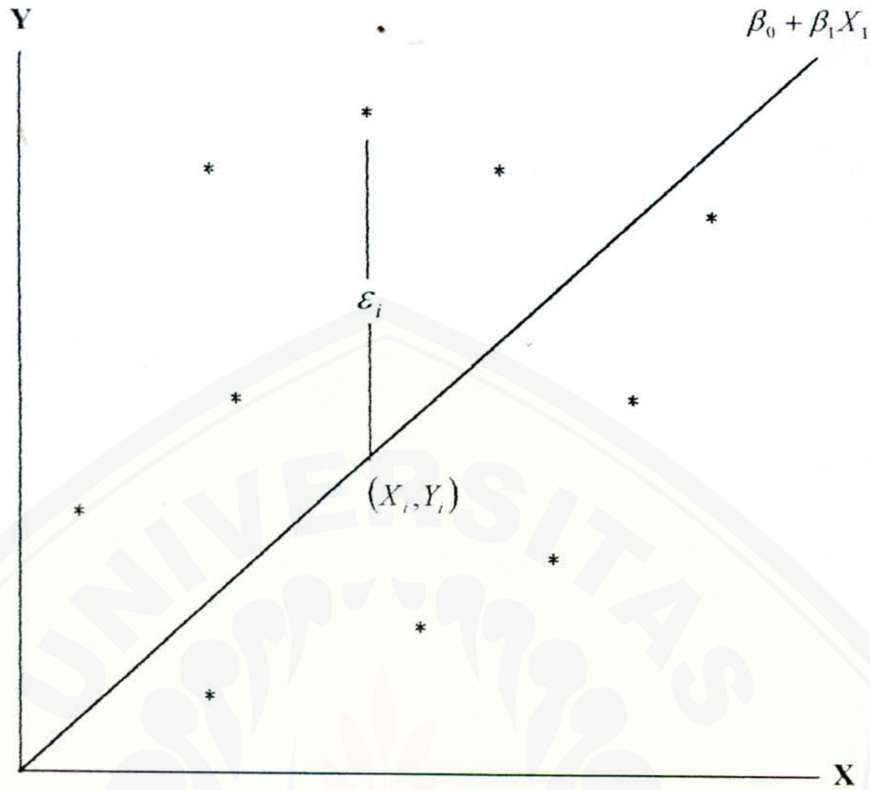
Salah satu langkah pokok dalam pemodelan statistika adalah melakukan pendugaan parameter yang belum diketahui (Tirta, 2000). Dalam model regresi (2.1) ada dua kelompok parameter yang belum diketahui yaitu parameter regresi β_0 dan β_1, serta parameter dispersi, σ^2. Adapun metode yang digunakan untuk melakukan pendugaan parameter dalam model regresi (2.1), adalah metode kuadrat terkecil.

2.2.1 Pendugaan Parameter Model Regresi Linier Sederhana dengan Metode Kuadrat Terkecil

Metode kuadrat terkecil digunakan untuk meminimumkan jumlah kuadrat selisih antara nilai variabel respon (Y_i) dengan nilai harapannya $E(Y_i)$. Dari model regresi (2.1), diketahui $E(Y_i) = \beta_0 + \beta_1 X_i$ dan $E(\varepsilon_i) = 0$, sehingga jumlah kuadrat selisih antara nilai variabel respon (Y_i) dengan nilai harapannya $E(Y_i)$ adalah

$$S = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2$$

Secara geometris S adalah jumlah kuadrat jarak vertikal dari titik (X_i, Y_i) ke garis lurus $\beta_0 + \beta_1 X$, seperti diilustrasikan pada Gambar 2.2.



Gambar 2.2 Jarak Titik (X_i, Y_i) ke Garis Lurus $\beta_0 + \beta_1 X$

Secara aljabar, S adalah jumlah kuadrat kesalahan, yaitu:

$$S = \sum_{i=1}^{n} \varepsilon_i^2$$

Penduga kuadrat terkecil untuk parameter β_0 dan β_1 diturunkan dengan meminimumkan S, dan nilai minimum S adalah mengukur kecocokan model. Keuntungan menggunakan metode kuadrat terkecil adalah tidak memerlukan asumsi secara detail (Dobson, 1983), misalnya asumsi untuk variabel kesalahan adalah berdistribusi normal. Dari model regresi (2.1), jumlah kuadrat kesalahan adalah

$$S = \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i\right)^2$$

Penduga kuadrat terkecil untuk parameter β_0 dan β_1 adalah penyelesaian dari

$$\frac{\partial S}{\partial \beta_0} = -2\sum_{i=1}^{n}\left(Y_i - \beta_0 - \beta_1 X_i\right) = 0 \qquad \text{.... (2.4)}$$

$$\frac{\partial S}{\partial \beta_1} = -2\sum_{i=1}^{n} X_i\left(Y_i - \beta_0 - \beta_1 X_i\right) = 0 \qquad \text{.... (2.5)}$$

Dari persamaan (2.4) diperoleh

$$\sum_{i=1}^{n} Y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} X_i = 0 \qquad \text{.... (2.6)}$$

dan dari persamaan (2.5) diperoleh

$$\sum_{i=1}^{n} X_i Y_i - \beta_0 \sum_{i=1}^{n} X_i - \beta_1 \sum_{i=1}^{n} X_i^2 = 0 \qquad \text{.... (2.7)}$$

Persamaan (2.6) dan (2.7) disebut persamaan normal kuadrat terkecil. Penyelesaian secara simultan persamaan (2.6) dan (2.7) akan diperoleh nilai penduga parameter, yaitu $\hat{\beta}_0$ dan $\hat{\beta}_1$. Untuk penyelesaian nilai penduga $\hat{\beta}_0$ adalah

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} \qquad \text{.... (2.8)}$$

untuk disubstitusikan ke dalam persamaan (2.7), menghasilkan

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i^2\right)^2}{n}} \qquad \text{.... (2.9)}$$

dengan $\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$ dan $\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ adalah mean Y_i dan X_i.

Dengan demikian, nilai penduga $\hat{\beta}_0$ dan $\hat{\beta}_1$ pada persamaan (2.8) dan (2.9) berhubungan dengan titik-titik turunan pertama persamaani (2.6) dan (2.7) adalah nol, yaitu pada saat jumlah kuadrat kesalahan adalah minimum. Persamaan garis regresi model regresi linier sederhana adalah:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X \qquad \text{.... (2.10)}$$

dengan $\hat{\beta}_0$ dan $\hat{\beta}_1$ adalah penduga kuadrat terkecil.

Secara notasi, ini baik untuk memberikan simbol khusus pada pembilang dan penyebut pada persamaan (2.9), yaitu:

$$S_{XX} = \sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i^2\right)^2}{n} = \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \qquad \text{.... (2.11)}$$

dan

$$S_{XY} = \sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n} = \sum_{i=1}^{n} Y_i \left(X_i - \overline{X}\right) \qquad \text{.... (2.12)}$$

Dengan demikian, untuk notasi baru ini, penduga $\hat{\beta}_1$ pada persamaan (2.9) menjadi:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i \left(X_i - \overline{X}\right)}{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2} = \frac{S_{XY}}{S_{XX}} \qquad \text{.... (2.13)}$$

2.2.2 Sifat-Sifat Kuadrat Terkecil Penduga Parameter

Model regresi linier sederhana menyatakan bahwa $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, dengan β_0 dan β_1 adalah konstanta yang tetap (parameter), sedangkan variabel X_i adalah variabel independen yang bersifat tetap dan diukur dengan mengabaikan kesalahan pengukuran. Variabel random kesalahan ε_i adalah bebas satu sama lain dan dengan asumsi mean sama dengan nol dan varians konstans σ^2. Dari persamaan garis regresi (2.10), maka penduga kuadrat terkecil $\hat{\beta}_0$ dan $\hat{\beta}_1$ sekarang merupakan variabel random. Dengan demikian, penduga kuadrat terkecil $\hat{\beta}_0$ dan $\hat{\beta}_1$ dapat dicari mean dan varians serta kovariansnya antara kedua penduga tersebut.

a. Mean Dan Varians Penduga $\hat{\beta}_0$

Dari persamaan (2.8), diperoleh bahwa:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

Dengan mensubstitusikan $\hat{\beta}_1$ dari persamaan (2.13), diperoleh

$$\hat{\beta}_0 = \bar{Y} - \frac{\sum_{i=1}^{n} Y_i (X_i - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \bar{X} \qquad \text{.... (2.14)}$$

Misalkan $b_i = \dfrac{(X_i - \bar{X})}{S_{XX}}$ dan $\bar{Y} = \dfrac{\sum_{i=1}^{n} Y_i}{n}$. Dengan demikian persamaan (2.14) menjadi:

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} Y_i}{n} - \sum_{i=1}^{n} (b_i Y_i) \bar{X}$$

$$= \sum_{i=1}^{n} \left(\frac{1}{n} - b_i \bar{X} \right) Y_i$$

Karena $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, maka diperoleh

$$\hat{\beta}_0 = \sum_{i=1}^{n} \left(\frac{1}{n} - b_i \bar{X} \right) (\beta_0 + \beta_1 X_i + \varepsilon_i)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \beta_0 + \frac{\beta_1}{n} \sum_{i=1}^{n} X_i + \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i$$

$$- \bar{X}\beta_0 \sum_{i=1}^{n} b_i - \bar{X}\beta_1 \sum_{i=1}^{n} b_i X_i - \bar{X} \sum_{i=1}^{n} b_i \varepsilon_i$$

Karena $\sum_{i}^{n} b_i = 0$ dan $\sum_{i}^{n} b_i X_i = 1$, maka

$$\hat{\beta}_0 = \frac{n}{n} \beta_0 + \beta_1 \bar{X} + \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i - \bar{X}\beta_0 (0) - \bar{X}\beta_1 \sum_{i=1}^{n} b_i X_i - \bar{X} \sum_{i=1}^{n} b_i \varepsilon_i$$

$$= \beta_0 + \beta_1 \bar{X} - \beta_1 \bar{X} + \sum_{i=1}^{n} \left(\frac{1}{n} \varepsilon_i - b_i \bar{X} \varepsilon_i \right)$$

$$= \beta_0 + \sum_{i=1}^{n} \left(\frac{1}{n} - b_i \bar{X} \right) (\varepsilon_i)$$

atau

$$\hat{\beta}_0 - \beta_0 = \sum_{i=1}^{n}\left(\frac{1}{n} - b_i\overline{X}\right)(\varepsilon_i)$$

dan

$$E\left(\hat{\beta}_0 - \beta_0\right) = E\left[\sum_{i=1}^{n}\left(\frac{1}{n} - b_i\overline{X}\right)(\varepsilon_i)\right].$$

Karena $\sum_{i=1}^{n}\left(\frac{1}{n} - b_i\overline{X}\right)$ adalah suatu konstanta, maka

$$E\left(\hat{\beta}_0 - \beta_0\right) = \sum_{i=1}^{n}\left(\frac{1}{n} - b_i\overline{X}\right)E(\varepsilon_i) \qquad \text{.... (2.15)}$$

Dengan demikian $E\left(\hat{\beta}_0 - \beta_0\right) = 0$, karena diasumsikan bahwa $E(\varepsilon_i) = 0$. Ini menunjukkan bahwa $\hat{\beta}_0$ adalah penduga takbias dari β_0. Varians $\hat{\beta}_0$ dinyatakan sebagai $E\left(\hat{\beta}_0 - \beta_0\right)^2$, dengan β_0 adalah nilai sebenarnya pada model regresi $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$. Diperoleh varians $\hat{\beta}_0$ adalah

$$\begin{aligned} var\{\hat{\beta}_0\} &= E\left(\hat{\beta}_0 - \beta_0\right)^2 \\ &= E\left[\sum_{i=1}^{n}\left(\frac{1}{n} - b_i\overline{X}\right)(\varepsilon_i)\right]^2 \quad \text{dari persamaan (2.15)} \end{aligned}$$

Misalkan $k_i = \left(\frac{1}{n} - b_i\overline{X}\right)$, maka

$$\begin{aligned} var\{\hat{\beta}_0\} &= E(k_i\varepsilon_i)^2 \\ &= E\left(k_1^2\varepsilon_1^2 + k_2^2\varepsilon_2^2 + \ldots + k_n^2\varepsilon_n^2 + \ldots + 2k_1k_2\varepsilon_1\varepsilon_2 + \right. \\ &\quad \left. 2k_2k_3\varepsilon_2\varepsilon_3 + \ldots + 2k_{n-1}k_n\varepsilon_{n-1}\varepsilon_n\right) \\ &= E\left(\sum_{i=1}^{n} k_i^2\varepsilon_i^2 + 2\sum_{i \neq j}\sum k_i k_j \varepsilon_i\varepsilon_j\right) \\ &= \sum_{i=1}^{n} k_i^2 E\left(\varepsilon_i^2\right) + 2\sum_{i \neq j}\sum k_i k_j E\left(\varepsilon_i\varepsilon_j\right) \end{aligned}$$

$$= \left[\frac{1}{n^2} - \frac{2}{n}b_i\overline{X} + b_i^2 X^2\right]E\left(\varepsilon_i^2\right)$$

Dari definisi $E\left(\varepsilon_i\varepsilon_j\right) = 0$, untuk $i \neq j$, maka

$$\sigma^2\{\hat{\beta}_0\} = \left(\frac{1}{n} - 0 + \frac{\overline{X}^2}{S_{XX}}\right) E(\varepsilon_i^2).$$

Karena $\sum_{i=1}^{n} b_i (X_i - \overline{X}) = 1$ dan $E(\varepsilon_i^2) = \sigma^2$, maka diperoleh

$$\sigma^2\{\hat{\beta}_0\} = \sigma^2 \left(\frac{1}{n} + \frac{\overline{X}^2}{S_{XX}}\right) \qquad \text{.... (2.16)}$$

dengan kesalahan baku $\hat{\beta}_0$ adalah

$$se\{\hat{\beta}_0\} = \sigma\{\hat{\beta}_0\} = \sigma \sqrt{\frac{1}{n} + \frac{\overline{X}^2}{S_{XX}}} \qquad \text{.... (2.17)}$$

b. Means Dan Varians Penduga $\hat{\beta}_1$

Dari persamaan (2.14), diketahui bahwa:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i (X_i - \overline{X})}{\sum_{i=1}^{n} (X_i - \overline{X})^2}$$

Misalkan $b_i = \frac{(X_i - \overline{X})}{S_{XX}}$ dan $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, maka

$$\hat{\beta}_1 = \sum_{i=1}^{n} b_i (\beta_0 + \beta_1 X_i + \varepsilon_i)$$

Dengan penguraian lebih lanjut, didapatkan seperti berikut:

$$\hat{\beta}_1 = \beta_0 \sum_{i=1}^{n} b_i + \beta_1 \sum_{i=1}^{n} b_i X_i + \sum_{i=1}^{n} b_i \varepsilon_i$$

Karena $\sum_{i=1}^{n} b_i = 0$ dan $\sum_{i=1}^{n} b_i X_i = 1$, maka

$$\hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} b_i \varepsilon_i$$

dan

$$E(\hat{\beta}_1 - \beta_1) = E\left(\sum_{i=1}^{n} b_i \varepsilon_i\right)$$

atau

$$E\left(\hat{\beta}_1 - \beta_1\right) = \sum_{i=1}^{n} b_i E(\varepsilon_i) \qquad \text{.... (2.18)}$$

karena $\sum_{i=1}^{n} b_i$ adalah suatu konstanta.

Dengan demikian $E\left(\hat{\beta}_1 - \beta_1\right) = 0$, karena dari definisi $E(\varepsilon_i)$ = 0. Oleh karena itu $\hat{\beta}_1$ adalah penduga takbias β_1. Jika persamaan (2.18) dikuadratkan, maka diperoleh varians penduga $\hat{\beta}_1$ adalah

$$E\left(\hat{\beta}_1 - \beta_1\right)^2 = E\left(\sum_{i=1}^{n} b_i \varepsilon_i\right)^2$$

$$= \sum_{i=1}^{n} b_i^2 E\left(\varepsilon_i^2\right)$$

Karena $\sum_{i=1}^{n} b_i\left(X_i - \overline{X}\right) = 1$ dan $E\left(\varepsilon_i^2\right) = \sigma^2$, maka diperoleh

$$var\left\{\hat{\beta}_1\right\} = \sigma^2\left\{\hat{\beta}_1\right\} = \sigma^2 \frac{1}{S_{XX}} \qquad \text{.... (2.19)}$$

dengan kesalahan baku $\hat{\beta}_1$ adalah

$$se\left\{\hat{\beta}_1\right\} = \sigma\left\{\hat{\beta}_1\right\} = \sigma\sqrt{\frac{1}{S_{XX}}} \qquad \text{.... (2.20)}$$

c. Kovarians antara $\hat{\beta}_0$ dan $\hat{\beta}_1$

Dengan menggunakan persamaan (2.15) dan (2.18), kovarians antara $\hat{\beta}_0$ dan $\hat{\beta}_1$ dapat didefinisikan sebagai berikut:

$$Cov\left(\hat{\beta}_0, \hat{\beta}_1\right) = E\left[\left(\hat{\beta}_0 - \beta_0\right)\left(\hat{\beta}_1 - \beta_1\right)\right]$$

$$= E\left[\left(\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i - \overline{X}\sum_{i=1}^{n} b_i \varepsilon_i\right)\left(\sum_{i=1}^{n} b_i \varepsilon_i\right)\right]$$

$$= E\left[\left(\frac{1}{n}\sum_{i=1}^{n} \varepsilon_i\left(\sum_{i=1}^{n} b_i \varepsilon_i\right)\right) - \overline{X}\left(\sum_{i=1}^{n} b_i \varepsilon_i\right)^2\right]$$

$$= \sum_{i=1}^{n} b_i \left(\sum_{i=1}^{n} \varepsilon_i \right)^2 \dot{-} \left(\overline{X} \sum_{i=1}^{n} b_i \varepsilon_i \right)^2$$

Karena $E(\varepsilon_i \varepsilon_j) = 0$, jika $i \neq j$ dan $E(\varepsilon_i^2) = \sigma^2$. Maka,

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = -\sigma^2 \frac{\overline{X}}{S_{XX}}, \text{ karena } \sum_{i=1}^{n} b_i (X_i - \overline{X}) = 1 \qquad \text{.... (2.21)}$$

2.3 Bentuk Matriks Model Regresi Linier Sederhana

Model regresi (2.1) sebelumnya, model dinyatakan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Misalkan observasi n yang tersedia, dan misalkan X_{ij} adalah observasi ke-i atau tingkat variabel X_j. Data tersebut akan terlihat seperti dalam Tabel 2.1 di bawah ini.

Tabel 2.1. Data Model Regresi dalam Bentuk Matriks

Observasi	i	Y	X_0	X_1
	1	Y_1	1	X_{11}
	2	Y_2	1	X_{12}
	.	.	.	.
	.	.	.	.
	.	.	.	.
	n	Y_n	1	X_{1n}

Dari Tabel 2.1, misalkan hubungan antara variabel respon (Y_i) dengan variabel (X_i) yang diasumsikan nilainya bersifat tetap, ditentukan oleh:

$$Y_1 = \beta_0 + \beta_1 X_1 + \varepsilon_1$$

$$Y_2 = \beta_0 + \beta_1 X_2 + \varepsilon_2 \qquad \text{.... (2.22)}$$

$$\vdots$$

$$Y_n = \beta_0 + \beta_1 X_n + \varepsilon_n$$

dengan variabel X_i adalah variabel independen yang nilainya bersifat tetap tidak random, dan variabel ε_i adalah variabel kesalahan dengan asumsi berdistribusi normal dengan mean sama dengan nol dan varians konstan σ^2, dan ε_i independen dengan ε_j, untuk $i \neq j$ (Tirta, 2000).

Hubungan variabel di atas dapat juga dituliskan dalam bentuk matriks dengan mendefinisikan matriks-matriks berikut:

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix};\ \mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ . & \\ . & \\ . & \\ 1 & X_n \end{bmatrix};\ \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix};\ \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix} \quad \text{.... (2.23)}$$

Sistem persamaan (2.23), selanjutnya dapat dinyatakan dalam bentuk matriks sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{.... (2.24)}$$

dengan:

$\mathbf{y}$ adalah vektor respon;

$\boldsymbol{\beta}$ adalah vektor parameter;

$\mathbf{X}$ adalah matriks variabel independen;

$\boldsymbol{\varepsilon}$ adalah vektor kesalahan dengan asumsi $\mathbf{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ dan $\mathbf{var}(\boldsymbol{\varepsilon}) = \boldsymbol{\sigma}^2\mathbf{I}$.

2.3.1 Pendugaan Parameter Model Regresi Linier Sederhana dalam Bentuk Matriks dengan Metode Kuadrat Terkecil

Misalkan $Y_1, \ldots, Y_n$ adalah variabel random dengan nilai ekpektasi

$$E(Y_i) = \mu_i(\boldsymbol{\beta}), \text{ untuk } i = 1, 2, \ldots, n.$$

dengan $\boldsymbol{\beta} = [\beta_0, \beta_1]^T$ adalah parameter yang diduga nilainya. Metode kuadrat terkecil digunakan untuk mencari penduga parameter $\hat{\boldsymbol{\beta}}$ dengan meminimumkan jumlah kuadrat kesalahannya, yaitu:

$$S = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [Y_i - \mu_i(\boldsymbol{\beta})]^2 \quad \text{.... (2.25)}$$

Dalam bentuk matriks persamaan (2.25) ditulis sebagai berikut:

$$\mathbf{S} = (\mathbf{y} - \boldsymbol{\mu})^T (\mathbf{y} - \boldsymbol{\mu}) \qquad \text{.... (2.26)}$$

dengan $\mathbf{y} = [Y_1, ..., Y_n]^T$ dan $\mu = [\mu_1, ..., \mu_n]^T$. Secara umum, penduga $\hat{\boldsymbol{\beta}}$ dicari dengan menurunkan $\mathbf{S}$ terhadap setiap elemen β_j dari $\boldsymbol{\beta}$ dan menyelesaikan secara simultan persamaan (2.26) dengan membuat $\frac{\partial \mathbf{S}}{\partial \beta_j} = 0$, untuk j = 0, 1. Dari persamaan (2.26), jika μ_i merupakan kombinasi linier dari parameter β_j, yakni $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ untuk matriks $\mathbf{X}$ berukuran $n \times 2$, maka persamaan (2.26) dapat ditulis menjadi:

$$\mathbf{S} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \qquad \text{.... (2.27)}$$

Dari persamaan (2.27), turunan $\mathbf{S}$ terhadap setiap elemen β_j dari $\boldsymbol{\beta}$ adalah

$$\frac{\partial \mathbf{S}}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^T\mathbf{y} + 2\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} \qquad \text{.... (2.28)}$$

Penduga kuadrat terkecil untuk $\boldsymbol{\beta}$ dicari dengan membuat $\frac{\partial \mathbf{S}}{\partial \boldsymbol{\beta}} = 0$, maka persamaan (2.28) menjadi:

$$-\mathbf{X}^T\mathbf{y} + \mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \mathbf{0} \qquad \text{.... (2.29)}$$

Persamaan (2.29) disebut persamaan normal kuadrat terkecil. Penduga kuadrat terkecil $\boldsymbol{\beta}$ adalah penyelesaian dari persamaan normal adalah

$$\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^T\mathbf{y} \qquad \text{.... (2.30)}$$

Dengan mengalikan kedua ruas pada persamaan (2.30) dengan invers $\mathbf{X}^T\mathbf{X}$, maka diperoleh penyelesaian penduga kuadrat terkecil untuk $\boldsymbol{\beta}$, yaitu:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \qquad \text{.... (2.31)}$$

2.4 Pengaruh Kesalahan Pengukuran pada Variabel X

Selama ini telah diasumsikan secara implisit bahwa variabel X pada model regresi linier sederhana diukur dengan mengabaikan kesalahan pengukuran, sehingga variabel X diasumsikan nilainya bersifat tetap (*fixed*). Untuk model regresi (2.1), diasumsikan bahwa data mengenai variabel X adalah akurat (Weisberg, 1985). Sayangnya, hal yang ideal ini sering tidak dijumpai dalam praktek (lapangan) untuk berbagai alasan, misalkan seperti contoh pemberian dosis pemupukan Nitrogen [halaman 1]. Karena variabel X mengandung kesalahan pengukuran, maka asumsi untuk model regresi linier harus dilengkapi. Neter *et al.*, (1990) kesalahan pengukuran δ_i didefinisikan sebagai berikut:

$$\delta_i = X_i^* - X_i \qquad ....(2.32)$$

dengan:

X_i^* adalah variabel yang diamati diasumsikan nilainya bersifat random;

X_i adalah variabel sebenarnya diasumsikan nilainya bersifat tetap;

δ_i adalah variabel kesalahan pengukuran dengan asumsi $E(\delta_i) = 0$ dan $\text{var}(\delta_i) = \sigma_\delta^2$.

Karena variabel yang diamati adalah X_i^*, maka model regresi (2.1) sebelumnya menjadi:

$$Y_i = \beta_0 + \beta_1(X_i^* - \delta_i) + \varepsilon_i \qquad .... (2.33)$$

Dengan menggantikan X dari persamaan (2.32) dan bentuk persamaan (2.33) dapat dituliskan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i^* + (\varepsilon_i - \beta_1\delta_i) \qquad .... (2.34)$$

dengan X_i^* adalah variabel yang diamati yang bersifat random dan $(\varepsilon_i - \beta_1\delta_i)$ adalah variabel random kesalahannya. Dari model (2.34) diasumsikan bahwa variabel kesalahan ε_i dan δ_i adalah tidak berkorelasi, sehingga $E(\varepsilon_i\delta_i) = 0$. Model (2.34) sekilas kelihatan sama dengan model regresi biasa (2.1), dengan variabel kesalahan $\gamma_i = \varepsilon_i - \beta_1\delta_i$. Menurut Neter *et al.*, (1990), variabel yang diamati X_i^* adalah variabel random dan berkorelasi dengan variabel kesalahan γ_i. Korelasi antara X_i^* dan γ_i dapat ditunjukkan sebagai berikut:

$$\begin{aligned} Cov(X_i^*, \gamma_i) &= E[(X_i^* - E(X_i^*))(\gamma_i - E(\gamma_i))] \\ &= E[(X_i^* - X_i)\gamma_i] \end{aligned}$$

$$= E\left[\left(X_i^* - X_i\right)\left(\varepsilon_i - \beta_1 \delta_i\right)\right]$$
$$= E\left(\delta_i \varepsilon_i - \beta_1 \delta_i^2\right)$$
$$= -\beta_1 \sigma_\delta^2$$

Jika $\beta_1 \neq 0$, maka variabel yang diamati X_i^* dan variabel kesalahan γ_i akan berkorelasi (Neter, *et al*., 1990).

Menurut Neter *et al*., (1990), dikatakan bahwa jika dalam model regresi linier ada hubungan antara X dan Y, maka kovarians untuk model regresi linier sederhana adalah tidak nol, karena ada korelasi antara kesalahan pengukuran dengan variabel yang diamati tadi.

2.5 Model Berkson

Berkson (1950) menunjukkan langkah penting di dalam menyelesaikan masalah pendugaan parameter pada model regresi linier sederhana dengan variabel X mengandung kesalahan pengukuran (Weisberg, 1985). Dari persamaan (2.32), Berkson (1950) mengasumsikan bahwa variabel yang diamati X_i^* diperlakukan sebagai variabel yang nilainya bersifat tetap (*fixed*), sedangkan variabel sebenarnya X_i diperlakukan sebagai variabel random (Neter *et al*., 1990). Sebagai contoh, dalam pemberian dosis pemupukan Nitrogen [halaman 1], diasumsikan bahwa berat dosis pupuk yang diamati X_i^* diperlakukan sebagai variabel yang nilainya bersifat tetap, maka penetapan berat dosis pemupukan X_i diperlakukan sebagai variabel random, karena dimungkinkan terjadi kesalahan pengamatan pada saat penimbangan atau timbangan tidak akurat dalam menunjukkan berat yang diamati tadi. Kesalahan pengukuran sebagaimana didefinisikan dalam persamaan (2.32):

$$X_i = X_i^* - \delta_i \qquad \ldots. (2.35)$$

dengan:

X_i^* adalah variabel yang diamati diasumsikan nilainya bersifat tetap;

X_i adalah variabel sebenarnya yang diasumsikan nilainya bersifat random;

δ_i adalah kesalahan pengukuran dengan asumsi $E(\delta_i) = 0$ dan $\text{var}(\delta_i) = \sigma_\delta^2$. Persamaan (2.34) sebelumnya, masih tetap digunakan untuk model Berkson ini, yaitu:

$$Y_i = \beta_0 + \beta_1 X_i^* + \gamma_i \qquad \text{.... (2.36)}$$

dengan:

Y_i adalah variabel respon;

β_0 dan β_1 adalah parameter yang tidak diketahui nilainya;

X_i^* adalah variabel yang diamati dengan asumsi nilainya bersifat tetap;

γ_i adalah variabel kesalahan gabungan dengan asumsi $E(\gamma_i) = 0$ dan $\text{var}(\gamma_i) = (\sigma_\delta^*)^2$ dengan $\gamma_i = \varepsilon_i - \beta_1\delta_i$ dan $(\sigma_\delta^*)^2 = \sigma_1^2 + \beta_1^2\sigma_2^2$.

III. METODOLOGI

3.1 Pendugaan Parameter Model Berkson dengan Metode Kuadrat Terkecil

Dari model Berkson (2.40), model Berkson dinyatakan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i^* + \gamma_i$$

dengan $\gamma_i = \varepsilon_i - \beta_1 \delta_i$, untuk i = 1, 2, ..., n. Pada model Berkson diasumsikan bahwa antara variabel kesalahan γ_i dengan variabel observasi X_i^* adalah saling independen (bebas), artinya tidak ada korelasi diantara kedua variabel tersebut, karena dalam model Berkson, variabel yang diamati X_i^* diasumsikan sebagai variabel yang bersifat tetap (*fixed*). Dengan demikian, prosedur pendugaan adalah sama, sebagaimana prosedur pendugaan pada model regresi biasa dengan menggunakan metode kuadrat terkecil.

Dari model Berkson, misalkan hubungan sebenarnya antara variabel respon (Y) dan variabel yang diamati (X_i^*) adalah sebuah garis lurus, dan nilai observasi (pengamatan) Y adalah sebuah variabel random. Metode kuadrat terkecil digunakan untuk meminimumkan jumlah kuadrat selisih antara variabel respon (Y) dan nilai harapannya $E(Y)$. Dari model Berkson (2.40), diketahui bahwa $E(Y_i) = \beta_0 + \beta_1 X_i^*$, dengan mengasumsikan bahwa $E(\gamma_i) = 0$, sehingga untuk model Berkson, jumlah kuadrat kesalahannya adalah:

$$S = \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i^*\right)^2 \qquad \text{.... (3.1)}$$

Pendugaan kuadrat terkecil parameter β_0 dan β_1 adalah penyelesaian dari

$$\frac{\partial S}{\partial \beta_0} = -2\sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i^*\right) = 0 \qquad \text{.... (3.2)}$$

$$\frac{\partial S}{\partial \beta_1} = -2\sum_{i=1}^{n} X_i^* \left(Y_i - \beta_0 - \beta_1 X_i^*\right) = 0 \qquad \text{.... (3.3)}$$

Dari persamaan (3.2) diperoleh

$$\sum_{i=1}^{n} Y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} X_i^* = 0 \qquad \text{.... (3.4)}$$

dan dari persamaan (3.3) diperoleh

$$\sum_{i=1}^{n} X_i^* Y_i - \beta_0 \sum_{i=1}^{n} X_i^* - \beta_1 \sum_{i=1}^{n} \left(X_i^*\right)^2 = 0 \qquad \text{.... (3.5)}$$

Persamaan (3.4) dan (3.5) disebut persamaan normal kuadrat terkecil model Berkson. Penyelesaian secara simultan persamaan (3.4) dan (3.5) akan diperoleh nilai penduga parameter, yaitu $\hat{\beta}_0$ dan $\hat{\beta}_1$. Untuk penyelesaian nilai penduga $\hat{\beta}_0$ adalah:

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}^* \qquad \text{.... (3.6)}$$

untuk disubstitusikan ke dalam persamaan (3.5), menghasilkan

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i^* Y_i - \frac{\sum_{i=1}^{n} X_i^* \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} \left(X_i^*\right)^2 - \frac{\left(\sum_{i=1}^{n} \left(X_i^*\right)^2\right)^2}{n}} \qquad \text{.... (3.7)}$$

dengan $\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$ dan $\overline{X} = \frac{\sum_{i=1}^{n} X_i^*}{n}$ adalah mean Y_i dan X_i^*.

Dengan demikian, nilai penduga $\hat{\beta}_0$ dan $\hat{\beta}_1$ pada persamaan (3.6) dan (3.7) berhubungan dengan titik-titik turunan pertama persamaan (3.4) dan (3.5) adalah nol, yaitu pada saat jumlah kuadrat kesalahan adalah minimum. Persamaan garis regresi model Berkson adalah:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X^* \qquad \text{.... (3.8)}$$

dengan $\hat{\beta}_0$ dan $\hat{\beta}_1$ adalah penduga kuadrat terkecil model Berkson.

3.2 Bentuk Matriks Model Berkson

Model Berkson (2.40), model dapat dinyatakan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i^* + \gamma_i$$

Misalkan observasi n yang tersedia, dan misalkan X_{ij}^{*} adalah observasi ke-i atau tingkat variabel X_{j}^{*}. Data tersebut akan terlihat seperti dalam Tabel 3.1 di bawah ini.

Tabel 3.1. Data Model Berkson dalam Bentuk Matriks

Observasi	i	Y	X_{0i}^{*}	X_{1i}^{*}
	1	Y_1	1	X_{11}^{*}
	2	Y_2	1	X_{12}^{*}
	.	.	.	.
	.	.	.	.
	.	.	.	.
	n	Y_n	1	X_{1n}^{*}

Dari Tabel 3.1, misalkan hubungan antara variabel respon (Y_i) dengan variabel (X_{ij}^{*}) yang diasumsikan nilainya bersifat tetap, ditentukan oleh:

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11}^{*} + \gamma_1 \\
Y_1 &= \beta_0 + \beta_1 X_{12}^{*} + \gamma_2 \\
&\;. \\
&\;. \\
&\;. \\
Y_1 &= \beta_0 + \beta_1 X_{1n}^{*} + \gamma_n
\end{aligned}
\qquad \text{.... (3.9)}
$$

dengan variabel X_{j}^{*} adalah variabel yang diamati yang nilainya bersifat tetap tidak random, dan variabel γ_i adalah variabel kesalahan dengan asumsi dengan mean sama dengan nol dan varians konstan $(\sigma_{\gamma}^{*})^2$, dan γ_i independen dengan γ_j, untuk $i \neq j$.

Hubungan variabel di atas dapat juga dituliskan dalam bentuk matriks dengan mendefinisikan matriks-matriks berikut:

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix}; \mathbf{X} = \begin{bmatrix} 1 & X_{11}^* \\ 1 & X_{12}^* \\ & . \\ & . \\ & . \\ 1 & X_{1n}^* \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}; \boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ \lambda_2 \\ . \\ . \\ . \\ \gamma_n \end{bmatrix} \qquad \text{.... (3.10)}$$

Sistem persamaan (3.10), selanjutnya dapat dinyatakan dalam bentuk matriks sebagai berikut:

$$\mathbf{y} = \mathbf{X}^* \boldsymbol{\beta} + \boldsymbol{\gamma} \qquad \text{.... (2.11)}$$

dengan:

$\mathbf{y}$ adalah vektor respon;

$\boldsymbol{\beta}$ adalah vektor parameter;

$\mathbf{X}^*$ adalah matriks variabel yang diamati;

$\boldsymbol{\gamma}$ adalah vektor kesalahan dengan asumsi $\mathbf{E}(\boldsymbol{\gamma}) = \mathbf{0}$ dan $\mathbf{var}(\gamma) = (\boldsymbol{\sigma}_{\gamma}^*)^2 \mathbf{I}$.

3.2.1 Pendugaan Parameter Model Berkson dalam Bentuk Matriks dengan Metode Kuadrat Terkecil

Misalkan $Y_1, \ldots, Y_n$ adalah variabel random dengan nilai ekpektasi sebagai berikut:

$$E(Y_i) = \mu_i(\boldsymbol{\beta}), \text{ untuk } i = 1, 2, \ldots, n.$$

dengan $\boldsymbol{\beta} = [\beta_0, \beta_1]^T$ adalah parameter yang diduga nilainya. Metode kuadrat terkecil digunakan untuk mencari penduga parameter $\hat{\boldsymbol{\beta}}$ dengan meminimumkan jumlah kuadrat kesalahannya, yaitu:

$$S = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [Y_i - \mu_i(\boldsymbol{\beta})]^2 \qquad \text{.... (3.12)}$$

Dalam bentuk matriks persamaan (3.12) ditulis sebagai berikut:

$$\mathbf{S} = (\mathbf{y} - \boldsymbol{\mu})^T (\mathbf{y} - \boldsymbol{\mu}) \qquad \text{.... (3.13)}$$

dengan $\mathbf{y} = [Y_1,\ldots,Y_n]^T$ dan $\mu = [\mu_1,\ldots,\mu_n]^T$. Secara umum, penduga $\hat{\boldsymbol{\beta}}$ dicari dengan menurunkan **S** terhadap setiap elemen β_j dari $\boldsymbol{\beta}$ dan menyelesaikan secara simultan persamaan (2.26) dengan membuat $\frac{\partial \mathbf{S}}{\partial \beta_j} = 0$, untuk j = 0, 1. Dari persamaan (3.12), jika μ_i merupakan kombinasi linier dari parameter β_j, yakni $\boldsymbol{\mu} = \mathbf{X}^*\boldsymbol{\beta}$ untuk matriks $\mathbf{X}^*$ berukuran $n \times 2$, maka persamaan (3.12) dapat ditulis menjadi:

$$\mathbf{S} = (\mathbf{y} - \mathbf{X}^*\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}^*\boldsymbol{\beta}) \qquad \ldots (3.13)$$

Dari persamaan (3.13), turunan **S** terhadap setiap elemen β_j dari $\boldsymbol{\beta}$ adalah

$$\frac{\partial \mathbf{S}}{\partial \boldsymbol{\beta}} = -2(\mathbf{X}^*)^T\mathbf{y} + 2(\mathbf{X}^*)^T\mathbf{X}^*\boldsymbol{\beta} \qquad \ldots (3.14)$$

Penduga kuadrat terkecil untuk $\boldsymbol{\beta}$ dicari dengan membuat $\frac{\partial \mathbf{S}}{\partial \boldsymbol{\beta}} = 0$, sehingga persamaan (3.14) menjadi:

$$-(\mathbf{X}^*)^T\mathbf{y} + (\mathbf{X}^*)^T\mathbf{X}^*\boldsymbol{\beta} = \mathbf{0} \qquad \ldots (3.15)$$

Persamaan (3.15) disebut persamaan normal kuadrat terkecil model Berkson. Penduga kuadrat terkecil $\boldsymbol{\beta}$ adalah penyelesaian dari persamaan normal adalah:

$$(\mathbf{X}^*)^T\mathbf{X}^*\boldsymbol{\beta} = (\mathbf{X}^*)^T\mathbf{y} \qquad \ldots (3.16)$$

Dengan mengalikan kedua ruas pada persamaan (3.16) dengan invers $\mathbf{X}^T\mathbf{X}$, maka diperoleh penyelesaian penduga kuadrat terkecil untuk $\boldsymbol{\beta}$, yaitu:

$$\hat{\boldsymbol{\beta}} = ((\mathbf{X}^*)^T\mathbf{X}^*)^{-1}(\mathbf{X}^*)^T\mathbf{y} \qquad \ldots (3.17)$$

3.3 Prosedur Simulasi

Tujuan dari simulasi ini adalah untuk mengetahui perbandingan mean dan standar deviasi penduga parameter antara model regresi biasa dengan model Berkson. Dalam program ini, prosedur simulasi yang digunakan ialah sebagai berikut ini.

1. Matriks parameter awal yang digunakan adalah (9,2), dan data sampel yang dibangkitkan adalah 100 data (n = 100).
2. Standar deviasi untuk membangkitkan data kesalahan pengukuran (δ) adalah 2,5 , sedangkan standar deviasi untuk membangkitkan data $\mathbf{Y}$ dan $\mathbf{Y}_1$ adalah 10.
3. Langkah ini adalah melakukan pendugaan parameter $\boldsymbol{\beta}$ dan menghitung standar kesalahan (s.e).
4. Langkah selanjutnya adalah melakukan 100 kali simulasi atau ulangan, kemudian dicari mean dan standar deviasi penduga parameter dari model regresi biasa maupun model Berkson.
5. Langkah terakhir adalah melakukan uji beda dua rata-rata penduga parameter antara model regresi biasa dengan model Berkson, dengan uji statistiknya menggunakan distribusi Z yang digunakan untuk melihat apakah ada perbedaan mean penduga parameter antara kedua model tersebut.

3.4 Pengujian Hipotesis Beda Dua Rata-Rata

Berdasarkan 100 kali simulasi, diperoleh mean penduga parameter, baik mean penduga parameter model regresi biasa maupun mean penduga parameter model Berkson. Selanjutnya dilakukan pengujian hiportesa beda dua rata-rata penduga parameter antara kedua model tersebut.

3.4.1 Prosedur Pengujian Hipotesis Beda Dua Rata-Rata Penduga $\hat{\beta}_0$

Untuk pengujian hipotesis beda dua rata-rata penduga $\hat{\beta}_0$ antara mean penduga $\hat{\beta}_0$ model regresi biasa dengan mean penduga $\hat{\beta}_0$ model Berkson, uji statistiknya menggunakan distribusi Z. Prosedur pengujian hipotesisnya ialah sebagai berikut ini.

1. Formulasi hipotesis:

   $H_0 : \mu_1 = \mu_2$

   $H_1 : \mu_1 \neq \mu_2$

   dengan:

μ_1 adalah mean penduga $\hat{\beta}_0$ model regresi biasa;

μ_2 adalah mean penduga $\hat{\beta}_0$ model Berkson.

2. Taraf nyata dan nilai Z tabelnya:

   $\alpha = 5\% = 0{,}05$; $\alpha / 2 = 0{,}025$.

   $Z_{\alpha/2} = 1{,}96$.

3. Kriteria pengujian:

   untuk $H_0 : \mu_1 = \mu_2$ dan $H_1 : \mu_1 \neq \mu_2$

   (a) H_0 diterima jika $-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$

   (b) H_0 ditolak jika $Z_0 > Z_{\alpha/2}$ atau $Z_0 < -Z_{\alpha/2}$

4. Uji statistik menggunakan distribusi Z:

$$Z_0 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

   dengan:

   $\overline{X}_1$ adalah nilai mean penduga $\hat{\beta}_0$ model regresi biasa;

   $\overline{X}_2$ adalah nilai mean penduga $\hat{\beta}_0$ model Berkson;

   s_1^2 adalah nilai standar deviasi penduga $\hat{\beta}_0$ model regresi biasa;

   s_2^2 adalah nilai standar deviasi penduga $\hat{\beta}_0$ model Berkson;

   n_1 dan n_1 adalah jumlah data penduga $\hat{\beta}_0$ baik model regresi biasa maupun model Berkson.

5. Kesimpulan:

   kesimpulan pengujian merupakan penerimaan atau penolakan H_0:

   (a) jika H_0 diterima, maka H_1 ditolak;

   (b) jika H_0 ditolak, maka H_1 diterima.

3.4.2 Prosedur Pengujian Hipotesis Beda Dua Rata-Rata Penduga $\hat{\beta}_1$

Untuk pengujian hipotesis beda dua rata-rata penduga $\hat{\beta}_1$, prosedur pengujiannya sama dengan prosedur pengujian hipotesis penduga $\hat{\beta}_0$. Prosedur pengujiannnya adalah sebagai berikut ini.

1. Formulasi hipotesis:

   $H_0 : \mu_1 = \mu_2$
   $H_1 : \mu_1 \neq \mu_2$

   dengan:

   μ_1 adalah mean penduga $\hat{\beta}_1$ model regresi biasa;

   μ_2 adalah mean penduga $\hat{\beta}_1$ model Berkson.

2. Taraf nyata dan nilai Z tabelnya:

   $\alpha = 5\% = 0{,}05$; $\alpha / 2 = 0{,}025$.

   $Z_{\alpha/2} = 1{,}96$.

3. Kriteria pengujian:

   untuk $H_0 : \mu_1 = \mu_2$ dan $H_1 : \mu_1 \neq \mu_2$

   (a) H_0 diterima jika $-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$

   (b) H_0 ditolak jika $Z_0 > Z_{\alpha/2}$ atau $Z_0 < -Z_{\alpha/2}$

4. Uji statistik menggunakan distribusi Z:

   $$Z_0 = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

   dengan:

   $\overline{X}_1$ adalah nilai mean penduga $\hat{\beta}_1$ model regresi biasa;

   $\overline{X}_2$ adalah nilai mean penduga $\hat{\beta}_1$ model Berkson;

   s_1^2 adalah nilai standar deviasi penduga $\hat{\beta}_1$ model regresi biasa;

   s_2^2 adalah nilai standar deviasi penduga $\hat{\beta}_1$ model Berkson;

n_1 dan n_1 adalah jumlah data penduga $\hat{\beta}_1$ baik model regresi biasa maupun model Berkson.

5. Kesimpulan:

kesimpulan pengujian merupakan penerimaan atau penolakan H_0:

a. Jika H_0 diterima, maka H_1 ditolak;

b. Jika H_0 ditolak, maka H_1 diterima.

V. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil pembahasan pada bab IV, memberikan kesimpulan sebagai berikut ini.

1. Berdasarkan salah satu simulasi diperoleh perbandingan nilai penduga $\hat{\beta}_0$ antara model regresi biasa dengan model Berkson adalah 5, 8664 dengan 8, 9135. Perbandingan nilai penduga $\hat{\beta}_1$ antara model regresi biasa dengan model Berkson adalah 2,2826 dengan 1,9932.
2. Berdasarkan 100 kali simulasi diperoleh perbandingan mean penduga $\hat{\beta}_0$ antara model regresi biasa dengan model Berkson adalah 8,9197 dengan 9,0532. Perbandingan mean penduga $\hat{\beta}_1$ antara model regresi biasa dengan model Berkson adalah 2,0121 dengan 0,0048.
3. Secara matematis, standar kesalahan (s.e) antara model regresi biasa dengan model Berkson adalah sama, masing-masing untuk $\hat{\beta}_0$ dan $\hat{\beta}_1$ adalah 2,1602 dengan 0,1741.
4. Berdasarkan 100 kali simulasi, secara matematis standar deviasi penduga $\hat{\beta}_0$ model Berkson lebih besar daripada standar deviasi penduga $\hat{\beta}_0$ model regresi biasa. Standar deviasi penduga $\hat{\beta}_1$ model Berkson juga lebih besar daripada standar deviasi penduga $\hat{\beta}_1$ model regresi biasa.
5. Dari hasil pengujian hipotesis dua rata-rata penduga $\hat{\beta}_0$ dengan taraf kepercayaan 95% diperoleh nilai uji statistik Z_0 sebesar –0, 4126. Karena nilai tersebut berada pada selang $-1{,}96 \le Z_0 \le 1{,}96$, maka hipotesis H_0 diterima. Jadi, secara statistik tidak ada perbedaan mean penduga $\hat{\beta}_0$ antara model regresi biasa dengan model Berkson. Dari hasil pengujian hipotesis dua rata-rata penduga $\hat{\beta}_1$ dengan taraf kepercayaan 95% diperoleh nilai uji statistik Z_0 sebesar 0,2779. Karena nilai tesebut berada pada selang $-1{,}96 \le Z_0 \le 1{,}96$,

maka hipotesis H_0 diterima. Jadi, secara statistik tidak ada perbedaan mean penduga $\hat{\beta}_1$ antara model regresi biasa dengan model Berkson.

5.2 Saran

Karena dalam skripsi ini hanya menggunakan standar deviasi untuk kesalahan pengukuran (*delta*) sebesar $\delta = 2{,}5$, maka dapat dicoba dengan standar deviasi untuk kesalahan pengukuran yang lebih besar, untuk melihat apakah secara statistik ada perbedaan mean penduga parameter antara model regresi biasa dengan model Berkson, dan untuk melihat apakah secara matematis standar deviasi model Berkson lebih besar daripada standar deviasi model regresi biasa.

DAFTAR PUSTAKA

Dobson, A.J. 1983. *Introduction to Statistical Modeling*. New York: Chapman & Hall

Hines, W.W & D.C, Montgomery. 1982. *Probability and Statistics in Engineering and Management Science*. New York: John Wiley & Sons.

Hodges, S.D & D.G, Moore. 1982. Data Uncertainties and Least Squares Regression. *Appl. Statist*, Vol. 21: 185-105.

Jorgensen, B. 1993. *The Theory of Linear Models*. New York: Chapman & Hall.

Montgomery, D.C & E.A. Peck. 1991. *Introduction to Linear Regression Analysis*. 2nd Ed. New York: John Wiley & Sons.

Neter, J., W. Wasserman & M.H, Kutner. 1990. *Applied Linear Statistical Models*. 3rd Ed. Illinois: Richard D. Irwin.

Ratih, R.T. 2000. Inferensia Pada Model-Model Linier Tergeneralisasi. Majalah *Matematika Dan Statistika FMIPA*. Universitas Jember. Vol. 1:30-40.

Tirta, I.M., dkk. 2000. *Model-Model Linier*. Jurusan Matematika FMIPA. Universitas Jember. Jember.

Weisberg, S. 1985. *Applied Linear Regression*. 2nd Ed. New York: John Wiley & Sons.

Lampiran 1. Script Program Perbandingan Mean dan Standar Deviasi Penduga Parameter antara Model Regresi Biasa dengan Model Berkson

```
#script program perbandingan Mean dan Standar Deviasi penduga
   parameter antara Model Regresi Biasa dengan Model Berkson
#mendefinisikan matriks beta, ukuran sampel, dan jumlah simulasi
options(echo=F,digit=4)
beta_matrix(c(9,2),2,1)
n_100
n.sim_100

#membuat matriks perhitungan
i_0
mat.mengabaikan.kesalahan_matrix(0,n.sim,4)
mat.mengandung.kesalahan_matrix(0,n.sim,4)
while(i<n.sim){
   i_i+1
   #membangkitkan data x (mengabaikan kesalahan pengukuran)
   x_matrix(1,n,2)
   x[,2]_seq(2,20,2)

   #membangkitkan data y (model regresi biasa)
   mu_x%*%beta
   sd.y_10
   y_rnorm(n,mu,sd.y)

   #membangkitkan data delta (kesalahan pengukuran)
   sd.delt_2.5
   delta_matrix(0,n,2)
   delta[,2]_rnorm(n,0,sd.delt)

   #data x.star (mengandung kesalahan pengukuran)
   x.star_x+delta

   #membangkitkan data y1 (model Berkson)
   mu.star_x.star%*%beta
   y1_rnorm(n,mu.star,sd.y)


   #estimasi parameter model regresi biasa
   b_solve(t(x)%*%x)%*%t(x)%*%y
   varb_sd.y^2*solve(t(x)%*%x)
   se.b0_sqrt(varb[1,1])
   se.b1_sqrt(varb[2,2])

   #estimasi parameter model Berkson
   b.star_solve(t(x)%*%x)%*%t(x)%*%y1
   varb.star_sd.y^2*solve(t(x)%*%x)
   se.b0.star_sqrt(varb.star[1,1])
   se.b1.star_sqrt(varb.star[2,2])

   #Matrik Output model regresi biasa
   mat.mengabaikan.kesalahan[i,1]_b[1,]
```



```
   mat.mengabaikan.kesalahan[i,2]_b[2,]
   mat.mengabaikan.kesalahan[i,3]_beta[1,]
   mat.mengabaikan.kesalahan[i,4]_beta[2,]

   #Matrik Output model Berkson
   mat.mengandung.kesalahan[i,1]_b.star[1,]
   mat.mengandung.kesalahan[i,2]_b.star[2,]
   mat.mengandung.kesalahan[i,3]_beta[1,]
   mat.mengandung.kesalahan[i,4]_beta[2,]
   }

#Mean dan standar deviasi penduga parameter model regresi biasa
   dan model Berkson
#Mean dan Standar Deviasi Penduga b0 model regresi biasa
mean.b0_mean(mat.mengabaikan.kesalahan[,1])
var.b0_var(mat.mengabaikan.kesalahan[,1])
sd.b0_sqrt(var(mat.mengabaikan.kesalahan[,1]))

#Mean dan Standar Deviasi Penduga b0 model Berkson
mean.b0.star_mean(mat.mengandung.kesalahan[,1])
var.b0.star_var(mat.mengandung.kesalahan[,1])
sd.b0.star_sqrt(var(mat.mengandung.kesalahan[,1]))

#Mean dan Standar Deviasi Penduga b1 model regresi biasa
mean.b1_mean(mat.mengabaikan.kesalahan[,2])
var.b1_var(mat.mengabaikan.kesalahan[,2])
sd.b1_sqrt(var(mat.mengabaikan.kesalahan[,2]))

#Mean dan Standar Deviasi Penduga b1 model Berkson
mean.b1.star_mean(mat.mengandung.kesalahan[,2])
var.b1.star_var(mat.mengandung.kesalahan[,2])
sd.b1.star_sqrt(var(mat.mengandung.kesalahan[,2]))

#Perbandingan Mean dan Standar Deviasi antara Model regresi Biasa
   dengan Model Berkson
for(j in 1:1){
   cat("\nMean, Standar Deviasi Penduga Parameter Model Regresi
   Biasa\n")
   cat("Komponen\t\t Mean Estimator\t\t Standar Deviasi\n")
   cat("Konstanta\t","\t",mean.b0,"\t",sd.b0,"\n")
   cat("Regresi\t","\t",mean.b1,"\t",sd.b1,"\n")

   cat("\nMean, Standar Deviasi Penduga Parameter Model
   Berkson\n")
   cat("Komponen\t\t Mean Estimator\t\t Standar Deviasi\n")
   cat("Konstanta\t","\t",mean.b0.star,"\t",sd.b0.star,"\n")
   cat("Regresi\t","\t",mean.b1.star,"\t",sd.b1.star,"\n")
   }

#Uji Beda Dua Rata-Rata Penduga Parameter antara Model Regresi
   Biasa dengan Model Berkson Berdasarkan 100 kali simulasi
#Uji beda dua rata-rata penduga b0 antara Model Regrtesi Biasa
   dengan Model Berkson
z0_(mean.b0-mean.b0.star)/sqrt((sd.b0.star^2+sd.b0^2)/n.sim)
print(z0)
if(-1.96<=z0<=1.96){
```



```
   cat("\nTerima H0 jika -1.96<=z1<=1.96\n")}
   else{cat("\nTolak H0 jika z1>1.96 atau z1<-1.96\n")}

#Uji beda dua rata-rata penduga b1 antara Model Regresi Biasa
   dengan Model Berkson
z1_(mean.b1-mean.b1.star)/sqrt((sd.b1.star^2+sd.b1^2)/n.sim)
print(z1)
if(-1.96<=z1<=1.96){
   cat("\nTerima H0 jika -1.96<=z1<=1.96\n")}
   else{cat("\nTolak H0 jika z1>1.96 atau z1<-1.96\n")}

#Grafik Perbandingan Nilai Penduga b0 antara Model Regresi Biasa
   dengan Model Berkson (berdasarkan 100 kali simulasi)
a_seq(1,n.sim,1)
plot(a,mat.mengabaikan.kesalahan[,1],ylab='Nilai Penduga
   b0',xlab='Simulasi',type='b',ylim=c(2.5,18),pch='o',lty=1)
points(a,mat.mengandung.kesalahan[,1],pch='*')
lines(a,mat.mengandung.kesalahan[,1],lty=2)
lines(a,mat.mengabaikan.kesalahan[,3])
legend(5,17.8,c("mengabaikan kesalahan pengukuran","mengandung
   kesalahan pengukuran"),lty = 1:2, pch ="o*")

#Grafik Perbandingan Nilai Penduga b1 antara Model Regresi Biasa
   dengan Model Berkson (berdasarkan 100 kali simulasi)
plot(a,mat.mengabaikan.kesalahan[,2],ylab='Nilai Penduga
   b1',xlab='Simulasi',type='b',ylim=c(1.3,2.5),pch='o',lty=1)
points(a,mat.mengandung.kesalahan[,2],pch='*')
lines(a,mat.mengandung.kesalahan[,2],lty=2)
lines(a,mat.mengabaikan.kesalahan[,4])
legend(3,1.5,c("mengabaikan kesalahan pengukuran","mengandung
   kesalahan pengukuran"),lty = 1:2, pch = "o*")

#Grafik Perbandingan Diagram Pencar dengan Garis Regresi antara
   Model Regresi Biasa dengan Model Berkson (diambil dari salah
   satu simulasi)
y.hat1_b[1,]+b[2,]*x[,2]
y.hat2_b.star[1,]+b.star[2,]*x[,2]
plot(x[,2],y,ylab='Y',xlab='data X',type='p', ylim=c(-
   5,80),pch='o')
points(x[,2],y1,pch='*')
lines(x[,2],y.hat1,lty=2)
lines(x[,2],y.hat2,lty=1)
legend(2,75,c("mengabaikan kesalahan pengukuran","mengandung
   kesalahan pengukuran"),lty = 1:2, pch ="o*")

#Perbandingan Penduga Parameter dan Standar Kesalahan antara Model
   Regresi biasa dengan Model Berkson (diambil dari salah satu
   simulasi)
for(k in 1:1){
   cat("\nNilai Penduga Parameter dan Standar Kesalahan Model
   Regresi Biasa\n")
   cat("Komponen\t\t Estimator\t\t s.e\n")
   cat("Konstanta\t","\t",b[1],"\t",se.b0,"\n")
   cat("Regresi\t","\t",b[2],"\t",se.b1,"\n")
```



```
cat("\nNilai Penduga Parameter dan Standar Kesalahan Model
Berkson\n")
cat("Komponen\t\t Estimator\t\t s.e\n")
cat("Konstanta\t","\t",b.star[1],"\t",se.b0.star,"\n")
cat("Regresi\t","\t",b.star[2],"\t",se.b1.star,"\n")
}
```

Lampiran 2. Report Program Perbandingan Mean dan Standar Deviasi Penduga Parameter antara Model Regresi Biasa dengan Model Berkson

```
> #script program perbandingan Mean dan Standar Deviasi penduga
parameter antara Model Regresi Biasa dengan Model Berkson
#mendefinisikan matriks beta, ukuran sampel, dan jumlah simulasi
options(echo = F, digit = 4)
> beta <- matrix(c(9, 2), 2, 1)
> n <- 100
> n.sim <- 100

#membuat matriks perhitungan
> i <- 0
> mat.mengabaikan.kesalahan <- matrix(0, n.sim, 4)
> mat.mengandung.kesalahan <- matrix(0, n.sim, 4)
> while(i < n.sim) {
      i <- i + 1
      #membangkitkan data x (mengabaikan kesalahan pengukuran)
      x <- matrix(1, n, 2)
      x[, 2] <- seq(2, 20, 2)

      #membangkitkan data y (model regresi biasa)
      mu <- x %*% beta
      sd.y <- 10
      y <- rnorm(n, mu, sd.y)

      #membangkitkan data delta (kesalahan pengukuran)
      sd.delt <- 2.5
      delta <- matrix(0, n, 2)
      delta[, 2] <- rnorm(n, 0, sd.delt)

      #data x.star (mengandung kesalahan pengukuran)
      x.star <- x + delta

      #membangkitkan data y1 (model Berkson)
      mu.star <- x.star %*% beta
      y1 <- rnorm(n, mu.star, sd.y)

      #estimasi parameter model regresi biasa
      b <- solve(t(x) %*% x) %*% t(x) %*% y
      varb <- sd.y^2 * solve(t(x) %*% x)
      se.b0 <- sqrt(varb[1, 1])
      se.b1 <- sqrt(varb[2, 2])

      #estimasi parameter model Berkson
      b.star <- solve(t(x) %*% x) %*% t(x) %*% y1
      varb.star <- sd.y^2 * solve(t(x) %*% x)
      se.b0.star <- sqrt(varb.star[1, 1])
      se.b1.star <- sqrt(varb.star[2, 2])

      #Matrik Output model regresi biasa
      mat.mengabaikan.kesalahan[i, 1] <- b[1, ]
```



```
        mat.mengabaikan.kesalahan[i, 2] <- b[2,  ]
        mat.mengabaikan.kesalahan[i, 3] <- beta[1,  ]
        mat.mengabaikan.kesalahan[i, 4] <- beta[2,  ]

        #Matrik Output model Berkson
        mat.mengandung.kesalahan[i, 1] <- b.star[1,  ]
        mat.mengandung.kesalahan[i, 2] <- b.star[2,  ]
        mat.mengandung.kesalahan[i, 3] <- beta[1,  ]
        mat.mengandung.kesalahan[i, 4] <- beta[2,  ]
}
#Mean dan standar deviasi penduga parameter model regresi biasa
dan model Berkson
#Mean dan Standar Deviasi Penduga b0 model regresi biasa
> mean.b0 <- mean(mat.mengabaikan.kesalahan[, 1])
> var.b0 <- var(mat.mengabaikan.kesalahan[, 1])
> sd.b0 <- sqrt(var(mat.mengabaikan.kesalahan[, 1]))

#Mean dan Standar Deviasi Penduga b0 model Berkson
> mean.b0.star <- mean(mat.mengandung.kesalahan[, 1])
> var.b0.star <- var(mat.mengandung.kesalahan[, 1])
> sd.b0.star <- sqrt(var(mat.mengandung.kesalahan[, 1]))

#Mean dan Standar Deviasi Penduga b1 model regresi biasa
> mean.b1 <- mean(mat.mengabaikan.kesalahan[, 2])
> var.b1 <- var(mat.mengabaikan.kesalahan[, 2])
> sd.b1 <- sqrt(var(mat.mengabaikan.kesalahan[, 2]))

#Mean dan Standar Deviasi Penduga b1 model Berkson
> mean.b1.star <- mean(mat.mengandung.kesalahan[, 2])
> var.b1.star <- var(mat.mengandung.kesalahan[, 2])
> sd.b1.star <- sqrt(var(mat.mengandung.kesalahan[, 2]))

#Perbandingan Mean dan Standar Deviasi antara Model regresi Biasa
dengan Model Berkson
> for(j in 1:1) {
      cat("\nMean, Standar Deviasi Penduga Parameter Model Regresi
         Biasa\n")
      cat("Komponen\t\t Mean Estimator\t\t Standar Deviasi\n")
      cat("Konstanta\t", "\t", mean.b0, "\t", sd.b0, "\n")
      cat("Regresi\t", "\t", mean.b1, "\t", sd.b1, "\n")
      cat("\nMean, Standar Deviasi Penduga Parameter Model
         Berkson\n")
      cat("Komponen\t\t Mean Estimator\t\t Standar Deviasi\n")
      cat("Konstanta\t", "\t", mean.b0.star, "\t", sd.b0.star,
         "\n")
      cat("Regresi\t", "\t", mean.b1.star, "\t", sd.b1.star, "\n")
}
Mean, Standar Deviasi Penduga Parameter Model Regresi Biasa
Komponen              Mean Estimator          Standar Deviasi
Konstanta             8.91970967623404        1.99263080639096
Regresi               2.01213099650717        0.160658389284237

Mean, Standar Deviasi Penduga Parameter Model Berkson
Komponen              Mean Estimator          Standar Deviasi
Konstanta             9.0531939650194         2.54873419620868
Regresi               2.00476254287128        0.210879750750335
```



```
#Uji Beda Dua Rata-Rata Penduga Parameter antara Model Regresi
Biasa dengan Model Berkson Berdasarkan 100 kali simulasi
#Uji beda dua rata-rata penduga b0 antara Model Regrtesi Biasa
dengan Model Berkson
> z0 <- (mean.b0 - mean.b0.star)/sqrt((sd.b0.star^2 +
         sd.b0^2)/n.sim)
> print(z0)
[1] -0.4125977
> if(-1.96 <= z0 <= 1.96) {
      cat("\nTerima H0 jika -1.96<=z1<=1.96\n")}
      else {
      cat("\nTolak H0 jika z1>1.96 atau z1<-1.96\n")}

Terima H0 jika -1.96<=z1<=1.96

#Uji beda dua rata-rata penduga b1 antara Model Regrtesi Biasa
dengan Model Berkson
> z1 <- (mean.b1 - mean.b1.star)/sqrt((sd.b1.star^2 +
         sd.b1^2)/n.sim)
> print(z1)
[1] 0.2779433
> if(-1.96 <= z1 <= 1.96) {
      cat("\nTerima H0 jika -1.96<=z1<=1.96\n")}
      else {
      cat("\nTolak H0 jika z1>1.96 atau z1<-1.96\n")}

Terima H0 jika -1.96<=z1<=1.96

#Grafik Perbandingan Nilai Penduga b1 antara Model Regresi Biasa
dengan Model Berkson (berdasarkan 100 kali simulasi)
> a <- seq(1, n.sim, 1)
> plot(a, mat.mengabaikan.kesalahan[, 1], ylab =
      "Nilai Penduga b0", xlab = "Simulasi", type = "b", ylim =
      c(2.5, 18), pch = "o", lty = 1)
> points(a, mat.mengandung.kesalahan[, 1], pch = "*")
> lines(a, mat.mengandung.kesalahan[, 1], lty = 2)
> lines(a, mat.mengabaikan.kesalahan[, 3])
> legend(5, 17.8, c("mengabaikan kesalahan pengukuran",
      "mengandung kesalahan pengukuran"), lty = 1:2, pch = "o*")

#Grafik Perbandingan Nilai Penduga b1 antara Model Regresi Biasa
dengan Model Berkson (berdasarkan 100 kali simulasi)
> plot(a, mat.mengabaikan.kesalahan[, 2], ylab =
      "Nilai Penduga b1", xlab = "Simulasi", type = "b", ylim =
      c(1.3, 2.5), pch = "o", lty = 1)
> points(a, mat.mengandung.kesalahan[, 2], pch = "*")
> lines(a, mat.mengandung.kesalahan[, 2], lty = 2)
> lines(a, mat.mengabaikan.kesalahan[, 4])
> legend(3, 1.5, c("mengabaikan kesalahan pengukuran",
      "mengandung kesalahan pengukuran"), lty = 1:2, pch = "o*")
```



```
#Grafik Perbandingan Diagram Pencar dengan Garis Regresi antara
Model Regresi Biasa dengan Model Berkson (diambil dari salah satu
simulasi)
> y.hat1 <- b[1,  ] + b[2,  ] * x[, 2]
> y.hat2 <- b.star[1,  ] + b.star[2,  ] * x[, 2]
> plot(x[, 2], y, ylab = "Y", xlab = "data X", type = "p",
     ylim = c(-5, 80), pch = "o")
> points(x[, 2], y1, pch = "*")
> lines(x[, 2], y.hat1, lty = 2)
> lines(x[, 2], y.hat2, lty = 1)
> legend(2, 75, c("mengabaikan kesalahan pengukuran",
      "mengandung kesalahan pengukuran"), lty = 1:2, pch = "o*")

#Perbandingan Penduga Parameter dan Standar Kesalahan antara Model
Regresi biasa dengan Model Berkson (diambil dari salah satu
simulasi)
> for(k in 1:1) {
      cat("\nNilai Penduga Parameter dan Standar Kesalahan Model
          Regresi Biasa\n")
      cat("Komponen\t\t Estimator\t\t s.e\n")
      cat("Konstanta\t", "\t", b[1], "\t", se.b0, "\n")
      cat("Regresi\t", "\t", b[2], "\t", se.b1, "\n")
      cat("\nNilai Penduga Parameter dan Standar Kesalahan Model
          Berkson\n")
      cat("Komponen\t\t Estimator\t\t s.e\n")
      cat("Konstanta\t", "\t", b.star[1], "\t", se.b0.star, "\n")
      cat("Regresi\t", "\t", b.star[2], "\t", se.b1.star, "\n")
}

Nilai Penduga Parameter dan Standar Kesalahan Model Regresi Biasa
Komponen             Estimator           s.e
Konstanta            5.8664282018795     2.16024689946929
Regresi              2.28261922073387    0.174077655955698

Nilai Penduga Parameter dan Standar Kesalahan Model Berkson
Komponen             Estimator           s.e
Konstanta            8.91350885861077    2.16024689946929
Regresi              1.99316459056419    0.174077655955698
```