BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis regresi digunakan untuk menggambarkan hubungan fungsional antara peubah respon *Y* dengan peubah prediktor *X* dalam suatu bentuk model regresi. Pada analisis regresi dibutuhkan data yang bersifat kuantitatif sehingga dapat didefinisikan hubungan antar peubah respon dan peubah prediktor. Pada suatu penelitian tidak jarang data yang diperoleh bersifat kualitatif. Data kualitatif dapat berbentuk kelompok, kelas atau tingkatan. Sebagai contoh ingin dilakukan analisis terhadap besar gaji berdasarkan lama bekerja dan jenis kelamin, peubah jenis kelamin bersifat kualitatif yaitu kelompok laki-laki dan perempuan maka analisis regresi tidak dapat dilakukan sehingga diperlukan suatu cara untuk memodelkan hubungan antara peubah respon dan prediktor kuantitatif dan kualitatif.

Data yang bersifat kualitatif menunjukan keberadaan, sifat (ciri) tertentu sehingga akan menimbulkan kelompok dalam data. Kelompok yang dihasilkan dapat diubah dari data kualitatif menjadi kuantitatif dengan cara membentuk suatu peubah buatan dalam sandi 0 dan 1 yang menyatakan ada atau tidaknya suatu sifat pada data, disebut peubah boneka. Analisis regresi dapat dilakukan pada peubah prediktor kualitatif melalui pembentukan peubah boneka. Analisis terhadap hubungan fungsional antara peubah respon dan peubah prediktor kualitatif dan kuantitatif disebut analisis regresi linier berkelompok.

Terdapat empat model pada analisis regresi linier berkelompok yaitu model regresi garis sejajar dengan perbedaan penduga kategori dasar, model regresi garis terpisah dengan perbedaan penduga kategori dasar, model regresi garis sejajar dan model regresi garis terpisah. Perbedaan keempat jenis model regresi linier berkelompok terletak pada pembentukan peubah boneka pada setiap kelompok, pembentukan elemen matriks **X** pada pendugaan model regresi dengan metode kuadrat terkecil dan interaksi antar peubah prediktor.

Keempat model regresi digunakan untuk mengetahui apakah peubah prediktor kuantitatif maupun kualitatif berpengaruh terhadap peubah respon dan apakah interaksi antarpeubah prediktor berpengaruh nyata terhadap peubah respon. Pada model regresi linier berkelompok dilakukan uji perbedaan garis regresi yang dihasilkan untuk mengetahui apakah garis yang dihasilkan secara statistik dapat dikatakan sama atau berbeda. Setelah dilakukan uji perbedaan garis regresi akan diketahui model mana yang tepat menggambarkan hubungan antara peubah respon dan peubah prediktor kuantitatif dan kualitatif. Pada penelitian ini akan dibahas mengenai hubungan antara panjang polong (cm) dan lokasi tanam (Malang dan Jombang) terhadap bobot polong segar (g) bersumber dari Listyorini (2008).

1.2 Rumusan Masalah

Rumusan masalah yang ingin dikemukakan berdasarkan latar belakang yaitu:

1. Bagaimana membentuk model regresi dengan peubah prediktor yang bersifat kuantitatif dan kualitatif (berkelompok)?
2. Bagaimana pengaruh interaksi antara peubah prediktor yang bersifat kuantitatif dan kualitatif?
3. Bagaimana menentukan model yang sesuai untuk menggambarkan hubungan peubah respon berdasarkan peubah prediktor kualitatif dan kuantitatif?

1.3 Batasan Masalah

Permasalahan dibatasi pada pembentukan empat model regresi linier berkelompok menggunakan metode kuadrat terkecil.

1.4 Tujuan

Tujuan penelitian ini adalah :

1. Mengetahui cara pembentukan keempat model regresi linier berkelompok.
2. Mengetahui pengaruh interaksi antar peubah prediktor terhadap peubah respon.
3. Menentukan model yang sesuai untuk menggambarkan hubungan peubah respon berdasarkan peubah prediktor yang bersifat kualitatif dan kuantitatif.

1.5 Manfaat

Hasil dari penelitian ini diharapkan dapat bermanfaat untuk membentuk model regresi dengan peubah prediktor kualitatif dan kuantitatif, mengetahui pengaruh interaksi yang terjadi antar peubah prediktor dan menentukan model paling tepat berdasarkan hasil pengujian perbedaan dua garis regresi.

BAB II
TINJAUAN PUSTAKA

2.1 Model Regresi Linier Sederhana

Model regresi linier sederhana merupakan model regresi dengan sebuah peubah prediktor *X* dan sebuah respon *Y*. Hubungan fungsional peubah respon *Y* dan peubah prediktor *X* dijelaskan oleh model regresi:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad i = 1, 2, ..., n \tag{2.1}$$

di mana: β_0 = intersep (jarak antara titik pusat salib sumbu (0,0) dengan titik potong garis regresi terhadap sumbu y)

β_1 = koefisien regresi

ε_i = galat pengamatan ke-i

n = banyaknya pengamatan

Dasar pendugaan terhadap parameter β_0 dan β_1 adalah asumsi bahwa galat menyebar normal dengan rata-rata nol, ragam σ^2 dan antar galat saling bebas, $\varepsilon_i \sim NIID(0, \sigma^2)$:

$$E(\varepsilon_i) = 0$$

$$V(\varepsilon_i) = \sigma^2$$

$$\operatorname{cov}(\varepsilon_i, \varepsilon_{i'}) = 0$$

$$\begin{aligned} E(Y_i) &= E(\beta_0) + E(\beta_1 X_i) + E(\varepsilon_i) \\ &= E(\beta_0 + \beta_1 X_i) + E(\varepsilon_i) \\ &= \beta_0 + \beta_1 X_i + 0 \end{aligned}$$

$$\hat{Y}_i = \beta_0 + \beta_1 X_i$$

$$\varepsilon_i = Y_i - \hat{Y}_i$$

$$\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_i)$$

Pendugaan parameter β_0 dan β_1 dilakukan dengan metode kuadrat terkecil (MKT) yaitu dengan memimimumkan jumlah kuadrat galat. Turunan parsial JKG terhadap parameter yang disamakan dengan nol menghasilkan persamaan normal:

$$JKG = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2 \tag{2.2}$$

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

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

$$0 = -2\sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)$$

$$0 = \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \hat{\beta}_0 - \sum_{i=1}^{n} \hat{\beta}_i X_i$$

$$0 = \sum_{i=1}^{n} Y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} X_i$$

$$n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i \tag{2.3}$$

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

$$0 = 2\sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)(-X_i)$$

$$0 = \sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} \hat{\beta}_0 X_i - \sum_{i=1}^{n} \hat{\beta}_i X_i^2$$

$$0 = \sum_{i=1}^{n} X_i Y_i - \hat{\beta}_0 \sum_{i=1}^{n} X_i - \hat{\beta}_1 \sum_{i=1}^{n} X_i^2$$

$$\hat{\beta}_0 \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i Y_i \qquad (2.4)$$

Persamaan normal (2.3) dan (2.4) berbentuk matriks:

$$\begin{bmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{bmatrix} \text{ atau}$$

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$$

di mana: $\mathbf{Y}_{n \times 1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_3 \end{bmatrix}$ $\mathbf{X}_{n \times 2} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$ dan $\hat{\boldsymbol{\beta}}_{2 \times 1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$

untuk memperoleh penduga bagi $\boldsymbol{\beta}$ maka setiap ruas persamaan normal dikalikan dengan $(\mathbf{X}'\mathbf{X})^{-1}$:

$$(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

karena $(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X}) = \mathbf{I}$,sehingga penduga bagi $\boldsymbol{\beta}$ adalah:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \qquad (2.5)$$

(Drapper dan Smith, 1992)

2.1.1 Pengujian Hipotesis Parameter Regresi Linier Sederhana

Pengujian dilakukan untuk mengetahui apakah terdapat hubungan kausal antara peubah prediktor dan peubah respon berdasarkan hipotesis:

H_0: $\beta_1 = 0$, lawan H_1: $\beta_1 \neq 0$

Sebaran penarikan contoh bagi $\hat{\beta}_1$ adalah $\hat{\beta}_1 \sim N(\beta_1, \text{var}(\hat{\beta}_1))$

Jika H_0 benar, statistik uji t:

$$\frac{\hat{\beta}_1}{Se(\hat{\beta}_1)} \sim t_{(n-2)} \tag{2.6}$$

di mana:

$$Se(\hat{\beta}_1) = \hat{\sigma}\sqrt{c_{jj}}$$

C_{jj} = elemen ke-*jj* matriks $(\mathbf{X'X})^{-1}$

Apabila statistik uji $t > t_{\alpha/2(n-2)}$ maka H_0 ditolak, peubah prediktor berpengaruh nyata terhadap peubah respon (Demaris, 2004).

2.2 Model Regresi Linier Berganda

Model regresi linier berganda memodelkan hubungan beberapa peubah prediktor dan sebuah peubah respon. Hubungan fungsional peubah respon *Y* dan *k* peubah prediktor ($X_1, X_2, ..., X_k$) dijelaskan oleh model regresi:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_k X_{ik} + \varepsilon_i \qquad i = 1, 2, ..., n$$

$$Y_i = \beta_0 + \sum_{j=1}^{k} \beta_j X_{ij} + \varepsilon_i \qquad j = 1, 2, ..., k \tag{2.7}$$

di mana:
- n = banyaknya pengamatan
- k = banyaknya prediktor
- $p = k+1$ = banyaknya parameter dalam model
- β_0 = intersep
- β_j = koefisien regresi prediktor ke-j
- ε_i = galat pengamatan ke-i

Pendugaan terhadap parameter β_0 dan β_j sama seperti pada regresi linier sederhana yaitu dengan menggunakan metode kuadrat terkecil melalui peminimuman jumlah kuadrat galat:

$$\varepsilon_i = Y_i - \hat{Y}_i$$

$$\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_{i1} + ... + \beta_k X_{ik})$$

$$JKG=\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}(Y_i-\beta_0-\sum_{j=1}^{k}\beta_j X_{ij})^2$$

menurunkan jumlah kuadrat galat terhadap $\beta_0,\ \beta_1,\ ...,\ \beta_k$ dan disamakan dengan nol menghasilkan *p* persamaan normal:

$$\frac{\partial JKG}{\partial \beta_0}=-2\sum_{i=1}^{n}(Y_i-\hat{\beta}_0-\sum_{j=1}^{k}\hat{\beta}_j X_{ij})=0$$

dan

$$\frac{\partial JKG}{\partial \beta_j}=-2\sum_{i=1}^{n}(Y_i-\hat{\beta}_0-\sum_{j=1}^{k}\hat{\beta}_j X_{ij})X_{ij}=0$$

atau

$$\begin{array}{ccccccccc}
n\hat{\beta}_0 & + & \hat{\beta}_1\sum_{i=1}^{n}x_{i1} & + & \hat{\beta}_2\sum_{i=1}^{n}x_{i2} & +\ldots+ & \hat{\beta}_k\sum_{i=1}^{n}x_{ik} & = & \sum_{i=1}^{n}y_i \\
\hat{\beta}_0\sum_{i=1}^{n}x_{i1} & + & \hat{\beta}_1\sum_{i=1}^{n}x_{i1}^2 & + & \hat{\beta}_2\sum_{i=1}^{n}x_{i1}x_{i2} & +\ldots+ & \hat{\beta}_k\sum_{i=1}^{n}x_{ik}x_{ik} & = & \sum_{i=1}^{n}x_{i1}y_i \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\hat{\beta}_0\sum_{i=1}^{n}x_{ik} & + & \hat{\beta}_1\sum_{i=1}^{n}x_{ik}x_{i1} & + & \hat{\beta}_2\sum_{i=1}^{n}x_{ik}x_{i2} & +\ldots+ & \hat{\beta}_k\sum_{i=1}^{n}x_{ik}^2 & = & \sum_{i=1}^{n}x_{ik}y_i
\end{array}$$

dalam bentuk matriks:

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik}x_{ik} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum_{i=1}^{n} x_{ik} & \sum_{i=1}^{n} x_{ik}x_{i1} & \sum_{i=1}^{n} x_{ik}x_{i2} & \cdots & \sum_{i=1}^{n} x_{ik}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1}y_i \\ \vdots \\ \sum_{i=1}^{n} x_{ik}y_i \end{bmatrix}$$

atau

$$(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{Y})$$

di mana:

$$\mathbf{Y}_{n \times 1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X}_{n \times p} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \text{ dan } \hat{\boldsymbol{\beta}}_{p \times 1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

Dengan cara yang sama akan diperoleh penduga bagi **β** menurut persamaan (2.5) (Drapper dan Smith, 1992).

2.2.1 Pengujian Hipotesis Parameter Regresi Linier Berganda

Pengujian koefisien regresi linier berganda dilakukan untuk mengetahui apakah antar peubah respon dan *k* peubah prediktor terdapat hubungan linier. Hipotesis yang melandasi pengujian adalah:

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0$

H_1 : paling tidak ada satu j di mana $\beta_j \neq 0$ $\qquad$ j = 1, 2, ..., k

Pengujian dilandasi pada uji F berdasarkan tabel 2.1.

Tabel 2.1. Analisis Ragam Pengujian Parameter Regresi Linier Berganda

Sumber keragaman	db	Jumlah kuadrat	Kuadrat tengah	Statistik uji F
Regresi	k	$\hat{\beta}' X' Y - n\bar{Y}^2$	$\frac{JK_{reg}}{db_{reg}}$	$\frac{KT_{reg}}{KT_{galat}}$ ~ $F_{k,(n-p)}$
Galat	n - p	$Y'Y - \hat{\beta}' X' Y$	$\frac{JK_{galat}}{db_{galat}}$	
Total	n-1	$Y'Y - n\bar{Y}^2$		

Tolak H_0 jika statistik uji $F > F_{\alpha(k,n-p)}$, paling tidak terdapat satu peubah prediktor (X_1, X_2, ..., X_k) yang berpengaruh nyata terhadap peubah respon (Hines dan Montgomery, 1990).

2.3 Pengukuran Kecocokan Model

Salah satu indikator kesesuaian data dengan model adalah koefisien determinasi yang disesuaikan. Besaran ini menunjukkan proporsi keragaman total dalam respon *Y* yang diterangkan oleh model.

$$R^2_{adj} = 1 - \frac{JKG/(n-p)}{JKT/(n-1)} = 1 - \frac{KTG}{KTT} \tag{2.8}$$

Semakin dekat R^2_{adj} dengan 1 maka semakin cocok data dengan model (Walpole dan Myers, 1986).

2.4 Peubah Boneka

Model regresi yang disajikan pada (2.1) dan (2.6) didasarkan pada peubah prediktor yang bersifat kuantitatif, padahal suatu penelitian dapat pula melibatkan peubah kualitatif. Sebagai contoh, diketahui bahwa jenis kelamin mempengaruhi prestasi belajar mahasiswa (IPK). Model regresi dengan peubah prediktor kualitatif, yaitu peubah boneka dengan penyandian 0 dan 1, adalah:

$$Y_i = \beta_0 + \beta_1 D_{i1} + \varepsilon_i \tag{2.9}$$

di mana Y_i = prestasi belajar mahasiswa ke-i

$$D_{i1} = \begin{cases} 1, & \text{jika mahasiswa ke-i perempuan} \\ 0, & \text{jika mahasiswa ke-i laki-laki} \end{cases}$$

Gujarati (1991) mencirikan model regresi yang mengandung peubah boneka:

1. Jika suatu peubah kualitatif terdiri dari *m* kategori, akan terdapat *m–1* peubah boneka. Pandang model berikut:

$$Y_i = \beta_0 + \beta_1 D_{i1} + \beta_2 D_{i2} + \beta_3 X_i + \varepsilon_i \qquad (2.10)$$

di mana:

$$D_{i1} = \begin{cases} 1, & \text{jika guru ke-i pria} \\ 0, & \text{jika guru ke-i wanita} \end{cases}$$

$$D_{i2} = \begin{cases} 1, & \text{jika guru ke-i wanita} \\ 0, & \text{jika guru ke-i pria} \end{cases}$$

dengan matriks data sebagai berikut :

$$\begin{array}{c} \\ P \\ P \\ P \\ W \\ W \\ W \end{array} \begin{array}{c} \begin{array}{ccccc} Y_i & & D_{i1} & D_{i2} & X_{i1} \end{array} \\ \begin{bmatrix} Y_1 & 1 & 1 & 0 & X_{11} \\ Y_2 & 1 & 1 & 0 & X_{21} \\ Y_3 & 1 & 1 & 0 & X_{31} \\ Y_4 & 1 & 0 & 1 & X_{41} \\ Y_5 & 1 & 0 & 1 & X_{51} \\ Y_6 & 1 & 0 & 1 & X_{61} \end{bmatrix} \end{array}$$

tampak bahwa D_1 dan D_2 mengalami kolinear secara sempurna yaitu $D_1 = 1 - D_2$ atau $D_2 = 1 - D_1$. Jika terjadi kolinieritas sempurna maka penduga MKT tidak dapat dilakukan oleh karena itu hanya digunakan sebanyak $m - 1$ peubah boneka untuk mengatasi kolinieritas sempurna sehingga matriks data yang digunakan adalah:

$$
\begin{array}{c}
\\
P \\ P \\ P \\ W \\ W \\ W
\end{array}
\begin{array}{c}
\begin{array}{cccc} Y_i & & D_{i1} & X_{i1} \end{array} \\
\begin{bmatrix}
Y_1 & 1 & 1 & X_{11} \\
Y_2 & 1 & 1 & X_{21} \\
Y_3 & 1 & 1 & X_{31} \\
Y_4 & 1 & 0 & X_{41} \\
Y_5 & 1 & 0 & X_{51} \\
Y_6 & 1 & 0 & X_{61}
\end{bmatrix}
\end{array}
$$

Penggunaan matriks data ini telah menghilangkan kolinieritas sempurna sehingga penduga MKT dapat dilakukan. Cara lain untuk mengatasi kolinieritas sempurna pada peubah boneka yaitu menggunakan peubah boneka sebanyak jumlah kategori yang ada tetapi intersep tidak digunakan dalam model ini, sehingga matriks data yang terbentuk adalah (O'Neill, 2013):

$$
\begin{array}{c}
\\
P \\ P \\ P \\ W \\ W \\ W
\end{array}
\begin{array}{c}
\begin{array}{cccc} Y_i & D_{i1} & D_{i2} & X_{i1} \end{array} \\
\begin{bmatrix}
Y_1 & 1 & 0 & X_{11} \\
Y_2 & 1 & 0 & X_{21} \\
Y_3 & 1 & 0 & X_{31} \\
Y_4 & 0 & 1 & X_{41} \\
Y_5 & 0 & 1 & X_{51} \\
Y_6 & 0 & 1 & X_{61}
\end{bmatrix}
\end{array}
$$

Matriks data ini tidak menggunakan intersep dan membuat peubah boneka untuk setiap kategori, sehingga setelah model regresi terbentuk akan didapatkan nilai rata-rata dari setiap kategori.

2. Penetapan nilai 0 dan 1 untuk dua kategori, misal pria dan wanita adalah tanpa dasar. Hal ini menunjukkan bahwa $D = 1$ untuk wanita dan $D = 0$ untuk pria atau sebaliknya.
3. Kelompok, kategori atau klasifikasi yang bernilai nol merupakan kategori dasar, kontrol atau perbandingan (*reference category*).
4. Koefisien β pada peubah boneka D disebut koefisien intersep diferensial karena menyatakan berapa banyak nilai unsur intersep dari kategori yang mendapat nilai 1 berbeda dari koefisien intersep dari kategori dasar.

2.5 Model Regresi Linier Berkelompok

Model regresi linier berkelompok memodelkan hubungan fungsional antara peubah respon dan peubah prediktor kualitaitf dan kuantitatif. Terdapat 4 jenis model regresi linier berkelompok yang akan dijelaskan pada subbab 2.5.1 sampai 2.5.4.

2.5.1 Model Regresi Garis Sejajar dengan Perbedaan Penduga pada Kategori Dasar

Pada model regresi garis sejajar dengan perbedaan penduga pada kategori dasar akan dijelaskan tiga model dengan perbedaan jumlah kelompok pada peubah prediktor kualitatif.

2.5.1.1 Model Regresi Satu Peubah Kuantitatif dan Satu Peubah Kualitatif dengan Dua Kelompok

Pandang persamaan (2.9) dengan tambahan dugaan bahwa lama belajar juga mempengaruhi prestasi belajar mahasiswa maka persamaan (2.9) akan menjadi:

$$Y_i = \beta_0 + \beta_1 D_{i1} + \beta_2 X_{i1} + \varepsilon_i \tag{2.11}$$

di mana: Y_i = prestasi belajar mahasiswa ke-i

X_{i1} = lama belajar mahasiswa ke-i (jam/hari)

$$D_{i1} = \begin{cases} 1, & \text{jika mahasiswa ke-i perempuan} \\ 0, & \text{jika mahasiswa ke-i laki-laki} \end{cases}$$

Dengan asumsi $E(\varepsilon_i) = 0$,

$$E(Y_i) = E(\beta_0 + \beta_1 D_{i1} + \beta_2 X_{i1} + \varepsilon_i)$$

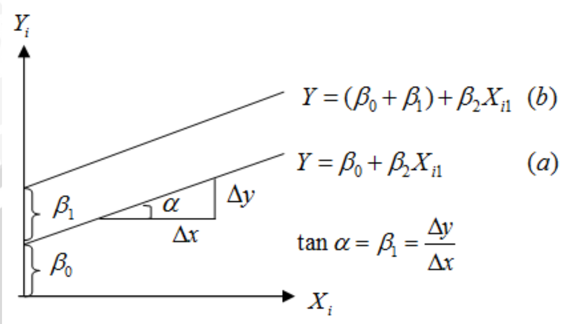
$$E(Y_i) = E(\beta_0 + \beta_1 D_{i1} + \beta_2 X_{i1}) + E(\varepsilon_i)$$

$$E(Y_i) = E(\beta_0 + \beta_1 D_{i1} + \beta_2 X_{i1}) \tag{2.12}$$

maka persamaan (2.12) dapat dibentuk menjadi persamaan prestasi belajar mahasiswa berdasarkan lama belajar untuk setiap jenis kelamin:

Laki-laki: $$E(Y_i \mid X_i, D_{i1} = 0) = \beta_0 + \beta_2 X_{i1} \tag{2.13}$$

Perempuan : $$E(Y_i \mid X_i, D_{i1} = 1) = (\beta_0 + \beta_1) + \beta_2 X_{i1} \tag{2.14}$$



Gambar 2.1. Garis regresi kelompok laki-laki (a) dan perempuan (b)

Model (2.13) dan (2.14) memperlihatkan bahwa prestasi belajar mahasiswa perempuan β_1 lebih tinggi dibanding prestasi mahasiswa laki-laki tetapi laju pertambahan prestasi belajar mahasiswa laki-laki dan perempuan adalah sama sehingga jenis kelamin tidak menentukan laju pertambahan prestasi belajar.

Pendugaan terhadap parameter β_0, β_1 dan β_2 dilakukan dengan metode kuadrat terkecil seperti pada persamaan (2.3) dan (2.4) sehingga diperoleh penduga bagi β pada (2.5), di mana:

$$\mathbf{Y}_{n\text{x}1} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X}_{n\text{x}3} = \begin{array}{c} \begin{matrix} & x_{i1} & D_{i1} \end{matrix} \\ \begin{bmatrix} 1 & x_{11} & D_{11} \\ 1 & x_{21} & D_{21} \\ 1 & x_{31} & D_{31} \\ 1 & x_{41} & D_{41} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & D_{n1} \end{bmatrix} \end{array} \quad \mathbf{X'Y}_{3\text{x}1} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \\ \sum_{i=1}^{n} D_{i1} y_i \end{bmatrix}$$

$$\mathbf{X'X}_{3\text{x}3} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} D_{i1} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1} D_{i1} \\ \sum_{i=1}^{n} D_{i1} & \sum_{i=1}^{n} x_{i1} D_{i1} & \sum_{i=1}^{n} D_{i1}^2 \end{bmatrix} \qquad \hat{\boldsymbol{\beta}}_{3\text{x}1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

2.5.1.2 Model Regresi Satu Peubah Kuantitatif dan Satu Peubah Kualitatif dengan lebih dari Dua Kelompok

Ingin diketahui hubungan antar lama bekerja dan tingkat pendidikan (SD, SMP, SMA) terhadap gaji karyawan. Peubah boneka yang digunakan ada 2 karena ada 3 kelompok tingkat pendidikan sehingga model yang terbentuk:

$$Y_i = \beta_0 + \beta_1 D_{i1} + \beta_2 D_{i2} + \beta_3 X_{i1} + \varepsilon_i \qquad (2.15)$$

di mana: Y_i = gaji karyawan ke-i

X_{i1} = lama bekerja karyawan ke-i (jam/hari)

$$D_{i1} = \begin{cases} 1, & \text{jika karyawan ke-i berpendidikan SMP} \\ 0, & \text{lainnya} \end{cases}$$

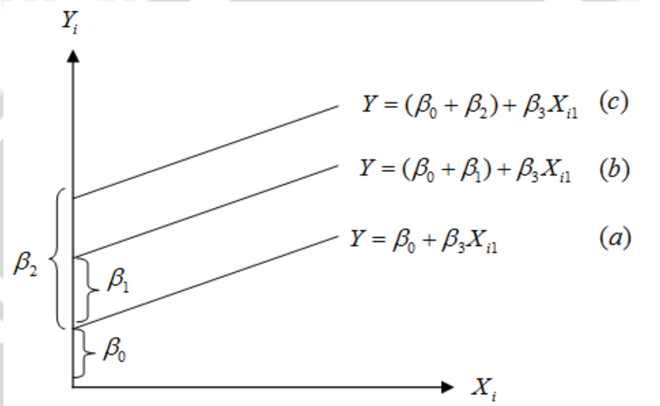
$$D_{i2} = \begin{cases} 1, & \text{jika karyawan ke-i berpendidikan SMA} \\ 0, & \text{lainnya} \end{cases}$$

Persamaan (2.15) memperlihatkan kelompok pendidikan SD sebagai kategori dasar sehingga β_0 merupakan intersep untuk kategori ini. Dengan mengasumsikan bahwa $E(\varepsilon_i) = 0$ seperti pada persamaan (2.12) maka persamaan gaji karyawan berdasarkan lama bekerja untuk tingkat pendidikan :

SD: $$E(Y_i | X_i, D_{i1} = 0, D_{i2} = 0) = \beta_0 + \beta_3 X_{i1} \qquad (2.16)$$

SMP: $$E(Y_i | X_i, D_{i1} = 1, D_{i2} = 0) = (\beta_0 + \beta_1) + \beta_3 X_{i1} \qquad (2.17)$$

SMA: $E(Y_i \mid X_i, D_{i1}=0, D_{i2}=1) = (\beta_0+\beta_2)+\beta_3 X_{i1}$ (2.18)



Gambar 2.2. Garis regresi pada kelompok SD (a), SMP (b) dan SMA (c)

Pendugaan parameter bagi β diperoleh dengan cara yang sama seperti pada persamaan (2.3) dan (2.4), di mana:

$$\mathbf{Y}_{n\text{x}1} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X}_{n\text{x}4} = \begin{array}{c} \begin{array}{cccc} & x_{i1} & D_{i1} & D_{i2} \end{array} \\ \begin{bmatrix} 1 & x_{11} & D_{11} & D_{12} \\ 1 & x_{21} & D_{21} & D_{22} \\ 1 & x_{31} & D_{31} & D_{32} \\ 1 & x_{41} & D_{41} & D_{42} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & D_{n1} & D_{n2} \end{bmatrix} \end{array} \quad \mathbf{X}'\mathbf{Y}_{4\text{x}1} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \\ \sum_{i=1}^{n} D_{i1} y_i \\ \sum_{i=1}^{n} D_{i2} y_i \end{bmatrix}$$

$$\hat{\boldsymbol{\beta}}_{4\text{x}1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} \quad \mathbf{X'X}_{4\text{x}4} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} D_{i1} & \sum_{i=1}^{n} D_{i2} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}D_{i1} & \sum_{i=1}^{n} x_{i1}D_{i2} \\ \sum_{i=1}^{n} D_{i1} & \sum_{i=1}^{n} x_{i1}D_{i1} & \sum_{i=1}^{n} D_{i1}^2 & \sum_{i=1}^{n} D_{i1}D_{i2} \\ \sum_{i=1}^{n} D_{i2} & \sum_{i=1}^{n} x_{i1}D_{i1} & \sum_{i=1}^{n} D_{i1}D_{i2} & \sum_{i=1}^{n} D_{i2}^2 \end{bmatrix}$$

2.5.1.3 Model Regresi Satu Peubah Kuantitatif dan Dua Peubah Kualitatif

Pandang persamaan (2.11) dengan mempertimbangkan bahwa daerah asal juga mempengaruhi prestasi belajar mahasiswa dan anggap bahwa ada dua kelompok daerah asal yaitu kota dan desa maka model regresi yang terbentuk adalah:

$$Y_i = \beta_0 + \beta_1 D_{i1} + \beta_2 D_{i2} + \beta_3 X_{i1} + \varepsilon_i \qquad (2.19)$$

di mana: Y_i = prestasi belajar mahasiswa ke-i

X_{i1} = lama belajar mahasiswa ke-i

$$D_{i1} = \begin{cases} 1, & \text{jika mahasiswa perempuan ke-i} \\ 0, & \text{lainnya} \end{cases}$$

$$D_{i2} = \begin{cases} 1, & \text{jika mahasiswa ke-i berasal dari kota} \\ 0, & \text{lainnya} \end{cases}$$

Berdasarkan asumsi $E(\varepsilon_i) = 0$ seperti pada persamaan (2.12) maka didapatkan persamaan prestasi belajar mahasiswa berdasarkan lama belajar untuk setiap jenis kelamin dan asal daerah:

1. Laki-laki dari desa:

$$E(Y_i \mid X_i, D_{i1} = 0, D_{i2} = 0) = \beta_0 + \beta_3 X_{i1} \qquad (2.20)$$

2. Perempuan dari desa:

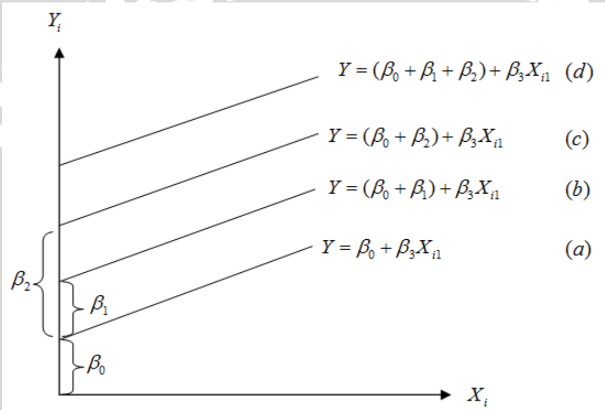
$$E(Y_i | X_i, D_{i1} = 1, D_{i2} = 0) = (\beta_0 + \beta_1) + \beta_3 X_{i1} \quad (2.21)$$

3. Laki-laki dari kota:

$$E(Y_i | X_i, D_{i1} = 0, D_{i2} = 1) = (\beta_0 + \beta_2) + \beta_3 X_{i1} \quad (2.22)$$

4. Perempuan dari kota:

$$E(Y_i | X_i, D_{i1} = 1, D_{i2} = 1) = (\beta_0 + \beta_1 + \beta_2) + \beta_3 X_{i1} \quad (2.23)$$



Gambar 2.3. Garis regresi pada kelompok laki-laki berasal dari desa (a), perempuan berasal dari desa (b), laki-laki berasal dari kota (c), perempuan berasal dari kota (d)

Pendugaan parameter $\beta_0, \beta_1, \beta_2$ dan β_3 model regresi dengan satu peubah kuantitatif dan dua peubah kualitatif sama seperti pada pendugaan parameter model (2.15) tetapi berbeda dalam interpretasi hasil sesuai pada persamaan (2.20) sampai (2.23) (Gujarati dan Porter, 2010).

2.5.2 Model Regresi Garis Terpisah dengan Perbedaan Penduga pada Kategori Dasar

Model regresi garis terpisah dengan perbedaan penduga dari kategori dasar merupakan model yang mengandung interaksi antara peubah prediktor. Pandang persamaan (2.11), di mana satu peubah prediktor bersifat kualitatif. Pengaruh interaksi pada model dapat digunakan dengan cara biasa dan menambahkan perkalian silang pada model. Persamaan regresi (2.11) dengan menambahkan interaksi pada lama belajar dan jenis kelamin adalah:

$$Y_i = \beta_0 + \beta_1 D_{i1} + \beta_2 X_{i1} + \beta_3 D_{i1} X_{i1} + \varepsilon_i \qquad (2.24)$$

di mana: Y_i = prestasi belajar mahasiswa ke-i

X_{i1} = lama belajar mahasiswa ke-i (jam/hari)

$$D_{i1} = \begin{cases} 1, & \text{jika mahasiswa ke-i perempuan} \\ 0, & \text{jika mahasiswa ke-i laki-laki} \end{cases}$$

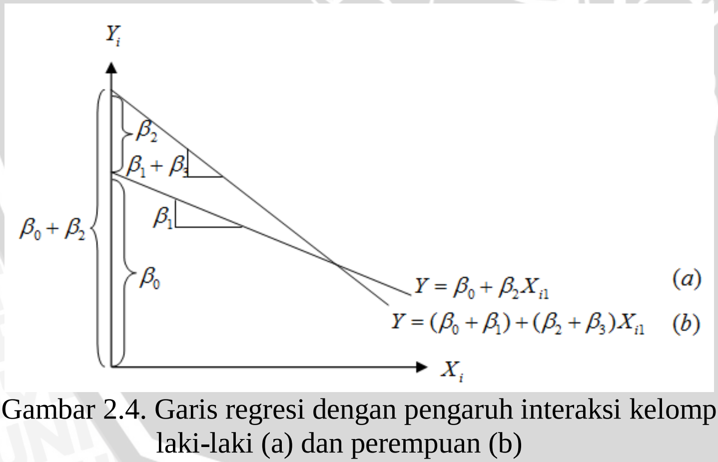
dengan mengasumsikan bahwa $E(\varepsilon_i) = 0$ seperti pada persamaan (2.12) maka diperoleh:

$$E(Y_i) = \beta_0 + \beta_1 D_{i1} + \beta_2 X_{i1} + \beta_3 D_{i1} X_{i1} \qquad (2.25)$$

Dari persamaan (2.25) didapatkan prestasi belajar mahasiswa berdasarkan lama belajar untuk setiap jenis kelamin:

Laki-laki: $$E(Y_i \mid X_i, D_{i1} = 0, D_{i1}X_{i1} = 0) = \beta_0 + \beta_2 X_{i1} \qquad (2.26)$$

Perempuan: $$E(Y_i \mid X_i, D_{i1} = 1, D_{i1}X_{i1} = X_{i1}) = (\beta_0 + \beta_1) + (\beta_2 + \beta_3) X_{i1}$$
(2.27)



Gambar 2.4. Garis regresi dengan pengaruh interaksi kelompok laki-laki (a) dan perempuan (b)

Pendugaan parameter model yang mengandung interaksi dilakukan dengan metode kuadrat terkecil seperti pada persamaan (2.3) dan (2.4) sehingga didapatkan $\hat{\beta}$ pada (2.5), di mana:

$$
\mathbf{X'Y}_{4\text{x}1} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \\ \sum_{i=1}^{n} D_{i1} y_i \\ \sum_{i=1}^{n} D_{i1} x_{i1} y_i \end{bmatrix}
\quad
\mathbf{X}_{n\text{x}4} = \begin{array}{c} \begin{array}{cccc} & x_{i1} & D_{i1} & D_{i1}x_{i1} \end{array} \\ \begin{bmatrix} 1 & x_{11} & D_{11} & D_{11}x_{11} \\ 1 & x_{21} & D_{21} & D_{21}x_{21} \\ 1 & x_{31} & D_{31} & D_{31}x_{31} \\ 1 & x_{41} & D_{41} & D_{41}x_{41} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & D_{n1} & D_{n1}x_{n1} \end{bmatrix} \end{array}
\quad
\hat{\boldsymbol{\beta}}_{4\text{x}1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}
$$

$$
\mathbf{X'X}_{4\text{x}4} = \begin{bmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} D_{i1} & \sum_{i=1}^{n} x_{i1} D_{i1} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1} D_{i1} & \sum_{i=1}^{n} x_{i1}^2 D_{i1} \\ \sum_{i=1}^{n} D_{i1} & \sum_{i=1}^{n} x_{i1} D_{i1} & \sum_{i=1}^{n} D_{i1}^2 & \sum_{i=1}^{n} x_{i1} D_{i1}^2 \\ \sum_{i=1}^{n} x_{i1} D_{i1} & \sum_{i=1}^{n} x_{i1}^2 D_{i1} & \sum_{i=1}^{n} x_{i1} D_{i1}^2 & \sum_{i=1}^{n} x_{i1}^2 D_{i1}^2 \end{bmatrix}
\quad
\mathbf{Y}_{n\text{x}1} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ y_n \end{bmatrix}
$$

(Kutner, 2005).

2.5.3 Model Regresi Garis Sejajar

Pandang persamaan (2.11), menurut O’Neill (2013) model regresi garis sejajar yang terbentuk yaitu:

$$Y_i = \beta_1 D_{i1(a)} + \beta_2 D_{i1(b)} + \beta_3 X_{i1} + \varepsilon_i \qquad (2.28)$$

di mana: Y_i = prestasi belajar mahasiswa ke-i

X_{i1} = lama belajar mahasiswa ke-i (jam/hari)

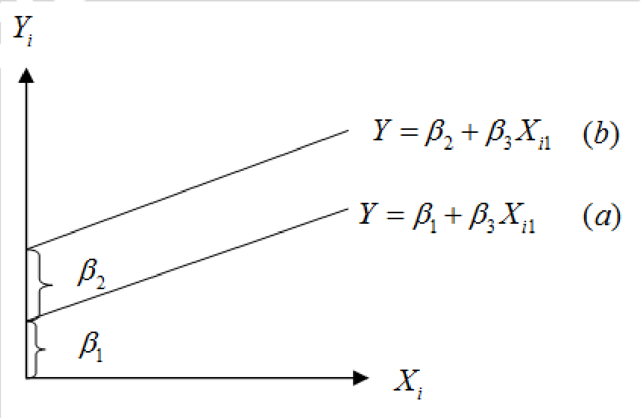
$$D_{i1(a)} = \begin{cases} 1, & \text{jika mahasiswa ke-i perempuan} \\ 0, & \text{jika mahasiswa ke-i laki-laki} \end{cases}$$

$$D_{i1(b)} = \begin{cases} 1, & \text{jika mahasiswa ke-i laki-laki} \\ 0, & \text{jika mahasiswa ke-i perempuan} \end{cases}$$

Dengan asumsi $E(\varepsilon_i) = 0$, maka persamaan (2.28) dapat dibentuk menjadi persamaan untuk prestasi belajar mahasiswa berdasarkan lama belajar pada setiap jenis kelamin:

Perempuan: $E(Y_i \mid X_i, D_{i1(a)} = 1, D_{i1(b)} = 0) = \beta_1 + \beta_3 X_{i1}$ (2.29)

Laki-laki: $E(Y_i \mid X_i, D_{i1(a)} = 0, D_{i1(b)} = 1) = \beta_2 + \beta_3 X_{i1}$ (2.30)



Gambar 2.5. Garis regresi kelompok perempuan (a) dan laki-laki (b)

Pendugaan parameter bagi β diperoleh dengan cara yang sama seperti pada persamaan (2.3) dan (2.4), di mana:

$$\mathbf{Y}_{n\text{x}1} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ y_n \end{bmatrix} \qquad \mathbf{X}_{n\text{x}3} = \begin{array}{c} \begin{array}{ccc} D_{i1(a)} & D_{i1(b)} & x_{i1} \end{array} \\ \begin{bmatrix} D_{11(a)} & D_{11(b)} & x_{11} \\ D_{21(a)} & D_{21(b)} & x_{21} \\ D_{31(a)} & D_{31(b)} & x_{31} \\ D_{41(a)} & D_{41(b)} & x_{41} \\ \vdots & \vdots & \vdots \\ D_{n1(a)} & D_{n1(b)} & x_{n1} \end{bmatrix} \end{array}$$

$$\mathbf{X'Y}_{3\text{x}1} = \begin{bmatrix} \sum_{i=1}^{n} D_{i1(a)} y_i \\ \sum_{i=1}^{n} D_{i1(b)} y_i \\ \sum_{i=1}^{n} x_{i1} y_i \end{bmatrix} \qquad \hat{\boldsymbol{\beta}}_{3\text{x}1} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$$

$$\mathbf{X'X}_{3\text{x}3} = \begin{bmatrix} \sum_{i=1}^{n} D_{i1(a)}^2 & \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} & \sum_{i=1}^{n} D_{i1(a)} x_{i1} \\ \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} & \sum_{i=1}^{n} D_{i1(b)}^2 & \sum_{i=1}^{n} D_{i1(b)} x_{i1} \\ \sum_{i=1}^{n} D_{i1(a)} x_{i1} & \sum_{i=1}^{n} D_{i1(b)} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 \end{bmatrix}$$

2.5.4 Model Regresi Garis Terpisah

Merujuk pada persamaan (2.11), model regresi garis terpisah adalah:

$$Y_i = \beta_1 D_{i1(a)} + \beta_2 D_{i1(b)} + \beta_3 D_{i1(a)} X_{i1} + \beta_4 D_{i1(b)} X_{i1} + \varepsilon_i \qquad (2.31)$$

di mana: Y_i = prestasi belajar mahasiswa ke-i

X_{i1} = lama belajar mahasiswa ke-i (jam/hari)

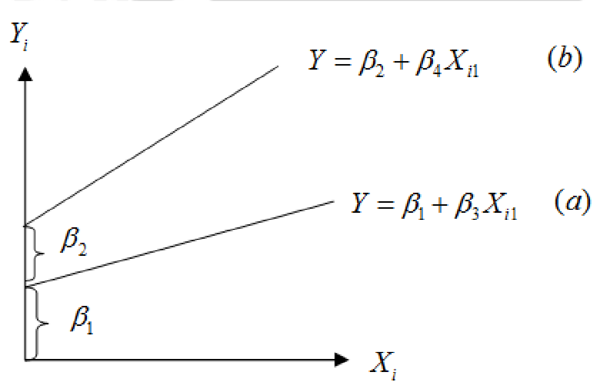
$$D_{i1(a)} = \begin{cases} 1, & \text{jika mahasiswa ke-i perempuan} \\ 0, & \text{jika mahasiswa ke-i laki-laki} \end{cases}$$

$$D_{i1(b)} = \begin{cases} 1, & \text{jika mahasiswa ke-i laki-laki} \\ 0, & \text{jika mahasiswa ke-i perempuan} \end{cases}$$

Dari persamaan (2.31) didapatkan prestasi belajar mahasiswa berdasarkan lama belajar untuk setiap jenis kelamin:

Perempuan: $E(Y_i \mid X_i, D_{i1(a)} = 1, D_{i1(b)} = 0) = \beta_1 + \beta_3 X_{i1}$ (2.32)

Laki-laki: $E(Y_i \mid X_i, D_{i1(a)} = 0, D_{i1(b)} = 1) = \beta_2 + \beta_4 X_{i1}$ (2.33)



Gambar 2.6. Garis regresi dengan interaksi kelompok perempuan (a) dan laki-laki (b)

Pendugaan parameter model garis terpisah dilakukan dengan metode kuadrat terkecil seperti persamaan (2.3) dan (2.4) sehingga didapatkan $\hat{\beta}$ pada (2.5), di mana:

$$\mathbf{Y}_{n\text{x}1} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ y_n \end{bmatrix} \qquad \mathbf{X}_{n\text{x}4} = \begin{array}{c} \begin{matrix} D_{i1(a)} & D_{i1(b)} & D_{i1(a)}x_{i1} & D_{i1(b)}x_{i1} \end{matrix} \\ \begin{bmatrix} D_{11(a)} & D_{11(b)} & D_{11(a)}x_{11} & D_{11(b)}x_{11} \\ D_{21(a)} & D_{21(b)} & D_{21(a)}x_{21} & D_{21(b)}x_{21} \\ D_{31(a)} & D_{31(b)} & D_{31(a)}x_{31} & D_{31(b)}x_{31} \\ D_{41(a)} & D_{41(b)} & D_{41(a)}x_{41} & D_{41(b)}x_{41} \\ \vdots & \vdots & \vdots & \vdots \\ D_{n1(a)} & D_{n1(b)} & D_{n1(a)}x_{n1} & D_{n1(b)}x_{n1} \end{bmatrix} \end{array}$$

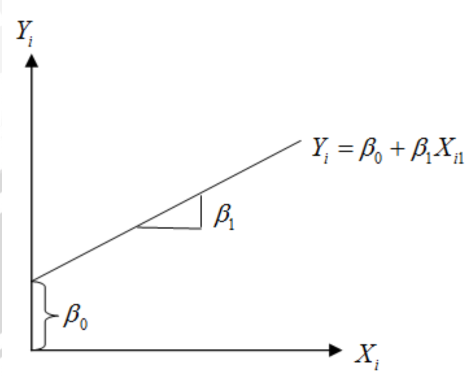
$$\mathbf{X'Y}_{4\text{x}1} = \begin{bmatrix} \sum_{i=1}^{n} D_{i1(a)} y_i \\ \sum_{i=1}^{n} D_{i1(b)} y_i \\ \sum_{i=1}^{n} D_{i1(a)} x_{i1} y_i \\ \sum_{i=1}^{n} D_{i1(b)} x_{i1} y_i \end{bmatrix} \qquad \hat{\boldsymbol{\beta}}_{4\text{x}1} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{bmatrix}$$

$$\mathbf{X'X}_{4\text{x}4} = \begin{bmatrix} \sum_{i=1}^{n} D_{i1(a)}^2 & \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} & \sum_{i=1}^{n} D_{i1(a)}^2 x_{i1} & \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} x_{i1} \\ \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} & \sum_{i=1}^{n} D_{i1(b)}^2 & \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} x_{i1} & \sum_{i=1}^{n} D_{i1(b)}^2 x_{i1} \\ \sum_{i=1}^{n} D_{i1(a)}^2 x_{i1} & \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} x_{i1} & \sum_{i=1}^{n} D_{i1(a)}^2 x_{i1}^2 & \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} x_{i1}^2 \\ \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} x_{i1} & \sum_{i=1}^{n} D_{i1(b)}^2 x_{i1} & \sum_{i=1}^{n} D_{i1(a)} D_{i1(b)} x_{i1}^2 & \sum_{i=1}^{n} D_{i1(b)}^2 x_{i1}^2 \end{bmatrix}$$

(O'Neill, 2013)

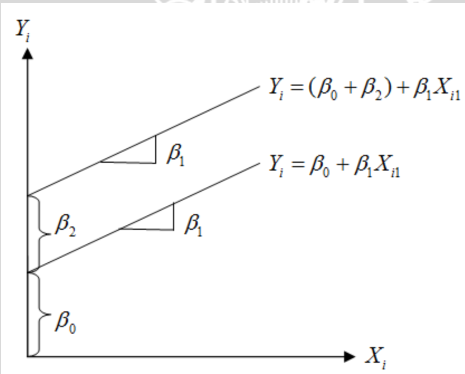
2.6 Perbandingan Dua Garis Regresi

Pandang peubah prediktor bersifat kualitatif dengan dua kelompok (L = Laki-laki dan P = Perempuan). Pertanyaan mendasar terhadap data seperti ini adalah, apakah kelompok satu (L) berbeda dari kelompok lain (P). Apakah penduga garis regresi cukup diwakili satu garis atau dua garis berbeda. Empat kemungkinan model regresi jika dipisahkan berdasarkan kelompok diilustrasikan sebagai berikut:



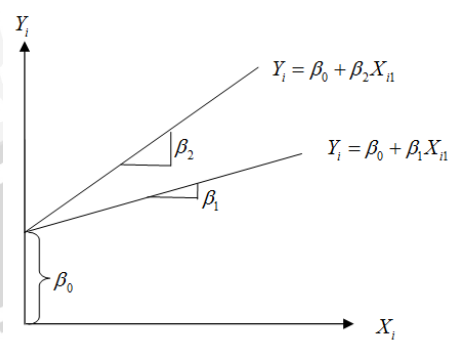
Gambar 2.7. Dua garis regresi identik

Pada garis regresi ini, hanya ada satu persamaan untuk dua kelompok.



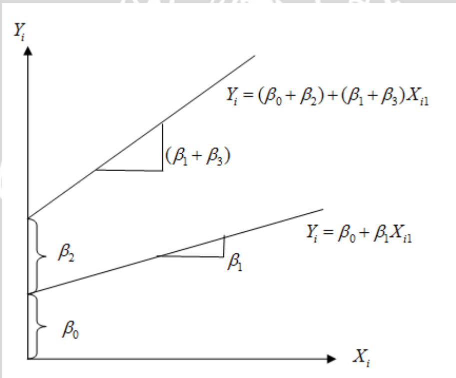
Gambar 2.8. Garis regresi dengan intersep berbeda, slope sama

Garis regresi kedua kelompok berbeda karena memiliki kemiringan sama tetapi intersep berbeda sehingga membentuk dua garis lurus sejajar.



Gambar 2.9. Garis regresi dengan intersep sama, slope berbeda

Garis regresi kedua kelompok berbeda karena memiliki kemiringan berbeda tetapi intersep sama sehingga membentuk dua buah garis.



Gambar 2.10. Garis regresi dengan intersep berbeda, slope berbeda

Garis regresi kedua kelompok berbeda karena memiliki kemiringan maupun intersep berbeda (Gujarati dan Porter, 2010).

2.7 Uji Perbedaan Dua Garis Regresi

Pengujian terhadap dua garis regresi dilakukan untuk mengetahui apakah dua garis regresi secara statistika dapat dikatakan sama atau berbeda. Terdapat dua macam uji yang digunakan untuk membandingkan dua buah garis regresi yaitu uji kesejajaran dan uji keberhimpitan. Pandang dua garis regresi yang akan dilakukan pengujian sebagai berikut:

$$Y_a = \beta_{01} + \beta_1 X_{i1} \tag{2.34}$$

$$Y_b = \beta_{02} + \beta_2 X_{i1} \tag{2.35}$$

Model penduga regresi gabungan persamaan (2.34) dan (2.35) adalah:

$$E(Y) = \beta_0 + \beta_1 D_{i1} + \beta_2 X_{i1} + \beta_3 D_{i1} X_{i1} \tag{2.36}$$

Berdasarkan susunan model penduga regresi gabungan dapat disusun hipotesis untuk uji kesejajaran dan uji keberhimpitan garis.

2.7.1. Pengujian Kesejajaran Garis Regresi

Pengujian kesejajaran dua garis regresi berdasarkan hipotesis:

H_0 : β_3 = 0 (kedua garis regresi sejajar), lawan

H_1 : $\beta_3 \neq 0$ (kedua garis regresi tidak sejajar)

Sebaran penarikan contoh bagi $\hat{\beta}_3$ adalah $\hat{\beta}_3 \sim N(\beta_3, \mathrm{var}(\hat{\beta}_3))$

Tabel 2.2. Analisis Ragam Uji Kesejajaran Garis Regresi

Sumber Keragaman	db	Jumlah kuadrat	Kuadrat tengah	Statistik uji F
Regresi	k	JKR	JKR / k	KTR / KTG
β_3(sej)	1	$JKR - JK_{\beta_0, \beta_1, \beta_2}$	$JKR - JK_{\beta_0, \beta_1, \beta_2}$	KT_{β_3} / KTG
Galat	n-p	$JKT - JKR$	$JKG / n - p$	
total	n-1	JKT		

Jika H_0 benar, statistik uji:

$$\frac{KT_{\beta_3}}{KTG} \sim F_{1,n-p}$$

Jika statistik uji $F > F_{\alpha(1,n-p)}$ maka H_0 ditolak yang berarti belum cukup bukti untuk menerima kenyataan bahwa kedua garis regresi sejajar.

2.7.2. Pengujian Keberhimpitan Garis Regresi

Hipotesis yang melandasi pengujian keberhimpitan adalah:

H_0 : $\beta_2 = \beta_3$ atau $\beta_2 - \beta_3 = 0$ (intersep kedua garis regresi sama) lawan,

H_1 : $\beta_2 \neq \beta_3$ atau $\beta_2 - \beta_3 \neq 0$ (intersep kedua garis regresi berbeda)

Sebaran penarikan contoh:

$$\hat{\beta}_2 \sim N(\beta_2, \operatorname{var}(\hat{\beta}_2))$$

$$\hat{\beta}_3 \sim N(\beta_3, \operatorname{var}(\hat{\beta}_3))$$

Sebaran penarikan contoh bagi $(\hat{\beta}_2 - \hat{\beta}_3)$ adalah:

$$(\hat{\beta}_2 - \hat{\beta}_3) \sim N(\beta_2 - \beta_3, \operatorname{var}(\hat{\beta}_2) + \operatorname{var}(\hat{\beta}_3))$$

Tabel 2.3. Analisis Ragam Uji Keberhimpitan Garis Regresi

Sumber Keragaman	db	Jumlah kuadrat	Kuadrat tengah	Statistik uji F
Regresi	k	JKR	JKR / k	KTR / KTG
β_3	1	$JKR - JK_{\beta_0,\beta_1,\beta_2}$	$JKR - JK_{\beta_0,\beta_1,\beta_2}$	KT_{β_3} / KTG
$\beta_2\ \beta_3$ (him)	2	$JKR - JK_{\beta_0,\beta_1}$	$\dfrac{(JKR - JK_{\beta_0,\beta_1})}{2}$	$KT_{\beta_2,\beta_3} / KTG$
Galat	n-p	$JKT - JKR$	$JKG / n - p$	
total	n-1	JKT		

Jika H_0 benar, statistik uji:

$$\frac{KT_{\beta_2,\beta_3}}{KTG} \sim F_{2,n-p}$$

Jika statistik uji $F > F_{\alpha(2,n-p)}$ maka H_0 ditolak yang berarti bahwa kedua garis regresi berhimpit (Santoso dan Kusnandi, 1992).

BAB III
METODE PENELITIAN

3.1. Sumber Data

Data yang digunakan adalah hasil penelitian Listyorini (2008): "Penampilan 12 Galur Harapan Kacang Panjang (*Vigna unguiculata* var. *sesquipedalis* L. Fruwirth) di Dua Lokasi Dataran Rendah". Penelitian dilaksanakan di dua lokasi di dataran rendah. Lokasi tanam pertama berada di Desa Jatikerto, Kecamatan Kromengan, Kabupaten Malang. Berada pada ketinggian ± 303m dpl, dengan suhu rata-rata harian 25°C. Lokasi tanam kedua di Desa Krembangan, Kecamatan Gudo, Kabupaten Jombang. Berada pada ketinggian tempat ± 94m dpl, dengan suhu rata-rata harian 27-34°C. Penelitian ini dilaksanakan pada bulan Februari sampai dengan Mei 2008, tujuannya ingin mengetahui hubungan panjang polong (cm) dan lokasi tanam terhadap bobot polong segar (g).

3.2. Metode Analisis

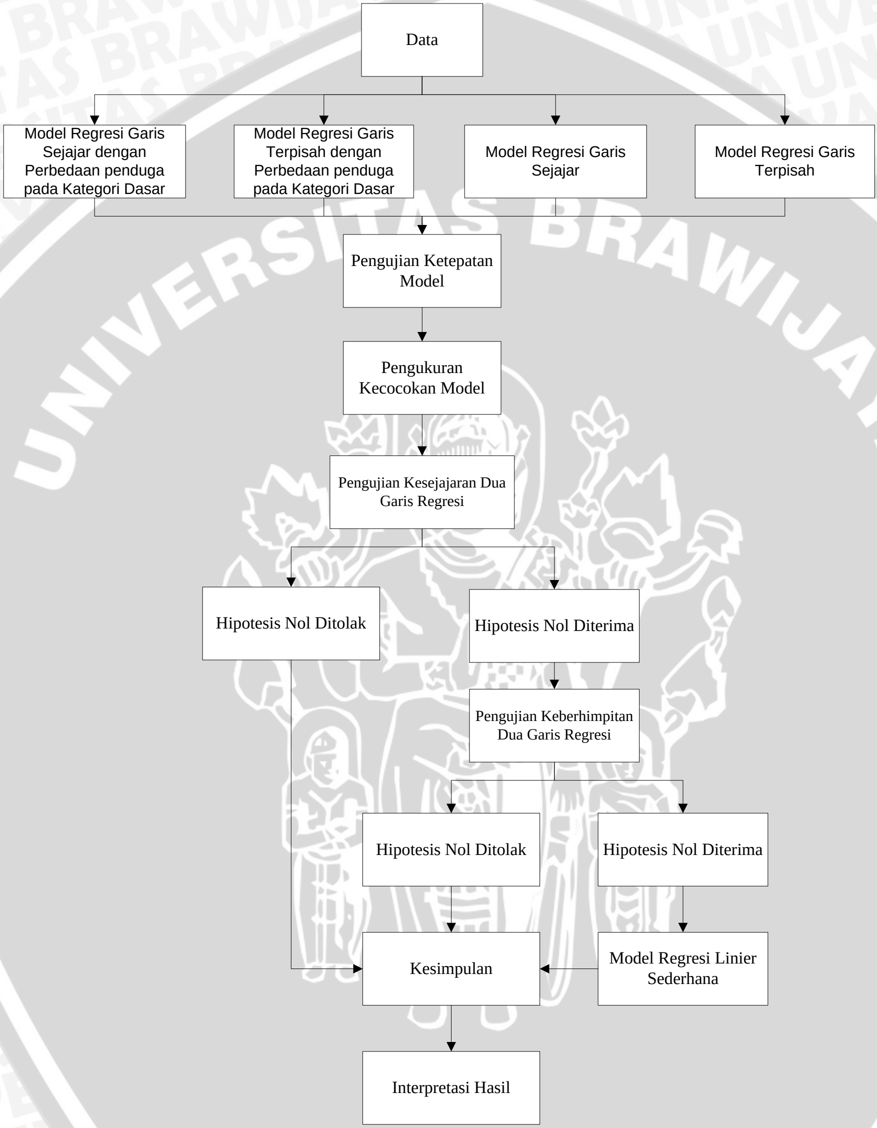
Berikut adalah langkah-langkah pengolahan data:

1. Membentuk 4 jenis model regresi linier berkelompok:
   a. Model regresi garis sejajar dengan perbedaan penduga kategori dasar berdasarkan persamaan (2.11).
   b. Model regresi garis terpisah dengan perbedaan penduga kategori dasar berdasarkan persamaan (2.24).
   c. Model regresi garis sejajar berdasarkan persamaan (2.28).
   d. Model regresi garis terpisah berdasarkan persamaan (2.31).
2. Melakukan pengujian ketepatan model yang telah terbentuk dengan statistik uji F seperti pada Tabel 2.1.
3. Melakukan pengukuran kecocokan model regresi yang tebentuk (2.8).
4. Melakukan uji kesejajaran antar dua regresi berdasarkan Tabel 2.2. Jika H_0 ditolak maka proses analisis berlanjut pada penarikan kesimpulan. Jika H_0 diterima maka disimpulkan bahwa kedua garis regresi sejajar dan berlanjut ke langkah berikutnya.
5. Melakukan uji keberhimpitan dua garis regresi berdasarkan Tabel 2.3. Jika belum cukup bukti untuk menerima bahwa

kedua garis regresi berhimpit maka dapat disimpulkan bahwa kedua garis regresi mempunyai laju pertambahan yang sama tetapi mempunyai rata-rata yang berbeda.

6. Jika pada uji keberhimpitan dua garis regresi H_0 diterima maka dapat disimpulkan bahwa gedua garis bergresi berhimpit sehingga pengaruh kelompok tidak nyata sehingga model yang terbentuk adalah model regresi linier sederhana.
7. Menarik kesimpulan dari model regresi yang terbentuk.
8. Melakukan interpretasi terhadap model yang terbentuk.

Berikut adalah prosedur analisis dalam penelitian ini:



Gambar 3.1. Langkah-langkah analisis regresi linier berkelompok