Prosiding Seminar Nasional FKIP Universitas Mataram
Mataram, 11-12 Oktober 2019

Original Research Paper

Aplikasi Metode Kuadrat Terkecil dan Aturan Cramer's dalam Menyelesaikan Permasalahan Analisis Regresi dan Korelasi

Marzuki[*]

Program Studi Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Mataram, Indonesia

*Corresponding Author: Marzuki, Program Studi Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Mataram, Indonesia;
Email: marzuki.fis@unram.ac.id

Abstract: Salah satu tujuan analisis data dalam statistika adalah untuk memperkirakan atau memprediksi besarnya efek (pengaruh) kuantitatif dari perubahan suatu kejadian terhadap kejadian lain. Untuk keperluan penilaian terhadap efek atau pengaruh tersebut, diperlukan suatu analisis yang membahas hubungan fungsional antara dua atau lebih variabel, yang dikenal dengan nama Analisis Regresi dan Korelasi. Analisis regresi dan korelasi dapat dengan mudah diselesaikan dengan menggunakan metode kuadrat terkecil (least square method) dan aturan Cramer's (Cramer's rule), atau disebut dengan istilah metode determinasi. Metode lain dapat juga digunakan metode inversi dalam mencari formula untuk koefisien-koefisien regresi.. Dengan sedikit melakukan manipulasi matematis dan memberikan pemaknaan baru terhadap suatu variabel, akan memunculkan formula-formula reduktif yang jauh lebih sederhana sehingga akan memudahkan mahasiswa dalam memahaminya.

Keywords: Analisis regresi dan korelasi, aturan Cramer, metode kuadrat terkecil.

Pendahuluan

Salah satu tujuan analisis data adalah untuk memperkirakan atau memprediksi besarnya efek (pengaruh) kuantitatif dari perubahan suatu kejadian terhadap kejadian lain. Untuk keperluan penilaian terhadap efek atau pengaruh tersebut, diperlukan suatu analisis yang membahas hubungan fungsional antara dua atau lebih variabel, yang dikenal dengan nama Analisis Regresi dan korelasi. Analisis Regresi dapat menunjukkan sifat hubungan atau bagaimana pengaruh variabel bebas (variabel yang mempengaruhi) terhadap variabel terikat (variabel yang dipengaruhi), sedangkan analisis korelasi menunjukkan keeratan (kuat-lemahnya) hubungan antara kedua jenis variabel itu (Ruseffendi, 1998; Sudjana, 2002).

Salah satu hal penting dalam analisis regresi dan korelasi adalah menentukan parameter-parameter regresi dan mencari koefisien korelasi. Dalam berbagai buku statistika, persoalan ini diselesaikan dengan rumus-rumus statistika yang sudah jadi, tanpa adanya tuntutan kepada mahasiswa untuk mengetahui atau menelusuri bagaimana rumus-rumus itu diperoleh. Oleh karena itulah penulis merasa perlu untuk menelaah kemunculan formula-formula statistika tersebut untuk menambah pemahaman kita tentang hal terkait.

Persoalan mengenai metode kuadrat terkecil dan aturan Cramers dibahas secara detail dalam kuliah Fisika Matematika. Dengan berbantuan kedua metode ini persoalan mengenai analisis regresi dan korelasi dapat diselesaikan dengan lebih mudah, sehingga berimplikasi pada tingkat pemahaman konsep mahasiswa tentang hal ini yang lebih mendalam.

Pembahasan

Analisis Model Regresi Linier dan Korelasi Sederhana

Analisis Model Regresi Linier Sederhana

Regresi merupakan persamaan matematis yang menunjukkan hubungan fungsional antara variabel bebas (X) dengan variabel terikat (Y), dalam hubungan Y = *f*(X). Regresi yang melibatkan variabel Y dengan hanya satu variabel terikat dinamakan regresi linier sederhana. Secara umum persamaan regresi linier sederhana adalah:

$$Y = A + BX, \quad \text{..............................(2.1)}$$

dimana A dan B adalah parameter-parameter regresi.

Hubungan di atas adalah hubungan matematis. Dalam kenyataannya Y tidak hanya dipengaruhi oleh X saja, melainkan masih banyak faktor lain yang tidak dimasukkan dalam

persamaan. Karena kesalahan itu tidak dapat dihilangkan samasekali, maka risiko itu betapapun kecilnya selalu ada. Risiko hanya bisa diperkecil dengan cara memperkecil kesalahan (*minimized error*). Dengan mempertimbangkan kesalahan pengganggu, ε, maka bentuk persamaan linier (2.1) menjadi: $Y = A + BX + \varepsilon$ . ...(2.2)
Berdasarkan persamaan (2.2), maka nilai Y bisa lebih besar atau lebih kecil dari $A + BX$, bergantung pada nilai ε, apakah positif ataukah negatif.

Dalam praktek, untuk melihat hubungan antara variabel X dan Y, kita mengumpulkan pasangan data (X,Y) sebagai hasil suatu obesrvasi. Apabila populasi dari seluruh pasangan nilai (X,Y) diketahui, kita dapat menghitung nilai *A* dan *B*. Dalam kenyataannya, kita tidak mengetahui nilai parameter tersebut, akan tetapi dapat diperkirakan dengan menggunakan data empiris yakni hasil observasi berdasarkan sampel yang ditarik dari populasi yang tidak terbatas. Untuk memperkirakan nilai *A* dan *B* digunakan **metode kuadrat terkecil** (Sudjana, 2002).

Jadi, model sebenarnya: $Y = A + BX + \varepsilon$ ditaksir dengan model perkiraan/taksiran: $Y = a + bX + e$, dimana *a*, *b*, dan *e* merupakan perkiraan atau taksiran atas *A*, *B*, dan ε.

Metode kuadrat terkecil merupakan suatu metode untuk menghitung nilai *a*, dan *b* (sebagai penduga dari *A* dan *B*), sedemikian rupa sehingga jumlah kuadrat galat (*error*) memiliki nilai terkecil (Supranto, J. 2001). Untuk setiap pasangan data (X,Y), dituliskan $Y_i = a + bX_i + e_i$, dimana i = 1,2,3,...,n, atau $e_i = Y_i - (a + bX_i)$, dengan *e* = *error*.

Jika $\hat{Y}_i = a + bX_i$ ....(2.3)
(dibaca Y topi, yaitu nilai perkiraan Y), maka $e_i = (Y_i - \hat{Y}_i) = Y_i - (a + bX_i)$.
Jumlah kuadrat dari e_i dengan $i = 1$ sampai n adalah :

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$
$$= \sum_{i=1}^{n}(Y_i - (a + bX_i))^2 \text{...... (2.4)}$$

Jadi, metode kuadrat terkecil adalah metode untuk menghitung nilai *a* dan *b*, sedemikian rupa sehingga $\sum_{i=1}^{n} e_i^2$ memiliki nilai terkecil (minimum). Dengan demikian, dalam Fisika Matematika (Fismat) hal ini merupakan persoalan nilai minimum, dimana turunan pertama dari $\sum_{i=1}^{n} e_i^2$ terhadap *a* dan juga terhadap *b* sama dengan nol.

$$\frac{\partial}{\partial a}\left(\sum_{i=1}^{n} e_i^2\right) = 0 \quad \text{dan} \quad \frac{\partial}{\partial b}\left(\sum_{i=1}^{n} e_i^2\right) = 0 \quad .$$

Berdasarkan persamaan (2.4) diperoleh:

$$\frac{\partial}{\partial a}\left(\sum_{i=1}^{n} e_i^2\right) = 0 ,$$
$$\frac{\partial}{\partial a}\sum_{i=1}^{n}\left(Y_i - (a + bX_i)\right)^2 = 0 ,$$
$$2\sum_{i=1}^{n}\left(Y_i - (a + bX_i)\right)^1 (-1) = 0 ,$$
$$\sum_{i=1}^{n}(a + bX_i) = \sum_{i=1}^{n} Y_i ,$$
$$na + b\sum X_i = \sum Y_i \quad \text{......................} \quad (2.5a)$$

dimana notasi $\sum$ mewakili $\sum_{i=1}^{n}$.
Berikutnya,

$$\frac{\partial}{\partial b}\left(\sum_{i=1}^{n} e_i^2\right) = 0 ,$$
$$\frac{\partial}{\partial b}\sum_{i=1}^{n}\left(Y_i - (a + bX_i)\right)^2 = 0 ,$$
$$2\sum_{i=1}^{n}\left(Y_i - (a + bX_i)\right)^1 (-X_i) = 0 ,$$
$$a\sum X_i + b\sum X_i^2 = \sum X_i Y_i \quad \text{..............(2.5b)}$$

Persamaan (2.5a) dan (2.5b) secara matriks dapat dituliskan menjadi:

$$\begin{pmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum Y_i \\ \sum X_i Y_i \end{pmatrix} \text{..........(2.6)}$$

Dengan menggunakan aturan Crame'r (Boas, M.L., 1983), melalui persamaan (2.6) diperoleh nilai a dan b sebagai berikut:

$$a = \frac{\begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}} \quad \text{dan} \quad b = \frac{\begin{vmatrix} n & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}}$$

Setelah diselesaikan, diperoleh:

$$a = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n\sum X_i^2 - (\sum X_i)^2} \text{, dan}$$
$$b = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2} \text{ . ... (2.7)}$$

Setelah nilai a dan b didapatkan, maka diperoleh persamaan garis regresi, yaitu persamaan yang digunakan untuk memprediksi nilai *Y* jika nilai *X* diketahui, seperti diungkapkan pada persamaan (2.3).

Nilai konstanta a dapat pula dihitung dengan cara lain seperti berikut ini: jika ruas kiri dan ruas kanan persamaan (2.5a) dibagi dengan *n*, diperoleh: $= \frac{\sum Y_i}{n} - b\frac{\sum X_i}{n}$. Oleh karena $\frac{\sum Y_i}{n} = \bar{Y}$ dan $\frac{\sum X_i}{n} = \bar{X}$, yaitu merupakan nilai rata-rata untuk masing-masing-masing Y dan X, maka diperoleh:

$$a = \bar{Y} - b\bar{X}, \quad \text{................................(2.8)}$$

namun dengan catatan nilai *b* harus dicari terlebih dahulu.

Jika didefinisikan $x_i = X_i - \bar{X}$ dan $y_i = Y_i - \bar{Y}$ dengan x dan y masing-masing merupakan deviasi nilai X dan Y terhadap nilai rata-ratanya, maka dapat ditunjukkan bahwa:

$$\sum x_i^2 = \sum(X_i - \bar{X})^2 = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$
$$\sum y_i^2 = \sum(Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}\text{, dan}$$
$$\sum x_i y_i = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} \quad \text{................(2.9)}$$

Dengan demikian maka formula untuk b pada persamaan (2.7) dalam ungkapan nilai deviasi mereduksi menjadi:

$b = \frac{\sum x_i y_i}{\sum x_i^2}$. ....(2.10).

Formula seperti pada persamaan (2.10) ini menjadi jauh lebih sederhana dan mudah untuk diingat, jika dibandingkan dengan (2.7). Bahkan kalau kita bekerja atau mengolah data dalam bentuk data deviasi, angka-angka yang terlibat menjadi jauh lebih kecil nilainya ketimbang dalam ungkapan data mentah (yang disimbolkan dengan huruf kapital). Akibatnya mahasiswa menjadi lebih mudah dalam mengolah datanya serta dapat memperkecil kemungkinan kesalahan perhitungan.

Jadi, dapatlah disimpulkan bahwa model persamaan baris regresi linier sederhana mempunyai bentuk: $\hat{Y}_i = a + bX_i$, dengan a dan b dapat dicari dengan rumus: $b = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$ dan $a = \bar{Y} - b\bar{X}$.

Analisis Korelasi Sederhana

Kalau pada analisis regresi ditujukan untuk mencari hubungan fungsional antar variabel, maka pada analisis korelasi mengkaji tingkat kekuatan hubungan antar variabel tersebut. Koefisien korelasi (r) merupakan akar dari koefisien determinasi. Koefisien dererminasi r^2 didefinisikan sebagai rasio *antara* selisih jumlah kuadrat deviasi nilai Y hasil observasi (Y_{obs}) terhadap reratanya dengan jumlah kuadrat deviasi Y_{obs} terhadap nilai Y hasil ekspektasi (Y_{exp}), *dan* jumlah kuadrat deviasi Y_{obs} terhadap reratanya. Secara matematis dituliskan sebagai berikut:

$r^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2 - \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$, ...(2.10)

(Spiegel. M. R. & $Larry\ J.S.$, 2002)

Persamaan ini dapat kita sederhanakan menjadi:

$r^2 = \frac{\sum_{i=1}^n y_i^2 - \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n y_i^2}$, atau

$r^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n y_i^2}$ .... (2.12)

Tampak dari (2.12) ini bahwa bila setiap $Y_i = \hat{Y}_i$ maka nilai $r^2 = 1$. Namun ini adalah kasus yang sangat ideal (hampir tidak mungkin) dimana hal ini terjadi bila hasil observasi selalu sama dengan ekspektasi.

Secara matematis, pembilang pada suku kedua ruas kanan persamaan (2.12) yaitu $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$, dapat dibuktikan bahwa untuk kasus regresi sederhana ini:

$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n y_i^2 - b \sum_{i=1}^n x_i y_i$,

sehingga didapat:

$r^2 = 1 - \frac{\sum_{i=1}^n y_i^2 - b \sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^2}$

$r^2 = 1 - 1 + \frac{b \sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^2}$

$r^2 = \frac{b \sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^2}$..................................(2.13)

Telah diketahui dari persamaan (2.10) bahwa $b = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$, maka dari persamaan (2.13) didapat:

$r^2 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \times \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^2}$

$r^2 = \frac{(\sum_{i=1}^n x_i y_i)^2}{\sum_{i=1}^n x_i^2 \times \sum_{i=1}^n y_i^2}$ atau

$r = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2 \times \sum_{i=1}^n y_i^2}}$. .... (2.14)

Persamaan (2.14) inilah yang dinamakan dengan koefisien korelasi, yang menyatakan derajat hubungan antara variabel X dan variabel Y.

Jika kita ingin menggunakan data skor mentah, dapat kita kembalikan ungkapan koefisien korelasi dalam skor deviasi ini ke dalam ungkapan dalam skor mentah dengan mengambil hubungan pada persamaan (2.9), sehingga didapat:

$r = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{[n \sum X_i^2 - (\sum X_i)^2][n \sum Y_i^2 - (\sum Y_i)^2]}}$......(2.15)

Analisis Regresi dan Korelasi Berganda

Analisis regresi linier berganda (dalam hal ini diambil 2 prediktor saja) merupakan analisis statistika untuk menyelidiki pengaruh variabel bebas X_1 dan X_2 secara secara sendiri sendiri maupun secara simultan (bersama-sama) terhadap variabel terikat (Y). Adapun bentuk persamaan regresi linier berganda dalam hal ini adalah:

$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2$, ...(3.1)

dengan b_0, b_1, dan b_2 merupakan koefisien-koefisien regresi.

Jumlah kuadrat dari e_i dengan $i = 1$ sampai n adalah :

$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n \left(Y_i - \hat{Y}_i\right)^2$

$= \sum_{i=1}^n (Y_i - (b_0 + b_1 X_{1i} + b_2 X_{2i}))^2$

....(3.2)

Sebagaimana analisis regresi linier sederhana, metode kuadrat terkecil adalah metode untuk menghitung nilai b_0, b_1, dan b_2 sedemikian rupa sehingga $\sum_{i=1}^{n} e_i^2$ memiliki nilai terkecil (minimum). Dengan demikian, turunan pertama dari $\sum_{i=1}^{n} e_i^2$ terhadap b_0, b_1, dan b_2 sama dengan nol, atau

$$\frac{\partial}{\partial b_0}\left(\sum_{i=1}^{n} e_i^2\right) = 0$$
$$\frac{\partial}{\partial b_1}\left(\sum_{i=1}^{n} e_i^2\right) = 0 \quad \text{............................(3.3)}$$
$$\frac{\partial}{\partial b_2}\left(\sum_{i=1}^{n} e_i^2\right) = 0$$

Dengan demikian, berdasarkan persamaan (3.2), dengan metode kuadrat terkecil ini diperoleh tiga buah persamaan sebagai berikut:

$$nb_0 + b_1 \sum X_1 + b_2 \sum X_2 = \sum Y$$
$$b_0 \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 = \sum X_1 Y$$
$$b_0 \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 = \sum X_2 Y$$
$$\text{..(3.4)}$$

Jika $x_i = X_i - \bar{X}$ dan $y_i = Y_i - \bar{Y}$, maka persamaan (3.4) akan mereduksi menjadi dua persamaan:

$$\left.\begin{aligned} \Sigma x_1 y &= b_1 \Sigma x_1^{\,2} + b_2 \Sigma x_1 x_2 \\ \Sigma x_2 y &= b_1 \Sigma x_1 x_2 + b_2 \Sigma x_2^{\,2} \end{aligned}\right\} \quad \text{............(3.5)}$$

Secara matriks dapat dituliskan sebagai berikut:

$$\begin{pmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_1 x_2 & \Sigma x_2^2 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \Sigma x_1 y \\ \Sigma x_2 y \end{pmatrix} \text{.....(3.6)}$$

Dengan menggunakan aturan Cramers (Boas, M.L., 1983) diperoleh nilai b_1 dan b_2, yaitu:

$$b_1 = \frac{\begin{vmatrix} \Sigma x_1 y & \Sigma x_1 x_2 \\ \Sigma x_2 y & \Sigma x_2^2 \end{vmatrix}}{\begin{vmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_1 x_2 & \Sigma x_2^2 \end{vmatrix}}; \quad b_2 = \frac{\begin{vmatrix} \Sigma x_1^2 & \Sigma x_1 y \\ \Sigma x_1 x_2 & \Sigma x_2 y \end{vmatrix}}{\begin{vmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_1 x_2 & \Sigma x_2^2 \end{vmatrix}}$$

Setelah diselesaikan, diperoleh:

$$b_1 = \frac{(\Sigma x_1 y).(\Sigma x_2^2) - (\Sigma x_2 y)(\Sigma x_1 x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2} \quad \text{.....(3.7a)}$$

$$b_2 = \frac{(\Sigma x_2 y).(\Sigma x_1^2) - (\Sigma x_1 y)(\Sigma x_1 x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2} \quad \text{....(3.7b)}$$

Sedangkan koefisien b_0, dapat dihitung melalui persamaan

$$b_0 = \bar{Y} - b_1 \bar{X}_1 - b_2 \bar{X}_2 . \quad \text{......................(3.8)}$$

Persamaan (3.8) ini diperoleh dari persamaan (3.4) baris pertama dengan cara membagi ruas kiri dan kanan dengan *n*.

Setelah nilai b_0, b_1, dan b_2 didapatkan kemudian disubstitusi ke persamaan (3.1), maka diperoleh persamaan garis regresi, yaitu persamaan yang digunakan untuk memprediksi nilai *Y* jika nilai X_1 dan X_2 diketahui.

Selanjutnya, untuk melihat besarnya pengaruh variabel bebas (secara sendiri-sendiri) terhadap variabel terikat, dapat diadopsi dari persamaan (2.14), yang secara umum dapat dituliskan sebagai berikut:

$$r_{iy}^2 = \frac{(\Sigma x_i y)^2}{\Sigma x_i^2 \cdot \Sigma y^2}, \text{ atau}$$

$$r_{iy} = \frac{\Sigma x_i y}{\sqrt{\Sigma x_i^2 \times \Sigma y^2}} \quad , \quad \text{..........................(3.9)}$$

dengan $i = 1$, dan 2.

Adapun untuk melihat besarnya pengaruh variabel bebas (secara bersama-sama) terhadap variabel terikat, dapat dimodifikasi dari persamaan (2.13), sehingga diperoleh:

$$R^2 = \frac{b_1 \Sigma x_1 y + b_2 \Sigma x_2 y}{\Sigma y^2} \text{ ....(3.10)}$$

Menguji Signifikansi Persamaan Garis Regresi

Setelah memperoleh persamaan garis regresi Y atas X (dalam regresi linier sederhana) ataupun regresi Y atas X_1 dan X_2 (dalam regresi berganda), langkah selanjutnya adalah melakukan uji signifikansi persamaan garis regresi, yaitu menguji apakah persamaan garis regresi yang diperoleh dapat dipakai untuk memprediksi nilai Y jika X diketahui. Untuk keperluan ini digunakan uji F, dengan kriteria, *persamaan garis regresi ganda dikatakan signifikan jika $F_{hitung} \geq F_{tabel}$*. Adapun untuk mencari nilai F_{hitung}, digunakan rumus (Sudjana, 2002):

$$F_{hit} = \frac{JK_{reg}/k}{JK_{res}/(n-k-1)} \text{ ....(4.1)}$$

dengan *k* adalah jumlah prediktor, *n* adalah jumlah sampel penelitian, JK_{reg} adalah jumlah kuadrat-kuadrat regresi, JK_{res} adalah jumlah kuadrat-kuadrat residu, yang masing-masing dicari dari persamaan berikut:

$$JK_{reg} = b\Sigma xy \text{ ....(4.2a)}$$

(untuk regresi sederhana) , dan

$$JK_{reg} = b_1 \Sigma x_1 y + b_2 \Sigma x_2 y \text{ , ...(4.2b)}$$

(untuk regresi berganda). Sedangkan

$$JK_{res} = \Sigma(Y - \hat{Y})^2 = \Sigma y^2 - JK_{reg} \text{ ....(4.3)}$$

Terkhusus untuk regresi linier berganda, setelah mengetahui signifikansi dari persamaan garis regresi yang diperoleh, langkah terakhir adalah menguji apakah kontribusi dari masing-masing variabel bebas terhadap variabel terikat cukup signifikan. Untuk keperluan ini dilakukan uji signifikansi koefisien-koefisien regresi (b_1 dan b_2), yaitu menggunakan uji *t* dengan rumus:

$$t_i = \frac{b_i}{S_{b_i}} \text{ ....(4.4)}$$

dengan i = 1, dan 2. Adapun kriterianya adalah: *kontribusi variabel bebas X_i dikatakan signifikan terhadap variabel terikat Y jika* $t_{hit} \geq t_{(n-k-1);(1-\frac{1}{2}\alpha)}$.

Nilai S_{b_i} dalam persamaan (4.4) dihitung dengan rumus:

$$S_{b_i} = \sqrt{\frac{S_{y.12}^2}{\Sigma x_i^2(1-r_{iy}^2)}} \text{ , ...(4.5)}$$

dengan i = 1 dan 2, dimana $S_{y.12}^2$ dirumuskan dengan:

$$S_{y.12}^2 = \frac{\Sigma(Y-\widehat{Y})^2}{n-k-1} = \frac{\Sigma y^2 - JK_{reg}}{n-k-1} \text{ , ... (4.6)}$$

dan r_{iy}^2 seperti pada persamaan (3.9).

Berdasarkan persamaan (3.10) dan (4.2b), didapat $JK_{reg} = (\Sigma y^2)R^2$. Sedangkan dari persamaan (4.3),

$$JK_{res} = \Sigma y^2 - JK_{reg}$$
$$= \Sigma y^2 - (\Sigma y^2)R^2$$
$$= \Sigma y^2(1-R^2) \text{ .}$$

Dengan demikian, seperti dalam beberapa literatur, nilai F hitung pada persamaan (4.1) dapat pula dirumuskan dengan:

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} \text{ ....(4.7)}$$

Model Regresi Nonlinier (Kuadratik)

Pada saat melakukan prediksi tentang model hubungan antara variabel bebas dan terikat menggunakan diagram pencar, seringkali orang mengambil bentuk linier. Padahal sebetulnya banyak sekali masalah fisika yang memiliki hubungan nonlinier. Beberapa model regresi nonlinier antara lain (Sudjana, 2002):

1. Model kuadratik
2. Model kubik
3. Model logaritmik
4. Model eksponensial
5. Dan sebagainya

Dalam tulisan ini misalnya diambil salah satu model nonlinier, yaitu model kuadratik. Salah satu contoh hubungan kuadratik misalnya gerak benda yang dilempar dengan sudut elevasi tertentu dan kecepatan tertentu pula. Tentu saja hal ini akan menjadi keliru kalau kita mengambil perkiraan model linier mengingat tidak sesuai dengan faktanya. Itulah sebabnya pada saat kita memplot model hubungan menggunakan diagram pencar, haruslah benar-benar dilakukan secara teliti dan akurat agar hasilnya seperti yang diharapkan.

Misalkan saja persamaan model perkiraan adalah: $\hat{Y} = a_0 + a_1X + a_2X^2$. Perkiraan nilai Y (yaitu $\hat{Y}_i$) untuk setiap data X_i adalah

$$\hat{Y}_i = a_0 + a_1X_i + a_2X_i^2 \text{. ....(5.1)}$$

Jika terdapat *n* buah data, maka diperoleh jumlah kuadrat selisih:

$$\Phi = \Sigma\left(Y_i - \hat{Y}_i\right)^2$$
$$\Phi = \Sigma\left(Y_i - (a_0 + a_1X_i + a_2X_i^2)\right)^2 \text{.....(5.2)}$$

Jumlah kuadrat selisih ini harus minimum terhadap nilai $a_0, a_1,$ dan a_2, yang kita kenal dengan metode kuadrat terkecil. Jadi ini adalah permasalahan nilai minimum, sehingga $\frac{\partial\Phi}{\partial a_0} = 0$, $\frac{\partial\Phi}{\partial a_1} = 0$, dan $\frac{\partial\Phi}{\partial a_2} = 0$. Dengan demikian jika persamaan (5.2) diturunkan terhadap $a_0, a_1,$ dan a_2 dan kemudian hasilnya dipersamakan dengan nol, maka diperoleh ada persamaan sebagai berikut:

$$na_0 + a_1\Sigma X_i + a_2\Sigma X_i^2 = \Sigma Y_i$$
$$a_0\Sigma X_i + a_1\Sigma X_i^2 + a_2\Sigma X_i^3 = \Sigma X_iY_i$$
$$a_0\Sigma X_i^2 + a_1\Sigma X_i^3 + a_1\Sigma X_i^4 = \Sigma X_i^2Y_i \text{ . ...(5.3)}$$

Ketiga persamaan ini dapat dituliskan secara matriks seperti berikut:

$$\begin{bmatrix} n & \Sigma X_i & \Sigma X_i^2 \\ \Sigma X_i & \Sigma X_i^2 & \Sigma X_i^3 \\ \Sigma X_i^2 & \Sigma X_i^3 & \Sigma X_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_iY_i \\ \Sigma X_i^2Y_i \end{bmatrix} \text{...(5.4)}$$

Nilai-nilai pada matriks (3x3) di ruas kiri dan pada matriks (3x1) di ruas kanan diperoleh dari data hasil pengamatan.

Tujuan dari persamaan matriks di atas sebenarnya adalah untuk mencari nilai-nilai parameter-parameter regresi $a_0, a_1,$ dan a_2. Dengan menggunakan aturan Cramer's, dapat diperoleh sebagai berikut:

$$a_0 = \frac{\begin{vmatrix} \Sigma Y_i & \Sigma X_i & \Sigma X_i^2 \\ \Sigma X_i Y_i & \Sigma X_i^2 & \Sigma X_i^3 \\ \Sigma X_i^2 Y_i & \Sigma X_i^3 & \Sigma X_i^4 \end{vmatrix}}{\begin{vmatrix} n & \Sigma X_i & \Sigma X_i^2 \\ \Sigma X_i & \Sigma X_i^2 & \Sigma X_i^3 \\ \Sigma X_i^2 & \Sigma X_i^3 & \Sigma X_i^4 \end{vmatrix}} \text{.................}(5.5a)$$

$$a_1 = \frac{\begin{vmatrix} n & \Sigma Y_i & \Sigma X_i^2 \\ \Sigma X_i & \Sigma X_i Y_i & \Sigma X_i^3 \\ \Sigma X_i^2 & \Sigma X_i^2 Y_i & \Sigma X_i^4 \end{vmatrix}}{\begin{vmatrix} n & \Sigma X_i & \Sigma X_i^2 \\ \Sigma X_i & \Sigma X_i^2 & \Sigma X_i^3 \\ \Sigma X_i^2 & \Sigma X_i^3 & \Sigma X_i^4 \end{vmatrix}} \text{.................}(5.5b)$$

dan

$$a_2 = \frac{\begin{vmatrix} n & \Sigma X_i & \Sigma Y_i \\ \Sigma X_i & \Sigma X_i^2 & \Sigma X_i Y_i \\ \Sigma X_i^2 & \Sigma X_i^3 & \Sigma X_i^2 Y_i \end{vmatrix}}{\begin{vmatrix} n & \Sigma X_i & \Sigma X_i^2 \\ \Sigma X_i & \Sigma X_i^2 & \Sigma X_i^3 \\ \Sigma X_i^2 & \Sigma X_i^3 & \Sigma X_i^4 \end{vmatrix}} \text{..................}(5.5c)$$

Parameter-parameter regresi $a_0, a_1,$ dan a_2 dapat juga dicari dengan metode inversi seperti berikut ini:

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} n & \Sigma X_i & \Sigma X_i^2 \\ \Sigma X_i & \Sigma X_i^2 & \Sigma X_i^3 \\ \Sigma X_i^2 & \Sigma X_i^3 & \Sigma X_i^4 \end{pmatrix}^{-1} \begin{pmatrix} \Sigma Y_i \\ \Sigma X_i Y_i \\ \Sigma X_i^2 Y_i \end{pmatrix}$$

....(5.6)

Setelah $a_0, a_1,$ dan a_2 dicari, maka selanjutnya disubstitusi ke persamaan (5.1), sehingga nantinya akan diperoleh persamaan garis regresi kuadratik.

Adapun koefisien korelasinya dapat dihitung dengan formula umum seperti pada persamaan (2.12), yaitu:

$r^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} y_i^2}$, dengan

$\hat{Y}_i = a_0 + a_1 X_i + a_2 X_i^2$.

Kesimpulan

Berdasarkan uraian di atas dapatlah disimpulkan bahwa, dengan banbtuan metode kuadrat terkecil dan aturan Cramers serta pendefinisian variabel deviasi $x_i = X_i - \bar{X}$ dan $y_i = Y_i - \bar{Y}$, diperoleh beberapa formula sederhana terkait Analisis Regresi dan Korelasi, sebagai berikut:

1. Untuk regresi linier sederhana dengan persamaan : $\hat{Y}_i = a + bX_i$, diperoleh $b = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$ dan $a = \bar{Y} - b\bar{X}$ serta koefisien korelasi dapat dicari dari $r^2 = \frac{b\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^2}$.
2. Untuk regresi berganda dua prediktor dengan persamaan $\hat{Y}_i = b_0 + b_1 X_{1i} + b_2 X_{2i}$, diperoleh formula:

   $$b_1 = \frac{(\Sigma x_1 y).(\Sigma x_2^2) - (\Sigma x_2 y)(\Sigma x_1 x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$$

   $$b_2 = \frac{(\Sigma x_2 y).(\Sigma x_1^2) - (\Sigma x_1 y)(\Sigma x_1 x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$$

   $$b_0 = \overline{Y} - b_1 \overline{X}_1 - b_2 \overline{X}_2 .$$

   Sedangkan koefisien korelasi (biasa) $r_{iy} = \frac{\Sigma x_i y}{\sqrt{\Sigma x_i^2 \times \Sigma y^2}}$, dengan $i = 1$dan 2, dan koefisien korelasi berganda dapat diperoleh dari: $R^2 = \frac{b_1 \Sigma x_1 y + b_2 \Sigma x_2 y}{\Sigma y^2}$.
3. Analisis regresi nonlinier kuadratik dengan persamaan $\hat{Y}_i = a_0 + a_1 X_i + a_2 X_i^2$, koefisien-koefisien regresinya dapat dengan mudah dihitung dengan bantuan metode kuadrat terkecil dan aturan Cramers (metode determinasi). Cara lain juga dapat digunakan untuk keperluan ini, yaitu metode inversi, yang nantinya akan mendapatkan nilai yang sama. Adapun koefisien korelasinya dapat ditentukan dari $r^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} y_i^2}$.

Daftar Pustaka

Adisoewignyo,W. 2010. *Statistika Inferensi*. Mataram: Mataram University Press.

Arfken, G. 1970. *Mathematical Methods for Fhysicist.* London: Academic Press, Inc.

Boas, M.L. 1983. *Mathematical Methods in the Fhysycal Science*. Second Edition. New York: John Wiley & Sons, Inc.

Rosdiana, dkk. 2004. *Matematika Fisika I,II.* Common Textbook (Edisi Revisi). Bandung: Jurusan Pendidikan Fisika, FPMIPA UPI.

Ruseffendi. 1998. *Statistika Dasar*. Bandung: IKIP Bandung.

Spiegel, M.R. & Larry, J.S. 2004. *Statistik*, Edisi Ketiga (Terjemahan oleh Wiwit Kastawan dan Irzam Harmein). Jakarta: Erlangga.

Sudjana. 2002. *Metode Statistika*. Bandung: Tarsito.

Supranto, J. 2001. *Statistika, Teori dan Aplikasi.* Jakarta: Erlangga.

Wospakrik, H.J. 1993. *Dasar-dasar Matematika untuk Fisika.* Bandung: Jurusan Fisika ITB.