Analisis Regresi

#14 Meeting

Regresi Polinomial

Ferdian Bangkit Wijaya, S.Stat., M.Si
NIP. 199005202024061001

Regresi Polinomial

Dalam analisis regresi, seringkali hubungan antara variabel independen (X) dan variabel dependen (Y) tidak bersifat linier (garis lurus), melainkan membentuk pola kurva (lengkungan). Untuk menangani pola data seperti ini, kita menggunakan Regresi Polinomial.

Bentuk Umum Model
Model regresi polinomial derajat ke-k dengan satu prediktor didefinisikan sebagai:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 + \cdots + \beta_k X_i^k + \varepsilon_i$$

Dimana:

- β_0: Intersep.
- β_1: Efek linier.
- β_2: Efek kuadratik (kelengkungan).
- β_3: Efek kubik (lengkungan ganda/bentuk S).

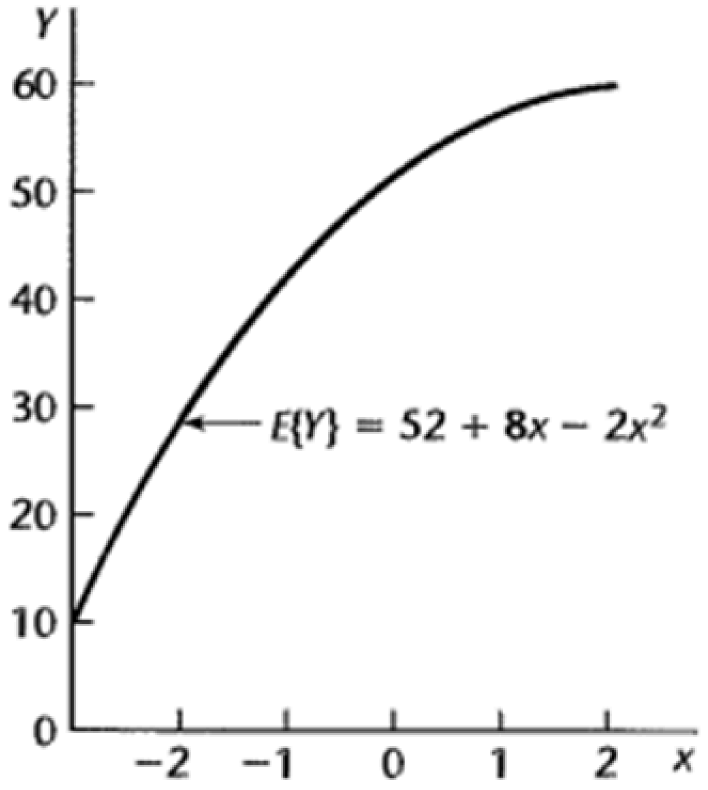
Regresi Polinomial

Model Derajat kedua dengan 1 Prediktor

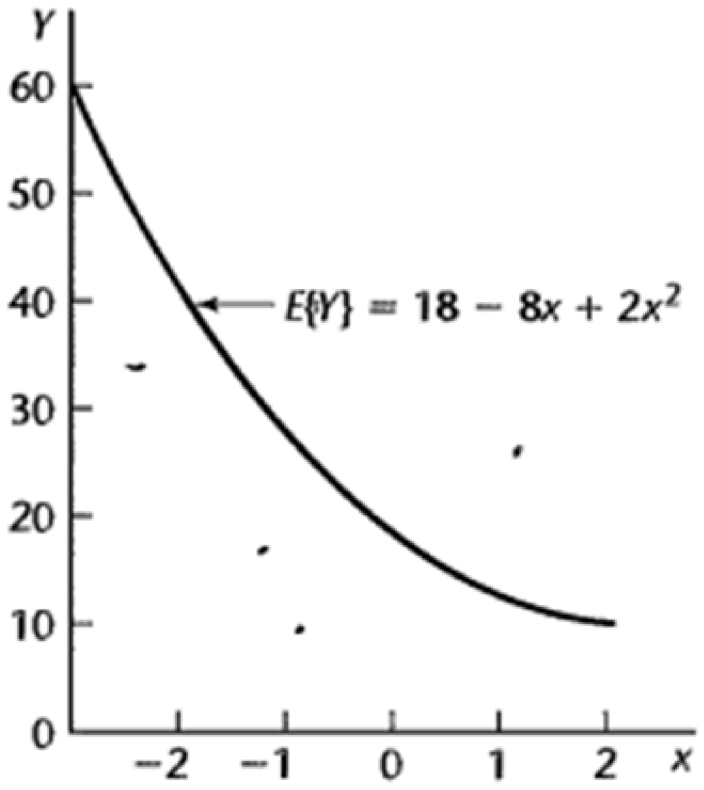
$$Y = \beta_0 + \beta_1 X + \beta_{11} X^2 + \varepsilon$$

Model Derajat kedua dengan 2 Prediktor

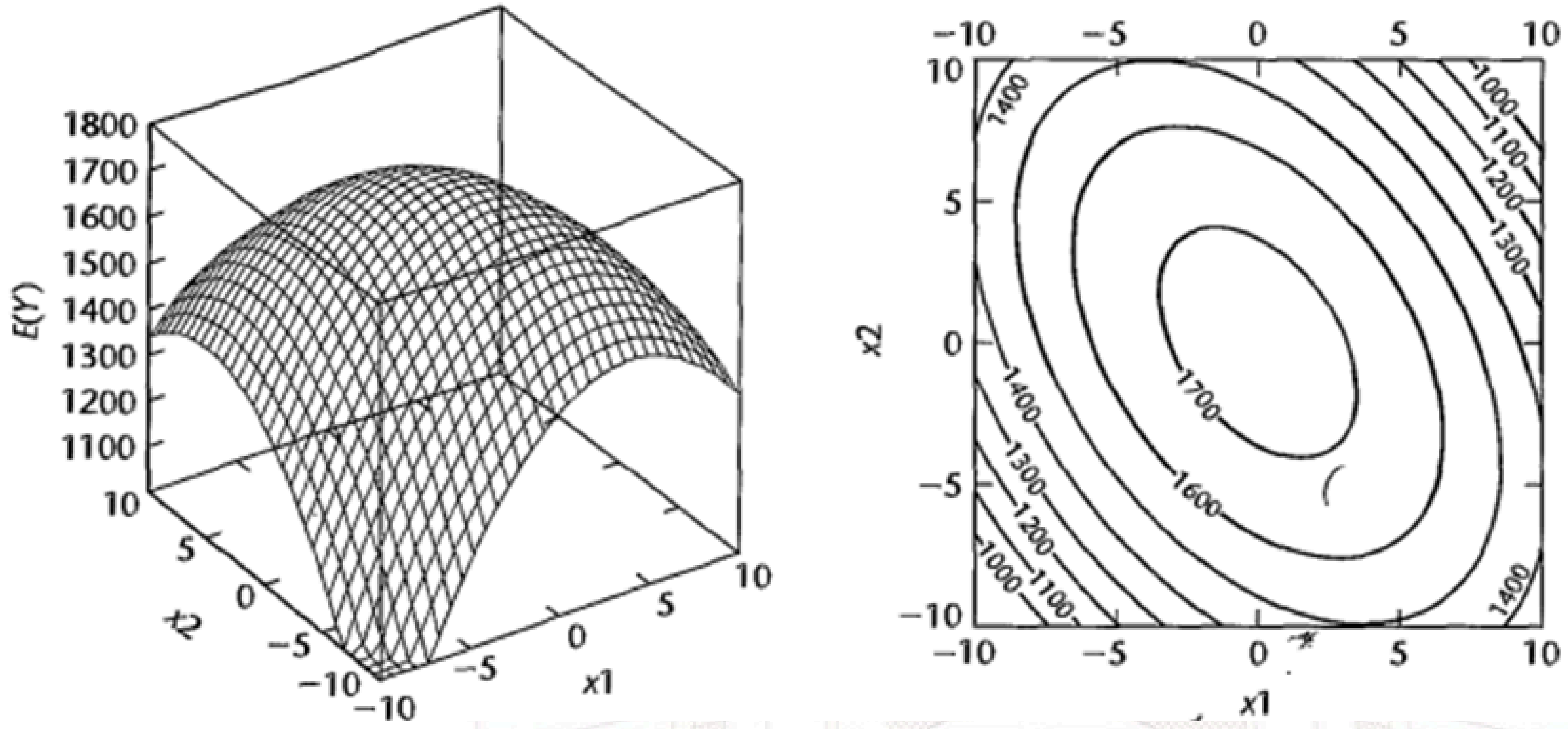
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} {X_1}^2 + \beta_{22} {X_2}^2 + \beta_{12} X_1 x_2 + \varepsilon$$



(a)



(b)



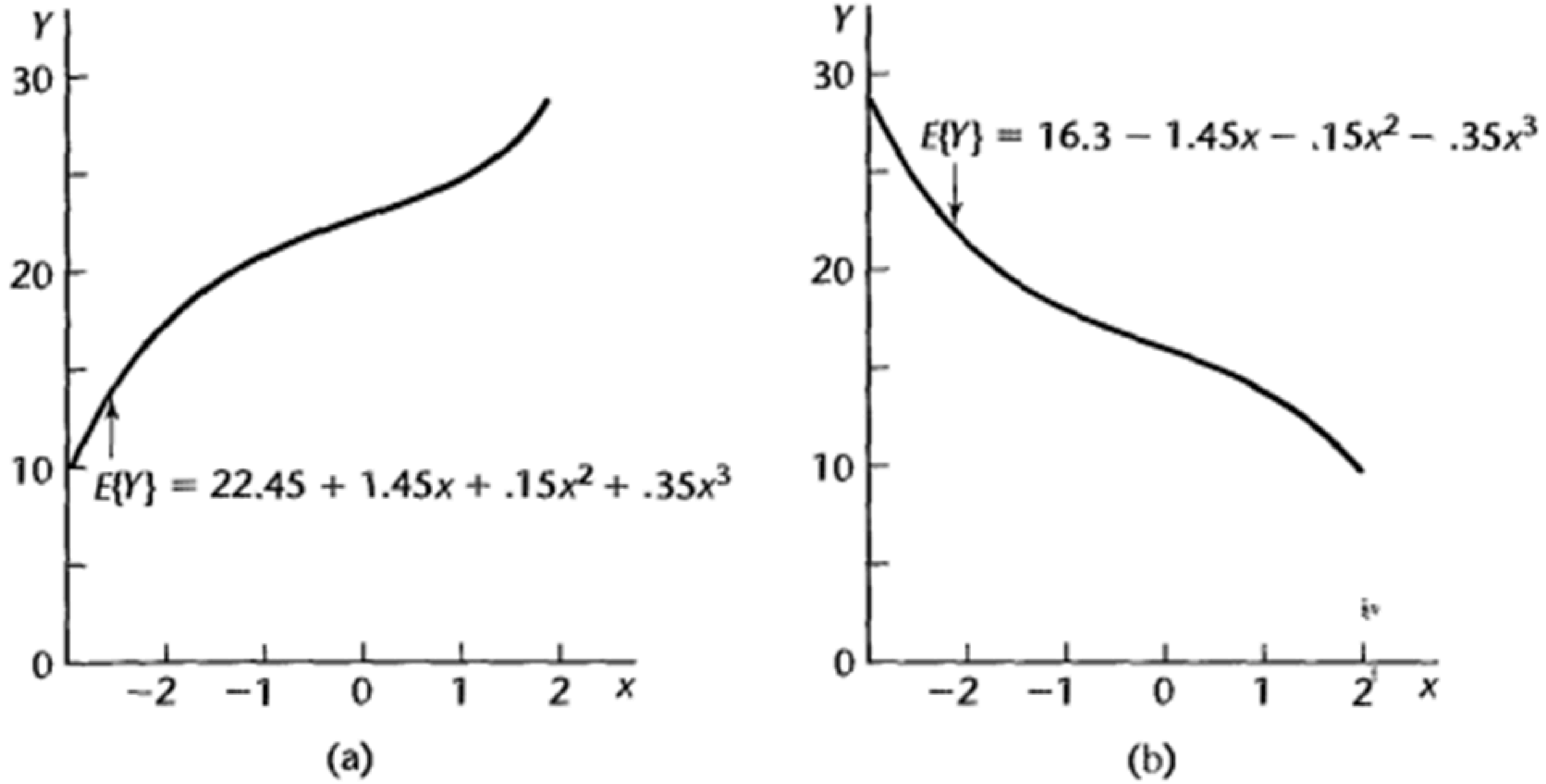
Regresi Polinomial

Model Derajat kedua dengan 3 Prediktor : 4 Dimensi

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} {X_1}^2 + \beta_{22} {X_2}^2 + \beta_{33} {X_3}^2 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + \varepsilon$$

Model Derajat ketiga dengan 1 Prediktor

$$\mathrm{Y} = \beta_0 + \beta_1 X + \beta_{11} X^2 + \beta_{111} X^3 + \varepsilon$$



Regresi Polinomial

Pendugaan Parameter

$$\boldsymbol{b} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$

$$Y = b_0 + b_1 X + b_{11} X^2 + \mathrm{e}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & {X_{11}}^2 \\ 1 & X_{12} & {X_{12}}^2 \\ \vdots & \vdots & \vdots \\ 1 & X_{1n} & {X_{1n}}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_{11} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Kasus :

Buatlah model regresi polinomial derajat 2 dari data di samping ini dan periksa asumsi nonmultikolinieritasnya.

Y	X
22	20
5	18
15	13
20	11
23	28
25	23
20	23
11	28
3	24
23	19
9	27
8	21
9	24
23	19
16	16

Regresi Polinomial

Pertimbangan Memilih Model

Dalam menentukan derajat polinomial, peneliti harus mempertimbangkan:

1. Derajat Model: Jangan memilih derajat terlalu tinggi (overfitting). Biasanya derajat 2 atau 3 sudah cukup.
2. Seleksi Peubah: Pastikan penambahan pangkat variabel memang signifikan secara statistik.
3. Ekstrapolasi: Hindari memprediksi nilai di luar rentang data, karena ekor kurva polinomial bisa bergerak drastis ke atas/bawah.
4. Hirarki Model: Jika X^2 masuk dalam model, maka X (pangkat 1) juga harus ada dalam model, terlepas signifikan atau tidak, untuk menjaga invarian terhadap transformasi linier.

Regresi Polinomial

Permasalahan Multikolinearitas :

Masalah pada Matriks (X'X)

Ketika kita memangkatkan X (menjadi X^2, X^3, ...), nilai-nilai tersebut akan berkorelasi sangat tinggi dengan nilai X aslinya.

- Contoh: Jika X = {1, 2, 3, ... , 10}, maka korelasi antara X dan X^2 akan mendekati 1.

Dampak:

1. Terjadi Multikolinearitas yang tinggi antar prediktor.
2. Matriks (X'X) menjadi ill-conditioned (hampir singular).
3. Invers matriks inv(X'X) menjadi tidak stabil secara numerik.
4. Varians penduga parameter menjadi sangat besar (estimasi tidak akurat).

Solusi: Model Polinomial Ortogonal (PO)

Untuk mengatasi masalah multikolinearitas tersebut, kita mentransformasi variabel prediktor X menjadi serangkaian polinomial yang saling bebas (Ortogonal).

Regresi Polinomial Ortogonal

Model Polinomial Ortogonal

$$Y = \alpha_0 P_0(X) + \alpha_1 P_1(X) + \alpha_2 P_2(X) + \cdots + \alpha_k P_k(X) + \varepsilon$$

P_u (X) adalah polinomial berderajat ke-u,
dimana u=0,1,2,…,k dan P_0 (X)=1

Dimana:

- α_j: Parameter model ortogonal (pengganti β).
- $P_j(X_i)$: Fungsi polinomial ortogonal derajat ke-j.
- Syarat Ortogonalitas: $\sum_{i=1}^{n} P_r(X_i)P_s(X_i) = 0$ untuk $r \neq s$.

Metode Pembentukan Polinomial ($P_u(X)$)

Polinomial ini dibentuk menggunakan:

1. **Deviasi Nilai ($X_i - \bar{X}$):** Untuk menyederhanakan perhitungan dan mengurangi korelasi awal.
2. **Proses Gram-Schmidt:** Algoritma aljabar linier untuk membentuk basis vektor yang ortogonal satu sama lain.
   - $P_0(X) = 1$
   - $P_1(X) = \lambda_1(X - \bar{X})$
   - $P_2(X) = \lambda_2[(X - \bar{X})^2 - \frac{\sum(X-\bar{X})^2}{n}]$ *(Nilai λ adalah konstanta pembobot)*

Regresi Polinomial Ortogonal

Pendugaan Parameter

$$Y = \alpha_0 P_0(X) + \alpha_1 P_1(X) + \alpha_2 P_2(X) + \cdots + \alpha_k P_k(X) + \varepsilon$$

$$X = \begin{bmatrix} P_0(x_1) & P_1(x_1) & \cdots & P_k(x_1) \\ P_0(x_2) & P_1(x_2) & \cdots & P_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ P_0(x_n) & P_1(x_n) & \cdots & P_k(x_n) \end{bmatrix}$$

Unsur selain diagonal utama bernilai 0 (nol)

$$\hat{\boldsymbol{\alpha}} = \begin{bmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_k \end{bmatrix} = (X^T X)^{-1} X^T Y$$

$$(X^T X) = \begin{bmatrix} \sum P_0^2 & 0 & \cdots & 0 \\ 0 & \sum P_1^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum P_k^2 \end{bmatrix}$$

$$\hat{\alpha}_j = \frac{\sum_{i=1}^{n} P_j(x_i) y_i}{\sum_{i=1}^{n} {P_j}^2(x_i)} \text{, j} = 0,1,2,\ldots,\text{k}$$

Regresi Polinomial Ortogonal

$$Y = \alpha_0 P_0(X) + \alpha_1 P_1(X) + \alpha_2 P_2(X) + \cdots + \alpha_k P_k(X) + \varepsilon$$

Penguraian Keragaman

Sumber Keragaman	Derajat Bebas	Jumlah kuadrat
$\hat{\alpha}_0$	1	$\hat{\alpha}_j \sum_{i=1}^{n} P_j(x_i) y_i$
$\hat{\alpha}_1$	1	
$\hat{\alpha}_2$	1	
$\vdots$	$\vdots$	
$\hat{\alpha}_k$	1	
Sisaan	$n-k-1$	$JKT - \sum_{j=1}^{k} \hat{\alpha}_j \sum_{i=1}^{n} P_j(x_i) y_i$
Total	n	$\sum_{i=1}^{n} (Y_i - \bar{Y})^2$

Uji Hipotesis Parsial

- **Hipotesis**

$$H_0: \alpha_j = 0$$
$$H_1: \alpha_j \neq 0$$

- **Statistik Uji**

$$F_{hitung} = \frac{\hat{\alpha}_j \sum_{i=1}^{n} P_j(x_i) y_i}{KTS}$$

- **Kriteria Penolakan H_0**

Tolak H_0 jika $F_{hitung} > F_{(\alpha;1;n-k-1)}$

Regresi Polinomial Ortogonal

Y	X
22	20
5	18
15	13
20	11
23	28
25	23
20	23
11	28
3	24
23	19
9	27
8	21
9	24
23	19
16	16

Rumus Umum (General Formula)

Untuk mencari polinomial derajat ke-k (P_k), rumusnya adalah:

$$P_k = P_1^k - \left(\frac{\sum(P_1^k \cdot P_{k-1})}{\sum P_{k-1}^2}\right) P_{k-1} - \left(\frac{\sum(P_1^k \cdot P_{k-2})}{\sum P_{k-2}^2}\right) P_{k-2} - \cdots - \left(\frac{\sum(P_1^k \cdot P_0)}{\sum P_0^2}\right) P_0$$

- P_1^k: Adalah nilai $(X - \bar{X})$ yang dipangkatkan dengan k.
- P_0 **(Polinomial Derajat 0)**:
  - Adalah vektor yang berisi angka **1**.
  - Nilainya selalu $P_0(x_i) = 1$ untuk setiap baris data.
  - Ini merepresentasikan komponen **Intersep** atau Rata-rata.
  - Implikasinya: $\sum P_0^2 = n$ (Jumlah data).
- $P_{k-1}, \ldots$: Polinomial ortogonal derajat sebelumnya yang sudah ditemukan.

Regresi Polinomial Ortogonal

$$P_k = P_1^k - \left(\frac{\sum(P_1^k \cdot P_{k-1})}{\sum P_{k-1}^2}\right) P_{k-1} - \left(\frac{\sum(P_1^k \cdot P_{k-2})}{\sum P_{k-2}^2}\right) P_{k-2} - \cdots - \left(\frac{\sum(P_1^k \cdot P_0)}{\sum P_0^2}\right) P_0$$

$$P_0(X) = 1$$

$$P_1(X) = (X - \bar{X})$$

$$P_2 = P_1^2 - \underbrace{\left(\frac{\sum P_1^3}{\sum P_1^2}\right) P_1} - \underbrace{\left(\frac{\sum(P_1^2 \cdot P_0)}{\sum P_0^2}\right) P_0}$$

$$P_2(X) = \underbrace{(X - \bar{X})^2}_{\text{Bentuk Dasar}} - \underbrace{\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^3}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right](X - \bar{X})}_{\text{Faktor Koreksi Kemiringan}} - \underbrace{\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}\right]}_{\text{Faktor Koreksi Rata-rata}}$$

Regresi Polinomial Ortogonal

No	Y	X	P1 (X−X̄)	P1^2	P1^3	P2 (P1^2+1.72P1 −24.46)	P2^2	P1 · Y	P2 · Y
1	22	20	-0.93	0.87	-0.81	-25.19	634.53	-20.53	-554.21
2	5	18	-2.93	8.6	-25.21	-20.9	436.81	-14.67	-104.51
3	15	13	-7.93	62.94	-499.37	24.84	617.02	-119	372.53
4	20	11	-9.93	98.67	-979.8	57.12	3262.69	-198.67	1142.36
5	23	28	7.07	49.94	352.92	37.64	1416.76	162.53	865.65
6	25	23	2.07	4.27	8.83	-16.63	276.55	51.67	-415.77
7	20	23	2.07	4.27	8.83	-16.63	276.55	41.33	-332.62
8	11	28	7.07	49.94	352.92	37.64	1416.76	77.73	414.01
9	3	24	3.07	9.4	28.87	-9.78	95.64	9.2	-29.35
10	23	19	-1.93	3.74	-7.23	-24.05	578.4	-44.47	-553.05
11	9	27	6.07	36.8	223.27	22.77	518.47	54.6	204.97
12	8	21	0.07	0	0	-24.34	592.43	0.53	-194.7
13	9	24	3.07	9.4	28.87	-9.78	95.64	27.6	-88.04
14	23	19	-1.93	3.74	-7.23	-24.05	578.4	-44.47	-553.05
15	16	16	-4.93	24.34	-119.98	-8.62	74.3	-78.93	-137.95
Total	232	314	0	366.93	-631.96	0	10871	-95.53	36.25

Regresi Polinomial Ortogonal

$$P_2 = P_1^2 - \underbrace{\left(\frac{\sum P_1^3}{\sum P_1^2}\right) P_1} - \underbrace{\left(\frac{\sum (P_1^2 \cdot P_0)}{\sum P_0^2}\right) P_0}$$

$\sum P_0^2 = n = \mathbf{15}$ (Karena $P_0 = 1$)

$\sum P_1^2 = \mathbf{366.93}$

$\sum P_1^3 = \mathbf{-631.96}$

$\sum (P_1^2 \cdot P_0) = \sum (P_1^2 \cdot 1) = \mathbf{366.93}$

$$\frac{\sum P_1^3}{\sum P_1^2} = \frac{-631.96}{366.93} = \mathbf{-1.72}$$

$$\frac{\sum (P_1^2 \cdot P_0)}{\sum P_0^2} = \frac{366.93}{15} = \mathbf{24.46}$$

$$\bar{X} = \frac{\sum X}{n} = \frac{314}{15} = 20,93333$$

$$P_2 = P_1^2 - (-1.72)P_1 - (24.46)(1)$$

$$P_2 = P_1^2 + 1.72P_1 - 24.46$$

$$\hat{\alpha}_j = \frac{\sum_{i=1}^{n} P_j(x_i) y_i}{\sum_{i=1}^{n} P_j^{\,2}(x_i)}, \text{j} = 0,1,2,\ldots,\text{k}$$

Parameter $\hat{\alpha}_0$ (Intersep/Rata-rata)

$$\hat{\alpha}_0 = \frac{\sum P_0 Y}{\sum P_0^2} = \frac{232}{15} = \mathbf{15.47}$$

Parameter $\hat{\alpha}_1$ (Efek Linier)

$$\hat{\alpha}_1 = \frac{\sum P_1 Y}{\sum P_1^2} = \frac{-95.53}{366.93} = \mathbf{-0.26}$$

Parameter $\hat{\alpha}_2$ (Efek Kuadratik)

$$\hat{\alpha}_2 = \frac{\sum P_2 Y}{\sum P_2^2} = \frac{36.25}{10.871} = \mathbf{0.0033}$$

Regresi Polinomial Ortogonal

Penyelesaian Pendugaan dengan Matrix

$$(X^T X) = \begin{bmatrix} \sum P_0^2 & 0 & \cdots & 0 \\ 0 & \sum P_1^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum P_k^2 \end{bmatrix}$$

$$\mathbf{X^T X} = \begin{bmatrix} 15 & 0 & 0 \\ 0 & 366,93 & 0 \\ 0 & 0 & 10.871 \end{bmatrix}$$

$$\mathbf{X^T Y} = \begin{bmatrix} \sum P_0 Y \\ \sum P_1 Y \\ \sum P_2 Y \end{bmatrix} = \begin{bmatrix} 232 \\ -95,53 \\ 36,25 \end{bmatrix}$$

$$(\mathbf{X^T X})^{-1} = \begin{bmatrix} \frac{1}{15} & 0 & 0 \\ 0 & \frac{1}{366,93} & 0 \\ 0 & 0 & \frac{1}{10.871} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0,0667} & 0 & 0 \\ 0 & \mathbf{0,0027} & 0 \\ 0 & 0 & \mathbf{0,000092} \end{bmatrix}$$

$$\hat{\alpha} = (\mathbf{X^T X})^{-1} \mathbf{X^T Y}$$

$$\begin{bmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{bmatrix} = \begin{bmatrix} 0,0667 & 0 & 0 \\ 0 & 0,0027 & 0 \\ 0 & 0 & 0,000092 \end{bmatrix} \times \begin{bmatrix} 232 \\ -95,53 \\ 36,25 \end{bmatrix}$$

$$\hat{\alpha} = \begin{bmatrix} 15,47 \\ -0,26 \\ 0,0033 \end{bmatrix}$$

Regresi Polinomial Ortogonal

Model Akhir

$$\hat{Y} = 15,47P_0 + (-0,26)P_1 + 0,0033P_2$$

$P_0 = 1$

$P_1 = (X - 20,93)$

$P_2 = (X - 20,93)^2 + 1,72(X - 20,93) - 24,46$

$$\hat{Y} = 22,16 - 0,39X + 0,0033X^2$$

Interpretasi Koefisien:

- $\beta_0 = 22,16$: Nilai dugaan Y ketika $X = 0$ (Intersep).
- $\beta_1 = -0,39$: Efek linier negatif (hubungan terbalik).
- $\beta_2 = 0,0033$: Efek kelengkungan positif (kurva sedikit cekung ke atas / *convex*), namun sangat lemah.

Regresi Polinomial Ortogonal

UNIVERSITAS SULTAN AGENG TIRTAYASA

UNTIRTA Jawara

Jujur Adil Wibawa Amanah Religius Akuntabel

ANOVA

Sumber Keragaman	Derajat Bebas	Jumlah kuadrat
$\hat{\alpha}_0$	1	$\hat{\alpha}_j \sum_{i=1}^{n} P_j(x_i) y_i$
$\hat{\alpha}_1$	1	
$\hat{\alpha}_2$	1	
$\vdots$	$\vdots$	
$\hat{\alpha}_k$	1	
Sisaan	$n-k-1$	$JKT - \sum_{j=1}^{k} \hat{\alpha}_j \sum_{i=1}^{n} P_j(x_i) y_i$
Total	n	$\sum_{i=1}^{n} (Y_i - \bar{Y})^2$

$$\textbf{JKT} = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = \sum Y^2 - n\bar{Y}^2$$

$$\sum Y^2 = 4.322$$

$$n\bar{Y}^2 = 3.588,27$$

$$\textbf{JKT} = 4.322 - 3.588,27 = \mathbf{733,73}$$

Jumlah Kuadrat Komponen ($\hat{\alpha}_j$)

$$\text{Jumlah Kuadrat } (\hat{\alpha}_j) = \hat{\alpha}_j \sum_{i=1}^{n} P_j(x_i) y_i$$

$$JK(\text{Linier}) = \hat{\alpha}_1 \times \sum P_1 Y = (-0,26) \times (-95,53) = \mathbf{24,84}$$

$$JK(\text{Kuadratik}) = \hat{\alpha}_2 \times \sum P_2 Y = (0,0033) \times (36,25) = \mathbf{0,12}$$

Regresi Polinomial Ortogonal

ANOVA

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat (JK)	Kuadrat Tengah (KT)	F-Hitung
$\hat{\alpha}_1$ **(Linier)**	1	**24,84**	24,84	$\frac{24{,}84}{59{,}06} = \mathbf{0,42}$
$\hat{\alpha}_2$ **(Kuadratik)**	1	**0,12**	0,12	$\frac{0{,}12}{59{,}06} = \mathbf{0,002}$
Sisaan	12	**708,77**	59,06	-
Total	15	**733,73**	-	-

Regresi Polinomial Ortogonal

UNIVERSITAS SULTAN AGENG TIRTAYASA

UNTIRTA Jawara
Jujur Adil Wibawa Amanah Religius Akuntabel

Uji Hipotesis Simultan

Hipotesis

$H_0 : \alpha_1 = \alpha_2 = 0$ (Model regresi tidak signifikan secara keseluruhan).

H_1 : Minimal ada satu $\alpha_j \neq 0$ (Model regresi signifikan).

$$JKR_{total} = JK(\hat{\alpha}_1) + JK(\hat{\alpha}_2)$$

$$JKR_{total} = 24,84 + 0,12 = \mathbf{24,96}$$

$$F_{hitung} = \frac{KTR}{KTS} = \frac{12,48}{59,06} = \mathbf{0,21}$$

Nilai Kritis (F-Tabel): Nilai F tabel untuk $df_1 = 2$ dan $df_2 = 12$ pada taraf nyata $\alpha = 5\%$ adalah **3,89**.

Kriteria Pengujian

Tolak H_0 jika $F_{hitung} > 3,89$.

Keputusan Statistik

Karena nilai $0,21 < 3,89$ **(Terima)** H_0

Kesimpulan Akhir Secara simultan, model regresi polinomial derajat 2 (gabungan linier dan kuadratik) **tidak berpengaruh signifikan** terhadap variabel respon Y.

Regresi Polinomial Ortogonal

Uji Hipotesis Parsial

A. Hipotesis

- $H_0 : \alpha_1 = 0$ (Komponen linier tidak signifikan).
- $H_1 : \alpha_1 \neq 0$ (Komponen linier signifikan).

B. Perhitungan Statistik Uji

$$F_{\text{hitung}} = \frac{JK(\hat{\alpha}_1)}{KTS} = \frac{24,84}{59,06} = \mathbf{0,42}$$

C. Keputusan

- Bandingkan: $F_{\text{hitung}}(0,42) < F_{\text{tabel}}(4,75)$.
- **Kesimpulan: Gagal Tolak** H_0.
- **Interpretasi:** Tidak terdapat bukti yang cukup bahwa variabel X berpengaruh secara linier terhadap Y.

Regresi Polinomial Ortogonal

UNIVERSITAS SULTAN AGENG TIRTAYASA

UNTIRTA Jawara
Jujur Adil Wibawa Amanah Religius Akuntabel

Uji Hipotesis Parsial

A. Hipotesis

- $H_0 : \alpha_2 = 0$ (Komponen kuadratik tidak signifikan).
- $H_1 : \alpha_2 \neq 0$ (Komponen kuadratik signifikan).

B. Perhitungan Statistik Uji

$$F_{\text{hitung}} = \frac{JK(\hat{\alpha}_2)}{KTS} = \frac{0,12}{59,06} = \mathbf{0,002}$$

C. Keputusan

- Bandingkan: $F_{\text{hitung}}(0,002) < F_{\text{tabel}}(4,75)$.
- **Kesimpulan: Gagal Tolak H_0.**
- **Interpretasi:** Penambahan derajat 2 (kuadratik) tidak memberikan kontribusi informasi apa pun terhadap model. Data tidak terbukti membentuk kurva.

Regresi Polinomial Ortogonal

Rekomendasi: Model regresi polinomial derajat 2 **tidak layak digunakan** untuk data ini. Variasi data Y terlalu acak (*random*) dan tidak dapat dijelaskan oleh variabel X dengan pendekatan polinomial ini. Sebaiknya peneliti mencari variabel prediktor lain atau memeriksa kualitas data.

SEE YOU NEXT WEEK !

Ferdian Bangkit Wijaya, S.Stat., M.Si
NIP. 199005202024061001
ferdian.bangkit@untirta.ac.id