

Pengujian Koefisien Regresi Baku (*Standardized Coefficients*) dengan Pendekatan Bayes pada Analisis Regresi Data Prevalensi *Stunting* di Indonesia

Nadia Tsabitah*, Suliadi

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Islam Bandung, Indonesia.

*nadiatsabitahhamzah@gmail.com, suliadi@gmail.com

Abstract. Standardized coefficients are linear regression coefficient that describe the strength and the direction of the linear relation between the dependent variable and the independent variable obtained when the two types of variables are standardized. Making confidence intervals for standardized coefficients using the frequency method is quite complicated, because the standardized coefficient is a nonlinear function of the random parameter estimates. This study applies Bayes method to the stunting prevalence in Indonesia. With this method, it is possible to obtain a credible interval of the standardized coefficient value with correlation matrix without use the complicated standard error values. The purpose of the study is to estimate the standardized coefficient to determine the effect on the prevalence of stunting in Indonesia. The result is the length of the credible interval generated by the Bayes method is narrower than the frequency one. The independent variables that have a direct significant effect on the prevalence of stunting in Indonesia based on the credible interval are the percentage of poor people with a credible interval ranged in $0,0071 < \beta_* < 0,6457$ and the average age at the first pregnancy for women with a credible interval ranged in $-1,2199 < \beta_* < -0.1007$.

Keywords: Bayes, Standardized Coefficient, Correlation Matrix, Credible Interval, Stunting.

Abstrak. Koefisien regresi yang dibakukan atau bisa disebut *standardized coefficients* merupakan koefisien regresi linier yang menggambarkan kekuatan dan arah hubungan linier antara variabel terikat (repons) dan variabel bebas (prediktor) yang diperoleh ketika kedua jenis variabel tersebut dibakukan. Pembuatan selang kepercayaan untuk koefisien regresi baku dengan menggunakan metode frekuentis tergolong rumit, karena koefisien regresi baku merupakan fungsi non-linier dari penduga parameter acak. Penelitian ini menerapkan metode bayes pada kasus prevalensi *stunting* di Indonesia. Dengan metode ini, bisa didapatkan selang kepercayaan dengan nilai koefisien regresi baku dengan matriks korelasi tanpa menggunakan nilai *standard error* yang rumit. Tujuan penelitian yaitu mengestimasi koefisien regresi yang dibakukan untuk mengetahui faktor-faktor yang mempengaruhi prevalensi *stunting* di Indonesia. Hasil pengujian yang dilakukan menunjukkan panjang selang kepercayaan yang dihasilkan dengan metode bayes lebih sempit daripada metode frekuentis. Variabel prediktor yang berpengaruh signifikan secara langsung terhadap prevalensi *stunting* berdasarkan selang kepercayaan yaitu, persentase penduduk miskin dengan selang kepercayaan $0,0071 < \beta_* < 0,6457$ dan rata-rata umur hamil pertama perempuan dengan selang kepercayaan $-1,2199 < \beta_* < -0.1007$.

Kata Kunci: Bayes, Koefisien regresi baku (*standardized coefficients*), Matriks Korelasi, Selang Kepercayaan (*credible interval*), *Stunting*.

1. Pendahuluan

Regresi linier berganda juga dapat digunakan untuk melihat variabel prediktor mana yang dinilai lebih berpengaruh dibanding variabel prediktor lainnya, yaitu dengan koefisien regresi baku. Koefisien regresi baku digunakan untuk membandingkan besarnya pengaruh koefisien regresi variabel bebas (prediktor), karena koefisien regresi baku bebas dari skala atau satuan dari variabel bebas (prediktor) dan variabel terikat (respons).

Koefisien regresi baku merupakan fungsi non-linier dari penduga parameter acak, maka untuk menciptakan selang kepercayaan (*credible interval*) menggunakan metode frekuentis tergolong rumit. Lu & Westfall (2019) mengemukakan algoritma pendugaan bayesian yang dapat menghasilkan selang kepercayaan (*credible interval*) dari koefisien regresi yang dibakukan (*standardized coefficients*).

Penelitian ini menerapkan metode di atas pada kasus prevalensi *stunting* di Indonesia. *Stunting* merupakan masalah gizi yang disebabkan oleh asupan gizi yang kurang dalam jangka waktu yang lama. *Stunting* merupakan masalah antar generasi yang terus menerus akan dihadapi bila dibiarkan. Menurut WHO (2018) penyebab terjadinya *stunting* pada anak terdapat 4 kategori besar yaitu faktor keluarga dan rumah tangga, makanan tambahan, menyusui dan infeksi. Faktor keluarga dan rumah tangga meliputi faktor maternal berupa nutrisi yang kurang, kehamilan pada usia remaja, kesehatan mental, dan lain-lain

Berdasarkan latar belakang yang telah diuraikan, maka tujuan dalam penelitian ini adalah:

1. Mengestimasi koefisien regresi baku untuk masing-masing faktor yang diasumsikan mempengaruhi prevalensi *stunting* di Indonesia
2. Menentukan faktor-faktor yang berpengaruh terhadap prevalensi *stunting* di Indonesia

2. Metodologi

Regresi Linear Berganda

Kurtner, et al., (2004) mengemukakan bahwa model regresi linier berganda untuk variabel respons y dengan jumlah pengamatan y_1, y_2, \dots, y_n dan variabel prediktor x_1, x_2, \dots, x_p dengan banyak pengamatan $x_{i1}, x_{i2}, \dots, x_{ip}$ adalah:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \text{ dengan } i = 1, 2, \dots, n.$$

Bila dituliskan dalam bentuk matriks adalah sebagai berikut:

$$Y = X\beta + \varepsilon$$

Menurut Kutner, et al., (2004) dengan metode kuadrat terkecil, penduga parameter pada regresi linier berganda dapat dinyatakan dengan

$$\hat{\beta}_j = (X'X)^{-1}X'y$$

Uji Asumsi Regresi

Asumsi yang harus dipenuhi pada model regresi dengan metode kuadrat terkecil adalah sebagai berikut:

1. Uji Normalitas

Hipotesis yang diuji adalah:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik uji yang digunakan menggunakan metode Shapiro-Wilk, dengan rumus sebagai berikut (NIST/SEMATECH,2012):

$$W = \frac{(\sum_{i=1}^n a_i y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Nilai kriteria ujinya adalah tolak H_0 jika tolak H_0 jika $W_{hitung} < W_{tabel(n,\alpha)}$ atau $p - value < \alpha$.

2. Pemeriksaan Multikolinearitas

Kurtner, et al., (2004) menyatakan cara untuk mendeteksi multikolinearitas dilihat dari nilai *Variance Inflating Factor* (VIF) dengan rumus sebagai berikut:

$$VIF = \frac{1}{(1-R_k^2)}$$

dimana $R_k^2 =$ koefisien determinasi. Apabila nilai VIF dari sebuah variabel prediktor bernilai lebih dari 10, dapat dikatakan terdapat multikolinieritas.

3. Uji Homogenitas

Hipotesis yang diuji adalah:

H_0 : Residual homogen

H_1 : Residual heterogen

Statistik uji yang digunakan adalah metode Breusch Pagan dengan rumus sebagai berikut (Kurtner, et al., 2004):

$$\chi^2 = \frac{JKR}{2} \div \left(\frac{JKS}{n} \right)^2$$

dimana

JKR = Jumlah Kuadrat Regresi

JKS = Jumlah Kuadrat Sisaan

n = Jumlah Pengamatan

Nilai kriteria ujinya adalah tolak H_0 jika $\chi^2_{hitung} > \chi^2_{tabel(df,\alpha)}$ atau $p - value < \alpha$.

4. Uji Koefisien Regresi Baku

Hipotesis yang diuji adalah:

H_0 : $\beta_{*j} = 0$ (parameter tidak signifikan)

H_1 : $\beta_{*j} \neq 0$ (parameter signifikan)

Statistik uji yang digunakan sebagai berikut:

$$\hat{\beta}_{*j} = \frac{s_j \hat{\beta}_j}{s_y}$$

Nilai kriteria ujinya adalah tolak H_0 jika $p - value < \alpha$.

5. Koefisien Regresi Baku/Dibakukan (*standardized*)

Koefisien regresi baku atau bisa disebut *standardized coefficients* merupakan nilai statistik dalam regresi linier yang menggambarkan kekuatan dan arah hubungan linier antara variabel terikat (repons) dan variabel bebas (prediktor).

Penduga Koefisien Regresi Baku dengan Metode Frekuentis

Menurut Kurtner, et al., (2004) penduga tak bias koefisien regresi yang dibakukan ($\hat{\beta}_*$) juga dapat dihitung menggunakan metode kuadrat terkecil, penduga tak bias koefisien regresi yang dibakukan untuk regresi linier berganda ($\hat{\beta}_{*j}$) didapat dengan:

$$\hat{\beta}_{*j} = \frac{s_j \hat{\beta}_j}{s_y}$$

Dimana $\hat{\beta}_j$ merupakan penduga parameter pada regresi linier berganda dan s_y^2 merupakan varians dari sampel pada variabel terikat (y), s_j^2 merupakan varians dari sampel pada variabel prediktor ke-j sebesar $\{S_{xx}\}_{jj}$ pada matriks kovarians sampel X.

Koefisien regresi yang dibakukan untuk variabel prediktor yang bersifat acak ($\beta_{*j}^{(u)}$) didapat dengan:

$$\beta_{*j}^{(u)} = \frac{\sigma_j \beta_j}{\sigma_y}$$

Dengan $Var(y_i) = \sigma_y^2 = \beta' \Sigma_{xx} \beta + \sigma^2$ dan $\sigma_j^2 = \{\Sigma_{xx}\}_{jj}$ yang merupakan kovarians untuk variabel prediktor ke j.

Koefisien regresi yang dibakukan untuk variabel prediktor yang bersifat tetap ($\beta_{*j}^{(c)}$) didapat dengan:

$$\beta_{*j}^{(c)} = \frac{s_j \beta_j}{\sigma_{y:x}}$$

Yuan & Chan (2011) menyarankan untuk menggunakan $\sigma_{y:x}^2 = \beta' S_{xx} \beta + \sigma^2$.

Penduga Standart Error untuk Koefisien Regresi Baku/Dibakukan

Untuk membangun selang kepercayaan dan menguji signifikansi dari β_{*j} dengan metode frekuentis dibutuhkan penduga dari *standart error*. Menurut Yuan & Chan (2011) persamaan

untuk mencari standard error koefisien regresi yang dibakukan pada variabel prediktor yang bersifat tetap atau yang bersifat acak dan tergolong sulit karena terdapat elemen penyusun persamaan-persamaan tersebut yang bersifat random sehingga banyak buku panduan ajar yang menuliskan rumus yang tidak tepat dalam mencari nilai *standart error*.

Menurut Yuan & Chan (2011) nilai standar error untuk variabel prediktor bersifat acak bernilai:

$$\hat{\sigma}^2_{(u)}(\hat{\beta}_{*j}) = \frac{s^2_j \{diag(S_{xx}^{-1})\}_{jj} \hat{\sigma}^2}{(n-3)s_y^2} + \frac{\hat{\beta}_j^2 [s_j^2 (\hat{\beta}' S_{xx} \hat{\beta}) - s_j^2 \hat{\sigma}^2 - \{(S_{xy})_j\}^2]}{(n-3)s_y^4}$$

Untuk nilai *standart error* koefisien regresi yang dibakukan dengan penduga koefisien regresi yang dibakukan variabel prediktor bersifat tetap bernilai:

$$\hat{\sigma}^2_{(c)}(\hat{\beta}_{*j}) = \frac{s^2_j \{diag(S_{xx}^{-1})\}_{jj} \hat{\sigma}^2}{ns_y^2} + \frac{s_j^2 \hat{\beta}_j^2 \hat{\sigma}^2 [3\hat{\sigma}^2 + 2(\hat{\beta}' S_{xx} \hat{\beta})]}{2ns_y^6}$$

Penduga Bayes

Selain dengan metode kuadrat terkecil, penduga koefisien regresi berganda linier juga dapat ditaksir dengan menggunakan metode bayesian.

Metode Bayes

Metode penduga bayes merupakan pendugaan dengan menggabungkan informasi yang terkandung dalam sampel dengan informasi yang telah tersedia sebelumnya yaitu pengetahuan subyektif mengenai distribusi probabilitas parameter. Teorema umum bayes adalah (Box & Tiao, 1973) :

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

Distribusi Prior

Dalam penggunaan teorema bayes, distribusi prior sangat dibutuhkan karena distribusi prior memberikan informasi yang sudah ada untuk parameter yang belum diketahui. Distribusi prior bisa dikelompokkan menjadi distribusi prior conjugate yang berarti distribusi posterior mengaju pada bentuk parametrik yang sama dengan distribusi priornya. Dan distribusi prior non conjugate tidak.

Distribusi Posterior

Distribusi posterior merupakan hasil dari perkalian likelihood dengan distribusi prior, sehingga penentuan parameter pada distribusi prior sangat berpengaruh. Distribusi posterior dapat dituliskan sebagai berikut:

$$f(\theta|y) = \frac{f(\theta, y)}{f(y)}$$

Penduga Selang Kepercayaan Koefisien Regresi Baku dengan Metode Bayes

Penduga koefisien regresi baku dengan metode bayes dibedakan menjadi dua cara, yaitu dengan menggunakan distribusi prior improper non-informatif dan menggunakan distribusi prior proper infomatif.

Penduga Koefisien Regresi Baku Metode Bayes dengan Distribusi Prior Improper Non informatif

Model regresi linier dengan kondisi variabel prediktor yang dibakukan akan menjadi:

$$y = \beta_0^x \mathbf{1} + X^s \beta^x + \varepsilon$$

dengan X^s adalah matriks dari variabel prediktor yang di bakukan yang bernilai $X^s = \{X - (1/n)11'X\}\{diag(S_{xx})\}^{-1/2}$ (bentuk matriks dari rumus koefisien regresi yang dibakukan untuk variabel prediktor). Menurut Lu & Westfall (2019), tetap dengan asumsi $\varepsilon \sim N_n(0, \sigma^2 I)$, misalkan $Z = [1: X^s]$ dan $\theta = [\beta_0: (\beta^x)']$. Distribusi data adalah normal multivariat sebagai berikut:

$$y|\theta, \sigma, X \sim N_n(Z\theta, \sigma^2 I)$$

dengan menggunakan prior non-informatif:

$$p(\theta, \sigma^2 | X) \propto \sigma^{-2}$$

sehingga didapat distribusi posterior:

$$\theta | y, X, \sigma^2 \sim N(\hat{\theta}, \sigma^2 (Z'Z)^{-1})$$

dengan $\hat{\theta}$ adalah penduga koefisien regresi baku yang didapatkan menggunakan:

$$\hat{\theta} = (Z'Z)^{-1}Z'y$$

Menggunakan keortogonalan 1 dan X^S serta sifat dari distribusi normal multivariat, distribusi posterior diatas bisa ditulis dengan:

$$\beta^x | y, X, \sigma^2 \sim N(\hat{\beta}^x, \sigma^2 R_{xx}^{-1} / (n - 1))$$

dengan $\hat{\beta}^x$ merupakan penduga koefisien regresi baku yang bernilai $(X^{S'}X^S)^{-1}X^{S'}y$. Distribusi posterior untuk σ^2 ditunjukkan dengan distribusi χ^2 -inverse (Gelman, et al., 2004):

$$\sigma^2 | y \sim Inv - \chi^2(n - k - 1, \sigma^2)$$

dimana

$$\sigma^2 = \frac{1}{(n - k - 1)} (y - X\hat{\beta})'(y - X\hat{\beta})$$

Menurut Lu & Westfall (2019), penduga koefisien regresi yang dibakukan untuk variabel prediktor bersifat tetap (*fixed*) dapat dihitung dengan menggunakan:

$$\beta_* = \frac{\beta^x}{(\beta^{x'}R_{xx}\beta^x + \sigma^2)^{1/2}}$$

Rumus diatas membutuhkan elemen-elemen seperti penduga varians (σ^2) dan penduga koefisien regresi baku hanya untuk variabel prediktor (β^x). Elemen-elemen tersebut jarang diketahui atau tidak tersedia. Oleh karena itu, untuk mendapatkan distribusi sampel posterior yang sama, digunakan matriks korelasi dari data (Y,X) dan sedikit modifikasi pada Persamaan (2.37) menjadi (Lu & Westfall (2019):

$$\beta_* = \frac{\beta^x / c}{\{(\beta^{x'} / c)R_{xx}(\beta^x / c) + (\sigma^2 / c^2)\}^{1/2}}$$

Nilai c merupakan sebuah nilai konstan dengan $c > 0$. (Lu & Westfall (2019) mengambil nilai $c = s_y$.

Penduga Koefisien Regresi Baku Metode Bayes dengan Distribusi Prior Proper Informatif

Distribusi prior proper informatif digunakan apabila terdapat beberapa informasi mengenai distribusi prior yang akan digunakan dalam koefisien regresi. Untuk koefisien regresi linier berganda dengan fungsi likelihood distribusi normal multivariat, digunakan distribusi prior *conjugate* yaitu distribusi normal-inverse-Gamma.

Selang Kepercayaan (Credible Interval)

Selang kepercayaan untuk bayesian merupakan sebuah rentang yang mengandung kemungkinan hasil penduga koefisien dengan persentase ketepatan tertentu. Pada penduga koefisien regresi baku, penduga selang kepercayaan untuk bayesian berasal dari distribusi posterior yang didapatkan dengan membangkitkan koefisien regresi baku (β_*). Metode untuk membangun selang kepercayaan untuk bayesian adalah *equal-tailed interval*. *Equal-tailed interval* atau bisa disebut *central interval* memiliki rentang interval:

$$I_\alpha = [Q_{\alpha/2}; Q_{1-\alpha/2}]$$

dengan Q merupakan kuantil-kuantil dari distribusi posterior. Menurut Zwillinger (2003) untuk mendapat penduga nilai kuantil secara manual, bisa dihitung dengan:

$$Q = q(n + 1)$$

dengan q merupakan kuantilnya atau proporsi di bawah nilai ke- i yang diinginkan, dan n merupakan banyaknya *item* pada *dataset*. Selang kepercayaan (*Credible Interval*) dari koefisien baku (*standardized coefficients*) dapat menunjukkan apakah variabel prediktor berpengaruh terhadap variabel respons secara signifikan dengan melakukan pengujian secara individual. Adapun hipotesis pengujian secara individual untuk koefisien baku β_{*j} dengan nilai $j = 1, 2, \dots, p$ adalah

$$H_0 : \beta_{*j} = 0$$

$$H_1 : \beta_{*j} \neq 0$$

Untuk kriteria uji yang digunakan yaitu, tolak H_0 apabila nilai rentang kuantil selang kepercayaan (*Credible Interval*) yang tidak melewati nilai nol. Hal tersebut menyimpulkan bahwa variabel prediktor tersebut berpengaruh terhadap variabel respons.

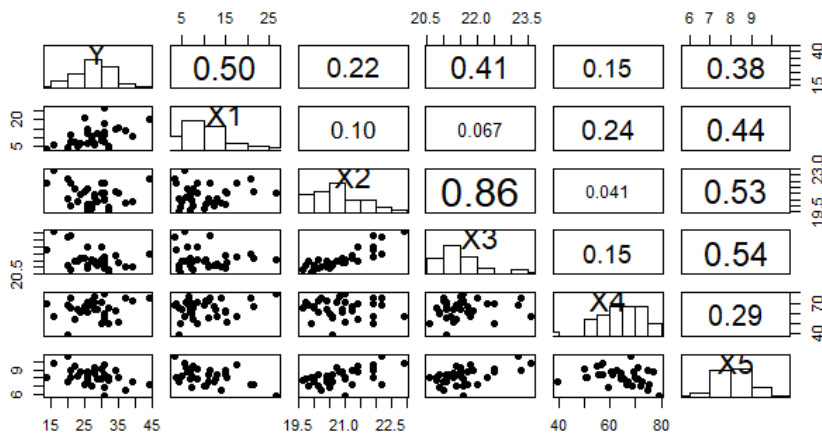
Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari hasil publikasi Badan Pusat Statistik (BPS) Indonesia yaitu Profil Statistik Kesehatan Tahun 2019 dan Laporan Akhir Penelitian Studi Status Gizi Balita Indonesia Tahun 2019. Penelitian ini menggunakan provinsi sebagai unit pengamatan yang terdiri dari 34 Provinsi pada tahun 2019. Variabel yang digunakan dalam penelitian ini adalah prevalensi *stunting* (Y), persentase penduduk miskin (X_1), rata-rata umur perkawinan pertama wanita usia 10 tahun keatas (X_2), rata-rata umur hamil pertama wanita usia 15-49 tahun (X_3), persentase anak usia 0-5 bulan yang diberikan ASI eksklusif (X_4), rata-rata lama sekolah perempuan (X_5).

3. Pembahasan dan Diskusi

Data yang digunakan pada penelitian ini dideskripsikan berdasarkan nilai rata-rata, nilai standar deviasi, nilai minimum, dan nilai maksimum. Karakteristik umum diketahui bahwa rata-rata prevalensi *stunting* di Indonesia (Y) sebesar 27,79% dengan standar deviasi sebesar 6,27%, prevalensi *stunting* terendah sebesar 14% dan prevalensi *stunting* tertinggi sebesar 44%.

Penyebaran data serta kemungkinan ada atau tidaknya hubungan setiap variabel penelitian berdasarkan provinsi ditampilkan dalam bentuk matriks plot. Matriks plot terdiri dari 3 komponen, segitiga atas merupakan nilai korelasi antar variabel, garis diagonal menunjukkan penyebaran data setiap variabel dengan histogram, dan segitiga bawah merupakan plot penyebaran data dan hubungan antar variabel.



Gambar 1. Matriks Plot

Analisis Regresi Linier Berganda dengan Metode OLS

Model koefisien regresi baku dengan metode OLS didapatkan dari pengujian menggunakan software R. Didapatkan model sebagai berikut:

$$Y = 0.3924X_1 + 0.3866X_2 - 0.7847X_3 + 0.2111X_4 + 0.6649X_5$$

Dengan nilai adjusted koefisien determinasi (\bar{R}^2) sebesar 0,3427.

Uji Normalitas

Didapatkan hasil output sebagai berikut:

W = 0.97938, p-value = 0.7531

Nilai statistik uji yang digunakan adalah $W = 0,9794$ dan $p\text{-value} = 0,7531$ Digunakan taraf nyata $\alpha = 5\%$ maka dapat dilihat dari hasil bahwa $p\text{-value} > \alpha$ atau $0,7531 > 0,05$ maka H_0 diterima. Artinya residual berdistribusi normal

Uji Multikolinearitas

Didapatkan hasil output sebagai berikut:

Tabel 1. Nilai VIF Pada Peubah Independen

Variabel	VIF
X ₁	1,7052*
X ₂	5,5582*
X ₃	5,4287*
X ₄	1,4712*
X ₅	2,2284*

Ket:* = Sig pada nilai VIF < 10

Multikolinearitas terjadi apabila nilai VIF > 10. Berdasarkan Tabel 1 nilai setiap VIF variabel bebas memiliki nilai kurang dari 10. Hal ini berarti tidak terjadi kasus multikolinearitas di semua variabel prediktor.

Uji Homogenitas

Didapatkan hasil output sebagai berikut:

BP = 10.335, df = 5, p-value = 0.06628
--

Nilai statistik uji yang digunakan adalah BP = 10,335 dan $p\text{-value} = 0,07$. Digunakan taraf nyata $\alpha = 5\%$ maka dapat dilihat dari hasil bahwa $p\text{-value} > \alpha$ atau $0,07 > 0,05$ maka H_0 diterima. Artinya residual homogen.

Uji Koefisien Regresi Baku

Dengan menggunakan software RStudio dan didapatkan hasil sebagai berikut.

Tabel 2. Nilai Koef. Regresi Baku dengan Metode OLS

Variabel	Koefisien	Std. Error	t value	p-value
X ₁	0,3924	0,1843	2,129	0,0422*
X ₂	0,3866	0,3327	1,162	0,2551
X ₃	-0,7847	0,3288	-2,386	0,0240*
X ₄	0,2111	0,1706	1,238	0,2262
X ₅	0,6649	0,2133	0,312	0,7576

Ket:* = Sig pada taraf nyata 5%

Digunakan taraf nyata $\alpha = 5\%$, maka berdasarkan Tabel 2 diatas terlihat bahwa hanya ada 2 variabel yang berpengaruh terhadap prevalensi stunting (Y), yaitu persentase penduduk miskin (X₁) dan variabel rata-rata umur hamil pertama wanita (X₃).

Distribusi Posterior Koefisien Regresi Baku

Dilakukan iterasi sebanyak 10.000 kali untuk menghasilkan distribusi posterior koefisien regresi yang dibakukan (*standardized coefficients*) dengan α sebesar 5%. Dibutuhkan nilai *adjusted R-squared* dan nilai penduga tak bias dari koefisien regresi yang dibakukan ($\hat{\beta}_*$) untuk membangun distribusi koefisien regresi yang dibakukan. Dengan nilai *adjusted R-squared* (\bar{R}^2) sebesar 0.3427 dan nilai penduga tak bias dari koefisien regresi baku pada Tabel 2, sehingga didapat distribusi posterior koefisien regresi baku dari masing-masing variabel prediktor dengan contoh sebagai berikut:

Tabel 3. Contoh Anggota Distribusi Posterior Koefisien Regresi yang Dibakukan

X ₁	X ₂	X ₃	X ₄	X ₅
0.4154	0.0807	-0.5575	0.2250	-0.0369
0.2458	0.5969	-0.8188	0.3225	0.1160
0.3701	0.2456	-0.6114	-0.0244	-0.0631
0.5637	0.3749	-0.7546	0.0336	0.2956
0.4035	0.0504	-0.2239	0.1337	-0.1710
0.6990	-0.3563	-0.1936	0.1211	0.3593

0.3647	0.2557	-0.3189	0.1931	-0.2358
0.3171	0.7782	-1.1578	0.3192	0.2146
0.5410	-0.2579	-0.1616	-0.0467	0.1506
0.4276	-0.1028	-0.6426	0.4771	0.3837

Selang Kepercayaan (*Credible Interval*)

Dengan α sebesar 5%, maka selang kepercayaan yang akan dihasilkan akan memiliki tingkat kepercayaan sebesar 95%.

Tabel 4. Hasil Selang Kepercayaan (*Credible Interval*) 95%

Variabel	β_*	2,5%	97,5%
X_1	0.3522	0.0071	0.6457
X_2	0.3452	-0.2584	0.913
X_3	-0.7018	-1.2199	-0.1007
X_4	0.1897	-0.1318	0.4728
X_5	0.0605	-0.3279	0.4321

Nilai koefisien regresi yang dibakukan diperoleh dari nilai *mean* dari selang kepercayaan dari masing-masing variabel prediktor (β_*).

Uji Individual dengan Pendekatan Selang Kepercayaan

Untuk mengetahui apakah terdapat pengaruh dari setiap variabel prediktor terhadap variabel respons secara signifikan, dilakukan pengujian secara individual dengan pendekatan selang kepercayaan.

Tabel 5. Perbandingan Nilai dan Selang Kepercayaan Koefisien Baku dengan OLS dan Bayes

Var.	OLS				Bayes			
	2,5%	97,5%	β_*	Sig.	2,5%	97,5%	β_*	Sig.
X_1	0.015	0.770	0.392	Sig.	0.007	0.646	0.352	Sig.
X_2	-0.295	1.068	0.387	Tidak Sig.	-0.258	0.913	0.345	Tidak Sig.
X_3	-1.458	-0.111	-0.787	Sig.	-1.220	-0.101	-0.702	Sig.
X_4	-0.138	0.561	0.211	Tidak Sig.	-0.132	0.473	0.190	Tidak Sig.
X_5	-0.370	0.503	-0.067	Tidak Sig.	-0.328	0.432	0.061	Tidak Sig.

Menurut Tabel 5 selang kepercayaan dari distribusi koefisien regresi yang dibakukan, pada taraf nyata 5%, terdapat 2 variabel prediktor yang cukup bukti berpengaruh signifikan secara langsung terhadap prevalensi *stunting*, yaitu persentase penduduk miskin (X_1) dan rata-rata umur hamil pertama (X_3), karena hanya kedua variabel tersebut yang mempunyai selang kepercayaan nilai rentang kuantil selang kepercayaan tidak melewati nilai nol.

Koefisien regresi yang dibakukan untuk variabel rata-rata umur hamil pertama (X_3) memiliki nilai mutlak terbesar, yaitu $|-0,7018| = 0,7018$ artinya rata-rata umur hamil pertama (X_3) bisa dikatakan variabel terpenting diantara lima variabel yang berpengaruh terhadap prevalensi *stunting* di Indonesia ($\beta_{*X5} < \beta_{*X4} < \beta_{*X2} < \beta_{*X1} < \beta_{*X3}$).

Tabel 5 menunjukkan perbandingan selang kepercayaan yang dihasilkan dari dua metode yaitu metode frekuentis dan bayes. Pada tingkat kepercayaan 95%, selang kepercayaan yang dihasilkan dengan metode bayes memiliki daerah cakupan yang lebih sempit atau panjang selang kepercayaan yang lebih pendek daripada selang kepercayaan yang dihasilkan dengan metode frekuentis.

4. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Faktor yang mempengaruhi prevalensi *stunting* di Indonesia tahun 2019 menggunakan pendekatan selang kepercayaan dengan metode bayes adalah variabel persentase penduduk miskin dengan selang kepercayaan $0,0071 < \beta_* < 0,6457$ dengan rata-rata sebesar 0,3522 dan variabel rata-rata umur hamil pertama wanita dengan selang kepercayaan $-1,2199 < \beta_* < -0,1007$ dengan rata-rata sebesar -0.7018.
2. Pada tahun 2019 prevalensi *stunting* di Indonesia memiliki rata-rata sebesar 27,79% dengan Provinsi Nusa Tenggara Timur sebagai provinsi dengan prevalensi *stunting* tertinggi sebesar 44% dan Provinsi Bali sebagai provinsi dengan prevalensi *stunting* terendah sebesar 14%.

Acknowledge

Penulis mengucapkan terima kasih kepada Bapak Suliadi, S.Si., M.Si. Ph.D. selaku dosen pembimbing yang telah memberikan bimbingan serta wawasan baru yang sangat membantu dalam menyelesaikan penelitian ini. Penulis juga berterima kasih kepada dosen-dosen pengampu mata kuliah pada jurusan statistika untuk segala ilmu yang diberikan untuk turut membantu penyusunan penelitian ini.

Daftar Pustaka

- [1] Badan Pusat Statistik. (2019). *Profil Statistik Kesehatan 2019*. Jakarta: Badan Pusat Statistik.
- [2] Box, G. E. P. and G. C. Tiao. (1973). *Bayesian Inference in Statistical Analysis*. Massachusetts: Addison-Wesley Publishing Company
- [3] Gelman, A., J. B. Carlin, H. S. Stern, & D. B. Rubin. (2004). *Bayesian Data Analysis*. Boca Raton (Florida): Chapman & Hall/CRC.
- [4] Kementerian Kesehatan Republik Indonesia. (2014). *Situasi dan Analisis ASI Eksklusif*. Jakarta: Pusat Data dan Informasi Kementerian Kesehatan RI.
- [5] Kurtner, M.H., C.J. Nachtsheim, J. Neter, & W. Li. (2004). *Applied Linear Statistical Models*. New York: The McGraw-Hill Companies, Inc.
- [6] Lu, Y. & P. Westfall. (2019). Simple and Flexible Bayesian Inferences for Standardized Regression Coefficients. *Journal of Applied Statistics*, 2254-2288.
- [7] NIST/SEMATECH. (2012). *E-Handbook of Statistical Methods*. (Online), (<https://doi.org/10.18434/M32189>, diakses 8 April 2021).
- [8] Yuan, K.-H. & W. Chan. (2011). Biases and Standard Errors of Standardized Regression. *Psychometrika*, 670 – 690.
- [9] Zwillinger, D. (2003). *Standard Mathematical Tables and Formulae*. New York: A CRC Press Company.
- [10] Utama Muhammad Bangkit Riksa, Hajarisman Nusar. (2021). *Metode Pemilihan Variabel pada Model Regresi Poisson Menggunakan Metode Nordberg*. *Jurnal Riset Statistika*, 1(1), 35-42.