

Model Regresi Logistik Biner Untuk Menentukan Keputusan Mahasiswa Statistika Universitas Negeri Padang dalam Memilih Bahan Bakar Minyak (BBM)

Afdhal Rezeki, Muhamad Fajri, Muhammad Tibri Syofyan
Mahasiswa Jurusan Statistika Universitas Negeri Padang
rezekiafdhal26@gmail.com, doangfajri45@gmail.com, tibri.work@gmail.com

ABSTRACT

Oil fuel is fuel derived and / or processed from petroleum. Oil fuels that are often used for motorized vehicles, especially two-wheeled motorized vehicles, are premium fuel oil, pertalite, and pertamax. Recently, the premium type of fuel oil will be phased out because this type of fuel has a low octane rating so it is not environmentally friendly. Elimination of premium fuel oil can be a consideration for consumers to choose other types of fuel oil. Many factors influence consumer decisions in choosing Oil Fuel apart from the price and fuel subsidies. In statistics, there are several methods that can be used to determine the factors that influence consumer decisions in choosing oil fuel, one of which is binary logistic regression. Binary logistic regression is a data analysis method used to find the relationship between the binary or dichotomous response variable (Y) and several predictor variables (X).

Keywords : *fuel, binary logistic regression model, categorical data analysis*

1. PENDAHULUAN

1.1. Latar Belakang

Di era globalisasi seperti sekarang ini, banyak orang yang menganggap bahwa kendaraan bermotor sudah menjadi kebutuhan primer. Hal ini dapat dilihat dari meningkatnya pengguna kendaraan bermotor khususnya kendaraan bermotor roda dua. Kenaikan pengguna kendaraan bermotor roda dua dapat mempengaruhi konsumsi bahan bakar. Bahan bakar minyak (BBM) adalah bahan bakar yang berasal dari minyak bumi.

Bahan Bakar Minyak yang sering digunakan untuk bahan bakar kendaraan bermotor roda dua antara lain premium, pertalite, dan pertamax. Beberapa waktu terakhir ini BBM jenis premium akan dihapuskan dikarenakan jenis bahan bakar ini memiliki oktan rendah sehingga tidak ramah lingkungan. BBM jenis premium mempunyai harga lebih murah dibandingkan dengan harga BBM jenis lainnya. Hal tersebut sangat mempengaruhi perekonomian masyarakat terutama mahasiswa. Penghapusan BBM jenis premium juga dapat menjadi salah satu pertimbangan bagi konsumen untuk memilih BBM jenis lainnya.

Banyak faktor yang mempengaruhi keputusan konsumen dalam memilih BBM selain dari harga dan subsidi BBM. Dalam statistika ada beberapa metode yang dapat digunakan untuk mengetahui faktor-faktor yang mempengaruhi keputusan konsumen dalam memilih BBM, salah satunya adalah regresi logistik biner. Regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (Y) yang bersifat biner atau dikotomis dengan beberapa variabel prediktor (X).

1.2. Tujuan

Tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor apa saja yang mempengaruhi keputusan mahasiswa statistika dalam memilih BBM menggunakan model regresi logistik biner.

2. TINJAUAN PUSTAKA

2.1. Bahan Bakar Minyak (BBM)

BBM adalah bahan bakar yang berasal dan/atau diolah dari minyak bumi. PT Pertamina adalah perusahaan minyak dan gas bumi yang dimiliki Pemerintah Indonesia. BBM yang sering digunakan untuk kendaraan bermotor, khususnya kendaraan bermotor roda dua adalahh BBM jenis premium, pertalite, dan pertamax. Premium merupakan bahan bakar mesin dengan angka oktan minimal 88 yang dapat digunakan pada kendaraan bermotor bensin dengan risiko kompresi rendah (dibawah 9,1). Pertalite merupakan bahan bakar gasoline yang memiliki angka oktan 90 yang dapat digunakan pada kendaraan dengan kompresi 9,1 hingga 10,1. Pertamax merupakan bahan bakar bensin dengan angka oktan minimal 92 berstandar international yang dapat digunakan pada kendaraan yang memiliki rasio 10:1 hingga 11:1.

2.2. Model Regresi Logistik Biner

Regresi logistik biner merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (Y) yang bersifat biner atau dikotomus dengan beberapa variabel prediktor (X). Model regresi logistik biner digunakan untuk menganalisis hubungan antara satu variabel respon dengan beberapa variabel prediktor, dengan variabel respon berupa data kualitatif dikotomi yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik. Model regresi logistik tersebut adalah

$$\pi(x_i) = \frac{e^{\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}}}{1 + e^{\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}}} \quad (1)$$

Persamaan (1) mempunyai bentuk yang tidak linier. Untuk membuat persamaan tersebut menjadi persamaan yang linier, maka digunakan transformasi log atau disebut juga transformasi logit. Berikut ini adalah logit dari $\pi(x_i)$

$$\ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} \quad (2)$$

$$\text{Jika } g(x_i) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}, \text{ maka } \pi(x_i) = \frac{e^{g(x_i)}}{1 + e^{g(x_i)}}$$

2.2.1. Pengujian Signifikansi Parameter

a. Uji Rasio Likelihood

Hipotesis :

$H_0 = \beta_0 = \dots = \beta_p = 0$ (Secara bersama-sama variabel prediktor tidak berpengaruh terhadap model)

$H_1 = \exists \beta_j \neq 0; j = 1, 2, \dots, p$ (Secara bersama-sama variabel prediktor berpengaruh terhadap model)

Statistik uji: $G = -2 \ln\left(\frac{l_0}{l_1}\right) = -2(\ln l_0 - \ln l_1) = -2(L_0 - L_1)$ atau

$$G = -2 \ln \frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\sum_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{1-y_i}}$$

dengan L_0 adalah log likelihood tanpa variabel prediktor

L_1 adalah log likelihood dengan variabel prediktor

Kriteria uji: Tolak H_0 jika $G > X^2_{(\alpha, p)}$.

- b. Uji Wald
 Hipotesis :
 $H_0 = \beta_j = 0, j = 1, 2, \dots, p$ (variabel prediktor tidak berpengaruh terhadap model)
 $H_1 = \beta_j \neq 0; j = 1, 2, \dots, p$ (variabel prediktor berpengaruh terhadap model)
 Statistik uji: $W = \left[\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right]^2$
 Kriteria uji: Tolak H_0 jika $W > X^2_{(\alpha, 1)}$ atau nilai sign $< \alpha$.
- c. Uji *Goodness of Fit*
 Hipotesis :
 H_0 = Model sesuai (tidak ada perbedaan antara observasi dan prediksi)
 H_1 = Model tidak sesuai (ada perbedaan antara observasi dan prediksi)
 Statistik uji: $\hat{C} = \sum_{b=1}^g \frac{(O_b - n'_b \bar{\pi}_b)^2}{n'_b \bar{\pi}_b (1 - \bar{\pi}_b)}$
 Kriteria uji: Tolak H_0 jika $C > X^2_{(\alpha, g-2)}$ atau nilai sign $< \alpha$

3. METODE PENELITIAN

3.1. Jenis, Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data mentah yang diperoleh dari Mahasiswa Statistika UNP melalui platform *google forms*. Data yang digunakan sebanyak 102 data.

Variabel penelitian yang digunakan untuk model regresi logistik biner dalam penelitian ini terdiri dari variabel respon dan variabel prediktor. Berikut merupakan tabel yang menjelaskan variabel respon (Y) dan variabel prediktor (X), dengan variabel-variabelnya sebagai berikut

Tabel 1. Variabel Penelitian Model Regresi Logistik Biner

No	Jenis Variabel	Keterangan	Tipe Data
1	Jenis BBM (Y)	0 = Pertamax 1 = Pertalite	Nominal
2	Jenis Kelamin (X ₁)	0 = Laki-Laki 1 = Perempuan	Nominal
3	Jenis Kendaraan (X ₂)	0 = Non Matik 1 = Matik	Nominal
4	Asal Daerah (X ₃)	0 = Luar Kota (Padang) 1 = Dalam Kota (Padang)	Nominal
5	Usia Kendaraan (X ₄)	0 = < 1 Tahun 1 = 1 - 5 Tahun 2 = > 5 Tahun	Ordinal
6	Pendapatan Tiap Bulan (X ₅)	0 = < Rp. 500.000 1 = Rp. 500.000 – Rp. 800.000 2 = > Rp. 800.000	Ordinal

3.2. Tahapan Analisis

Tahapan analisis yang digunakan sebagai berikut

- 1) Mengidentifikasi variabel-variabel yang digunakan.
- 2) Mengumpulkan data primer yaitu dengan memberikan g-form (*google form*) kepada responden penelitian.
- 3) Menganalisis data dengan menggunakan analisis regresi logistik dan bantuan *software* statistika.
- 4) Menyimpulkan hasil analisis.

4. HASIL DAN PEMBAHASAN

4.1. Model Regresi Logistik Biner

a. Uji Rasio Likelihood

Hipotesis: $H_0 : \beta_0 = \dots = \beta_7 = 0$

$$H_0 : \exists \beta_j \neq 0, j = 1, 2, \dots, 7$$

Taraf Signifikansi $\alpha = 5\%$

Tabel 2. Uji Signifikansi Secara Keseluruhan

<i>-2 Log likelihood</i>	<i>Cox & Snell R Square</i>	<i>Nagelkerke R Square</i>
133.607	0.068	0.091

Statistik uji: $G = -2 \ln \left(\frac{l_0}{l_1} \right) = 133.607$

Keputusan : Karena nilai $G = 133.607 > X^2_{(0.05,7)} = 14.0671$ maka H_0 ditolak

Kesimpulan: Berdasarkan keputusan yang diperoleh, pada taraf signifikansi 5% dapat disimpulkan bahwa terdapat minimal ada satu $\beta_j \neq 0$ atau secara bersama-sama variabel prediktor berpengaruh terhadap model.

b. Uji Wald

Hipotesis :

$$H_0 = \beta_j = 0; j = 1, 2, \dots, 7$$

$$H_1 = \beta_j \neq 0; j = 1, 2, \dots, 7$$

Taraf signifikansi $\alpha = 5\%$

Statistik uji: $W = \left[\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right]^2$

Tabel 3. Nilai Statistik Uji Wald

	B	SE	Wald	df	Sig	Exp(B)
X1(1)	0.513	0.549	0.872	1	0.350	1.671
X2(1)	-0.792	0.669	1.402	1	0.236	0.453
X3(1)	-0.019	0.526	0.001	1	0.971	0.981
X4			2.119	2	0.347	
X4(1)	1.508	1.212	1.548	1	0.213	4.518
X4(2)	-0.182	0.443	0.169	1	0.681	0.834
X5			0.217	2	0.897	
X5(1)	-0.060	0.592	0.010	1	0.919	0.942
X5(2)	-0.240	0.600	0.159	1	0.690	0.787
Constant	-0.324	0.727	0.198	1	0.656	0.723

Kriteria uji: Tolak H_0 jika $W > X^2_{(\alpha,1)}$

Kesimpulan: Dari nilai uji statistik Wald tersebut dapat disimpulkan bahwa H_0 diterima. Ini berarti bahwa semua variabel prediktor tidak berpengaruh terhadap model.

c. Uji *Goodness of Fit*

Hipotesis :

H_0 = Model sesuai (tidak ada perbedaan antara observasi dan prediksi)

H_1 = Model tidak sesuai (ada perbedaan antara observasi dan prediksi)

Statistik uji *Hosmer-Lemeshow*: $\hat{C} = \sum_{b=1}^g \frac{(o_b - n'_b \bar{\pi}_b)^2}{n'_b \bar{\pi}_b (1 - \bar{\pi}_b)}$

Kriteria uji: Tolak H_0 jika $C > X^2_{(\alpha, g-2)}$ atau nilai sign $< \alpha$

Tabel 4. Tabel Kontingensi Uji *Hosmer-Lemeshow*

	Y = Pertalite		Y = Pertamina		Total
	Observasi	Prediksi	Observasi	Prediksi	
1	9	7.670	1	2.330	10
2	3	5.921	6	3.079	9
3	6	6.137	4	3.863	10
4	11	8.943	5	7.057	16
5	2	4.659	7	4.341	9
6	7	5.131	3	4.869	10
7	4	4.091	4	3.909	8
8	5	5.428	6	5.572	11
9	5	4.209	4	4.791	9
10	3	2.811	7	7.189	10

Kesimpulan: Dari Tabel 4 diatas, dapat dilihat bahwa dari sepuluh langkah pengamatan untuk keputusan pemilihan jenis BBM Pertamina (0) dan Pertalite (1). Nilai yang diamati maupun nilai yang diprediksi, tidak mempunyai perbedaan yang terlalu ekstrim. Ini menunjukkan bahwa model regresi logistik yang digunakan dalam penelitian ini mampu memprediksi nilai observasinya.

Tabel 5. Tabel Uji *Hosmer-Lemeshow*

Chi-Square	df	Sig.
11.197	8	0.191

Berdasarkan Tabel 5, dapat dilihat bahwa nilai $C = 11.197 < X^2_{(\alpha, g-2)} = 15,5073$ atau nilai sign = 0.191 > nilai taraf (α) = 0.05 maka H_0 diterima. Hal ini berarti model regresi layak digunakan dalam analisis selanjutnya karena tidak ada perbedaan yang nyata antara klasifikasi yang diprediksi dengan klasifikasi yang diamati

Setelah dilakukan uji signifikansi paramter diatas, maka model logistik dapat dibentuk dengan menggunakan nilai taksiran parameter pada Tabel 3. Model yang terbentuk adalah :

$$g(x_i) = -0.324 + 0,513 X1(1) - 0,792 X2(1) - 0,019 X3(1) + 1,508 X4(1) - 0,182 X4(2) - 0,060 X5(1) - 0,240 X5(2)$$

Persamaan menunjukkan bahwa nilai intersep = - 0.324. Artinya: $g(x_i) = - 0.324$, pada saat tidak dipengaruhi oleh semua variabel.

5. KESIMPULAN

Penelitian ini meneliti tentang pengaruh jenis kelamin, jenis kendaraan, asal daerah, usia kendaraan, dan pendapatan tiap bulan terhadap keputusan pemilihan jenis BBM. Variabel penelitian yang digunakan adalah variabel dependen, yaitu jenis BBM. Sedangkan variabel independen yang digunakan adalah jenis kelamin, jenis kendaraan, asal daerah, usia kendaraan, dan pendapatan tiap bulan. Analisis dilakukan dengan menggunakan analisis regresi logistik biner. Data sampel mahasiswa Statistika Universitas Negeri Padang sebanyak 102 pengamatan yang diambil secara non-probabilitas sampling.

Berdasarkan hasil dari analisis data beserta pembahasan yang telah dilakukan, maka diperoleh model sebagai berikut:

$$g(x_i) = -0.324 + 0,513 X1(1) - 0,792 X2(1) - 0,019 X3(1) + 1,508 X4(1) \\ - 0,182 X4(2) - 0,060 X5(1) - 0,240 X5(2)$$

Pada pengujian yang dilakukan, diperoleh variabel prediktor yang belum signifikan berpengaruh dalam model. Hal ini disebabkan karena belum didapatkan suatu variabel prediktor yang secara signifikan mempengaruhi model. Sedangkan dengan pengujian secara simultan diperoleh bahwa dengan pemakaian variabel prediktor secara bersamaan dapat mempengaruhi model.

DAFTAR PUSTAKA

- [1] Agresti, A. 1990. *Categorical Data Analysis*. John Wiley & Sons, Inc, New York.
- [2] Hosmer, D., & Lemeshow, S. (2000). *Applied Logistic Regressions*. USA: John Wiley & Sons.
- [3] Kotimah, Muinah Kusnul dan Sri Pingit Wulandari, "Model Regresi Logistik Biner Stratifikasi Pada Partisipasi Ekonomi Perempuan Di Provinsi Jawa Timur". *Jurnal Sains dan Seni Pomits*, Volume 3, Nomor 1, (2014): 2337-3520. Institut Teknologi Sepuluh November. (diakses 18 Desember 2020)
- [4] PT. Pertamina. <https://www.pertamina.com/id/fuel-retail> [diakses 18 Desember 2020]