

# IDENTIFIKASI DISTRIBUSI JUMLAH KECELAKAAN LALU LINTAS DI DEPOK DAN PENDUGAAN PARAMETERNYA MENGGUNAKAN METODE BAYES

Tari Adriana Musana, Ferra Yanuar, Yudiantri Asdi

*Program Studi Matematika,  
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,  
Kampus UNAND Limau Manis Padang, Indonesia,  
tariadriana0103@gmail.com*

**Abstrak.** Kecelakaan lalu lintas adalah masalah yang umum terjadi dalam penyelenggaraan sistem transportasi. Menurut data dari Polresta Depok, diperoleh informasi bahwa jumlah kecelakaan lalu lintas mengalami penurunan di tahun 2018. Rata-rata dari jumlah kecelakaan di suatu ruas jalan perlu diketahui sebagai upaya untuk mengurangi jumlah kecelakaan. Berdasarkan data jumlah kecelakaan lalu lintas dapat ditentukan distribusi tertentu, yakni distribusi Binomial Negatif. Pada distribusi Binomial Negatif akan dilakukan pendugaan parameter dispersi ( $k$ ) diketahui dan parameter skala ( $\theta$ ). Pendugaan parameter untuk parameter  $\theta$  dari distribusi Binomial Negatif menggunakan metode Bayes. Hasil simulasi data menunjukkan bahwa distribusi Beta( $\alpha, \beta$ ) sebagai distribusi prior konjugat lebih baik dari pada distribusi Gamma dan distribusi prior Jeffrey.

**Kata Kunci :** *Jumlah Kecelakaan Lalu Lintas, distribusi Binomial Negatif, Metode Bayes.*

## 1. Pendahuluan

Kecelakaan lalu lintas adalah masalah yang umum terjadi dalam penyelenggaraan sistem transportasi. Faktor-faktor yang menyebabkan kecelakaan lalu lintas adalah perilaku berlalu lintas, cuaca, kondisi kendaraan dan kondisi jalan [6].

Pada tugas akhir ini akan ditentukan distribusi dari data jumlah kecelakaan lalu lintas, kemudian akan dilakukan pendugaan parameter untuk memperoleh nilai duga rata-rata terjadinya kecelakaan pada ruas jalan tertentu, disimbolkan dengan parameter  $\theta$ . Metode penduga parameter tersebut menggunakan metode Bayes. Rata-rata dari jumlah kecelakaan di suatu ruas jalan perlu diketahui sebagai upaya untuk mengurangi jumlah kecelakaan.

## 2. Landasan Teori

### 2.1. Distribusi Binomial Negatif

**Definisi 2.1.** [?] Suatu peubah acak  $X$  dikatakan memiliki distribusi Binomial Negatif dengan parameter  $k$  dan  $\theta$ , bila memiliki fungsi kepekatan peluang dalam bentuk

$$f(x; k, \theta) = \binom{x-1}{k-1} \theta^k (1-\theta)^{x-k} \quad x = k, k+1, \dots \quad (2.1)$$

dimana  $x$  menyatakan banyaknya ulangan yang diperlukan untuk mendapatkan  $k$  keberhasilan dan  $\theta$  menyatakan probabilitas sukses dalam suatu percobaan konstan. Nilai dari parameter  $k$  merupakan bilangan bulat positif.

## 2.2. Distribusi Beta

**Definisi 2.2.** [?] Suatu peubah acak  $X$  dikatakan memiliki distribusi Beta ditulis  $X \sim \text{Beta}(\alpha, \beta)$ , jika fungsi kepekatan peluangnya adalah

$$f(x) = \begin{cases} \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} & ; 0 < x < 1 \text{ dan } \alpha, \beta > 0, \\ 0 & ; \text{lainya.} \end{cases} \quad (2.2)$$

## 2.3. Distribusi Gamma

**Definisi 2.3.** [?] Misalkan  $X$  suatu peubah acak kontinu berdistribusi Gamma dengan parameter  $\alpha > 0$  dan  $\beta > 0$ , maka bentuk fungsi kepekatan peluangnya adalah

$$f(x, \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & \text{untuk } x > 0, \\ 0 & \text{lainya.} \end{cases} \quad (2.3)$$

## 2.4. Distribusi Uniform

**Definisi 2.4.** [?] Suatu peubah acak  $X$  dikatakan distribusi Uniform pada interval  $a \leq x \leq b$  jika fungsi kepekatan peluang dari distribusi Uniform, yaitu:

$$f(x, a, b) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b, \\ 0 & \text{lainya.} \end{cases} \quad (2.4)$$

## 2.5. Pendugaan Parameter dengan Metode Bayes

### 2.5.1. Fungsi Likelihood

**Definisi 2.5.** [?] Fungsi kepekatan peluang bersama peubah acak  $X_1, \dots, X_n$  yang dihitung pada  $x_1, \dots, x_n$  yang dinyatakan dalam bentuk  $f(x_1, \dots, x_n | \theta)$  dikatakan sebagai fungsi likelihood. Untuk  $x_1, \dots, x_n$  tetap, fungsi likelihood adalah fungsi dari parameter  $\theta$  yang dinotasikan dengan  $L(\theta)$ . Jika  $X_1, \dots, X_n$  merupakan sampel acak yang saling bebas dari  $f(x|\theta)$  maka:

$$L(\theta) = f(x_1; \theta) f(x_2; \theta) \cdots f(x_n; \theta) = \prod_{i=1}^n f(x_i; \theta). \quad (2.5)$$

### 2.5.2. Distribusi Posterior

**Definisi 2.6.** [?] Fungsi kepekatan peluang bersyarat dari  $\theta$  diberikan pengamatan sampel  $x = (x_1, \dots, x_n)$  disebut fungsi kepekatan peluang posterior, dan diberikan oleh

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{\int_{-\infty}^{\infty} f(x|\theta)f(\theta) d\theta} \quad (2.6)$$

## 2.6. Metode Jeffrey

**Aturan Jeffrey** Distribusi prior  $f(\theta)$  dikatakan distribusi prior noninformatif dari parameter  $\theta$  jika distribusi prior tersebut proporsional dengan akar dari informasi Fisher  $I(\theta)$ [4].

$$f(\theta) \propto \sqrt{I(\theta)} \quad (2.7)$$

dengan informasi Fisher dari parameter untuk suatu peubah acak  $x = (x_1, \dots, x_n)$  didefinisikan dengan:

$$I(\theta) = -E \left[ \frac{\partial^2 \ln f(x|\theta)}{\partial \theta^2} \right] \quad (2.8)$$

Dimana  $f(x|\theta)$  merupakan fungsi kepadatan peluang dari  $x$ .

## 2.7. Monte Carlo Markov Chain (MCMC)

Monte Carlo Markov Chain (MCMC) adalah metode untuk membangkitkan nilai sampel secara random dari distribusi posterior kompleks dengan membangun rantai Markov yang memiliki distribusi limit mendekati distribusi posterior kompleks [10].

## 2.8. Uji Kolmogorov-Smirnov

Uji Kolmogorov-Smirnov adalah suatu uji statistik yang digunakan untuk menguji hipotesis tentang kecocokan distribusi data (yang diobservasi) dengan distribusi teoritis tertentu. Statistik uji yang digunakan adalah:

$$D = \sup |F_0(x) - S_N(x)| \quad (2.9)$$

Hipotesis uji Kolmogorov-Smirnov adalah sebagai berikut:

$H_0$  : Data mengikuti distribusi tertentu.

$H_1$  : Data tidak mengikuti distribusi tertentu.

Aturan keputusannya adalah tolak  $H_0$  jika  $p - value < \alpha$  atau tolak  $H_0$  jika  $D > D_{tabel}$  [12]. Nilai  $D_{tabel}$  dapat diperoleh dengan rumus  $D_{tabel} = \frac{1,36}{\sqrt{N}}$  [12].

## 3. Pembahasan

### 3.1. Identifikasi Jenis Distribusi dari Data Kasus

Pertama, akan dilakukan analisis apakah data berdistribusi Poisson atau tidak dengan memilih taraf nyata  $\alpha = 0,05$  dan hipotesis yang digunakan adalah:

$H_0$  : Data jumlah kecelakaan lalu lintas di suatu ruas jalan tertentu berdistribusi Poisson.

$H_1$  : Data jumlah kecelakaan lalu lintas di suatu ruas jalan tertentu tidak berdistribusi Poisson.

Setelah dilakukan perhitungan uji Kolmogorov-Smirnov dengan menggunakan program SPSS 16 diperoleh nilai D yaitu 0,304 dan  $p - value$  yaitu 0,000 yang disajikan pada lampiran 1. Sedangkan nilai  $D_{tabel}$  adalah 0,1034. Karena diperoleh  $p - value < \alpha$  ( $0,000 < 0,05$ ) dan  $D_{tabel} < D$  ( $0,1034 < 0,304$ ) maka tolak

$H_0$ . Hal ini berarti pada tingkat kepercayaan 95% sampel yang digunakan pada penelitian ini yaitu data jumlah kecelakaan lalu lintas di suatu jalan tersebut tidak berdistribusi Poisson. Karena data mengalami overdispersi, maka data dimodelkan ulang dengan model Binomial Negatif.

### 3.2. Inferensi Bayesian dari Distribusi Binomial Negatif

#### 3.2.1. Fungsi Likelihood dari Distribusi Binomial Negatif

Fungsi Likelihood distribusi Binomial Negatif dengan fungsi kepekatan peluang sesuai persamaan 4.2.1

$$f(y|\theta) = \frac{\theta^{kn}(1-\theta)^{\sum_{i=1}^n y_i}}{(\Gamma(k))^n} \prod_{i=1}^n \frac{\Gamma(y_i+k)}{\Gamma(y_i+1)}. \quad (3.1)$$

#### 3.2.2. Pendugaan Titik Distribusi Prior Konjugat

Jika  $X_1, \dots, X_n$  merupakan peubah acak berdistribusi Beta dua parameter yaitu  $\alpha$  dan  $\beta$ , dengan  $X_i \sim Beta(\alpha, \beta)$  maka fungsi kepekatan peluangnya, yaitu:

$$f(\theta) = \frac{1}{(\alpha, \beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}. \quad (3.2)$$

Dari persamaan (4.2.6) dan persamaan (4.2.7) diperoleh distribusi posterior sebagai berikut:

$$f(\theta|y) = \frac{\theta^{kn+\alpha-1} (1-\theta)^{\sum_{i=1}^n y_i + \beta - 1}}{\left( kn + \alpha, \sum_{i=1}^n y_i + \beta \right)} \quad (3.3)$$

Atau dapat dinyatakan bahwa:

$$\theta|y \sim Beta \left( kn + \alpha, \sum_{i=1}^n y_i + \beta \right). \quad (3.4)$$

#### 3.2.3. Pendugaan Titik Distribusi Prior Non Konjugat

Peubah acak  $\theta$  berdistribusi Gamma dua parameter yaitu  $\alpha$  dan  $\beta$ , dinotasikan dengan  $\theta \sim Gamma(\alpha, \beta)$ . Berdasarkan Definisi 2.8.1 fungsi kepekatan peluang distribusi Gamma, yaitu:

$$f(\theta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \theta^{\alpha-1} e^{-\frac{\theta}{\beta}}. \quad (3.5)$$

Dari persamaan (4.2.13) dan persamaan (4.2.14) diperoleh distribusi posterior sebagai berikut:

$$f(\theta|y) = \frac{\frac{\theta^{kn+\alpha-1} ((1-\theta)^{\sum_{i=1}^n y_i} e^{-\frac{\theta}{\beta}})}{\Gamma(k)^n \beta^\alpha \Gamma(\alpha)} \prod_{i=1}^n \frac{\Gamma(y_i+k)}{\Gamma(y_i+1)}}{\int_0^1 \frac{\theta^{kn+\alpha-1} ((1-\theta)^{\sum_{i=1}^n y_i} e^{-\frac{\theta}{\beta}})}{\Gamma(k)^n \beta^\alpha \Gamma(\alpha)} \prod_{i=1}^n \frac{\Gamma(y_i+k)}{\Gamma(y_i+1)} d\theta} \quad (3.6)$$

3.2.4. *Pendugaan Titik Distribusi Prior Noninformatif*

Misalkan  $X_1, \dots, X_n$  contoh acak yang berdistribusi Binomial Negatif dua parameter yaitu parameter dispersi,  $k$  diketahui dan parameter skala,  $\theta$  yang dinotasikan dengan  $X_i \sim BN(k, \theta)$ . Misalkan  $v = (k, \theta)$ , dimana  $v$  adalah suatu peubah acak dengan dua parameter yaitu  $k$  dan  $\theta$  yang saling bebas. Sehingga fungsi kepekatan peluangnya adalah  $f(v) = f(k, \theta) = f(k)f(\theta)$ . Nilai prior noninformatif  $f(\theta)$  diperoleh dengan menggunakan pendekatan Jeffrey. Berdasarkan aturan Jeffrey, diperoleh

$$f(\theta) = \frac{1}{\theta(1-\theta)} \sqrt{kn(1+\theta^2) - 3\theta nk} \tag{3.7}$$

Selanjutnya akan dicari nilai prior noninformatif  $f(k)$ , karena  $k$  nilainya diketahui atau dengan kata lain  $k$  merupakan konstanta, sehingga  $f(k) = c$  untuk suatu  $c$  (konstan). Dengan demikian, diperoleh prior noninformatif bersama  $f(v)$ , yaitu:

$$f(v) = f(\theta)f(k) = \frac{\sqrt{kn(1+\theta^2) - 3\theta nk}}{\theta(1-\theta)}. \tag{3.8}$$

Dari persamaan (4.2.18) dan persamaan (4.2.19) diperoleh distribusi posterior sebagai berikut:

$$f(v|y) = \frac{\frac{\theta^{kn-1}((1-\theta)^{\sum_{i=1}^n y_i - 1} \sqrt{kn(1+\theta^2) - 3\theta nk})}{\Gamma(k)^n} \prod_{i=1}^n \frac{\Gamma(y_i+k)}{\Gamma(y_i+1)}}{\int_0^1 \int_0^1 \frac{\theta^{kn-1}((1-\theta)^{\sum_{i=1}^n y_i - 1} \sqrt{kn(1+\theta^2) - 3\theta nk})}{\Gamma(k)^n} \prod_{i=1}^n \frac{\Gamma(y_i+k)}{\Gamma(y_i+1)} dk d\theta} \tag{3.9}$$

**3.3. Inferensi Bayesian untuk Data Kasus Inferensi Bayesian untuk Data Kasus**

Hasil perhitungan pada subbagian 3.2 akan digunakan untuk menduga parameter dispersi,  $k$  diketahui dan parameter skala,  $\theta$  dengan menggunakan ketiga prior yang dipilih pada penelitian ini untuk data jumlah kecelakaan pada suatu ruas jalan tertentu di Depok.

Distribusi prior	$\hat{\theta}$	$\sigma$	Lebar selang
Beta(2,4)	0,0649	0,007489	0,029378
Gamma(2,4)	0,8911	0,1354	0,530768
Jeffrey	1,215	0,1916	0,751072.

**Tabel 1.** Tabel Estimasi dengan Berbagai Distribusi Prior

Dari Tabel 1 diperoleh informasi bahwa distribusi priorkonjugat (dalam kasus ini yaitu distribusi Beta) merupakan penduga bayes yang terbaik dibanding dua penduga dengan prior berbeda lainnya karena menghasilkan  $\sigma$  dan lebar selang yang lebih kecil.

**4. Kesimpulan**

Berdasarkan hasil dan pembahasan yang dilakukan maka diperoleh kesimpulan, yaitu:

- (1) Hasil yang diperoleh dari uji Kolmogorov-Smirnov dengan  $\alpha = 0,05$  adalah data jumlah kecelakaan lalu lintas di Depok tidak berdistribusi Poisson dan nilai variannya lebih besar dari meannya, sehingga data mengalami overdispersi. Oleh sebab itu data dimodelkan ulang dengan distribusi Binomial Negatif untuk mengatasi masalah tersebut.
- (2) Estimasi parameter  $\theta$  pada Distribusi Binomial Negatif dengan menggunakan distribusi Beta ( $\alpha, \beta$ ) menghasilkan distribusi posterior yaitu distribusi  $Beta\left(kn + \alpha, \sum_{i=1}^n y_i + \beta\right)$  dan  $\hat{\theta} = \frac{(kn+\alpha)}{\left(\sum_{i=1}^n y_i + \beta\right)}$ .
- (3) Estimasi parameter  $\theta$  pada Distribusi Binomial negatif dengan menggunakan distribusi Beta (2,4) sebagai prior konjugat menghasilkan  $\hat{\theta}$  sebesar 0,0649 dan  $\sigma$  sebesar 0,007489, distribusi Gamma (2,4) sebagai prior non konjugat menghasilkan  $\hat{\theta}$  sebesar 0,8911 dan  $\sigma$  sebesar 0,1354, prior Jeffrey sebagai priornon-informatif menghasilkan  $\hat{\theta}$  sebesar 1,215 dan  $\sigma$  sebesar 0,1916.
- (4) Dari hasil pendugaan parameter dengan distribusi Beta (2,4) sebagai prior konjugat, distribusi Gamma(2,4) sebagai prior non konjugat dan prior Jeffrey sebagai prior non informatif diperoleh informasi bahwa distribusi prior Beta menghasilkan dugaan standard error,  $\sigma$  dan lebar selang yang lebih kecil dibanding prior non konjugat dan noninformatif. Dapat disimpulkan bahwa distribusi prior Beta merupakan distribusi prior terbaik [3].

## 5. Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Dr. Maiyastri, Dr. Lyra Yulianti, Zulakmal, M.Si yang telah memberikan kritikan dan masukan sehingga makalah ini dapat diselesaikan dengan baik.

## Daftar Pustaka

- [1] Bain, L.J and Engelhardt, M.1992. *Introduction to Probability and Mathematical Statistics Second Edition*. Duxbury Press, California.
- [2] M.Berger, James O. 1993. *Statistical Decision Theory and Bayesian Analysis Second Edition*. Duxubury Press, California
- [3] Bolstad, W.M 2007. *Introduction to Bayesian Statistical, Second Edition*. A John Willey & Sons, Amerika.
- [4] Box, G. E. P and Tiao, G. C. 1973. *Bayesian Inference in Statistical Analysis*. Philippines: Addison-Wesley Publishing Company, Inc.
- [5] Casella, G and R.L. Berger. 2001. *Statistical Inference Second Edition*. Pasific Grove, California.
- [6] Fajar, M. 2015. *Analisis Kecelakaan Lalu Lintas Jalan Raya Di Kota Semarang Menggunakan Metode K-Means Clustering*. Skripsi S1. Fakultas Teknik Universitas Negeri Semarang.
- [7] Hilbe, J. 2011. *Negative Binomial Regression Second Edition*. Cambridge University Press.
- [8] Mulyaningsih, D. 2016. *Estimasi Parameter Distribusi Binomial Negatif Dengan Pendekatan Bayesian Menggunakan Monte Carlo Markov Chain Berda-*

sarkan *Algoritma Metropolis Hasting*. Skripsi S1. Fakultas Sains dan Teknologi Universitas Airlangga.

- [9] Norman L. Johnson, Samuel Kotz, and Adrianne W. Kemp, 1992.
- [10] Nzoufras, I. 2009. *Bayesian Modeling Using Winbugs*. John Wiley and Sons, New Jersey.
- [11] Spiegel, M.R, J.J. Schiller and R.A. Srinivasan. 2009. *Probability and Statistics Third Edition*. The McGraw-Hill Companies Inc., New York.
- [12] Siegel, S. 1985. *Statistik Nonparametrik*. PT Gramedia, Jakarta.
- [13] Shafira. 2011. *Penaksiran Parameter Distribusi Binomial Negatif Pada Kasus Overdispersi*. Skripsi S1. Fakultas Matematika Dan Pengetahuan Alam Program Studi Matematika Universitas Indonesia.
- [14] Walpole, R.E and R.H. Myers. 1995. *Probability and Statistics for Engineers Scientists Eighth Edition*. Pearson Education Inc, New of America.