

Distribution of the Difference of Two Independent Poisson Random Variables and Its Application to the Literate Population Data

Atina Ahdika

Program Studi Statistika, Universitas Islam Indonesia
Jalan Kaliurang Km 14,5 Sleman Yogyakarta
atina.a@uii.ac.id

ABSTRACT

Study about the distribution of the function of a random variable, in particular the distribution of the sum of two or more random variables, has been done widely. While studies on the distribution of the difference between two random variables still rarely done, moreover for discrete random variables. This work focuses on the derivation of the distribution of two independent Poisson random variables. The methods used to determine the distribution are method of moment generating function and probability mass function. The properties of the distribution are derived through the characteristics of the k -th moment; the first up to fourth moments in particular. Furthermore, the distribution of the difference of two independent Poisson random variables is compared to the Poisson Difference (PD) distribution of Alzaid and Omair (2010) or also known as Skellam distribution. Parameters of the distribution are estimated using method of moment and maximum likelihood estimation. Application to the literate population data is carried out to get a better understanding about the distribution.

Keywords: Poisson distribution, Poisson Difference distribution, moment generating function, probability mass function, method of moment, maximum likelihood estimation.

ABSTRAK

Kajian mengenai distribusi dari suatu fungsi peubah acak, khususnya mengenai distribusi dari jumlah dua atau lebih peubah acak, sudah banyak dilakukan. Sedangkan kajian mengenai distribusi dari selisih dua peubah acak masih jarang dilakukan, terlebih pada peubah acak-peubah acak diskrit. Penelitian ini berfokus pada penurunan distribusi selisih dua peubah acak Poisson yang saling bebas. Metode yang digunakan untuk menentukan distribusi tersebut adalah metode fungsi pembangkit momen dan metode fungsi massa peluang. Karakteristik-karakteristik dari distribusi diturunkan melalui sifat-sifat dari momen ke- k ; khususnya dari momen pertama hingga momen keempat. Selanjutnya, distribusi dari selisih dua peubah acak Poisson yang saling bebas ini akan dibandingkan dengan distribusi *Poisson Difference (PD)* yang diturunkan oleh Alzaid dan Omair (2010) atau dikenal juga dengan nama distribusi Skellam. Parameter-parameter dari distribusi diestimasi menggunakan metode momen dan maksimum likelihood. Aplikasi pada data penduduk melek huruf dilakukan untuk mendapatkan pemahaman yang lebih baik mengenai distribusi tersebut.

Kata Kunci: distribusi Poisson, distribusi *Poisson Difference*, fungsi pembangkit momen, fungsi massa peluang, metode momen, metode maksimum likelihood.

Pendahuluan

Salah satu konsep dasar yang harus dipahami dan berperan penting dalam ilmu statistika adalah distribusi peluang suatu data. Pemahaman mengenai hal

tersebut sangat penting untuk diketahui ketika seseorang harus berhadapan dengan serangkaian data dan menentukan langkah yang tepat untuk menangani data yang tersedia.

Berdasarkan ruang sampelnya, distribusi peluang suatu data dibagi menjadi dua, yaitu distribusi kontinu dan distribusi diskrit. Ada banyak kajian yang berkaitan dengan distribusi-distribusi tersebut, salah satunya adalah menentukan bentuk distribusi dari suatu peubah acak bila ia dikenai suatu fungsi tertentu. Fungsi dari peubah acak yang paling sederhana dan paling masuk akal (bermakna) ketika diterapkan dalam suatu data adalah penjumlahan atau pengurangan dari dua peubah acak. Secara khusus, kajian mengenai distribusi jumlah dua atau lebih peubah acak sudah banyak dipelajari dan mudah diturunkan. Sedangkan distribusi selisih peubah acak masih jarang dilakukan. Pada umumnya, kejadian dalam kehidupan sehari-hari berkaitan dengan kejadian yang terhitung (diskrit). Berdasarkan kedua hal tersebut, maka dalam penelitian ini akan dilakukan formulasi terhadap distribusi dari selisih dua peubah acak diskrit, khususnya selisih dua peubah acak Poisson yang saling bebas. Selain itu akan diturunkan pula karakteristik-karakteristik distribusi, penaksiran parameter, serta aplikasi dari distribusi tersebut pada data real.

Pada penelitian terdahulu, J.O. Irwin (1937) telah membahas mengenai distribusi dari selisih dua peubah acak

Poisson saling bebas dengan parameter yang sama. Kemudian terdapat pula pembahasan mengenai distribusi serupa yang dikenal dengan nama distribusi *Poisson Difference (PD)* yang diturunkan oleh Alzaid dan Omair (2010) atau dikenal juga dengan nama distribusi Skellam, di mana distribusi tersebut melibatkan peubah acak ketiga di dalamnya. Untuk itu, di dalam paper ini akan dibahas juga kaitan antara distribusi selisih dua peubah acak Poisson yang saling bebas dengan distribusi PD tersebut.

Terdapat beberapa aplikasi dari distribusi *Poisson Difference* ini; Karlis dan Ntzoufras (2000) mengaplikasikan distribusi tersebut untuk memodelkan perbedaan banyaknya gol dalam permainan sepak bola, Hwang *et al* (2007) mengukur perbedaan intensitas piksel dalam kamera. Pada paper ini akan dilakukan aplikasi dari distribusi selisih dua peubah acak Poisson saling bebas pada persentase melek huruf di atas 15 tahun berdasarkan golongan umur dan daerah tempat tinggal untuk melihat peluang kenaikan serta penurunan persentasenya tiap tahun.

Tujuan Penelitian

1. Memformulasikan distribusi selisih dua peubah acak Poisson yang saling bebas dan menurunkan karakteristik-karakteristiknya.
2. Mempelajari kaitan distribusi tersebut dengan distribusi PD/distribusi Skellam.
3. Melakukan estimasi parameter distribusi selisih dua peubah acak Poisson yang saling bebas menggunakan metode momen dan maksimum likelihood.
4. Mengaplikasikan konsep distribusi selisih dua peubah acak Poisson yang saling bebas ke dalam data real, yaitu data persentase melek huruf penduduk di atas 15 tahun berdasarkan golongan umur dan daerah tempat tinggal.

Landasan Teori

Kebebasan Peubah Acak

Peubah acak merupakan fungsi yang memetakan anggota ruang sampel ke himpunan bilangan riil: $X : S \mapsto R$

Dua peubah acak X_1 dan X_2 dikatakan saling bebas jika dan hanya jika:

$$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1) \cdot P(X_2 = x_2)$$

Distribution of the Difference of Two Independent Poisson Random Variables and Its Application to the Literate Population Data
(Atina Ahdika)

Distribusi Poisson

Peubah acak X dikatakan berdistribusi Poisson jika X menyatakan banyaknya hasil (sukses) pada selang waktu tertentu (daerah tertentu). Sifat-sifat distribusi Poisson adalah:

1. Banyaknya hasil yang terjadi dalam suatu selang waktu atau daerah tertentu tidak terpengaruh (bebas dari) apa yang terjadi pada selang waktu atau daerah lain yang terpisah.
2. Peluang terjadinya suatu hasil (tunggal) dalam selang waktu yang amat pendek atau dalam daerah yang kecil sebanding dengan panjang selang waktu atau besarnya daerah dan tidak bergantung pada banyaknya hasil yang terjadi di luar selang waktu atau daerah tersebut.
3. Peluang terjadinya lebih dari satu hasil dalam selang waktu yang pendek atau daerah yang sempit tersebut dapat diabaikan.

Suatu peubah acak X berdistribusi Poisson dengan parameter θ atau dinotasikan dengan $X \sim POI(\theta)$ jika fungsi massa peluangnya adalah

$$P(X = x) = e^{-\theta} \frac{\theta^x}{x!}, \quad x = 0, 1, 2, \dots, \theta > 0$$

Sedangkan fungsi distribusi kumulatifnya adalah

$$P(X \leq x) = \sum_{x=0}^k e^{-\theta} \frac{\theta^x}{x!}$$

Berikut adalah karakteristik dari distribusi Poisson:

1. Nilai mean dan variansinya sama,yaitu $E(X)=Var(X)=\theta$
2. Nilai skewness (kemiringan data) lebih besar dari nol, yang berarti bahwa data menceng ke kanan.
3. Nilai kurtosis (kelancipan data) lebih besar dari tiga, yang berarti bahwa nilai data-datanya mengumpul di sekitar meannya.
4. Fungsi pembangkit peluangnya adalah $M_X(t) = e^{\theta(e^t - 1)}$

Distribusi Poisson Difference (PD)

Alzaid dan Omair (2010) mendefinisikan distribusi Poisson Difference sebagai berikut

Definisi 1. Misalkan terdapat pasangan peubah acak (X, Y) yang bisa dituliskan sebagai $X = W_1 + W_3$ dan $Y = W_2 + W_3$ dengan $W_1 \sim POI(\theta_1)$ saling bebas dengan $W_2 \sim POI(\theta_2)$ dan W_3 sebarang distribusi, maka fungsi massa peluang dari $Z = X - Y$ diberikan

$$P(Z = z) = e^{-(\theta_1 + \theta_2)} \left(\frac{\theta_1}{\theta_2} \right)^{z/2} I_z(2\sqrt{\theta_1 \theta_2})$$

$$z = \dots, -1, 0, 1, \dots$$

dengan

$$I_y(x) = \left(\frac{x}{2} \right)^y \sum_{k=0}^{\infty} \frac{\left(\frac{x^2}{4} \right)^k}{k!(y+k)!}$$

$I_y(x)$ merupakan fungsi Bessel jenis pertama termodifikasi.

Pembahasan

Distribusi Selisih Dua Peubah Acak Poisson yang Saling Bebas

Pada bagian ini akan diformulasikan mengenai distribusi selisih dari dua peubah acak Poisson yang saling bebas.

Misalkan $X_1 \sim POI(\theta_1)$ saling bebas dengan $X_2 \sim POI(\theta_2)$. Misalkan $Z = X_1 - X_2$, maka distribusi dari peubah acak Z dapat ditentukan dengan menggunakan kedua teknik berikut:

Teknik Fungsi Pembangkit Momen

Fungsi pembangkit momen dari Z adalah sebagai berikut

$$\begin{aligned} M_Z(t) &= E(e^{tZ}) = E(e^{t(X_1 - X_2)}) \\ &= E(e^{tX_1} \cdot e^{-tX_2}) = E(e^{tX_1}) \cdot E(e^{-tX_2}) \\ &= e^{-(\theta_1 + \theta_2) + \theta_1 e^t + \theta_2 e^{-t}} \end{aligned}$$

Bentuk fungsi pembangkit momen tersebut tidak sama dengan fungsi pembangkit momen dari distribusi Poisson.

Teknik Fungsi Massa Peluang

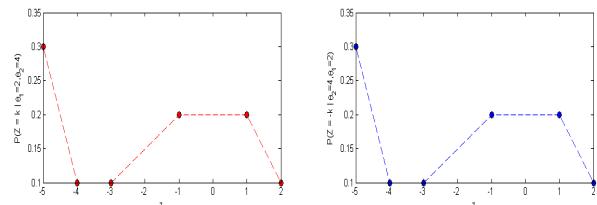
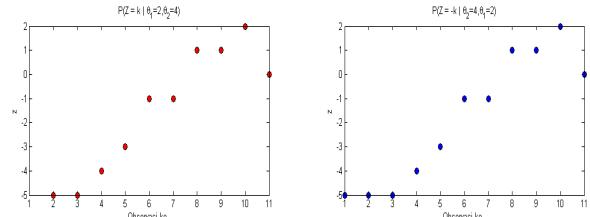
Fungsi massa peluang dari peubah acak Z diperoleh dari penurunan berikut

$$\begin{aligned} P(Z = k) &= P(X_1 - X_2 = k) \\ &= \sum_{n=0}^{\infty} P(X_1 - X_2 = k \mid X_2 = n) P(X_2 = n) \\ &= \sum_{n=0}^{\infty} f_{X_1}(n+k) f_{X_2}(n) \\ &= \sum_{n=0}^{\infty} e^{-\theta_1} \frac{\theta_1^{n+k}}{(n+k)!} e^{-\theta_2} \frac{\theta_2^n}{n!} \\ &= e^{-(\theta_1+\theta_2)} \sum_{n=0}^{\infty} \frac{\theta_1^{n+k} \theta_2^n}{(n+k)! n!} \end{aligned}$$

Jika $Z = -k$, maka fungsi peluangnya adalah:

$$\begin{aligned} P(Z = -k) &= P(X_1 - X_2 = -k) \\ &= \sum_{n=0}^{\infty} P(X_1 - X_2 = -k \mid X_2 = n) P(X_2 = n) \\ &= \sum_{n=0}^{\infty} f_{X_1}(n-k) f_{X_2}(n) \\ &= \sum_{n=0}^{\infty} e^{-\theta_1} \frac{\theta_1^{n-k}}{(n-k)!} e^{-\theta_2} \frac{\theta_2^n}{n!} \\ &= e^{-(\theta_1+\theta_2)} \sum_{n=0}^{\infty} \frac{\theta_1^{n-k} \theta_2^n}{n! (n-k)!} \end{aligned}$$

Selanjutnya, terdapat sifat $P(Z = k \mid \theta_1, \theta_2) = P(Z = -k \mid \theta_2, \theta_1)$ yang ditunjukkan oleh hasil simulasi berikut:



Gambar 1. Simulasi Numerik

Sifat $P(Z = k \mid \theta_1, \theta_2) = P(Z = -k \mid \theta_2, \theta_1)$, sehingga:

$$\frac{P(Z = k \mid \theta_1, \theta_2)}{P(Z = -k \mid \theta_1, \theta_2)} = \frac{P(Z = -k \mid \theta_2, \theta_1)}{P(Z = -k \mid \theta_1, \theta_2)} = \left(\frac{\theta_1}{\theta_2}\right)^k$$

Berdasarkan hubungan tersebut, maka fungsi massa peluang dari distribusi selisih dua peubah acak Poisson yang saling bebas dapat diturunkan dan diperoleh yaitu sebagai berikut:

$$\begin{aligned}
 P(Z=k | \theta_1, \theta_2) &= P(Z=-k | \theta_1, \theta_2) \left(\frac{\theta_1}{\theta_2} \right)^k \\
 &= e^{-(\theta_1+\theta_2)} \sum_{n=0}^{\infty} \frac{\theta_1^{n-k} \theta_2^n}{n!(n-k)!} \left(\frac{\theta_1}{\theta_2} \right)^k \\
 &= e^{-(\theta_1+\theta_2)} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{k}{2}} \sum_{n=0}^{\infty} \frac{\theta_1^{n-k} \theta_2^n}{n!(n-k)!} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{k}{2}} \\
 &= e^{-(\theta_1+\theta_2)} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{k}{2}} \sum_{n=0}^{\infty} \frac{\theta_1^n \theta_2^{n+k}}{(n+k)! n!} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{k}{2}} \\
 &= e^{-(\theta_1+\theta_2)} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{k}{2}} \sum_{n=0}^{\infty} \frac{((\theta_1 \theta_2)^{1/2})^{2n+k}}{n!(n+k)!} \\
 &= e^{-(\theta_1+\theta_2)} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{k}{2}} I_k(2\sqrt{\theta_1 \theta_2})
 \end{aligned}$$

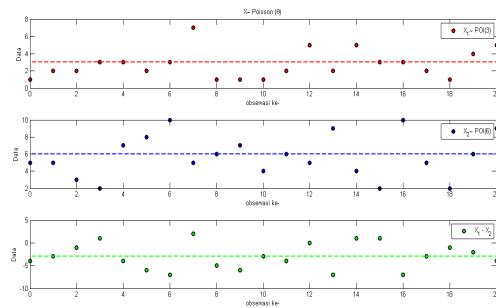
dengan $I_z(x)$ adalah fungsi Bessel jenis pertama termodifikasi.

Fungsi massa peluang tersebut sama dengan fungsi massa peluang distribusi *Poisson Difference* pada Definisi 1. Perbedaan dan kaitan antara distribusi selisih dua peubah acak Poisson saling bebas dengan distribusi *Poisson Difference* ini akan dibahas pada bab selanjutnya.

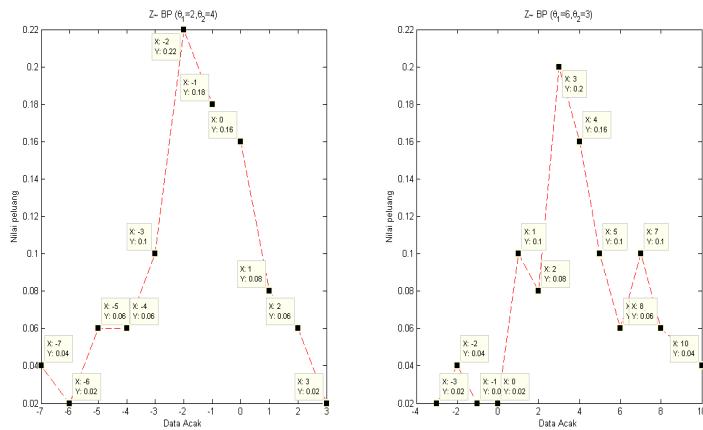
Jika dilihat dari bentuknya, fungsi peluang tersebut memiliki bentuk yang berbeda dengan fungsi peluang dari distribusi Poisson.

Berdasarkan sifat unik dari suatu distribusi dan hasil perhitungan menggunakan dua teknik tersebut, maka dapat disimpulkan bahwa distribusi selisih dua peubah acak Poisson yang saling bebas bukan merupakan distribusi Poisson. Selanjutnya kita akan menyebut distribusi tersebut sebagai distribusi Beda Poisson saling bebas dengan parameter (θ_1, θ_2) .

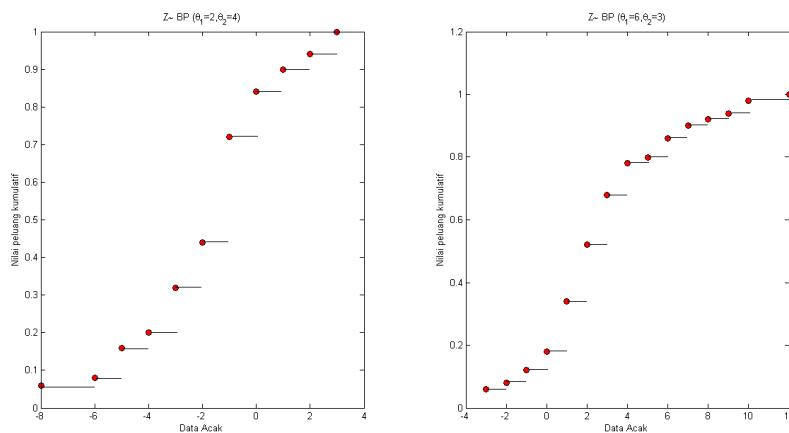
Berikut adalah grafik sebaran data, fungsi massa peluang, serta fungsi distribusi kumulatif dari distribusi Beda Poisson saling bebas.



Gambar 2. Sebaran Data Berdistribusi Beda Poisson Saling Bebas



Gambar 3. Fungsi Massa Peluang Distribusi Beda Poisson Saling Bebas dengan $\theta_1 < \theta_2$ dan $\theta_1 > \theta_2$



Gambar 4. Fungsi Distribusi Kumulatif Beda Poisson Saling Bebas dengan $\theta_1 < \theta_2$ dan $\theta_1 > \theta_2$

Gambar 2 menunjukkan bahwa sebaran data berdistribusi Beda Poisson saling bebas berada di sekitar nilai selisih parameternya, yaitu $\theta_1 - \theta_2$. Berbeda dengan distribusi Poisson, pada distribusi Beda Poisson saling bebas terdapat juga kemungkinan data bernilai negatif. Selanjutnya pada Gambar 3, ditampilkan pola fungsi massa peluang dari distribusi Beda Poisson saling bebas dengan $\theta_1 < \theta_2$

dan $\theta_1 > \theta_2$. Berdasarkan grafik tersebut dapat dilihat bahwa nilai peluang terbesarnya (modus) berada pada data dengan nilai di sekitar $\theta_1 - \theta_2$. Selain itu, dari Gambar 4 dapat dilihat bahwa kenaikan nilai peluang terbesar berada pada data yang nilainya di sekitar $\theta_1 - \theta_2$. Dengan demikian, dapat disimpulkan bahwa distribusi Beda Poisson saling bebas bersifat unimodal.

Selanjutnya, akan ditentukan karakteristik-karakteristik dari distribusi Beda Poisson saling bebas melalui sifat-sifat momen ke- k dari distribusi tersebut. Momen ke- k dari distribusi Beda Poisson saling bebas diperoleh dengan menggunakan turunan ke- k dari fungsi pembangkit momen $X_1 - X_2$ pada saat $t = 0$, yaitu

$$M^k|_{X_1-X_2}(t)|_{t=0} = \frac{d^k}{dt^k}|_{t=0} \left(e^{-(\theta_1+\theta_2)+\theta_1 t + \theta_2 e^{-t}} \right)$$

Dengan teknik tersebut, diperoleh momen pertama hingga momen keempat sebagai berikut:

1. Momen ke-1

$$E(X_1 - X_2) = M'|_{X_1-X_2}(0) = \theta_1 - \theta_2$$

2. Momen ke-2

$$\begin{aligned} E((X_1 - X_2)^2) &= M''|_{X_1-X_2}(0) \\ &= (\theta_1 + \theta_2) + (\theta_1 - \theta_2)^2 \end{aligned}$$

3. Momen ke-3

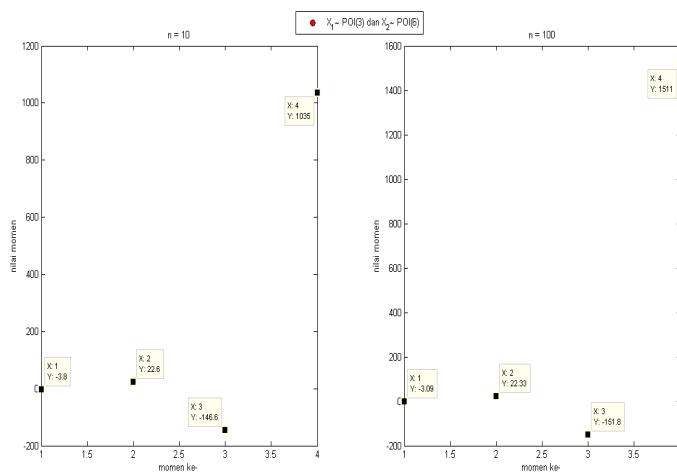
$$\begin{aligned} E((X_1 - X_2)^3) &= M'''|_{X_1-X_2}(0) \\ &= (\theta_1 - \theta_2) + 3(\theta_1^2 - \theta_2^2) \\ &\quad + (\theta_1 - \theta_2)^3 \end{aligned}$$

4. Momen ke-4

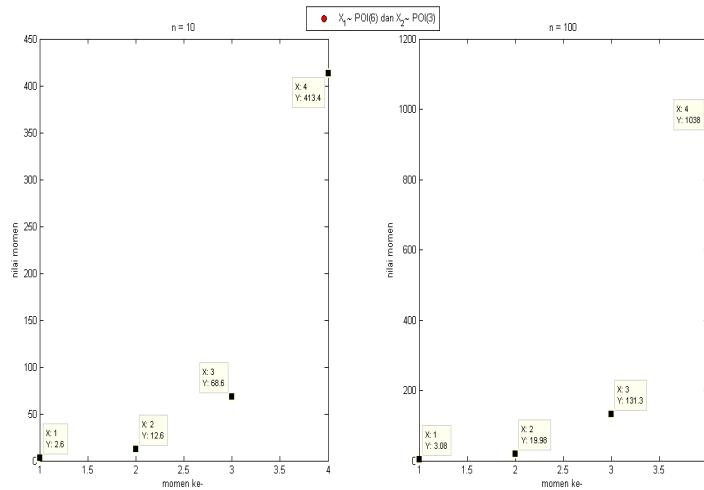
$$\begin{aligned} E((X_1 - X_2)^4) &= M^{iv}|_{X_1-X_2}(0) \\ &= (\theta_1 + \theta_2) + 6(\theta_1^2 + \theta_2^2) \\ &\quad + 6(\theta_1^2 - \theta_2^2)(\theta_1 - \theta_2) + (\theta_1 - \theta_2)^2 \\ &\quad + (\theta_1 - \theta_2)^4 \end{aligned}$$

Secara visual, momen ke- k dari distribusi Beda Poisson saling bebas ditampilkan oleh Gambar 5 berikut.

Berdasarkan Gambar 5 dapat dilihat bahwa untuk $\theta_1 > \theta_2$ momennya akan selalu bernilai positif, sedangkan untuk $\theta_1 < \theta_2$ memungkinkan nilai momen yang negatif.



(a)



(b)

Gambar 5. Momen ke- k Distribusi Beda Poisson Saling Bebas; (a) $\theta_1 < \theta_2$, (b) $\theta_1 > \theta_2$

Dari keempat momen tersebut diperoleh karakteristik dari distribusi Beda Poisson saling bebas yang meliputi mean, variansi, skewness, serta kurtosis dari $Z = X_1 - X_2$ yaitu

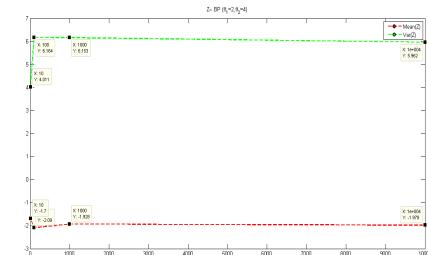
1. Mean

$$E(Z) = E(X_1 - X_2) = \theta_1 - \theta_2$$

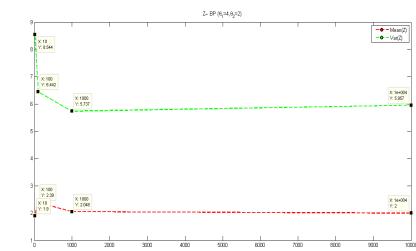
2. Variansi

$$\begin{aligned} Var(Z) &= Var(X_1 - X_2) \\ &= E((X_1 - X_2)^2) - (E(X_1 - X_2))^2 \\ &= \theta_1 + \theta_2 \end{aligned}$$

Berbeda dengan distribusi Poisson, mean dan variansi dari distribusi Beda Poisson saling bebas tidak memiliki nilai yang sama. Secara visual, mean dan variansi dari distribusi Beda Poisson saling bebas ditunjukkan oleh Gambar 8.



(a)



(b)

Gambar 6. Mean vs Variansi; (a) $\theta_1 < \theta_2$, (b) $\theta_1 > \theta_2$

Gambar 6 merupakan plot mean dan variansi distribusi Beda Poisson saling bebas dengan parameter $\theta_1 < \theta_2$ dan $\theta_1 > \theta_2$ menggunakan sampel acak sebanyak $n = 10, 100, 1000, 10000$. Dari plot

bisa dilihat bahwa baik mean maupun variansi sampel distribusi Beda Poisson saling bebas akan semakin mendekati nilai parameternya untuk n yang semakin membesar. Nilai meannya akan selalu di bawah variansinya baik $\theta_1 < \theta_2$ maupun $\theta_1 > \theta_2$ (*underdispersion*).

3. Skewness

$$\beta_1 = E\left(\left(\frac{Z - \mu_Z}{\sigma}\right)^3\right) = \frac{\theta_1 - \theta_2}{(\theta_1 + \theta_2)^{\frac{3}{2}}}$$

4. Kurtosis

$$\beta_1 = E\left(\left(\frac{Z - \mu_Z}{\sigma}\right)^4\right) = \frac{1+3(\theta_1 + \theta_2)}{\theta_1 + \theta_2}$$

Nilai skewness dan kurtosis distribusi Beda Poisson saling bebas antara suatu sampel dengan sampel yang lain berbeda-beda tergantung dengan besarnya θ_1 dan θ_2 .

Kaitan antara Distribusi Beda Poisson Saling Bebas dengan Distribusi Poisson Difference

Berdasarkan Definisi 1, terdapat peubah acak ketiga pada distribusi *Poisson Difference* yaitu W_3 dengan sebarang distribusi, sehingga dalam hal ini kita tidak bisa mengatakan bahwa peubah acak X dan peubah acak Y saling bebas.

Untuk peubah acak Poisson yang tidak saling bebas, misalkan $X \sim POI(\theta_1)$ dan $Y \sim POI(\theta_2)$, R. Vernic (1997) menyatakan fungsi peluang bivariatnya, yaitu

$$P(X = x, Y = y) = e^{-(\theta_1 + \theta_2 + \theta_3)} \frac{\theta_1^x}{x!} \frac{\theta_2^y}{y!} \sum_{k=0}^{\min(x,y)} \binom{x}{k} \binom{y}{k} k! \left(\frac{\theta_3}{\theta_1 \theta_2}\right)^k$$

Sedangkan fungsi peluang bersama untuk dua peubah acak Poisson yang saling bebas adalah

$$P(X = x, Y = y) = e^{-(\theta_1 + \theta_2)} \frac{\theta_1^x}{x!} \frac{\theta_2^y}{y!}$$

Dari fungsi peluang bivariat Poisson dan fungsi peluang bersama dua peubah acak Poisson yang saling bebas, kita bisa melihat peranan dari korelasi di antara kedua peubah acak tersebut (dinotasikan sebagai $f(\cdot)$) yaitu sebagai berikut

$$P(X = x, Y = y) = P(X = x)P(Y = y)f(\cdot) \\ = e^{-(\theta_1 + \theta_2)} \frac{\theta_1^x}{x!} \frac{\theta_2^y}{y!} f(\cdot)$$

Dengan demikian nilai dari $f(\cdot)$ dapat diperoleh, yaitu

$$f(\cdot) = e^{-\theta_3} \sum_{k=0}^{\min(x,y)} \binom{x}{k} \binom{y}{k} k! \left(\frac{\theta_3}{\theta_1 \theta_2}\right)^k$$

Ambil salah satu kasus, misalkan $\min(x, y) = x$, maka fungsi $f(\cdot)$ apabila dijabarkan akan menjadi

$$\begin{aligned} f(\cdot) &= e^{-\theta_3} \sum_{k=0}^{\min(x,y)} \binom{x}{k} \binom{y}{k} k! \left(\frac{\theta_3}{\theta_1 \theta_2} \right)^k \\ &= e^{-\theta_3} \sum_{k=0}^x \frac{x!}{k!(x-k)!} \frac{y!}{k!(y-k)!} k! \left(\frac{\theta_3}{\theta_1 \theta_2} \right)^k \\ &= e^{-\theta_3} \sum_{k=0}^x \frac{x!}{k!(x-k)!} \frac{y!}{(y-k)!} \left(\frac{\theta_3}{\theta_1 \theta_2} \right)^k \\ &= e^{-\theta_3} \left(1 + xy \left(\frac{\theta_3}{\theta_1 \theta_2} \right) + \dots + \frac{y!}{(y-x)!} \left(\frac{\theta_3}{\theta_1 \theta_2} \right)^x \right) \end{aligned}$$

Pada distribusi Bivariat Poisson, nilai korelasinya adalah

$$\rho_{X,Y} = \frac{\theta_3}{\sqrt{(\theta_1 + \theta_3)(\theta_2 + \theta_3)}}$$

Jika nilai $\theta_3 = 0$ maka $\rho_{X,Y} = 0$ dan $f(\cdot) = 1$. Ini berarti bahwa jika korelasinya nol, maka fungsi Bivariat Poisson merupakan fungsi peluang bersama dari dua peubah acak Poisson yang saling bebas.

Berdasarkan hal tersebut, maka distribusi Beda Poisson saling bebas merupakan kasus khusus dari distribusi *Poisson Difference* di mana peubah acak ketiganya, yaitu W_3 , berdistribusi sebarang dengan $\theta_3 = 0$.

Penaksiran Parameter

Pada tulisan ini, metode yang digunakan untuk mengestimasi parameter distribusi Beda Poisson saling bebas adalah metode momen dan maksimum likelihood.

Metode Momen

Misalkan Z_1, Z_2, \dots, Z_n adalah sampel acak dari distribusi Beda Poisson saling bebas yang identik dan independen, maka

$$\begin{aligned} E(Z) &= \bar{Z} = \theta_1 - \theta_2 \\ Var(Z) &= S^2 = \theta_1 + \theta_2 \end{aligned}$$

Dengan mensubstitusikan kedua persamaan tersebut diperoleh nilai penaksir untuk θ_1 dan θ_2 yaitu

$$\begin{aligned} \hat{\theta}_{1MM} &= \frac{1}{2}(S^2 + \bar{Z}) \\ \hat{\theta}_{2MM} &= \frac{1}{2}(S^2 - \bar{Z}) \end{aligned}$$

Metode Maksimum Likelihood

Misalkan Z_1, Z_2, \dots, Z_n adalah sampel acak dari distribusi Beda Poisson saling bebas yang identik dan independen, maka fungsi likelihoodnya adalah

$$\begin{aligned} L(\theta_1, \theta_2) &= \prod_{i=1}^n P(Z_i = z_i) \\ &= \prod_{i=1}^n \left(e^{-\theta_1 - \theta_2} \left(\frac{\theta_1}{\theta_2} \right)^{\frac{z_i}{2}} I_{z_i} \left(2\sqrt{\theta_1 \theta_2} \right) \right) \end{aligned}$$

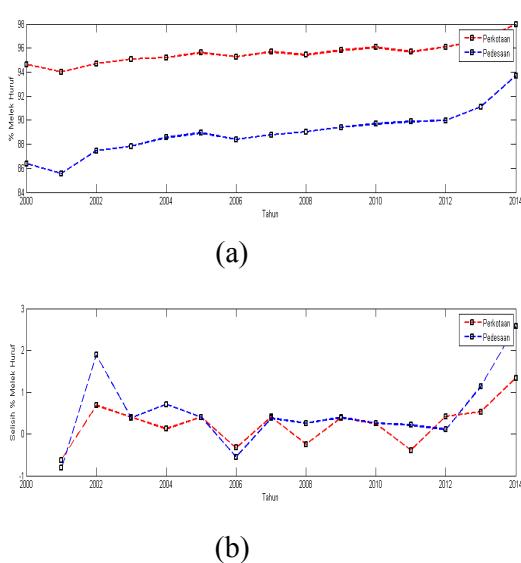
Selanjutnya fungsi likelihood tersebut di- \ln -kan sebagai berikut

$$\begin{aligned}\ln L(\theta_1, \theta_2) &= \ln \prod_{i=1}^n \left(e^{-\theta_1 - \theta_2} \left(\frac{\theta_1}{\theta_2} \right)^{z_i} I_{z_i}(2\sqrt{\theta_1 \theta_2}) \right) \\ &= \sum_{i=1}^n -\theta_1 - \theta_2 + \sum_{i=1}^n \frac{z_i}{2} (\ln \theta_1 - \ln \theta_2) \\ &\quad + \sum_{i=1}^n \ln I_{z_i}(2\sqrt{\theta_1 \theta_2})\end{aligned}$$

Kemudian fungsi $\ln L(\theta_1, \theta_2)$ tersebut diturunkan terhadap θ_1 dan θ_2 sehingga diperoleh penaksiran parameter untuk kedua parameter tersebut yaitu

$$\hat{\theta}_{1MLE} = \hat{\theta}_{2MLE} + \bar{z}$$

Aplikasi



Gambar 7. Persentase dan Selisih: Persentase Melek Huruf (a), Penduduk di Atas 15 Tahun (b)

Distribusi Beda Poisson saling bebas ini diaplikasikan pada data persentase penduduk berumur 15 tahun ke atas yang melek huruf menurut golongan umur dan daerah tempat tinggal dari tahun 2000 sampai 2014. Data ini diperoleh dari repositori data Badan Pusat Statistik (BPS) Indonesia.

Gambar 7 berikut menunjukkan plot data persentase dan selisih persentase penduduk melek huruf pada waktu t dan $t-1$. Selisih persentase tersebut menunjukkan kenaikan maupun penurunan persentase banyaknya penduduk melek huruf dari tahun ke tahun.

Tabel 1. Statistik Deskriptif Selisih Persentase Melek Huruf Penduduk di Atas 15 Tahun

Daerah	Mean	Variansi	Maks Selisih %	Min Selisih %
Kota	0.2379	0.2579	1.34	-0.63
Desa	0.5221	0.7602	2.57	-0.81

Statistik deskriptif dari data selisih persentase melek huruf disajikan dalam Tabel 1.

Berdasarkan statistik deskriptif tersebut, terlihat bahwa secara umum di pedesaan rata-rata kenaikan persentase melek huruf dari tahun ke tahun lebih besar daripada di perkotaan. Namun penurunan persentase terbesar juga

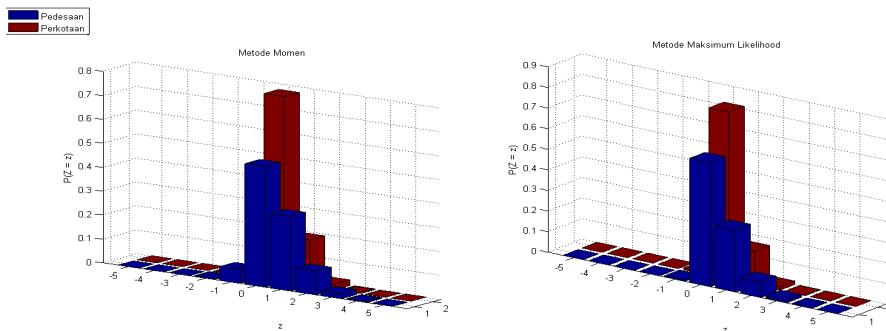
terjadi di pedesaan, tepatnya pada tahun 2006.

Selanjutnya, parameter dari distribusi data tersebut diestimasi menggunakan metode momen dan maksimum likelihood dengan bantuan *software* MATLAB dan diperoleh hasil

Tabel 2. Estimasi Parameter Distribusi Selisih Persentase Melek Huruf Penduduk di Atas 15 Tahun

Daerah	$\hat{\theta}_{1MM}$	$\hat{\theta}_{2MM}$	$\hat{\theta}_{1MLE}$	$\hat{\theta}_{2MLE}$
Kota	0.2479	0.01	0.201	0.01
Desa	0.6412	0.119	0.501	0.01

Dengan menggunakan estimasi parameter tersebut kemudian dihitung peluang selisih persentase melek huruf penduduk di atas 15 tahun untuk daerah perkotaan dan pedesaan.



Gambar 8. Sebaran Peluang Selisih Persentase Melek Huruf Penduduk di Atas 15 Tahun di Perkotaan dan Pedesaan

Berdasarkan Gambar 8, dapat dilihat bahwa besar peluang untuk selisih persentase melek huruf penduduk di atas 15 tahun baik menggunakan metode momen maupun maksimum likelihood memberikan hasil sebaran peluang yang hampir sama. Peluang tertinggi terjadi ketika kenaikan ataupun penurunan persentase melek huruf berada di sekitar nilai nol, artinya dari tahun ke tahun mulai tahun 2000 sampai 2014 tidak terjadi kenaikan maupun penurunan

persentase melek huruf yang cukup berarti.

Namun, secara umum, persentase kenaikan melek huruf memiliki peluang yang lebih besar dibandingkan penurunannya. Hal ini ditunjukkan dengan lebih tingginya peluang dari z yang bernilai positif daripada z negatif. Dalam hal ini peubah acak z menyatakan besar selisih persentase penduduk melek huruf. Selain itu, bila dibandingkan antara perkotaan dengan pedesaan, ternyata peluang kenaikan

persentase melek huruf di pedesaan lebih besar daripada di perkotaan. Hal ini mungkin saja terjadi karena saat ini pembangunan sumber daya manusia di pedesaan lebih diprioritaskan daripada di perkotaan.

Kesimpulan

Berdasarkan pembahasan dan analisis yang dilakukan, dapat disimpulkan beberapa hal sebagai berikut

1. Distribusi selisih dua peubah acak Poisson yang saling bebas tidaklah berdistribusi Poisson namun memiliki jenis distribusi lain yang dalam paper ini dinamakan distribusi Beda Poisson saling bebas.
2. Distribusi Beda Poisson saling bebas merupakan kasus khusus dari distribusi *Difference Poisson* (PD) /distribusi Skellam di mana parameter dari peubah acak ketiganya bernilai nol, $\theta_3 = 0$.
3. Estimasi parameter dapat dilakukan dengan menggunakan metode momen dan maksimum likelihood dan penyelesaiannya dapat dilakukan secara numerik dengan bantuan MATLAB.
4. Distribusi Beda Poisson saling bebas ini diaplikasikan pada data besarnya

persentase melek huruf penduduk di atas 15 tahun di perkotaan dan pedesaan. Berdasarkan hasil estimasi, diperoleh kesimpulan bahwa penaksiran parameter dan peluang selisih persentase melek huruf baik menggunakan metode momen maupun maksimum likelihood memberikan hasil yang sama baiknya.

Pustaka

- Alzaid, A., Omair, M. 2010. *On The Poisson Difference Distribution Inference and Applications*. Bull. Malays. Math. Sci. Soc. (2) **33**(1), 17-45.
- Badan Pusat Statistik (BPS). <http://www.bps.go.id/Subjek/view/id/28#subjekViewTab3|accordion-daftar-subjek1>. Diakses pada 15 Desember 2015.
- Irwin, M. 2006. *Moment Generating Function. Statistics 110*. Harvard : Summer 2006.
- Karlis, D., Ntzoufras, I. 2000. *Discrete Distribution with Applications in Sports*. Department of Statistics Athena University of Economics and Business, Technical Report.
- Karlis, D., Ntzoufras, I. 2006. *Bayesian Analysis of The Differences of Count Data*. Wiley InterScience.
- Ross, S. 2007. *Introduction to Probability Models, 9th Edition*. Berkeley California: University of California. Hal. 32, 51, 53, 66, 307.

Vernic, R. 1997. *On The Bivariate Generalized Poisson Distribution.* ASTIN Bulletin, Vol. 27, Issue 01, May 1997, pp 23-32.

Walpole, R., Myers, R. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan, Edisi ke-4 (Terjemah).* Bandung: Penerbit ITB.

Hwang, Y., et al. 2007. *Sensor Noise Modeling Using the Skellam Distribution: Application to the Color Edge Detection.* IEEE Conference on Computer Vision and Pattern Recognition, pp 1 – 8.