

Analisis Korespondensi Jumlah Penderita Penyakit Menular di Kota Bengkulu

Priska Julianti¹, Sigit Nugroho² dan Jose Rizal²

¹ Mahasiswa Jurusan Matematika FMIPA Universitas Bengkulu

² Staf Pengajar Jurusan Matematika FMIPA Universitas Bengkulu

ABSTRAK

Analisis korespondensi digunakan untuk menganalisis baris dan kolom secara bersamaan dari suatu tabel kontingensi dua arah dalam ruang vektor berdimensi dua. Keunggulan analisis ini adalah menampilkan hasil analisisnya dalam bentuk grafik yang mudah untuk dipahami dan diinterpretasikan. Data yang digunakan berbentuk kategori dan diskrit.

Tujuan penelitian ini adalah menerapkan teori Analisis Korespondensi guna mendapatkan gambaran beberapa penyakit menular seperti TB paru, PNEUMONIA, rabies, DBD, diare, dan malaria di Kecamatan yang ada di Kota Bengkulu, dengan pendekatan jarak Kai Kuadrat. Hasil Analisis Korespondensi ini dipetakan dalam ruang vektor berdimensi dua dengan mereduksi dimensi data berdasarkan nilai *inersianya*. Kecamatan yang memiliki kecenderungan terhadap jenis penyakit TB Paru adalah Sungai Serut dengan jarak Kai Kuadrat 0.028, Kecamatan yang cenderung terhadap jenis penyakit PNEUMONIA adalah Gading Cempaka dengan jarak Kai Kuadrat sebesar 0.003, dan jenis penyakit DBD cenderung banyak ditemukan di Kecamatan Ratu Agung dengan jarak Kai Kuadratnya 0.042. Dalam pereduksian dimensi informasi yang hilang sebesar $L=7.03\%$.

Kata kunci : Analisis Korespondensi, Tabel kontingensi dua arah, nilai *inersia*, penyakit menular.

PENDAHULUAN

Penyakit menular merupakan suatu wabah penyakit yang dapat menulari banyak orang pada kurun waktu yang relatif singkat atau cepat. Dengan banyaknya jenis penyakit menular saat ini, tentunya banyak faktor yang mempengaruhi penyebarannya. Sehingga diduga kondisi suatu daerah atau kawasan pemukiman mempengaruhi tingkat penyebaran penyakit menular.

Kota Bengkulu memiliki 8 kecamatan yaitu kecamatan Muara Bangkahulu, Gading Cempaka, Teluk Segara, Selebar, Kampung Melayu, Ratu Agung, Ratu Samban dan Sungai Serut. Informasi mengenai penyakit menular yaitu DBD, diare, malaria, PNEUMONIA, rabies, TB paru dan jumlah penderita penyakit menular untuk masing-masing kecamatan telah tersedia di Dinas Kesehatan Kota Bengkulu. Tulisan ini berisi penelitian mengenai analisis statistik dalam melihat hubungan antara lingkungan dan jenis penyakit. Khususnya akan dilihat kecamatan mana yang paling dominan terhadap penyakit menular dengan menggunakan Analisis Korespondensi.

TINJAUAN PUSTAKA

Menurut Greenacre dalam Sudarsono dan Latra (2005) *Analyses des Correspondances* atau Analisis Korespondensi adalah teknik analisis data yang memperagakan baris dan kolom secara bersamaan dari suatu tabel kontingensi dua arah dalam ruang vektor berdimensi dua. Dalam Analisis Korespondensi ada beberapa asumsi yang harus dipenuhi, yaitu

1. Ukuran jarak Kai Kuadrat antar titik-titik (nilai kategori) analogi dengan konsep korelasi antar variabel.
2. Variabel kolom yang tepat di variabel kategori baris diasumsikan homogen.
3. Analisis Korespondensi hanya digunakan untuk dua atau tiga variabel.
4. Analisis Korespondensi adalah sebuah teknik nonparametrik yang tidak memerlukan pengujian asumsi seperti kenormalan, autokorelasi, multikolinearitas, heteroskedastisitas, linieritas sebelum melakukan analisis selanjutnya.
5. Dimensi yang terbentuk dalam Analisis Korespondensi disebabkan dari kontribusi titik-titik dari dimensi yang terbentuk dan penamaan dari dimensinya subjektif dari kebijakan, pendapat dan *error*.
6. Dalam Analisis Korespondensi variabel yang digunakan yaitu variabel diskrit yang mempunyai banyak kategori.

Tabel Kontingensi Dua Arah

Jika X dan Y adalah dua peubah yang masing-masing mempunyai sebanyak a dan b kategori, maka dapat dibentuk suatu matriks data pengamatan \mathbf{P} dengan ukuran $a \times b$, Dengan $p_{ij} \geq 0$ menyatakan frekuensi dari sel ke (i, j) .

Tabel Kontingensi Dua Arah

	Y_1	...	Y_j	...	Y_b	Total
X_1	p_{11}	...	p_{1j}	...	p_{1b}	$p_{1.}$
\vdots						
X_i	p_{i1}	...	p_{ij}	...	p_{ib}	$p_{i.}$
\vdots						
X_a	p_{a1}	...	p_{aj}	...	p_{ab}	$p_{a.}$
Total	$p_{.1}$...	$p_{.j}$...	$p_{.b}$	$p_{..}$

keterangan :

$$p_{i.} = \sum_{j=1}^b p_{ij} \quad i = 1, 2, \dots, a \quad \text{peluang marginal } X$$

$$p_{.j} = \sum_{i=1}^a p_{ij} \quad j = 1, 2, \dots, b \quad \text{peluang marginal } Y$$

$$p_{..} = \sum_i \sum_j p_{ij} \quad \text{Jumlah total frekuensi dari matriks } \mathbf{P}$$

p_{ij} adalah frekuensi pengamatan ke i baris pada j kolom

Profil Baris dan Profil Kolom

Profil adalah proporsi dari setiap baris atau kolom Matriks Korespondensi yaitu setiap frekuensi pengamatan baris ke- i dan kolom ke- j dibagi dengan jumlah setiap total baris dan kolomnya masing-masing. $\mathbf{R} = \mathbf{D}_r^{-1}\mathbf{K}$ disebut profil baris (*row profile*). $\mathbf{C} = \mathbf{D}_c^{-1}\mathbf{K}'$ disebut sebagai profil kolom (*column profile*). \mathbf{K} disebut matriks korespondensi dengan $k_{ij} = \frac{p_{ij}}{p_{..}}$ merupakan elemen-elemen setiap matriks korespondensi.

Untuk menampilkan profil-profil baris dan profil-profil kolom tersebut kedalam ruang dimensi *Euclid* yang berdimensi dua digunakan pendekatan jarak Kai Kuadrat yaitu :

$$\chi^2 = \sum_{i=1}^a \sum_{j=1}^b \frac{(p_{ij} - m_{ij})^2}{m_{ij}}$$

Penguraian Nilai Singular (*Singular Value Decomposition*)

Untuk mereduksi dimensi data berdasarkan keragaman data (nilai *eigen/inersia*) terbesar dengan mempertahankan informasi optimum, diperlukan penguraian nilai singular. Teorema Dekomposisi Nilai Singular, yaitu misalkan \mathbf{A} matriks berukuran $m \times n$ maka ada matriks diagonal Σ berukuran $r \times r$ dan $r \leq \min\{m, n\}$, matriks orthogonal \mathbf{U} berukuran $m \times m$, matriks orthogonal \mathbf{V} berukuran $n \times n$, sehingga $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}'$ dengan Σ adalah matriks berukuran $m \times n$ yang mempunyai bentuk $\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$, $\lambda_1^2 \geq \dots \geq \lambda_m^2$ adalah nilai inersia dari $\mathbf{U}'\mathbf{U}$. Berdasarkan Teorema Dekomposisi Nilai Singular tersebut, Matriks yang akan di *singular value decomposition* adalah matriks $\mathbf{U} = \mathbf{D}_r^{-1/2}(\mathbf{K} - \mathbf{rc}')\mathbf{D}_c^{-1/2}$ yang hasilnya adalah :

\mathbf{A} adalah matriks berukuran $(a \times m)$

\mathbf{B} adalah matriks berukuran $(b \times m)$

dan Λ merupakan suatu matriks yang elemen-elemennya adalah nilai singular, dimana nilai singular adalah akar dari nilai *inersianya*.

Penguraian Nilai Singular Umum

Koordinat dari baris dan kolomnya ditentukan dengan menggunakan GSVD dari matriks $(\mathbf{K} - \mathbf{rc}')$ hasilnya $\mathbf{A}\mathbf{\Lambda}\mathbf{B}'$, dengan \mathbf{A} adalah matriks berukuran $a \times m$, \mathbf{B} adalah matriks berukuran $b \times m$, $\mathbf{\Lambda}$ adalah matriks diagonal yang mempunyai unsur-unsur diagonalnya nilai singular dari matriks $\mathbf{K} - \mathbf{rc}'$, dimana berlaku $\mathbf{A}'\mathbf{D}_r^{-1}\mathbf{A} = \mathbf{I}_m$ dan $\mathbf{B}'\mathbf{D}_c^{-1}\mathbf{B} = \mathbf{I}_m$.

Tiap himpunan titik dapat dihubungkan dengan sumbu utama dari himpunan titik lainnya yaitu:

	Rumusan dari koordinat baris	Rumusan untuk koordinat kolom
Analisis profil baris	$F = \mathbf{D}_r^{-1}\mathbf{A}\mathbf{\Lambda}$	$G = \mathbf{D}_c^{-1}\mathbf{B}$
Analisis profil kolom	$F = \mathbf{D}_r^{-1}\mathbf{A}$	$G = \mathbf{D}_c^{-1}\mathbf{B}\mathbf{\Lambda}$
Analisis keduanya (baris dan kolom)	$F = \mathbf{D}_r^{-1}\mathbf{A}\mathbf{\Lambda}$	$G = \mathbf{D}_c^{-1}\mathbf{B}\mathbf{\Lambda}$

Sumber : Bee- Leng Lee

Nilai Inersia

Nilai *inersia* menunjukkan kontribusi dari baris ke- i pada *inersia* total. Sedangkan dimaksud *inersia* total adalah jumlah bobot kuadrat jarak titik-titik ke pusat, massa dan jarak. Jumlah bobot kuadrat koordinat titik-titik dalam sumbu utama ke- d pada tiap-tiap himpunan yaitu λ_d^2 yang dinotasikan dengan λ_d . Nilai ini disebut sebagai *inersia* utama ke- d . Persamaan *inersia* utama baris dan kolom serta pusatnya dapat dinyatakan sebagai berikut:

$$\textit{inersia} \text{ utama baris adalah } F'\mathbf{D}_r F = \mathbf{\Lambda}$$

bukti $F'\mathbf{D}_r F = \mathbf{\Lambda}$, akan ditunjukkan sebagai berikut :

$$\begin{aligned} F'\mathbf{D}_r F &= (\mathbf{D}_r^{-1}\mathbf{A}\mathbf{\Lambda})' \mathbf{D}_r (\mathbf{D}_r^{-1}\mathbf{A}) \\ &= \mathbf{\Lambda}' \mathbf{A}' (\mathbf{D}_r^{-1})^{-1} \mathbf{A} \\ &= \mathbf{\Lambda}' \mathbf{A}' \mathbf{D}_r^{-1} \mathbf{A}, \text{ dengan menggunakan persamaan } \mathbf{A}' \mathbf{D}_r^{-1} \mathbf{A} = \mathbf{I}_m, \text{ didapatkan } \mathbf{\Lambda}' \mathbf{I}_m = \mathbf{\Lambda}' \end{aligned}$$

Karena matriks $\mathbf{\Lambda}'$ adalah simetris sehingga $\mathbf{\Lambda}' = \mathbf{\Lambda}$ jadi $F'\mathbf{D}_r F = \mathbf{\Lambda}$.

$$\textit{inersia} \text{ utama kolom adalah } G'\mathbf{D}_c G = \mathbf{\Lambda}$$

bukti $G'\mathbf{D}_c G = \mathbf{\Lambda}$, akan ditunjukkan sebagai berikut :

$$\begin{aligned} G'\mathbf{D}_c G &= (\mathbf{D}_c^{-1}\mathbf{B}\mathbf{\Lambda})' \mathbf{D}_c (\mathbf{D}_c^{-1}\mathbf{B}) \\ &= \mathbf{\Lambda}' \mathbf{B}' (\mathbf{D}_c^{-1})^{-1} \mathbf{B} \end{aligned}$$

$= \Lambda' \mathbf{B}' \mathbf{D}_r^{-1} \mathbf{B}$, dengan menggunakan $\mathbf{B}' \mathbf{D}_c^{-1} \mathbf{B} = \mathbf{I}_m$, didapatkan $\Lambda' \mathbf{I}_m = \Lambda'$. Karena matriks Λ' adalah simetris sehingga $\Lambda' = \Lambda$ jadi $G' \mathbf{D}_c G = \Lambda$.

Besaran relatif untuk mengukur besarnya kehilangan informasi dapat dirumuskan sebagai berikut:

$$L = 1 - \frac{\sum_{i=1}^d \lambda_i^2}{\sum_{i=1}^m \lambda_i^2}$$

Uji Kesesuaian Kai Kuadrat (*Test of Goodness of Fit*)

Uji yang sesuai untuk mengetahui ada tidaknya hubungan antara dua variabel kategori yang berupa tabel kontingensi, adalah *Pearson Chi-Square test*.

Koefisien Kontingensi

Untuk melihat keeratan hubungan atau kecenderungan antara variabel satu dengan yang lainnya. Dengan menggunakan rumusan koefisien kontingensi sebagai berikut :

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}}$$

keterangan:

- χ^2 = Statistik uji Kai Kuadrat
- N = banyaknya populasi sampel
- Nilainya $0 \leq C < 1$

METODE PENELITIAN

Jenis data yang digunakan adalah data sekunder berupa jumlah penderita penyakit menular terbanyak, hasil pencatatan setiap bulan pada tahun 2006 di Kota Bengkulu.

PEMBAHASAN

Analisis Korespondensi ini digunakan untuk memperoleh informasi mengenai posisi variabel penyakit menular dengan kecamatan. Analisis ini merupakan analisis *multivariate* yang mereduksi matriks data dengan ruang dimensi tinggi menjadi ruang dimensi dua. Untuk memperoleh informasi yang lengkap (100%) seperti informasi awal, dibutuhkan lima dimensi. Total keragaman yang dapat disumbang oleh dimensi pertama 73,63%, bila ditambah dengan dimensi kedua total keragaman yang dapat dijelaskan sebesar 92,97% dan 98,06% bila menggunakan tiga dimensi secara bersamaan, dan dimensi keempat menjelaskan keragaman sebesar 99,39%. Penambahan variasi yang dapat dijelaskan dengan ditambahkan dimensi kedua adalah sebesar 19,33% sedangkan penambahan dimensi ketiga menambah sebesar 5,09% atau lebih kecil daripada penambahan dimensi kedua.

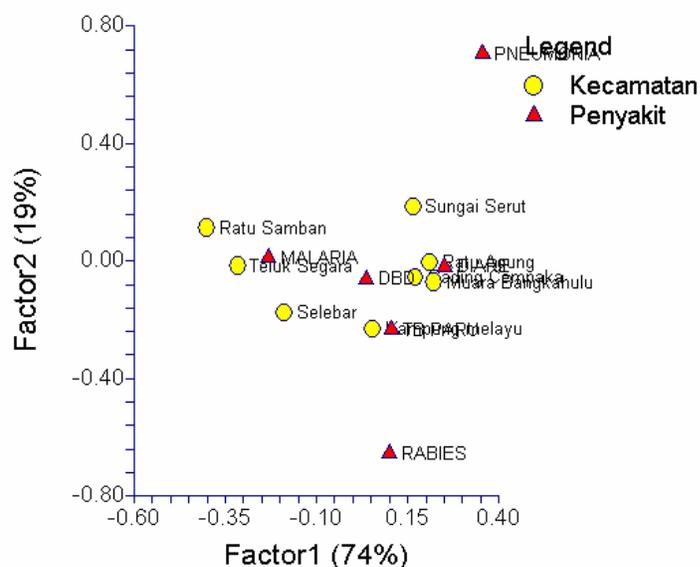
Oleh karena itu pemetaan yang disajikan menggunakan dua dimensi. Besarnya informasi yang hilang dalam pereduksian dimensi data dalam analisis ini sebesar $L=7.03\%$. Nilai singular dari setiap dimensi diatas adalah akar dari setiap *inersia* atau *eigenvalue* dari setiap dimensinya dapat dilihat dari hasil nilai *singular value decomposition*.

Analisis Tabel Kontingensi

Factor	Singular value	inertia	Individual percent	Cumulative percent	Bart Chart
1	0.24062	0.057914	73.63	73.63	
2	0.1233	0.015207	19.33	92.97	
3	0.0632	0.004003	5.09	98.06	
4	0.0316	0.001048	1.33	99.39	
5	0.0224	0.000482	0.61	100.00	
Total		0.078654			

Sumber: Analisis dengan NCSS 2000

Correspondence Plot



Plot di atas dapat dilihat bahwa penyakit menular malaria cenderung berada di Kecamatan Teluk Segara dan mempunyai kedekatan pada Kecamatan Ratu Samban. Sedangkan penyakit menular TB paru cenderung pada Kecamatan Kampung Melayu yang mempunyai kedekatan atau kemiripan dengan Kecamatan Selebar. Dan diare cenderung berada di Kecamatan Muara Bangkahulu yang mempunyai kedekatan dan kemiripan

dengan Kecamatan Ratu Agung dan Gading Cempaka. Dan penyakit menular DBD dan diare mempunyai kedekatan ataupun kemiripan. Jika semakin kecil jarak Kai_Kuadrat antar dua variabel Kecamatan dan kategori penyakit menular itu, maka kedua variabel itu memiliki suatu kemiripan dan kedekatan.

PENUTUP

Kesimpulan

Dari hasil penelitian yang dilakukan dengan judul Analisis Korespondensi jumlah Penderita Penyakit menular di Kota Bengkulu, ada beberapa kesimpulan sebagai berikut :

1. Berdasarkan grafik data dari jumlah penderita penyakit menular di Kota Bengkulu pada tahun 2006, tercatat bahwa ada sebanyak 16283 penderita penyakit menular yaitu DBD, diare, malaria, PNEUMONIA, rabies, dan TB Paru
2. Kecamatan yang mempunyai kemiripan dan kedekatan jenis penyakit menular DBD tersebar banyak di kecamatan Ratu Agung dengan jarak Kai Kuadrat 0.042, Ratu Samban jarak kai Kuadratnya 0.049, Teluk Segara jarak Kai Kuadratnya adalah 0.382 dan Muara Bangkahulu dengan jarak Kai Kuadrat 0.069. Sehingga solusi yang tepat untuk masing-masing Kecamatan itu yaitu kebijakan yang tepat dalam pemberantasan mengenai penyakit menular DBD di Kecamatan tersebut.
3. Kecamatan yang cenderung terhadap jenis penyakit TB Paru adalah Sungai Serut dengan jarak Kai Kuadratnya 0.028. Sehingga solusi yang tepat untuk Kecamatan Sungai Serut yaitu kebijakan yang mengarah pada penanggulangan mengenai penyakit menular TB paru.
4. Kecamatan yang cenderung terhadap jenis penyakit PNEUMONIA adalah Gading Cempaka dengan jarak Kai Kuadratnya 0.003. Sehingga solusinya yaitu suatu kebijakan dalam penanggulangan penyakit menular PNEUMONIA.

SARAN

Beberapa saran yang dapat diberikan oleh penulis sebagai bahan penelitian lanjutan adalah sebagai berikut :

1. Variabel yang digunakan dalam penelitian ini hanya dua variabel diskrit yang berkategori nominal serta data tercatat untuk tiap bulannya dalam satu tahun, oleh karena itu bagi peneliti yang ingin menggunakan Analisis Korespondensi sebaiknya dicoba untuk variabel yang lebih dari dua variabel dan berkategori banyak.
2. Dalam penulisan ini pembahasan mengenai data hilang belum banyak dibicarakan oleh peneliti, karenanya ini dapat menjadi salah satu bahan pertimbangan bagi pengembangan Analisis Korespondensi.