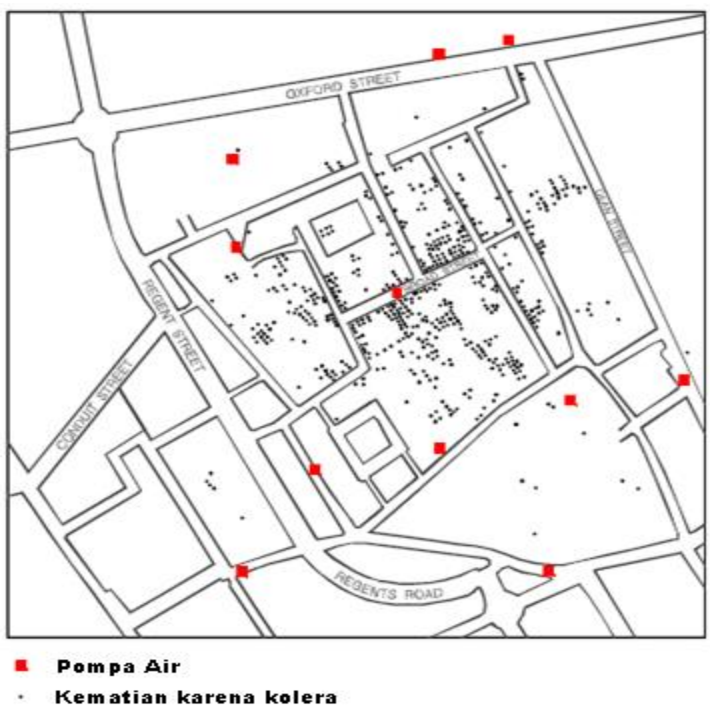
BAB 8

SEBARAN DUA TIITIK ATAU LEBIH

Muhammad Nur Aidi

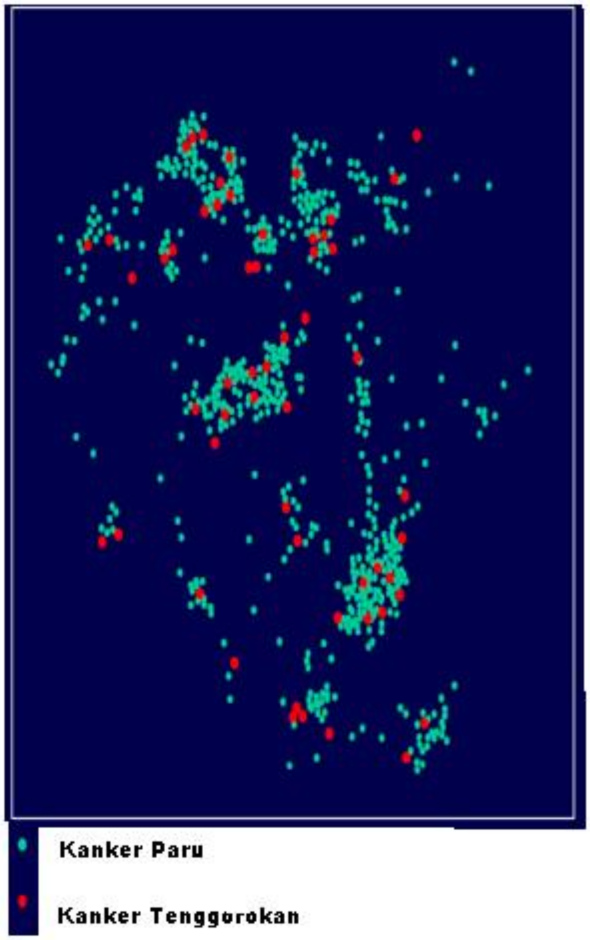
Pada pembiaraan sebelumnya kita membahas sebarang satu jenis/tipe titik dalam ruang. Pembahasan pada topik tersebut mendeteksi apakah titik-titik tersebut menyebar dalam ruang secara reguler, acak atau cluster/bergerombol. Metode yang ditempuh adalah dengan teknik kuadran atau teknik tetangga terdekat. Metode lainnya adalah menguji apakah titik dalam ruang tersebut mempunyai sebaran Poison, Binomial atau Binomial Negatif. Pengujian dilakukan dengan Khi-Kuadrat.

Bagaimana seandainya yang dibicarakan adalah sebaran dua tipe atau lebih titik-titik dalam ruang. Misalkan sebaran orang-orang yang mati karena kolera serta letak pompa air (Gambar 8.1).



Gambar 8.1. Sebaran Lokasi Penduduk Terkenan Kolera serta Sumber Air

Contoh lain adalah Sebaran spasial penderita penyakit kanker paru-paru dengan penderita kanker tenggorokan (Gambar 8.2)



Gambar 8.2. Sebaran Penduduk Terkena Kanker Paru dan Kanker Tenggorokan

Pada kasus di atas kita akan menemui berbagai variasi pola dari sebaran spasial dari dua tipe sebaran titik. Di dunia nyata, hal yang terpenting dari pola tersebut adalah kesamaaan pola antara dua sebaran titik tesebut. Jika sebaran spasial dua tipe titik tersebut menunjukkan pola yang sama, maka dapat diduga bahwa dua-duanya berhubungan antar sesamanya, baik secara langsung maupun tidak langsung. Kesamaan antara dua sebaran titik tersebut ada dua bentuk

a. Satu sebaran titik adalah penyebab keberadaan sebaran titik lainnya
b. Dua sebaran titik mempunyai penyebab yang sama

Untuk kasus sebaran kematian karena kolera dengan sebaran keberadaan pompa air dapat diduga bahwa penyakit kolera disebabkan sumber air pompa yang ada di lokasi sudah tercemar bibit kolera. Sedangkan sebaran spasial keberadaan penyakit kanker paru dengan kanker tenggorokan disebabkan oleh faktor lingkungan yang kurang baik pada lokasi bersangkutan.

Untuk kasus yang lain, kita akan menemui kesamaan antara dua distribusi spasial pada

a. Plankton dan predatornya
b. Lokasi yang tercemar limbah berbahaya dan sebuah penyakit yang diakibatkannya
c. Pertemuan jalan bebas hambatan dengan supermarket-supermarket
d. Gedung bioskop dan rumah makan.

Jika sebuah disribusi adalah komplemen dari distribusi lainnya, ini menunjukkan sebuah hubungan antara distribusi tersebut. Hubungan ini menunjukkan bahwa sebuah himpunan dari obyek-obyek spasial menyerang himpunan obyek-obyek lainnya. Sebuah contoh ini adalah distribusi kejahatan dengan kantor polisi. Kantor polisi mengurangi kejadian kriminalitas.

Spasial similaritas antara dua distribusi dapat dijelaskan secara kunatitatif oleh spasial prosimitas antara distribusi-distribusi tersebut.

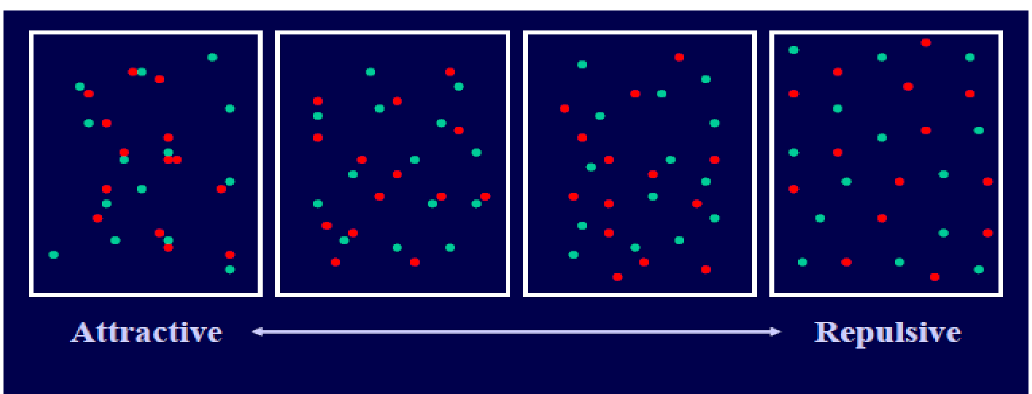
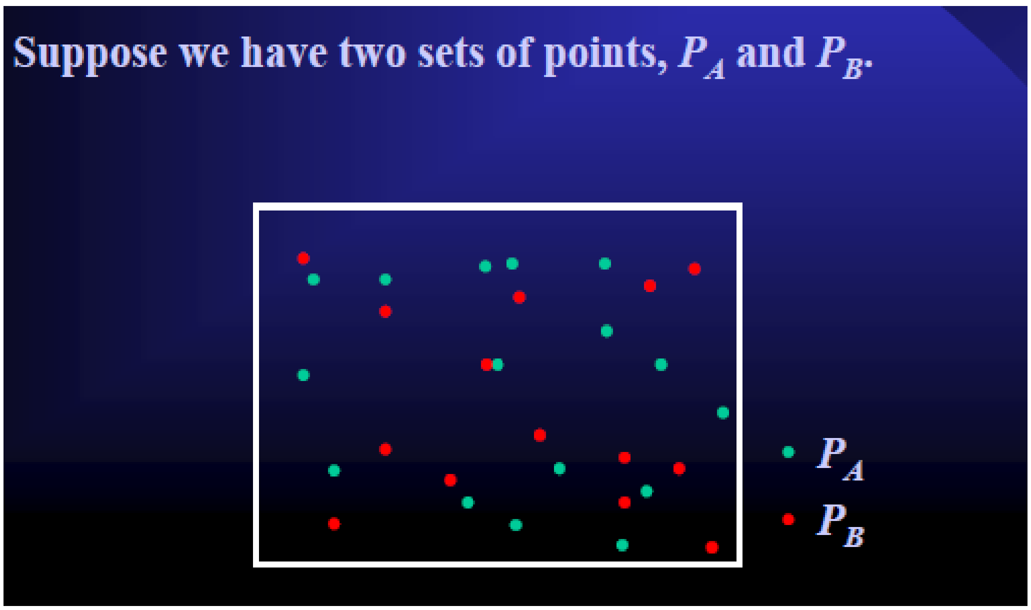


Figure: Relationship between two distributions of points

8.1. Metode Kuadran

Metode kuadran adalah sebuah statistika uji untuk mengevaluasi similaritas antara dua distribusi. Statistik ini digunakan untuk menganalisis tidak hanya untuk sebuah himpunan titik tetapi jua untuk hubungan antara dua himpunan titik-titik.

Sebagai contoh kita mempunyai dua himpunan titik-titik, katakan Pa dan Pb yang digambarkan pada Gambar 8.3 berikut.



Gambar 8.3. Sebaran Dua Himpunan Titik

Hipotesis yang dikembangkan keadaan di atas adalah

H0= Titik-titik pada Himpunan P_A dan P_B adalah menyebar secaraindependen. Mereka tidak berkorelasi secara spasial antara sesamanya.

H1= Titik-titik pada Himpunan P_A dan P_B adalah menyebar secara tidak independen. Mereka berkorelasi secara spasial antara sesamanya.

Metode kuadran mempertimbangkan apakah apakah dua distribusi tersebut independen secara spasial. Metode ini tidak menjawab secara langsung apakah dua distribusi tersebut bergerombol atau memisah secara spasial.

Dalam metode kuadran, pertama kita merubah data titik pada data raster. Kita selanjutnya mengklasifikasi sel-sel menjadi empat kategori dan menghitung banyaknya sel untuk setiap kategori yakni :

a. Sel yang berisi dua himpunan titik-titik P_A dan P_B
b. Sel yang hanya berisi himpunan titik-titik P_A
c. Sel yang hanya himpunan titik-titik P_B
d. Sel kosong

Dari banyaknya sel-sel tersebut, kita menghitung statistik Khi-kuadran, χ^2, dan melakukan uji.

C_{AB}=Banyaknya sel yang berisi kedua himpunan P_A dan P_B
C_{A0}=Banyaknya sel yang hanya berisi himpunan P_A
C_{0B}=Banyaknya sel yang hanya berisi himpunan P_B
C_{00}=Banyaknya sel kosong

$C_A = C_{AB} + C_{A0}$
$C_B = C_{AB} + C_{0B}$
$C = C_{AB} + C_{A0} + C_{0B} + C_{00}$

Hubungan antara dua peubah dapat ditunjukkan melalui table 2 x 2 berikut :

C_{AB}	C_{0B}	C_B
C_{A0}	C_{00}	$C-C_B$
C_A	$C-C_A$	C

Jika PA dan PB independen secara spasial maka tabel 2 x 2 tersebut akan proporsional, sebagai contoh adalah berikut

10	20	30
30	60	90
40	80	120

Khi-kuadrat test membandingkan antara data hasil observasi dengan nilai harapan dari model teori. Pada kasus kita, kita menggunakan pola proporsional sebagai distribusi teori dari banyaknya sel.

C_{AB}=Banyaknya sel yang berisi kedua himpunan P_A dan P_B
C_{A0}=Banyaknya sel yang hanya berisi himpunan P_A

C_{0B}=Banyaknya sel yang hanya berisi himpunan P_B
C_{00}=Banyaknya sel kosong

$C_A = C_{AB} + C_{A0}$
$C_B = C_{AB} + C_{0B}$
$C = C_{AB} + C_{A0} + C_{0B} + C_{00}$

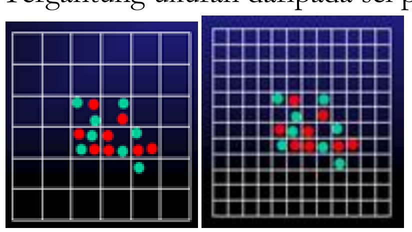
Y_{AB} adalah nilai harapan banyaknya sel yang berisi titik-titik P_A dan P_B
Y_{A0} adalah nilai harapan banyaknya sel yang berisi hanya titik-titik P_A
Y_{0B} adalah nilai harapan banyaknya sel yang berisi hanya titik-titik P_B
Y_{00} adalah nilai harapan banyaknya sel koson
$Y_{AB} = \frac{C_A C_B}{C}$
$Y_{A0} == \frac{C_A(C-C_B)}{C}$
$Y_{0B} == \frac{C_B(C-C_A)}{C}$
$Y_{00} == \frac{(C-C_B)(C-C_A)}{C}$

$$\chi^2 = \frac{(C_{AB} - Y_{AB})^2}{Y_{AB}} + \frac{(C_{A0} - Y_{A0})^2}{Y_{A0}} + \frac{(C_{0B} - Y_{0B})^2}{Y_{0B}} + \frac{(C_{00} - Y_{00})^2}{Y_{00}}$$

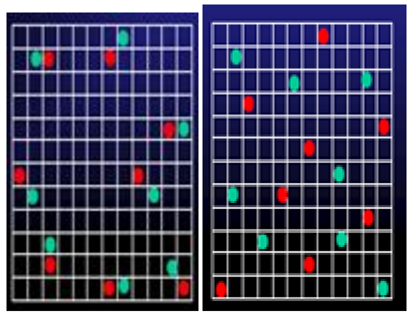
Jika P_A dan P_B adalah independen secara spasial, maka statistik uji tersebut akan mengikuti sebaran χ^2 dengan derajat bebas 1

Ada beberapa kelemahan metode kuadran yang digunakan pada dua distribusi titik secara spasial yang mirip dengan kelemahan metode kuadran untuk melihat pola spasial pada satu populasi titik, antara lain :

a. Tergantung ukuran daripada sel pada metode kuadran

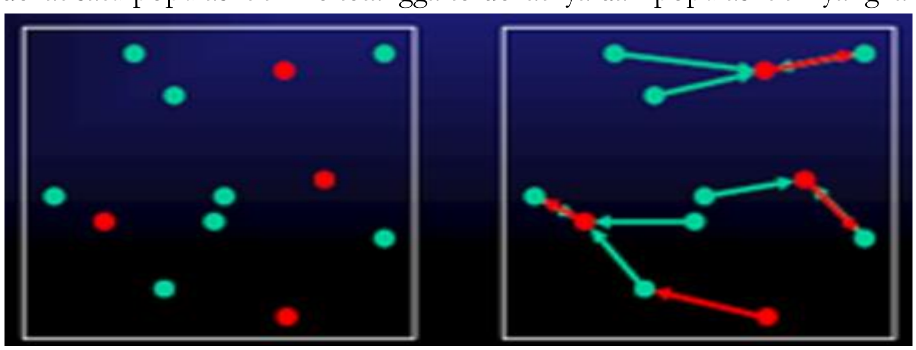


b. Metode kuadran tidak dapat membedakan beberapa perbedaan distribusi.



8.2. Metode Silang-Tetangga Terdekat

Metode Silang tetangga terdekat merupakan pengembanga dari metode tetangga terdekat sederhana (*ordinary*) yang digunakan untuk mempelajari homogenitas distribusi titik. Metode silang tetangga terdekat menghitung jarak terdekat satu populasi titik ke tetangga terdekatnya dari populasi titik yang lain.



Gambar 8. 4. Proses Penghitungan Jarak dengan Metode Silang Tetangga Terdekat

Seperti pada metode tetangga terdekat sederhana, metode silang tetangga terdekat mempunyai dua metode statistik untuk memngukur pola gerombol dari distribusi titik secara spasial. Hal ini disebabkan ada dua perlakuan distribusi titik secara spasial yang erat kaitannya dengan kesamaan (similarity) antara dua distribusi, yakni satu distribusi disebabkan disebabkan distribusi lainnya atau dua distribusi mempunyai penyebab yang sama.

8.2.1. Pengujian Dua Arah

Jika kita ingin mengetahui apakah dua distribusi mempunyai penyebab yang sama, kita memperlakukan distribusi-distribusi tersebut adalah equivalen. Sebagai contoh kejadian kasus sebaran spasial penderita kangker tenggorokan dengan kanker paru-paru, sebaran restoran Hamburger dengan rumah makan cepat saji. Pengujian dua arah ini juga sangat berguna ketika dua distribusi saling mempengaruhi. Hal ini mengasumsikan saling mempengaruhi secara spasial antara distribusi-distribusi tersebut. Contoh lain : Stasiun pompa bensin dengan restoran pada jalan bebas hambatan, toko obat dengan *groceries*

Katakan ada dua himpunan titik yakni P_A dan P_B pada sebuah daerah S

d_A : Jarak dari titik ke i pada himpunan titik P_A ke titik terdekat dari himpunan titik P_B

d_B : Jarak dari titik ke i pada himpunan titik P_B ke titik terdekat dari himpunan titik P_A

n_A: jumlah titik pada Himpunan P_A

n_B: jumlah titik pada Himpunan P_B

Formulasi dari Silang tetangga terdekat adalah sebagai berikut

$$V = \frac{1}{n_A + n_B}\left(\sum_{i=1}^{n_A} d_{Ai} + \sum_{i=1}^{n_B} d_{Bi}\right)$$

Nilai V kecil apabila dua distribusi tersebut menggerombol secara spasial dan V besar apabila dua distribusi tersebut saling terpisah. Hipotesis yang dikembangkan adalah

H0 : Dua distribusi independen secara spasial. Setiap distribusi mengikuti sebaran poison secara independen.

H1: Dua distribusi berkorelasi secara spasial. Distribusi-distribusi tersebut secara spasial saling mempengaruhi.

Jika V adalah signifikan kecil atau besar, kita menerima hipotesis satu, dan kita katakan bahwa distribusi-distribusi tersebut secara spasial saling mempengaruh. Sebaliknya, menerima hipotesis nol, kita katakan distribusi-distribusi tersebut independen.

Di bawah hipotesis nol, jika titik-titik tersebar secara acak pada ruang yang tak berbatas, distribusi peluang dari statistik V yang didapatkan mempunyai sebaran normal dengan nilai

$$N\left(\frac{1}{n_A+n_B}\left(\frac{n_A}{2\sqrt{\lambda_a}}+\frac{n_B}{2\sqrt{B}}\right),\frac{n_A\lambda_A(4-\pi n_A)+n_B\lambda_B(4-\pi n_B)+2\pi n_A n_B\sqrt{\lambda_A\lambda_B}}{4\pi\lambda_A\lambda_B(n_A-n_B)}\right)$$

λ_A adalah kepekatan peluang himpunan P_A($=\frac{n_A}{A}$)

λ_B adalah kepekatan peluang himpunan P_B($=\frac{n_B}{A}$)

A=Luas wlayah

8.2.2. Pengujian Satu Arah

Pengujian satu arah digunakan ketika hanya satu distribusi mempunyai pengaruh terhaap distribusi lainnya. Ini berarti ada pengaruh satu arah secara spasial. Hal ini menandakan kita tidak memperlakukan dua distribusi secara sama. Sebagai contoh pada kasus ini adalah distribusi tempat sumber air dengan distribusi penyakit kolera, distribusi tempat pembuangan limbah berbahaya dengan distribusi spasial kasus leukemia.

Test hipotesis satu arah, silang jarak tertangga terdekat adalah berbeda. Jika kita tertarik sebuah himpunan titik P_A adalah menggerombol secara spasial di sekitar himpunan titik P_B, kita menghitung Silang jarak tetanga terdekat dari himpunan titik P_A

$$V_A=\frac{1}{n_A}\sum_{i=1}^{n_A}d_{Ai}$$

d_{Ai} adalah jarak dari titik ke i dari P_A ke tetangga terdekatnya dari himpunan P_B

Hipotesis yang dikembangkan adalah

H0= Titik-titik P_A adalah independen secara spasial dari titik-titik P_B.
Titik-titik P_A mengikuti sebaran poison homogen

H1= Titik-titik P_A dipengaruhi oleh titik-titik P_B. Titik-titik berkecenderungan bertempat tidak jauh dari titik P_B

Untuk menguji hipotesis ini, kita menetapkan lokasi titik-titik P_B dan menganggap distribusi himpunan titik PA adalah acak.

8.3. Kasus Anak Kekurangan Gizi dengan Ibu Kekurangan Gizi

Pada kasus ini kita melihat data tentang jumlah ibu yang kurang gizi dengan jumlah anak yang kekurangan gizi di beberapa desa. Data disajikan pada Tabel 8.1.

Tabel 8.1. Sebaran Jumlah Ibu dan Anak yang Kekurangan Gizi

Desa	Jumlah anak kurang Gizi	Jumlah Ibu kurang Gizi	Desa	Jumlah anak kurang Gizi	Jumlah Ibu kurang Gizi
A	1	0	P	3	1
B	1	1	Q	1	0
C	1	1	R	3	0
D	1	0	S	2	1
E	3	2	T	2	1
F	0	0	U	2	0
G	2	1	V	0	0
H	4	0	W	2	0
I	0	0	X	0	0
J	3	1	Y	2	2
K	3	1	Z	2	1
L	1	2	AA	0	0
M	2	0	AB	1	2
N	5	2	AC	1	1
O	2	2	AD	0	0

Pertanyaan :

a. Apakah ada korelasi antara banyaknya Anak kurang gizi dengan banyak ibu kurang gizi

Hipotesis : H_0 : Tidak ada korelasi antara banyaknya anak kurang gizi dengan banyaknya ibu kurang gizi

H_1 : Ada korelasi antara banyaknya anak kurang gizi dengan banyaknya ibu kurang gizi

Pearson correlation of Anak and Ibu (r)= 0,392 P-Value = 0,032

Taraf nyata α = 5%=0,05, Karena nilai P-value = 0,032 < α = 0,05 maka H_0 ditolak artinya ada korelasi antara banyaknya anak kurang gizi dengan banyaknya ibu kurang gizi.

b. Apakah ada asosiasi secara spasial antara banyaknya Anak kurang gizi dengan banyak ibu kurang gizi.

Desa	Jumlah anak kurang gizi	Jumlah ibu kurang gizi
A	1	0
B	1	1
C	1	1
D	1	0
E	1	1
F	0	0
G	1	1
H	1	0
I	0	0
J	1	1
K	1	1
L	1	1
M	1	0
N	1	1
O	1	1
P	1	1
Q	1	0
R	1	0

S	1	1
T	1	1
U	1	0
V	0	0
W	1	0
X	0	0
Y	1	1
Z	1	1
AA	0	0
AB	1	1
AC	1	1
AD	0	0

Hipotesis : H_0 : Tidak ada asosiasi spasial antara banyaknya anak kurang gizi dengan banyaknya ibu kurang gizi

H_1 : Ada asosiasi spasial antara banyaknya anak kurang gizi dengan banyaknya ibu kurang gizi

Statistik uji: $\chi^2 = \frac{(ad-bc)^2 N}{(a+d)(b+c)(a+c)(b+d)}$

	Anak Kurang Gizi		
Ibu Kurang Gizi	**Tidak ada**	**Ada**	**Total**
Tidak ada	6 (a)	8 (b)	14
ada	0 (c)	16 (d)	16
Total	**6**	**24**	**30**

$$\chi^2 = \frac{(ad-bc)^2 N}{(a+d)(b+c)(a+c)(b+d)} = \frac{(6*16-0*8)^2*30}{(14)(16)(6)(24)} = 8.57142857$$

p-val = 0.003 atau $\chi^2_{(0.05,1)=}$3.841459149

Karena P –value < taraf nyata 5%, maka tolak Ho.

Atau

Karena nilai $\chi^2 > \chi^2_{(0.05,1)}$, maka tolak Ho

Artinya pada taraf nyata 5% ada asosiasi antara banyaknya Anak kurang gizi dengan banyak ibu kurang gizi dengan nilai asosiasi sebesar 8.571.

8.4. Daftar Pustaka

1. Ludwig, J.A, Reynold, J.F. 1988. Statistical Ecology. A Primer on Method and Computing. John Wiley and Sons. New York.
2. Rogers, A. 1974. Statistical Analysis of Spatial Dispersion. London : Pion Limited
3. Ross, S. 1989. A First Course in Probability. Macmillian Publishing Company. New York
4. Thomas, R. W. 1977. An Introduction to Quadrat Analysis. Norwich : Geo Abstracts Lt