BAB II

TINJAUAN PUSTAKA

2.1 Pendahuluan

Dalam bab ini akan membahas mengenai teori-teori tentang statistika nonparametrik, korelasi parsial tau Kendall dan korelasi parsial menurut Ebuh GU dan Oyeka ICA.

2.2 Statistika Nonparametrik

Istilah nonparametrik pertama kali diciptakan oleh Jacob Wolfowitz, pada tahun 1942 (Kvam dkk, 2007: 1). Analisis nonparametrik adalah analisis yang modelnya tidak menetapkan syarat-syarat mengenai parameter-parameter populasi yang merupakan induk sampel penelitiannya (Siegel, 1992: 38). Anggapan-anggapan tertentu dikaitkan dengan sejumlah besar tes-tes statistik nonparametrik, yaitu bahwa obsevasi-observasinya dapat independen atau berpasangan dan bahwa variabel yang diteliti pada dasarnya memiliki kontinuitas. Pada statistik nonparametrik digunakan untuk menganalisis dengan data yang minimal berukuran skala ordinal.

2.3 Korelasi Parsial Tau Kendall

Koefisien korelasi parsial adalah ukuran yang digunakan untuk mempelajari hubungan antara dua variabel acak pada korelasi biasa dengan mengkaji pengaruh beberapa variabel lain dengan variabel yang terakhir dianggap konstan (Gibbons, 2003: 483). Koefisien korelasi parsial merupakan generalisasi dari koefisien korelasi tau Kendall (Siegel, 1992: 265). Jika terlihat ada korelasi antara dua variabel, selalu terdapat kemungkinan bahwa korelasi ini adalah akibat dari asosiasi antara masing-masing kedua variabel itu dengan suatu variabel ketiga. Secara statistik, masalah ini dapat diatasi dengan metode korelasi parsial. Dalam korelasi parsial, akibat-akibat variasi yang disebabkan oleh suatu variabel ketiga terhadap hubungan antara variabel X dan Y dihilangkan. Dengan kata lain, korelasi antara X dan Y ditemukan dengan variabel ketiga yaitu Z dan variabel Z dijaga agar konstan. Dalam merancang suatu eksperimen, orang dapat memilih antara mengadakan kontrol terhadap eksperimen itu sendiri untuk menghapuskan pengaruh variabel ketiga tersebut atau menggunakan metode-metode statistik untuk menghapuskan pengaruhnya. Oleh karenanya untuk menentukan hubungan langsung kedua variabel satu dengan yang lain, pengaruh perbedaan pada variabel ketiga Z harus dikontrol. Namun jika kontrol eksperimental ini tidak mungkin dijalankan, maka dapat diterapkan kontrol statistik. Dengan teknik korelasi parsial ini kita dapat menjaga agar variabel ketiga Z terhadap hubungan antara variabel X dan Y itu konstan.

Maka dari itu, suatu metode kontrol statistik yang dapat digunakan adalah korelasi *rank* Kendall. Untuk menggunakan metode korelasi nonparametrik ini, kita harus memiliki data yang diukur sekurang-kurangnya dalam skala ordinal.

Untuk memudahkan pemahaman, misalkan kita mendapatkan *ranking* untuk 4 subjek pada tiga variabel X, Y dan Z. Kita akan menentukan korelasi antara X dan Y jika Z dibuat konstan. *Ranking-ranking* itu adalah:

Tabel 2.1 Contoh *ranking* data

Subjek	***Ranking* Z**	***Ranking* X**	***Ranking* Y**
a	1	3	2
b	2	1	1
c	3	2	3
d	4	4	4

Sekarang kita lihat pasangan *ranking* yang mungkin dalam tiap variabel, kita mengetahui terdapat $C\binom{4}{2}$ pasangan yang mungkin. Setelah mengatur *ranking-ranking* pada Z dalam urutan wajar, kita perhatikan setiap pasangan yang mungkin dalam *ranking* X, *ranking* Y dan *ranking* Z. kemudian berilah tanda (+) untuk tiap-tiap pasangan yang didalamnya *ranking* yang lebih rendah mendahului *ranking-ranking* yang lebih tinggi, tanda (-) untuk tiap-tiap pasangan yang didalamnya *ranking* yang lebih tinggi mendahului *ranking-ranking* yang lebih rendah.

Tabel 2.2 Contoh Tanda +1 dan -1

Pasangan	**Z**	**X**	**Y**
(a,b)	+1	-1	-1
(a,c)	+1	-1	-1
(a,d)	+1	+1	+1
(b,c)	+1	+1	+1
(b,d)	+1	+1	+1
(c,d)	+1	+1	+1

Selanjutnya informasi yang telah didapatkan diringkas kedalam tabel kontingensi 2 x 2. Kemudian perhatikanlah ketiga tanda di bawah (a,b). untuk himpunan pasangan *rank* itu, baik X maupun Y diberi tanda (-) sedangkan Z diberi tanda plus (+). Dengan demikian X maupun Y “tidak sesuai” dengan Z. Meringkas informasi dengan menempatkan pasangan (a,b) dalam sel D pada Tabel 2.3. Selanjutnya perhatikan pasangan (a,c) disini tanda yang dimiliki Y sesuai dengan tanda Z, tetapi tanda X tidak sesuai dengan tanda Z. Oleh sebab itu, pasangan (a,c) ditempatkan dalam sel C pada Tabel 2.3. Dalam setiap kasus pasangan-pasangan yang lain, baik tanda Y maupun tanda X sama dengam tanda Z. Oleh sebab itu, keempat pasangan ini dimasukkan dalam sel A dalam Tabel 2.3.

Tabel 2.3 Tabel Kontingensi banyaknya tanda +1 dan -1 untuk Perhitungan Korelasi Parsial

	Pasangan Y bertanda sama dengan Z	Pasangan Y bertanda tidak sama dengan Z	Jumlah
Pasangan X betanda sama dengan Z	A 4	B 0	(A+B) 4
Pasangan X bertanda tidak sama dengan Z	C 1	D 1	(C+D) 2
Jumlah	(A+C) 5	(B+D) 1	$C\binom{n}{2}$ 6

Rumus korelasi parsial tau Kendall untuk Tabel 2.3 adalah:

$$\tau_{xy.z} = \frac{AD - BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} \qquad (2.1)$$

2.3 Korelasi Parsial Menurut Ebuh GU dan Oyeka ICA

Ambil variabel X, Y dan Z dimana x_i, y_i dan z_i adalah masing-masing pengamatan ke-i dalam sampel acak ukuran n yang memiliki skala pengukuran minimal skala ordinal, untuk i = 1,2 ... n. Sedangkan d_{ix} adalah *ranking* untuk x_i dari populasi X, d_{iy} adalah *ranking* untuk y_i dari populasi Y dan d_{iz} adalah *ranking* untuk z_i dari populasi Z, dengan $i = 1,2,...,n$.

Dibawah ini merupakan tabel struktur data untuk korelasi parsial dengan tiga variabel yaitu X, Y dan Z dengan variabel Z dikonstankan.

Tabel 2.4 Struktur Data

No. Subjek	Variabel Z	Variabel X	Variabel Y
1	z_1	x_1	y_1
2	z_2	x_2	y_2
3	z_3	x_3	y_3
4	z_4	x_4	y_4
⋮	⋮	⋮	⋮
n	z_n	x_n	y_n

Kemudian untuk menentukan urutan wajar, maka ketiga variabel tersebut diberi *ranking* mulai dari 1 sampai n. Kemudian urutkan *ranking* tersebut berdasarkan ranking Z.

Tabel 2.5 Struktur Data untuk *Ranking* Variabel X, Y dan Z

No. Subjek	***Ranking* Z**	***Ranking* X**	***Ranking* Y**
1	d_{1z}	d_{1x}	d_{1y}
2	d_{2z}	d_{2x}	d_{2y}
3	d_{3z}	d_{3x}	d_{3y}
4	d_{4z}	d_{4x}	d_{3y}
⋮	⋮	⋮	⋮
n	d_{nz}	d_{nx}	d_{ny}

Untuk memperkirakan koefisien korelasi antara pengamatan dari populasi X dan pengamatan dari populasi Z kita mendefinisikan.

$$U_{jk:(x.z)} = \begin{cases} 1, jika\ d_{jx} < d_{kx} \\ 0, jika\ d_{jx} = d_{kx} \\ -1, jika\ d_{jx} > d_{kx} \end{cases} \qquad (2.2)$$

Tabel 2.6 Struktur Data untuk Nilai $U_{jk:(x.z)}$

Urutan *ranking* untuk variabel X

	d_{kx}							
	No. Subjek		1	2	3	4	…	n
		d_{1x}	d_{2x}	d_{3x}	d_{4x}	…	d_{nx}	
	1	d_{1x}	0					
d_{jx}	2	d_{2x}		0				
	3	d_{3x}			0			
	4	d_{4x}				0		
	⋮	⋮					0	
	n	d_{nx}						0

Tanda 0, +1 dan -1

No. pengamatan yang diurutkan sesuai dengan urutan *ranking* pada variabel Z

Keterangan:

$j =$ nomor pengamatan untuk baris , $j = 1,2,...,n$

$k =$ nomor pengamatan untuk kolom , $k = 1,2,...,n$; $j \neq k$

Ranking ditetapkan pada pengamatan ke-k dari populasi X yang telah ada sebelumnya dan *ranking* untuk pengamatan ke-j dari populasi yang sama ketika pengamatan disusun berdasarkan urutan wajar atau dari *ranking* pengamatan sesuai dari populasi Z.

$$\pi_x^+ = P(U_{jk;(x.z)} = 1)\ ;\ \pi_x^0 = P(U_{jk;(x.z)} = 0)\ ;\ \pi_x^- = P(U_{jk;(x.z)} = -1) \qquad (2.3)$$

Keterangan:

$\pi_x^+ = P(U_{jk:x.z} = 1) =$ probabilitas konkordan untuk variabel X

$\pi_x^- = P(U_{jk:x.z} = -1) =$ probabilitas diskordan untuk variabel X

$$\pi_x^+ = \frac{f_x^+}{n(n-1)};\ \pi_x^0 = \frac{f_x^0}{n(n-1)};\ \pi_x^- = \frac{f_x^-}{n(n-1)} \qquad (2.4)$$

Dimana f_x^+, f_x^0 dan f_x^- yang masing-masing merupakan jumlah dari 1, 0 dan -1 dalam distribusi frekuensi dari angka-angka di $U_{jk;(x.z)}$, $j = 1,2,...,n$; $k = 1,2,...,n$; $j \neq k$.

Dengan cara yang sama untuk variabel Y definisikan terlebih dahulu seperti pada persamaan 2.3. Kemudian data disusun seperti pada tabel 2.6 dan hitung probabilitas konkordan dan diskordan seperti pada persamaan 2.4.

Untuk memperkirakan koefisien korelasi tau Kendall antara X dan Y dengan mengkonstankan variabel Z adalah dengan cara mengurutkan *ranking* dari variabel X dan *ranking* variabel Y mengikuti pasangannya, dengan menghilangkan variabel Z.

Tabel 2.7 Struktur Data untuk *Ranking* Variabel X dan Y

No. Subjek	***Ranking* X**	***Ranking* Y**
1	d_{1x}	d_{1y}
2	d_{2x}	d_{2y}
3	d_{3x}	d_{3y}
4	d_{4x}	d_{3y}
$\vdots$	$\vdots$	$\vdots$
n	d_{nx}	d_{ny}

Dapat didefinisikan bahwa

$$U_{jk:(xy)} = \begin{cases} 1, jika\ d_{jy} < d_{ky} \\ 0, jika\ d_{jy} = d_{ky} \\ -1, jika\ d_{jy} > d_{ky} \end{cases} \qquad (2.5)$$

Tabel 2.8 Struktur Data untuk Nilai $U_{jk:(xy)}$

Urutan *ranking* untuk variabel Y

	d_{ky}							
	No. Subjek		1	2	3	4	…	n
		d_{1y}	d_{2y}	d_{3y}	d_{4y}	…	d_{ny}	
	1	d_{1y}	0					
d_{jy}	2	d_{2y}		0				
	3	d_{3y}			0			
	4	d_{4y}				0		
	⋮	⋮					0	
	n	d_{ny}						0

Tanda 0, +1 dan -1

No. pengamatan yang diurutkan sesuai dengan urutan *ranking* pada variabel Z

Keterangan:

$j =$ nomor pengamatan untuk baris , $j = 1,2,...,n$

$k =$ nomor pengamatan untuk kolom , $k = 1,2,...,n$; $j \neq k$

Ranking ditetapkan pada pengamatan ke-k dari populasi Y yang telah ada sebelumnya dan *ranking* untuk pengamatan ke-j dari populasi yang sama ketika pengamatan disusun berdasarkan urutan wajar atau dari *ranking* pengamatan sesuai dari populasi X.

$$\pi_{xy}^{+} = P(U_{jk;(xy.z)} = 1) \text{ ; } \pi_{xy}^{0} = P(U_{jk;(xy.z)} = 0) \text{ ; } \pi_{xy}^{-} = P(U_{jk;(xy.z)} = -1) \qquad (2.6)$$

Keterangan:

$\pi_{xy}^{+} = P(U_{jk:(xy.z)} = 1) =$ probabilitas konkordan untuk variabel XY

$\pi_{xy}^{-} = P(U_{jk:(xy.z)} = -1) =$ probabilitas diskordan untuk variabel XY

$$\pi_{xy}^{+} = \frac{f_{xy}^{+}}{n(n-1)} \text{ ; } \pi_{xy}^{0} = \frac{f_{xy}^{0}}{n(n-1)} \text{ ; } \pi_{xy}^{-} = \frac{f_{xy}^{-}}{n(n-1)} \qquad (2.7)$$

Dimana f_{xy}^{+}, f_{xy}^{0} dan f_{xy}^{-} yang masing-masing merupakan jumlah dari 1, 0 dan -1 dalam distribusi frekuensi dari angka-angka di $U_{jk;(xy)}$, $j = 1,2,...,n$; $k = 1,2,...,n$; $j \neq k$.

Maka estimasi nonparametrik koefisien korelasi parsial antara variabel X dan Y saat pengamatan dari variabel Z dianggap konstan adalah

$$r_{xy.z} = \frac{2\left(\pi_{xy}^{+} - \pi_{xy}^{-}\right) - 4\left(\pi_{x}^{+} - \pi_{x}^{-}\right)\left(\pi_{y}^{+} - \pi_{y}^{-}\right)}{\sqrt{\left(1 - 4\left(\pi_{x}^{+} - \pi_{x}^{-}\right)^{2}\right)\left(1 - 4\left(\pi_{y}^{+} - \pi_{y}^{-}\right)^{2}\right)}} \qquad (2.8)$$

Rumusan hipotesis untuk uji keberartian korelasi parsial adalah sebagai berikut:

$H_0 : \rho_{xy.z} = 0$, Tidak ada hubungan antara variabel X dan Y dengan variabel Z konstan.

$H_1 : \rho_{xy.z} \neq 0$, Ada hubungan antara antara variabel X dan Y dengan variabel Z konstan.

Dengan statistik ujinya adalah:

$$r_{xy.z} = \sqrt{\frac{\chi^2}{n}}$$

$$\chi^2 = r_{xy.z}{}^2 n \qquad (2.9)$$

Kriteria pengujian adalah tolak H_0 jika $\chi^2_{hitung} > \chi^2_{(1),\alpha}$

Keterangan:

α = taraf nyata