

# PEMODELAN COPULA: STUDI BANDING KUANTIFIKASI AUTOKORELASI

Fachrur Rozi

Jurusan Matematika, Fakultas Sains dan Teknologi  
Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang  
e-mail: fachrurkibar@yahoo.com

## Abstrak

Dalam tulisan ini akan dijelaskan beberapa metode kuantifikasi dependensi antara dua variabel acak (bivariat) dan perbandingan antara metode kuantifikasi dependensi tersebut. Selain itu akan diperkenalkan teori copula dalam kaitannya dengan kuantifikasi dependensi pada data time series, yang biasa disebut autokorelasi, khususnya kuantifikasi autokorelasi Kendall's tau melalui copula.

**Kata kunci:** autokorelasi, copula, kendall's tau.

## 1. Pendahuluan

Dalam kehidupan sehari-hari, sering kita dipertemukan dengan fenomena hubungan antara beberapa karakteristik yang diduga mempunyai keterkaitan antara karakteristik yang satu dengan karakteristik yang lain. Dalam ilmu statistika keterkaitan ini sering disebut dependensi (keterhubungan) antara variabel yang satu dengan variabel yang lain. Jogdeo (1982) mengatakan:

“Hubungan ketergantungan (dependensi) antara beberapa variabel acak adalah salah satu persoalan yang sangat banyak dipelajari dalam ilmu probabilitas dan statistika. ....”

Dalam tulisan ini akan dijelaskan beberapa metode kuantifikasi dependensi antara dua variabel acak dan perbandingan antara metode kuantifikasi dependensi tersebut. Salah satu hal yang bisa dikatakan baru adalah memperkenalkan teori *copula* kaitannya dengan dependensi antara dua variabel acak, khususnya dependensi pada data time series, yang biasa disebut autokorelasi. Dalam hal ini *copula* hadir sebagai perluasan metode dalam memodelkan dependensi antar variabel acak, *copula* akhir-akhir ini banyak dikembangkan dalam bidang biostatistika, ilmu aktuaria, dan keuangan.

Pembahasan pada tulisan ini hanya dilakukan untuk dependensi antara dua variabel, khususnya pada kasus dependensi data *time series*, yang biasa disebut autokorelasi lag-1. Dengan demikian, teori-teori yang dijelaskan lebih ditekankan pada dependensi antara dua variabel (bivariat). Dalam membandingkan kuantifikasi dependensi dilakukan dengan membandingkan hasil simulasi dari kuantifikasi dependensi yang yang diperoleh.

## 2. Copula dan Sifat-sifatnya

Pada bagian ini, akan dijelaskan mengenai definisi *copula* dan sifat-sifat dasarnya sebagai teori dasar yang akan digunakan dalam pembahasan selanjutnya.

**Definisi 1.** *Copula dua dimensi (2-copula) adalah fungsi  $C: I^2 \rightarrow I$  yang memenuhi sifat-sifat:*

a. Untuk setiap  $\vec{u} = (x, y)^t \in I^2$ , maka berlaku

$$C(x, 0) = 0 = C(0, y) \text{ dan } C(x, 1) = x ; C(1, y) = y. \quad (2.1)$$

b. Untuk setiap  $\vec{u} = (x_1, y_1)^t, \vec{v} = (x_2, y_2)^t \in I^2$  sedemikian sehingga  $\vec{u} \leq \vec{v}$ , maka berlaku

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0. \tag{2.2}$$

Himpunan dari semua copula dua dimensi didefinisikan sebagai  $\mathcal{C}_2$ .

Mengingat,

$$\begin{aligned} C(u, v) &= C(u, v) - C(u, 0) - C(0, v) + C(0, 0) \\ &= V_C([0, u] \times [0, v]) \end{aligned} \tag{2.3}$$

maka hal ini akan menunjukkan bahwa  $C(u, v)$  sebagai pengaitan suatu bilangan di  $I$  terhadap persegi panjang  $[0, u] \times [0, v]$ .

### 2.1. Copula dan Variabel Acak

**Teorema 1. (Teorema Sklar)** Misalkan  $H$  adalah fungsi distribusi gabungan dari variable  $X$  dan  $Y$ , dengan  $F$  dan  $G$  masing-masing adalah fungsi distribusi marginal dari  $X$  dan  $Y$ . Maka terdapat sebuah copula  $C$  sedemikian sehingga untuk setiap  $x, y \in \bar{R}$  berlaku

$$H(x, y) = C(F(x), G(y)) = C(u, v), \tag{2.4}$$

dengan  $u = F(x)$  dan  $v = G(y)$ .

Jika  $F$  dan  $G$  kontinu, maka copula  $C$  tunggal, jika  $F$  dan  $G$  tidak kontinu, maka copula  $C$  tunggal pada  $\text{Range}(F) \times \text{Range}(G)$ .

Sebaliknya, misalkan  $C$  adalah sebuah copula,  $F$  dan  $G$  masing-masing adalah fungsi distribusi marginal dari  $X$  dan  $Y$ . Maka terdapat fungsi distribusi gabungan  $H$  sedemikian sehingga untuk setiap  $x, y \in \bar{R}$  berlaku

$$H(x, y) = C(F(x), G(y)) = C(u, v).$$

Sebagai konsekuensi dari teorema Sklar, jika  $X$  dan  $Y$  adalah variabel acak dengan fungsi distribusi gabungan  $H$  dan mempunyai fungsi distribusi marginal masing-masing adalah  $F$  dan  $G$ , maka untuk setiap  $x, y \in \bar{R}$  berlaku

$$\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y)). \tag{2.5}$$

### 2.2. Copula Empiris

**Definisi 2.** Misalkan  $\{(x_k, y_k)\}_{k=1}^n$  adalah sampel berukuran  $n$  dari distribusi bivariat yang kontinu. Copula empiris adalah fungsi  $C_n$  yang didefinisikan sebagai

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{\text{banyaknya pasangan data } (x, y) \text{ dalam sampel sehingga } x \leq x_{(i)} \text{ dan } y \leq y_{(j)}}{n}$$

di mana  $x_{(i)}$  dan  $y_{(j)}$ ,  $1 \leq i, j \leq n$ , adalah statistik urutan dari sampel. Copula frekuensi empiris adalah fungsi  $c_n$  yang didefinisikan sebagai

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} 1/n, & \text{jika } (x_{(i)}, y_{(j)}) \text{ adalah elemen dari } \{(x_k, y_k)\}_{k=1}^n, \\ 0, & \text{lainnya.} \end{cases} \tag{2.6}$$

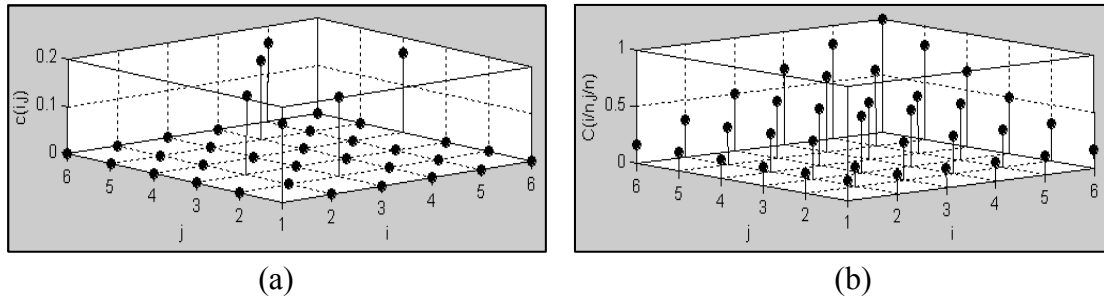
**Ilustrasi:**

Misalkan diberikan sampel acak  $\{(x_k, y_k)\}_{k \in N}$  dari variabel acak  $(X, Y)$ ,

TABEL 1. Data Sampel Acak

$K$	1	2	3	4	5	6
$X$	1.3	2.4	3.2	1.7	4.3	0.8
$Y$	2.5	3.5	2.6	2.1	3.2	1.8

Berdasarkan definisi 2, maka grafik dari *copula* frekuensi empiris  $c\left(\frac{i}{n}, \frac{j}{n}\right)$  dan  $C\left(\frac{i}{n}, \frac{j}{n}\right)$ ,  $i, j = 1, 2, \dots, 6$  dari variabel  $X$  dan  $Y$  adalah:



GAMBAR 1. (a) Grafik *copula* frekuensi empiris  $c\left(\frac{i}{n}, \frac{j}{n}\right)$ ; (b) Grafik *copula* frekuensi empiris  $C\left(\frac{i}{n}, \frac{j}{n}\right)$

### 3. Dependensi Autokorelasi

Berdasarkan pembatasan masalah pada pendahuluan, pembahasan akan difokuskan dalam mempelajari perbandingan kuantifikasi dependensi dalam data deret waktu (*time series*), yang biasanya disebut dengan istilah autokorelasi, dan lebih khusus lagi untuk kuantifikasi autokorelasi lag-1. Bentuk klasik yang umum digunakan dalam kuantifikasi autokorelasi ini adalah koefisien autokorelasi dari fungsi autokorelasi (Autocorrelation Coefficient Function/ACF). Bentuk kuantifikasi dependensi yang lain adalah bentuk modern yang menggunakan konsep konkordan, yang salah satunya kuantifikasi autokorelasi dengan pendekatan *copula*.

#### 3.1. Autokorelasi

Suatu himpunan hasil pengamatan yang dilakukan berdasarkan urutan waktu, biasa disebut *time series*. Data *time series* yang dibahas dalam tulisan ini adalah berkaitan dengan data *time series* diskrit. Adapun suatu fenomena statistika yang berkembang dalam kaitannya dengan runtutan waktu yang sesuai dengan aturan probabilistik di sebut proses stokastik. Jadi analisa dari *time series*, dianggap sebagai realisasi dari proses stokastik. Salah satu proses yang khusus dalam proses stokastik, disebut proses stasioner (Box & Jenkins, 1976).

**Definisi 3.** Misalkan  $Z = (Z_t)_{t \in T}$  adalah proses stokastik pada ruang probabilistik  $(\Omega, F, P)$ . Maka  $X$  dikatakan stasioner kuat, jika untuk setiap  $m \in N$ ,  $\{t_1, t_2, \dots, t_m\} \subset T$  dan setiap  $h > 0$  dengan  $\{t_1 + h, t_2 + h, \dots, t_m + h\} \subset T$ , kita punya

$$P(Z_{t_1+h} < z_1, \dots, Z_{t_m+h} < z_m) = P(Z_{t_1} < z_1, \dots, Z_{t_m} < z_m) \quad (3.1)$$

untuk setiap  $z_1, z_2, \dots, z_m \in R$ .

Jika  $m = 1$ , asumsi kestasioneran berakibat bahwa fungsi distribusi peluang  $f(z_t)$  adalah sama untuk setiap  $t \in T$ , dan cukup menuliskan  $f(z)$ . Oleh karena itu proses stasioner mempunyai mean konstan:

$$\mu = E[Z_t] = E[Z] = \int_{-\infty}^{\infty} Zf(z).dz \quad (3.2)$$

dan variansi konstan:

$$\sigma_z^2 = E\left[(Z_t - \mu)^2\right] = E\left[(Z - \mu)^2\right] = \int_{-\infty}^{\infty} (Z - \mu)^2 f(z) dz . \quad (3.3)$$

Dalam praktek, mean dan variansi dari proses stationer ditaksir dengan

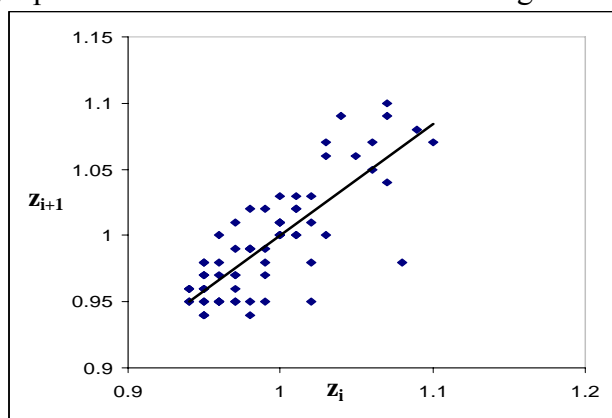
$$\hat{\mu} = \bar{z} = \frac{1}{N} \sum_{t=1}^N z_t \quad (3.4)$$

dan

$$\hat{\sigma}_z^2 = \frac{1}{N} \sum_{t=1}^N (z_t - \bar{z})^2 \quad (3.5)$$

di mana  $z_1, z_2, \dots, z_N$  adalah pengamatan/sampel *time series*.

Asumsi kestasioneran juga berakibat bahwa fungsi distribusi peluang gabungan  $f(z_{t_1}, z_{t_2})$  adalah sama untuk setiap  $t_1, t_2$  yang mana merupakan interval konstan yang terpisah. Sehingga, karakteristik distribusi gabungan ini dapat diduga dengan memplot diagram pencar dari data pasangan  $(z_{t_1}, z_{t_2}) = (z_t, z_{t+k})$  yang merupakan bagian dari data *time series* yang dipisahkan oleh interval konstan atau lag  $k$ .



GAMBAR 2. Contoh diagram pencar untuk data *time series* dengan lag-1

Dalam hal ini kita dapat berbicara mengenai kovariansi dari  $z_t$  dan  $z_{t+k}$ , yang dipisahkan oleh  $k$  interval waktu diskrit, yang disebut autokovariansi pada lag  $k$  yang didefinisikan oleh:

$$\gamma_k = \text{cov}[Z_t, Z_{t+k}] = E\left[(Z_t - \mu)(Z_{t+k} - \mu)\right] \quad (3.6)$$

Sehingga, dari definisi autokovariansi, kita bisa definisikan kuantifikasi autokorelasi pada lag  $k$ , didefinisikan oleh:

$$\begin{aligned} \rho_k &= \frac{E\left[(Z_t - \mu)(Z_{t+k} - \mu)\right]}{\sqrt{E\left[(Z_t - \mu)^2\right]E\left[(Z_{t+k} - \mu)^2\right]}} \\ &= \frac{E\left[(Z_t - \mu)(Z_{t+k} - \mu)\right]}{\sigma_z^2} \end{aligned} \quad (3.7)$$

Dalam praktek, kuantifikasi autokorelasi di atas ditaksir oleh:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2} \quad (3.8)$$

Bentuk terakhir pada persamaan (3.8) merupakan bentuk klasik dari dependensi time series yang disebut koefisien autokorelasi (*Autocorrelation Function*).

### 3.2. Kendall's Tau dan Copula

Pada bagian ini akan dijelaskan bentuk kuantifikasi dependensi lain selain koefisien autokorelasi, salah satunya adalah statistik *Kendall's tau*, yaitu kuantifikasi dependensi yang didasarkan atas data rangking.

**Definisi 4.** Misalkan  $(X_1, Y_1)$  dan  $(X_2, Y_2)$  dua vektor acak yang independen dan berdistribusi identik pada ruang peluang  $(\Omega, A, P)$ . Kendall's  $\tau$  didefinisikan sebagai.

$$\tau \equiv \tau_{X,Y} = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0) \quad (3.9)$$

Selanjutnya kita menyebut bentuk di atas adalah definisi Kendall's  $\tau$  untuk populasi. Jadi, Kendall's  $\tau$  adalah perbedaan antara peluang dari konkordan dan peluang dari diskordan.

Dalam prakteknya, kita dapat mendefinisikan ukuran dependensi Kendall's  $\tau$  berdasarkan sampel. Misalkan  $\{(x_1, y_1), \dots, (x_n, y_n)\}$ ,  $n \geq 2$  adalah sampel berukuran  $n$  dari vektor acak kontinu  $(X, Y)$ . Setiap pasang sampel  $\{(x_i, y_i), (x_j, y_j)\}$ ,  $i, j \in \{1, \dots, n\}$ ,  $i \neq j$  merupakan suatu diskordan atau konkordan. Maka jelas terdapat  $\binom{n}{2}$  pasangan berbeda dari sampel yang ada. Misalkan  $K$  menyatakan banyaknya pasangan konkordan, dan  $D$  menyatakan banyaknya pasangan diskordan. Maka Kendall's  $\tau$  untuk sampel didefinisikan menjadi

$$\hat{\tau} = \frac{K - D}{K + D} = \frac{K - D}{\binom{n}{2}} \quad (3.10)$$

Dengan definisi Kendall's  $\tau$  di atas, kita dapat menunjukkan bahwa *copula* mempunyai hubungan dengan Kendall's  $\tau$ , untuk menunjukkan hubungan tersebut, sebelumnya perlu didefinisikan terlebih dahulu suatu fungsi konkordan  $Q$ , yang menyatakan perbedaan peluang dari konkordan dan peluang diskordan antara dua vektor  $(X_1, Y_1)$  dan  $(X_2, Y_2)$  dari variabel acak kontinu dengan fungsi distribusi gabungan (yang mungkin) berbeda  $H_1$  dan  $H_2$ , tetapi dengan fungsi distribusi marginal yang sama  $F$  dan  $G$ . Kemudian akan ditunjukkan bahwa fungsi konkordan ini bergantung pada distribusi dari  $(X_1, Y_1)$  dan  $(X_2, Y_2)$  melalui *copula* mereka.

**Teorema 2.** Misalkan  $(X_1, Y_1)$  dan  $(X_2, Y_2)$  adalah dua vektor random dengan fungsi distribusi gabungan masing-masing  $H_1$  dan  $H_2$ , di mana  $X_i \sim F$  dan  $Y_i \sim G$ ,  $i = 1, 2$ . Lebih lanjut, misalkan  $C_1$  dan  $C_2$  menyatakan *copula* dari  $(X_1, Y_1)$  dan  $(X_2, Y_2)$ , sedemikian sehingga  $H_1(x, y) = C_1(F(x), G(y))$  dan  $H_2(x, y) = C_2(F(x), G(y))$ .

Jika  $Q$  menyatakan perbedaan antara peluang dari konkordan dan peluang diskordan dari  $(X_1, Y_1)$  dan  $(X_2, Y_2)$ , yang didefinisikan sebagai

$$Q = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0), \quad (3.11)$$

maka kita peroleh:

$$Q = Q(C_1, C_2) = 4 \iint_{\mathcal{I}^2} C_2(u, v) dC_1(u, v) - 1. \quad (3.12)$$

Berdasarkan definisi fungsi konkordan pada teorema 2, maka kita dapat mendefinisikan Kendall's  $\tau$  untuk  $X$  dan  $Y$  melalui copula dengan teorema berikut (Nelsen, 1999):

**Teorema 3.** Misalkan  $X$  dan  $Y$  variabel acak kontinu dengan copula  $C$ . Maka Kendall's  $\tau$  untuk  $X$  dan  $Y$  diberikan oleh

$$\tau_{X,Y} \equiv \tau_C = Q(C, C) = 4 \cdot \iint_{I^2} C(u, v) dC(u, v) - 1. \quad (3.20)$$

Perhatikan bahwa bentuk integral yang ada pada persamaan (3.20) dapat diinterpretasikan sebagai ekspektasi dari fungsi  $C(U, V)$ , di mana  $U$  dan  $V$  variabel acak yang berdistribusi  $U(0,1)$ , atau dengan kata lain

$$\tau_C = 4 \cdot E_C [C(U, V)] - 1. \quad (3.21)$$

**Teorema 4.** Misalkan  $X$  dan  $Y$  variabel acak kontinu dengan fungsi distribusi gabungan  $H$ , dan misalkan

$$T = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{y'} \int_{-\infty}^{x'} [h(x, y)h(x', y') - h(x, y')h(x', y)] dx dy dx' dy'. \quad (3.22)$$

Maka Kendall's  $\tau$  untuk  $X$  dan  $Y$  diberikan oleh  $\tau_{X,Y} = 2T$ .

Berdasarkan definisi copula empiris, dan dengan memperhatikan kembali definisi Kendall's  $\tau$  untuk populasi untuk suatu variabel acak kontinu  $X$  dan  $Y$  dengan copula  $C$  seperti pada persamaan 3.22 dengan modifikasi yaitu

$$\tau = 2 \int_0^1 \int_0^1 \int_0^{v'} \int_0^{u'} [c(u, v)c(u', v') - c(u', v)c(u, v')] dudvdu' dv', \quad (3.23)$$

Maka teorema berikut (Nelsen, 1999) akan menjelaskan bentuk turunan koefisien Kendall's  $\tau$  untuk sampel.

**Teorema 5.** Misalkan  $C_n$  dan  $c_n$  masing-masing adalah fungsi copula empirik dan fungsi frekuensi copula empirik untuk sampel  $\{(x_k, y_k)\}_{k=1}^n$ . Jika  $t$  adalah koefisien Kendall's  $\tau$  untuk sampel, maka

$$t = \frac{2n}{n-1} \sum_{i=2}^n \sum_{j=2}^n \sum_{p=1}^{i-1} \sum_{q=1}^{j-1} \left[ c_n \left( \frac{i}{n}, \frac{j}{n} \right) c_n \left( \frac{p}{n}, \frac{q}{n} \right) - c_n \left( \frac{i}{n}, \frac{q}{n} \right) c_n \left( \frac{p}{n}, \frac{j}{n} \right) \right]. \quad (3.24)$$

### 3.3. Kendall's Tau untuk Autokorelasi

Pada subbab ini, akan dijelaskan pendekatan Kendall's  $\tau$  untuk kuantifikasi autokorelasi dari data *time series*, dan dikhususkan untuk autokorelasi lag-1 (*first-order serial dependence*).

Misalkan diberikan barisan variabel acak  $X_1, X_2, \dots, X_n$ ,  $n \geq 3$  yang merupakan data *time series* dan  $R_1, R_2, \dots, R_n$  adalah barisan ranking yang bersesuaian dengan barisan variabel acak, maka ukuran autokorelasi lag-1 secara khusus didasarkan pada data pasangan

$$(R_1, R_2), (R_2, R_3), \dots, (R_{n-1}, R_n), \quad (3.25)$$

dan mungkin menambah dengan  $(R_n, R_1)$ , dalam kasus barisan variabel acak tersebut bersiklus.

Misalkan diberikan barisan variabel acak  $X_1, X_2, \dots, X_n$ ,  $n \geq 3$  yang bersesuaian dengan barisan ranking  $R_1, R_2, \dots, R_n$  dari barisan acak. Maka kuantifikasi Kendall's  $\tau$  untuk autokorelasi lag-1 untuk kasus barisan bersiklus dapat didefinisikan sebagai:

$$\tau_n = 1 - \left( 2N / \binom{n}{2} \right) = 1 - \frac{4D}{n \cdot (n-1)}, \quad (3.26)$$

di mana  $D$  menyatakan banyaknya diskordan, atau

$$\begin{aligned}
 D &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \{I(R_i < R_j, R_{i+1} > R_{j+1}) + I(R_i > R_j, R_{i+1} < R_{j+1})\} \\
 &= \sum_{i=1}^n \sum_{j=1}^n (I(R_i < R_j, R_{i+1} > R_{j+1}))
 \end{aligned}
 \tag{3.27}$$

di mana  $I(A)$  menyatakan fungsi indikator dari himpunan  $A$ . Untuk kasus barisan yang tidak bersiklus, dengan mensubstitusikan  $n-1$  untuk  $n$  pada persamaan (3.26) dan (3.27).

### 3.4. Pengujian Keberartian Autokorelasi

Setelah melakukan kuantifikasi autokorelasi, untuk dapat membandingkan hasil yang telah diperoleh dari masing-masing ukuran dependensi yang telah dijelaskan pada subbab sebelumnya, perlu dilakukan pengujian keberartian dependensi dari masing-masing ukuran dependensi, artinya hasil kuantifikasi autokorelasi harus diuji keberartiannya.

### 3.5. Pengujian untuk Koefisien Autokorelasi Klasik

Misalkan  $Z = (Z_1, Z_2, \dots, Z_n)$ ,  $n \in N$  adalah proses stationer dengan waktu diskrit berukuran  $n$ , dan  $\hat{\rho}_1$  adalah taksiran untuk koefisien autokorelasi lag-1, maka Arcana (2005) mengatakan bahwa pada tingkat keberartian  $\alpha$  pada tingkat keberartian  $\alpha$ , nilai  $\hat{\rho}_1$  dikatakan berarti untuk pengujian dua arah jika

$$|\hat{\rho}_1| > \left| \frac{Z_{\alpha/2}}{\sqrt{n}} \right|, \tag{3.28a}$$

sedangkan, untuk pengujian satu arah, jika

$$\hat{\rho}_1 > \frac{Z_{1-\alpha}}{\sqrt{n}} \text{ atau } \hat{\rho}_1 < \frac{Z_{\alpha}}{\sqrt{n}}, \tag{3.28b}$$

di mana  $Z_{\alpha}$  adalah suatu nilai sehingga  $P(Z < Z_{\alpha}) = \alpha$  dengan  $Z \sim N(0,1)$ .

### 3.6. Pengujian untuk Koefisien Autokorelasi Kendall's $\tau$

Misalkan  $Z = (Z_1, Z_2, \dots, Z_n)$ ,  $n \in N$  adalah proses stationer dengan waktu diskrit berukuran  $n$ , dan  $\hat{\tau}_n$  adalah taksiran Kendall's  $\tau$  untuk koefisien autokorelasi lag-1, maka Genest & Ferguson (1999) mendefinisikan

$$T_n = \frac{\hat{\tau}_n - E[\hat{\tau}_n]}{\sqrt{Var(\hat{\tau}_n)}}, \tag{3.29}$$

di mana  $E[\hat{\tau}_n] = 0$  adalah mean dari  $\hat{\tau}_n$  yang didefinisikan,

$$E[\hat{\tau}_n] = \frac{2}{3(n-1)}, \quad n \geq 3, \tag{3.30}$$

untuk kasus proses stasioner yang bersiklus maupun tidak bersiklus. Dan  $Var(\hat{\tau}_n)$  adalah variansi dari  $\hat{\tau}_n$  yang berbeda untuk kasus proses stasioner yang bersiklus dan tidak bersiklus. Untuk kasus proses stasioner yang bersiklus  $Var(\hat{\tau}_n) = 0$ , untuk  $n = 3$ , dan

$$Var(\hat{\tau}_n) = \frac{20n^3 - 14n^2 - 98}{45n^2(n-1)^2}, \text{ untuk } n \geq 4. \tag{3.31}$$

Sedangkan untuk kasus proses stasioner yang tidak bersiklus  $Var(\hat{\tau}_n) = 8/9$ , untuk  $n = 3$ , dan

$$Var(\hat{\tau}_n) = \frac{20n^3 - 74n^2 + 54n + 148}{45(n-1)^2(n-2)^2}, \text{ untuk } n \geq 4. \quad (3.32)$$

Selanjutnya untuk pengujian keberartian nilai  $\hat{\tau}_n$ , Genest & Ferguson (1999) mengatakan pada tingkat keberartian  $\alpha$ , nilai  $\hat{\tau}_n$  dikatakan berarti untuk pengujian dua arah, jika

$$|T_n| > |t_{\alpha,n}|, \quad (3.33a)$$

sedangkan, untuk pengujian satu arah, jika

$$T_n > t_{\alpha,n} \text{ atau } T_n < -t_{\alpha,n} \quad (3.33b)$$

di mana  $t_{\alpha,n}$  adalah suatu nilai sehingga  $P(T_n > t_{\alpha,n}) = \alpha$  dengan  $T_n \sim t_n$ .

## 4. Simulasi dan Hasil Perbandingan Kuantifikasi Autokorelasi

### 4.1. Desain Simulasi

Simulasi ini dilakukan untuk membandingkan hasil kuantifikasi autokorelasi klasik dan kuantifikasi autokorelasi Kendall's tau melalui *copula* empirik. Adapun desain dari simulasi yang dilakukan adalah membangun barisan data yang mengikuti model proses stasioner, simulasi ini dilakukan untuk barisan data berukuran  $n=10, 15, \text{ dan } 20$ . Cara membangun barisan data dalam simulasi ini adalah sebagai berikut:

Misalkan  $e_i, i=1,2,\dots,n$  adalah data yang dibangkitkan secara acak dari distribusi normal  $(0,1)$ , maka desain

$$X_1 = (1+\theta^2)^{-1/2} e_1 \quad \text{dan} \quad X_i = \theta X_{i-1} + e_i, \quad i=2,\dots,n \quad (4.1)$$

di mana nilai  $\theta$  disimulasikan untuk beberapa nilai  $\theta = (1/2)^j, j=1,2,\dots,5$  dan  $\theta=0$ .

Selanjutnya, kuantifikasi autokorelasi dari barisan data yang ada dilakukan dengan metode yang telah dijelaskan sebelumnya. Percobaan ini lakukan sebanyak 500 kali untuk setiap ukuran  $n$ , kemudian dilakukan pengujian keberartian autokorelasi berdasarkan hipotesis nol yang menyatakan bahwa tidak terdapat autokorelasi yang berarti pada barisan data simulasi.

Perbandingan terhadap hasil kuantifikasi autokorelasi untuk masing-masing metode dilakukan dengan menghitung prosentase penolakan hipotesis nol. Sehingga dalam simulasi ini diperlukan data acak berdistribusi normal  $(0,1)$  sebanyak  $500 \times 6 \times (10+15+20) = 135000$  data.

### 4.2. Hasil Simulasi

Setelah dilakukan simulasi berdasarkan desain di atas, maka hasilnya dapat dilihat pada tabel 2. Dari tabel tersebut, dapat dikatakan secara keseluruhan, jika nilai  $\theta$  semakin mendekati nol, maka prosentase penolakan hipotesis nol yang menyatakan bahwa autokorelasi pada model proses stasioner cenderung menurun.

Dalam batas nilai  $\theta$  yang digunakan dalam simulasi ini, perbandingan antara kuantifikasi autokorelasi klasik dengan kuantifikasi autokorelasi Kendall's tau, dapat dikatakan mengenai beberapa hal, yaitu:

1. Untuk nilai  $\theta = 1/2, 1/4, 1/8$ , prosentase penolakan hipotesis nol berdasarkan kuantifikasi autokorelasi Kendall's tau melalui *copula* lebih tinggi dibanding prosentase penolakan hipotesis nol berdasarkan kuantifikasi autokorelasi klasik.



2. Sebaliknya untuk nilai  $\theta = \frac{1}{16}, \frac{1}{32}, 0$ , prosentase penolakan hipotesis nol berdasarkan kuantifikasi autokorelasi klasik lebih tinggi dibanding prosentase penolakan hipotesis nol berdasarkan kuantifikasi autokorelasi Kendall's tau melalui *copula*.
3. Hal ini menunjukkan bahwa untuk nilai  $\theta$  tertentu kuantifikasi autokorelasi Kendall's tau melalui *copula* dikatakan lebih baik dibanding dengan kuantifikasi autokorelasi klasik, sedangkan untuk  $\theta$  yang mendekati nol, kuantifikasi autokorelasi klasik dikatakan lebih baik dibanding dengan kuantifikasi autokorelasi Kendall's tau melalui *copula*.

TABEL 2. Prosentase Penolakan Hipotesis Non pada Simulasi Pengujian Autokorelasi

Statistic	N	Theta					
		1/2	1/4	1/8	1/16	1/32	0
<i>Autocorrelation Function</i>	10	15.4%	3.0%	2.0%	0.8%	1.6%	0.2%
<i>Kendall' Tau dg Copula</i>	10	22.6%	8.4%	6.2%	3.6%	3.4%	2.8%
<i>Autocorrelation Function</i>	15	32.6%	11.8%	4.0%	4.0%	3.2%	0.6%
<i>Kendall' Tau dg Copula</i>	15	38.8%	18.4%	8.0%	7.2%	6.6%	4.6%
<i>Autocorrelation Function</i>	20	49.4%	15.4%	6.0%	3.2%	4.0%	1.2%
<i>Kendall' Tau dg Copula</i>	20	54.8%	20.8%	10.6%	7.4%	5.6%	5.2%

### 5. Kesimpulan dan Saran

Berdasarkan penjelasan teori dan hasil simulasi yang dilakukan pada bagian sebelumnya, penulis mencoba mengambil beberapa kesimpulan sebagai berikut:

1. Kendall's tau melalui *copula* adalah metode alternatif yang dapat digunakan dalam kuantifikasi dependensi antara dua variabel acak, lebih khusus dapat digunakan dalam kuantifikasi autokorelasi *lag-1*.
2. Pada model  $X_i = \theta.X_{i-1} + e_i, i = 1, 2, \dots, n$ , untuk batas  $\theta$  tertentu, kuantifikasi autokorelasi Kendall's tau melalui *copula* dikatakan lebih baik dibanding dengan kuantifikasi autokorelasi klasik, sedangkan untuk  $\theta$  yang mendekati nol, kuantifikasi autokorelasi klasik dikatakan lebih baik dibanding dengan kuantifikasi autokorelasi Kendall's tau melalui *copula*.
3. Untuk simulasi, perlu adanya penelitian lebih lanjut mengenai pemilihan model *time series* yang lain serta penentuan batas nilai  $\theta$  yang digunakan dalam model.

### Daftar Pustaka

- Arcana, I Nyoman. (2005), Batch Process, How to Measure a Process Capability with A Better Way: A Case Study at Soft Drink Factory, *J. Int. Conf. of App. Math. (ICAMs)*, Bandung, Indonesia, pp 5-6.
- Box, G.E., dan Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco, pp 23-28.
- Genest, C., dan Ferguson, T., (1999), Kendall's Tau for Autocorrelation, *Departement of Statistics Papers of University of California*, Los Angeles, pp 1-9, 11, 12.
- Jogde, K., (1982), Concepts of Dependence, in *Encyclopedia of Statistical Sciences*, Vol.1, S. Kotz dan N.L, Johnson, editor, John Wiley & Sons, New York.
- Nelsen, B. Roger, (1999), *An Introduction to Copula*, Spinger-Verlag, New York.
- Schmitz, V., (2003), Copulas and Stochastic Processes, *Disertasi Program Doktor*, Shaker Verlag, Aachen.