Algoritma Rumus Statify

Modul Analisis *Time Series*

Deskripsi

Modul ini dibuat untuk membantu pengguna dalam memahami algoritma, rumus, serta konsep yang digunakan dalam analisis data runtun waktu (*time series*) dalam Aplikasi Berbasis Web Statify. Sebagai tambahan informasi, modul ini dikembangkan menggunakan bahasa pemrograman Rust. Analisis deret waktu yang tersedia pada modul ini masih terbatas pada analisis deret waktu univariat yang dilengkapi fitur untuk persiapan analisis awal, uji asumsi data deret waktu, hingga pemodelan data deret waktu dengan daftar fitur sebagai berikut:

1. *Smoothing*
2. *Decomposition*
3. *Autocorrelation*
4. *Unit Root Test*
5. *Box-Jenkins Model*

Smoothing

Deskripsi

Smoothing adalah suatu teknik dalam analisis deret waktu yang bertujuan untuk melakukan peramalan dengan menggunakan data periode sebelumnya sebagai dasar untuk meramal data di masa depan. Pada modul ini metode *smoothing* yang tersedia terdiri dua metode utama yaitu, metode *Moving Average* dan Exponential Smoothing. Selanjutnya metode *Moving Average* terdiri atas *Simple Moving Average* dan *Double Moving Average* sedangkan *Exponential Smoothing* terdiri atas metode *Simple Exponential Smooting*, *Double Exponential Smoothing*, *Holt's Exponential Smoothing*, *Winter's Exponential Smoothing*.

Metode *Smoothing*

1. ***Simple Moving Average***

   Simple Moving Average adalah salah satu metode *smoothing* yang merata-ratakan sejumlah data dalam rentang tertentu untuk meramal data di masa mendatang. Berikut rumus untuk menghitung *simple moving average* pada fitur ini.

   $$M_t = \hat{Y}_{t+1} = \frac{1}{d}\sum_{i=0}^{d-1} Y_{t-i}$$

   Dimana

   $$t = d, d+1, d+2, \dots, n$$

   Keterangan:
   M_t = Nilai *moving average* periode ke-t
   $\hat{Y}_{t+1}$ = Data ramalan satu periode selanjutnya setelah t
   Y_t = Data awal periode ke-t
   d = Jumlah jarak yang digunakan untuk menghitung rata-rata (*average*)
   n = Jumlah periode data

2. ***Double Moving Average***

   Double Moving Average adalah pengembangan dari metode *simple moving average*, yakni melakukan *moving average* sebanyak dua kali lalu menghitung nilai tren yang menyerupai bentuk regresi linear sederhana dalam menghitung data setelahnya.

   $$M_t = \frac{1}{d}\sum_{i=0}^{d-1} Y_{t-i}$$

   $$M'_t = \frac{1}{d}\sum_{i=0}^{d-1} M_{t-i}$$

   $$a_t = 2M_t - M'_t$$

   $$b_t = \frac{2}{d-1}(M_t - M'_t)$$

   $$\hat{Y}_{t+p} = a_t + b_t p$$

   Keterangan:
   M_t = Nilai *moving average* pertama periode ke-t

M'_t = Nilai *moving average* kedua periode ke-t
a_t = Nilai selisih antara dua *moving average* periode ke-t
b_t = Nilai perubahan atau gradien antara *moving average* periode ke-t
$\hat{Y}_{t+p}$ = Data ramalan p periode selanjutnya setelah t
Y_t = Data awal periode ke-t
d = Jumlah jarak yang digunakan untuk menghitung rata-rata (*average*)

3. ***Simple Exponential Smoothing***

Simple Exponential Smoothing adalah metode *smoothing* yang melakukan peramalan menggunakan pembobotan nilai sekarang dan nilai hasil peramalan masa sebelumnya. Pembobotan ini menggunakan satu parameter yang umumnya sering disebut parameter alpa. Berikut rumus untuk melakukan proses *simple exponential smoothing*.

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$$

Keterangan:
$\hat{Y}_{t+1}$ = Data ramalan baru satu periode selanjutnya
α = Konstanta *smoothing* ($0 < \alpha < 1$)
Y_t = Data awal periode ke-t
$\hat{Y}_t$ = Data ramalan lama satu periode sebelum data ramalan baru

4. ***Double Exponential Smoothing***

Double Exponential Smoothing adalah metode peramalan yang didasarkan pada metode *brown* yang meramal data deret waktu dengan menggunakan *trend* linear yang memiliki kesamaan konsep dengan *double moving average*. Berikut rumus yang digunakan pada *double exponential smoothing*.

$$A_t = \alpha Y_t + (1-\alpha)A_{t-1}$$
$$A'_t = \alpha A_t + (1-\alpha)A'_{t-1}$$

$$a_t = 2A_t - A'_t$$
$$b_t = \frac{\alpha}{1-\alpha}(A_t - A'_t)$$

$$\hat{Y}_{t+p} = a_t + b_t p$$

Keterangan:
A_t = Nilai *exponential smoothing* pertama periode ke-t
A'_t = Nilai *exponential smoothing* kedua periode ke-t
α = Konstanta *smoothing* ($0 < \alpha < 1$)
a_t = Nilai selisih antara dua *exponential smoothing* periode ke-t
b_t = Nilai perubahan atau gradien antara *exponential smoothing* periode ke-t
$\hat{Y}_{t+p}$ = Data ramalan p periode selanjutnya setelah t
Y_t = Data awal periode ke-t

5. ***Holt's Exponential Smoothing***

Holt's Exponential Smoothing adalah metode *smoothing* yang penghitungannya menggunakan dua parameter. Parameter pertama digunakan untuk menghitung *exponential smoothing* dasar untuk setiap periodenya sedangkan parameter kedua digunakan untuk menghitung nilai tren dari *exponential smoothing* sebelumnya untuk setiap periode. Berikut rumus yang digunakan untuk menghitung *holt's exponential smoothing*.

Exponential smoothing awal

$$A_t = \alpha Y_t + (1-\alpha)(A_{t-1} - T_{t-1})$$

Estimasi *trend*

$$T_t = \beta(A_t - A_{t-1}) + (1-\beta)T_{t-1}$$

Peramalan p periode selanjutnya

$$\hat{Y}_{t+p} = A_t + pT_t$$

Keterangan:

A_t = Nilai *exponential smoothing* periode ke-t

T_t = Nilai estimasi *trend* periode ke-t

α = Konstanta *smoothing* ($0 < \alpha < 1$)

β = Konstanta *trend* ($0 < \beta < 1$)

$\hat{Y}_{t+p}$ = Data ramalan p periode selanjutnya setelah t

Y_t = Data awal periode ke-t

6. *Winter's Exponential Smoothing*

Winter's Exponential Smoothing adalah suatu pengembangan dari teknik *exponential smoothing* yang sebelumnya yang menambahkan faktor dan parameter musim dalam penghitungannya. Berikut adalah rumus untuk menghitung *winter's exponential smoothing*.

exponential smoothing awal

$$A_t = \alpha \frac{Y_t}{S_{t-L}} + (1-\alpha)(A_{t-1} - T_{t-1})$$

Estimasi *trend*

$$T_t = \beta(A_t - A_{t-1}) + (1-\beta)T_{t-1}$$

Estimasi musim

$$S_t = \gamma \frac{Y_t}{A_t} + (1-\gamma)S_{t-L}$$

Peramalan p periode selanjutnya

$$\hat{Y}_{t+p} = (A_t + pT_t)S_{t-L+P}$$

Dikarenakan penghitungan menggunakan panjang periode musim maka untuk nilai awal yang periodenya kurang dari periode musim (t < L) perlu dilakukan inisiasi sebagai berikut.

exponential smoothing awal

$$A_1 = Y_1;\ A_t = \alpha \frac{Y_t}{1.0} + (1-\alpha)(A_{t-1} - T_{t-1})$$

Estimasi *trend*

$$T_1 = 0.0;\ T_t = \beta(A_t - A_{t-1}) + (1-\beta)T_{t-1}$$

Estimasi musim

$$S_1 = 1.0;\ S_t = \gamma \frac{Y_t}{A_t} + (1-\gamma)1.0$$

Keterangan:

A_t = Nilai *exponential smoothing* periode ke-t

T_t = Nilai estimasi *trend* periode ke-t
S_t = Nilai estimasi musim periode ke-t
α = Konstanta *smoothing* (0 < α < 1)
β = Konstanta *trend* (0 < β < 1)
γ = Konstanta musim (0 < γ < 1)
L = Panjang periode musim
$\hat{Y}_{t+p}$ = Data ramalan p periode selanjutnya setelah t
Y_t = Data awal periode ke-t

Evaluasi *Smoothing*

1. ***Mean Square Error* (MSE)**

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2$$

Keterangan:
n = Jumlah periode data
Y_t = Data awal periode ke-t
F_t = Data hasil peramalan (*forecasting*) ke-t

2. ***Root Mean Square Error* (RMSE)**

$$RMSE = \sqrt{MSE}$$

3. ***Mean Absolute Errot* (MAE)**

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|Y_t - F_t|$$

4. ***Mean Percentage Error* (MPE)**

$$MPE = \frac{1}{n}\sum_{t=1}^{n}\left[\left(\frac{Y_t - F_t}{Y_t}\right) \times 100\right]$$

5. ***Mean Absolute Percentage Error* (MAPE)**

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\left(\frac{Y_t - F_t}{Y_t}\right) \times 100\right|$$

Referensi

Hanke, J. E., & Wichern, D. W. (1995). *Business forecasting* (5th ed.). Pearson Prentice Hall.

Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1983). *Forecasting: Methods and applications*. New York: John Wiley and Sons.

Decomposition

Deskripsi

Decomposition (Dekomposisi) adalah suatu teknik peramalan data runtun waktu dengan memisahkan data runtun waktu menjadi empat komponen, di antaranya adalah komponen *trend*, *cyclical*, *seasonal*, dan *irrengular*. Komponen *trend* adalah suatu komponen data runtun waktu yang menggambarkan suatu pola garis *trend*. Lalu Komponen *seasonal* sendiri adalah suatu komponen data runtun waktu yang menggambarkan pola dalam kurun waktu tertentu yang terjadi secara berulang, biasanya terlihat dalam data kurun waktu bulanan, triwulanan, ataupun mingguan. Sedangkan komponen *cyclical* merupakan komponen yang menunjukkan pola dalam kurun waktu tertentu namun dalam jangka waktu yang lebih panjang sehingga dalam praktisnya komponen *cyclical* sering dianggap menjadi satu kesatuan dengan *trend* atau disebut komponen *trend-cycle*. Terakhir, komponen *irrengular* adalah komponen data runtun waktu yang sifatnya acak dan tidak memiliki pola karena merupakan sisa pengurangan dari ketiga komponen sebelumnya sehingga komponen ini sering disebut juga *error*. Dengan demikian, dekomposisi data runtun waktu memiliki bentuk umum sebagai berikut.

Bentuk Multiplikatif

$$Y_t = T_t \times S_t \times I_t$$

Bentuk Aditif

$$Y_t = T_t + S_t + I_t$$

Pada modul ini fitur dekomposisi menggunakan metode dekomposisi klasik yang dijelaskan di Makridakis, Spyros G., 1997.

Algoritma *Decomposition*

1. *Multiplicative Decomposition*

Berikut tahapan dekomposisi data runtun waktu bentuk multiplikatif.

1.1. Menghitung *Moving Average*

Pada tahapan ini data akan dimuluskan (*smoothing*) menggunakan teknik *simple moving average* jika periode data bernilai ganjil dan *centered moving average* jika periode data bernilai genap. Proses *smoothing* data ini bertujuan untuk menghilangkan komponen *trend-cycle* pada data untuk sementara waktu sebagai bentuk persiapan perhitungan komponen *seasonal*.

Ganjil

$$TCtmp_t = \frac{1}{p} \sum_{i=\ k-t-1}^{p+t} Y_i$$

Dimana

$$t = k+1, k+2, \dots, n - \ k$$
$$k = \frac{p-1}{2}$$

Keterangan:

$TCtmp_t$ = Komponen *trend-cycle* sementara periode ke-t
p = Periode musim data
Y_i = Data periode ke-i

Genap

$$TCtmp_t = \frac{1}{2}\left(\frac{1}{p}\sum_{i=k-t-1}^{p+t} Y_i + \frac{1}{p}\sum_{i=k-t}^{p+t+1} Y_i\right)$$

Dimana

$$t = k+1, k+2, \ldots, n - k$$
$$k = \frac{p}{2}$$

Keterangan:
$TCtmp_t$ = Komponen *trend-cycle* sementara periode ke-t
p = Periode musim data
Y_i = Data periode ke-i

Dilanjutkan

$$Ywtctmp_t = \begin{cases} \dfrac{Y_t}{TCtmp_t}; t = k+1, k+2, \ldots, n - k \\ 0; t\ lainnya \end{cases}$$

Dimana

$$k = \begin{cases} \dfrac{p-1}{2}; p\ ganjil \\ \dfrac{p}{2}; p\ genap \end{cases}$$

Keterangan:
$Ywtctmp_t$ = Data tanpa komponen *trend-cycle* sementara periode ke-t

1.2. Menghitung Komponen *Seasonal*

Pada tahapan ini data tanpa komponen *trend-cycle* akan diolah untuk mendapatkan *seasonal indices* (indeks musiman) yang dapat disebut juga sebagai komponen *seasonal*.

Menghitung Rata-Rata Data Per Indeks Musim

$$S_i = \frac{1}{d}\sum_{j=0}^{\frac{n}{p}-1} Ywtctmp_{j\times p+i}$$

Dimana

$$i = 1, 2, \ldots, p$$

$$d = \begin{cases} \left(\sum_{j=0}^{\frac{n}{p}-1} 1\right) - \frac{p-1}{2}; i = 1, p; p\ ganjil \\ \left(\sum_{j=0}^{\frac{n}{p}-1} 1\right) - \frac{p}{2}; i = 1, p; p\ genap \\ \left(\sum_{j=0}^{\frac{n}{p}-1} 1\right); i\ lainnya \end{cases}$$

Keterangan:
S_i = Rata-rata data musim ke i

Menghitung Rata-Rata Musim Keseluruhan

$$\bar{S} = \frac{1}{p}\sum_{i=1}^{p} S_i$$

$\bar{S}$ = Rata-rata musim keseluruhan

Menghitung Indeks Musim

$$SI_i = \frac{S_i}{\bar{S}} \times p$$

SI_i = Indeks musim ke-i

Selanjutnya akan dibentuk data komponen *seasonal* dengan mengiterasikan indeks musim sehingga jumlah periode komponen *seasonal* sama dengan jumlah periode data awal.

$$SC_t = \begin{cases} SI_p; t \equiv 0\ (mod\ p) \\ SI_{t\ mod\ p}; t\ lainnya \end{cases}$$

Keterangan:
SC_t = Komponen *seasonal* periode ke-t

Menghitung Data Tanpa Komponen *Seasonal* (*Deseasonalize*)

$$Ywsc_t = \frac{Y_t}{SC_t}$$

Keterangan:
Yws_t = Data tanpa komponen *seasonal* periode ke-t

1.3. Menghitung Komponen *Trend-Cycle* Menggunakan Teknik *Local Regression*

Pada tahapan ini data tanpa komponen *seasonal* akan diregresikan terhadap periode waktu data berdasarkan pilihan metode regresi dari pengguna. Pada proses regresi ini periode waktu akan dijadikan sebagai variabel bebas sedangkan data tanpa komponen *seasonal* akan dijadikan variabel terikat. Berikut tiga bentuk regresi yang tersedia pada fitur ini.

Regresi Linear Sederhana

$$Ywsc_t = a + bt$$

Regresi Eksponensial

$$Ywsc_t = e^{a+bt}$$

Hasil estimasi dijadikan komponen *trend-cycle* seperti berikut.

$$TC_t = \hat{Y}wsc_t$$

Keterangan:
TC_t = Komponen *trend-cycle* periode ke-t
$\hat{Y}wsc_t$= Estimasi data tanpa komponen *seasonal* periode ke-t

1.4. Menghitung Komponen *Irrengular*

Pada tahapan ini data tanpa komponen *seasonal* akan dibagi dengan komponen *trend-cycle* untuk memperoleh komponen *irrengular* atau *error*.

Komponen *Irrengular*

$$IC_t = \frac{Ywsc_t}{TC_t}$$

Keterangan:
IC_t = Komponen *irrengular* periode ke-t

1.5. Menghitung *Forecasting* Menggunakan Komponen *Seasonal* dan *Trend-Cycle*

Pada tahapan ini komponen *seasonal* dan *trend-cycle* akan dikalikan untuk memperoleh hasil ramalan (*forecasting*) sebagai berikut.

$$F_t = SC_t \times TC_t$$

Keterangan:
F_t = *Forecasting* periode ke-t

2. Additive Decomposition

Berikut tahapan dekomposisi data runtun waktu bentuk aditif.

2.1. Menghitung Komponen *Trend-Cycle* Menggunakan Teknik *Moving Average*

Pada tahapan ini data akan dimuluskan (*smoothing*) menggunakan teknik *simple moving average* jika periode data bernilai ganjil dan *centered moving average* jika periode data bernilai genap. Proses *smoothing* data ini bertujuan untuk mendapatkan komponen *trend-cycle*.

Ganjil

$$TC_t = \frac{1}{p} \sum_{i=\,k-t-1}^{p+t} Y_i$$

Dimana

$$t = k+1, k+2, \ldots, n- \ k$$
$$k = \frac{p-1}{2}$$

Keterangan:
TC_t = Komponen *trend-cycle* periode ke-t
p = Periode musim data
Y_i = Data periode ke-i

Genap

$$TC_t = \frac{1}{2}\left(\frac{1}{p}\sum_{i=\,k-t-1}^{p+t} Y_i + \frac{1}{p}\sum_{i=\,k-t}^{p+t+1} Y_i\right)$$

Dimana

$$t = k+1, k+2, \dots, n-\ k$$

$$k = \frac{p}{2}$$

Keterangan:
TC_t = Komponen *trend-cycle* periode ke-t
p = Periode musim data
Y_i = Data periode ke-i

Dikarenakan terdapat data yang hilang pada proses *moving average* maka perlu dilakukan inisiasi lanjutan untuk data di awal dan di akhir sebagai berikut.

$$TC_t = \begin{cases} \dfrac{1}{t}\displaystyle\sum_{i=1}^{t} Y_i; t < k; \\ \dfrac{1}{n-t}\displaystyle\sum_{i=n-t}^{n} Y_i; t > n-k \end{cases}$$

$$TC_1 = TC_1 - \frac{1}{2}(TC_1 - TC_2)$$

$$TC_n = TC_{n-2} - \frac{1}{2}(TC_{n-2} - TC_{n-1})$$

Dimana

$$t = 1,2, \dots, n$$

$$k = \begin{cases} \dfrac{p-1}{2}; p\ ganjil \\ \dfrac{p}{2}; p\ genap \end{cases}$$

Dilanjutkan

$$Ywtc_t = \begin{cases} Y_t - TC_t; t =\ k+1, k+2, \dots, n-\ k \\ 0; t\ lainnya \end{cases}$$

Dimana

$$k = \begin{cases} \dfrac{p-1}{2}; p\ ganjil \\ \dfrac{p}{2}; p\ genap \end{cases}$$

Keterangan:
$Ywtc_t$ = Data tanpa komponen *trend-cycle* periode ke-t

2.2. Menghitung Komponen *Seasonal*

Pada tahapan ini data tanpa komponen *trend-cycle* akan diolah untuk mendapatkan *seasonal indices* (indeks musiman) yang dapat disebut juga sebagai komponen *seasonal*.

Menghitung Indeks Musim dengan Menghitung Rata-Rata Per Musim

$$SI_i = \frac{1}{d}\sum_{j=0}^{\frac{n}{p}-1} Ywtc_{j\times p+i}$$

Dimana

$$i = 1, 2, .., p$$

$$d = \begin{cases} \left(\sum_{j=0}^{\frac{n}{p}-1} 1\right) - \frac{p-1}{2}; i = 1, p; p\ ganjil \\ \left(\sum_{j=0}^{\frac{n}{p}-1} 1\right) - \frac{p}{2}; i = 1, p; p\ genap \\ \left(\sum_{j=0}^{\frac{n}{p}-1} 1\right); i\ lainnya \end{cases}$$

Keterangan:
SI_i = Indeks musim ke-i

Selanjutnya akan dibentuk data komponen *seasonal* dengan mengiterasikan indeks musim sehingga jumlah periode komponen *seasonal* sama dengan jumlah periode data awal.

$$SC_t = \begin{cases} SI_p; t \equiv 0\ (mod\ p) \\ SI_{t\ mod\ p}; t\ lainnya \end{cases}$$

Keterangan:
SC_t = Komponen *seasonal* periode ke-t

2.3. Menghitung Komponen *Irrengular*

Pada tahapan ini data tanpa komponen *trend-cycle* akan dibagi dengan komponen *seasonal* untuk memperoleh komponen *irrengular* atau *error*.

Komponen *Irrengular*

$$IC_t = Ywtc_t - SC_t$$

Keterangan:
IC_t = Komponen *irrengular* periode ke-t

2.4. Menghitung *forecasting* Menggunakan Komponen *Seasonal* dan *Trend-Cycle*

Pada tahapan ini komponen *seasonal* dan *trend-cycle* akan dikalikan untuk memperoleh hasil ramalan (*forecasting*) sebagai berikut.

$$F_t = SC_t + TC_t$$

Keterangan:
F_t = Forecasting periode ke-t

3. Menghitung Evaluasi *Forecasting*

Pada tahapan ini hasil *forecasting* data dari hasil perkalian komponen *seasonal* dan *trend-cycle* akan dievaluasi menggunakan beberapa indikator, antara lain.

***Mean Square Error* (MSE)**

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2$$

Keterangan:
n = Jumlah periode data
Y_t = Data awal periode ke-t
F_t = Data hasil peramalan (*forecasting*) ke-t

***Root Mean Square Error* (RMSE)**

$$RMSE = \sqrt{MSE}$$

***Mean Absolute Error* (MAE)**

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|Y_t - F_t|$$

***Mean Percentage Error* (MPE)**

$$MPE = \frac{1}{n}\sum_{t=1}^{n}\left[\left(\frac{Y_t - F_t}{Y_t}\right) \times 100\right]$$

***Mean Absolute Percentage Error* (MAPE)**

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\left(\frac{Y_t - F_t}{Y_t}\right) \times 100\right|$$

Referensi

Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1983). *Forecasting: Methods and applications*. New York: John Wiley and Sons.

Peck, E., Vining, G., & Montgomery, D. (2012). *Introduction to Linear Regression Analysis*. Wiley.

Autocorrelation

Deskripsi

Jika pada materi pengantar statistik kita mengetahui bahwa korelasi adalah suatu alat ukur statistik yang digunakan untuk melihat kekuatan hubungan antara dua variabel, namun pada data runtun waktu analisis univariat, suatu variabel runtun waktu tidak akan dibandingkan dengan variabel runtun waktu lainnya melainkan perbandingannya akan dilakukan terhadap pada variabel runtun waktu yang sama tetapi pada periode waktu yang berbeda atau dalam analisis ini disebut autokorelasi (*autocorrelation*). Ukuran ini sering digunakan dalam penentuan orde parameter AR (*Autoregressive*) dan MA (*Moving Average*) dalam membentuk *Box-Jenkins Model* atau sering disebut model ARIMA. Selain itu, ukuran autokorelasi ini juga sering dilakukan secara bersamaan dengan autokorelasi parsial (*partial autocorrelation*) untuk mengidentifikasi hubungan setiap data runtun waktu antar periode waktu secara kondisional. Berikut tahapan menghitung ukuran autokorelasi dan autokorelasi parsial dengan tambahan statistik uji *ljung-box*.

Ukuran Statistik

1. **Autokorelasi**

   Sampel Autokorelasi

   $$r_k = \frac{\sum_{t=1}^{n-k}(Y_t - \bar{Y})\,(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n-k}(Y_t - \bar{Y})}$$

   Keterangan:
   n = Jumlah periode data
   r_k = Sampel autokorelasi *lag* ke-k
   Y_t = Data periode ke-t
   $\bar{Y}$ = Rata-rata data

   Standard *Error* Sampel Autokorelasi

   $$var(r_k) = \frac{1}{n}\left(1 + 2\sum_{i=1}^{q} r_i\right)$$

   Dimana

   $$k = 1,2, \dots, q$$

   Keterangan:
   $var(r_k)$ = Varians sampel autokorelasi *lag* ke-k
   q = Jumlah *lag*

2. **Statistik Uji *Ljung-Box***

   $$Q_k = n(n+2)\sum_{i=1}^{k} \frac{{r_i}^2}{n-i}$$

   $$prob = \; Q_k \sim \chi^2_{k-p-d}$$

   Keterangan:
   Q_k = Statistik uji *ljung-box lag* ke-k

3. **Autokorelasi Parsial**

Sampel Autokorelasi Parsial

$$\hat{\phi}_{11} = r_1$$

$$\hat{\phi}_{22} = \frac{(r_2 - {r_1}^2)}{(1 - {r_1}^2)}$$

$$\hat{\phi}_{kj} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\hat{\phi}_{k-1,k-j}; k = 2, \dots, j; j = 1,2, \dots, k-1$$

$$\hat{\phi}_{kk} = \frac{\left(r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j}\right)}{\left(1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j\right)}; k = 3, \dots$$

Standard Error Sampel Autokorelasi Parsial

Dengan asumsi bahwa model AR(p) memiliki jumlah orde $p \leq k - 1$, sehingga $\hat{\phi}_{kk} \cong N\left(0, \frac{1}{n}\right)$ dikatakan dalam Quenouville, 1949.

Maka

$$var(\hat{\phi}_{kk}) \cong \frac{1}{n}$$

4. Nilai Batas Bartlett

Autokorelasi

$$r_k \pm Z_{\frac{\alpha}{2}}\, s.e(r_k)$$

Autokorelasi Parsial

$$\hat{\phi}_{kk} \pm Z_{\frac{\alpha}{2}}\, s.e(\hat{\phi}_{kk})$$

Referensi

Bartlett, M. S. (1946). On the theoretical specification of sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, Series B, 8*, 27–27.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (Revised ed.). San Francisco: Holden-Day.

Cryer, J. D. & Chan, Kung-Sik(2008). *Time Series Analysis With Aplication in R* (2th ed.). New York: Springer Science.

Quenouville, M. H. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Series B, 11*, 68–68.

Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1983). *Forecasting: Methods and applications*. New York: John Wiley and Sons.

Wei, William W. S. (2005). *Time series analysis: Univariate and multivariate methods* (2nd ed.). Addison Wesley.

Unit Root Test

Deskripsi

Unit Root Test adalah suatu uji untuk memeriksa kestasioneran data runtun waktu. Uji ini penting dilakukan karena untuk analisis lanjutan dan pemodelan data runtun waktu harus memenuhi asumsi stasioneritas. Terdapat beberapa metode dalam melakukan uji ini antara lain *Augmented Dickey-Fuller*, *Phillips-Peron*, dan *Kwiatkowski-Phillips-Schmidt-Shin*. Namun, pada modul ini Uji Unit Root masih terbatas pada uji *dickey-fuller* dan *augmented dickey-fuller*. Uji *dickey-fuller* adalah uji menggunakan statistik uji *tau* (τ) untuk menguji kestasioneran, dengan bentuk umum rumus sebagai berikut.

$$\tau = \frac{\gamma}{se(\gamma)}$$

Nilai γ diperoleh dari koefisien slope dari persamaan berikut.

$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t$$

Dengan demikian dapat dikatakan bahwa γ adalah koefisien *slope* data runtun waktu hasil regresi dari data *differencing* terhadap data runtun waktu satu periode sebelumnya. Sedangkan Augmented Dickey-Fuller adalah bentuk lebih kompleks dari uji *Dickey-Fuller* yang menambah faktor *lag* (p) data *differencing* ke dalam persamaan sebelumnya sehingga dapat ditulis seperti berikut.

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

Selain itu, uji hipotesis dari Uji *Dickey-Fuller* dan Uji *Augmented Dickey-Fuller* adalah sebagai berikut.

$$H_0 : \gamma = 0 \ (tidak\ stasioner)$$

$$H_1 : \gamma < 0 \ (stasioner)$$

Pada penjelasan sebelumnya bentuk persamaan yang ditampilkan dalam uji ini masih satu jenis bentuk dari beberapa bentuk persamaan yang tersedia, berikut beberapa bentuk persamaan yang tersedia pada uji *Dickey-Fuller* dan *Augmented Dickey-Fuller*.

Persamaan *Dickey-Fuller*

Persamaan tanpa konstanta

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

Persamaan dengan konstanta

$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t$$

Persamaan dengan tren waktu

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \varepsilon_t$$

Persamaan *Augmented Dickey-Fuller*

Persamaan tanpa konstanta

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

Persamaan dengan konstanta

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

Persamaan dengan tren waktu

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

Selanjutnya dalam uji juga akan dihitung nilai kritis dan probabilitas yang metode penghitungan didasarkan pada jurnal *Mac Kinnon*, 1994.

Rumus Nilai Kritis

$$C(\alpha) = \beta_0 + \frac{\beta_1}{N} + \frac{\beta_2}{N^2} + \frac{\beta_3}{N^3}$$

Keterangan:

α = Nilai signifikansi

Tabel Nilai Kritis *Mac Kinnon*

Persamaan	$\boldsymbol{\alpha}$	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_3}$
tanpa konstanta	0.01	-2.56574	-2.2358	-3.627	0
	0.05	-1.941	-0.2686	-3.365	31.223
	0.10	-1.61682	0.2656	-2.714	25.364
dengan konstanta	0.01	-3.43035	-6.5393	-16.786	-79.433
	0.05	-2.86154	-2.8903	-4.234	-40.04
	0.10	-2.56677	-1.5384	-2.809	0
dengan tren waktu	0.01	-3.95877	-9.0531	-28.428	-134.155
	0.05	-3.41049	-4.3904	-9.036	-45.374
	0.10	-3.12705	-2.5856	-3.925	-22.38

Rumus Nilai Probabilitas

$$p = \phi(\hat{\gamma}_0 + \hat{\gamma}_1 \tau + \hat{\gamma}_2 \tau^2 + \hat{\gamma}_3 \tau^3)$$

Keterangan:

ϕ = *Cumulative Distribution Function Normal Standard*

$\tau nc(1)$ = Persamaan tanpa konstanta satu variabel

$\tau c(1)$ = Persamaan dengan konstanta satu variabel

$\tau ct(1)$ = Persamaan dengan tren waktu variabel

Tabel *Mac Kinnon Probability Value* untuk $\tau < \tau^*$

Statistic	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2 \times 10^2$	τ*	τmin
τnc(1)	0.6344	1.2378	3.2496	-1.04	-19.04
τc(1)	2.1659	1.4412	3.8269	-1.61	-18.83
τct(1)	3.2512	1.6047	4.9588	-2.89	-16.18

Tabel *Mac Kinnon Probability Value* untuk $\tau > \tau^*$

Statistic	$\hat{\gamma}_0$	$\hat{\gamma}_1 \times 10$	$\hat{\gamma}_2 \times 10$	$\hat{\gamma}_3 \times 10^2$	τ*	τmax
τnc(1)	0.4797	9.3557	-0.6999	3.3066	-1.04	-
τc(1)	1.7339	9.3202	-1.2745	-1.0368	-1.61	2.74
τct(1)	2.5261	6.1654	-3.7956	-6.0285	-2.89	0.7

Selanjutnya pengujian koefisien regresi hingga ukuran penyeleksian kriteria regresi juga ditampilkan dengan rumus sebagai berikut.

Uji Koefisien Regresi

1. **Statistik uji-t**

$$t_i = \frac{\beta_i}{se(\beta_i)}$$

2. **Nilai Probabilitas uji-t**

$$p_i = 2 \times (1 - F_t(|t_i|, df)$$

Dimana

$$k = 1,2, \ldots, k$$

$$df = n - k$$

Keterangan:

β_i = Parameter ke-i

$se(\beta_i)$ = *Standard error* parameter ke-i

p_i = Nilai probabilitas uji-t parameter ke-i

t_i = Nilai statistik uji-t parameter ke-i

$F_t()$ = *Cumulative Distribution Function Distribusi Student-T*

df = Derajat bebas

n = Jumlah observasi data

k = Jumlah parameter

Ukuran Penyeleksian Kriteria Regresi

1. ***Sum of Squared Residual* (SS_{Res})**

$$SS_{Res} = \begin{cases} \sum_{i=1}^{n} e^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \text{ ; } untuk\ regresi\ linear\ sederhana \\ e'e = \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right); untuk\ regresi\ linear\ berganda \end{cases}$$

Keterangan:

y = Vektor variabel terikat (n x 1)

X= Matriks variabel bebas (n x p)

$\hat{\beta}$ = Vektor estimasi parameter (p x 1)

2. ***Standard Error of Regression***

$$\sigma = \sqrt{MS_{Res}} = \begin{cases} \sqrt{\frac{SS_{Res}}{n-k-1}}; untuk\ persamaan\ \tau c\ dan\ \tau ct \\ \sqrt{\frac{SS_{Res}}{n-k}}; untuk\ persamaan\ \tau nc \end{cases}$$

3. ***Coeficient Determination* (R)**

$$R^2 = 1 - \frac{SS_{Res}}{SS_T}$$

Dimana

$$SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

4. ***Coeficient Determination Adjusted* (R-adj)**

$$R^2_{Adj} = \begin{cases} 1 - \dfrac{SS_{Res}/(n-k-1)}{SS_T/(n-1)} ; untuk\ persamaan\ \tau c\ dan\ \tau ct \\ 1 - \dfrac{SS_{Res}/(n-k)}{SS_T/n} ; untuk\ persamaan\ \tau nc \end{cases}$$

5. ***Log Likelihood***

$$\ln L(y, X, \beta, \sigma^2) = -\frac{n}{2}\ln(2\pi) - n\ln(\sigma) - \frac{1}{2\sigma^2} SS_{Res}$$

Dimana

$$\sigma^2 \approx \tilde{\sigma}^2 = \frac{SS_{Res}}{n}$$

6. ***F-Statistic***

$$F_0 = \frac{MS_R}{MS_{Res}}$$

Dimana

$$MS_R = \frac{SS_R}{k}$$

$$SS_R = SS_T - SS_{Res}$$

7. **Probabilitas *F-Statistic***

$$p = 1 - P(F_0 > F_{\propto;k;n-k-1})$$

8. ***Mean Dependent Variable***

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

9. ***Standard Deviation Dependent Variable***

$$se(y) = \frac{\sum_{i=1}^{n}(y_i - \bar{y})}{n-1}$$

10. *Akaike Info Criterion* (AIC)

$$AIC = -2\ln L + 2p$$

$$Mean(AIC) = \frac{AIC}{n}$$

Keterangan:

$\ln L$ = Nilai *log likelihood*

p = Jumlah parameter

11. *Schwarz Bayesian Criterion* (SBC)

$$SBC = -2\ln L + p\ln n$$

$$Mean(SBC) = \frac{SBC}{n}$$

12. *Hannan-Quinn Criterion* (HQC)

$$HQC = -2\ln L + 2p\ln(\ln n)$$

$$Mean(HQC) = \frac{HQC}{n}$$

13. Durbin-Watson Statistic

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} {e_t}^2}$$

Dimana

$$e_t = y_t - \hat{y}_t$$

Keterangan:

e_t = Residual observasi ke-t

y_t = Data awal observasi ke-t

$\hat{y}_t$ = Data estimasi observasi ke-t

Referensi

Enders, W. (2015). *Applied Econometric Time Series* (4th ed.). Wiley

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin Heidelberg.

MacKinnon, J. G. (1994). Approximate asymptotic distribution functions for unit-root and cointegration tests. *Journal of Business & Economic Statistics*, 12(2), 167-176. https://doi.org/10.1080/07350015.1994.10510005

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, 11(6), 601-618. https://doi.org/10.1002/(SICI)1099-1255(199611)11:6

MacKinnon, J. G. (2010). Critical values for cointegration tests (Queen's Economics Department Working Paper No. 1227). *Queen's University, Department of Economics*. http://qed.econ.queensu.ca/working_papers/papers/qed_wp_1227.pdf

Mohamad. (2016, October 19). *Appendix E: Hannan-Quinn Information Criterion (HQC)*. NumXL Support. Retrieved March 20, 2025, from https://support.numxl.com/hc/en-us/articles/215531183-Appendix-E-Hannan-Quinn-Information-Criterion-HQC

Peck, E., Vining, G., & Montgomery, D. (2012). *Introduction to Linear Regression Analysis*. Wiley.

Box-Jenkins Model

Deskripsi

Box-Jenkins Model adalah salah satu model peramalan dalam analisis data deret waktu. Model ini pertama kali diperkenalkan pada tahun 1970, oleh George Box dan dan Gwilym Jenkins. Model ini dikembangkan dengan menggabungkan beberapa model deret waktu lainnya, yaitu *Autoregressive* (AR) Model dan *Moving Avarage* (MA) Model. Selain itu, Model ini juga mempertimbangkan kestasioneran data yang dinotasikan dengan I. Dengan demikian model ini disebut juga sebagai *Autoregressive Integrated Moving Average* (ARIMA) Model. Dalam fitur ini parameter model diestimasi menggunakan dua tahapan. Pada tahapan pertama dilakukan inisiasi parameter menggunakan beberapa ketentuan sebagai berikut.

1. Parameter intersep (μ) diinisiasi menggunakan nilai mean dari data aktual.
2. Parameter AR (ϕ) diinisiasi menggunakan Algoritma *Durbin-Levinson*.
3. Parameter MA (θ) diinisiasi menggunakan Algoritma *Innovation*.

Setelah proses inisiasi parameter, tahap kedua dilanjutkan dengan estimasi parameter menggunakan metode *Conditional Least Square* (CLS) dengan meminimalkan fungsi *Conditional Sum of Square* (CSS). Pada tahap ini juga dilakukan proses optimasi menggunakan metode *Limited-Memory Broyden-Fletcher-Goldfarb-Shanno* (LBFGS) sehingga memperoleh parameter estimasi yang konvergen. Sebagai catatan, pada fitur ini proses diferensiasi hanya terbatas pada diferensiasi data sebanyak dua kali karena persamaan *forecasting* ARIMA yang memiliki pola yang semakin kompleks jika dijabarkan. Berikut penjelasan rinci mengenai tahapan estimasi parameter Model ARIMA.

Estimasi Parameter

1. **Inisiasi Parameter**

   Parameter Intersep

   Parameter intersep diestimasi menggunakan persamaan berikut.

$$E(\bar{Y}) = \mu = \bar{Y} = \frac{1}{n}\sum_{k=1}^{n-1} Y_t$$

   Keterangan:
   $\bar{Y}$ = Mean Data Aktual
   n = Jumlah Observasi Data Aktual
   Y_t = Data Aktual Observasi ke-t

   Algoritma *Durbin-Levinson*

   Berikut persamaan yang digunakan untuk menginisiasi parameter AR pada Model ARMA.

$$\phi_{kk} = \left[\gamma(k) - \sum_{j=1}^{k-1} \phi_{k-1,j}\, \gamma(k-j)\right] v_{k-1}^{-1};\ k = 1, 2, \ldots, p$$

$$v_k = v_{k-1}[1 - \phi_{kk}^2];\ k = 0, 1, 2, \dots, p$$

Dengan inisiasi awal sebagai berikut

$$\phi_{11} = \frac{\gamma(1)}{\gamma(0)}$$
$$v_0 = \gamma(0)$$

Dimana

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-|h|} \left(Y_{t+|h|} - \bar{Y}_n\right)(Y_t - \bar{Y}_n)$$

$$h = 0, 1, 2, \dots, n - 1$$

Keterangan:
$\hat{\gamma}(h)$ = Kovarians data aktual
ϕ_{kk} = Autokorelasi parsial ke-k atau parameter AR ke-k
p = Jumlah parameter AR

Algoritma *Innovation*
Berikut persamaan yang digunakan untuk menginisiasi parameter MA pada model MA dan ARMA.

$$v_0 = k(1,1)$$

$$\theta_{n,n-k}(t) = v_k^{-1} \left(k(n + 1,\ \ k + 1) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j \right), 0 \leq k < n$$

$$v_n = k(n + 1, n + 1) - \sum_{j=0}^{k-1} \theta_{n,n-j}^2\, v_j$$

Keterangan:
$\theta_{n,n-k}$ = Parameter MA ke-n,n-k
n = Jumlah parameter MA.
$k(i,j)$ = Kovarians data = $\gamma(|i - j|)$

2. **Fungsi *Conditional Sum of Square* (CSS)**
Pada tahapan ini nilai residual setiap model dihitung dengan cara yang berbeda yang hasilnya akan disubstitusikan pada persamaan umum *Conditional Sum of Square* yang dinotasikan dengan lambang f sebagai berikut.

$$f = S_c(\beta, \mu) = \sum_{t=1}^{n} e_t^2$$

Keterangan:
β = Parameter model

e_t = Residual ke-t

Rumus Residual Model AR

$$e_t = (Y_t - \mu) - \sum_{i=1}^{p} \phi_i (Y_{t-i} - \mu)$$

Rumus Residual Model MA

$$e_t = (Y_t - \mu) + \sum_{i=1}^{q} \theta_i (e_{t-i})$$

Dengan $e_0 = e_{-1} = e_{-2} = \cdots = e_{-q} = 0$

Rumus Residual Model ARMA

$$e_t = (Y_t - \mu) - \sum_{i=1}^{p} \phi_i (Y_{t-i} - \mu) + \sum_{i=1}^{q} \theta_i (e_{t-i})$$

Dengan $e_p = e_{p-1} = e_{p-2} = \cdots = e_{p+1-q} = 0$

Setelah itu, berdasarkan persamaan di atas akan dibentuk fungsi gradien g yang meminimkan fungsi $S_c(\beta, \mu)$ sebagai berikut.

$$g = \frac{\partial S_c(\beta, \mu)}{\partial \beta}$$

dan

$$g = \frac{\partial S_c(\beta, \mu)}{\partial \mu}$$

Dengan menggunakan fungsi f dan g serta parameter AR (ϕ), MA (θ), dan intersep (μ) akan dioptimalkan menggunakan fungsi *Limited-Memory Broyden-Fletcher-Goldfarb-Shanno* (L-BFGS) dengan iterasi maksimal dua ratus kali dengan rincian kerja dari halaman web berikut https://chokkan.org/software/liblbfgs/.

Estimasi *Standard Error* Parameter

1. *Standard Error* Parameter Intersep

Standard error parameter intersep dihitung menggunakan persamaan berikut.

$$Var(\bar{Y}) = \frac{\gamma_0}{n} \left[1 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \rho_k \right]$$

$$se(\bar{Y}) = \sqrt{Var(\bar{Y})}$$

Dengan ketentuan

$$\rho_k = \begin{cases} 0; k > p + q \\ \dfrac{\gamma_k}{\gamma_0}; 1 \leq k \leq p + q \end{cases}$$

Keterangan:

ρ_k = Autokorelasi observasi ke-t

p = Jumlah parameter AR
q = Jumlah parameter MA
γ_k = Autokovarians observasi ke-t, dapat dinotasikan juga sebagai $\hat{\gamma}(k)$
γ_0 = Varians data aktual

2. ***Standard Error* Parameter AR dan MA**
Standard error parameter AR dan MA dihitung menggunakan persamaan berikut.

$$Var(\beta) = 2\sigma_e^2 \begin{bmatrix} \frac{\partial^2 S_c(\beta)}{\partial \beta_1^2} & \frac{\partial^2 S_c(\beta)}{\partial \beta_2 \beta_1} & \cdots & \frac{\partial^2 S_c(\beta)}{\partial \beta_k \beta_1} \\ \frac{\partial^2 S_c(\beta)}{\partial \beta_1 \beta_2} & \frac{\partial^2 S_c(\beta)}{\partial \beta_2^2} & \cdots & \frac{\partial^2 S_c(\beta)}{\partial \beta_k \beta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 S_c(\beta)}{\partial \beta_1 \beta_k} & \frac{\partial^2 S_c(\beta)}{\partial \beta_2 \beta_k} & \cdots & \frac{\partial^2 S_c(\beta)}{\partial \beta_k^2} \end{bmatrix}^{-1}$$

$$se(\beta) = \sqrt{Var(\beta)}$$

Dimana

$$\sigma_e^2 = \frac{S_c(\beta)}{df}$$

$$df = n - 2p - q - 1$$

Keterangan:
$S_c(\beta)$ = Nilai *Conditional Sum of Square* (CSS)
σ_e^2 = Varians residual
df = Derajat bebas CSS

Uji Koefisien Regresi

1. **Statistik Uji-t**

$$t_i = \frac{\beta_i}{se(\beta_i)}$$

2. **Nilai Probabilitas Uji-t**

$$p_i = 2 \times (1 - F_t(|t_i|, df)$$

Dimana

$$k = 1,2, \dots, k$$

$$df = n - k$$

Keterangan:

β_i = Parameter ke-i

$se(\beta_i)$ = *Standard error* parameter ke-i

p_i = Nilai probabilitas uji t parameter ke-i

t_i = Nilai statistik uji t parameter ke-i

$F_t()$ = *Cumulative Distribution Function Distribusi Student-T*

df = Derajat bebas

n = Jumlah observasi data

k = Jumlah parameter

Ukuran Penyeleksian Model

1. ***Sum of Squared Residual* (SS_{Res})**

$$SS_{Res} = S_c(\beta)$$

2. ***Standard Error of Regression***

$$\sigma = \sqrt{MS_{Res}} = \sqrt{\frac{SS_{Res}}{n-p-q-1}}$$

3. ***Coeficient Determination* (R)**

$$R^2 = 1 - \frac{SS_{Res}}{SS_T}$$

Dimana

$$SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

4. ***Coeficient Determination Adjusted* (R-adj)**

$$R^2_{Adj} = 1 - \frac{SS_{Res}/(n-p-q-1)}{SS_T/(n-1)}$$

5. ***Log Likelihood***

$$\ln L(y, X, \beta, \sigma^2) = -\frac{n}{2}\ln(2\pi) - n\ln(\sigma) - \frac{1}{2\sigma^2}SS_{Res}$$

Dimana

$$\sigma^2 \approx \tilde{\sigma}^2 = \frac{SS_{Res}}{n}$$

6. *F-Statistic*

$$F_0 = \frac{MS_R}{MS_{Res}}$$

Dimana

$$MS_R = \frac{SS_R}{k}$$

$$SS_R = SS_T - SS_{Res}$$

7. Probabilitas *F-Statistic*

$$p = 1 - P(F_0 > F_{\propto;k;n-k-1})$$

8. *Mean Dependent Variable*

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

9. *Standard Deviation Dependent Variable*

$$se(y) = \frac{\sum_{i=1}^{n}(y_i - \bar{y})}{n-1}$$

10. *Akaike Info Criterion* (AIC)

$$AIC = -2\ln L + 2p$$

$$Mean(AIC) = \frac{AIC}{n}$$

Keterangan:

$\ln L$ = Nilai *log likelihood*

p = Jumlah parameter

11. *Schwarz Bayesian Criterion* (SBC)

$$SBC = -2\ln L + p\ln n$$

$$Mean(SBC) = \frac{SBC}{n}$$

12. *Hannan-Quinn Criterion* (HQC)

$$HQC = -2\ln L + 2p\ln(\ln n)$$

$$Mean(HQC) = \frac{HQC}{n}$$

13. Durbin-Watson Statistic

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} {e_t}^2}$$

Dimana

$$e_t = y_t - \hat{y}_t$$

Keterangan:

e_t = Residual observasi ke-t

y_t = Data awal observasi ke-t

$\hat{y}_t$ = Data estimasi observasi ke-t

Forecasting

Berikut rumus yang digunakan untuk melakukan *forecasting* berdasarkan penggunaan diferensiasi pada data.

1. **Tanpa Diferensiasi**
   Persamaan *forecasting* ARIMA (p,0,q) tanpa diferensiasi jika ditulis menggunakan *backshift* operator adalah sebagai berikut.
   $$\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)Y_t = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right)e_t$$
   Jika dijabarkan dan persamaan memiliki parameter konstan (intersep) atau sering dinotasikan sebagai μ, bentuk persamaan *forecasting* menjadi seperti berikut.
   $$Y_t = \mu + \sum_{i=1}^{p} \phi_i Y_{t-i} - \sum_{i=1}^{q} \theta_i e_{t-i}$$

2. **Diferensiasi Satu Kali**
   Persamaan *forecasting* ARIMA (p,1,q) dengan diferensiasi data satu kali jika ditulis menggunakan *backshift* operator adalah sebagai berikut.
   $$\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)(1 - B)Y_t = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right)e_t$$
   Jika dijabarkan dan persamaan memiliki parameter konstan (intersep) atau sering dinotasikan sebagai μ, bentuk persamaan forecasting menjadi seperti berikut.
   $$Y_t = \mu + (1 + \phi_1)Y_{t-1} - \sum_{i=2}^{p}(\phi_{i-1} - \phi_i)Y_{t-i} - \phi_p Y_{t-p-1} - \sum_{i=1}^{q} \theta_i e_{t-i}$$

3. **Diferensiasi Dua Kali**

   Persamaan *forecasting* ARIMA (p,2,q) dengan diferensiasi data dua kali jika ditulis menggunakan *backshift* operator adalah sebagai berikut.

   $$\left(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p\right)(1-B)^2 Y_t=\left(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q\right) e_t$$

   Jika dijabarkan dan persamaan memiliki parameter konstan (intersep) atau sering dinotasikan sebagai μ, bentuk persamaan forecasting menjadi seperti berikut.

   $$Y_t=\mu+(2+\phi_1) Y_{t-1}+(\phi_2-2\phi_1+1) Y_{t-2}+\sum_{i=3}^{p}(\phi_i-\phi_{i-1}-\phi_{i-2}) Y_{t-i} - \left(2\phi_p+\phi_{p-1}\right) Y_{t-p-1}-\phi_p Y_{t-p-2}-\sum_{i=1}^{q} \theta_i e_{t-i}$$

Evaluasi *Forecasting*

Berikut ukuran yang digunakan untuk mengevaluasi data hasil *forecasting*.

1. ***Mean Square Error* (MSE)**

   $$MSE=\frac{1}{n} \sum_{t=1}^{n}(Y_t-F_t)^2$$

   Keterangan:
   n = Jumlah Periode Data
   Y_t = Data Awal Periode ke-t
   F_t = Data Hasil Peramalan (*Forecasting*) ke-t

2. ***Root Mean Square Error* (RMSE)**

   $$RMSE=\sqrt{MSE}$$

3. ***Mean Absolute Error* (MAE)**

   $$MAE=\frac{1}{n} \sum_{t=1}^{n}|Y_t-F_t|$$

4. ***Mean Percentage Error* (MPE)**

   $$MPE=\frac{1}{n} \sum_{t=1}^{n}\left[\left(\frac{Y_t-F_t}{Y_t}\right) \times 100\right]$$

5. ***Mean Absolute Percentage Error* (MAPE)**

   $$MAPE=\frac{1}{n} \sum_{t=1}^{n}\left|\left(\frac{Y_t-F_t}{Y_t}\right) \times 100\right|$$

Referensi

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control* (*Revised ed*.). San Francisco: Holden-Day.

Brockwell, P. J., & Davis, R. A. (1996). *Introduction to time series and forecasting* (1st ed.). Springer.

Cryer, J. D., & Chan, Kung-Sik (2008). *Time Series Analysis With Application in R* (2nd ed.). New York: Springer Science.

Lütkepohl, H. (2005). New *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin Heidelberg.

Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1983). *Forecasting: Methods and applications*. New York: John Wiley and Sons.

Mohamad. (2016, October 19). *Appendix E: Hannan-Quinn Information Criterion (HQC)*. NumXL Support. Retrieved March 20, 2025, from https://support.numxl.com/hc/en-us/articles/215531183-Appendix-E-Hannan-Quinn-Information-Criterion-HQC

Okazaki, N. (2002-2014). *libLBFGS: A library of Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS)* [Software]. Retrieved from https://chokkan.org/software/liblbfgs/

Peck, E., Vining, G., & Montgomery, D. (2012). *Introduction to Linear Regression Analysis*. Wiley.

Wei, William W. S. (2005). *Time Series Analysis: Univariate and Multivariate Methods* (2nd ed.). Addison Wesley.