Q1

8

Series : $x_t = \beta_1 + \beta_2 t + W_t$

Moving average : $V_t = \frac{1}{2q+1} \sum_{j=-q}^{q} x_{t-j}$

$$E[V_t] = \frac{1}{2q+1} E\left(\sum_{j=-q}^{q} [X_{t-j}] \right) \quad -(1)$$

Now $\sum_{j=-q}^{q} x_{t-j} = \sum_{j=-q}^{q} \beta_1 + \sum_{j=-q}^{q} \beta_2 (t-j) + \sum_{j=-q}^{q} W_{t-j}$

$= (2q+1)\beta_1 + \beta_2 \left((2q+1)t - \sum_{j=-q}^{q} j \right) + \sum_{j=-q}^{q} W_{t-j}$

$= (2q+1)\beta_1 + (2q+1)\beta_2 t + \sum_{j=-q}^{q} W_{t-j}$

Put back in eq (1)

$$E[V_t] = \frac{1}{(2q+1)} \left[(2q+1)\beta_1 + (2q+1)t\beta_2 + E\left(\sum_{j=-q}^{q} W_{t-j} \right) \right]$$

0.5

As W_t is white noise $\Rightarrow E\left(\sum_{j=-q}^{q} W_{t-q} \right) = 0$

since $E[W_{t-j}] = 0 \quad \forall j$.

so $E[V_t] = \beta_1 + \beta_2 t + 0$

$= \beta_1 + \beta_2 t$

Hence Proved

Autocovariance Function of V_t :→

$$\gamma(h) = Cov(V_t, V_{t+h}) = E[(V_t - E[V_t])(V_{t+h} - E[V_{t+h}])]$$

$$V_t - E[V_t] = \frac{1}{(2q+1)} \sum_{j=-q}^{q} W_{t-j}$$

$$V_{t+h} - E[V_{t+h}] = \frac{1}{(2q+1)} \sum_{k=-q}^{q} W_{t+h-k}$$

$$Cov(V_t, V_{t+h}) = \frac{1}{(2q+1)^2} \sum_{j=-q}^{q} \sum_{k=-q}^{q} Cov(W_{t-j}, W_{t+h-k})$$

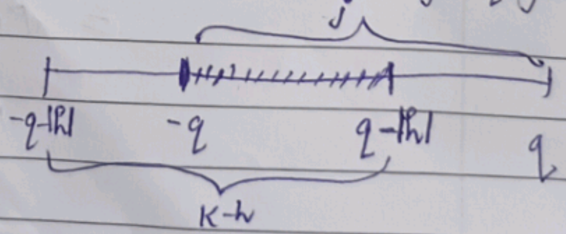
now given W_t is white noise

~~$Cov(W_{t-j}, W_{t+h-k}) = \sigma_w^2$ if $t-j = t+h-k$~~

~~covariance is non-zero only when $j = k-h$~~

$$Cov(W_{t-j}, W_{t+h-k}) \begin{cases} = \sigma_w^2 & \text{if } t-j = t+h-k \\ = 0 & \text{else} \end{cases}$$

number of terms when $j = k-h$ will be $(2q+1) - |h|$ as j and k both goes from $-q$ to q so if $-q \le j \le q$ then $-q \le k-h \le q$



$j = k-h$ in region

so $(2q+1) - |h|$ cases are

$$\gamma(h) = \begin{cases} \dfrac{\sigma_w^2}{(2q+1)^2}(2q+1-|h|) & \text{if } |h| < 2q+1 \\ 0 & \text{elsewise} \end{cases}$$

1

Q2) $X_t = U_1 \sin(2\pi w_0 t) + U_2 \cos(2\pi w_0 t)$

To show series is weakly stationary, we can prove that mean and auto covariance function are constant. Variance is also constant.

Mean of X_t: $E[X_t] = E[U_1 \sin(2\pi w_0 t) + U_2 \cos(2\pi w_0 t)]$

$E[X_t] = E[U_1] \sin(2\pi w_0 t) + E[U_2] \cos(2\pi w_0 t)$

(as expectation is linear)

using $E[U_1] = E[U_2] = 0$

$E[X_t] = 0$ —(1)

so $E[X_t]$ is constant wrt t

Autocov function $\gamma(h)$:

$\gamma(h) = E[X_t X_{t+h}]$

$= E[(U_1 \sin(2\pi w_0 t) + U_2 \cos(2\pi w_0 t))(U_1 \sin(2\pi w_0(t+h)) + U_2 \cos(2\pi w_0(t+h)))]$

1

$= E[U_1^2] \sin(2\pi w_0 t) \sin(2\pi w_0(t+h)) + E[U_2^2] \cos(2\pi w_0 t) \cos(2\pi w_0(t+h))$

$= \sigma^2 [\sin(2\pi w_0 t) \sin(2\pi w_0(t+h)) + \cos(2\pi w_0 t) \cos(2\pi w_0(t+h))]$

$= \sigma^2 \cos(2\pi w_0(t+h-t)) = \sigma^2 \cos(2\pi w_0 h)$

(only depends on h)

~~mean and Auto cov. function didn't depend on t so X_t is stationary.~~

<u>Var of X_t</u>

$$Var(X_t) = E[X_t^2] - (E(X_t))^2$$

$$= E[X_t^2] - 0 \quad \text{(from ①)}$$

$$X_t^2 = (U_1 \sin(2\pi w_0 t) + U_2 \cos(2\pi w_0 t))^2$$
$$= U_1^2 \sin^2(2\pi w_0 t) + U_2^2 \cos^2(2\pi w_0 t) + 2U_1U_2 \sin(2\pi w_0 t)\cos(2\pi w_0 t)$$

$$E[X_t^2] = E[U_1^2]\sin^2(2\pi w_0 t) + E(U_2^2)\cos^2(2\pi w_0 t) + 2E[U_1]E[U_2]\sin(2\pi w_0 t)\cos(2\pi w_0 t)$$

($\because U_1, U_2$ are independent)

$$= E[U_1^2]\sin^2(2\pi w_0 t) + E[U_2^2]\cos(2\pi w_0 t)$$

$(\because E[U_1] = E[U_2] = 0)$

$$= \sigma^2(\sin^2(2\pi w_0 t) + \cos^2(2\pi w_0 t))$$
$$= \sigma^2$$

(const. w.r.t. t)

So, mean, var, and autocovariance function are constant wrt t
$\Rightarrow X_t$ is weakly stationary.

__/__/__

Q3

$$Y_t = W_t - \theta W_{t-1} + U_t$$

(a) ACF of Y_t

for $h=0$:→ $\gamma(0) = Var(Y_t) = Var(W_t - \theta W_{t-1} + U_t)$

now as W_t, W_{t-1} and U_t are uncorrelated

so $\gamma(0) = Var(W_t) + \theta^2 Var(W_{t-1}) + Var(U_t)$

$$= \sigma_W^2 + \theta^2 \sigma_W^2 + \sigma_U^2 = (1+\theta^2)\sigma_W^2 + \sigma_U^2$$

for $h=1$:→

$$\gamma(1) = Cov(Y_t, Y_{t+1})$$

substitute Y_t, Y_{t+1}

$$\gamma(1) = Cov(W_t - \theta W_{t-1} + U_t,\ W_{t+1} - \theta W_t + U_{t+1})$$

$$\left\{\begin{array}{l} Cov(W_t, W_{t+1}) = 0 \quad (\because W_t, W_{t+1} \text{ are independent}) \\ Cov(W_t, W_t) = \sigma_W^2 \\ Cov(W_{t-1}, W_t) = 0 \end{array}\right. \longrightarrow (1)$$

also terms having U_t and W_t are zero because U_t and W_t are independent.

$$\Rightarrow \gamma(1) = -\theta \sigma_W^2$$

for $h=-1$:→ using symmetry of cov. function

$$\gamma(-1) = -\theta \sigma_W^2$$

for $|h| \geq 2$:→ In this case W_t and W_{t+h} are independent for all $|h| \geq 2$

so covariance will be 0.

$$\gamma(h) = 0 \qquad \text{for } |h| \geq 2$$

<u>Autocorrelation function $\rho_Y(h)$:-</u>

$$\rho_Y(h) = \frac{\gamma(h)}{\gamma(0)}$$

$$\rho_Y(h) = \begin{cases} 1 & \text{if } h = 0 \\ \dfrac{-\theta\sigma_W^2}{(1+\theta^2)\sigma_W^2 + \sigma_V^2} & \text{if } h = \pm 1 \\ 0 & \text{if } |h| \geq 2 \end{cases}$$

($\because$ By simply putting values from ⓐ part)

<u>CCF : $\rho_{X,Y}(h)$</u>

Formula: $$\rho_{X,Y}(h) = \frac{\gamma_{X,Y}(h)}{\sqrt{\gamma_X(0)}\sqrt{\gamma_Y(0)}}$$

Now as $X_t = W_t \Rightarrow Var(X_t) = Var(W_t) = \sigma_W^2$

<u>$h=0$:→</u>

$$Cov(X_t, Y_t) = cov(W_t, W_t - \theta W_{t-1} + V_t)$$
$$= cov(X_t, Y_t) + 0$$
($\because$ using ① from ⓐ)
$$= \sigma_W^2$$

~~$h = 1$:→~~

$$\cancel{Cov(X_t, Y_{t+1}) = cov(W_t, W_{t+1} - \theta W_t + V_{t+1})}$$
$$\cancel{= cov(W_t, W_{t+1}) - \theta\, cov(W_t, W_t)}$$
$$\cancel{+ cov(W_t, U_{t+1})}$$

(3)

$h=1:\rightarrow$

$$Cov(X_{t+1}, Y_t) = Cov(W_{t+1}, W_t - \theta W_{t-1} + U_t)$$

as W_{t+1} is independent of W_t, W_{t-1}, U_t

$$Cov(X_{t+1}, Y_t) = 0$$

$h=-1:\rightarrow$

$$Cov(X_{t-1}, Y_t) = Cov(W_{t-1}, W_t - \theta W_{t-1} + U_t)$$

$$= 0 - \theta Cov(W_{t-1}, W_{t-1})$$

($\because$ using (1) from (a) part)

$$= -\theta \sigma_w^2$$

$|h| \geq 2:\rightarrow$

Covariance is 0 because white noise independent at lag $|h| \geq 2$.

$$Cov(X_{t+h}, Y_t) = 0 \quad \text{for } |h| \geq 2$$

Now apply formula $\rho_{X,Y}(h) = \dfrac{Cov(X_t, Y_{t+h})}{\sqrt{Var(X_t) \cdot Var(Y_t)}}$

1.5

$$\rho_{X,Y}(h) = \begin{cases} \dfrac{\sigma_w}{\sqrt{\sigma_w^2(1+\theta^2)+\sigma_v^2}} & \text{if } h=0 \\ \dfrac{-\theta\sigma_w}{\sqrt{\sigma_w^2(1+\theta^2)+\sigma_v^2}} & \text{if } h=-1 \\ 0 & h=1, \pm 2, \pm 3 \ldots \end{cases}$$

__/__/__

(4) $X_t = \sin(2\pi U t)$

For time series to be weakly stationary we can show that mean, variance and autocovariance function does not depend on t.

Mean :

$$E[X_t] = E[\sin(2\pi U t)] = \int_0^1 \sin(2\pi u t)\, du$$

($\because U \sim \text{Uniform}(0,1)$
and $f_U(u) = 1$ for $u \in (0,1)$)

Now integral of sine function over a full period is zero.

$$\Rightarrow E[X_t] = 0 \quad \forall t$$

Variance :→

$$Var(X_t) = E[X_t^2] - (E[X_t])^2$$
$$= E[\sin^2(2\pi U t)] - 0$$

$$\left(\because \sin^2(x) = \frac{1-\cos(2x)}{2}\right)$$

$$E[\sin^2(2\pi U t)] = \int_0^1 \sin^2(2\pi u t)\, du = \int_0^1 \frac{1-\cos(4\pi u t)}{2}\, du$$

again we know $\cos(4\pi u t)$ over full period have integral equal to 0.

$E[\sin^2(2\pi Ut)] = 1/2 = Var(X_t)$

Auto Covariance of X_t :→

$$Cov(X_t, X_{t+h}) = E[X_t X_{t+h}] - E[X_t]E[X_{t+h}]$$

$$= E[X_t X_{t+h}] - 0 \quad (\because mean = 0)$$

$$= E[\sin(2\pi Ut)\sin(2\pi U(t+h))]$$

(Now $\because \sin A \sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)]$)

$$= \frac{1}{2} E[\cos(2\pi Uh)]$$

$$= \frac{1}{2} \cdot \int_0^1 \cos(2\pi uh)\, du$$

Integral of $\cos(2\pi uh)$ over $u \in (0,1)$ (full cycle) will be zero unless $h = 0$. For $h=0$ it is 1.

If $h=0$ $\quad cov(X_t, X_{t+h}) = 1/2$

else if $h \neq 0$ $\quad cov(X_t, X_{t+h}) = 0$

so, autocov. only depend on lag h.

so mean, var and auto covariance does not depend on t.

$\Rightarrow X_t$ is weakly stationary

1

(4b)

To prove X_t is not strictly stationary we will check whether X_t distribution depends on time t.

A time series is strictly stationary if joint pdf does not depend on time at which series is observed.

So if CDF of X_t depends on t

$\Rightarrow$ X_t is not strictly stationary.

0

$$X_t = \sin(2\pi U t) \qquad U \sim unif(0,1)$$

$$P(X_t \le 1/2) = P\left(\sin(2\pi U t) \le 1/2\right)$$

As $\sin(\theta) = \frac{1}{2}$ when $\theta = \pi/6$

$$\Rightarrow \quad 2\pi U t \le \frac{\pi}{6}$$

$$\Rightarrow \quad U \le \frac{1}{12t}$$

By CDF of uniform function

$$P(X_t \leq 1/2) = P\left(U \leq \frac{1}{12t}\right) = \frac{1}{12t}$$

Thus CDF $F_{X_t}(x)$ depends on t.

$\Rightarrow X_t$ is not strictly stationary

Hence Proved

Q.6

Gausian Model for Regression is

$$\underset{\sim}{y} = X \underset{\sim}{\beta_1} + \underset{\sim}{\varepsilon_1} \qquad \theta_1 = (\beta_1', \sigma_1^2)'$$

where $E(\underset{\sim}{\varepsilon_1}) = \underset{\sim}{0}$; $Var(\underset{\sim}{\epsilon_1}) = \sigma_1^2 I$

$$\underset{\sim}{y} = X \underset{\sim}{\beta_2} + \underset{\sim}{\epsilon_2} \qquad \theta_2 = (\beta_2', \sigma_2^2)'$$

(I am using X for notation instead of Z)

where $E(\underset{\sim}{\varepsilon_2}) = 0$ and $Var(\underset{\sim}{\epsilon_2}) = \sigma_2^2 I$

$$I(\theta_1 ; \theta_2) = T^{-1} E_1 \log \frac{f(y; \theta_1)}{f(y; \theta_2)}$$

$$f(y; \theta_1) = \frac{1}{(2\pi)^{T/2} (\sigma_1^2)^{T/2}} e^{-\frac{1}{2\sigma_1^2}(\underset{\sim}{y} - X\underset{\sim}{\beta_1})^T(\underset{\sim}{y} - X\underset{\sim}{\beta_1})}$$

$\underset{\sim}{y} \in R^T$

$$f(y; \theta_2) = \frac{1}{(2\pi)^{T/2} (\sigma_2)^{T/2}} e^{-\frac{1}{2\sigma_2^2}(\underset{\sim}{y} - X\underset{\sim}{\beta_2})^T(\underset{\sim}{y} - X\underset{\sim}{\beta_2})}$$

$\underset{\sim}{y} \in R^T$

Now,
$$\frac{f(y; \theta_1)}{f(y; \theta_2)} = \frac{\exp\left(-\frac{1}{2\sigma_1^2}(\underset{\sim}{y} - X\underset{\sim}{\beta_1})^T(\underset{\sim}{y} - X\underset{\sim}{\beta_1})\right) \frac{1}{(\sigma_1^2)^{T/2}}}{\exp\left(-\frac{1}{2\sigma_2^2}(\underset{\sim}{y} - X\underset{\sim}{\beta_2})^T(\underset{\sim}{y} - X\underset{\sim}{\beta_2})\right) \frac{1}{(\sigma_2^2)^{T/2}}}$$

__/__/___

$$\log \frac{f(y;\theta_1)}{f(y;\theta_2)} = \frac{T}{2}\log\frac{\sigma_2^2}{\sigma_1^2} - \frac{1}{2\sigma_1^2}(\underset{\sim}{y}-X\underset{\sim}{\beta_1})^T(\underset{\sim}{y}-X\underset{\sim}{\beta_1})$$

$$+ \frac{1}{2\sigma_2^2}(\underset{\sim}{y}-X\underset{\sim}{\beta_2})^T(\underset{\sim}{y}-X\underset{\sim}{\beta_2})$$

Expectation under θ_1 :-

As $y \sim N(X\beta_1, \sigma_1^2 I)$. $y = X\beta_1 + \varepsilon$

$\varepsilon \sim N(0, \sigma_1^2 I)$

$E(\varepsilon'\varepsilon) = T\sigma_1^2$ — (1)

$$E_1\left[\log\frac{f(y;\theta_1)}{f(y;\theta_2)}\right] = \frac{T}{2}\log\frac{\sigma_2^2}{\sigma_1^2} - E_1\left[\frac{1}{2\sigma_1^2}(\underset{\sim}{y}-X\underset{\sim}{\beta_1})^T(\underset{\sim}{y}-X\underset{\sim}{\beta_1})\right]$$

$$+ E_1\left[\frac{1}{2\sigma_2^2}(\underset{\sim}{y}-X\underset{\sim}{\beta_2})^T(\underset{\sim}{y}-X\underset{\sim}{\beta_2})\right].$$

$$= \frac{T}{2}\log\frac{\sigma_2^2}{\sigma_1^2} - \frac{T}{2} + E_1\left[\frac{1}{2\sigma_2^2}((X\underset{\sim}{\beta_1}+\varepsilon-X\underset{\sim}{\beta_2})^T(X\beta_1+\varepsilon-X\beta_2)\right]$$

$\because$ using eq (1)

$$= \frac{T}{2}\log\frac{\sigma_2^2}{\sigma_1^2} - \frac{T}{2} + E_1\left[\frac{1}{2\sigma_2^2}((X(\underset{\sim}{\beta_1}-\underset{\sim}{\beta_2}))^T(X(\beta_1-\underset{\sim}{\beta_2}))+2(X(\beta_1-\beta_2))^T\varepsilon + \varepsilon^T\varepsilon)\right]$$

$$= \frac{T}{2}\log\frac{\sigma_2^2}{\sigma_1^2} - \frac{T}{2} + \frac{1}{2\sigma_2^2}((\underset{\sim}{\beta_1}-\underset{\sim}{\beta_2})^T X^T X(\underset{\sim}{\beta_1}-\underset{\sim}{\beta_2}) + T\sigma_1^2)$$

Hence, we get

$$\Rightarrow I(\theta_1;\theta_2) = \frac{1}{2}\left(\frac{\sigma_1^2}{\sigma_2^2} - \log\frac{\sigma_1^2}{\sigma_2^2} - 1\right)$$

1

$$+\frac{1}{2}\,\frac{(\beta_1-\beta_2)^T X^T X(\beta_1-\beta_2)}{T\sigma_2^2}$$

Hence proved

(Note: I have used X in notation instead of Z)

~~[illegible]~~, ~~[illegible]~~ ~~[illegible]~~

__/__/__

Q(59) <u>Proof that Autocovariance function is Non-neg. definite :→</u>

Consider stationary process (X_t) with mean ($\mu = E[X_t]$).

Autocov. at lag h is

$$\gamma(h) = E[(X_t - \mu)(X_{t+h} - \mu)]$$

consider set of real no. ($c_1, c_2 \ldots c_n$) and corresponding set of time points ($t_1, \ldots, t_n$)

Quadratic form is

$$Z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \gamma(t_j - t_i)$$

(By def^n of autocov.)

$$= E\left[\left(\sum_{i=1}^{n} c_i (X_{t_i} - \mu)\right)^2\right]$$

Expected value of squared random variable is always non negative

$\Rightarrow Z \geq 0$

<u>Hence proved</u>

—/—/—

(5b) Verification that Sample Autocovariance is Non negative Definite

Sample Autocovariance is estimator of true auto cov. based on observed data
Let $(X_1 \ldots X_n)$ be stationary Process with lag h.

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X})$$

$\bar{X}$ is sample mean

lets consider set of real no.
$(c_1, \ldots c_n)$
Quadratic expression is:

$$\hat{Z} = \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \hat{\gamma}(t_j - t_i)$$

1.5

(using def^n of auto cov.)

$$\hat{Z} = \frac{1}{n} \sum_{t=1}^{n-h} \left(\sum_{i=1}^{n} c_i (X_{t_i} - \bar{X}) \right)^2$$

$\hat{Z} \geq 0$

(because sum of squares is always non negative)

Hence Proved

Question 7

Part (a): Estimated Average Annual Increase in Logged Earnings per Share

If the model is correct, the coefficient of the time_sequence (which represents time) in the regression model will give us the estimated average annual increase in logged earnings per share. In the regression model: The coefficient beta is the estimate for the average annual increase in logged earnings. After fitting the model using the code below, the summary of the model will display this value.

```
library(astsa)
```

```
## Warning: package 'astsa' was built under R version 4.3.2
```

```
data(jj)
trend = time(jj) - 1970 # center time
time = time(jj)
X = log(jj) # log casting of data
Q = factor(cycle(jj)) # quarter separation
model = lm(X~0 + trend + Q) # no intercept
sum <- summary(model)
# coefficient of trend is beta i.e. estimated avg annual increase in logged earnings
paste0("estimated avg annual increase in logged earnings is: ", coef(model)["trend"])
```

```
## [1] "estimated avg annual increase in logged earnings is: 0.167172164378767"
```

X 0

The coefficient of time_adjusted gives the estimated average annual increase in logged earnings per share. In this case, it is approximately 0.167, which represents a 16.7% average annual increase.

Part (b): Change in Logged Earnings from Q3 to Q4 and Percentage Change

In this part, we calculate the change from Q3 to Q4 and compute the percentage increase or decrease using the respective coefficients

```
# Extract the coefficients for Q3 and Q4
cq3 <- coef(model)["Q3"]
cq4 <- coef(model)["Q4"]
# Calculate the percentage change
percent_change <- (exp(cq4 - cq3) - 1) * 100
paste0("The average logged earning decrease from third quarter to fourth by: ", -percent_change)
```

```
## [1] "The average logged earning decrease from third quarter to fourth by: 23.5671557610801"
```

x 0

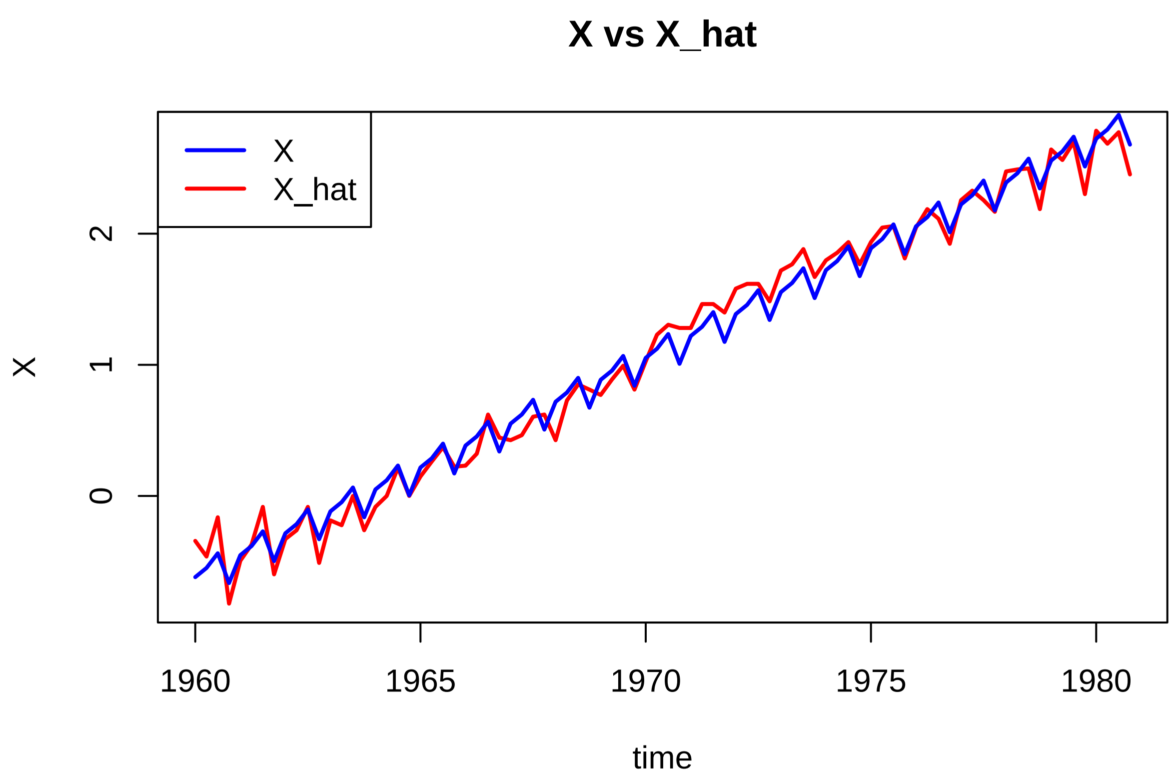
Explanation:

We extract the coefficients for Q3 and Q4 from the model output. These coefficientsrepresent the logged earnings for the third and fourth quarters, respectively. We calculate the percentage change between the two quartersThis transformation ensures that the change is calculated in non-logarithmic terms. The result shows that the earnings decrease by approximately 23.57% from Q3 to Q4. Conclusion: The analysis reveals that the average logged earnings decrease by 23.57% from the third quarter to the fourth quarter. This suggests a notable drop in earnings as we transition from Q3 to Q4, potentially reflecting seasonal or other cyclical factors.

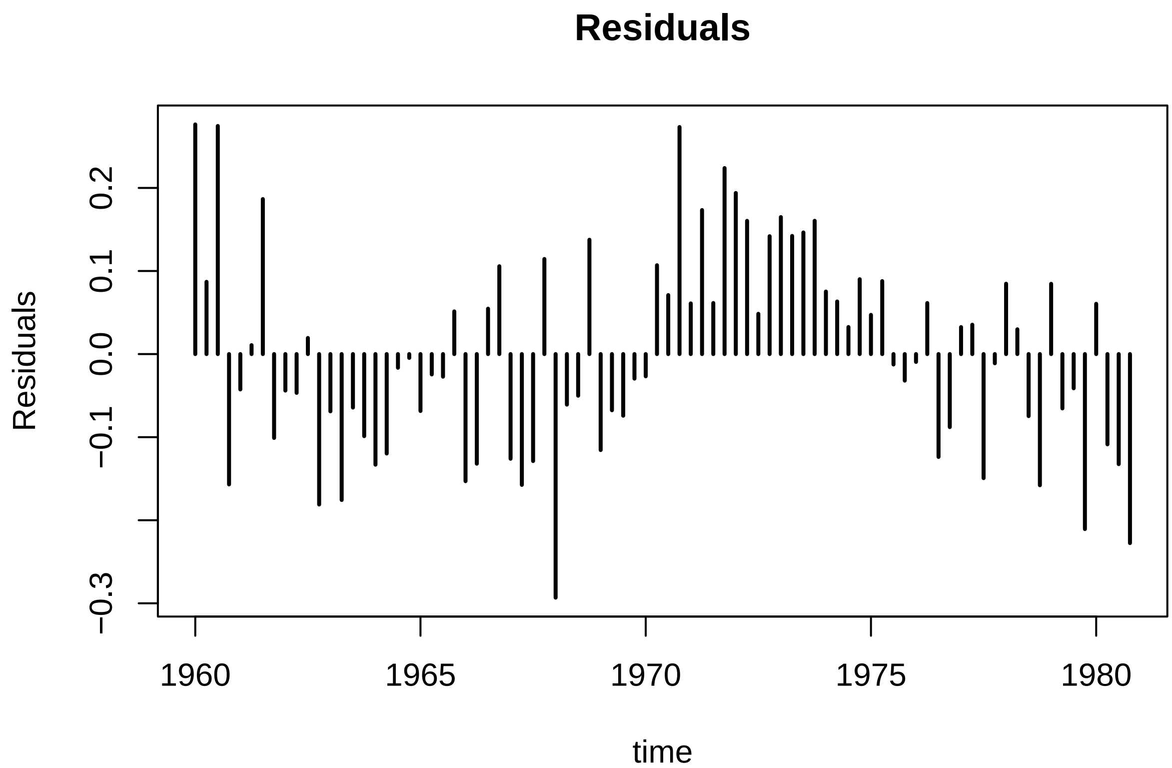
Part (c): Graph the Data, Fitted Values, and Residuals

In this part, we plot the actual logged earnings data along with the fitted values, and we also plot the residuals to determine if they resemble white noise.

```
X_hat<- fitted(model)
# Plot the data and fitted values
time <- as.vector(time)
plot(time, X, type = "l", col = "red", lwd = 2, ylab = "X", main = "X vs X_hat")
lines(time, X_hat, col = "blue", lwd = 2) # Fitted values
legend("topleft", legend = c("X", "X_hat"), col = c("blue", "red"), lwd = 2)
```




```
# Residual Calculation
residuals <- residuals(model)
plot(time, residuals, type = "h", col = "black", lwd = 2, ylab = "Residuals", main = "Residuals")
```



Conclusion: The model captures the primary trends in the Johnson & Johnson data, estimating a steady annual growth in logged earnings and highlighting a notable decrease in earnings between Q3 and Q4. While the residuals do not perfectly resemble white noise, the model seems to fit the data well enough to provide meaningful insights.

0.5