pstat274_hw01_aoxu

AO XU

2022-10-04

Problem 4

```
polyroot(c(1,-2))
```

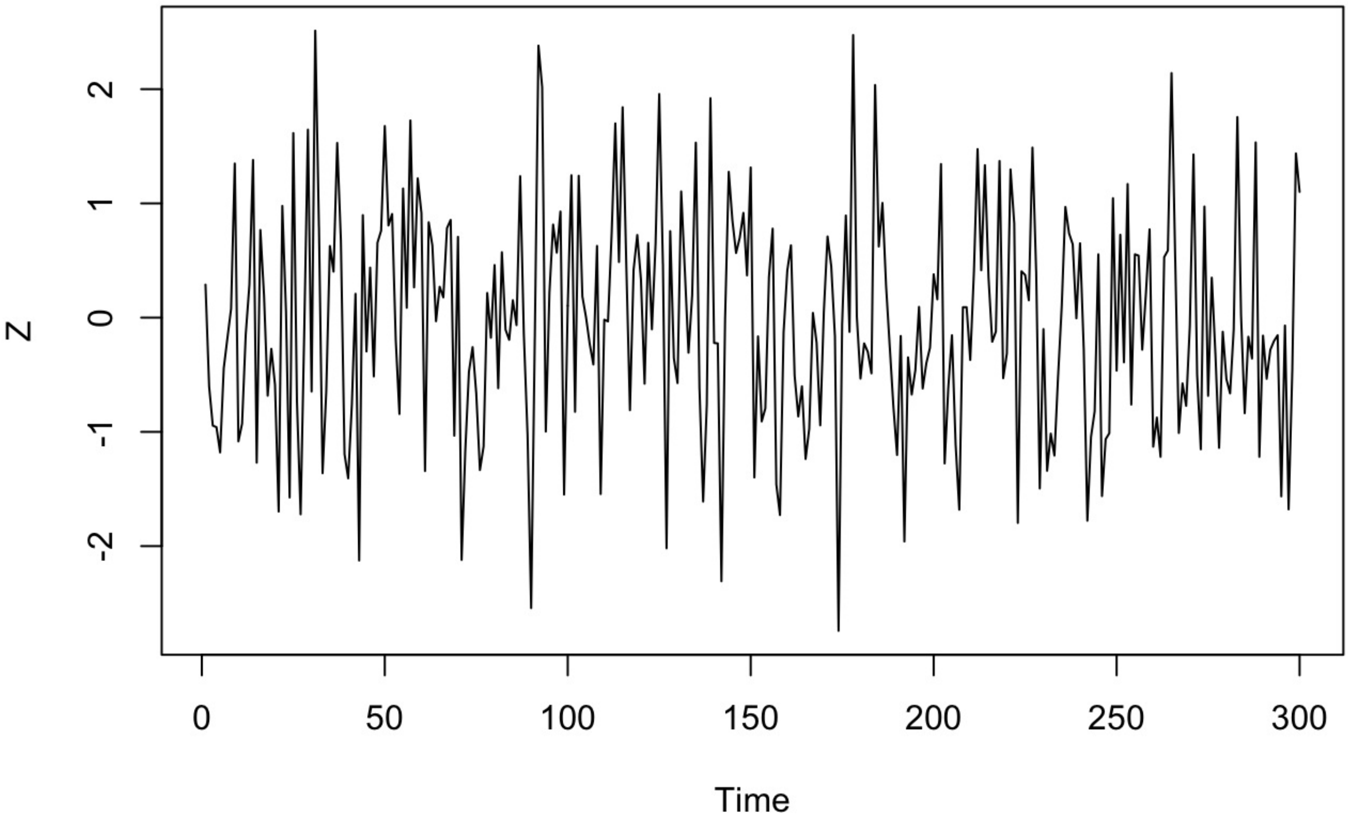
```
## [1] 0.5+0i
```

```
polyroot(c(1,-0.45,0.05))
```

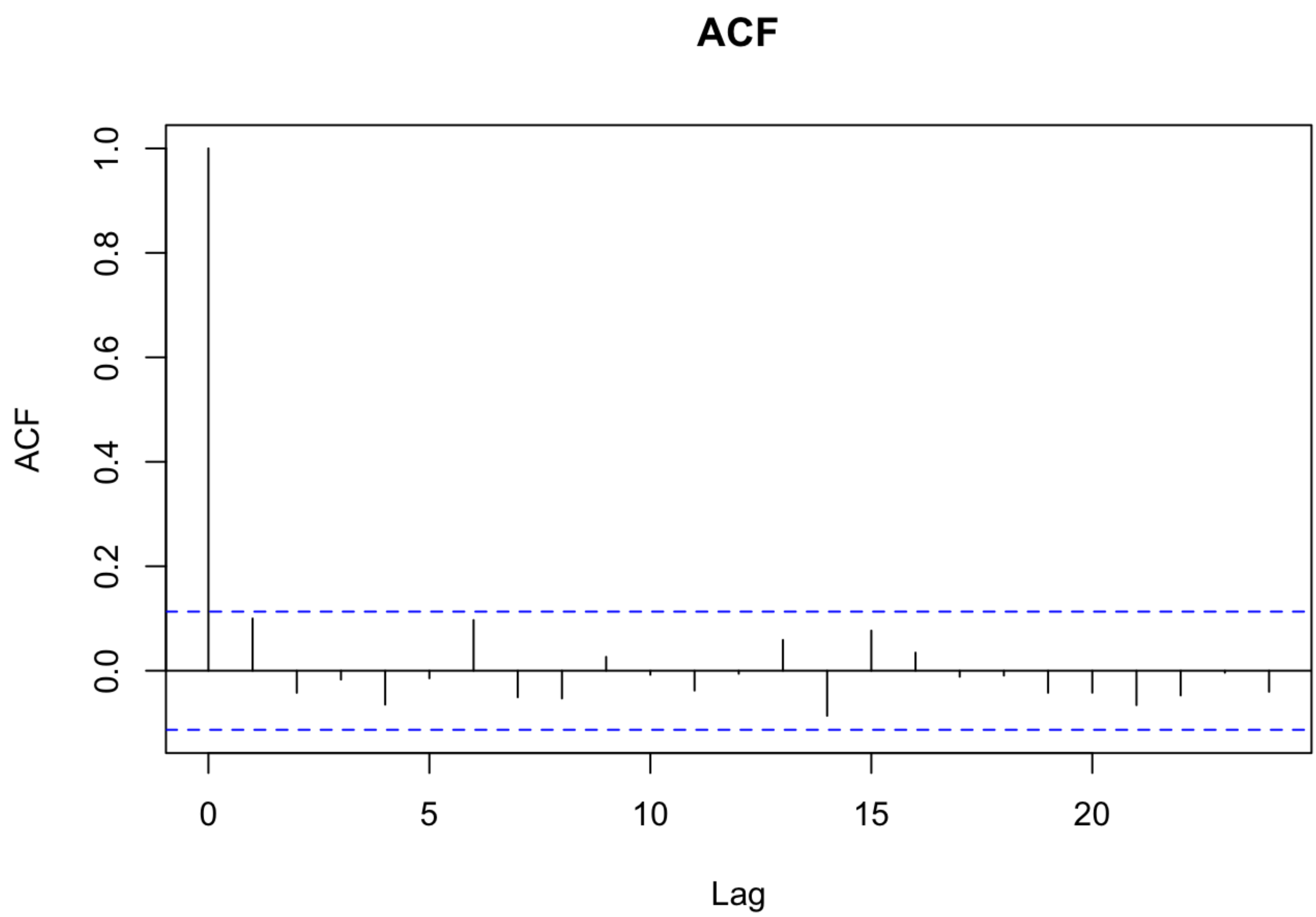
```
## [1] 4+0i 5-0i
```

Problem 6

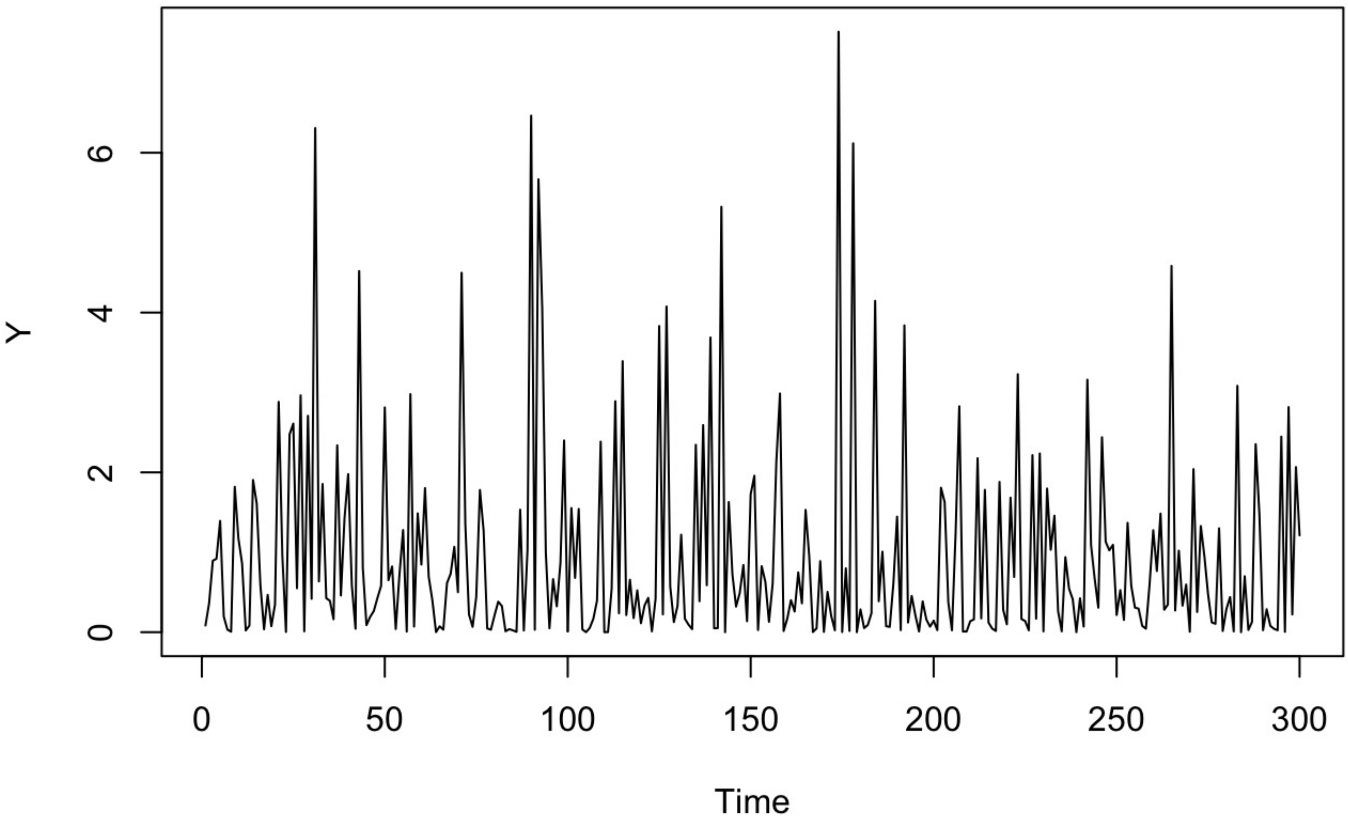
```
# a
Z<- rnorm(300)
plot.ts(Z, xlab="Time", ylab="Z")
```



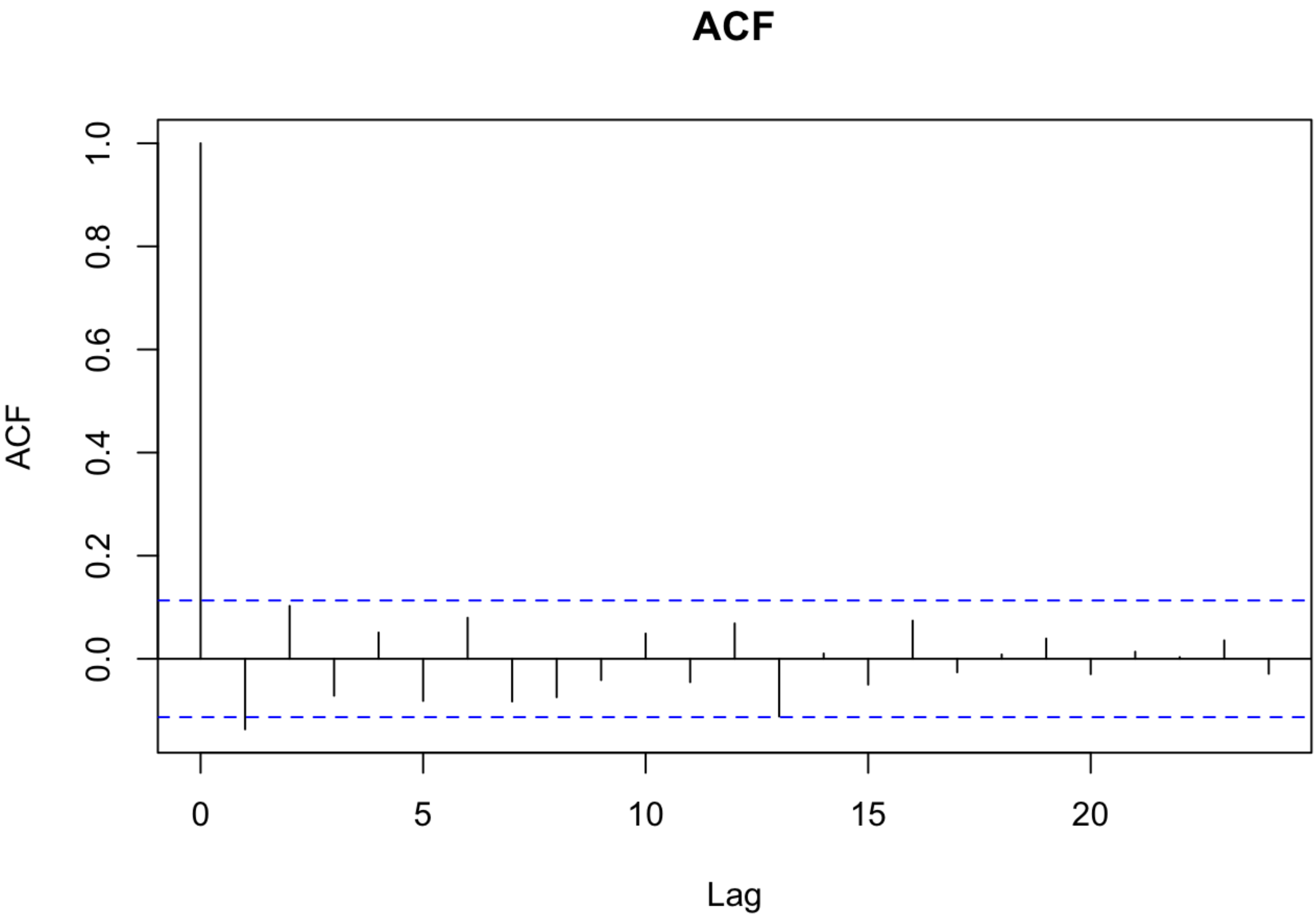

```
acf(Z, main="ACF")
```




```
# b
Y<- Z^2
plot.ts(Y, xlab="Time", ylab="Y")
```




```
acf(Y, main="ACF")
```



6c

c1) There’s a difference between the plots of graphs of time series Z and Y that the graph of z is symmetrically distributed around 0 while the graph of Y is all above 0.

But both of they are stationary since both of them have no trend or seasonality.

c2) There is no noticeable difference in the plots of acf functions ρZ and ρY. I would describe Y as a non-Gaussian white noise sequence based on my plots since it has 0 constant mean, variance and it is uncorrelated.

```
# d
mean(Y)
```

```
## [1] 0.9301439
```

```
mean(Z)
```

```
## [1] -0.03541707
```

PSTAT 174/274: Homework # 1.

This homework is based on Lectures 1–2. Please study material of week 1 *before* starting working on this problems. *Good Luck!*

1. *Understanding deterministic and stochastic trends.* You are given the following statements about time series:

I. Stochastic trends are characterized by explainable changes in direction.

II. Deterministic trends are better suited to extrapolation than stochastic trends.

III. Deterministic trends are typically attributed to high serial correlation with random error.

Determine which statements are false. Explain.

A. I only B. II only C. III only D. I, II, and III

E. The answer is not given by (A), (B), (C), or (D).

2. *Random walk and stationarity. In this question we introduce random walk with non-zero mean.*
A random walk is expressed as $X_1 = Z_1$, $X_t = X_{t-1} + Z_t$, $t = 2, 3, \ldots$, where $Z_t \sim WN(\mu_Z, \sigma_Z^2)$, that is, $E(Z_t) = \mu_Z$, $Var(Z_t) = \sigma_Z^2$, and $Cov(Z_t, Z_s) = 0$ for $t \neq s$. Determine which statements are true with respect to a random walk model; show calculations and provide complete explanations.

I. If $\mu_Z \neq 0$, then the random walk is nonstationary in the mean.
(*Hint: Nonstationary in the mean means that the mean changes with time.*)

II. If $\sigma_Z^2 = 0$, then the random walk is nonstationary in the variance.
(*Hint: Nonstationary in the variance means that the variance changes with time.*)

III. If $\sigma_Z^2 > 0$, then the random walk is nonstationary in the variance.

3. *Calculation of sample acf.* You are given the following stock prices of company CAS:

Day	Stock Price
1	538
2	548
3	528
4	608
5	598
6	589
7	548
8	514
9	501
10	498

Calculate the sample autocorrelation at lag 3.

Hints:
(i) We are given a sample of size $n = 10$ *to estimate autocorrelation at lag 3:* $\rho(3) = Cor(X_1, X_4) = \frac{\gamma(3)}{\gamma(0)}$,
– for definition of autocorrelation at lag 3 see Week 1 slide 52 or (2.1.3) on p. 6 of Lecture Notes.
(ii) General formulas for calculating sample mean and covariance are given on slide 38 of week 1 and in §1.2 on p. 4 of Lecture notes for week 1. To estimate $\rho(3) = Cor(X_1, X_4)$ *we have:*

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t, \quad \hat{\rho}_3 = \frac{\hat{\gamma}(3)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-3}(x_t - \bar{x})(x_{t+3} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}.$$

4. *Polyroot command in R. Recall from algebra, that a function* $f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$ *is called a polynomial function of order* n*. Roots of a polynomial function* f *are solutions of the equation* $f(z) = 0$*. Roots of a quadratic equation* $ax^2 + bx + c = 0$ *are given by the formula* $x_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$.

Let $f(z) = 1 - 2z$ and $g(z) = 1 - 0.45z + 0.05z^2$. Find their roots, show calculations. Check your answers using R command *polyroot*:

$>$ *polyroot*$(c(1, -2))$
$>$ *polyroot*$(c(1, -0.45, 0.05))$. (*Do not forget to include your output!*)

5. *Model identification.* You are given the following information about a MA(1) model with coefficient $|\theta_1| < 1$: $\rho_1 = -0.4$, $\rho_k = 0$, $k = 2, 3, \ldots$. Determine the value of θ_1.

6. *Gaussian White Noise and its square.* Let $\{Z_t\}$ be a Gaussian white noise, that is, a sequence of i.i.d. normal r.v.s each with mean zero and variance 1. Let $Y_t = Z_t^2$.
(a) Using R generate 300 observations of the Gaussian white noise Z. Plot the series and its acf.
(b) Using R, plot 300 observations of the series $Y = Z_t^2$. Plot its acf.
(c) Analyze graphs from (a) and (b).
– Can you see a difference between the plots of graphs of time series Z and Y? From the graphs, would you conclude that both series are stationary (or not)?
– Is there a noticeable difference in the plots of acf functions ρ_Z and ρ_Y? Would you describe Y as a non-Gaussian white noise sequence based on your plots?
Provide full analysis of your conclusions.
(d) Calculate the second-order moments of Y: $\mu_Y(t) = E(Y_t)$, $\sigma_Y^2(t) = Var(Y_t)$, and $\rho_Y(t, t+h) = Cor(Y_t, Y_{t+h})$. Do your calculations support your observations in (c)?

Hints: (i) Slides 65 and 68 of week 1 have R commands to generate MA(1) time series. White Noise is a MA(1) process with coefficient $\theta_1 = 0$*. Here is a more direct code to generate WN* $\{Z_t\} \sim N(0,1)$:

```
Z <- rnorm(300)
plot.ts(Z, xlab = "", ylab = "")
acf(Z, main = "ACF")
```

(ii) Useful for part (d): For $X \sim N(0, \sigma^2)$, $E(X^4) = 3(\sigma^2)^2$.

The following two problems are for students enrolled in PSTAT 274 ONLY

G1. Let $\{Z_t\}$ be *Gaussian* white noise, i.e. $\{Z_t\}$ is a sequence of i.i.d. normal r.v.s each with mean zero and variance 1. Define

$$X_t = \begin{cases} Z_t, & \text{if } t \text{ is even;} \\ (Z_{t-1}^2 - 1)/\sqrt{2}, & \text{if } t \text{ is odd} \end{cases}$$

Show that $\{X_t\}$ is WN$(0,1)$ (that is, variables X_t and $X_{t+k}, k \geq 1$, are uncorrelated with mean zero and variance 1) but that X_t and X_{t-1} are **not** i.i.d.

G2. If $\{X_t\}$ and $\{Y_t\}$ are uncorrelated stationary sequences, i.e., if X_r and Y_s are uncorrelated for every r and s, show that $\{X_t + Y_t\}$ is stationary with autocovariance function equal to the sum of the autocovariance functions of $\{X_t\}$ and $\{Y_t\}$.

1. *Understanding deterministic and stochastic trends.* You are given the following statements about time series:

I. Stochastic trends are characterized by explainable changes in direction.
II. Deterministic trends are better suited to extrapolation than stochastic trends.
III. Deterministic trends are typically attributed to high serial correlation with random error.

Determine which statements are false. Explain.

A. I only B. II only C. III only D. I, II, and III

E. The answer is not given by (A), (B), (C), or (D).

Solution: (E)

(I) is wrong since stochastic trends are caused by random variation and could not be characterized by explainable changes in direction.

(II) is true since deterministic trend is a function estimated by regression, changes slowly, and can be extrapolated.

(III) is wrong since stochastic trends are typically attributed to high serial correlation with random error.

So (I) and (III) are wrong, so we need to choose (E).

2. *Random walk and stationarity. In this question we introduce random walk with non-zero mean.* A random walk is expressed as $X_1 = Z_1$, $X_t = X_{t-1} + Z_t$, $t = 2, 3, \dots$, where $Z_t \sim WN(\mu_Z, \sigma_Z^2)$, that is, $E(Z_t) = \mu_Z$, $Var(Z_t) = \sigma_Z^2$, and $Cov(Z_t, Z_s) = 0$ for $t \neq s$. Determine which statements are true with respect to a random walk model; show calculations and provide complete explanations.

I. If $\mu_Z \neq 0$, then the random walk is nonstationary in the mean.
(*Hint: Nonstationary in the mean means that the mean changes with time.*)

II. If $\sigma_Z^2 = 0$, then the random walk is nonstationary in the variance.
(*Hint: Nonstationary in the variance means that the variance changes with time.*)

III. If $\sigma_Z^2 > 0$, then the random walk is nonstationary in the variance.

Answer:

Solution: (I) $\boxed{\text{True.}}$

If $\mu_Z \neq 0$, then $E(X_1) = E(Z_1) = \mu_Z \neq 0$

$E(X_2) = E(X_1 + Z_2) = E(X_1) + E(Z_2) = 2\mu_Z \neq 0$

$E(X_3) = E(X_2 + Z_3) = E(X_2) + E(Z_3) = 3\mu_Z \neq 0$

So, we could get $E(X_t) = t\mu_Z \neq 0$

Then, we get the random walk is nonstationary in the mean.

Answer:

(II) $\boxed{\text{False}}$

If $\sigma_Z^2 = 0$, then $var(X_1) = var(Z_1) = \sigma_Z^2 = 0$

$var(X_2) = var(X_1 + Z_2) = var(X_1) + var(X_2) + 2cov(X_1, Z_2)$
$= \sigma_Z^2 + \sigma_Z^2 + 2\times 0 = 2\sigma_Z^2 = 0$

$var(X_3) = var(X_2 + Z_3) = var(X_2) + var(Z_3) +$
$cov(X_2, X_3) = 2\mu_Z^2 + \mu_Z^2 = 3\mu_Z^2 = 0$

$\Rightarrow var(X_t) = t\sigma_Z^2 = 0$ where $t = 1, 2, \dots$, $\sigma_Z^2 = 0$

Then the variance of random walk is nonstationary in the variance is false.

Answer:

(III) $\boxed{\text{True.}}$

From (II), we get $var(X_t) = t\sigma_Z^2$, if $t>0$, then $var(X_t)$ changes as t changes, so the random walk is non-stationary in the variance is true.

3. *Calculation of sample acf.* You are given the following stock prices of company CAS:

Day	Stock Price
1	538
2	548
3	528
4	608
5	598
6	589
7	548
8	514
9	501
10	498

Calculate the sample autocorrelation at lag 3.

Hints:
(i) We are given a sample of size $n = 10$ to estimate autocorrelation at lag 3: $\rho(3) = Cor(X_1, X_4) = \frac{\gamma(3)}{\gamma(0)}$,
– for definition of autocorrelation at lag 3 see Week 1 slide 52 or (2.1.3) on p. 6 of Lecture Notes.
(ii) General formulas for calculating sample mean and covariance are given on slide 38 of week 1 and in §1.2 on p. 4 of Lecture notes for week 1. To estimate $\rho(3) = Cor(X_1, X_4)$ we have:

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t, \quad \hat{\rho}_3 = \frac{\hat{\gamma}(3)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-3}(x_t - \bar{x})(x_{t+3} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}.$$

Solution:

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t = \frac{1}{10}(538+548+528+608+598+589+548+514+501+498) = 547$$

$$\hat{\rho}_3 = \frac{\hat{\gamma}(3)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-3}(X_t-\bar{X})(X_{t+3}-\bar{X})}{\sum_{t=1}^{n}(X_t-\bar{X})^2}$$

$$= \frac{(538-547)(608-547)+(548-547)(598-547)+(528-547)(589-547)+(608-547)(548-547)+(598-547)(514-547)+(589-547)(501-547)+(548-547)(498-547)}{(538-547)^2+(548-547)^2+(528-547)^2+(608-547)^2+(598-547)^2+(589-547)^2+(548-547)^2+(514-547)^2+(501-547)^2+(498-547)^2}$$

$$= -0.347$$

4. *Polyroot command in R. Recall from algebra, that a function* $f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$ *is called a polynomial function of order* n*. Roots of a polynomial function* f *are solutions of the equation* $f(z) = 0$*. Roots of a quadratic equation* $ax^2 + bx + c = 0$ *are given by the formula* $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.

Let $f(z) = 1 - 2z$ and $g(z) = 1 - 0.45z + 0.05z^2$. Find their roots, show calculations. Check your answers using R command *polyroot*:

$> polyroot(c(1, -2))$
$> polyroot(c(1, -0.45, 0.05))$. (*Do not forget to include your output!*)

Solution: $f(z) = 1 - 2z = 0 \Rightarrow z = 0.5$

$g(z) = 1 - 0.45z + 0.05z^2 = 0$

$\Rightarrow z_{1,2} = \frac{-(-0.45) \pm \sqrt{0.45^2 - 4 \cdot 0.05 \cdot 1}}{2 \cdot 0.05}$

$= \frac{0.45 \pm 0.05}{0.1}$

$\Rightarrow z_1 = 5, \ z_2 = 4$

By using R, it's checked to be right.

5. *Model identification.* You are given the following information about a MA(1) model with coefficient $|\theta_1| < 1$: $\rho_1 = -0.4$, $\rho_k = 0$, $k = 2, 3, \ldots$. Determine the value of θ_1.

Solution:

MA(1) $X_t = Z_t + \theta_1 Z_{t-1}$ where $Z_t \sim WN(0, \sigma_Z^2)$ and $|\theta_1| < 1$

Since $\rho_1 = \frac{\gamma_1}{\gamma_2} = \frac{\theta_1}{1+\theta_1^2} = -0.4$,

then $0.4\theta_1^2 + \theta_1 + 0.4 = 0$

$2\theta_1^2 + 5\theta_1 + 2 = 0$

$(2\theta_1 + 1)(\theta_1 + 2) = 0$

$\Rightarrow \theta_1 = -0.5$ or -2

Answer:

Since $|\theta_1| < 1$, then $\theta_1 = \boxed{-0.5}$

6. *Gaussian White Noise and its square.* Let $\{Z_t\}$ be a Gaussian white noise, that is, a sequence of i.i.d. normal r.v.s each with mean zero and variance 1. Let $Y_t = Z_t^2$.

(a) Using R generate 300 observations of the Gaussian white noise Z. Plot the series and its acf.

(b) Using R, plot 300 observations of the series $Y = Z_t^2$. Plot its acf.

(c) Analyze graphs from (a) and (b).

– Can you see a difference between the plots of graphs of time series Z and Y? From the graphs, would you conclude that both series are stationary (or not)?

– Is there a noticeable difference in the plots of acf functions ρ_Z and ρ_Y? Would you describe Y as a non-Gaussian white noise sequence based on your plots?

Provide full analysis of your conclusions.

(d) Calculate the second-order moments of Y: $\mu_Y(t) = E(Y_t)$, $\sigma_Y^2(t) = Var(Y_t)$, and $\rho_Y(t, t+h) = Cor(Y_t, Y_{t+h})$. Do your calculations support your observations in (c)?

Hints: (i) Slides 65 and 68 of week 1 have R commands to generate MA(1) time series. White Noise is a MA(1) process with coefficient $\theta_1 = 0$. Here is a more direct code to generate WN $\{Z_t\} \sim N(0, 1)$:

```
Z <- rnorm(300)
plot.ts(Z, xlab = "", ylab = "")
acf(Z, main = "ACF")
```

(ii) Useful for part (d): For $X \sim N(0, \sigma^2)$, $E(X^4) = 3(\sigma^2)^2$.

(d) $Z \sim N(0,1)$, $\mu_Y(t) = E(Y_t) = E(Z_t^2) = var(Z_t) - [E(Z_t)]^2$

$= \sigma^2 = 1$

$\sigma_Y^2(t) = var(Y_t) = E(Y_t^2) - (E(Y_t))^2 = E(Z_t^4) - 1^2 = 3 - 1 = 2$

$\rho_Y(t, t+h) = Cor(Y_t, Y_{t+h}) = Cor(Z_t^2, Z_{t+h}^2)$

$= \dfrac{Cov(Z_t^2, Z_{t+h}^2)}{\sqrt{var(Z_t^2) Var(Z_{t+h}^2)}} = 0$

Yes, it supports my observations in (c). since $\mu_Y(t)$ is 1 which is the same with my observation, and it's non-Gaussian WN due to constant μ, constant var and it's uncorrelated.

G1. Let $\{Z_t\}$ be *Gaussian* white noise, i.e. $\{Z_t\}$ is a sequence of i.i.d. normal r.v.s each with mean zero and variance 1. Define

$$X_t = \begin{cases} Z_t, & \text{if } t \text{ is even;} \\ (Z_{t-1}^2 - 1)/\sqrt{2}, & \text{if } t \text{ is odd} \end{cases}$$

Show that $\{X_t\}$ is $\mathrm{WN}(0,1)$ (that is, variables X_t and $X_{t+k}, k \geq 1$, are uncorrelated with mean zero and variance 1) but that X_t and X_{t-1} are **not** i.i.d.

Solution:

$Z_t \overset{\text{iid}}{\sim} N(0,1)$

$$\mu_X(t) = E(X_t) = \begin{cases} E(Z_t) = 0 & \text{if } t \text{ is even} \\ \left(E(Z_{t-1}^2) - 1\right)/\sqrt{2} = 0 & \text{if } t \text{ is odd} \end{cases}$$

$$\mathrm{var}(X_t) = \begin{cases} \mathrm{var}(Z_t) = 1 & \text{if } t \text{ is even} \\ \frac{1}{2}\mathrm{var}(Z_{t-1}^2) = \frac{1}{2} \times \left(E(Z_{t-1}^4) - E(Z_{t-1}^2)^2\right) = \frac{1}{2}(3-1) = 1 & \text{if } t \text{ is odd.} \end{cases}$$

Then, $\{X_t\}$ is $WN(0,1)$

$$\mathrm{Cor}(X_t, X_{t+k}) = \frac{\mathrm{Cov}(X_t, X_{t+k})}{\sqrt{\mathrm{Var}(X_t)\mathrm{Var}(X_{t+k})}}$$

$$\Rightarrow \mathrm{Cov}(X_t, X_{t+k}) = E(X_t X_{t+k}) - E(X_t)E(X_{t+k})$$

Case 1: X_t, X_{t+k} are both odd or even, they are i.i.d then $\mathrm{Cov}(X_t, X_{t+k}) = 0$

Case 2: X_t, X_{t+k}: one is even, one is odd,

$$\mathrm{Cov}(X_t, X_{t+k}) = E\left(\frac{Z_t(Z_{t-1}^2 - 1)}{\sqrt{2}}\right) - E(Z_t)E\left(\frac{Z_{t-1}^2 - 1}{\sqrt{2}}\right) = 0$$

So X_t and X_{t+k} are uncorrelated since $\mathrm{Cov}(X_t, X_{t+k}) = 0$.

For X_t and X_{t-1}, it's clear that one is even case the other must be odd case. then X_t and X_{t-1} are not iid.

G2. If $\{X_t\}$ and $\{Y_t\}$ are uncorrelated stationary sequences, i.e., if X_r and Y_s are uncorrelated for every r and s, show that $\{X_t+Y_t\}$ is stationary with autocovariance function equal to the sum of the autocovariance functions of $\{X_t\}$ and $\{Y_t\}$.

Proof:

$E(X_t+Y_t)=\mu_x+\mu_y$

$Var(X_t+Y_t)=Var(X_t)+Var(Y_t)=\sigma_x^2+\sigma_y^2$

$$\gamma_{x+y}(h)=Cov(X_{t+h}+Y_{t+h},\ X_t+Y_t)$$
$$=Cov(X_{t+h},X_t)+Cov(Y_{t+h},Y_t)+0$$
$$=\gamma_x(h)+\gamma_y(h)$$

Since X and Y are uncorrelated, and mean and autocovariance functions are free of t, then the process $\{X_t+Y_t\}$ is stationary.