MA 641 Time Series Analysis
Homework 1

Nicolas Jorquera

Problem 1

The Cauchy-Schwarz inequality says that for any random variables X and Y, $[E(XY)]^2 \leq E(X^2)E(Y^2)$. If $E(X^2) = 0$, then $P(X = 0) = 1$ and $E(XY) = 0$ so the inequality holds. So assume $E(X^2) > 0$ and $E(Y^2) > 0$. If $E(X^2) = \infty$ or $E(Y^2) = \infty$, then the right-hand side of the inequality is ∞, and the inequality holds.

Part A

Assume $0 < E(X^2) < \infty$ and $0 < E(Y^2) < \infty$. For any constants a and b, $0 \leq E[(aX + bY)^2] = a^2E(X^2) + 2E(Y^2) + 2abE(XY)$, and $0 \leq E[(aX - bY)^2] = a^2E(X^2) + b^2E(Y^2) - 2abE(XY)$.
Using these inequalities with $a = \sqrt{E(Y^2)}$ and $b = \sqrt{E(X^2)}$, write the inequality with $E(XY)$ on one side, and show the Cauchy-Schwarz inequality holds.

As mentioned, the **Cauchy-Schwarz inequality** states that for any two vectors A and B, the square of the dot product of A and B is less than or equal to the product of the dot product A with itself and the dot product B with itself. In this case the inequality can be written as shown above: $[E(XY)]^2 \leq E(X^2)E(Y^2)$. The **Expected Value**; represented by $\mathbb{E}$ is the mean value of a random variable over a large number of experiments or trials.

The problem is using two inequalities involving the expectation of random variables, X and Y, as well as constants a and b. These inequalities are given by:

$$0 \leq E[(aX + bY)^2] = a^2E(X^2) + b^2E(Y^2) + 2abE(XY)$$
$$0 \leq E[(aX - bY)^2] = a^2E(X^2) + b^2E(Y^2) - 2abE(XY)$$

In each inequality, the squared quantity must always be non-negative, hence we can write $0 \leq$ on the left. Then, we can substitute $a = \sqrt{\mathbb{E}(Y^2)}$ and $b = \sqrt{\mathbb{E}(X^2)}$ into the above equations; which results in:

$$0 \leq E[(\sqrt{E(Y^2)}X + \sqrt{E(X^2)}Y)^2] = E(Y^2)E(X^2) + E(X^2)E(Y^2) + 2\sqrt{E(X^2)E(Y^2)}E(XY)$$
$$0 \leq E[(\sqrt{E(Y^2)}X - \sqrt{E(X^2)}Y)^2] = E(Y^2)E(X^2) + E(X^2)E(Y^2) - 2\sqrt{E(X^2)E(Y^2)}E(XY)$$

Here, the right side of both equations simplifies to $2E(X^2)E(Y^2)$ because the $+2\sqrt{E(X^2)E(Y^2)}E(XY)$ and $-2\sqrt{E(X^2)E(Y^2)}E(XY)$ terms cancel out when we add the two inequalities together.

$$\begin{aligned} 0 &\leq 2E(X^2)E(Y^2) + 2E(X^2)E(Y^2) \\ &\leq 4E(X^2)E(Y^2) \\ &\leq E(X^2)E(Y^2) \end{aligned}$$

Since the square of any real number is non-negative, we can further simplify this to:

$$0 \leq [E(XY)]^2 \leq E(X^2)E(Y^2)$$

Which results in the Cauchy-Schwarz inequality for expectations

Part B

Let $U = X - E(X)$ and $W = Y - E(Y)$. Assume all expectations exist. Then $E(U, W) = \text{Cov}(X, Y)$. Use the Cauchy-Schwarz inequality to show that $-1 \leq \text{corr}(X, Y) \leq 1$.

We know that the **Variance** $Var(Z)$ is the expectation of the squared deviation of a random variable from its mean, denoted as $\mathbb{E}[(Z - \mathbb{E}(Z))^2]$. We also know that **Covariance** $Cov(X, Y)$ is a measure of how much two random variables change together, denoted as $E[(X - \mathbb{E}(X))(Y - E(Y))]$. Lastly, **Correlation** $corr(X, Y)$ is a measure of the linear relationship between two random variables, which is the covariance divided by the product of their standard deviations, (*given below*) and always lies between -1 and 1.

$$corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}$$

We are given: $U = X - E(X)$, $W = Y - E(Y)$ and $Cov(X, Y) = E(UW)$. Then by **Cauchy-Schwarz Inequality:**

$$[E(UW)]^2 \leq E(U^2)E(W^2)$$

$$[E(Cov(X, Y))]^2 \leq E[(X - E(X))^2]E[(Y - E(Y))^2]$$

This simplifies to:

$$[E(Cov(X, Y))]^2 \leq Var(X)Var(Y)$$

We can divide both sides by $Var(X)Var(Y)$:

$$\frac{[E(Cov(X, Y))]^2}{Var(X)Var(Y)} \leq \frac{Var(X)Var(Y)}{Var(X)Var(Y)}Var(X)Var(Y)$$

Since we already know that the expression on the left above is the square root of $corr(X, Y)$; we can square both sides to show:

$$[corr(X, Y)]^2 \leq 1$$

Since the square of the correlation is less than or equal to 1, it follows that the correlation itself is between -1 and 1 (*These are the restrictions of a squared product inequality*):

$$-1 \leq corr(X, Y) \leq 1$$

Proving the boundedness of the correlation coefficient using the Cauchy-Schwarz inequality.

Problem 2

The first difference is $W_t = Y_t – Y_{t-1}$. The second difference is $Z_t = W_t - W_{t-1}$. Suppose the model is $Z_t = e_t$ where the e_t are iid with $E(e_t) = 0$ and $Var(e_t) = \sigma^2$. Write the model in the form $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$.

This problem is asking us to transform one model of a time series into another, and we are given a time series Y_t, from which two other time series are constructed: W_t and Z_t which are the first and second differences of Y_t respectively. The first difference of the time series is usually used to remove linear trends from the series, while the second difference is used to remove quadratic trends. This is because:

> *"For a discrete time-series, the* ***second-order difference*** *represents the curvature of the series at a given point in time. If the second-order difference is positive then the time-series is curving upward at that time, and if it is negative then the time series is curving downward at that time."*

We are also given that the second difference series Z_t is a simple random noise process, i.e., $Z_t = e_t$ where the e_t are independently and identically distributed (iid) with mean 0 and variance σ^2. This makes our assumptions easier since all the e_t are random variables that have the same probability distribution and are mutually independent.

Starting with $Z_t = e_t$, we have

$$Z_t = W_t - W_{t-1} = e_t$$

which gives

$$W_t = W_{t-1} + e_t$$

We also know that $W_t = Y_t - Y_{t-1}$, so

$$Y_t - Y_{t-1} = W_{t-1} + e_t$$

Now we can substitute $W_{t-1} = Y_{t-1} - Y_{t-2}$ into this, giving

$$Y_t - Y_{t-1} = (Y_{t-1} - Y_{t-2}) + e_t$$

If we rearrange this, we get

$$Y_t = Y_{t-1} + Y_{t-1} - Y_{t-2} + e_t$$

which simplifies to

$$Y_t = 2Y_{t-1} - Y_{t-2} + e_t$$

Therefore, we have $\phi_1 = 2$ and $\phi_2 = -1$. Therefore the given model $Z_t = e_t$ is equivalent to the model $Y_t = 2Y_{t-1} - Y_{t-2} + e_t$.

Problem 3

Let e_t be a series of independent and identically distributed random variables where $e_t = 1$ represents Heads and $e_t = -1$ represents Tails in a coin flip. Given the R output of rbinom(10,1,0.5), which yields the sequence 0 0 0 0 0 0 1 1 0 0, redefine the 0's as -1's and define $Y_t = \sum_{i=1}^{t} e_i$.

Part A

Plot Y_t on the vertical axis versus time t on the horizontal axis.

Since we are creating a time series Y_t, defined as the cumulative sum of e_t up to time t; we know that Y_t is a random walk. A **Random Walk** is a mathematical object, known as a stochastic or random process, that describes a path that consists of a succession of random steps on some mathematical space such as the integers or real numbers. *Below is the code used to find the solution:*

```r
# Original vector from rbinom function
original_vector <- c(0, 0, 0, 0, 0, 0, 1, 1, 0, 0)

# Convert 0's to -1's to represent e_t
e_t <- ifelse(original_vector == 0, -1, 1)

# Calculate Y_t as the cumulative sum of e_t
Y_t <- cumsum(e_t)

# Plot Y_t against time t
plot(Y_t, type="o", xlab="Time t", ylab="Y_t", main="Random Walk")
```

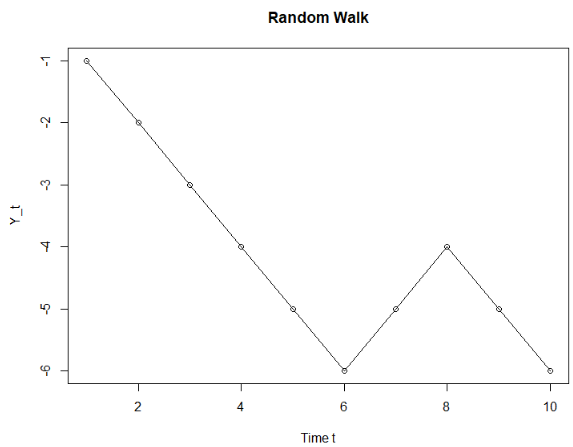

Figure 1: Y_t vs Time (t)

Part B

Consider the process Y_t, a random walk with $E(Y_t) = 0$. Does the plot from part a) appear to show a trend, or are the Y_t values evenly distributed about the horizontal axis?

From observing the plot, the Y_t values do not scatter about the horizontal axis in a roughly even band. Instead, they seem to show a downward trend. However, because this is a random walk with only one realization (*One sample set*), this apparent trend isn't enough to make a conclusion.

Problem 4

Suppose $Y_t = \sum_{s=-\infty}^{\infty} a_s e_{t-s}$ where $\sum_{s=-\infty}^{\infty} a_s^2 < \infty$ and the e_t are iid with $E(e_t) = 0$ and $Var(e_t) = \sigma^2$.

Part A

Assume that the expectation may be taken inside the sum. Find $E(Y_t)$.

We are given that a specific mathematical model for a process Y_t is defined as a weighted sum of independent and identically distributed (iid) random variables e_{t-s} (which have mean 0 and variance σ^2). The weights are the a_s terms, and it's given that the square of these terms sum to a finite number. For Part A we are assuming that we can exchange the order of the sum and the expectation operator:

To calculate $E(Y_t)$, we can apply the expectation to each term in the sum individually, due to **linearity of expectation -** the expected value of the sum of random variables is equal to the sum of their individual expected values, regardless of whether they are independent.

$$E(Y_t) = E(\sum_{s=-\infty}^{\infty} a_s e_{t-s}) = \sum_{s=-\infty}^{\infty} a_s E(e_{t-s})$$

Since it's given that $E(e_t) = 0$, it follows that $E(Y_t) = 0$ for all t, so the mean is constant over time.

Part B

It can be shown that $Cov(Y_t, Y_{t-k}) = \sigma^2 \sum_{i=-\infty}^{\infty} a_{k+i} a_i$. Is the process Y_t stationary? Explain briefly.

A **Stationary Process** in Time Series Analysis is a model whose statistical properties such as mean, variance, and autocorrelation are all constant over time. For a process to be stationary, it has to satisfy the following conditions:

1. The mean/ variance must be constant over time, which implies $E(Y_t) = \mu$ for all t.
2. The covariance between the values at different times depends only on the difference between the times, not on the actual times themselves.

Since we have proved in Part A that the mean is constant, below we will explore the covariance given:

$$Cov(Y_t, Y_{t-k}) = \sigma^2 \sum_{i=-\infty}^{\infty} a_{k+i} a_i$$

Here we can see that the covariance between Y_t and Y_{t-k} depends only on the lag k, and not on the specific time t. Therefore, we can conclude that the process Y_t is stationary.

Problem 5

To show that a process Y_t is not stationary, try to show that $E(Y_t)$ depends on t. If this fails, try to show that $Var(Y_t)$ depends on t. If this fails, show that $\gamma_{t,t-k} = Cov(Y_t, Y_{t-k})$ depends on t.

Part A

Let $Y_t = \sum_{i=1}^{t} e_t$ where the e_t are iid with $E(e_t) = \mu > 0$ and $V(e_t) = \sigma^2$. Show that Y_t is not stationary.

Since the expectation of Y_t is:

$$E(Y_t) = E(\sum_{i=1}^{t} e_t) = \sum_{i=1}^{t} E(e_t) = \sum_{i=1}^{t} \mu = t\mu$$

we can see that since the expected value $E(Y_t)$ depends on t, the process Y_t is not stationary.

Part B

Let $Y_t = \sum_{i=1}^{t} e_t$ where the e_t are iid with $E(e_t) = 0$ and $V(e_t) = \sigma^2$. Show that Y_t is not stationary.

In this case we'll examine the Variance: $Var(Y_t) = Var(\sum_{i=1}^{t} e_t) = \sum_{i=1}^{t} Var(e_t) = \sum_{i=1}^{t} \sigma^2 = t\sigma^2$ Since the variance $Var(Y_t)$ depends on t, the process Y_t is not stationary.

Problem 6

If X and Y are dependent but $Var(X) = Var(Y)$, find $Cov(X + Y, X - Y)$.

In this problem we are given dependent variables, (meaning their outcome influences each other); and need to find the covariance of these two dependent variables. As mentioned above, the formula for covariance of two random variables X and Y is:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$$

where $E[X]$ and $E[Y]$ are the expected values of X and Y, respectively. Some important properties to remember are:

$$Cov(X, X) = Var(X)$$
$$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$$
$$Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$$

Because X and Y have the same variance: $Var(X) = Var(Y)$ $(given)$, from the first property above we also know that: $Cov(X, X) = Cov(Y, Y)$.

To solve this problem:

$$\begin{aligned} Cov(X + Y, X - Y) &= Cov(X, X - Y) + Cov(Y, X - Y) \\ &= Cov(X, X) - Cov(X, Y) + Cov(Y, X) - Cov(Y, Y) \end{aligned}$$

Because X and Y are dependent we can simplify; knowing: $Cov(X, Y) = Cov(Y, X)$, and $Var(X) = Var(Y)$:

$$\begin{aligned} Cov(X + Y, X - Y) &= Var(X) - Cov(X, Y) + Cov(X, Y) - Var(Y) \\ &= Var(X) - Var(Y) = 0 \end{aligned}$$

Therefore, $Cov(X + Y, X - Y) = 0$, despite X and Y being dependent. This shows that dependence doesn't necessarily imply non-zero compliance.

Problem 7

Suppose $Y_t = X$ for all t where $E(X) = \mu$ and $V(X) = \sigma^2$.

Part A

Show that Y_t is stationary.

As mentioned; to show that Y_t is stationary, we need to prove that the mean and variance are constant, and that the autocovariance depends only on the lag, not on time. We are given that $E(Y_t) = E(X) = \mu$ and $V(Y_t) = V(X) = \sigma^2$ for all t; which proves that the mean and variance are constant.

Part B

Find the autocovariance function γ_k for Y_t.

The autocovariance function is defined as $\gamma_k = Cov(Y_t, Y_{t-k})$. Since $Y_t = X$ for all t, we have $Y_t = Y_{t-k} = X$ for any t and k. Therefore, $\gamma_k = Cov(X, X) = \sigma^2$ for all $k \neq 0$, and $\gamma_k = 0$ for $k = 0$.

Problem 8

Use R to answer the following questions:

Part A

Simulate a completely random process of length 48 with independent, normal values. Plot the time series plot. Does it look "random"? Repeat this exercise several times with a new simulation each time.

A **normal distribution** is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent.

```
for(i in 1:5){
  Y <- rnorm(48)
  plot(Y, type="l", main="Normal Distribution")
}
```

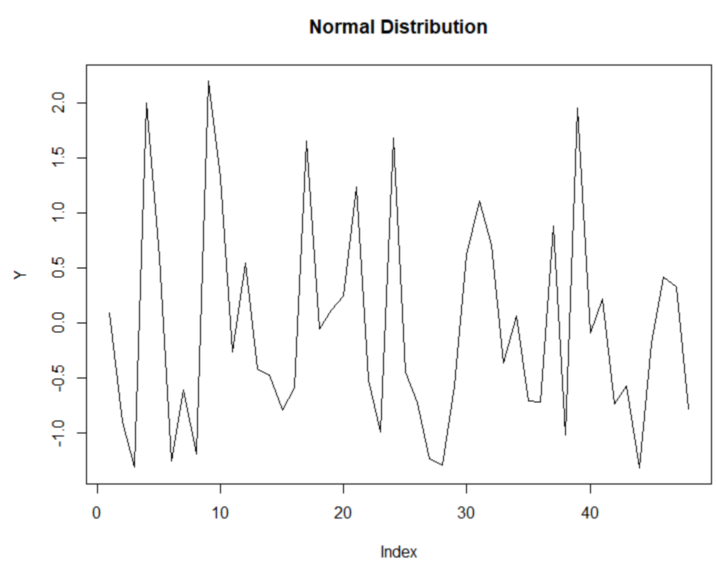

Figure 2: Y_t vs Time (t)

For all 5 plots generated, the data results in no discernible pattern or trend in the data. Each graph also looks distinct, which is expected since it is a random process.

Part B

Simulate a completely random process of length 48 with independent, chi-square distributed values, each with 2 degrees of freedom. Display the time series plot. Does it look "random" and non-normal? Repeat this exercise several times with a new simulation each time.

A **chi-square distribution** is the distribution of a sum of the squares of k independent standard normal random variables. It is a particular case of the gamma distribution and is one of the most widely used probability distributions in inferential statistics, notably in hypothesis testing or in construction of confidence intervals.

```
for(i in 1:5){
  Y <- rchisq(48, df=2)
  plot(Y, type="l", main="Chi-square Distribution")
}
```

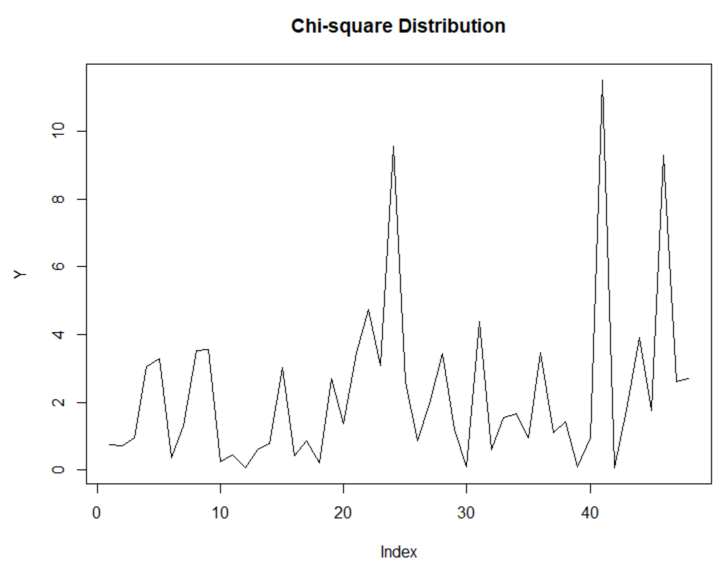

Figure 3: Y_t vs Time (t)

For all 5 plots generated, the data results in no discernible pattern or trend in the data. Each graph also looks distinct, which is expected since it is a random process. However unlike Part A, the distribution of data is not symmetric. The chi-square distribution is skewed to the right; however it makes no discernible difference in the outcome.

Part C

Simulate a completely random process of length 48 with independent, t-distributed values each with 5 degrees of freedom. Construct the time series plot. Does it look "random" and non-normal? Repeat this exercise several times with a new simulation each time.

A **T-distribution**, or Student's T-distribution, is a probability distribution that is used to estimate population parameters when the sample size is small and/or when the population variance is unknown. The **degrees of freedom** of a distribution is the number of values in the final calculation of a statistic that are free to vary.

```
for(i in 1:5){
  Y <- rt(48, df=5)
  plot(Y, type="l", main="T-Distribution")
}
```

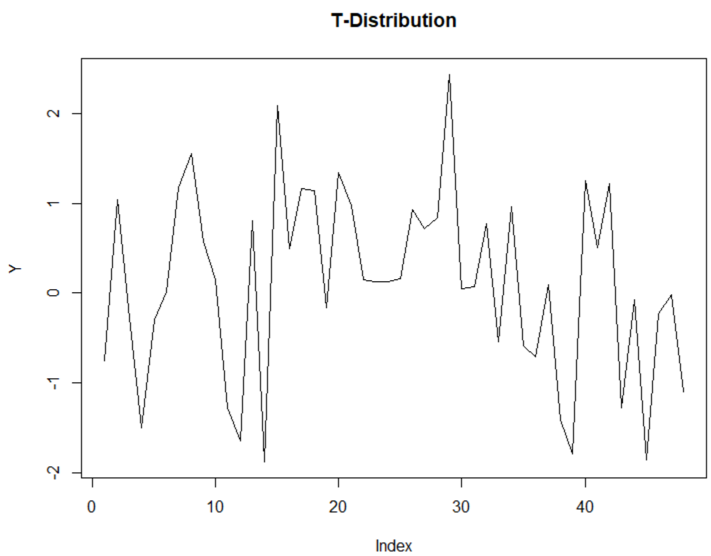

Figure 4: Y_t vs Time (t)

For all 5 plots generated, the data results in no discernible pattern or trend in the data. Each graph also looks distinct, which is expected since it is a random process. However unlike Part A, the distribution of data is not symmetric. The T-distribution has heavier tails than the normal distribution; however it makes no discernible difference in the outcome.