IE313: Time Series Analysis
Problem Sets 1

Student Name: Nguyen Minh Duc
Student ID: 20202026

Problem 1. Suppose $Y_t = 5 + 2t + X_T$, where $\{X_t\}$ is a zero-mean stationary series with autocovariance function γ_k.

(a) Find the mean function for $\{Y_t\}$.

(b) Find the autocovariance function for $\{Y_t\}$.

(c) Is $\{Y_t\}$ stationary? Why or why not?

Solution.

(a) **Find the mean function for $\{Y_t\}$.**

$$E[Y_t] = E[5 + 2t + X_t] = E[5] + E[2t] + E[X_t] = 5 + 2t.$$

(b) **Find the autocovariance function for $\{Y_t\}$.**

$$\begin{aligned}
Cov[Y_t, Y_{t-k}] &= Cov[5 + 2t + X_t, 5 + 2(t-k) + X_{t-k}] \\
&= Cov[5 + 2t, 5 + 2(t-k)] + Cov[5 + 2t, X_{t-k}] + Cov[X_t, 5 + 2(t-k)] + Cov[X_t, X_{t-k}] \\
&= 0 + 0 + 0 + \gamma_k \\
&= \gamma_k
\end{aligned}$$

(c) **Is $\{Y_t\}$ stationary? Why or why not?**
Note that $E[Y_t] = 5 + 2t$, which is not constant. Therefore, $\{Y_t\}$ is not stationary.

Problem 2. Let $\{Y_t\}$ be stationary with autocovariance function γ_k. Let $\overline{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$. Show that

$$\begin{aligned}
Var\left(\overline{Y}\right) &= \frac{\gamma_0}{n} + \frac{2}{n}\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\gamma_k \\
&= \frac{1}{n}\sum_{k=-n+1}^{n-1}\left(1 - \frac{|k|}{n}\right)\gamma_k
\end{aligned}$$

Solution.

$$
\begin{aligned}
Var\left(\overline{Y}\right) &= Var\left(\frac{1}{n}\sum_{t=1}^{n} Y_t\right) \\
&= \frac{1}{n^2} Var\left(\sum_{t=1}^{n} Y_t\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov\left(Y_i, Y_j\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var(Y_i) + \sum_{i\neq j} Cov\left(Y_i, Y_j\right)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} \gamma_0 + 2\sum_{i<j} Cov\left(Y_i, Y_j\right)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} \gamma_0 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov\left(Y_i, Y_j\right)\right)
\end{aligned}
$$

Let's change the variable j into $k = j - i$. Then k would go from 1 to $n - i$. The above expression becomes

$$
\begin{aligned}
Var\left(\overline{Y}\right) &= \frac{1}{n^2}\left(\sum_{i=1}^{n} \gamma_0 + 2\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} Cov\left(Y_i, Y_{i+k}\right)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} \gamma_0 + 2\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \gamma_k\right) \\
&= \frac{1}{n^2}\left(n\gamma_0 + 2\sum_{k=1}^{n-1}\sum_{i=1}^{n-k} \gamma_k\right) \\
&= \frac{1}{n^2}\left(n\gamma_0 + 2\sum_{k=1}^{n-1} (n-k)\gamma_k\right) \\
&= \frac{\gamma_0}{n} + \frac{2}{n^2}\sum_{k=1}^{n-1} (n-k)\gamma_k \\
&= \frac{\gamma_0}{n} + \frac{2}{n}\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)\gamma_k \qquad (1) \\
&= \frac{\gamma_0}{n} + \frac{1}{n}\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)\gamma_k + \frac{1}{n}\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)\gamma_k
\end{aligned}
$$

Note that $\gamma_k = Cov(Y_{t+k}, Y_t) = Cov(Y_t, Y_{t+k}) = Cov(Y_t, Y_{t-(-k)}) = \gamma_{-k}$. The last expression becomes

$$
\begin{aligned}
Var\left(\overline{Y}\right) &= \frac{\gamma_0}{n} + \frac{1}{n}\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)\gamma_k + \frac{1}{n}\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)\gamma_{-k} \\
&= \frac{1}{n}\left(1 - \frac{0}{n}\right)\gamma_0 + \frac{1}{n}\sum_{k=1}^{n-1} \left(1 - \frac{|k|}{n}\right)\gamma_k + \frac{1}{n}\sum_{k=-n+1}^{-1} \left(1 - \frac{|k|}{n}\right)\gamma_k \\
&= \frac{1}{n}\sum_{k=-n+1}^{n-1} \left(1 - \frac{|k|}{n}\right)\gamma_k \qquad (2)
\end{aligned}
$$

From (1) and (2), the desired formula is proven.

Problem 3. Suppose $Y_t = \beta_0 + \beta_1 t + X_t$, where $\{X_t\}$ is a zero-mean stationary series with autocovariance function γ_k and β_0 and β_1 are constants.

(a) Show that $\{Y_t\}$ is not stationary but that $W_t = \nabla Y_t = Y_t - Y_{t-1}$ is stationary.

(b) In general, show that if $Y_t = \mu_t + X_t$, where $\{X_t\}$ is a zero-mean stationary series and μ_t is a polynomial in t of degree d, then $\nabla^m Y_t = \nabla\left(\nabla^{m-1} Y_t\right)$ is stationary for $m \geq d$ and nonstationary for $0 \leq m < d$.

Solution.

(a) **Show that $\{Y_t\}$ is not stationary but that $W_t = \nabla Y_t = Y_t - Y_{t-1}$ is stationary.**

Note that $E(Y_t) = E(\beta_0 + \beta_1 t + X_t) = E(\beta_0) + E(\beta_1 t) + E(X_t) = \beta_0 + \beta_1 t$, which is not constant. Therefore, $\{Y_t\}$ is not stationary.

However, $W_t = \nabla Y_t = Y_t - Y_{t-1}$ is stationary. Considering its mean function

$$\begin{aligned} E(W_t) = E\left(Y_t - Y_{t-1}\right) &= E\left[\beta_0 + \beta_1 t + X_t - \beta_0 - \beta_1(t-1) - X_{t-1}\right] \\ &= E\left[\beta_1 + X_t - X_{t-1}\right] \\ &= E\left[\beta_1\right] + E\left[X_t\right] - E\left[X_{t-1}\right] \\ &= \beta_1 = \text{const} \end{aligned} \tag{3}$$

Also, the autocovariance function of W_t is

$$\begin{aligned} \gamma_{t,s}^W = Cov\left(W_t, W_s\right) &= Cov\left(Y_t - Y_{t-1}, Y_s - Y_{s-1}\right) \\ &= Cov\left(\beta_1 + X_t - X_{t-1}, \beta_1 + X_s - X_{s-1}\right) \\ &= Cov\left(X_t - X_{t-1}, X_s - X_{s-1}\right) \\ &= Cov\left(X_t, X_s\right) - Cov\left(X_t, X_{s-1}\right) - Cov\left(X_{t-1}, X_s\right) + Cov\left(X_{t-1}, X_{s-1}\right) \\ &= \gamma_{|t-s|} - \gamma_{|t-s+1|} - \gamma_{|t-s-1|} + \gamma_{|t-s|}. \end{aligned}$$

Note that

$$\begin{aligned} \gamma_{t,t-k}^W &= \gamma_{|t-t+k|} - \gamma_{|t-t+k+1|} - \gamma_{|t-t+k-1|} + \gamma_{|t-t+k|} \\ &= \gamma_{|k|} - \gamma_{|k+1|} - \gamma_{|k-1|} + \gamma_{|k|} \\ &= \gamma_{|0-k|} - \gamma_{|0-k-1|} - \gamma_{|0-k+1|} + \gamma_{|0-k|} \\ &= \gamma_{0,k}^W \quad \forall t, k. \end{aligned} \tag{4}$$

From (3) and (4), W_t is indeed stationary.

(b) **In general, show that if $Y_t = \mu_t + X_t$, where $\{X_t\}$ is a zero-mean stationary series and μ_t is a polynomial in t of degree d, then $\nabla^m Y_t = \nabla\left(\nabla^{m-1} Y_t\right)$ is stationary for $m \geq d$ and nonstationary for $0 \leq m < d$.**

To begin with, let's define some auxiliary terms.
Let $\deg_t(f)$ be the degree of a polynomial f in t; $\mathcal{P}_n$ be the set of polynomials of degree n.

Lemma 1. *For any $f \in \mathcal{P}_0$, i.e., $f(t) = \beta_0 \ \forall t$, where β_0 is a constant, $\nabla f(t) = 0$.*

Proof. $\nabla f(t) = f(t) - f(t-1) = \beta_0 - \beta_0 = 0$. □

Lemma 2. *For any $n \in \mathbb{N}^+$ and $f \in \mathcal{P}_n$, $\nabla f(t) \in \mathcal{P}_{n-1}$, i.e., the difference operator reduces the degree of any polynomial by one.*

Proof. Let $f(t) = \sum_{i=0}^{n} \beta_i t^i$, where β_i are constants for $0 \leq i \leq n$. Then, the first-order difference of $f(t)$ is

$$
\begin{aligned}
\nabla f(t) = f(t) - f(t-1) &= \sum_{i=0}^{n} \beta_i t^i - \sum_{i=0}^{n} \beta_i (t-1)^i \\
&= \beta_0 + \sum_{i=1}^{n} \beta_i t^i - \beta_0 - \sum_{i=1}^{n} \beta_i (t-1)^i \\
&= \sum_{i=1}^{n} \beta_i t^i - \beta_i (t-1)^i \\
&= \sum_{i=1}^{n} \beta_i \left(t^i - (t-1)^i\right) \\
&= \sum_{i=1}^{n} \beta_i \left(t^i - \sum_{j=0}^{i} (-1)^j \binom{i}{j} t^j\right) \\
&= \sum_{i=1}^{n} \beta_i \left(t^i - t^i - \sum_{j=0}^{i-1} (-1)^j \binom{i}{j} t^j\right) \\
&= -\sum_{i=1}^{n} \beta_i \sum_{j=0}^{i-1} (-1)^j \binom{i}{j} t^j.
\end{aligned}
$$

The degree of $\nabla f(t)$ is

$$
\begin{aligned}
\deg(f(t)) = \deg_t &\left(-\sum_{i=1}^{n} \beta_i \sum_{j=0}^{i-1} (-1)^j \binom{i}{j} t^j\right) \\
&= \max_{1 \le i \le n} \left\{ \deg_t \left(\beta_i \sum_{j=0}^{i-1} (-1)^j \binom{i}{j} t^j\right)\right\} \\
&= \max_{1 \le i \le n} \left\{ \max_{0 \le j \le i-1} \left\{ \deg_t \left(\beta_i (-1)^j \binom{i}{j} t^j\right)\right\}\right\} \\
&= \max_{1 \le i \le n} \left\{ \max_{0 \le j \le i-1} \left\{ \deg_t (j) \right\}\right\} \\
&= \max_{1 \le i \le n} \{i-1\} \\
&= n-1.
\end{aligned}
$$

Therefore, $\nabla f(t) \in \mathcal{P}_{n-1}$. □

Corollary 2.1. *For any natural number $n, m \in \mathbb{N}$ and $f \in \mathcal{P}_n$, $\nabla^m f(t) \in \mathcal{P}_{n-m}$ for $n \ge m$, and $\nabla^m f(t) = 0$ for $n < m$.*

Proof. We will prove by strong induction on m. Let $f \in \mathcal{P}_n$, i.e., $f(t) = \sum_{i=0}^{n} \beta_i t^i$, where β_i are constants for $0 \le i \le n$.
Base cases:

- When $m = 0$, $\nabla^0 f(t) = f(t) \in \mathcal{P}_n$. The corollary holds.
- When $m = 1$, by Lemma 2, $\nabla f(t) \in \mathcal{P}_{n-1}$ for $n \in \mathbf{N}^+$, i.e., $n \ge 1 = m$.
  For $n < m \implies n = 0$. By Lemma 1, $\nabla f(t) = 0$. The corollary holds.

Inductive case: Assume that the corollary holds for all $0 \le m \le k$, i.e., $\nabla^m f(t) \in \mathcal{P}_{n-m}$ for $n \ge m$, and $\nabla^m f(t) = 0$ when $n < m$. Considering $m = k+1$: $\nabla^{k+1} f(t) = \nabla\left(\nabla^k f(t)\right)$.
For $n \ge k+1$, by the Induction Hypothesis, $\nabla^k f(t) \in \mathcal{P}_{n-k}$, hence, $\nabla\left(\nabla^k f(t)\right) \in \mathcal{P}_{n-k-1} = \mathcal{P}_{n-(k+1)}$.
For $n < k+1$, $\nabla^{k+1} f(t) = \nabla\left(\nabla^k f(t)\right)$. If $n = k$, by the Induction Hypothesis, $\nabla^k f(t) \in \mathcal{P}_{n-k} = \mathcal{P}_0$, which means $\nabla^k f(t) = \alpha$, where α is a constant. Now, $\nabla^{k+1} f(t) = \nabla(\alpha) = 0$ by Lemma 1. If

$n < k$, $\nabla^k f(t) = 0$, hence, $\nabla^{k+1} f(t) = \nabla(0) = 0$.
Thus, the corollary also holds for $m = k+1$. This completes the mathematical induction that for any natural number $n, m \in \mathbb{N}$ and $f \in \mathcal{P}_n$, $\nabla^m f(t) \in \mathcal{P}_{n-m}$ for $n \geq m$, and $\nabla^m f(t) = 0$ for $n < m$. □

Lemma 3. *For any two random processes $\{F_t\}$, $\{G_t\}$, and a constant $\alpha \in \mathbb{R}$,*

- $\nabla(F_t + G_t) = \nabla F_t + \nabla G_t$,
- $\nabla(cF_t) = c\nabla F_t$.

Proof. This is trivial, as

- $\nabla(F_t + G_t) = (F_t + G_{t-1}) - (F_{t-1} + G_{t-1}) = (F_t - F_{t-1}) + (G_t - G_{t-1}) = \nabla F_t + \nabla G_t$,
- $\nabla(cF_t) = cF_t - cF_{t-1} = c(F_t - f_{t-1}) = c\nabla F_t$.

Hence, the difference operator is a linear operator. □

In general, the m^{th} order difference is a linear operator.

Corollary 3.1. *For any two random processes $\{F_t\}$, $\{G_t\}$, a natural number $m \in \mathbb{N}$, and a constant $\alpha \in \mathbb{R}$,*

- $\nabla^m(F_t + G_t) = \nabla^m F_t + \nabla^m G_t$,
- $\nabla^m(cF_t) = c\nabla^m F_t$.

Proof. Both of them can be proven by Induction on m.
Base cases:

- When $m = 0$, $\nabla^0(F_t + G_t) = F_t + G_t = \nabla^0 F_t + \nabla^0 G_t$, and $\nabla^0(cF_t) = cF_t = c\nabla^0 F_t$. The corollary holds for $m = 0$.
- When $m = 1$, by Lemma 3, the corollary hold for $m = 1$.

Inductive case: Assume the corollary is true for $m = k$, we need to prove it is also true for $m = k+1$.
Consider $\nabla^{k+1}(F_t + G_t)$,

$$\begin{aligned}
\nabla^{k+1}(F_t + G_t) &= \nabla\left(\nabla^k(F_t + G_t)\right) \\
&= \nabla\left(\nabla^k F_t + \nabla^k G_t\right) \ (\because \text{ Induction Hypothesis}) \\
&= \nabla\left(\nabla^k F_t\right) + \nabla\left(\nabla^k G_t\right) \ (\because \text{ Lemma 3}) \\
&= \nabla^{k+1} F_t + \nabla^{k+1} G_t. \qquad (5)
\end{aligned}$$

Also,

$$\begin{aligned}
\nabla^{k+1}(cF_t) &= \nabla\left(\nabla^k(cF_t)\right) \\
&= \nabla\left(c\nabla^k F_t\right) \ (\because \text{ Induction Hypothesis}) \\
&= c\nabla\left(\nabla^k F_t\right) \ (\because \text{ Lemma 3}) \\
&= c\nabla^{k+1} F_t. \qquad (6)
\end{aligned}$$

From (5) and (6), the corollary also holds for $m = k+1$. This completes the mathematical induction that the m^{th} order difference is a linear operator. □

The last lemma before getting to the actual proof of the theorem is about an explicit formula for the m^{th} order difference of any random process.

Lemma 4. *For any natural number $m \in \mathbb{N}$, and any random process $\{F_t\}$,*

$$\nabla^m F_t = \sum_{i=0}^{m} (-1)^i \binom{m}{i} F_{t-i}.$$

Proof. This can also be proven by Induction on m.
Base cases:

- For $m = 0$, $\nabla^0 F_t = F_t = (-1)^0 \binom{0}{0} F_t$. The lemma is true for $m = 0$.
- For $m = 1$, $\nabla F_t = F_t - F_{t-1} = (-1)^0 \binom{1}{0} F_t + (-1)^1 \binom{1}{1} F_{t-1}$. The lemma is true for $m = 1$.

Inductive case: Assume that the lemma is true for $m = k$, we need to prove that it also holds for $m = k + 1$. Consider $\nabla^{k+1} F_t$,

$$
\begin{aligned}
\nabla^{k+1} F_t &= \nabla \left(\nabla^k F_t\right) \\
&= \nabla \left(\sum_{i=0}^{k} (-1)^i \binom{k}{i} F_{t-i} \right) \quad (\because \text{ Induction Hypothesis}) \\
&= \sum_{i=0}^{k} (-1)^i \binom{k}{i} F_{t-i} - \sum_{i=0}^{k} (-1)^i \binom{k}{i} F_{t-i-1} \\
&= (-1)^0 \binom{k}{0} F_{t-0} + \sum_{i=1}^{k} (-1)^i \binom{k}{i} F_{t-i} + \sum_{i=0}^{k} (-1)^{i+1} \binom{k}{i} F_{t-(i+1)} \\
&= (-1)^0 \binom{k}{0} F_{t-0} + \sum_{i=1}^{k} (-1)^i \binom{k}{i} F_{t-i} + \sum_{i=1}^{k+1} (-1)^i \binom{k}{i-1} F_{t-i} \\
&= (-1)^0 \binom{k}{0} F_{t-0} + \sum_{i=1}^{k} (-1)^i \binom{k}{i} F_{t-i} + \sum_{i=1}^{k} (-1)^i \binom{k}{i-1} F_{t-i} + (-1)^{k+1} \binom{k}{k} F_{t-(k+1)} \\
&= (-1)^0 \binom{k}{0} F_{t-0} + \sum_{i=1}^{k} (-1)^i F_{t-i} \left[\binom{k}{i} + \binom{k}{i-1} \right] + (-1)^{k+1} \binom{k}{k} F_{t-(k+1)} \\
&= (-1)^0 \binom{k+1}{0} F_{t-0} + \sum_{i=1}^{k} (-1)^i \binom{k+1}{i} F_{t-i} + (-1)^{k+1} \binom{k+1}{k+1} F_{t-(k+1)} \\
&= \sum_{i=0}^{k+1} (-1)^i \binom{k+1}{i} F_{t-i}.
\end{aligned}
$$

Thus, the lemma is also true for $m = k + 1$. This completes the mathematical induction for the general formula of the m^{th} order difference. □

Just one more Lemma to finish up the prerequisites.

Lemma 5. *For any natural number $m \in \mathbb{N}$, and any zero-mean random process $\{X_t\}$, the expected value $E\left[\nabla^m X_t\right] = 0$.*

Proof. This is trivial,

$$
E\left[\nabla^m X_t\right] = E\left[\sum_{i=0}^{m} (-1)^i \binom{m}{i} X_{t-i} \right] = \sum_{i=0}^{m} (-1)^i \binom{m}{i} E\left[X_{t-i}\right] = 0.
$$

□

Now, we are ready to prove the main theorem.

Theorem. *If $Y_t = \mu_t + X_t$, where $\{X_t\}$ is a zero-mean stationary series and μ_t is a polynomial in t of degree d, then $\nabla^m Y_t = \nabla\left(\nabla^{m-1} Y_t\right)$ is stationary for $m \geq d$ and nonstationary for $0 \leq m < d$.*

Proof. Firstly, let's consider the case when $0 \leq m < d$.

$$\begin{aligned}
E\left[\nabla^m Y_t\right] &= E\left[\nabla^m\left(\mu_t + X_t\right)\right] \\
&= E\left[\nabla^m \mu_t + \nabla^m X_t\right] \ (\because \text{ Corollary 3.1}) \\
&= E\left[\nabla^m \mu_t\right] + E\left[\nabla^m X_t\right] \\
&= E\left[\nabla^m \mu_t\right] + 0 \ (\because \text{ Lemma 5}) \\
&= E\left[\sum_{i=0}^{m-d} \beta_i t^i\right] \ (\because \text{ Corollary 2.1}) \\
&= \sum_{i=0}^{m-d} \beta_i t^i.
\end{aligned}$$

Note that this is a polynomial of degree $m - d > 0$, which has non-constant terms, hence, $E\left[\nabla^m Y_t\right]$ is not constant. Therefore, $\nabla^m Y_t$ is nonstationary for $0 \leq m < d$.
Lastly, let's consider the other case where $m \geq d$.

$$E\left[\nabla^m Y_t\right] = E\left[\nabla^m \mu_t\right].$$

By Corollary 2.1, $\nabla^m \mu_t = \begin{cases} 0, & \text{if } m > d \\ \beta_0, & \text{if } m = d \end{cases}$, which is constant in both cases. The remaining part is the autocovariance function

$$\begin{aligned}
Cov\left[\nabla^m Y_t, \nabla^m Y_s\right] &= Cov\left[\sum_{i=0}^{m}(-1)^i\binom{m}{i}Y_{t-i}, \sum_{i=0}^{m}(-1)^i\binom{m}{i}Y_{s-i}\right] \ (\because \text{ Lemma 4}) \\
&= \sum_{i=0}^{m}\sum_{j=0}^{m} Cov\left[(-1)^i\binom{m}{i}Y_{t-i}, (-1)^j\binom{m}{j}Y_{s-j}\right] \\
&= \sum_{i=0}^{m}\sum_{j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j} Cov\left[Y_{t-i}, Y_{s-j}\right] \\
&= \sum_{i=0}^{m}\sum_{j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}\gamma_{t-i-s+j}.
\end{aligned}$$

Consider $Cov\left[\nabla^m Y_t, \nabla^m Y_{t+k}\right]$,

$$\begin{aligned}
Cov\left[\nabla^m Y_t, \nabla^m Y_{t+k}\right] &= \sum_{i=0}^{m}\sum_{j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}\gamma_{t-i-t-k+j} \\
&= \sum_{i=0}^{m}\sum_{j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}\gamma_{j-i-k}. \qquad (7)
\end{aligned}$$

Also $Cov\left[\nabla^m Y_0, \nabla^m Y_k\right]$,

$$\begin{aligned}
Cov\left[\nabla^m Y_0, \nabla^m Y_k\right] &= \sum_{i=0}^{m}\sum_{j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}\gamma_{0-i-k+j} \\
&= \sum_{i=0}^{m}\sum_{j=0}^{m}(-1)^{i+j}\binom{m}{i}\binom{m}{j}\gamma_{j-i-k}. \qquad (8)
\end{aligned}$$

From (7) and (8), $Cov\left[\nabla^m Y_t, \nabla^m Y_{t+k}\right] = Cov\left[\nabla^m Y_0, \nabla^m Y_k\right]$, and together with the fact that $E\left[\nabla^m Y_t\right]$ is constant, we can conclude that $\nabla^m Y_t$ is stationary for $m \geq d$. $\square$

Problem 4. The data file wages [data(wages) in TSA package of R] contains monthly values of the average hourly wages (in dollars) for workers in the U.S. apparel and textile products industry from July 1981 through June 1987.

Solution.

(a) Display and interpret the time series plot for these data.

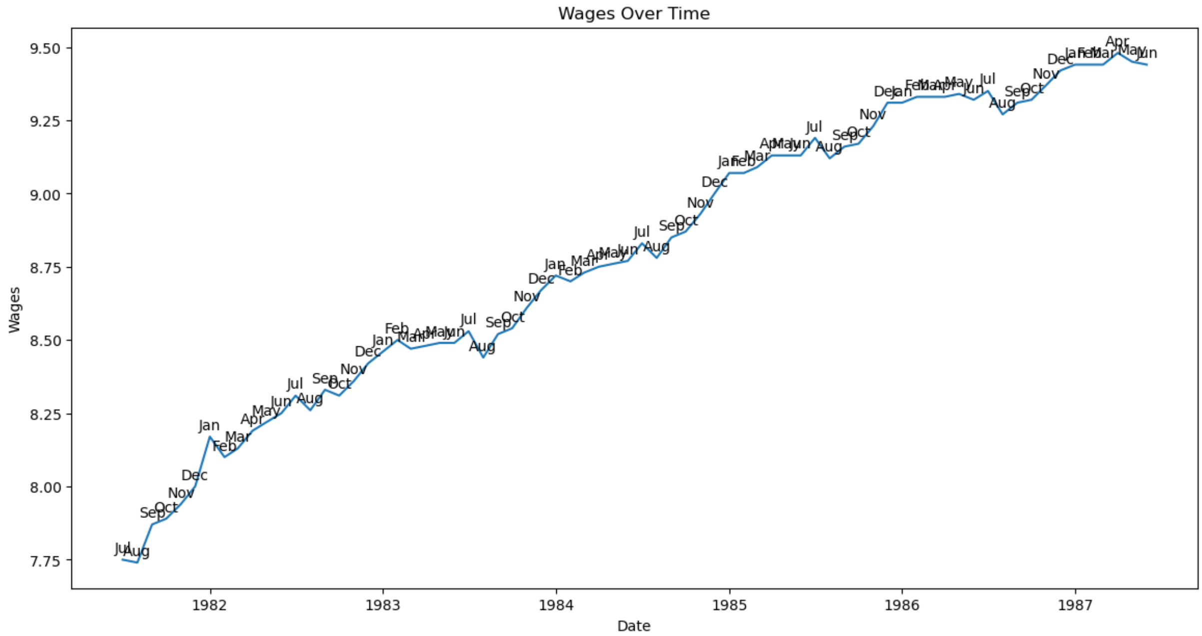

Figure 1: Time series plot for the wages data.

Please refer to Figure 1 for the visualization. As one can observe, there is a positive trend across the timeline and a notable seasonality pattern: it seems like August is the time that the workers receive the least amount of money in a year. However, From September to December, wages increase dramatically, which can be influenced by big holidays during Fall such as Halloween, Thanksgiving, and Christmas.

(b) Use least squares to fit a linear time trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

Coefficient	**Estimate**	**Std. Error**	***t*-value**	$Pr(>\|t\|)$
Intercept	7.9314	0.020	403.314	< 0.001
Time(wages)	0.0234	0.000	50.028	< 0.001

Residual standard error	0.083 with 70 degree of freedom
Multiple R-Squared	0.973
Adjusted R-Squared	0.972
***F*-statistic**	2503. with 1 and 70 df; p-value = 1.58e-56

Table 1: The linear regression output

Please refer to Table 1 for the regression results. The Residual standard error is quite small, 0.083, hence, it indicates that our model is a good fit for the data. The R^2 score of this regression is 0.973, which is quite good. About 97% of the unobserved stochastic component in the wages series can be explained by the linear trend. As for the parameters, note that the standard deviations are small and both p-values are also extremely small, hence, the results are statistically significant.

(c) Construct and interpret the time series plot of the standardized residuals from part (b).

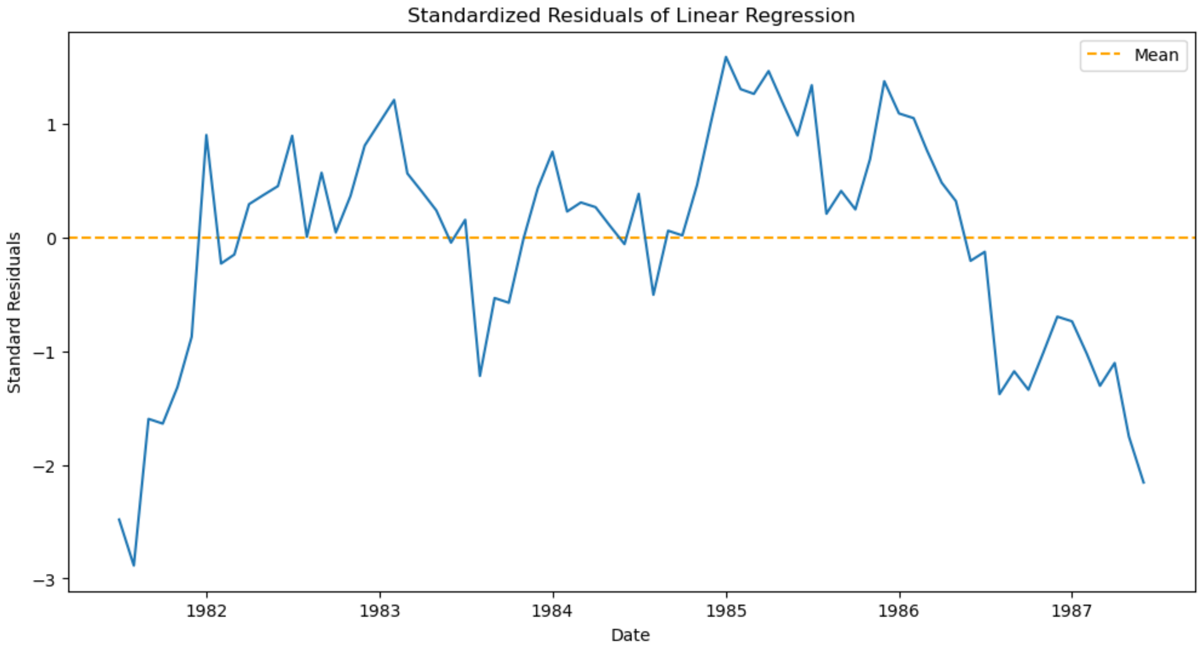

Figure 2: Time series plot for the standardized residuals of Linear Regression.

Please refer to Figure 2 for the visualization of the time series. The plot is not well-centered around the mean, and we can still see some repeating patterns and some correlations related to the time. Therefore, the plot suggests that these standardized residuals of linear regression are not white noise.

(d) Use least squares to fit a quadratic time trend to the wages time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

Coefficient	**Estimate**	**Std. Error**	***t*-value**	$Pr(>\lvert t\rvert)$
Intercept	7.7974	0.021	364.127	< 0.001
Time(wages)	0.0343	0.001	25.328	< 0.001
Time(wages)^2	-0.0001	1.8e-05	-8.282	< 0.001

Residual standard error	0.059 with 69 degree of freedom
Multiple R-Squared	0.986
Adjusted R-Squared	0.986
***F*-statistic**	2494. with 2 and 69 df; p-value = 4.53e-65

Table 2: The quadratic regression output

Please refer to Table 2 for the regression results. The Residual standard error is quite small, 0.059, hence, it indicates that our model is a good fit for the data. The R^2 score of this regression is 0.986, which is quite good. About 98% of the unobserved stochastic component in the wages series can be explained by the quadratic trend. All of the statistics suggest that this quadratic model fits the data better than the linear model. As for the parameters, the standard deviations are small and all of the p-values are also extremely small, hence, the results are statistically significant. Note that the coefficient of the second-order term is very small (-0.0001) suggesting that the quadratic term does not contribute much compared to the linear term.

(e) Construct and interpret the time series plot of the standardized residuals from part (d).

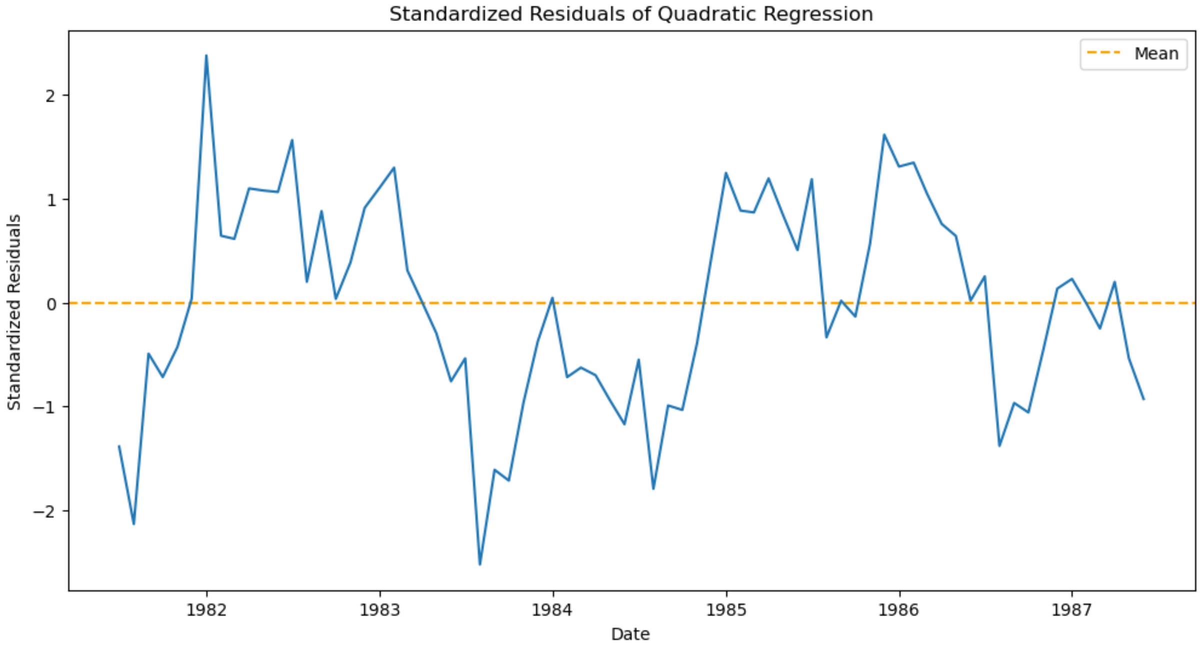

Figure 3: Time series plot for the standardized residuals of Quadratic Regression.

Please refer to Figure 3 for the visualization of the time series. The residuals of quadratic regression look more like random noise compared to the linear ones. The residuals are generally centered around zero, and there is no clear trend, but there are still some recurring patterns.

(f) Consider the residuals from the least squares fit of a quadratic time trend.

The results can be found above.

(g) Perform the runs test on the standardized residuals and interpret the results.

The number of runs	15
Expected number of runs	37
z-score	-5.222
p-value	1.767

Table 3: The Wald–Wolfowitz runs test output

Please refer to Table 3 for the runs test results. As one can observe, the number of actual runs is much smaller compared to the expected number of runs, indicating that neighboring residuals are positively dependent, hence, not a white noise, which confirms our suspicion in part (e). The p-value, which is 1.767, is quite high, hence, the independence could not be rejected under the usual confidence interval.

(h) Calculate and interpret the sample autocorrelations for the standardized residuals.

Please refer to Figure 4 for the visualization of the sample autocorrelations for the standardized residuals. The x-axis denotes the lag, and the y-axis denotes the value of the ACF. Immediately, we can see a decreasing trend in the magnitude of the correlation as the lag increases, i.e., there is a large correlation among neighboring residuals. There is also a seasonality trend that once in about every 6 time steps a bell curve is formed. Moreover, There are a lot of lags whose autocorrelation exceeds two standard errors, for example, lag 1, lag 2, lag 18, and lag 20. This is not the expected behavior of a white noise process, hence, it further reinforces our suspicion in part (e).

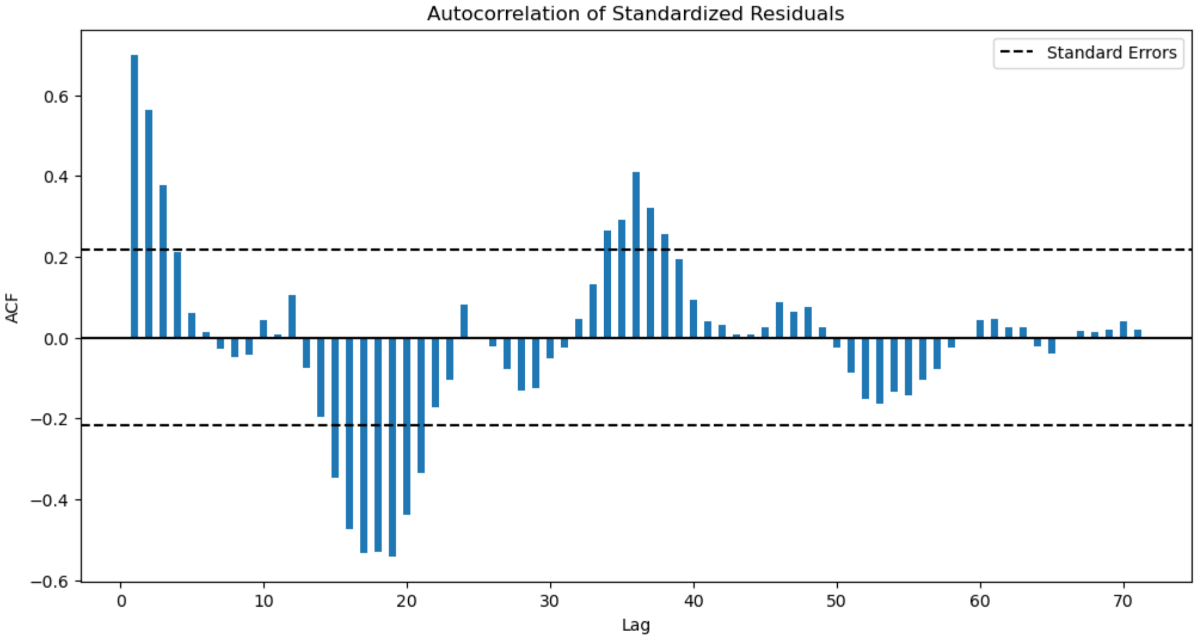

Figure 4: Plot for the sample autocorrelations for the standardized residuals.

(i) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.

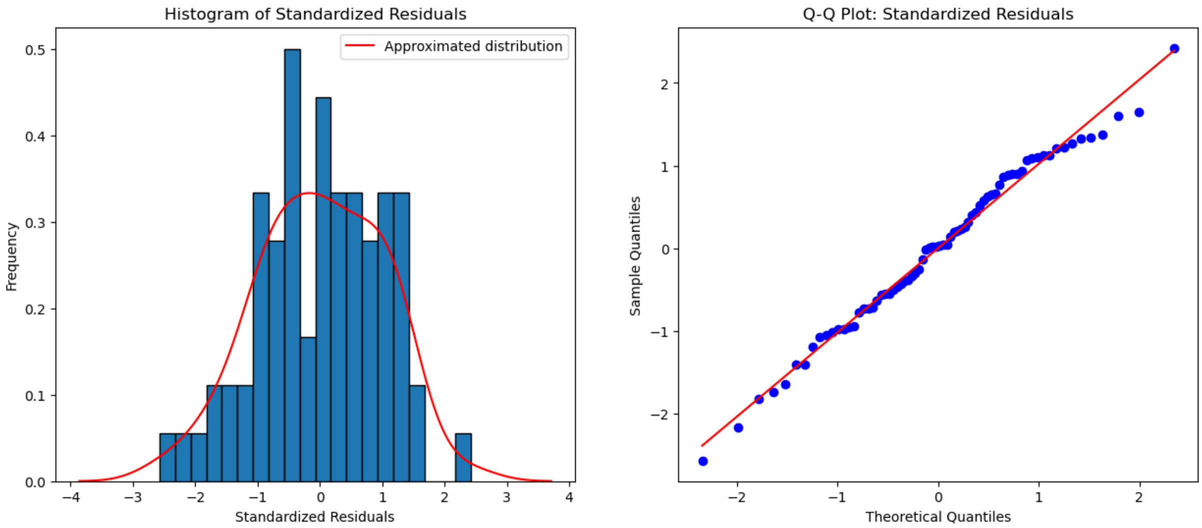

Figure 5: Histograms and normal probability plots for the standardized residuals.

Please refer to Figure 5 for the visualization of the required plots. The histogram provides a rough estimation of the distribution of the standardized residuals, which appears to be roughly a bell curve, suggesting a normal distribution. The distribution is slightly right-skew as it appears to have a longer tail on the right side of the mean, which is close to 0 in this case. As for the Q-Q plot, the data points resemble the reference straight-line $y = x$, further ensuring the normality. Note that there is a big deviation and outliers from the line at the tails (especially the right tail), suggesting a right-skew behavior.