Time Series Analysis (STA 5015)
Chapter 1 Solution

1. (a) The sample space $S = \{(1,g),(0,g),(1,f),(0,f),(1,s),(0,s)\}$.

   (b) If A is the event that the patient is in serious condition, the patient will possibly have insurance or not. Thus, the description is $A = \{(0,s),(1,s)\}$.

   (c) An uninsured patient is possibly in condition g, f or s. Thus, $B = \{(0,g),(0,f),(0,s)\}$.

   (d) $B^c \cup A$ is the event that the patient is insured or in serious condition, $B^c \cup A = \{(1,s),(1,g),(1,f),(0,s)\}$.

2. $P(X=0) = P(\text{1st no sell \& 2nd no sell})$
   $= P(\text{1st no sell})\,P(\text{2nd no sell}) = (1-.3)(1-.6) = .28$
   $P(X=500) = P(\text{1st standard and 2nd no sell} \cup \text{1st no sell and 2nd standard})$
   $= .3(1/2)(.4) + (.7)(.6)(1/2) = .27$
   $P(X=1000) = P(\text{both stndadard} \cup \text{1st deluxe \& 2nd no} \cup \text{1st no \& 2nd deluxe})$
   $= (.3)(1/2)(.6)(1/2) + .3(1/2)(.4) + (.7)(.6)(1/2) = .045 + .27 = .315$
   $P(X=1500) = P(\text{1st standard and 2nd deluxe} \cup \text{1st deluxe and 2nd standard})$
   $= (.3)(1/2)(.6)(1/2) + (.3)(1/2)(.6)(1/2) = .09$
   $P(X=2000) = P(\text{1st deluxe and 2nd deluxe})$
   $= (.3)(1/2)(.6)(1/2) = .045$

3. (a) Recall that $P(X=x) = F_X(x) - F_X(x-)$, where $F_X(x-)$ is the left-limit. Therefore,

$$P(X=1) = F(1) - F(1-) = 1/2 + (1-1)/4 - 1/4 = 1/4,$$
$$P(X=2) = F(2) - F(2-) = 11/12 - (1/2 + (2-1)/4) = 2/12,$$
$$P(X=3) = F(3) - F(3-) = 1 - 11/12 = 1/12.$$

   (b) Observe that $P(1/2 < X < 3/2) = P(X < 3/2) - P(X \le 1/2) = F(3/2-) - F(1/2) = 1/2 + (3/2-1)/4 - (1/2)/4 = 1/2$.

4. (a) In order to become a density function, it must satisfy i) non-negative and ii) add-up to 1. From the first condition, we have that $C(4x - 2x^2) > 0$ for $0 < x < 2$, hence $C > 0$. From the second condition, we have

$$\int_0^2 C(4x - 2x^2)dx = 1.$$

   Solving this gives $C = 3/8$.

   (b) From the definition of expected value, note that

$$E(X^{-1}) = \int_0^2 x^{-1}\frac{3}{8}(4x - 2x^2)dx = \int_0^2 \frac{3}{8}(4-2x)dx = \frac{3}{2}.$$

5. (a) We need to check i) $f(x,y) \geq 0$ and ii) $\int_0^2 \int_0^1 f(x,y)dxdy = 1$. The first condition holds because $x \in (0,1)$ and $y \in (0,2)$, hence x^2 and xy are positive. To check the second condition, observe that

$$\int_0^2 \int_0^1 \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) dxdy = \frac{6}{7}\int_0^2 \left(\frac{1}{3}x^3 + \frac{1}{4}x^2 y\Big|_0^1\right) dy$$

$$= \frac{6}{7}\int_0^2 \left(\frac{1}{3} + \frac{1}{4}y\right) dy = \frac{6}{7}\left(\frac{1}{3}y + \frac{1}{8}y^2\Big|_0^2\right) = \frac{6}{7}(2/3 + 1/2) = 1.$$

(b) $f(x) = \dfrac{6}{7}\displaystyle\int_0^2 \left(x^2 + \frac{xy}{2}\right) dy = \frac{1}{7}(12x^2 + 6x).$

(c) Note that

$$P(X > Y) = \int_0^1 \int_y^1 f(x,y)dxdy = \frac{15}{56}$$

or you can use

$$P(X > Y) = \int_0^1 \int_0^x f(x,y)dydx = \frac{15}{56}.$$

Recall that interchanging integral comes from the Fubini's theorem.

(d) Observe that

$$P\left(Y > \frac{1}{2} \,\Big|\, X < \frac{1}{2}\right) = \frac{P(Y > \frac{1}{2} \cap X < \frac{1}{2})}{P\left(X < \frac{1}{2}\right)},$$

and

$$P\left(Y > \frac{1}{2} \cap X < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^2 \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) dydx = \frac{69}{448}$$

$$P\left(X < \frac{1}{2}\right) = \int_0^{1/2} \left(\frac{12}{7}x^2 + \frac{6}{7}x\right) dx = \frac{5}{28}.$$

Thus, $P\left(Y > \frac{1}{2} | X < \frac{1}{2}\right) = 69/80 = .8625.$

(e) $E(Y) = \displaystyle\int_0^2 \int_0^1 y\frac{6}{7}\left(x^2 + \frac{xy}{2}\right) dydx = 8/7.$

6. It is enough to show that the MGF of $\mathbf{Y}$ agrees with that of $MVN(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$ by the uniqueness of MGF. Note that

$$M_{\mathbf{Y}}(t) := E(e^{t'\mathbf{Y}}) = E(e^{t'(\mathbf{a}+\mathbf{B}\mathbf{X})}) = \exp(t'\mathbf{a})E(e^{t'\mathbf{B}\mathbf{X}})$$

because $\mathbf{a}$ is a constant vector. Furthermore, $t'\mathbf{B}$ becomes $1 \times n$ vector so that we can let $k' = t'\mathbf{B}$ for some $n \times 1$ column vector k. Then,

$$M_{\mathbf{Y}}(t) = \exp(t'\mathbf{a})E(e^{k'\mathbf{X}}) = \exp(t'\mathbf{a})M_{\mathbf{X}}(k), \quad (1)$$

where $M_{\mathbf{X}}(\cdot)$ is the MGF of $\mathbf{X}$. Since the MGF of $\mathbf{X}$ is given by

$$M_{\mathbf{X}}(k) := E(e^{k'\mathbf{X}}) = \exp\left(k'\boldsymbol{\mu} + \frac{k'\boldsymbol{\Sigma}k}{2}\right), \quad (2)$$

plug-in (2) to (1) and using $k' = t'\mathbf{B}$ leads to

$$M_{\mathbf{Y}}(t) = \exp(t'\mathbf{a}) \exp\left(k'\boldsymbol{\mu} + \frac{k'\boldsymbol{\Sigma}k}{2}\right) = \exp(t'\mathbf{a}) \exp\left(k'\boldsymbol{\mu} + \frac{k'\boldsymbol{\Sigma}k}{2}\right)$$

$$= \exp(t'\mathbf{a}) \exp\left(t'\mathbf{B}\boldsymbol{\mu} + \frac{t'\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}'t}{2}\right) = \exp\left(t'(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}) + \frac{t'\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}'t}{2}\right).$$

This is indeed the MGF of MVN with mean vector $\mathbf{a} + \mathbf{B}\boldsymbol{\mu}$ and variance-covariance matrix $\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}'$. Hence, by the uniqueness of MGF, we showed that $\mathbf{Y} \sim \mathbf{MVN}(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$ as required.

7. DIY

8. Problem 1.10 From the definition of operator ∇, it is easily seen that

$$\nabla m_t = m_t - m_{t-1} = a + bt + ct^2 - (a + b(t-1) + c(t-1)^2)$$

$$= a + bt + ct^2 - (a + bt - b + ct^2 - 3ct - c) = (b - c) + 2ct.$$

Therefore, ∇m_t is a polynomial of degree 1. Similarly,

$$\nabla^2 m_t = \nabla(\nabla m_t) = \nabla m_t - \nabla m_{t-1} = 2ct - c - (2c(t-1) - c) = 2ct - c - (2ct - 3c) = 2c,$$

which means that $\nabla^2 m_t$ is a polynomial of degree 0 (constant), and

$$\nabla^3 m_t = \nabla(\nabla^2 m_t) = \nabla^2 m_t - \nabla^2 m_{t-1} = 2c - 2c = 0,$$

hence $\nabla^3 m_t = 0$

9. Note that

$$\sum_{j=-q}^{q} a_n m_{t-j} = \sum_{j=-q}^{q} (2q+1)^{-1}(c_0 + c_1(t-j)) = (2q+1)^{-1} \left\{ \sum_{j=-q}^{q} (c_0 + c_1 t) + \sum_{j=-q}^{q} -j \right\}$$

$$= c_0 + c_1 t - \frac{1}{2q+1} \sum_{j=-q}^{q} j = c_0 + c_1 t = m_t$$

since $\sum_{j=-q}^{q} j = 0$.

10. Problem 1.12

   (a) We will first show sufficiency. Suppose that

$$m_t = \sum_j a_j m_{t-j} \tag{3}$$

   for all $m_t = c_0 + c_1 t + \ldots + c_k t^k$. If $m_t = c_0 \neq 0$, then (3) implies that

$$c_0 = \sum_j a_j c_0.$$

Therefore

$$\sum_j a_j = 1. \tag{4}$$

Next, suppose that $m_t = c_1 t, c_1 \neq 0$, then we have from (3) that

$$c_t t = \sum_j a_j c_1 (t-j) = c_1 \sum_j a_j (t-j).$$

Therefore,

$$t = \sum_j a_j t - \sum_j a_j j$$

implies together with (4) that

$$\sum_j j a_j = 0.$$

By iterating above reasoning till $m_t = c_k t^k$, we have that

$$\sum a_j =, \quad \sum_j j^r a_j = 0, \quad \text{for all } r = 1, \dots, k.$$

Now, conversely, suppose that

$$\sum a_j = 1, \quad \sum_j j^r a_j = 0, \quad \text{for all } r = 1, \dots, k,$$

holds. Then,

$$\begin{aligned}
\sum_j a_j m_{t-j} &= \sum_j a_j (c_0 + c_1(t-j) + \dots + c_k (t-j)^k) \\
&= c_0 \sum_j a_j + c_1 \sum_j a_j (t-j) + \dots + c_k \sum_j a_j (t-j)^k \\
= c_0 \sum_j a_j + c_1 (\sum_j a_j t - \sum_j a_j j) &+ c_2 (t^2 \sum_j a_j - 2t \sum_j a_j + \sum_j a_j j^2) + \dots \\
+ c_k \left(\binom{k}{k} t^k \sum_j a_j (-j)^0 \right. &+ \left. \binom{k}{k-1} t^{k-1} \sum_j a_j (-j)^1 + \dots + \binom{k}{0} t^0 \sum_j a_j (-j)^k \right) \\
&= c_0 + c_1 t + \dots + c_k t^k = m_t
\end{aligned}$$

by using $\sum_j j^r a_j = 0$ for all $r = 1, \dots, k$ from the assumption.

Remark. In the above expansion, we have used the Binomial theorem

$$(x+y)^n = \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y + \dots + \binom{n}{n} x^0 y^n$$

$$= \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

(b) As the result of part (a), it is sufficient to check that Spencer's 15 points moving average filter given by

$$(a_0, a_1, \ldots, a_7) = \frac{1}{320}(74, 67, 46, 21, 3, -5, -6, -3), \quad a_j = a_{-j}, j = -1, \ldots, -7$$

satisfies

$$\sum_{j=-7}^{7} a_j = 1, \quad \sum_{j=-7}^{7} j^r a_j = 0, \quad r = 1, 2, 3.$$

This is a straightforward calculation, so we conclude that Spencer's 15 points moving average *preserves cubic trends without distortion.*

11. CS-inequality

(a) There are couple of ways to show

$$|x_1 y_1 + \ldots + x_n y_n|^2 \le (x_1^2 + \ldots + x_n^2)(y_1^2 + \ldots + y_n^2). \tag{5}$$

Maybe the easiest (but tedious) proof is to use induction. For $n = 1$,

$$LHS = |x_1 y_1|^2 = x_1^2 y_1^2 = RHS.$$

Suppose that (5) holds for $n = k - 1$, $k \ge 2$. Then, for $n = k$, observe that

$$\begin{aligned}
LHS &= |x_1 y_1 + \ldots + x_k y_k|^2 = |x_1 y_1 + \ldots + x_{k-1} y_{k-1} + x_k y_k|^2 \\
&= |x_1 y_1 + \ldots + x_{k-1} y_{k-1}|^2 + 2(x_1 y_1 + \ldots + x_{k-1} y_{k-1}) x_k y_k + x_k^2 y_k^2 \\
&\le (x_1^2 + \ldots + x_{k-1}^2)(y_1^2 + \ldots + y_{k-1}^2) + 2(x_1 y_1 + \ldots + x_{k-1} y_{k-1}) x_k y_k + x_k^2 y_k^2
\end{aligned}$$

from the induction assumption. It further leads to

$$= (x_1^2 + \ldots + x_k^2)(y_1^2 + \ldots + y_k^2) - x_k^2(y_1^2 + \ldots + y_{k-1}^2) - y_k^2(x_1^2 + \ldots + x_{k-1}^2) + 2(x_1 y_1 + \ldots + x_{k-1} y_{k-1}) x_k y_k$$

$$= (x_1^2 + \ldots + x_k^2)(y_1^2 + \ldots + y_k^2) - (x_k y_1 - x_1 y_k)^2 - (x_k y_2 - x_2 y_k)^2 - \ldots - (x_k y_{k-1} - x_{k-1} y_k)^2$$

$$\le (x_1^2 + \ldots + x_k^2)(y_1^2 + \ldots + y_k^2) = RHS.$$

Therefore, (5) holds for any $n \ge 1$ as required.
More succinct proof is to use

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (x_i y_j - x_j y_i)^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i^2 y_j^2 - 2 x_i x_j y_i y_j + x_j^2 y_i^2)$$

$$= \sum_{i=1}^{n} x_i^2 \sum_{j=1}^{n} y_j^2 - 2 \left(\sum_{i=1}^{n} x_i y_i \right)^2 + \sum_{j=1}^{n} x_j^2 \sum_{i=1}^{n} y_i^2 \ge 0$$

Hence, we have that

$$\left| \sum_{i=1}^{n} x_i y_i \right|^2 \le \sum_{i=1}^{n} x_i^2 \sum_{j=1}^{n} y_j^2$$

(b) An inner product space is a vector space V equipped with an inner product $\langle \cdot, \cdot \rangle$ on the space V. An inner product on the vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ ($\mathbb{R}$ is the field of real numbers) that assigns to each pair of vectors $\mathbf{u}, \mathbf{v}$ a scalar (real number) such that for α a scalar and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ the following properties hold:

i. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ with equality if and only if $\mathbf{u} = \mathbf{0}$
ii. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
iii. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
iv. $\langle \alpha \mathbf{u}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle$

(c) From the hint, note that the first axiom of inner-product gives that

$$\langle x - ty, x - ty \rangle \geq 0 \quad \text{for all } t.$$

The left-hand side becomes (again by using axioms *iii*) and *iv*))

$$\langle x, x \rangle - t\langle x, y \rangle - t\langle y, x \rangle + t^2 \langle y, y \rangle \geq 0 \quad \text{for all } t.$$

Using axiom *ii*) and rearranging terms gives that

$$t^2 \langle y, y \rangle - 2t\langle x, y \rangle + \langle x, x \rangle \geq 0 \quad \text{for all } t. \tag{6}$$

If $\langle y, y \rangle > 0$, then it means that the quadratic equation (6) has no root or one real root, hence discriminant must be

$$\langle x, y \rangle^2 - \langle y, y \rangle \langle x, x \rangle \leq 0$$

If $\langle y, y \rangle = 0$, then from axiom *i*) we have that $y = \mathbf{0}$ and it implies that (you can show this by using axioms) $\langle x, \mathbf{0} \rangle = 0$. Thus, CS inequality holds for any vector x and y.
Since we have proved CS inequality for *any* inner-product,

$$|\text{Cov}(Z, W)| \leq \sqrt{\text{Var}(Z)\text{Var}(W)} \tag{7}$$

follows immediately by simply taking $Z = X - E(X), W = Y - E(Y)$ for the inner-product given by $\langle Z, W \rangle = E(ZW)$.

12. From the definition of non-negative matrix, we need to show that $a'\Gamma_1 a \geq 0$ for any vector $a = (a_1, a_2) \in \mathbb{R}^2$. Note that

$$(a_1, a_2)\Gamma_1 \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = (a_1, a_2) \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = (2a_1 + a_2, a_1 + 2a_2) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$
$$= 2(a_1^2 + a_1 a_2 + a_2^2) = 2\left\{ \left(a_1 + \frac{a_2}{2} \right)^2 + \frac{3a_2^2}{4} \right\} \geq 0.$$

Therefore, Γ_1 is a non-negative definite matrix. Similarly for Γ_2, observe that

$$(a_1, a_2)\Gamma_2 \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = (a_1, a_2) \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

$$= (a_1 - 2a_2, -2a_1 + 4a_2) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 2(a_1^2 + a_1 a_2 + a_2^2)$$
$$= (a_1 - 2a_2)^2 \geq 0.$$

Therefore, Γ_2 is a non-negative definite matrix.

13. Problem 1.4

(a) $X_t = a + bZ_t + cZ_{t-2}$ is a weakly stationary process since

$$E(X_t) = E(a + bZ_t + cZ_{t-2}) = a + bE(Z_t) + cE(Z_{t-2}) = a$$

$$\text{Cov}(X_{t+h}, X_t) = \text{Cov}(a + bZ_{t+h} + cZ_{t+h-2}, a + bZ_t + cZ_{t-2})$$
$$= b^2\text{Cov}(Z_{t+h}, Z_t) + bc\text{Cov}(Z_{t+h}, Z_{t-2}) + c^2\text{Cov}(Z_{t+h-2}, Z_{t-2})$$
$$= \begin{cases} bc\sigma^2 & h = -2 \text{ or } h = 2, \\ (b^2 + c^2)\sigma^2 & h = 0, \\ 0, & \text{other wise} \end{cases}$$

do not depend on t.

(b) $X_t = Z_1 \cos(ct) + Z_2 \sin(ct)$ is a weakly stationary process since

$$E(X_t) = E(Z_1 \cos(ct) + Z_2 \sin(ct)) = E(Z_1)\cos(ct) + E(Z_2)\sin(ct) = 0$$

$$\text{Cov}(X_{t+h}, X_t) = \text{Cov}(Z_1 \cos(c(t+h)) + Z_2 \sin(c(t+h)), Z_1 \cos(ct) + Z_2 \sin(ct))$$
$$= \cos(c(t+h))\cos(ct)\text{Cov}(Z_1, Z_1) + \sin(c(t+h))\sin(ct)\text{Cov}(Z_2, Z_2)$$
$$= \cos(c(t+h) - ct)\sigma^2 = \cos(ch)\sigma^2$$

do not depend on t

(c) $X_t = Z_t \cos(ct) + Z_{t-1} \sin(ct)$ is not a stationary process since

$$E(X_t) = E(Z_t \cos(ct) + Z_{t-1} \sin(ct)) = E(Z_t)\cos(ct) + E(Z_{t-1})\sin(ct) = 0,$$

however,

$$\text{Cov}(X_{t+h}, X_t) = \text{Cov}(Z_{t+h} \cos(c(t+h)) + Z_{t+h-1} \sin(c(t+h)), Z_t \cos(ct) + Z_{t-1} \sin(ct))$$

$$= \cos(c(t+h))\cos(ct)\text{Cov}(Z_{t+h}, Z_t) + \cos(c(t+h))\sin(ct)\text{Cov}(Z_{t+h}, Z_{t-1})$$
$$+ \sin(c(t+h))\cos(ct)\text{Cov}(Z_{t+h-1}, Z_t) + \sin(c(t+h))\sin(ct)\text{Cov}(Z_{t+h-1}, Z_{t-1})$$
$$= \begin{cases} \sigma^2, & h = 0 \\ \cos(c(t-1))\sin(ct)\sigma^2, & h = -1 \\ \sin(c(t+1))\cos(ct)\sigma^2, & h = 1 \\ 0, & \text{o.w} \end{cases}$$

depends on t.

(d) $X_t = a + bZ_0$ is a weakly stationary process because

$$E(X_t) = E(a + bZ_0) = a + bE(Z_0) = a$$

$$\mathrm{Cov}(X_{t+h}, X_t) = \mathrm{Cov}(a + bZ_0, a + bZ_0) = b^2\mathrm{Cov}(Z_0, Z_0) = b^2\sigma^2$$

do not depend on t.

(e) $X_t = Z_0 \cos(ct)$ is not a weakly stationary process since

$$E(X_t) = E(Z_0 \cos(ct)) = E(Z_0)\cos(ct) = 0,$$

however,

$$\mathrm{Cov}(X_{t+h}, X_t) = \mathrm{Cov}(Z_0 \cos(c(t+h)), Z_0 \cos(ct)) = \cos(c(t+h))\cos(ct)\sigma^2$$

depends on t.

(f) $X_t = Z_t Z_{t-1}$ is a stationary process because

$$E(X_t) = E(Z_t Z_{t-1}) = 0$$

$$\mathrm{Cov}(X_{t+h}, X_t) = \mathrm{Cov}(Z_{t+h}Z_{t+h-1}, Z_t Z_{t-1}) = E(Z_{t+h}Z_{t+h-1}Z_t Z_{t-1}) = \begin{cases} \sigma^4, & h = 0 \\ 0, & \text{o.w} \end{cases}$$

14. Problem 1.8

We will show that $i)E(X_t) = 0$, $ii)\mathrm{Var}(X_t) = 1$ and $iii)\mathrm{Cov}(X_{t+h}, X_t) = 0$ for $h \neq 0$.

$i)$ If t is even, then $E(X_t) = E(Z_t) = 0$. When t is odd,

$$E(X_t) = E\left(\frac{Z_{t-1}^2 - 1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}(E(Z_{t-1}^2) - 1) = 0.$$

Thus, $E(X_t) = 0$ for all t.

$ii)$ When t is even, $\mathrm{Var}(X_t) = E(Z_t^2) = 1$. If t is odd, then

$$\mathrm{Var}(X_t) = E\left(\frac{Z_{t-1}^2 - 1}{\sqrt{2}}\right)^2 = \frac{1}{2}E(Z_{t-1}^4 - 2Z_{t-1}^2 + 1) = \frac{1}{2}(E(Z_{t-1}^4) - 2E(Z_{t-1}^2) + 1)$$

$$= \frac{1}{2}(3 - 2 + 1) = 1$$

(Note that for $Z \sim \mathcal{N}(0,1)$, $E(Z^4) = 3$ and $E(Z^3) = 0$.)

$iii)$ For $h \neq 0$, observe that

$$\mathrm{Cov}(X_{t+h}, X_t) = \begin{cases} E(Z_{t+h}Z_t), & \text{if } t+h \text{ even, } t \text{ even} \\ E(Z_{t+h}(Z_{t-1}^2 - 1)/\sqrt{2}), & \text{if } t+h \text{ even, } t \text{ odd} \\ E((Z_{t+h-1}^2 - 1)/\sqrt{2}Z_t), & \text{if } t+h \text{ odd, } t \text{ even} \\ E((Z_{t+h-1}^2 - 1)/\sqrt{2}(Z_{t-1}^2 - 1)/\sqrt{2}), & \text{if } t+h \text{ odd, } t \text{ odd.} \end{cases}$$

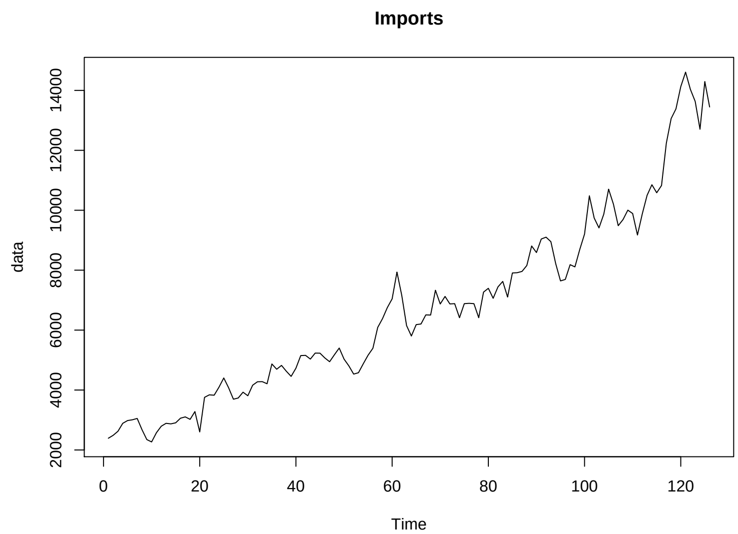
In all cases considered, $\mathrm{Cov}(X_{t+h}, X_t) = 0$ since $h \neq 0$. For example,

$$E(Z_{t+h}(Z_{t-1}^2 - 1)/\sqrt{2}) = \begin{cases} E((Z_{t-1}^3 - Z_{t-1})/\sqrt{2}) = 0 & h = 1 \\ \frac{1}{\sqrt{2}}E(Z_{t+1})E(Z_{t-1}^2) - 1) = 0 & h = \pm 3, \pm 5, \ldots \end{cases}$$

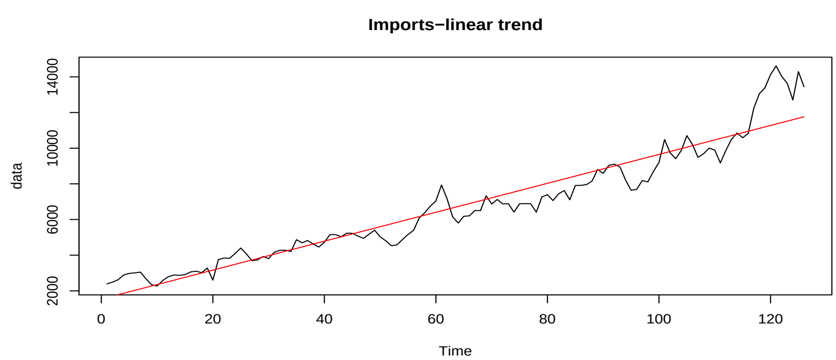
Therefore, $\{X_t\}$ is WN(0,1) but not i.i.d. because for odd t, $X_t = (Z_{t-1}^2 - 1)/\sqrt{2}$ and $X_{t-1} = Z_{t-1}$ (($t-1$) is even) are clearly dependent.

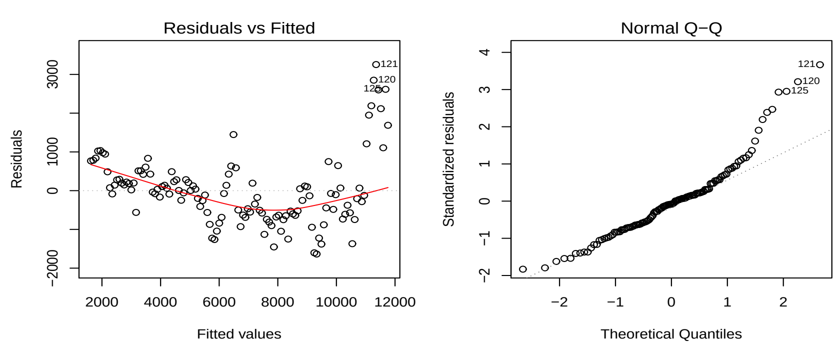
15. The following data analysis is more open-ended. Hence, if you have properly applied indicated methods and discussed your opinion about the results, you would have full credit.

    (a) We can find some increasing trend and couple of outliers nearby t=60,120.

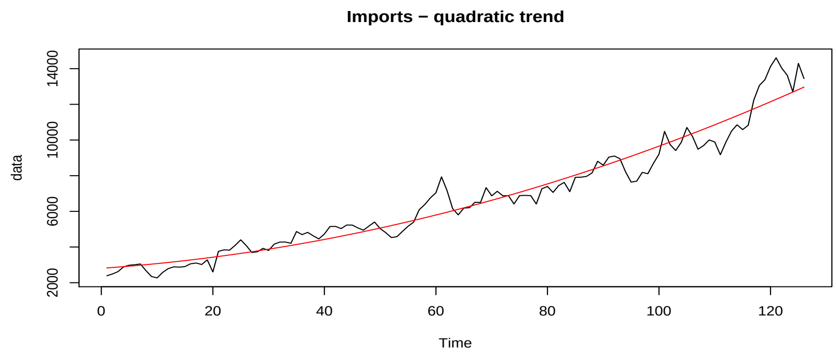


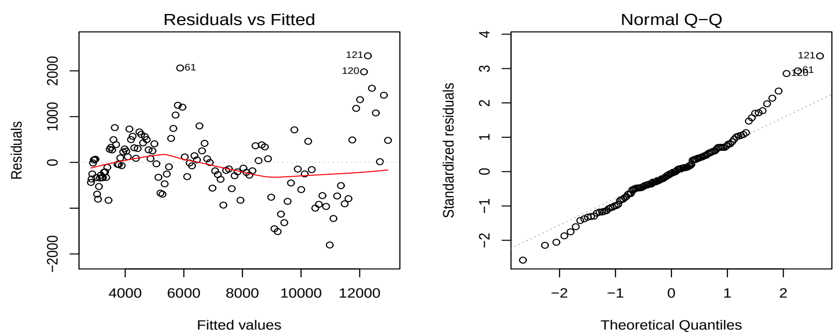
    (b) Fitting a linear trend model seems not fully explain the data. The plot in the bottom left shows residual plot, and we can see some quadratic trend left.



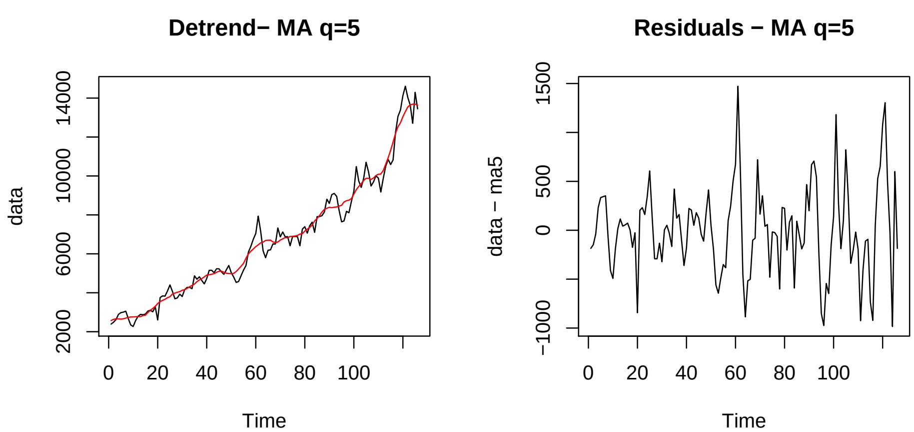


Therefore, tried quadratic trend model and it looks more reasonable than the linear model.

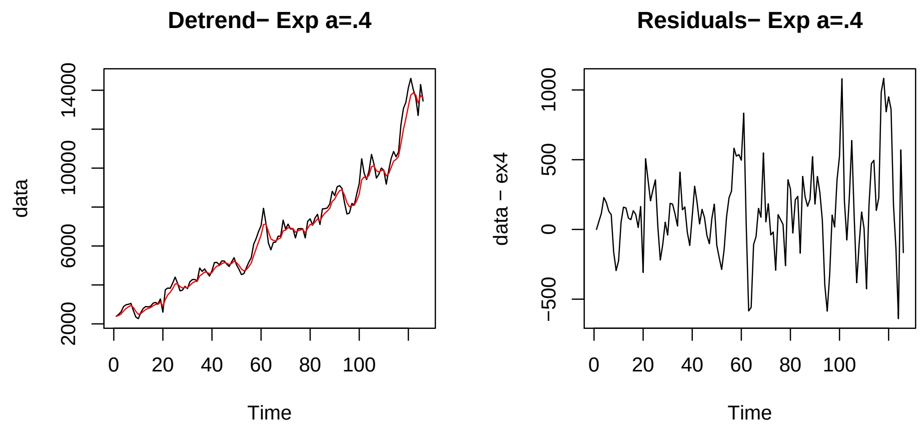




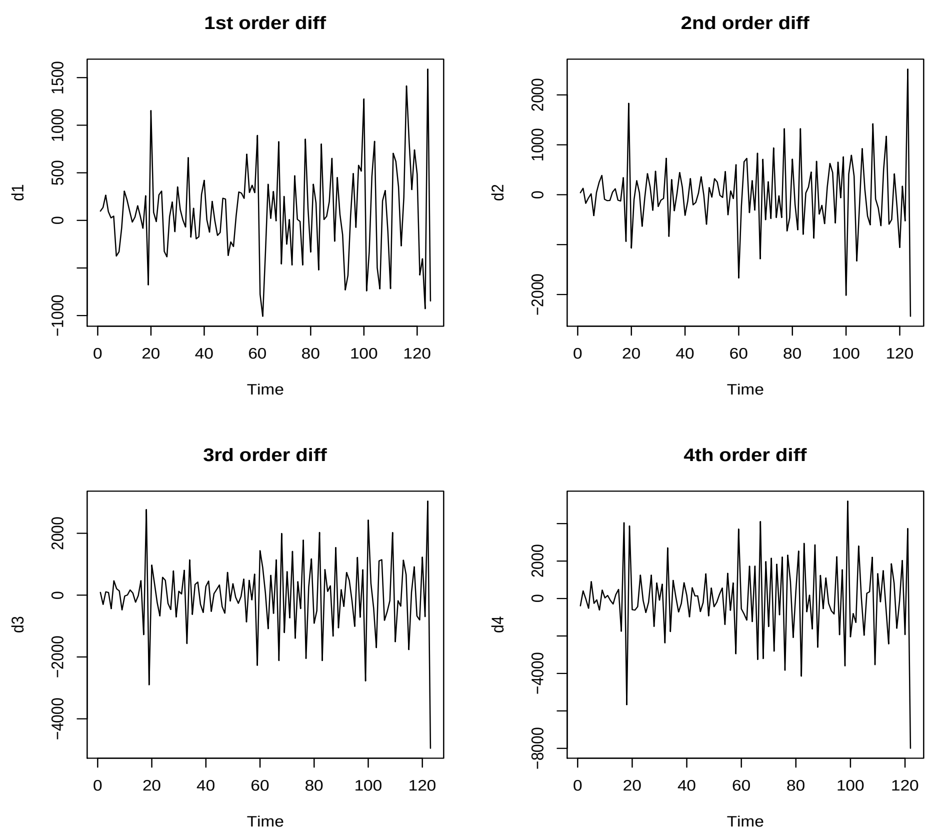
(c) If you apply 5 point moving average, the estimated trend and residuals are following. No clear trend left.



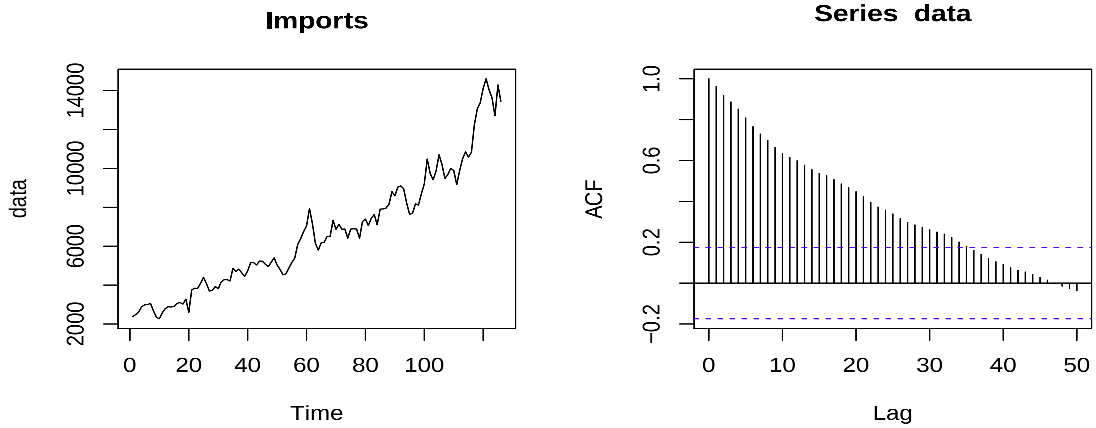
(d) By checking the graph, we can find out exponential smoothing with $a = .4$ undersmooth the trend compared to 5 point moving average.



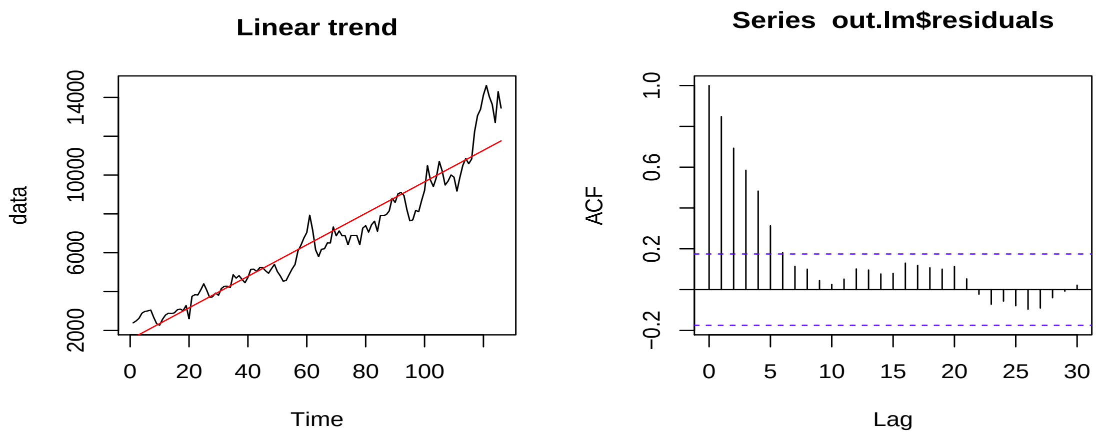
(e) Results are given in the below figure when you applied 1st to 4th order differencing. From this figure, it seems that 1st order differencing is enough to remove trend because higher oder differencing does not show any improvement.



(f) From the correlogram in the below we observe very slowly decaying and almost linearly decaying SACFs. Together with time plot, it may due to increasing trend.



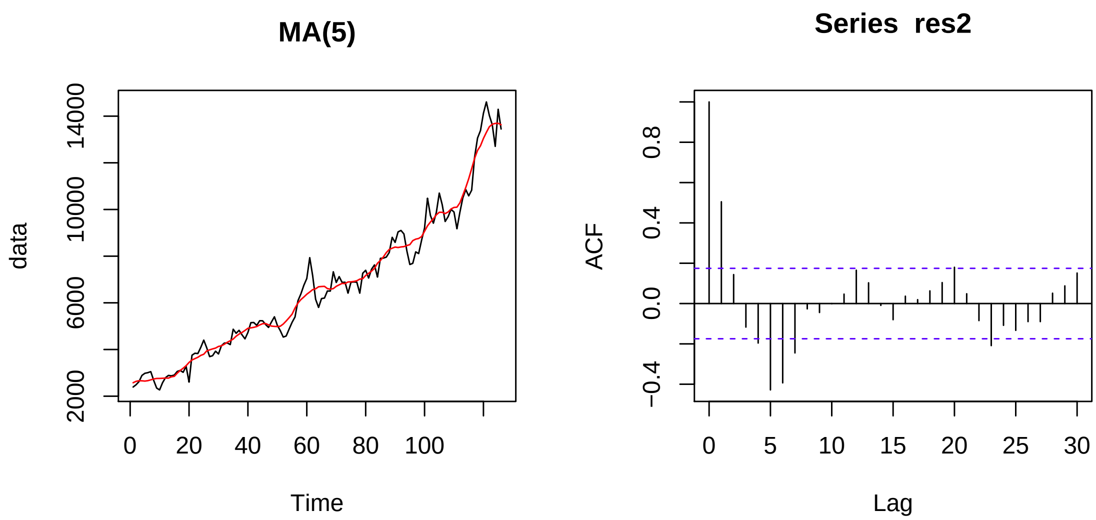
(g) Once you fit linear trend and observe correlogram, you can see that such slowly decaying SACFs in (a) is disappeared. It is positively correlated on smaller lags, but you can see that it looks stationary process.



Formal testing for independence are given in the below. All three tests reject the null hypotheses of independence, so we conclude that the errors are not i.i.d.

```
Null hypothesis: Residuals are iid noise.
Test                        Distribution Statistic   p-value
Ljung-Box Q                Q ~ chisq(20)    266.68         0 *
McLeod-Li Q                Q ~ chisq(20)    185.23         0 *
Turning points T   (T-82.7)/4.7 ~ N(0,1)        61         0 *
```

(h) If we remove trend by moving average and draw correlogram, you can also observe that very slowly decaying SACFs are diminished. Formal tests also reject the null hypotheses of i.i.d errors, hence we see that errors are (weakly) correlated.




```
Null hypothesis: Residuals are iid noise.
Test                        Distribution Statistic   p-value
Ljung-Box Q                Q ~ chisq(20)    110.27         0 *
McLeod-Li Q                Q ~ chisq(20)      38.6    0.0075 *
Turning points T   (T-82.7)/4.7 ~ N(0,1)        60         0 *
```

Program code.

```
setwd("C:\\ designate your own working directory")
library(itsmr)
data = scan("imports.txt")

plot.ts(data);
title("Imports")

# Can observe linear trend
# Method 1 Simple linear regression
n = length(data);
x = seq(from=1, to = n, by=1);
out.lm = lm(data ~ 1 + x);
summary(out.lm)


layout(matrix(c(1,1,2,3), 2, 2, byrow = TRUE))
plot.ts(data);
title("Imports-linear trend")
lines(out.lm$fitted.values, col="red")
# residual diagnostics
plot(out.lm, which=c(1,2))
```

```
##quadratic trend
n = length(data);
x = seq(from=1, to = n, by=1);
x1 = x^2
out.lm2 = lm(data ~ 1 + x + x1);
summary(out.lm2)


layout(matrix(c(1,1,2,3), 2, 2, byrow = TRUE))
plot.ts(data);
title("Imports - quadratic trend")
lines(out.lm2$fitted.values, col="red")
# residual diagnostics
plot(out.lm2, which=c(1,2))


# Detrend by smoothing
ma5 = smooth.ma(data, 5);

par(mfrow=c(1,2))
plot.ts(data);
lines(ma5, col="red")
title("Detrend- MA q=5")
plot.ts(data-ma5);
title("Residuals - MA q=5")


# Apply exponential smoothing
ex4 = smooth.exp(data, .4)

par(mfrow=c(1,2))
plot.ts(data);
lines(ex4, col="red")
title("Detrend- Exp a=.4")
plot.ts(data-ex4);
title("Residuals- Exp a=.4")

## Apply differncing
par(mfrow=c(2,2))
d1=diff(data)
plot.ts(d1); title("1st order diff");
d2=diff(d1)
plot.ts(d2); title("2nd order diff");
d3=diff(d2)
plot.ts(d3); title("3rd order diff");
d4=diff(d3)
```

```
plot.ts(d4); title("4th order diff");


par(mfrow=c(1,2));
plot.ts(data);
title("Imports")
acf(data, 30);

# Method 1 Simple linear regression
n = length(data);
x = seq(from=1, to = n, by=1);
out.lm = lm(data ~ 1 + x);
summary(out.lm)


plot.ts(data);
title("Linear trend")
lines(out.lm$fitted, col="red")
acf(out.lm$residuals, 30)

test(out.lm$residuals)

# Method 2 MA(5)
# Detrend by smoothing
ma5 = smooth.ma(data, 5);
res2 = data - ma5;

plot.ts(data);
title("MA(5)")
lines(ma5, col="red")
acf(res2, 30)

test(res2)
```