1.4 Let $\{Z_t\}$ be a sequence of independent normal random variables, each with mean 0 and variance σ^2, and let a, b, and c be constants. Which, if any, of the following processes are stationary? For each *stationary* process specify the mean and autocovariance function.

a. $X_t = a + bZ_t + cZ_{t-2}$

b. $X_t = Z_1 \cos(ct) + Z_2 \sin(ct)$

c. $X_t = Z_t \cos(ct) + Z_{t-1} \sin(ct)$

d. $X_t = a + bZ_0$

e. $X_t = Z_0 \cos(ct)$

f. $X_t = Z_t Z_{t-1}$

$Z_t \sim N(0, \sigma^2)$

a. $X_t = a + bZ_t + CZ_{t-2}$

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

$Var(X_t) = b^2\sigma^2 + c^2\sigma^2$

$$\begin{cases} h=0: & \gamma(h) = Var(X_t) = (b^2+c^2)\sigma^2 \\ h=-2: & \gamma(h) = Cov(CZ_{t-2}, bZ_{t-2}) = bc\sigma^2 \\ h=2: & \gamma(h) = Cov(bZ_t, CZ_t) = bc\sigma^2 \\ \text{Otherwise}: & \gamma(h) = 0, \end{cases}$$

$\therefore$ It depends on h, stationary

b. $X_t = Z_1 \cos(ct) + Z_2 \sin(ct)$

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

$\gamma_X(t, t+h) = Cov(X_t, X_{t+h})$

$= Cov(Z_1\cos(ct) + Z_2\sin(ct), Z_1\cos(c(t+h)) + Z_2\sin(c(t+h)))$

$= \cos(ct)\cos(c(t+h))\sigma^2 + \sin(ct)\sin(c(t+h))\sigma^2$

$= \sigma^2[\cos(ct)\cos(c(t+h)) + \sin(ct)\sin(c(t+h))]$

$\because \cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta$

$\therefore \underbrace{\gamma_X(t,t+h)}_{\downarrow} = \sigma^2(\cos(ch))$

do not depend on t, $\{X_t: t\in \mathbb{Z}\}$ is weakly Stationary.

C. $X_t = Z_t\cos(ct) + Z_{t-1}\sin(ct)$

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

$\gamma_X(t,t+h) = Cov(X_t, X_{t+h})$

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

$$\begin{cases} h=0, & \gamma_X(t,t) = \sigma^2\cos^2(ct) + \sigma^2\sin^2(ct) = \sigma^2 \\ h=1, & \gamma_X(t,t+1) = \sigma^2\cos(ct)\sin(c(t+1)) \\ h=-1, & \gamma_X(t,t-1) = \sigma^2\sin(ct)\cos(c(t-1)) \end{cases}$$

$\therefore$ $c=\pm k\pi$, $k\in\mathbb{Z}$, does not depend on t, depends on h, $\rightarrow$ weakly Stationary

$c\neq\pm k\pi$, $k\in\mathbb{Z}$, depends on h,

$\therefore$ not Stationary

d. $X_t = a + bZ_0$

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

$\gamma_X(t, t+h) = Cov(a+bZ_0, a+bZ_0) = b^2\sigma^2$

Since X_t does not depend on t, it is a weakly stationary process

e. $X_t = Z_0 \cos(ct)$

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

$\gamma_X(t, t+h) = Cov(Z_0\cos(ct), Z_0\cos(c(t+h)))$

$= \sigma^2 \cos(ct)\cos(c(t+h))$

$\therefore$ if $c = \pm k\pi$, $k \in \mathbb{Z}$, X_t does not depend on t,

it is a stationary process.

if $c \neq \pm k\pi$, not stationary

f. $X_t = Z_t Z_{t-1}$ due to the independence

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

$\gamma_X(t, t+h) = Cov(Z_t Z_{t-1}, Z_{t+h} Z_{t+h-1})$

$\therefore h = 0$, $\gamma_X(t, t+h) = \sigma^4$

$h \neq 0$, $\gamma_X(t, t+h) = 0$.

$\therefore$ It is a weakly stationary process because it depends on h.

1.5 Let $\{X_t\}$ be the moving-average process of order 2 given by

$$X_t = Z_t + \theta Z_{t-2},$$

where $\{Z_t\}$ is WN(0, 1).

a. Find the autocovariance and autocorrelation functions for this process when $\theta = 0.8$.

b. Compute the variance of the sample mean $(X_1+X_2+X_3+X_4)/4$ when $\theta = 0.8$.

c. Repeat (b) when $\theta = -0.8$ and compare your answer with the result obtained in (b).

a. $X_t = Z_t + 0.8Z_{t-2}$, $E(X_t) = E(Z_t) + 0.8E(Z_{t-2}) = 0$

$\gamma_X(t,t+h) = Cov(Z_t + 0.8Z_{t-2}, Z_{t+h} + 0.8Z_{t-2+h})$

$h=0$, $\gamma_X(t,t) = Cov(Z_t + 0.8Z_{t-2}, Z_t + 0.8Z_{t-2}) = 1.64$

$h=2$, $\gamma_X(t,t+2) = Cov(Z_t + 0.8Z_{t-2}, Z_{t+2} + 0.8Z_t) = 0.8$

$h=-2$, $\gamma_X(t,t-2) = 0.8$

$|h|>2$, $\gamma_X(t,t+h) = 0$

$\therefore \rho(h) = \begin{cases} 1, & h=0 \\ \frac{0.8}{1.64}, & |h|=2 \\ 0, & |h|>2 \end{cases}$

b. $Var[(X_1+X_2+X_3+X_4)/4] = \frac{1}{16}Var(X_1+X_2+X_3+X_4)$

$= \frac{1}{16}[Var(X_1)+Var(X_2)+Var(X_3)+Var(X_4)+2Cov(X_1,X_3)+2Cov(X_2,X_4)]$

$= \frac{1}{16}(4\gamma_X(0)+4\gamma_X(2)) = \frac{\gamma_X(0)+\gamma_X(2)}{4} = \frac{1.64+0.8}{4} = 0.61$

c. $\theta = -.8$, $\gamma_X(0) = 1.64$ $\gamma_X(2) = -0.8$

$Var(\frac{1}{4}(\ldots)) = \frac{1.64-0.8}{4} = 0.21$

1.8 Let $\{Z_t\}$ be IID N(0, 1) noise and define

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

a. Show that $\{X_t\}$ is WN(0, 1) but not iid(0, 1) noise.

b. Find $E(X_{n+1}|X_1, \ldots, X_n)$ for n odd and n even and compare the results.

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

t even. $E(Z_t) = 0$ $\quad$ t odd: $E\left[(Z_{t-1}^2 - 1)/\sqrt{2}\right] = \frac{1}{\sqrt{2}} E(Z_{t-1}^2 - 1) = 0$

t even. $\gamma_X(t,t) = E(Z_t^2) = 1$

t odd. $\gamma_X(t,t) = E\left(\frac{(Z_{t-1}^2 - 1)^2}{2}\right) = \frac{1}{2} E(Z_{t-1}^4 - 2Z_{t-1}^2 + 1) = \frac{1}{2}(3-2+1) = 1$

$h=1$

t even: $\gamma_X(t,t+1) = Cov(Z_t, (Z_t^2 - 1)/\sqrt{2}) = \frac{1}{\sqrt{2}} E(Z_t^3 - Z_t) = 0$

t odd: $\gamma_X(t,t+1) = Cov\left[(Z_{t-1}^2 - 1)/\sqrt{2}, Z_{t+1}\right] = E\left[(Z_{t-1}^2-1)/\sqrt{2}\right] E(Z_{t+1}) = 0$

$|h| \geq 2$, $\gamma_X(t, t+h) = 0$

$\therefore \begin{cases} \gamma(h) = 0 & |h| \geq 1 \\ \gamma(h) = 1, & h = 0 \end{cases}$

$\therefore$ It depends on h,

$\therefore X_t \sim WN(0,1)$

b. if n odd

$E(X_{n+1}|X_1, \ldots, X_n) = E(Z_{n+1}|Z_0, Z_2, \cdots Z_{n-1}) = E(Z_{n+1}) = 0$

if n even

$E(X_{n+1}|X_1, \ldots, X_n) = E(Z_{n+1}|Z_0, Z_2, \ldots Z_n) = E\left[(Z_n^2 - 1)/\sqrt{2}\right]$

$= \frac{X_n^2 - 1}{\sqrt{2}}$

1.11 Consider the simple moving-average filter with weights $a_j = (2q+1)^{-1}$, $-q \le j \le q$.

a. If $m_t = c_0 + c_1 t$, show that $\sum_{j=-q}^{q} a_j m_{t-j} = m_t$.

b. If Z_t, $t = 0, \pm 1, \pm 2, \ldots$, are independent random variables with mean 0 and variance σ^2, show that the moving average $A_t = \sum_{j=-q}^{q} a_j Z_{t-j}$ is "small" for large q in the sense that $EA_t = 0$ and $\mathrm{Var}(A_t) = \sigma^2/(2q+1)$.

a. $\sum_{j=-q}^{q} a_j m_{t-j} = \sum_{j=-q}^{q} (2q+1)^{-1}[C_0 + C_1(t-j)]$

$= \frac{1}{2q+1}\left[C_0(2q+1) + C_1 \sum_{j=-q}^{q}(t-j)\right]$

$= C_0 + \frac{C_1}{2q+1}\sum_{j=-q}^{q}(t-j)$

$= C_0 + \frac{C_1}{2q+1}\left(t(2q+1) - \sum_{j=-q}^{q} j\right)$

$= C_0 + C_1 t - \frac{C_1}{2q+1}\underbrace{\sum_{j=-q}^{q} j}_{=0}$

$= C_0 + C_1 t$

$= m_t$

b. $Z_t \sim N(0, \sigma^2)$

$E(A_t) = E\left(\sum_{j=-q}^{q}(2q+1)^{-1} Z_{t-j}\right)$

$= \sum_{j=-q}^{q}(2q+1)^{-1} E(Z_{t-j}) = 0$

$\mathrm{Var}(A_t) = (2q+1)^{-2}\underline{\sum_{j=-q}^{q}\sigma^2}$

$= (2q+1)^{-2}(2q+1)\sigma^2$

$= \frac{\sigma^2}{2q+1}$

2.21 Let $X_1,\ X_2,\ X_4,\ X_5$ be observations from the MA(1) model

$$X_t = Z_t + \theta Z_{t-1}, \quad \{Z_t\} \sim \mathrm{WN}\left(0, \sigma^2\right).$$

a. Find the best linear estimate of the missing value X_3 in terms of X_1 and X_2.
b. Find the best linear estimate of the missing value X_3 in terms of X_4 and X_5.

a. $\hat{X}_3 = \alpha_1 X_1 + \alpha_2 X_2$

minimize $E(\hat{X}_3 - X_3)^2$, $\Rightarrow$

$$\alpha_1 = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2} \qquad \alpha_2 = \frac{\rho(1) - \rho(1)\rho(2)}{1 - \rho(1)^2}$$

$$\gamma_X(0) = Cov(X_t, X_t) = Cov(Z_t + \theta Z_{t-1}, Z_t + \theta Z_{t-1}) = \sigma^2 + \theta^2\sigma^2$$

$$\gamma_X(1) = Cov(X_t, X_{t+1})$$
$$= Cov(Z_t + \theta Z_{t-1}, Z_{t+1} + \theta Z_t) = \theta\sigma^2$$

$$\gamma_X(-1) = Cov(Z_{t-1} + \theta Z_{t-2}, Z_t + \theta Z_{t-1}) = \theta\sigma^2$$

$\therefore |h| = 1$. $\gamma_X(1) = \theta\sigma^2$

$|h| > 1$, $\gamma_X(t, t+h) = 0$. $\quad \therefore \gamma_X(2) = 0$

$$\therefore \gamma_X(h) = \begin{cases} (1+\theta^2)\sigma^2, & h = 0 \\ \theta\sigma^2, & |h| = 1 \\ 0, & |h| > 1 \end{cases}$$

$$\rho(h) = \begin{cases} 1, & h = 0 \\ \frac{\theta}{1+\theta^2}, & |h| = 1 \\ 0, & |h| > 1 \end{cases}$$

$\therefore$ For MA(1) model.

$$\alpha_1 = \frac{-\left(\frac{\theta}{1+\theta^2}\right)^2}{1-\left(\frac{\theta}{1+\theta^2}\right)^2} = -\frac{\theta^2}{1+\theta^2+\theta^4}$$

$$\alpha_2 = \frac{\frac{\theta}{1+\theta^2}}{1-\left(\frac{\theta}{1+\theta^2}\right)^2} = \frac{\theta+\theta^3}{1+\theta^2+\theta^4}$$

b. $\hat{X}_3 = \beta_1 X_5 + \beta_2 X_4$

$\beta_1 = \alpha_1 \qquad \beta_2 = \alpha_2$

2. Let X and Y are independent random variables with means μ_x, μ_y and variances σ_x^2 and σ_y^2 respectively. Find an expression for the correlation of XY and Y in terms of these means and variances.

$$Corr(XY, Y) = \frac{Cov(XY, Y)}{\sqrt{Var(XY)}\sqrt{Var(Y)}}$$

$Var(Y) = E(Y^2) - (E(Y))^2$

$$Cov(XY, Y) = E(XY^2) - E(XY)E(Y)$$
$$= E(X)E(Y^2) - E(X)(E(Y))^2$$
$$= \mu_x(\sigma_y^2 + \mu_y^2) - \mu_x \mu_y^2 = \mu_x \sigma_y^2$$

$$Var(XY) = E(X^2Y^2) - (E(XY))^2 = E(X^2)E(Y^2) - [E(X)E(Y)]^2$$
$$= (\sigma_x^2 + \mu_x^2)(\sigma_y^2 + \mu_y^2) - \mu_x^2\mu_y^2$$
$$= \sigma_x^2\sigma_y^2 + \sigma_x^2\mu_y^2 + \mu_x^2\sigma_y^2$$

$$\therefore Corr(XY, Y) = \frac{\mu_x \sigma_y}{\sigma_y\sqrt{\sigma_x^2\sigma_y^2 + \sigma_x^2\mu_y^2 + \sigma_y^2\mu_x^2}}$$

3. Let X_1, X_2 and X_3 and uncorrelated random variables, each with mean μ and variance σ^2. Find, in terms of μ and σ^2, $Cov(X_1 + X_2, X_2 + X_3)$ and $Cov(X_1 + X_2, X_1 - X_2)$.

$Cov(X_1+X_2,\ X_2+X_3) = Cov(X_2, X_2) = \sigma^2$

$Cov(X_1+X_2,\ X_1-X_2) = Cov(X_1, X_1) + Cov(X_2, -X_2)$

$= \sigma^2 - \sigma^2 = 0$

4. Suppose X and Z are independent. $Var(X) = 1$ and $Var(Z) = 0.01$. $Y = X + Z$. Compute the correlation between X and Y.

$Var(Y) = Var(X+Z) = 1.01$

$Cov(X, Y) = Cov(X, X+Z) = Var(X) = 1$

$Corr(X, Y) = \dfrac{Cov(X, Y)}{\sqrt{Var(X)\, Var(Y)}} = \dfrac{1}{\sqrt{1.01}} = 0.995$

5. If $Var(X) = 1, Var(Y) = 2$ and $Cov(X, Y) = -1$. If $U = 3X - 2Y$ and $V = X + 2Y$, find $Var(U), Var(V)$ and $Cov(U, V)$.

$\because Var(cA+dB) = c^2 Var(A) + d^2 Var(B) + 2cd\, Cov(A, B)$

$\therefore Var(U) = Var(3X-2Y) = 9+8-12\cdot(-1) = 29$

$Var(V) = Var(X+2Y) = 1+8-4 = 5$

$Cov(U, V) = Cov(3X-2Y,\ X+2Y)$

$= 3Cov(X, X) + 6Cov(X, Y) - 2Cov(X, Y)$

$- 4Cov(Y, Y)$

$= 3-6+2-8 = -9$

6. The residents of Strange Town withdraw money from a cash machine according to the following probability function (X):

x	50	100	200
$p(x)$	0.3	0.5	0.2

$Var(N) = E(N) = \theta$

The number of customers per day has the distribution $N \sim Poisson(\theta)$. Let $T_N = X_1 + X_2 + ... + X_N$ be the total amount of money withdrawn in a day, where each X_i has the probability mass function above, and $X_1, X_2, \ldots$ are independent of each other and of N.

a. Find the Expectation and Variance of T_N.

b. Suppose $\theta = 50$. Verify your answers in part (a) using R by generating 10000 samples of T_N and computing the sample mean and sample variance of those values.

a. $E(X_i) = 50 \times 0.3 + 100 \times 0.5 + 200 \times 0.2$

$= 15 + 50 + 40 = 105$

$Var(X_i) = (105-50)^2 \times 0.3 + (105-100)^2 \times 0.5$

$+ (105-200)^2 \times 0.2 = 2725$

$E(T_N|N) = E(X_1) + \cdots + E(X_N) = 105N$

$Var(T_N|N) = Var(X_1) + \cdots + Var(X_N) = 2725N$

$E(T_N) = E(E(T_N|N)) = E(105N)$

$= 105E(N) = 105\theta$

$Var(T_N) = Var(E(T_N|N)) + E(Var(T_N|N))$

$= Var(105N) + E(2725N)$

$= 105^2 Var(N) + 2725 E(N)$

$= 13750\theta$

b. R-simulation.

```{r}
#Define a function to generate single value of T_N given N
tn.func <- function(n){
sum(sample(c(50, 100, 200), n, replace=T, prob=c(0.3, 0.5, 0.2)))
}
# Generate 10,000 random values of N, using theta=50:
N <- rpois(10000, 50)
# Generate 10,000 random values of T_N, conditional on N:
TN <- sapply(N, tn.func)
# Find the sample mean of T_N values, which should be close
# to 105 * 50 = 5250:
mean(TN)
# Find the sample variance of T_N values, which should be close
# to 13750 * 50 = 687500:
var(TN)
#Note that the sample variance is often some distance from the true variance, even when the sample size is 10,000.
# It shows the difficulty with the estimation of variance.
```

```
[1] 5239.25
[1] 693180.3
```