Homework 5

1. Problem 4.5 of Prof. Kim's note.
   Solution:
   Given the joint pdf

$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } |x|+|y| \le 1/\sqrt{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

   (a) From the joint pdf we have

$$\int_{\frac{1}{\sqrt{2}}}^{-\frac{1}{\sqrt{2}}} \int_{-\frac{1}{\sqrt{2}}+|y|}^{\frac{1}{\sqrt{2}}-|y|} c \text{ d}x \text{ d}y = \int_{-\frac{1}{\sqrt{2}}}^{-\frac{1}{\sqrt{2}}} 2c(\frac{1}{\sqrt{2}} - |y|) \text{ d}y = 4c(\frac{1}{2} - \frac{1}{4}) = c. \tag{2}$$

   We know that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$, therefore $c = 1$.

   (b) We know that $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)$. For $|x| \le \frac{1}{\sqrt{2}}$ we have

$$f_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) \text{ d}y = \int_{-\frac{1}{\sqrt{2}}+|x|}^{\frac{1}{\sqrt{2}}-|x|} 1 \text{ d}y = 2(\frac{1}{\sqrt{2}} - |x|) = \sqrt{2} - 2|x|. \tag{3}$$

   Similarly $f_X(x) = 0$ for $|x| > \frac{1}{\sqrt{2}}$.
   Similarly $f_Y(y) = \sqrt{2} - 2|y|$, for $|y| \le \frac{1}{\sqrt{2}}$ and 0 otherwise.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \begin{cases} \frac{1}{\sqrt{2}-2|y|} & \text{if } |x| \le \frac{1}{\sqrt{2}} - |y|, |y| \le \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

   (c) $f_{X,Y}(0,0) = 1$ whereas $f_X(0)f_Y(0) = \sqrt{2}^2 = 2$. Therefore, X and Y are not independent.

   (d) Let us find the CDF of $Z = |X| + |Y|$. Consider $0 \le z \le \frac{1}{\sqrt{2}}$

$$\begin{aligned} P(Z \le z) = P(|X|+|Y| \le z) &= \int_{\{(x,y):|x|+|y|\le z\}} 1 \text{ d}x \text{ d}y = \int_{-z}^{z}\int_{-z+|y|}^{z-|y|} \text{d}x \text{ d}y \\ &= \int_{-z}^{z} 2(z-|y|) \text{ d}y = 4z^2 - 2z^2 = 2z^2. \end{aligned}$$

   Therefore, pdf of Z is given by

$$f_Z(z) = \begin{cases} 4z & \text{if } 0 \le z \le \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

2. Problem 4.11 of Prof. Kim's note.
   Solution:

   (a) Consider $x = 2$, then the couple will have one child (first child being a girl) or two children (first a boy and then a girl or boy). Therefore $p_{Y|X}(1 \mid 2) = \frac{1}{2}$ and $p_{Y|X}(2 \mid 2) = 2\frac{1}{4} = \frac{1}{2}$.
   Similarly, for $x = 3$, we have $p_{Y|X}(1 \mid 3) = \frac{1}{2}$, $p_{Y|X}(2 \mid 3) = \frac{1}{4}$, and $p_{Y|X}(3 \mid 3) = \frac{1}{4}$.
   Finally for $x = 4$, we have $p_{Y|X}(1 \mid 4) = \frac{1}{2}$, $p_{Y|X}(2 \mid 4) = \frac{1}{4}$, $p_{Y|X}(3 \mid 4) = \frac{1}{8}$, and $p_{Y|X}(4 \mid 4) = \frac{1}{8}$.

x \ y	2	3	4
1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
2	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
3	0	$\frac{1}{4}$	$\frac{1}{8}$
4	0	0	$\frac{1}{8}$

Table 1: Table for $p_{Y|X}(y \mid x)$

(b) We know that $p_Y(y) = \sum_{k=2}^{4} p_X(k) p_{Y|X}(y \mid k) = \sum_{k=2}^{4} \frac{p_{Y|X}(y|k)}{3}$. Therefore we get

$$p_Y(y) = \begin{cases} \frac{1}{2}, & \text{if } y = 1, \\ \frac{1}{3}, & \text{if } y = 2, \\ \frac{1}{8}, & \text{if } y = 3, \\ \frac{1}{24}, & \text{if } y = 4. \end{cases} \tag{6}$$

3. *Fair vs Unfair Coin*:

(a) Consider a coin with the probability of head $p = 0.3$. Provide an estimate for the probability that after 1000 flipping of this coin we observe between 250 and 350 heads.

(b) Repeat the above part for a fair coin with the probability of head $p = 0.5$, i.e., provide an estimate for the probability that after 1000 flipping of this coin we observe between 250 and 350 heads.

(c) If a third party mixes the two coins, how can you distinguish between the two coins?

Solution:

i. Let $X_k = 1$ if the coin comes head and $X_k = 0$ if the coin comes tail. Then, $\mathbb{E}[X_k] = 0.3$ and $\mathbf{Var}(X_k) = 0.21$. Then,

$$\begin{aligned} \mathbf{P}(250 \le X_1 + \cdots + X_{1000} \le 350) &= \mathbf{P}(-50 \le S_{1000} - 1000\mu \le 50) \\ &= \mathbf{P}(-\frac{50}{10\sqrt{10}} \le \frac{S_{1000} - 1000\mu}{10\sqrt{10}} \le \frac{50}{10\sqrt{10}}) \\ &= \mathbf{P}(-\frac{\sqrt{10}}{2} \le \frac{S_{1000} - 1000\mu}{10\sqrt{10}} \le \frac{\sqrt{10}}{2}) \\ &\approx \mathbf{P}(-1.5811 \le X \le 1.5811) \end{aligned} \tag{7}$$

for a random variable $X \sim \mathcal{N}(0, 0.21)$. Using the command `normcdf` in MATLAB, $\mathbf{P}(-1.5811 \le X \le 1.5811) \approx 0.9994$.

ii. Let $X_k = 1$ if the coin comes head and $X_k = 0$ if the coin comes tail. Then, $\mathbb{E}[X_k] = 0.5$ and $\mathbf{Var}(X_k) = 0.25$. Then,

$$\begin{aligned} \mathbf{P}(250 \le X_1 + \cdots + X_{1000} \le 350) &= \mathbf{P}(-250 \le S_{1000} - 1000\mu \le -150) \\ &= \mathbf{P}(-\frac{250}{10\sqrt{10}} \le \frac{S_{1000} - 1000\mu}{10\sqrt{10}} \le -\frac{150}{10\sqrt{10}}) \\ &= \mathbf{P}(-\frac{5\sqrt{10}}{2} \le \frac{S_{1000} - 1000\mu}{10\sqrt{10}} \le -\frac{3\sqrt{10}}{2}) \\ &\approx \mathbf{P}(-7.9056 \le X \le -4.7434) \end{aligned} \tag{8}$$

for a random variable $X \sim \mathcal{N}(0, 0.25)$. Using the command `normcdf` in MATLAB, $\mathbf{P}(-7.9056 \le X \le -4.7434) \approx 1.19 \times 10^{-21}$.

iii. We randomly flip the coin. If the number of heads is between 250 and 350, it is the first coin ($p = 0.4$). Otherwise, it is the second coin ($p = 0.5$).

4. Consider the simple pandemics model $X_{k+1} = W_k X_k$ where $\{W_k\}$ is an i.i.d. sequence that is uniformly distributed over $[0.4, 3]$.

   (a) Find the CDF and PDF of $\log W_k$.

   (b) Find $\mathbb{E}[\log W_k]$. What can you say about $\lim_{k\to\infty} X_k$?

   (c) Let $X_1 = 1$. Generate 100 sample paths $\{X_k\}$ for $1 \le k \le 200$ and plot them using the regular linear plot (command:`plot`) and log-plot (command:`semilogy`) and store all the 100 sample paths for a later use.

   (d) Theoretically explain your observation.

   (e) For a level $\alpha \ge 1$, define the random variable $T_\alpha = \min\{k \ge 1 \mid X_k \ge \alpha\}$. Such random variables are called *stopping times*. In this case, it is the first time that the pandemics hits α people. Show that T_α is indeed a random variable.

   (f) Let $\alpha = 10^4$. Familiarize yourself with the command `histogram`. Plot the histogram of T_{10^4} for the 100 sample paths.

   (g) Based on the observed data, provide an estimate for $\mathbb{E}[T_{10^4}]$.

Solution:

(a) Let $Y = \log W$. Since $\log 0.4 < Y$, we have $F_Y(y) = 0$ for $y < \log 0.4$. For $\log 0.4 \le y$, we can write

$$F_Y(y) = \mathbf{P}[Y \le y] = \mathbf{P}[\log(W) \le y] = \mathbf{P}[W \le \exp(y)] = \frac{\exp(y) - 0.4}{2.6}.$$

Also, since $Y \le \log 3$, we have $F_Y(y) = 1$ for $y \ge \log 3$. Therefore,

$$F_Y(y) = \begin{cases} 0, & \text{if } y < \log 0.4 \\ \frac{\exp(y)-0.4}{2.6}, & \text{if } \log 0.4 \le y < 3 \\ 1, & \text{if } \log 3 \le y \end{cases}$$

In this case

$$f_Y(y) = \begin{cases} \frac{\exp(y)}{2.6}, & \text{if } 0.4 \le y < 3 \\ 0, & \text{otherwise} \end{cases}.$$

(b) We have

$$\mathbb{E}[\log W_k] = \mathbb{E}[Y] = \int_{\log 0.4}^{\log 3} y\frac{\exp(y)}{2.6}\, dy = \frac{(\log 3 - 1)3 - (\log 0.4 - 1)0.4}{2.6} = 0.409.$$

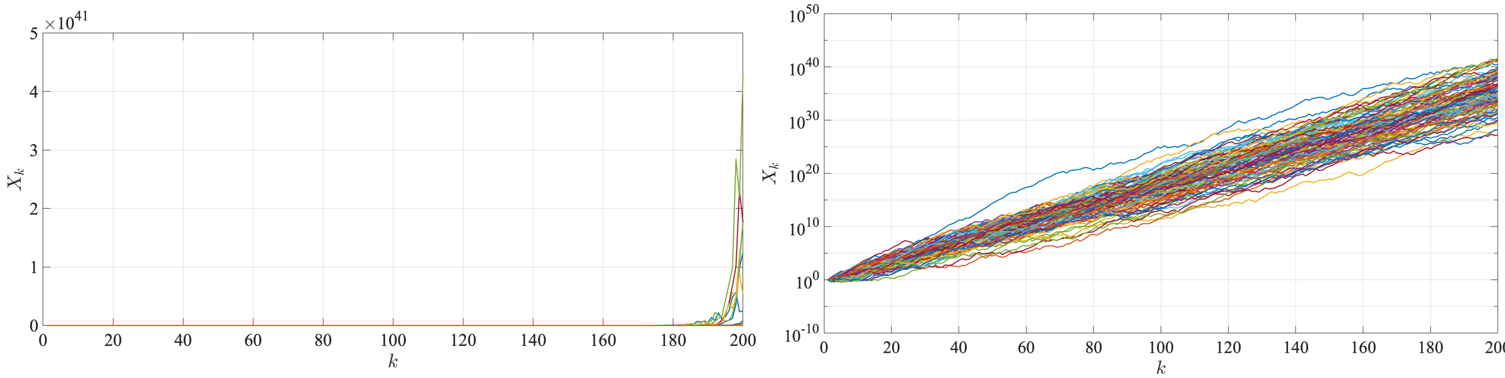
(c)

(d) Since $\mathbb{E}[\log W_k] > 0$, X_k goes exponentially to infinity, i.e., for almost all ω, we have $T(\omega, \gamma)$ for $0 < \gamma < \mathbb{E}[\log W_k]$ such that $X_k \ge e^{\gamma k}$ for all $k \ge T(\omega, \gamma)$.

(e) For $b \in \mathbb{R}$, we have

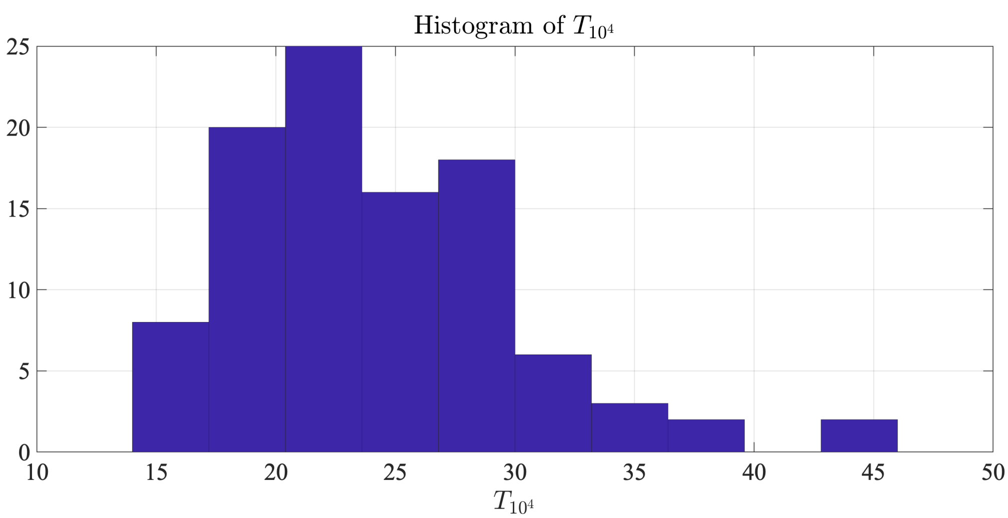
$$\begin{aligned} T_\alpha^{-1}((-\infty, b]) &= \{\omega \mid \min_k\{X_k(\omega) \ge \alpha\} \le b\} \\ &= \bigcup_{n=1}^{\lfloor b \rfloor} \{\omega \mid \min_k\{X_k(\omega) \ge \alpha\} = n\} \\ &= \bigcup_{n=1}^{\lfloor b \rfloor} \{\omega \mid X_k(\omega) < \alpha \text{ for } k < n, X_n(\omega) \ge \alpha\} \\ &= \bigcup_{n=1}^{\lfloor b \rfloor} \bigcap_{k=1}^{n-1} \{X_k < \alpha\} \cap \{X_n(\omega) \ge \alpha\} \in \mathcal{F}, \end{aligned}$$

where follows from the fact that X_ks are random variables.



(a) linear plot

(b) log plot



(f)

(g) $\mathbb{E}[T_{10^4}] = 24.26$.

```
clear all;
close all;
T=200;
N=100;
a4=10^4;
w=[ones(N,1) (3-0.4)*rand(N,T-1)+0.4];
for i=1:N
    x(i,:)=cumprod(w(i,:));
    St4(i)=find(x(i,:)>=a4,1);
end
plot(x')
xlabel('$k$','Interpreter','latex')
ylabel('$X_k$','Interpreter','latex')
figure
semilogy(x')
xlabel('$k$','Interpreter','latex')
ylabel('$X_k$','Interpreter','latex')
figure
hist(St4)
xlabel('$T_{10^4}$','Interpreter','latex')
title('Histogram of $T_{10^4}$','Interpreter','latex')
mean(St4)
```

5. For a random vector $\mathbf{X} = (X_1, \ldots, X_n)$, we define its covariance matrix to be the $n \times n$ matrix C with

$$C_{ij} = \mathbb{E}[(X_i - \bar{X}_i)(X_j - \bar{X}_j)],$$

where $\bar{X}_i = \mathbb{E}[X_i]$. Show that a covariance matrix C is always positive semi-definite. Can we say that it is always positive definite?

hint: An $n \times n$ symmetric matrix A is called positive semi-definite if $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$. If the inequality holds as a strict inequality for all non-zero x, it is called positive definite.

Solution: Covariance matrix C is calculated by the formula,

$$C \triangleq \mathbb{E}\left[(\mathbf{X} - \overline{\mathbf{X}})(\mathbf{X} - \overline{\mathbf{X}})^T\right]$$

For an arbitrary real vector $\mathbf{u}$, we can write,

$$\begin{aligned} \mathbf{u}^T C\mathbf{u} &= \mathbf{u}^T \mathbb{E}\left[(\mathbf{X} - \overline{\mathbf{X}})(\mathbf{X} - \overline{\mathbf{X}})^T\right]\mathbf{u} \\ &= \mathbb{E}\left[\mathbf{u}^T(\mathbf{X} - \overline{\mathbf{X}})(\mathbf{X} - \overline{\mathbf{X}})^T\mathbf{u}\right] \\ &= \mathbb{E}\left[Y^2\right], \end{aligned}$$

where $Y \triangleq (\mathbf{X} - \overline{\mathbf{X}})^T\mathbf{u}$ is a zero-mean random variable. Since $Y^2 \geq 0$, therefore, $\mathbb{E}\left[Y^2\right] \geq 0$ and hence, $u^T Cu \geq 0$ for all $u \in \mathbb{R}^n$ and the covariance matrix C is positive semi-definite.

6. In the class, we looked at the LMMSE estimator of X given $Y = X + Z$, where $X, Z \sim \mathcal{N}(0,1)$ are independent. We found out that in this case the LMMSE estimator is $\hat{X} = \frac{1}{2}Y$. Now, suppose that we have k independent measurements $Y_i = X + Z_i$ for $i = 1, \ldots, k$ of X where $X, Z_1, \ldots, Z_k \sim \mathcal{N}(0,1)$ are all independent. Find the LMMSE estimator of X given $Y_1, \ldots, Y_k$ using the provided formula

$$\hat{X} = \mathbf{Cov}(X, Y)\mathbf{Cov}^{-1}(Y, Y)(Y - \bar{Y}) + \bar{X}.$$

You don't need to prove the above identity.

Solution: We know $\bar{X} = 0$ and $\bar{Y}_i = 0$ for all $i \in [k]$. Since $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix}$ we have $\bar{Y} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$.

Now $\mathbf{Cov}(X, Y_i) = \mathbb{E}[X^2 + XZ_i] = 1$. Therefore $\mathbf{Cov}(X, Y) = \begin{bmatrix}1 & \dots & 1\end{bmatrix}$.

Similarly,

$$\mathbf{Cov}(Y_i, Y_j) = \mathbb{E}[(X + Z_i)(X + Z_j)] = \mathbb{E}[X^2] + \mathbb{E}[XZ_i] + \mathbb{E}[XZ_j] + \mathbb{E}[Z_iZ_j] = 1 + \mathbb{E}[Z_iZ_j].$$

We know $\mathbb{E}[Z_iZ_j] = 0$ when $i \neq j$ and $\mathbb{E}[Z_iZ_j] = \mathbb{E}[Z_i^2] = 1$, when $i = j$. Therefore for $i, j \in [k]$, $\mathbf{Cov}(Y_i, Y_j) = 1$ for $i \neq j$ and $\mathbf{Cov}(Y_i, Y_j) = 2$ for $i = j$. Therefore,

$$\mathbf{Cov}(Y, Y) = \begin{bmatrix} 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ \vdots & & & & \vdots \\ 1 & 1 & \cdots & 1 & 2 \end{bmatrix}.$$

Taking the inverse we get

$$\mathbf{Cov}^{-1}(Y, Y) = \frac{1}{k+1} \begin{bmatrix} k & -1 & -1 & \cdots & -1 \\ -1 & k & -1 & \cdots & -1 \\ \vdots & & & & \vdots \\ -1 & -1 & \cdots & -1 & k \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \hat{X} &= \frac{1}{k+1} \begin{bmatrix}1 & \dots & 1\end{bmatrix} \begin{bmatrix} k & -1 & -1 & \cdots & -1 \\ -1 & k & -1 & \cdots & -1 \\ \vdots & & & & \vdots \\ -1 & -1 & \cdots & -1 & k \end{bmatrix} Y \\ &= \frac{1}{k+1} \sum_{i=1}^{k} Y_i. \end{aligned} \tag{9}$$

7. Suppose that $\{X_k\}$ is an i.i.d. random process with finite mean $\mathbb{E}[X_k] = \mu$. By the strong law of large numbers we know that

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu \quad \text{a.s.},$$

where $S_n = \sum_{k=1}^{n} X_k$. Show that if we further have a finite variance, i.e., $\mathbf{Var}(X_k) = \sigma^2 < \infty$, then $\lim_{n\to\infty} \frac{S_n}{n} \xrightarrow{L_2} \mu$. In other words,

$$\lim_{n\to\infty} \|\frac{S_n}{n} - \mu\|^2 = \lim_{n\to\infty} \mathbf{Var}(\frac{S_n}{n}) = 0.$$

This is known as the weak law of large numbers.

Solution: The independence of the random variables implies no correlation between them, and we have that

$$\|\frac{S_n}{n} - \mu\|^2 = \mathbf{Var}\left(\frac{S_n}{n}\right) = \mathbf{Var}\left(\frac{1}{n}(X_1 + \cdots + X_n)\right) = \frac{1}{n^2}\,\mathrm{Var}\,(X_1 + \cdots + X_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Therefore, $\lim_{n\to\infty} \mathbf{Var}(\frac{S_n}{n}) = 0$.