

UNIVERSITY OF CALIFORNIA, SAN DIEGO
Electrical & Computer Engineering Department
ECE 250 - Winter Quarter 2020
Random Processes

Solutions to P.S. #4

1. *Optical communication channel.* Let the signal input to an optical channel be given by

$$X = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ 10 & \text{with probability } \frac{1}{2}. \end{cases}$$

The conditional pmf of the output of the channel $Y|\{X = 1\} \sim \text{Poisson}(1)$, i.e., Poisson with intensity $\lambda = 1$, and $Y|\{X = 10\} \sim \text{Poisson}(10)$.

Show that the MAP rule reduces to

$$D(y) = \begin{cases} 1, & y < y^* \\ 10, & \text{otherwise.} \end{cases}$$

Find y^* and the corresponding probability of error.

Solution: We have

$$p_{X|Y}(x|y) = \frac{p_{Y|X}(y|x)p_X(x)}{\sum_{x \in \mathcal{X}} p_{Y|X}(y|x)p_X(x)}.$$

For $x = 1$, this gives

$$\begin{aligned} p_{X|Y}(1|y) &= \frac{p_{Y|X}(y|1)p_X(1)}{\sum_{x \in \mathcal{X}} p_{Y|X}(y|x)p_X(x)} \\ &= \frac{\frac{e^{-1}}{y!}}{\frac{e^{-1}}{y!} + \frac{e^{-10}10^y}{y!}} \\ &= \frac{e^{-1}}{e^{-1} + e^{-10}10^y}. \end{aligned}$$

Similarly,

$$p_{X|Y}(10|y) = \frac{e^{-10}10^y}{e^{-1} + e^{-10}10^y}.$$

Thus, the MAP rule will give $D(y) = 1$ if $p_{X|Y}(1|y) > p_{X|Y}(10|y)$, i.e., if $e^{-1} > e^{-10}10^y$,

i.e., if $10^y < e^9$, i.e., if $y < \frac{9}{\log_e 10}$.

Thus, the MAP rule is given by

$$D(y) = \begin{cases} 1, & y < \frac{9}{\log_e 10}, \\ 10, & \text{otherwise.} \end{cases}$$

The probability of error is given by

$$\begin{aligned} & P\left(Y > \frac{9}{\log_e 10} \middle| X = 1\right)P(X = 1) + P\left(Y < \frac{9}{\log_e 10} \middle| X = 10\right)P(X = 10) \\ &= \frac{1}{2} \left[P\left(\text{Poisson}(1) > \frac{9}{\log_e 10}\right) + P\left(\text{Poisson}(10) < \frac{9}{\log_e 10}\right) \right] \\ &= \frac{1}{2} \left[1 - \sum_{j=0}^3 \frac{e^{-1}}{j!} + \sum_{j=0}^3 \frac{e^{-10} 10^j}{j!} \right] \\ &= .0147 \end{aligned}$$

2. *Iocane or Sennari.* An absent-minded chemistry professor forgets to label two identically looking bottles. One bottle contains a chemical named “Iocane” and the other bottle contains a chemical named “Sennari”. It is well known that the radioactivity level of “Iocane” has the $U[0, 1]$ distribution, while the radioactivity level of “Sennari” has the $\text{Exp}(1)$ distribution.

- (a) Let X be the radioactivity level measured from one of the bottles. What is the optimal decision rule (based on the measurement X) that maximizes the chance of correctly identifying the content of the bottle?
- (b) What is the associated probability of error?

Solution:

- (a) Let $Y = \begin{cases} 0, & \text{if Iocane bottle is chosen} \\ 1, & \text{if Sennari bottle is chosen.} \end{cases}$

Then, $P(Y = 1) = P(Y = 0) = 1/2$, $X|\{Y = 0\} \sim U[0, 1]$ and $X|\{Y = 1\} \sim \text{Exp}(1)$.

Thus,

$$\begin{aligned} P(Y = 0|X = x) &= \frac{f_{X|Y}(x|0)P(Y = 0)}{\sum_{y \in \{0,1\}} f_{X|Y}(x|y)P(Y = y)} \\ &= \frac{\mathbb{1}_{[0,1]}(x)}{\mathbb{1}_{[0,1]}(x) + \mathbb{1}_{[0,\infty)}(x)e^{-x}} \\ &= \begin{cases} 0, & x > 1 \\ \frac{1}{1 + e^{-x}}, & x \in [0, 1]. \end{cases} \end{aligned}$$

Thus, $P(Y = 0|X = x) > P(Y = 1|X = x) \iff P(Y = 0|X = x) > 1/2 \iff \{e^{-x} < 1\} \cup \{x \in (0, 1)\}$.

Thus, the optimal decision rule is to declare “Iocane” if $X \in (0, 1)$ (or $[0, 1]$), and to declare “Sennari” otherwise.

(b) The probability of error is given by

$$\begin{aligned} & P(X > 1|Y = 0)P(Y = 0) + P(X < 1|Y = 1)P(Y = 1) \\ &= \frac{1}{2} \left[P(U[0, 1] > 1) + P(\text{Exp}(1) < 1) \right] \\ &= \frac{1}{2} \left[\int_0^1 e^{-x} dx \right] \\ &= \frac{1 - e^{-1}}{2}. \end{aligned}$$

3. *Radar signal detection.* The signal for a radar channel $S = 0$ if there is no target and a random variable $S \sim N(0, P)$ if there is a target. Both occur with equal probability. Thus

$$S = \begin{cases} 0, & \text{with probability } \frac{1}{2} \\ X \sim N(0, P), & \text{with probability } \frac{1}{2}. \end{cases}$$

The radar receiver observes $Y = S + Z$, where the noise $Z \sim N(0, N)$ is independent of S . Find the optimal decoder for deciding whether $S = 0$ or $S = X$ and its probability of error? Provide your answer in terms of intervals of y and provide the boundary points of the intervals in terms of P and N .

Solution: Since Z is independent of X and $Z \sim \mathcal{N}(0, N)$ and $X \sim \mathcal{N}(0, P)$, we have that if $S = X$, then $Y = S + Z = X + Z \sim \mathcal{N}(0, P + N)$.

$$\text{Let } W = \begin{cases} 0, & S = 0 \\ 1, & S = X. \end{cases}$$

Then, $Y|\{W = 0\} \sim \mathcal{N}(0, N)$, and $Y|\{W = 1\} \sim \mathcal{N}(0, P + N)$.

Thus,

$$\begin{aligned} P(W = 0|Y = y) &= \frac{f_{Y|W}(y|0)P(W = 0)}{f_{Y|W}(y|0)P(W = 0) + f_{Y|W}(y|1)P(W = 1)} \\ &= \frac{f_{Y|W}(y|0)}{f_{Y|W}(y|0) + f_{Y|W}(y|1)} \\ &= \frac{\frac{1}{\sqrt{2\pi N}} \exp\left[-\frac{y^2}{2N}\right]}{\frac{1}{\sqrt{2\pi N}} \exp\left[-\frac{y^2}{2N}\right] + \frac{1}{\sqrt{2\pi(P+N)}} \exp\left[-\frac{y^2}{2(P+N)}\right]}. \end{aligned}$$

Thus,

$$\begin{aligned}
& \mathbf{P}(W = 0|Y = y) > \mathbf{P}(W = 1|Y = y) \\
& \iff \mathbf{P}(W = 0|Y = y) > 1/2 \\
& \iff \frac{1}{\sqrt{2\pi N}} \exp\left[-\frac{y^2}{2N}\right] > \frac{1}{\sqrt{2\pi(P+N)}} \exp\left[-\frac{y^2}{2(P+N)}\right] \\
& \iff \exp\left[-\frac{y^2}{2(P+N)} + \frac{y^2}{2N}\right] < \sqrt{1 + \frac{P}{N}} \\
& \iff y^2 < \frac{N(P+N)}{P} \log\left(1 + \frac{P}{N}\right).
\end{aligned}$$

Thus, the optimal decoding rule is to declare

$$\begin{aligned}
S = 0 \text{ if } y \in & \left(-\sqrt{N\left(1 + \frac{N}{P}\right) \log\left(1 + \frac{P}{N}\right)}, \sqrt{N\left(1 + \frac{N}{P}\right) \log\left(1 + \frac{P}{N}\right)}\right), \text{ and} \\
S = X \text{ otherwise.}
\end{aligned}$$

Denoting $\sqrt{N\left(1 + \frac{N}{P}\right) \log\left(1 + \frac{P}{N}\right)}$ by y^* , the probability of error is given by

$$\begin{aligned}
& \mathbf{P}\left(Y \notin (-y^*, y^*) \middle| W = 0\right) \mathbf{P}(W = 0) + \mathbf{P}\left(Y \in (-y^*, y^*) \middle| W = 1\right) \mathbf{P}(W = 1) \\
& = \frac{1}{2} \left[\mathbf{P}\left(\mathcal{N}(0, N) \notin (-y^*, y^*)\right) + \mathbf{P}\left(\mathcal{N}(0, P+N) \in (-y^*, y^*)\right) \right] \\
& = \frac{1}{2} \left(2Q(y^*/\sqrt{N}) + \Phi\left(\frac{y^*}{\sqrt{P+N}}\right) - \Phi\left(\frac{-y^*}{\sqrt{P+N}}\right) \right) \\
& = Q\left(\frac{y^*}{\sqrt{N}}\right) + \Phi\left(\frac{y^*}{\sqrt{P+N}}\right) - \frac{1}{2}. \\
& = \frac{1}{2} + Q\left(\frac{y^*}{\sqrt{N}}\right) - Q\left(\frac{y^*}{\sqrt{P+N}}\right).
\end{aligned}$$

4. *Mean and variance.* Let X and Y be random variables with joint pdf

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

Define the random variable $Z = |X| + |Y|$. Find the mean and variance of Z without first finding the pdf of Z .

Solution: We have

$$\begin{aligned}
\mathbf{E}(Z) &= \mathbf{E}(|X|) + \mathbf{E}(|Y|), \\
\text{Var}(Z) &= \mathbf{E}(Z^2) - (\mathbf{E}(Z))^2 = \mathbf{E}(X^2) + 2\mathbf{E}(|XY|) + \mathbf{E}(Y^2) - (\mathbf{E}(|X|) + \mathbf{E}(|Y|))^2.
\end{aligned}$$

So to find the mean and variance of Z without first finding the pdf of Z , all we need to know are $E(|X|)$, $E(|Y|)$, $E(X^2)$, $E(Y^2)$, and $E(|XY|)$. We have

$$\begin{aligned} E(|X|) &= \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} \int_{-\frac{1}{\sqrt{2}}+|x|}^{\frac{1}{\sqrt{2}}-|x|} |x| f_{X,Y}(x,y) dy dx = \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} |x|(\sqrt{2} - 2|x|) dx = \frac{\sqrt{2}}{6}, \\ E(X^2) &= \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} \int_{-\frac{1}{\sqrt{2}}+|x|}^{\frac{1}{\sqrt{2}}-|x|} x^2 f_{X,Y}(x,y) dy dx = \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} x^2(\sqrt{2} - 2|x|) dx = \frac{1}{12}, \\ E(|XY|) &= \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} \int_{-\frac{1}{\sqrt{2}}+|x|}^{\frac{1}{\sqrt{2}}-|x|} |xy| f_{X,Y}(x,y) dy dx = \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} |x| \left(\frac{1}{\sqrt{2}} - |x| \right)^2 dx = \frac{1}{24}. \end{aligned}$$

By symmetry, $E(|Y|) = E(|X|) = \frac{\sqrt{2}}{6}$, $E(Y^2) = E(X^2) = \frac{1}{12}$. Hence, we have

$$\begin{aligned} E(Z) &= E(|X|) + E(|Y|) = \frac{\sqrt{2}}{6} + \frac{\sqrt{2}}{6} = \frac{\sqrt{2}}{3}, \\ \text{Var}(Z) &= E(Z^2) - (E(Z))^2 = E(X^2) + 2E(|XY|) + E(Y^2) - (E(|X|) + E(|Y|))^2 \\ &= \frac{1}{12} + 2 \times \frac{1}{24} + \frac{1}{12} - \left(\frac{\sqrt{2}}{3} \right)^2 = \frac{1}{36}. \end{aligned}$$

5. *Random phase signal.* Let $Y(t) = \sin(\omega t + \Theta)$ be a sinusoidal signal with random phase $\Theta \sim \text{Unif}[-\pi, \pi]$. Assume here that ω and t are constants. Find the mean and variance of $Y(t)$. Do they depend on t ?

Solution: We have

$$\begin{aligned} E(Y(t)) &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sin(\omega t + \theta) d\theta = 0, \\ \text{Var}(Y(t)) &= E(Y^2(t)) - (E(Y(t)))^2 = E(Y^2(t)) = \int_{-\pi}^{\pi} \frac{1}{2\pi} \sin^2(\omega t + \theta) d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2} - \frac{1}{2} \cos(2\omega t + 2\theta) d\theta \\ &= \frac{1}{2}. \end{aligned}$$

So neither the mean or variance depend on t .

6. *Iterated expectation.* Let Λ and X be two random variables with

$$\Lambda \sim f_{\Lambda}(\lambda) = \begin{cases} \frac{5}{3}\lambda^{\frac{2}{3}}, & 0 \leq \lambda \leq 1 \\ 0, & \text{otherwise,} \end{cases}$$

and $X|\{\Lambda = \lambda\} \sim \text{Exp}(\lambda)$. Find $E(X)$.

Solution: Since $X|(\Lambda = \lambda) \sim \text{Exp}(\lambda)$, we have

$$E(X|\Lambda = \lambda) = \frac{1}{\lambda}.$$

Thus,

$$E[X] = E(E(X|\Lambda)) = E\left(\frac{1}{\Lambda}\right) = \int_{-\infty}^{\infty} \frac{1}{\lambda} f_{\Lambda}(\lambda) d\lambda = \int_0^1 \frac{1}{\lambda} \cdot \frac{5}{3} \lambda^{\frac{2}{3}} d\lambda = \frac{5}{2}.$$

7. *Sum of packet arrivals.* Consider a network router with two types of incoming packets, wireline and wireless. Let the random variable $N_1(t)$ denote the number of *wireline* packets arriving during time $(0, t]$ and let the random variable $N_2(t)$ denote the number of *wireless* packets arriving during time $(0, t]$. Suppose $N_1(t)$ and $N_2(t)$ are independent Poisson with pmfs

$$\begin{aligned} P\{N_1(t) = n\} &= \frac{(\lambda_1 t)^n}{n!} e^{-\lambda_1 t} & \text{for } n = 0, 1, 2, \dots \\ P\{N_2(t) = k\} &= \frac{(\lambda_2 t)^k}{k!} e^{-\lambda_2 t} & \text{for } k = 0, 1, 2, \dots \end{aligned}$$

Let $N(t) = N_1(t) + N_2(t)$ be the total number of packets arriving at the router during time $(0, t]$.

- (a) Find the mean $E(N(t))$ and variance $\text{Var}(N(t))$ of the total number of packet arrivals.
- (b) Find the pmf of $N(t)$.
- (c) Let the random variable Y be the time to receive the first packet of either type. Find the pdf of Y .
- (d) What is the probability that the first received packet is wireless?

Solution:

- (a) Since N_1 and N_2 are independent,

$$EN = EN_1 + EN_2 = \lambda_1 t + \lambda_2 t = (\lambda_1 + \lambda_2)t.$$

and

$$\text{Var}(N) = \text{Var}(N_1) + \text{Var}(N_2) = (\lambda_1 + \lambda_2)t.$$

- (b) In fact, as discussed in class, $N(t)$ is Poisson itself. To see this,

$$\begin{aligned} P\{N = n\} &= \sum_{k=0}^n P\{N_1 = k\} P\{N_2 = n - k\} \\ &= \sum_{k=0}^n \frac{(\lambda_1 t)^k}{k!} e^{-\lambda_1 t} \frac{(\lambda_2 t)^{n-k}}{(n-k)!} e^{-\lambda_2 t} \\ &= \sum_{k=0}^n \binom{n}{k} (\lambda_1 t)^k (\lambda_2 t)^{n-k} \frac{e^{-(\lambda_1 + \lambda_2)t}}{n!} \\ &= \frac{((\lambda_1 + \lambda_2)t)^n}{n!} e^{-(\lambda_1 + \lambda_2)t} \end{aligned}$$

for $n = 0, 1, 2, \dots$

- (c) Since $N(t)$ is Poisson with parameter $(\lambda_1 + \lambda_2)t$, the time to the first packet is exponential with parameter $\lambda_1 + \lambda_2$ (refer to the note below to see why). Therefore,

$$f_Y(y) = \begin{cases} (\lambda_1 + \lambda_2)e^{-(\lambda_1 + \lambda_2)y}, & y \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Alternatively, we can let X_1 and X_2 denote the times until first packets of each type and see that $Y = \min(X_1, X_2)$.

Note: Let $N(t)$ denote the number of packets arriving during time $(0, t]$ and let X be the time of the arrival of the first packet. Then

$$P\{X > t\} = P\{N(t) = 0\}$$

If $N(t)$ is a Poisson r.v. with parameter λt , then for $t \geq 0$ we have

$$P\{X > t\} = P\{N(t) = 0\} = e^{-\lambda t}$$

Therefore,

$$F_X(t) = P\{X \leq t\} = 1 - P\{X > t\} = 1 - e^{-\lambda t}$$

and

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, the time of the arrival of the first packet is exponential with parameter λ .

- (d) Let X_1 and X_2 denote the times until first packets of each type. According to the note in the previous part, X_1 and X_2 are exponential random variables with parameters λ_1 and λ_2 respectively. Therefore, the probability that the first packet is wireless is simply $P\{X_2 < X_1\} = \frac{\lambda_2}{\lambda_1 + \lambda_2}$ (refer to *First available teller* problem).

8. *Conditioning on an event.* Let X be a r.v. with pdf

$$f_X(x) = \begin{cases} 2(1-x) & \text{for } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and let the event $A = \{X \geq 1/3\}$. Find $f_{X|A}(x)$, $E(X|A)$, and $\text{Var}(X|A)$.

Solution: Note that

$$F_{X|A}(x|A) = P(X \leq x|X \in A) = \frac{\int_{\{-\infty, x\} \cap A} f_X(x) dx}{P(X \in A)}.$$

Therefore,

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{P\{X \in A\}} & x \in A \\ 0 & \text{otherwise.} \end{cases}$$

Thus for the given X and A ,

$$f_{X|A}(x) = \begin{cases} \frac{9}{2}(1-x) & x \in A \\ 0 & \text{otherwise.} \end{cases}$$

Now, using this conditional pdf

$$\mathbb{E}(X|A) = \int_{1/3}^1 \frac{9}{2}x(1-x) dx = \frac{5}{9}.$$

And

$$\mathbb{E}(X^2|A) = \int_{1/3}^1 \frac{9}{2}x^2(1-x) dx = \frac{1}{3}.$$

Thus, $\text{Var}(X|A) = \mathbb{E}(X^2|A) - (\mathbb{E}(X|A))^2 = 2/81$.

9. *Inequalities.* Label each of the following statements with $=$, \leq , or \geq . Justify each answer.

- (a) $\frac{1}{\mathbb{E}(X^2)}$ vs. $\mathbb{E}\left(\frac{1}{X^2}\right)$.
- (b) $(\mathbb{E}(X))^2$ vs. $\mathbb{E}(X^2)$.
- (c) $\text{Var}(X)$ vs. $\text{Var}(E(X|Y))$.
- (d) $\mathbb{E}(X^2)$ vs. $\mathbb{E}((E(X|Y))^2)$.

Solution:

- (a) We have, by the Cauchy-Schwartz inequality,

$$\begin{aligned} \mathbb{E}[X^2] \cdot \mathbb{E}\left[\frac{1}{X^2}\right] &\geq \left(\mathbb{E}\left[X \cdot \frac{1}{X}\right]\right)^2 \\ &= 1, \end{aligned}$$

$$\text{hence } \frac{1}{\mathbb{E}[X^2]} \leq \mathbb{E}\left[\frac{1}{X^2}\right].$$

- (b) We have

$$\begin{aligned} 0 &\leq \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\ &= \mathbb{E}[X^2] - 2(\mathbb{E}[X])^2 + (\mathbb{E}[X])^2 \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2, \end{aligned}$$

hence $(\mathbb{E}[X])^2 \leq \mathbb{E}[X^2]$.

(Note : this can also be shown directly by applying Jensen's inequality on the convex function $f(x) = x^2$.)

(c) We have

$$\begin{aligned}
\text{Var}(X) - \text{Var}(\mathbb{E}[X|Y]) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 - \mathbb{E}\left[\left(\mathbb{E}[X|Y]\right)^2\right] + \left(\mathbb{E}[\mathbb{E}[X|Y]]\right)^2 \\
&= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 - \mathbb{E}\left[\left(\mathbb{E}[X|Y]\right)^2\right] + (\mathbb{E}[X])^2 \\
&= \mathbb{E}[\mathbb{E}[X^2|Y]] - \mathbb{E}\left[\left(\mathbb{E}[X|Y]\right)^2\right] \\
&= \mathbb{E}\left[\mathbb{E}[X^2|Y] - \left(\mathbb{E}[X|Y]\right)^2\right] \\
&= \mathbb{E}[\text{Var}(X|Y)] \\
&\geq 0,
\end{aligned}$$

since $\text{Var}(X|Y) \geq 0$.

(d) Similar to the previous part, we have

$$\begin{aligned}
\mathbb{E}[X^2] - \mathbb{E}\left[\left(\mathbb{E}[X|Y]\right)^2\right] &= \mathbb{E}[\mathbb{E}[X^2|Y]] - \mathbb{E}\left[\left(\mathbb{E}[X|Y]\right)^2\right] \\
&= \mathbb{E}\left[\mathbb{E}[X^2|Y] - \left(\mathbb{E}[X|Y]\right)^2\right] \\
&= \mathbb{E}[\text{Var}(X|Y)] \\
&\geq 0,
\end{aligned}$$

since $\text{Var}(X|Y) \geq 0$.

10. *Correlation coefficient.* Let X and Y have correlation coefficient $\rho_{X,Y}$.

- (a) What is the correlation coefficient between X and $3Y$?
- (b) What is the correlation coefficient between $2X$ and $-5Y$?

Solution:

(a) The correlation coefficient between X and $3Y$ is given by

$$\begin{aligned}
\frac{\mathbb{E}[X \cdot 3Y] - \mathbb{E}[X]\mathbb{E}[3Y]}{\sqrt{\text{Var}(X) \cdot \text{Var}(3Y)}} &= \frac{3\mathbb{E}[XY] - 3\mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{9 \text{Var}(X) \text{Var}(Y)}} \\
&= \rho_{X,Y}.
\end{aligned}$$

(b) The correlation coefficient between $2X$ and $-5Y$ is given by

$$\begin{aligned}
\frac{\mathbb{E}[2X \cdot (-5Y)] - \mathbb{E}[2X]\mathbb{E}[-5Y]}{\sqrt{\text{Var}(2X) \cdot \text{Var}(-5Y)}} &= -\frac{10\mathbb{E}[XY] - 10\mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{100 \text{Var}(X) \text{Var}(Y)}} \\
&= -\rho_{X,Y}.
\end{aligned}$$

11. *Cauchy–Schwartz inequality.*

- (a) Prove the following inequality: $(E(XY))^2 \leq E(X^2)E(Y^2)$. (Hint: Use the fact that for any real t , $E((X + tY)^2) \geq 0$.)
- (b) Prove that equality holds if and only if $X = cY$ for some constant c . Find c in terms of the second moments of X and Y .
- (c) Use the Cauchy–Schwartz inequality to show the correlation coefficient $|\rho_{X,Y}| \leq 1$.
- (d) Prove the *triangle inequality*: $\sqrt{E((X + Y)^2)} \leq \sqrt{E(X^2)} + \sqrt{E(Y^2)}$.

Solution:

- (a) We have, for every $t \in \mathbb{R}$, $E[(X + tY)^2] \geq 0$, i.e., $\min_{t \in \mathbb{R}} (t^2 E[Y^2] + 2tE[XY] + E[X^2]) \geq 0$.

From calculus, we see that the expression on the left attains its minimum value when $t = -\frac{E[XY]}{E[Y^2]}$, and this minimum value is given by

$E[X^2] - \frac{(E[XY])^2}{E[Y^2]}$. Thus, since $E[Y^2]$ is non-negative, we have

$$E[X^2]E[Y^2] - (E[XY])^2 \geq 0, \text{ i.e.}$$

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

- (b) For the “if” part, we see that if $X = cY$ for some constant c , then $E[X^2] = c^2 E[Y^2]$ and $E[XY] = cE[Y^2]$, thus

$$\begin{aligned} (E[XY])^2 &= c^2 (E[Y^2])^2 \\ &= c^2 E[Y^2]E[Y^2] \\ &= E[X^2]E[Y^2]. \end{aligned}$$

Thus, equality holds in this case.

For the “only if” part, we see from part (a) that equality will hold only if for some real t , $E[(X + tY)^2] = 0$.

Since $(X + tY)^2$ is non-negative, this implies that $X + tY = 0$, i.e., $X = -tY$. Writing $c = -t$, the result follows.

- (c) Writing $X_1 = X - E[X]$ and $Y_1 = Y - E[Y]$, we have, from the Cauchy–Schwartz

inequality,

$$\begin{aligned}
& (\mathbb{E}[X_1 Y_1])^2 \leq \mathbb{E}[X_1^2] \mathbb{E}[Y_1^2] \\
& \implies \frac{(\mathbb{E}[X_1 Y_1])^2}{\mathbb{E}[X_1^2] \mathbb{E}[Y_1^2]} \leq 1 \\
& \implies \frac{(\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])])^2}{\mathbb{E}[(X - \mathbb{E}[X])^2] \mathbb{E}[(Y - \mathbb{E}[Y])^2]} \leq 1 \\
& \implies \frac{(\text{Cov}(X, Y))^2}{\text{Var}(X) \text{Var}(Y)} \leq 1 \\
& \implies \rho_{X,Y}^2 \leq 1 \\
& \implies |\rho_{X,Y}| \leq 1.
\end{aligned}$$

(d) We have $(\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2] \mathbb{E}[Y^2]$ and thus,

$$\mathbb{E}[XY] \leq |\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2] \mathbb{E}[Y^2]}.$$

Thus,

$$\begin{aligned}
\mathbb{E}[(X + Y)^2] &= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\mathbb{E}[XY] \\
&\leq \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\sqrt{\mathbb{E}[X^2] \mathbb{E}[Y^2]} \\
&= \left(\sqrt{\mathbb{E}[X^2]} + \sqrt{\mathbb{E}[Y^2]} \right)^2,
\end{aligned}$$

and taking square roots, the result follows.