Solutions to Homework Set #4
(Prepared by TA Halyun Jeong)

1. Read Sections 4.3, 4.6, 5.7, 5.9, 6.5 in the text. Try to work on all examples.

2. *Two envelopes.* An amount A is placed in one envelope and the amount $2A$ is placed in another envelope. The amount A is fixed but unknown to you. The envelopes are shuffled and you are given one of the envelopes at random. Let X denote the amount you observe in this envelope. Designate by Y the amount in the other envelope. Thus

$$(X, Y) = \begin{cases} (A, 2A), & \text{with probability } \frac{1}{2}, \\ (2A, A), & \text{with probability } \frac{1}{2}. \end{cases}$$

You may keep the envelope you are given, or you can switch envelopes and receive the amount in the other envelope.

(a) Find $E(X)$ and $E(Y)$.
(b) Find $E\left(\frac{X}{Y}\right)$.
(c) Suppose you switch. What is the expected amount you receive?

Solution:

(a) The expected amount in the first envelope is

$$\begin{aligned} E(X) &= \sum_{x \in \mathcal{X}} x p_X(x) \\ &= \frac{1}{2}A + \frac{1}{2}(2A) \\ &= \frac{3}{2}A. \end{aligned}$$

Since you are given one of the envelopes at random, the expectation is the same for the envelope you are *not* given. Thus the expected amount in the second envelope is $E(Y) = \frac{3}{2}A$.

(b) The expected factor by which the amount in the second envelope exceeds the amount in the first is

$$\begin{aligned} E\left(\frac{Y}{X}\right) &= \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \left(\frac{y}{x}\right) p_{XY}(x,y) \\ &= \frac{1}{2}\frac{A}{2A} + \frac{1}{2}\frac{2A}{A} \\ &= \frac{1}{4} + 1 \\ &= \frac{5}{4}. \end{aligned}$$

(c) If you switch, the expected amount you will receive is $E(Y) = \frac{3}{2}A$.

3. *Tall trees.* Suppose that the average height of trees on campus is 20 feet. Argue that no more than half of the tree population is taller than 40 feet.

Solution: The average height of the trees in the population is 20 feet. So $\frac{1}{n}\sum_{i=1}^{n} h_i = 20$, where n is the population size and h_i is the height of the i-th tree. If more than half of the population is at least 40 feet tall, then the average will be greater than $\frac{1}{2} \cdot 40 = 20$ feet since each person is at least 0 feet tall. Thus no more than half of the population is 40 feet tall.

Alternatively, we can use the Markov inequality with respect to the fraction of population to obtain the same result.

4. Let Λ and X be two random variables with

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

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,

$$
\begin{aligned}
E[X] &= E(E(X|\Lambda)) \\
&= E\left(\frac{1}{\Lambda}\right) \\
&= \int_{-\infty}^{\infty} \frac{1}{\lambda} f_\Lambda(\lambda)\, d\lambda \\
&= \int_0^1 \frac{1}{\lambda} \cdot \frac{5}{3} \lambda^{\frac{2}{3}}\, d\lambda \\
&= \frac{5}{3} \int_0^1 \lambda^{-\frac{1}{3}}\, d\lambda \\
&= \frac{5}{3} \left[\frac{\lambda^{\frac{2}{3}}}{\frac{2}{3}} \right]_0^1 \\
&= \frac{5}{2}(1-0) \\
&= \frac{5}{2}.
\end{aligned}
$$

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

   (a) $\frac{1}{E(X^2)}$ vs. $E\left(\frac{1}{X^2}\right)$.

   (b) $(E(X))^2$ vs. $E(X^2)$.

   (c) $\text{Var}(X)$ vs. $\text{Var}(E(X|Y))$.

   (d) $E(X^2)$ vs. $E((E(X|Y))^2)$.

   Solution:

   (a) Since $g(x) = 1/x$ is convex, by Jensen's inequality we have

   $$\frac{1}{E(X^2)} \leq E\left(\frac{1}{X^2}\right).$$

   (b) We have

   $$\text{Var}(X) = E((X-(EX))^2) = E(X^2) - (E(X))^2 \geq 0.$$

   Thus, $E(X^2) \geq (EX)^2$. (Or we can use Jensen's inequality with the convex function $g(x) = x^2$.

   (c) We will, in fact, show that

   $$\text{Var}(X) = \text{Var}(E(X|Y)) + E[\text{Var}(X|Y)]$$

   so that $\text{Var}(X) \geq \text{Var}(E(X|Y))$ (since $E[\text{Var}(X|Y)] \geq 0$). To show this, consider that the conditional variance, by definition, is equal to:

   $$\text{Var}(X|Y) = E[(X - E(X|Y))^2|Y]$$

By following the same steps as in the unconditional case, we may write:

$$\text{Var}(X|Y) = E(X^2|Y) - (E(X|Y))^2$$

and by taking expectations:

$$E[\text{Var}(X|Y)] = E[E(X^2|Y)] - E[(E(X|Y))^2] = E(X^2) - E[(E(X|Y))^2]$$

Also by the definition of the variance:

$$\text{Var}(E(X|Y)) = E[(E(X|Y))^2] - (EX)^2$$

since $EX = E[E(X|Y)]$. By adding the previous two expressions:

$$\text{Var}(X) = \text{Var}(E(X|Y)) + E[\text{Var}(X|Y)]$$

(d) From above, we know that $\text{Var}(X) \geq \text{Var}(E(X|Y))$. But $\text{Var}(X) = E(X^2) - (EX)^2$ and $\text{Var}(E(X|Y)) = E[(E(X|Y))^2] - (EX)^2$, so we have

$$E(X^2) - (EX)^2 \geq E[(E(X|Y))^2] - (EX)^2,$$

which implies $E(X^2) \geq E[(E(X|Y))^2]$.

6. Let X and Y be two random variables with joint pdf

$$f(x,y) = \begin{cases} x+y, & 0 \leq x \leq 1,\, 0 \leq y \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the MMSE estimator of X given Y.

(b) Find the corresponding MSE.

(c) Find the pdf of $Z = E(X|Y)$.

Solution:

(a) We first calculate the marginal pdf of Y, by a direct integration. For $y < 0$ or $y > 1$, we integrate over 0. For $0 \leq y \leq 1$,

$$\begin{aligned} f_Y(y) &= \int_0^1 (x+y)\,dx \\ &= y + \frac{1}{2} \end{aligned}$$

Please note the limits of the integration, which are derived from the definition of the joint pdf. Thus,

$$f_Y(y) = \begin{cases} y + \frac{1}{2}, & 0 \leq y \leq 1 \\ 0 \quad\quad , & \text{otherwise.} \end{cases}$$

Now we can calculate the conditional pdf:

$$
\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{XY}(x,y)}{f_Y(y)} \\
&= \frac{x+y}{\frac{1}{2}+y}
\end{aligned}
$$

for $0 \le x, y \le 1$. Therefore, for $0 \le y \le 1$:

$$
\begin{aligned}
E[X|Y=y] &= \int_0^1 x f_{X|Y}(x|y)\, dx \\
&= \int_0^1 \frac{x^2+yx}{y+\frac{1}{2}}\, dx \\
&= \frac{\frac{1}{3}+\frac{y}{2}}{\frac{1}{2}+y},
\end{aligned}
$$

hence

$$
E[X|Y] = \frac{\frac{1}{3}+\frac{Y}{2}}{\frac{1}{2}+Y}.
$$

(b) The MSE is given by

$$
\begin{aligned}
E(\mathrm{Var}(X|Y)) &= EX^2 - E((E(X|Y))^2) \\
&= \int_0^1 x^2\left(x+\frac{1}{2}\right) dx - \int_0^1 \left(\frac{\frac{1}{3}+\frac{y}{2}}{\frac{1}{2}+y}\right)^2 \left(y+\frac{1}{2}\right) dy \\
&= \frac{5}{12} - \left(\frac{1}{3}+\frac{\ln(3)}{144}\right) \\
&= \frac{1}{12} - \frac{\ln(3)}{144}.
\end{aligned}
$$

(c) Since $E[X|Y=y] = \frac{\frac{1}{3}+\frac{y}{2}}{\frac{1}{2}+y} = \frac{1}{2} + \frac{1}{6(1+2y)}$, we have
$\frac{5}{9} \le E[X|Y=y] \le \frac{2}{3}$ for $0 \le y \le 1$.
From the part (a), we know

$$
Z = E[X|Y] = \frac{\frac{1}{3}+\frac{Y}{2}}{\frac{1}{2}+Y}.
$$

Thus, $\frac{5}{9} \le Z \le \frac{2}{3}$.

We first find the cdf $F_Z(z)$ of Z and then differentiate it to get the pdf.

$$\begin{aligned} F_Z(z) &= \mathrm{P}\{Z \le z\} \\ &= \mathrm{P}\Big\{\frac{\frac{1}{3}+\frac{Y}{2}}{\frac{1}{2}+Y} \le z\Big\} \\ &= \mathrm{P}\{2+3Y \le 3z+6zY\} \\ &= \mathrm{P}\Big\{Y \ge \frac{3z-2}{3-6z}\Big\} \\ &= 1-\mathrm{P}\Big\{Y < \frac{3z-2}{3-6z}\Big\} \end{aligned}$$

For $0 \le \frac{3z-2}{3-6z} \le 1$, i.e., $\frac{5}{9} \le z \le \frac{2}{3}$, we have

$$\begin{aligned} F_Z(z) &= 1-\int_{\frac{3z-2}{3-6z}}^{0} f_Y(y)dx \\ &= 1-\int_{\frac{3z-2}{3-6z}}^{0} (y+\frac{1}{2})dx \end{aligned}$$

By differentiating with respect to z, we get

$$\begin{aligned} f_Z(z) &= -(\frac{3z-2}{3-6z}+\frac{1}{2}) \cdot \frac{d}{dz}(\frac{3z-2}{3-6z}) \\ &= -\frac{-1}{6(1-2z)} \cdot \frac{-1}{3(1-2z)^2} \\ &= \frac{1}{18(2z-1)^3} \end{aligned}$$

for $\frac{5}{9} \le z \le \frac{2}{3}$. Otherwise $f_Z(z) = 0$.

7. *Orthogonality.* Let $\hat{X}$ be the minimum MSE estimate of X given Y.
   (a) Show that for any function $g(y)$, $E((X-\hat{X})g(Y)) = 0$, i.e., the error $(X-\hat{X})$ and $g(Y)$ are orthogonal.
   (b) Show that

   $$\mathrm{Var}(X) = E(\mathrm{Var}(X|Y)) + \mathrm{Var}(\hat{X}).$$

   Provide a geometric interpretation for this result.

Solution:

(a) We use iterated expectation and the fact that $E(g(Y)|Y) = g(Y)$.

$$\begin{aligned} E\left((X-\hat{X})g(Y)\right) &= E\left[E\left((X-\hat{X})g(Y)|Y\right)\right] \\ &= E\left[E\left((X-E(X|Y))g(Y)\right)|Y\right] \\ &= E\left(g(Y)E\left((X-E(X|Y))|Y\right)\right) \\ &= E\left(g(Y)\left(E(X|Y)-E(X|Y)\right)\right) \\ &= 0. \end{aligned}$$

(b) First we write

$$E(\mathrm{Var}(X|Y)) = E(X^2) - E((E(X|Y))^2),$$

and

$$\begin{aligned}\mathrm{Var}(E(X|Y)) &= E((E(X|Y))^2) - (E(E(X|Y)))^2 \\ &= E((E(X|Y))^2) - (E(X))^2.\end{aligned}$$

Adding the two terms completes the proof.

Interpretation: If we view X, $E(X|Y)$, and $X - E(X|Y)$ as vectors with "norms" $\sqrt{\mathrm{Var}(X)}$, $\sqrt{\mathrm{Var}(E(X|Y))}$, and $\sqrt{E(\mathrm{Var}(X|Y))}$, respectively, then this result is a "Pythagoras theorem", where the signal, the error and the estimate are the sides of a right triangle (estimate and error being orthogonal).

8. *Additive shot noise channel.* Consider an additive noise channel $Y = X + Z$, where the signal $X \sim N(0,1)$, and the noise $Z|\{X = x\} \sim N(0,|x|)$, i.e., the variance of the noise increases linearly with the absolute value of the signal.

(a) Find $E(Z^2)$. We are expecting a numerical answer here.

(b) Find the best linear MSE estimate of X given Y.

Solution:

(a) Since $Z|\{X = x\} \sim N(0,|x|)$,

$$\begin{aligned}E(Z|X) &= 0 \\ \mathrm{Var}(Z|X = x) &= E(Z^2|X = x) - E(Z|X = x)^2 = |x|.\end{aligned}$$

Therefore,

$$\begin{aligned}E(Z^2) = E(E(Z^2|X)) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x| e^{-\frac{x^2}{2}}\, dx \\ &= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} |x| e^{-\frac{x^2}{2}}\, dx = -\frac{\sqrt{2}}{\sqrt{\pi}} e^{-\frac{x^2}{2}} |_0^\infty = \sqrt{\frac{2}{\pi}}.\end{aligned}$$

(b) From the the best linear estimate formula,

$$\hat{X} = \frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}(Y - E(Y)) + E(X).$$

We only need to determine the various values. Therefore,

$$\begin{aligned} E(X) &= 0, \\ {\sigma_X}^2 &= 1, \\ E(Y) &= E(X+Z) = E(X) + E(Z) = E(X) + E(E(Z|X)) = 0 + E(0) = 0, \\ E(XZ) &= E(E(XZ|X)) = E(XE(Z|X)) = E(X \cdot 0) = E(0) = 0, \end{aligned}$$

$$\begin{aligned} {\sigma_Y}^2 &= E(Y^2) - (E(Y))^2 = E((X+Z)^2) - 0^2 \\ &= E(X^2) + E(Z^2) + 2E(XZ) - 0 \\ &= 1 + \sqrt{\frac{2}{\pi}} + 2 \times 0 - 0 = 1 + \sqrt{\frac{2}{\pi}}, \end{aligned}$$

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E((X - E(X))(Y - E(Y))) = E(XY) \\ &= E(X(X+Z)) = E(X^2) + E(XZ) = 1 + 0 = 1. \end{aligned}$$

Using all of the above, we get

$$\hat{X} = \frac{1}{1 + \sqrt{\frac{2}{\pi}}} Y.$$

9. *Jointly Gaussian random variables.* Let X and Y be jointly Gaussian random variables with pdf

$$f_{X,Y}(x,y) = \frac{1}{\pi\sqrt{1/12}} e^{-2x^2 - 8y^2 - 4xy + 12x + 24y - 24}.$$

(a) Find $E(X)$, $E(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, and $\mathrm{Cov}(X,Y)$.

(b) Find the minimum MSE estimate of X given Y and its MSE.

Solution:

(a) We can write the joint pdf for X and Y jointly Gaussian as

$$f_{X,Y}(x,y) = \frac{\exp\left(-\left[a\,(x-\mu_X)^2 + b\,(y-\mu_Y)^2 + c\,(x-\mu_X)\,(y-\mu_Y)\right]\right)}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}},$$

where

$$\begin{aligned} a &= \frac{1}{2(1-\rho_{X,Y}^2)\sigma_X^2}, \\ b &= \frac{1}{2(1-\rho_{X,Y}^2)\sigma_Y^2}, \\ c &= \frac{-2\rho_{X,Y}}{2(1-\rho_{X,Y}^2)\sigma_X\sigma_Y}. \end{aligned}$$

By inspection of the given $f_{X,Y}(x,y)$ we find that

$$a = 2, \qquad b = 8, \qquad c = 4,$$

and we get three equations in three unknowns

$$\begin{aligned}
\rho_{X,Y} &= -\frac{c}{2\sqrt{ab}} \\
&= -\frac{4}{2\sqrt{2 \times 8}} \\
&= -\frac{1}{2}, \\
\sigma_X^2 &= \frac{1}{2(1-\rho_{X,Y}^2)a}, \\
&= \frac{1}{3} \\
\sigma_Y^2 &= \frac{1}{2(1-\rho_{X,Y}^2)b}. \\
&= \frac{1}{12}.
\end{aligned}$$

To find μ_X and μ_Y, we solve the equations

$$\begin{aligned}
2a\mu_X x + c\mu_Y x &= 12x, \\
2b\mu_Y y + c\mu_X y &= 24y,
\end{aligned}$$

and find that

$$\begin{aligned}
\mu_X &= 2, \\
\mu_Y &= 1.
\end{aligned}$$

Finally

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \rho_{X,Y}\sigma_X\sigma_Y \\
&= -\frac{1}{12}.
\end{aligned}$$

(b) X and Y are jointly Gaussian random variables. Thus, the minimum MSE estimate of X given Y is linear

$$\begin{aligned}
E(X|Y) &= \frac{\rho_{X,Y}\sigma_X}{\sigma_Y}(Y-\mu_Y) + \mu_X \\
&= -(Y-1) + 2 \\
&= 3 - Y. \\
\mathrm{MMSE} &= \mathrm{Var}(X|Y) \\
&= (1-\rho_{XY}^2)\sigma_X^2 \\
&= \frac{1}{4}.
\end{aligned}$$

10. Consider a channel with the observation $Y = XZ$, where the signal X and the noise Z are uncorrelated Gaussian random variables. Let $E(X) = 1$, $E(Z) = 3$, $\sigma_X^2 = 3$, and $\sigma_Z^2 = 8$.

    (a) Find the best MSE linear estimate of X given Y.

    (b) Suppose your friend from Caltech tells you that he was able to derive an estimator with a lower MSE. Your friend from UCLA disagrees, saying that this is not possible because the signal and the noise are Gaussian, and hence the linear MSE estimator will also be the best MSE estimator. Could your UCLA friend be wrong?

Solution:

(a) We know that the best linear estimate is given by the formula

$$\hat{X} = \frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}(Y - E(Y)) + E(X).$$

We only need to determine the various values. Note that X and Z Gaussian and uncorrelated implies they are independent. Therefore,

$$\begin{aligned}
E(Y) &= E(XZ) = E(X)E(Z) = 3 \\
E(XY) &= E(X^2Z) = E(X^2)E(Z) = (\sigma_X^2 + E^2(X))E(Z) = 12 \\
E(Y^2) &= E(X^2Z^2) = E(X^2)E(Z^2) = (\sigma_X^2 + E^2(X))(\sigma_Z^2 + E^2(Z)) = 68 \\
\sigma_Y^2 &= E(Y^2) - E^2(Y) = 59 \\
\frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2} &= \frac{E(XY) - E(X)E(Y)}{\sigma_Y^2} = \frac{9}{59}.
\end{aligned}$$

Using all of the above, we get

$$\hat{X} = \frac{9}{59}Y + \frac{32}{59}.$$

(b) The fact that the best linear estimate equals the best MMSE estimate when input and noise are independent Gaussians is only known to be true for *additive* channels. For multiplicative channels this need not be the case in general. Therefore, your UCLA friend could be wrong.