UNIVERSITY OF CALIFORNIA, SAN DIEGO

Electrical & Computer Engineering Department

ECE 250 - Winter Quarter 2020

Random Processes

Solutions to P.S. #5

1. *Neural net.* Let $Y = X + Z$, where the signal $X \sim \mathrm{U}[-1,1]$ and noise $Z \sim \mathrm{N}(0,1)$ are independent.

   (a) Find the function $g(y)$ that minimizes

   $$\mathrm{MSE} = \mathsf{E}\left[(\mathrm{sgn}(X) - g(Y))^2\right],$$

   where

   $$\mathrm{sgn}(x) = \begin{cases} -1 & x \le 0 \\ +1 & x > 0. \end{cases}$$

   (b) Plot $g(y)$ vs. y.

Solution: The minimum MSE is achieved when $g(Y) = \mathsf{E}(\mathrm{sgn}(X) \,|\, Y)$. We have

$$g(y) = \mathsf{E}(\mathrm{sgn}(X) \,|\, Y = y) = \int_{-\infty}^{\infty} \mathrm{sgn}(x) f_{X|Y}(x|y)\, dx\,.$$

To find the conditional pdf of X given Y, we use

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

Since X and Z are independent,

$$f_{Y|X}(y|x) = f_Z(y-x) \ \Rightarrow \ Y \,|\, \{X = x\} \sim \mathcal{N}(x, 1)\,.$$

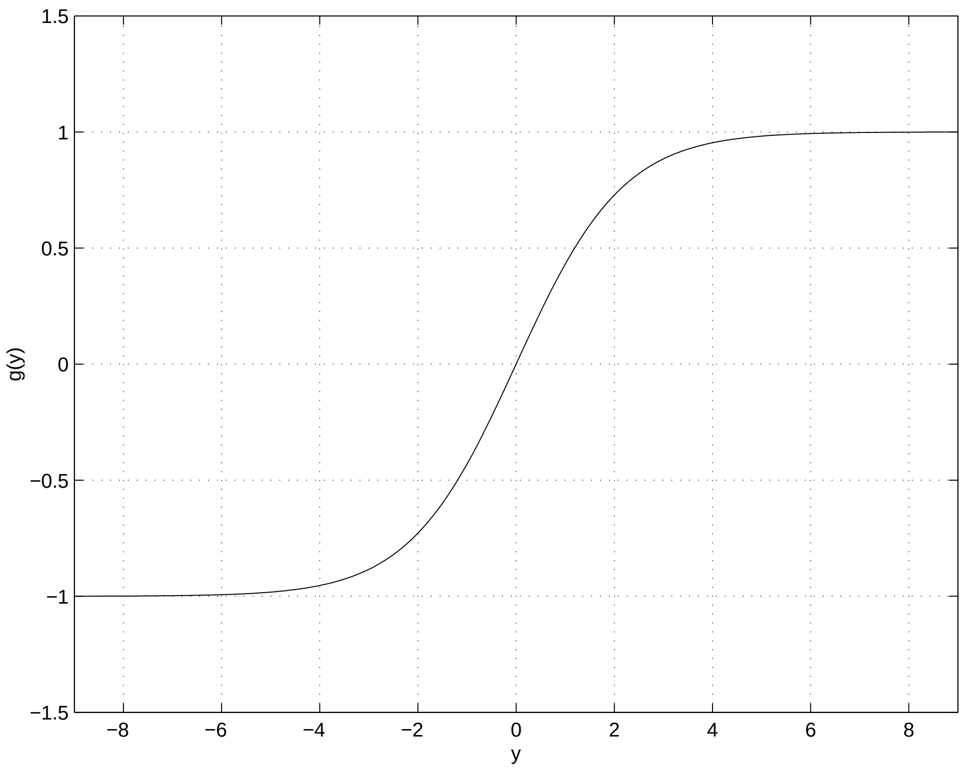
To find $f_Y(y)$ we integrate $f_{Y|X}(y|x) f_X(x)$ over x:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x)\, dx = \int_{-1}^{1} \frac{1}{2\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}\, dx \\ &= \frac{1}{2} \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}\, dx = \frac{1}{2} \left(\int_{-1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}\, dx - \int_{1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}\, dx \right) \\ &= \tfrac{1}{2} (Q(y-1) - Q(y+1))\,. \end{aligned}$$

Combining the above results, we get

$$
\begin{aligned}
g(y) &= \int_{-\infty}^{\infty} \operatorname{sgn}(x) f_{X|Y}(x|y)\, dx \;=\; \int_{-1}^{1} \operatorname{sgn}(x) \frac{\frac{1}{2\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}}{f_Y(y)}\, dx \\
&= \frac{1}{2 f_Y(y)} \left(-\int_{-1}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}\, dx + \int_{0}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}}\, dx \right) \\
&= \frac{Q(y+1) - 2Q(y) + Q(y-1)}{Q(y-1) - Q(y+1)}\,.
\end{aligned}
$$

The plot is shown below. Note the sigmoidal shape corresponding to the common neural network activation function.



2. *Additive shot noise channel.* Consider an additive noise channel $Y = X + Z$, where the signal $X \sim \mathcal{N}(0, 1)$, and the noise $Z|\{X = x\} \sim \mathcal{N}(0, x^2)$, i.e., the noise power increases linearly with the signal squared.

   (a) Find $E(Z^2)$.

   (b) Find the best linear MSE estimate of X given Y.
   [Hint: To compute $\mathrm{Var}(Y)$, use the law of conditional variance.]

Solution:

(a) $\mathsf{E}[Z^2|X=x]=x^2$, and thus

$$\begin{aligned}\mathsf{E}[Z^2] &= \mathsf{E}[\mathsf{E}[Z^2|X]]\\ &= \mathsf{E}[X^2]\\ &= 1.\end{aligned}$$

(b) Given $X=x$, $Y=x+Z\sim\mathcal{N}(x,x^2)$.

Thus, $\mathsf{E}[Y|X]=X$ and $\mathrm{Var}(Y|X)=X^2$. Hence,

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

We also have $\mathsf{E}[Y]=\mathsf{E}[\mathsf{E}[Y|X]]=0$.

Moreover, $f_{Z|X}(z|x)=\dfrac{1}{\sqrt{2\pi x^2}}\exp\Big[-\dfrac{z^2}{2x^2}\Big]$, and hence

$$\begin{aligned}f_{X,Z}(x,z) &= f_X(x)f_{Z|X}(z|x)\\ &= \frac{1}{2\pi|x|}\exp\Big[-\frac{1}{2}\Big(x^2+\frac{z^2}{x^2}\Big)\Big].\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathrm{Cov}(X,X+Z)\\ &= \mathrm{Var}(X)+\mathrm{Cov}(X,Z)\\ &= 1+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xz\frac{1}{2\pi|x|}\exp\Big[-\frac{1}{2}\Big(x^2+\frac{z^2}{x^2}\Big)\Big]dxdz\\ &= 1+0\\ &= 1,\end{aligned}$$

where the integral becomes zero since for each x, the integrand is an odd function of z.

Thus, the best linear MSE estimate of X given Y is

$$\begin{aligned}\mathsf{E}[X]+\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}(Y-\mathsf{E}[Y]) &= 0+\frac{1}{2}(Y-0)\\ &= \frac{Y}{2}.\end{aligned}$$

3. *Additive uniform noise channel.* Let the signal

$$X=\begin{cases}+1, & \text{with probability } \frac{1}{2}\\ -1, & \text{with probability } \frac{1}{2},\end{cases}$$

and the noise $Z\sim\mathrm{Unif}[-2,2]$ be independent random variables. Their sum $Y=X+Z$ is observed.

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

Solution:

We can easily find the piecewise constant density of Y

$$f_Y(y) = \begin{cases} \frac{1}{4} & |y| \le 1 \\ \frac{1}{8} & 1 < |y| \le 3 \\ 0 & \text{otherwise} \end{cases}$$

One way to do this is to use the law of total probability:

$$\begin{aligned} f_Y(y) &= f_{Y|X}(y|1)p_X(1) + f_{Y|X}(y|{-1})p_X(-1) \\ &= \frac{1}{2}f_Z(y-1) + \frac{1}{2}f_Z(y+1) \end{aligned}$$

Alternatively, since X and Z are independent, the density of Y is the convolution of the densities of X and Z. Since X is discrete, its density is expressed using impulse functions. So, we have

$$\begin{aligned} f_Y(y) &= f_X(y) * f_Z(y) \\ &= \left(\frac{1}{2}\delta(y-1) + \frac{1}{2}\delta(y+1)\right) * f_Z(y) \\ &= \frac{1}{2}f_Z(y-1) + \frac{1}{2}f_Z(y+1) \end{aligned}$$

The conditional probabilities of X given Y can be found by using Bayes' rule:

$$p_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)}{f_Y(y)}p_X(x).$$

This yields

$$\mathsf{P}\{X = +1 \,|\, Y = y\} = \begin{cases} 0 & -3 \le y < -1 \\ \frac{1}{2} & -1 \le y \le +1 \\ 1 & +1 < y \le +3 \end{cases}$$

$$\mathsf{P}\{X = -1 \,|\, Y = y\} = \begin{cases} 1 & -3 \le y < -1 \\ \frac{1}{2} & -1 \le y \le +1 \\ 0 & +1 < y \le +3 \end{cases}$$

Thus the best MSE estimate is

$$g(Y) = \mathsf{E}(X \,|\, Y) = \begin{cases} -1 & -3 \le Y < -1 \\ 0 & -1 \le Y \le +1 \\ +1 & +1 < Y \le +3 \end{cases}$$

The minimum mean square error is

$$\begin{aligned}
\mathsf{E}_Y(\mathrm{Var}(X \mid Y)) &= \mathsf{E}_Y(\mathsf{E}(X^2 \mid Y) - (\mathsf{E}(X \mid Y))^2) = \mathsf{E}(1 - g(Y)^2) \\
&= 1 - \mathsf{E}(g(Y)^2) = 1 - \int_{-\infty}^{\infty} g(y)^2 f_Y(y)\, \delta y \\
&= 1 - \left(\int_{-3}^{-1} 1 \cdot \tfrac{1}{8}\, \delta y + \int_{-1}^{1} 0 \cdot \tfrac{1}{4}\, \delta y + \int_{+1}^{+3} 1 \cdot \tfrac{1}{8}\, \delta y \right) \\
&= 1 - \int_{-3}^{-1} \tfrac{1}{8}\, \delta y - \int_{1}^{3} \tfrac{1}{8}\, \delta y = 1 - \tfrac{1}{4} - \tfrac{1}{4} = \tfrac{1}{2}\,.
\end{aligned}$$

(b) Now suppose we use a decoder to decide whether $X = +1$ or $X = -1$ so that the probability of error is minimized. Find the optimal decoder and its probability of error. Compare the optimal decoder's MSE to the minimum MSE of part (a).

Solution:

From the posterior probabilities computed in the previous part, we see that

$p_{X|Y}(+1|y) > p_{X|Y}(-1|y)$ if $y \in (1, 3]$, and $p_{X|Y}(+1|y) < p_{X|Y}(-1|y)$ if $y \in [-3, -1)$.

Thus the decision rule should be

$$g^*(y) = \begin{cases} +1, & y \in (1, 3] \\ -1, & y \in [-3, -1). \end{cases}$$

For $y \in [-1, 1]$, the posteriors are equal, and hence, let us arbitrarily choose $g^*(y) = +1$ for $y \in [-1, 1]$.

For this decoding rule, there is no error if $X = +1$, and so, the probability of error is given by

$$\begin{aligned}
\mathsf{P}\Big(\{ \text{ error } \}\Big) &= \mathsf{P}\Big(Y \notin [-3, -1) \Big| X = -1\Big) \mathsf{P}(X = -1) \\
&= \frac{1}{2} \times \frac{1}{2} \\
&= \frac{1}{4}.
\end{aligned}$$

Since there is no error if $X = +1$, the decoder's MSE is given by

$$\begin{aligned}
\mathsf{E}[(g^*(Y) - X)^2 | X = -1] \mathsf{P}(X = -1) &= \mathsf{E}[(g^*(Y) + 1)^2 | X = -1] \mathsf{P}(X = -1) \\
&= 4\mathsf{P}\Big(Y \in [-1, 1] \Big| X = -1\Big) \mathsf{P}(X = -1) \\
&= 4 \times \frac{1}{2} \times \frac{1}{2} \\
&= 1.
\end{aligned}$$

As expected, this is larger than the minimum MSE.

4. *Linear estimator.* Consider a channel with the observation $Y = XZ$, where the signal X and the noise Z are uncorrelated jointly Gaussian random variables. Let $E[X] = 1$, $E[Z] = 2$, $\sigma_X^2 = 5$, 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 best 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

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

Note that X and Z jointly Gaussian and uncorrelated implies they are independent. Therefore,

$$\begin{aligned}
\mathsf{E}(Y) &= \mathsf{E}(XZ) = \mathsf{E}(X)\mathsf{E}(Z) = 2, \\
\mathsf{E}(XY) &= \mathsf{E}(X^2Z) = \mathsf{E}(X^2)\mathsf{E}(Z) = (\sigma_X^2 + \mathsf{E}^2(X))\mathsf{E}(Z) = 12, \\
\mathsf{E}(Y^2) &= \mathsf{E}(X^2Z^2) = \mathsf{E}(X^2)\mathsf{E}(Z^2) = (\sigma_X^2 + \mathsf{E}^2(X))\,(\sigma_Z^2 + \mathsf{E}^2(Z)) = 72, \\
\sigma_Y^2 &= \mathsf{E}(Y^2) - \mathsf{E}^2(Y) = 68, \\
\frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2} &= \frac{\mathsf{E}(XY) - \mathsf{E}(X)E(Y)}{\sigma_Y^2} = \frac{5}{34}.
\end{aligned}$$

Using all of the above, we get

$$\hat{X} = \frac{5}{34}Y + \frac{12}{17}.$$

(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. In the following, we prove Y is not Gaussian by contradiction.
Suppose Y is Gaussian, then $Y \sim N(2, 68)$. We have

$$f_Y(y) = \frac{1}{\sqrt{2\pi \times 68}} e^{-\frac{(y-2)^2}{2\times 68}}.$$

On the other hand, as a function of two random Variables, Y has pdf

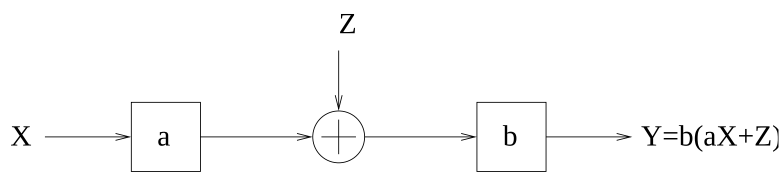
$$f_Y(y) = \int_{-\infty}^{\infty} f_X(x) f_Z\left(\frac{y}{x}\right) dx.$$

But these two expressions are not consistent, because

$$\begin{aligned} f_Y(0) &= \int_{-\infty}^{\infty} f_X(x) f_Z\left(\frac{0}{x}\right) dx = f_Z(0) \int_{-\infty}^{\infty} f_X(x) dx = f_Z(0) \\ &= \frac{1}{\sqrt{2\pi \times 8}} e^{-\frac{(0-2)^2}{2\times 8}} \\ &\neq \frac{1}{\sqrt{2\pi \times 68}} e^{-\frac{(0-2)^2}{2\times 68}} = f_Y(0)\,, \end{aligned}$$

which is a contradiction. Hence, X and Y are not joint Gaussian, and we might be able to derive an estimator with a lower MSE.

5. *Additive-noise channel with path gain.* Consider the additive noise channel shown in the figure below, where X and Z are zero mean and uncorrelated, and a and b are constants.



Find the MMSE linear estimate of X given Y and its MSE in terms only of σ_X, σ_Z, a, and b.

Solution: By the theorem of MMSE linear estimate, we have

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

Since X and Z are zero mean and uncorrelated, we have

$$\begin{aligned} \mathsf{E}(X) &= 0, \\ \mathsf{E}(Y) &= b(a\mathsf{E}(X) + \mathsf{E}(Z)) = 0, \\ \mathrm{Cov}(X,Y) &= \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y) = \mathsf{E}(Xb(aX+Z)) = a\,b\,\sigma_X^2, \\ \sigma_Y^2 &= \mathsf{E}(Y^2) - (\mathsf{E}(Y))^2 = \mathsf{E}(b^2(aX+Z)^2) = b^2a^2\sigma_X^2 + b^2\sigma_Z^2. \end{aligned}$$

Hence, the best linear MSE estimate of X given Y is given by

$$\hat{X} = \frac{a\,\sigma_X^2}{b\,a^2\sigma_X^2 + b\,\sigma_Z^2} Y.$$

6. *Worst noise distribution.* Consider an additive noise channel $Y = X+Z$, where the signal $X \sim \mathcal{N}(0,P)$ and the noise Z has zero mean and variance N. Assume X and Z are independent. Find a distribution of Z that maximizes the minimum MSE of estimating X given Y, i.e., the distribution of the worst noise Z that has the given mean and variance. You need to justify your answer.

Solution: Let us calculate the MSE of the linear MMSE estimator of X given Y, which will serve as an upper bound on the minimum MSE of estimating X given Y.

We have, since X and Z are independent,

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{Var}(X+Z) \\
&= \mathrm{Var}(X) + \mathrm{Var}(Z) \\
&= P + N, \text{ and}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathrm{Cov}(X, X+Z) \\
&= \mathrm{Var}(X) + \mathrm{Cov}(X,Z) \\
&= \mathrm{Var}(X) \\
&= P.
\end{aligned}$$

Thus, the MSE of the linear MMSE estimator of X given Y is

$$\begin{aligned}
\mathrm{Var}(X) - \frac{\Big(\mathrm{Cov}(X,Y)\Big)^2}{\mathrm{Var}(Y)} &= P - \frac{P^2}{P+N} \\
&= \frac{PN}{P+N}.
\end{aligned}$$

Thus, the minimum MSE of estimating X given Y is upper-bounded by $\dfrac{PN}{P+N}$, and we also know that the linear MMSE estimator is the same as the MMSE estimator when X and Y are jointly Gaussian.

X and Y are jointly Gaussian if $Z \sim \mathcal{N}(0, N)$, and this is thus the noise distribution that makes the MSE of the linear MMSE estimator the same as the overall minimum MSE.

Thus, the distribution $Z \sim \mathcal{N}(0, N)$ maximizes the minimum MSE of estimating X given Y.

7. *Image processing.* A pixel signal $X \sim \mathrm{U}[-k, k]$ is digitized to obtain

$$\tilde{X} = i + \frac{1}{2}, \text{ if } i < X \le i+1,\ i = -k,\ -k+1, \ldots,\ k-2,\ k-1.$$

To improve the the visual appearance, the digitized value $\tilde{X}$ is dithered by adding an independent noise Z with mean $\mathsf{E}(Z) = 0$ and variance $\mathrm{Var}(Z) = N$ to obtain $Y = \tilde{X} + Z$.

(a) Find the correlation of X and Y.

(b) Find the best linear MSE estimate of X given Y. Your answer should be in terms only of k, N, and Y.

Solution:

(a) From the definition of $\tilde{X}$, we know $\mathrm{P}\{\tilde{X} = i + \frac{1}{2}\} = \mathrm{P}\{i < X \le i+1\} = \frac{1}{2k}$. By the law of total expectation, we have

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y) = \mathsf{E}(X(\tilde{X}+Z)) = \mathsf{E}(X\tilde{X}) \\
&= \sum_{i=-k}^{k-1} \mathsf{E}[X\tilde{X} \mid i < X \le i+1] P(i < X \le i+1) \\
&= \sum_{i=-k}^{k-1} \int_i^{i+1} x(i+\frac{1}{2})\frac{1}{2k} dx = \frac{1}{8k}\sum_{i=-k}^{k-1}(2i+1)^2 = \frac{1}{4k}\sum_{i=1}^{k}(2i-1)^2 \\
&= \frac{4k^2-1}{12}.
\end{aligned}$$

Since, $\sum_{i=1}^{k} i^2 = k(k+1)(2k+1)/6$.

(b) We have

$$\begin{aligned}
\mathsf{E}(X) &= 0, \\
\mathsf{E}(Y) &= \mathsf{E}(\tilde{X}) + \mathsf{E}(Z) = 0, \\
\sigma_Y^2 &= \mathrm{Var}\,\tilde{X} + \mathrm{Var}\,Z = \sum_{i=-k}^{k-1}(i+\frac{1}{2})^2\frac{1}{2k} + N = \frac{1}{4k}\sum_{i=0}^{k}(2i+1)^2 + N = \frac{4k^2-1}{12} + N.
\end{aligned}$$

Then, the best linear MMSE estimate of X given Y is given by

$$\begin{aligned}
\hat{X} &= \frac{\mathrm{Cov}(X,Y)}{\sigma_Y^2}(Y - \mathsf{E}(Y)) + \mathsf{E}(X) = \frac{\frac{4k^2-1}{12}}{\frac{4k^2-1}{12}+N}Y \\
&= \frac{4k^2-1}{4k^2-1+12N}Y.
\end{aligned}$$

8. *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 have

$$\begin{aligned}
\mathsf{E}[(X-\hat{X})g(Y)] &= \mathsf{E}[Xg(Y)] - \mathsf{E}[\mathsf{E}[X|Y]g(Y)] \\
&= \mathsf{E}[Xg(Y)] - \mathsf{E}[\mathsf{E}[Xg(Y)|Y]] \\
&= \mathsf{E}[Xg(Y)] - \mathsf{E}[Xg(Y)] \quad \text{(by the law of iterated expectation)} \\
&= 0.
\end{aligned}$$

(b) We have

$$\begin{aligned}
\mathsf{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\hat{X}) &= \mathsf{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(\mathsf{E}[X|Y]) \\
&= \mathsf{E}\Big[E[X^2|Y]\Big] - \mathsf{E}\Big[\Big(\mathsf{E}[X|Y]\Big)^2\Big] + \mathsf{E}\Big[\Big(\mathsf{E}[X|Y]\Big)^2\Big] - \Big(\mathsf{E}[\mathsf{E}[X|Y]]\Big)^2 \\
&= \mathsf{E}[X^2] - (\mathsf{E}[X])^2 \\
&= \mathrm{Var}(X).
\end{aligned}$$

9. *Difference from sum.* Let X and Y be two random variables. Let $Z = X + Y$ and let $W = X - Y$. Find the best linear estimate of W given Z as a function of $E(X), E(Y), \sigma_X, \sigma_Y, \rho_{XY}$ and Z.

Solution: By the theorem of MMSE linear estimate, we have

$$\hat{W} = \frac{\mathrm{Cov}(W,Z)}{\sigma_Z^2}(Z - \mathsf{E}(Z)) + \mathsf{E}(W).$$

Here we have

$$\begin{aligned}
\mathsf{E}(W) &= \mathsf{E}(X) - \mathsf{E}(Y), \\
\mathsf{E}(Z) &= \mathsf{E}(X) + \mathsf{E}(Y), \\
\sigma_Z^2 &= \sigma_X^2 + \sigma_Y^2 + 2\rho_{XY}\sigma_X\sigma_Y, \\
\mathrm{Cov}(W,Z) &= \mathsf{E}(WZ) - \mathsf{E}(W)\mathsf{E}(Z) = \mathsf{E}((X-Y)(X+Y)) - (\mathsf{E}(X) - \mathsf{E}(Y))(\mathsf{E}(X) + \mathsf{E}(Y)) \\
&= \mathsf{E}(X^2) - \mathsf{E}(Y^2) - (\mathsf{E}(X))^2 + (\mathsf{E}(Y))^2 = \sigma_X^2 - \sigma_Y^2.
\end{aligned}$$

So the best linear estimate of W given Z is

$$\hat{W} = \frac{\sigma_X^2 - \sigma_Y^2}{\sigma_X^2 + \sigma_Y^2 + 2\rho_{XY}\sigma_X\sigma_Y}(Z - \mathsf{E}(X) - \mathsf{E}(Y)) + \mathsf{E}(X) - \mathsf{E}(Y).$$

10. *Sum of exponentials via transforms.* Let $X \sim \mathrm{Exp}(\lambda)$ and $Y \sim \mathrm{Exp}(\mu)$ be independent exponential random variables, where λ and μ are positive constants. Using transform methods, evaluate the probability density function of $Z = X + Y$ if $\mu \neq \lambda$ and if $\mu = \lambda$.

Solution: The characteristic functions are given by:

$$\Phi_X(\omega) = \frac{\lambda}{\lambda - j\omega}, \quad \Phi_Y(\omega) = \frac{\mu}{\mu - j\omega}.$$

and

$$\Phi_Z(\omega) = \frac{\lambda\mu}{(\lambda - j\omega)(\mu - j\omega)}$$

Case $\mu \neq \lambda$:

By partial fraction expansion:

$$\Phi_Z(\omega) = \frac{\mu}{\mu - \lambda} \cdot \frac{\lambda}{\lambda - j\omega} - \frac{\lambda}{\mu - \lambda} \cdot \frac{\mu}{\mu - j\omega}$$

The inverse transform gives:

$$f_Z(z) = \frac{\mu}{\mu - \lambda}\lambda e^{-\lambda z}u(z) - \frac{\lambda}{\mu - \lambda}\mu e^{-\mu z}u(z),$$

where $u(z)$ is the unit step function

$$u(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0. \end{cases}$$

Case $\mu = \lambda$:

$$\Phi_Z(\omega) = \frac{\lambda^2}{(\lambda - j\omega)^2}.$$

The inverse transform gives:

$$f_Z(z) = \lambda^2 z \lambda e^{-\lambda z}u(z).$$

11. *Moment theorem application.* The discrete random variable X has probability mass function given by

$$p_X(n) = \left(\frac{1}{2}\right)^n, \ n = 1, 2, \ldots.$$

Evaluate the mean and variance of X using the moment theorem.

Solution: The characteristic function is given by:

$$\begin{aligned}\Phi_X(\omega) &= \sum_{n=1}^{\infty} p_X(n)e^{j\omega n} \\ &= \sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n e^{j\omega n} \\ &= \sum_{n=1}^{\infty}\left(\frac{1}{2}e^{j\omega}\right) n \\ &= \frac{e^{j\omega}}{2 - e^{j\omega}}.\end{aligned}$$

Using the moment theorem,

$$\mathsf{E}[X] = \frac{1}{j}\frac{d}{d\omega}\Phi_X(\omega)\Big|_{\omega=0} = \frac{2e^{j\omega}}{(2 - e^{j\omega})^2}\Big|_{\omega=0} = 2.$$

Similarly,

$$\mathsf{E}[X^2] = \left(\frac{1}{j}\right)^2 \frac{d^2}{d\omega^2}\Phi_X(\omega)\Big|_{\omega=0} = \frac{2e^{j\omega}(2 + e^{j\omega})}{(2 - e^{j\omega})^3}\Big|_{\omega=0} = 6.$$

So,

$$\mathrm{Var}(X) = \mathsf{E}[X^2] - (\mathsf{E}[X])^2 = 2.$$