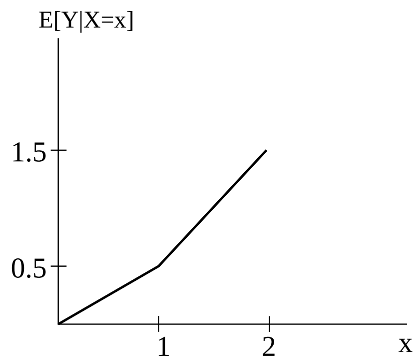
Recitation 24 Solutions

1. (a) $c = \frac{2}{3}$.

   (b) Here we are trying to choose a $g(X)$ that minimizes the conditional mean squared error $\mathbf{E}[(Y - g(X))^2|X]$. As shown in Section 8.3 in the text, this estimator is $g(X) = \mathbf{E}[Y|X]$.

$$g(x) = \mathbf{E}[Y|X = x] \quad = \quad \begin{cases} \frac{1}{2}x & 0 \le x < 1 \\ x - \frac{1}{2} & 1 \le x \le 2 \\ \text{Undefined} & \text{Otherwise} \end{cases}$$

A plot of $g(x)$:



   (c) $\mathbf{E}[g^*(X)] = \mathbf{E}[\mathbf{E}[Y|X]] = \mathbf{E}[Y]$, and note that $f_{X,Y}(x,y) = c = \frac{2}{3}$.

$$\begin{aligned} \mathbf{E}[Y] &= \int_x \int_y y f_{X,Y}(x,y) dy dx \\ &= \int_0^1 \int_0^x y\, \frac{2}{3} dy dx + \int_1^2 \int_{x-1}^x y\, \frac{2}{3} dy dx \\ &= \frac{7}{9} \end{aligned}$$

$$\begin{aligned} \mathrm{var}(g^*(X)) &= \mathrm{var}(\mathbf{E}[Y|X]) = \mathbf{E}[\mathbf{E}[Y|X]^2] - (\mathbf{E}[\mathbf{E}[Y|X]])^2 \\ &= \int_0^2 \mathbf{E}[Y|X = x]^2 f_X(x) dx - (\mathbf{E}[Y])^2 \\ &= \int_0^1 (\frac{1}{2}x)^2 \cdot \frac{2}{3} x dx + \int_1^2 (x - \frac{1}{2})^2 \cdot \frac{2}{3} dx - (\frac{7}{9})^2 \\ &= \frac{103}{648} \\ &= 0.159 \end{aligned}$$

where $f_X(x) = \frac{2}{3}x$ for $0 \le x \le 1$, $f_X(x) = \frac{2}{3}$ for $1 \le x \le 2$, and 0 otherwise.

(d) $\mathbf{E}[(Y - g^*(X))^2]$ and $\mathbf{E}[\text{var}(Y|X)]$ are the same thing.

$$\begin{aligned}\mathbf{E}[(Y - g^*(X))^2] &= \int_x \int_y (y - \mathbf{E}[Y|X=x])^2 f_{Y|X}(y|x) f_X(x) dy dx \\ &= \int_x \text{var}(Y|X=x) f_X(x) dx \\ &= \mathbf{E}[\text{var}(Y|X)]\end{aligned}$$

For any given value of X, $f_{Y|X}(y|x)$ is uniform. When $0 \le x \le 1$, $f_{Y|X}(y|x)$ is uniform over $0 \le y \le x$. When $1 \le x \le 2$, $f_{Y|X}(y|x)$ is uniform over $x - 1 \le y \le x$. Thus,

$$\text{var}(Y|X=x) = \begin{cases} \frac{(x-0)^2}{12} = \frac{x^2}{12} & 0 \le x < 1 \\ \frac{(x-(x-1))^2}{12} = \frac{1}{12} & 1 \le x \le 2 \\ \text{undefined} & \text{otherwise} \end{cases}$$

$$\begin{aligned}\mathbf{E}[\text{var}(Y|X)] &= \int_x \text{var}(Y|X=x) f_X(x) dx \\ &= \int_0^1 \frac{x^2}{12} \cdot \frac{2}{3} x dx + \int_1^2 \frac{1}{12} \cdot \frac{2}{3} dx \\ &= \frac{5}{72}\end{aligned}$$

(e) By the law of total variance, we have $\text{var}(Y) = \mathbf{E}[\text{var}(Y|X)] + \text{var}(\mathbf{E}[Y|X])$. Using the answers to (b) and (c),

$$\begin{aligned}\text{var}(Y) &= \mathbf{E}[\text{var}(Y|X)] + \text{var}(\mathbf{E}[Y|X]) \\ &= \frac{5}{72} + \frac{103}{648} = \frac{37}{162}\end{aligned}$$

Of course, you can always find $f_Y(y)$ first and then calculate the variance in the usual way; it's just that in this problem we happen to have both components in the sum above.

(f) The optimal linear estimate is given by,

$$l^*(X) = \mathbf{E}[Y] + \frac{\text{cov}(X,Y)}{\text{var}(X)}[X - \mathbf{E}[X]].$$

In part(c) we calculated, $\mathbf{E}[Y] = \frac{7}{9}$. In order to calculate $\text{var}(X)$ we first calculate $\mathbf{E}[X^2]$ and $\mathbf{E}[X]^2$.

$$\begin{aligned}\mathbf{E}[X^2] &= \int_0^1 x^3 \frac{2}{3} dx + \int_1^2 x^2 \frac{2}{3} dx, \\ &= \frac{31}{18}, \\ \mathbf{E}[X] &= \int_0^1 x^2 \frac{2}{3} dx + \int_1^2 x \frac{2}{3} dx, \\ &= \frac{11}{9}\end{aligned}$$

The $\mathrm{var}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{37}{162}$. To determine $\mathrm{cov}(X,Y)$ we need to evaluate $\mathbf{E}[XY]$.

$$\begin{aligned}\mathbf{E}[YX] &= \int_x \int_y xy f_{X,Y}(x,y)dydx \\ &= \int_0^1 \int_0^x yx\,\frac{2}{3}dydx + \int_1^2 \int_{x-1}^x yx\,\frac{2}{3}dydx \\ &= \frac{41}{36}\end{aligned}$$

Therefore $\mathrm{cov}(X,Y) = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = \frac{61}{324}$. Therefore,

$$l^*(X) = \frac{7}{9} + \frac{61}{74}[X - \frac{11}{9}].$$

The linear LMS estimator is unbiased, therefore expectation is given by,

$$\mathbf{E}[l^*(X)] = \mathbf{E}[Y] = \frac{7}{9}.$$

The $\mathrm{var}(l^*(X)) = \frac{\mathrm{cov}(X,Y)^2}{\mathrm{var}(X)} = 0.155$

(g) The LMS estimator is the optimal estimator among all classes of estimators that minimize the mean squared error. The linear LMS estimator therefore performs worse or equal to the LMS estimator, i.e., we expect $\mathbf{E}[(Y - l^*(X))^2] > \mathbf{E}[(Y - g^*(X))^2]$.

$$\begin{aligned}\mathbf{E}[(Y - l^*(X))^2] &= \sigma_Y^2(1-\rho^2), \\ &= \sigma_Y^2(1 - \frac{\mathrm{cov}(X,Y)^2}{\sigma_X^2\sigma_Y^2}), \\ &= \frac{37}{162}\left(1 - (\frac{61}{74})^2\right), \\ &= 0.073.\end{aligned}$$

This value is larger than $\frac{5}{72} = 0.06$.

2. (a) Let X be the number of detected photons. From Bayes' rule, we have

$$\begin{aligned}\mathbf{P}(\text{transmitter is on} \mid X = k) &= \frac{\mathbf{P}(X = k \mid \text{transmitter is on}) \cdot \mathbf{P}(\text{transmitter is on})}{\mathbf{P}(X = k)} \\ &= \frac{\mathbf{P}(\Theta + N = k)\cdot p}{\mathbf{P}(N = k)\cdot(1-p) + \mathbf{P}(\Theta + N = k)\cdot p}.\end{aligned}$$

The PMFs of Θ and $\Theta + N$ are

$$p_\Theta(\theta) = \frac{\lambda^\theta e^{-\lambda}}{\theta!}, \qquad p_{\Theta+N}(n) = \frac{(\lambda+\mu)^n e^{-(\lambda+\mu)}}{n!}.$$

Thus, using part (a) we obtain

$$\begin{aligned}\mathbf{P}(\text{transmitter is on} \mid X = k) &= \frac{p \cdot \frac{(\lambda+\mu)^k e^{-(\lambda+\mu)}}{k!}}{p\cdot\frac{(\lambda+\mu)^k e^{-(\lambda+\mu)}}{k!} + (1-p)\cdot\frac{\mu^k e^{-\mu}}{k!}} \\ &= \frac{p(\lambda+\mu)^k e^{-\lambda}}{p(\lambda+\mu)^k e^{-\lambda} + (1-p)\mu^k}.\end{aligned}$$

(b) We calculate $\mathbf{P}(\text{transmitter is on} \mid X = k)$ and decide that the transmitter is on if and only if this probability is at least 1/2; equivalently, if and only if

$$p(\lambda+\mu)^k e^{-\lambda} \geq (1-p)\mu^k.$$

(c) Let S be the number of transmitted photons, so that S is equal to Θ with probability p, and is equal to 0 with probability $1-p$. The linear LMS estimator is

$$\hat{S} = \mathbf{E}[S] + \frac{\text{cov}(S,X)}{\sigma_X^2}(X - \mathbf{E}[X]).$$

We calculate all the terms in the preceding expression.
Since Θ and N are independent, S and N are independent as well. We have

$$\begin{aligned}
\mathbf{E}[S] &= p\mathbf{E}[\Theta] = p\lambda, \\
\mathbf{E}[S^2] &= p\mathbf{E}[\Theta^2] = p(\lambda^2+\lambda), \\
\sigma_S^2 &= \mathbf{E}[S^2](\mathbf{E}[S])^2 = p(\lambda^2+\lambda) - (p\lambda)^2.
\end{aligned}$$

It follows that

$$\mathbf{E}[X] = \mathbf{E}[S] + \mathbf{E}[N] = (p\lambda+\mu),$$

and

$$\begin{aligned}
\sigma_X^2 &= \sigma_S^2 + \sigma_N^2 \\
&= p(\lambda^2+\lambda) - (p\lambda)^2 + \mu \\
&= (p\lambda+\mu) + p(1-p)\lambda^2.
\end{aligned}$$

Finally, we calculate $\text{cov}(S,X)$:

$$\begin{aligned}
\text{cov}(S,X) &= \mathbf{E}\left[(S-\mathbf{E}[S])(X-\mathbf{E}[X])\right] \\
&= \mathbf{E}\left[(S-\mathbf{E}[S])(S-\mathbf{E}[S]+N-\mathbf{E}[N])\right] \\
&= \mathbf{E}\left[(S-\mathbf{E}[S])(S-\mathbf{E}[S])\right] + \mathbf{E}\left[(S-\mathbf{E}[S])(N-\mathbf{E}[N])\right] \\
&= \sigma_S^2 + \mathbf{E}\left[(S-\mathbf{E}[S])(N-\mathbf{E}[N])\right] \\
&= \sigma_S^2 \\
&= p(\lambda^2+\lambda) - (p\lambda)^2,
\end{aligned}$$

where we have used the fact that $S - \mathbf{E}[S]$ and $N - \mathbf{E}[N]$ are independent, and that $\mathbf{E}\left[S - \mathbf{E}[S]\right] = \mathbf{E}\left[N - E[N]\right] = 0$.