9.1.10 Solved Problems

Problem 1

Let $X \sim N(0,1)$. Suppose that we know

$$Y \mid X = x \quad \sim \quad N(x,1).$$

Show that the posterior density of X given $Y = y$, $f_{X|Y}(x|y)$, is given by

$$X \mid Y = y \quad \sim \quad N\left(\frac{y}{2}, \frac{1}{2}\right).$$

Solution

Our goal is to show that $f_{X|Y}(x|y)$ is normal with mean $\frac{y}{2}$ and variance $\frac{1}{2}$. Therefore, it suffices to show that

$$f_{X|Y}(x|y) = c(y)\exp\left\{-\left(x-\frac{y}{2}\right)^2\right\},$$

where $c(y)$ is just a function of y. That is, for a given y, $c(y)$ is just the normalizing constant ensuring that $f_{X|Y}(x|y)$ integrates to one. By the assumptions,

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(y-x)^2}{2}\right\},$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}.$$

Therefore,

$$
\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)} \\
&= (\text{a function of } y) \cdot f_{Y|X}(y|x)f_X(x) \\
&= (\text{a function of } y) \cdot \exp\left\{-\frac{(y-x)^2+x^2}{2}\right\} \\
&= (\text{a function of } y) \cdot \exp\left\{-\left(x-\frac{y}{2}\right)^2+\frac{y^2}{4}\right\} \\
&= (\text{a function of } y) \cdot \exp\left\{-\left(x-\frac{y}{2}\right)^2\right\}.
\end{aligned}
$$

Problem 2

We can generalize the result of Problem 9.1 using the same method. In particular, assuming

$$X \sim N(\mu, \tau^2) \quad \text{and} \quad Y \mid X = x \quad \sim \quad N(x, \sigma^2),$$

it can be shown that the posterior density of X given $Y = y$ is given by

$$X \mid Y = y \quad \sim \quad N\left(\frac{y/\sigma^2 + \mu/\tau^2}{1/\sigma^2 + 1/\tau^2}, \frac{1}{1/\sigma^2 + 1/\tau^2}\right).$$

In this problem, you can use the above result. Let $X \sim N(\mu, \tau^2)$ and

$$Y \mid X = x \quad \sim \quad N(x, \sigma^2).$$

Suppose that we have observed the random sample Y_1, Y_2, $\cdots$, Y_n such that the Y_i's are i.i.d. and have the same distribution as Y.

a. Show that the posterior density of X given $\overline{Y}$ (the sample mean) is

$$X \mid \overline{Y} \quad \sim \quad N\left(\frac{n\overline{Y}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2}, \frac{1}{n/\sigma^2 + 1/\tau^2}\right).$$

b. Find the MAP and the MMSE estimates of X given $\overline{Y}$.

Solution

a. Since $Y \mid X = x \quad \sim \quad N(x, \sigma^2)$, we conclude

$$\overline{Y} \mid X = x \quad \sim \quad N\left(x, \frac{\sigma^2}{n}\right).$$

Therefore, we can use the posterior density given in the problem statement (we need to replace σ^2 by $\frac{\sigma^2}{n}$). Thus, the posterior density of X given $\overline{Y}$ is

$$X \mid \overline{Y} \quad \sim \quad N\left(\frac{n\overline{Y}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2}, \frac{1}{n/\sigma^2 + 1/\tau^2}\right).$$

b. To find the MAP estimate of X given $\overline{Y}$, we need to find the value that maximizes the posterior density. Since the posterior density is normal, the maximum value is obtained at the mean which is

$$\hat{X}_{MAP} = \frac{n\overline{Y}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2}.$$

Also, the MMSE estimate of X given $\overline{Y}$ is

$$\hat{X}_M = E[X|\overline{Y}] = \frac{n\overline{Y}/\sigma^2 + \mu/\tau^2}{n/\sigma^2 + 1/\tau^2}.$$

Problem 3

Let $\hat{X}_M$ be the MMSE estimate of X given Y. Show that the MSE of this estimator is

$$MSE = E\big[\mathrm{Var}(X|Y)\big].$$

Solution

We have

$$\begin{aligned}
\mathrm{Var}(X|Y) &= E[(X - E[X|Y])^2|Y] \qquad \big(\text{by definition of } \mathrm{Var}(X|Y)\big)\\
&= E[(X - \hat{X}_M)^2|Y].
\end{aligned}$$

Therefore,

$$\begin{aligned}
E[\mathrm{Var}(X|Y)] &= E\big[E[(X - \hat{X}_M^2)|Y]\big]\\
&= E[(X - \hat{X}_M)^2] \qquad (\text{by the law of iterated expectations})\\
&= MSE \qquad (\text{by definition of MSE}).
\end{aligned}$$

Problem 4

Consider two random variables X and Y with the joint PMF given in Table 9.1.

Table 9.1: Joint PMF of X and Y for Problem 4

	$Y=0$	$Y=1$
$X=0$	$\frac{1}{5}$	$\frac{2}{5}$
$X=1$	$\frac{2}{5}$	0

a. Find the linear MMSE estimator of X given Y, ($\hat{X}_L$).
b. Find the MMSE estimator of X given Y, ($\hat{X}_M$).
c. Find the MSE of $\hat{X}_M$.

Solution

Using the table we find out

$$\begin{aligned}
P_X(0) &= \frac{1}{5} + \frac{2}{5} = \frac{3}{5}, \\
P_X(1) &= \frac{2}{5} + 0 = \frac{2}{5}, \\
P_Y(0) &= \frac{1}{5} + \frac{2}{5} = \frac{3}{5}, \\
P_Y(1) &= \frac{2}{5} + 0 = \frac{2}{5}.
\end{aligned}$$

Thus, the marginal distributions of X and Y are both $Bernoulli(\frac{2}{5})$. Therefore, we have

$$\begin{aligned}
EX &= EY = \frac{2}{5}, \\
\mathrm{Var}(X) &= \mathrm{Var}(Y) = \frac{2}{5} \cdot \frac{3}{5} = \frac{6}{25}.
\end{aligned}$$

a. To find the linear MMSE estimator of X given Y, we also need $\mathrm{Cov}(X, Y)$. We have

$$EXY = \sum x_i y_j P_{XY}(x, y) = 0.$$

Therefore,

$$\begin{aligned}\mathrm{Cov}(X,Y) &= EXY - EXEY\\ &= -\frac{4}{25}.\end{aligned}$$

The linear MMSE estimator of X given Y is

$$\begin{aligned}\hat{X}_L &= \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}(Y-EY)+EX\\ &= \frac{-4/25}{6/25}\left(Y-\frac{2}{5}\right)+\frac{2}{5}\\ &= -\frac{2}{3}Y+\frac{2}{3}.\end{aligned}$$

Since Y can only take two values, we can summarize $\hat{X}_L$ in the following table.

Table 9.2: The linear MMSE estimator of X given Y for Problem 4

	$Y=0$	$Y=1$
$\hat{X}_L$	$\frac{2}{3}$	0

b. To find the MMSE estimator of X given Y, we need the conditional PMFs. We have

$$\begin{aligned}P_{X|Y}(0|0) &= \frac{P_{XY}(0,0)}{P_Y(0)}\\ &= \frac{\frac{1}{5}}{\frac{3}{5}} = \frac{1}{3}.\end{aligned}$$

Thus,

$$P_{X|Y}(1|0) = 1-\frac{1}{3} = \frac{2}{3}.$$

We conclude

$$X|Y=0 \quad \sim \quad Bernoulli\left(\frac{2}{3}\right).$$

Similarly, we find

$$\begin{aligned}P_{X|Y}(0|1) &= 1,\\ P_{X|Y}(1|1) &= 0.\end{aligned}$$

Thus, given $Y=1$, we have always $X=0$. The MMSE estimator of X given Y is

$$\hat{X}_M = E[X|Y].$$

We have

$$E[X|Y=0] = \frac{2}{3},$$
$$E[X|Y=1] = 0.$$

Thus, we can summarize $\hat{X}_M$ in the following table.

Table 9.3: The MMSE estimator of X given Y for Problem Problem 4

	$Y=0$	$Y=1$
$\hat{X}_M$	$\frac{2}{3}$	0

We notice that, for this problem, the MMSE and the linear MMSE estimators are the same. In fact, this is not surprising since here, Y can only take two possible values, and for each value we have a corresponding MMSE estimator. The linear MMSE estimator is just the line passing through the two resulting points.

c. The MSE of $\hat{X}_M$ can be obtained as

$$\begin{aligned} MSE &= E[\tilde{X}^2] \\ &= EX^2 - E[\hat{X}^2_M] \\ &= \frac{2}{5} - E[\hat{X}^2_M]. \end{aligned}$$

From the table for $\hat{X}_M$, we obtain $E[\hat{X}^2_M] = \frac{4}{15}$. Therefore,

$$MSE = \frac{2}{15}.$$

Note that here the MMSE and the linear MMSE estimators are equal, so they have the same MSE. Thus, we can use the formula for the MSE of $\hat{X}_L$ as well:

$$\begin{aligned} MSE &= (1-\rho(X,Y)^2)\mathrm{Var}(X) \\ &= \left(1 - \frac{\mathrm{Cov}(X,Y)^2}{\mathrm{Var}(X)\mathrm{Var}(Y)}\right)\mathrm{Var}(X) \\ &= \left(1 - \frac{(-4/25)^2}{6/25 \cdot 6/25}\right)\frac{6}{25} \\ &= \frac{2}{15}. \end{aligned}$$

Problem 5

Consider Example 9.9 in which X is an unobserved random variable with $EX = 0$, $\text{Var}(X) = 4$. Assume that we have observed Y_1 and Y_2 given by

$$Y_1 = X + W_1,$$
$$Y_2 = X + W_2,$$

where $EW_1 = EW_2 = 0$, $\text{Var}(W_1) = 1$, and $\text{Var}(W_2) = 4$. Assume that W_1, W_2 , and X are independent random variables. Find the linear MMSE estimator of X given Y_1 and Y_2 using the vector formula

$$\hat{\mathbf{X}}_L = \mathbf{C_{XY}}\mathbf{C_Y}^{-1}(\mathbf{Y} - E[\mathbf{Y}]) + E[\mathbf{X}].$$

Solution

Note that, here, X is a one dimensional vector, and $\mathbf{Y}$ is a two dimensional vector

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X + W_1 \\ X + W_2 \end{bmatrix}.$$

We have

$$\mathbf{C_Y} = \begin{bmatrix} \text{Var}(Y_1) & \text{Cov}(Y_1, Y_2) \\ \text{Cov}(Y_2, Y_1) & \text{Var}(Y_2) \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 8 \end{bmatrix},$$
$$\mathbf{C_{XY}} = \begin{bmatrix} \text{Cov}(X, Y_1) & \text{Cov}(X, Y_2) \end{bmatrix} = \begin{bmatrix} 4 & 4 \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \hat{\mathbf{X}}_L &= \begin{bmatrix} 4 & 4 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 4 & 8 \end{bmatrix}^{-1} \left(\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) + 0 \\ &= \begin{bmatrix} \frac{2}{3} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \\ &= \frac{2}{3} Y_1 + \frac{1}{6} Y_2, \end{aligned}$$

which is the same as the result that we obtained using the orthogonality principle in Example 9.9.

Problem 6

Suppose that we need to decide between two opposing hypotheses H_0 and H_1. Let C_{ij} be the cost of accepting H_i given that H_j is true. That is

C_{00}: The cost of choosing H_0, given that H_0 is true.

C_{10}: The cost of choosing H_1, given that H_0 is true.

C_{01}: The cost of choosing H_0, given that H_1 is true.

C_{11}: The cost of choosing H_1, given that H_1 is true.

It is reasonable to assume that the associated cost to a correct decision is less than the cost of an incorrect decision. That is, $c_{00} < c_{10}$ and $c_{11} < c_{01}$. The average cost can be written as

$$\begin{aligned} C =& \sum_{i,j} C_{ij} P(\text{choose } H_i | H_j) P(H_j) \\ =& C_{00} P(\text{choose } H_0 | H_0) P(H_0) + C_{01} P(\text{choose } H_0 | H_1) P(H_1) \\ & + C_{10} P(\text{choose } H_1 | H_0) P(H_0) + C_{11} P(\text{choose } H_1 | H_1) P(H_1). \end{aligned}$$

Our goal is to find the decision rule such that the average cost is minimized. Show that the decision rule can be stated as follows: Choose H_0 if and only if

$$f_Y(y|H_0)P(H_0)(C_{10} - C_{00}) \geq f_Y(y|H_1)P(H_1)(C_{01} - C_{11}) \qquad (9.8)$$

Solution

First, note that

$$\begin{aligned} P(\text{choose } H_0 | H_0) &= 1 - P(\text{choose } H_1 | H_0), \\ P(\text{choose } H_1 | H_1) &= 1 - P(\text{choose } H_0 | H_1). \end{aligned}$$

Therefore,

$$\begin{aligned} C =& C_{00} \left[1 - P(\text{choose } H_1 | H_0)\right] P(H_0) + C_{01} P(\text{choose } H_0 | H_1) P(H_1) \\ & + C_{10} P(\text{choose } H_1 | H_0) P(H_0) + C_{11} \left[1 - P(\text{choose } H_0 | H_1)\right] P(H_1) \\ =& (C_{10} - C_{00}) P(\text{choose } H_1 | H_0) P(H_0) + (C_{01} - C_{11}) P(\text{choose } H_0 | H_1) P(H_1) \\ & + C_{00} p(H_0) + C_{11} p(H_1). \end{aligned}$$

The term $C_{00}p(H_0) + C_{11}P(H_1)$ is constant (i.e., it does not depend on the decision rule). Therefore, to minimize the cost, we need to minimize

$$D = P(\text{choose } H_1|H_0)P(H_0)(C_{10} - C_{00}) + P(\text{choose } H_0|H_1)P(H_1)(C_{01} - C_{11}).$$

The above expression is very similar to the average error probability of the MAP test (Equation 9.8). The only difference is that we have $p(H_0)(C_{10} - C_{00})$ instead of $P(H_0)$, and we have $p(H_1)(C_{01} - C_{11})$ instead of $P(H_1)$. Therefore, we can use a decision rule similar to the MAP decision rule. More specifically, we choose H_0 if and only if

$$f_Y(y|H_0)P(H_0)(C_{10} - C_{00}) \geq f_Y(y|H_1)P(H_1)(C_{01} - C_{11}).$$

Problem 7

Let

$$X \sim N(0, 4) \quad \text{and} \quad Y \mid X = x \quad \sim \quad N(x, 1).$$

Suppose that we have observed the random sample Y_1, Y_2, $\cdots$, Y_{25} such that the Y_i's are i.i.d. and have the same distribution as Y. Find a 95% credible interval for X, given that we have observed

$$\overline{Y} = \frac{Y_1 + Y_2 + \ldots + Y_n}{n} = 0.56$$

Hint: Use the result of Problem 9.2.

Solution

By part (a) of Problem 9.2, we have

$$\begin{aligned} X \mid \overline{Y} \quad &\sim \quad N\left(\frac{25(0.56)/1 + 0/4}{25/1 + 1/4}, \frac{1}{25/1 + 1/4}\right) \\ &= N\big(0.5545, 0.0396\big). \end{aligned}$$

Therefore, we choose the interval in the form of

$$\big[0.5545 - c, 0.5545 + c\big].$$

We need to have

$$\begin{aligned} P\bigg(0.5545 - c \leq X \leq 0.5545 + c|\overline{Y} = 0.56\bigg) &= \Phi\left(\frac{c}{\sqrt{0.0396}}\right) - \Phi\left(\frac{-c}{\sqrt{0.0396}}\right) \\ &= 2\Phi\left(\frac{c}{\sqrt{0.0396}}\right) - 1 = 0.95 \end{aligned}$$

Solving for c, we obtain

$$c = \sqrt{0.0396}\Phi^{-1}(0.975) \approx 0.39$$

Therefore, the 95% credible interval for X is

$$\left[0.5545 - 0.39, 0.5545 + 0.39\right] \approx [0.1645, 0.9445].$$
