EE7401 Probability and Random Processes
RA 3 Solutions

Note that there can be many possible solutions. This is just one possible approach.

1.

(a) Let $N = \sum_{i=1}^{1000} X_i$. We want to find $\mathbb{E}[N]$. We have

$$\begin{aligned}\mathbb{E}[N] &= \sum_{i=1}^{1000} \mathbb{E}[X_i] \\ &= \sum_{i=1}^{1000} 10^{-8} \\ &= 10^{-5}.\end{aligned}$$

(b) By the Markov inequality,

$$\begin{aligned}\mathbb{P}(N \geq 10) &\leq \frac{1}{10}\mathbb{E}[N] \\ &= 10^{-6}.\end{aligned}$$

2.

(a) We have

$$\begin{aligned}p_{Y|X}(y \mid x) &= p_{Z|X}(y - x \mid x) \\ &= \begin{cases} \frac{1}{4} & -2 \leq y - x \leq 2, \\ 0 & \text{otherwise} \end{cases}\end{aligned} \tag{1}$$

and

$$\begin{aligned}p_Y(y) &= \frac{1}{2}p_{Y|X}(y \mid 1) + \frac{1}{2}p_{Y|X}(y \mid -1) \\ &= \begin{cases} \frac{1}{4} & -1 \leq y \leq 1, \\ \frac{1}{8} & -3 \leq y < -1,\ 1 < y \leq 3, \\ 0 & \text{otherwise.} \end{cases}\end{aligned}$$

Using

$$p_{X|Y}(x \mid y) = \frac{p_{Y|X}(y \mid x) p_X(x)}{p_Y(y)},$$

we have

$$\begin{aligned} p_{X|Y}(1 \mid y) &= \begin{cases} 0 & -3 \le y < -1, \\ \frac{1}{2} & -1 \le y \le 1, \\ 1 & 1 < y \le 3, \end{cases} \\ p_{X|Y}(-1 \mid y) &= \begin{cases} 1 & -3 \le y < -1, \\ \frac{1}{2} & -1 \le y \le 1, \\ 0 & 1 < y \le 3. \end{cases} \end{aligned} \tag{2}$$

The MMSE estimate of X given Y is

$$\begin{aligned} \mathbb{E}[X \mid Y = y] &= 1 \cdot p_{X|Y}(1 \mid y) + (-1) \cdot p_{X|Y}(-1 \mid y) \\ &= \begin{cases} -1 & -3 \le y < -1, \\ 0 & -1 \le y \le 1, \\ 1 & 1 < y \le 3, \end{cases} \end{aligned}$$

and its MSE is

$$\begin{aligned} \mathbb{E}[\operatorname{var}(X \mid Y)] &= \mathbb{E}\big[\mathbb{E}\big[X^2 \mid Y\big] - \mathbb{E}[X \mid Y]^2\big] \\ &= 1 - \mathbb{E}\big[\mathbb{E}[X \mid Y]^2\big] \\ &= 1 - \int_{-3}^{-1} \frac{1}{8} \, \mathrm{d}y - \int_{1}^{3} \frac{1}{8} \, \mathrm{d}y \\ &= \frac{1}{2}. \end{aligned}$$

(b) Let the MAP decoder be $D(y)$. It takes value 1 if $p_{X|Y}(1 \mid y) \ge p_{X|Y}(-1 \mid y)$ and -1 otherwise. From (2), we have

$$D(y) = \begin{cases} 1 & -1 \le y \le 3, \\ -1 & -3 \le y < -1. \end{cases}$$

The error probability is

$$\begin{aligned} \mathbb{P}(D(Y) \ne X) &= \mathbb{P}(-3 \le Y < -1 \mid X = 1) p_X(1) + \mathbb{P}(-1 \le Y \le 3 \mid X = -1) p_X(-1) \\ &= \frac{1}{4}, \end{aligned}$$

where we have made use of (1). Considering the MAP decoder as an estimator, its MSE is

$$\begin{aligned}
&\mathbb{E}\big[(D(Y)-X)^2\big] \\
&= \mathbb{E}\big[(D(Y)-X)^2 \,\big|\, D(Y)=X\big]\mathbb{P}(D(Y)=X) \\
&\qquad\qquad + \mathbb{E}\big[(D(Y)-X)^2 \,\big|\, D(Y)\neq X\big]\mathbb{P}(D(Y)\neq X) \\
&= 4\cdot\frac{1}{4} \\
&= 1.
\end{aligned}$$

As expected, the MSE of the MAP decoder is greater than the MMSE found in (a).

Remark 1. *Note that when defining the MAP decoder $D(y)$, it does not matter which value we set $D(y)$ to be in the case where $p_{X|Y}(1 \mid y) = p_{X|Y}(-1 \mid y)$. To see this, recall that we want to minimize the probability of error, which is equivalent to maximizing*

$$\mathbb{P}(D(Y)=X) = \int_{-\infty}^{\infty} p_{X|Y}(D(y) \mid y) p_Y(y)\,\mathrm{d}y. \tag{3}$$

For each y, we choose $D(y)$ such that $p_{X|Y}(D(y) \mid y)$ is maximized. From (3)*, the maximum probability is*

$$\mathbb{P}(D(Y)=X) = \int_A p_{X|Y}(1 \mid y)p_Y(y)\,\mathrm{d}y + \int_B p_{X|Y}(-1 \mid y)p_Y(y)\,\mathrm{d}y + \int_C p_{X|Y}(D(y) \mid y)p_Y(y)\,\mathrm{d}y,$$

where

- $A = \{y \,:\, p_{X|Y}(1 \mid y) > p_{X|Y}(-1 \mid y)\}$,
- $B = \{y \,:\, p_{X|Y}(1 \mid y) < p_{X|Y}(-1 \mid y)\}$ *and*
- $C = \{y \,:\, p_{X|Y}(1 \mid y) = p_{X|Y}(-1 \mid y)\}$. *The error probability does not change regardless of how we choose $D(y)$ on C.*

This is true in general for any Bayesian inference problem: Find a function $D(\cdot)$ of an observed Y to infer X based on minimizing $\mathbb{E}[\ell(D(Y),X)]$, where $\ell(\cdot,\cdot)$ is a non-negative risk *function. From*

$$\mathbb{E}[\ell(D(Y),X)] = \mathbb{E}[\mathbb{E}[\ell(D(Y),X) \,|\, Y]],$$

this is equivalent to minimizing $\mathbb{E}[\ell(D(y),X) \,|\, Y=y]$ for each y. If $\arg\min_{D(\cdot)} \mathbb{E}[\ell(D(y),X) \,|\, Y=y]$ is non-unique, we are choose any of the minimizers as they all give the same $\mathbb{E}[\ell(D(y),X) \,|\, Y=y]$ value.