Sakshi Suman (01/29/2022) suman.sak@northeastern.edu

Problem 1. *Assume* $\vec{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ *and* $\vec{b} \in \mathbb{R}^m$. *Let* $f(\vec{x}) = \vec{b}^T A\vec{x}$. *Find* ∇f.

Solution. We know that,

$$P(X = x, Y = y, Z = z) = P(Y = y) \cdot P(Z = z \mid Y = y) \cdot P(X = x \mid Z = z, Y = y)$$

Therefore,

$$\begin{aligned}\sum_{z \in \mathscr{Z}} P(X = x \mid Y = y, Z = z) \cdot P(Z = z \mid Y = y) &= \sum_{z \in \mathscr{Z}} \frac{P(X = x, Y = y, Z = z)}{P(Y = y)} \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= P(X = x \mid Y = y)\end{aligned}$$

□

Problem 2. *Let X be a random variable on* $\mathscr{X} = \{a, b, c\}$ *with the probability mass function* $p(x)$. *Let* $p(a) = 0.1$, $p(b) = 0.2$, *and* $p(c) = 0.7$ *and some function* $f(x)$ *be*

$$f(x) = \begin{cases} 10 & x = a \\ 5 & x = b \\ \frac{10}{7} & x = c \end{cases}$$

a) (5 points) What is $\mathbb{E}[f(X)]$ *?*
b) (5 points) What is $\mathbb{E}[1/p(X)]$?
c) (5 points) For an arbitrary finite set $\mathscr{X}$ *with n elements and arbitrary* $p(x)$ *on* $\mathscr{X}$, *what is* $\mathbb{E}[1/p(X)]$ *?*

Solution.

$$f(x) = \begin{cases} 10 & x = a \\ 5 & x = b \\ \frac{10}{7} & x = c \end{cases}$$

a)

$$\begin{aligned}\mathbb{E}[f(x)] &= \sum_{x \in \mathscr{X}} f(X = x) \times p(X = x) \\ &= 10 \times 0.1 + 5 \times 0.2 + \frac{10}{7} \times 0.7 \\ &= \boxed{3}\end{aligned}$$

b)

$$\begin{aligned}\mathbb{E}\left[\frac{1}{p(X)}\right] &= \sum_{x \in \mathscr{X}} \frac{1}{p(X = x)} \times p(X = x) \\ &= 1 + 1 + 1 \\ &= \boxed{3}\end{aligned}$$

c) Let,

$$\mathscr{X} = \{x_1, x_2, \ldots, x_n\}$$

$$\mathbb{E}\left[\frac{1}{p(x)}\right] = \sum_{x\in\mathscr{X}} \frac{1}{p(X=x)} \times p(X=x)$$
$$= \sum_{i=1}^{n} 1$$
$$= \boxed{n}$$

□

Problem 3. *A biased four-sided die is rolled and the down face is a random variable X described by the following* pmf

$$p(x) = \begin{cases} x/10 & x = 1,2,3,4 \\ 0 & \textit{otherwise} \end{cases}$$

Given the random variable X a biased coin is flipped and the random variable Y is 1 or zero according to whether the coin shows heads or tails. The conditional pmf is

$$p(y \mid x) = \left(\frac{x+1}{2x}\right)^{y}\left(1-\frac{x+1}{2x}\right)^{1-y}$$

where $y \in \{0,1\}$.
a) (5 points) Find the expectation $\mathbb{E}[X]$ *and variance* $V[X]$.
b) (5 points) Find the conditional pmf $p(x \mid y)$.
c) (5 points) Find the conditional expectation $\mathbb{E}[X \mid Y=1]$; *i.e., the expectation with respect to the conditional* pmf $p_{X|Y}(x \mid 1)$.

Solution. Given,

$$p(x) = \begin{cases} x/10 & x = 1,2,3,4 \\ 0 & \text{otherwise} \end{cases}$$

$$p(y \mid x) = \left(\frac{x+1}{2x}\right)^{y}\left(1-\frac{x+1}{2x}\right)^{1-y}$$

where $y \in \{0, 1\}$.
a)

$$\begin{aligned}
\mathbb{E}[x] &= \sum_{x\in\{1,2,3,4\}} x \cdot p(X = x) \\
&= 1 \times \frac{1}{10} + 2 \times \frac{2}{10} + 3 \times \frac{3}{10} + 4 \times \frac{4}{10} \\
&= \frac{30}{10} \\
&= \boxed{3}
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \sum_{x\in\mathscr{X}} x^2 p(X = x) \\
&= 1^2 \times \frac{1}{10} + 2^2 \times \frac{2}{10} + 3^2 \times \frac{3}{10} + 4^2 \times \frac{4}{10} \\
&= 10
\end{aligned}$$

$$\begin{aligned}
\mathbb{V}[X] &= \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 \\
&= 10 - 9 \\
&= \boxed{1}
\end{aligned}$$

b)

$$\begin{aligned}
p(x \mid y) &= \frac{p(y \mid x)p(x)}{p(y)} \\
&= \frac{p(y \mid x) \times p(x)}{\sum_{x\in\mathscr{X}} p(y \mid x) \times p(x)} \\
&= \frac{\left(\frac{x+1}{2x}\right)^y \left(1 - \frac{x+1}{2x}\right)^{1-y} \times \frac{x}{10}}{\left(\frac{2}{2}^y \times 0^{1-y} \times \frac{1}{10} + \left(\frac{3}{4}\right)^y \left(\frac{1}{4}\right)^{1-y} \times \frac{2}{10} + \left(\frac{2}{3}\right)^y \left(\frac{1}{3}\right)^{1-y} \times \frac{3}{10} + \left(\frac{5}{8}\right)^y \left(\frac{3}{8}\right)^{1-y} \times \frac{4}{10}\right)} \\
&= \frac{2 \times \left(\frac{x+1}{2x}\right)^y \left(1 - \frac{x+1}{2x}\right)^{1-y} \times x}{(1^y \times 0^{1-y} \times 2 + 3^y \times 1^{1-y} \times 4 + 4^y \times 2^{1-y} \times 6 + 5^y \times 3^{1-y} \times 8)}
\end{aligned}$$

Writing it in piecewise form,

$$p(x \mid y) = \begin{cases} \frac{x-1}{6} & y = 0 \\ \frac{x+1}{14} & y = 1 \end{cases}$$

c)

$$\mathbb{E}[X \mid Y = 1] = \sum_{x\in\mathscr{X}} x \cdot p_{X|Y}(x \mid 1)$$

$$p_{X|Y}(x \mid 1) = \frac{\left(\frac{x+1}{2x}\right)^1 \times \left(1 - \frac{x+1}{2x}\right)^0 \times \frac{x}{10}}{1^1 \times 0^0 \times \frac{1}{10} + \left(\frac{3}{4}\right)^1 \times \left(\frac{1}{4}\right)^0 \times \frac{2}{10} + \left(\frac{4}{6}\right)^1 \times \left(\frac{2}{6}\right)^0 \times \frac{3}{10} + \left(\frac{5}{8}\right)^1 \times \left(\frac{3}{8}\right)^0 \times \frac{4}{10}}$$
$$= \frac{x+1}{14}$$

$$\begin{aligned}\mathbb{E}[X \mid Y = 1] &= \sum_{x \in \mathscr{X}} x \times \frac{x+1}{14} \\ &= 1 \times \frac{2}{14} + 2 \times \frac{3}{14} + 3 \times \frac{4}{14} + 4 \times \frac{5}{14} \\ &= \boxed{\frac{20}{7}}\end{aligned}$$

□

Problem 4. *Suppose that the data set* $\mathscr{D} = \{1, 0, 1, 1, 1, 0, 1, 1, 1, 0\}$ *is an i.i.d. sample form a Bernoulli distribution*

$$p(x \mid \alpha) = \alpha^x (1-\alpha)^{1-x} \quad 0 < \alpha < 1$$

with an unknown parameter α.
a) (5 points) Calculate the log-likelihood of the data $\mathscr{D}$ *when* $\alpha = \frac{1}{e}$*; i.e., find* $\log p(\mathscr{D} \mid \alpha = \frac{1}{e})$*. The parameter e is the Euler number. Write the final expression as compactly as you can.*
b) (10 points) Compute the maximum likelihood estimate of α*. Show all your work.*
c) (10 points) Suppose the prior distribution for α *is the uniform distribution on (0,1). Compute the Bayes estimator for* α*. Note that* $\int_0^1 v^m (1-v)^r dv = \frac{m!r!}{(m+r+1)!}$

Solution. a)

$$\begin{aligned}p(\mathscr{D} \mid \alpha) &= p(\{x_i\}_{i=1}^n \mid \alpha) \\ &= \prod_{i=1}^n p(x_i \mid \alpha) \\ &= \alpha^{\sum_{i=1}^n x_i} \cdot (1-\alpha)^{n - \sum_{i=1}^n x_i}\end{aligned}$$

$$ll(\mathscr{D}, \alpha) = \ln \alpha \cdot \sum_{i=1}^n x_i + \ln(1-\alpha) \cdot \left(n - \sum_{i=1}^n x_i\right)$$
$$ll\left(\mathscr{D}, \frac{1}{e}\right) = 3 \ln(e-1) - 10$$

b)

$$\begin{aligned}\frac{\partial ll(\mathscr{D}, \alpha)}{\partial \alpha} &= \frac{\sum_{i=1}^n x_i}{\alpha} - \frac{n - \sum_{i=1}^n x_i}{1-\alpha} \\ &= 0\end{aligned}$$

$$\implies (1-\alpha_{ML}) \cdot \sum_{i=1}^{n} x_i - \left(n - \sum_{i=1}^{n} x_i\right) \cdot \alpha_{ML} = 0$$

$$\alpha_{ML} = \frac{\sum_{i=1}^{n} x_i}{n} = \boxed{\frac{7}{10}}$$

c)

$$\begin{aligned}
p(\mathscr{D} \mid \alpha) &= \alpha^{\sum_{i=1}^{n} x_i} \cdot (1-\alpha)^{n-\sum_{i=1}^{n} x_i} \\
p(\alpha) &= 1 \\
p(\alpha \mid \mathscr{D}) &= \frac{p(\mathscr{D} \mid \alpha) \cdot p(\alpha)}{p(\mathscr{D})} \\
p(\mathscr{D}) &= \int_0^1 p(\mathscr{D} \mid \alpha)\, p(\alpha)\; d\alpha \\
&= \int_0^1 \alpha^7 (1-\alpha)^3\; d\alpha \\
&= \frac{7! \cdot 3!}{11!}
\end{aligned}$$

$$\begin{aligned}
\alpha_B &= \int_0^1 \alpha\, p(\alpha \mid \mathscr{D})\; d\alpha \\
&= \int_0^1 \frac{\alpha \times \alpha^7 \times (1-\alpha)^3 \times 1}{p(\mathscr{D})}\, d\alpha \\
&= \frac{\int_0^1 \alpha^8 \times (1-\alpha)^3\; d\alpha}{\frac{7! \cdot 3!}{11!}} \\
&= \boxed{\frac{2}{3}}
\end{aligned}$$

□

Problem 5. *Let* $\mathscr{D} = \{x_i\}_{i=1}^{n}$ *be an i.i.d. sample from*

$$p(x) = \begin{cases} e^{-(x-\theta_0)} & x \geq \theta_0 \\ 0 & \textit{otherwise} \end{cases}$$

Determine θ_{ML} *- the maximum likelihood estimate of* θ_0*.*

Solution.

$$\begin{aligned}
p(x) &= \begin{cases} e^{-(x-\theta_0)} & x \geq \theta_0 \\ 0 & \text{otherwise} \end{cases} \\
\mathscr{D} &= \{x_i\}_{i=1}^{n}
\end{aligned}$$

For $\theta_0 \leq min\left(\{x_i\}_{i=1}^n\right)$,

$$\begin{aligned} p\left(\mathscr{D} \mid \theta_0\right) &= p\left(\{x_i\}_{i=1}^n \mid \theta_0\right) \\ &= \prod_{i=1}^n p\left(x_i \mid \theta_0\right) \\ &= e^{-\left(\sum_{i=1}^n x_i - n\cdot\theta_0\right)} \end{aligned}$$

For $\theta_0 > min\left(\{x_i\}_{i=1}^n\right)$,

$$p\left(\mathscr{D} \mid \theta_0\right) = 0$$

This is a strictly increasing function w.r.t. θ_0 until $\theta_0 = min\left(\{x_i\}_{i=1}^n\right)$ and then falls to 0 after that.

Thefore, the maximum likelihood

$$\theta_{ML} = min\left(\{x_i\}_{i=1}^n\right)$$

□