Instructions

1. Submit your answers as a pdf file on Canvas.
2. We recommend students type answers with LaTeX or word processors. A scanned handwritten copy would also be accepted. If writing by hand, write as clearly as possible. No credit may be given to unreadable handwriting.
3. Write out all steps required to find the solutions so that partial credit may be awarded.
4. The first question is meant for undergraduate students only while the last question is meant for graduate students only. Both these questions carry 5 pts each. Each of the other three questions is for 10 pts. There is no extra credit for answering additional questions than what is required.
5. We generally encourage collaboration with other students. You may discuss the questions and potential directions for solving them with another student. However, you need to write your own solutions and code separately, and not as a group activity. Please list the students you collaborated with.

Questions

1. **[Undergraduate Students Only]** Provide answers to the following operations or write "invalid" if not possible: [5 pts]

   (a) $\begin{bmatrix} 3 & 5 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 7 \end{bmatrix}$

   (b) $\begin{bmatrix} 3 & 5 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 7 \end{bmatrix}$

   (c) $\begin{bmatrix} 3 & 5 & 4 \\ 4 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 & 2 \\ 1 & 1 & 2 \end{bmatrix}^\intercal$

   (d) $\begin{bmatrix} 3 & 5 & 4 \end{bmatrix}^\intercal \begin{bmatrix} 4 & 3 \end{bmatrix}^\intercal$

   (e) $\begin{bmatrix} 3 & 5 & 4 \end{bmatrix}^\intercal \begin{bmatrix} 4 & 3 \end{bmatrix}$

   Solution.
   Solution goes here. ■

2. The entropy of a discrete random variable X is defined as (use base e for all log operations unless specified otherwise):

$$H(X) = -\sum_{x \in X} P(x) \log P(x)$$

   (a) Compute the entropy of the distribution $P(x) = \text{Multinomial}([0.2, 0.3, 0.5])$. [3 pts]

   (b) Compute the entropy of the uniform distribution $P(x) = \frac{1}{m} \forall x \in [1, m]$. [3 pts]

   (c) Consider the entropy of the joint distribution P(X, Y):

$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(x,y)$$

   How does this entropy relate to H(X) and H(Y), (i.e. the entropies of the marginal distributions) when X and Y are independent? [4 pts]

   Solution.
   a) $H(X) = -(0.2 \cdot \log(0.2) + 0.3 \cdot \log(0.3) + 0.5 \cdot \log(0.5)) = -(-1.03) = 1.03$

   b) $H(X) = -\sum_{i=1}^{m} P(x_i) \log P(x_i) = -\sum_{i=1}^{m} \frac{1}{m} \log \frac{1}{m} = -m \cdot \frac{1}{m} \log \frac{1}{m} = -\log \frac{1}{m} = -\log(1) + \log(m) = \log(m)$

   c)
   $H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(x,y)$
   $= -\sum_{x \in X} \sum_{y \in Y} P(x)P(y) \log \left[P(x)P(y)\right]$

$= -\sum_{x\in X}\sum_{y\in Y} P(x)P(y)\big[\log P(x) + \log P(y)\big]$
$= -\sum_{x\in X}\sum_{y\in Y} P(x)P(y)\log P(x) + P(x)P(y)\log P(y)$
$= -\sum_{x\in X}\sum_{y\in Y} P(x)P(y)\log P(x) - \sum_{x\in X}\sum_{y\in Y} P(x)P(y)\log P(y)$
$= -\sum_{x\in X} P(x)\log P(x)\sum_{y\in Y} P(y) - \sum_{y\in Y} P(y)\log P(y)\sum_{x\in X} P(x)$
$= -\sum_{x\in X} P(x)\log P(x) - \sum_{y\in Y} P(y)\log P(y)$
$= H(X) + H(Y)$

∴ H(X, Y) = H(X) + H(Y) if X and Y are independent ∎

3. You are investigating articles from the New York Times and from Buzzfeed. Some of the articles contain *fake* news, while others contain *real* news (assume that there are only two types of news).
   Note: for the following questions, write your answer using up to 3 significant figures.

   (a) Fake news only accounts for 5% of all articles in all newspapers. However, it is known that 30% of all fake news comes from Buzzfeed. In addition, Buzzfeed generates 25% of all news articles. What is the probability that a randomly chosen Buzzfeed article is fake news? [3 pts]

   (b) Suppose that 15% of all fake news comes from the New York Times (NYT). Furthermore, suppose that 60% of all real news comes from the NYT. Under all assumptions so far, what is the probability that a randomly chosen NYT article is fake news? [3 pts]

   (c) Mike is an active reader of the New York Times: Mike reads 80% of all NYT articles. However, he also has a suspicion that the NYT is a bad publisher, and he believes that 25% of all NYT articles are fake news. Furthermore, the NYT generates 30% of all news articles. Under all assumptions so far, what is the probability that a randomly chosen article (from all newspapers) will be from the NYT, will be read by Mike and will be *believed* to be fake news? [4 pts]

Solution.
a) $P(\text{fake}|\text{Buzzfeed}) = \frac{P(\text{Buzzfeed}|\text{fake})P(\text{fake})}{P(\text{Buzzfeed})} = \frac{0.3\cdot 0.05}{0.25} = 0.06$

There is a 6% chance the news article is fake news.

b) $P(\text{fake}|\text{NYT}) = \frac{P(\text{NYT}|\text{fake})P(\text{fake})}{P(\text{NYT})} = \frac{P(\text{NYT}|\text{fake})P(\text{fake})}{P(\text{NYT}|\text{fake})P(\text{fake})+P(\text{NYT}|\text{real})P(\text{real})} = \frac{0.15\cdot 0.05}{0.15\cdot 0.05+0.6\cdot 0.95} = \frac{0.0075}{0.5775} = 0.013$

There is a 1.3% chance the news article is fake news.

c) $P(\text{NYT}, \text{read}, \text{believed}) = P(\text{NYT})P(\text{read}|\text{NYT})P(\text{believed}|\text{NYT}) = 0.3\cdot 0.8\cdot 0.25 = 0.06$

There is a 6% chance the article is produced by the NYT, read by Mike, and believed to be fake news. ∎

4. Suppose we have a probability density function (pdf) defined as:

$$f(x,y) = \begin{cases} C(x^2+2y), & 0 < x < 1 \text{ and } 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

   (a) Find the value of C. [2pts]
   (b) Find the marginal distribution of X and Y. [4pts]
   (c) Find the joint cumulative density function (cdf) of X and Y. [4pts]

Solution.
a)
$\int_0^1 \int_0^1 C(x^2+2y)dydx$
$= \int_0^1 \int_0^1 Cx^2 + 2Cydydx$
$= \int_0^1 (Cx^2y + Cy^2)|_0^1 dx$
$= \int_0^1 Cx^2(1) + C(1)^2 - Cx^2(0) - C(0)^2 dx$
$= \int_0^1 Cx^2 + Cdx$
$= \frac{1}{3}Cx^3 + Cx|_0^1$
$= \frac{1}{3}C + C$
Since f(x, y) is a pdf, we know its integral is 1. Therefore, we need $\frac{1}{3}C + C = 1$:
$\frac{1}{3}C + C = 1$
$\implies \frac{4}{3}C = 1$
$\implies C = \frac{3}{4}$

$\therefore$ We get $C = \frac{3}{4}$.

b)
$f_X(x) = \int_0^1 \frac{3}{4}(x^2+2y)dy$
$= \frac{3}{4}\int_0^1 (x^2+2y)dy$
$= \frac{3}{4}(x^2y + y^2)|_0^1$
$= \frac{3}{4}(x^2+1)$
$= \frac{3}{4}x^2 + \frac{3}{4}$
$= \begin{cases} \frac{3}{4}x^2 + \frac{3}{4}, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$

$f_Y(y) = \int_0^1 \frac{3}{4}(x^2+2y)dx$
$= \frac{3}{4}\int_0^1 (x^2+2y)dx$
$= \frac{3}{4}(\frac{1}{3}x^3 + 2yx)|_0^1$
$= \frac{3}{4}(\frac{1}{3} + 2y)$
$= \frac{1}{4} + \frac{3}{2}y$
$= \begin{cases} \frac{1}{4} + \frac{3}{2}y, & 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$

c)

We know when $x < 0$ and $y < 0$, f(x, y) = 0, so $F_{XY} = 0$.

We also know that when $x \geq 1$ and $y \geq 1$, $F_{XY} = 1$.

For $0 \leq x \leq 1$ and $0 \leq y \leq 1$, we need to calculate the following:

$F_{XY}(x, y) = \int_0^y \int_0^x \frac{3}{4}(x^2 + 2y)dxdy$
$= \frac{3}{4}\int_0^y \int_0^x (x^2 + 2y)dxdy$
$= \frac{3}{4}\int_0^y (\frac{1}{3}x^3 + 2xy)|_0^x dy$
$= \frac{3}{4}\int_0^y (\frac{1}{3}x^3 + 2xy)dy$
$= \frac{3}{4}(\frac{1}{3}x^3 y + xy^2)|_0^y$
$= \frac{1}{4}x^3 y + \frac{3}{4}xy^2$

Now, when $0 \leq x \leq 1$ and $y \geq 1$, we get $F_{XY}(x, y) = F_{XY}(x, 1) = \frac{1}{4}x^3 + \frac{3}{4}x$.

Additionally, when $x \geq 1$ and $0 \leq y \leq 1$, we get $F_{XY}(x, y) = F_{XY}(1, y) = \frac{1}{4}y + \frac{3}{4}y^2$.

So, the joint cdf is as follows:

$$F_{XY}(x, y) = \begin{cases} 0, & x < 0 \text{ or } y < 0, \\ \frac{1}{4}x^3 y + \frac{3}{4}xy^2, & 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1, \\ \frac{1}{4}y + \frac{3}{4}y^2, & x \geq 1 \text{ and } 0 \leq y \leq 1, \\ \frac{1}{4}x^3 + \frac{3}{4}x, & 0 \leq x \leq 1 \text{ and } y \geq 1, \\ 1, & x \geq 1 \text{ and } x \geq 1, \end{cases}$$

■

5. [**Graduate Students Only**] A 2-D Gaussian distribution is defined as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

Compute the following integral:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(x, y)\ (5x^2y^2 + 3xy + 1)\ dx\ dy$$

Hint: Think in terms of the properties of probability distribution functions. [5 pts]

Solution.
$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma^2} \exp(-\frac{x^2+y^2}{2\sigma^2})(5x^2y^2 + 3xy + 1)dxdy$
$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{x^2}{2\sigma^2}) \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{y^2}{2\sigma^2})(5x^2y^2 + 3xy + 1)dxdy$
$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{y^2}{2\sigma^2}) \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{x^2}{2\sigma^2})(5x^2y^2 + 3xy + 1)dxdy$

Let $G(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{x^2}{2\sigma^2})$ and $G(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{y^2}{2\sigma^2})$ represent the marginal distributions of X and Y respectively, with both being Gaussians:

$= \int_{-\infty}^{\infty} G(y) \int_{-\infty}^{\infty} G(x)(5x^2y^2 + 3xy + 1)dxdy$
$= \int_{-\infty}^{\infty} G(y)\big[\int_{-\infty}^{\infty} G(x)5x^2y^2dx + \int_{-\infty}^{\infty} G(x)3xydx + \int_{-\infty}^{\infty} G(x)dx\big]dy$
$= \int_{-\infty}^{\infty} G(y)\big[5y^2 \int_{-\infty}^{\infty} x^2G(x)dx + 3y \int_{-\infty}^{\infty} xG(x)dx + \int_{-\infty}^{\infty} G(x)dx\big]dy$

Now, $\mu = E[X] = \int_{-\infty}^{\infty} xf(x)dx$ and $\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$ by definition of the mean (μ) and variance (σ^2). Additionally, the mean of both Gaussians $G(x)$ and $G(y)$ are 0 by simple comparison of them and the univariate Gaussian equation.

$= \int_{-\infty}^{\infty} G(y)\big[5y^2\sigma^2 + 3y \int_{-\infty}^{\infty} xG(x)dx + \int_{-\infty}^{\infty} G(x)dx\big]dy$ (definition of variance)
$= \int_{-\infty}^{\infty} G(y)\big[5y^2\sigma^2 + 3y\mu + \int_{-\infty}^{\infty} G(x)dx\big]dy$ (definition of mean)
$= \int_{-\infty}^{\infty} G(y)\big[5y^2\sigma^2 + 3y\mu + 1\big]dy$
$= \int_{-\infty}^{\infty} G(y)5y^2\sigma^2dy + \int_{-\infty}^{\infty} G(y)3y\mu dy + \int_{-\infty}^{\infty} G(y)dy$
$= 5\sigma^2 \int_{-\infty}^{\infty} y^2G(y)dy + 3\mu \int_{-\infty}^{\infty} yG(y)dy + \int_{-\infty}^{\infty} G(y)dy$
$= 5\sigma^4 + 3\mu \int_{-\infty}^{\infty} yG(y)dy + \int_{-\infty}^{\infty} G(y)dy$ (definition of variance)
$= 5\sigma^4 + 3\mu^2 + \int_{-\infty}^{\infty} G(y)dy$ (definition of mean)
$= 5\sigma^4 + 3\mu^2 + 1$
$= 5\sigma^4 + 1$ ($\mu = 0$ as discussed previously)

So, the value of the integral is $5\sigma^4 + 1$. ■