1. (20 points) Flaws in a certain type of fabric appear on the average of one in 200 square feet. Let $X$ be the number of flaws that appear in 500 square feet of drapery material. Assume that $X$ has the Poisson distribution.

(a) State the value of $\lambda$ for this Poisson distribution, and write down the probability mass function (PMF) for $X$.

(b) Find $P(X > 0)$.

\[ a) \quad X \sim \text{Poisson}(\lambda = \frac{500}{200} = \frac{5}{2}) \]

\[ P(X = n) = e^{-\lambda} \frac{\lambda^n}{n!} \]

\[ = e^{-5/2} \frac{(5/2)^n}{n!} \quad \text{for} \quad n = 0, 1, 2, \ldots \]

\[ b) \quad P(X > 0) = 1 - P(X = 0) \]

\[ = 1 - e^{-5/2} \frac{(5/2)^0}{0!} \]

\[ = 1 - e^{-5/2} \]
2. (20 points) Jointly distributed random variables $X$ and $Y$ have the following joint PMF:

$$f(x, y) = \frac{x + 2y}{C}, \quad x \in \{1, 2, 3\}, \ y \in \{1, 2\},$$

where $C$ is a constant.

(a) Find the value of $C$.

(b) Find the marginal distributions for $X$ and $Y$.

(c) Are $X$ and $Y$ independent random variables? Justify your answer.

\[
\begin{array}{ccc|c|c|c|c}
\text{ } & 1 & 2 & 3 & \sum \frac{x+2y}{C} = \frac{30}{C} \\
\hline
1 & 1/30 & 4/30 & 5/30 & 12/30 \\
2 & 5/30 & 6/30 & 7/30 & 18/30 \\
\hline
\end{array}
\]

\[
\begin{array}{ccc|c|c|c|c}
\text{Marginal for } X & 1 & 2 & 3 & \sum P(X=x) = 1 \\
\hline
7(x=x) & 8/30 & 10/30 & 12/30 \\
\hline
\end{array}
\]

\[
\begin{array}{ccc|c|c|c|c}
\text{Marginal for } Y & 1 & 2 & \sum P(Y=y) = 1 \\
\hline
P(Y=y) & 12/30 & 14/30 \\
\hline
\end{array}
\]

\[
\begin{array}{ccc|c}
\text{Note:} & P(X=1,Y=1) = 3/30 \\
\hline
P(X=1)P(Y=1) = 8 \cdot 12 = 30 \\
\end{array}
\]

\[
\begin{array}{cc}
c) \ X \neq Y \text{ are not independent, since } P(X=x,Y=y) \neq P(X=x)P(Y=y) \\
\end{array}
\]
3. (20 points) A continuous random variable $X$ has probability density function (PDF)

$$p(x) = \begin{cases} 
Cx(1-x) & \text{if } 0 \leq x \leq 1, \\
0 & \text{otherwise}
\end{cases}$$

where $C$ is a constant.

(a) Find the value of $C$.

(b) Find the expected value of $X$.

(c) Find $E[X^t]$, where $t$ is an arbitrary real number.

\[a) \quad 1 = \int_0^1 Cx(1-x) \, dx = C \left( \frac{1}{2}x^2 - \frac{1}{3}x^3 \right) \bigg|_0^1 = C \cdot \frac{1}{6} \]

So $C = 6$, and the PDF is $p(x) = 6x(1-x)$, $x \in [0,1]$

\[b) \quad E(X) = \int_0^1 x \cdot 6x(1-x) \, dx = 6 \left( \frac{1}{3}x^3 - \frac{1}{4}x^4 \right) \bigg|_0^1 = 6 \cdot \frac{1}{12} = \frac{1}{2} \]

\[c) \quad E(X^t) = \int_0^1 x^t \cdot 6x(1-x) \, dx \]

\[= \quad 6 \int_0^1 (x^{t+1} - x^{t+2}) \, dx \]

\[= \quad 6 \left( \frac{x^{t+2}}{t+2} - \frac{x^{t+3}}{t+3} \right) \bigg|_0^1 \]

\[= \quad 6 \left( \frac{1}{t+2} - \frac{1}{t+3} \right) \]
4. (20 points)
(a) A continuous random variable $X$ has probability density function

$$p(x) = \begin{cases} xe^{-x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the moment generating function (MGF) for $X$.

(b) A continuous random variable $Y$ has CDF

$$F(y) = \frac{1}{1 + e^{-y}}, \quad -\infty < y < \infty.$$ 

Find the probability density function for $Y$.

\[
\text{a) } M(t) = E(e^{tX}) = \int_0^\infty e^{tx} xe^{-x} \, dx \\
= \int_0^\infty x e^{x(t-1)} \, dx \\
= \left( x \frac{1}{t-1} e^{x(t-1)} - \frac{1}{(t-1)^2} e^{x(t-1)} \right) \bigg|_0^\infty \\
= \left( 0 - 0 \right) - \left( 0 - \frac{1}{(t-1)^2} \right) \quad \text{if } t < 1 \\
= \frac{1}{(t-1)^2} \quad \text{for } t < 1.
\]

\[
\text{b) } p(y) = F'(y) = \frac{d}{dy} \left( \frac{e^{-y}}{1 + e^{-y}} \right) = \frac{e^{-y}}{(1 + e^{-y})^2}
\]
5. (20 points) Suppose $X$ and $Y$ are jointly distributed discrete random variables, each taking values in \{0, 1\}, and having the joint probability mass function $f$ given by

\[
f(0, 0) = 1/2, \quad f(1, 0) = 1/6, \quad f(1, 1) = 1/3
\]

(a) Find the correlation between $X$ and $Y$.

(b) Find the equation of the least squares regression line for predicting $Y$ from $X$.

Note: $X \sim \text{Bern}(p = \frac{1}{2})$, $Y \sim \text{Bern}(p = \frac{1}{3})$

So

\[
M_X = 1/2, \quad \sigma_X^2 = \frac{1}{2} \cdot \frac{1}{2} = 1/4
\]

\[
M_Y = 1/3, \quad \sigma_Y^2 = \frac{1}{3} \cdot \frac{1}{3} = 2/9
\]

Also

\[
E(XY) = 1 \cdot 1 \cdot \frac{1}{3} = \frac{1}{3}
\]

So

\[
\text{Cov}(X, Y) = \frac{1}{3} - \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{6}
\]

and the correlation between $X$ and $Y$ is

\[
\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\frac{1}{6}}{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3}} = \frac{1}{2}
\]

\[
\text{So the least-squares regression line is}
\]

\[
y = \frac{1}{3} x
\]

\[
a = M_Y - b M_X = \frac{1}{3} - \frac{2}{3} \left( \frac{1}{2} \right) = 0
\]
6. (20 points) Let $X$ have the exponential distribution with mean $\theta$, and define a new random variable $Y$ by

$$Y = 1 - e^{-X/\theta}$$

Show that $Y$ has the continuous uniform distribution on $[0, 1]$. (Hint: Find $G(y) = P(Y \leq y) = P(1 - e^{-X/\theta} \leq y)$ for $0 < y < 1$, and interpret your result.)

Let $y \in (0, 1)$. Then

$$P(Y \leq y) = P(1 - e^{-X/\theta} \leq y)$$

$$= P(e^{-X/\theta} \geq 1 - y)$$

$$= P\left(-\frac{X}{\theta} \geq \ln(1 - y)\right)$$

$$= P\left(X \leq -\theta \ln(1 - y)\right)$$

Since the CDF of the $\text{exp} (\theta)$ distribution is

$$F(x) = P(X \leq x) = 1 - e^{-x/\theta}$$

we have

$$P(Y \leq y) = F\left(-\theta \ln(1 - y)\right)$$

$$= 1 - \exp\left(\frac{-\theta \ln(1 - y)}{\theta}\right)$$

$$= 1 - \exp\left(\ln(1 - y)\right) = 1 - (1 - y) = y$$

So, for $0 < y < 1$

$$P(Y \leq y) = y$$

which is the uniform $(0, 1)$ CDF.