MH1820 Introduction to Probability and Statistical Methods

Tutorial 4 (Week 5) $\boxed{\text{Solution}}$

Problem 1 (Poisson distribution)

(a) In the past, computer tape was used to store files, and flaws occured on these tapes. Suppose that flaws occur on the average of one flaw per 1200 feet. Let X be the number of flaws in a 4800-foot tape.

(i) What is the PMF of X?

(ii) What is the probability that $2 \le X \le 4$?

(b) Let X have a Poisson distribution such that $3\mathbb{P}(X = 1) = \mathbb{P}(X = 2)$. Find $\mathbb{P}(X = 4)$.

$\boxed{\textbf{Solution}}$ (a) The expected number of flaws in a 4800-foot tape is $4 \times 1 = 4$, that is $\lambda = \mathbb{E}[X] = 4$.

(i) The PMF of X is given by $p(x) = \dfrac{e^{-4}4^x}{x!}, \quad x = 0, 1, 2, \ldots$.

(ii)

$$\begin{aligned}
\mathbb{P}(2 \le X \le 4) &= \sum_{i=2}^{4} \mathbb{P}(X = i) \\
&= \frac{e^{-4}4^2}{2!} + \frac{e^{-4}4^3}{3!} + \frac{e^{-4}4^4}{4!} \\
&= 0.5373.
\end{aligned}$$

(b) Suppose $X \sim Poisson(\lambda)$. From the equation given, we have

$$3\frac{e^{-\lambda}\lambda^1}{1!} = \frac{e^{-\lambda}\lambda^2}{2!} \Longrightarrow 6\lambda - \lambda^2 = 0 \Longrightarrow \lambda(6 - \lambda) = 0 \Longrightarrow \lambda = 6.$$

So

$$\mathbb{P}(X = 4) = \frac{e^{-6}6^4}{4!} = 0.1339.$$

$\square$

Problem 2 (Continuous random variables, PDF and CDF)

Let X be a random variable with PDF given by

$$f(x) = \frac{x^2}{9} \text{ for } 0 \le x \le 3 \text{ and } f(x) = 0 \text{ otherwise.}$$

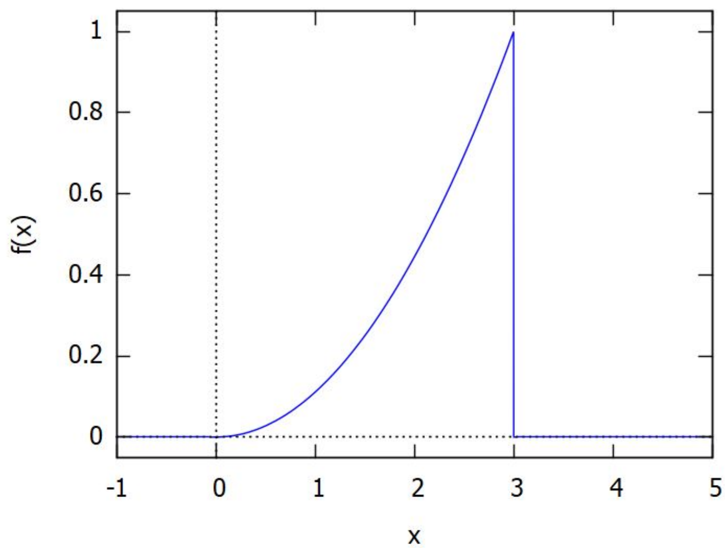
(a) Draw a graph of f.

(b) Compute the CDF F of X.

(c) Compute $\mathbb{E}[X]$, $\mathbb{E}[X^2]$, and $\mathrm{Var}[X]$.

(d) Find $\mathbb{P}(X \geq 1)$, $\mathbb{P}(2 \leq X \leq 3)$, and $\mathbb{P}(X^2 \leq 2)$.

Solution (a) Plot of the PDF $f(x)$:



(b) Let $F(x)$ denote the CDF of X. For $x < 0$, we have $F(x) = 0$. For $0 \leq x \leq 3$, we get

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)dt = \int_0^x \frac{t^2}{9} dt \\ &= \left[\frac{t^3}{27}\right]_0^x = \frac{x^3}{27}. \end{aligned}$$

Since $F(3) = 1$, we have $F(x) = 1$ for $x \geq 3$. In summary,

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^3}{27}, & 0 \leq x \leq 3 \\ 1, & x > 3. \end{cases}$$

(c) We find

$$\begin{aligned} \mathbb{E}[X] &= \int_0^3 x f(x)\, dx = \int_0^3 \frac{x^3}{9} dx = \left[\frac{x^4}{36}\right]_0^3 = \frac{9}{4}, \\ \mathbb{E}[X^2] &\int_0^3 x^2 f(x)\, dx = \int_0^3 \frac{x^4}{9} dx = \left[\frac{x^5}{45}\right]_0^3 = \frac{27}{5}, \\ \mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{27}{5} - \left(\frac{9}{4}\right)^2 = \frac{27}{80}. \end{aligned}$$

(d) Using the formula

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$$

we compute

$$\mathbb{P}(X \geq 1) = \int_1^\infty f(x)dx = \int_1^3 \frac{x^2}{9}dx = \left[\frac{x^3}{27}\right]_1^3 = \frac{26}{27},$$

$$\mathbb{P}(2 \leq X \leq 3) = \int_2^3 \frac{x^2}{9}dx = \left[\frac{x^3}{27}\right]_2^3 = \frac{19}{27},$$

$$\mathbb{P}(X^2 \leq 2) = \mathbb{P}(-\sqrt{2} \leq X \leq \sqrt{2}) = \int_{-\sqrt{2}}^{\sqrt{2}} f(x)\,dx = \int_0^{\sqrt{2}} \frac{x^2}{9}dx = \left[\frac{x^3}{27}\right]_0^{\sqrt{2}} = \frac{2\sqrt{2}}{27}.$$

Alternatively, we can use the CDF of X (was computed in part (b)) to find these probabilities. For instance,

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X \leq 1) = 1 - F(1) = 1 - \frac{1}{27} = \frac{26}{27}.$$

□

Problem 3 (Continuous random variables, PDF and CDF)
Suppose X is a random variable whose PDF is defined by

$$f(x) = \begin{cases} x, & 0 \leq x < 1, \\ \frac{c}{x^3}, & 1 \leq x < \infty, \\ 0, & \text{elsewhere.} \end{cases}$$

Find (i) the value of the constant c so that $f(x)$ is a PDF and (ii) the mean of X.

Solution (i) For $f(x)$ to be a PDF, we must have $\int_{-\infty}^{\infty} f(x)\,dx = 1$. Thus,

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= 1 \\ \int_0^1 x\,dx + \int_1^\infty cx^{-3}\,dx &= 1 \\ \left[\frac{x^2}{2}\right]_0^1 + \left[\frac{cx^{-2}}{-2}\right]_1^\infty &= 1 \\ \frac{1}{2} + \frac{c}{2} &= 1 \\ c &= 1. \end{aligned}$$

(ii) The mean is

$$
\begin{aligned}
\mathbb{E}[x] &= \int_{\infty}^{\infty} x f(x)\, dx \\
&= \int_0^1 x \cdot x\, dx + \int_1^{\infty} x \cdot \frac{1}{x^3}\, dx \\
&= \left[\frac{x^3}{3}\right]_0^1 + \left[-x^{-1}\right]_1^{\infty} \\
&= \frac{1}{3} + 1 = \frac{4}{3}.
\end{aligned}
$$

□

Problem 4 (Continuous random variables, expected value) A trader receives a bonus if the total losses X (in \$100,000) of his investment is less than 1.5. The bonus equals $0.05 - \frac{X}{30}$ if $X < 1.5$, and equals 0 otherwise. Suppose X has PDF $f(x) = \frac{3}{x^4}$ for $x > 1$ and $f(x) = 0$ for $x \leq 1$. Find the expected value of the bonus.

Hint: $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$, where $f(x)$ is the PDF of X.

Solution The bonus is given by $g(X) = 0.05 - \frac{x}{30}$ for $x < 1.5$ and 0 otherwise. Hence,

$$
\begin{aligned}
\mathbb{E}[g(X)] &= \int_{-\infty}^{\infty} g(x) f(x)\, dx \\
&= \int_{-\infty}^{1.5} \left(0.05 - \frac{x}{30}\right) f(x)\, dx + \int_{1.5}^{\infty} 0 \cdot f(x)\, dx \\
&= \int_{-\infty}^{1} \left(0.05 - \frac{x}{30}\right) \cdot 0\, dx + \int_{1}^{1.5} \left(0.05 - \frac{x}{30}\right) \cdot \frac{3}{x^4}\, dx \\
&= \int_{1}^{1.5} \left(0.05 - \frac{x}{30}\right) \cdot \frac{3}{x^4}\, dx \\
&= \int_{1}^{1.5} \left(0.15x^{-4} - 0.1x^{-3}\right)\, dx \\
&= [0.15x^{-3}/(-3) - 0.1x^{-2}/(-2)]_1^{1.5} \\
&= 0.0074074.
\end{aligned}
$$

So the expected bonus is $0.0074074 \times \$100,000 = \$740.74$. □

Problem 5 (Exponential distribution)

(a) A certain type of aluminium screen has, on the average, 3 flaws in a 100-foot roll.

  (i) What is the probability that the first 40 feet in a roll contain no flaws?

  (ii) What assumption did you make to solve part (i)?

(b) Cars arrive at a tollbooth at a mean rate of 5 cars every 10 minutes according to a Poisson distribution. Let X be the time in minutes that the toll collector will have to wait before collecting the first toll. Find $\mathbb{E}[X]$ and the standard deviation of X.

$\boxed{\textbf{Solution}}$ (a) (i) The average flaws per foot is $\lambda = \frac{3}{100} = 0.03$. Let X be the length (in feet) of aluminimum screen rolled out before the first flaw is observed. Assume that $X \sim Exp(\theta = \frac{1}{\lambda} = \frac{100}{3})$. Then

$$\begin{aligned}\mathbb{P}(X > 40) &= \int_{40}^{\infty} 0.03e^{-0.03x}\,dx \\ &= 0.03\left[\frac{e^{-0.03x}}{-0.03}\right]_{40}^{\infty} \\ &= e^{-0.03\times 40} = 0.3012.\end{aligned}$$

(a) (ii) We have assumed that X follows an exponential distribution with $\theta = \frac{100}{3}$.
(b) The cars arrive at the tollbooth on an average of $\lambda = \frac{5}{10} = \frac{1}{2}$ cars per minute. Let X be the waiting time (in minutes) until the first toll. Then $X \sim Exp(\theta = 2)$. Hence,

$$\mathbb{E}[X] = \theta = 2.$$

$$\text{Var}[X] = \theta^2 = 4 \Longrightarrow \text{standard deviation} = \sqrt{4} = 2.$$

□

Problem 6 (Uniform distribution)

(a) Let X be uniformly distributed over the interval $[a, b]$, i.e. $X \sim U(a, b)$. Find $\mathbb{E}[X]$.

(b) Let X be uniformly distributed over the interval $[0, 1]$. Let $Y = X^2$. Find the CDF and hence the PDF of Y.

  Hint: Differentiating CDF will give the PDF.

$\boxed{\textbf{Solution}}$ (a) The PDF of X is $f(x) = \frac{1}{b-a}$ for $a \leq x \leq b$ and $f(x) = 0$ elsewhere. Hence,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\,dx = \int_a^b \frac{x}{b-a}\,dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.$$

(b) Let $F(y)$ be the CDF of $Y = X^2$. Note that $F(y) = 0$ if $y < 0$ since $Y = X^2$ cannot take onnegative value. Let $0 \leq y \leq 1$.

$$
\begin{aligned}
F(y) &= \mathbb{P}(Y \le y) \\
&= \mathbb{P}(X^2 \le y) \\
&= \mathbb{P}(-\sqrt{y} \le X \le \sqrt{y}) \\
&= \int_{-\sqrt{y}}^{\sqrt{y}} 1\, dx \\
&= \int_{0}^{\sqrt{y}} 1\, dx \quad \text{since } X \ge 0 \\
&= \sqrt{y}.
\end{aligned}
$$

Hence, the CDF of Y is given by

$$
F(y) = \begin{cases} 0, & y \le 0 \\ \sqrt{y}, & 0 < y \le 1 \\ 1, & y > 1. \end{cases}
$$

Differentiating $F(y)$ with respect to y, we will get the PDF $f(y)$ of Y as follows:

$$
f(y) = \begin{cases} 0, & y \le 0 \\ \frac{1}{2\sqrt{y}}, & 0 < y \le 1 \\ 0, & y > 1. \end{cases}
$$

□

Problem 7 (Poisson distribution) A store selling newspapers orders only $n = 4$ of a certain newspaper because the manager does not get many requests for that publication. If the number of requests per day follows a Poisson distribution with mean 3,

(i) What is the expected value of the number sold per day?

(ii) What is the minimum number that the manager should order so that the chance of having more requests than available newspapers is less than 0.05?

You may use the CDF table for Poisson distribution attached at the end of this worksheet for calculations.

Hint: Let X be the number of requests per day, and Y be the number of newspapers sold per day. For part (i), we are interested in finding $\mathbb{E}[Y]$. How can we get the PMF for Y based on what we know about X? Remember that the store only has 4 newspapers.

Solution

Let Y be the number of the newspapers sold per day. Let X be the number of requests per day. It is given that $X \sim Poisson(3)$. Then

$$
\mathbb{P}(Y = 0) = \mathbb{P}(X = 0) = \frac{e^{-3}3^0}{0!} = 0.05,
$$

$$\mathbb{P}(Y=1) = \mathbb{P}(X=1) = \frac{e^{-3}3^1}{1!} = 0.149,$$
$$\mathbb{P}(Y=2) = \mathbb{P}(X=2) = \frac{e^{-3}3^2}{2!} = 0.224,$$
$$\mathbb{P}(Y=3) = \mathbb{P}(X=3) = \frac{e^{-3}3^3}{3!} = 0.224,$$
$$\mathbb{P}(Y=4) = \mathbb{P}(X \geq 4) = 1 - \mathbb{P}(X \leq 3) = 1 - 0.647 = 0.353.$$

The reason $\mathbb{P}(Y = 4) = \mathbb{P}(X \geq 4)$ is because the store only ordered 4 news papers, so it can only sell 4 of them whenever the number of request is 4 or more. Hence, the PMF of Y is as follows:

y	0	1	2	3	4
$p(y)$	0.05	0.149	0.224	0.224	0.353

(i)
$$\mathbb{E}[Y] = 0(0.05) + 1(0.149) + 2(0.224) + 3(0.224) + 4(0.353) = 2.681.$$

(ii) Let N be the number of the newspapers ordered. We are interested in the probability $\mathbb{P}(X > N) = 1 - \mathbb{P}(X \leq N)$ for different N. Indeed, according to the question, we want this probability to be less than 0.05. From the CDF table for Poisson distribution, we have

N	$\mathbb{P}(X > N) = 1 - \mathbb{P}(X \leq N)$
0	$1 - 0.05 = 0.95$
1	$1 - 0.199 = 0.801$
2	$1 - 0.423 = 0.577$
3	$1 - 0.647 = 0.353$
4	$1 - 0.815 = 0.185$
5	$1 - 0.916 = 0.084$
6	$1 - 0.966 = 0.034$

Notice that $\mathbb{P}(X > 6) < 0.05$ for the first time as N increases. So the minimum number of newspapers to be ordered is 6 so that the chance of having more requests than the newspapers ordered (i.e. $X > N$) is less than 0.05.

$\square$

Problem 8 (Monty Hall Problem - Revisited)
There are 3 doors, behind which are two goats and a car. A contestant picked a door (call it Door 1). The contestant is hoping for the car, of course. Monty Hall, the game show host, examines the other doors (Door 2 & 3) and opens one with a goat. (If both doors have goats, he picks randomly.)
Let H be the event that the car is behind Door 1. Let E be the event that the game host opens Door 3 with a goat after Door 1 has been picked.

(i) Calculate the probability $\mathbb{P}(H)$, and conditional probabilities $\mathbb{P}(E|H)$ and $\mathbb{P}(E|\overline{H})$.

(ii) Calculate the conditional probabilities $\mathbb{P}(H|E)$ and $\mathbb{P}(\overline{H}|E)$.

(iii) Suppose that the game host opens Door 3 with a goat. Explain why the contestant would have a higher chance of winning the car if s/he switches to Door 2 from Door 1.

Solution

(i) The event H does not depend on any prior information/conditions. The car is equally likely to be behind a door. So $\mathbb{P}(H) = \frac{1}{3}$.

Suppose that the car is behind Door 1 (i.e. assuming that H had occured). Since Door 1 has been picked by the contestatnt, there is a 50-50 chance that Door 3 will be picked by the game host. So $\mathbb{P}(E|H) = 0.5$.

Suppose we know that the car is not behind Door 1 (i.e. the event $\overline{H}$ had occured). There is a 50-50 chance that the car is behind Door 2 (and Door 3 respectively). The game host will open Door 3 only if the car is behind Door 2. So again $\mathbb{P}(E|\overline{H}) = 0.5$.

(ii) By Bayes' theorem, we have

$$\begin{aligned}
\mathbb{P}(H|E) &= \frac{\mathbb{P}(E|H)\mathbb{P}(H)}{\mathbb{P}(E)} \\
&= \frac{\mathbb{P}(E|H)\mathbb{P}(H)}{\mathbb{P}(E|H)\mathbb{P}(H) + \mathbb{P}(E|\overline{H})\mathbb{P}(\overline{H})} \\
&= \frac{0.5(\frac{1}{3})}{0.5(\frac{1}{3}) + 0.5(\frac{2}{3})} \\
&= \frac{1}{3}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}(\overline{H}|E) &= 1 - \mathbb{P}(H|E) \\
&= 1 - \frac{1}{3} = \frac{2}{3}.
\end{aligned}$$

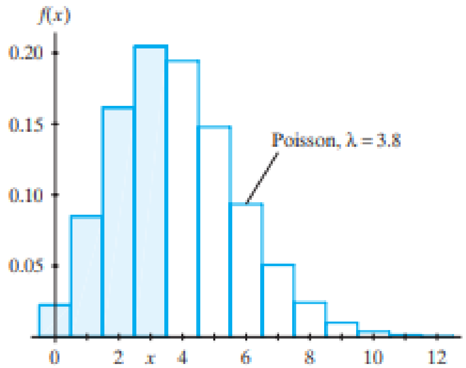
(iii) Suppose that the game host opens Door 3 to show a goat (which can only happen after Door 1 has been picked). Then the probability that the car is behind Door 1 is $\mathbb{P}(H|E) = \frac{1}{3}$, which is the probability of winning if the contestant does not switch door from Door 1.

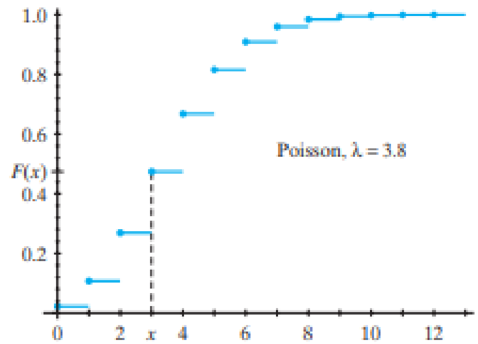
On the other hand, the probability that the car is not behind Door 1 (and hence must be behind Door 2) is $\mathbb{P}(\overline{H}|E) = \frac{2}{3}$, which is the probability of winning if the contestant switches to Door 2.

□

Answer Keys. 1(a)(ii). 0.5373 1(b). 0.1339 2(c). 9/4, 27/5, 27/80 2(d). 26/27, 19/27, $2\sqrt{2}/27$ 3(i). 1 3(ii). 4/3 4. \$740.74 5(a)(i). 0.3012 5(b). 2, 2 6(a). $(a+b)/2$ 7(i). 2.681 7(ii). 6

Table III The Poisson Distribution





$$F(x) = P(X \le x) = \sum_{k=0}^{x} \frac{\lambda^k e^{-\lambda}}{k!}$$

$\lambda = E(X)$

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.905	0.819	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	0.995	0.982	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2	1.000	0.999	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.996
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	0.333	0.301	0.273	0.247	0.223	0.202	0.183	0.165	0.150	0.135
1	0.699	0.663	0.627	0.592	0.558	0.525	0.493	0.463	0.434	0.406
2	0.900	0.879	0.857	0.833	0.809	0.783	0.757	0.731	0.704	0.677
3	0.974	0.966	0.957	0.946	0.934	0.921	0.907	0.891	0.875	0.857
4	0.995	0.992	0.989	0.986	0.981	0.976	0.970	0.964	0.956	0.947
5	0.999	0.998	0.998	0.997	0.996	0.994	0.992	0.990	0.987	0.983
6	1.000	1.000	1.000	0.999	0.999	0.999	0.998	0.997	0.997	0.995
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

x	2.2	2.4	2.6	2.8	3.0	3.2	3.4	3.6	3.8	4.0
0	0.111	0.091	0.074	0.061	0.050	0.041	0.033	0.027	0.022	0.018
1	0.355	0.308	0.267	0.231	0.199	0.171	0.147	0.126	0.107	0.092
2	0.623	0.570	0.518	0.469	0.423	0.380	0.340	0.303	0.269	0.238
3	0.819	0.779	0.736	0.692	0.647	0.603	0.558	0.515	0.473	0.433
4	0.928	0.904	0.877	0.848	0.815	0.781	0.744	0.706	0.668	0.629
5	0.975	0.964	0.951	0.935	0.916	0.895	0.871	0.844	0.816	0.785
6	0.993	0.988	0.983	0.976	0.966	0.955	0.942	0.927	0.909	0.889
7	0.998	0.997	0.995	0.992	0.988	0.983	0.977	0.969	0.960	0.949
8	1.000	0.999	0.999	0.998	0.996	0.994	0.992	0.988	0.984	0.979
9	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.996	0.994	0.992
10	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000