MAST20006 Probability for Statistics /MAST90057 Elements of Probability

Assignment 3, Semester 1 2020

Due time: 4pm, Friday May 8.

- ## New completion process

  Note this assignment is being handled using a similar process to that now planned for the final exam so you can start to become familiar with it.

  To complete this assignment, you need to write your solutions into the blank answer spaces following each question in this assignment PDF.

  If you have a printer (or can access one), then you must print out the assignment template and handwrite your solutions into the answer spaces.

  If you do not have a printer (**NB you should have one for the exam**), but you can figure out how to annotate a PDF using an iPad/Android tablet/Graphics tablet or using Adobe Acrobat, then annotate your answers directly onto the assignment PDF and save a copy for submission.

  Failing both of these methods, you may handwrite your answers as normal on blank paper and then scan for submission (but note that you will thereby miss valuable practice for the exam process).

  Scan your assignment to a PDF file using your mobile phone then upload by going to the Assignments menu on Canvas and submit the PDF to the GradeScope tool by first selecting your PDF file and then clicking on 'Upload PDF'.

- The **strict** submission deadline is **4pm Melbourne time on Friday 8 May**. You have longer than of the normal one week to complete this assignment. Consequently late assignments will **NOT** be accepted. We recommend you submit at least a day before the due date to avoid any technical delays. If there are extenuating, eg medical, circumstances, contact the Tutorial Coordinator.

- There are 9 questions, of which 2 randomly chosen questions will be marked. Note you are expected to submit answers to **all** questions, otherwise **a mark penalty will apply**.

- Working and reasoning **must** be given to obtain full credit. Give clear and concise explanations. Clarity, neatness, and style count.

1. The number of breakdowns of a computer network follows a Poisson process with rate $\alpha = 0.2$ breakdowns per week. This means the number of breakdowns during a period of t weeks is a Poisson random variable with parameter $\lambda = 0.2t$.

   (a) What is the probability that exactly 3 breakdowns are to occur during a 10-week period?

Solution:

Let X be the number of breakdowns in 10 weeks. Then $X \overset{d}{=} Poi(2)$. Hence,

$$P(X = 3) = \frac{2^3 e^{-2}}{3!} = 0.180447.$$

   (b) What is the probability that at least 2 breakdowns are to occur in next 10 weeks?

Solution:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \frac{2^0 e^{-2}}{0!} - \frac{2^1 e^{-2}}{1!} \\ &= 1 - 3e^{-2} \\ &= 0.5939942. \end{aligned}$$

   (c) How many breakdowns are expected to occur in 52 weeks?

Solution:

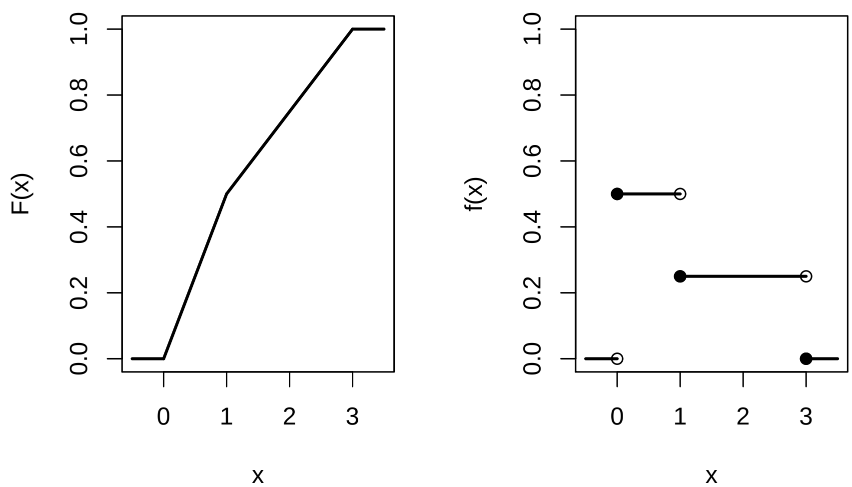
We expect $0.2 \times 52 = 10.4$ breakdowns to occur.

2. A random variable X has a cdf given by

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}x, & 0 \le x < 1 \\ \frac{1}{4}x + \frac{1}{4}, & 1 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

(a) Sketch the graph of $F(x)$.

Solution:



(b) Find the pdf $f(x)$ of X and sketch its graph.

Solution:

$$f(x) = \begin{cases} \frac{1}{2}, & 0 \le x < 1, \\ \frac{1}{4}, & 1 \le x < 3. \end{cases}$$

(c) Find $P(X \leq \frac{1}{2})$.

Solution:

$$P(X \leq 0.5) = F(0.5) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

(d) Find $P(X > \frac{1}{3})$.

Solution:

$$P(X > 1/3) = 1 - F(1/3) = 1 - \frac{1}{2} \times \frac{1}{3} = \frac{5}{6}.$$

(e) Find $P(X \leq \frac{4}{3})$.

Solution:

$$P(X \leq 4/3) = F(4/3) = \frac{1}{4} \times \frac{4}{3} + \frac{1}{4} = \frac{7}{12}.$$

(f) What is $P(X = 1)$?

Solution:

$P(X = 1) = 0$ as X is a continuous r.v..

(g) Find $E(X)$, the expectation of X.

Solution:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x \cdot \frac{1}{2} dx + \int_1^3 x \cdot \frac{1}{4} dx \\ &= \frac{1}{4} + (\frac{1}{8} \cdot 3^2 - \frac{1}{8} \cdot 1^2) \\ &= \frac{5}{4}. \end{aligned}$$

(h) Find the 75th percentile of the distribution. Namely, find the value of $\pi_{0.75}$ so that $P(X \leq \pi_{0.75}) = F(\pi_{0.75}) = 0.75$.

Solution:

Solving $F(\pi_{0.75}) = 0.75$, which is $\frac{1}{4}x + \frac{1}{4} = 0.75$, we get $\pi_{0.75} = 2$ which is the 75th percentile.

(i) Find the conditional probability $P(X > \frac{5}{4} | X > \frac{3}{4})$.

Solution:

$$\begin{aligned} P(X > 5/4 | X > 3/4) &= \frac{1 - F(5/4)}{1 - F(3/4)} \\ &= \frac{1 - (\frac{1}{4} \times \frac{5}{4} + \frac{1}{4})}{1 - \frac{1}{2} \times \frac{3}{4}} \\ &= \frac{7}{10}. \end{aligned}$$

3. A random variable X has a pdf given by

$$f(x) = \begin{cases} x, & 0 \le x < 1, \\ 2 - x, & 1 \le x < 2. \end{cases}$$

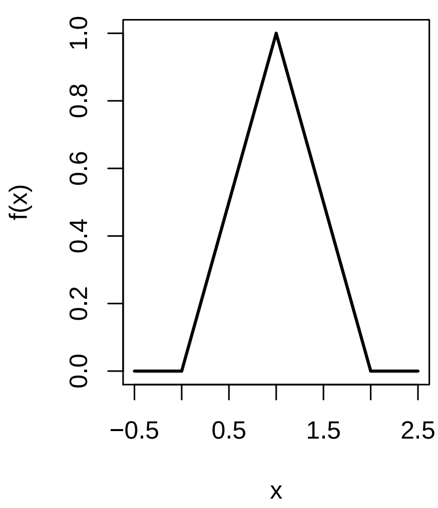
(a) Find the cdf $F(x)$ of X.

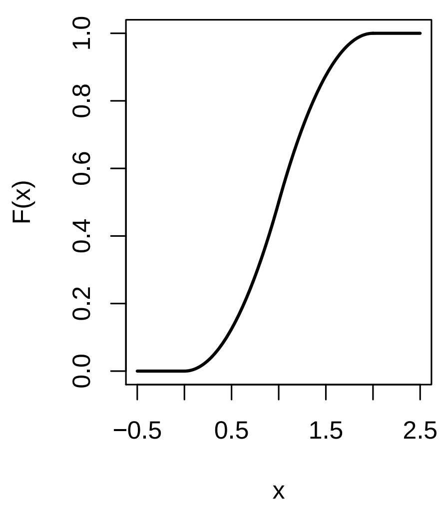
Solution:

$$F(x) = \begin{cases} 0, & x < 0, \\ \int_0^x t dt = \frac{1}{2}x^2, & 0 \le x < 1, \\ \frac{1}{2} + \int_1^x (2-t)dt = 2x - \frac{1}{2}x^2 - 1, & 1 \le x < 2, \\ 1, & x \ge 2. \end{cases}$$

(b) Sketch the graphs of $f(x)$ and $F(x)$.

Solution:





(c) Find $E(X)$, the expectation of X.

Solution:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x \cdot x dx + \int_1^2 x(2-x) dx \\ &= \frac{1}{3} + (3 - \frac{7}{3}) \\ &= 1. \end{aligned}$$

(d) Find $E(X^2)$, the second moment of X. Then find the variance σ^2 of X.

Solution:

$$\begin{aligned} E(X^2) &= \int_0^1 x^2 \cdot x dx + \int_1^2 x^2(2-x) dx \\ &= \frac{1}{4} + (\frac{14}{3} - \frac{15}{4}) = \frac{7}{6}. \end{aligned}$$

So $\mathrm{Var}(X) = \frac{7}{6} - 1^2 = \frac{1}{6}$.

(e) Find $M(t) = E(e^{tX})$, the mgf of X. Justify why it is well defined for any real finite value of t.

Solution:

$$\begin{aligned} M(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_0^1 e^{tx} x dx + \int_1^2 e^{tx}(2-x)dx \\ &= \left(\frac{1}{t} x e^{tx}|_0^1 - \int_0^1 \frac{1}{t} e^{tx} dt\right) \\ &\quad + \left(\frac{1}{t}(2-x)e^{tx}|_1^2 + \int_1^2 \frac{1}{t} e^{tx} dt\right) \\ &= \left(\frac{e^t}{t} - \frac{e^t - 1}{t^2}\right) + \left(-\frac{e^t}{t} + \frac{e^{2t} - e^t}{t^2}\right) \\ &= \left(\frac{e^t - 1}{t}\right)^2, \quad -\infty < t < \infty. \end{aligned}$$

Using the Maclaurin series expansion of e^t, we have

$$\frac{e^t - 1}{t} = (1/t)\left(\sum_{k\geq 0} \frac{t^k}{k!} - 1\right) = (1/t)\sum_{k\geq 1} \frac{t^k}{k!} = \sum_{k\geq 1} \frac{t^{k-1}}{k!},$$

so $M(t)$ can also be written as

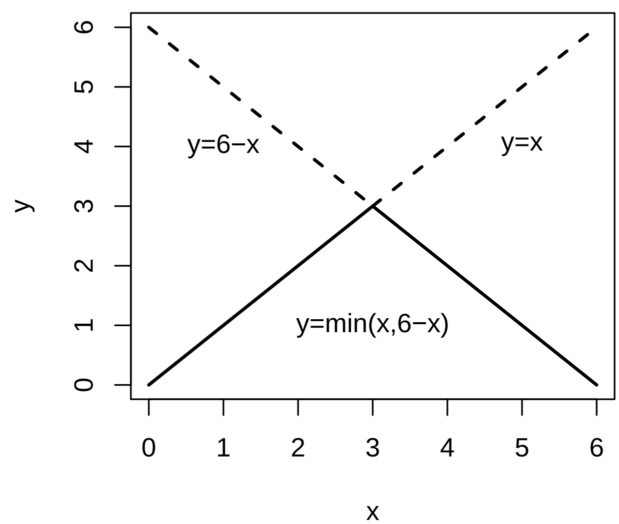
$$M(t) = \left(1 + \sum_{k\geq 1} \frac{t^k}{(k+1)!}\right)^2,$$

which is well defined for every (finite) value of t.

4. Suppose a value is chosen "at random" in the interval [0,6]. In other words, x is an observed value of a random variable $X \stackrel{d}{=} U(0,6)$. The random variable X divides the interval [0,6] into two subintervals, the lengths of which are X and $6-X$ respectively. Denote by $Y = \min(X, 6-X)$, the length of the shorter one of the two intervals.

   (a) Draw a picture which illustrates the value of Y as a function of the possible values of X

Solution:



(b) Describe the support of Y and find the probability $P(Y > y)$ for any given y.

Solution:

From the above plot we see the support of Y is $[0, 3]$.

So, when $0 \leq y \leq 3$,

$$\begin{aligned} P(Y > y) &= P(\min(X, 6 - X) > y) \\ &= (\{X > y\} \cap \{6 - X > y\}) \\ &= P(y < X < 6 - y\}) \\ &= \frac{6 - y - y}{6} \\ &= \frac{3 - y}{3} \end{aligned}$$

When $y < 0$, $P(Y > y) = 1$, and when $y > 3$, $P(Y > y) = 0$.

(c) Find both the cdf and the pdf of Y. Can you tell the name of the distribution of Y?

Solution:

The cdf of Y is

$$F_Y(y) = 1 - P(Y > y) = \begin{cases} 0, & y < 0 \\ \frac{y}{3}, & 0 \le y \le 3 \\ 1, & y > 3. \end{cases}$$

The pdf of Y is $f_Y(y) = F_Y'(y) = \frac{1}{3}, 0 \le y \le 3$.

Therefore, $Y \stackrel{d}{=} U(0, 3)$.

5. The moment-generating function of a r.v. X is

$$M(t) = \frac{e^{8t} - e^{4t}}{4t}, \quad t \ne 0, \quad \text{and} \quad M(0) = 1.$$

(a) Find $E(X)$.

Solution:

The mgf $M(t)$ is the same as that of $U(4, 8)$ distribution.

So $X \stackrel{d}{=} U(4, 8)$.

Therefore

$$E(X) = \frac{4 + 8}{2} = 6.$$

(b) Find Var(X).

Solution:

$$\text{Var}(X) = \frac{(8-4)^2}{12} = \frac{4}{3}.$$

(c) Find $P(4.5 < X < 7.5)$.

Solution:

$$P(4.5 < X < 7.5) = \int_{4.5}^{7.5} \frac{1}{8-4} dx = 0.75.$$

6. (*Q3.3-8*) Telephone calls to the general office of a university department on a working day follows a Poisson process with the rate of occurrence being 2 every 30 minutes. Let X denote the waiting time (in minutes, counted from 10 am) until the first call that arrives.

   (a) What is the pdf of X? Name the pdf if it has a name.

Solution:

The rate of occurrence parameter $\lambda = \frac{2}{30}$.
So the mean parameter $\theta = \frac{1}{\lambda} = \frac{30}{2} = 15$.
X has an $exp(15)$ distribution with pdf $f(x) = \frac{1}{15}e^{-x/15}, \quad x > 0$.

   (b) Find $P(X > 10)$.

Solution:

$$P(X > 10) = \int_{10}^{\infty} \frac{1}{15} e^{-x/15} dx = e^{-2/3}.$$

(c) Find $P(X > 15|X > 0)$.

Solution:

$$
\begin{aligned}
P(X > 15|X > 0) &= \frac{P(\{X > 15\} \cap \{X > 0\})}{P(X > 0)} = \\
&= \frac{P(X > 15)}{P(X > 0)} \\
&= \frac{e^{-15/15}}{e^{-0/15}} \\
&= e^{-1}.
\end{aligned}
$$

7. The shelf life (in hours) of a certain perishable packaged food is a random variable whose pdf is given by

$$f(x) = \begin{cases} \frac{2\times 10^4}{(x+100)^3}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$$

   (a) Find the probability that the shelf life is at least 200 hours.

Solution:

$$\begin{aligned} P(X \geq 200) &= \int_{200}^{\infty} \frac{2\times 10^4}{(x+100)^3} dx \\ &= -\frac{10^4}{(x+100)^2}\Big|_{200}^{\infty} \\ &= \frac{10^4}{300^2} \\ &= \frac{1}{9} \end{aligned}$$

.

(b) Find the probability that the shelf life is between 80 and 120 hours.

Solution:

$$
\begin{aligned}
P(80 \le X \le 120) &= \int_{80}^{120} \frac{2 \times 10^4}{(x+100)^3} dx \\
&= -\frac{10^4}{(x+100)^2}\Big|_{80}^{120} \\
&= -\frac{10^4}{220^2} + \frac{10^4}{180^2} \\
&= 0.102.
\end{aligned}
$$

(c) Find the mean hours of the shelf life.

Solution:

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} x f(x) dx = \int_{0}^{\infty} \frac{2 \times 10^4 x}{(x+100)^3} dx \\
&= \int_{0}^{\infty} \left(\frac{2 \times 10^4}{(x+100)^2} - \frac{2 \times 10^6}{(x+100)^3} \right) dx \\
&= \left(-\frac{2 \times 10^4}{x+100} + \frac{10^6}{(x+100)^2} \right) \Big|_{0}^{\infty} \\
&= \frac{2 \times 10^4}{0+100} - \frac{10^6}{(0+100)^2} \\
&= 100.
\end{aligned}
$$

8. In a certain city, the daily consumption of electric power in millions of kilowatt hours can be treated as a random variable having a gamma distribution with $\alpha = 3$ and $\theta = 2$.

   If the power plant of this city has a daily capacity of 12 million kilowatt-hours, what is the probability that this power supply will be inadequate on any given day?

   Note that there is no need to do an integration to answer the question.

Solution:

It is known that probabilities involving gamma$(\alpha = 3, \theta = 2)$ may be computed from the probabilities involving Poisson$(\lambda = \frac{1}{2}x)$. That is,

$$\begin{aligned} P(X > 12) &= \sum_{k=0}^{2} \frac{(12/2)^k}{k!} e^{-12/2} \\ &= e^{-6} + 6e^{-6} + 18e^{-6} \\ &= 25e^{-6} \\ &= 0.06197. \end{aligned}$$

9. Cars arrive at a toll booth at a mean rate of 5 cars every 10 minutes according to a Poisson process. What is the probability that the toll collector will have to wait longer than 28.41 minutes before collecting the 10th toll?

   Use two different approaches to solve this question by using the relationship between the gamma, Poisson, and chi-square distributions.

Solution:

Let W be the waiting time. Note that the rate of occurrence is $\lambda = 5/10$ so $\theta = 2$. Thus

$$W \stackrel{d}{=} gamma(\alpha = 10, \theta = 2) = \chi^2(20),$$

and $P(W > 28.41) = 0.1$ by Maple, R or chi-square probability table (28.41 is the upper 10th percentile of a chi-square distribution with 20 degrees of freedom).

Using the relationship between gamma and Poisson distributions, and letting $Y \stackrel{d}{=} Poi(\lambda \times 28.41) = Poi(14.205)$, we also get

$$\begin{aligned} P(W > 28.41) &= P(Y \le 9) \\ &= 0.1 \end{aligned}$$

using Maple or R.