Probability $\&$ Statistics (MA2670), Term 1 - 2nd Tutorial

Dr. Dalia Chakrabarty; dalia.chakrabarty@brunel.ac.uk
Department of Mathematics, Brunel University London

- While waiting at a street corner, the time that is observed to elapse between noticing two consecutive smokers, is a r.v. X that is distributed according to the probability density function (pdf):

$$f_X(x) = \lambda x \exp(-x) \text{ for } x > 0,$$

$$f_X(x) = 0 \text{ otherwise .}$$

  1. Compute λ.
  2. Find the probablity distribution of X.
  3. What is the probability that an observer - who notices the 1st smoker - will see another smoker in 2 to 5 minutes? In at least 7 minutes?

- Consider the function

$$\begin{aligned} f_X(x) &= 1/2 - (1/4)|x-3| \quad 1 \leq x \leq 5 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

  1. Show that the function is a pdf of r.v. X.
  2. Find the cdf of X. Show that it is continuous.

- Consider the function

$$\begin{aligned} g(x) &= K(2/3)^x, \text{ for } x = 1, 2, 3, \ldots \\ &= 0 \text{ elsewhere .} \end{aligned}$$

  Can $g(x)$ be a probability mass function?

- Let r.v. X have a pdf:

$$\begin{aligned} f_X(x) &= 2/x^2 \quad 1 < x < 2 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

1. Find the cdf and pdf of the r.v. $Y = X^2$.

- Let measurements be taken of a variable. The error of this measurement is the r.v. X, that is known to have a pdf:

$$
\begin{aligned}
f_X(x) &= 1/2 \quad -1 < x < 1 \\
&= 0 \quad \text{otherwise}
\end{aligned}
$$

1. Find the cdf and pdf of the magnitude of this error r.v.

- 1. As $f_X(\cdot)$ is a pdf,
$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$
Then
$$\int_{-\infty}^{0} f_X(x)dx + \int_{0}^{\infty} f_X(x)dx = 1.$$
But for $x \in (-\infty, 0]$, $f_X(x) = 0$. Then using the form of the pdf over non-negative x, we get
$$\int_{0}^{\infty} \lambda x \exp(-x)dx = 1,$$
where λ is a constant. So, by integrating by parts we get
$$\lambda \int x \exp(-x)dx = \lambda[-x \exp(-x) - \exp(-x)],$$
on which if we insert the limits of the integration, we realise that
$$\lambda[-x \exp(-x) - \exp(-x)]\Big|_0^{\infty} = 1 \Longrightarrow \lambda[1] = 1.$$
Thus, $\lambda = 1$.

  2. The probability distribution of X or its cdf is:
$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{0}^{x} \lambda t \exp(-t)dt$$
since density at negative values of X is 0.
Then at
$$x > 0, \quad F_X(x) = \lambda[-t \exp(-t) - \exp(-t)]\Big|_0^{x} = 1 - x \exp(-x) - \exp(-x).$$
Again, at
$$x \leq 0, \quad F_X(x) = 0.$$

  3. $\Pr(2 < X < 5) = F_X(5-) - F_X(2)$. But 5 is not at the edge of (one of the two) a branch of the cdf of X, i.e. cdf is continuous at 5, i.e. $F_X(5-) = F_X(5)$. So $\Pr(2 < X < 5) = F_X(5) - F_X(2) = -5\exp(-5) - \exp(-5) + 2\exp(-2) + \exp(-2) \approx 0.37$.
Again, $\Pr(X \geq 7) = 1 - \Pr(X < 7) = 1 - F_X(7-) = 1 - F_X(7) =$. since cdf is continuous at 7. So $\Pr(X \geq 7) = (7 + 1)e^{-7} \approx 0.007$.

- 1. We start by considering the branch of the function $f_X(x)$ over $[1, 5]$. We realise that $|x - 3| = x - 3 \forall x \in [3, 5)$, and $|x - 3| = -x + 3 \forall x \in [1, 3)$. So the function is clarified as

$$\begin{aligned} f_X(x) = 0 \quad x < 1. & \\ &= 1/2 + (1/4)|x - 3| \quad 1 \leq x < 3 \\ &= 1/2 - (1/4)|x - 3| \quad 3 \leq x < 5 \\ &= 0 \quad x \geq 5 \end{aligned}$$

   We can see that over the 4 branches of this function, $f_X(x) \geq 0$ always.
   If you plot $f_X(x)$ across x you see that the function forms an isosceles triangle over $x \in [1, 5]$ s.t. mid-point of the base of this triangle is at $x = 3$ and height of the triangle is 1/2. Then area of this triangle is: $(1/2) \times (5 - 1) \times (1/2) = 1$. For $x \notin [1, 5]$, $f_X(x) = 0$. Thus, area under $f_X(x)$ bounded by the X-axis and $-\infty$ and ∞ is 1. In other words, integral of $f_X(x)$ over the real-line is 1.
   Thus $f_X(x)$ is a pdf.

  2. We compute the cdf of X over the respective branch of the pdf. Thus, for $x < 1$, $F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = 0$.
   Thus, for $1 \leq x < 3$, $F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{1}^{3} f_X(t)dt = \int\limits_{1}^{x} (1/2 + t/4 - 3/4)dt = -(1/4)(x - 1) + (x^2 - 1)/8 = -x/4 + (1/8) + (x^2)/8$.
   Thus, for $3 \leq x < 5$, $F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{1}^{5} f_X(t)dt = [-x/4 + (1/8) + (x^2)/8]|_{x=3} + \int\limits_{3}^{x} (1/2 - t/4 + 3/4)dt = 4/8 + (5/4)(x - 3) - (x^2 - 9)/8 = 5x/4 + (9 - 30 + 4/8) - (x^2)/8 = 5x/4 - (17/8) - (x^2)/8 =$.
   Thus, for $x \geq 5$, $F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{1}^{5} f_X(t)dt = [5x/4 - (17/8) - (x^2)/8]|_{x=5} + 0 = 8/8 = 1$.
   Now collate the cdf over the 4 branches of this function.

  3. We know that a cdf is by definition right-continuous. So to check for con-

tinuity, we check for left-continuity, and we do so at the edges of each interval that defines each branch of the cdf.
Now, $F_X(1-) = 0$ from the first branch, while $F_X(1) = -1/4 + 1/8 + 1/8 = 0$ from the 2nd branch. Thus, $F_X(1-) = F_X(1)$.
Now, $F_X(3-) = 4/8$ from the 2nd branch, while $F_X(3) = 30/8 - 17/8 - 9/8 = 1/2$ from the 3rd branch. Thus, $F_X(3-) = F_X(3)$.
Now, $F_X(5-) = 1$ from the 3rd branch, while $F_X(5) = 1$ from the 4th branch. Thus, $F_X(5-) = F_X(5)$.
Thus, the cdf is left-continuous as well.
Thus, the cdf is continuous.

- For $g(x)$ to be a pmf,
–$g(x) \geq 0 \; \forall x$,
–$\sum_{x \in \mathbb{N}} g(x) = 1$.
Here, $g(x) = K \times (2/3)^x$, for $x = 1, 2, ...$ Then $g(x) > 0$ for $K > 0$.
Also, $\sum_x K(2/3)^x = K \times (2/3)[1 + (2/3) + (2/3)^2 + \ldots]$. The infinite series is easily identified as a geometric series with common ratio 2/3.
We recall that sum of an infinite geometric series: $1 + r + r^2 + \ldots$, with common ratio $r < 1$, is $1/(1-r)$.
So here, we get $\sum_x K(2/3)^x = K \times (2/3)/(1/3)$. So if the sum is set to 1, $K = 1/2$, i.e. the given function is a pmf for $K = 1/2$.

- Let the cdf of $Y = X^2$, computed at the number y be $G_Y(y)$. Let pdf of Y be $g_Y(y)$. Then $G_Y(y) = \Pr(Y \leq y) = \Pr(X^2 \leq y)$.
Now, $X^2 \leq y \implies |X| \leq |\sqrt{y}$.
$\implies -\sqrt{y} \leq X \leq \sqrt{y}$.
Then
$$G_Y(y) = \Pr(-\sqrt{y} \leq X \leq \sqrt{y}).$$
–For $x < 1$, $f_X(x) = 0$ and $F_X(x) = 0$, i.e. for $y < 1$, $G_Y(y) = \Pr(-1 \leq X \leq 1) = 0$.
–For $1 \leq x < 2$, $f_X(x) = 2x^{-2}$,i.e. for $1 < y < 4$, $G_Y(y) = \Pr(-\sqrt{y} \leq X \leq \sqrt{y}) = \int\limits_{t=-\sqrt{y}}^{1} f_X(t)dt + \int\limits_{t=1}^{\sqrt{y}} 2t^{-2}dt = \int\limits_{t=1}^{\sqrt{y}} 2t^{-2}dt = (-2/t)|_1^{\sqrt{y}} = 2 - 2/\sqrt{y}$.
–For $x \geq 2$, i.e. for $y \geq 4$, $G_Y(y) = 1$.
Collate the last 3 branches of the function $G_Y(y)$ to give the cdf of Y.

So the pdf of Y is:

$$g_Y(y) = dG_Y(y)/dy = y^{-3/2} \text{ for } 1 \leq y < 4.$$

Also, for $y \notin [1,4)$, $g_Y(y) = 0$.
Collating the last 2 statements gives us the pdf of Y.

- By magnitude of X is implied $|X|$. So we first compute the cdf $G_Y(\cdot)$ of the r.v. $Y := |X|$. We consider the range of $(-\infty,0), [0,1), [1,\infty)$ for Y

  By definition $Y = |X| \geq 0$, i.e. there is zero probability for Y to be less than equal to any negative y, i.e. cdf of Y is 0 at $y < 0$.
  In other words, cdf $G_Y(y) = \Pr(|Y| \leq y) = 0 \forall y < 0$.
  When $-1 < X < 1$, $0 \leq |X| 1$. Then for any $x \in (-1,1)$, $G_Y(x) = \Pr(-x < X < x) = \int\limits_{t=-x}^{x} f_X(t)dt = (1/2)t|_{t=-x}^{x} = x$. Thus, for $y := |x|$, if $0 \leq y < 1$, then $G_Y(y) = y$.
  When $x \geq 1, y := |x| \geq 1$. Then $G_Y(y) = \Pr(-x < X < x) = \int\limits_{t=-\infty}^{-1} f_X(t)dt + \int\limits_{t=-1}^{1} f_X(t)dt + \int\limits_{t=1}^{\infty} f_X(t)dt = 0 + (1/2) \times t|_{t=-1}^{1} + 0 = 1$. Thus, $G_Y(y) = 1 \forall y \geq 1$.
  Thus the 3 branches of the cdf of $Y := |X|$ are as given in the 3 statements above:
  —$G_Y(y) = \Pr(|Y| \leq y) = 0 \, \forall \, y < 0$.
  —$G_Y(y) = y \, \forall \, y : 0 \leq y < 1$.
  —$G_Y(y) = 1 \, \forall \, y \geq 1$.

  The pdf of Y has no contribution from the intervals defined by the 1st and 3rd branches of the cdf, but for $0 \leq y < 1$, pdf $g_Y(y)$ is $dG_Y(y)/dy = 1$. Thus the pdf is:
  —$g_Y(y) = 1 \, \forall \, y \in [0,1]$; —$g_Y(y) = 0$ elsewhere.