Andre Guimaraes Duarte 07/18/16

MSAN 504: Homework 1

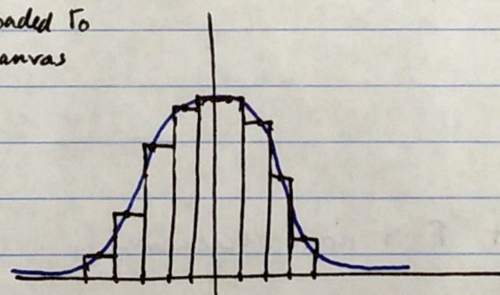
1) We have $X \sim Bin(3; 1/2)$. Therefore, we can write:

$$p(x) = \begin{cases} \binom{3}{0}\left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^3 & \text{if } x=0 \\ \binom{3}{1}\left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^2 & \text{if } x=1 \\ \binom{3}{2}\left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^1 & \text{if } x=2 \\ \binom{3}{3}\left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^0 & \text{if } x=3 \end{cases} \iff p(x) = \begin{cases} 1/8 & \text{if } x=0 \text{ or } x=3 \\ 3/8 & \text{if } x=1 \text{ or } x=2 \end{cases}$$

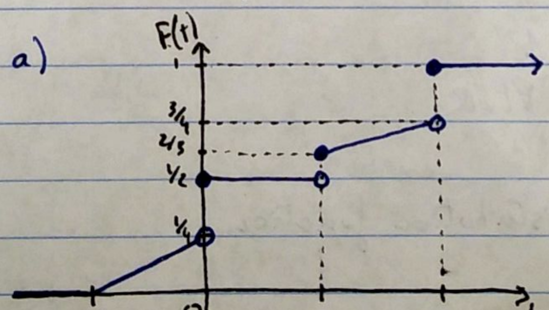
2) We use R to plot dnorm for values ranging from $x=-10$ to $x=10$, and verify that $p(x) > 0 \ \forall x$. To verify that the area under the curve is equal to 1, we can use the integrate function (integrate(dnorm, -10, 10) which gives us the result 1 with absolute error $< 7{,}4 \cdot 10^{-5}$).

We can also use the rectangles method to approximate the area under the curve instead of using a built-in function (see appended R code). The idea behind this method is shown below. The more steps we have in the horizontal axis, the more accurate this method will be.

↳ uploaded to Canvas



3) a)



b) From the distribution function F, it is possible to determine the density function f:

$$f(x) = \begin{cases} 1/4 & \text{if } -1 \leq x < 0 \\ 1/4 & \text{if } x = 0 \\ 1/6 & \text{if } x = 1 \\ 1/12 & \text{if } 1 < x < 2 \\ 1/4 & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

This could have been done without this step, but only with the cdf:

$\mathbb{P}(X<1) = 1/2$

$\mathbb{P}(X=1) = F(1) - F(1^-) = 2/3 - 1/2 = 1/6$

$\mathbb{P}(1 \leq X < 2) = \mathbb{P}(X<2) - \mathbb{P}(X<1) = \left(\frac{2}{12} + \frac{7}{12}\right) - \frac{1}{2} = \frac{1}{4}$

$\mathbb{P}(X>2) = 1 - \mathbb{P}(X \leq 2) = 1 - 1 = 0$

$\mathbb{P}(X = 3/2) = 0$ (because F is continuous from 1 to 2)

$\mathbb{P}(1 < X \leq 6) = \mathbb{P}(X \leq 6) - \mathbb{P}(X \leq 1) = 1 - 2/3 = 1/3$

Therefore:

$$\mathbb{P}(X<1) = \int_{-\infty}^{1^-} f(x)\,dx = \int_{-1}^{0} \frac{1}{4}\,dx + \frac{1}{4} \cancel{\frac{1}{6}} = \frac{1}{4} + \frac{1}{4} \cancel{\frac{1}{6}} = \cancel{\frac{2}{3}}\ \frac{1}{2}$$

$$\mathbb{P}(X=1) = \frac{1}{6}$$

$$\mathbb{P}(1 \leq X < 2) = \int_{1}^{2^-} f(x)\,dx = \frac{1}{6} + \int_{1}^{2} \frac{1}{12}\,dx = \frac{1}{6} + \frac{1}{12} = \frac{1}{4}$$

$$\mathbb{P}(X>2) = 0$$

$$\mathbb{P}\left(X = \frac{3}{2}\right) = 0 \ \cancel{\frac{1}{12}}$$

$$\mathbb{P}(1 < X \leq 6) = \int_{1^+}^{6} f(x)\,dx = \int_{1}^{2} \frac{1}{12}\,dx + \frac{1}{4} = \frac{1}{12} + \frac{1}{4} = \frac{1}{3}$$

4) $$F(t) = \begin{cases} \frac{t}{1+t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

- The plot of $F(t)$ shows that F is non-decreasing.
- When $t \to -\infty$, $F(t) = 0$
- When $t \to +\infty$, $F(t) \to 1$
- $F(t)$ is continuous for $\forall t \in \mathbb{R}$

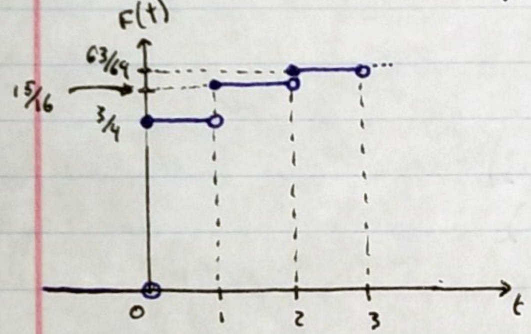
Therefore, $F(t)$ is a distribution function.

Andre Guimaraes Duarte

5) $p(x) = \left(\frac{3}{4}\right) \cdot \left(\frac{1}{4}\right)^x$ for $x = 0, 1, 2, \ldots$

We have $F(t) = \sum_{\substack{x \in \mathbb{N} \\ x \leq t}} \frac{3}{4} \cdot \left(\frac{1}{4}\right)^x$. $F(t)$ is the sum of a geometric series with first term $\frac{3}{4}$ and ratio $\frac{1}{4}$. Therefore, we can write:

$$F(t) = \frac{3}{4} \cdot \frac{1 - \left(\frac{1}{4}\right)^t}{1 - \frac{1}{4}} = 1 - \frac{1}{4^t}$$



6) $\mathbb{E}[X(11-X)] = \mathbb{E}[11.X - X^2] = 11\,\mathbb{E}[X] - \mathbb{E}[X^2]$

$$\mathbb{E}[X] = \sum_{i=1}^{10} i \cdot \frac{1}{10} = \frac{1}{10} \sum_{i=1}^{10} i = \frac{1}{\cancel{10}} \cdot \frac{\cancel{10}(1+10)}{2} = \frac{11}{2} = 5{,}5$$

$$\mathbb{E}[X^2] = \sum_{i=1}^{10} i^2 \cdot \frac{1}{10} = \frac{1}{10} \sum_{i=1}^{10} i^2 = \frac{1}{\cancel{10}} \cdot \frac{\cancel{10}(10+1)(2.10+1)}{6} = 38{,}5$$

$$\Rightarrow \mathbb{E}[X(11-X)] = 11.\,5{,}5 - 38{,}5 = 16{,}5$$

7) From $F(x)$, we can deduce the ~~density~~ mass function:

$$p(x) = \begin{cases} 3/8 & \text{if } x = -3 \\ 3/8 & \text{if } x = 0 \\ 2/8 & \text{if } x = 6 \\ 0 & \text{otherwise} \end{cases}$$

Therefore, we have:

$$\mathbb{E}[X] = -3 \cdot \frac{3}{8} + 0 \cdot \frac{3}{8} + 6 \cdot \frac{2}{8} = \frac{3}{8}$$

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \left(-3 - \frac{3}{8}\right)^2 \cdot \frac{3}{8} + \left(0 - \frac{3}{8}\right)^2 \cdot \frac{3}{8} + \left(6 - \frac{3}{8}\right)^2 \cdot \frac{2}{8} \simeq 12{,}23$$

$$\sigma_X = \sqrt{\mathrm{Var}(X)} \simeq 3{,}50$$

8) $\mathrm{Var}(-3x+5) = \mathbb{E}\left[\left((-3x+5) - \mathbb{E}[-3x+5]\right)^2\right]$

$\cancel{= \mathbb{E}[((-3x+5) - \mathbb{E}[-3x+5]}$

$\mathbb{E}[-3x+5] = -3\mathbb{E}[x] + \mathbb{E}[5] = -3 + 5 = 2$

$$\begin{aligned} \Rightarrow \mathrm{Var}(-3x+5) &= \mathbb{E}\left[(-3x+5-2)^2\right] \\ &= \mathbb{E}\left[(-3x+3)^2\right] \\ &= \mathbb{E}\left[9(x-1)^2\right] \\ &= 9\,\mathbb{E}\left[(x-1)^2\right] \\ &= 9\,\mathbb{E}\left[x^2-2x+1\right] \\ &= 9\left(\mathbb{E}[x^2-2x] + \mathbb{E}[1]\right) \\ &= 9\left(\mathbb{E}[x(x-2)] + 1\right) \\ &= 9(3+1) \\ &= 36 \end{aligned}$$

9) Let x be the ~~event~~ random variable "a student was born in April or in October". We have $X \sim \mathrm{Bin}(6; \frac{1}{6})$. Therefore:

$$p(x=3) = \binom{6}{3}\left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^3 \simeq 0{,}054$$

10) Let x be the random variable "there is a foreign immigrant". We have $X \sim \mathrm{Bin}(80; 0{,}025)$. It is easier to calculate the probabilities that there are exactly 0 and 1 illegal immigrant, and then subtract these from 1.

$$p(x=0) = \binom{80}{0}(0{,}025)^0(0{,}975)^{80} \simeq 0{,}132$$
$$p(x=1) = \binom{80}{1}(0{,}025)^1(0{,}975)^{79} \simeq 0{,}271$$

$$\Rightarrow p(x \geq 2) = 1 - \left(p(x=0) + p(x=1)\right) \simeq 0{,}597$$

Andre Guimaraes Duarte

11) The probability that at least n bulbs are requires is 1 minus the probability that at most n-1 bulbs are required. Let x be the random variable "x bulbs are required". $X \sim Geo(p)$.

Therefore:

$$p(x \geq n) = 1 - p(x \leq n-1)$$
$$= 1 - \sum_{i=1}^{n-1} p(1-p)^{i-1}$$

12) a) For f to be a proper density function, we need two things:

i) $f(x) > 0 \iff c(x-1)(2-x) \geq 0$ for $1 < x < 2$

If $1<x<2$, then $\underbrace{(x-1)}_{>0}\underbrace{(2-x)}_{>0} > 0$

Therefore, $c \geq 0$

ii) $\int_1^2 c(x-1)(2-x)\,dx = 1$

We have $\int_1^2 c(x-1)(2-x)\,dx = c\int_1^2 (-x^2+3x-2)\,dx$

$$= c\left[-\frac{x^3}{3} + \frac{3x^2}{2} - 2x\right]_1^2$$

$$= c\left[\left(-\frac{8}{3} + \frac{12}{2} - 4\right) - \left(-\frac{1}{3} + \frac{3}{2} - 2\right)\right]$$

$$= \frac{c}{6}$$

Therefore, $\frac{c}{6} = 1 \iff \underline{c = 6}$

b) The distribution function F can be obtained by integrating f:

$$F(t) = \begin{cases} \int 6(t-1)(2-t)\,dt & \text{if } 1<t<2 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} -2t^3 + 9t^2 - 12t & \text{if } 1<t<2 \\ 0 & \text{otherwise} \end{cases}$$

(3)

c) The probability that a student will finish in less than 75 minutes (1,25h) is:

$$\mathbb{P}(X \leq 1,25) = \cancel{F(1,25)} = \int_1^{1,25} f(x)\,dx = \Big[-2t^3 + 9t^2 - 12t\Big]_1^{1,25} \simeq 0,156$$

The probability that a student will finish between 1,5 and 2 hours is:

$$\mathbb{P}(1,5 \leq X \leq 2) = \int_{1,5}^{2} f(x)\,dx = \Big[-2t^3 + 9t^2 - 12t\Big]_{1,5}^{2} = 0,5$$

13) $\mu_x = \int_1^2 x \cdot 6(x-1)(2-x)\,dx$

$= 6\int_1^2 (-x^3 + 3x^2 - 2x)\,dx$

$= 6\Big[-\frac{x^4}{4} + x^3 - x^2\Big]_1^2$

$= 6\Big[(-4+8-4) - (-\frac{1}{4}+1-1)\Big]$

$= \frac{6}{4}$

$\mu = 1,5h$

$Var(x) = \int_1^2 (x-\frac{3}{2})^2 \cdot 6(x-1)(2-x)\,dx$

$= 6\int_1^2 (x^2 - 3x + \frac{9}{4})(-x^2 + 3x - 2)\,dx$

$= 6\int_1^2 \Big(-x^4 + 6x^3 - \frac{53}{4}x^2 + \frac{51}{4}x - \frac{9}{2}\Big)\,dx$

$= 6\Big[-\frac{x^5}{5} + \frac{6}{4}x^4 - \frac{53}{4 \cdot 3}x^2 + \frac{51}{4 \cdot 2}x^2 - \frac{9}{2}x\Big]_1^2$

$Var(x) = 0,05\ h^2$

$\Rightarrow \sigma_x = \sqrt{0,05} \simeq 0,224\,h$

Andre Guimaraes Duarte

14) Suppose we have a function unif(a, b) that generates a value from a uniform distribution between a and b.

Additionally, the cdf of a geometric ~~distribution~~ random variable is $F(t) = 1 - (1-p)^t$, so $F^{-1}(t) = \left| \frac{\ln(1-t)}{\ln(1-p)} \right|$ for $t \in (0,1)$.

Pseudo-code to simulate realizations from a geometric random variable with success parameter p is as follows.

```
p <- SUCCESS_RATE
x <- vector()
for (i in NUMBER_OF_TRIALS) {
    u <- unif(0,1)
    x_ <- abs( ln(1-u) / ln(1-p) )

    x.append(x_)
}
```

15) a) In order to prove that this pmf is a proper pmf, we need to show that $p(x) > 0$ for $x = 0, 1, 2, \dots$

$$p(x) > 0 \iff \frac{\lambda^x e^{-\lambda}}{x!} > 0$$

Since $\lambda > 0$, $\lambda^x > 0 \ \forall x \in \mathbb{N}$.
Also, $e^{-\lambda} > 0$ and $x! > 0 \ \forall x \in \mathbb{N}$.
Therefore, $p(x) > 0 \ \forall x \in \mathbb{N}$.

In addition, we need to prove that $\sum_{x \in \mathbb{N}} p(x) = 1$

$\sum_{x \in \mathbb{N}} p(x) = e^{-\lambda} \sum_{x \in \mathbb{N}} \frac{\lambda^x}{x!}$ and we know that $e^{\lambda} = \sum_{x \in \mathbb{N}} \frac{\lambda^x}{x!}$ by definition of the exponential.

So $\sum_{x \in \mathbb{N}} p(x) = e^{-\lambda} \cdot e^{\lambda} = 1$

So the pmf $p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ for $x \in \mathbb{N}$ is a proper pmf.

b) We want to prove the following recursive relationship:

$$p(x+1) = \frac{\lambda}{x+1} p(x).$$

First, let's prove it for the first value ($x=0$).

$$p(0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda}$$

$$p(1) = \frac{\lambda^1 e^{-\lambda}}{1!} = \lambda e^{-\lambda} = \frac{\lambda}{0+1} p(0)$$

Next, we suppose this relationship is true up to a ~~term~~ step x. Let's prove that it's true for the next step $(x+1)$.

$$p(x+2) = \frac{\lambda^{x+2} e^{-\lambda}}{(x+2)!} = \frac{\lambda \cdot \lambda^{x+1} e^{-\lambda}}{(x+2)(x+1)!} = \frac{\lambda}{x+2} p(x+1)$$

We have shown that the recursive relationship is true for the step $x+1$. Therefore, it must be true for every $x \in \mathbb{N}$.

16) See appended R code (uploaded to Canvas)