Solutions to problem set 3, 24 February, 2025

1. We have ordered an item online, and we know that the delivery is between 8 : 00 and 10 : 00, at a time which is uniformly distributed on the interval $[8, 10]$: the probability that it is between a and b is proportional to $b - a$, if $8 \leq a \leq b \leq 10$. Let X be the time of the delivery (it is a random element of $[8, 10]$, hence this is a random variable).

   (a) Calculate $\mathbb{P}(X \leq 9)$, that is, the probability that the delivery is before 9.

   Since the probabilities are proportional to the length of the interval (we have $a = 8.b = 9$):

$$\mathbb{P}(X \leq 9) = \frac{9-8}{10-8} = 1/2.$$

   (b) Calculate the probability $\mathbb{P}(8.5 < X < 9)$, that is, the probability is between 8 : 30 and 9 : 00.

   Similarly for $a = 8.5.b = 9$ we obtain:

$$\mathbb{P}(8.5 < X < 9) = \frac{9-8.5}{10-8} = 1/4.$$

   (c) Calculate the probability $\mathbb{P}(X > 9.75)$.

   Similarly

$$\mathbb{P}(X > 9.75) = \frac{10-9.75}{10-8} = 1/8.$$

   (d) Draw the curve $\mathbb{P}(X \leq t)$ (the cumulative distribution function of X) as a function of t, where $t \in \mathbb{R}$ is a real number.

   We have three different cases.

   If $t < 8$, then $X \leq t$ cannot happen: $\mathbb{P}(X \leq t) = 0$.

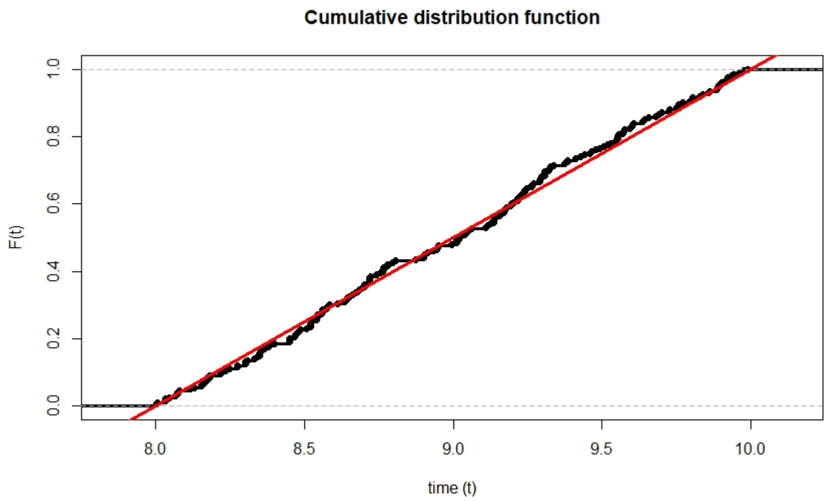
   Similarly, if $t > 10$, then $\mathbb{P}(X \leq t)$ for sure, hence $\mathbb{P}(X \leq t) = 1$.

   If $8 \leq X \leq 10$. then we can calculate similarly to case (a):

$$\mathbb{P}(X \leq t) = \frac{t-8}{10-8} = \frac{t-8}{2} = \frac{t}{2} - 4.$$

   Program R: mean and empirical cumulative distribution function for a sample of size $n = 200$ (the proportion of observations not larger than t) and the theoretical cumulative distribution function of the uniform distribution:

```
sample200<-runif(200, 8, 10)
mean(sample200)
[1] 8.997175
plot(ecdf(sample200), lwd="3", xlab="time (t)", ylab="F(t)", main="Cumulative distribution function")
abline(-4, 0.5, lwd="3", col="red")
```

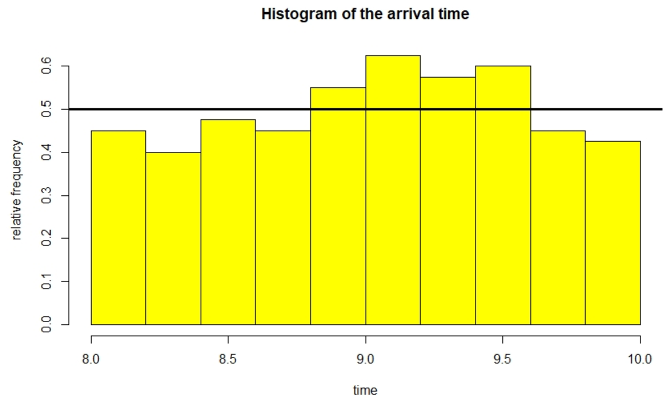


(e) Does there exist a function $f : \mathbb{R} \to \mathbb{R}_+$, such that $\mathbb{P}(a < X < b) = \int_a^b f(x)\, dx$ is satisfied for all real numbers $a < b$? If we take $a \to -\infty$, then we obtain $\mathbb{P}(X \le b) = \int_{-\infty}^b f(x)\, dx$. Hence function f can be obtained as the derivative of the function $\mathbb{P}(X \le b)$ with respect to b. This is

$$f(x) = \begin{cases} 0; & \text{if } x < 8 \text{ or } x > 10; \\ \frac{1}{2}. & \text{if } 8 \le x \le 10. \end{cases}$$

It is easy to check that this function f satisfies the condition.
Histogram:

```
> sample200<-runif(200, 8, 10)
> hist(sample200, col="yellow", freq=F, main="Histogram of the arrival time", xlab="time",
ylab="relative frequency")
> abline(0.5, 0, lwd="3")
```



2. The density function of the random variable X has the following form:

$$f(x) = \begin{cases} 0, & x \le 0; \\ c \cdot x, & 0 \le x \le 1; \\ 0, & x > 1. \end{cases}$$

Determine the value of c. What is the probability that X is between $1/4$ and $1/2$? What is the probability that it is between $1/2$ and $3/4$? Determine the cumulative distribution function of X. and also its expectation and variance.
For every density function we have

$$\int_{-\infty}^{\infty} f(x)\, dx = 1.$$

Now

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^1 c \cdot x\, dx = c \int_0^1 x\, dx = c \left[\frac{x^2}{2}\right]_{x=0}^1 = c \cdot \frac{1}{2} = 1 \quad \Rightarrow \quad c = 2.$$

In general, if f is the density function of X. then

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

holds for every $a < b$. Hence the probability that X is between $1/4$ and $1/2$:

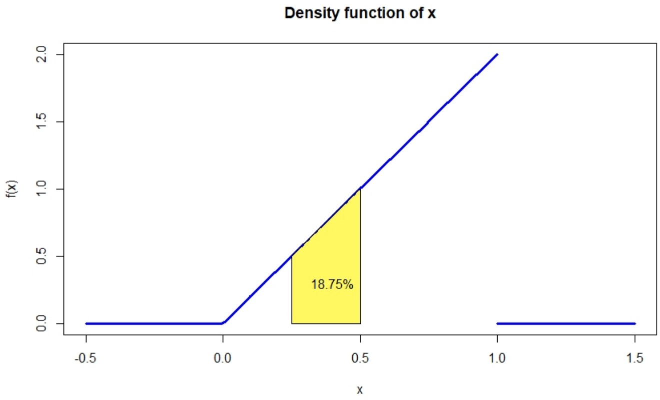
$$\mathbb{P}(1/4 \le X \le 1/2) = \int_{1/4}^{1/2} f(x)\, dx = \int_{1/4}^{1/2} 2x\, dx = \left[x^2\right]_{x=1/4}^{1/2} = \left(\frac{1}{2}\right)^2 - \left(\frac{1}{4}\right)^2 = \frac{3}{16} = 18.75\%.$$

Similarly:

$$\mathbb{P}(1/2 \le X \le 3/4) = \int_{1/2}^{3/4} f(x)\, dx = \int_{1/2}^{3/4} 2x\, dx = \left[x^2\right]_{x=1/2}^{3/4} = \left(\frac{3}{4}\right)^2 - \left(\frac{1}{2}\right)^2 = \frac{5}{16} = 31.25\%.$$

As for the cumulative distribution function, by the definition of the density function, we have $F(t) = \mathbb{P}(X \le t) = \int_{-\infty}^t f(x)\, dx$. This function is 0, if $t < 0$, because we integrate function 0. On the other hand, if $0 < t < 1$, then

$$F(t) = \int_{-\infty}^t f(x)\, dx = \int_0^t 2x dx = \left[x^2\right]_{x=0}^t = t^2.$$



Finally, if $t > 1$, then

$$F(t) = \int_{-\infty}^{t} f(x)\,dx = \int_{0}^{1} 2x dx = \left[x^2\right]_{x=0}^{1} = 1.$$

By definition, the expectation of this random variable can be calculated as follows:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_{0}^{1} 2x^2\,dx = \frac{2}{3}.$$

Then, the expectation of X^2 is as follows:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx = \int_{0}^{1} 2x^3\,dx = \frac{1}{2}.$$

Hence we obtain the variance of X by

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{1}{2} - \frac{4}{9} = \frac{5}{18} = 0.278.$$

3. Let the cumulative distribution function of X be given by:

$$F(t) = \mathbb{P}(X \leq t) = \begin{cases} 0, & \text{if } t \leq 0; \\ 25t^2, & \text{if } 0 \leq t \leq 1/5; \\ 1, & \text{if } t \geq 1/5. \end{cases}$$

(a) Determine the density function of X.
(b) What is the probability that X is between 0.1 and 0.15?
(c) Determine the expectation of X.
(d) Determine the standard deviation of X.

If it exists, the density function is the derivative of the cumulative distribution function:

$$f(x) = \begin{cases} 0. & \text{if } x \leq 0; \\ 50 \cdot x. & \text{if } 0 \leq x \leq 1/5; \\ 0. & \text{if } x \geq 1/5. \end{cases}$$

It is easy to check that $\int_{-\infty}^{t} f(x)dx = F(t)$ is satisfied for every t.
Based on the definition of the cumulative distribution function:

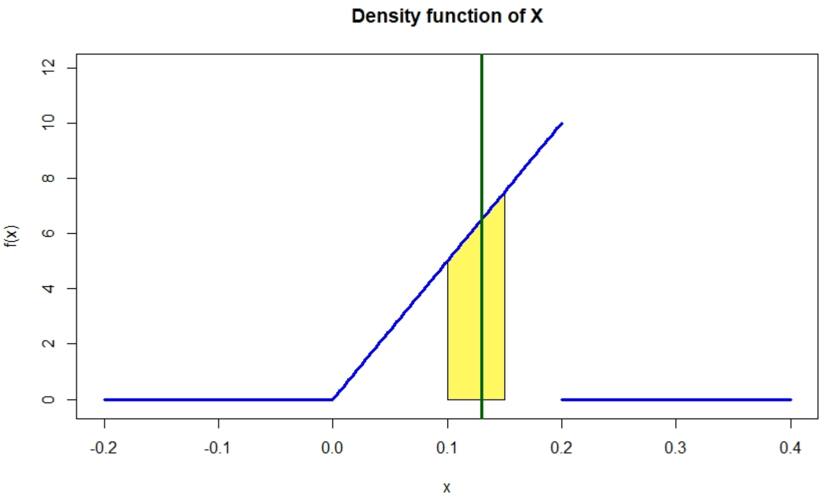
$$\mathbb{P}(0.1 \leq X \leq 0.15) = \mathbb{P}(X \leq 0.15) - \mathbb{P}(X \leq 0.1) = F(0.15) - F(0.1) = 25 \cdot 0.15^2 - 25 \cdot 0.1^2 = 31.25\%.$$

Or:

$$\mathbb{P}(0.1 \leq X \leq 0.15) = \int_{0.1}^{0.15} f(x)\,dx.$$

Based on the definition of expectation:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_{0}^{1/5} 50 \cdot x^2 dx = 50 \cdot \left[\frac{x^3}{3}\right]_{x=0}^{1/5} = 50 \cdot \frac{1}{5^3 \cdot 3} = 0.13.$$



We first calculate the expectation of the square of X:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int_0^{1/5} 50 \cdot x^3 dx = 50 \cdot \left[\frac{x^4}{4}\right]_{x=0}^{1/5} = 50 \cdot \frac{1}{5^4 \cdot 4} = 0.02.$$

We obtain that

$$s.d.(X) = \sqrt{\mathbb{E}(X^2) - \mathbb{E}(X)^2} = \sqrt{0.02 - 0.13^2} = 0.06.$$

4. Suppose that X has uniform distribution on the interval $[0,1]$. that is, its density function is 1 within $[0,1]$ and 0 otherwise. a) Find the cumulative distribution function and density function of X^2.

$$\mathbb{P}(X^2 \le t) = \mathbb{P}(X \le \sqrt{t}) = \begin{cases} 0, & t \le 0; \\ \sqrt{t}, & 0 < t < 1; \\ 1, & t \ge 1. \end{cases}$$

The density function of X^2 can be determined by differentiating this:

$$f_{X^2}(t) = \begin{cases} 0, & t \le 0; \\ \frac{1}{2\sqrt{t}}, & 0 < t < 1; \\ 0, & t \ge 1. \end{cases}$$

b) Find the cumulative distribution function and density function of X^5.
Similarly we have

$$\mathbb{P}(X^5 \le t) = \mathbb{P}(X \le t^{1/5}) = \begin{cases} 0, & t \le 0; \\ t^{1/5}, & 0 < t < 1; \\ 1, & t \ge 1. \end{cases}$$

The density function of X^5 can be determined by differentiating this:

$$f_{X^5}(t) = \begin{cases} 0, & t \le 0; \\ \frac{1}{4} \cdot t^{-4/5}, & 0 < t < 1; \\ 0, & t \ge 1. \end{cases}$$

c) Determine the expectation and variance of X.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} s \cdot f(s)ds = \int_a^b s \cdot \frac{1}{b-a} ds = \frac{1}{b-a}\left[\frac{s^2}{2}\right]_{s=a}^{b} = \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2},$$

because the primitive function of x is $\frac{x^2}{2}$, and $b^2 - a^2 = (b-a)(b+a)$.
As for the expectation of X^2:

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} s^2 \cdot f(s)ds = \int_a^b s^2 \cdot \frac{1}{b-a} ds = \frac{1}{b-a}\left[\frac{s^3}{3}\right]_{s=a}^{b} \\ &= \frac{1}{b-a} \cdot \frac{b^3-a^3}{3} = \frac{a^2+ab+b^2}{3},\end{aligned}$$

because the primitive function of x^2 is $\frac{x^3}{3}$, és $b^3 - a^3 = (b-a)(a^2+ab+b^2)$.
Hence the variance of X is as follows:

$$\begin{aligned}Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 &= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12} \quad \Rightarrow \quad s.d.(X) = \frac{b-a}{\sqrt{12}}.\end{aligned}$$

$\square$

Definition 1 *The* **cumulative distribution function** *of a random variable X is a function $F : \mathbb{R} \to [0,1]$, where*

$$F(t) = \mathbb{P}(X \le t)$$

holds for every real number t.
A function $f : \mathbb{R} \to \mathbb{R}$ is **density function** *of a random variable X if*

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

holds for every real numbers $a < b$.

Then (only if the density function exists) we have

$$F(t) = \mathbb{P}(X \le t) = \int_{-\infty}^{t} f(x)\, dx \qquad \text{and} \qquad f(t) = F'(t).$$

where the latter equation is for "almost all" real numbers t.
For every density function f we have $\int_{-\infty}^{\infty} f(x)\, dx = 1$ and $f(x) \ge 0$ for "almost all" real numbers t.
The **expectation** of a discrete random variable is defined by

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j). \quad \text{provided that } \sum_{j=1}^{\infty} |x_j| \cdot \mathbb{P}(X = x_j) < \infty.$$

where $x_1, x_2, \ldots$ are the possible values of X.
Continued on the next page

Definition 2 *Let X be a random variable with density function f. Then the expectation of X is given by*

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx.$$

The standard deviation of X is given by

$$s.d.(X) = \sqrt{\mathbb{E}(X^2) - \mathbb{E}(X)^2}.$$

where

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx.$$

Definition 3 (Exponential distribution) *A random variable X has exponential distribution with parameter λ is its density function is given by*

$$f(x) = \begin{cases} 0, & x \le 0; \\ \lambda e^{-\lambda x}, & x \ge 0. \end{cases}$$

If X has exponential distribution with parameter λ. then

$$\mathbb{E}(X) = s.d.(X) = \frac{1}{\lambda}.$$

---

5. Suppose that the response time of a server (in seconds) has exponential distribution with parameter $\lambda = 2$.

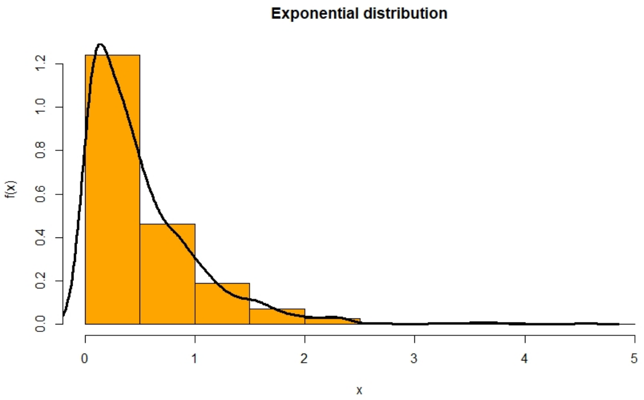
(a) Determine the cumulative distribution function and the density function of the response time. The distribution function:

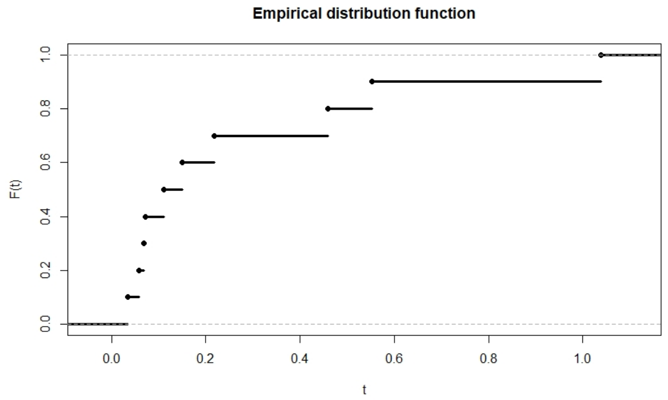
$$F(t) = \begin{cases} 0, & t \leq 0; \\ 1 - e^{-2t}, t \geq 0. \end{cases}$$

The density function:

$$f(x) = \begin{cases} 0, & x \leq 0; \\ 2e^{-2x}, x \geq 0. \end{cases}$$

```
expsample=rexp(1000, rate=2)
hist(expsample, col="orange", main="Exponential distribution", xlab="x", ylab="f(x)",
freq=FALSE)
lines(density(expsample), lwd="3")
plot(ecdf(expsample[1:10]), lwd="3", xlab="t", ylab="F(t)", main="Empirical distribution
function")
```





(b) What is the probability that the response time is more than 0.5 seconds?

$$\mathbb{P}\left(X \geq \frac{1}{2}\right) = 1 - \mathbb{P}\left(X \leq \frac{1}{2}\right) = 1 - F\left(\frac{1}{2}\right) = 1 - (1 - e^{-1/2 \cdot 2}) = e^{-1} = 36.8\%.$$

```
> pexp(0.5, rate=2, lower.tail=F)
[1] 0.3678794
```

(c) the response time is at least 1 second?

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X \leq 1) = 1 - F(1) = 1 - (1 - e^{-2 \cdot 1}) = e^{-2} = 13.5\%.$$

```
> pexp(1, rate=2, lower.tail=F)
[1] 0.1353353
```

(d) given that the response time is at least 0.5 second. what is probability that it is at least 2.5 seconds? Let X be the response time.

$$\mathbb{P}(X \geq 2.5 | X \geq 0.5) = \frac{\mathbb{P}(X \geq 2.5)}{\mathbb{P}(X \geq 0.5)} = \frac{e^{-\lambda \cdot 2.5}}{e^{-\lambda \cdot 0.5}} = e^{-\lambda \cdot 2} = e^{-4} = 1.8\%.$$

Notice that

$$\mathbb{P}(X \geq 2) = 1 - \mathbb{P}(X \leq 2) = 1 - F(2) = 1 - (1 - e^{-2 \cdot 2}) = e^{-4} = 1.8\%.$$

```
> pexp(2, rate=2, lower.tail=F)
[1] 0.01831564
```

(e) What is the probability that the response time is between 1 and 2 seconds?

$$\mathbb{P}(1 \leq X \leq 2) = \mathbb{P}(X \leq 2) - \mathbb{P}(X \leq 1) = F(2) - F(1) = (1 - e^{-4}) - (1 - e^{-2}) = e^{-2} - e^{-4} = 11.7\%$$

```
> pexp(2, rate=2)-pexp(1, rate=2)
[1] 0.1170196
```

(f) For which t is it true that the probability that the response time is at most t is equal to $1/2$?

$$P(X < t) = \mathbb{P}(X \leq t) = F(t) = 1 - e^{-2t} = \frac{1}{2} \Leftrightarrow e^{-2t} = \frac{1}{2} \Leftrightarrow -2t = \log\frac{1}{2} \Leftrightarrow t = \frac{\log 2}{2} = 0.35.$$

(g) Find the expectation and standard deviation of X.

$$\mathbb{E}(X) = s.d.(X) = \frac{1}{2} = 0.5.$$

(h) Generate a sample of 1000 random variables from the exponential distribution with parameter $\lambda = 2$. Make a histogram, and find the mean, standard deviation and median of the sample. Compare the results with the values calculated above.

```
expsample=rexp(1000, rate=2)
hist(expsample, col="orange", main="Exponential distribution", xlab="x", ylab="f(x)",
freq=FALSE)
> mean(expsample)
[1] 0.5207195
> sd(expsample)
[1] 0.5241174
```