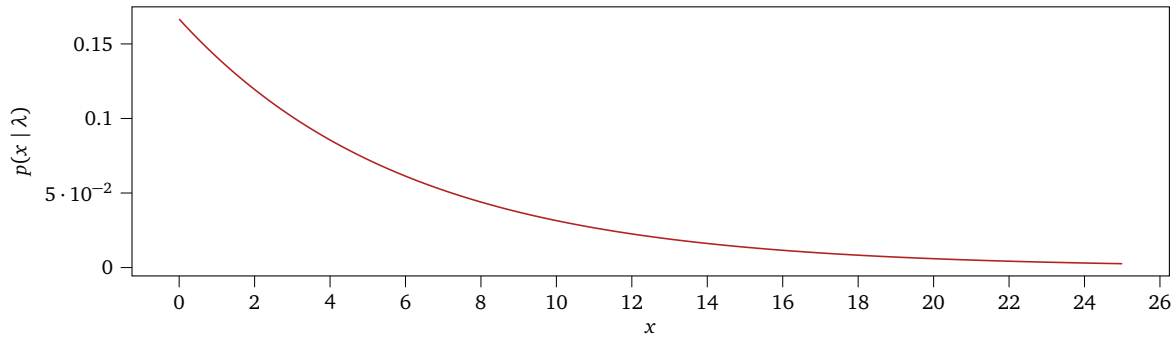
Data Literacy – Winter Term 2020/21

# Exercise Sheet #7

due on Tuesday, 22 December 2020, 12 noon sharp

In the feedback, several people noted that they had trouble with the notion of $p$-values and estimation. Thus, this week's exercise sheet offers additional opportunities to get accustomed to concepts of probability and hypothesis tests.




1. **EXAMple:** Consider a situation in which events can happen in any interval of length $\delta t \in \mathbb{R}_+$ with probability $\lambda \cdot \delta t$ for $\lambda \in \mathbb{R}_+$ (i.e. the probability is proportional to duration of the time step. This is known as a Poisson process). It can be shown that in such situations, the distribution of the length $x \in [0, \infty)$ of *time intervals between two events* ("waiting times") is given by the *exponential distribution* (cf. figure above), defined by the probability density function

$$p(x \mid \lambda) = \lambda e^{-\lambda x}.$$

   Compute

   (a) the **cumulative density function** (CDF) of $p(x \mid \lambda)$

   (b) the **mean** of $p(x \mid \lambda)$

   (c) the **variance** of $p(x \mid \lambda)$

2. **Theory Question:** There were several questions regarding *one-* and *two-sided* tests after the lecture. This exercise might help show why it is not always clear whether to use a one- or two-sided test.

   You're waiting at a bus stop in the town of *Poissingen*. A sign put up by the bus operator claims that busses arrive here independently at random (a Poisson process), with, on average, one bus every six minutes (see Exercise 1, $\lambda = \frac{1}{6\,\text{min}}$, the pdf is plotted above). Knowing that there's no schedule, you just showed up at your leisure, not aiming for a particular point in time.

   (a) What is your *expected* wait for a bus after you arrive at the stop?

   (b) You have been waiting for a bus for 20 minutes now. Which $p$-value would you assign to the null hypothesis that the bus operator's claim is correct? Do you want to use a one- or two-sided test?

   (c) The next morning, you're on your way to the bus stop again. Exactly 1 second after you arrive at the bus stop, the bus shows up. Are you *more* or *less* surprised about this event than you were about yesterday's wait? What is the $p$-value you would assign for the null hypothesis under this observation (just this one, not including yesterday's)? Do you want to change your answer to (b)? If so, what is the correct way to do a two-sided test for the situation described in (b) (there is no unique answer to this, just think about it)?

3. **Practical Question:** You can find this week's practical exercise in `Exercise_07.ipynb`