Exercise Sheet 7

May 31th 2023

Due to the holiday submission of the homework assignments until June 9th, 09:00 am online in TUM-moodle in groups of two. Please put the *full names* and *student IDs* of you *and* your partner on all parts of your submission. The solution will be discussed in the classes one week after and published on moodle after some time.

Homework

Problem H 28 - Exponential Distribution [5 pts.]

The waiting time T at the checkout of the canteen is 4 minutes on average. Assume that T can be modeled by an exponential distribution.

a) State the probability density function of T, its expected value and its variance.

b) Find the probability for the following events:

(i) to wait exactly 2 minutes,

(ii) to wait more than 4 minutes,

(iii) to wait between 3 and 5 minutes,

(iv) after having waited 4 minutes, to wait at least another 4 minutes.

c) Explain how to simulate the waiting time by using uniformly distributed random numbers. Include explicit mathematical formulas to your explanation.

Solution:

a) The probability density function of an exponential random variable is fully determined by the parameter $\lambda = 1/\mathbb{E}(T) = 1/4$ (in min^{-1}), so it is given as

$$f_T(T) = \begin{cases} \frac{1}{4} \cdot e^{-T/4} & \text{if } T \geq 0 \\ 0 & \text{if } T < 0 \end{cases} .$$

Clearly, the expected value is $\mathbb{E}(T) = 4$ min, the variance $\text{Var}(T) = 1/\lambda^2 = 16\ \text{min}^2$. ✓

b) For the following: ✓✓✓.

(i) To find the probability the event has to be identified with an Borel set, i.e. an interval. In this case, it is

$$Pr(T = 4) = Pr(4 \leq T \leq 4) = \int_4^4 f_T(T)dT = 0.$$

It is important to understand that the probability of observing an exact value of the continuous waiting time is equal to zero.

(ii)
$$Pr(T > 4) = 1 - Pr(T \leq 4) = 1 - F_T(4) = e^{-4/4} = e^{-1} \approx 0.37.$$

(iii)
$$Pr(3 \leq T \leq 5) = F_T(5) - F_T(3) = e^{-3/4} - e^{-5/4} \approx 0.19.$$

(iv) An exponential variable is memoryless, so we have

$$Pr(T \geq 4 + 4|T \geq 4) = Pr(T \geq 4) = 1 - F_T(4) = e^{-4/4} \approx 0.37.$$

c) By the inverse transform sampling method, a random number r_0 from Unif$(0,1)$ can be transformed to a sample T_0 from f_T via the inverse of the cumulative distribution function. For this setup, we need to calculate

$$T_0 = F_T^{-1}(r_0) = -\frac{1}{\lambda}\ln(1 - r_0) = -4\ln(1 - r_0). \checkmark$$

Since also $(1 - r_0) \sim$ Unif$(0,1)$ we may instead calculate

$$T_0 = -4\ln(r_0).$$

Problem H 29 - Probability measures for the continuous case **[4 pts.]**

Let $d > 0$ be a positive constant and X be a continuous random variable with the probability density function

$$f_X(x) = \begin{cases} d \cdot (1 - x^6) & \text{if } x \in [-1, 1] \\ 0 & \text{else} \end{cases}.$$

a) Show that f_X is (*Borel*) measurable. *(Hint: Recall the properties of measurable functions that were explained in the lecture prior to the definition of the Lebesgue integral.)*

b) Determine the constant d.

Solution:

a) By definition, a function f_X is Borel measurable if and only if the inverse image of any Borel set again is a Borel set. In particular, every indicator function of a Borel set A,

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{else} \end{cases}$$

is Borel measurable. Further, every continuous function is Borel measurable, hence $d \cdot (1 - x^6)$ is Borel measurable. Finally, the product of Borel measurable functions is Borel measurable as well. As we can express the function of interest as $f_X(x) = d \cdot (1 - x^6) \cdot g_{[-1,1]}$ and as the interval $[-1, 1]$ is a Borel set by definition, $f_X(x)$ is Borel measurable. ✓✓

b) The function $f_X(x)$ is an admissible probability density if f_X is integrable, if it maps to non-negative real numbers and, in particular, satisfies $\int_{-\infty}^{\infty} f_X(x)dx = 1$.
In general, a function is integrable if it is Borel measurable and if it has a finite integral. We have shown already that f_X is Borel measurable and, obviously, have $f_X(x) \geq 0$ for all $x \in \mathbb{R}$. So, consider now the integral $\int_{-\infty}^{\infty} f_X(x)dx$. Since there is only even powers of x in $f_X(x)$, $f_X(x)$ is an even function, i.e. its graph is symmetric with respect to the y-axis. From this reason, we calculate

$$\int_{-\infty}^{\infty} f_X(x)dx = 2\int_0^1 d \cdot (1 - x^6)dx = 2d \cdot \left[x - \frac{x^7}{7}\right]_0^1 = d \cdot \frac{12}{7}.$$

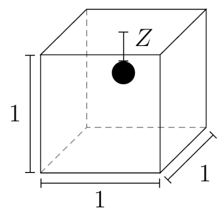
So, by the normalization condition we obtain $d = \frac{7}{12}$. ✓✓

Problem H 30 - Distance in a cube **[4 pts.]**

For a molecular simulation, an atom is modeled as a point (x, y, z) in three-dimensional space $\mathbb{R}^3$. Let us consider that the atom is constrained to the cubic box $[0, 1]^3 \subset \mathbb{R}^3$ and uniformly takes any position in that space. Further define by Z its distance to the closest surface of the cube.

a) Calculate the probability $Pr(Z > c)$ as a function of $c \in \mathbb{R}$.

b) Determine the probability distribution of Z.



Solution:

a) First, observe that $Pr(Z \geq 0) = 1$ since there are no negative distances, and $Pr(Z > \frac{1}{2}) = 0$ because no point may be farther than $\frac{1}{2}$ from a surface of the cube $[0,1]^3$. Now let us ask for which points a certain surface is the closest. Without loss of generality we focus on the surface $(0 \leq x \leq 1, 0 \leq y \leq 1, z = 0)$. The set of all points for which this surface is the closest is a (right square) pyramid with the surface $(0 \leq x \leq 1, 0 \leq y \leq 1, z = 0)$ and the apex at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ that we denote by Π_1. A point $(x, y, z) \in \Pi_1$ has a distance larger than c to the base surface of Π_1 if its z coordinate is within $c < z < \frac{1}{2}$. This can be thought of as cutting Π_1 at $z = c$, yielding a smaller pyramid. Similarly, we find that a point in Π_2 defined by the base surface $(0 \leq x \leq 1, 0 \leq y \leq 1, z = 1)$ and apex $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ will have a distance larger than c to the surface if its z coordinate is within $\frac{1}{2} < z < 1-c$. It is straightforward to extend this consideration to the other four surfaces defining pyramids Π_3, Π_4, Π_5 and Π_6, respectively, which brings analogous ranges for the x and y coordinates. By consequence, a point with a distance larger than c to any surface of the cube has coordinates within $c < x < 1 - c$, $c < y < 1 - c$ and $c < z < 1 - c$. $\checkmark$

Geometrically this means to merge the six cuts of the pyramids defining a cube of edge lengths $(1 - 2c)$ centered around $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. (One could also find the ratio of one of the smaller pyramids to the larger one and then extent this to the total cube by symmetry.) Another way to find this cube is to observe that the distance to a certain surface is perpendicular to the surface, i.e. parallel to one of the axes, and thus depends on only one of the coordinates x, y, z. From the symmetry of each two surfaces, the ranges $c < \cdots < 1 - c$ follow.

As the atom is distributed uniformly over $[0,1]^3$ the desired probability is equal to the volume of this cube $[c, 1-c]^3$, i.e. $Pr(Z > c) = (1 - 2c)^3$. Alternatively, one may calculate the integral with respect to the three coordinates

$$\begin{aligned} Pr(Z > c) &= Pr((x, y, z) \in [c, 1-c]^3) \\ &= Pr(c < x < 1 - c, c < y < 1 - c, c < z < 1 - c) \\ &= \int_c^{(1-c)} \int_c^{(1-c)} \int_c^{(1-c)} 1 \cdot dx dy dz = (1 - 2c)^3. \end{aligned}$$

In total, we have

$$Pr(Z > c) = \begin{cases} 1 & \text{if } c < 0 \\ (1 - 2c)^3 & \text{if } 0 \leq c \leq \frac{1}{2} \\ 0 & \text{if } c > \frac{1}{2} \end{cases} . \checkmark$$

b) The function $c \mapsto Pr(Z > c)$ is continuous for all c and piecewise differentiable. Hence, the probability distribution simply follows from calculating the derivate of the cumulative distribution $Pr(Z \leq c)$. In the range $0 < c < \frac{1}{2}$ this yields

$$\frac{d}{dc} Pr(Z \leq c) = \frac{d}{dc}(1 - Pr(Z > c)) = -3(1 - 2c)^2 \cdot (-2) = 6 \cdot (1 - 2c)^2. \checkmark$$

For arbitrary c we thus find the distribution $f_z : \mathbb{R} \to [0, \infty)$

$$f_z(c) = \begin{cases} 6 \cdot (1 - 2c)^2 & \text{if } 0 \leq c \leq \frac{1}{2} \\ 0 & \text{else.} \end{cases} . \checkmark$$

Problem H 31 - Axle fracture **[6 pts.]**

Mechanical engineers model the axle of 2 meters length of a car by the interval $[-1, 1]$. In case of an accident the axis cracks at a random position X in that interval. From theoretical considerations the engineers found that this point of fracture follows the continuous density

$$f_X(x) = \begin{cases} ax^4 - ax^2 + b & \text{if } x \in [-1,1] \\ 0 & \text{else} \end{cases}$$

with two positive constants $a, b > 0$.

a) Draw a sketch of $f_X(x)$ (for an arbitrary but fixed choice of a, b).

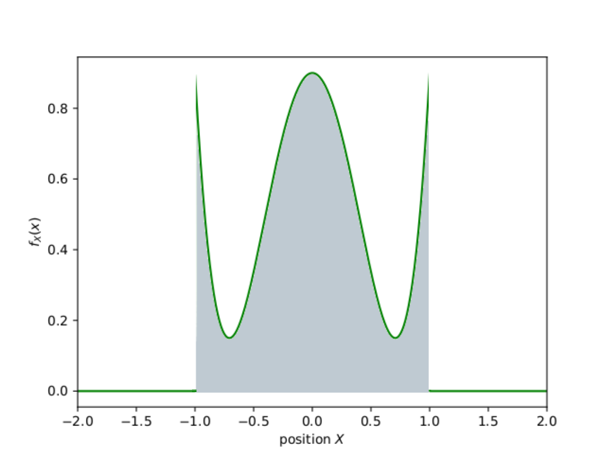
b) Find a relation on a and b such that f_X is a valid probability density function. Explain your reasoning at your sketch.

c) Fix the constants to $a = 3$ and $b = \frac{9}{10}$ from now on. Find the *cumulative* distribution function $F_X(x)$ of the position X as well as the probability that the fracture is on the rightmost third of the axis.

d) At which position is the axis expected to crack? What is the variance of X?

Solution:

a) For the sketch ✓



b) $f_X(x)$ is a valid probability density function if it satisfies the normalization condition

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

As $f_X(x)$ is non-negative in the interval $[-1,1]$ only, we get

$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} f_X(x)dx \\
&= \int_{-1}^{1} (ax^4 - ax^2 + b)dx \\
&= \left[\frac{a}{5}x^5 - \frac{a}{3}x^3 + bx\right]_{-1}^{1} \\
&= 2\left(\frac{a}{5} - \frac{a}{3} + b\right) \\
&= -\frac{4a}{15} + 2b \overset{!}{=} 1.
\end{aligned}$$

So, we have the relation $b = \frac{1}{2} + \frac{2a}{15}$. Graphically, this follows from the demand that the grey region has an area of 1. ✓
Note, that additionally $f_X(x)$ has to be non-negative over its whole domain, so in particular the minimum values in $[-1,1]$ may not be smaller than 0. Differentiation yields the demand

$$\frac{d}{dx}f_X(x) = 4ax^3 - 2ax \overset{!}{=} 0$$

by which the minimum values are obtained at $x = \pm\frac{\sqrt{2}}{2}$. So,

$$f_X\left(\pm\frac{\sqrt{2}}{2}\right) = \left(a\frac{1}{4} - a\frac{1}{2} + b\right) = \left(-\frac{a}{4} + b\right) \geq 0$$

and thus $b \geq \frac{a}{4}$.

c) From the lecture we know that the cumulative distribution is obtained by

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$

Again, we use that the probability density is equal to 0 if $x < -1$ and find

$$\begin{aligned}
F_X(x) &= \int_{-1}^{x} f_X(t)dt \\
&= \left[\frac{3}{5}t^5 - t^3 + \frac{9}{10}x\right]_{-1}^{x} \\
&= \frac{3}{5}x^5 - x^3 + \frac{9}{10}x - \left(-\frac{3}{5} + 1 - \frac{9}{10}\right) \\
&= \frac{3}{5}x^5 - x^3 + \frac{9}{10}x + \frac{1}{2}. \checkmark
\end{aligned}$$

The probability of interest is expressed as $Pr(\frac{1}{3} \leq X \leq 1)$. We use the cumulative distribution function of X found right now to obtain

$$
\begin{aligned}
Pr\left(\frac{1}{3} \leq X \leq 1\right) &= F_X(1) - F_X\left(\frac{1}{3}\right) \\
&= 1 - \frac{3}{5}\left(\frac{1}{3}\right)^5 + \left(\frac{1}{3}\right)^3 - \frac{9}{10}\left(\frac{1}{3}\right) - \frac{1}{2} \\
&= \frac{19}{81} \approx 23.46\%. \checkmark
\end{aligned}
$$

d) The expected position of the crack can be calculated by

$$
\mathbb{E}(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)dt.
$$

In our case, the integrand $t \cdot f_X(t)$ is a polynomial function containing odd exponents only, i.e. it is rotationally symmetric with respect to the origin. So, we have

$$
\begin{aligned}
\mathbb{E}(X) &= \int_{-1}^{1} t \cdot f_X(t)dt \\
&= \int_{-1}^{0} t \cdot f_X(t)dt + \int_{0}^{1} t \cdot f_X(t)dt \\
&= -\int_{0}^{1} t \cdot f_X(t)dt + \int_{0}^{1} t \cdot f_X(t)dt = 0. \checkmark
\end{aligned}
$$

The variance is computed by

$$
\begin{aligned}
\mathrm{Var}(X) &= \int_{-\infty}^{\infty} (t - \mathbb{E}(X))^2 \cdot f_X(t)dt \\
&= \int_{-1}^{1} t^2 \cdot f_X(t)dt \\
&= \int_{-1}^{1} \left(3t^6 - 3t^4 + \frac{9}{10}t^2\right) dt \\
&= \left[\frac{3}{7}t^7 - \frac{3}{5}t^5 + \frac{3}{10}t^3\right]_{-1}^{1} \\
&= 2\left(\frac{3}{7} - \frac{3}{5} + \frac{3}{10}\right) \\
&= \frac{9}{35}. \checkmark
\end{aligned}
$$