9th problem set for Probability and Statistics — April 16th

- $Exp(\lambda)$ has density $\lambda e^{-\lambda x}$, distribution function $1 - e^{-\lambda x}$, mean $1/\lambda$, and variance $1/\lambda^2$.
- $N(0,1)$ has density $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, distribution function Φ, mean 0, and variance 1.
- $N(\mu, \sigma^2)$ has density $\frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma})$, distribution function $\Phi(\frac{x-\mu}{\sigma})$, mean μ, and variance σ^2.

x	-4	-3	-2	-1	0	1	2	3	4
$\Phi(x)$	0.00003	0.00135	0.02275	0.15866	0.500000	0.84135	0.97725	0.99865	0.99997

Further values see `https://en.wikipedia.org/wiki/Standard_normal_table` – section Cumulative.

From each chapter, try at least one example! If you get stuck, there are some hints at the end.

Uniform distribution

1. Mr. Chen visited Prague and at a uniformly random time (0:00-24:00), he appears in the Old Town Square. Every hour from 9:00 to 23:00, 12 apostle figures appear on the astronomical clock.

(a) What is the probability that Mr. Chen will see the apostles without waiting for more than 15 minutes?

(b) What if Mr. Chen arrives at the Old Town Square at a uniformly random time after noon, i.e., 12:00–24:00?

Exponential distribution

2. Let's assume that at the post office counter, the time to serve one customer follows an exponential distribution with a mean of 4 minutes.

(a) What is the parameter λ, what is the distribution function?

(b) What is the probability that we will wait more than 4 minutes?

(c) What is the probability that we will wait between 3 and 5 minutes?

3. We say that a random variable X (or its distribution) is *memoryless* if

$$P(X > s + t \mid X \geq s) = P(X > t)$$

for $s, t \geq 0$. In other words, the time we have already waited does not affect the time we will still wait. In the third tutorial, we saw that the geometric distribution is memoryless. Show that the exponential distribution is also memoryless. More is true: it is the only continuous memoryless distribution on positive numbers (and the geometric distribution is the only discrete one without memory), but you don't have to prove that.

4. The duration of an oral exam follows an exponential distribution with a mean of 20 minutes. Two students are scheduled, Adam at 10:00, Beatrice at 10:20. If Adam's exam takes longer, Beatrice's exam will start when Adam finishes, otherwise it will start exactly at 10:20.

(a) What is the probability that when Beatrice arrives, Adam's exam is already over?

(b) What is the expected value of the time Beatrice will have to wait for Adam's exam to finish, if Adam is not done yet when she arrives?

(c) What is the expected time when Beatrice's examination will start?

(d) What is the expected time when Beatrice will finish her examination?

Normal distribution

5. Let $Z \sim N(0,1)$. Use the Φ function table to verify the 3σ rule, i.e., calculate

(a) $P(|Z| \leq 1)$

(b) $P(|Z| \leq 2)$

(c) $P(|Z| \leq 3)$

(d) Rewrite what this means for a random variable $X \sim N(\mu, \sigma^2)$.

6. We will model the amount of snow that will lie on the ground in a Krkonoše ski resort, on New Year's eve. We will use normal distribution with a mean of 40 (centimeters) and a standard deviation of 10.

(a) What is the probability that the model will give us a negative value for the snow cover?

(b) What is the probability that the snow cover will be between 50 and 70 cm?

Working with distribution Functions

7. We break a one-meter stick into two pieces, at a uniformly random point. Let X be the length of the longer piece.

(a) What is the distribution of X?

(b) Determine $\mathbb{E}(X)$.

8. For a certain problem, we have two algorithms, A and B. Algorithm C consists of randomly choosing which of the algorithms A or B to use – A has a probability of p, and B has a probability of $1-p$, and then using this algorithm. We interpret the running time of A, B, C as random variables, denoted X, Y, Z.

(a) Determine F_Z using F_X, F_Y.

(b) If X, Y are continuous, determine f_Z using f_X, f_Y.

Hints

2, 3: Use the formula for the distribution function of the exponential distribution.
4: b: similar to exercise 3, c: law of total expectation, d: linearity
5: You always need to subtract two suitable values in the table on the first page.
6: Convert to a statement about the random variable $N(0,1)$.
7: First calculate the distribution function of X, then its density.
8: Law of total probability also applies here.

8th homework, due next week

9. We're throwing darts at a target – a circle with a radius of 1. Let's assume that each point in the target has an equal probability of being hit, more precisely, the probability of each subset is proportional to its area. Let X be the distance from the center. (a) Find the distribution function F_X. (b) Find the density function f_X. (c) Determine $\mathbb{E}(X)$, $\mathrm{var}(X)$, σ_X.

More practice problems

10. The median lifespan of a hard disk is 7 years – approximately half of them fail after 7 years. Let's assume that this time is described by a random variable with an exponential distribution. (This is not a realistic assumption, rather a first approximation, see for example `https://www.backblaze.com/blog/how-long-do-disk-drives-last/`.)

(a) What is the probability that the disk fails within the first three years?

(b) What is the probability that it lasts at least 10 years?

(c) After how much time do 10% of the disks fail?

(d) What percentage of disks fail after 50 years? (Some pacemakers use plutonium-238 as a power source. `https://en.wikipedia.org/wiki/Plutonium-238#Nuclear_powered_pacemakers`)

11. $X \sim N(0,1)$, $Y \sim N(1,4)$.

(a) Find a linear function $f(t) = a \cdot t + b$, so that $f(Y)$ has the same distribution as X.

(b) Calculate $P(X \leq 1)$, $P(X > 2)$.

(c) Calculate $P(Y < 0)$, $P(Y > 2)$.

12. Franta's long jump measured 9 meters, exceeding the world record by 5 cm. However, during the measurement, an error was made with a distribution of $N(0, 0.01)$. What is the probability that the record was actually broken?

13. Let $X \sim N(0,1)$ and $Y = |X|$. Determine $\mathbb{E}(Y)$ and $\mathrm{var}(Y)$.

14. Plutonium-238 has a half-life of 87.7 years. Its decay will be modeled using an exponential distribution: for each atom, the time until it decays is considered as a random variable with the distribution $Exp(\lambda)$.

(a) What is λ?
(b) What is the mean lifetime of a plutonium-238 atom?
(c) After how long do 90% of the atoms decay?
(d) What percentage of atoms decay after 50 years?

15. The time to see a meteor is exponentially distributed with a mean of 1 (minute).

(a) What is the probability that we will have to wait more than 5 minutes?
(b) What is the probability that we will see one within at most one minute?
(c) * What is the distribution of the time until we see the second meteor? The third one, ... (Assuming that individual meteors are independent.)

16. Find an analogy for the "3 sigma rule,"i.e., calculate $P(|X - \mathbb{E}(X)| < c \cdot \sigma_X)$ $(c = 1, 2, 3)$, if

(a) X has a uniform distribution, (b) $X \sim Exp(1)$, (c) $X \sim Exp(2)$.