Winter term 2024/25
U. von Luxburg
E. Günther/ K. Frohnapfel

Presence sheet 10

Mathematics for Machine Learning

Tutorial of **Week 11 (07.01. - 10.01.2025)**

Exercise 1 (Convergence).

a) Consider the sequence of random variables $X_1, X_2, ...$ that follow the distribution

$$P\left(X_n = \frac{1}{n}\right) = 1 - \frac{1}{n}, \qquad P(X_n = n) = \frac{1}{n}.$$

Decide whether $X_n \xrightarrow{P} 0$.

b) Consider the sequence of random variables $X_1, X_2, ...$ on $([0,1], \mathcal{B}([0,1]), \lambda)$, where λ is the Lebesgue-measure, given by

$$X_n(\omega) = \begin{cases} 1, & \omega \leq \frac{1}{n} \\ 0, & \omega > \frac{1}{n}. \end{cases}$$

Decide whether $X_n \xrightarrow{a.s.} 0$.

Solution:

a) Let $\varepsilon > 0$. We have

$$P(|X_n| > \varepsilon) = \begin{cases} 1, & \frac{1}{n} > \varepsilon \\ \frac{1}{n}, & \frac{1}{n} \leq \varepsilon. \end{cases}$$

So, for n big enough we have $P(|X_n| > \varepsilon) = \frac{1}{n} \to 0$.

b) Let $\omega \in (0,1]$. We know $X_n(\omega) = 0$ for $n \geq 1/\omega$. Hence $X_n(\omega) \to 0$ for every $\omega \in (0,1]$. Since $P((0,1]) = 1$, we have $X_n \xrightarrow{a.s.} 0$.

Exercise 2 (Modeling). A company has a large number of regular users logging onto its website. On average, 4 users every hour fail to connect to the company's website at their first attempt.

a) What probability distribution could be used to model the number of failed connections? What goes wrong if the number of users is not that large?

b) Within a period of two hours, what is the probability of all users connecting successfully?

c) Within a period of two hours, what is the probability of at least 4 users failing to connect?

Solution:

a) A Poisson distribution is a suitable model. Note that in a Poisson distribution, each natural number has some positive probability to occur. If the number of users is rather small, high numbers are impossible, which contradicts our model. If the number of users is big enough, this is neglectable as the probability for large n is very small in a Poisson distribution.

b) For the two hour period we still use the Poisson model. Since the expected value is linear, we expect $\lambda = 8$ failures within the two hours. We compute

$$P(X = 0) = \mathrm{e}^{-8} \cdot \frac{8^0}{0!} \approx 0.0003.$$

c) Again using $\lambda = 8$, we compute

$$P(X \geq 4) = P(X \in \{0, 1, 2, 3\}) = \mathrm{e}^{-8} \sum_{k=1}^{3} \frac{8^k}{k!} \approx 0.003 \cdot \left(1 + 32 + \frac{512}{6} + \frac{4096}{24}\right) = 0.867.$$

Exercise 3 (**Some surprising result**).
There are two sealed envelopes, each of them contains some unknown amount of money. We call the amounts of money a and b. It holds $a, b \in \mathbb{N}$ and $a < b$. You choose an envelope and look at the amount of money inside. Afterwards, you may decide to either keep it or take the other one and keep that.
Consider the following strategy: You throw a coin until it shows tails. If the amount of tries you needed is larger than the amount of money in your envelope, you switch to the other one, otherwise you stay with the one you initially picked. Prove that, then, the probability of receiving the higher amount of money is strictly greater than 1/2.

Solution: Let X be the random variable indicating the amount of money in the initially picked envelope and Y the amount of money in the envelope after the potential switch. Obviously $P(X = a) = P(X = b) = 1/2$. Let further r be the number of tries needed to get tails. We then know applying the law of total probability

$$\begin{aligned}
P(Y = b) &= P(Y = b \mid r < a) \cdot P(r < a) \\
&\quad + P(Y = b \mid a < r < b) \cdot P(a < r < b) \\
&\quad + P(Y = b \mid b < r) \cdot P(b < r) \\
&= \underbrace{P(X = a)}_{=0.5} \cdot P(r < a) + 1 \cdot \underbrace{P(a < r < b)}_{\neq 0} + \underbrace{P(X = b)}_{=0.5} \cdot P(b < r) \\
&> 0.5(P(r < a) + P(a < r < b) + P(b < r)) \\
&= 0.5.
\end{aligned}$$

Exercise 4 (**Deciding about convergence based on distribution**).
Consider a sequence of random variables $X, X_1, X_2, ...$ that are Bernoulli distributed with $p = 1/2$. Which of the following is true? Prove or give a counterexample!

a) $(X_n)_{n \in \mathbb{N}}$ converges to some Ber(1/2)-distributed random variable stochastically.

b) $(X_n)_{n \in \mathbb{N}}$ does not converge to some Ber(1/2)-distributed random variable stochastically.

c) One cannot determine whether $(X_n)_{n \in \mathbb{N}}$ converges to some Ber(1/2)-distributed random variable stochastically.

Solution: Answer c) is correct. If we have $X = X_1 = X_2 = ...$ it obviously holds $X_n \xrightarrow{P} X$. If we define $X_n = \mathbb{1}_{[0,0.5]}$ and $X = \mathbb{1}_{[0.5,1]}$, we have $P(|X_n - X| > 0.5) = 1$ for all $n \in \mathbb{N}$.

Exercise 5 (**Again convergence**).
Consider a sequence of random variables $X_1, X_2, ...$ that satisfy

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N : \quad P(|X_n| \geq \varepsilon) < \varepsilon.$$

Prove that $X_n \xrightarrow{P} 0$.

Solution: By definition, X_n converges to 0 stochastically iff

$$\forall \varepsilon > 0 : P(|X_n| > \varepsilon) \overset{n\to\infty}{\longrightarrow} 0.$$

This is the case iff

$$\forall \varepsilon > 0 \quad \forall \delta > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N : \quad P(|X_n| > \varepsilon) < \delta.$$

Now consider some $\varepsilon, \delta > 0$. If $\varepsilon \geq \delta$, we now that for n big enough

$$P(|X_n| > \varepsilon) \leq P(|X_n| > \delta) < \delta$$

for n big enough. If $\varepsilon < \delta$ We know

$$P(|X_n| > \varepsilon) < \varepsilon < \delta$$

for n big enough.