

NAME: NEPTUN-CODE:

Probability theory 1 exam, 2024. Jan. 16.

Working time: 100 min. Only simple, non-programmable calculators are allowed, standard normal distribution table on the other side.

The achievable maximum score (with the Bonus exercise) is 110 points, but we consider 100 points as 100%.

- T. 1.** (a) (5 points) State and prove the law of total probability! (b) (6 points) State and prove Bayes' theorem. (c) (5 points) Assume that $E_4 \subseteq E_3 \subseteq E_2 \subseteq E_1$. Write down and prove the relation that expresses $\mathbb{P}(E_4)$ in terms of $\mathbb{P}(E_1)$, $\mathbb{P}(E_2 | E_1)$, $\mathbb{P}(E_3 | E_2)$, and $\mathbb{P}(E_4 | E_3)$.
- T. 2.** (a) (2+2+5 points) Define the GEO(p) distribution based on its intuitive meaning, and use it to calculate its weight function. Calculate the moment generating function of GEO(p).
(b) (2+3 points) Define the negative binomial distribution based on its intuitive meaning and use it to calculate its weight function.
(c) (3 points) Calculate the moment generating function of the negative binomial distribution. *Hint:* What is the moment generating function of the sum of independent random variables?
- T. 3.** (a) (4 points) Let the pair of variables (X, Y) be jointly absolutely continuous, denote their joint density function by $f(x, y)$. Define the following notions (i.e., write a formula for them using f): the probability density function f_Y of the marginal distribution Y , the conditional density function $f_{X|Y}$ of X given $Y = y$, the conditional expectation of X given $Y = y$.
(b) (13 points) The signal X emitted by a radio antenna is a random variable with $\mathcal{N}(\mu, 1)$ distribution. The received signal Y is obtained by adding a noise Z (independent from X) to the signal X , where Z has $\mathcal{N}(0, 1)$ distribution. Determine the conditional density function of X given $Y = y$, moreover calculate the conditional expectation of X given $Y = y$.
- P. 1.** (17 points) Let $f : [0, 1] \rightarrow [0, 1]$ be a continuous function, and let $I = \int_0^1 f(x)dx$. Let U_1, U_2, \dots, U_n and V_1, V_2, \dots, V_n be i.i.d. with uniform distribution on $[0, 1]$. Let X_n denote the number of indices $i \in \{1, \dots, n\}$ for which $V_i < f(U_i)$. We estimate I by X_n/n . How big should we choose n if we want the error of our estimation to be greater than 0.05 with probability at most 0.01, if
(a) (10 points) $f(x) = x^4$?
(b) (7 points) we apply this method to an unknown continuous function $f : [0, 1] \rightarrow [0, 1]$?
- P. 2.** (16 points) The chroniclers of ancient Egypt distinguished two types of years: scarce years and abundant years. Successive years are scarce or abundant, independently from each other. On the long run, half of the years are scarce and half are abundant. Observing the Egyptian chronicle for n years, let us denote by X_n the number of scarce years that are followed by a scarce year and let us denote by Y_n the number of abundant years that are followed by an abundant year. Calculate the limit of the correlation coefficient of X_n and Y_n as $n \rightarrow \infty$!
- P. 3.** Let X and Y have bivariate normal distribution. Assume that $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 0$, $\text{Var}(X) = 1$, $\text{Var}(Y) = 4$ and $\text{Cov}(X, Y) = 0$. Below, M_i , $i = 1, 2, 3, 4$ denote planar domains, and the task is to calculate $\mathbb{P}((X, Y) \in M_i)$. *Hint:* centralization (since $\mathbb{E}(X) \neq 0$) + normalization = standardization, and then draw M_i !
(a) (5 points) $M_1 = \{(x, y) : |x| \leq 1, |y| \geq 1\}$
(b) (6 points) $M_2 = \{(x, y) : 0 \leq x \leq 2, y \leq x - 1\}$
(c) (6 points) $M_3 = \{(x, y) : 4x^2 - 8x + y^2 \leq 0\}$
- Bonus:** (10 points) $M_4 = \{(x, y) : x \leq 1, y \geq 0, y \leq 2x\}$

