Statistics for Data Science — Exercise 4

Wintersemester 2023/24

Please complete these problems before the exercise session on Tuesday 14 November, 2023, 8:30. Please be prepared to present your solutions to any problems that you completed successfully.

1. Let $X \sim \mathcal{N}(0,1)$. Compute $\mathbb{E}[|X|]$ and $\mathrm{Var}(|X|)$, where $|\cdot|$ denotes the absolute value.

2. The *moment-generating function* of a real-valued random variable X is defined as the expected value
$$M_X(t) = \mathbb{E}[\mathrm{e}^{tX}], \quad t \in \mathbb{R}.$$
The moment-generating function gets its name from the fact that it can be used to find the *moments* of a random variable. Since $\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$ for all $x \in \mathbb{R}$, then by linearity of the expected value there holds
$$M_X(t) = \mathbb{E}[\mathrm{e}^{tX}] = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{E}[X^k], \quad t \in \mathbb{R},$$
so one can simply read off the value of the k^{th} moment $\mathbb{E}[X^k]$ from the coefficients of the series expansion for $M_X(t)$.
   (a) Compute the moment-generating function of $X \sim \mathcal{N}(0,1)$.
   (b) Develop the series expansion for the moment-generating function you obtained in part (a) and find $\mathbb{E}[X^{1000}]$ for $X \sim \mathcal{N}(0,1)$.

3. Let $X \sim \mathcal{N}(0,1)$. Show that the random variable X^2 has the probability density function
$$f(t) = \begin{cases} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{1}{2}t} & \text{if } t > 0, \\ 0 & \text{if } t \le 0. \end{cases} \tag{1}$$
Hint: Differentiate the expression for $\mathbb{P}(X^2 \le t)$ with respect to $t \in \mathbb{R}$.

4. Let $X \sim \mathcal{N}(0,1)$ and $Y \sim \mathcal{N}(0,1)$ be independent random variables. Show that the random variable $X^2 + Y^2$ has the probability density function
$$f(t) = \begin{cases} \frac{1}{2} \mathrm{e}^{-\frac{1}{2}t} & \text{if } t > 0, \\ 0 & \text{if } t \le 0. \end{cases}$$
Hint: Notice that both X^2 and Y^2 have the probability density function (1). Compute the probability density function for the sum of these two continuous random variables. In your calculations, you will need to simplify an integral of the form
$$\int_0^t \frac{\mathrm{d}x}{\sqrt{tx - x^2}}, \quad t > 0.$$

If you are having trouble computing the value of this integral, consider using the substitution $u = \frac{2}{t}x - 1$.

Remark. The random variables in exercises 3 and 4 are special cases of the random variable $Z_n = X_1^2 + \cdots + X_n^2$, where $X_i \sim \mathcal{N}(0,1)$ are independent for all $i \in \{1, \ldots, n\}$. The random variable Z_n is said to follow the $\chi^2(n)$ distribution with parameter $n \in \mathbb{N}$, and we write $Z_n \sim \chi^2(n)$. The probability density functions in exercises 3 and 4 correspond to $n = 1$ and $n = 2$, respectively.