Statistics for Data Science — Exercise 3
Wintersemester 2023/24
Please complete these problems before the exercise session on Tuesday 7 November, 2023, 8:30. Please be prepared to present your solutions to any problems that you completed successfully.

1. Let X be a continuous real-valued random variable with the probability density function $f\colon \mathbb{R} \to \mathbb{R}$,
$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{4x\mathrm{e}^{-x^2}}{(1+\mathrm{e}^{-x^2})^2} & \text{if } x \geq 0. \end{cases}$$
Use inverse transform sampling to draw a sample from this distribution. Use sample size $n = 10^5$ and visualize the sample as a histogram. You can validate your results by plotting the probability density function f alongside your (appropriately normalized) histogram.
Hint: In Python, you can use `numpy.random.uniform(a,b,size=n)` to draw a sample containing n entries from the uniform distribution $\mathcal{U}(a,b)$.

2. Let $X \sim \mathcal{N}(0,1)$ and define $Y = g(X)$, where $g(x) = \arctan x$ for $x \in \mathbb{R}$.
   (a) Draw a sample $X \sim \mathcal{N}(0,1)$ of size $n = 10^5$ and plot $g(X)$ as a histogram. That is, apply the mapping g *directly* to the sample drawn from the normal distribution.
   (b) Derive the probability density function of random variable Y. Plot this alongside the histogram you obtained in part (a).

   Hint: In Python, you can use `numpy.random.normal(m,sigma,size=n)` to draw a sample containing n entries from the Gaussian distribution $\mathcal{N}(m, \texttt{sigma}^2)$.

3. Let $X, Y \sim \mathcal{U}(0,1)$ be independent random variables and define $Z = \max(X, Y^2)$.
   (a) Derive the probability density function of Z.
   (b) Draw a sample of size $n = 10^5$ from the probability distribution of Z and visualize the sample as a histogram. Plot the probability density you obtained in part (a) alongside the histogram.

4. Let (X_1, X_2) be a joint random variable and let us assume that
$$(\log X_1, \log X_2) \sim \mathcal{N}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \right).$$
Derive the joint probability density of (X_1, X_2).
Hint: Recall that the bivariate Gaussian distribution $Y \sim \mathcal{N}(\mu, C)$ for vector $\mu \in \mathbb{R}^2$ and a symmetric, positive definite matrix $C \in \mathbb{R}^{2\times 2}$ has the probability density function
$$f_Y(y) = \frac{1}{2\pi\sqrt{\det C}} \exp\left(-\frac{1}{2}(y-\mu)^{\mathrm{T}} C^{-1} (y-\mu) \right), \quad y \in \mathbb{R}^2.$$