Random number generation

1. Let X be a random variable on $\mathbb{R}$ with density f_X and invertible distribution function F_X. Here f_X, F_X, and the inverse F_X^{-1} are assumed to be known. Let $\mathsf{I} = (a, b)$ be an interval such that $\mathbb{P}(X \in \mathsf{I}) > 0$.
   (a) Find the conditional distribution function $F_{X|X\in\mathsf{I}}(x) = \mathbb{P}(X \leq x|X \in \mathsf{I})$ and the density $f_{X|X\in\mathsf{I}}(x)$ of X given that $X \in \mathsf{I}$.
   (b) Find the inverse $F^{-1}_{X|X\in\mathsf{I}}$. How can this be used for simulating X conditionally on $X \in \mathsf{I}$?

2. Let X have *standard Cauchy* distribution with density

$$f(x) = \frac{1}{\pi(1 + x^2)}, \quad x \in \mathbb{R}.$$

   (a) Show how to simulate X using the inverse method.
   (b) Set $Y = 1/X$ and find the distribution of Y.

3. During Lecture 2 it was discussed how to estimate π using Buffon's needle. However, the presented computer simulation-based approach needed uniformly distributed random numbers between $-\pi/2$ and $\pi/2$ as input. Maybe this could be dealt with using the following approach?
   (a) Construct a random number generator generating random variables X with density function

$$f(x) = \frac{2}{\pi\sqrt{1 - x^2}}, \quad 0 < x < 1,$$

   without using the value of π, e.g., by using rejection sampling.
   (b) Let Z be Bernoulli variable taking the values -1 and 1 with equal probability, i.e. $\mathbb{P}(Z = 1) = \mathbb{P}(Z = -1) = 1/2$, and derive the distribution of $U = Z \arcsin(X)$. Check your result in MATLAB. Conclusion?

4. In Buffon's needle problem we actually need to simulate, for θ being uniformly distributed between $-\pi/2$ and $\pi/2$, $\cos(\theta)$ rather than θ itself. For this purpose, let U_1 and U_2 be independent and uniformly distributed between 0 and 1, and make the transformation

$$\begin{cases} Y_1 = \dfrac{U_1}{\sqrt{U_1^2 + U_2^2}}, \\ Y_2 = \sqrt{U_1^2 + U_2^2}. \end{cases}$$

   (a) Find the distribution function of $\cos(\theta)$.
   (b) Find the joint density of Y_1 and Y_2.
   (c) Find the conditional distribution function of Y_1 given $Y_2 \leq 1$. Conclusion? Propose an algorithm simulating random variables with the same distribution as $\cos(\theta)$.

5. The following method can be used for simulating two independent and $N(0, 1)$-distributed random variables: generate two independent variables U_1 and U_2 with uniform distribution between 0 and 1 and put

$$\begin{cases} X_1 = \sqrt{-2\log(U_1)}\cos(2\pi U_2), \\ X_2 = \sqrt{-2\log(U_1)}\sin(2\pi U_2). \end{cases}$$

   Show that X_1 and X_2 are independent and $N(0, 1)$-distributed.

Power production of a wind turbine

6. The amount $P(v)$ of electrical power produced by a wind turbine at a given wind speed v is described by the *power curve* of the turbine. For a 2 MW turbine we approximate the power curve with

$$P(v) = \begin{cases} 0, & v < 4, \\ \cos\left(v\frac{\pi}{11} + \frac{7\pi}{11}\right) + 1, & 4 \leq v < 15, \\ \frac{7}{4} + v\frac{1}{30} - v^2\frac{1}{900}, & 15 \leq v < 25, \\ 0, & 25 \leq v. \end{cases}$$

Of course the wind speed at a wind turbine will depend on the location of the turbine and vary during the year. Meteorological records can be used for estimating the distribution of winds in a given area. We plan to build a wind turbine at a site in northern Europe; from records we see that such stochastic wind speeds V can be modeled by a Weibull distribution

$$f(v) = \frac{k}{\lambda}\left(\frac{v}{\lambda}\right)^{k-1} \exp\left\{-\left(\frac{v}{\lambda}\right)^k\right\}, \quad v \geq 0, \tag{1}$$

with $k = 1.5$ and $\lambda = 10.5$ (in MATLAB, Weibull-distributed random numbers can be generated using `wblrnd`; in addition, the probability density and cumulative distribution functions of the Weibull distribution can be evaluated using `wblpdf` and `wblcdf`).

We now wish to investigate the potential of building a wind turbine at the site in question.

(a) Create an approximate 95% confidence interval for the expected amount of power generated by the wind turbine using draws from a (possibly truncated) Weibull distribution (here you may have use of your solution to Problem 1).

(b) Compare the above result to an approximate 95% confidence interval created by means of importance sampling based on some convenient instrumental distribution.

(c) At the site in question it is observed that extreme wind speeds occur more often than allowed by the Weibull distribution (1). It is thus proposed to replace the Weibull distribution with a distribution with somewhat heavier tail, namely

$$\tilde{f}(v) \propto v^{1/2-\beta v}, \quad v \geq 0, \tag{2}$$

where $\beta = 0.035$. Estimate the expected amount of power generated by the wind turbine when wind speeds are modeled by the distribution (2).

Two important characteristics of power plants are the *capacity factor*, or the ratio of the actual output over a time period and the maximum possible output during that time; and the *availability factor*, or the amount of time that electricity is produced during a given period divided by the length of the period. Wind turbines typically have a capacity factor of 20–40% and an availability greater than 90%. Does this seem like a good site to build a wind turbine?

The delta method

7. Recall that a sequence $(X_n)_{n\in\mathbb{N}}$ of real-valued random variables is said to converge *in probability* to some other real-valued random variable X, denoted $X_n \xrightarrow{\mathbb{P}} X$, if for all $\varepsilon > 0$,

$$\mathbb{P}\left(|X_n - X| \geq \varepsilon\right) \to 0 \quad \text{as } n \to \infty.$$

A fundamental result in probability says that convergence in probability implies convergence in distribution ($\xrightarrow{\text{d.}}$), but that these two convergence types are equivalent in the case where the limiting variable X is a deterministic constant. Moreover, *Slutsky's lemma* states that if $X_n \xrightarrow{\text{d.}} X$ and $Y_n \xrightarrow{\mathbb{P}} c$, where $c \in \mathbb{R}$ is a constant, then it holds in general that

(i) $X_n + Y_n \xrightarrow{\text{d.}} X + c$,

(ii) $X_n Y_n \xrightarrow{\text{d.}} cX$,

(iii) $X_n / Y_n \xrightarrow{\text{d.}} X/c$ (if $c \neq 0$).

The aim of this exercise is to use Slutsky's lemma to establish the *delta method* discussed in Lecture 2. For this purpose, let $(X_i)_{i\in\mathbb{N}}$ be i.i.d random variables with common distribution (the target distribution) and let $\phi : \mathbb{R} \to \mathbb{R}$ be an objective function such that $\sigma^2(\phi) = \mathbb{V}(\phi(X_1)) < \infty$. As usual, we let $\tau_n = \sum_{i=1}^n \phi(X_i)/n$ denote the basic Monte Carlo estimator of $\tau = \mathbb{E}(\phi(X_1))$.

We proceed stepwise.

(a) Let X be some nonnegative random variable and show that for all $\varepsilon > 0$,

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{1}{\varepsilon}\mathbb{E}(X).$$

(b) Use (a) to show that $\tau_n \xrightarrow{\mathbb{P}} \tau$ as $n \to \infty$ (in other words, τ_n is an asymptotically *consistent* estimator).

(c) Let $\varphi : \mathbb{R} \to \mathbb{R}$ be continuous in τ. Show that $\varphi(\tau_n) \xrightarrow{\mathbb{P}} \varphi(\tau)$.

(d) Now let $\varphi : \mathbb{R} \to \mathbb{R}$ be two times continuously differentiable in τ. Use Slutsky's lemma to show that

$$\sqrt{n}(\varphi(\tau_n) - \varphi(\tau)) \xrightarrow{\text{d.}} N(0, \varphi'(\tau)^2\sigma^2(\phi)) \quad \text{as } n \to \infty.$$