Simulation of random variables

Exercise 1 1. *Write a Python function in order to generate N values of a given linear congruential generator (the input parameters should be the multiplier a, the increment b, the modulus N and the seed y_0).*

2. *Test the following generators:*

$$y_{n+1} = 65539\, y_n \quad [2^{31}]$$
$$y_{n+1} = (129\, y_n + 907633385) \quad [2^{32}]$$
$$y_{n+1} = 16807\, y_n \quad [2^{31} - 1].$$

In each case, generate 10 000 numbers uniformly distributed on $[0, 1]$ and use these values to plot an histogram.

3. *Choose one of these generators and apply the Chi-square test with the following null hypothesis:* the distribution of the data is uniform. *In order to achieve this statistical test, split the interval $I = [0, 1]$ into 20 classes and compute the observed and the expected frequencies for each class.*

Exercise 2 1. *Write a Python function which permits to generate a random variable taking the values $x_1, \ldots, x_n$ with the corresponding probabilities $p_1, \ldots, p_n$.*

2. *Using a sample of size* 10 000, *verify that the observed frequencies are close to the probabilities p_1,...p_n.*

Exercise 3 (Poisson distribution) *Let $(U_i)_{i\geq 1}$ a sequence of independent uniformly distributed random variables.*

1. *Show that the random variable X defined by*

$$X = \sum_{i\geq 1} 1_{\{U_1 U_2 \ldots U_i \geq e^{-\lambda} > U_1 U_2 \ldots U_{i+1}\}}$$

has a Poisson distribution of parameter λ.

2. *Describe the link between the distribution of the stopping time*

$$N = \inf\{i \geq 1 : \prod_{j=1}^{i} U_j < e^{-\lambda}\}$$

and a Poisson distribution.

3. *Write a Python function for the generation of n random variables with a Poisson distribution.*

Exercise 4 1. *We aim to generate a random variable X with the Cauchy distribution*

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

Generate X using the inverse transform sampling. Is it possible to generate such a random variable using a classical distribution function (gaussian, exponential, uniform,...) as proposal distribution in an acceptance-rejection method ?

2. *Propose an acceptance-rejection algorithm in order to generate a random variable with density*

$$p(x) = \frac{2}{\sqrt{\pi}} \sqrt{x} e^{-x} 1_{\{x \geq 0\}}.$$

Use an exponential distribution $\mathcal{E}(\lambda)$ as proposal distribution. Find the optimal choice for the parameter λ.

Exercise 5 *Using the acceptance-rejection method, generate a random variable whose distribution corresponds to the triangular density on $[-a, a]$:*

$$f(x) = \frac{1}{a}\left(1 - \frac{|x|}{a}\right)_+.$$

On a same figure, represent both the histogram of the sample and the pdf.

Exercise 6 1. *Recall the classical method used to generate a gaussian r.v.*

2. *Propose an acceptance-rejection method using the proposal distribution $(\lambda/2)\exp(-\lambda|x|)$ in order to generate a gaussian r.v.*

3. *Let X and Y two independent exponentially distributed r.v. $\mathcal{E}(1)$.*

   a. *Find the conditional distribution of X given $\{Y > (1 - X)^2/2\}$.*

   b. *Let Z a r.v. with such a distribution (a) and let S an independent random variable taking the values ± 1 with probability $1/2$. Find the distribution of SZ.*

   c. *Deduce a new method for the simulation of a gaussian r.v. $\mathcal{N}(0, 1)$*

Exercise 7 *Let F a cumulative distribution function which admits the inverse function F^{-1}. Let X a random variable with distribution F.*

1. *Explain how to use a rejection method in order to simulate the conditional distribution of X given $X > m$? Estimate the efficiency of the method. What happens in particular as m becomes large?*

2. *Let U uniformly distributed on $[0, 1]$, we define:*

$$Z = F^{-1}(F(m) + (1 - F(m))U).$$

Compute the cumulative distribution function of Z and deduce a simulation method in order to generate X given $X > m$. Compare the efficiency of both methods presented so far.

3. *Generalize this method in order to simulate the conditional distribution of X given $a < X < b$.*

4. *We aim to simulate the conditional distribution of the gaussian r.v. X (distribution: $\mathcal{N}(\mu, \sigma^2)$) given $X > m$. Show that we can restrict the study to the centered and normalized case (it suffices to change the value of m).*

5. *In order to solve the previous question, propose an acceptance-rejection method based on the proposal distribution $\theta e^{-\theta(x-m)} 1_{\{x>m\}}$. What is the suitable choice for θ?*