LAB 4. Continuous random variables.

Part I. Some experiments.

Remember from Lab 5 that generating a uniformly distributed random variable U over the interval $(0,1)$ is equivalent to generating a random number in the interval $(0,1)$.

ASSIGNMENTS:

♣ **1. Estimation of areas - the Monte Carlo procedure.** Consider a pair (x, y) of numbers, each chosen independently at random from the interval $[0,1]$. Then we can interpret this pair (x, y) as the coordinates of a point chosen at random from the unit square. The point is equally likely to fall in subsets of equal area. Since the total area of the square is 1, the probability of the point falling in a specific subset E of the unit square should be equal to the area of E. Thus, we can estimate the area of any subset of the unit square by estimating the probability that a point chosen at random from this square falls in the subset.

a) Use this method to estimate the area of the region E under the curve $y = x^2$ in the unit square, by choosing a large number of points (x, y) at random and record what fraction of them fall in the region $E = \{(x, y) \ : \ y \le x^2\}$.

b) Estimate the area of the circle centered at $(0.5, 0.5)$ of radius 0.5 and deduce an estimated value of π.

c) Estimate the area of the region E under the curve $y = \dfrac{1}{x+1}$ and deduce and estimated value of $\ln(2)$?

♣ **2. Buffon's needle.** Suppose that we take a table and draw across the top surface a set of parallel lines a unit distance apart. We then drop a common needle of unit length at random on this surface and observe whether or not the needle lies across one of the lines. We can describe the possible outcomes of this experiment by coordinates as follows: Let d be the distance from the center of the needle to the nearest line. Next, let L be the line determined by the needle, and define θ as the acute angle that the line L makes with the set of parallel lines. Using this description, we have $0 \le d \le 1/2$, and $0 \le \theta \le \pi/2$. Moreover, we see that the needle lies across the nearest line if and only if the hypotenuse of the triangle (see Figure 1) is less than half the length of the needle, that is,

$$\frac{d}{\sin\theta} \le \frac{1}{2}$$

Now we assume that when the needle drops, the pair (θ, d) is chosen at random from the rectangle $0 \le \theta \le \pi/2$, $0 \le d \le 1/2$. Therefore, the probability of the event $E =$ "The needle intersects a line" is equal to the probability that the randomly generated pair of numbers (θ, d) satisfies $d \le 1/2 \sin(\theta)$ and represents a fraction of the area of the rectangle. Using this experiment, give an estimation of π.

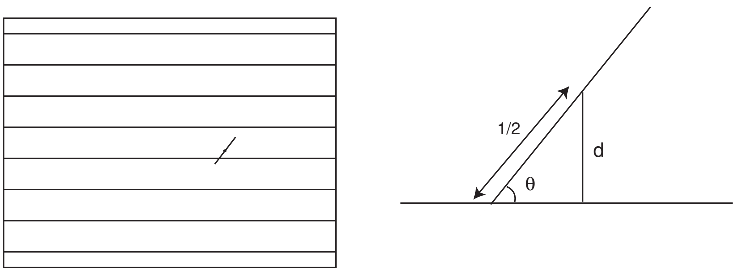

Figure 1: Buffon's needle experiment

Part II. Simulating continuous random variables.

The Inverse Transformation Method

Let U be a uniform random variable over the interval $[0, 1]$. For any continuous distribution function F, the random variable

$$X = F^{-1}(U)$$

has the distribution function F.

ASSIGNMENT:

♣ **3. Simulating an exponentially distributed random variable.**

In this case, $F(x) = 1 - e^{-\lambda x}$ and $F^{-1}(u) = -\dfrac{\ln(1-u)}{\lambda}$. Hence, if U is uniformly distributed, the random variable $X = F^{-1}(U)$ will be exponentially distributed, $X \sim Exp(\lambda)$. Write a procedure which generates the exponentially distributed random variable X and test it.

♣ **4. Simulating the Pareto distribution.** The Pareto distribution with parameter $a = 3$ has PDF $f(x) = a/x^{a+1}$ for $x \geq 1$ (and 0 otherwise). This distribution is often used in statistical modeling.

(a) Find the CDF of a Pareto r.v. with parameter $a = 3$; find the inverse of this CDF (both on paper).

(b) Suppose that for a simulation you want to run in R, you need to generate i.i.d. Pareto(3) r.v.s. You have a computer that knows how to generate i.i.d. $Unif(0, 1)$ r.v.s but does not know how to generate Pareto r.v.s. Show how to do this in R.

♣ **5. Simulating the Logistic distribution.** Let $U \sim Unif(0, 1)$ and

$$X = \log \frac{U}{1-U}.$$

Then X has the Logistic distribution. Write a function that simulates the Logistic distribution and use it to compute an approximation for $E[X]$.