

4 The Inverse CDF Method

Let $U \sim U(0, 1)$. Suppose $F(x)$ is a well-defined CDF which is invertible. Then the random variable $X = F^{-1}(U)$ has CDF $F(x)$.

Proof:

Example 4.1: Suppose we obtain two observations from a $U(0, 1)$ distribution: 0.1 and 0.85. Use these values to obtain two observations from $X \sim \text{Exp}(2)$.

Example 4.2: Write some R code to sample n observations from an $Exp(2)$ distribution. Do not use the built-in `rexp` function. Check your code by comparing a histogram of some simulated realisations against the PDF for an $Exp(2)$ distribution. Also check your code against results from the built-in `rexp` function.

Example 4.3: The *Pareto distribution*, named after the Italian civil engineer, economist, and sociologist Vilfredo Pareto, is a power-law probability distribution that is used in the description of social, scientific, geophysical, actuarial, and many other types of observable phenomena. It was originally used to describe the distribution of wealth in a society, fitting the trend that a large portion of wealth is held by a small fraction of the population.

If $X \sim \text{Pareto}(\sigma)$, then it has cumulative distribution function

$$F_X(x; \sigma) = 1 - (\sigma/x)^{\log_4 5}, \quad x \geq \sigma, \sigma > 0.$$

Write a function in R to produce realisations from this distribution, and test your function out for $\sigma = \{1, 2, 5\}$.

Example 4.4: The *Weibull distribution* is often used in survival data analysis to model lifetimes (e.g. the lifetimes of patients receiving a particular drug). If $X \sim \text{Weibull}(\lambda, \kappa)$, then it has cumulative distribution function

$$F_X(x; \lambda, \kappa) = 1 - e^{-(x/\lambda)^\kappa}, \quad x \geq 0, \lambda, \kappa > 0.$$

Write a function in R to produce realisations from this distribution, and test your function out (a recent study suggested that, for a particular cancer drug, $\lambda = 2$ and $\kappa = 1.5$ for survival times in months after starting this drug).

Example 4.5: Patients arriving at a doctor's surgery have appointment times 15 minutes apart. The random variable L measures how late a patient is for his/her appointment, with negative values denoting early arrives. The PDF of L is

$$f_L(\ell) = \begin{cases} 0.0025(\ell + 20), & -20 \leq \ell \leq 0, \\ 0.05e^{-\ell/10}, & \ell > 0. \end{cases}$$

- (a) Sketch $f_L(\ell)$.
- (b) Show that the CDF of L is

$$F_L(\ell) = \begin{cases} 0, & \ell < -20, \\ 0.00125(\ell + 20)^2, & -20 \leq \ell \leq 0, \\ 1 - \frac{e^{-\ell/10}}{2}, & \ell > 0. \end{cases}$$

- (c) Suppose that 0.205 and 0.713 are two realized values of $U \sim U(0, 1)$. Show that the corresponding realized values of L are -7.19 and 5.55 under the inverse CDF simulation method.

