Chapter 2
Generating Random Numbers

To perform a Monte Carlo approximation, we have to generate random variables (rv.) on a computer according to a given df. F. In this chapter, we will discuss some commonly used procedures and their application under R.

Since most of the widely used distributions are implemented in R, random variables according to these distributions can easily be generated directly in R through the corresponding built-in R functions. In the first section of this chapter, we will give a brief overview on those distributions which are implemented in the R `stats` package.

However, if a specific distribution is needed which is neither supported by R itself nor by any additional package, one can try the "quantile transformation method" or the "method of rejection". Both approaches are considered in this chapter. For a detailed discussion of random number generation, we refer to Devroye (1986) and Ripley (1987). In Eubank and Kupresanin (2011, Chapter 4), this is also considered in the R-context.

2.1 Distributions in the R-Package Stats

The standard R-package `stats` contains several standard probability distributions. We can list them from a R-workspace by typing the command

```
help(distributions)
```

For all these distributions, the corresponding cumulative distribution function, density function, quantile function, and random generation function are implemented and can be called by

- dxxx(...)—density function;
- pxxx(...)—distribution function;

G. Dikta and M. Scheer, *Bootstrap Methods*,
https://doi.org/10.1007/978-3-030-73480-0_2

- qxxx(. . .)—quantile function; and
- rxxx(. . .)—random number generator function.

In the notation above, “xxx” is the name in R of the corresponding distribution and (. . .) a placeholder for the required parameters of the function call. The following example lists some calls regarding a normal distribution with expected value $\mu = 2$ and variance $\sigma^2 = 4$, here abbreviated as $\mathcal{N}(2, 4)$.

R-Example 2.1 Note that in the corresponding function calls under R, the standard deviation (sd=2) is used while in the notation $\mathcal{N}(2, 4)$ the variance $\sigma^2 = 4$ is given. The R name “xxx” of the normal distribution is “norm”.

```
#call the help for rnorm
help(rnorm)
#density function at x = 2
dnorm(x = 2, mean = 2, sd = 2)

  ## [1] 0.1994711

#distribution function at q = 2
pnorm(q = 2, mean = 2, sd = 2)

  ## [1] 0.5

#0.5-quantile
qnorm(p = 0.5, mean = 2, sd = 2)

  ## [1] 2

#3 normal random variables
rnorm(n = 3, mean = 2, sd = 2)

  ## [1] 1.008104 3.768502 2.064348
```

2.2 Uniform df. on the Unit Interval

A rv. U is *uniformly* distributed on the interval $[a, b]$, where $-\infty < a < b < \infty$, if

$$\mathbb{P}(U \le u) = \begin{cases} 0 & : u < a \\ (u-a)/(b-a) & : a \le u \le b \\ 1 & : u > b. \end{cases}$$

We denote this distribution here by $UNI(a, b)$ and use UNI to abbreviate UNI $(0, 1)$, the *standard uniform distribution*, which is also referred to as the *uniform distribution*.

The uniform distribution is the most important one in generating rv. As we will see in the next section, we can generate a rv. $X \sim F$ for every df. F if we can generate a rv. $U \sim UNI$.

There is a large literature on generating sequences of independent and uniformly distributed rvs. which we will not discuss here. Eubank and Kupresanin (2011, Chapter 4) is a good reference for pseudo-random number generators (PRNG), which specifically addresses R.

Remark 2.2 In this manuscript, we usually take “Mersenne twister” as PRNG. If a normally distributed rv. is to be created, this is done using the “inversion” method. For reasons of reproducibility, a starting value (“set.seed”) is set before each simulation. This seed also contains the name of the PRNG used and the name of the method for generating normal distributed rvs. A typical call looks like

```
set.seed(123,kind ="Mersenne-Twister",normal.kind ="Inversion")
```

With the `simTool` package, simulations can also be run in parallel. In this case, “L’Ecuyer-CMRG” is set globally as PRNG!

2.3 The Quantile Transformation

The following theorem says that we can generate a rv. X according to an arbitrary df. F if we apply a certain transformation to a generated rv. $U \sim UNI$.

Theorem 2.3 *Let F be the df. of a rv. X and define for $0 < u < 1$*

$$F^{-1}(u) = \inf\{x \in \mathbb{R} \mid F(x) \geq u\} \tag{2.1}$$

the quantile function (qf.). If $U \sim UNI$ then:

$$X := F^{-1}(U) \sim F.$$

Proof At first note that the qf. equals the inverse function of F if F is strictly increasing. If this is not the case, F^{-1} is still well defined and therefore qf. is a generalized inverse of an increasing function.

We have to show that

$$\mathbb{P}(F^{-1}(U) \leq x) = F(x), \quad \forall x \in \mathbb{R}.$$

For this choose $x \in \mathbb{R}$ and $0 < u < 1$ arbitrarily. Then the following equivalence holds:

$$F^{-1}(u) \leq x \iff u \leq F(x) \tag{2.2}$$

"$\Leftarrow$:" If $u \leq F(x)$, apply the definition of F^{-1} to get $F^{-1}(u) \leq x$.
"$\Rightarrow$:" Assume now $F^{-1}(u) \leq x$ and continue indirectly. For this assume further that $u > F(x)$. Since F is continuous from above there exists $\varepsilon > 0$ such that $u > F(x+\varepsilon)$. Apply the definition of F^{-1} to get $F^{-1}(u) \geq x + \varepsilon$. This contradiction leads to $F^{-1}(u) \leq x$.
Now, apply (2.2) to get for arbitrary $x \in \mathbb{R}$:

$$\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x) = \mathbb{P}(X \leq x),$$

where the second equality follows from $U \sim UNI$. This finally proves the theorem. □

Example 2.4 Let $U \sim UNI$ and

$$F(x) := \begin{cases} 0 & : x \leq 0 \\ 1 - \exp(-\alpha\, x) & : x > 0 \end{cases}$$

the df. of the exponential distribution with parameter $\alpha > 0$, abbreviated by $EXP(\alpha)$. Calculate the inverse of F to get

$$F^{-1}(u) = -\frac{\ln(1-u)}{\alpha}.$$

The last theorem guarantees that $F^{-1}(U) \sim EXP(\alpha)$.

R-Example 2.5 This example shows the generation of 1000 $EXP(2)$ variables with R based on the quantile transformation derived in Example 2.4.

```r
gen.exp <- function(n, alpha){
  #n - number of observations
  #alpha - distribution parameter

  return(-log(1 - runif(n)) / alpha)
}

# set the seed for the pseudo random number generator
# for reproducible results
set.seed(123,kind ="Mersenne-Twister",normal.kind ="Inversion")

# generate 1000 EXP(2) random variables
obs <- gen.exp(n = 1000, alpha = 2)

# draw a histogram with 50 cells
hist(obs, breaks = 50, freq = FALSE,
     main = "Histogram of 1000 EXP(2)",
     xlab = "", ylab = "density",
```

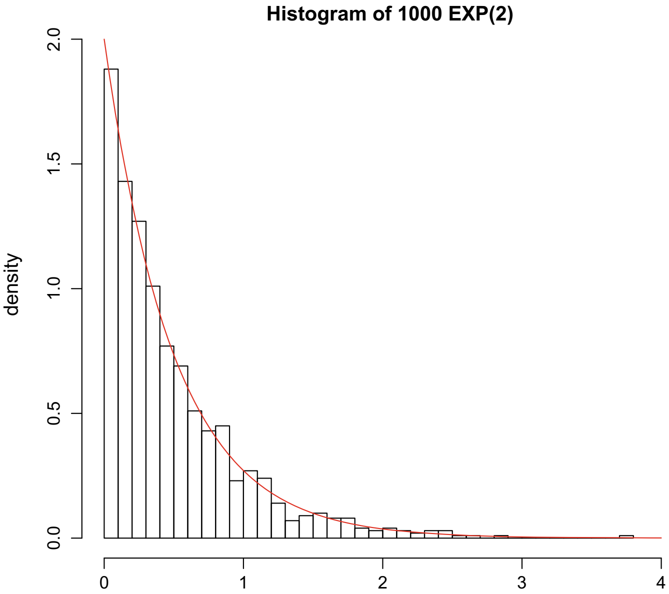

Fig. 2.1 Histogram of 1000 *EXP*(2) distributed random variables and the *EXP*(2)-density

```
      xlim=c(0,4),
      ylim=c(0,2))

# add the density function of a EXP(2) distributed random
# variable to the plot
curve(dexp(x, rate = 2), add = TRUE, col = "red")
```

In the first statement, the R-function "gen.exp" is defined with two parameters; n and $alpha$, which implements the result derived in Example 2.4. It returns a vector of n independent realizations of the $EXP(alpha)$ distribution. In the second statement, the seed is set for the pseudo-random number generator which is here "Mersenne-Twister" to obtain reproducible results. "gen.exp" is applied with $n = 1000$ and $alpha = 2$. The resulting vector is stored in the variable "obs" in statement three. With the fourth statement a histogram of the generated variables is produced and with the last statement this histogram is overlaid with the true density function of the $EXP(2)$ distribution, see Fig. 2.1.

In the following lemma, some further properties of the quantile function are listed:

Lemma 2.6 *Let F be an arbitrary df. and denote by* F^{-1} *the corresponding quantile function. We have for* $x, x_1, x_2 \in \mathbb{R}$ *and* $0 < u < 1$*:*

1. $F(x) \geq u \iff F^{-1}(u) \leq x.$
2. $F(x) < u \iff F^{-1}(u) > x.$

3. $F(x_1) < u \leq F(x_2) \iff x_1 < F^{-1}(u) \leq x_2.$

Proof (i) Already shown under (2.2) of Theorem 2.3.
(ii) Consequence of part (i).
(iii) Consequence of part (i) and (ii).

□

Lemma 2.7 *Let F be an arbitrary df. and $0 < u < 1$. Then*

$$F \circ F^{-1}(u) \geq u.$$

If $u \in F(\mathbb{R})$ the inequality above changes to an equality.

Proof The inequality can be obtained from Lemma 2.6 (ii), since $F \circ F^{-1}(u) < u$ would result in the obvious contradiction $F^{-1}(u) > F^{-1}(u)$.

Now assume in addition that $u \in F(\mathbb{R})$, i.e., there exists $x \in \mathbb{R}$ such that $u = F(x)$. Therefore, by definition of F^{-1}, we get $F^{-1}(u) \leq x$. Applying F to both sides of this inequality, the monotony of F implies that $F \circ F^{-1}(u) \leq F(x) = u$. Thus, $F \circ F^{-1}(u) > u$ is not possible and according to the first part of the proof we get $F \circ F^{-1}(u) = u$. □

Corollary 2.8 *Let X be a rv. with continuous df. F. Then*

$$F(X) \sim UNI.$$

Proof According to Theorem 2.3, we can assume that

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

where $U \sim UNI$. Thus, it remains to show that $F \circ F^{-1}(U) \sim UNI$. For this choose $0 < u < 1$ arbitrarily. Then continuity of F and the last lemma leads to

$$\mathbb{P}(F \circ F^{-1}(U) \leq u) = \mathbb{P}(U \leq u) = u$$

which proves the corollary. □

We finalize the section by another inequality of the quantile function.

Lemma 2.9 *For each df. F and $x \in \mathbb{R}$, we have*

$$F^{-1} \circ F(x) \leq x.$$

If in addition x fulfills the extra condition that for all $y < x$, $F(y) < F(x)$ holds, then the inequality above changes to an equality.

Proof If $F^{-1} \circ F(x) > x$ for $x \in \mathbb{R}$, then Lemma 2.6 (ii) immediately yields the contradiction $F(x) < F(x)$. Thus, the inequality stated above is correct.
Now, assume the extra condition of the lemma for the point $x \in \mathbb{R}$. According to the part just shown, we have to prove that $F^{-1}(F(x)) < x$ cannot be correct. Assuming that this inequality is correct, Lemma 2.7 implies

$$F(x) \leq F \circ F^{-1}(F(x)) < F(x)$$

which is obviously a contradiction. □

2.4 The Normal Distribution

Theorem 2.3 of the last section shows how the quantile function can be used to generate a rv. according to a given df. F. However, the quantile function might be difficult to calculate. Therefore, the procedure suggested under Theorem 2.3 is only used in standard situations where F is invertible and the inverse can easily be obtained. In those cases where it is not possible to calculate F^{-1} directly, other procedures should be applied.
In the case of the standard normal distribution, i.e., the rv. $X \sim \mathcal{N}(0, 1)$, the df. Φ has the density ϕ with

$$\Phi(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} \phi(t)\,\mathrm{d}t = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t$$

which can be obtained only numerically. Thus, quantile transformation is not applicable to generate such a rv.
As the next lemma will show, we can generate a rv. $Z \sim \mathcal{N}(\mu, \sigma^2)$, i.e., Z has df. F with

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \mathrm{d}t, \tag{2.3}$$

through a linear transformed rv. $X \sim \mathcal{N}(0, 1)$.

Lemma 2.10 *Let $X \sim \mathcal{N}(0, 1)$. Then $Z := \sigma \cdot X + \mu$ is distributed according to $\mathcal{N}(\mu, \sigma^2)$.*

Proof Let $\mu \in \mathbb{R}$, $\sigma > 0$, and $z \in \mathbb{R}$ be given. Then

$$\begin{aligned} \mathbb{P}(Z \leq z) &= \mathbb{P}(\sigma X + \mu \leq z) = \mathbb{P}(X \leq (z-\mu)/\sigma) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(z-\mu)/\sigma} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t. \end{aligned}$$

Now, differentiate both sides w.r.t. z to obtain by the chain rule and the Fundamental Theorem of Calculus the density function

$$f(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right).$$

But f is precisely the density function of a rv. which is $\mathcal{N}(\mu, \sigma^2)$ distributed. □

In the next theorem, the *Box-Muller* algorithm to generate $\mathcal{N}(0, 1)$ distributed rv. is given.

Theorem 2.11 Box-Muller algorithm. *Let $U, V \sim UNI$ be two independent rv. uniformly distributed on the unit interval. Then the rv.*

$$X = \sqrt{-2\log(U)}\cos(2\pi V), \quad Y = \sqrt{-2\log(U)}\sin(2\pi V)$$

are independent from one another and both are $\mathcal{N}(0, 1)$ distributed.

Proof The proof is omitted here but can be found in Box and Muller (1958). □

2.5 Method of Rejection

As already discussed in the last section, quantile transformation is not always applicable in practise. In this section, we discuss a method which is applicable in a situation where the df. F has a density function f.

Theorem 2.12 Method of Rejection. *Let F, G be df. with probability density functions f, g. Furthermore, let $M > 0$ be such that*

$$f(x) \le Mg(x), \qquad \forall x \in \mathbb{R}.$$

To generate a rv. $X \sim F$ perform the following steps:

(i) Generate $Y \sim G$.
(ii) Generate $U \sim UNI$ independent of Y.
(iii) If $U \le f(Y)/(M \cdot g(Y))$, return Y. Else reject Y and start again with step (i).

Proof We have to prove that $X \sim F$. Note first that

$$\mathbb{P}(X \le x) = \mathbb{P}\left(Y \le x \,\middle|\, U \le \frac{f(Y)}{M \cdot g(Y)}\right) = \frac{\mathbb{P}\left(Y \le x,\, U \le \frac{f(Y)}{M \cdot g(Y)}\right)}{\mathbb{P}\left(U \le \frac{f(Y)}{M \cdot g(Y)}\right)}.$$

For the numerator on the right-hand side, we obtain by conditioning w.r.t. Y

$$\begin{aligned}\mathbb{P}\left(Y \le x,\, U \le \frac{f(Y)}{M \cdot g(Y)}\right) &= \int_{-\infty}^{x} \mathbb{P}\left(U \le \frac{f(Y)}{M \cdot g(Y)} \,\middle|\, Y = y\right) G(\mathrm{d}y) \\ &= \int_{-\infty}^{x} \mathbb{P}\left(U \le \frac{f(y)}{M \cdot g(y)}\right) G(\mathrm{d}y),\end{aligned}$$

where the last equality follows from the independence of U and Y. Since $U \sim UNI$, the last integral is equal to

$$\int_{-\infty}^{x} \frac{f(y)}{M \cdot g(y)}\, g(y)\mathrm{d}y = \frac{1}{M}\int_{-\infty}^{x} f(y)\,\mathrm{d}y = \frac{F(x)}{M}.$$

Since the denominator is the limit of the numerator for $x \to \infty$ and $F(x) \to 1$ for $x \to \infty$, the denominator must be identical to $1/M$. This finally proves the theorem. □

Generally, one chooses the rv. $Y \sim G$ in such a way that Y can be easily generated by quantile transformation. The constant $M > 0$ should then be chosen as small as possible to minimize the cases of rejection.
In the following example, we apply the rejection method to generate a rv. $X \sim \mathcal{N}(0, 1)$. For the df. G, we choose the Cauchy distribution given under Exercise 2.16.

Example 2.13 At first, we have to find a proper constant M

$$\frac{f(x)}{g(x)} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \Big/ \left(\frac{1}{\pi(1+x^2)}\right) = \sqrt{\pi/2}\exp\left(-\frac{x^2}{2}\right)(1+x^2).$$

The function $\exp\left(-\frac{x^2}{2}\right)(1+x^2)$ is symmetric around 0 and has a global maximum at $x = 1$. Thus, the constant

$$M := \frac{2\sqrt{\pi/2}}{\sqrt{\mathrm{e}}} = \sqrt{\frac{2\pi}{\mathrm{e}}}$$

can be used.

R-Example 2.14 The results of the last example can be implemented in R like

```
set.seed(123,kind ="Mersenne-Twister",normal.kind ="Inversion")
gen.norm.rm <- function(n){
  # n - number of observations

  # constant used during the method of rejection
  M = sqrt(2 * pi * exp(-1))

  # actual method of rejection, returning one observation
  MethodOfRejection <- function() {
    repeat{
      Y = rcauchy(1)
      if(runif(1) <= dnorm(Y) / (M * dcauchy(Y)))
        return(Y)
    }
  }
```

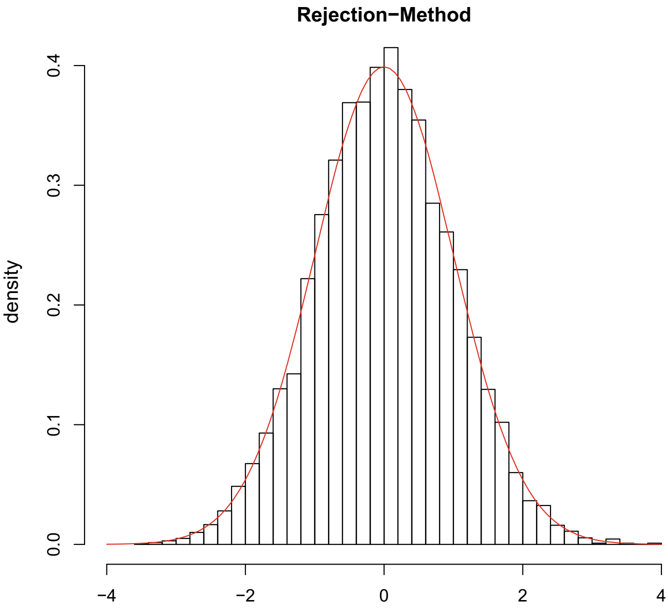

Fig. 2.2 Histogram of 10000 $\mathcal{N}(0, 1)$ rvs. generated with the rejection method and the $\mathcal{N}(0, 1)$-density

```
  # calling MethodOfRejection n times
  replicate(n, MethodOfRejection())
}
obs <- gen.norm.rm(n = 10000)
hist(obs, breaks = 50, freq = FALSE, xlab = "", xlim=c(-4,4),
                ylab = "density",
        main = "Rejection-Method")
curve(dnorm(x), col = "red", add = TRUE)
```

In the source code above, we define the function “gen.norm.rm” which returns a vector of n independent standard normal rvs. by applying the rejection method as described in Example 2.13. The function is called with $n = 10000$ and the result is stored in the variable “obs”. The last two lines produce the histogram in Fig. 2.2. Within “gen.norm.rm” the functions “rcauchy”, “runif”, “dnorm”, and “dcauchy” from the `stats` library are called. For the meaning of these functions, compare Sect. 2.1.

2.6 Generation of Random Vectors

In this section, we discuss the generation of two-dimensional random vectors (X, Y). If the variables are independent of one another, the methods stated above can be used. Thus, the remaining difficulty is the generation of dependent rv.

In the case of a known regression function, e.g.,

$$Y = f(X) + \varepsilon,$$

where X is independent of ε and $\mathbb{E}(\varepsilon) = 0$, we can also use the methods described above. To be precise, we generate the rv. X and ε independent of one another and substitute the results into the right-hand side of the regression equation above to obtain the rv. Y. Finally, (X, Y) is returned.

If no regression function is given but the regular conditional df. of Y given $X = x$ is known for each $x \in \mathbb{R}$, i.e.,

$$\mathscr{B}^* \ni B \longrightarrow \mathbb{P}(Y \in B \mid X = x),$$

where $\mathscr{B}^*$ denotes the Borel sets, then the *Rosenblatt Transformation* can be applied. For this transformation, let $G(y \mid x) := \mathbb{P}(Y \leq y \mid X = x)$ denote the conditional df. of Y given $X = x$ and F the df. of X. Then

$$(X, Y) \sim (F^{-1}(U), G^{-1}(V \mid F^{-1}(U))), \tag{2.4}$$

where U, V are independent rv. which are uniformly distributed on the unit interval.

Proof The proof is based on some standard operations of conditional distributions.

$$\begin{aligned}
\mathbb{P}\Big(F^{-1}(U) \leq t, G^{-1}\big(V \mid F^{-1}(U)\big) \leq y\Big) &= \int_{-\infty}^{t} \mathbb{P}(G^{-1}(V \mid x) \leq y)\, F(dx) \\
&= \int_{-\infty}^{t} \mathbb{P}(V \leq G(y \mid x))\, F(dx) \\
&= \int_{-\infty}^{t} G(y \mid x)\, F(dx) \\
&= \int_{-\infty}^{t} \mathbb{P}(Y \leq y \mid X = x)\, F(dx) \\
&= \mathbb{P}(X \leq t, Y \leq y).
\end{aligned}$$

□

2.7 Exercises

Exercise 2.15 Assume that $X_1, \ldots, X_n$ is an i.i.d. sequence of rvs. with common continuous df. F and let F_n denote the associated empirical distribution function; compare with (1.9).

(i) Determine $n\, F_n(X_i)$ and $F_n^{-1}(i/n)$, for $1 \leq i \leq n$.
(ii) Find the distribution of $F_n^{-1}(U)$ given the observations $X_1, \ldots, X_n$, if $U \sim UNI$ is independent of the sequence.
(iii) Implement a R-function to generate rvs. according to F_n.

Exercise 2.16 The density function of the Cauchy distribution is defined by

$$\mathbb{R} \ni x \longrightarrow f(x) := \frac{1}{\pi(1+x^2)}.$$

Determine the corresponding df. F and F^{-1}.

Exercise 2.17 The Weibull distribution to the parameter (α, β), where $\alpha > 0$ and $\beta > 0$, abbreviated by $WEIB(\alpha, \beta)$, possess the df.

$$F(x) := \begin{cases} 1 - \exp(-(x/\alpha)^{\beta}) & : x \geq 0 \\ 0 & : \text{otherwise} \end{cases}$$

(i) Use the quantile transformation to define a procedure for generating Weibull distributed rvs.
(ii) Implement your Weibull generator in R.
(iii) Generate 10000 independent $WEIB(2, 2)$ variables in R with your generator and visualize the result in a histogram together with the corresponding density function.

Exercise 2.18 Let f, g be the pdfs. in the rejection method and $M > 0$ the corresponding constant. Determine the probability that the rejection method succeeds in the first step, i.e., no rejection.

References

Box GEP, Muller ME (1958) A note on the generation of random normal deviates. Ann Math Stat 29(2):610–611

Devroye L (1986) Non-uniform random variate generation. Springer, New York

Eubank R, Kupresanin A (2011) Statistical computing in C++ and R. Chapman and Hall/CRC Press, New York

Ripley BD (1987) Stochastic simulation. Wiley series in probability and mathematical statistics. Wiley, New York