Chapter Four
SPECIALIZED ALGORITHMS

1. INTRODUCTION.

1.1. Motivation for the chapter.

The main techniques for random variate generation were developed in chapters II and III. These will be supplemented in this chapter with a host of other techniques: these include historically important methods (such as the Forsythe-von Neumann method), methods based upon specific properties of the uniform distribution (such as the polar method for the normal density), methods for densities that are given as convergent series (the series method) and methods that have proven particularly successful for many distributions (such as the ratio-of-uniforms method).

To start off, we insert a section of exercises requiring techniques of chapters II and III.

1.2. Exercises.

1. Give one or more reasonably efficient methods for the generation of random variates from the following densities (which should be plotted too to gain some insight):

Density	Range for x	Range for the parameter(s)
$(\pi\log(\frac{1}{x}))^{-\frac{1}{2}}$	$0<x<1$	
$2\sqrt{\frac{1}{\pi}\log(\frac{1}{x})}$	$0<x<1$	
$\frac{4}{\pi^2 x}\log(\frac{1+x}{1-x})$	$0<x<1$	
$\frac{8}{\pi^2(1-x^2)}\log(\frac{1}{x})$	$0<x<1$	
$\frac{2e^{2a}}{\sqrt{2\pi}}e^{-x^2-\frac{a^2}{x^2}}$	$x>0$	$a>0$
$\frac{4x^2}{\sqrt{\pi}}e^{-x^2}$	$x\geq 0$	
$\sqrt{\frac{\theta\pi}{x}}e^{-\theta x}$	$x>0$	$\theta>0$

2. Write short and fast programs for generating random variates with the densities given in the table below. In the programs, use only uniform [0,1] and/or uniform [-1,1] random variates.

Density	Range for x	Range for the parameter(s)
$\frac{n}{n-1}(1-x^{n-1})$	$0\leq x\leq 1$	$n\geq 2$, n integer
$\frac{1}{2x^4}e^{-\frac{1}{x}}$	$x>0$	
$\frac{2}{e^{\pi x}+e^{-\pi x}}$	$x\in R$	
$\frac{4\log(2x-1)}{\pi^2(x-1)x}$	$x>1$	

3. Write one-line generators (i.e., assignment statements) for generating random variates with densities as described below. You can use log,exp,cos,atan,max,min and functions that generate uniform [0,1] and normal random variates.

Density	Range of x	Range of the parameter(s)
$\frac{(-\log x)^n}{n!}$	$0<x<1$	n positive integer
$\frac{1}{2}e^{-\lvert x\rvert}$	$x\in R$	
$\frac{2}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{n-1}e^{-\frac{x^2}{2}}$	$x>0$	n positive integer
$\frac{1}{2+e^x+e^{-x}}$	$x\in R$	
$a-(2a-2)x$	$0\leq x\leq 1$	$1\leq a\leq 2$

In number 2 we recognize the Laplace density. Number 4 is the logistic density.

4. Show how one can generate a random variate of one's choice having a density f on $[0,\infty)$ with the property that $\lim_{x \downarrow 0} f(x)=\infty$, $f(x)>0$ for all x.
5. Give random variate generators for the following simple densities:

Density	Range for x
$\frac{6}{\pi^2}\frac{x}{e^x-1}$	$x>0$
$\frac{12}{\pi^2}\frac{x}{e^x+1}$	$x>0$
$\frac{6}{\pi^2}\frac{\log(\frac{1}{x})}{1-x}$	$0<x<1$
$\frac{12}{\pi^2}\frac{\log(1+x)}{x}$	$0<x<1$
$\frac{\arctan(x)}{Gx}$	$0<x<1$
$\frac{\log(\frac{1}{x})}{G(1+x^2)}$	$0<x<1$
$\frac{2\tan(x)}{\pi x}$	$x\geq 0$
$\frac{2}{\pi}(\frac{\sin(x)}{x})^2$	$x\geq 0$

Here G is Catalan's constant (0.9159655941772190...).

6. Find a direct method (i.e., one not involving rejection of any kind) for generating random variates with distribution function $F(x)=1-e^{-ax-bx^2-cx^3}$ $(x\geq 0)$, where $a,b,c>0$ are parameters.
7. Someone shows you the rejection algorithm given below. Find the density of the generated random variate. Find the dominating density used in the rejection method, and determine the rejection constant.

```
REPEAT
    Generate iid uniform [0,1] random variates U1,U2,U3.
UNTIL U3(1+U1U2)≤1
RETURN X←-log(U1U2)
```

8. Find a simple function of two iid uniform [0,1] random variates which has distribution function $F(x)=1-\frac{\log(1+x)}{x}$ $(x>0)$. This distribution function is important in the theory of records (see e.g. Shorrock, 1972).
9. Give simple rejection algorithms with good rejection constants for generating discrete random variates with distributions determined as follows:

p_n	Range for n
$\frac{4}{\pi}\arctan(\frac{1}{2n^2})$	$n \geq 1$
$\frac{8}{\pi}\frac{1}{(4n+1)(4n+3)}$	$n \geq 0$
$\frac{8}{\pi^2}\frac{1}{(2n+1)^2}$	$n \geq 0$
$\frac{4}{\pi}\arctan(\frac{1}{n^2+n+1})$	$n \geq 1$

10. **The hypoexponential distribution.** Give a uniformly fast generator for the family of hypoexponential densities given by

$$f(x) = \frac{\lambda\mu}{\mu-\lambda}(e^{-\lambda x} - e^{-\mu x}) \quad (x > 0) ,$$

where $\mu > \lambda > 0$ are the parameters of the distribution.

2. THE FORSYTHE-VON NEUMANN METHOD.

2.1. Description of the method.

In 1951, von Neumann presented an ingenious method for generating exponential random variates which requires only comparisons and a perfect uniform [0,1] random variate generator. The exponential distribution is entirely obtained by manipulating the outcomes of the comparisons. Forsythe (1972) later generalized the technique to other distributions, albeit at the expense of simplicity since the method requires more than just comparisons. The method was then applied with a great deal of success in normal random variate generation (Ahrens and Dieter, 1973; Brent, 1974) and even in beta and gamma generators (Atkinson and Pearce, 1976). Unfortunately, in the last decade, most of the algorithms based on the Forsythe-von Neumann method have been surpassed by other algorithms partially due to the discovery of the alias and acceptance-complement methods. The method is expensive in terms of uniform [0,1] random variates unless special "tricks" are used to reduce the number. In addition, for general distributions, there is a tedious set-up step which makes the algorithm virtually inaccessible to the average user.

Just how comparisons can be manipulated to create exponentially distributed random variables is clear from the following Theorem.

Theorem 2.1.

Let $X_1, X_2, \ldots$ be iid random variables with distribution function F. Then:

(i) $P(x \geq X_1 \geq \cdots \geq X_{k-1} < X_k) = \frac{F(x)^{k-1}}{(k-1)!} - \frac{F(x)^k}{k!}$ (all x).

(ii) If the random variable K is determined by the condition $x \geq X_1 \geq \cdots \geq X_{K-1} < X_K$, then $P(K \text{ odd}) = e^{-F(x)}$, all x.

(iii) If Y has distribution function G and is independent of the X_i's, and if K is defined by the condition $Y \geq X_1 \geq \cdots \geq X_{K-1} < X_K$, then

$$P(Y \leq x \mid K \text{ odd}) = \frac{\int_{-\infty}^{x} e^{-F(y)} \, dG(y)}{\int_{-\infty}^{+\infty} e^{-F(y)} \, dG(y)} \quad \text{(all } x\text{)}.$$

Proof of Theorem 2.1.

For fixed x,

$$P(x \geq X_1 \geq \cdots \geq X_k) = \frac{1}{k!} P(\max_{i \leq k} X_i \leq x) = \frac{F(x)^k}{k!} .$$

Thus,

$$\begin{aligned} & P(x \geq X_1 \geq \cdots \geq X_{k-1} < X_k) \\ &= P(x \geq X_1 \geq \cdots \geq X_{k-1}) - P(x \geq X_1 \geq \cdots \geq X_k) \\ &= \frac{F(x)^{k-1}}{(k-1)!} - \frac{F(x)^k}{k!} . \end{aligned}$$

Also,

$$P(K \text{ odd}) = (1 - \frac{F(x)}{1!}) + (\frac{F(x)^2}{2!} - \frac{F(x)^3}{3!}) + \cdots = e^{-F(x)} .$$

Part (iii) of the theorem finally follows from the following equalities:

$$P(Y \leq x, K \text{ odd}) = \int_{-\infty}^{x} P(K \text{ odd} \mid Y = y) \, dG(y) = \int_{-\infty}^{x} e^{-F(y)} \, dG(y) ,$$

$$P(K \text{ odd}) = \int_{-\infty}^{+\infty} e^{-F(y)} \, dG(y) . \blacksquare$$

We can now describe Forsythe's method (Forsythe, 1972) for densities f which can be written as follows:

$$f(x) = cg(x)e^{-F(x)} ,$$

where g is a density, $0 \leq F(x) \leq 1$ is some function (not necessarily a distribution function), and c is a normalization constant.

Forsythe's method

```
REPEAT
    Generate a random variate X with density g .
    W ← F(X)
    K ← 1
    Stop ← False (Stop is an auxiliary variable for getting out of the next loop.)
    REPEAT
        Generate a uniform [0,1] random variate U .
        IF U > W
            THEN Stop ← True
            ELSE W ← U ,K ← K +1
    UNTIL Stop
UNTIL K odd
RETURN X
```

We will first verify with the help of Theorem 2.1 that this algorithm is valid. First, for fixed $X=x$, we have for the first iteration of the outer loop,

$$P(K \text{ odd}) = e^{-F(x)} .$$

Thus, at the end of the first iteration,

$$P(X \leq x, K \text{ odd}) = \int_{-\infty}^{x} e^{-F(y)} g(y)\, dy .$$

Arguing as in the proof of the properties of the rejection method, we deduce that:

(i) The returned random variate X satisfies

$$P(X \leq x) = \int_{-\infty}^{x} ce^{-F(y)} g(y)\, dy .$$

Thus, it has density $ce^{-F(x)}g(x)$.

(ii) The expected number of outer loops executed before halting is $\frac{1}{p}$ where p is the probability of exit, i.e. $p = P(K \text{ odd}) = \int_{-\infty}^{+\infty} e^{-F(y)} g(y)\, dy$.

(iii) In any single iteration,

$$E(K) = \int \left\{ 1(1-\frac{F(x)}{1!})+2(\frac{F(x)}{1!}-\frac{F(x)^2}{2!})+ \cdots \right\} g(x)\, dx$$

$$= \int \left\{ 1+\frac{F(x)}{1!}+\frac{F(x)^2}{2!}+ \cdots \right\} g(x)\, dx$$

$$= \int e^{F(x)} g(x)\, dx .$$

(iv) If N is the total number of uniform [0,1] random variates required, then (by Wald's equation)

$$E(N) = \frac{1+E(K)}{p} = \frac{1+\int e^{F(x)} g(x)\, dx}{\int e^{-F(x)} g(x)\, dx} .$$

In addition to the N uniform random variates, we also need on the average $\frac{1}{p}$ random variates with density g. It should be mentioned though that g is often uniform on [0,1] so that this causes no major drawbacks. In that case, the total expected number of uniform random variates needed is at least equal to $||f||_\infty$ (this follows from Letac's lower bound). From (iv) above, we deduce that

$$2 \leq E(N) \leq \frac{1+e}{\frac{1}{e}} = e+e^2 .$$

Observe that Forsythe's method does not require any exponentiation. There are of course about $\frac{1}{p}$ evaluations of F. If we were to use the rejection method with as dominating density g, then p would be exactly the same as here. Per iteration, we would also need a g-distributed random variate, one uniform random variate, and one computation of e^{-F}. In a nutshell, we have replaced the latter evaluation by a (usually) cheaper evaluation of F and some additional uniform random variates. If exponential random variates are cheap, then we can in the rejection method replace the e^{-F} evaluation by an evaluation of F if we replace also the uniform random variate by the exponential random variate. In such situations, it seems very unlikely that Forsythe's method will be faster.

One of the disadvantages of the algorithm shown above is that F must take values in [0,1], yet many common densities such as the exponential and normal densities when put in a form useful for Forsythe's method, have unbounded F such as $F(x)=x$ or $F(x)=\frac{x^2}{2}$. To get around this, the real line must be broken up into pieces, and each piece treated separately. This will be documented further on. It should be pointed out however that the rejection method for $f = ce^{-F} g$ puts no restrictions on the size of F.

2.2. Von Neumann's exponential random variate generator.

A basic property of the exponential distribution is given in Lemma 2.1:

Lemma 2.1.

An exponential random variable E is distributed as $(Z-1)\mu+Y$ where Z,Y are independent random variables and $\mu>0$ is an arbitrary positive number: Z is geometrically distributed with

$$P(Z=i) = \int_{(i-1)\mu}^{i\mu} e^{-x}\,dx = e^{-(i-1)\mu}-e^{-i\mu} \qquad (i \geq 1),$$

and Y is a truncated exponential random variable with density

$$f(x) = \frac{e^{-x}}{1-e^{-\mu}} \qquad (0 \leq x \leq \mu).$$

Proof of Lemma 2.1.

Straightforward. ■

If we choose $\mu=1$, then Forsythe's method can be used directly for the generation of Y. Since in this case $F(x)=x$, nothing but uniform random variates are required:

von Neumann's exponential random variate generator

REPEAT

  Generate a uniform [0,1] random variate Y. Set $W \leftarrow Y$.

  $K \leftarrow 1$

  Stop $\leftarrow$ False

  REPEAT

    Generate a uniform [0,1] random variate U.

    IF $U > W$

      THEN Stop $\leftarrow$ True

      ELSE $W \leftarrow U, K \leftarrow K+1$

  UNTIL Stop

UNTIL K odd

Generate a geometric random variate Z with $P(Z=i)=(1-\frac{1}{e})(\frac{1}{e})^{i-1}$ $(i \geq 1)$.

RETURN $X \leftarrow (Z-1)+Y$

The remarkable fact is that this method requires only comparisons, uniform random variates and a counter. A quick analysis shows that $p=P(K\text{ odd})=\int_0^1 e^{-x}\,dx = 1-\frac{1}{e}$. Thus, the expected number of uniform random variates needed is

$$E(N) = \frac{1+\int_0^1 e^{x}\,dx}{\int_0^1 e^{-x}\,dx} = \frac{e^2}{e-1} .$$

This is a high bottom line. Von Neumann has noted that to generate Z, we need not carry out a new experiment. It suffices to count the number of executions of the outer loop: this is geometrically distributed with the correct parameter, and turns out to be independent of Y.

2.3. Monahan's generalization.

Monahan (1979) generalized the Forsythe-von Neumann method for generating random variates X with distribution function

$$F(x) = \frac{H(-G(x))}{H(-1)}$$

where

$$H(x) = \sum_{n=1}^{\infty} a_n x^n ,$$

$1 = a_1 \geq a_2 \geq \cdots \geq 0$ is a given sequence of constants, and G is a given distribution function.

Theorem 2.2. (Monahan, 1979)

The following algorithm generates a random variate X with distribution function F:

Monahan's algorithm

REPEAT
  Generate a random variate X with distribution function G.
  $K \leftarrow 1$
  Stop $\leftarrow$ False
  REPEAT
    Generate a random variate U with distribution function G.
    Generate a uniform [0,1] random variate V.
    IF $U \leq X$ AND $V \leq \frac{a_{K+1}}{a_K}$
      THEN $K \leftarrow K+1$
      ELSE Stop $\leftarrow$ True
  UNTIL Stop
UNTIL K odd
RETURN X

The expected number of random variates with distribution function G is

$$\frac{1+H(1)}{-H(-1)} .$$

Proof of Theorem 2.2.

We define the event A_n by $[X = \max(X, U_1, \ldots, U_n), Z_1 = \cdots = Z_n = 1]$, where the U_i's refer to the random variates U generated in the inner loop, and the Z_i's are Bernoulli random variables equal to consecutive values of $I_{[V \leq \frac{a_{i+1}}{a_i}]}$.

Thus,

$$P(X \leq x, A_n) = a_n G(x)^n ,$$

$$P(X \leq x, A_n, A_{n+1}^c) = a_n G(x)^n - a_{n+1} G(x)^{n+1} .$$

We will call the probability that X is accepted p_0. Then

$$p_0 = P(K \text{ odd}) = \sum_{n=1}^{\infty} a_n (-1)^{n+1} = H(-1) .$$

Thus, the returned X has distribution function

$$F(x) = P(X \leq x) = \frac{\sum_{n=1}^{\infty} a_n G(x)^n (-1)^{n+1}}{p_0} = \frac{H(-G(x))}{H(-1)} .$$

The expected number of G-distributed random variates needed is $E(N)$ where

$$E(N) = \frac{1}{p_0} \sum_{n=1}^{\infty} (n+1) P(A_n A_{n+1}^c)$$

$$= \sum_{n=1}^{\infty} (n+1) \frac{a_n - a_{n+1}}{p_0}$$

$$= \frac{1 + \sum_{n=1}^{\infty} a_n}{p_0}$$

$$= \frac{1+H(1)}{-H(-1)} . \blacksquare$$

Example 2.1.

Consider the distribution function

$$F(x) = 1 - \cos\left(\frac{\pi x}{2}\right) \quad (0 \leq x \leq 1) .$$

To put this in the form of Theorem 2.2, we choose another distribution function, $G(x) = x^2$ $(0 \leq x \leq 1)$, and note that

$$F(x) = \frac{H(-G(x))}{H(-1)}$$

where

$$H(x) = x + \frac{\pi^2}{48}x^2 + \frac{\pi^4}{5760}x^3 + \cdots + \frac{\pi^{2i-2}}{2^{2i-3}(2i)!}x^i + \cdots .$$

One can easily show that $p_0 = H(-1) = \frac{8}{\pi^2}$, while $E(N)$ is approximately 2.74. Also, all the conditions of Theorem 2.2 are satisfied. Random variates with this distribution function can of course be obtained by the inversion method too, as $\frac{2}{\pi}\arccos(U)$ where U is a uniform [0,1] random variate. Monahan's algorithm avoids of course any evaluation of a transcendental function. The complete algorithm can be summarized as follows, after we have noted that

$$\frac{a_{n+1}}{a_n} = \left(\frac{\pi}{2}\right)^2 \frac{1}{(2n+2)(2n+1)} :$$

REPEAT
  Generate $X \leftarrow \max(U_1, U_2)$ where U_1, U_2 are iid uniform [0,1] random variates.
  $K \leftarrow 1$
  Stop ← False
  REPEAT
    Generate U, distributed as X.
    Generate a uniform [0,1] random variate V.
    IF $U \le X$ AND $V \le \frac{(\frac{\pi}{2})^2}{4K^2+6K+2}$
      THEN $K \leftarrow K+1$
      ELSE Stop ← True
  UNTIL Stop
UNTIL K odd
RETURN X ■

2.4. An example: Vaduva's gamma generator.

We will apply the Forsythe-von Neumann method to develop a gamma generator when the parameter a is in $(0,1]$. Vaduva (1977) suggests handling the part of the gamma density on $[0,1]$ separately. This part is

$$f(x) = c(ax^{a-1})e^{-x} \qquad (0<x \leq 1),$$

where c is a normalization constant. This is in the form $cg(x)e^{-F(x)}$ for a density g and a $[0,1]$-valued function F. Random variates with density $g(x)=ax^{a-1}$ can be generated as $U^{\frac{1}{a}}$ where U is a uniform $[0,1]$ random variate. Thus, we can proceed as follows:

Vaduva's generator for the left part of the gamma density

```
REPEAT
    Generate a uniform [0,1] random variate U. Set X←U^(1/a).
    W←X
    K←1
    Stop ← False
    REPEAT
        Generate a uniform [0,1] random variate U.
        IF U>W
            THEN Stop ← True
            ELSE W←U, K←K+1
    UNTIL Stop
UNTIL K odd
RETURN X
```

Let N be the number of uniform $[0,1]$ random variates required by this method. Then, as we have seen,

$$E(N) = \frac{1+\int_0^1 ax^{a-1}e^{x}\,dx}{\int_0^1 ax^{a-1}e^{-x}\,dx}.$$

Lemma 2.2.

For Vaduva's partial gamma generator shown above, we have

$$2 \le E(N) \le (2+a(e-1))e^{\frac{a}{a+1}} \le \sqrt{e}\,(e+1)\,,$$

and

$$\lim_{a \downarrow 0} E(N) = 2\,.$$

Proof of Lemma 2.2.

First, we have

$$1 = \int_0^1 ax^{a-1}\,dx \ge \int_0^1 ax^{a-1}e^{-x}\,dx$$

$$= E(e^{-Y}) \quad \text{(where } Y \text{ is a random variable with density } ax^{a-1})$$

$$\ge e^{-E(Y)} \quad \text{(by Jensen's inequality)}$$

$$= e^{\frac{-a}{a+1}}\,.$$

Also,

$$1 \le \int_0^1 ax^{a-1}e^{x}\,dx$$

$$= 1+\frac{a}{a+1}+\frac{a}{2!(a+2)}+\cdots \quad \text{(by expansion of } e^{x}\text{)}$$

$$\le 1+a\,(1+\frac{1}{2!}+\frac{1}{3!}+\cdots)$$

$$= 1+a\,(e-1)\,.$$

Putting all of this together gives us the first inequality. Note that the supremum of the upper bound for $E(N)$ is obtained for $a=1$. Also, the limit as $a \downarrow 0$ follows from the inequality. ■

What is important here is that the expected time taken by the algorithm remains uniformly bounded in a. We have also established that the algorithm seems most efficient when a is near 0. Nevertheless, the algorithm seems less efficient than the rejection method with dominating density g developed in Example II.3.3. There the rejection constant was

$$c = \frac{1}{\int_0^1 ax^{a-1}e^{-x}\,dx}$$

which is known to lie between 1 and $e^{\frac{a}{a+1}}$. Purely on the basis of expected number of uniform random variates required, we see that the rejection method has $2 \leq E(N) \leq 2e^{\frac{a}{a+1}} \leq 2\sqrt{e}$. This is better than for Forsythe's method for all values of a. See also exercise 2.2.

2.5. Exercises.

1. Apply Monahan's theorem to the exponential distribution where $H(x)=e^x-1$, $G(x)=x$, $0<x<1$, and $F(x)=\frac{(1-e^{-x})}{1-\frac{1}{e}}$. Prove that $p_0=1-\frac{1}{e}$ and that $E(N)=\frac{e}{e-1}$ (Monahan, 1979).
2. We can use decomposition to generate gamma random variates with parameter $a \leq 1$. The restriction of the gamma density to [0,1] is dealt with in the text. For the gamma density restricted to $[1,\infty)$ rejection can be used based upon the dominating density $g(x)=e^{1-x}$ $(x \geq 1)$. Show that this leads to the following algorithm:

   REPEAT

   Generate an exponential random variate E. Set $X \leftarrow 1+E$.

   Generate a uniform [0,1] random variate U. Set $Y \leftarrow U^{-\frac{1}{1-a}}$.

   UNTIL $X \leq Y$

   RETURN X

   Show that the expected number of iterations is $\frac{1}{\int_1^\infty e^{1-x} x^{a-1}\,dx}$, and that this varies monotonically from 1 (for $a=1$) to $\frac{1}{\int_1^\infty \frac{e^{1-x}}{x}\,dx}$ (as $a \downarrow 0$).
3. Complicated densities are often cut up into pieces, and each piece is treated separately. This usually yields problems of the following type: $f(x)=ce^{-F(x)}$ $(a \leq x \leq b)$, where $0 \leq F(x) \leq F^* \leq 1$, and F^* is usually much smaller than 1. This is another way of putting that f varies very little on $[a,b]$. Show that the expected number of uniform random variates

needed in Forsythe's algorithm does not exceed $e^{F*}+e^{2F*}$. In other words, this approaches 2 very quickly as $F* \downarrow 0$.

3. ALMOST-EXACT INVERSION.

3.1. Definition.

A random variate with absolutely continuous distribution function F can be generated as $F^{-1}(U)$ where U is a uniform [0,1] random variate. Often, F^{-1} is not feasible to compute, but can be well approximated by an easy-to-compute strictly increasing absolutely continuous function ψ. Of course, $\psi(U)$ does not have the desired distribution unless $\psi=F^{-1}$. But it is true that $\psi(Y)$ has distribution function F where Y is a random variate with a nearly uniform density. The density h of Y is given by

$$h(y) = f(\psi(y))\psi'(y) ,$$

where f is the density corresponding to F. The almost-exact inversion method can be summarized as follows:

Almost-exact inversion

Generate a random variate Y with density h.
RETURN $\psi(Y)$

The point is that we gain if two conditions are satisfied: (i) ψ is easy to compute; (ii) random variates with density h are easy to generate. But because we can choose ψ from among wide classes of transformations, it should be obvious that this freedom can be exploited to make generation with density h easier. Marsaglia (1977, 1980, 1984) has made the almost-exact inversion method into an art. His contributions are best explained in a series of examples and exercises, including generators for the gamma and t distributions.

Just how one measures the goodness of a certain transformation ψ depends upon how one wants to generate Y. For example, if straightforward rejection from a uniform density is used, then the smallness of the rejection constant

$$c = \sup_y h(y)$$

would be a good measure. On the other hand, if h is treated via the mixture method and h is decomposed as

$$h(y) = pI_{[0,1]}(y) + (1-p)r(y) ,$$

then the probability p is a good measure, since the residual density r is normally difficult. A value close to 1 is highly desirable here. Note that in any case,

$$p \leq \inf_{y \in [0,1]} h(y) .$$

Thus, ψ will often be chosen so as to minimize c or to maximize p, depending upon the generator for h.

All of the above can be repeated if we take a convenient non-uniform distribution as our starting point. In particular, the normal density seems a useful choice when the target densities are the gamma or t densities. This generalization too will be discussed in this section.

3.2. Monotone densities on $[0,\infty)$.

Nonincreasing densities f on the positive real line have sometimes a shape that is similar to that of $\frac{\theta}{(1+\theta x)^2}$ where $\theta > 0$ is a parameter. Since this is the density of the distribution function $\frac{\theta x}{1+\theta x}$, we could look at transformations ψ defined by

$$\psi(y) = \frac{y}{\theta(1-y)} .$$

In this case, h becomes:

$$h(y) = f\left(\frac{y}{\theta(1-y)}\right)\frac{1}{\theta(1-y)^2} \qquad (0 \leq y \leq 1) .$$

For example, for the exponential density, we obtain

$$h(y) = e^{-\frac{y}{\theta(1-y)}} \frac{1}{\theta(1-y)^2} \qquad (0 \leq y \leq 1) .$$

Assume that we use rejection from the uniform density for generation of random variates with density h. This suggests that we should try to minimize $\sup h$. By elementary computations, one can see that h is maximal for $1-y = \frac{1}{2\theta}$, and that the maximal value is

$$4\theta e^{\frac{1}{\theta}-2} ,$$

which is minimal for $\theta=1$. The minimal value is $\frac{4}{e}$=1.4715177.... The rejection algorithm for h requires the evaluation of an exponent in every iteration, and is therefore not competitive. For this reason, the composition approach is much more likely to produce good results.

3.3. Polya's approximation for the normal distribution.

In this section, we will illustrate the composition approach. The example is due to Marsaglia (1984). For the inverse F^{-1} of the absolute normal distribution function F, Polya (1949) suggested the approximation

$$\psi(y) = \sqrt{-\theta \log(1-y^2)} \quad (0 \le y \le 1) ,$$

where he took $\theta = \frac{\pi}{2}$. Let us keep θ free for the time being. For this transformation, the density $h(y)$ of Y is

$$h(y) = \frac{1}{\sqrt{2\pi}} \frac{\theta y (1-y^2)^{\frac{\theta}{2}-1}}{\sqrt{-\theta \log(1-y^2)}} \quad (0 \le y \le 1) .$$

Let us now choose θ so that $\inf_{[0,1]} h(y)$ is maximal. This occurs for $\theta \approx 1.553$ (which is close to but not equal to Polya's constant, because our criterion for closeness is different). The corresponding value p of the infimum is about 0.985. Thus, random variates with density h can be generated as shown in the next algorithm:

Normal generator based on Polya's approximation

Generate a uniform [0,1] random variate U.

IF $U \le p$ (p is about 0.985 for the optimal choice of θ)

THEN RETURN $\psi(\frac{U}{p})$ (where $\psi(y) = \sqrt{-\theta \log(1-y^2)}$)

ELSE

Generate a random variate Y with residual density $\frac{(h(y)-p)}{(1-p)}$ $(0 \le y \le 1)$.

RETURN $\psi(Y)$

The details, such as a generator for the residual density, are delegated to exercise 3.5. It is worth pointing out however that the uniform random variate U is used in the selection of a mixture density and in the returned variate $\psi(\frac{U}{p})$. For this reason, it is "almost" true that we have one normal random variate per uniform random variate.

3.4. Approximations by simple functions of normal random variates.

In analogy with the development for the uniform distribution, we can look at other common distributions such as the normal distribution. The question now is to find an easy to compute function ψ such that $\psi(Y)$ has the desired density, where now Y is nearly normally distributed. In fact, Y should have density h given in the introduction:

$$h(y) = f(\psi(y))\psi'(y) \quad (y \in R) .$$

Usually, the purpose is to maximize p in the decomposition

$$h(y) = p\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\right) + (1-p)r(y)$$

where r is a residual density. Then, the following algorithm suggested by Marsaglia (1984) can be used:

Marsaglia's almost-exact inversion algorithm

Generate a uniform [0,1] random variate U.
IF $U \le p$
  THEN Generate a normal random variate Y.
  ELSE Generate a random variate Y with residual density r.
RETURN $\psi(Y)$

For the selection of ψ, one can either look at large classes of simple functions or scan the literature for transformations. For popular distributions, the latter route is often surprisingly efficient. Let us illustrate this for the gamma (a) density. In the table shown below, several choices for ψ are given that transform normal random variates in nearly gamma random variates (and hopefully nearly normal random variates into exact gamma random variates).

Method	$\psi(y)$	Reference
	$a+y\sqrt{a}$	Central limit theorem
Freeman-Tukey	$\frac{(y+\sqrt{4a})^2}{4}$	Freeman and Tukey (1950)
Fisher	$\frac{(y+\sqrt{4a-1})^2}{4}$	
Wilson-Hilferty	$a\left(\frac{y}{\sqrt{9a}}+1-\frac{1}{9a}\right)^3$	Wilson and Hilferty (1931)
Marsaglia	$a-\frac{1}{3}+py\sqrt{a}+\frac{y^2}{3}$, $p=1-\frac{0.16}{a}$	Marsaglia (1984)

In this table we omitted on purpose more complicated and often better approximations such as those of Cornish-Fisher, Severo-Zelen and Peizer-Pratt. For a comparative study and a bibliography of such approximations, the reader should consult Narula and Li (1977). Bolshev (1959, 1963) gives a good account of how one can obtain normalizing transformations in general. Note that our table contains only simple polynomial transformations. For example, Marsaglia's quadratic transformation is such that

$$h(y) = p\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\right) + (1-p) r(y) ,$$

where $p = 1 - \frac{0.16}{a}$. For example, when $a = 16$, we have $p = 0.99$. See exercise 3.1 for more information.

The Wilson-Hilferty transformation was first used by Greenwood (1974) and later by Marsaglia (1977). We first verify that h now is

$$h(y) = cz^{3a-1} e^{-az^3} \quad (z = \frac{y}{\sqrt{9a}} + 1 - \frac{1}{9a} \geq 0) ,$$

where c is a normalization constant. The algorithm now becomes:

Gamma generator based upon the Wilson-Hilferty approximation

Generate a random variate Y with density h.

RETURN $X \leftarrow \psi(Y) = a\left(\frac{Y}{\sqrt{9a}} + 1 - \frac{1}{9a}\right)^3$

Generation from h is done now by rejection from a normal density. The details require careful analysis, and it is worthwhile to do this once. The normal density used for the rejection differs slightly from that used by Marsaglia (1977). The story is told in terms of inequalities. We have

Lemma 3.1.

Assume that $a > \frac{1}{3}$. Define $z = \frac{y}{\sqrt{9a}} + 1 - \frac{1}{9a}$, and $z_0 = \left(\frac{3a-1}{3a}\right)^{\frac{1}{3}}$. Define the density $h(y) = cz^{3a-1} e^{-az^3}, z \geq 0$ (note: this is a density in y, not in z), where c is a normalization constant. Then, the following inequality is valid for $z \geq 0$:

$$\frac{z^{3a-1} e^{-az^3}}{z_0^{3a-1} e^{-az_0^3}} \leq e^{-\frac{(z-z_0)^2}{2\sigma^2}} ,$$

where $\sigma^2 = \frac{1}{9a\left(1-\frac{1}{3a}\right)^{\frac{1}{3}}}$.

Proof of Lemma 3.1.

The proof is based upon the Taylor series expansion. We will write $e^{g(z)}$ instead of $h(y)$ for notational convenience. Thus,

$$g(z) = -az^3+(3a-1)\log z + \log c \ .$$

This function is majorized by a quadratic polynomial in z for this will give us a normal dominating density. In such situations, it helps to expand the function about a point z_0. This point should be picked in such a way that it corresponds to the peak of g because doing so will eliminate the linear term in Taylor's series expansion. Note that

$$g'(z) = -3az^2+\frac{3a-1}{z} \ ,$$

$$g''(z) = -6az-\frac{3a-1}{z^2} \ ,$$

$$g'''(z) = -6a+\frac{6a-2}{z^3} \ .$$

We see that $g'(z)=0$ for $z=z_0$. Thus, by Taylor's series expansion,

$$g(z) = g(z_0)+\frac{1}{2}(z-z_0)^2 g''(\xi) \ ,$$

where ξ is in the interval $[z,z_0]$ (or $[z_0,z]$). We obtain our result if we can show that

$$\sup_{\xi \geq 0} g''(\xi) \leq -\frac{1}{\sigma^2} \ .$$

But when we look at g''', we notice that it is zero for $z=(\frac{3a-1}{3a})^{\frac{1}{3}}$. It is not difficult to verify that for this value, g'' attains a maximum on the positive half of the real line. Thus,

$$\sup_{\xi \geq 0} g''(\xi) \leq -9a\,(1-\frac{1}{3a})^{\frac{1}{3}} \ .$$

This concludes the proof of Lemma 3.1. ∎

The first version of the rejection algorithm is given below.

First version of the Wilson-Hilferty based gamma generator

[SET-UP]

Set $\sigma^2 \leftarrow \frac{1}{9a\,(1-\frac{1}{3a})^{\frac{1}{3}}}$, $z_0 = (\frac{3a-1}{3a})^{\frac{1}{3}}$.

[GENERATOR]

REPEAT

Generate a normal random variate N and a uniform [0,1] random variate U.

Set $Z \leftarrow z_0 + \sigma N$

UNTIL $Z \geq 0$ AND $U e^{-\frac{(Z-z_0)^2}{2\sigma^2}} \leq (\frac{Z}{z_0})^{3a-1} e^{-a(Z^3-z_0^3)}$

RETURN $X \leftarrow aZ^3$

Note that we have used here the fact that $z = \frac{y}{\sqrt{9a}} + 1 - \frac{1}{9a}$. There are two things left to the designer. First, we need to check how efficient the algorithm is. This in effect boils down to verifying what the rejection constant is. Then, we need to streamline the algorithm. This can be done in several ways. For example, the acceptance condition can be replaced by

UNTIL $Z \geq 0$ AND $-E - \frac{(Z-z_0)^2}{2\sigma^2} \leq (3a-1)\log(\frac{Z}{z_0}) - a\,(Z^3 - z_0^3)$

where E is an exponential random variate. Also, $\frac{(Z-z_0)^2}{2\sigma^2}$ is nothing but $\frac{N^2}{2}$. Additionally, we could add a squeeze step by using sharp inequalities for the logarithm. Note that $\frac{Z}{z_0} = 1 + \frac{\sigma N}{z_0}$, so that for large values of a, Z is close to z_0 which in turn is close to 1. Thus, inequalities for the logarithm should be sharp near 1. Such inequalities are given for example in the next Lemma.

Lemma 3.2.

Let $x \in [0,1)$. Then the following series expansion is valid:

$$\log(1-x) = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \cdots .$$

Thus, for $k \geq 1$,

$$-\sum_{i=1}^{k} \frac{1}{i} x^i \geq \log(1-x) \geq -\sum_{i<k} \frac{1}{i} x^i - \frac{1}{k} \frac{x^k}{1-x} .$$

Furthermore, for $x \leq 0$, and k odd,

$$-\sum_{i=1}^{k+1} \frac{1}{i} x^i \leq \log(1-x) \leq -\sum_{i=1}^{k} \frac{1}{i} x^i .$$

Proof of Lemma 3.2.

We note that in all cases,

$$-\log(1-x) = \sum_{i=1}^{k} \frac{1}{i} x^i + \frac{x^k}{k(1-\xi)^k}$$

where ξ is between 0 and x. The bounds are obtained by looking at the k-th term in the sums. Consider first $0 \leq \xi \leq x < 1$. Then, the k-th term is at least equal to $\frac{x^k}{k}$. If $x \leq \xi \leq 0$ and k is odd, then the same is true. If however k is even, then the k-th term is majorized by $\frac{x^k}{k}$.

We also note that for $0 \leq x < 1$,

$$-\log(1-x) = x + \frac{1}{2}x^2 + \cdots \leq x + \cdots + \frac{1}{k}x^k (1+x+x^2+x^3+ \cdots)$$

$$= \sum_{i=1}^{k} \frac{1}{k} \frac{x^i}{1-x} . \blacksquare$$

Let us return now to the algorithm, and use these inequalities to avoid computing the logarithm most of the time by introducing a quick acceptance step.

Second version of the Wilson-Hilferty based gamma generator

[SET-UP]

Set $\sigma^2 \leftarrow \frac{1}{9a\left(1-\frac{1}{3a}\right)^{\frac{1}{3}}}$, $z_0 = \left(\frac{3a-1}{3a}\right)^{\frac{1}{3}}$, $z_1 \leftarrow a-\frac{1}{3}$.

[GENERATOR]

REPEAT

Generate a normal random variate N and an exponential random variate E.

Set $Z \leftarrow z_0+\sigma N$ (auxiliary variate)

Set $X \leftarrow aZ^3$ (variate to be returned)

$W \leftarrow \frac{\sigma N}{Z}$ (note that $W = 1-\frac{z_0}{Z}$)

Set $S \leftarrow -E-\frac{N^2}{2}+(X-z_1)$

Accept $\leftarrow [S \leq (3a-1)(W+\frac{1}{2}W^2+\frac{1}{3}W^3)]$ AND $[Z \geq 0]$

IF NOT Accept

THEN Accept $\leftarrow [S \leq -(3a-1)\log(1-W)]$ AND $[Z \geq 0]$

UNTIL Accept

RETURN X

In this second version, we have implemented most of the suggested improvements. The algorithm is only applicable for $a > \frac{1}{3}$ and differs slightly from the algorithms proposed in Greenwood (1974) and Marsaglia (1977). Obvious things such as the observation that $(W+\frac{1}{2}W^2+\frac{1}{3}W^3)$ should be evaluated by Horner's rule, are not usually shown in our algorithms. There are two quantities that should be analyzed:

(i) The expected number of iterations before halting.

(ii) The expected number of computations of the logarithm in the acceptance step (a comparison with (i) will show us how efficient the squeeze step is).

Lemma 3.3.

The expected number of iterations of the algorithm given above (or its rejection constant) is

$$(\frac{\sqrt{a}\, a^{a-1}}{\Gamma(a)})(\frac{3a-1}{3a})^{a-\frac{1}{2}} e^{-a+\frac{1}{3}} \sqrt{2\pi} .$$

For $a \geq \frac{1}{2}$, this is less than $e^{\frac{1}{6a}}$. It tends to 1 as $a \to \infty$ and to ∞ as $a \downarrow \frac{1}{3}$.

Proof of Lemma 3.3.

The area under the dominating curve for h is

$$\int_{-\infty}^{\infty} h(z_0) e^{-\frac{(z-z_0)^2}{2\sigma^2}} dy$$

where we recall that $z=\frac{y}{\sqrt{9a}}+1-\frac{1}{9a}$, $z_0=(\frac{3a-1}{3a})^{\frac{1}{3}}$. Since $dy=\sqrt{9a}\ dz$, we see that this equals

$$h(z_0)\sqrt{2\pi}\sqrt{9a}\ \sigma$$

$$= c z_0^{3a-1} e^{-az_0^3} \sqrt{2\pi} \frac{1}{(1-\frac{1}{3a})^{\frac{1}{6}}}$$

$$= (\frac{\sqrt{a}\, a^{a-1}}{\Gamma(a)})(\frac{3a-1}{3a})^{a-\frac{1}{3}} e^{-a+\frac{1}{3}} \sqrt{2\pi} (\frac{3a}{3a-1})^{\frac{1}{6}} .$$

Here we used the fact that the normalization constant c in the definition of h is $\frac{\sqrt{a}\, a^{a-1}}{\Gamma(a)}$, which is verified by noting that

$$\int_{z \geq 0} z^{3a-1} e^{-az^3}\, dy = \frac{\Gamma(a)}{\sqrt{a}\, a^{a-1}} .$$

The remainder of the proof is based upon simple facts about the Γ function: for example, the function stays bounded away from 0 on $[0,\infty)$. Also, for $a>0$,

$$\Gamma(a) = (\frac{a}{e})^a \sqrt{\frac{2\pi}{a}} e^{\frac{\theta}{12a}} ,$$

where $0 \leq \theta \leq 1$. We will also need the elementary exponential inequalities

$$e^{-px} \geq (1-x)^p \geq e^{-\frac{px}{1-x}} \quad (p \geq 0, 0 \leq x \leq 1) .$$

Using this in our expression for the rejection constant gives an upper bound

$$\frac{\sqrt{a}\, a^{a-1} e^{a} \sqrt{2\pi a}\, e^{-a+\frac{1}{3}}}{a^{a}\sqrt{2\pi}} \left(\frac{3a-1}{3a}\right)^{a-\frac{1}{2}}$$

$$= e^{\frac{1}{3}} \left(1-\frac{1}{3a}\right)^{a-\frac{1}{2}}$$

$$\leq e^{\frac{1}{3}-(a-\frac{1}{2})(3a)^{-1}}$$

$$= e^{\frac{1}{6a}} ,$$

which is $1+\frac{1}{6a}+O(\frac{1}{a^2})$ as $a \to \infty$. ■

From Lemma 3.3, we conclude that the algorithm is not uniformly fast for $a \in (\frac{1}{3},\infty)$. On the other hand, since the rejection constant is $1+\frac{1}{6a}+O(\frac{1}{a^2})$ as $a \to \infty$, it should be very efficient for large values of a. Because of this good fit, it does not pay to introduce a quick rejection step. The quick acceptance step on the other hand is very effective, since asymptotically, the expected number of computations of a logarithm is $o(1)$ (exercise 3.1). In fact, this example is one of the most beautiful applications of the effective use of the squeeze principle.

3.5. Exercises.

1. Consider the Wilson-Hilferty based gamma generator developed in the text. Prove that the expected number of logarithm calls is $o(1)$ as $a \to \infty$.
2. For the same generator, give all the details of the proof that the expected number of iterations tends to ∞ as $a \downarrow \frac{1}{3}$.
3. For Marsaglia's quadratic gamma-normal transformation, develop the entire comparison-based algorithm. Prove the validity of his claims about the value of p as a function of a. Develop a fixed residual density generator based upon rejection for

$$r^*(x) = \sup_{a \geq a_0} r(x) .$$

Here a_0 is a real number. This helps because it avoids setting up constants each time. See Marsaglia (1984) for graphs of the residual densities r.

4. **Student's t-distribution.** Consider the t-density

$$f(x) = \frac{1}{\sqrt{\pi a}} \frac{\Gamma(\frac{a+1}{2})}{\Gamma(\frac{a}{2})} \frac{1}{(1+\frac{x^2}{a})^{\frac{a+1}{2}}} .$$

Find the best constant p if f is to be decomposed into a mixture of a normal and a residual density (p is the weight of the normal density). Repeat the same thing for $h(y)$ if we use almost-exact inversion with transformation

$$\psi(y) = y + \frac{y+y^3}{4a} .$$

Compare both values of p as a function of a. (This transformation was suggested by Marsaglia (1984).)

5. Work out all the details of the normal generator based on Polya's approximation.

6. Bolshev (1959, 1963) suggests the following transformations which are supposed to produce nearly normally distributed random variables based upon sums of iid uniform [0,1] random variates. If X_n is $\sqrt{\frac{3}{n}}\sum_{i=1}^{n} U_i$ where the U_i's are iid uniform [0,1] random variates, then

$$Y_n = X_n - \frac{1}{20n}(3X_n - X_n^{\ 3})$$

and

$$Z_n = X_n - \frac{41}{13440n^2}(X_n^{\ 5} - 10X_n^{\ 3} + 15X_n)$$

are nearly normally distributed. Use this to generate normal random variates. Take $n = 1,2,3$.

7. Show that the rejection constant of Lemma 3.3 is at most $\left(\frac{e^2}{3a-1}\right)^{\frac{1}{6}}$ when $\frac{1}{3} < a \le \frac{1}{2}$.

8. For the gamma density, the quadratic transformations lead to very simple rejection algorithms. As an example, take $s = a - \frac{1}{2}, t = \sqrt{\frac{s}{2}}$. Prove the following:

A. The density of $X = s\left(\sqrt{\frac{Z}{s}} - 1\right)$ (where Z is gamma (a) distributed) is

$$f(x) = c\left(1+\frac{x}{s}\right)^{2a-1} e^{-2x} e^{-\frac{x^2}{s}} \quad (x \ge -s)$$

where $c = 2s^{a-1}e^{-s^2}/\Gamma(a)$.

B. We have

$$f(x) \le c e^{-\frac{x^2}{s}} .$$

C. If this inequality is used to generate random variates with density f, then the rejection constant, $c\sqrt{\pi s}$, is $\sqrt{\frac{2\pi}{e}}$ at $a = 1$, and tends to

$\sqrt{2}$ as $a \uparrow \infty$. Prove also that for all values $a > \frac{1}{2}$, the rejection constant is bounded from above by $\sqrt{2} e^{\frac{1}{4a}}$.

D. The raw almost-exact inversion algorithm is:

Almost-exact inversion algorithm for gamma variates

REPEAT
  Generate a normal random variate N and an exponential random variate E.
  $X \leftarrow tN$
UNTIL $X \geq s$ AND $E - 2X + 2s \log(1+\frac{X}{s}) \geq 0$
RETURN $s(1+\frac{X}{s})^2$

E. Introduce quick acceptance and rejection steps in the algorithm that are so accurate that the expected number of evaluations of the logarithm is $o(1)$ as $a \uparrow \infty$. Prove the claim.

Remark: for a very efficient implementation based upon another quadratic transformation, see Ahrens and Dieter (1982).

4. MANY-TO-ONE TRANSFORMATIONS.

4.1. The principle.

Sometimes it is possible to exploit some distributional properties of random variables. Assume for example that $\psi(X)$ has an easy density h, where X has density f. When ψ is a one-to-one transformation, X can then be generated as $\psi^{-1}(Y)$ where Y is a random variate with the easy density h. A point in case is the inversion method of course where the easy density is the uniform density. There are important examples in which the transformation ψ is many-to-one, so that the inverse is not uniquely defined. In that case, if there are k solutions $X_1, \ldots, X_k$ of the equation $\psi(X)=Y$, it suffices to choose among the X_i' s. The probabilities however depend upon Y. The usefulness of this approach was first realized by Michael, Schucany and Haas (1976), who gave a comprehensive description and discussion of the method. They were motivated by a simple fast algorithm for the inverse gaussian family based upon this approach.

By far the most important case is $k=2$, which is the one that we shall deal with here. Several important examples are developed in subsections.

Assume that there exists a point t such that ψ' is of one sign on $(-\infty,t)$ and on (t,∞). For example, if $\psi(x)=x^2$, then $\psi'(x)=2x$ is nonpositive on $(-\infty,0)$ and nonnegative on $(0,\infty)$, so that we can take $t=0$. We will use the notation

$$x = l(y), \; x = r(y)$$

for the two solutions of $y=\psi(x)$: here, l is the solution in $(-\infty,t)$, and r is the solution in (t,∞). If ψ satisfies the conditions of Theorem I.4.1 on each interval, and X has density f, then $\psi(X)$ has density

$$h(y) = |l'(y)| f(l(y)) + |r'(y)| f(r(y)) .$$

This is quickly verified by computing the distribution function of $\psi(X)$ and then taking the derivative. Vice versa, given a random variate Y with density h, we can obtain a random variate X with density f by choosing $X=l(Y)$ with probability

$$\frac{|l'(Y)| f(l(Y))}{h(Y)} ,$$

and choosing $X=r(Y)$ otherwise. Note that $|l'(y)| = 1/|\psi'(l(y))|$. This, the method of Michael, Schucany and Haas (1976), can be summarized as follows:

Inversion of a many-to-one transformation

Generate a random variate Y with density h.

Generate a uniform $[0,1]$ random variate U.

Set $X_1 \leftarrow l(Y)$, $X_2 \leftarrow r(Y)$

IF $U \leq \dfrac{1}{1+\dfrac{f(X_2)}{f(X_1)}\left|\dfrac{\psi'(X_1)}{\psi'(X_2)}\right|}$

THEN RETURN $X \leftarrow X_1$

ELSE RETURN $X \leftarrow X_2$

It will be clear from the examples that in many cases the expression in the selection step takes a simple form.

4.2. The absolute value transformation.

The transformation $y = | x-t |$ for fixed t satisfies the conditions of the previous section. Here we have $l(y)=t-y$, $r(y)=t+y$. Since $| \psi' |$ remains constant, the decision is extremely simple. Thus, we have

Generate a random variate Y with density $h(y)=f(t-y)+f(t+y)$.
Generate a uniform [0,1] random variate U.
IF $U \le \frac{f(t-Y)}{f(t-Y)+f(t+Y)}$
  THEN RETURN $X \leftarrow t-Y$
  ELSE RETURN $X \leftarrow t+Y$

If f is symmetric about t, then the decisions $t-Y$ and $t+Y$ are equally likely. Another interesting case occurs when h is the uniform density. For example, consider the density

$$f(x) = \frac{1+\cos x}{\pi} \quad (0 \le x \le \pi) .$$

Then, taking $t = \frac{\pi}{2}$, we see that

$$h(y) = f(t-y)+f(t+y) = \frac{2}{\pi} \quad (0 \le y \le \frac{\pi}{2}) .$$

Thus, we can generate random variates with this density as follows:

Generate two iid uniform [0,1] random variates U,V.
Set $Y \leftarrow \frac{\pi V}{2}$.
IF $U \le \frac{1+\cos Y}{2}$
  THEN RETURN $X \leftarrow Y$
  ELSE RETURN $X \leftarrow \pi - Y$

Here we have made use of additional symmetry in the problem. It should be noted that the evaluation of the cos can be avoided altogether by application of the series method (see section 5.4).

4.3. The inverse gaussian distribution.

Michael, Schucany and Haas (1976) have successfully applied the many-to-one transformation method to the **inverse gaussian distribution**. Before we proceed with the details of their algorithm, it is necessary to give a short introductory tour of the distribution (see Folks and Chhikara (1978) for a survey).

A random variable $X \geq 0$ with density

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} e^{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}} \qquad (x \geq 0)$$

is said to have the inverse gaussian distribution with parameters $\mu>0$ and $\lambda>0$. We will say that a random variate X is $I(\mu,\lambda)$. Sometimes, the distribution is also called Wald's distribution, or the first passage time distribution of Brownian motion with positive drift.

The densities are unimodal and have the appearance of gamma densities. The mode is at

$$\mu\left(\sqrt{1+\frac{9\mu^2}{4\lambda^2}}-\frac{3\mu}{2\lambda}\right) .$$

The densities are very flat near the origin and have exponential tails. For this reason, all positive and negative moments exist. For example, $E(X^{-a})=E(X^{a+1})/\mu^{2a+1}$, all $a \in R$. The mean is μ and the variance is $\frac{\mu^3}{\lambda}$. The main distributional property is captured in the following Lemma:

Lemma 4.1. (Shuster, 1968)

When X is $I(\mu,\lambda)$, then

$$\frac{\lambda(X-\mu)^2}{\mu^2 X}$$

is distributed as the square of a normal random variable , i.e. it is chi-square with one degree of freedom.

Proof of Lemma 4.1.

Straightforward. ∎

Based upon Lemma 4.1, we can apply a many-to-one transformation

$$\psi(x) = \frac{\lambda(x-\mu)^2}{\mu^2 x} .$$

Here, the inverse has two solutions, one on each side of μ. The solutions of $\psi(X)=Y$ are

$$X_1 = \mu+\frac{\mu^2 Y}{2\lambda}-\frac{\mu}{2\lambda}\sqrt{4\mu\lambda Y+\mu^2 Y^2}$$

$$X_2 = \frac{\mu^2}{X_1}$$

One can verify that

$$\frac{f(X_2)}{f(X_1)} = \left(\frac{X_1}{\mu}\right)^3 ,$$

$$\frac{\psi'(X_1)}{\psi'(X_2)} = -\left(\frac{\mu}{X_1}\right)^2 .$$

Thus, X_1 should be selected with probability $\frac{\mu}{\mu+X_1}$. This leads to the following algorithm:

Inverse gaussian distribution generator of Michael, Schucany and Haas

Generate a normal random variate N.

Set $Y \leftarrow N^2$

Set $X_1 \leftarrow \mu+\frac{\mu^2 Y}{2\lambda}-\frac{\mu}{2\lambda}\sqrt{4\mu\lambda Y+\mu^2 Y^2}$

Generate a uniform [0,1] random variate U.

IF $U \le \frac{\mu}{\mu+X_1}$

THEN RETURN $X \leftarrow X_1$

ELSE RETURN $X \leftarrow \frac{\mu^2}{X_1}$

This algorithm was later rediscovered by Padgett (1978). The time-consuming components of the algorithm are the square root and the normal random variate generation. There are a few shortcuts: a few multiplications can be saved if we replace Y by μY at the outset, for example. There are several exercises about the inverse gaussian distribution following this sub-section.

4.4. Exercises.

1. **First passage time distribution of drift-free Brownian motion.** Show that as $\mu \to \infty$ while λ remains fixed, the $I(\mu,\lambda)$ density tends to the density

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}}\, e^{-\frac{\lambda}{2x}} \quad (x \geq 0),$$

which is the one-sided stable density with exponent $\frac{1}{2}$, or the density for the first passage time of drift-free Brownian motion. Show that this is the density of the inverse of a gamma $(\frac{1}{2}, \frac{2}{\lambda})$ random variable (Wasan and Roy, 1967). This is equivalent to showing that it is the density of $\frac{\lambda}{N^2}$ where N is a normal random variable.

2. This is a further exercise about the properties of the inverse gaussian distribution. Show the following:
   (i) If X is $I(\mu,\lambda)$, then cX is $I(c\,\mu, c\,\lambda)$.
   (ii) The characteristic function of X is $e^{\frac{\lambda}{\mu}\left(1-\sqrt{1-\frac{2i\mu^2 t}{\lambda}}\right)}$.
   (iii) If X_i, $1 \leq i \leq n$, are independent $I(\mu_i, c\,\mu_i^2)$ random variables, then $\sum_{i=1}^{n} X_i$ is $I(\sum \mu_i, c\,(\sum \mu_i)^2)$. Thus, if the X_i's are iid $I(\mu,\lambda)$, then $\sum X_i$ is $I(n\,\mu, n^2 \lambda)$.
   (iv) Show that when N_1, N_2 are independent normal random variables with variances σ_1^2 and σ_2^2, then $\frac{N_1 N_2}{\sqrt{N_1^2+N_2^2}}$ is normal with variance σ_3^2 determined by the relation $\frac{1}{\sigma_3} = \frac{1}{\sigma_1} + \frac{1}{\sigma_2}$.
   (v) The distribution function of X is

$$F(x) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu}-1\right)\right) + e^{\frac{2\lambda}{\mu}}\, \Phi\left(-\sqrt{\frac{\lambda}{x}}\left(1+\frac{x}{\mu}\right)\right),$$

   where Φ is the standard normal distribution function (Ziganglrov, 1962).

5. THE SERIES METHOD.

5.1. Description.

In this section, we consider the problem of the computer generation of a random variable X with density f where f can be approximated from above and below by sequences of functions f_n and g_n. In particular, we assume that:

(i) $\lim_{n \to \infty} f_n = f$;
$\lim_{n \to \infty} g_n = f$.
(ii) $f_n \leq f \leq g_n$.
(iii) $f \leq ch$ for some constant $c \geq 1$ and some easy density h .

The sequences f_n and g_n should be easy to evaluate, while the dominating density h should be easy to sample from. Note that f_n need not be positive, and that g_n need not be integrable. This setting is common: often f is only known as a series, as in the case of the Kolmogorov-Smirnov distribution or the stable distributions, so that random variate generation has to be based upon this series. But even if f is explicitly known, it can often be expanded in a fast converging series such as in the case of a normal or exponential density. The series method described below actually avoids the exact evaluation of f all the time. It can be thought of as a rejection method with an infinite number of acceptance and rejection conditions for squeezing. Nearly everything in this section was first developed in Devroye (1980).

The series method

REPEAT
  Generate a random variate X with density h .
  Generate a uniform [0,1] random variate U.
  $W \leftarrow Uch(X)$
  $n \leftarrow 0$
  REPEAT
    $n \leftarrow n+1$
    IF $W \leq f_n(X)$ THEN RETURN X
  UNTIL $W > g_n(X)$
UNTIL False

The fact that the outer loop in this algorithm is an infinite loop does not matter, because with probability one we will exit in the inner loop (in view of $f_n \to f$, $g_n \to f$). We have here a true rejection algorithm because we exit when $W \leq Uch(X)$. Thus, the expected number of outer loops is c , and the choice of the dominating density h is important. Notice however that the time should be measured in terms of the number of f_n and g_n evaluations. Such analysis will be given further on. While in many cases, the convergence to f is so fast that the expected number of f_n evaluations is barely larger than c, it is true that there are examples in which this expected number is ∞. It is also worth observing that the squeeze steps are essential here for the correctness of the algorithm. They actually form the algorithm.

In the remainder of this section, we will give three important special cases of approximating series. The series method and its variants will be illustrated with the aid of the exponential, Raab-Green and Kolmogorov-Smirnov distributions further on.

Assume first that f can be written as a convergent series

$$f(x) = \sum_{n=1}^{\infty} s_n(x) \leq ch(x)$$

where

$$\left| \sum_{i=n+1}^{\infty} s_i(x) \right| \leq R_{n+1}(x)$$

is a known estimate of the remainder, and h is a given density. In this special instance, we can rewrite the series method in the following form:

The convergent series method

REPEAT
  Generate a random variate X with density h.
  Generate a uniform [0,1] random variate U.
  $W \leftarrow Uch(X)$
  $S \leftarrow 0$
  $n \leftarrow 0$
  REPEAT
    $n \leftarrow n+1$
    $S \leftarrow S+s_n(X)$
  UNTIL $|S-W| > R_{n+1}(X)$
UNTIL $S \leq W$
RETURN X

Assume next that f can be written as an alternating series

$$f(x) = ch(x)(1-a_1(x)+a_2(x)-a_3(x)+ \cdots)$$

where a_n is a sequence of functions satisfying the condition that $a_n(x) \downarrow 0$ as $n \rightarrow \infty$, for all x, c is a constant, and h is an easy density. Then, the series method can be written as follows:

The alternating series method

REPEAT
  Generate a random variate X with density h.
  Generate a uniform $[0,c]$ random variate U.
  $n \leftarrow 0, W \leftarrow 0$
  REPEAT
    $n \leftarrow n+1$
    $W \leftarrow W + a_n(X)$
    IF $U \geq W$ THEN RETURN X
    $n \leftarrow n+1$
    $W \leftarrow W - a_n(X)$
  UNTIL $U < W$
UNTIL False

This algorithm is valid because f is bounded from above and below by two converging sequences:

$$1+\sum_{j=1}^{k}(-1)^j a_j(x) \leq \frac{f(x)}{ch(x)} \leq 1+\sum_{j=1}^{k+1}(-1)^j a_j(x) , \ k \text{ odd} .$$

That this is indeed a valid inequality follows from the monotonicity of the terms (consider the terms pairwise). As in the ordinary series method, f is never fully computed. In addition, h is never evaluated either.

A second important special case occurs when

$$f(x) = ch(x)\, e^{-a_1(x)+a_2(x)-\cdots}$$

where c, h, a_n are as for the alternating series method. Then, the alternating series method is equivalent to:

The alternating series method; exponential version

```
REPEAT
    Generate a random variate X with density h .
    Generate an exponential random variate E .
    n ←0, W ←0
    REPEAT
        n ←n +1
        W ←W + a_n (X)
        IF E ≥ W THEN RETURN X
        n ←n +1
        W ←W - a_n (X)
    UNTIL E < W
UNTIL False
```

5.2. Analysis of the alternating series algorithm.

For the four versions of the series method defined above, we know that the expected number of iterations is equal to the rejection constant, c . In addition, there is a hidden contribution to the time complexity due to the fact that the inner loop, needed to decide whether $Uch(X) \leq f(X)$, requires a random number of computations of a_n . The computations of a_n are assumed to take a constant time independent of n - if they do not, just modify the analysis given in this section slightly. In all the examples that will follow, the a_n computations take a constant time.

In Theorem 5.1, we will give a precise answer for the alternating series method.

Theorem 5.1.

Consider the alternating series method for a density f decomposed as follows:

$$f(x) = ch(x)(1-a_1(x)+a_2(x)-\cdots),$$

where $c \geq 1$ is a normalization constant, h is a density, and $a_0 \equiv 1 \geq a_1 \geq a_2 \geq \cdots \geq 0$. Let N be the total number of computations of a factor a_n before the algorithm halts. Then,

$$E(N) = c\int_0^\infty [\sum_{i=0}^\infty a_i(x)]\, h(x)\, dx .$$

Proof of Theorem 5.1.

By Wald's equation, $E(N)$ is equal to c times the expected number of a_n computations in the first iteration. In the first iteration, we fix $X=x$ with density h. Then, dropping the dependence on x, we see that for the odd terms a_n, we require

1 with probability $1-a_1$
2 with probability a_1-a_2
3 with probability a_2-a_3
4 with probability a_3-a_4
...

computations of a_n. The expected value of this is

$$\sum_{i=1}^\infty i(a_{i-1}-a_i) = \sum_{i=0}^\infty a_i .$$

Collecting these results gives us Theorem 5.1. ■

Theorem 5.1 shows that the expected time complexity is equal to the oscillation in the series. Fast converging series lead to fast algorithms.

5.3. Analysis of the convergent series algorithm.

As in the previous section, we will let N be the number of computations of terms s_n before the algorithm halts. We have:

Theorem 5.2.

For the convergent series algorithm of section 5.1,

$$E(N) \leq 2\int(\sum_{n=1}^{\infty} R_n(x))\, dx .$$

Proof of Theorem 5.2.

By Wald's equation, $E(N)$ is equal to c times the expected number of s_n computations in the first global iteration. If we fix X with density h, then if N is the number of s_n computations in the first iteration alone,

$$P(N>n \mid X) \leq \frac{2R_{n+1}(X)}{ch(X)} .$$

Thus,

$$E(N \mid X) = \sum_{n=0}^{\infty} P(N>n \mid X)$$

$$\leq \sum_{n=0}^{\infty} \frac{2R_{n+1}(X)}{ch(X)} .$$

Hence, turning to the overall number of s_n computations,

$$E(N) \leq c \sum_{n=1}^{\infty} \int h(x) \frac{2R_n(x)}{ch(x)}\, dx$$

$$= 2\int(\sum_{n=1}^{\infty} R_n(x))\, dx . \blacksquare$$

It is important to note that a series converging at the rate $\frac{1}{n}$ or slower cannot yield finite expected time. Luckily, many important series, such as those of all the remaining subsections on the series method converge at an exponential rather than a polynomial rate. In view of Theorem 5.2, this virtually insures the finiteness of their expected time. It is still necessary however to verify whether the expected time statements are not upset in an indirect way through the dependence of $R_n(x)$ upon x: for example, the bound of Theorem 5.2 is infinite when $\int R_n(x)\, dx = \infty$ for some n.

5.4. The exponential distribution.

It is known that for all odd k and all $x>0$,

$$\sum_{j=0}^{k-1}(-1)^j\frac{x^j}{j!} \geq e^{-x} \geq \sum_{j=0}^{k}(-1)^j\frac{x^j}{j!} .$$

We will apply the alternating series method to the truncated exponential density

$$f(x)=\frac{e^{-x}}{1-e^{-\mu}} \quad (0\leq x\leq\mu) ,$$

where $1\geq\mu>0$ is the truncation point. As dominating curve, we can use the uniform density (called h) on $[0,\mu]$. Thus, in the decomposition needed for the alternating series method, we use

$$c=\frac{\mu}{1-e^{-\mu}} ,$$

$$h(x)=\frac{1}{\mu}I_{[0,\mu]}(x) ,$$

$$a_n(x)=\frac{x^n}{n!} .$$

The monotonicity of the a_n's is insured when $|x|\leq 1$. This forces us to choose $\mu\leq 1$. The expected number of a_n computations is

$$E(N)=c\int_0^\mu\sum_{j=0}^{\infty}\frac{x^j}{j!}\frac{1}{\mu}\,dx$$

$$=c\,\frac{e^\mu-1}{\mu}$$

$$=\frac{e^\mu-1}{1-e^{-\mu}} .$$

For example, for $\mu=1$, the value e is obtained. But interestingly, $E(N)\downarrow 1$ as $\mu\downarrow 0$. The truncated exponential density is important, because standard exponential random variates can be obtained by adding an independent properly scaled geometric random variate (see for example section IV.2.2 on the Forsythe-von Neumann method or section IX.2 about exponential random variates). The algorithm for the truncated exponential density is given below:

A truncated exponential generator via the alternating series method

REPEAT

  Generate a uniform $[0,\mu]$ random variate X.

  Generate a uniform $[0,1]$ random variate U.

  $n \leftarrow 0, W \leftarrow 0, V \leftarrow 1$ (V is used to facilitate evaluation of consecutive terms in the alternating series.)

  REPEAT

    $n \leftarrow n+1$

    $V \leftarrow \frac{VX}{n}$

    $W \leftarrow W+V$

    IF $U \geq W$ THEN RETURN X

    $n \leftarrow n+1$

    $V \leftarrow \frac{VX}{n}$

    $W \leftarrow W-V$

  UNTIL $U < W$

UNTIL False

The alternating series method based upon Taylor's series is not applicable to the exponential distribution on $[0,\infty)$ because of the impossibility of finding a dominating density h based upon this series. In the exercise section, the ordinary series method is applied with a family of dominating densities, but the squeezing is still based upon the Taylor series for the exponential density.

5.5. The Raab-Green distribution.

The density

$$f(x) = \frac{1+\cos(x)}{2\pi} \qquad (|x| \leq \pi)$$

$$= \frac{1}{\pi}(1-\frac{1}{2}\frac{x^2}{2!}+\frac{1}{2}\frac{x^4}{4!}- \cdots)$$

was suggested by Raab and Green (1961) as an approximation for the normal density. The series expansion is very similar to that of the exponential function. Again, we are in a position to apply the alternating series method, but now with $h(x)=\frac{1}{2\pi}$ $(|x| \leq \pi)$, $c=2$ and $a_n(x)=\frac{1}{2}\frac{x^{2n}}{2n!}$. It is easy to verify that $a_n \downarrow 0$ as $n \rightarrow \infty$ for all x in the range:

$$\frac{a_{n+1}(x)}{a_n(x)} = \frac{x^2}{(2n+2)(2n+1)} \leq \frac{\pi^2}{12} \qquad (n \geq 1) .$$

Note however that a_1 is not smaller than 1, which was a condition necessary for the application of Theorem 5.1. Nevertheless, the alternating series method remains formally valid, and we have:

A Raab-Green density generator via the alternating series method

REPEAT

Generate a uniform $[-\pi,\pi]$ random variate X.

Generate a uniform $[0,1]$ random variate U.

$n \leftarrow 0, W \leftarrow 0, V \leftarrow 1$ (V is used to facilitate evaluation of consecutive terms in the alternating series.)

REPEAT

$n \leftarrow n+1$

$V \leftarrow \frac{VX^2}{(2n)(2n-1)}$

$W \leftarrow W+V$

IF $U \geq W$ THEN RETURN X

$n \leftarrow n+1$

$V \leftarrow \frac{VX^2}{(2n)(2n-1)}$

$W \leftarrow W-V$

UNTIL $U < W$

UNTIL False

The drawback with this algorithm is that c, the rejection constant, is 2. But this can be avoided by the use of a many-to-one transformation described in section IV.4. The principle is this: if (X,U) is uniformly distributed in $[-\frac{\pi}{2},\frac{\pi}{2}]\times[0,2]$, then we can exit with X when $U \leq 1+\cos(X)$ and with $\pi\,\text{sign}X-X$ otherwise, thereby avoiding rejections altogether. With this improvement, we obtain:

An improved Raab-Green density generator based on the alternating series method

Generate a uniform $[-\frac{\pi}{2},\frac{\pi}{2}]$ random variate X.

Generate a uniform [0,1] random variate U.

$n \leftarrow 0, W \leftarrow 0, V \leftarrow 1$ (V is used to facilitate evaluation of consecutive terms in the alternating series.)

REPEAT

 $n \leftarrow n+1$

 $V \leftarrow \frac{VX^2}{(2n)(2n-1)}$

 $W \leftarrow W+V$

 IF $U \geq W$ THEN RETURN X

 $n \leftarrow n+1$

 $V \leftarrow \frac{VX^2}{(2n)(2n-1)}$

 $W \leftarrow W-V$

 IF $U \leq W$ THEN RETURN $\pi\,\mathrm{sign}X-X$

UNTIL False

This algorithm improves over the algorithm of section IV.4 for the same distribution in which the cos was evaluated once per random variate. We won't give a detailed time analysis here. It is perhaps worth noting that the probability that the UNTIL step is reached, i.e. the probability that one iteration is completed, is about 2.54%. This can be seen as follows: if $N*$ is the number of completed iterations, then

$$P(N*>i) = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \frac{1}{2}\frac{x^{4i}}{(4i)!}\,dx = \frac{1}{\pi}\frac{(\frac{\pi}{2})^{4i+1}}{(4i+1)!} .$$

and thus

$$E(N*) = \sum_{i=0}^{\infty} \frac{1}{\pi}\frac{(\frac{\pi}{2})^{4i+1}}{(4i+1)!} .$$

In particular, $P(N*>1)=\frac{\pi^4}{3840}\approx 0.0254$. Also, $E(N*)$ is about equal to $1+2P(N*>1)\approx 1.0254$ because $P(N*>2)$ is extremely small.

5.6. The Kolmogorov-Smirnov distribution.

The **Kolmogorov-Smirnov distribution function**

$$F(x) = \sum_{n=-\infty}^{\infty} (-1)^n e^{-2n^2x^2} \quad (x \geq 0)$$

appears as the limit distribution of the Kolmogorov-Smirnov test statistic (Kolmogorov (1933); Smirnov (1939); Feller (1948)). No simple procedure for inverting F is known, hence the inversion method is likely to be slow. Also, both the distribution function and the corresponding density are only known as infinite series. Thus, exact evaluation of these functions is not possible in finite time. Yet, by using the series method, we can generate random variates with this distribution extremely efficiently. This illustrates once more that generating random variates is simpler than computing a distribution function.

First, it is necessary to obtain convenient series expansions for the density. Taking the derivative of F, we obtain the density

$$f(x) = 8 \sum_{n=1}^{\infty} (-1)^{n+1} n^2 x e^{-2n^2x^2} \quad (x \geq 0) ,$$

which is in the format of the alternating series method if we take

$$ch(x) = 8xe^{-2x^2} ,$$

$$a_n(x) = (n+1)^2 e^{-2x^2((n+1)^2-1)} \quad (n \geq 0) .$$

There is another series for F and f which can be obtained from the first series by the theory of theta functions (see e.g. Whittaker and Watson , 1927):

$$F(x) = \frac{\sqrt{2\pi}}{x} \sum_{n=1}^{\infty} e^{-\frac{(2n-1)^2\pi^2}{8x^2}} \quad (x > 0) ;$$

$$f(x) = \frac{\sqrt{2\pi}}{x} \sum_{n=1}^{\infty} [\frac{(2n-1)^2\pi^2}{4x^3} - \frac{1}{x}] e^{-\frac{(2n-1)^2\pi^2}{8x^2}} \quad (x > 0) .$$

Again, we have the format needed for the alternating series method, but now with

$$ch(x) = \frac{\sqrt{2\pi}\pi^2}{4x^4} e^{-\frac{\pi^2}{8x^2}} \quad (x > 0) ,$$

$$a_n(x) = \begin{cases} \frac{4x^2}{\pi^2} e^{-\frac{(n^2-1)\pi^2}{8x^2}} & (n \text{ odd} , x > 0) \\ (n+1)^2 e^{-\frac{((n+1)^2-1)\pi^2}{8x^2}} & (n \text{ even} , x > 0) \end{cases} .$$

We will refer to this series expansion as the second series expansion. In order for the alternating series method to be applicable, we must verify that the a_n's satisfy the monotonicity condition. This is done in Lemma 5.1:

Lemma 5.1.

The terms a_n in the first series expansion are monotone $\downarrow$ for $x > \sqrt{\frac{1}{3}}$. For the second series expansion, they are monotone $\downarrow$ when $x < \frac{\pi}{2}$.

Proof of Lemma 5.1.

In the first series expansion, we have

$$\log\left(\frac{a_{n-1}(x)}{a_n(x)}\right) = -2\log\left(1+\frac{1}{n}\right)+2(2n+1)x^2$$

$$\geq -\frac{2}{n}+2(2n+1)x^2 \geq -2+6x^2 > 0 .$$

For the second series expansion, when n is even,

$$\frac{a_n(x)}{a_{n+1}(x)} = \frac{(n+1)^2\pi^2}{4x^2} \geq \frac{\pi^2}{4x^2} > 1 .$$

Also,

$$\log\left(\frac{a_{n-1}(x)}{a_n(x)}\right) = -\log\left(\frac{(n+1)^2\pi^2}{4x^2}\right)+\frac{n\pi^2}{2x^2} = y-2\log(n+1)-\log\left(\frac{y}{2}\right)$$

where $y=\frac{\pi^2}{2x^2}$. The last expression is increasing in y for $y \geq 2$ and all $n \geq 2$. Thus, it is not smaller than $2n-2\log(n+1) \geq 0$. ■

We now give the algorithm of Devroye (1980). It uses the mixture method because one series by itself does not yield easily identifiable upper and lower bounds for f on the entire real line. We are fortunate that the monotonicity conditions are satisfied on $(\sqrt{\frac{1}{3}},\infty)$ and on $(0,\frac{\pi}{2})$ for the two series respectively. Had these intervals been disjoint, then we would have been forced to look for yet another approximation. We define the breakpoint for the mixture method by $t \in (\sqrt{\frac{1}{3}},\frac{\pi}{2})$. The value 0.75 is suggested. Define also $p=F(t)$.

Generate a uniform [0,1] random variate U.
IF $U<p$
THEN RETURN a random variate X with density $\frac{f}{p}, 0<x<t$.
ELSE RETURN a random variate X with density $\frac{f}{1-p}, t<x$.

For generation in the two intervals, the two series expansions are used. Another constant needed in the algorithm is $t'=\frac{\pi^2}{8t^2}$. We have:

Generator for the leftmost interval

```
REPEAT
    REPEAT
        Generate two iid exponential random variates, E_0, E_1.
        E_0 ← E_0 / (1 - 1/(2t'))
        E_1 ← 2E_1
        G ← t' + E_0
        Accept ← [(E_0)^2 ≤ t' E_1 (G + t')]
        IF NOT Accept
            THEN Accept ← [G/t' - 1 - log(G/t') ≤ E_1]
    UNTIL Accept
    X ← π / √(8G)
    W ← 0
    Z ← 1/(2G)
    P ← e^(-G)
    n ← 1
    Q ← 1
    Generate a uniform [0,1] random variate U.
    REPEAT
        W ← W + ZQ
        IF U ≥ W THEN RETURN X
        n ← n + 2
        Q ≤ P^(n^2 - 1)
        W ← W - n^2 Q
    UNTIL U < W
UNTIL False
```

Generator for the rightmost interval

REPEAT

  Generate an exponential random variate E.

  Generate a uniform [0,1] random variate U.

  $X \leftarrow \sqrt{t^2+\frac{E}{2}}$

  $W \leftarrow 0$

  $n \leftarrow 1$

  $Z \leftarrow e^{-2X^2}$

  REPEAT

    $n \leftarrow n+1$

    $W \leftarrow W+n^2 Z^{n^2-1}$

    IF $U \geq W$ THEN RETURN X

    $n \leftarrow n+1$

    $W \leftarrow W-n^2 Z^{n^2-1}$

  UNTIL $U \leq W$

UNTIL False

The algorithms are both straightforward applications of the alternating series method, but perhaps a few words of explanation are in order regarding the algorithms used for the dominating densities. This is done in two lemmas.

Lemma 5.2.

The random variable $\sqrt{t^2+\frac{E}{2}}$ (where E is an exponential random variable and $t>0$) has density

$$cxe^{-2x^2} \quad (x \geq t),$$

where $c>0$ is a normalization constant.

Proof of Lemma 5.2.

Verify that the distribution function of the random variable is $1-e^{-2(x^2-t^2)}$ $(x \geq t)$. Taking the derivative of this distribution function yields the desired result. ∎

Lemma 5.3.

If G is a random variable with truncated gamma $(\frac{3}{2})$ density $c\sqrt{y}\,e^{-y}$ $(y \geq t' = \frac{\pi^2}{8t^2})$, then $\frac{\pi}{\sqrt{8G}}$ has density

$$\frac{c}{x^4} e^{-\frac{\pi^2}{8x^2}} \quad (0 < x \leq t) ,$$

where the c 's stand for (possibly different) normalization constants, and $t > 0$ is a constant. A truncated gamma $(\frac{3}{2})$ random variate can be generated by the algorithm:

Truncated gamma generator

REPEAT

Generate two iid exponential random variates, E_0, E_1.

$E_0 \leftarrow \frac{E_0}{1-\frac{1}{2t'}}$

$E_1 \leftarrow 2E_1$

$G \leftarrow t' + E_0$

Accept $\leftarrow [(E_0)^2 \leq t' E_1 (G + t')]$

IF NOT Accept

THEN Accept $\leftarrow [\frac{G}{t'} - 1 - \log(\frac{G}{t'}) \leq E_1]$

UNTIL Accept

RETURN G

Proof of Lemma 5.3.

The Jacobian of the transformation $y = \frac{\pi^2}{8x^2}$ is $\frac{4\pi}{(8y)^{\frac{3}{2}}}$. This gives the distributional result without further work if we argue backwards. The validity of the rejection algorithm with squeezing requires a little work. First, we start from the inequality

$$y \cdot \leq e^{\frac{y}{t'}} \frac{t'}{e} \quad (y \geq t') ,$$

which can be obtained by maximizing $ye^{\frac{-y}{t'}}$ in the said interval. Thus,

$$\sqrt{y}\,e^{-y} \le \sqrt{\frac{t'}{e}}\,e^{-(1-\frac{1}{2t'})y} \qquad (y \ge t') .$$

The upper bound is proportional to the density of $t'+\frac{E}{1-\frac{1}{2t'}}$ where E is an exponential random variate. This random variate is called G in the algorithm. Thus, if U is a uniform random variate, we can proceed by generating couples G,U until

$$e^{\frac{G}{2t'}}\sqrt{\frac{t'}{e}}\,U \le \sqrt{G} .$$

This condition is equivalent to

$$\frac{G}{t'}-1-\log(\frac{G}{t'}) \le 2E_1$$

where E_1 is another exponential random variable. A squeeze step can be added by noting that $\log(1+u) \ge \frac{2u}{2+u}$ $(u \ge 0)$ (exercise 5.1). ■

All the previous algorithms can now be collected into one long (but fast) algorithm. For generalities on good generators for the tail of the gamma density, we refer to the section on gamma variate generation. In the implementation of Devroye (1980), two further squeeze steps were added. For the rightmost interval, we can return X when $U \ge 4e^{-6t^2}$ (which is a constant). For the leftmost interval, the same can be done when $U \ge \frac{4t^2}{\pi^2}$. For $t=0.75$, we have $p \approx 0.373$, and the quick acceptance probabilities are respectively ≈ 0.86 and ≈ 0.77 for the latter squeeze steps.

Related distributions.

The **empirical distribution function** $F_n(x)$ for a sample $X_1, \ldots, X_n$ of iid random variables is defined by

$$F_n(x) = \sum_{i=1}^{n} \frac{1}{n} I_{[X_i \le x]}$$

where I is the indicator function. If X_i has distribution function $F(x)$, then the following goodness-of-fit statistics have been proposed by various authors:

(1) The asymmetrical Kolmogorov-Smirnov statistics $K_n^{+}=\sqrt{n}\,\sup(F_n-F)$, $K_n^{-}=\sqrt{n}\,\sup(F-F_n)$.

(ii) The Kolmogorov-Smirnov statistic $K_n = \max(K_n^+, K_n^-)$.

(iii) Kuiper's statistic $V_n = K_n^+ + K_n^-$.

(iv) von Mises' statistic $W_n^2 = n \int (F_n - F)^2 dF$.

(v) Watson's statistic $U_n = n \int (F_n - F - (\int (F_n - F) dF))^2 dF$.

(vi) The Anderson-Darling statistic $A_n^2 = n \int \frac{(F_n - F)^2}{F(1-F)} dF$.

For surveys of the properties and applications of these and other statistics, see Darling (1955), Barton and Mallows (1965), and Sahler (1968). The limit random variables (as $n \to \infty$) are denoted with the subscripts ∞. The limit distributions have characteristic functions that are infinite products of characteristic functions of gamma distributed random variables except in the case of A_∞. From this, we note several relations between the limit distributions. First, $2K_\infty^{+2}$ and $2K_\infty^{-2}$ are exponentially distributed (Smirnov, 1939; Feller, 1948). K_∞ has the Kolmogorov-Smirnov distribution function discussed in this section (Kolmogorov, 1933; Smirnov, 1939; Feller, 1948). Interestingly, V_∞ is distributed as the sum of two independent random variables distributed as K_∞ (Kuiper, 1960). Also, as shown by Watson (1961, 1962), U_∞ is distributed as $\frac{1}{\pi}\sqrt{K_\infty}$. Thus, generation for all these limit distributions poses no problems. Unfortunately, the same cannot be said for A_∞ (Anderson and Darling, 1952) and W_∞ (Smirnov, 1937; Anderson and Darling, 1952).

5.7. Exercises.

1. Prove the following inequality needed in Lemma 5.3: $\log(1+u) \geq \frac{2u}{2+u}$ $(u > 0)$.
2. **The exponential distribution.** For the exponential density, choose a dominating density h from the family of densities

$$\frac{na^n}{(x+a)^{n+1}} \quad (x > 0),$$

where $n \geq 1$ and $a > 0$ are design parameters. Show the following:

(i) h is the density of $a(U^{-\frac{1}{n}} - 1)$ where U is a uniform [0,1] random variable. It is also the density of $a(\max^{-1}(U_1, \ldots, U_n) - 1)$ where the U_i's are iid uniform [0,1] random variables.

(ii) Show that the rejection constant is $c = (\frac{n+1}{e})^{n+1} \frac{e^a a^{-n}}{n}$, and show that this is minimal when $a = n$.

(iv) Show that with $a = n$, we have $c = \frac{1}{e}(1+\frac{1}{n})^{n+1} \to 1$ as $n \to \infty$.

(v) Give the series method based upon rejection from h (where $a=n$ and $n \geq 1$ is an integer). Use quick acceptance and rejection steps based upon the Taylor series expansion.

(vi) Show that the expected time of the algorithm is ∞ when $n=1$ (this shows the danger inherent in the use of the series method). Show also that the expected time is finite when $n \geq 2$.

(Devroye, 1980)

3. Apply the series method for the normal density truncated to $[-a,a]$ with rejection from a uniform density. Since the expected number of iterations is

$$\frac{2a}{\sqrt{2\pi}(F(a)-F(-a))}$$

where F is the normal distribution function, we see that it is important that a be small. How would you handle the tails of the distribution ? How would you choose a for the combined algorithm ?

4. In the study of spectral phenomena, the following densities are important:

(i) $f_1(x) = \frac{1}{\pi}(\frac{\sin(x)}{x})^2$ (the Fejer-de la Vallee Poussin density);

(ii) $f_2(x) = \frac{3}{\pi}(\frac{\sin(x)}{x})^4$ (the Jackson-de la Vallee Poussin density) .

These densities have oscillating tails. Using the fact that

$$\frac{\sin(x)}{x} = 1-\frac{x^2}{3!}+\frac{x^4}{5!}- \cdots ,$$

and that $\frac{\sin(x)}{x}$ falls between consecutive partial sums in this series, derive a good series algorithm for random variate generation for f_1 and f_2. Compare the expected time complexity with that of the obvious rejection algorithms.

5. **The normal distribution.** Consider the series method for the normal density based upon the dominating density $h(x)=\min(a,\frac{1}{16ax^2})$ where $a>0$ is a parameter. Show the following:

(i) If (U,V) are iid uniform $[-1,1]$ random variates, then $\frac{V}{4aU}$ has density h.

(ii) Show that

$$e^{-\frac{x^2}{2}} \leq \max(\frac{1}{a},\frac{32a}{e})h(x)$$

and deduce that the best constant a is $\sqrt{\frac{e}{32}}$.

(iii) Prove that the following algorithm is valid:

Normal generator via the series method

REPEAT

Generate two iid uniform $[-1,1]$ random variates V_1, V_2 and a uniform $[0,1]$ random variate U.

$X \leftarrow \sqrt{\frac{2}{e}} \frac{V_1}{V_2}$

IF $|X| \leq \sqrt{\frac{2}{e}}$

THEN $W \leftarrow \sqrt{\frac{32}{e}} U - 1$

ELSE $W \leftarrow \frac{U}{\sqrt{8e} X^2} - 1$

$n \leftarrow 0, Y \leftarrow \frac{X^2}{2}, P \leftarrow -1$

REPEAT

$n \leftarrow n+1$

$P \leftarrow -\frac{PY}{n}$

$W \leftarrow W+P$

IF $W \leq 0$ THEN RETURN X

$n \leftarrow n+1$

$P \leftarrow -\frac{PY}{n}$

$W \leftarrow W+P$

UNTIL $W > 0$

UNTIL False

(iv) Show that in this algorithm, the expected number of iterations is $\frac{4}{\sqrt{\pi e}}$. (An iteration is defined as a check of the UNTIL False statement or a permanent return.)

6. Erdos and Kac (1946) encountered the following distribution function on $[0,\infty)$:

$$F(x) = \frac{4}{\pi} \sum_{j=0}^{\infty} (-1)^j \frac{1}{2j+1} e^{-(2j+1)^2 \pi^2/(8x^2)} \quad (x > 0) .$$

This shows some resemblance to the Kolmogorov-Smirnov distribution function. Apply the series method to obtain an efficient algorithm for generating random variates with this distribution function. Furthermore, show the identity

$$F(x) = \sum_{j=-\infty}^{\infty} (-1)^j (\Phi((2j+1)x) - \Phi((2j-1)x)) ,$$

where Φ is the normal distribution function (Grenander and Rosenblatt, 1953), which can be of some help in the development of your algorithm.

6. REPRESENTATIONS OF DENSITIES AS INTEGRALS.

6.1. Introduction.

For most densities, one usually first tries the inversion, rejection and mixture methods. When either an ultra fast generator or an ultra universal algorithm is needed, we might consider looking at some other methods. But before we go through this trouble, we should verify whether we do not already have a generator for the density without knowing it. This occurs when there exists a special distributional property that we do not know about, which would provide a vital link to other better known distributions. Thus, it is important to be able to decide which distributional properties we can or should look for. Luckily, there are some general rules that just require knowledge of the shape of the density. For example, by Khinchine's theorem (given in this section), we know that a random variable with a unimodal density can be written as the product of a uniform random variable and another random variable, which turns out to be quite simple in some cases. Khinchine's theorem follows from the representation of the unimodal density as an integral. Other representations as integrals will be discussed too. These include a representation that will be useful for generating stable random variates, and a representation for random variables possessing a Polya type characteristic function. There are some general theorems about such representations which will also be discussed. It should be mentioned though that this section has no direct link with random variate generation, since only probabilistic properties are exploited to obtain a convenient reduction to simpler problems. We also need quite a lot of information about the density in question. Thus, were it not for the fact that several key reductions will follow for important densities, we would not have included this section in the book. Also, representing a density as an integral really boils down to defining a continuous mixture. The only novelty here is that we will actually show how to track down and invent useful mixtures for random variate generation.

6.2. Khinchine's and related theorems.

By far the most important class of densities is the class of unimodal densities. Thus, it is useful to have some integral representations for such densities. Formally, a distribution is called convex on a set A of the real line if for all $x, y \in A$,

$$F(\lambda x+(1-\lambda)y) \leq \lambda F(x)+(1-\lambda)F(y) \quad (0 \leq \lambda \leq 1) .$$

It is concave if the inequality is reversed. It is **unimodal** if it is convex on $(-\infty,0]$ and concave on $[0,\infty)$, and in that case the point 0 is called a mode of the distribution. The rationale for this definition becomes obvious when translated to the density (if it exists). We will not consider other possible locations for the mode to keep the notation simple.

Theorem 6.1. Khinchine's theorem.

A random variable X is unimodal if and only if X is distributed as UY where U,Y are independent random variables: U is uniformly distributed on $[0,1]$ and Y is another random variable not necessarily possessing a density. If Y has distribution function G on $[0,\infty)$, then UY has distribution function

$$F(x) = \int_0^1 G\left(\frac{x}{u}\right) du .$$

Proof of Theorem 6.1.

We refer to Feller (1971, p. 158) for the only if part. For the if part we observe that $P(UY \leq x \mid U=u) = \frac{G(x/u)}{u}$, and thus, integrating over $[0,1]$ with respect to du gives us the result. ∎

To handle the corollaries of Khinchine's theorem correctly, we need to recall the definition of an absolutely continuous function f on an interval $[a,b]$: for all $\epsilon>0$, there exists a $\delta>0$ such that for all nonoverlapping intervals $(x_i,y_i), 1 \leq i \leq n$, and all integers n,

$$\sum_{i=1}^{n} \mid x_i - y_i \mid < \delta$$

implies

$$\sum_{i=1}^{n} \mid f(x_i) - f(y_i) \mid < \epsilon .$$

When f is absolutely continuous on $[a,b]$, its derivative f' is defined almost everywhere on $[a,b]$. Also, it is the indefinite integral of its derivative:

$$f(x) - f(a) = \int_a^x f'(u)\, du \qquad (a \leq x \leq b) .$$

See for example Royden (1968). Thus, Lipschitz functions are absolutely continuous. And if f is a density on $[0,\infty)$ with distribution function F, then F is absolutely continuous,

$$F(x) = \int_0^x f(u)\, du ,$$

and

$$F'(x) = f(x) \text{ almost everywhere .}$$

A density f is called monotone on $[0,\infty)$ (or, in short, monotone) when f is nonincreasing on $[0,\infty)$ and f vanishes on $(-\infty,0)$. However, it is possible that $\lim_{x \downarrow 0} f(x)=\infty$.

Theorem 6.2.

Let X be a random variable with a monotone density f. Then

$$\lim_{x \to \infty} xf(x) = \lim_{x \downarrow 0} xf(x) = 0 .$$

If f is absolutely continuous on all closed intervals of $(0,\infty)$, then f' exists almost everywhere,

$$f(x) = -\int_x^\infty f'(u)\, du ,$$

and X is distributed as UY where U is a uniform $[0,1]$ random variable, and Y is independent of U and has density

$$g(x) = -xf'(x) \quad (x>0) .$$

Proof of Theorem 6.2.

Assume that $\limsup_{x \to \infty} xf(x) \geq 2a > 0$. Then there exists a subsequence $x_1 < x_2 < \cdots$ such that $x_{i+1} \geq 2x_i$ and $x_i f(x_i) \geq a > 0$ for all i. But

$$1 = \int_0^\infty f(x)\, dx \geq \sum_{i=1}^\infty (x_{i+1}-x_i) f(x_{i+1}) \geq \sum_{i=1}^\infty \frac{1}{2} x_{i+1} f(x_{i+1}) = \infty ,$$

which is a contradiction. Thus, $\lim_{x \to \infty} xf(x)=0$.

Assume next that $\limsup_{x \downarrow 0} xf(x) \geq 2a > 0$. Then we can find $x_1 > x_2 > \cdots$ such that $x_{i+1} \leq \frac{x_i}{2}$ and $x_i f(x_i) \geq a > 0$ for all i. Again, a contradiction is obtained:

$$1 = \int_0^\infty f(x)\, dx \geq \sum_{i=1}^\infty (x_i - x_{i+1}) f(x_i) \geq \sum_{i=1}^\infty \frac{1}{2} x_i f(x_i) = \infty .$$

Thus, $\lim_{x \downarrow 0} xf(x)=0$. This brings us to the last part of the Theorem. The first two statements are trivially true by the properties of absolutely continuous functions. Next we show that g is a density. Clearly, $f' \leq 0$ almost everywhere. Also, xf is absolutely continuous on all closed intervals of $(0,\infty)$. Thus, for $0 < a < b < \infty$, we have

$$bf(b)-af(a) = \int_a^b f(x)\,dx + \int_a^b xf'(x)\,dx .$$

By the first part of this Theorem, the left-hand-side of this equation tends to 0 as $a \downarrow 0, b \rightarrow \infty$. By the monotone convergence theorem, the right-hand side tends to $1+\int_0^\infty xf'(x)dx$, which proves that g is indeed a density. Finally, if Y has density g, then UY has density

$$\int_x^\infty \frac{g(u)}{u}\,du = -\int_x^\infty f'(u)\,du = f(x) .$$

This proves the last part of the Theorem. ∎

The extra condition on f in Theorem 6.2 is needed because some monotone densities have $f'=0$ almost everywhere (think of staircase functions). The extra condition in Theorem 6.2 not present in Khinchine's theorem essentially guarantees that the mixing Y variable has a density too. In general, Y needs to have distribution function

$$1-xf(x)-\int_x^\infty f(u)\,du \qquad (x>0) .$$

(exercise 6.9). We also note that Theorem 6.2 has an obvious extension to unimodal densities.

For monotone f that are absolutely continuous on all closed intervals of $(0,\infty)$, the following generator is thus valid:

Generator for monotone densities based on Khinchine's theorem

Generate a uniform [0,1] random variate U.
Generate a random variate Y with density $g(x)=-xf'(x), x>0$.
RETURN $X \leftarrow UY$

Example 6.1. The exponential power distribution (EPD).

Subbotin (1923) introduced the following symmetric unimodal densities:

$$f(x) = \left(2\Gamma(1+\frac{1}{\tau})\right)^{-1} e^{-|x|^\tau} ,$$

where $\tau>0$ is a parameter. This class contains the normal ($\tau=2$) and Laplace ($\tau=1$) densities, and has the uniform density as a limit ($\tau\rightarrow\infty$). By Theorem 6.2, and the symmetry in f, it is easily seen that

$$X \leftarrow VY^{\frac{1}{\tau}}$$

has the given density where V is uniformly distributed on $[-1,1]$ and Y is gamma$(1+\frac{1}{\tau},1)$ distributed. In particular, a normal random variate can be obtained as $V\sqrt{2Y}$ where Y is gamma $(\frac{3}{2})$ distributed, and a Laplace random variate can be obtained as $V(E_1+E_2)$ where E_1,E_2 are iid exponential random variates. Note also that X can be generated as $SY^{1/\tau}$ where Y is gamma $(\frac{1}{\tau})$ distributed. For direct generation from the EPD distribution by rejection, we refer to Johnson (1979). ■

Example 6.2. The Johnson-Tietjen-Beckman family of densities.

Another still more flexible family of symmetric unimodal densities was proposed by Johnson, Tietjen and Beckman (1980):

$$f(x) = \frac{1}{2\Gamma(\alpha)} \int_{x^{\frac{1}{\tau}}}^{\infty} u^{\alpha-\tau-1} e^{-u} \, du ,$$

where $\alpha>0$ and $\tau>0$ are shape parameters. An infinite peak at 0 is obtained whenever $\alpha\leq\tau$. The EPD distribution is obtained for $\alpha=\tau+1$, and another distribution derived by Johnson and Johnson (1978) is obtained for $\tau=\frac{1}{2}$. By Theorem 6.2 and the symmetry in f, we observe that the random variable

$$X \leftarrow VY^{\tau}$$

has density f whenever V is uniformly distributed on $[-1,1]$ and Y is gamma (α) distributed. For the special case $\tau=1$, the gamma-integral distribution is obtained which is discussed in exercise 6.1. ■

Example 6.3. Simple relations between densities.

In the table below, a variety of distributional results are given that can help for the generation of some of them.

Density of Y	Density of UY (U is uniform on [0,1])
Exponential	Exponential-integral ($\int_x^\infty \frac{e^{-u}}{u} \, du$)
Gamma (2)	Exponential
Beta(2,b)	Beta(1,b +1)
Rayleigh ($xe^{-x^2/2}$)	$\int_x^\infty e^{-u^2/2} \, du$
Uniform [0,1]	$-\log(x)$
$(1+a)x^a$ ($x \in [0,1]$) ($a > 0$)	$\frac{a+1}{a}(1-x^a)$
Maxwell ($\frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$)	Normal

■

There are a few other representation theorems in the spirit of Khinchine's theorem. For particular forms, one could consult Lux (1978) and Mikhailov (1965). For the stable distribution discussed in this section, we will need:

Theorem 6.3.

Let U be a uniform [0,1] random variable, let E be an exponential random variable, and let $g : [0,1] \rightarrow [0,\infty)$ be a given function. Then $\frac{E}{g(U)}$ has distribution function

$$F(x) = 1 - \int_0^1 e^{-xg(u)} \, du$$

and density

$$f(x) = \int_0^1 g(u) e^{-xg(u)} \, du .$$

Proof of Theorem 6.3.

For $x > 0$,

$$P\left(\frac{E}{g(U)} > x\right) = P(E > xg(U)) = E(e^{-xg(U)}) = \int_0^1 e^{-xg(u)}\, du \ .$$

The derivative with respect to x is $-f(x)$ where f is defined above. ■

Finally, we mention a useful theorem of Mikhailov's about convolutions with exponential random variables:

Theorem 6.4. (Mikhailov, 1965)

If Y has density f and E is an exponential random variable independent of Y, then $E+Y$ has density

$$h(x) = \int_0^\infty e^{-u} f(x+u)\, du = \int_{-\infty}^{x} f(u) e^{u-x}\, du \ .$$

Furthermore, if g is an absolutely continuous density on $[0,\infty)$ with $g(0)=0$ and $g+g' \geq 0$, then $X \leftarrow E+Y$ has density g where now Y has density $g+g'$, and E is still exponentially distributed.

Proof of Theorem 6.4.

The first statement is trivial. For part two, we note that $g+g'$ is indeed a density since $g+g' \geq 0$ and $\int_0^\infty (g+g')=1$. (This follows from the fact that g is absolutely continuous and has $g(0)=0$.) But then, by partial integration, X has density

$$\int_{-\infty}^{x} (h(u)+h'(u)) e^{u-x}\, du = h(x) \ . ■$$

6.3. The inverse-of-f method for monotone densities.

Assume that f is monotone on $[0,\infty)$ and continuous, and that its inverse f^{-1} can be computed relatively easily. Since f^{-1} itself is a monotone density, we can use the following method for generating a random variate with density f :

The inverse-of-f method for monotone densities

Generate a random variate Y with density f^{-1}.
Generate a uniform [0,1] random variate U.
RETURN $X \leftarrow Uf^{-1}(Y)$

The correctness of the algorithm follows from the fact that (Y,X) is uniformly distributed under the curve of f^{-1}, and thus that (X,Y) is uniformly distributed under the curve of f .

Example 6.4.

If Y is exponentially distributed, then Ue^{-Y} has density $-\log(x)$ $(0<x\leq 1)$ where U is uniformly distributed on [0,1]. But by the well-known connection between exponential and uniform distributions, we see that the product of two iid uniform [0,1] random variables has density $-\log(x)$ $(0<x\leq 1)$. ■

Example 6.5.

If Y has density

$$f^{-1}(y) = \left(\log\left(\frac{2}{\pi y^2}\right)\right)^{\frac{1}{2}} \quad \left(0\leq y \leq \sqrt{\frac{\pi}{2}}\right),$$

and U is uniformly distributed on [0,1], then $X \leftarrow Uf^{-1}(Y)$ has the halfnormal distribution. ■

6.4. Convex densities.

The more we know about a density, the easier it is to generate random variates with this density. There are for example a multitude of tools available for monotone densities, ranging from very specific methods based upon Khinchine's theorem to black box or universal methods. In this section we look at an even smaller class of densities, the convex densities. We will consider the class C_+ of convex densities on $[0,\infty)$, and the class C of densities that are convex on $[0,\infty)$ and on $(-\infty,0)$. Thus, C_+ is a subclass of the monotone densities dealt with in the previous section.

Convex densities are absolutely continuous on all closed subintervals of $(0,\infty)$, and possess monotone right and left derivatives everywhere that are equal except possibly on a countable set. If the second derivative f'' exists, then f is convex if $f'' \geq 0$. We will give one useful representation for convex densities.

Theorem 6.5. (Mixture of triangles)

For every $f \in C_+$, we have

$$f(x) = \int_0^\infty \frac{2}{u}(1-\frac{x}{u})_+ \, dF(u) ,$$

where F is a distribution function with $F(0)=0$ defined by:

$$F(u) = 1+\frac{u^2}{2}f'(u)-(uf(u)+\int_u^\infty f) \quad (u>0) ,$$

where f' is the right-hand derivative of f (which exists on $[0,\infty)$). If F is absolutely continuous, then it has density

$$g(u) = \frac{1}{2}uf''(u) \quad (u>0) .$$

Proof of Theorem 6.5.

We have to show first that if V,Y are independent random variables, where V has a triangular density $2(1-x)_+$ and Y has distribution function F, then $X \leftarrow VY$ has density f. But for $x>0$,

$$\int_x^\infty f = \int_x^\infty (1-\frac{x}{u})^2 \, dF(u)$$

$$= \int_x^\infty dF(u)-2x\int_x^\infty \frac{dF(u)}{u}+x^2\int_x^\infty \frac{dF(u)}{u^2} ,$$

$$f(x) = \int_x^\infty \frac{2}{u}(1-\frac{x}{u})dF(u) = 2\int_x^\infty \frac{dF(u)}{u}-2x\int_x^\infty \frac{dF(u)}{u^2} ,$$

and

$$-f'(x) = 2\int_x^\infty \frac{dF(u)}{u^2} \; .$$

In our case, it can be verified that the interchange of integrals and derivatives is allowed. Substitute the value of f' in the right-hand sides of the equalities for f and $\int_x^\infty f$. Then check that

$$xf(x)+\int_x^\infty f(u)\,du = \int_x^\infty dF(u)+\frac{x^2}{2}f'(x)$$

and this gives us the first result. If F is absolutely continuous, then taking the derivative gives its density, $\frac{x^2}{2}f''(x)$. ∎

This theorem states that for $f \in C_+$, we can use the following algorithm:

Generator for convex densities

Generate a triangular random variate V (this can be done as $\min(U_1,U_2)$ where the U_i's are iid uniform [0,1] random variates).

Generate a random variate Y with distribution function $F(u) = 1+\frac{u^2}{2}f'(u)-(uf(u)+\int_u^\infty f)$ $(u>0)$. (If F is absolutely continuous, then Y has density $\frac{x^2}{2}f''(x)$.)

RETURN $X \leftarrow VY$

6.5. Recursive methods based upon representations.

Representations of densities as integrals lead sometimes to properties of the following kind: assume that three random variables X,Y,Z have densities f,g,h which are related by the decomposition

$$g(x) = ph(x)+(1-p)f(x) \; .$$

Assume that X is distributed as $\psi(Y,U)$ for some function ψ and a uniform [0,1] random variable U independent of Y (this is always the case). Then, we have with probability p, $X \approx \psi(Z, U)$ and with probability $1-p$, $X \approx \psi(\psi(Y', U'), U)$ where (Y', U') is another pair distributed as (Y, U). (The notation $\approx$ is ued for "is distributed as".) This process can be repeated until we reach a substitution by Z. We assume that Z has an easy density h. Notice that we never need to actually generate from g! Formally, we have , starting with Z:

Recursive generator

```
Generate a random variate Z with density h, and a uniform [0,1] random variate U.
X←ψ(Z,U)
REPEAT
    Generate a uniform [0,1]random variate V.
    IF V≤p
        THEN RETURN X
        ELSE
            Generate a uniform [0,1] random variate U.
            X←ψ(X,U)
UNTIL False
```

The expected number of iterations in the REPEAT loop is $\frac{1}{p}$ because the number of V-variates needed is geometrically distributed with parameter p. This algorithm can be fine-tuned in most applications by discovering how uniform variates can be re-used.

Let us illustrate how this can help us. We know that for the gamma density with parameter $a \in (0,1)$,

$$f(x) = \frac{x^{a-1}e^{-x}}{\Gamma(a)} \quad (x > 0):$$

$$g(x) = -xf'(x) = ah(x) + (1-a)f(x),$$

where h is the gamma $(a+1)$ density. This is a convenient decomposition since the parameter of h is greater than one. Also, we know that a gamma (a) random variate is distributed as UY where U is a uniform [0,1] random variate and Y has density $-xf'(x)$ (apply Theorem 6.2). Recall that we have seen several fast gamma generators for $a \geq 1$ but none that was uniformly fast over all a. The previous recursive algorithm would boil down to generating X as

$$Z \prod_{i=1}^{L} U_i$$

where Z is gamma $(a+1)$ distributed, L is geometric with parameter a, and the U_i's are iid uniform [0,1] random variates. Note that this in turn is distributed as Ze^{-G_L} where G_L is a gamma (L) random variate. But the density of G_L is

$$\sum_{i=1}^{\infty} a(1-a)^{i-1} \frac{x^{i-1} e^{-x}}{(i-1)!} = e^{-ax} \quad (x>0) .$$

Thus, we have shown that the following generator is valid:

A gamma generator for a < 1

Generate a gamma ($a+1$) random variate Z.

Generate an exponential random variate E.

RETURN $X \leftarrow Ze^{-\frac{E}{a}}$

The recursive algorithm does not require exponentiation, but the expected number of iterations before halting is $\frac{1}{a}$, and this is not uniformly bounded over (0,1). The algorithm based upon the decomposition as $Ze^{-\frac{E}{a}}$ on the other hand is uniformly fast.

Example 6.6. Stuart's theorem.

Without knowing it, we have proved a special case of a theorem of Stuart's (Stuart, 1962): if Z is gamma (a) distributed, and Y is beta ($b, a-b$) distributed and independent of Z, then $ZY, Z(1-Y)$ are independent gamma (b) and gamma($a-b$) random variables. If we put $b=1$, and formally replace a by $a+1$ then it is clear that $ZU^{\frac{1}{a}}$ is gamma (a) distributed, where U is a uniform [0,1] random variable. ■

There are other simple examples. The von Neumann exponential generator is also based upon a recursive relationship. It is true that an exponential random variate E is with probability $1-\frac{1}{e}$ distributed as a truncated exponential random variate (on [0,1]) , and that E is with probability $\frac{1}{e}$ distributed as $1+E$. This recursive rule leads precisely to the exponential generator of section IV.2.

6.6. A representation for the stable distribution.

The standardized stable distribution is best defined in terms of its characteristic function ϕ:

$$\log\phi(t) = \begin{cases} -\mid t \mid^{\alpha} e^{-i \frac{\pi}{2} \bar{\alpha} \delta \operatorname{sgn}(t)} & (\alpha \neq 1) \\ -\mid t \mid (1+i\ \delta \frac{2}{\pi} \operatorname{sgn}(t) \log(\mid t \mid)) & (\alpha = 1) \end{cases}$$

Here $\delta \in [-1,1]$ and $\alpha \in (0,2]$ are the shape parameters of the stable distribution, and $\bar{\alpha}$ is defined by $\min(\alpha,2-\alpha)$. We omit the location and scale parameters in this standard form. To save space, we will say that X is stable(α,δ) when it has the above mentioned characteristic function. This form of the characteristic function is due to Zolotarev (1959). By far the most important subclass is the class of symmetric stable distributions which have $\delta=0$: their characteristic function is simply

$$\phi(t) = e^{-\mid t \mid^{\alpha}} .$$

Despite the simplicity of this characteristic function, it is quite difficult to obtain useful expressions for the corresponding density except perhaps in the special cases $\alpha=2$ (the normal density) and $\alpha=1$ (the Cauchy density). Thus, it would be convenient if we could generate stable random variates without having to compute the density or distribution function at any point. There are two useful representations that will enable us to apply Theorem 6.4 with a slight modification. These will be given below.

Theorem 6.6. (Ibragimov and Chernin, 1959; Kanter, 1975)

For $\alpha<1$, the density of a stable$(\alpha,1)$ random variable can be written as

$$f(x) = \frac{\alpha x^{\frac{1}{\alpha-1}}}{(1-\alpha)\pi} \int_0^{\pi} g(u) e^{-g(u) x^{\frac{\alpha}{\alpha-1}}} du ,$$

where

$$g(u) = \left(\frac{\sin(\alpha u)}{\sin(u)}\right)^{\frac{1}{1-\alpha}} \frac{\sin((1-\alpha)u)}{\sin(\alpha u)} .$$

When U is uniformly distributed on $[0,1]$ and E is independent of U and exponentially distributed, then

$$\left(\frac{g(\pi U)}{E}\right)^{\frac{1-\alpha}{\alpha}}$$

is stable$(\alpha,1)$ distributed.

Proof of Theorem 6.6.

For the first statement, we refer to Ibragimov and Chernin (1959). The latter statement is an observation of Kanter's (1975) which is quite easily verified by computing the distribution function of $(\frac{g(\pi U)}{E})^{\frac{1-\alpha}{\alpha}}$, and noting that it is equal to

$$\frac{1}{\pi}\int_0^{\pi} e^{-g(u)x^{\frac{\alpha}{\alpha-1}}}\, du \ .$$

Taking the derivative gives us the density f . ∎

The second part of the proof uses a slight extension of Theorem 6.4. This representation allows us to generate stable(α,1) random variates quite easily - in most computer languages, one line of computer code will suffice! There are two problems however. First, we are stuck with the evaluation of several trigonometric functions and of two powers. We will see some methods of generating stable random variates that do not require such costly operations, but they are much more complicated. Our second problem is that Theorem 6.6 does not cover the case $\delta \neq 1$. But this is easily taken care of by the following Lemma for which we refer to Feller (1971):

Lemma 6.1.

A. If X and Y are iid stable(α,1), then $Z \leftarrow pX - qY$ is stable(α,δ) where

$$p^{\alpha} = \sin(\frac{\pi\overline{\alpha}(1+\delta)}{2})/\sin(\pi\overline{\alpha}) \ ,$$

$$q^{\alpha} = \sin(\frac{\pi\overline{\alpha}(1-\delta)}{2})/\sin(\pi\overline{\alpha}) \ .$$

B. If X is stable($\frac{\alpha}{2}$,1) and N is independent of X and normally distributed, then $N\sqrt{2X}$ is stable(α,0), all $\alpha \in (0,2]$.

Using this Lemma and Theorem 6.6, we see that we can generate all stable random variates with either $\alpha < 1$ or $\delta = 0$. To fill the void, Chambers, Mallows and Stuck (1976) proposed to use a representation of Zolotarev's (1966):

Theorem 6.7. (Zolotarev, 1966; Chambers, Mallows and Stuck,1976)

Let E be an exponential random variable, and let U be a uniform $[-\frac{\pi}{2},\frac{\pi}{2}]$ random variable independent of E. Let further $\gamma = -\frac{\pi\delta\bar{\alpha}}{2\alpha}$. Then, for $\alpha \neq 1$,

$$X \leftarrow \frac{\sin(\alpha(U-\gamma))}{(\cos U)^{\frac{1}{\alpha}}} \left(\frac{\cos(U-\alpha(U-\gamma))}{E}\right)^{\frac{1-\alpha}{\alpha}}$$

is stable(α,δ) distributed. Also,

$$X \leftarrow \frac{2}{\pi}\left(\left(\frac{\pi}{2}+\delta U\right)\tan(U)-\delta\log\left(\frac{\pi E\cos(U)}{\pi+2\delta U}\right)\right)$$

is stable(1,δ) distributed.

We leave the determination of the integral representation of f to the reader. It is noteworthy that Theorem 6.7 is a true extension of Theorem 6.6 (just note that for $\alpha<1,\delta=1$, we obtain $\gamma=-\frac{\pi}{2}$. There are three special cases worth noting:

(i) A stable(2,0) random variate can be generated as $\sqrt{E}\frac{\sin(2U)}{\cos(U)} = 2\sqrt{E}\sin(U)$. This is the well-known Box-Muller representation of $\sqrt{2}$ times a normal random variate (see section V.4).

(ii) A stable(1,0) random variate can be obtained as $\tan(U)$, which yields the inversion method for generating Cauchy random variates.

(iii) A stable($\frac{1}{2}$,1) random variate can be obtained as

$$\frac{1}{4E\sin^2\left(\frac{U}{2}-\frac{\pi}{4}\right)},$$

which is distributed in turn as

$$\frac{1}{4E\cos^2(U)},$$

which is in turn distributed as $\frac{1}{2N^2}$ where N is normally distributed.

6.7. Densities with Polya type characteristic functions.

This section is added because it illustrates that representations offer unexpected help in many ways. It is frustrating to come across a distribution with a very simple characteristic function in one's research, and not be able to generate random variates with this characteristic function, at least not without a lot of work. But we do know of course how to generate random variates with some characteristic functions such as normal, uniform and exponential random variates. Thus, if we can find a representation of the characteristic function ϕ in terms of one of these simpler characteristic functions, then there is hope of generating random variates with characteristic function ϕ. By this process, we can take care of quite a few characteristic functions, even some for which the density is not known in a simple analytic form. This will be illustrated now for the class of Polya characteristic functions, i.e. real even continuous functions ϕ with $\phi(0)=1$, $\lim_{t\to\infty}\phi(t)=0$, convex on $(0,\infty)$. This class is important both from a practical point of view (it contains many important distributions) and from a didactical point of view. The examples that we will consider in this subsection are listed in the table below.

Characteristic function $\phi(t)$	Name
$e^{-\lvert t\rvert^\alpha}, 0<\alpha\leq 1$	Symmetric stable distribution
$\frac{1}{1+\lvert t\rvert^\alpha}, 0<\alpha\leq 1$	Linnik's distribution
$(1-\lvert t\rvert)^\alpha, \lvert t\rvert\leq 1, \alpha\geq 1$	
$1-\lvert t\rvert^\alpha, \lvert t\rvert\leq 1, 0<\alpha\leq 1$	

The second entry in this table is the characteristic function of a unimodal density for $\alpha\in(0,2]$ (Linnik (1962), Lukacs (1970, pp. 96-97)), yet no simple form for the density is known. We are now ready for the representation.

Theorem 6.8. (Girault, 1954; Dugue and Girault, 1955)

Every Polya characteristic function ϕ can be decomposed as follows:

$$\phi(t) = \int_0^\infty \left(1-\left|\frac{t}{s}\right|\right)_+ dF(s) \quad (t>0),$$

$$\phi(t) = -\phi(-t) \quad (t<0),$$

where F is a distribution function with $F(0)=0$ and defined by

$$F(s) = 1-\phi(s)+s\,\phi'(s) \quad (s>0).$$

Here ϕ' is the right-hand derivative of ϕ (which exists everywhere). If F is absolutely continuous, then it has density

$$g(s) = s\,\phi''(s) \quad (s>0).$$

From this, it is a minor step to conclude:

Theorem 6.9. (Devroye, 1984)

If ϕ is a Polya characteristic function, then $X \leftarrow \frac{Y}{Z}$ has this characteristic function when Y, Z are independent random variables: Z has the distribution function F of Theorem 6.8, and Y has the Fejer-de la Vallee Poussin (or: FVP) density

$$\frac{1}{2\pi}\left(\frac{\sin(\frac{x}{2})}{\frac{x}{2}}\right)^2 .$$

Theorem 6.9 uses Theorem 6.8 and the fact that the FVP density has characteristic function $(1-|t|)_+$. There are but two things left to do now: first, we need to obtain a fast FVP generator because it is used for all Polya type distributions. Second, it is important to demonstrate that the distribution function F in the various examples is often quite simple and easy to handle.

Remark 6.1. A generator for the Fejer-de la Vallee Poussin density.

Notice that if X has density

$$\frac{1}{\pi}\left(\frac{\sin(x)}{x}\right)^2 ,$$

then $2X$ has the FVP density. In view of the oscillating behavior of this density, it is best to proceed by the rejection method or the series method We note first that $\sin(x)$ is bounded from above and below by consecutive terms in the series expansion

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots ,$$

and that it s bounded in absolute value by 1. Thus, the density f of X is bounded as follows:

$$f(x) \le \frac{4}{\pi} h(x) ,$$

where $h(x) = \min(\frac{1}{4}, \frac{1}{4x^2})$, which is the density of V^B, where V is a uniform $[-1,1]$ random variable, and B is ± 1 with equal probability. The rejection constant of $\frac{4}{\pi}$ in this inequality is usually quite acceptable. Thus, we have:

FVP generator based upon rejection

REPEAT
    Generate iid uniform [-1,1] random variates U, X.
    IF $U < 0$
        THEN
            $X \leftarrow \frac{1}{X}$
            Accept $\leftarrow [\,|U| \le \sin^2(X)]$
        ELSE Accept $\leftarrow [\,|U|\,X^2 \le \sin^2(X)]$
UNTIL Accept
RETURN $2X$

The expected time can be reduced by the judicious use of squeeze steps. First, if $|X|$ is outside the range $[0,\frac{\pi}{2}]$, it can always be reduced to a value within that range (as far as the value of $\sin^2(X)$ is concerned). Then there are two cases:

(i) If $|X| \le \frac{\pi}{4}$, we can use

$$X - \frac{X^3}{6} \le \sin(X) \le X \ .$$

(ii) If $|X| \in (\frac{\pi}{4}, \frac{\pi}{2}]$, then we can use the fact that $\sin(X) = \cos(\frac{\pi}{2} - X) = \cos(Y)$, where Y now is in the range of (i). The following inequalities will be helpful:

$$1 - \frac{Y^2}{2} \le \sin(X) \le 1 - \frac{Y^2}{2} + \frac{Y^4}{24} \ . \blacksquare$$

Example 6.7. The symmetric stable distribution.

In Theorem 6.9, Z has density g given by

$$g(s) = (\alpha^2 s^{2\alpha-1} + \alpha(1-\alpha) s^{\alpha-1}) e^{-s} \quad (s > 0) \ .$$

But we note that Z^α has density

$$\alpha(s e^{-s}) + (1-\alpha)(e^{-s}) \quad (s > 0) \ ,$$

which is a mixture of a gamma (2) and an exponential density. Thus, Z is distributed as

$$(E_1+E_2 I_{[U<\alpha]})^{\frac{1}{\alpha}}$$

where E_1, E_2 and U are independent random variables: E_1 and E_2 have an exponential density, and U is uniformly distributed on [0,1]. It is also worth observing that if we use $U_1, \ldots$ for iid uniform [0,1] random variables, then Z is distributed as

$$(E_1+\max(E_2+\log(\alpha),0))^{\frac{1}{\alpha}}$$

and as

$$\log^{\frac{1}{\alpha}}(\max(\frac{\alpha}{U_1 U_2},\frac{1}{U_1})) . \blacksquare$$

Example 6.8. Linnik's distribution

We verify that Z in Theorem 6.9 has density g given by

$$g(s) = ((\alpha^2+\alpha)s^{2\alpha-1}+(\alpha-\alpha^2)s^{\alpha-1})(1+s^{\alpha})^{-3} \quad (s>0) .$$

It is perhaps easier to work with the density of Z^{α}:

$$\frac{s(\alpha+1)+(1-\alpha)}{(1+s)^3} \quad (s>0) .$$

The latter density has distribution function $1-\frac{1+\alpha}{1+s}+\frac{\alpha}{(1+s)^2}$, and this is easy to invert. Thus, a random variate Z can be generated as

$$(\frac{\alpha+1-\sqrt{(\alpha+1)^2-4\alpha U}}{2U}-1)^{\frac{1}{\alpha}} ,$$

where U is a uniform [0,1] random variate. If speed is extremely important, the square root can be avoided if we use the rejection method for the density of Z^{α}, with dominating density $(1+s)^{-2}$, which is the density of $\frac{1}{U}-1$. A little work shows that Z can be generated as follows:

REPEAT

Generate iid uniform [0,1] random variates U, V.

$X \leftarrow \frac{1}{U}-1$

UNTIL $2\alpha U \leq V$ (Now, X is distributed as Z^α.)

RETURN $X^{\frac{1}{\alpha}}$

The expected number of iterations is $1+\alpha$. ■

Example 6.9. Other examples.

Assume that $\phi(t) = (1-|t|)_+^\alpha$ for $\alpha>1$. Then $\phi(s)-s\,\phi'(s)$ is absolutely continuous. Thus, the random variable Z of Theorem 6.9 has beta $(2,\alpha-1)$ density $g(s)=\alpha(\alpha-1)s(1-s)^{\alpha-2}$ $(0\leq s\leq 1)$.

There are situations in which the distribution function F of Theorems 6.8 and 6.9 is not absolutely continuous. To illustrate this, take $\phi(t)=(1-|t|^\alpha)_+$, and note that $F(s) = (1-\alpha)s^\alpha$ $(0\leq s\leq 1)$. Also, $F(1)=1$. Thus, F has an atom of weight α at 1, and it has an absolutely continuous part of weight $1-\alpha$ with support on (0,1). The absolutely continuous part has density $\alpha s^{\alpha-1}$ $(0\leq s\leq 1)$, which is the density of $U^{\frac{1}{\alpha}}$ where U is uniform on [0,1]. Thus,

$$Z = \begin{cases} 1 & \text{with probability } \alpha \\ U^{\frac{1}{\alpha}} & \text{with probability } 1-\alpha \end{cases} .$$

Here we can use the standard trick of recuperating part of the uniform [0,1] random variate used to make the "with probability α" choice. ■

A. f is convex if and only if $a,b \geq 1$. It is concave if and only if $a,b \leq 1$.
B. Y^b has density f, where Y is beta $(b, a+1)$ distributed.
C. $(\frac{Y}{Y+Z})^b$ has density f where Y is gamma (b) distributed, and Z is gamma $(a+1)$ distributed and independent of Y.

6. This is a continuation of exercise 5 for the special case $b=1$. The density is $f(x)=(a+1)(1-x)^a$ $(0 \leq x \leq 1)$. From the previous exercise we recall that a random variate with this distribution can be obtained as $1-U^{\frac{1}{a+1}}$ and as $\frac{E}{E+G_{a+1}}$ where U is a uniform [0,1] random variate, E is an exponential random variate, and G_{a+1} is a gamma $(a+1)$ random variate independent of E. Both these methods require costly operations. The following rejection algorithms are usually faster:

Rejection method #1, recommended for a > 1

REPEAT
    REPEAT
        Generate two iid exponential random variates, E_1, E_2.
        $X \leftarrow \frac{E_1}{a}$
    UNTIL $X \leq 1$
    Accept $\leftarrow [E_2(1-X)-aX^2 \geq 0]$
    IF NOT Accept THEN Accept $\leftarrow [aX+E_2+a\log(1-X) \geq 0]$
UNTIL Accept
RETURN X

Rejection method #2, recommended for a < 1

REPEAT
    Generate two iid uniform [0,1] random variates, U, X.
UNTIL $U \leq (1-X)^a$
RETURN X

Show that the rejection algorithms are valid. Show furthermore that the expected number of iterations is $\frac{a+1}{a}$ and $a+1$ respectively. (Thus, a uniformly fast algorithm can be obtained by using the first method for $a \geq 1$

6.8. Exercises.

1. **The gamma-integral distribution.** We say that X is GI(a) (has the gamma-integral distribution with parameter $a>0$) when its density is

$$f(x)=\int_x^\infty \frac{u^{a-2}e^{-u}}{\Gamma(a)}\,du \qquad (x>0) .$$

This distribution has a few remarkable properties: it decreases monotonically on $[0,\infty)$. It has an infinite peak at 0 when $a\leq 1$. At $a=1$, we obtain the exponential-integral density. When $a>1$, we have $f(0)=\frac{1}{a-1}$. For $a=2$, the exponential density is obtained. When $a>2$, there is a point of inflection at $a-2$, and $f'(0)=0$. For $a=3$, the distribution is very close to the normal distribution. In this exercise we are mainly interested in random variate generation. Show the following:

A. X can be generated as UY where U is uniformly distributed on [0,1] and Y is gamma (a) distributed.

B. When a is integer, X is distributed as G_Z where Z is uniformly distributed on $1, \ldots, a-1$, and G_Z is a gamma (Z) random variate. Note that X is distributed as $-\log(U_1 \cdots U_Z)$ where the U_i's are iid uniform [0,1] random variates. Hint: use induction on a.

C. As $a\rightarrow\infty$, $\frac{X}{a}$ tends in distribution to the uniform [0,1] density.

D Compute all moments of the GI(a) distribution. (Hint: use Khinchine's theorem.)

2. The density of the energy spectrum of fission neutrons is

$$f(x)=\frac{1}{\sqrt{\pi ab}}e^{-(a+x)b}\sinh\left(\frac{2\sqrt{ax}}{b}\right) \qquad (x>0) ,$$

where $a,b>0$ are parameters. Recall that $\sinh(x)=\frac{1}{2}(e^x-e^{-x})$. Apply Theorem 6.4 for designing a generator for this distribution(Mikhailov, 1965).

3. How would you compute $f(x)$ with seven digits of accuracy for the exponential-integral density of Example 6.3? Prove also that for the same distribution, $F(x)=(1-e^{-x})+xf(x)$ where F is the distribution function.

4. If U,V are iid uniform [0,1] random variables, then for $0<a<1$, $UV^{\frac{1}{1-a}}$ has density $x^{-a}-1$ $(0<x<1)$.

5. In the next three exercises, we consider the following class of monotone densities on [0,1]:

$$f(x)=\frac{\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(b+1)}(1-x^{\frac{1}{b}})^a \qquad (0\leq x\leq 1) ,$$

where $a,b>0$ are parameters. The coefficient will be called B. The mode of the density occurs at $x=0$, and $f(0)=B$. Show the following:

and the second method for $a < 1$.)

7. Continuation of exercise 5 for $b = \frac{1}{2}$. The density we are considering here can be written as follows:

$$f(x) = B(1-x^2)^a \quad (0 \le x \le 1) .$$

(Here $B = \frac{2}{\sqrt{\pi}} \frac{\Gamma(a+\frac{3}{2})}{\Gamma(a+1)}$.) From exercise 5 we recall that a random variate with this density can be generated as $\frac{N}{\sqrt{N^2+2G_{a+1}}}$ where N is a normal random variate, and G_{a+1} is a gamma $(a+1)$ random variate independent of N.

A. Show that we can also use $|2Y-1|$ where Y is beta $(a+1,a+1)$ distributed.

B. Show that if we keep generating iid uniform [0,1] random variates U,X until $U \le (1-X^2)^a$, then X has density f, the expected number of iterations is B, and B increases monotonically from 1 $(a=0)$ to ∞ $(a \to \infty)$.

C. Show that the following rejection algorithm is valid and has rejection constant $\frac{\Gamma(a+\frac{3}{2})}{\sqrt{a}\,\Gamma(a+1)}$ (which tends monotonically to 1 as $a \to \infty$):

Rejection from a normal density

REPEAT

Generate independent normal and exponential random variates N,E.

$X \leftarrow \frac{|N|}{\sqrt{2a}}$, $Y \leftarrow X^2$

Accept $\leftarrow [Y \le 1]$AND$[1-Y(1+\frac{aY}{E}) \ge 0]$

IF NOT Accept THEN Accept $\leftarrow [Y \le 1]$ AND $[aY+E+a\log(1-Y) \ge 0]$

UNTIL Accept

RETURN X

Hint: use the inequalities $-\frac{x}{1-x} \le \log(1-x) \le -x \quad (0 < x < 1)$.

8. **The exponential power distribution.** Show that if S is a random sign, and $G_{\frac{1}{\tau}}$ is a gamma $(\frac{1}{\tau})$ random variate, then $S(G_{\frac{1}{\tau}})^{\frac{1}{\tau}}$ has the exponential

power distribution with parameter τ, that is, its density is of the form $ce^{-|x|^\tau}$ where c is a normalization constant.

9. Extend Theorem 6.2 by showing that for all monotone densities, it suffices to take Y with distribution function

$$F(x) = 1-\int_x^\infty f(u)\,du - xf(x) \quad (x \in R).$$

10. Extend Theorem 6.5 to all convex densities in C.

11. **The Pareto distribution.** Let E, Y be independent random variables, where E is exponentially distributed, and Y has density g on $[0,\infty)$. Give an integral form for the density and distribution function of $X = E/Y$. Random variables of this type are called exponential scale mixtures. Show that when Y is gamma (a), then $1+E/Y$ is Pareto with parameter a, i.e. $1+E/Y$ has density a/x^{a+1} $(x>1)$ (see e.g. Harris, 1968).

12. Develop a uniformly fast generator for the family of densities

$$f(x) = C_n \left(\frac{\sin(x)}{x}\right)^n ,$$

where $n \geq 1$ is an integer parameter, and C_n is a constant depending upon n only.

7. THE RATIO-OF-UNIFORMS METHOD.

7.1. Introduction.

The rejection method has one big drawback: densities with infinite tails have to be handled with care; often, tails have to be cut off and treated separately. In many cases, this can be avoided if the ratio-of-uniforms method is used. This method is particularly well suited for bell-shaped densities with tails that decrease at least as fast as x^{-2}. The ratio-of-uniforms method was first proposed by Kinderman and Monahan (1977), and later applied to a variety of distributions such as the t distribution (Kinderman and Monahan,1979) and the gamma distribution (Cheng and Feast, 1979).

Because the resulting algorithms are short and often fast, and because we have yet another beautiful illustration of the rejection and squeeze principles, we will devote quite a bit of space to this method. The treatment will be systematic and simple: we are not looking for the most general form of algorithm but for one that is easy to understand.

We begin with

Theorem 7.1. (Kinderman and Monahan, 1977)

Let $A=\{(u,v):0\leq u\leq\sqrt{f(\frac{v}{u})}\}$ where $f\geq 0$ is an integrable function. If (U,V) is a random vector uniformly distributed over A, then $\frac{V}{U}$ has density $\frac{1}{c}f$ where $c=\int f=2$ area (A).

Proof of Theorem 7.1.

Define (X,Y) by $X=U,Y=\frac{V}{U}$. The Jacobian of the transformation $u=x,v=xy$ is x. The density of (U,V) is $I_A(u,v)/(c/2)$. Thus, the density of (X,Y) is x times $I_A(x,yx)/(c/2)=xI_{[0,f(y)]}(x)/(c/2)$. The density of $Y=\frac{V}{U}$ is the marginal density computed as

$$\int_0^{\sqrt{y}}\frac{x}{(c/2)}dx=\frac{f(y)}{c}\ .\blacksquare$$

But we already know how to generate uniformly distributed random vectors: it suffices to enclose the area A by a simple set such as a rectangle, in which we know how to generate uniform random vectors, and to apply the rejection principle. Thus, it is important to verify what A looks like in general. First, A is a subset of $[0,\infty)\times R$. It is symmetric about the u-axis if f is symmetric about 0. It vanishes in the negative v-quadrant when f is the density of a nonnegative random variable. But what interests us more than anything else are conditions insuring that $A\subseteq[0,b)\times[a_-,a_+]$ for some finite constants $b\geq 0,a_-\leq 0,a_+\geq 0$. It helps to note that the boundary of A can be found parametrically by $\{(u(x),v(x)):x\in R\}$ where

$$u(x)=\sqrt{f(x)}\ ,$$
$$v(x)=x\sqrt{f(x)}\ .$$

Thus, A can be enclosed in a rectangle if and only if

(i) $f(x)$ is bounded;
(ii) $x^2f(x)$ is bounded.

Basically, this includes all bounded densities with subquadratic tails, such as the normal, gamma, beta, t and exponential densities. From now on, the enclosing rectangle will be called $B=[0,b)\times[a_-,a_+]$. For the sake of simplicity, we will only treat densities satisfying (i) and (ii) in this section.

The ratio-of-uniforms method

[SET-UP]

Compute b, a_-, a_+ for an enclosing rectangle. Note that $b \geq \sup \sqrt{f(x)}, a_- \leq \inf x\sqrt{f(x)}, a_+ \geq \sup x\sqrt{f(x)}$.

[GENERATOR]

REPEAT

Generate U uniformly on $[0,b]$, and V uniformly on $[a_-,a_+]$.

$X \leftarrow \frac{V}{U}$

UNTIL $U^2 \leq f(X)$

RETURN X

By Theorem II.3.2, (U,V) is uniformly distributed in A. Thus, the algorithm is valid, i.e. X has density proportional to the function f. We can also replace f by cf for any constant c. This allows us to eliminate all annoying normalization constants. In any case, the expected number of iterations is

$$\frac{b(a_+-a_-)}{\text{area } A} = \frac{2b(a_+-a_-)}{\int_{-\infty}^{\infty} f(x)\,dx} .$$

This will be called the rejection constant. Good densities are densities in which A fills up most of its enclosing rectangle. As we will see from the examples, this is usually the case when f puts most of its mass near zero and has monotonically decreasing tails. Roughly speaking, most bell-shaped f are acceptable candidates.

The acceptance condition $U^2 \leq f(X)$ cannot be simplified by using logarithmic transformations as we sometimes did in the rejection method - this is because U is explicitly needed in the definition of X. The next best thing is to make sure that we can avoid computing f most of the time. This can be done by introducing one or more quick acceptance and quick rejection steps. Typically, the algorithm takes the following form.

The ratio-of-uniforms method with two-sided squeezing

[SET-UP]

Compute b, a_-, a_+ for an enclosing rectangle. Note that $b \geq \sup\sqrt{f(x)}, a_- \leq \inf x\sqrt{f(x)}, a_+ \geq \sup x\sqrt{f(x)}$.

[GENERATOR]

REPEAT

Generate U uniformly on $[0,b]$, and V uniformly on $[a_-,a_+]$.

$X \leftarrow \frac{V}{U}$

IF [Quick acceptance condition]

THEN Accept ← True

ELSE IF [Quick rejection condition]

THEN Accept ← False

ELSE Accept ← [Acceptance condition ($U^2 \leq f(X)$)]

UNTIL Accept

RETURN X

In the next sub-section, we will give various quick acceptance and quick rejection conditions for the distributions listed in this introduction, and analyze the performance for these examples.

7.2. Several examples.

We will need various inequalities in the design of squeeze steps. The following Lemma can be useful in this respect.

Lemma 7.1.

(i) $-x \geq \log(1-x) \geq -\frac{x}{1-x} \quad (0 \leq x < 1)$.

(ii) $-x-\frac{x^2}{2} \geq \log(1-x)$

$\geq -x-\frac{x^2}{2(1-x)} \quad (0 \leq x < 1)$.

(iii) $\log(x) \leq x-1 \quad (x > 0)$.

(iv) $x-\frac{x^2}{2} \leq \log(1+x)$

$\leq x-\frac{x^2}{2}+\frac{x^3}{3} \leq x \quad (0 < x < 1)$.

(v) $\frac{2x+3x^2}{2(1+x)^2} \leq \log(1+x)$

$\leq \frac{2x+3x^2+x^3}{2(1+x)^2}$

$= x-\frac{x^2}{2(1+x)} \quad (x \geq 0)$.

(vi) Reverse the inequalities in (v) when $-1 < x \leq 0$.

Proof of Lemma 7.1.

Parts (i) through (iv) were obtained in Lemma IV.3.2. By the Taylor series for $g(x)=(1+x)\log(1+x)$, we see that

$$g(x) = g(0)+xg'(0)+\frac{x^2}{2}g''(\xi)$$

for some ξ between 0 and x. But $g(0)=0, g'(x)=\log(1+x)-1, g'(0)=1, g''(x)=\frac{1}{1+x}$. Thus, for $x > 0$,

$$x+\frac{x^2}{2(1+x)} \leq g(x) \leq x+\frac{x^2}{2} .$$

This proves (v) and (vi). ■

For various densities, we list quick acceptance and rejection conditions in terms of u, v, x. When used in the algorithm, these running variables should be replaced by the random variates U, V, X of course. Other useful quantities such as the rejection constant and values for b, a_-, a_+ are listed too.

Example 7.1. The normal density.

All of the above is summarized in the table given below:

$f(x)$	$e^{-\frac{x^2}{2}}$ $(x \in R)$
$b = \sup \sqrt{f(x)}$	1
$a_+ = \sup x \sqrt{f(x)}, a_- = \inf x \sqrt{f(x)}$	$\sqrt{\frac{2}{e}}, -\sqrt{\frac{2}{e}}$
area (A)	$\sqrt{\frac{\pi}{2}}$
Rejection constant	$\frac{4}{\sqrt{\pi e}}$
Acceptance condition	$x^2 \leq -4\log u$
Quick acceptance condition	$x^2 \leq 4(-cu+1+\log c)$ $(c>0)$
	$x^2 \leq 4-4u$
	$x^2 \leq 6-8u+2u^2$
	$x^2 \leq \frac{44}{6}-12u+6u^2-\frac{4}{3}u^3$
Quick rejection condition	$x^2 > 4(\frac{c}{u}-1-\log c)$ $(c>0)$
	$x^2 \geq \frac{4}{u}-4$
	$x^2 \geq \frac{2}{u}-2u$

The table is nearly self-explanatory. The quick acceptance and rejection conditions were obtained from the acceptance condition and Lemma 7.1. Most of these are rather straightforward. The fastest experimental results were obtained with the third entries in both lists. It is worth pointing out that the first quick acceptance and rejection conditions are valid for all constants $c>0$ introduced in the conditions, by using inequalities for $\log(uc)$ given in Lemma 7.1. The parameter c should be chosen so that the area under the quick acceptance curve is maximal, and the area under the quick rejection curve is minimal. ■

Example 7.2. The exponential density.

In analogy with the normal density, we present the following table.

$f(x)$	e^{-x} $(x \in R)$
$b = \sup \sqrt{f(x)}$	1
$a_+ = \sup x \sqrt{f(x)}, a_- = \inf x \sqrt{f(x)}$	$\frac{2}{e}, 0$
area (A)	$\frac{2}{e}$
Rejection constant	$\frac{4}{e}$
Acceptance condition	$x \le -2\log u$
Quick acceptance condition	$x \le 2(1-u)$
Quick rejection condition	$x \ge \frac{2}{u} - 2$
	$x \ge \frac{2}{eu} - \frac{(u-\frac{1}{e})^2}{u}$

It is insightful to draw A and to construct simple quick acceptance and rejection conditions by examining the shape of A. Since A is convex, several linear functions could be useful. ∎

Example 7.3. The t distribution.

The ratio-of-uniforms method has led to some of the fastest known algorithms for the t distribution. In this section, we omit, as we can, the normalization constant of the t density with parameter a, which is

$$\frac{\Gamma(\frac{a+1}{2})}{\sqrt{\pi a}\,\Gamma(\frac{a}{2})} .$$

Since for large values of a, the t density is close to the normal density, we would expect that the performance of the algorithm would be similar too. This is indeed the case. For example, as $a \to \infty$, the rejection constant tends to $\frac{4}{\sqrt{\pi e}}$, which is the value for the normal density.

$f(x)$	$\frac{1}{(1+\frac{x^2}{a})^{\frac{a+1}{2}}}$ $(x \in R)$
$b=\sup\sqrt{f(x)}$	1
$a_+=\sup x\sqrt{f(x)}, a_-=\inf x\sqrt{f(x)}$	$\frac{\sqrt{2a}(a-1)^{\frac{a-1}{4}}}{(a+1)^{\frac{a+1}{4}}}, \frac{\sqrt{2a}(a-1)^{\frac{a-1}{4}}}{(a+1)^{\frac{a+1}{4}}}$
area (A)	$2\frac{\sqrt{2a}(a-1)^{\frac{a-1}{4}}}{(a+1)^{\frac{a+1}{4}}}$
Rejection constant	$4\frac{\sqrt{2a}(a-1)^{\frac{a-1}{4}}}{(a+1)^{\frac{a+1}{4}}} \frac{\Gamma(\frac{a+1}{2})}{\sqrt{\pi a}\,\Gamma(\frac{a}{2})}$
Acceptance condition	$x^2 \le a(u^{-\frac{4}{a+1}}-1)$
Quick acceptance condition	$x^2 \le 5-4u(1+\frac{1}{a})^{\frac{a+1}{4}}$
Quick rejection condition	$x^2 \ge -3+\frac{4}{u}(1+\frac{1}{a})^{-\frac{a+1}{4}}$ (only valid for $a \ge 3$)

We observe that the ratio-of-uniforms method can only be useful when $a \ge 1$ for otherwise A would be unbounded. The quick acceptance and rejection steps follow from inequalities obtained by Kinderman and Monahan (1979). The corresponding algorithm is known in the literature as algorithm TROU: one can show that the expected number of iterations is uniformly bounded over $a \ge 1$, and that it varies from $\frac{4}{\pi}$ at $a=1$ to $\frac{4}{\sqrt{\pi e}}$ as $a \to \infty$.

There are two important special cases. For the Cauchy density ($a=1$), the acceptance condition is $u^2 \le \frac{1}{1+x^2}$, or, put differently, $u^2+v^2 \le 1$. Thus, we obtain the result that if (U,V) is uniformly distributed in the unit circle, then $\frac{V}{U}$ is Cauchy distributed. Without squeeze steps, we have:

A Cauchy generator based upon the ratio-of-uniforms method

REPEAT

Generate iid uniform [-1,1] random variates U,V.

UNTIL $U^2+V^2 \le 1$

RETURN $X \leftarrow \frac{V}{U}$

For the t density with 3 degrees of freedom ($a=3$),

$$\frac{2}{\pi\sqrt{3}} \frac{1}{(1+\frac{x^2}{3})^2} ,$$

the acceptance condition is $\frac{x^2}{3} \le \frac{1}{u}-1$, or $v^2 \le 3u(1-u)$. Thus, once again, the acceptance region A is ellipsoidal. The unadorned ratio-of-uniforms algorithm is:

t3 generator based upon ratio-of-uniforms method

REPEAT
    Generate U uniformly on [0,1].
    Generate V uniformly on $[-\frac{\sqrt{3}}{2},\frac{\sqrt{3}}{2}]$.
UNTIL $V^2 \le 3U(1-U)$
RETURN $X \leftarrow \frac{V}{U}$

This is equivalent to

t3 generator based upon ratio-of-uniforms method

REPEAT
    Generate iid uniform [-1,1] random variates U,V.
UNTIL $U^2+V^2 \le 1$
RETURN $X \leftarrow \sqrt{3}\frac{V}{1+U}$

Both the Cauchy and $t\,3$ generators have obviously rejection constants of $\frac{4}{\pi}$, and should be accelerated by the judicious use of quick acceptance and rejection conditions that are linear in their arguments. ■

Example 7.4. The gamma density.

In this example, we consider the centered gamma (a) density with mode at the origin,

$$f(x) = c\frac{e^{a-1}}{(a-1)^{a-1}}(x+a-1)^{a-1}e^{-(x+a-1)} \quad (x+a-1 \geq 0) .$$

Here c is a normalization constant equal to $\frac{(a-1)^{a-1}}{e^{a-1}\Gamma(a)}$ which will be dropped.The table with facts is given below. Notice that the expected number of iterations is $\frac{4}{e}$ at $a=1$, and $\frac{4}{\sqrt{\pi e}}$ as $a \to \infty$, just as for the t density.

$f(x)$	$\frac{e^{a-1}}{(a-1)^{a-1}}(x+a-1)^{a-1}e^{-(x+a-1)} \quad (x+a-1 \geq 0)$
$b = \sup\sqrt{f(x)}$	1
$a_+ = \sup x\sqrt{f(x)}, a_- = \inf x\sqrt{f(x)}$	$z_+\sqrt{f(z_+)}$ where $z_+ = 1+\sqrt{2a-1}$, $z_-\sqrt{f(z_-)}$ where $z_- = 1-\sqrt{2a-1}$
area (A)	$a_+ - a_-$
Rejection constant	$2c(a_+ - a_-)$
Acceptance condition	$u \leq (\frac{e(x+a-1)}{a-1})^{\frac{a-1}{2}} e^{-\frac{x+a-1}{2}}$
	$2\log u + x \leq (a-1)\log(1+\frac{x}{a-1})$
Quick acceptance condition	$(x+a-1)^2(-2u^2+8u-6) \leq -x^2(2x+a-1) \quad (x \geq 0)$
	$(x+a-1)(-2u^2+8u-6) \leq -x^2 \quad (x \leq 0)$
Quick rejection condition	$(x+a-1)(2u^2-2) \geq -ux^2 \quad (x \geq 0)$
	$(a-1)(2u^2-2) \geq -ux^2 \quad (x \leq 0)$

We leave the verification of the inequalities implicit in the quick acceptance and rejection steps to the readers. All one needs here is Lemma 7.1. Timings with this algorithm have shown that good speeds are obtained for a greater than 5. The algorithm is uniformly fast for $a \in [1,\infty)$. The ratio-of-uniforms algorithms of Cheng and Feast (1979), Robertson and Walls (1980) and Kinderman and Monahan (1979) are different in conception. ■

7.3. Exercises.

1. For the quick acceptance and rejection conditions for Student's t distribution, the following inequality due to Kinderman and Monahan (1979) was used:

$$5-4(1+\frac{1}{a})^{\frac{a+1}{4}} u \leq a(u^{-\frac{4}{a+1}}-1) \leq -3+\frac{4(1+\frac{1}{a})^{-\frac{a+1}{4}}}{u} \quad (u \geq 0) .$$

The upper bound is only valid for $a \geq 3$. Show this. Hint: first show that the middle expression $g(u)$ is convex in u. Thus,

$$g(u) \geq g(z)+(u-z)g'(z) .$$

Here z is to be picked later. Show that the area under the quick acceptance curve is maximal when $z=(1+\frac{1}{a})^{-\frac{a+1}{4}}$, and substitute this value. For the lower bound, show that $g(u)$ as a function of $\frac{1}{u}$ is concave, and argue similarly.

2. Barbu (1982) has pointed out that when (U,V) is uniformly distributed in $A=\{(u,v):0\leq u \leq f(u+v)\}$, then $U+V$ has a density which is proportional to f. Similarly, if in the definition of A, we replace $f(u+v)$ by $(f(\frac{v}{\sqrt{u}}))^{\frac{2}{3}}$, then $\frac{V}{\sqrt{U}}$ has a density which is proportional to f. Show this.
3. Prove the following property. Let X have density f and define $Y=\sqrt{f(X)}\max(U_1,U_2)$ where U_1,U_2 are iid uniform [0,1] random variables. Define also $U=Y, V=XY$. Then (U,V) is uniformly distributed in $A=\{(u,v):0\leq u \leq \sqrt{f(\frac{v}{u})}\}$. Note that this can be useful for rejection in the (u,v) plane when rectangular rejection is not feasible.
4. In this exercise, we study sufficient conditions for convergence of performances. Assume that f_n is a sequence of densities converging in some sense to a density f as $n \to \infty$. Let b_n, a_{+n}, a_{-n} be the defining constants for the enclosing rectangles in the ratio-of-uniforms method. Let b, a_+, a_- be the constants for f. Show that the rejection constants converge,i.e.

$$\lim_{n \to \infty} b_n(a_{+n}-a_{-n}) = b(a_+-a_-)$$

when

$$\sup_x \left| \frac{f_n(x)}{f(x)}-1 \right| = o(1) ,$$

or when

$$\sup_x x^2 \left| f_n(x)-f(x) \right| = o(1) .$$

5. Give an example of a bounded density on $[0,\infty)$ for which the region A is unbounded in the v-direction, i.e. $b=\infty$.
6. Let f be a mixture of nonoverlapping uniform densities of varying widths and heights. Draw the region A.
7. From general principles (such as exercise 4), prove that the rejection constant for the t distribution tends to the rejection constant for the normal density as $a \to \infty$.

8. Prove that all the quick acceptance and rejection inequalities used for the gamma density are valid.