stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK
(DEPARTMENT OF MATHEMATICAL STATISTICS)

SN 9/79 DECEMBER

C. VAN PUTTEN & I. VAN DER TWEEL

ON GENERATING RANDOM VARIABLES

Statal Report 2

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

1980 Mathematics subject classification: Primary 62E25
Secondary 65C10,68J99

On generating random variables

by

C. van Putten & I. van der Tweel

ABSTRACT

Methods and algorithms are presented to generate random variables and (ordered) random samples by means of a random number generator. The random variables may be distributed according to several discrete and continuous probability distributions, including some multivariate distributions.

This report tries to give an easy accessible survey of procedures, many of which can be found in the literature. It also contains some procedures which do not seem to have been published before.

KEY WORDS & PHRASES: *Random numbers, Monte Carlo methods*

CONTENTS

1. INTRODUCTION 1

2. MATHEMATICAL PRELIMINARIES 4

3. DISCRETE DISTRIBUTIONS 6

3.1. THE DISCRETE UNIFORM DISTRIBUTION. 6

3.2. THE ALTERNATIVE DISTRIBUTION 7

3.3. THE LOGARITHMIC SERIES DISTRIBUTION. 7

3.4. THE GEOMETRIC DISTRIBUTION 8

3.5. THE NEGATIVE BINOMIAL DISTRIBUTION 9

3.6. THE BINOMIAL DISTRIBUTION. 10

3.7. THE POISSON DISTRIBUTION 11

3.8. THE HYPERGEOMETRIC DISTRIBUTION. 13

3.9. A RANDOM PERMUTATION 14

3.10. A RANDOM SAMPLE WITHOUT REPLACEMENT FROM A FINITE POPULATION . 15

3.11. AN ORDERED RANDOM SAMPLE WITHOUT REPLACEMENT FROM A FINITE POPULATION 16

3.12. AN ORDERED RANDOM SAMPLE WITH REPLACEMENT FROM A FINITE POPULATION 17

3.13. THE MULTIVARIATE HYPERGEOMETRIC DISTRIBUTION 18

3.14. THE MULTINOMIAL DISTRIBUTION 19

3.15. THE INFINITE DIMENSIONAL MULTINOMIAL DISTRIBUTION. 20

3.16. AN ORDERED RANDOM SAMPLE FROM A POISSON DISTRIBUTION 21

4. CONTINUOUS DISTRIBUTIONS 23

4.1. THE UNIFORM DISTRIBUTION 23

4.2. THE EXPONENTIAL DISTRIBUTION 24

4.3. THE GUMBEL DISTRIBUTION. 24

4.4. THE WEIBULL DISTRIBUTION 25

4.5. THE CAUCHY DISTRIBUTION. 26

4.6. THE LAPLACE DISTRIBUTION 26

4.7. THE LOGISTIC DISTRIBUTION. 27

4.8. THE NORMAL DISTRIBUTION. 28

4.9. THE BETA DISTRIBUTION. 29

4.10. THE GAMMA DISTRIBUTION 31

4.11. THE CHI-SQUARE DISTRIBUTION. 33

4.12. THE MULTIVARIATE NORMAL DISTRIBUTION 33
4.13. AN ORDERED RANDOM SAMPLE FROM A UNIFORM DISTRIBUTION 34
5. DISCUSSION ON POSSIBLE LIMITATIONS WHEN USING A PSEUDO RANDOM NUMBER GENERATOR 36

REFERENCES 39

1. INTRODUCTION

A random variable can be used to represent, in a mathematical model, an experiment rendering, each time it is performed, a (numerical) value, which cannot be predicated perfectly. Mathematically the simplest and most fundamental situation is a completely unpredictable experiment, the result of which belongs to a known finite set, A say, represented by a random variable having equal probability to obtain any of the elements of A, a so called discrete uniform random variable. If $A = \{k/2^n \mid k = 1,2,\ldots,2^n\}$ and n is large, this experiment may be used to approximate the one corresponding to a continuous uniform random variable with density equal to 1 on the interval $(0,1]$.

When generating realizations of random variables this situation is basic, for any random variable (with a known distribution function) can be obtained by a suitable transformation of such a uniform random variable. Thus to obtain a realization of any random variable it suffices to generate a realization of a uniform random variable and to apply the proper transformation to it.

In applications on a binary digital computer using an approximation of the kind just mentioned is the best one can do to generate (realizations of) random variables. In this case n might be chosen equal to the number of bits used to represent an integer number in the computer.

For the solution of the first part of the problem of generating random variables, i.e. the construction of a device producing realizations of a uniformly distributed random variable, a so called random number generator, the reader is referred to Statal Report 1 or any other description of such generators in the literature.

This report will only deal with the second part of the problem, i.e. the determination of suitable transformations. Theoretically a solution to this problem is provided by the following. Let $\underline{x}$ be a random variable having cumulative distribution function F. If $\underline{u}$ is uniformly distributed over the interval $(0,1]$, then $F^{-1}(\underline{u})$ and $\underline{x}$ have the same probability distribution, where F^{-1} may be defined as

$$F^{-1}(u) = \begin{cases} \inf \{x \mid x \in \mathbb{R},\ F(x) \geq u\} & \text{when } 0 < u \leq 1 \\ \infty & \text{when the set above is empty.} \end{cases}$$

Note, however, that the event $\{F^{-1}(\underline{u}) = \infty\}$ has probability zero. This proposition is proved in Subsection 2.1.

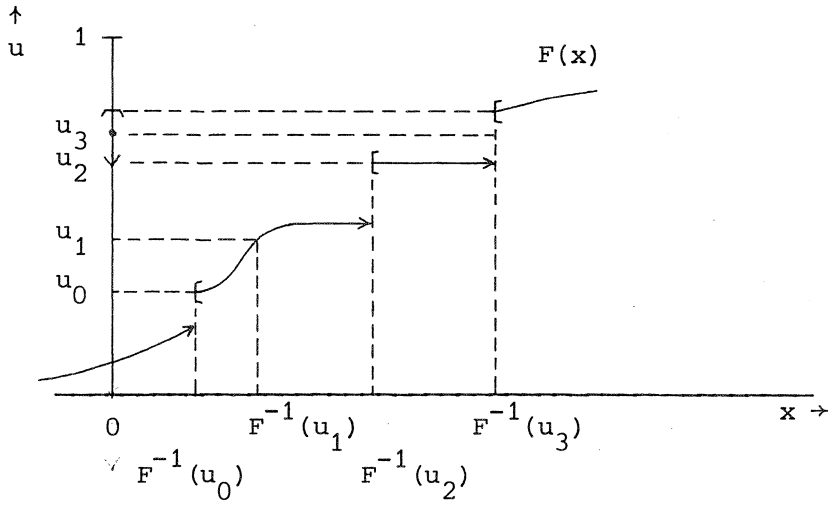

Fig.1 Illustration of F^{-1}

Thus if u is a realization of $\underline{u}$, obtained by a random number generator, then $F^{-1}(u)$ is a realization of $\underline{x}$. In cases where the computation of F^{-1} is simple, this may be a very useful and effective method, as for example when $\underline{x}$ has an exponential distribution. In other cases sometimes other methods may be more effective in the sense that they require less computing time. The underlying principles of two methods might be outlined as follows.

1. Let $(\underline{u}_1,\ldots,\underline{u}_n)$ be a random sample of $\underline{u}$, i.e. $\underline{u}_1,\ldots,\underline{u}_n$ are independent random variables, each having the same distribution as $\underline{u}$. Define a function ϕ, easy to compute, such that $\phi(\underline{u}_1,\ldots,\underline{u}_n)$ and $\underline{x}$ have the same distribution.
2. Look for a random variable $\underline{y}$ with cumulative distribution function G with the properties that G resembles F in a way to be made more precise and that G^{-1} can be computed easily. Define a subset S of $\mathbb{R}^2$, depending

on G and F, such that $\underline{y}$ and $\underline{x}$ have the same distribution if $(\underline{y},\underline{u}) \in S$, where $\underline{u}$ is uniformly distributed and independent of $\underline{y}$.

The second method is called the rejection method (cf. CHENG [4]), because a realization y of $\underline{y}$ is rejected if not $(y,u) \in S$, in which case one has to produce new realizations of $(\underline{y},\underline{u})$ until this condition is satisfied. All methods described so far are exact in that they generate exactly the distribution wanted. Apart from these there are methods that only approximate the right distribution. As an example a normally distributed random variable could be approximated by using the central limit theorem. However, since for every distribution considered in this report a satisfactory exact method is available no such method is used.

A random sample can be obtained by successively generating its elements independently of each other. A realization of a random sample without replacement from a finite population*) can also be generated stepwise but now each step consists of producing a realization of the conditional probability distribution given the results of the previous steps. If the ordering of the observations of a random sample is irrelevant, sometimes it is more efficient to create an ordered random sample*). An example of this is given at the end of Section 3.

Section 2 contains some mathematical prerequisites for the algorithms presented in the succeeding sections. In Section 3 discrete distributions are considered, the description of how to generate random variables having continuous distributions is given in Section 4. In Section 5 some restrictions are discussed that might be desirable when using the algorithms driven by a pseudo random number generator.

Primary the algorithms are meant for use on large computers, using high level programming languages, or on small programmable calculators, on which the natural logarithm, in the sequel denoted by log, the exponential function, sine, cosine, tangent, entier, square root, etc., are standard functions.

---

*) Strictly speaking this is no random sample, since a random sample of some random variable $\underline{x}$ is defined to be a random vector with independent components each having the same probability distribution as $\underline{x}$.

Some general references are: BURY [3] and JOHNSON & KOTZ [7,8] on the theory of probability distributions; HAMMERSLEY & HANDSCOMB [6], KNUTH [10] and, more recently, YAKOWITZ [12] on Monte Carlo methods.

2. MATHEMATICAL PRELIMINARIES

In this section we collect some theorems which will be useful in the following sections.

2.1. THE PROBABILITY DISTRIBUTION OF $F^{-1}(\underline{u})$

Let F be the cumulative distribution function of a random variable $\underline{x}$, let its inverse F^{-1} be defined on the unit interval by

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

and let $\underline{u}$ be uniformly distributed over the unit interval. Then $F^{-1}(\underline{u})$ and $\underline{x}$ have the same probability distribution.

PROOF. By the definition of F^{-1} we have $F(x) \geq u \Rightarrow F^{-1}(u) \leq x$, by the monotonicity of F $F^{-1}(u) < x \Rightarrow F(x) \geq u$ and by the right continuity of F $F^{-1}(u) = x \Rightarrow F(x) \geq u$. Thus $P\{F^{-1}(\underline{u}) \leq x\} = P\{\underline{u} \leq F(x)\} = F(x)$. □

REMARK. There are other ways of defining F^{-1} such that this proposition is true.

2.2. THE PROBABILITY DISTRIBUTION OF A TRANSFORMED CONTINUOUS RANDOM VARIABLE

Let $\underline{x}$ be a p-dimensional random variable with density f and let A be an open subset of $\mathbb{R}^p$ such that $P\{\underline{x} \in A\} = 1$. If $\phi = (\phi_1,\ldots,\phi_p)$ is a one-to-one transformation from A to $\mathbb{R}^p$ with continuous first partial derivatives in A such that the Jacobian $J_\phi = \det\left(\frac{\partial \phi_i}{\partial x_j}\right)$ does not vanish on A, then the density g of $\underline{y} = \phi(\underline{x})$ is given by

$$g(y) = f(\phi^{-1}(y))\,\mathrm{abs}(J_{\phi^{-1}}(y)) \quad \text{for every } y \in \phi(A).$$

PROOF. See BICKEL & DOKSUM [1]. □

2.3. THE REJECTION METHOD

Suppose the cumulative distribution function F has a density f and G is a cumulative distribution function with density g such that $\sup_{x\in A} \frac{f(x)}{g(x)} \le M < \infty$, where $A = \{x \mid x \in \mathbb{R},\ f(x) > 0\}$. Let $\underline{y}$ be a random variable with distribution function G and let $\underline{u}$ be a random variable uniformly distributed over (0,1] and independent of $\underline{y}$. Then we have

$$P\left\{\frac{1}{M}\frac{f(\underline{y})}{g(\underline{y})} \ge \underline{u}\right\} = \frac{1}{M} \quad \text{and} \quad P\left\{\underline{y} \le y \mid \frac{1}{M}\frac{f(\underline{y})}{g(\underline{y})} \ge \underline{u}\right\} = F(y)$$

for every $y \in \mathbb{R}$.

PROOF. We notice that $M \ge 1$, for

$$1 = \int_A f(x)dx = \int_A \frac{f(x)}{g(x)} \cdot g(x)dx \le M \int_A g(x)dx \le M.$$

Since $\underline{y}$ and $\underline{u}$ are independent we have

$$P\left\{\frac{1}{M}\frac{f(\underline{y})}{g(\underline{y})} \ge \underline{u}\right\} = \int_{\mathbb{R}} \left(\int_{(0,\frac{1}{M}\frac{f(y)}{g(y)}]} du \right) dG(y) =$$

$$= \int_{\mathbb{R}} \frac{1}{M}\frac{f(y)}{g(y)} g(y)dy = \frac{1}{M}.$$

For every $y_0 \in \mathbb{R}$ we derive in the same way

$$P\left\{\underline{y} \le y_0,\ \frac{1}{M}\frac{f(\underline{y})}{g(\underline{y})} \ge \underline{u}\right\} = \int_{(-\infty,y_0]} \frac{1}{M}\frac{f(y)}{g(y)} g(y)dy = \frac{1}{M} F(y_0)$$

and hence the result follows. □

REMARK. This theorem is a special case of the following result.

Let μ_F be the probability measure on $(\mathbb{R},\mathcal{B})$ induced by a distribution function F, where $\mathcal{B}$ are the Borel sets of $\mathbb{R}$, and suppose $\mu_F << \mu_G$ for some distribution function G with a bounded Radon-Nikodym derivative ϕ, say

$0 \le \phi(x) \le M < \infty$ for all $x \in \mathbb{R}$. If $\underline{y}$ and $\underline{u}$ are as above, we have

$$P\{\frac{1}{M}\,\phi(\underline{y}) \ge \underline{u}\} = \frac{1}{M} \quad \text{and} \quad P\{\underline{y} \le y \mid \frac{1}{M}\,\phi(\underline{y}) \ge \underline{u}\} = F(y)$$

for every $y \in \mathbb{R}$.

PROOF. A trivial modification of the previous proof. □

3. DISCRETE DISTRIBUTIONS

In this section procedures are given to generate discretely distributed random variables and samples from finite populations. Throughout the rest of this report it is assumed that a random number generator simulating a random variable $\underline{u}$, uniformly distributed over the (real) interval (0,1], is available. We also assume that it is possible to produce a sequence of independent random variables $\underline{u}_1,\ldots,\underline{u}_n$ each having the same distribution as $\underline{u}$, where n is an integer that may be chosen sufficiently large for our purposes. $u_1,\ldots,u_n$ will always denote realizations of $\underline{u}_1,\ldots,\underline{u}_n$.

The samples from finite populations in 3.9 - 3.13 are generated by a recursive procedure. If a sample of size k is wanted, the first element is generated by some method and the remaining problem is to generate a sample of size k-1, etc. To generate large samples from discretely distributed random variables, the Subsections 3.14 and 3.15 may be useful. As an example the latter is applied in Subsection 3.16 to generate an ordered sample from a Poisson distribution.

3.1. THE DISCRETE UNIFORM DISTRIBUTION

Let $\underline{x}$ be uniformly distributed over the set of integers $\{a,a+1,\ldots,b\}$ $(a \le b)$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = x\} = \begin{cases} \dfrac{1}{b-a+1} & \text{if } x \in \{a,a+1,\ldots,b\} \\ 0 & \text{otherwise.} \end{cases}$$

3.1.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = [(b-a+1)(1-\underline{u})] + a$$

where [y] denotes the *entier* of y, i.e. the greatest integer less than or equal to y.

3.1.2. Remark

When applying this method on a digital computer, sometimes not every value of b-a+1 will give satisfactory results because of the approximating nature of the device implemented to generate $\underline{u}$; for a discussion on this, the reader is referred to Section 5.

3.2. THE ALTERNATIVE DISTRIBUTION

Let $\underline{x}$ have an alternative distribution with parameter p, $0 \leq p \leq 1$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = 0\} = 1 - p \quad \text{and} \quad P\{\underline{x} = 1\} = p.$$

3.2.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \begin{cases} 1 & \text{if } 0 < \underline{u} \leq p \\ 0 & \text{if } p < \underline{u} \leq 1. \end{cases}$$

3.3. THE LOGARITHMIC SERIES DISTRIBUTION

Let $\underline{x}$ have a logarithmic series distribution with parameter θ, $0 < \theta < 1$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = x\} = \begin{cases} -\dfrac{\theta^x}{x \log(1-\theta)} & \text{if } x \in \{1,2,\ldots\} \\ \\ 0 & \text{otherwise.} \end{cases}$$

3.3.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \begin{cases} 1 & \text{if } (\underline{u}-1)\log(1-\theta) < \theta \\ y & \text{if } \sum_{i=1}^{y-1} \frac{\theta^i}{i} \le (\underline{u}-1)\log(1-\theta) < \sum_{i=1}^{y} \frac{\theta^i}{i} \text{ and } y \ge 2. \end{cases}$$

3.3.2. The algorithm

To compute a realization y of $\underline{y}$ we may proceed as follows.

1. Introduce auxiliary variables p,s and initialize p := s := $-\theta/\log(1-\theta)$.*)
2. For each i = 1,2,... while $1-u \ge s$ do (p := $p\theta/(1+1/i)$; s := s+p).
3. y := i. □

Note that the value assigned to y is the value of i for which $1-u < s$ is satisfied for the first time. This algorithm uses only one realization u of $\underline{u}$.

3.3.3. Remark.

$1 - \underline{u}$ has a uniform distribution over the interval [0,1), so it is impossible for $\underline{y}$ to be infinite.

3.4. THE GEOMETRIC DISTRIBUTION

Let $\underline{x}$ have a geometric distribution with parameter p, $0 < p \le 1$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = x\} = \begin{cases} p(1-p)^{x-1} & \text{if } x \in \{1,2,\ldots\} \text{ and } 0 < p < 1 \\ 1 & \text{if } x = 1 \text{ and } p = 1 \\ 0 & \text{otherwise.} \end{cases}$$

*) The symbol := means that the value of the right side is assigned to the variable on the left side.

3.4.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \begin{cases} \left[\frac{\log \underline{u}}{\log(1-p)}\right] + 1 & \text{if } 0 < p < 1 \\ 1 & \text{if } p = 1. \end{cases}$$

PROOF. Suppose $0 < p < 1$. Then we have for every $y \in \{1,2,\ldots\}$
$P\{\underline{y} = y\} = P\{y - 1 \leq \frac{\log \underline{u}}{\log(1-p)} < y\} = P\{(1-p)^y < \underline{u} \leq (1-p)^{y-1}\} =$
$= (1-p)^{y-1}(1 - (1-p))$. □

3.4.2. Remark.

The previous proof employs the fact that the geometric distribution is related to the exponential distribution.

3.5. THE NEGATIVE BINOMIAL DISTRIBUTION

Let $\underline{x}$ have a negative binomial distribution with parameters k and p, $k \in \{1,2,\ldots\}$, $0 < p \leq 1$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = x\} = \begin{cases} \binom{x-1}{k-1} p^k (1-p)^{x-k} & \text{if } x \in \{k,k+1,\ldots\} \text{ and } 0 < p < 1 \\ 1 & \text{if } x = k \text{ and } p = 1 \\ 0 & \text{otherwise.} \end{cases}$$

3.5.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \sum_{i=1}^{k} \underline{y}_i,$$

where $\underline{y}_1,\ldots,\underline{y}_k$ are independent random variables each having a geometric distribution with parameter p, e.g. generated by 3.4.

3.5.2. The algorithm

To compute a realization y of $\underline{y}$ we may proceed as follows if $0 < p < 1$.

1. Define $a := \log(1-p)$.
2. $y := k + \sum_{i=1}^{k} \left[\frac{\log u_i}{a}\right]$. □

$u_1,\ldots,u_k$ are realizations of k independent, uniformly distributed random variables $\underline{u}_1,\ldots,\underline{u}_k$.

3.6. THE BINOMIAL DISTRIBUTION

Let $\underline{x}$ have a binomial distribution with parameters n and p, $n \in \{1,2,\ldots\}$, $0 \le p \le 1$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = x\} = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{if } x \in \{0,1,\ldots,n\} \text{ and } 0 < p < 1 \\ 1 & \text{if } (x=0 \text{ and } p=0) \text{ or } (x=n \text{ and } p=1) \\ 0 & \text{otherwise.} \end{cases}$$

3.6.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \begin{cases} 0 & \text{if } p = 0 \text{ or } (0 < p \le \frac{1}{2} \text{ and } \underline{y}_1 > n) \\ \max\{i \mid \sum_{j=1}^{i} \underline{y}_j \le n\} & \text{if } 0 < p \le \frac{1}{2} \text{ and } \underline{y}_1 \le n \\ n - \max\{i \mid \sum_{j=1}^{i} \underline{y}_j \le n\} & \text{if } \frac{1}{2} < p < 1 \text{ and } \underline{y}_1 \le n \\ n & \text{if } p = 1 \text{ or } (\frac{1}{2} < p < 1 \text{ and } \underline{y}_1 > n), \end{cases}$$

where $\underline{y}_1, \underline{y}_2, \ldots$ are independent random variables each having a geometric distribution with parameter p if $0 < p \le \frac{1}{2}$ and $1 - p$ if $\frac{1}{2} < p < 1$.

PROOF. a) Suppose $0 < p \le \frac{1}{2}$, then we have for every $y \in \{1,2,\ldots,n\}$ $P\{\underline{y} \ge y\} = P\{\sum_{j=1}^{y} \underline{y}_j \le n\}$. $\sum_{j=1}^{y} \underline{y}_j$ has a negative binomial distribution with parameters y and p. Thus the result is obtained by using a well known relation of the negative binomial and the binomial distribution (see JOHNSON & KOTZ [7]).

b) Suppose $\frac{1}{2} < p < 1$. Arguing as above we see that $n - \underline{y}$ has a binomial distribution with parameters n and $1 - p$, thus $\underline{y}$ has a binomial distribution with parameters n and p. □

3.6.2. The algorithm

To compute a realization y of $\underline{y}$ we may proceed as follows if $0 < p < 1$.

1. Define a by if $p \leq \frac{1}{2}$ then $a := \log(1-p)$ else $a := \log p$.
2. Introduce the auxiliary variable s and initialize $s := n$.
3. For each $i = -1,0,1,2,\ldots$ while $s \geq 0$ do $s := s-1 - \left\lceil \frac{\log u_i}{a} \right\rceil$.
4. If $p \leq \frac{1}{2}$ then $y := i$ else $y := n - i$. □

$u_{-1}, u_0, \ldots$ are realizations of independent, uniformly distributed random variables $\underline{u}_{-1}, \underline{u}_0, \ldots$. After the third step i is supposed to have the value for which $s < 0$ for the first time.

3.6.3. Remarks

The distribution of $\underline{x}$ could also be obtained as the distribution of a sum of n independent random variables each having an alternative distribution with parameter p. This would involve no computations of the logarithm but more random numbers would have to be produced, especially when p is small. Which method is to be preferred depends on the properties of the computer used. On pocket calculators the method presented here might be favourable.

An other alternative is presented by the following algorithm, using the F^{-1} method.

1. Define a by if $p \leq \frac{1}{2}$ then $a := p/(1-p)$ else $a := 1/p - 1$.
2. Introduce the auxiliary variables s and t and initialize if $p \leq \frac{1}{2}$ then $t := (1-p)^n$ else $t := p^n$; $s := u - t$.
3. For each $i = 0,1,2,\ldots,n$ while $s > 0$ do $(t:=ta(n-i)/(i+1);s:=s-t)$.
4. If $p \leq \frac{1}{2}$ then $y := i$ else $y := n - i$. □

3.7. THE POISSON DISTRIBUTION

Let $\underline{x}$ have a Poisson distribution with parameter μ, $\mu > 0$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = x\} = \begin{cases} e^{-\mu}\mu^x/x! & \text{if } x \in \{0,1,\ldots\} \\ 0 & \text{otherwise.} \end{cases}$$

3.7.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ can be generated by the F^{-1} method according to

$$\underline{y} = \min\{x \mid P\{\underline{x} \leq x\} \geq \underline{u}\}.$$

3.7.2. The algorithm

To compute a realization y of $\underline{y}$ we may proceed as follows.

1. Introduce auxiliary variables s, t and w and initialize s := 0, t := 1, w := ue^{μ}.
2. For each i = 1,2,... while s < w do (t:=tμ/i;s:=s+t).
3. y := i - 1. □

u is a realization of a uniform distributed random variable $\underline{u}$. After the second step i is supposed to have the value for which s ≥ w for the first time.

3.7.3. Remarks

If more realizations are wanted, i.e. one wants to generate a realization of a random sample, and if enough memory is available, one can use a more efficient method, a so called *table search method*, which is a multiple F^{-1} method. The method is based on the property that the sum of independent random variables having a Poisson distribution with possibly different parameters, is again Poisson distributed with parameter equal to the sum of those parameters.

In principle, the table used, T say, has two entries, a line number i and an argument j, such that $T[i,j] = P\{\underline{v}_i \leq j\}$, where $\underline{v}_i$ has a Poisson distribution with parameter 2^i. One can considerably reduce the size of T by removing the tails of the distributions of the $\underline{v}_i$. Then, of course, one has to create an emergency procedure to restore them when they are needed. The method assumes that this table is computed in advance.

To generate a realization from a Poisson distribution with parameter μ, first μ is decomposed according to $\mu = \sum_i d_i 2^i + \mu_0$, where $d_i \in \{0,1\}$ and $0 \leq \mu_0 < 2$. Then for each i such that $d_i = 1$ a realization from a Poisson distribution with parameter 2^i is generated, using the i-th line of T. If $\mu_0 > 0$, then some other method, e.g. 3.7.1, is used for a realization from a Poisson distribution with parameter μ_0. Finally all realizations are added.

The following algorithm uses this method.

1. Compute $d_1,\ldots,d_m$ and μ_0 satisfying $\mu = \sum_{i=1}^{m} d_i 2^i + \mu_0$, such that

$d_1,\ldots,d_m \in \{0,1\}$ and $0 \le \mu_0 < 2$.

2. For each i = 1,2,...,m do step 3.
3. If $d_i = 1$ then (perform a line search to find y_i, the smallest integer v such that $P\{\underline{v}_i \le v\} \ge u_i$) else $y_i := 0$.
4. If $\mu_0 > 0$ then (generate a realization, y_0 say, from a Poisson distribution with parameter μ_0, for instance by applying algorithm 3.7.2) else $y_0 := 0$.
5. $y := \sum_{i=1}^{m} y_i d_i + y_0$. □

$u_1,\ldots,u_m$ are realizations of independent, uniformly distributed random variables $\underline{u}_1,\ldots,\underline{u}_m$. To do the line search in step 4, several methods can be used, such as a search from the left to the right or a search starting from the mode, $\hat{m}_i$ say, which has to be marked in advance in this case, and then going either to the left, if $P\{\underline{v}_i \le \hat{m}_i\} > u_i$, or the right, if $P\{\underline{v}_i \le \hat{m}_i\} < u_i$. m is the number of lines of T. T can be used if $\mu < 2^{m+1}$, otherwise T has to be extended or an other method has to be used.

3.8. THE HYPERGEOMETRIC DISTRIBUTION

Let $\underline{x}$ have a hypergeometric distribution with parameters n, r and N, $N \in \{1,2,\ldots\}$, $n,r \in \{0,1,\ldots,N\}$, i.e. the probability distribution of $\underline{x}$ is given by

$$P\{\underline{x} = x\} = \binom{n}{x}\binom{N-n}{r-x}\Big/\binom{N}{r} \quad \text{for every integer } x.$$

3.8.1. The method

Suppose a population Π of N elements is divided into two subsets, i.e. A, consisting of r elements and its complement B, thus consisting of N-r elements. Then the number of elements which belong to A, say $\underline{y}$, in a random sample without replacement of size n has the same distribution as $\underline{x}$. This sample may be obtained by successively selecting n elements of Π at random. When in this procedure i ($0 \le i \le n-1$) elements of Π have been selected and a (depending on i) of them have turned out to be an element of A, then the next element to be sampled will be one of A with probability $\frac{r-a}{N-i}$. Thus $\underline{y}$ is obtained as the sum of n *dependent* random variables with an alternative

distribution which in their turn may be considered to be functions of *independent* uniform random variables.

3.8.2. The algorithm

To compute a realization y of $\underline{y}$ we may proceed as follows if $1 \leq n \leq r \leq N-r \leq N-n$.

1. Introduce the auxiliary variable a and initialize a := 0.
2. For each i = N,N-1,...,N-n+1 perform the steps 3 and 4.
3. If i = r then (y:=a+n-N+i-1; goto step 6).
4. If $iu_i \leq r$ then (a:=a+1; if r=1 then (y:=a; goto step 6) else r:=r-1).
5. y := a.
6. Stop. □

By a := a+1 we mean that the value of a is increased by 1, etc.. $u_1,\ldots,u_n$ are realizations of independent uniform random variables as defined at the beginning of Section 3.

3.8.3. Remark

The condition $1 \leq n \leq r \leq N-r \leq N-n$ doesn't cause a loss of generality, because it can always be satisfied by interchanging the roles of n, r, N-r and N-n if necessary. It is imposed to increase the efficiency of the algorithm.

3.9. A RANDOM PERMUTATION

Let $(\underline{x}_1,\ldots,\underline{x}_n)$ be a random permutation of the set {1,2,...,n}, i.e. the values of $(\underline{x}_1,\ldots,\underline{x}_n)$ are the permutations of {1,2,...,n} and every permutation has the same probability $\frac{1}{n!}$.

3.9.1. The method

A random sample $(\underline{y}_1,\ldots,\underline{y}_n)$ with the same distribution as $(\underline{x}_1,\ldots,\underline{x}_n)$ is generated in the following way. First the value of $\underline{y}_n$ is generated. To do this let $\underline{i}_n$ be uniformly distributed over {1,2,...,n}, then $\underline{y}_n$ is set equal to $\underline{i}_n$ and $\underline{i}_n$ is deleted from {1,2,...,n}, the remaining set is denoted by Π_{n-1}. Thus the problem is reduced to generating a random permutation

$(\underline{y}_1,\ldots,\underline{y}_{n-1})$ from Π_{n-1}. Proceeding as above $\underline{y}_{n-1}$ is generated (equal to the $\underline{i}_{n-1}$-th element of Π_{n-1}), etc..

3.9.2. The algorithm

To compute a realization $(y_1,\ldots,y_n)$ of $(\underline{y}_1,\ldots,\underline{y}_n)$ we may proceed as follows.

1. Introduce auxiliary variables $a_1,\ldots,a_n$ and initialize for every $i \in \{1,2,\ldots,n\}$ $a_i := i$.
2. For each $k = n,n-1,\ldots,2$ perform the steps 3 and 4.
3. Generate a realization i_k of $\underline{i}_k$, uniformly distributed over $\{1,2,\ldots,k\}$, e.g. by 3.1.
4. If $i_k < k$ then exchange the values of a_{i_k} and a_k.
5. Finally the random permutation is given by $(y_1,\ldots,y_n) = (a_1,\ldots,a_n)$. □

3.9.3. Remark

Obviously the random permutation can also be generated in the reversed order, first $\underline{y}_1$, next $\underline{y}_2$, etc.; the way it is presented here gives a slightly more efficient algorithm.

3.10. A RANDOM SAMPLE WITHOUT REPLACEMENT FROM A FINITE POPULATION

Let $\{\underline{x}_1,\ldots,\underline{x}_n\}$ be a random sample without replacement from the set $\{1,2,\ldots,n\}$, $k \leq n$, i.e. the values of $\{\underline{x}_1,\ldots,\underline{x}_k\}$ are the subsets of size k of $\{1,2,\ldots,n\}$ and every subset has the same probability $\binom{n}{k}^{-1}$.

3.10.1. The method

A random sample $\{\underline{y}_1,\ldots,\underline{y}_k\}$ with the same distribution as $\{\underline{x}_1,\ldots,\underline{x}_k\}$ is generated in a way which is essentially the same as the one of 3.9; instead of n now only k elements are drawn from $\{1,2,\ldots,n\}$, thus the procedure of 3.9 is terminated after k steps.

3.10.2. The algorithm

To compute a realization $\{y_1,\ldots,y_k\}$ of $\{\underline{y}_1,\ldots,\underline{y}_k\}$ we may proceed as follows.

1. Introduce auxiliary variables $a_1,\ldots,a_n$ and initialize for every $i \in \{1,2,\ldots,n\}$ $a_i := i$.
2. For each $m = 1,2,\ldots,k$ perform the steps 3, 4 and 5.
3. Generate a realization i_m of $\underline{i}_m$, uniformly distributed over $\{1,2,\ldots,n-m+1\}$.
4. $y_m := a_{i_m}$.
5. If $i_m < n-m+1$ then $a_{i_m} := n-m+1$. □

3.11. AN ORDERED RANDOM SAMPLE WITHOUT REPLACEMENT FROM A FINITE POPULATION

Let $(\underline{x}_1,\ldots,\underline{x}_k)$ be an ordered random sample without replacement from the set $\{1,2,\ldots,n\}$, $k \leq n$, i.e. the values of $(\underline{x}_1,\ldots,\underline{x}_k)$ are the vectors $(x_1,\ldots,x_n)$ satisfying $1 \leq x_1 < x_2 < \ldots < x_k \leq n$ and every vector has the same probability $\binom{n}{k}^{-1}$.

3.11.1. The method

A random sample $(\underline{y}_1,\ldots,\underline{y}_k)$ with the same distribution as $(\underline{x}_1,\ldots,\underline{x}_k)$ is generated in the following way. First generate a random variable $\underline{a}_n$ with an alternative distribution with parameter k/n to decide whether or not n has to be included in the sample. If so, assign the value n to $\underline{y}_k$ and the problem left is to generate the same kind of sample of size $k-1$ from the set $\{1,2,\ldots,n-1\}$; if not, a sample of size k from the set $\{1,2,\ldots,n-1\}$ has to be generated. Thus a procedure is obtained, consisting of n steps or less if the contents of the rest of the sample is already known before the n-th step, rendering an ordered sample $(\underline{y}_1,\ldots,\underline{y}_k)$.

PROOF. See KAAS [9]. □

3.11.2. The algorithm

To compute a realization $(y_1,\ldots,y_k)$ of $(\underline{y}_1,\ldots,\underline{y}_k)$ we may proceed as follows.

1. For each $m = n,n-1,\ldots,1$ while $k \geq 1$ perform the step 2.
2. If $u_m \leq k/m$ then $(y_k := m; k := k-1)$. □

3.12. AN ORDERED RANDOM SAMPLE WITH REPLACEMENT FROM A FINITE POPULATION

Let $(\underline{x}_1,\ldots,\underline{x}_k)$ be a random sample with replacement from the set $\{1,2,\ldots,n\}$, i.e. $(\underline{x}_1,\ldots,\underline{x}_k)$ has a discrete uniform distribution over the set $\{1,2,\ldots,n\}^k$. If $\underline{\pi}$ is a permutation of $\{1,2,\ldots,k\}$ such that $\underline{x}_{\underline{\pi}(1)} \leq \underline{x}_{\underline{\pi}(2)} \leq \ldots \leq \underline{x}_{\underline{\pi}(k)}$, then $(\underline{x}_{\underline{\pi}(1)},\ldots,\underline{x}_{\underline{\pi}(k)})$ is called an ordered sample with replacement from $\{1,2,\ldots,n\}$.

3.12.1. The method

To generate a sample $(\underline{y}_1,\ldots,\underline{y}_k)$ with the same distribution as $(\underline{x}_{\underline{\pi}(1)},\ldots,\underline{x}_{\underline{\pi}(k)})$ we observe its close relation to a random vector $(\underline{z}_1,\ldots,\underline{z}_n)$ with a multinomial distribution. Indeed the correspondence is established by defining $\underline{z}_i$ to be the number of i's in $(\underline{x}_1,\ldots,\underline{x}_k)$. From $(\underline{z}_1,\ldots,\underline{z}_n)$ the $\underline{y}_1,\ldots,\underline{y}_k$ may be defined by $\underline{y}_i = m$ if $\underline{z}_0 + \ldots + \underline{z}_{m-1} < i \leq \underline{z}_0 + \ldots + \underline{z}_{m-1} + \underline{z}_m$, for every $i \in \{1,2,\ldots,k\}$, where $\underline{z}_0 = 0$, introduced for convenience. Thus it suffices to generate $\underline{z}_1,\ldots,\underline{z}_n$. The probability distribution of $\underline{z}_n$ is binomial with parameters k and 1/n and $(\underline{z}_1,\ldots,\underline{z}_{n-1})$ given $\underline{z}_n$ again has a multinomial distribution. So after having generated $\underline{z}_n$, e.g. by 3.6, the problem is reduced to the same one with k and n replaced by $k-\underline{z}_n$ and n-1, respectively.

PROOF. See KAAS [9]. □

3.12.2. The algorithm

To compute a realization $(y_1,\ldots,y_k)$ of $(\underline{y}_1,\ldots,\underline{y}_k)$ we may proceed as follows.

1. For each m = n,n-1,...,1 while $k \geq 1$ perform the steps 2, 3 and 4.
2. Generate a realization z_m of $\underline{z}_m$ having a binomial distribution with parameters k and 1/m.
3. For each j satisfying $k-z_m+1 \leq j \leq k$ do $y_j := m$.
4. $k := k - z_m$. □

3.13. THE MULTIVARIATE HYPERGEOMETRIC DISTRIBUTION

Let $(\underline{x}_1,\ldots,\underline{x}_k)$ have a multivariate hypergeometric distribution with parameters n, $r_1,\ldots,r_k$ and N, $N \in \{k,k+1,\ldots\}$, $n,r_1,\ldots,r_k \in \{1,2,\ldots,N-1\}$, $\sum_{i=1}^{k} r_i = N$, i.e. the probability distribution of $(\underline{x}_1,\ldots,\underline{x}_n)$ is given by

$$P\{(\underline{x}_1,\ldots,\underline{x}_k) = (x_1,\ldots,x_k)\} =$$

$$= \begin{cases} \left[\prod_{i=1}^{k} \binom{r_i}{x_i}\right] \Big/ \binom{N}{n} & \text{if } x_i \in \{0,1,\ldots\} \text{ for } 1 \le i \le k \text{ and } \sum_{i=1}^{k} x_i = n \\ 0 & \text{otherwise.} \end{cases}$$

3.13.1. The method

Suppose a population Π of N elements is divided into k disjoint subsets $A_1,\ldots,A_k$ consisting of $r_1,\ldots,r_k$ elements respectively, $\sum_{i=1}^{k} r_i = N$. If in a random sample without replacement $\underline{y}_i$ is the number of elements belonging to A_i, for every $i \in \{1,2,\ldots,k\}$, then the random vector $(\underline{y}_1,\ldots,\underline{y}_k)$ has the same distribution as $(\underline{x}_1,\ldots,\underline{x}_k)$. This sample may be obtained by successively selecting n elements of Π at random. When in this procedure j $(0 \le j \le n-1)$ elements of Π have been selected and $a_1,\ldots,a_k$, depending on j, $(\sum_{i=1}^{k} a_i = j)$, belong to $A_1,\ldots,A_k$ respectively, then the next element to be sampled will be one of A_i with probability $(r_i-a_i)/(N-j)$, $i \in \{1,2,\ldots,k\}$.

The following algorithm is an immediate generalization of that of Subsection 3.8.

3.13.2. The algorithm

To compute a realization $(y_1,\ldots,y_k)$ of $(\underline{y}_1,\ldots,\underline{y}_k)$ we may proceed as follows.

1. Introduce auxiliary variables $a_1,\ldots,a_k$, s and t and initialize for every $i \in \{1,2,\ldots,k\}$ $a_i := 0$.
2. For each $j = N,N-1,\ldots,N-n+1$ perform the steps 3, 4 and 5.
3. $s := 0$, $t := j\,u_j$.
4. For each $i = 1,2,\ldots$ while $s < t$ do $s := s+r_i$.

5. $i := i-1$; $r_i := r_i-1$; $a_i := a_i+1$.
6. For each i, $1 \le i \le k$, assign the value of a_i to y_i. □

We remark that after the execution of step 4 $i-1$ is supposed to be equal to the smallest integer m such that $\sum_{\ell=1}^{m} r_\ell \ge t$.

3.14. THE MULTINOMIAL DISTRIBUTION

Let $(\underline{x}_1,\ldots,\underline{x}_k)$ have a multinomial distribution with parameters $n, p_1,\ldots,p_k$, $n \in \{1,2,\ldots\}$, $p_i > 0$ for $1 \le i \le k$, $\sum_{i=1}^{k} p_i = 1$, i.e. the probability distribution of $(\underline{x}_1,\ldots,\underline{x}_k)$ is given by

$$P\{(\underline{x}_1,\ldots,\underline{x}_k) = (x_1,\ldots,x_k)\} =$$

$$= \begin{cases} n! \prod_{i=1}^{k} \frac{p_i^{x_i}}{x_i!} & \text{if } x_i \in \{0,1,\ldots\} \text{ for } 1 \le i \le k \text{ and } \sum_{i=1}^{k} x_i = n \\ 0 & \text{otherwise.} \end{cases}$$

3.14.1. The method

To generate a random vector $(\underline{y}_1,\ldots,\underline{y}_k)$ with the same distribution as $(\underline{x}_1,\ldots,\underline{x}_k)$, we use an ordered sample from a uniform distribution $(\underline{v}_1,\ldots,\underline{v}_n)$, generated, e.g., by the method of Subsection 4.13. If $s_i = \sum_{j=1}^{i} p_j$ for every (integer) i, $1 \le i \le k$, we define $\underline{y}_i$ to be the number of observations $\underline{v}$ in the interval $(s_{i-1}, s_i]$, with $s_0 = 0$, i.e. for every i satisfying $1 \le i \le k$ we have

$$\underline{y}_i = \#\{j \mid s_{i-1} < \underline{v}_j \le s_i\}.$$

An *ordered* sample is used to obtain a more efficient algorithm.

3.14.2. The algorithm

To compute a realization $(y_1,\ldots,y_k)$ of $(\underline{y}_1,\ldots,\underline{y}_k)$ we may proceed as follows.

1. Introduce auxiliary variables t, s, r, i and initialize $t := 0$, $s := 1-p_k$,

$r := \log s$, $i := k$.

2. Initialize for every j, $2 \le j \le k$, $y_j = 0$.
3. For each $m = n, n-1, \ldots, 1$ perform the steps 4 and 5 as long as step 6 is not performed.
4. $t := t + (\log u_m)/m$.
5. If $t \le r$ then [if $i=2$ then ($y_1 = m-1$; goto step 6) else ($i := i-1$; $s := s-p_i$; $r := \log s$; goto step 5)] else $y_i := y_i + 1$.
6. Stop. □

3.15. THE INFINITE DIMENSIONAL MULTINOMIAL DISTRIBUTION

Let $(\underline{x}_1, \underline{x}_2, \ldots)$ have an infinite dimensional multinomial distribution with parameters n, $p_1, p_2, \ldots$, $n \in \{1,2,\ldots\}$, $p_i > 0$ for $i \ge 1$, $\sum_{i=1}^{\infty} p_i = 1$, i.e. the probability distribution of $(\underline{x}_1, \underline{x}_2, \ldots)$ is given by

$$P\{(\underline{x}_1, \underline{x}_2, \ldots) = (x_1, x_2, \ldots)\} =$$

$$= \begin{cases} n! \prod_{i=1}^{\infty} \frac{p_i^{x_i}}{x_i!} & \text{if } x_i \in \{0,1,\ldots\} \text{ for } i \ge 1 \text{ and } \sum_{i=1}^{\infty} x_i = n \\ 0 & \text{otherwise.} \end{cases}$$

3.15.1. The method

To generate a random sequence $(\underline{y}_1, \underline{y}_2, \ldots)$ with the same distribution as $(\underline{x}_1, \underline{x}_2, \ldots)$, we use the method of Subsection 3.14. Like in that section an ordered sample from a uniform distribution, $(\underline{v}_1, \ldots, \underline{v}_n)$, is generated. $\underline{v}_n$ determines the random integer $\underline{k}$ such that $\underline{y}_{\underline{k}} > 0$ and $\underline{y}_i = 0$ if $i > \underline{k}$. Using $\underline{k}$, the algorithm 3.14.2 can be applied.

3.15.2. The algorithm

To compute a realization $(y_1, y_2, \ldots)$ of $(\underline{y}_1, \underline{y}_2, \ldots)$ we may proceed as follows.

1. Introduce auxiliary variables $r_2, r_3, \ldots$, s, t and v and initialize $s := p_1$, $t := (\log u_n)/n$, $v := e^t$.
2. If $s \ge v$ then ($y_1 := n$; $k := 1$; goto step 9).
3. For each $i = 2,3,\ldots$ while $s < v$ do ($r_i := \log s$; $s := s+p_i$).

4. i := i-1; y_i := 1; k := i.
5. Initialize for every j, 1 ≤ j ≤ k-1, y_j := 0.
6. For each m = n-1,n-2,...,1 perform the steps 7 and 8.
7. t := t + (log u_m)/m.
8. If t > r_i then y_i := y_i+1 else[if i=2 then (y_1:=m; goto step 9) else (i:=i-1; goto step 8)].
9. Stop. □

3.15.3. Remarks

Let $\{w_1,w_2,...\}$ be the set of values taken by a random variable $\underline{w}$, then the previous procedure may be used to generate the numbers of occurrences of $w_1,w_2,...$ in a random sample of $\underline{w}$ with sample size n, by defining $p_i = P\{\underline{w} = w_i\}$ for i = 1,2,... .

In step 3 the number of variables r_i needed is unknown, a priori. If this should cause any difficulties when implementing the algorithm, one might use the following modification of it.

1*. Introduce auxiliary variables r, s, t and v and initialize s := p_1, t := (log u_n)/n, v := e^t.
2. If s ≥ v then (y_1:=n;k:=1; goto step 9).
3*. For each i = 2,3,... while s < v do s := s + p_i.
4*. i := i-1; y_i := 1; k := i; s := s-p_i; r := log s.
5. Initialize for every j, 1 ≤ j ≤ k-1, y_j := 0.
6. For each m = n-1,n-2,...,1 perform the steps 7 and 8.
7. t := t + (log u_m)/m.
8*. If t > r then y_i := Y_i+1 else[if i=2 then (y_1:=m; goto step 9) else (i:=i-1;s:=s-p_i;r:=log s; goto step 8)].
9. Stop. □

3.16. AN ORDERED SAMPLE FROM A POISSON DISTRIBUTION

Let $\underline{x}_1,...,\underline{x}_k$ be independent random variables, each having a Poisson distribution with parameter μ, μ > 0. If $\underline{\pi}$ is a permutation of {1,2,...,k} such that $\underline{x}_{\underline{\pi}(1)} \leq \underline{x}_{\underline{\pi}(2)} \leq ... \leq \underline{x}_{\underline{\pi}(k)}$, then $(\underline{x}_{\underline{\pi}(1)},...,\underline{x}_{\underline{\pi}(k)})$ is called an ordered sample from a Poisson distribution.

3.16.1. The method

To generate a sample $(\underline{y}_1,\ldots,\underline{y}_k)$ with the same distribution as $(\underline{x}_{\underline{\pi}(1)},\ldots,\underline{x}_{\underline{\pi}(k)})$, we use the modified algorithm of 3.15.

3.16.2. The algorithm

To compute a realization $(y_1,\ldots,y_k)$ of $(\underline{y}_1,\ldots,\underline{y}_k)$ we may proceed as follows.

1. Introduce auxiliary variables r, s, t, v, and w and initialize s := 1, $t := (\log u_k)/k$, $v := e^t$, w := 0.
2. For each i = 1,2,... while $s/e^{\mu} < v$ do (w:=w+log i;s:=s+exp(i log μ-w)).
3. If i = 1 then (y_1:=0,...,y_k:=0); goto step 8).
4. i := i-1; $y_k := i$; s := s - exp(i log μ-w); w := w - log i; r := log s - μ.
5. For each m = k-1,k-2,...,1 perform the steps 6 and 7.
6. $t := t + (\log u_m)/m$.
7. If t > r then $y_m := i$ else[if i=1 then (y_1:=0,...,y_m:=0; goto step 8) else (i:=i-1;s:=s-exp(i log μ-w);w:=w-log i; r:=log s-μ; goto step 7)].
8. Stop. □

3.16.3. Remark

The algorithm can be stated also in the following form, which might be slightly faster if k is large, compared to $\sqrt{\mu}$.

1. Introduce auxiliary variables m, r, s, t, v, and w and initialize m := k, s := 1, $t := (\log u_k)/k$, $v := e^t$, w := 0.
2. For each i = 1,2,... while $s/e^{\mu} < v$ do (w:=w+log i;s:=s+exp(i log μ-w)).
3. For each ℓ = i-1,i-2,...,1 perform the steps 4 and 5.
4. s := s - exp(ℓ log μ-w); w := w - log ℓ; r := log s-μ.
5. If t > r then (y_m:=ℓ; if m=1 then goto step 7; m:=m-1;t:=t+(log u_m)/m; goto step 5).
6. $y_1 := 0,\ldots,y_m := 0$.
7. Stop. □

Step 3 is supposed to be a dummy statement if i = 1.

4. CONTINUOUS DISTRIBUTIONS

In this section procedures are given to generate random variables with continuous probability distributions and ordered samples of some of them. If an inverse F^{-1} of the cumulative distribution function F can be computed easily, the corresponding random variable is generated by $F^{-1}(\underline{u})$. This is done in the Subsections 4.1 - 4.7. Sometimes a simple real valued function ϕ exists such that $\phi(\underline{u}_1,\ldots,\underline{u}_n)$ has the probability distribution wanted. An example of this can be found in 4.8. If both methods fail or are not satisfactory, one may try to find an auxiliary probability distribution slightly differing from the original one and such that one of the methods just mentioned can be applied to it. A random variable generated to have this distribution is transformed by only *accepting* it if it satisfies a certain condition, under which its conditional probability distribution is equal to the one wanted. Informally one might say that one chooses an "easy" distribution in the "neighbourhood" and then makes a correction to it by conditioning. In order to be an efficient procedure the probability that the condition of this rejection method is satisfied has to be close to 1. This is the reason why one should stay in the "neighbourhood". The rejection method is applied in 4.9 - 4.11. For mathematical details the reader is referred to Subsection 2.3.

To generate ordered samples from continuous distributions we use the fact that an ordered sample from an exponential distribution has spacings, i.e. distances between successive elements, which are independent and exponentially distributed with known parameters.

4.1. THE UNIFORM DISTRIBUTION

Let $\underline{x}$ have a uniform distribution with parameters a and b, $a < b$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \begin{cases} 1 & \text{if } a < x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

4.1.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = (b-a)\underline{u} + a.$$

4.2. THE EXPONENTIAL DISTRIBUTION

Let $\underline{x}$ have an exponential distribution with parameter λ, $\lambda > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

4.2.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = -\frac{1}{\lambda} \log \underline{u}.$$

PROOF. $P\{\underline{y} \leq y\} = P\{\underline{u} \geq e^{-\lambda y}\} = 1 - e^{-\lambda y}$ for every $y \geq 0$. □

4.2.2. Remark

The cumulative distribution function F is given by $F(x) = 1 - e^{-\lambda x}$ if $x \geq 0$, so its inverse is $F^{-1}(u) = -\frac{1}{\lambda} \log(1-u)$ if $0 \leq u < 1$. By observing that $\underline{u}$ and $1 - \underline{u}$ have the same probability distribution, we see that a minor modification of the F^{-1} method leads to $\underline{y}$ defined above.

4.3. THE GUMBEL DISTRIBUTION

Let $\underline{x}$ have a Gumbel distribution with parameters μ and σ, $\sigma > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \frac{1}{\sigma} \exp\left\{-\frac{x-\mu}{\sigma} - e^{-\frac{x-\mu}{\sigma}}\right\} \quad \text{for every } x \in \mathbb{R}.$$

4.3.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \mu - \sigma \log(-\log \underline{u}) \quad \text{if } \underline{u} < 1,$$

by which we mean that from a sequence $\underline{u}_1, \underline{u}_2, \ldots$ of independent uniform random variables the first $\underline{u}_i$ is satisfying $\underline{u}_i < 1$ is used to compute $\underline{y}$.

PROOF. The cumulative distribution function F of $\underline{x}$ is given by $F(x) = \exp\{-e^{-\frac{x-\mu}{\sigma}}\}$ for every real x and hence its inverse is $F^{-1}(u) = \mu - \sigma \log(-\log u)$ if $0 < u < 1$. The result follows by observing that the first $\underline{u}_i$ satisfying $\underline{u}_i < 1$ has a uniform distribution over the interval (0,1). □

4.4. THE WEIBULL DISTRIBUTION

Let $\underline{x}$ have a Weibull distribution with parameters μ, σ and α, $\sigma > 0$, $\alpha > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \begin{cases} \frac{\alpha}{\sigma}\left(\frac{x-\mu}{\sigma}\right)^{\alpha-1} \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{\alpha}\right\} & \text{if } x > \mu \\ 0 & \text{otherwise.} \end{cases}$$

4.4.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \mu + \sigma(-\log \underline{u})^{1/\alpha}.$$

PROOF. Imitate the proof of 4.2. □

4.5. THE CAUCHY DISTRIBUTION

Let $\underline{x}$ have a Cauchy distribution with parameters μ and σ, $\sigma > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \frac{1}{\sigma\pi} \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{for every } x \in \mathbb{R}.$$

4.5.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \mu + \sigma \tan((\underline{u}-\tfrac{1}{2})\pi) \quad \text{if } \underline{u} < 1.$$

An explanation of how to handle this condition is given in 4.3.1.

PROOF. For every real x we have $F(x) = \frac{1}{\pi} \operatorname{artan}\left(\frac{x-\mu}{\sigma}\right) + \frac{1}{2}$ and hence $F^{-1}(u) = \mu + \sigma \tan((u-\frac{1}{2})\pi)$ if $0 < u < 1$. □

4.6. THE LAPLACE DISTRIBUTION

Let $\underline{x}$ have a Laplace distribution with parameters μ and σ, $\sigma > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \frac{1}{2\sigma} \exp\left\{-\frac{|x-\mu|}{\sigma}\right\} \quad \text{for every } x \in \mathbb{R}.$$

4.6.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \begin{cases} \mu + \sigma \log(2\underline{u}) & \text{if } \underline{u} \leq \frac{1}{2} \\ \mu - \sigma \log(2(1-\underline{u})) & \text{if } \frac{1}{2} < \underline{u} < 1. \end{cases}$$

PROOF. The cumulative distribution function of $\underline{x}$ is given by $F(x) = \frac{1}{2} \exp\{-\frac{|x-\mu|}{\sigma}\}$ if $x \leq \mu$ and $F(x) = 1 - \frac{1}{2} \exp\{-\frac{|x-\mu|}{\sigma}\}$ if $x > \mu$ and hence $F^{-1}(u) = \mu + \sigma \log(2u)$ if $u \leq \frac{1}{2}$ and $F^{-1}(u) = \mu - \sigma \log(2(1-u))$ if $\frac{1}{2} < u < 1$. □

4.6.2. Remarks

Two alternative procedures are the following.

1. Generate $\underline{y}_1$ according to

$$\underline{y}_1 = \mu + \sigma\phi(\underline{u}_1 - \tfrac{1}{2})\log \underline{u}_2,$$

where, as stated at the beginning of Section 3, $\underline{u}_1$ and $\underline{u}_2$ are independent random variables, both having a uniform distribution over the interval $(0,1]$ and the function ϕ is defined by $\phi(a) = 1$ if $a \leq 0$ and $\phi(a) = -1$ if $a > 0$.

PROOF. For every $y_1 \leq \mu$ we have $P\{\underline{y}_1 \leq y_1\} = P\{\underline{u}_1 \leq \frac{1}{2}, \underline{u}_2 \leq \exp\{(y_1-\mu)/\sigma\} = \frac{1}{2}\exp\{-|y_1-\mu|/\sigma\}$; for every $y_1 > \mu$ we have $P\{\underline{y}_1 \leq y_1\} = P\{\underline{y}_1 \leq \mu\} + P\{\mu < \underline{y}_1 \leq y_1\} = P\{\underline{u}_1 \leq \frac{1}{2}\} + P\{\underline{u}_1 > \frac{1}{2}, \underline{u}_2 \geq \exp\{-(y_1-\mu)/\sigma\} = \frac{1}{2} + \frac{1}{2}(1 - \exp\{-|y_1-\mu|/\sigma\})$. □

2. Generate $\underline{y}_2$ according to

$$\underline{y}_2 = \mu + \sigma \log(\underline{u}_1/\underline{u}_2).$$

PROOF. For every $y_2 \leq \mu$ we have $P\{\underline{y}_2 \leq y_2\} = P\{\underline{u}_1 \leq \underline{u}_2 \exp\{-\frac{|y_2-\mu|}{\sigma}\}\} = \frac{1}{2}\exp\{-\frac{|y_2-\mu|}{\sigma}\}$. The last equality is easily obtained by using 2.3. For every $y_2 > \mu$ we have $P\{\underline{y}_2 \leq y_2\} = 1 - P\{\underline{y}_2 > y_2\} = 1 - P\{\underline{u}_2 < \underline{u}_1 \exp\{-\frac{|y_2-\mu|}{\sigma}\} = 1 - \frac{1}{2}\exp\{-|y_2-\mu|/\sigma\}$. □

4.7. THE LOGISTIC DISTRIBUTION

Let $\underline{x}$ have a logistic distribution with parameters μ and σ, $\sigma > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right)\Big/\left[1 + \exp\left(-\frac{x-\mu}{\sigma}\right)\right]^2 \quad \text{for every } x \in \mathbb{R}.$$

4.7.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = \mu - \sigma \log(1/\underline{u} - 1) \quad \text{if } \underline{u} < 1.$$

PROOF. For every real x we have $F(x) = [1 + \exp(-\frac{x-\mu}{\sigma})]^{-1}$. Thus $\underline{y}$ is obtained by computing $F^{-1}(\underline{u})$ if $\underline{u} < 1$. An explanation of the condition $\underline{u} < 1$ is given in 4.3.1. □

4.8. THE NORMAL DISTRIBUTION

Let $\underline{x}$ have a normal distribution with parameters μ and σ, $\sigma > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad \text{for every } x \in \mathbb{R}.$$

4.8.1. The method

Two independent random variables $\underline{y}_1$ and $\underline{y}_2$ each having the same distribution as $\underline{x}$ are simultaneously generated according to

$$\begin{cases} \underline{y}_1 = \mu + \sigma\sqrt{-2\log \underline{u}_1}\, \cos(2\pi \underline{u}_2) \\ \underline{y}_2 = \mu + \sigma\sqrt{-2\log \underline{u}_1}\, \sin(2\pi \underline{u}_2). \end{cases}$$

PROOF. From the defining relations of $\underline{y}_1$ and $\underline{y}_2$ it follows that

$$\underline{u}_1 = \exp\left\{-\tfrac{1}{2}\left(\frac{\underline{y}_1-\mu}{\sigma}\right)^2 - \tfrac{1}{2}\left(\frac{\underline{y}_2-\mu}{\sigma}\right)^2\right\}$$

and

$$\underline{u}_2 = \frac{1}{2\pi} \operatorname{artan} \frac{\underline{y}_2-\mu}{\underline{y}_1-\mu} + \theta(\underline{y}_1 - \mu, \underline{y}_2 - \mu),$$

where θ is defined by $\theta(a,b) = 0$ if $a > 0$ and $b > 0$, $\theta(a,b) = \frac{1}{2}$ if $a < 0$, $\theta(a,b) = 1$ if $a > 0$ and $b < 0$. Hence the simultaneous probability density g of $\underline{y}_1$ and $\underline{y}_2$ can be computed by 2.2. For every $0 < u_1 < 1$, $0 < u_2 \leq 1$ and corresponding values of y_1, y_2 we have

$$g(y_1,y_2)\ \operatorname{abs}\left|\frac{\partial(y_1,y_2)}{\partial(u_1,u_2)}\right| = 1$$

and

$$\left|\frac{\partial(y_1,y_2)}{\partial(u_1,u_2)}\right| = \begin{vmatrix} -\dfrac{\sigma\cos(2\pi u_2)}{u_1\sqrt{-2\log u_1}} & -\dfrac{\sigma\sin(2\pi u_2)}{u_1\sqrt{-2\log u_1}} \\ -2\pi\sigma\sqrt{-2\log u_1}\sin(2\pi u_2) & 2\pi\sigma\sqrt{-2\log u_1}\cos(2\pi u_2) \end{vmatrix}$$

$$\Rightarrow g(y_1,y_2) = \frac{u_1}{\sigma^2 2\pi} = f(y_1)f(y_2). \quad \square$$

4.8.2. Remark

$\sqrt{-2\log \underline{u}_1}$, $2\pi\underline{u}_2$ can be considered to be the polar coordinates of the rectangular coordinates $(\underline{y}_1-\mu)/\sigma$, $(\underline{y}_2-\mu)/\sigma$. This may be used in an algorithm if a transformation from polar to rectangular coordinates is available on the computer or calculator to be used.

4.9. THE BETA DISTRIBUTION

Let $\underline{x}$ have a beta distribution with parameters α and β, $\alpha > 0$, $\beta > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise,} \end{cases}$$

where Γ is the gamma function, $\Gamma(\alpha) = \int_0^\infty v^{\alpha-1}e^{-v}dv$.

4.9.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated by applying a rejection method in the following way. Reject $\underline{w} = \underline{u}_1^{1/\alpha} + \underline{u}_2^{1/\beta}$ as long as $\underline{w} > 1$. If $\underline{w} \le 1$ then $\underline{y}$ is determined by $\underline{y} = \underline{u}_1^{1/\alpha}/(\underline{u}_1^{1/\alpha}+\underline{u}_2^{1/\beta})$. This means that from a sequence $\underline{u}_1,\underline{u}_2,\ldots$, the first pair $(\underline{u}_{2i-1},\underline{u}_{2i})$ of which the corresponding $\underline{w}$ satisfies $\underline{w} \le 1$ is used to generate $\underline{y}$.

PROOF. For every $0 < w_1 \le 1$ we have $P\{\underline{u}_1^{1/\alpha} \le w_1\} = w_1^\alpha$. Define $\underline{w}_1 = \underline{u}_1^{1/\alpha}$ and $\underline{w}_2 = \underline{u}_2^{1/\beta}$, then the cumulative distribution function G of $\underline{y}$ is, if $0 < y < 1$

$G(y) = P\{\underline{w}_1/(\underline{w}_1+\underline{w}_2) \le y \mid \underline{w}_1 + \underline{w}_2 \le 1\} = P\{\underline{w}_1(1-y) \le y\,\underline{w}_2,\ \underline{w}_1 + \underline{w}_2 \le 1\}/$
$/P\{\underline{w}_1 + \underline{w}_2 \le 1\} = CP\{\underline{w}_1 \le \min(\frac{y}{1-y}\underline{w}_2, 1-\underline{w}_2)\}$, where C is a constant, i.e.

independent of y. Therefore

$$G(y) = C\int_0^1 \left\{\min\left(\frac{y}{1-y}w_2, 1-w_2\right)\right\}^{\alpha} \beta w_2^{\beta-1} dw_2, \quad \frac{y}{1-y}w_2 \le 1-w_2 \iff w_2 \le 1-y,$$

$$\Rightarrow$$

$$G(y) = C\int_0^{1-y} \left(\frac{y}{1-y}w_2\right)^{\alpha} \beta w_2^{\beta-1} dw_2 + C\int_{1-y}^1 (1-w_2)^{\alpha}\beta w_2^{\beta-1} dw_2.$$

From this we see that G is differentiable at y, so the density of G may be computed as follows.

$$g(y) = \frac{dG(y)}{dy} = \frac{C\beta}{\alpha+\beta}\frac{d}{dy}(y^{\alpha}(1-y)^{\beta}) + C\beta y^{\alpha}(1-y)^{\beta-1}$$

$$= \frac{\alpha\beta}{\alpha+\beta} C y^{\alpha-1}(1-y)^{\beta-1}.$$

Thus $\underline{x}$ and $\underline{y}$ have the same probability density. □

4.9.2. Remarks

When using a rejection method, like above, the natural thing to ask is how many pairs $(\underline{u}_{2i-1}, \underline{u}_{2i})$ have to be generated to generate $\underline{y}$. The number of pairs has a geometric distribution with parameter p equal to $p = P\{\underline{w}_1+\underline{w}_2 \le 1\} = \int_0^1 (1-w)^{\alpha}\beta w^{\beta-1} dw = \Gamma(\alpha+1)\Gamma(\beta+1)/\Gamma(\alpha+\beta+1)$. Thus the expected number is $1/p = \Gamma(\alpha+\beta+1)/\Gamma(\alpha+1)/\Gamma(\beta+1)$; if e.g. $\alpha = \beta = 1$ then $1/p = 2$. 1/p may become very large as α and β increase. As a consequence this method will become less efficient. An alternative can be given by using gamma distributed random variables (see 4.10). Let $\underline{w}_1$ have a gamma distribution with parameters 1 and α, $\alpha > 1$ and let $\underline{w}_2$ be independent of $\underline{w}_1$ and let it have a gamma distribution with parameters 1 and β, $\beta > 1$, then $\underline{w}_1/(\underline{w}_1+\underline{w}_2)$ has a beta distribution with parameters α and β.

PROOF. Let $\underline{z}_1$ and $\underline{z}_2$ be defined by $\underline{z}_1 = \underline{w}_1/(\underline{w}_1+\underline{w}_2)$ and $\underline{z}_2 = \underline{w}_1 + \underline{w}_2$. The transformation from $(\underline{w}_1, \underline{w}_2)$ to $(\underline{z}_1, \underline{z}_2)$ corresponds to a bijective map from $\mathbb{R}^+ \times \mathbb{R}^+$ to $(0,1) \times \mathbb{R}^+$ ($\mathbb{R}^+$ is the set of positive real numbers). The inverse transformation is $\underline{w}_1 = \underline{z}_1\underline{z}_2$, $\underline{w}_2 = \underline{z}_2(1-\underline{z}_1)$, thus the density h of $\underline{z}_1$ and $\underline{z}_2$, computed by 2.2, is

$$h(z_1,z_2) = C(z_1z_2)^{\alpha-1}e^{-z_1z_2}\,(z_2(1-z_1))^{\beta-1}e^{-z_2(1-z_1)}\ \operatorname{abs}\left|\frac{\partial(w_1,w_2)}{\partial(z_1,z_2)}\right|,$$

where C is a constant, i.e. independent of z_1 and z_2. Hence

$$h(z_1,z_2) = C\, z_1^{\alpha-1}(1-z_1)^{\beta-1}z_2^{\alpha+\beta+1}e^{-z_2}$$

for every $(z_1,z_2) \in (0,1) \times \mathbb{R}^+$. From this we conclude that $\underline{z}_1$ has a beta distribution with parameters α and β. □

Note that this proof does not require $\alpha > 1$ and $\beta > 1$, it is just that otherwise the beta distribution is used to generate gamma distributed random variables, as will be seen in the next subsection.

4.10. THE GAMMA DISTRIBUTION

Let $\underline{x}$ have a gamma distribution with parameters σ and α, $\sigma > 0$, $\alpha > 0$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \begin{cases} \dfrac{1}{\sigma\Gamma(\alpha)}\left(\dfrac{x}{\sigma}\right)^{\alpha-1}\exp\{-\dfrac{x}{\sigma}\} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\Gamma(\alpha) = \int_0^\infty v^{\alpha-1}e^{-v}dv$ is the gamma function.

4.10.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated by applying a rejection method in the following way.

If $0 < \alpha < 1$ then first a random variable $\underline{w}$ is generated (e.g. by using 4.9) with a beta distribution with parameters α and $1-\alpha$, next $\underline{y}$ is computed according to

$$\underline{y} = -\,\sigma\,\underline{w}\,\log\,\underline{u}.$$

(Writing $\underline{u}$ implicity implies that $\underline{u}$ is independent of all other random variables, especially $\underline{w}$, unless stated otherwise.)

If $\alpha \geq 1$ then first a sequence of random variables $\underline{z}_1,\underline{z}_2,\ldots$ is generated from an auxiliary distribution with cumulative distribution

function H with $H(z) = z^\tau/(\alpha^\tau+z^\tau)$ if $z \geq 0$, $\tau = \sqrt{2\alpha-1}$, as proposed by CHENG [4]. This is done according to $\underline{z}_i = H^{-1}(1 - \underline{u}_{2i-1}) = \alpha(1/\underline{u}_{2i-1}-1)^{1/\tau}$ if $\underline{u}_{2i-1} < 1$. Next $\underline{i}$ is defined to be the smallest i satisfying

$$\frac{1}{M}\frac{f(\underline{z}_i)}{h(\underline{z}_i)} \geq \underline{u}_{2i},$$

where h is the density of H, $h(z) = \tau\alpha^\tau z^{\tau-1}/(\alpha^\tau+z^\tau)^2$ if $z > 0$, and $M = \sup_{z>0} \frac{f(z)}{h(z)} = 4\alpha^\alpha e^{-\alpha}/[\sqrt{2\alpha-1}\ \Gamma(\alpha)]$. Finally $\underline{y}$ is computed,

$$\underline{y} = \sigma\underline{z}_i.$$

PROOF. a. Suppose $0 < \alpha < 1$. Define random variables $\underline{v}$, $\underline{y}_1$ and $\underline{y}_2$ by $\underline{v} = -\log \underline{u}$, $\underline{y}_1 = \underline{v}\ \underline{w}$ and $\underline{y}_2 = \underline{v} - \underline{v}\ \underline{w}$. By using 2.2 we can compute the simultaneous probability density g of $\underline{y}_1$ and $\underline{y}_2$. For $y_1 > 0$ and $y_2 > 0$ we have, since $\underline{v}$ has an exponential distribution with parameter 1

$$g(y_1,y_2) = Ce^{-(y_1+y_2)}\left(\frac{y_1}{y_1+y_2}\right)^{\alpha-1}\left(\frac{y_2}{y_1+y_2}\right)^{-\alpha} \operatorname{abs}\begin{vmatrix} 1 & \frac{y_2}{(y_1+y_2)^2} \\ 1 & \frac{-y_1}{(y_1+y_2)^2} \end{vmatrix},$$

where C is a constant, i.e. independent of y_1 and y_2. Hence

$$g(y_1,y_2) = C\, y_1^{\alpha-1} e^{-y_1}\, y_2^{-\alpha} e^{-y_2} \quad \text{for all } y_1 > 0,\ y_2 > 0.$$

From this we see that $\underline{y}_1$ has a gamma distribution with parameters 1 and α.

b. Suppose $\alpha \geq 1$. See 2.3 and CHENG [4]. □

4.10.2. The algorithm

If $\alpha \geq 1$ we may proceed as follows to compute a realization y of $\underline{y}$.

1. Define constants $b = 1 - (\log 4)/\alpha$ and $c = \sqrt{2\alpha-1}$.
2. Let i be an integer variable with initial value 0.
3. i := i+1.
4. If $u_{2i-1} = 1$ then go to step 3.
5. Introduce the auxiliary variables r and t and assign $r := \log(1/u_{2i-1}-1)/c$, $t := \exp(r)$.

6. If $r - t + b - \log(u_{2i}u_{2i-1}(1-u_{2i-1}))/\alpha < 0$ then goto step 3 else
   $y := \sigma\alpha t$. □

4.10.3. Remark

If α has an integer value, a random variable, having a gamma distribution with parameters σ and α, also can be obtained as the sum of α independent random variables, each having an exponential distribution with parameter $1/\sigma$.

4.11. THE CHI-SQUARE DISTRIBUTION

Let $\underline{x}$ have a chi-square distribution with parameter ν, $\nu \in \mathbb{N}$, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = \begin{cases} \frac{1}{2\Gamma(\nu/2)} \left(\frac{x}{2}\right)^{\nu/2-1} \exp\{-\frac{x}{2}\} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\Gamma(\alpha) = \int_0^\infty v^{\alpha-1}e^{-v}dv$ is the gamma function.

4.11.1. The method

A random variable $\underline{y}$ with the same distribution as $\underline{x}$ is generated according to 4.10.1 with parameters $\sigma=2$ and $\alpha=\nu/2$.

4.11.2. Remark

If ν is even, a chi-square distributed random variable can be obtained as the sum of independent, exponentially distributed random variables, see 4.10.3.

4.12. THE MULTIVARIATE NORMAL DISTRIBUTION

Let $\underline{x}$ have a p-dimensional multivariate normal distribution with parameters $\mu \in \mathbb{R}^p$ and $\Sigma \in \mathbb{R}^{p\times p}$, Σ positive definite, i.e. the probability density of $\underline{x}$ is given by

$$f(x) = (2\pi)^{-p/2}|\Sigma|^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\} \quad \text{for every } x \in \mathbb{R}^p,$$

where $|\Sigma|$ is the determinant of the matrix Σ and $(x-\mu)'$ is the transpose of the vector $x - \mu$.

4.12.1. The method

Let $\Sigma = LL'$ be a decomposition of the covariance matrix Σ and suppose $\underline{z}_1,\ldots,\underline{z}_p$ are independent random variables, which are normally distributed with parameters 0 and 1, possibly generated according to Subsection 4.8, then a random vector $\underline{y}$ having the same distribution as $\underline{x}$ is generated according to

$$\underline{y} = L\underline{z} + \mu,$$

where $\underline{z} = (\underline{z}_1,\ldots,\underline{z}_p)'$.

4.12.2. Remark

The matrix L of the decomposition $\Sigma = LL'$ can be chosen to be lower triangular by using the Choleski decomposition method. A proof of this can be found in many books, e.g. BROYDEN [2].

4.13. AN ORDERED SAMPLE FROM A UNIFORM DISTRIBUTION

Let $\underline{x}_1,\ldots,\underline{x}_k$ be independent random variables, each uniformly distributed over the interval $(0,1]$. If $\underline{\pi}$ is a permutation of $\{1,2,\ldots,k\}$ such that $\underline{x}_{\underline{\pi}(1)} \leq \underline{x}_{\underline{\pi}(2)} \leq \ldots \leq \underline{x}_{\underline{\pi}(k)}$, then $(\underline{x}_{\underline{\pi}(1)},\ldots,\underline{x}_{\underline{\pi}(k)})$ is called an ordered sample from a (standard) uniform distribution.

4.13.1. The method

To generate a sample $(\underline{y}_1,\ldots,\underline{y}_k)$ with the same distribution as $(\underline{x}_{\underline{\pi}(1)},\ldots,\underline{x}_{\underline{\pi}(k)})$, we use the property that the spacings, i.e. the differences of successive observations, of an ordered sample from an exponential distribution are independently and again exponentially distributed, however with different parameter values.

For every integer i satisfying $1 \leq i \leq k$ let $\underline{s}_i$ be a random variable with an exponential distribution with parameter k-i+1 such that $\underline{s}_1,\ldots,\underline{s}_k$ are independent. These variables might be generated by using Subsection 4.2. Next we define $\underline{z}_i = \sum_{j=1}^{i} \underline{s}_j$ for every i, $1 \leq i \leq k$, then $(\underline{z}_1,\ldots,\underline{z}_k)$ is an ordered sample from an exponential distribution with parameter 1. Finally $(\underline{y}_1,\ldots,\underline{y}_k)$ is defined by $\underline{y}_i = \exp\{-\underline{z}_{k-i+1}\}$ for $1 \leq i \leq k$.

PROOF. See DAVID [5]. □

4.13.2. The algorithm

To compute a realization $(y_1,\ldots,y_k)$ of $(\underline{y}_1,\ldots,\underline{y}_k)$ we may proceed as follows.

1. Introduce the auxiliary variable s and initialize s := 0.
2. For each m = k,k-1,...,1 perform the steps 3 and 4.
3. s := s + (log u_m)/m.
4. $y_m := e^s$. □

4.13.3. Remarks

A realization $(y_1,\ldots,y_k)$ of an ordered sample from an exponential distribution with parameter λ can be generated by the following algorithm.

1. Introduce the auxiliary variable s and initialize s := 0.
2. For each m = k,k-1,...,1 perform the steps 3 and 4.
3. s := s - (log u_m)/m.
4. $y_{m-k+1} := s/\lambda$. □

One can generate an ordered sample from an arbitrary probability distribution, with cumulative distribution function F, by using the algorithm 4.3.12 with the fourth step replaced by

4^*. $y_m := \phi(e^s)$.

Here ϕ is a monotonous function, defined on the unit interval, such that F is the cumulative distribution function of $\phi(\underline{u})$, if $\underline{u}$ is uniformly distributed over the interval (0,1]. Thus F^{-1} is an example of a function ϕ like that.

5. DISCUSSION ON POSSIBLE LIMITATIONS WHEN USING A PSEUDO RANDOM NUMBER GENERATOR

A device to simulate a uniformly distributed random variable only approximately can have the property of complete unpredictability. There are rather good approximations, like a good lottery with a very large number of tickets. Curiously, devices exist which can be implemented in a digital computer and which give satisfactory results in a simulation. Some properties of such so called *pseudo random number generators* are:

1. Each time they are activated they render a number, called a *realization*, belonging to some fixed set, $N_m = \{0,1,\ldots,m-1\}$ say, where m is a parameter of the generator, in general closely connected to the way an integer number is represented in the computer.
2. Successive realizations are successive elements of a finite cyclic sequence, $r_1,r_2,\ldots,r_{n-1},r_n,r_1,\ldots$ say, where n may be very large.

Two examples of such generators are the linear congruential generator, e.g. see KNUTH [10], and the generator of POHL [11]. Thus a pseudo random number generator is *perfectly predictable*! An advantage of this is, e.g., the possibility of tracing errors in a simulation computer program. From the properties stated above it will be clear that certain precautions are necessary for the pseudo random number generator to look like a (genuine) random number generator. "To look like" in this context may be interpreted in the way that one can not discriminate between simulation results based on realizations of the real generator and the pseudo generator, if the only information available consists of those simulation results. As an example to illustrate this, consider the following two sequences, one produced by tossing a coin and one produced by a pseudo random number generator and method 3.1.1.

0,1,1,0,1,1,0,1,0,0,...
0,1,1,0,1,0,1,1,1,1,...

The question is whether one can recognize the producing devices by such sequences.

The precautions to be taken will depend on the kind of simulation experiment in which the pseudo random number generator is to be used. In

the following we shall consider two situations in which a linear congruential generator is used.

If the generator has been constructed carefully, n = m and the best one can say is that n of its successive realizations "look like" a realization of a random permutation of N_m, since every number in N_m occurs exactly once in each cycle. This implies that we accept a random permutation of N_m to be a mathematical model for the linear congruential generator. The problem is whether or not a sequence of independent, uniformly distributed random variables can be a reasonable model for this generator. The limitations we are about to introduce, will reduce the differences between realizations of the two models in a particular situation.

CASE 1. One has to simulate a sample of size k from a random variable x, uniformly distributed over $\{1,\ldots,N\}$, $(\underline{x}_1,\ldots,\underline{x}_k)$ say. Apply the algorithm of 3.1. This induces a partition $\{C_1,\ldots,C_N\}$ of the set of m possible realizations such that each C_i contains about m/N elements and the event $\{\underline{x} = i\}$ corresponds to the generator producing a result in C_i. Assume that the sample is generated stepwise and let the condition be imposed that in each step the ratio of the probabilities of obtaining a result in C_i in both models does not differ more than some small constant, η say, $0 < \eta < 1$, from 1, uniformly in i. This implies

$$\left| \frac{m/N - (k-1)}{m - (k-1)} \Big/ \frac{1}{N} - 1 \right| \leq \eta,$$

which is the same as $(N-1+\eta)(k-1) \leq m\eta$. For the inverse ratio we obtain $(N-1+N\eta)(k-1) \leq m\eta$. Both relations are implied by $Nk \leq m\eta/(1+\eta)$. Since η is supposed to be small, in this situation

$$Nk \leq m\eta$$

could be a useful condition. Let all numbers be represented w.r.t. some base, b say, $b > 1$, then from this condition the following rule of thumb can be derived by applying $^b\log$ to both sides.

(no. of digits necessary to represent the no. of categories)
+ (no. of digits necessary to represent the sample size)
≤ (no. of digits given by the pseudo random number generator)
- (no. of digits necessary to represent $1/\eta$). □

CASE 2. Consider the same situation as before and define n_i = no. of sample observations in C_i. In the first model, in which the pseudo random number generator is represented by a random permutation, the variance of $\underline{n}_i$ will be smaller than in the second model, in which the generator is represented by a sequence of independent, uniformly distributed random variables. A condition that one might want to impose is that the ratio of the variances in the first and second model does not differ more than some small amount, ε say, $0 < \varepsilon < 1$, from 1.

In the first model $\underline{n}_i$ has a hypergeometric probability distribution with variance

$$\mathrm{var}_1\, \underline{n}_i \overset{*)}{\approx} \frac{m}{N} \times \frac{m(N-1)}{N} \times k \times (m-k) \times [m^2(m-1)]^{-1} = \frac{(N-1)k(m-k)}{N^2(m-1)}.$$

$\underline{n}_i$ in the second model is binomially distributed with variance

$$\mathrm{var}_2\, \underline{n}_i \approx k \times \frac{m/N}{m} \times \frac{m(N-1)/N}{m} = \frac{k(N-1)}{N^2}.$$

The ratio is $\mathrm{var}_1\, \underline{n}_i / \mathrm{var}_2\, \underline{n}_i \approx (m-k)/(m-1)$. The condition $(m-k)/(m-1) \geq 1-\varepsilon$ is equivalent to $k \leq m - (m-1)(1-\varepsilon)$; it is implied by

$$k \leq m\varepsilon. \qquad \square$$

These are just two special cases to indicate what kind of limitations might be introduced. In general, before doing a simulation experiment one has to analyse the specific situation at hand and one has to introduce conditions under which the differences between the simulation and the ideal, theoretical situation will not be "too obvious". Finally these conditions have to be reformulated in terms of the parameters of the simulation experiment.

---

*) " ≈ " because the no. of elements of C_i is only approximately equal to m/N.

REFERENCES

[1] BICKEL, P.J. & K.A. DOKSUM (1977), *Mathematical Statistics: Basic Ideas and Selected Topics,* Holden-Day, San Francisco, pp. 10-11.

[2] BROYDEN, C.G. (1975), *Basic Matrices,* The Macmillan Press LTD, London, p. 59.

[3] BURY, K.V. (1975), *Statistical Models in Applied Science,* John Wiley & Sons, New York.

[4] CHENG, R.C.H. (1977), *The Generation of Gamma Variables with Non-integral Shape Parameter,* Appl. Stat. 26, 1, pp. 71-74.

[5] DAVID, H.A. (1970), *Order Statistics,* J. Wiley & Sons, New York, p. 17.

[6] HAMMERSLEY, J. & P. HANDSCOMB (1964), *Monte Carlo Methods,* Methuen and Company, London.

[7] JOHNSON, N.L. & S. KOTZ (1969), *Distributions in Statistics: Discrete Distributions,* Houghton Mifflin Company, Boston.

[8] JOHNSON, N.L. & S. KOTZ (1970), *Distributions in Statistics: Continuous Univariate Distributions,* I, II, Houghton Mifflin Company, Boston.

[9] KAAS, R. (1977), *The Complexity of Drawing an Ordered Random Sample,* Techn. Report 51/77, Mathematical Centre, Amsterdam.

[10] KNUTH, D.E. (1969), *The Art of Computer Programming,* Volume 2, Semi-numerical algorithms, Addison Wesley, Reading Mass.

[11] POHL, P. (1976), *Description of MCV, a Pseudo Random Number Generator,* Scand. Act. J., pp. 1-14.

[12] YAKOWITZ, S.J. (1977), *Computational Probability and Simulation,* Addison-Wesley, Reading, Massachusetts.