12. Random drawing

12.1 Pseudo-random number streams

All random drawing procedures of SIMULA 67 are based on the technique of obtaining "basic drawings" from the uniform distribution in the interval <0,1>.

A basic drawing will replace the value of a specified integer variable, say U, by a new value according to an implementation defined algorithm. As an example, the following algorithm may be suitable for binary computers:

$$U_{i+1} = \text{remainder } ((U_i \times 5^{2p+1}) \div 2^n)$$

where U_i is the i'th value of U, n is an integer related to the size of a computer word and p is a positive integer. It can be proved that, if U_o is a positive odd integer, the same is true for all U_i and the sequence U_o, U_1, U_2, ... is cyclic with period 2^{n-2}. (The last two bits of U remain constant, while the other n-2 take on all possible combinations).

The real numbers $u_i = U_i \times 2^{-n}$ are fractions in the range <0,1>. The sequence u_1, u_2, ... is called a "stream" of pseudo-random numbers, and u_i (i = 1,2, ...) is the result of the i'th basic drawing in the stream U. A stream is completely determined by the initial value U_o of the corresponding integer variable. Nevertheless, it is a "good approximation" to a sequence of truly random drawings.

12.2 Random drawing procedures

The following procedures all perform a random drawing of some kind. Unless it is explicitly stated otherwise, the drawing is effected by means of one single basic drawing, i.e. the procedure has the side effect of advancing the specified stream by one step. The necessary type conversions are effected for the actual parameters, with the exception of the last one. The latter must always be an integer variable specifying a pseudo-random number stream.

1. Boolean procedure draw (a,U); name U; real a;
   integer U;

   The value is true with the probability a, false with the probability 1 - a. It is always true if a ≥ 1 and always false if a ≤ 0.

2. integer procedure randint (a,b,U); name U;
   integer a,b,U;

   The value is one of the integers a, a+1, ..., b-1, b with equal probability. If b < a, the call constitutes an error.

3. real procedure uniform (a,b,U); name U; real a,b;
   integer U;

   The value is uniformly distributed in the interval [a,b>. If b < a, the call constitutes an error.

4. real procedure normal (a,b,U); name U;
   real a,b; integer U;

   The value is normally distributed with mean a and standard deviation b. An approximation formula may be used for the normal distribution function.

(See M. Abramowitz & I. A. Stegun (ed): Handbook of Mathematical Functions, National Bureau of Standard Applied Mathematics Series No. 55, p. 952 and C. Hastings formula (26.2.23) on p. 933.)

5. <u>real</u> <u>procedure</u> negexp (a,U); <u>name</u> U; <u>real</u> a; <u>integer</u> U;

The value is a drawing from the negative exponential distribution with mean 1/a, defined by -ln(u)/a, where u is a basic drawing. This is the same as a random "waiting time" in a Poisson distributed arrival pattern with expected number of arrivals per time unit equal to a.

6. <u>integer</u> <u>procedure</u> Poisson (a,U); <u>name</u> U; <u>real</u> a; <u>integer</u> U;

The value is a drawing from the Poisson distribution with parameter a. It is obtained by n+1 basic drawings, u_i, where n is the function value. n is defined as the smallest non-negative integer for which

$$\prod_{i=0}^{n} u_i < e^{-a}$$

The validity of the formula follows from the equivalent condition

$$\sum_{i=0}^{n} -\ln(u_i)/a > 1$$

where the left hand side is seen to be a sum of "waiting times" drawn from the corresponding negative exponential distribution.

When the parameter a is greater than some implementation defined value, for instance 20.0, the value may be approximated by entier (normal (a,sqrt (a),U)+0.5) or, when this is negative, by zero.

7. <u>real</u> <u>procedure</u> Erlang (a,b,U); <u>name</u> U; <u>integer</u> U;
   <u>real</u> a,b;

The value is a drawing from the Erlang distribution with mean 1/a and standard deviation $1/(a\sqrt{b})$. It is defined by b basic drawings u_i, if b is an integer value,

$$-\sum_{i=1}^{b} \frac{\ln(u_i)}{ab}$$

and by c+1 basic drawings u_i otherwise, where c is equal to entier (b),

$$-\left(\sum_{i=1}^{c} \frac{\ln(u_i)}{ab}\right) - \left(\frac{(b-c)\ \ln\ (u_{c+1})}{ab}\right)$$

both a and b must be greater than zero.
The last formula represents an approximation.

8. <u>integer</u> <u>procedure</u> discrete (A,U); <u>name</u> U;
   <u>real</u> <u>array</u> A; <u>integer</u> U;

The one-dimensional array A, augmented by the element 1 to the right, is interpreted as a step function of the subscript, defining a discrete (cumulative) distribution function. The array is assumed to be of type <u>real</u>.

The function value is an integer in the range [lsb, usb+1], where lsb and usb are the lower and upper subscript bounds of the array. It is defined as the smallest i such that A[i]> u, where u is a basic drawing and A [usb+1] = 1.

9. <u>real</u> <u>procedure</u> linear (A,B,U); <u>name</u> U;
   <u>real</u> <u>array</u> A,B; <u>integer</u> U;

The value is a drawing from a (cumulative) distribution function F, which is obtained by linear interpolation in a non-equidistant table defined by A and B, such that A[i] = F(B[i]).

It is assumed that A and B are one-dimensional <u>real</u> arrays of the same length, that the first and last elements of A are equal to 0 and 1 respectively and that A[i] $\geq$ A[j] and B [i]> B [j] for i > j. If any of these conditions are not satisfied, the effect is implementation defined.

The steps in the function evaluation are:

1. draw a uniform <0,1> random number, u.

2. determine the lowest value of i, for which
   A[i-1] $\leq$ u $\leq$ A[i]

3. compute D= A[i] - A[i-1]

4. if D= 0: linear = B[i-1]
   if D$\neq$ 0: linear = B[i-1] +

$$\frac{(B[i] - B[i-1])}{D} \quad (u-A[i-1])$$

10. <u>integer</u> <u>procedure</u> histd (A,U); <u>name</u> U; <u>real</u> <u>array</u> A;
<u>integer</u> U;

The value is an integer in the range [lsb,usb], where lsb and usb are the lower and upper subscript bounds of the one-dimensional array A. The latter is interpreted as a histogram defining the relative frequencies of the values.