# CHAPTER 7.

# RANDOM DRAWING.

## 7.1 Pseudo-random Number Streams.

All random drawing procedures of SIMULA use the same 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 the following algorithm.

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

where $U_i$ is the i'th value of U.

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 the period $2^{n-2}$. (The last two bits of U remain constant, while the other n-2 take on all possible combinations.) In UNIVAC 1107 we have n = 35. p is chosen equal to 6.

The real numbers $u_i = U_i \times 2^{-n}$ are fractions in the range $\langle 0,1 \rangle$. 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.

By reversing the sign of the initial value $U_0$ of a stream variable the antithetic drawings $1 - u_1$, $1 - u_2$, ------ are obtained. In certain situations it can be proved that means obtained from samples based on antithetic drawings have a smaller variance than those obtained from uncorrelated streams. This can be used to reduce the sample size required to obtain reliable estimates.

## 7.2 Random Drawing Procedures.

The following procedures all perform a random drawing of some kind. Unless otherwise is explicitly stated the drawing is effected by means of <u>one single</u> 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 <u>integer variable</u> specifying a pseudo-random number stream. All parameters except the last one are called by value.

1. <u>Boolean</u> <u>procedure</u> draw (a, U); <u>real</u> a; <u>integer</u> U;

   The value is <u>true</u> with the probability a, false with the probability 1 - a. It is always <u>true</u> if $a \geq 1$, always <u>false</u> if $a \leq 0$.

2. <u>integer</u> <u>procedure</u> randint (a, b, U); <u>integer</u> a, b, U;

   The value is one of the integers a, a + 1, ---, b - 1, b with equal probability. It is assumed that $b \geq a$.

3. <u>real</u> <u>procedure</u> uniform (a, b, U); <u>real</u> a, b; <u>integer</u> U;

   The value is uniformly distributed in the interval [a, b>. It is assumed that b > a.

4. <u>real</u> <u>procedure</u> normal (a,b,U); <u>real</u> a,b,; <u>integer</u> U;

The value is normally distributed with mean a and standard deviation b. An approximation formula is 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> psnorm (a, b, c, U); <u>real</u> a, b; <u>integer</u> c, U;

The value is formed as the sum of c basic drawings, suitably transformed so as to approximate a drawing from the normal distribution. The following formula is used:

$$a + b\left(\left(\sum_{i=1}^{c} u_i\right) - c/2\right)\sqrt{12/c}$$

This procedure is faster, but less accurate than the preceding one. c is assumed $\leq 12$.

6. <u>real</u> <u>procedure</u> negexp (a, 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.

7. <u>integer</u> <u>procedure</u> Poisson (a, 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

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

The validity of the formula follows from the equivalent condition

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

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 20.0, the value is approximated by integer (normal (a,sqrt(a),u)) or, when this is negative, by zero.

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

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,

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

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

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

Both a and b must be greater than zero.

9. ***integer procedure*** discrete (A, U); ***array*** A; ***integer*** 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 ***real***.

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.

10. ***real procedure*** linear (A, B, U); ***array*** A, B; ***integer*** 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 ***real*** arrays of the same length, that the first and last elements of A are equal to 0 and 1 respectively and that A [i] ≥ A [j] and B [i] > B [j] for i > j.

11. ***integer procedure*** histd (A, U); ***array*** A; ***integer*** 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.

This procedure is more time-consuming than the procedure discrete, where the cumulative distribution function is given, but it is more useful if the frequency histogram is updated at run time.