Denis Lorrain
Centre d'Applications Musicales
de l'Informatique (CAMI)
Faculté de Musique
Université de Montréal
C. P. 6128
Montréal, Québec
Canada

A Panoply of Stochastic 'Cannons'

1.0 Introduction

The term *stochastic* was introduced in music by Xenakis (1963). It is a synonym of *random*, but it has been preferred as a more scientific, less commonplace term, used by Xenakis to qualify his use of probability theory. Earlier, this use set it apart from the then frequent use of "random" sections in contemporary music, which are in fact improvised sections—random only from the composer's point of view! *Cannon* is a term Xenakis coined to name algorithms "shooting" random values according to a specific probability distribution. (In French we say *tirer au hasard* for "choose at random." literally "shoot at random.")

In 1954, Xenakis set forth two criticisms of serial composition, which was then undeniably the most widely used technique among avant-garde composers: "The serial system is brought back into question in its two foundations which bear the seed of their own destruction and transcendence" (1971*a*, p. 120).

His first criticism is general. To begin with, Xenakis stated, serial manipulations are nothing but particular cases in the vast domain of combinatorial analysis. Moreover, he claimed, serialism remains encrusted in its historical inheritance while, especially through electroacoustics, new areas are open: "Why twelve and not thirteen or *n* sounds? Why not the continuity of the frequency spectrum? Of the timbre spectrum? Of the intensity spectrum and of durations?" (1971*a*, p. 120).

Xenakis's second criticism concerns a contradiction of serial processes with regard to their sound result: "Upon audition, the enormous complexity prevents one from following the tangle of lines, and has as macroscopic effect an unreasoned and fortuitous dispersion of sounds in the whole sound spectrum" (1971*a*, p. 120).

He concluded: "The macroscopic effect could thus be controlled by the mean of the motions of the *n* objects chosen by us. Hence follows the introduction of the idea of probability, which again implies combinatorial analysis in this precise case" (1971*a*, p. 120). This introduces the concept of overall control of complex sound events rather than their analytic and mechanistic elaboration, exactly as in science, where statistical methods are a tool for the overall prehension of exceedingly complex phenomena; this conforms to our intuitive idea of randomness. But it must be added that Xenakis had, on the other hand, a certain musical intuition about sound masses, probably anterior to their rational justification (Xenakis 1963, pp. 19–20).

In the course of a long practice of serial techniques, Gottfried Michael Koenig comes to the same conclusion: "The greater the extent to which parameters adhere to certain arrangements and the greater the extent to which musical meaning is to depend on the perception of these arrangements, the more unpredictable, 'random,' are the effects of one parameter on the other: the various characteristics coalesce into 'sounds' whose order is unequivocally defined neither by the course of an individual parameter nor by the polyphony of all of them" (1971, p. 22). The ambition of systematic composition, "A formalism completely describing in two directions [past and future events] each instant of the work" (Koenig 1971, p. 9), upsets itself. Besides, technically, complex determinisms generate an excess of possibilities, a maze of ambiguities. Hence, for instance, Koenig's concept of *potential* and *actual form:* a piece does not necessarily realize all possibilities in a domain, but they are available: the composer chooses (Koenig 1971, pp. 65–66).

When the composer does not wish to choose throughout, he or she arrives at what serialists have called *randomness*, negotiating a compromise between their conception of musical composition and Cage's ideas. In certain sections of a piece, the composer may supply basic material, on which an ad hoc solution must be "improvised." Or, at the macrostructural level, the composer may make available choices among prefabricated musical modules. In any case, the fundamental question of choice is only postponed; but the work gains generality—attractive combinations of varied perspectives on the soundscape (Murray Schafer) accessible to the piece.

But compositional use of randomness is not the same as randomness as a metaphysical principle; as an impersonal technique intended to let sounds and silence be, simply, which Cage practiced beforehand (Daniel 1978). In a sense, Cage wishes to liberate the sounds from music and musicians, whereas Xenakis "is voluntarily in keeping, not with the tradition, but with the history of his art" (Revault d'Allonnes 1973, p. 235).

Stochastic music is more than a technical solution; it is the fruit of a long tradition of rational thought. Indeed, probability theory is a formal structure based on the fundamental question, What can happen in our world, in such and such a given situation? Granted a basic hypothesis (homogeneity or not, varied concepts of mean, etc.), probability theory provides choosing tools, adapted to the hypothesis—even *inferred* from it, and we find our conception of artistic creation in choice.

The stochastic cannons or automata presented here might be used to control various dimensions of a compositional process, such as the elements of pitch, register, duration, amplitude, and so forth. Or they might be employed in the composition of sound microstructure, in particular in the synthesis of sound-wave forms.

I am deeply indebted to Xenakis: my understanding of probability theory has come through him—the case of a nonmathematician introducing to another some mathematical knowledge. In the following pages you will find methods for generating random variables which are not new, original, or exclusive in themselves. For instance, you may read others (Knuth; Maurin 1975) that refer to a large literature. This paper, however, has been written by a musician with the intention of gathering a panoply for musical purposes. My aim is to give musicians with even less mathematical knowledge than I useful tools. A particular aim is to bring many computer applications in musical composition out of the eternal and poor use of random values directly from the computer's omnipotent black box: the uniform distribution pseudo-random generator.

2.0 Preliminary Definitions

2.1 Probability

A probability is a fraction of a *sample space:* the set of all possible outcomes in a given situation. Outcomes could be card configurations in a card game, the states of a machine, the colors of cars passing on the street, or the 88 keys of a piano. The probability of one specific event amongst the whole sample space (an event is a subset of the sample space) is

$$\frac{\text{number of ways the event can happen}}{\text{number of outcomes in the sample space}}$$

Clearly, the value of this fraction must be between 0 and 1, for

$$0 = \frac{\text{number of ways an event } \textit{absent} \text{ from the sample space can happen}}{\text{number of outcomes in the sample space}}, \quad (1)$$

$$\text{and } 1 = \frac{\text{number of outcomes in the sample space}}{\text{number of outcomes in the sample space}}.$$

A probability is thus a ratio comparing two sets of events: the sample space and a subset of the latter—an event, measuring the proportion of a specific event included in the sample space. The larger the probability of an event, the more likely it is to happen, provided that something happens.

In simple cases, probabilities are quite straightforward. For instance, the probability of obtaining an odd result when throwing a die is 3/6 (the event is three outcomes out of six). But if two dice are thrown, some thinking is necessary. Here is a famous problem, "Banach's Match Boxes" (Feller 1966*a*, p. 166). A smoker has two match boxes of n matches each; when he needs a match he selects a box at random; when he takes the last match in a box, what is the probability that the other still contains x matches? Complex combinatorial calculations often must be performed merely to determine how many ways can the event happen and how many events are there in the sample space.

This intuitive approach to probability can suffice for our purpose, and will be completed as needed. Except as concrete examples for some basic notions, we will not deal here with real situations or mechanisms; we shall rather describe algorithms—imaginary machines—conforming to some classical schemes of probability theory. For thorough and formal definitions of probability and elements of combinatorial analysis, see Calot (1967, chapters 1 and 2), or Feller (1966*a*, chapters 1 and 2). Simple introductions to probability theory can be found in college mathematical textbooks; Jacquard (1976) is clear and quite complete, as is Lipschutz (1965).

2.2 Random Variables

A *random variable* (denoted X hereafter) is a variable that assumes specific values randomly (denoted x_s hereafter) corresponding to specific events out of the sample space. Thus if the sample space is an urn containing 12 colored balls: 3 red, 5 blue and 4 white, the random variable X is the color of a ball picked at random: it can assume three values {red, blue, white} which can be *abstracted* and *ordered* as $\{x_1, x_2, x_3\}$. Moreover, we know the probabilities P of these values:

$$P\{X = x_1\} = \frac{3}{12},$$

$$P\{X = x_2\} = \frac{5}{12}, \text{ and}$$

$$P\{X = x_3\} = \frac{4}{12}, \tag{1}$$

which are read: "the probability that X is x_1 equals" and so forth.

From the definitions of probabilities and random variables, it follows that the sum of the probabilities of all eventual values of a random variable X equals 1 for s covering the range of values of x,

$$\sum P\{X = x_s\} = 1 = P\{\text{any event in the sample space}\}. \tag{2}$$

In the above example:

$$\sum_{s=1}^{3} P\{X = x_s\} = \frac{3}{12} + \frac{5}{12} + \frac{4}{12} = 1.$$

2.3 Probability Distributions

The concept of random variable leads to that of probability distribution as a function of x: for s covering the range of values of x, the function

$$f(x_s) = P\{X = x_s\}$$

is called the *probability distribution* of the random variable X. We have now only stated that $P\{X = x_s\}$ is a function of x_s. The usefulness of this statement is perhaps not apparent for the time being, but it will prove to be of extreme importance later on, when we shall make use of authentic functions of x to describe probability distributions, particularly for continuous random variables (Section 2.6). We can at least use this terminology in statements such as:

$$0 \leqslant f(x_s) \leqslant 1,$$

equivalent to Eq. (1), and

$$\sum f(x_s) = 1, \tag{3}$$

equivalent to Eq. (2).

Probability distribution can be visualized in a two-dimensional graph as any function of x; a graph of probability distribution is called a *histogram*. Fig. 1 shows the histogram of the urn example in Section 2.2; each vertical bar has its height scaled to the probability it illustrates.

2.4 Distribution Functions

The above probability distribution $f(x)$ must be distinguished from the *distribution function* $F(x)$ of X, defined as

$$F(x_s) = \sum_{r \leq s} f(x_r) = P\{X \leqslant x_s\} \tag{4}$$

Fig. 1. *Histogram of urn example in Section 2.2.*

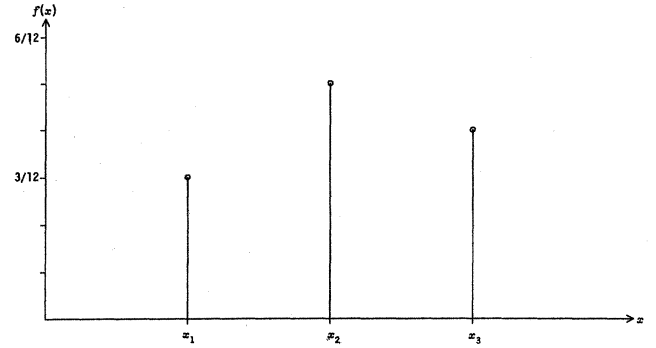
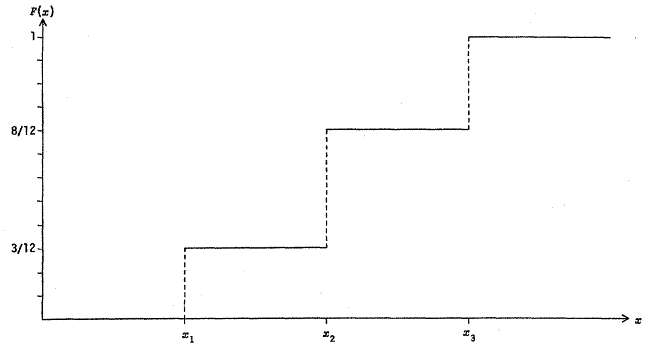


Fig. 2. *Distribution function of urn example in Section 2.2.*



or, the sum of the probabilities of events not exceeding x_s in order. Stated otherwise: $F(x_s)$ is the probability of X assuming one of the values

$$\{x_1, x_2, \ldots, x_s\},$$

forming a subset of the sample space. Equation (3) in Section 2.3 represented the particular case when $F(x_s) = 1$ because s covers the entire sample space. In our urn example,

$$F(x_2) = \sum_{r=1}^{2} f(x_r) = \frac{3}{12} + \frac{5}{12} = \frac{2}{3},$$

two times out of three X will be x_1 or x_2 (the ball will not be white).

This concept of distribution function may seem pointless now but it will prove fundamental to our discussion of algorithms. A distribution function can be graphed: for our urn example, Fig. 2 shows clearly the cumulative definition of $F(x)$.

2.5 Discrete Variables

A random variable X is said to be discrete if it can take values from an ordered set of a finite number of events. We have only dealt with discrete variables so far in order to establish a solid foundation for the next paragraph concerning continuous variables. Concretely, discrete variables are those which can assume only absolutely distinguishable values, such as x_1, x_2, or x_3, with no alternative in between. This is visible in their histogram made of different vertical bars in Fig. 1 and their stepwise distribution function in Fig. 2.

For a complete theory of discrete variables, see Feller (1966*a*) or appropriate chapters in Calot (1967). Gnedenko and Khintchine (1969) provide a very good introduction to discrete probability theory that is easy to read. College textbooks such as these suggested are often limited to discrete variables.

2.6 Continuous Variables

A random variable X is continuous when its distribution function $F(x)$ is continuous. This implies that its probability distribution $f(x)$—although it may include some discontinuities—is not, as for a discrete variable, a collection of discrete values $[f(x_1), f(x_2), \ldots, f(x_n)]$. Rather, its distribution is an $f(x)$ defined over a range of real numbers which can be assumed by X. Figure 3 shows the histogram of a continuous distribution (the normal distribution); this random variable can assume an infinity of values—any real number.

In the case of discrete variables, Eq. (3) in Section 2.3 stated that the sum of the probabilities of all events in a sample space must equal one. Graphically, this means that the sum of the histogram's vertical bars must be scaled to one (Fig. 1). In the case of continuous variables, it is the area bounded by the probability distribution and the horizontal axis which equals one. This can be intuitively understood by imagining this area as being filled with an infinity of vertical bars corresponding to the infinity of real values assumable by the random variable, as in the shaded area of Fig. 4. Thus when passing from discrete to continuous variables,

Fig. 3. Histogram of a continuous distribution (normal distribution).

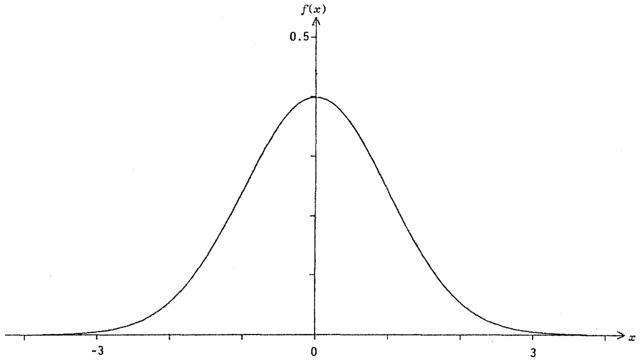
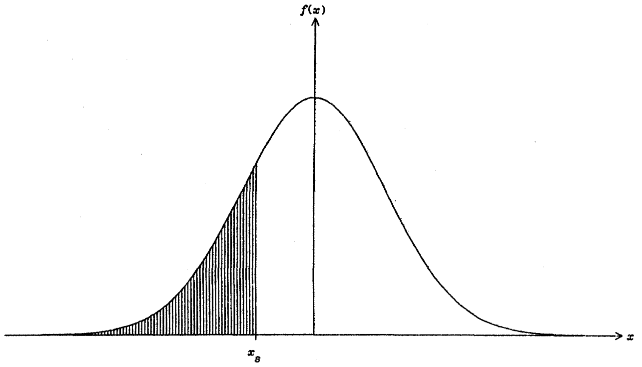


Fig. 4. Distribution function as the area under the probability distribution: the shaded area is $F(x_s) = P\{X \leq x_s\}$.



we substitute the integration of Eq. (5) for the summation of Eq. (3):

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{5}$$

In the same manner, the definition of a *continuous distribution function* is

$$F(x_s) = \int_{-\infty}^{x_s} f(x)dx = P\{X \leq x_s\}, \tag{6}$$

instead of Eq. (4) (Section 2.4). This is illustrated in Fig. 4.

In the case of discrete variables, probabilities of events correspond to vertical bars of scaled height in the histogram (Fig. 1). For continuous variables, because of the infinity of possible values, one must abandon the idea of a probability associated with one specific event; indeed, this probability is zero:

$$\frac{\text{one specific event}}{\text{an infinity of outcomes in the sample space}} = \frac{1}{\infty} = 0.$$

Thus the "height" of $f(x_s)$ is not directly the probability of x_s. We must be satisfied with knowing the probability of X assuming a value in a certain interval called *differential* (dx), which can be as "thin" as we wish, but not null. Thus the probability distribution of a continuous variable X is $f(x)$, but the probability of a specific x_s is zero:

$$P\{X = x_s\} = \frac{1}{\infty} = 0.$$

We must introduce the differential in order to handle a "tangible" probability:

$$P\{X = \text{in the interval } dx \text{ around } x_s\} = f(x_s)dx.$$

This is illustrated in Fig. 5: the area under $f(x_s)$ is null (the mathematical point $f(x_s)$ has no "width") and we must approximate $P\{X \approx x_s\}$ with a small differential rectangle of width dx. With this mathematical provision, a continuous histogram can be read directly: the higher $f(x_s)$, the more probable is $\{X \approx x_s \text{ in } dx\}$.

For an introduction to differential and integral calculus, and an understanding of these transitions from discontinuities to continuity, one can study Calot (1967) with profit or Guénon (1946) for an historical and philosophical point of view. Feller (1966*b*) and Calot contain complete theories of continuous variables.

2.6.1 Continuous Uniform Variables

The *continuous uniform* variable, for which we will henceforth use the special symbol U, is one that can take any real value u between zero and one. It is defined as follows:

$$f(u) = 0 \text{ if } u < 0$$
$$f(u) = 1 \text{ if } 0 < u < 1$$
$$f(u) = 0 \text{ if } u > 1.$$

Figure 6 shows its histogram. The rectangular area under $f(u)$ and the U axis measures one by one: the total area is thus one, in accordance with Eq. (5) in Section 2.6. Figure 7 shows the continuous uniform distribution function: notice how $F(u)$ is indeed continuous (although $f(u)$ contains two discontinuities), and how it increases from zero to one as u also

Fig. 5. *The area of the rectangle measuring* $f(x_s)$ *by* dx *is the probability of* X *assuming a value in the interval* dx *around* x_s.

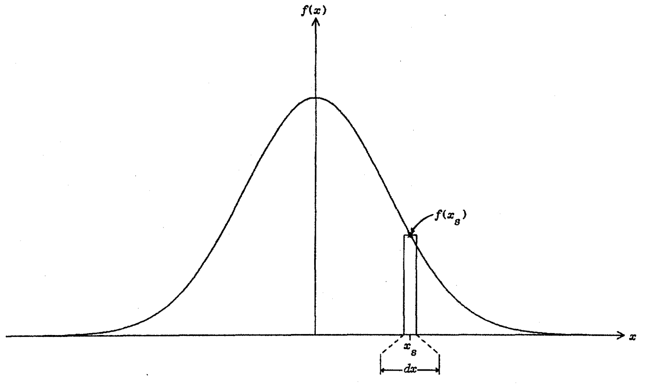
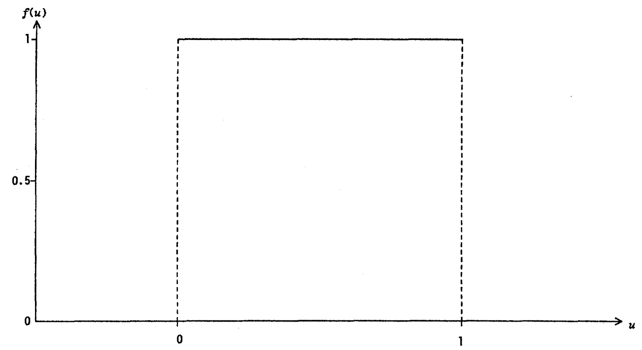


Fig. 6. *Histogram of the continuous uniform distribution* U.



increases in the same interval: in this case, the distribution function is simply equal to the variable, that is

$$F(u_s) = \int_{-\infty}^{u_s} f(u)du = u_s = P\{U \leq u_s\}. \qquad (7)$$

We shall make a constant use of this fact in our cannons, since the continuous uniform variable U (see Section 4.0, for more details) is the standard "pseudorandom number generator" available in computers as library function, under names such as **ran, ranf, rand,** and so forth, and all our algorithms will rely on it.

2.7 The Continuous Structure of Musical Parameters

In order to establish the continuous structure of the musical parameters we intend to work with, we will start with a consideration of pitch, and we shall assume the mere capability of perceiving distinct pitches. From a collection of discrete pitches comes forth some terms for qualifying different sensations (the usual "low" and "high"). Then follows the notion of pitch differences—distances between pitch pairs—together with the ability to compare these in terms of size. In comparing, we perform mental operations such as abstracting pitch differences from chronological pitch successions, transferring differences between various higher or lower pitch pairs, inverting the direction of differences, from "toward high" to "toward low" or vice versa, perhaps using reiterations of a small unit difference (e.g., semitone) to measure larger ones, and so forth.

By these operations we could intuitively arrive at a *total ordering* of pitch sensations, since we could evaluate the relative size and direction of differences between any two pitches. Moreover, we could axiomatically define tempered pitch scales with merely a reference pitch and one arbitrary unit difference, as Xenakis has shown following Peano's axiomatization of numbers (Xenakis 1971*a*, p. 61).

Of course, musicians take for granted this totally ordered structure of pitches, and have long since gained a great ease in the manipulation of pitch intervals. If we define addition ($\oplus$) of two intervals as making the arrival point of the first and the starting point of the second interval correspond to the same pitch, and consider the result of the operation as the interval from the starting point of the first to the arrival point of the second operand, it can be shown that in such usual musical operations, pitch intervals have the mathematical structure of a *commutative group.* (See Grossman and Magnus, Chapters 1 and 2 for group definitions):

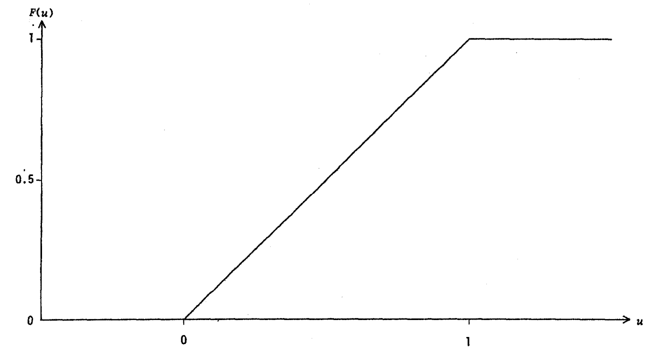
(a) stability: addition of any two intervals yields another interval;

(b) associativity: for any three intervals, $(d_1 \oplus d_2) \oplus d_3 = d_1 \oplus (d_2 \oplus d_3)$;

(c) there exists an identity element I (prime interval, or unison) such that $d \oplus I = I \oplus d = d$;

(d) any interval d has its inverse d^{-1} (the same

Fig. 7. Distribution function of the continuous uniform distribution U.



size of interval, but inverted in direction) such that

$$d \oplus d^{-1} = d^{-1} \oplus d = I;$$

(e) commutativity comes to boot:

$$d_1 \oplus d_2 = d_2 \oplus d_1.$$

Usually this group of intervals of pitch sensations has the same structure (i.e., is isomorphous to that of rational numbers under arithmetical addition). Ancient and traditional music was based on varied scales and thus handled rational numbers; and we can only gain in theoretical generality if we extend this domain to the whole set of real numbers, of which the rationals are a subset, to all possible pitch intervals. If the starting situation were tempered scales, such as twelve-tone equal-temperament, the connection would be similar: we can label the intervals with the number of unit intervals they represent in size and direction, in which case we form a group isomorphous to that of integers under arithmetical addition (a subgroup of the preceding). We can expand this situation to cover rational, irrational, and real numbers as well.

If we leave the domain of perception for that of physical phenomena, and speak in terms of frequency, the extension to real numbers is as feasible. Nontempered scales are built from a group of rational ratios under arithmetical multiplication and tempered intervals can be considered as irrational ratios (for instance, involving roots of two if they possess an octave modulo); from here also the extension to real numbers follows naturally.

In such ways, we gradually fill up the gaps between integers and rationals, tending toward continuity. Leibnitz stated, "It follows, from the actual division, that in a part of matter, however small it may be, there exists a sort of world consisting of innumerable creatures" (Guénon 1946, p. 55). We arrive at intervals smaller than our perceptual difference threshold, vanishing into continuity, through speculative subdivision. But perhaps, conversely, we could consider continuity as given, and look at historical scales as landmark systems arbitrarily extracted from it to satisfy our need for ease of manipulation, hierarchy, order, and comparative norms: "A continuous set is not the result of the parts into which it is divisible, but, on the contrary, is independent of them, and consequently, that it is given to us as a whole by no means implies the actual existence of these parts" (Guénon 1946, p. 60).

In any case we can claim that the intervals, and pitches themselves as embodiments of intervals, are isomorphous to real numbers, and to a straight line isomorphous to the latter: this musical parameter is continuous and totally ordered. Some other parameters possess the same structure (e.g., amplitude, except those based on qualitative definitions such as timbre). Time itself, metrical time, in the mental operations we can perform on it, exhibits the same structure deeply rooted in our psychology, as shown by Piaget (1973, pp. 74–83), where the notion of logical "grouping" is developed. Leonardo da Vinci understood much of this, which is evident from a small text from around 1500 that I quote with pleasure: "If one applies to the point the terms reserved for time, it must be compared to the instant, and the line to the length of a great duration of time. And as the points constitute the beginning and end of the said line, so the instants form the origin and the term of a certain portion of given time. And if a line is divisible to infinity, it is not impossible that a portion of time is also. And if the divided parts of the line can offer a certain proportion between them, it is the same for the parts of time" (da Vinci 1942, p. 76).

This justification may seem unnecessary to the reader, but it has not been in our musical tradition very long. Consider for instance, in 1916, the very laborious explanations of Russolo about his noise instruments (*intonarumori*) capable of "a complete enharmonic system where each tone possesses all possible mutations by subdividing in an indefinite number of fractions" (Russolo 1975, p. 83). Russolo finally introduces, in the context of a discussion on notation problems, a rather vague notion of "dynamic continuity" (Russolo 1975, p. 87) meaning pitch continuity. It is only quite recently that such notions have been accepted in musical practice. Perhaps now they are taken for granted.

In the following pages, we will stand on the musically rather abstract ground of the group of real numbers, under addition, since we will have no reason to venture out of the perceptual domain. But since we have established the correspondence of real values to musical parameters or characteristics, music will always be "virtually" present, and actual applications will be preferably left to the reader's own imagination.

2.8 Cannons

> *"In the final analysis, randomness, like beauty, is in the eye of the beholder."*—R. W. Hamming (1973)

Probability theory normally finds its practical use in statistics. We can analyze the results of tests and see if they conform to some distribution if abnormal configurations imply faults in the testing environment. Probability distributions are studied as mathematical models of populations—"natural" events or series of events happening in everyday life as well as in complex scientific experiments. But our purpose is the opposite: we wish to synthesize populations conforming to probability distributions. For this we shall need to transform the formula of probability theory in such a way that we make the random variable a function of its distribution function; this will be clarified later. We are interested in starting from a definition of a population, a histogram, chosen because of formal, aesthetic, or other musical reasons, in order to give shape to a certain musical characteristic. Then we want to synthesize values fitting it, eventually combining these values into sound events.

Nothing guarantees that our synthetic population, which will be composed of a limited number of random values, will conform with precision and elegance to the intended histogram and embody it in an ideal manner. Random values are random, nevertheless, and do not necessarily follow our imagination and intentions. Conformity of a population to a given probability distribution can only be achieved for a very large number of sample values. This is known as the *weak law of large numbers:* "If, in a trial, the probability of an event is p, and if we repeat the trial a large number of times, the ratio between the number of times the event happens and the total number of trials—that is, the frequency f of the event—tends more and more towards p. . . . if the number of trials is large enough, it becomes highly improbable that the difference between f and p will be greater than a given value, as small as it may be" (Vessereau 1953, p. 31).

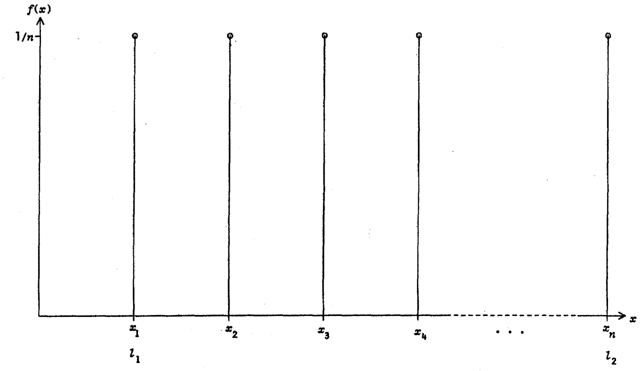
Thus we must be ready to accept some "strange" random values, or successions of values, or perhaps choose among some populations the one best embodying our platonistic Idea of the distribution. "This brings up a philosophical point. Do we really want genuine random numbers, or do we want a set of homogenized, guaranteed, and certified numbers whose effect is random but at the same time we do not run the risk of the fluctuations of a truly random source? . . . we usually find that we want to get the security of a large [number of samples] by taking [as few] as we can" (Hamming 1973, p. 143).

2.9 Notation Conventions

In the following pages we shall apply all the preceding notions and continue using the same symbols, reviewed here:

X is a random variable
x represents values of X
x_s is one specific value of X
U is the continuous uniform variable (random real number in the range [0, 1])
u represents values of U
u_s is one specific value of U, standing for a call to the U cannon in statements such as "u_s: = ran(0)"
$f(x)$ is the probability distribution of X
$P\{X = x_s\} = f(x_s)$

Fig. 8. *Histogram of the discrete uniform distribution.*



in discrete cases the probability of X takes the value x_s

$= f(x_s)dx$

in continuous cases the probability of X takes a value within the interval dx around x_s

$F(x)$ is the distribution function of X

$F(x_s) = \sum^{i \leq s} x_i = P\{X \leq x_s\}$ is the probability of X taking a value smaller or equal to x_s in discrete cases and

$$= \int_{-\infty}^{x_s} f(x) d_x = \mathrm{P}\{X \leq x_s\}$$

in continuous cases

Where a cannon consists in the application of a single formula, this formula only will be given. When complete algorithms are implied, they will be given in Pascal, a common and easily translatable procedural language.[1]

We will examine a number of different techniques for generating stochastic data according to specific formulas. First, discrete distributions will be discussed, including the discrete uniform distribution. In addition, binomial, Poisson, and permutational algorithms will be detailed. Then continuous distributions will be covered, including the uniform distribution; and a number of weighted functions, including: linear, exponential, gamma, bilateral exponential (or First Law of Laplace), Cauchy, hyperbolic cosine, logistic, arc sine, Gauss-Laplace, and beta distributions.

3.0 Types of Discrete Distributions

3.1 The Discrete Uniform Distribution

This distribution concerns n equiprobable events. It is equivalent to the random choice of one of n integers between (and including) two limits ℓ_1 and ℓ_2 ($\ell_1 < \ell_2$). The probability distribution is constant for all possible outcomes:

$$f(x) = \frac{1}{n},$$

where $n = (\ell_2 - \ell_1) + 1$, the number of possible integers in the sample set. A histogram is shown in Fig. 8. To choose an item, we could use the general method described for other discrete variables, but since all probabilities are equal the case is simple and we can directly make X a linear function of U:

$$X = nU + \ell_1,$$

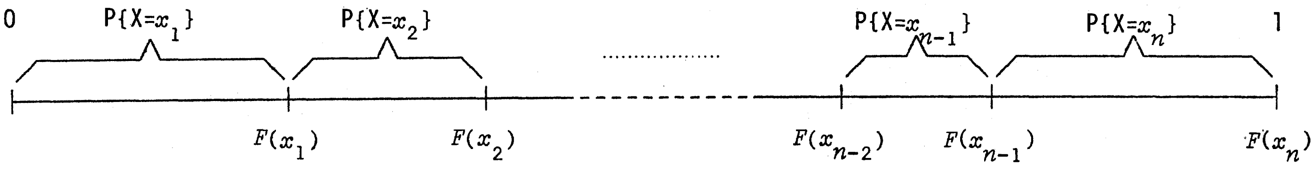
and rely on the computer's integer arithmetic to truncate this real X to integers. The preceding equation could be paraphrased as: given a real u in the interval [0, 1], we "stretch" this interval to $[0, n]$, and then transfer it to $[\ell_1, n + \ell_1]$. When truncated, the result will be an integer in $[\ell_1, \ell_2]$. Algorithm performs the task:

```
function inrect (l1,l2: integer): integer;
var n: integer; u: real;
begin
  n: = (l2 - l1) + 1;
  u: = ran(0);
  inrect: = trunc(n*u) + l1
end; {inrect}
```

Algorithm (1)

1. In the original manuscript, the author supplied algorithms written in Fortran. This code has been converted to Pascal by the editor in the interest of readability. The code is Standard Pascal with the addition of the exponentiation operator "**", and a **ran( )** function. Any errors in the code are the responsibility of the editor.

Fig. 9. $P\{X = x\}$ as segments on [0,1].



3.2 A General Method for Obtaining other Distributions

When the probabilities assigned to n discrete events are not equal, it is necessary to use a *cumulative table* of probabilities. Such a table contains the n following values:

(1) $P\{X = x_1\}$
(2) $P\{X = x_1\} + P\{X = x_2\}$
$\vdots$
(n) $P\{X = x_1\} + P\{X = x_2\} + \ldots + P\{X = x_n\}$,

which are equivalent to:

(1) $F(x_1)$
(2) $F(x_2)$
. . . .
(n) $F(x_n)$.

Since $F(x_1) > 0$ and $F(x_n) = 1$, all these values can be assimilated to points dividing a real segment [0, 1] into smaller segments equal to the different probabilities involved (Fig. 9). After obtaining a u_s from our U cannon, we need only look into which segment it falls in order to point at the chosen x_s:

$$X = x_1 \text{ if } 0 \leq u_s < F(x_1)$$
$$X = x_2 \text{ if } F(x_1) \leq u_s < F(x_2)$$
$$\vdots$$
$$X = x_n \text{ if } F(x_{n-1}) \leq u_s < 1.$$

In the following cannons, this simple method will be used: scanning $F(x)$ from $F(x_1)$ to $F(x_n)$ until the proper x_s is found. Knuth (pp. 101–102) can be consulted for other algorithms.

3.3 Other Discrete Variables

3.3.1 A Choice Between Two Alternatives

The following algorithm makes a choice between two alternatives of given (most likely different) probabilities. Only $P\{X = x_1\}$ need be specified, since $P\{X = x_2\} = 1 - P\{X = x_1\}$. The following parameters are transmitted from the calling program:

x1: the alternative x_1
x2: the second alternative x_2, and
p1: $P\{X = x_1\}$.

```
function alter2(x1,x2,px1: real): real;
begin {alter2}
  u: = ran(0);
  if u < px1 then
    alter2: = x1
    else
    alter2: = x2
end; {alter2}
```

Algorithm (2)

A particular version of the preceding cannon may be useful for choosing between +1 and −1 as equiprobable signs in equations involving a + or − alternative. Such cases are frequent (Xenakis 1963, chapter 4) in formulas of the type

$$z_i = z_{i-1} \pm x$$

where x is a random interval obtained from an always-positive random distribution. The preceding equation could be programmed:

z(i):= z(i − 1) + (xsigne(0)**call to a cannon)*

with the following:

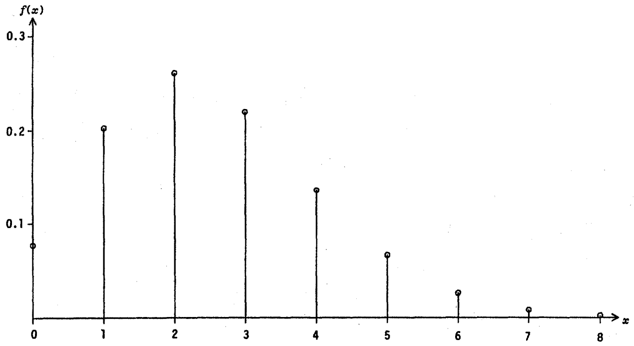


Fig. 10. Histogram of the binomial distribution B(50,0.05).

```
function xsigne(ibidon: real): real;
begin {xsigne}
  u: = ran(0);
  if u < .5 then xsigne: = 1.0
  else
  xsigne: = -1.0
end; {xsigne}
```

Algorithm (3)

3.3.2 The Binomial Distribution

The choices between two alternatives described in Section 3.3.1 are called *Bernouilli trials.* A succession of n Bernouilli trials synthesizes a population conforming to the binomial distribution:

$$B(n,p)$$

where p is the probability of one of the alternatives, labeled *success.* This distribution looks at a succession of Bernouilli trials, and has as variable the number of successes obtained in n trials:

$$P\{B(n_1p) = x\} = \binom{n}{x} p^x (1 - p)^{n-x}$$
$$\text{for } x = 0, 1, 2, \ldots, n$$

where

$$\binom{n}{x} = \frac{n!}{x!(n - x)!}$$

It is thus an analytic view of choices between two alternatives.

It is perhaps not necessary to devise a binomial cannon, since we can seize the problem by straightforward choices rather than by resorting to analytic distribution. But if such an algorithm were needed, it should be similar to that of the Poisson distribution (Section 3.3.3). First, a cumulative table $F(x)$ is prepared:

$$F(0) = P\{B(n_1p) = 0\}$$
$$F(1) = P\{B(n_1p) = 0\} + P\{B(n_1p) = 1\}$$
$$\vdots$$
$$F(n) = P\{B(n_1p) = 0\} + P\{B(n_1p) = 1\} + \ldots + P\{B(n_1p) = n\} = 1$$

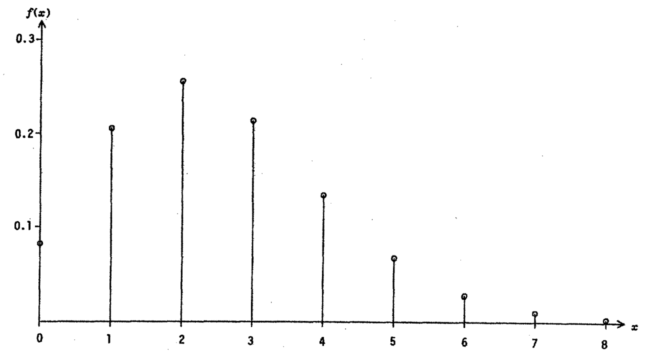
Then u_s is used to point at the resulting x_s. We then decide that among the next n events there should be x_s successes, and of course $n - x_x$ occurrences of the other alternative. Another solution would be to use the Poisson cannon directly, since Poisson distributions are approximations to binomial distributions. A histogram of $B(50, 0.50)$ is shown in Fig. 10.

It can be shown (Calot 1967, pp. 310–316; Feller 1966*a*, Chapter 7) that when n becomes large and p is not small, for example, when np and $n(1 - p)$ are somewhat greater than 15 or 20, with better approximations when p is near ½, the binomial distribution $B(n,p)$ is quite equivalent to a Gauss-Laplace distribution of mean np and standard deviation $\sqrt{np(1 - p)}$. In such cases we could more easily use the Gauss-Laplace cannon (Section 4) and round the result to the nearest integer. Calot (1967, pp. 445–449) shows interesting graphical comparisons between binomial and Gauss-Laplace histograms, while Vessereau (1953, chapter 3) presents a good introduction to the family of binomial, Gauss-Laplace and Poisson distributions. Gnedenko and Khintchine (1969), treating exclusively discrete variables, also approach the Gauss-Laplace distribution through the binomial.

3.3.3 The Poisson Distribution

On the other hand, if n becomes large and p small, it can be shown (Calot 1967, pp. 307–309; Feller 1966*a*, p. 153) that the binomial distribution can be approximated by the Poisson distribution, which has the advantage of discarding n from the calculations. This approximation is good for values of:

Fig. 11. Histogram of the Poisson distribution for $\delta = 2.5$, quite similar to the binomial $B(50,0.05)$ of Fig. 10.



$$n > 50 \text{ and}$$
$$p < 0.1,$$

such that the product $np = \delta$ is in the order of a few units. This parameter δ is called the average density of the Poisson distribution. This relation between Poisson and binomial distributions implies that, for a large δ (somewhat greater than 20), Poisson also approximates a Gauss-Laplace distribution of mean δ and standard deviation $\sqrt{\delta}$ (Calot 1967, pp. 330–331; Feller 1966*a*, p. 190). A Poisson algorithm could, under such conditions, be used as an integer Gauss-Laplace cannon. However, the Poisson distribution is quite cumbersome to program, so this is not very advantageous.

The Poisson distribution is thus, as the binomial, an analytic point of view on a series of Bernouilli trials: it gives the probability of obtaining x successes in a large number of trials, when the average density of successes is δ:

$$P\{X = x\} = \frac{e^{-\delta}\delta^x}{x!}, \quad \text{for } x = 0, 1, 2, \ldots.$$

Note that n, the number of trials, is not present in the formula. A histogram of a Poisson distribution of density 2.5 is shown in Fig. 11.

We are giving an algorithm for this distribution because it has attained some popularity through its use by Xenakis (1963, p. 35) as an analytic approach to phenomena appearing in the time dimension (*cf.* also Truax, in *Computer Music Journal,* Vol. 1, No 3 (1977) pp. 30–39). In such cases, the Poisson distribution concerns the probability of finding x points per time unit on a time axis, when the average density of points per time unit is δ (small enough) and the distribution of points conforms to a homogeneous exponential distribution (Section 4.3.2). It is as if, from a binomial point of view, we regarded time as a succession of a very large number of small intervals; a success would then be the presence of a point occupying an interval. Our binomial-Gauss-Laplace-Poisson family, which includes in addition a negative-binomial distribution (Calot 1967, p. 321; Feller 1966*a*, p. 164) that we do not study here, is thus also related to gamma distributions, of which the exponential is a member (Calot 1967, pp. 332–337, 354–356; Feller 1966*b*, pp. 11–12). ". . . the remarkable fact that there exists a few distributions of great universality which occur in a surprisingly great variety of problems. The three principal distributions, with ramifications throughout probability theory, are the binomial, the [Gauss-Laplace] . . . , and the Poisson distribution" (Feller 1966*a*, p. 156).

The following algorithm is devised to handle several Poisson distributions of independent parameters:

```
procedure poissoninit (var i, j, n, itot, nmax: integer;
                        val, d: real);
type
  realarray = array [1..itot,1..nmax] of real;
var
  xfac: real; xk: integer; tab: realarray;
begin {poissoninit}
  xfac:=1;
  for j:=1 to n-1 do
  begin {loop}
    xk:=j-1;
    if xk<=1 then
      val:=((d**xk)/xfac)*exp(-d)
    else
      begin
      xfac:=xfac*xk; val:=((d**xk)/xfac)*exp(-d)
      end
    if xk = 0 then
      tab[i,1]:=val
    else
      tab[i,j]:=tab[i,(j-1)+val]
  end; {loop}
```

```
  tab[i,n]:=1.0;
end; {poissoninit}
```

Algorithm (4)

```
procedure poisson(var i,n,nbr,itot,nmax: integer);
var u: real;
begin {poisson}
  u:=ran(0);
  for j:=1 to n do
  begin {loop}
    if u<=tab[i,j] then
    begin
      nbr:=j-1;      {transmit value}
      j:=n+1         {terminate loop}
    end
  end {loop}
end; {poisson}
```

Algorithm (5)

The statement "poissoninit" begins a section of initialization of the tab tables of the different $F(x)$ required. Tab has two dimensions (itot,nmax) and is considered as itot vectors, each containing the $F(x)$ of the i-th Poisson distribution of a given δ_i:

$$\text{tab}(i,n) = \sum_{x=0}^{n-1} \frac{e^{-\delta_i}\delta_i^x}{x!}$$

Such vectors need more or less length, according to the density δ_i. One must figure nmax such that, for δ_{max}, the greatest density used,

$$P\{X = \text{nmax}\} = \frac{e^{-\delta_{max}}\,\delta^{\text{nmax}}}{\text{nmax}!}$$

is very small and can be neglected: indeed, the algorithm can never output a result greater than nmax − 1.

A call to poissoninit includes the following parameters:

- i:the reference number for the distribution with density δ_i
- n(≤nmax):the effectively used length of the tab vector for the i-th density, such that $P\{X = N\}$ can be neglected
- d:the average density δ_i
- tab:the name of the array declared in the main program
- itot:the total number of different densities used
- nmax:the length of the longest tab vector

The calling program must call "poissoninit" with the proper parameters in order to initialize all tables. After this preparation, the calls are made to proceed with "poisson." There we get a u_s from **ran**(0) and use it to point at an x_s. A new parameter is included in the call:

> nbr transmits x_s, the result, to the calling program: there should be nbr events.

All other parameters are as above.

3.3.4 A Choice Between Several Alternatives

If the sample space contains more than two events of given different probabilities, the following algorithm is needed. It simply scans the cumulative table of the probabilities in order to designate a resulting x_s by a u_s.

```
type
  realarray = array[1 . . n] of real;
:
function altern(var boul, prob: realarray):real
var u, som:real;
begin{altern}
  som:=0; u:=ran(0); i:=1;
  while i <=n do
  begin {loop}
    som:=som + prob[i];
    if u >= som then
    begin {assign values and exit}
  altern := boul[i];
      i:=n {exit loop}
    end; {assign}
 i:=i+1
  end{loop}
 {not yet successful, assign last value}
 altern:=boul[n]
end;{altern}
```

Algorithm (6)

The parameters of the call are

boul: the name of the calling program's vector containing the list of different alternatives:

$$\text{boul(i)} = x_i$$

prob: the name of the calling program's vector containing the list of different probabilities, such as
prob(i) $= P\{X = x_i\}$ and

$$\sum_{I=1}^{N} \text{prob(i)} = 1$$

n: the dimension of boul and prob.

altern returns the chosen boul(i) $= x_s$.

The preceding algorithms were particular cases of altern: inrect when all probabilities are equal, and alter2 when only two alternatives are valid.

The multinomial distribution is the analytic approach of this problem (Calot 1967, p. 337; Feller 1966*a*, p. 167). An algorithm using the multinomial distribution would be very cumbersome, since it should handle whole vectors of events: in n events there should be x_{s1} events x_1, x_{s2} events x_2, and so on. It would prove quite impractical in face of the simplicity of the direct synthesis of the events.

3.3.5 Exhaustive Trials, Permutations

A typically serial procedure would be to choose events from a sample space without repetition of any event before exhaustion of the sample space (Koenig 1971, pp. 15–17, 32–33), thus achieving a permutation of the events of the sample space. This idea is of course valid even if some of the available events are identical—twelve-tone rows are but a particular case: the classical analogy is that of picking balls of different colors from an urn without putting them back. The initial configuration is of N balls, n_2 of color x_1, n_2 of color x_2, and so on, so that for the first pick

$$P\{X = x_s\} = \frac{n_s}{N}.$$

Of course, after the first ball, of color x_s, is out of the urn,

$$P\{X = x_s\} \text{ becomes } \frac{n_s - 1}{N - 1},$$

and so forth (Gnedenko and Khintchine 1969, pp. 103–114).

There is an analytic approach to this problem: the *hypergeometric* distribution (Calot 1967, p. 316; Feller 1966*a*, pp. 43–47), which would be of little use here. A synthetic approach would be to modify the above altern algorithm in order for it to adjust the prob list of probabilities after the result of each successive choice, or to have the main program handle these modifications. But since most likely the sample space will be of a rather small number of events, it is more practical to use the permutation algorithm, Algorithm 7 (Koenig, pp. 70–73) and to read the permuted array one item at a time as the result of an exhaustive series of trials.

```
type
  integerarray=array[1 . . n] of integer;
  integer2array=array[1 . . 2, . . n] of integer;
:
.
procedure permut(varlensmb:integerarray; itrav:
                         integer2array;n,iopt:integer);
var
  ic, i, ix, ir:integer;
begin {permut}
  ic:=n;
  for i:=1 to n do
  itrav[1,i]:=i;
  repeat
    ix:=inrect(1,ix); {call to function inrect (cf.3.1)}
    ir:=itrav[1,ix];
    itrav[2,ic]:=lensmb[ir];
    itrav[1,ix]:=itrav[1,ic];
    ic:=ic-1;
  until ic=0;
  if iopt=1 then
  begin
    for i:=1 to n do
    lensmb[i]:=itrav[2,i]
  end
end;
```

Algorithm (7)

The calling program should have declared the two arrays lensmb [0 . . n] and itrav[0 . . 2, 0 . . n]. In the calling sequence,

lensmb is the name of the set of events (array) to be permuted
itrav is a working array
n is the dimension of lensmb and the second dimension of itrav
iopt is a switch:
iopt = 1 $\Rightarrow$ the permutation is returned in lensmb; that is, lensmb is modified
iopt $\neq$ 1 $\Rightarrow$ lensmb is not modified, and the calling program finds the permuted events only in itrav[2,j].

In the algorithm, ic is a counter, itrav[1,j] is used to prevent repetitions of the same event, and itrav[2,j] stores the permutation while it is prepared. Permut uses the inrect cannon (Section 3.1). A calling program may well use several lensmb sets, but only one work array itrav[2,n_{max}] is necessary, where n_{max} is the dimension of the largest lensmb.

4.0 Types of Continuous Distributions

4.1 Uniform Distributions

Sol per te le mie ore son generate.
—Leonardo da Vinci

As we have seen, the continuous uniform distribution in the range [0,1] is the basis of all our algorithms, source of all randomness. Of course, continuity must be understood as limited to the precision of the computer words; but current precisions of eight or more decimal digits are quite sufficient for our purposes.

It is taken for granted that a typical U cannon (or pseudo-random number generator) is available; otherwise, one can be programmed with some knowledge or found in the literature. U cannon algorithms are discussed from a general point of view by Hamming (1973, pp. 136–142), Knuth (pp. 1–100), and Maurin (1975, chapter 2). Practical examples can be found in IBM's Scientific Subroutine Package (p. 77) (for 32-bit words), and in Koenig's *Computer Composition* (p. 69) (similar, but for 27(!)-bit words). A complete Fortran algorithm that is not machine-dependent can be found in Schrage's article. We must be aware of the fact that, being the *deus ex machina* of all stochastic applications, the U cannon deserves the utmost care. Particularly in cases making use of large quantities of source random numbers, only reliable and tested algorithms should be used for the endeavor to be of some consistency.

It is remarkable that, in accordance with the very concept of the algorithm, these random generators are rigorously deterministic in generating each random value through some manipulations of the preceding random value:

$$u_s \rightarrow \quad \text{black box} \quad \rightarrow u_{s+1}.$$

They embody exactly the old epicurian concept of randomness as a subjective point of view on a complex combination of deterministic causalities standing out of our comprehension: "Randomness is an uncertain cause with regard to persons, times and places" (Aetius 1976, p. 72).

In order to obtain a continuous distribution in another range than [0,1], we use a linear function of U, exactly as for the inrect algorithm (Section 3.1), but without truncation:

$$x_s = (r_2 - r_1)u_2 + r_1,$$

where r_1 and r_2 are two real limits ($r_1 < r_2$), generates a *rectangular continuous* distribution in $[r_1,r_2]$.

4.2 A General Method for Obtaining Other Distributions

In order to generate a continuous variable X from the uniform variable U, the general method is the following. We can equate the distribution function of U (Eq. 7, Section 2.6)

$$F(u_s) = u_s,$$

to the $F(x)$ of the desired distribution:

$$u_s = F(x_s).$$

We then have u_s as a function of x_s, and this can be algebraically inverted into x_s as a function of u_s. Thus u_s is transformed into an x_s conforming to the desired distribution:

$$x_s - F^{-1}(u_s)$$

(Knuth, pp. 102–103; Maurin 1975, chapter 3;

Hamming 1973, pp. 142–143; Calot 1967, chapter 15). This can be intuitively understood as using the u_s in [0,1] to designate a fraction of the area under the desired probability distribution, and looking at the corresponding x_s (see Figure 4). This is simply a more general, continuous application of the discrete method described in Section 3.2.

However, some probability distributions have distribution functions $F(x)$ which cannot be integrated in order to be handled in this manner. Such cases need special algorithms; two (the Gauss-Laplace and beta distributions) will be studied in Section 4.4.

4.3 Directly Obtainable Continuous Distributions

4.3.1 The Linear Distribution

This distribution is described by Xenakis for generating intervals of pitch, time, *etc.* (1963, p. 27, 219) as the probability of picking a segment of length x from a line of length g when the two points delimiting the chosen segment are randomly designated (with rectangular distribution) on the line $[0,g]$:

$$P\{X = x\} = \frac{2}{g}\left(1 - \frac{x}{g}\right)dx \quad (0 \leq x \leq g).$$

Its histogram is shown in Fig. 12. This distribution is equivalent to the right half of the triangular distribution described by Feller (1966*b*, p. 50).

The cannon is prepared thus:

$$F(x_s) = \int_0^{x_s} \frac{2}{g}\left(1 - \frac{x}{g}\right)dx$$

$$= \frac{2x_s}{g} = \frac{x_s^2}{g}.$$

By equalizing with $F(u_s) = u_s$ we arrive at the quadratic equation

$$\frac{x_s^2}{g^2} - \frac{2x_s}{g} + u_s = 0,$$

which has two roots:

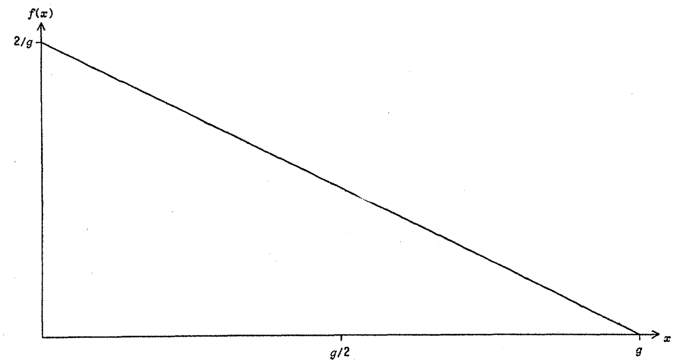


Fig. 12. Histogram of the linear distribution of parameter g.

$$x_s = g(1 \pm \sqrt{1 - u_s}),$$

of which only

$$x_s = g(1 - \sqrt{1 - u_s})$$

is useful here, since the other would always make $x_s > g$, and we must stay in the interval $0 \leq x_s \leq g$. The term $(1 - u_s)$ being merely a uniform random number in [0,1], symmetric to u_s, we can simplify finally to:

$$x_s = g(1 - \sqrt{u_s}).$$

As visible on the histogram (Fig. 12), small values are favored, and this because of the relation:

$$\sqrt{u_s} > u_s,$$

which tends to make $1 - \sqrt{u_s}$ always nearer zero.

4.3.2 The Exponential Distribution

As mentioned in Section 3.3.3. about the Poisson distribution, the exponential distribution has been used by Xenakis (1963, pp. 26, 169, 171, 215) to synthesize stochastic time intervals. Its theoretical basis makes it suitable for homogeneous situations (Feller 1966*b*, p. 8). From a musical point of view, the exponential distribution is particularly suited for time intervals, being at once homogeneous and sufficiently varied. But of course others can be used to introduce variations or radically different configurations (Section 5).

The exponential distribution of average density $\delta > 0$ (average interval $= 1/\delta$) is

$$P\{X = x\} = \delta e^{-\delta x}dx \ (x \geq 0).$$

Fig. 13. *Histograms of exponential distributions for different values of δ.*

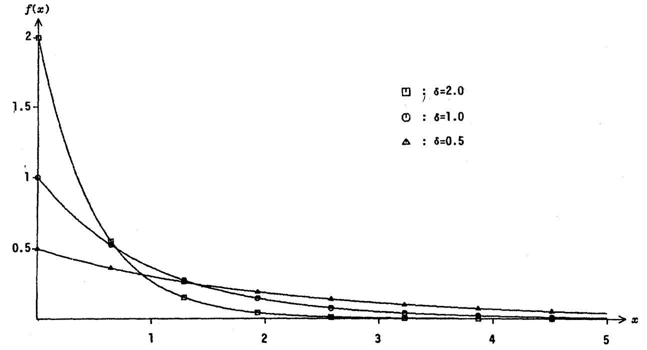
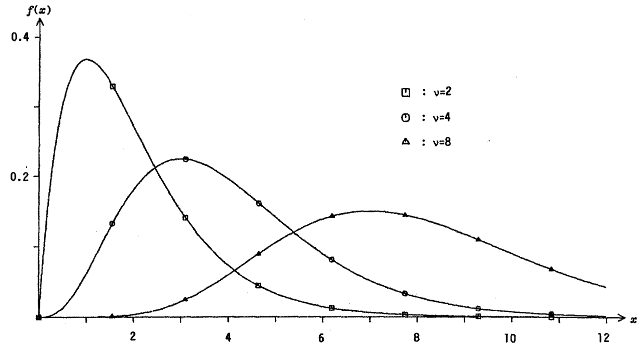


Fig. 14. *Histograms of gamma distributions for different values of ν.*



The distribution function is

$$F(x_s) = \int_0^{x_s} \delta e^{-\delta x} dx = 1 - e^{-\delta x_s}.$$

Through equalization with $F(u_s)$, we have

$$u_s = 1 - e^{-\delta x_s},$$

inversed into

$$x_s = \frac{-\ln(1 - u_s)}{\delta},$$

and finally:

$$x_s = \frac{-\ln(u_s)}{\delta}.$$

Histograms of exponential distributions for three different densities are shown in Fig. 13.

4.3.3 Gamma Distributions

The gamma distribution γ_ν, of parameter $\nu > 0$, is

$$P\{X = x\} = \frac{1}{\Gamma(\nu)} e^{-x} x^{\nu-1} dx \ (x \geq 0),$$

where $\Gamma(\nu)$ is the Eulerian *gamma function*

$$\Gamma(\nu) = \int_0^{\infty} e^{-x} x^{\nu-1} dx.$$

The exponential distribution of density 1 (e^{-x}) is a particular case of the gamma distribution for $\nu = 1$. Histograms of gamma distributions for three different ν are shown in Fig. 14. They indicate the particular interest of the distribution, its asymmetry. The highest probabilities (*mode,* in statistical terminology) are for values near $\nu - 1$, whereas the mean is ν (Calot 1967, p. 350). In rhythmical applications, it can produce a sort of *easing* or *rubato* compared to the exponential (see Section 5). It is quite meaningless to use values of ν much greater than 10: as ν increases, the gamma distribution tends toward a Gauss-Laplace distribution (Calot 1967, pp. 353–354).

A gamma cannon formula allowing any real value for ν cannot be arrived at directly. If necessary, an algorithm for any real ν can be found in Jöhnk's article (pp. 13–15), but it is not so simple. It requires the gamma cannon for integer ν parameters described here, plus a beta cannon (see Section 4.4.2). Otherwise, an algorithm analogous to the first beta cannon described here and in Section 4.4.2 would be necessary. But it can be shown (Calot 1967, pp. 200–201) that for gamma variables of parameters ν and ω, the following equality holds:

$$\gamma_v + \gamma_w = \gamma_{v+w}.$$

Provided we are satisfied with integer valued ν parameters, we can thus synthesize values of γ_ν by summation of ν independent γ_1 variables:

$$\gamma_\nu = \sum_{i=1}^{\nu} \gamma_1.$$

This restriction is not drastic, since we can always afterward multiply the γ_ν variable obtained by some factor τ in order to set the mode at $\tau(\nu - 1)$ and the mean at $\tau\nu$, or even devise a linear function $\tau\gamma_\nu + \mu$ (Calot 1967, p. 352).

The cannon formula for γ_1 (exponential with $\delta = 1$) being

$$x_s = -\ln(u_s),$$

we need only add successive x_s's ν times in a loop. Since the following equality holds:

$$\ln a + \ln b = \ln(ab),$$

the algorithm can avoid calling the **alog** function ν times:

```
function gamma(nu:real):real;
var sum:real;
begin {gamma}
  sum:=1.0;
  for i:=1 to nu do
    sum:=sum*ran(0);
  gamma:=-alog(sum);
end;{gamma}
```

Algorithm (8)

This method is described by Knuth (p. 115) for the chi-square distribution, closely related to the gamma, and by Jöhnk (pp. 11–12).

4.3.4 First Law of Laplace

The *bilateral exponential* distribution (Feller 1966*b*, p. 49) is described by Calot (1967, pp. 356–357) as the First Law of Laplace (I adopt Calot's less technical term out of personal admiration for this mathematician). The First Law of Laplace is defined as:

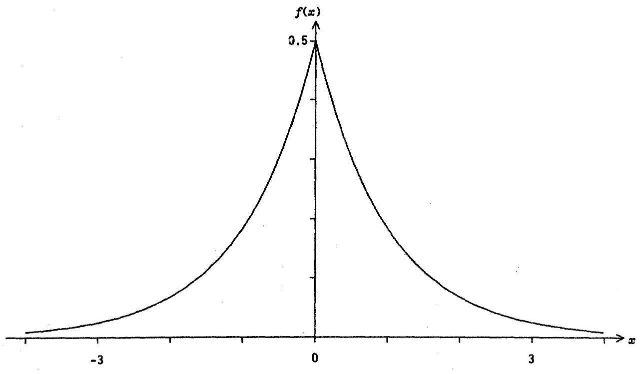
$$P\{X = x\} = \frac{e^{-|x|}}{2}\,dx \qquad -\infty < x < \infty.$$

A histogram of this distribution is shown in Fig. 15. It is a symmetrical distribution around a mean, but of quite a different shape than the more common normal distribution (Section 4.4.1).

The distribution function is as follows:

$$F(x_s) = \begin{cases} \dfrac{e^{x_s}}{2} & \text{for } x \leqslant 0 \\[2ex] 1 - \dfrac{e^{-x_s}}{2} & \text{for } x \geqslant 0. \end{cases}$$

Fig. 15. Histogram of the First Law of Laplace (with our parameters $\tau = 1$ and $\mu = 0$).



By equalizing with u_s and splitting the outcomes as

$$x_s < 0 \text{ if } u_s < 0.5 \text{ and}$$
$$x_s > 0 \text{ if } u_s > 0.5,$$

the two following formulas perform the distribution:

$$u_s < 0.5 \Rightarrow x_s = \ln(2u_s)$$
$$u_s > 0.5 \Rightarrow x_s = -\ln(2 - 2u_s).$$

In the following algorithm, we have introduced two parameters in order to make it directly generate a linear function of X:

$$\text{plapla(xmu,tau)} = \tau X + \mu.$$

μ can be a mean other than zero, and τ is a parameter controlling the dispersion (range, or "horizontal spread") of the distribution. Thus our algorithm generates the following random variable:

$$P\{X = \mu + x\} = \frac{1}{2\tau}\, e^{\frac{-|x|}{\tau}} dx, \text{ for any } x.$$

```
function plapla(var xmu,tau: real):real;
var
  u: real
begin {plapla}
  u: =ran(0)*2.0;
  if u > 1.0 then
    begin
      u:=2.0-u;
      plapla:=(-tau*alog(u))+xmu
    end
    else
```

```
    plapla:=(tau*alog(u))+xmu
end;{plapla}
```

Algorithm (9)

4.3.5 Cauchy Distribution

The Cauchy distribution is another symmetrical distribution centered around zero, but it has the particularity of having no mean and generating very heterogeneous values. The histogram of the Cauchy distribution (Fig. 16) is similar to that of the Gauss-Laplace distribution, but it approaches the horizontal axis so slowly that values of X extremely distant from the mode are quite probable. No finite mean can be computed; this may be difficult to accept intuitively, but the mathematical definition of the mean, or expectation, of a distribution can make it clear (Calot 1967, pp. 213; Feller 1966*b* pp. 117–118).

With a parameter $\tau > 0$ scaling the dispersion of the variable, the Cauchy distribution is

$$P\{X = x\} = \frac{\tau}{\pi(\tau^2 + x^2)}\,dx \quad \text{for any } x.$$

The distribution function is

$$F(x_s) = \int_{-\infty}^{x_s} \frac{\tau}{\pi(\tau^2 + x^2)}\,dx = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\frac{x_s}{\tau},$$

from which we obtain the cannon formula:

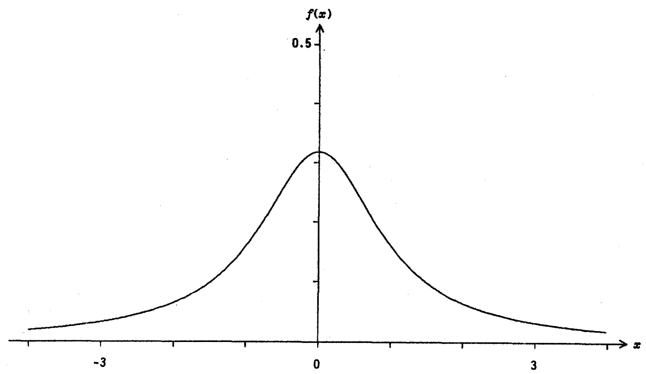
$$x_s = \tau\left\{\tan\left[\pi\left(u_s - \frac{1}{2}\right)\right]\right\}.$$

The cannon works by taking the tangent of an angle between $-\pi/2$ and $\pi/2$. In order to optimize our algorithm, we may as well take the tangent of an angle between 0 and π, thus implementing this final formula (for a call with iopt $\neq$ 1):

$$x_s = \tau[\tan(\pi u_s)].$$

This formula makes intuitively tangible the absence of mean of the Cauchy distribution: we know that the tangent of an angle near $\pi/2$ is extremely large; since πu_s can designate such an angle with as much probability as any other, very large values of X are quite probable, and can bring out of balance the distribution's eventual tendency toward a mean.

Fig. 16. Histogram of the Cauchy distribution (τ = 1, IOPT $\neq$ 1).



A modified version of the Cauchy distribution is generated when the call parameter iopt is equal to 1: only positive values are generated, producing, in fact, the following distribution:

$$P\{X = x\} = \frac{2\tau}{\pi(\tau^2 + x^2)}\,dx \text{ for } x \geqslant 0.$$

This version has at least a definite rhythmic interest: it can generate very dense but irregular aggregates of time points, separated by relatively enormous intervals (see Section 5). For the composer, it may involve the risk of losing some listeners because of a gap of a few days between two successive sounds, but it can very well be used with some poetic precautions! Nevertheless, the symmetrical version has been used by Xenakis (1971*b*, chapter 9) as a source of radical dissymmetries in microcomposition applications (stochastic sound wave forms).

```
function cauchy(var tau,:real; iopt: integer):real;
constant
  pi=3.1415927;
var
  u:real;
begin {cauchy}
  u:=ran(0);
  if iopt=1 then
    u:=u/2.0;
  u:=pi*u;
  cauchy:=tau*(sin(u)/cos(u))
end:{cauchy}
```

Algorithm (10)

Fig. 17. Histogram of the hyperbolic cosine distribution.

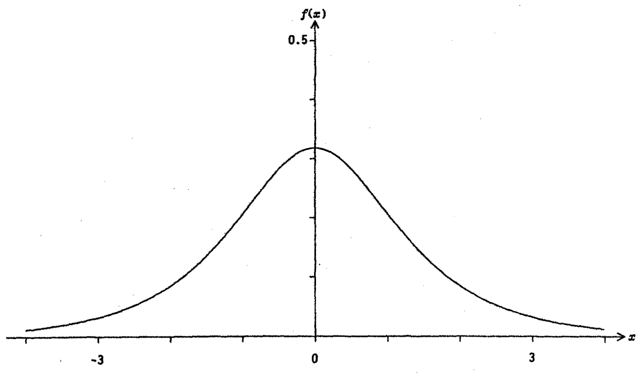
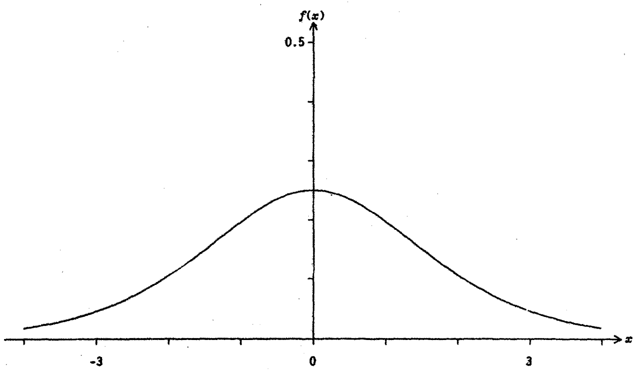


Fig. 18. Histogram of the logistic distribution for $\beta = 0$ (mean 0) and $\alpha = 1$.



4.3.6 The Hyperbolic Cosine Distribution

This is another symmetrical distribution which can be of some use. Figure 17 shows its histogram. As with the Cauchy distribution, although it is centered on zero, this distribution has no mean. It is defined as:

$$P\{X = x\} = \frac{1}{\pi \cosh x}\, dx \text{ for any } x,$$

$$\text{where } \cosh = \frac{e^{x} + e^{-x}}{2}.$$

The distribution function is:

$$F(x_s) = \frac{1}{\pi}\int_{-\infty}^{x_s} \frac{1}{\cosh x}\, dx$$

$$= \frac{2}{\pi}\tan^{-1}(e^{x_s}),$$

we can easily devise a cannon using the formula

$$x_s = \ln\left[\tan\left(\frac{\pi u_s}{2}\right)\right].$$

4.3.7 The Logistic Distribution

With two parameters α and β, the logistic distribution is defined as:

$$P\{X = x\} = \frac{\alpha e^{(-\alpha x-\beta)}}{(1 + e^{(-\alpha x-\beta)})^2}\, dx \text{ for any } x \text{ or } \beta, \text{ and } \alpha > 0.$$

Figure 18 shows its histogram: still another symmetrical distribution, of mean $-\beta/\alpha$ and mode:

$$f(-\beta/\alpha) = \alpha/4.$$

The two parameters α and β thus control the mean and dispersion of the distribution (dispersion is inversely proportional to α). The distribution function (Feller 1966*b*, p. 52) is

$$F(x_s) = \frac{1}{1 + e^{(-\alpha x_s-\beta)}},$$

which can be inverted into the cannon formula

$$x_s = \frac{-\beta - \ln(1/u_s - 1)}{\alpha}.$$

Both the hyperbolic cosine and the logistic distributions have been used by Xenakis in microcomposition (sound synthesis) (1971*b*, chapter 9).

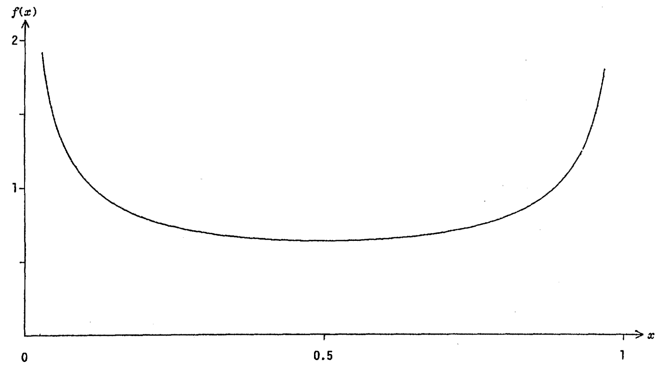
4.3.8 The Arc Sine Distribution

This distribution is the same as the $\beta(0.5,0.5)$ distribution (see Section 4.4.2), but it is still useful since it involves a simpler cannon. The histogram is shown in Fig. 19. The arc sine distribution is defined for $0 < X < 1$: the generated values will be in the interval $[0,1]$, with stronger probabilities for values near 0 and 1. As well as for the uniform variable, a random variable can be made a linear function of the arc sine distribution A:

$$X = (r_2 - r_2)A + r_1$$

in order to cover an interval $[r_1,r_2]$. The distribution is

Fig. 19. Histogram of the arc sine distribution.



$$P\{X = x\} = \frac{1}{\pi\sqrt{x(1-x)}}\,dx \text{ for } 0 < x < 1.$$

The distribution function is:

$$F(x_s) = \int_0^{x_s} \frac{1}{\pi\sqrt{x(1-x)}}\,dx$$

$$= \frac{1}{2} - \frac{1}{\pi}\sin^{-1}(1 - 2x_s),$$

which gives the following cannon:

$$x_s = \frac{1 - \sin[\pi(u_s - ½)]}{2}. \quad (8)$$

Jacquard (1976, p. 50), gives us another formula:

$$F(x_s) = \frac{2}{\pi}\sin^{-1}\sqrt{x_s},$$

which amounts to the same thing. This second formula yields the cannon

$$x_s = \left[\sin\left(\frac{\pi u_s}{2}\right)\right]^2. \quad (9)$$

An algorithm using Eq. (9) would involve three multiplications, against which Eq. (8), with two multiplications and two subtractions, can save a few microseconds—life is so short!

4.4 Two Further Distributions

In the following cases, the distributions do not yield a simple cannon formula, and we must rely on approximation algorithms or indirect methods.

4.4.1 Gauss-Laplace Distribution

Strictly speaking, Calot (1967, p. 362) restricts the denomination *normal distribution* to the function of mean 0 and standard deviation 1, defined as

$$P\{X = x\} = \frac{1}{\sqrt{2\pi}}\,e^{\frac{-x^2}{2}}dx \text{ for any } x,$$

and names *Gauss-Laplace* the concrete applications of the normal distribution with any mean μ and standard deviation σ, which can be seen either as a linear function of a normal variable:

$$\textit{Gauss-Laplace} = \sigma(\textit{normal}) + \mu,$$

or as the complete distribution

$$P\{X = x\} = \frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \text{ for any } x.$$

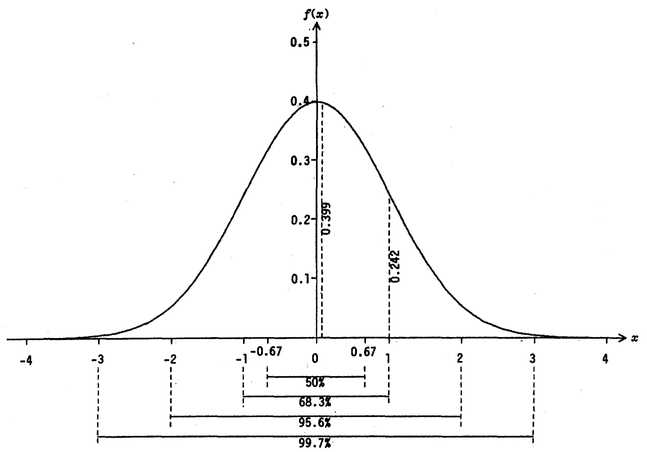
I have adopted this terminology in the present discussion.

The theoretical importance of the normal distribution has been mentioned (Section 3.3.3). It is the classical bell-shaped distribution shown in Figure 3. Statistically, it accounts for a great number of phenomena because of its very fundamental theoretical basis and role, which explain its "natural look": relatively strong probabilities for the mean and near the mean, rounded mode, smooth vanishing of the probabilities for far fetched values, and so forth. Figure 20 analyzes some characteristics of this important distribution. Note for instance that the probability of values outside the range $[-2\sigma,2\sigma]$ is ≈ 0.04, and outside of $[-3\sigma,3\sigma]$ it is less than 0.003. Tables of values for $f(x)$ and $F(x)$ can be found in any book on probabilities.

In spite of its nobility, we may use the Gauss-Laplace distribution freely. For instance, Xenakis has used it for durations (1963 p. 174) and for his well-known glissandi textures (*Pithoprakta* (1955–56), in particular (1963, p. 27)), through an analogy with the Maxwell-Boltzmann distribution for three-dimensional velocities of molecules in a gas (see also Feller, 1966*b*, pp. 29–32).

Because of mathematical properties (it is not possible to get the integral out of $F(x)$, so to speak), no simple cannon formula can be programmed. The Gauss-Laplace distribution function has to be approximated. The algorithm described here is

Fig. 20. Histogram of the normal distribution ($\mu = 0$, $\sigma = 1$). Some important characteristics are shown. Indicated percentages are of the area under $f(x)$ (Feller 1966*a*, p. 178).



sketched in Hamming's book (1973, p. 143), and developed in IBM's Scientific Subroutine Package (p. 77). It is based on the following formula (from the *central limit theorem;* see *Les Probabilitiés* [Jacquard 1976, pp. 91–92], for instance, where it is expressed in a form similar to the following):

$$X = \frac{\sum_{i=1}^{k} u_i - \frac{k}{2}}{\sqrt{k/12}}$$

where the u_i are, as usual, uniform values in [0,1] and X approaches a truly normal distribution as k approaches infinity. A compromise is to use $k = 12$: the algorithm then simplifies into

$$X = \sum_{i=1}^{12} u_i - 6.$$

Adjustment to the required μ and σ are done as

$$\text{gauss} = \sigma X + \mu.$$

```
function gauss(var xmu, sigma:real):real;
var
  s:real;
  i:integer;
begin {gauss}
 s:=0.0;
  for i:=1 to 12 do
     s:=s+ran(0);
  gauss:=(sigma*(s-6.0))+xmu;
end;{gauss}
```

Algorithm (11)

The efficiency of this algorithm is poor: twelve calls to the U cannon are required for each generation; but its simplicity is quite appealing. In cases where execution speed is critical, it may be preferable to use tables of the distribution function $F(x)$, in a procedure similar to the first algorithm described in Section 4.4.2 for beta distributions. Such tables can be computed in the algorithm, of course, or given as data from manuals (Calot 1967, p. 455; Feller 1966*a*, p. 176). Methods more effective than straightforward scanning of these tables can be devised; we can rely more on linear interpolation in order to reduce the size of the table and searching time.

4.4.2 Beta Distributions

The beta distribution is defined with two positive parameters a and b, for a random variable in [0,1]:

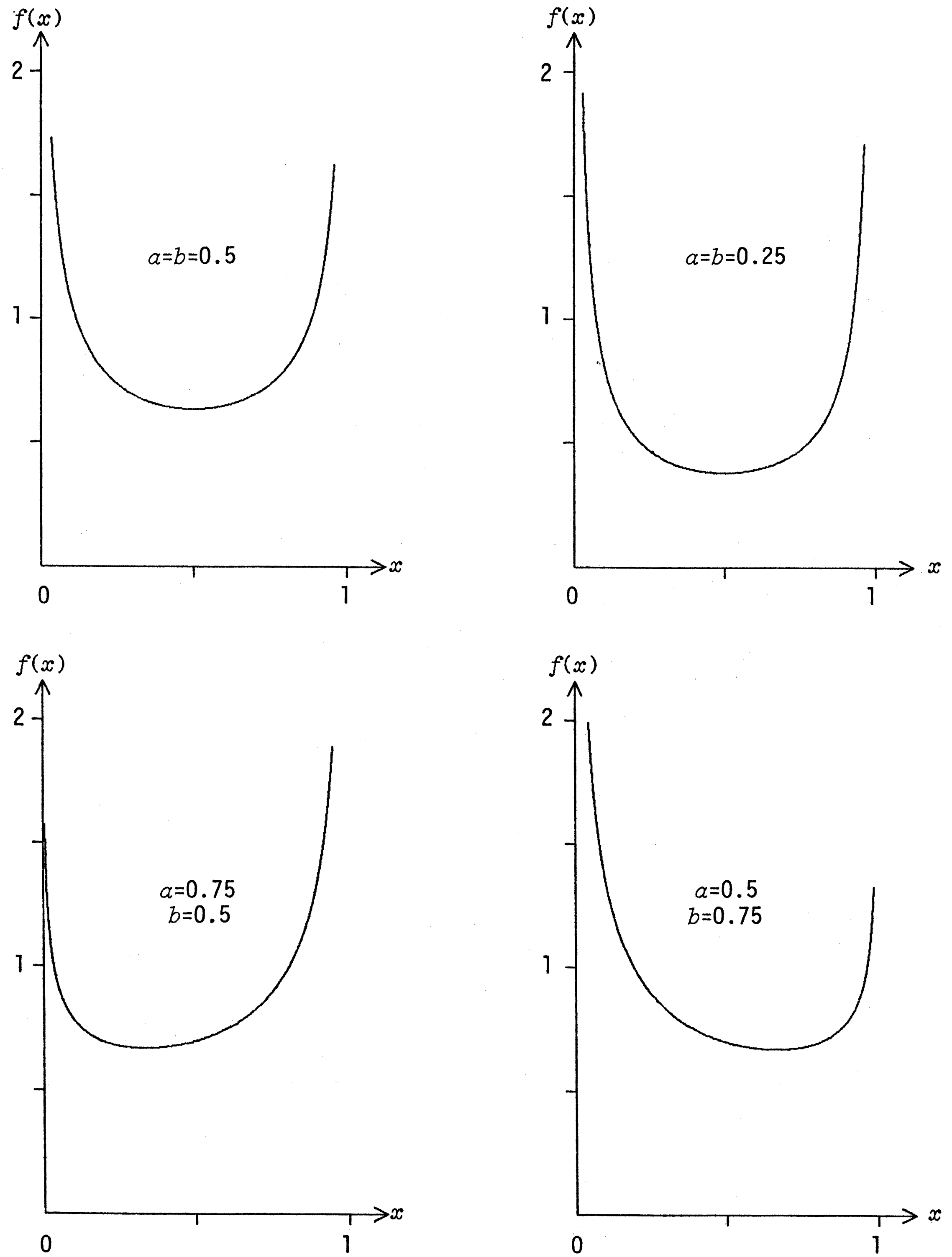
$$P\{X = x\} = \frac{1}{B(a,b)} x^{a-1}(1 - x)^{b-1}dx$$

for $0 < x < 1$
and $a > 0$, $b > 0$,
where $B(a,b)$ is the Eulerian beta function:

$$B(a,b) = \int_0^1 x^{a-1}(1 - x)^{b-1}dx.$$

These distributions are closely related to the gamma distribution. For different parameter values, histograms of the beta distribution can take different shapes: more or less symmetrical bell shapes, exponential from left to right or from right to left (see Calot 1967, p. 358; Jacquard 1976, p. 97 for graphs). However, we will limit ourselves to cases where $a < 1$ and $b < 1$, because of the very interesting distributions they imply: more or less symmetrical "U"-shaped histograms. Figure 21 shows some cases: the "depth" of the distribution is inversely proportional to a near 1, and to b near 0. For $a = b = 1$, we have the *continuous uniform* distribution as a special case. For $a = b = 0.5$, we have the *arc sine* distribution (Section 4.3.8).

Fig. 21. Beta distributions for four different (a,b) parameter pairs.



In order to devise a cannon, we can use tables of $F(x)$. The following algorithm can handle several beta distributions of different parameter pairs (a,b); for the i-th pair $(a,b)_i$, it prepares an approximation table of $F(x)$, for x incremented by steps of 0.05 from 0 to 1:

$$\text{tabl}(i,n) \approx F(0.05^*[n - 1])$$

$$\approx \frac{1}{B(a,b)} \int_0^{0.05(n-1)} x^{a-1}(1 - x)^{b-1}dx \quad \text{for } (a,b)_i. \tag{10}$$

In order to compute the factor $1/B(a,b)$ we use the property:

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}$$

(see Section 4.3.3 for a definition of $\Gamma[a]$). We can thus use the fgamma(2) function following to compute this constant, and then approximate $F(x)$ step by step:

$$F(x_s) \approx \sum_{i=1}^{x_s/dx} f(idx)dx$$

We use a microincrement $dx = 0.000125$, entering in table(i,n) only the values for an increment of 0.05, as in Eq. (10). This initialization is done in the first part of the program (procedure beinit), and is of course quite time-consuming, but the resulting tables can be stored for quick reference in actual use.

An actual call for a random value uses the procedure beta. As in discrete cases, we there use a u_s to point between two successive entries of the $F(x)$ table. A linear interpolation is used to simulate continuity.

```
program betauser(input,output);
constant
  ntot=10; {ten a,b pairs}
type
  real2array=array[0 . . ntot,o . . 21] of real;
.
.
.
procedure beinit(var nordr,ntot:integer;
  xa,xb:real; tabl:real2array);
constant
  dx=.000125;
var
  ab,sbeta,som,x,a,b,px:real;
begin {beinit}
  ab:=xa+xb;
  sbeta:=(fgamma(ab)/fgamma(xa))*fgamma(xb);
    {see fgamma code below}
  tabl[nordr,1]:=0.0;
  som:=0.0;
  x:=-.0000625;
  a:=xa-1.0;
  b:=xb-1.0;
  for i:=2 to 20 do
  begin {outerloop}
    for j:=1 to 400 do
    begin {inner loop}
      x:=x+dx;
      px:=(x**a)*((1.0-x)**b)*sbeta;
      som:=som+(px*dx)
    end; {innerloop}
    tabl[nordr,i]:=som
  end; {outerloop}
  tabl[nordr,21]:=1.0
end {beinit}
end. {betauser}
```

Algorithm (12)

```
procedure beta(var nordr,ntot:integer;
  xx:real; tabl:real2array);
var
  u,z,ds,db:real;
  k:integer;
begin {beta}
  u:=ran(0);
  for k:=2 to 21 do
  begin {loop}
    if u < tabl[nordr,k] then
    begin {innerblock}
      z:=tabl[nordr,k-1]:
      ds:=u-z;
      db:=tabl[nordr,k]-z;
      xx:=0.05*((ds/db)+(k-2)
    end {innerblock}
  end {loop}
end; {beta}
```

Algorithm (13)

The "beinit" call parameters are:

nordr the reference number for the i-th distribution with parameter pair $(a,b)_i$
xa,xb the given parameter pair
tabl the name of the array [ntot,21] declared in the calling program
ntot the number of different parameter pairs used in the calling program(s).

When using the procedure "beta":

nordr, tabl, ntot: as for "beinit" xx returns the generated random value x_s.

The following function: "fgamma(2)" is called by "beinit." It is an approximation for computing $\Gamma(z)$, (Hastings 1955, p. 157) in a configuration intended for:

$$\Gamma(1 + z), \text{ where } 0 \leq z \leq 1.$$

But, by the property:

$$\Gamma(z) = \frac{\Gamma(1 + z)}{z}$$

we can use it for $0 \leq z \leq 1$:

```
type
  realarray7=array[1 . . 7] of real;
.
.
.
function fgamma(var z:real):real;
type
  flag=(on,off);
var
  x,r:real;     i:integer;     f:flag;
  c:realarray7;
begin {fgamma}
  c[1]:=-.57710166;     c[2]:=.98585399;
  c[3]:=-.87642182;     c[4]:=.8328212;
  c[5]:=-.5684729;      c[6]:=.25482049;
  c[7]:=-.514993;
  x:=z;
  f:=off;
  if x >=1.0 then
  begin {large x}
    x:=x-1.0;
    f:=on;
  end; {large x}
 r:=1.0;
  for i:=1 to 7 do
  r:=r+(c[i]*(x**i);
  if f=off then
  fgamma:=r/x
  else
  fgamma:=r
end; {fgamma}
```

Algorithm (14)

Once the $F(x)$ tables are computed, Algorithm (14) is quite efficient: it requires one call to the U cannon, a search in the table, and linear interpolation. However, it certainly lacks elegance, and has been described mainly as an example procedure for cases where the computing of tables and linear interpolation would be unavoidable or more efficient, for example, in extensive use of Gauss-Laplace (Section 4.4.1). A very astute beta cannon has been devised by Jöhnk (pp. 9–10), valid for any positive real parameters a and b, and without the theoretical shortcoming of approximation:

```
function beta(var a,b:real):real;
var
  ea,eb,y1,y2,s:real;
begin {beta}
  ea:=1.0/a;
  eb:=1.0/b;
  repeat
    y1:=ran(0)**ea;
    y2:=ran(0)**eb;
    s:=y1+y2
  until s <=1.0;
  beta:=y1/s
end; {beta}
```

Algorithm (15)

The efficiency of Algorithm (15) is interesting, at least for the parameter values in which we are interested ($a < 1$ and $b < 1$). In such cases, "s:=y1+y2;" has a fair chance of not being greater than 1, since the exponents ea and eb are greater than 1. Thus y1 and y2 are even smaller than the u variables generating them. For $a = b = 0.75$, 3.15 calls to ran are necessary on average; for $a = b = 0.5$, 2.55 calls are required.

5.0 Some Comparisons and Applications

L'action n'est possible que dans une certaine insouciance et la vie n'est qu'un acte de confiance en nous-mêmes et dans la bienveillance des hasards.—Rémy de Gourmont

In order to compare the five symmetrical distributions with one central mode that we have described (of Laplace, Cauchy, hyperbolic cosine, logistic, and Gauss-Laplace) we can adjust them to a common characteristic. For instance, we can scale their dispersions by means of multiplying factors. Let us take as starting point

$$y = F(3) = 0.99865$$

of the normal distribution ($\mu = 0, \sigma = 1$): the probability of values greater than 3 is $1 - y = 0.00135$. We can center the other distributions on zero as well, and scale them to conform to this condition. This can be done easily by replacing x_s by 3, and u_s by $y(=0.99865)$ in the cannon formulas, and solving for the scaling factor.

This gives, for the first law of Laplace:

$$-\tau \ln(2 - 2y) = 3 \rightarrow \tau = 0.50723$$

for the Cauchy distribution:

$$\tau \tan[\pi(y - ½)] = 3 \rightarrow \tau = 0.01272$$

for the hyperbolic cosine:

$$\tau \ln[\tan(\frac{\pi y}{2})] = 3 \rightarrow \tau = 0.48732$$

logistic (with $\beta = 0$, in order to have the mean at 0):

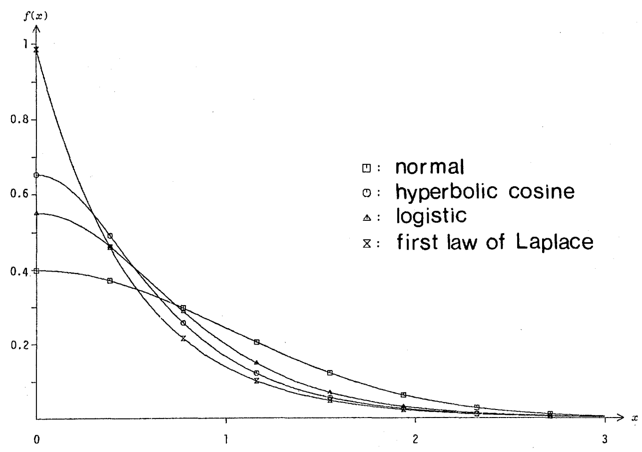
$$\frac{-\ln(\frac{1}{y} - 1)}{\alpha} = 3 \rightarrow \alpha = 2.2021.$$

Figure 22 shows the right halves of the histograms, with the proper factors. The Cauchy distribution is not shown because its representation is impossible on the same scale as the others: with the above τ, it gives

$$f(0) = 25.02436,$$
$$f(1) = 0.00405, \text{ etc.:}$$

a vertiginously abrupt mode! Otherwise, the figure speaks for itself: each distribution has a certain "personality," a certain demeanor. Their differences can, for instance, be considered in modulating various behaviors of pitch clouds, more or less centered on their mean. Of course, large numbers of random values are necessary for this kind of process to have any significant overall relation to the histograms.

Fig. 22. Right halves of the histograms of four distributions, all scaled so that F(3) = 0.99865.



5.1 Application of the Gamma Distribution

Random variables generated by the gamma, exponential, and linear distributions can be applied, for instance, to time intervals. Each distribution again has a personality. In order to make the comparison clear, all sequences of random intervals should be set to the same average density: we choose $e = 2.71828\ldots$ points per second. The mean of the linear distribution of parameter g is $g/3$. Setting $g = 3/e$ will generate the average value $1/e$ as interval between points, that is: e points per second. The exponential cannon can be called directly with $\delta = e$. For the gamma distributions, we know that the mean of γ_ν is ν; in order to set this mean to $1/e$,

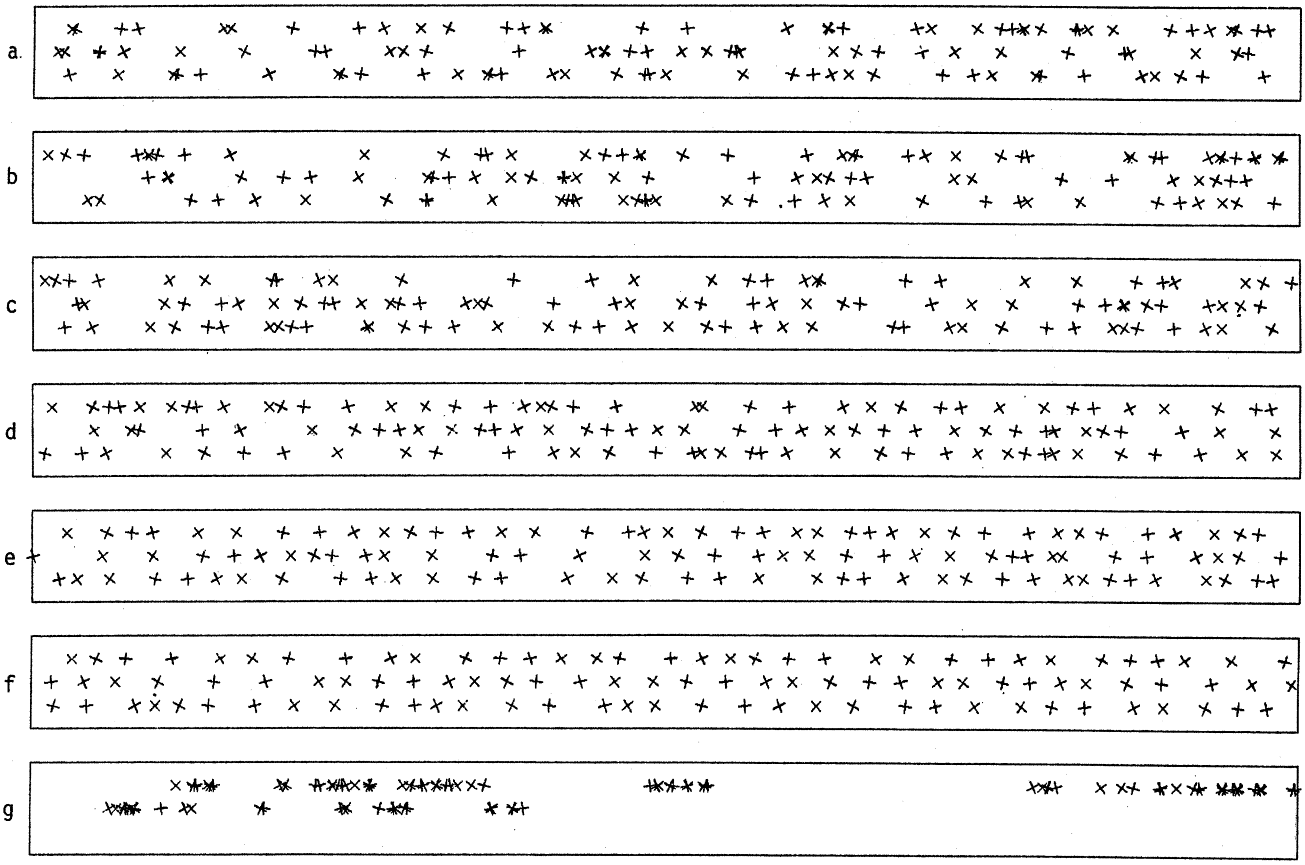


Fig. 23. The same average density generated by different distributions: (a) linear, (b) γ_1 (exponential), (c) γ_2, (d) γ_4, (e) γ_8, (f) γ_{16}, (g) scaled Cauchy (see text). An "x" represents a "point"; rotations of this symbol are used to render somewhat visible clusters of two or more neighboring points, superimposed because of the figure's exiguity. There are three lines of points per rectangle, representing a total of 40 units.

we must set:

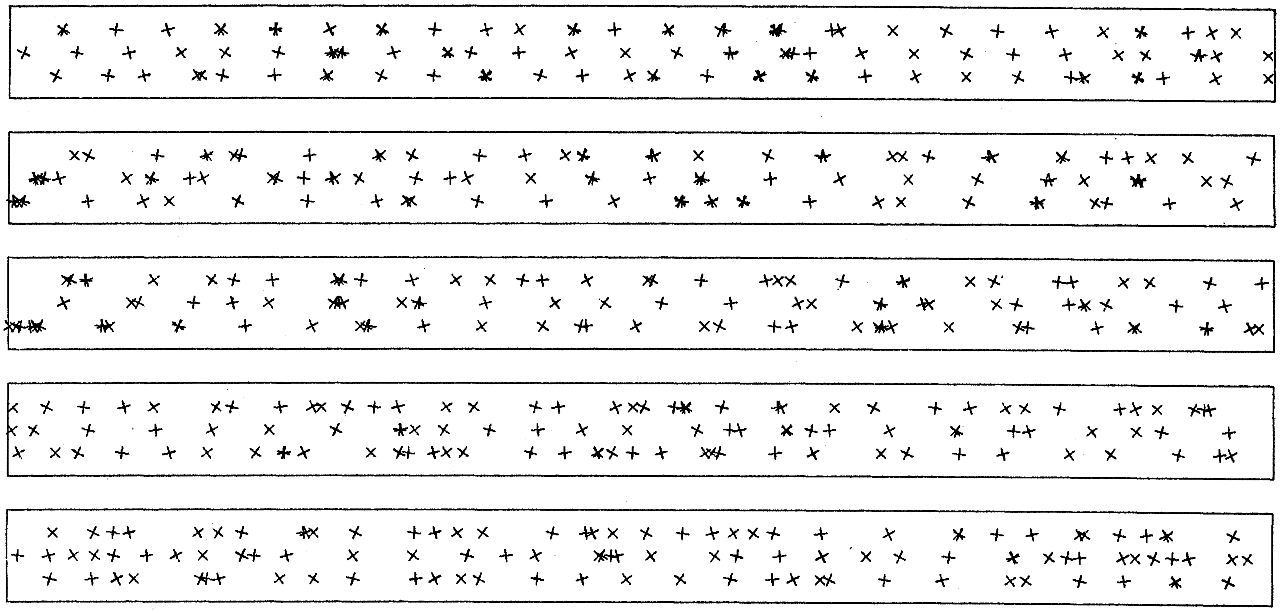
$$X = \frac{\gamma_\nu}{\nu e}.$$

We have chosen to test these distributions (including the exponential as γ_1) with $\nu = 1, 2, 4, 8, 16$. This exponential scaling of ν amounts to a rather linear rhythmical regularization from the perceptual point of view, obviously due to the fact that the gamma distribution loses its asymmetry for large ν's. Figure 23, from (b) to (f), shows results of this comparison. Each rectangle represents 40 units, and the average density per unit is each time near e (less than 2% difference). From γ_1(b) to γ_{16}(f), the increase of symmetry is clear. The linear distribution (a) stands somewhere between the γ_1 and the γ_2. For the sake of comparison, (g) shows a Cauchy distribution of intervals (obtained with iopt = 1 in the algorithm in Section 4.3.5) which has been scaled *a posteriori* to give the same density of e points per unit. This distribution's irregularity seems truly unpredictable, and no romantic effort to control it by its parameter τ can overcome the mathematical fact that it has no predictable mean. The example shown has been chosen as an illustration of a sort of maximum acceptable asymmetry; more would be difficult to use in such a context.

5.2 Application of the Beta Distribution

In rhythmic applications also, the beta distributions, including the continuous uniform as the special case $\beta(1,1)$, show an interesting progression. The mean of

Fig. 24. The same average density generated by different beta distributions: (a) $\beta(0.125,0.125)$, (b) $\beta(0.25,0.25)$, (c) $\beta(0.5,0.5)$, (d) $\beta(0.75,0.75)$, (e) $\beta(1,1)$, equivalent to the continuous uniform. An "x" represents a "point"; rotations of this symbol are used to render somewhat visible clusters of two or more neighboring points, superimposed because of the figure's exiguity. There are three lines of points per rectangle, representing a total of 40 units.



a $\beta(a,b)$ distribution is

$$a/(a+b)$$

Since here also we would like to compare the different distributions with a common average density of e points per unit, a multiplying factor must be applied to the random values generated in [0,1] by the cannons, in order to bring the mean interval to $1/e$:

$$X = \beta(a,b)\left[\frac{a+b}{ae}\right].$$

We then have cases (Figure 24) of time intervals in the finite range $[0,(a+b)/ae]$.

In our examples, since we have always set $a = b$, the actual range is $[0,2/e]$. As compared with asymmetrical cases of Fig. 23, a greater homogeneity is achieved even for (a), due to the fact that a sort of basic pulsation is present, equal to the longest interval, $2/e$, and the random values can be perceived as subdivisions of this quasi-tempo. In (a), since $\beta(0.125, 0.125)$ greatly favors values near 0 and 1, we have many configurations of this type:

with different numbers of short notes. From (a) to (e) this characteristic becomes less important, since mean values become more and more probable, until reaching equiprobability in (e).

6.0 Conclusion

The preceding comparisons and examples are merely suggested as starting points for composition. Even with these stochastic generators specified, a central problem in compositional design remains in applying these automata to some music processes. There is no limit to the use of stochastic variables in music. For instance, the points of Fig. 23 could be read as lists of selected pitches on a frequency axis, and they could then be used as material for further development. The durations of sounds could be controlled by some symmetrically centered distribution, or by some time interval distribution—perhaps different from the ones controlling attack times.

Going beyond these direct applications of static stochastic processes, one may devise stochastic transitions between different distributions. One distribution may be used to control the parameters of another distribution, creating a hierarchical stochastic system.

If probability is an answer to certain problems of compositional choice and selection, there still remains a lot to choose—the music is still left to be made!

7.0 Acknowledgments

I must once again express my deep gratitude to Iannis Xenakis for reading this manuscript and suggesting some modifications. I would also like to thank Michel Decoust and Jean-Claude Risset for providing the occasion of putting nearly all the content of this paper into practice in my tape piece, *Les portes du sombre Dis* (IRCAM, Paris, 1979).

References

1. Aetius. 1976. From *Epicure et les épicuriens.* Paris: P.U.F.
2. Calot, G. 1967. *Cours de calcul des probabilités.* Paris: Dunod.
3. Charles, Daniel. 1978. *Gloses sur John Cage,* coll. 10/18 no. 1212. Paris: U.G.E.
4. Da Vinci, Leonardo. 1942. Translated from *Carnets,* vol. 1. Paris: Gallimard.
5. Feller, William. 1966*a. An introduction to probability theory and its applications,* vol. I. New York: John Wiley.
6. Feller, William. 1966*b. An introduction to probability theory and its applications,* vol. II. New York: John Wiley.
7. Gnedenko, B. V., and Khintchine, A. La. 1969. *Introduction à la théorie des probabilités,* coll. Science-poche no. 9. Paris: Dunod.
8. Grossman, Israël, and Magnus, Wilhelm. *Groups and their graphs,* New Mathematical Library, no. 14. New York: Random House.
9. Guénon, René. 1946. *Les principes du calcul infinitésimal,* Paris: Gallimard.
10. Hamming, R. W. 1973. *Numerical methods for scientists and engineers.* New York: McGraw-Hill.
11. Hastings, C. Jr. 1955. *Approximations for digital computers.* Princeton: Princeton University Press.
12. Jacquard, A. 1976. *Les probabilités,* coll. Que Sais-Je, no. 1571. Paris: P.U.F.
13. Jöhnk, M. D. Erzeugung von betaverteilten und gammaverteilten Zufallzahlen. *Metrika* 8(1):5–15.
14. Kaegi, Werner. 1971. Musique et technologie dans l'Europe de 1970. *Musique et Technologie, La Revue Musicale* no. 268–269. Paris: Richard-Masse.
15. Kleppner, Daniel, and Ramsey, Norman. 1965. *Quick calculus,* New York: John Wiley.
16. Knuth, D. E. *The art of computer programming,* vol. II. Reading, Massachusetts: Addison-Wesley.
17. Koenig, G. M. 1971. *Observations on compositional theory.* Utrecht, The Netherlands: Instituut voor Sonologie.
18. Koenig, G. M. *Computer composition.* Utrecht, The Netherlands: Instituut voor Sonologie.
19. Lipschutz, S. 1965. *Probability.* Schaum's Outline Series. New York: McGraw-Hill.
20. Maurin, J. 1975. *Simulation déterministe du hasard,* Paris: Masson.
21. Piaget, Jean. 1973. *Le développement de la notion de temps chez l'enfant.* Paris: P.U.F.
22. Revault d'Allonnes, Olivier. 1973. *La création artistique et les promesses de liberté.* Paris: Klincksiek.
23. Russolo, Luigi. 1975. *L'art des bruits,* coll. Avant-gardes. Lausanne: L'Age d'Homme.
24. Schrage, Linus. A more portable Fortran random number generator. *ACM Transactions on Mathematical Software* 5(2):132.
25. *Scientific subroutine package.* IBM Corporation.
26. Vessereau, A. 1953. *La statistique,* coll. Que Sais-Je, no. 281. Paris: P.U.F.
27. Xenakis, Iannis. 1963. *Musiques formelles,* La Revue Musicale, no. 253–254. Paris: Richard-Masse.
28. Xenakis, Iannis. 1971*a. Musique architecture,* coll. Mutations-orientations no. 11, Casterman.
29. Xenakis, Iannis. 1971*b. Formalized music.* Bloomington, Indiana: Indiana University Press.