ISTITUTO NAZIONALE DI FISICA NUCLEARE

Sezione di Milano

INFN/AE-94/16
19 Luglio 1994

I. Manno:

INTRODUCTION INTO THE MONTE-CARLO METHOD

Servizio Documentazione
dei Laboratori Nazionali di Frascati

INFN - Istituto Nazionale di Fisica Nucleare
Sezione di Milano

INFN/AE-94/16
19 Luglio 1994

INTRODUCTION INTO THE MONTE-CARLO METHOD

I. Manno
INFN.Sezione di Milano - Via Celoria, 16 - 20133 Milano - Italy

ABSTRACT

This article describes the basic principles of the Monte-Carlo method. The use of the method is shown on examples taken from the field of experimental particle physics.

Contents

1 Introduction

Although the Monte-Carlo method was known long time before, it came into general use only with the advent of the electronic computers because it needs a lot of calculations. J. von Neumann and S. Ulam are considered as the inventors of the Monte-Carlo method. The random number plays fundamental role in the method. Monte-Carlo is famous for its casinos and the roulett is nothing else than a random number generator. This explanes the name of the method. The method includes all numerical methods which simulate processes depending on random variables. Usually these calculations are too complex to solve them analytically. The calculation are carried out by means of pseudorandom numbers which have very similar properties to that of the true random numbers. The Monte-Carlo method may be used to solve problems which do not depend on random variables but may be described with a random model like the Monte-Carlo integration.

As an example we calculate the $\pi = 3.14159$ by means of random process. We drop a needle, whose length is l, on a plane where parallel straight lines are drawn (Fig. 1a). The distance between two neighbouring lines is l. The angle between the needle and the sraight lines is ϑ $(0 \leq \vartheta < \pi)$. At a given angle ϑ, the probability that the needle falls on a line is given by $p = x/l$, where $x = l \sin \vartheta$ is the projection of the needle. The probability of all possible events is 1. The average over the whole $(0 \leq \vartheta < \pi)$ interval is

$$\frac{k}{N} = \frac{\int_0^\pi \sin \vartheta d\vartheta}{\int_0^\pi d\vartheta} = \frac{2}{\pi} ,$$

where k is the number of 'fortunate' events, N is the whole number of events and k/N is the relative frequency (see Section 2.1). The probability $p(\vartheta)$ that at angle ϑ the needle falls on a line is shown in (Fig. 1b). The value of π is given by

$$\pi = \frac{2N}{k} .$$

The error of the Monte-Carlo calculation may be estimated with the error of the binomial distribution (see Section 2.3.3):

$$\sigma_\pi \sim \frac{2.37}{\sqrt{N}}.$$

2 Probability

2.1 Basic Concepts of Probability

A random event is an event which has more than one possible outcome. A probability may be associated with each outcome. The outcome of a random event is not predictable, only the probabilities of the possible outcomes are known. In contrast, an event which has only one possible outcome is certain, the probability of the outcome is unity.

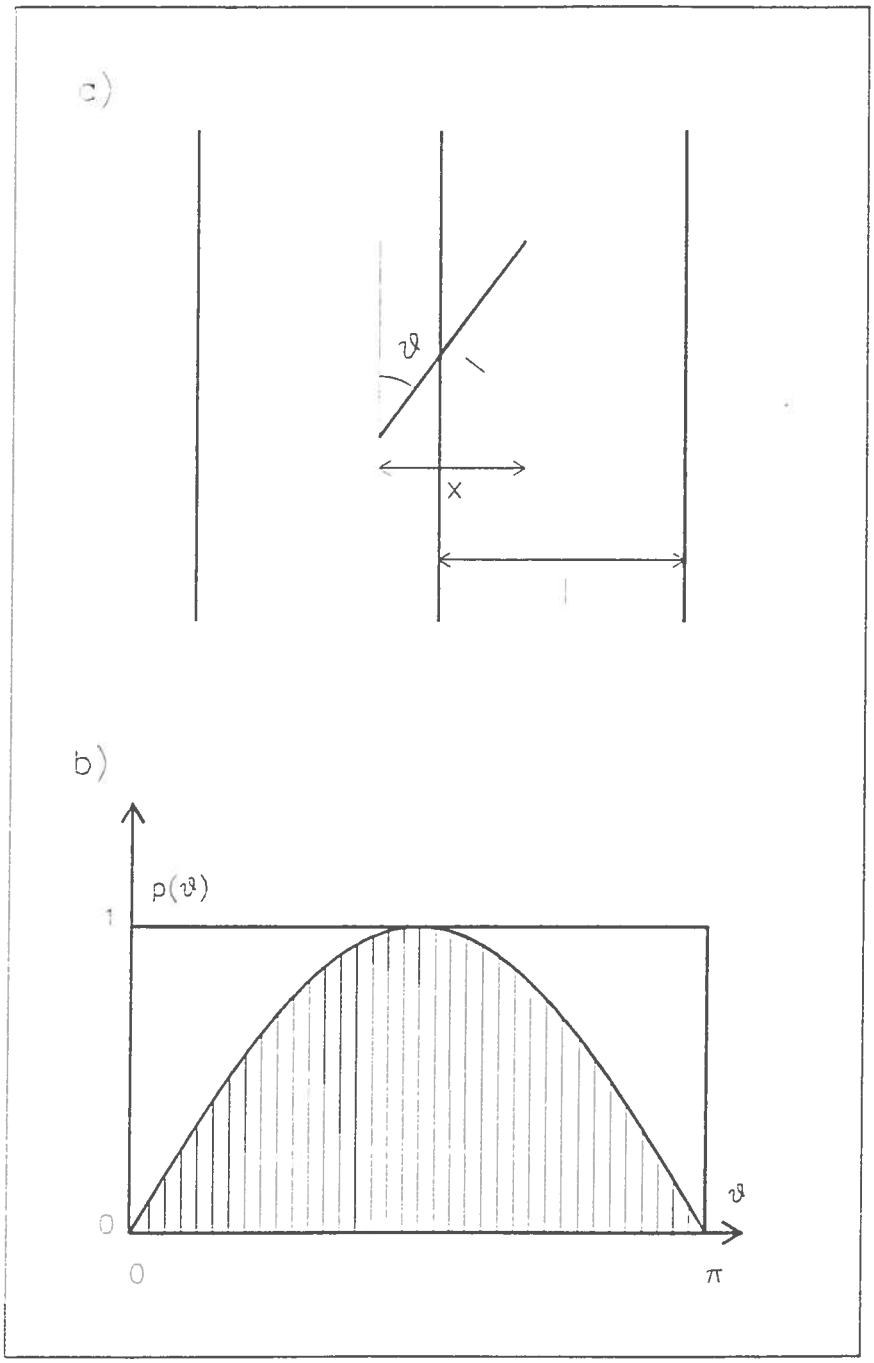

Figure 1: Calculation of π with Monte-Carlo method

2.1.1 Probability and Frequency

Here are some definitions of the probability without an attempt to giving its exact mathematical definition.

- Probability is the numerical value of the chance of occurrence of one or more possible results of an unpredictable event.

- The set of all possible exclusive elementary events x_i is denoted by Ω. The probability of occurence of x_i, $P(x_i)$, should have the followig properties:

  1. $P(x_i) \geq 0$ for all i ,
  2. $P(x_i \text{ or } x_j) = P(x_i) + P(x_j)$, x_i and x_j are exclusive ,
  3. $P(\Omega) = 1$.

- Consider the throwing of a die. Suppose that having thrown the die N times, where we find that k times as a result of the throw we obtained '5'. The probability of getting '5' as the results of a throw can be taken to be

$$p = \frac{k}{N} , \tag{1}$$

  i.e. the number k of 'fortunate' events divided by the whole number N of events. This ratio is called relative frequency.

The relative frequency may be determined by means of theoretical considerations in many cases. For example in the case of a die due to its symmetry one expects that each of the six sides has the same probability 1/6.

2.1.2 Law of Addition of Probabilities

The events are called exclusive when the occurrence of one of them implies that none of the others occurs. The result of one throw of a die is one of the six possible numbers. There are two exclusive events A and B. Consider the event C which we suppose to happen if one of the two events A or B happens. The probability of the event C is given by

$$P(C) = P(A) + P(B) . \tag{2}$$

The probabilities are given approximately by the relative frequencies:

$$P(A) \sim \frac{k_A}{n} \qquad P(B) \sim \frac{k_B}{n} .$$

The event C happens $k_C = k_A + k_B$ times and the whole number of the events is n.

$$P(C) \sim \frac{k_C}{n} = \frac{k_A + k_B}{n} = \frac{k_A}{n} + \frac{k_B}{n} \sim P(A) + P(B) .$$

2.1.3 Law of Multiplication of Probabilities

Suppose that having thrown a die two times. The event A happens, if we get odd number as the result of the first throw. The event B happens, if we get '1' as the result of the second throw. The event C happens if the result of the first throw is A and the result of the second throw is B. The probability of the event C is given by

$$P(C) = P(A)P(B) \; . \tag{3}$$

The probabilities are given approximately by the relative frequencies:

$$P(A) \sim \frac{k_A}{n} \qquad P(B) \sim \frac{k_B}{n} \; .$$

The event C happens $k_C = k_A k_B$ times and the whole number of the events is nn.

$$P(C) \sim \frac{k_C}{nn} = \frac{k_A k_B}{nn} = \frac{k_A}{n}\frac{k_B}{n} \sim P(A)P(B) \; .$$

2.1.4 Conditional Probability

There are two random events A and B. The probability that A happens when B happens is given by the conditional probability of A given B written $P(A|B)$:

$$P(A|B) \sim \frac{k_{AB}}{k_B} = \frac{k_{AB}/n}{k_B/n} \sim \frac{P(AB)}{P(B)} \; , \tag{4}$$

where the probabilities are approximated with the relative frequencies. This is probability that an elementary event, known to belong to the set of B is also the member of the set A.

The events A and B are said to be independent if

$$P(A|B) = P(A)$$

which means that the occurrence of B is irrelevant to the occurrence of A.

2.2 Random Variables

With a random event A may be associated a random variable ξ, which takes different possible numerical values $x_1, x_2, ...$ corresponding to different outcomes. The corresponding probabilities $P(x_1), P(x_2), ...$ form a probability distribution of the random variable.

2.2.1 Discrete Random Variable

The random variable ξ and its probability distribution is called discrete, when it can take its value from a finit or infinit set of discrete values $x_1, x_2, ...$ The distribution of the random variable is given by the table:

$$\xi = \begin{Bmatrix} x_1, x_2, ..., x_n, \cdots \\ p_1, p_2, ..., p_n, \cdots \end{Bmatrix} \tag{5}$$

where $x_1, x_2, ..., x_n, ...$ are the possible values of the random variable, while $p_1, p_2, ..., p_n, ...$ are their probabilities. The probability that ξ takes the value x_i is given by

$$P(\xi = x_i) = P(x_i) = p_i \ .$$

Generally the values $x_1, x_2, ..., x_n, ...$ may be optional. The probabilities p_1, p_2, ..., p_n, ... should satisfy the following conditions:

- All should be positive:

$$p_i > 0$$

- Their sum should be equal to 1:

$$\sum_i p_i = 1 \ .$$

In other words it is certain that in each case ξ takes one of the values $x_1, x_2, ..., x_n, ...$ The mathematical expectation value $E(\xi)$ of the random variable ξ is given by

$$E(\xi) = \frac{\sum_i x_i p_i}{\sum_i p_i} = \sum_i x_i p_i \ . \tag{6}$$

The variance $D^2(\xi)$ of the random variable ξ is given by

$$D^2(\xi) = E((\xi - E(\xi))^2) = \sum_i (x_i - E(\xi))^2 p_i \ . \tag{7}$$

There are two discrete random variables ξ and η. Their possible values are $k_1, k_2, ...$ and $l_1, l_2, ...$ The probability distribution of the random variables ξ and η is $h(\xi, \eta)$. The probability distribution of ξ is given by

$$p(\xi) = \sum_i h(\xi, l_i) \ .$$

The expectation value of ξ is

$$E(\xi) = \sum_j k_j p(k_j) = \sum_j k_j \sum_i h(k_j, l_i) \ .$$

The probability distribution of η is given by

$$q(\eta) = \sum_j h(k_j, \eta) \ .$$

The expectation value of η is

$$E(\eta) = \sum_i l_i q(l_i) = \sum_i l_i \sum_j h(k_j, l_i) \ .$$

The random variables ξ and η are called independent of each other when

$$h(\xi, \eta) = p(\xi) q(\eta) \ .$$

The expectation value of a function $f(\xi,\eta)$ is given by

$$E(f(\xi,\eta)) = \sum_i \sum_j f(k_j,l_i)h(k_j,l_i) \ .$$

The expectation value of $f(\xi,\eta) = \xi + \eta$ is

$$E(\xi+\eta) = \sum_j \sum_i (k_j + l_i)h(k_j,l_i) =$$

$$\sum_j \sum_i k_j h(k_j,l_i) + \sum_j \sum_i l_i h(k_j,l_i) =$$

$$\sum_j k_j \sum_i h(k_j,l_i) + \sum_i l_i \sum_j h(k_j,l_i) =$$

$$\sum_j k_j p(k_j) + \sum_i l_i q(l_i) = E(\xi) + E(\eta)$$

Generally for the product of two random variables

$$E(\xi\eta) \neq E(\xi)E(\eta) \ .$$

If the random variables are independent, then $E(\xi\eta) = E(\xi)E(\eta)$:

$$E(\xi\eta) = \sum_j \sum_i k_j l_i h(k_j,l_i) =$$

$$\sum_j \sum_i k_j l_i p(k_j) q(l_i) =$$

$$\sum_j k_j p(k_j) \sum_i l_i q(l_i) = E(\xi)E(\eta)$$

The ξ and η are two independent random variables. The following relations are true for them

$$E(\xi+\eta) = E(\xi) + E(\eta) \ , \tag{8}$$

$$E(\xi\eta) = E(\xi)E(\eta) \ , \tag{9}$$

$$E(a+b\xi) = a + bE(\xi) \ , \tag{10}$$

$$D^2(\xi+\eta) = D^2(\xi) + D^2(\eta) \qquad \text{and} \tag{11}$$

$$D^2(a+b\xi) = b^2 D^2(\xi) \ , \tag{12}$$

where a and b are constants.

2.2.2 Continuous Random Variable

Suppose there is a radioactive source at the center of a rectangular coordinate system. The particles are emitted isotropically by the source. The angle φ of the projection of the fly direction on the xy-plane may have any value from the $(0, 2\pi)$ interval (Fig. 2a).

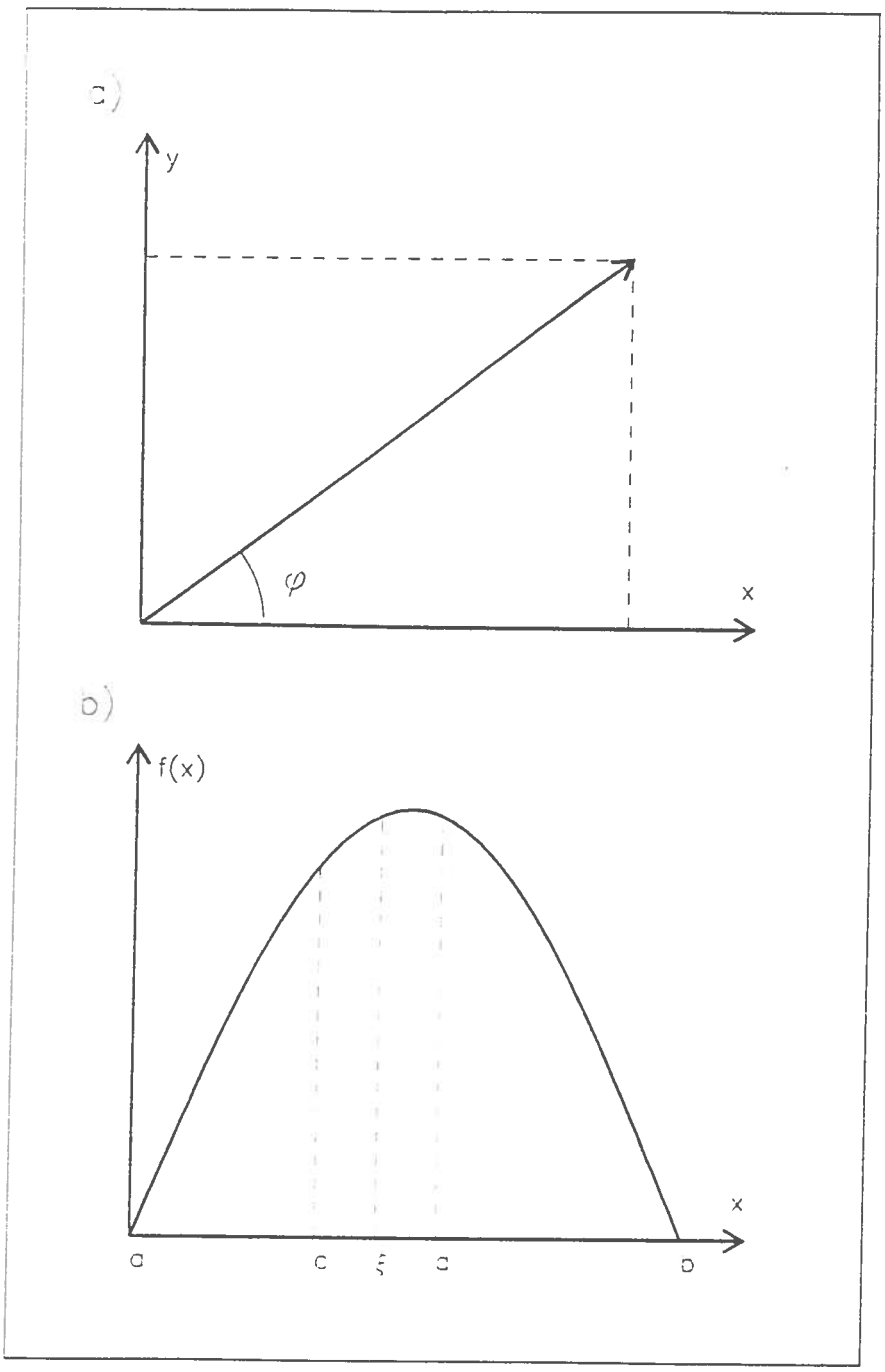

Figure 2: Continuous Random Variable

The random variable ξ is called continuous if it may have any value from the (a,b) interval. The probability that one observation is from the (c,d) subinterval of the (a,b) interval is given by

$$P(c \leq \xi \leq d) = \int_c^d f(x)dx \ , \tag{13}$$

where $f(x)$ the probability density function (Fig. 2b).

The values of a and b may be optional. There are cases where $a = -\infty$ and $b = +\infty$. The probability density function $f(x)$ should satisfy the following conditions:

- It should be positive:
$$f(x) > 0.$$
- Its integral over the whole interval should be 1:
$$\int_a^b f(x)dx = 1 \ .$$

The cumulative distribution function is given by

$$F(x) = \int_a^x f(x')dx' \ .$$

It has the followig properties:

- $F(a) = 0$,
- $F(b) = 1$ and
- $0 \leq F(x) \leq 1$, if $a \leq x \leq b$.

The function $F(x)$ is a monotone increasing function:

$$F(x_1) < F(x_2), \quad \text{if} \quad x_1 < x_2 \ .$$

The probability that the random variable ξ is in the (c,d) interval is given by

$$P(c < \xi < d) = F(d) - F(c) \ .$$

The mathematical expectation value $E(\xi)$ of the random variable ξ is given by

$$E(\xi) = \frac{\int_a^b xf(x)dx}{\int_a^b f(x)dx} = \int_a^b xf(x)dx \ .$$

The variance $D^2(\xi)$ of the random variable ξ are given by

$$D^2(\xi) = \int_a^b (x - E(\xi))^2 f(x)dx \ .$$

As in the case of the discrete random variable for two independent random variables ξ and η the following relations are true:

$$E(\xi+\eta) = E(\xi) + E(\eta) \ ,$$

$$E(\xi\eta) = E(\xi)E(\eta) \ ,$$

$$E(a+b\xi) = a + bE(\xi) \ ,$$

$$D^2(\xi+\eta) = D^2(\xi) + D^2(\eta) \qquad \text{and}$$

$$D^2(a+b\xi) = b^2 D^2(\xi) \ ,$$

where a and b are constants.

2.3 Probability Distributions

2.3.1 Joint Distribution of Two Random Variables

There are two random variables ξ and η. They form a random vector variable (ξ,η), whose possible values are

$$\xi = x_1, x_2, ..., x_n, ... \qquad \text{and} \qquad \eta = y_1, y_2, ..., y_m, ...$$

The values of the random vector variable are points on the xy-plane. The set of all possible values is Ω. A subset of Ω is E. In the discrete case the probability that the values of ξ and η are x_i and y_j respectively is given by

$$P((\xi = x_i, \eta = y_j) = p_{ij} \ , \qquad i = 1,2,... \qquad \text{and} \qquad j = 1,2,... \ . \tag{14}$$

The probability that the value of (ξ,η) is in the subset E is given by

$$P((\xi,\eta) \in E) = \sum_{(x_i,y_j)\in E} p_{i,j}.$$

The probability that the values are in Ω is given by

$$P((\xi,\eta) \in \Omega) = \sum_{(x_i,y_j)\in \Omega} p_{i,j} = 1 \ .$$

In the continuous case

$$P((\xi,\eta) \in E) = \int_E \int f(x,y)dxdy, \tag{15}$$

where $f(x,y)$ is the probability density function. As in the discrete case

$$P((\xi,\eta) \in \Omega) = \int_\Omega \int f(x,y)dxdy = 1 \ .$$

2.3.2 Uniform Distribution

The discrete random variable ξ is called uniform on the $x_1, x_2, ..., x_n$ numerical values, if these values are the possible outcomes of the random variable ξ and each outcome has the same probability:

$$P(\xi = x_i) = \frac{1}{n}\,, \qquad i = 1, 2, ..., n. \tag{16}$$

The mathematical expectation value and the variance of the random variable ξ are given by

$$E(\xi) = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{and}$$

$$D^2(\xi) = \frac{1}{n}\sum_{i=1}^{n}\Big(x_i - E(\xi)\Big)^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n^2}\Big(\sum_{i=1}^{n} x_i\Big)^2\,.$$

The continuous random variable ξ is called uniform on the (a, b) interval, if its probability density function is

$$f(x) = \begin{cases} 1/(b-a)\,, & \text{if } a < x < b; \\ 0\,, & \text{otherwise.} \end{cases} \tag{17}$$

The mathematical expectation value and the variance of the random variable ξ are given by

$$E(\xi) = \int_a^b \frac{x}{b-a}dx = \frac{a+b}{2} \qquad \text{and}$$

$$D^2(\xi) = \int_a^b \frac{x^2}{b-a}dx - \Big(\frac{a+b}{2}\Big)^2 = \frac{(b-a)^2}{12}\,.$$

2.3.3 Binomial Distribution

Any random process with exactly two possible outcomes (alternative) is a Bernoulli process. The formula of Bernoulli may be derived using the laws of addition and multiplication of the probabilities. Let us consider an observation which has only two outcomes a and b. The probability of the outcome a is p, while the probability of the outcome b is $q = 1 - p$. If the process is repeated n times independently, then the probability of obtaining exactly $\xi = k$ times the outcome a is given by

$$f(k; n, p) = P_k^{(n)} = P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k},\ k = 0, 1, 2, ..., n\,. \tag{18}$$

The sum of the probabilities, using the binomial theorem, is

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + 1 - p)^n = 1\,.$$

The binomial distribution is shown in Fig. 3. The mathematical expectation value $E(\xi)$ and the variance $D^2(\xi)$ of the random variable ξ are given by

$$E(\xi) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} =$$

$$np\sum_{r=0}^{n-1}\binom{n-1}{r}p^r(1-p)^{n-1-r}=np \qquad \text{and}$$

$$D^2(\xi)=\sum_{k=0}^{n}(k-np)^2\binom{n}{k}p^k(1-p)^{n-k}=$$

$$\sum_{k=0}^{n}(k(k-1)+k)\binom{n}{k}p^k(1-p)^{n-k}-(np)^2=np(1-p)\ .$$

2.3.4 Poisson Distribution

The observed result of a Poisson process is a non-negative integer number n. The parameter μ is any non-negative real number. The Poisson distribution describes the population of events in any interval of x (e.g. space or time) whenever: (a) the number of events in any interval of x is independent of that in any non-overlapping interval; (b) in any small Δx, the probability of one event is $\lambda\Delta x$ and the probability of two or more vanishes at least as fast as $(\Delta x)^2$, as $\Delta x \to 0$; and (c) λ does not depend on x. Then $\mu \equiv \lambda x$ (Fig. 4).

In the case of detection of particles, the mean value of the particles detected during the time T is n. For each particle the probability that it is detected in the subinterval (t_0, t_0+t) is given by $p = t/T$. The probability that in this interval (t_0, t_0+t) exactly k particles are detected is given by the binomial distribution

$$f(k;n,p)=\binom{n}{k}\Big(\frac{t}{T}\Big)^k\Big(1-\frac{t}{T}\Big)^{n-k}\ ,$$

where $p = t/T$, $np = nt/T$, n/T is the intensity if the particles and $nt/T = \mu$ is the mean value of the detected particles in the interval (t_0, t_0+t).

The Poisson distribution may be derived from the binomial distribution:

$$f(k;n,p)=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k},\ k=1,2,...,n\ .$$

Let us increase n $(n \to \infty)$ and decrease p $(p \to 0)$ so that

$$np=\mu>0.$$

In the case of a fix k value $f(k;n,p)$ has a limit value which is derived below:

$$f(k;n,\mu)=P_k^{(n)}=\binom{n}{k}\Big(\frac{\mu}{n}\Big)^k\Big(1-\frac{\mu}{n}\Big)^{n-k}=$$

$$=\frac{n(n-1)(n-2)...(n-k+1)}{k!}\Big(\frac{\mu}{n}\Big)^k\Big(1-\frac{\mu}{n}\Big)^n\Big(1-\frac{\mu}{n}\Big)^{-k}=$$

$$=\frac{\mu^k}{k!}\Big(1-\frac{\mu}{n}\Big)^n\frac{n(n-1)(n-2)...(n-k+1)}{n^k}\Big(1-\frac{\mu}{n}\Big)^{-k}.$$

If k is fix even number and $n \to \infty$, then:

$$\lim_{n\to\infty}\frac{n(n-1)(n-2)...(n-k+1)}{n^k}=\Big(1-\frac{1}{n}\Big)\Big(1-\frac{2}{n}\Big)...\Big(1-\frac{k-1}{n}\Big)\to 1,$$

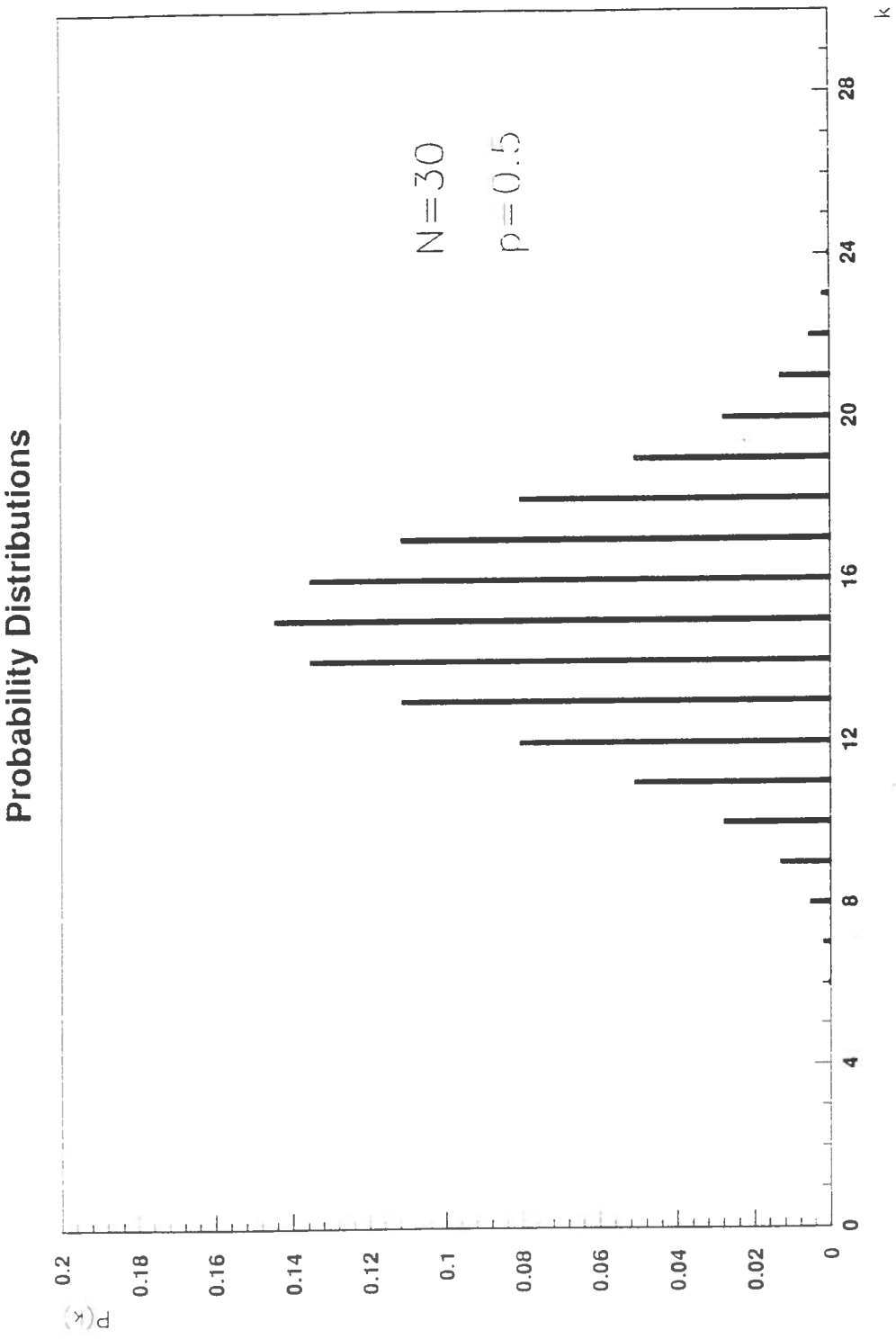

Figure 3: Binomial Distribution

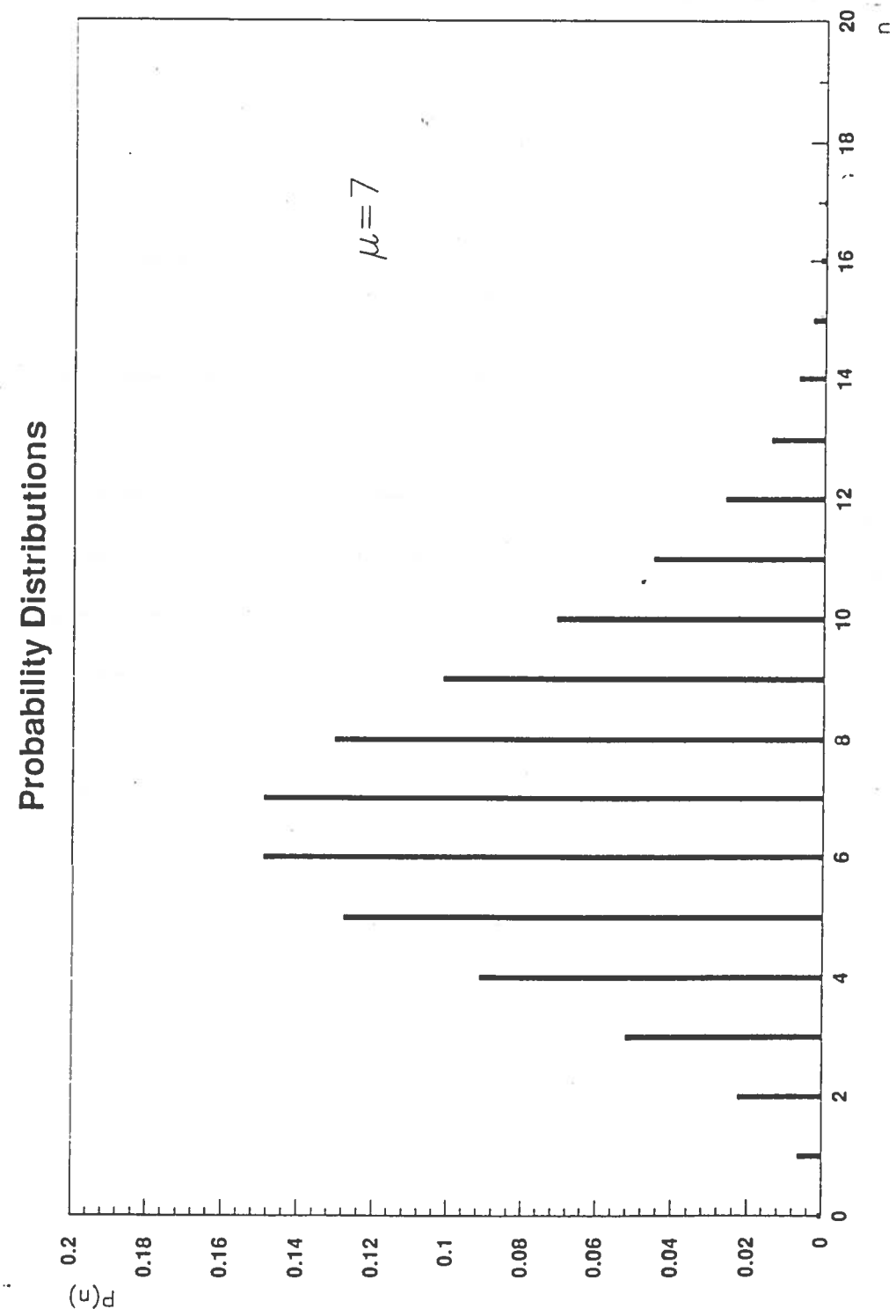

Figure 4: Poisson Distribution

$$\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = e^{-\mu} \qquad \text{and} \qquad \lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^{-k} \to 1 \ .$$

Accordig to these result the Poisson distribution with parameter μ is given by

$$f(k;\mu) = \lim_{n\to\infty} f(k,n;p) = \lim_{n\to\infty} P^{(n)}(\xi = k) = \frac{\mu^k}{k!}e^{-\mu}, \qquad k = 1,2... \tag{19}$$

$$\sum_{k=0}^{\infty}\frac{\mu^k}{k!}e^{-\mu} = e^{-\mu}\sum_{k=0}^{\infty}\frac{\mu^k}{k!} = e^{-\mu}e^{\mu} = 1$$

The mathematical expectation value $E(\xi)$ and the variance $D^2(\xi)$ of the random variable ξ are given by

$$E(\xi) = \sum_{k=0}^{\infty} k\frac{\mu^k}{k!}e^{-\mu} = \mu\sum_{k=0}^{\infty}\frac{\mu^{k-1}}{(k-1)!}e^{-\mu} = \mu \qquad \text{and}$$

$$D^2(\xi) = \sum_{k=0}^{\infty} k^2\frac{\mu^k}{k!}e^{-\mu} - \mu^2 = \mu^2\sum_{k=0}^{\infty}\frac{\mu^{k-2}}{(k-2)!}e^{-\mu} + \mu - \mu^2 =$$

$$\sum_{k=0}^{\infty}(k+(k-1)k)\frac{\mu^k}{k!}e^{-\mu} + \mu - \mu^2 = \mu.$$

2.3.5 Exponential Distribution

The time which passes until a radioactive decay, the number of radioactive atoms, the life time of the parts of an equipment, the propagation of neutrons in a substance, the attenuation of radiation etc. are described by exponential distribution (Fig. 5). The probability of the decay of a radioactive atom in a time interval depends only on the properties of the atom and the length of the time interval and does not depend on the past of the atom. The probability that a particle has not decayed until time t_1 and it will be alive until time t_2 $(t_1 < t_2)$ is denoted by $P(t_1;t_2)$. According to the conditional probability:

$$P(t_1;t_2) = \frac{P(0;t_2)}{P(0;t_1)} \ .$$

The probability depends only on the length of the interval:

$$P(t_2;t_1) = P(t_2 - t_1)$$

and so

$$P(t_2 - t_1)P(t_1) = P(t_2) \ .$$

The solution of this equation is the exponential function:

$$P(t) = e^{-\lambda t} \ .$$

The probability that the particle decays during the time interval t is given by the cumulative distribution function:

$$F(t) = 1 - e^{-\lambda t} \ .$$

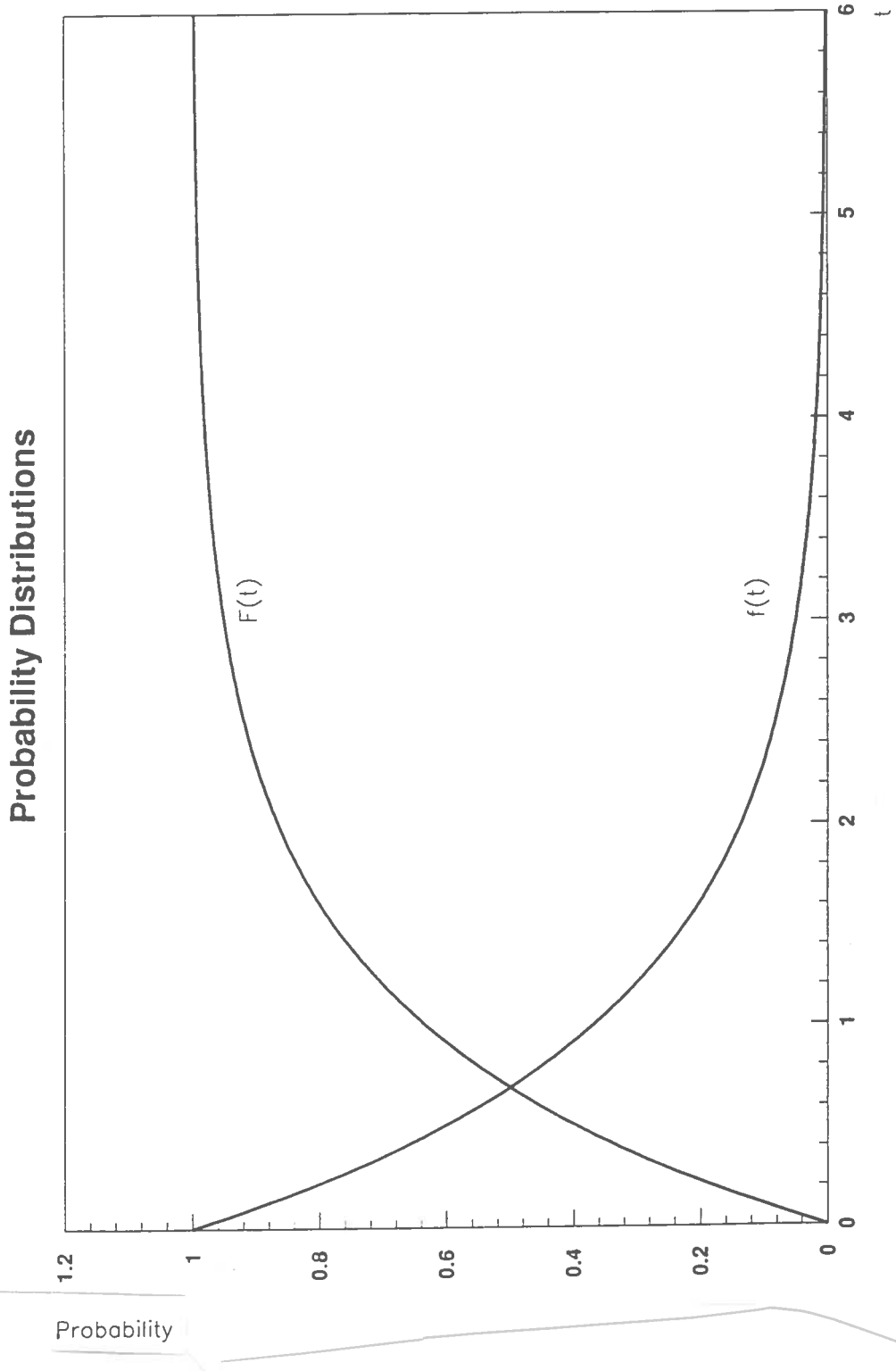

Figure 5: Exponential Distribution

The probability density function $f(t)$ of the decay time may be derived from the cumulative distribution function:

$$f(t) = \lambda e^{-\lambda t} \ . \tag{20}$$

The mathematical expectation value $E(\xi)$ and the variance $D^2(\xi)$ of the random variable ξ, which is distributed according to the exponential distribution are given by

$$E(\xi) = \lambda \int_0^\infty x e^{-\lambda x} dx = \frac{1}{\lambda} \qquad \text{and}$$

$$D^2(\xi) = E(\xi^2) - (E(\xi))^2 = \frac{1}{\lambda^2}.$$

2.3.6 Gamma Distribution

There are n independent random variables $\xi_1, \xi_2, ..., \xi_n$. Each random variable ξ_i is distributed according to an exponential distribution, whose parameter is λ. The random variable

$$\eta = \xi_1 + \xi_2 + ... + \xi_n$$

is distributed according to the gamma distribution with n degrees of freedom (Fig. 6):

$$f(x; \lambda, n) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x}, \qquad 0 < x < \infty \ . \tag{21}$$

The mathematical expectation value $E(\eta)$ and the variance $D^2(\eta)$ of the random variable η are given by

$$E(\eta) = \frac{n}{\lambda} \qquad \text{and}$$

$$D^2(\eta) = \frac{n}{\lambda^2} \ .$$

The continuous form of the gamma distribution with p degrees of freedom is given by

$$f(x; \lambda, p) = \frac{\lambda^p x^{p-1}}{\Gamma(p)} e^{-\lambda x}, \qquad 0 < x < \infty \ , \tag{22}$$

where $\Gamma(p)$ is the gamma function (see Appendix A). The mathematical expectation value $E(\eta)$ and the variance $D^2(\eta)$ of the random variable η are given by

$$E(\eta) = \frac{p}{\lambda} \qquad \text{and}$$

$$D^2(\eta) = \frac{p}{\lambda^2} \ .$$

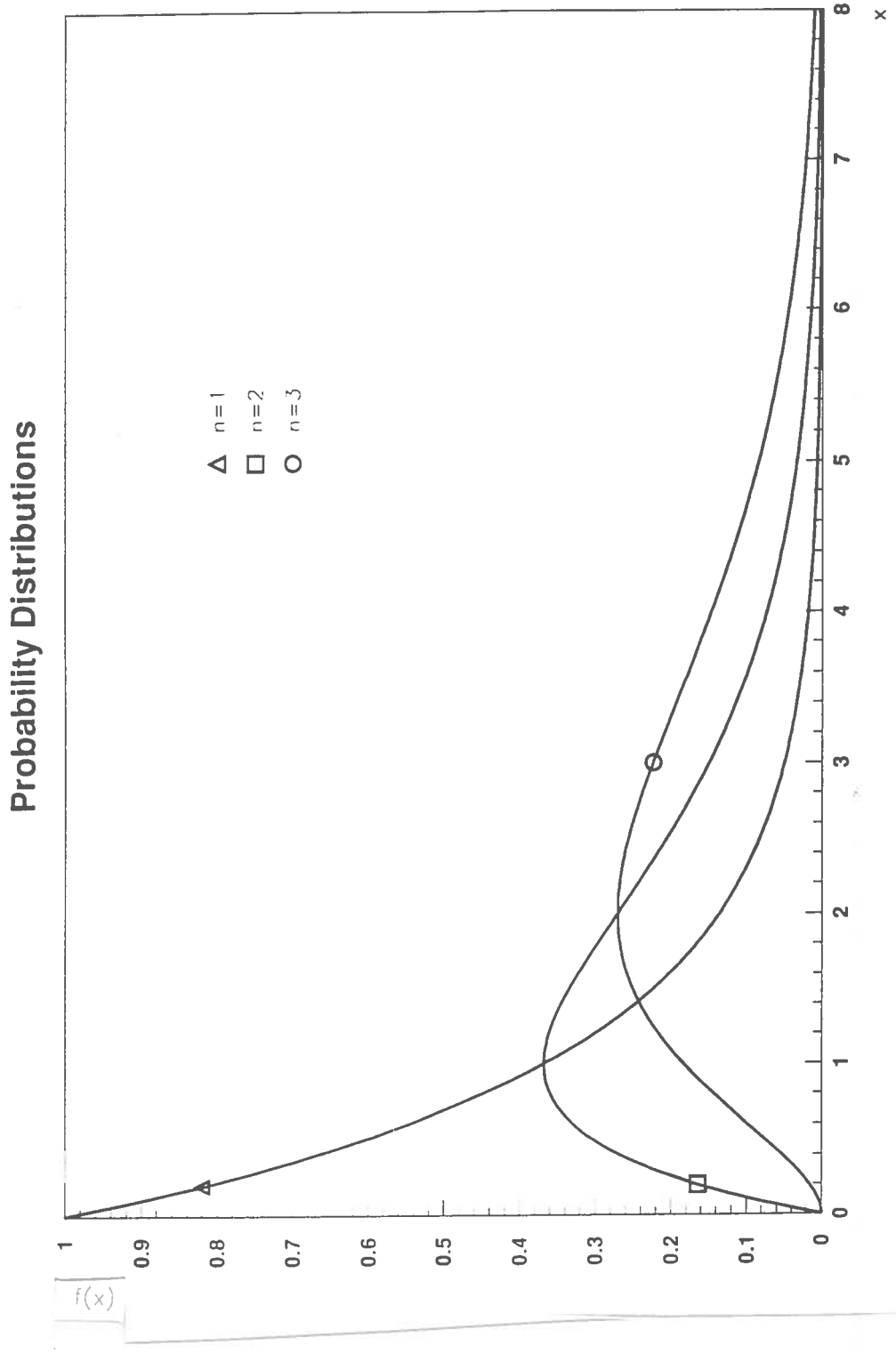

Figure 6: Gamma Distribution

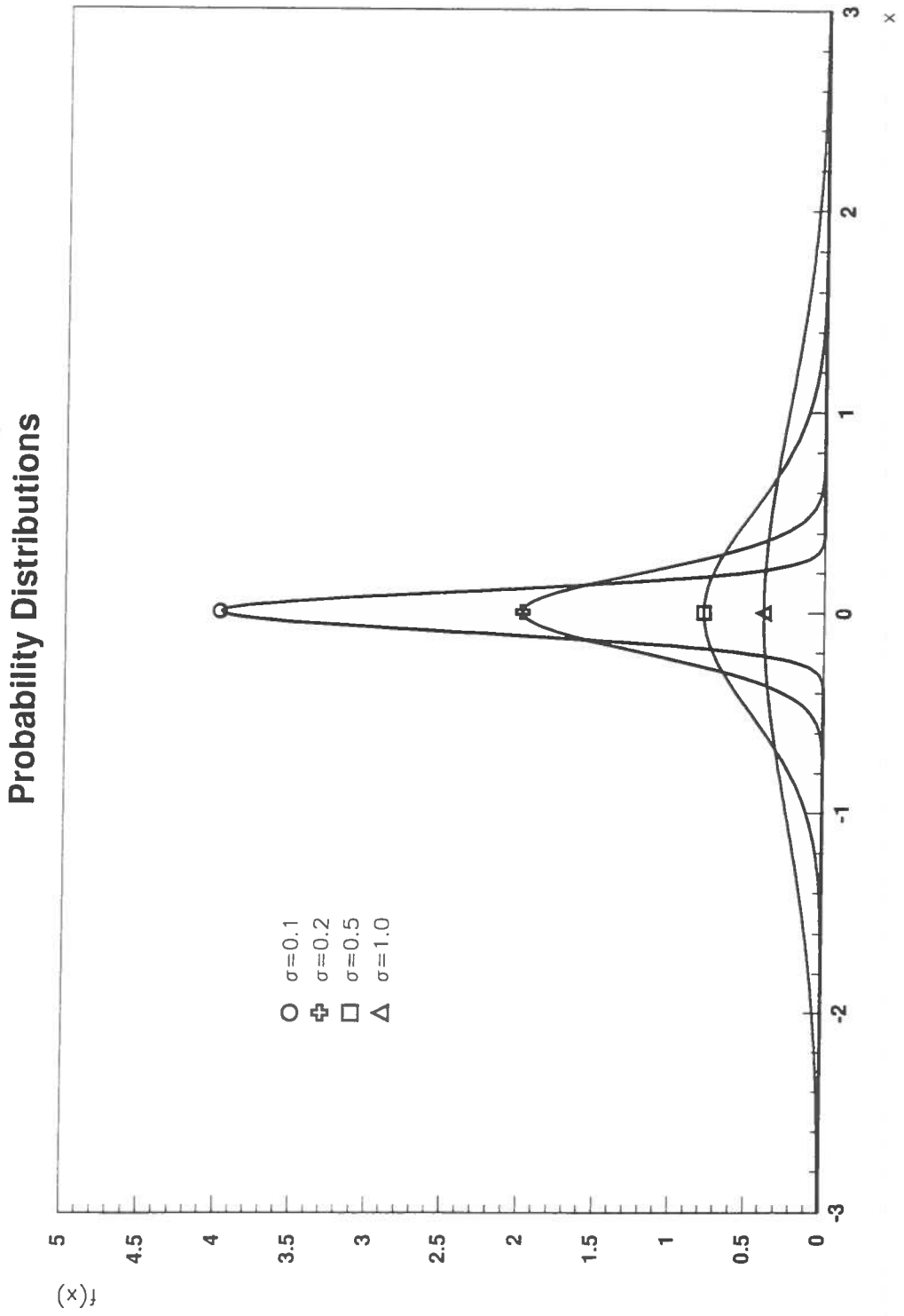

Figure 7: Gaussian Distribution

2.3.7 Gaussian Distribution

The probability density function of the Gaussian or normal distribution is given by

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} , \tag{23}$$

where μ and σ are real and positive constants respectively (Fig. 7). The integral of the probability density function should be equal to 1:

$$I = \int_{-\infty}^{\infty} f(x;\mu,\sigma)dx = 1 .$$

The proof of this is presented below. One has to use the

$$I = \int_{0}^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

equation. I^2 can be written in the following form:

$$I^2 = \int_{0}^{\infty} e^{-x^2} dx \int_{0}^{\infty} e^{-y^2} dy = \int_{0}^{\infty}\int_{0}^{\infty} e^{-(x^2+y^2)} dxdy .$$

Using the $x = r\cos\varphi$, $y = r\sin\varphi$ transformation whose Jacobi determinant is:

$$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r.$$

$$I^2 = \int_{0}^{\pi/2}\int_{0}^{\infty} re^{-r^2} drd\varphi = \frac{\pi}{2}\int_{0}^{\infty} re^{-r^2} dr = \frac{\pi}{4} .$$

Using this result and introducing the new variable $u = \frac{(x-\mu)}{\sqrt{2}\sigma}$, $du = \frac{dx}{\sqrt{2}\sigma}$ the result is given by

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-u^2} du = 1 .$$

The shape of the distribution is independent of the parameter μ and depends on the parameter σ.

$$\max f(x) = f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}.$$

The maximum increases decreasing the σ, but the area below the curve $f(x)$ is unity (Fig. 7). The mathematical expectation value $E(\xi)$ and the variance $D^2(\xi)$ of the random variable ξ are given by

$$E(\xi) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x\exp^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu \qquad \text{and}$$

$$D^2(\xi) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} (x-\mu)^2\exp^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2.$$

The probability that one outcome is in the interval $(\mu - 3\sigma, \mu + 3\sigma)$ is given by

$$P(\mu - 3\sigma < x < \mu + 3\sigma) = \int_{\mu-3\sigma}^{\mu+3\sigma} f(x)dx = 0.997 \ .$$

The normal distribution whose expectation value is $\mu = 0$ and variance $\sigma = 1$ is denoted by $N(0,1)$.

2.3.8 χ^2 Distribution

The $\xi_1, \xi_2, ..., \xi_n$ are independent Gaussian distributed random variables, the sum of them

$$\chi^2 = \sum_{i=1}^{n} (\xi_i - \mu_i)^2/\sigma_i^2$$

is distributed as a χ^2 with n degrees of freedom (Fig. 8). The probability density distribution is given by

$$f(\chi^2; n) = \frac{(\chi^2)^{\frac{n}{2}-1} e^{-\frac{\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \ , \tag{24}$$

where:

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

is the Gamma function (see Appendix A). The mathematical expectation value $E(\chi^2)$ and the variance $D^2(\chi^2)$ of the random variable χ^2 are given by

$$E(\chi^2) = nE(\xi_i^2) = n \qquad \text{and}$$

$$D^2(\chi^2) = nD^2(\xi_i^2) = n(E(\xi_i^4) - (E(\xi_i^2))^2) = n(3-1) = 2n \ .$$

2.3.9 Student's t Distribution

The $\xi_1, \xi_2, ..., \xi_n$ and ξ are independent Gaussian distributed random variables with expectation value $\mu = 0$ and variance $\sigma^2 = 1$. We then define

$$z = \sum_{i=1}^{n} \xi_i^2 \qquad \text{and}$$

$$t = \xi\sqrt{\frac{n}{z}} \ .$$

The variable z belongs to a $\chi^2(n)$ distribution. Then t is distributed according to the Student's t distribution with n degrees of freedom (Fig. 9):

$$f(t; n) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}, \qquad -\infty < x < \infty \ . \tag{25}$$

The mathematical expectation value $E(t)$ and the variance $D^2(t)$ of the random

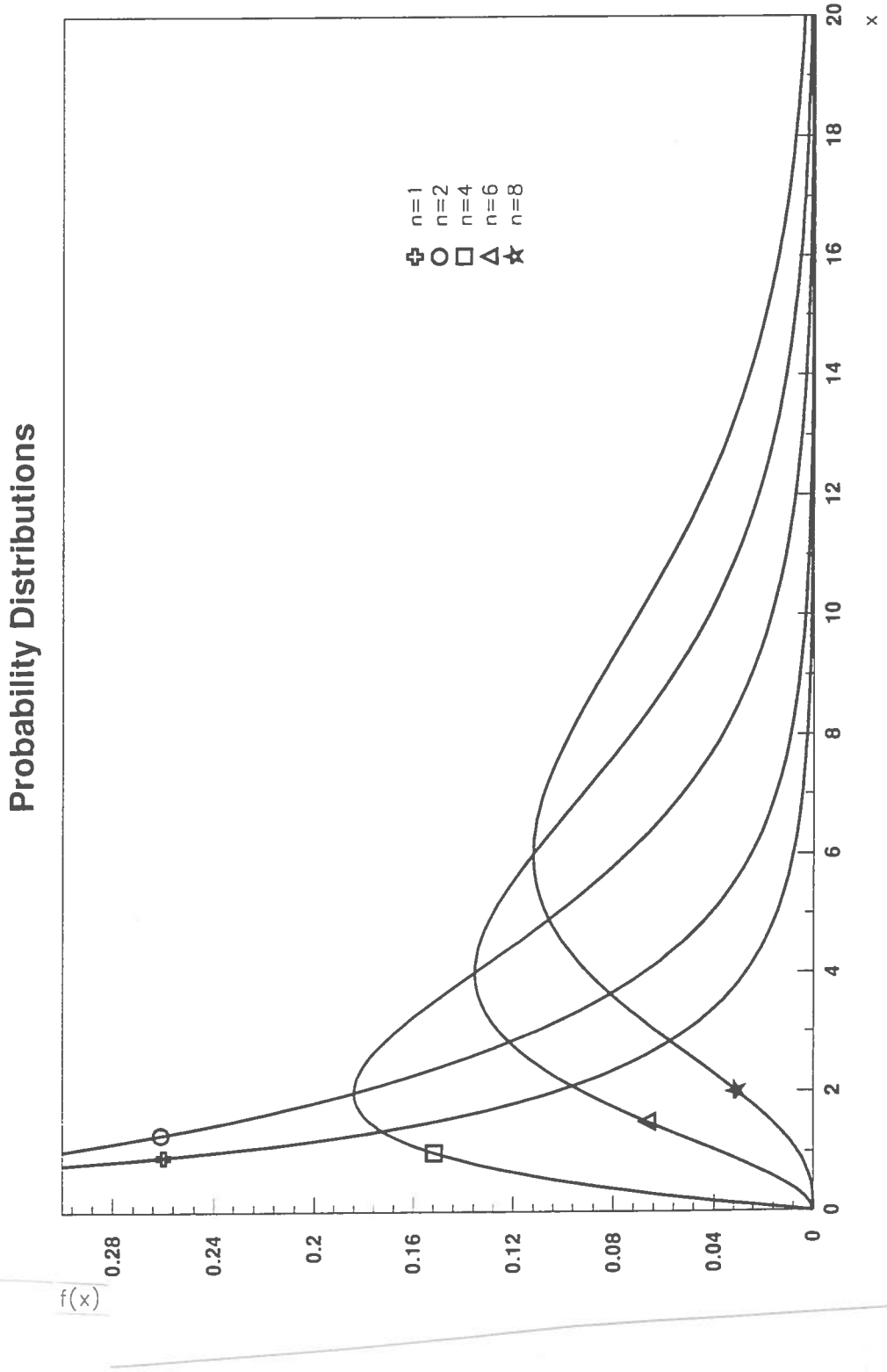

Figure 8: χ^2 Distribution

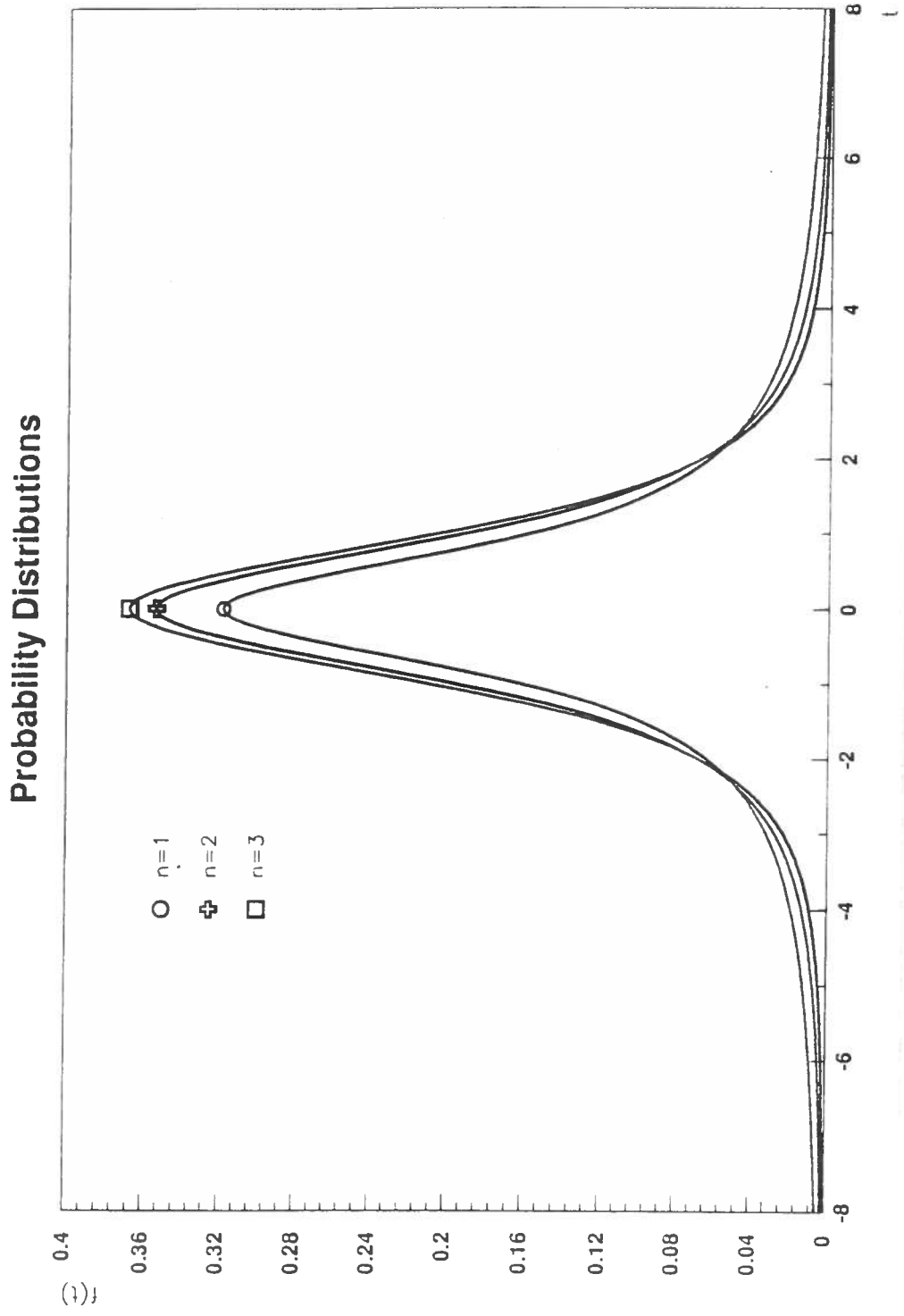

Figure 9: Student's t Distribution

variable t are qiven by

$$E(t) = 0, \qquad \text{for} \qquad n > 1 \qquad \text{and}$$

$$D^2(t) = \frac{n}{n-2}, \qquad \text{for} \qquad n > 2 \ .$$

Student's t distribution resembles a Gaussian distribution with wide tails. As $n \to \infty$ the distribution approches a Gaussian, and if $n = 1$, the distribution is Cauchy or Breit-Wigner. The mean value is finit for $n > 1$ and the variance is finit for $n > 2$, So for $n = 1$ or $n = 2$ the random variable t does not obey the central limit theorem (see Section 2.4).

2.4 Center Limit Theorem

The random variables $\xi_1, \xi_2, ..., \xi_N$ are independent and have the same probability distribution. The mathematical expectation values and the variances are given by

$$E(\xi_1) = E(\xi_2) = ... = E(\xi_N) = m \qquad \text{and}$$

$$D^2(\xi_1) = D^2(\xi_2) = ... = D^2(\xi_N) = b^2 \ .$$

The random variable ρ_N is the sum of these random variables:

$$\rho_N = \xi_1 + \xi_2 + ... + \xi_N .$$

The mathematical expectation value $E(\rho_N)$ and the variance $D^2(\rho_N)$ of the random variable ρ_N are given by

$$E(\rho_N) = E(\xi_1 + \xi_2 + ... + \xi_N) = Nm \qquad \text{and}$$

$$D^2(\rho_N) = D^2(\xi_1 + \xi_2 + ... + \xi_N) = Nb^2 \ .$$

Let us take the Gaussian distribution with parameters $\mu = Nm$ and $\sigma^2 = Nb^2$ and introduce the new random variable:

$$\rho'_N = \frac{\rho_N - \mu}{\sigma} \ .$$

The center limit theorem states, that for any (c, d) interval:

$$\lim_{N \to \infty} P(c \leq \rho'_N \leq d) = \frac{1}{\sqrt{2\pi}} \int_c^d e^{\frac{-x^2}{2}} dx. \tag{26}$$

This is the reason that the Gaussian distribution plays an important role in many fields. According to the Gaussian distribution the probability that ρ_N is in the interval $(Nm - 3b\sqrt{N}, Nm + 3b\sqrt{N})$ is given by

$$P(Nm - 3b\sqrt{N} < \rho_N < Nm + 3b\sqrt{N}) \sim 0.997 \ ,$$

$$P\Big(m - \frac{3b}{\sqrt{N}} < \frac{\rho_N}{N} < m + \frac{3b}{\sqrt{N}}\Big) \sim 0.997 \qquad \text{or}$$

$$P\Big(\Big|\frac{1}{N}\sum_{j=1}^{N} \xi_j - m\Big|\Big) < \frac{3b}{\sqrt{N}} \sim 0.997.$$

3 Random Numbers

In the Monte-Carlo method the random numbers play fundamental role. There are three different types of random numbers:

1. True random numbers. These random numbers are random in the sense of statistics. Any part of the series of the random numbers is independent of the previous ones. These series are unrepeatable.
2. Pseudorandom numbers. These random numbers are generated with some kind of algorithm, so each number depends on the previous numbers but in such a way that any short part of the series resembles the true random numbers in many respects.
3. Quasi random numbers. These series are absolutly not random numbers, but taking a long part of the series for some problems they are better than the true random numbers.

3.1 True Random Numbers

These random number are not predictable. The series of true random numbers are unrepeatable. These random numbers can be generated only by random physical processes. Such physical processes are, for example: the radioactive decay, the thermic noise of electronic equipments, the occurrence of cosmic particles, etc. If one uses true random numbers in the Monte-Carlo method, then the results are exact.

Generation of true random numbers with radioactive source. We count the radioactive decays in a fix interval Δt. The bits of the binary random number is constructed in the followin way.

- Odd number of decays $\rightarrow$ bit 1.
- Even number of decays $\rightarrow$ bit 0.

There is a series of random bits, but $P(0) \neq P(1) \neq 1/2$, $P(0) = p$ and $P(1) = 1-p$. The systematic error of the random number generator can be eliminated with the method as follows. We take consecutive bits b_{2i} and b_{2i+1} which form a bit pair. The bits of the binary numbers are selected as follows:

1. If $b_{2i} = b_{2i+1}$, then take the next bit pair and go to 1.
2. If $b_{2i} \neq b_{2i+1}$, then take b_{2i} as the next bit of the corrected random number and go to 1.

Using this method we lose about 75% of the bits, but the probability of 0 and 1 is the same:

$$p(0) = P(0)P(1) = p(1-p) \qquad \text{and} \qquad p(1) = P(1)P(0) = (1-p)p \ .$$

Problems concerning the true random numbers:

- They need special hardware.
- It takes a long time to generate them.
- The generators may have systematic errors.

3.2 Pseudorandom Numbers

In the Monte-Carlo method the pseudorandom numbers are used generally. They should have similar properties as the true random numbers. They are generated with numerical algorithms, so the series of the pseudorandom numbers are predictable. Most of the algorithms have the form:

$$x_{n+1} = f(x_n) \; . \tag{27}$$

One has to select the function very carefully. The function shown in Fig. 10a is obviously wrong. The points whose coordinates are consecutive numbers of the series

$$(x_1, x_2), \qquad (x_3, x_4), \qquad (x_5, x_6), \qquad \ldots$$

are all on the curve of the function and so are not distributed uniformly on the unit square ($0 < x < 1,\ 0 < y < 1$). The curve of the function should cover the unit square as uniformly as possible (Fig. 10b). The function

$$y = \{gx\} \; , \tag{28}$$

where g is a big number and the fractional part of z is denoted by $\{z\}$. In Fig. 10b g is equal to 28.

Note: It is better to use simply, but well understood function

3.2.1 J. von Neumann's Mid-Square Method

The short description of the mid-square method is the following:

1. Take a four-digit fraction, for example: $\gamma_0 = 0.1234$.
2. Calculate the square of this number: $\gamma_0^2 = 0.01522756$.
3. Take the four consecutive digits from the center as a new four-digit fraction ($\gamma_1 = 0.5227$). Go to 1 to continue.

The beginning of the series is shown in Table 1. The series may be wrong if we do not select the first number carefully (see Table 2).

3.2.2 Multiplicative Congruent Random Number Generator

This random number generator, which is widely spread in Monte-Carlo calculations is based on the procedure:

$$u_{n+1} = (au_n + b) \bmod c \; , \tag{29}$$

where a, b and c are parameters which one has to select very carefully. The properties of these random number generators are:

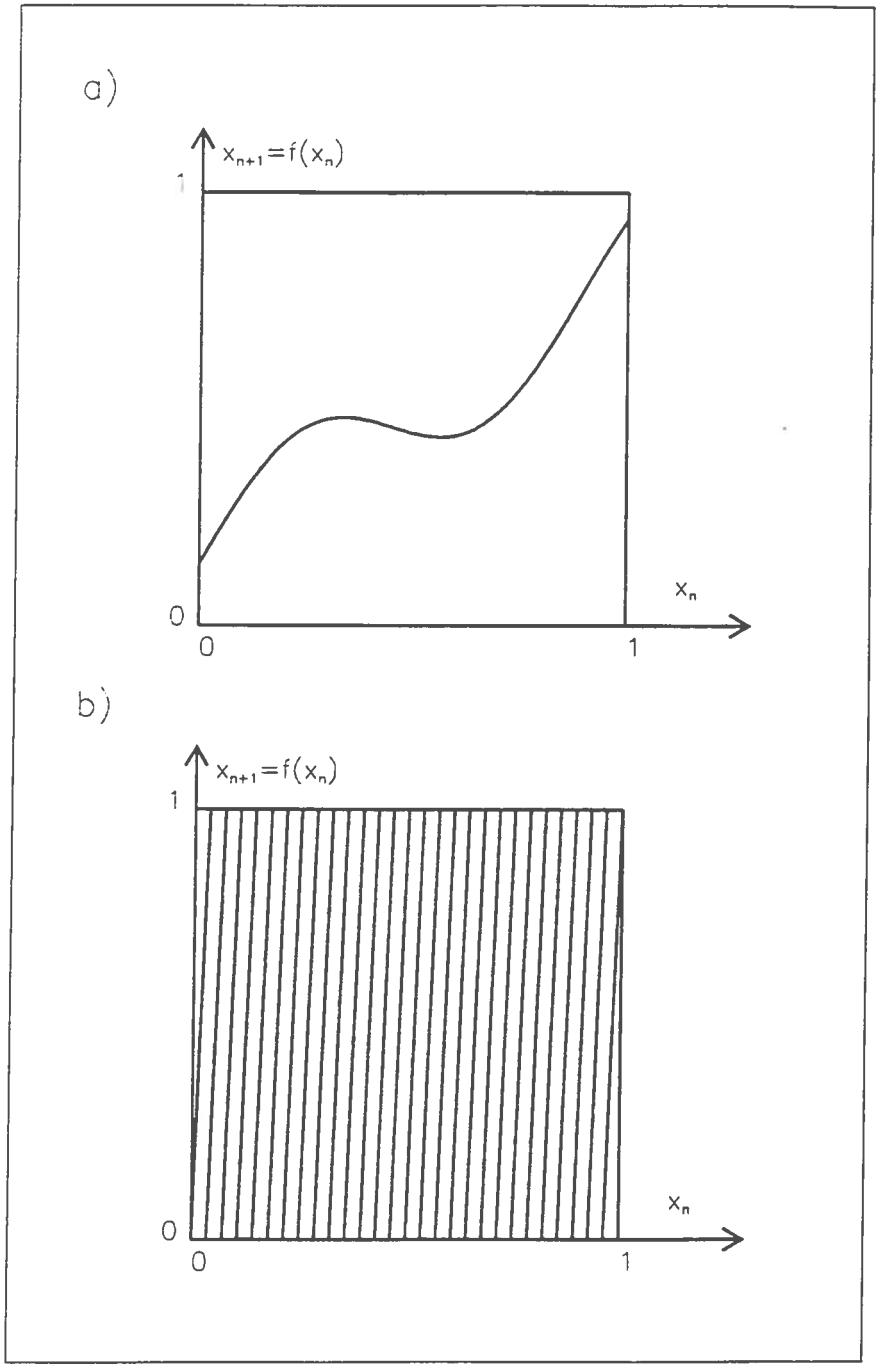

Figure 10: Multilicative Congruent Generator

	$\gamma_0 = 0.1234$
$\gamma_0^2 = 0.01522756$	$\gamma_1 = 0.5227$
$\gamma_1^2 = 0.27321529$	$\gamma_2 = 0.3215$
$\gamma_2^2 = 0.10336225$	$\gamma_3 = 0.3362$
$\gamma_3^2 = 0.11303044$	$\gamma_4 = 0.3030$
$\gamma_4^2 = 0.09180900$	$\gamma_5 = 0.1809$
$\gamma_5^2 = 0.03272481$	$\gamma_6 = 0.2724$
$\gamma_6^2 = 0.07420176$	$\gamma_7 = 0.4201$
$\gamma_7^2 = 0.17648401$	$\gamma_8 = 0.6484$
$\gamma_8^2 = 0.42042256$	$\gamma_9 = 0.0422$
.	
.	
.	

Table 1: J. von Neumann's Mid-Square Method.

2413
1681
4624
3844
7056
0025
0004

Table 2: Wrong series of the mid-square method.

- They are wide-spread and are tested very carefully.
- The constants a, b and c should be selected very carefully.
- The length of the series is less or equal to c.
- $c = 2^m$, where m is the length of the computer word.
- The case $b = 0$ is fast and good enough.
- The long period length might be wrong.

3.2.3 Uniform Random Number Generator

The random numbers of different distributions are generated using the random numbers distributed uniformly on $(0, 1)$. In the first step the uniform random number generator generates random numbers which are distributed uniformly between 0 and the biggest binary integer number which can be represented in a

computer word. (If the length of the computer word is 32 bits, then the biggest integer binary number is $2^{32}-1$.) These integer binary numbers are transformed into the $(0,1)$ interval. (The calculations are faster if one may use integer random numbers.)

The random numbers which are generated in the interval $(0, 2^{32}-1)$ are repeated at the least after 2^{32} numbers. The modern random number generators use the function:

$$f(x) = (ax + b) \bmod c \ , \tag{30}$$

where $c = 2^t$, t is the number of the bits of the computer word and x_0, a and b are integer binary numbers which may be represented in the computer word. Some possible values are given below:

- The value of the first random number x_0 may be optional. It is worth trying $x_0 = 1$ to test the algorithm.
- The value of a should satisfy the following conditions:
  1. $a \bmod 8 = 5$.
  2. $c/100 < a < c - \sqrt{c}$.
  3. The binary digits of a must not have obvious regularity.
- The constant c should be an odd number which satisfy the condition:

$$\frac{b}{c} \sim \frac{1}{2} - \frac{1}{6}\sqrt{3} \sim 0.21132 \ .$$

It is worth to note that the least significant bits are not too "random", so one should use the most significant bits.

3.3 Quasi Random Numbers

There are many Monte-Carlo calculations whose results are indifferent to the correlations between the consecutive random numbers. The quasi random numbers play an important role in the Monte-Carlo integrations. With random numbers, which are random in the statistical sense, the error of the results is proportional to $1/\sqrt{N}$. Using the quasi random numbers one may have errors which are proportional to $1/N$. There are quasi random numbers where the difference of two consecutive numbers is constant. The correlation here is obvious, but the distribution of the numbers is more uniform than that of the true random numbers.

Two quasi random number generators were studied profoundly. These random number generators are used in the Monte-Carlo integration. One may reach with them faster convergence than $1/N$.

3.3.1 Richtmyer formula

The value of the i-th random number of the j-th generator is calculated with the formula:

$$r_{ij} = iS_j \bmod 1 \ , \tag{31}$$

where S_j is the square root of the j-th prim number. So the difference between two consecutive numbers is S_j. This means a strong correlation.

3.3.2 Van der Corput formula

This procedure takes the integer number of a given numerical system whose base is b. In the binary system $b = 2$. Then it reverse the order of the digits and constructs a fraction puting the decimal point before the number as it is shown in the Table 3. One may decrease the short range correlations with 'mixing' methods.

decimal	binary	binary fraction	decimal fraction
1	0001	0.1000	0.5000
2	0010	0.0100	0.2500
3	0011	0.1100	0.7500
4	0100	0.0010	0.1250
5	0101	0.1010	0.6250
6	0110	0.0110	0.3750
7	0111	0.1110	0.8750
8	1000	0.0001	0.0625

Table 3: Van der Corput formula.

4 Transformations

To generate the values of random variables there were built a special equipments at the very beginning of the use of the Monte-Carlo method. To generate the values of the discrete random variable ξ, whose probability distribution is given by the table

$$\xi = \begin{Bmatrix} x_1, & x_2, & x_3, & x_4 \\ 0.5, & 0.25, & 0.125, & 0.125 \end{Bmatrix}, \tag{32}$$

one has to build a roulett as it is shown in Fig. 11a. It was find out very soon that it is enough to generate random numbers distributed uniformly on the interval $(0, 1)$. One may generate random number of any probability distribution using these uniform random numbers. The methods with calculates the values of a given random variable ξ using one or more random numbers distributed uniformly on $(0, 1)$ are called the generation of the random number ξ.

4.1 General Method

The probability density function $f(x)$ of the random variable ξ may have any value from the (a, b) interval. The cumulative distribution function is given by

$$F(x) = \int_a^x f(x')dx'$$

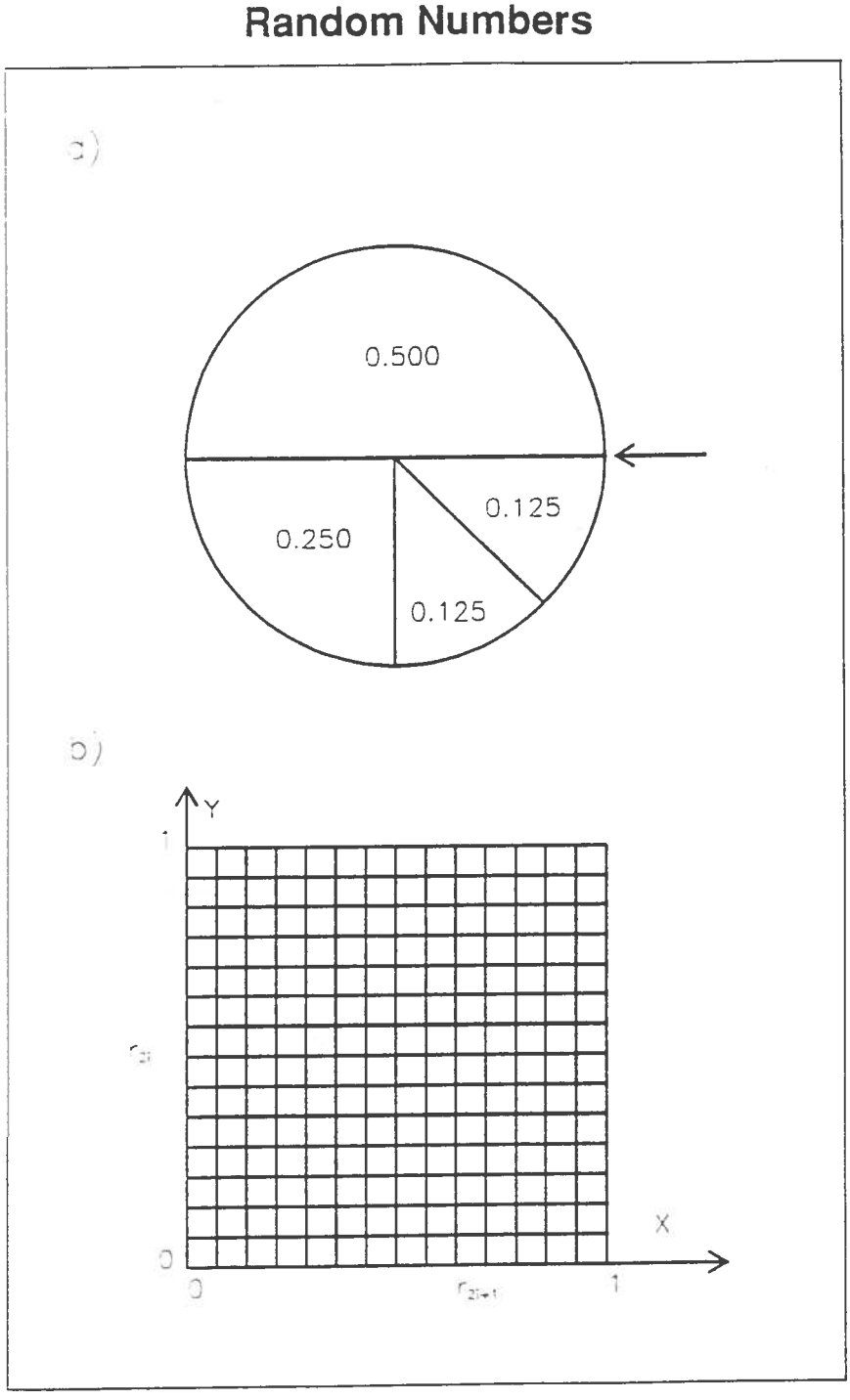

Figure 11: Generation and Test

and its inverse function is $x = F^{-1}(y)$. The values of the random variable ξ are generated with the equation

$$x = F^{-1}(u) ,$$

where u is a random number distributed uniformly on $(0,1)$. The so generated x values are distributed according to the $f(x)$ probability density function. The proof of this and a lot of examples are given in Section 6.

5 Test of Random Numbers

It is not rare that unforseen correlations introduce non-negligible errors in the Monte-Carlo simulations. A useful test for this is to recompute the same results with different algorithms for the pseudorandom numbers.

5.1 General Method

1. Take n consecutive random numbers $\bar{u} = (u_1, u_2, ..., u_n)$
2. Construct a function of these random numbers.
3. Calculate the value of the function in the case of true random numbers.
4. Confront the results.

5.2 A Useful Test

Take pairs of random numbers. Suppose they are the rectangular coordinates (x, y) of points on the xy-plane. Study the distribution of the points on a grid of quadrats.

$$\chi^2 = \sum_{quadrats} \frac{(n_k - \mu_k)^2}{\mu_k^2} , \tag{33}$$

where μ_k is the mathematical expectation value and n_k is the number of points in the k-th quadrat (Fig. 11b).

6 Monte-Carlo Techniques

6.1 Discrete Random Variable

The probability distribution of the discrete random variable ξ is given by the table:

$$\xi = \begin{Bmatrix} x_1, x_2, ..., x_n, ... \\ p_1, p_2, ..., p_n, ... \end{Bmatrix} .$$

We devide the $(0,1)$ interval with the $n+1$ points, $y_0 = 0$, $y_1 = p_1$, $y_2 = p_1 + p_2$, ..., $y_n = \sum_{i=1}^{n} p_i = 1$, into n subintervals. So the length of the i-th subinterval is equal to p_i. We generate a random number u which is distributed uniformly on $(0,1)$. The probability that the value u is in the (y_{i-1}, y_i) subinterval is given by

$$P(y_{i-1} < u \leq y_i) = p_i . \tag{34}$$

To generate the values of the discrete random variable ξ one may follow the following procedure.

Generation of Discrete Random Variable

1. Initialize by storing the possible values x_i in an array X and storing the y_i, $i = 1, 2, ..., n$ values in an array Y.

$$y_0 = 0, \qquad \text{and}$$

$$y_i = \sum_{j=1}^{i} p_j, \qquad i = 1, 2, ..., n \ .$$

2. Generate a random number u and find the smallest y_i value which is equal to or greater than u. Use the fast binary search (see Appendix B).
3. Use the index of the array element y_i, determined in the previous step, to return the x_i value of the random variable ξ. On the next call go to 2.

6.2 Continuous Random Variable

The continuous random variable ξ may have any value from the (a, b) interval. Its probability density function is $f(x)$. The cumulative distribution function $F(x)$ is given by

$$F(x) = \int_a^x f(x')dx' \ .$$

The cumulative distribution function has the following properties:

$$F(a) = 0, \quad F(b) = 1 \qquad \text{and} \qquad F'(x) = f(x) > 0.$$

So the cumulative distribution function is a monotone increasig function:

$$F(x_1) < F(x_2) \ , \qquad \text{if} \qquad x_1 < x_2 \ .$$

Using the uniform random number generator one may generate the values of the random variable ξ with the equation

$$F(\xi) = \int_a^{\xi} f(x)dx = u \ , \tag{35}$$

where u is the value of the random number distributed uniformly on $(0, 1)$. Each $F = u$ line $(0 \leq u \leq 1)$ and the $F(x)$ function have only one intersection point (Fig. 12a). The abscissa of this intersection point is one value of the randon variable ξ. So the $F(x) = u$ equation has only one solution. Let us take a subinterval (c, d) of the interval (a, b) so $a \leq c \leq d \leq b$ (Fig. 12b). The values of the function $F(x)$ are in the interval $F(c) < F(x) < F(d)$ if the values of ξ are in the (c, d) subinterval $c < x < d$ and vice versa. And so

$$P(c < \xi < d) = P(F(c) < u < F(d)) \ .$$

The random number u is uniform on $(0, 1)$, therefore

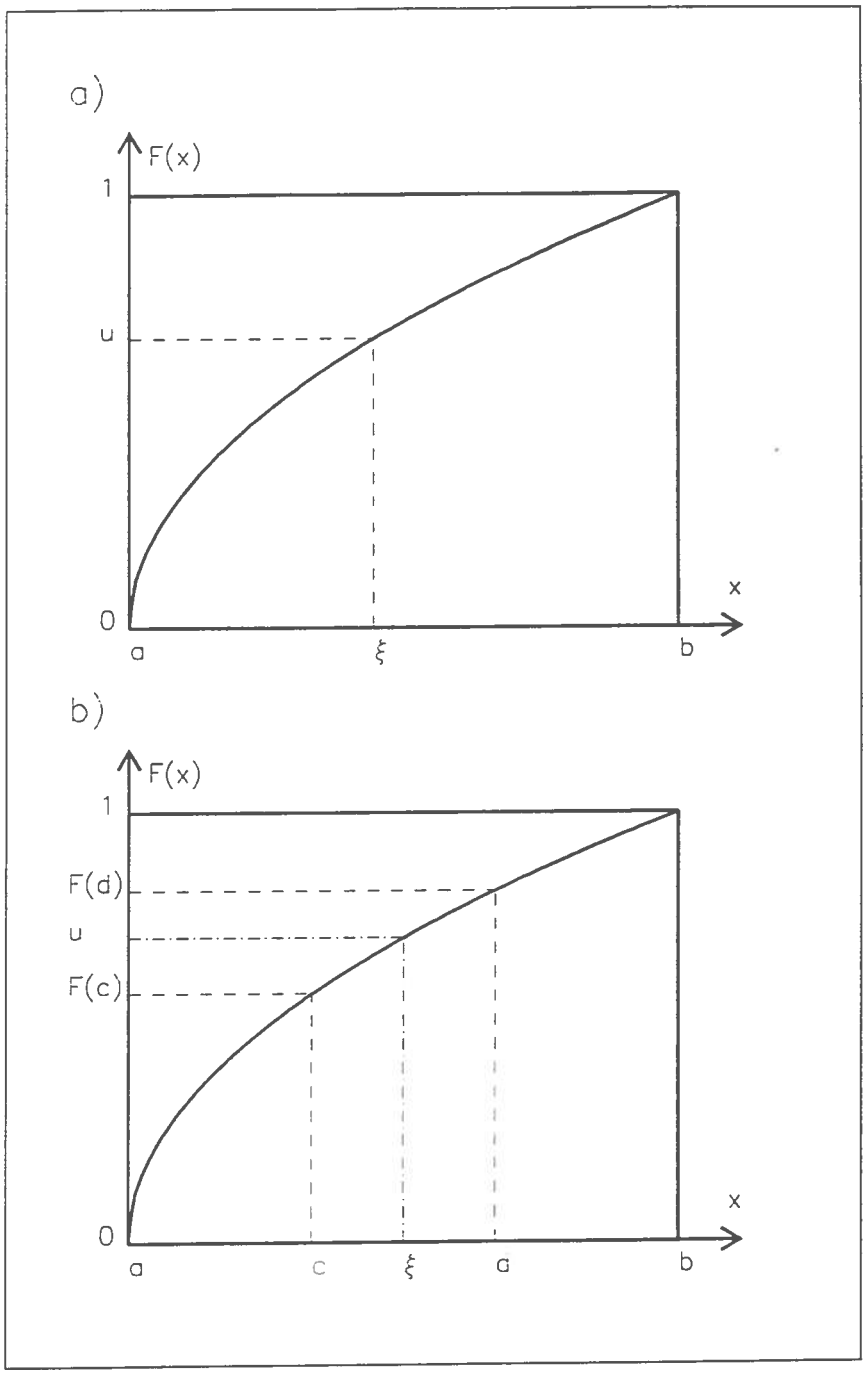

Figure 12: Continuous Random Variable

$$P(F(c) < u < F(d)) = F(d) - F(c) = \int_c^d f(x)dx \ .$$

Finally from these equations we get

$$P(c < \xi < d) = \int_c^d f(x)dx \ .$$

This means that the distribution of the generated values is described by the $f(x)$ probability density function.

Generation of Continuous Random Variable The random variable ξ is continuous and may have any value from the (a, b) interval and its probability density function $f(x)$ may be integrated analytically. To generate a value of the random variable one has to solve the equation

$$F(\xi) = \int_a^\xi f(x)dx = u \ ,$$

where u is a random number distributed uniformly on $(0, 1)$. The value of ξ is given by

$$\xi = F^{-1}(u) \ .$$

When the (c, d) interval is only a subinterval of the whole (a, b) interval from which ξ may have its values, then one has to solve the equation

$$\frac{1}{\int_c^d f(x)dx} \int_c^\xi f(x)dx = \frac{1}{F(d) - F(c)} \int_c^\xi f(x)dx = u \ .$$

The value of ξ is given by

$$\xi = F^{-1}(F(c) + (F(d) - F(c))u) \ .$$

Linear Interpolation The value u of the uniform random variable may be transformed into the value of the random variable ξ with the algorithm described below. The vector $\bar{x}$ has $n + 1$ elements: $x_0, x_1, x_2, ..., x_n$. The value assigned to the i-th element is given by the solution of the equation

$$\frac{i}{n} = \int_a^{x_i} f(x')dx' \ .$$

Obviously x_0 is equal to a and x_n is equal to b. To get the next value of the random variable ξ one has to generate the next value of the of the uniform random variable u and multiply it by n:

$$r = nu \ .$$

the integer part of r is: $i = [r]$ and the fractional part is $\Delta = r - [r]$. One get the next value of the random variable x by interpolating between the elements of the vector $\bar{x}$:

$$x = x_i + (x_{i+1} - x_i)\Delta \ .$$

The accuracy of the interpolation depends on the number of the elements of the vector $\bar{x}$ and on the shape of the probability density function $f(x)$.

6.2.1 J. von Neumann's Acceptance-Rejection Method

Very commonly an analytic form of $F(x)$ is unknown or too complex to work with. We suppose that in the (a, b) interval any given value of x the probability density function $f(x)$ can be computed and the function has an upper limit

$$f(x) \leq M \ ,$$

so one may enclose the whole function in a rectangle (Fig. 13a). One may generate the values of the random variable ξ which has the probability density function $f(x)$ with the following algorithm:

1. Generate two random numbers u_1 and u_2 distributed uniformly on (0,1). Calculate the coordinates which are distributed uniformly on the rectangle:
$$x = a + (b - a)u_1 \qquad \text{and} \qquad y = Mu_2 \ .$$

2. If $y \leq f(x)$, then return $\xi = x$, which is distributed according to the $f(x)$ distribution. If $y > f(x)$, then go to 1.

The generated points $Q(x, y)$ are distributed uniformly on the rectangle. The area of the rectangle is $M(b - a)$. The probability that the point is below the function $f(x)$ is given by the ratio of the areas (Fig. 13a):

$$\frac{\int_a^b f(x)dx}{M(b-a)} = \frac{1}{M(b-a)} \ .$$

The probability that the point is in the subinterval (x_1, x_2) below the function $f(x)$ is given by

$$\frac{\int_{x_1}^{x_2} f(x)dx}{M(b-a)} \ .$$

The ratio of this value and the value of the all possible values is given by

$$\frac{\int_{x_1}^{x_2} f(x)dx}{M(b-a)} : \frac{1}{M(b-a)} = \int_{x_1}^{x_2} f(x)dx \ ,$$

this is the probability to have an event in the subinterval (x_1, x_2).

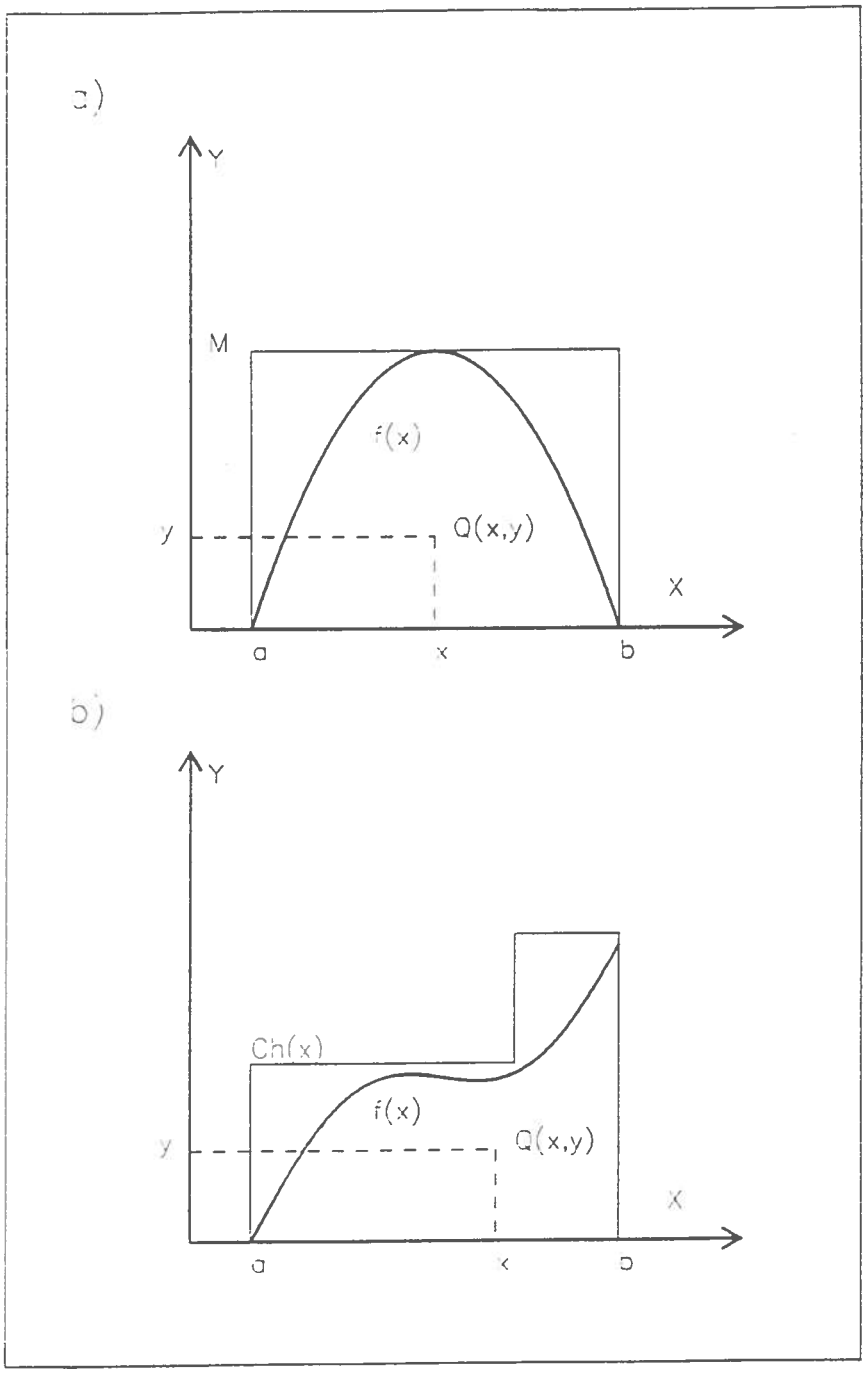

Figure 13: Acceptance and Rejection Method

6.2.2 Importance Sampling Method

The random variable ξ has the probability density function $f(x)$ and may have any value from the interval (a, b). The function $h(x)$ is uniform or is a normalized sum of uniform distributions:

$$\int_a^b h(x)dx = 1 \ ,$$

where the function $h(x)$ may have any value from (a, b). One may generate the values of ξ with the algorithm described below:

1. Generate the abscissa x according to the probability density function $h(x)$ (see Section 6.2).
2. Calculate the ordinate of the point.

$$y = Ch(x)u_2 \ .$$

3. Calculate $f(x)$ and test if $f(x) \leq Ch(x)u_2$. If so accept x; if no reject x and try again (go to 1).

If we regard x and $u_2Ch(x)$ as the abscissa and ordinate of a point (Fig. 13b) in the two-dimensional plot, then we accept the points which fall under $f(x)$. The efficiency is the ratio of the areas, which is equal to $1/C$; therefore we must keep C as close to 1.0 as possible.

This method is called importance sampling, because we generate more trials of x in the region where $f(x)$ is most important.

The random variable ξ has the probability density function $g(x)$. We take a random variable η whose probability density function $f(x)$ is approximated good with the probability density function $g(x)$. Both the cumulative density function $y = F(x)$ and the inverse of the cumulative distribution function $x = F^{-1}(y)$ exist. To generate the values of the random variable ξ one has to use the following procedure:

1. Generate one value of x of the random variable η, with the usual methods, which has the probability density function $f(x)$.
2. Generate a random number z which is distributed uniformly on the $(0, (g(x)/f(x))_{max})$ interval.
3. If $z \leq g(x)/f(x)$, then use the value x which is distributed according to $g(x)$. If $z > g(x)/f(x)$, then reject x and try again (go to 1).

The proof of this method one may find in Section 7.

6.2.3 Two-Dimensional Acceptance-Rejection Method

The two-dimensional random vector variable (ξ, η), whose probability density function is $f(x, y)$, may have any value from the rectangular region $(a < x < b,\ c < y < d)$. The maximum value of the probability density function is M:

$$f(x, y) \leq M, \qquad a < x < b \qquad \text{and} \qquad c < y < d \ .$$

One may generate the values of the random vector variable (ξ, η) whose probability density function is $f(x, y)$ with the following procedure:

1. Generate three random numbers u_1, u_2 and u_3 which are distributed uniformly on $(0, 1)$. Calculate the random values:

$$x = a + (b - a)u_1 \qquad y = c + (d - c)u_2 \qquad \text{and} \qquad z = Mu_3 \ .$$

2. If $z \leq f(x, y)$, the return $(\xi = x, \eta = y)$, which is distributed according to the $f(x, y)$. If $z > f(x, y)$, then reject x and y and try again (go to 1).

6.3 Algorithms

6.3.1 Uniform Distribution

The available random number generators sacrifice randomness in favour of speed. It is not rare that unforseen correlations will introduce non-negligible errors in the results. A useful test for this is to recompute the same result with different algorithm for the pseudorandom numbers. To prove the performance of an existing generator one may use the Bays-Durham algorithm:

1. Initialize by generating and storing N random numbers in an array v, using the available random number generator. Generate a new random number u and save it.
2. Use the u as an address $j = 1 + [Nu]$ to select v_j as the random number to be returned. Also save this v_j as u for the next call. Replace v_j in the array with a new random number using the available generator. On the next call, go to 2.

6.3.2 Uniform Distribution on Interval

Generation of uniform random numbers on the (a, b) interval. The probability density function of the random variable η, whose distribution is uniform on the (a, b) interval (see Section 6.3.1), is given by

$$f(x) = \frac{1}{b - a} \ , \qquad a < x < b \ .$$

To generate a value of the random variable one has to use the equation

$$\int_a^\eta \frac{1}{b - a} dx = u \ ,$$

where u is the uniform random variable on $(0, 1)$. Finally the value of η is given by

$$\eta = a + (b - a)u \ . \tag{36}$$

6.3.3 Uniform Distribution on Square

The coordinates of the $Q(x,y)$ points are independent, so the relation of the probability density functions is given by

$$f(x,y) = h(x)g(y) \ .$$

The probality density functions are $h(x) = 1$ and $g(y) = 1$ and so $f(x,y) = 1$. This means that the points $Q(x,y)$ are distributed uniformly on the unit square $(0 < x < 1,\ 0 < y < 1)$.

6.3.4 Uniform Distribution in Circle

Using the polar coordinates r and φ, the surface element is given by

$$dS = rdrd\varphi \ .$$

The point $P(x,y)$ should have distributed in the circle uniformly. The probability that a uniformly distributed point is on the surface element $dS = rdrd\varphi$ is given by

$$f(r,\varphi)dS = f(r,\varphi)rdrd\varphi = \frac{dS}{\int_0^{2\pi}\int_0^R rdrd\varphi} = \frac{rdrd\varphi}{R^2\pi} \ .$$

the random variables r and φ are independent, $f(r,\varphi) = f_r(r)f_\varphi(\varphi)$:

$$\frac{rdrd\varphi}{R^2\pi} = \frac{d\varphi}{2\pi}\frac{2rdr}{R^2} \ .$$

To get one value of the random variable φ, one has to solve the equation

$$\frac{1}{2\pi}\int_0^\varphi d\varphi = u \ ,$$

where u is distributed uniformly on $(0,1)$:

$$\varphi = 2\pi u \ .$$

To generate one value of the random variable r one has to solve the equation

$$\frac{2}{R^2}\int_0^r rdr = u \ ,$$

where u is distributed uniformly on $(0,1)$:

$$r = \sqrt{R^2 u} \ .$$

The rectangular coordinates of the point are given by

$$x = r\cos\varphi \qquad \text{and} \qquad y = r\sin\varphi \ . \tag{37}$$

There are faster algorithms which do not use the square root, sine and cosine functions

1. Generate two random numbers u_1 and u_2 which are uniformly distributed on $(0,1)$. Construct the random numbers $v_1 = 2u_1 - 1$ and $v_2 = 2u_2 - 1$ which are uniformly distributed on $(-1,1)$.
2. Calculate $r^2 = v_1^2 + v_2^2$ and if $r^2 > 1$, then go to 1. If $r^2 \leq 1$, then the coordinates of the point are given by

$$x = Rv_1 \qquad \text{and} \qquad y = Rv_2 \ .$$

In this algorithm the generated points (u_1, u_2) are distributed uniformly on a unit square, so about 21.5% is outside the circle.

6.3.5 Random Direction on Plane

The start point of the unit vector is in the origin of the rectangular coordinate system. The end point moves on a circle. If the direction is distributed uniformly, then the end point is distributed uniformly on the circle too. The probability that the end point is on the arc element ds is given by $ds/(2\pi)$. Let us take the azimuthal angle φ $(0 \leq \varphi < 2\pi)$. The arc element is

$$d\varphi = ds \ .$$

The probability density function $f(\varphi)$ is given by

$$f(\varphi)d\varphi = \frac{d\varphi}{2\pi} \ .$$

One may generate the random angle φ uniformly distributed on $(0, 2\pi)$ with the random number u distributed uniformly on (0,1) with the equation:

$$\varphi = 2\pi u \ ,$$

The components of the unit vector are given by

$$x = \cos\varphi \quad \text{and} \quad y = \sin\varphi \ . \tag{38}$$

6.3.6 Sine and Cosine of Random Angle

On the $(0, 2\pi)$ interval uniformly distributed angle (α) is generated with a random number which is distributed uniformly on $(0,1)$ with the equation

$$\alpha = 2\pi u \ .$$

The sine (S) and the cosine (C) of the uniformly distributed angle are given by

$$S = \sin\alpha \quad \text{és} \quad C = \cos\alpha \ . \tag{39}$$

Faster algorithms do not use the circular function because they are slow procedures:

1. Generate the random numbers u_1 and u_2. They are uniform on $(0,1)$. Then $v_1 = 2u_1 - 1$ is uniform on $(-1,1)$ and $v_2 = u_2$ is uniform on $(0,1)$.
2. If $r^2 = v_1^2 + v_2^2 > 1$, go to 1. Otherwise the sine (S) and the cosine (C) of a random angle are given by

$$S = \frac{2v_1v_2}{r^2} \qquad \text{and} \qquad C = \frac{(v_1^2 - v_2^2)}{r^2} \ .$$

The random angle is uniform on $(0, 2\pi)$.

6.3.7 Random Directions

The start point of the unit vector is in the origin of the rectangular coordinate system. If the direction of the unit vector is distributed uniformly, then the end point of the unit vector is distributed uniformly on a spherical surface ($r = 1$). the probability that the end point is in the surface element dS is $dS/(4\pi)$. In the polar coordinate system the surface element is given by

$$dS = \sin\vartheta d\vartheta d\varphi$$

where $0 \leq \vartheta \leq \pi$ and $0 \leq \varphi < 2\pi$. The probability density function of the end point $Q(\vartheta, \varphi)$ of the unit vector is given by

$$f(\varphi, \vartheta)d\varphi d\vartheta = \frac{dS}{4\pi} = \frac{\sin\vartheta d\vartheta d\varphi}{4\pi} \ .$$

Knowing the joint distribution function of ϑ and φ

$$f(\vartheta, \varphi) = \frac{\sin\vartheta}{4\pi} \ ,$$

one can determine easily the probability density function of ϑ and φ:

$$f_\vartheta(\vartheta) = \int_0^{2\pi} f(\varphi, \vartheta)d\varphi = \frac{\sin\vartheta}{2} \qquad \text{and} \qquad f_\varphi(\varphi) = \int_0^{\pi} f(\varphi, \vartheta)d\vartheta = \frac{1}{2\pi} \ .$$

The equation $f(\varphi, \vartheta) = f_\varphi(\varphi) f_\vartheta(\vartheta)$ shows that ϑ and φ are independent. We may see also that φ is distributed uniformly on $(0, 2\pi)$. To generate φ we use the quation

$$\varphi = 2\pi u \ ,$$

where u is distributed uniformly on $(0, 1)$. To generate ϑ we use the equation

$$\frac{1}{2}\int_0^{\vartheta} \sin x dx = u \ ,$$

where u is distributed uniformly on $(0, 1)$. From the equation ϑ is given by

$$\cos\vartheta = 1 - 2u \qquad \text{and}$$

$$\vartheta = \arccos(1 - 2u) \ .$$

The components of the unit vector are given by

$$x = \sin\vartheta\cos\varphi, \qquad y = \sin\vartheta\sin\varphi \qquad \text{and} \qquad z = \cos\vartheta \ . \tag{40}$$

6.3.8 Binomial Distribution

The random variable k, which is distributed according to the binomial distribution (see Section 2.3.3), one may generate with the following algorithm:

1. Construct an array B whose elements are

$$B_0 = 0 \ ,$$

$$B_{i+1} = \sum_{k=0}^{i} f(k;n,p) \ ,$$

where:

$$f(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k} \ .$$

The relation of two consecutive elements is given by

$$f(0;n,p) = (1-p)^n \qquad \text{and}$$

$$f(k+1;n,p) = f(k;n,p)\frac{n-k}{k+1}\frac{p}{1-p} \ .$$

2. Generate a random number u which is distributed uniformly on $(0,1)$.
3. Search for the subinterval

$$B_i < u \leq B_{i+1}$$

and return the $k = i-1$ value which is distributed according to the binomial distribution. On the next call go to 2.

One may organize the algorithm in such a way that the elements of the array B are calculated only if they are needed.

6.3.9 Poisson Distribution

The random variable n, which is distributed according to the Poisson distribution (see Section 2.3.4), one may generate with the following algorithm:

1. Generate a random number u which is distributed uniformly on $(0,1)$.
2. Search for the subinterval

$$P_i < u \leq P_{i+1}$$

If necessary, fill the array P up to the index $i+1$ with the corresponding values

$$P_0 = 0 \ ,$$

$$P_{n+1} = \sum_{k=0}^{n} f(k;\mu) \ ,$$

where:

$$f(n;\mu) = \frac{\mu^n e^{-\mu}}{n!} \ .$$

The relation of two consecutive elements is given by

$$f(0;\mu) = e^{-\mu} \qquad \text{and}$$

$$f(n+1;\mu) = f(n;\mu)\frac{\mu}{n+1} \ .$$

Return the $n = i-1$ value which is distributed according to the Poisson distribution. On the next call go to 2.

In this algorithm the elements of the array P are calculated only if they are needed.

6.3.10 Exponential Distribution

To generate the value x of the random variable ξ, which is distributed according to the exponential distribution (see Section 2.3.5) one may use the equation

$$\int_0^x \lambda e^{-\lambda y} dy = u \ ,$$

where u is a random number distributed uniformly on $(0,1)$. The value of ξ is given by

$$x = -\frac{1}{\lambda} \ln u \ .$$

The distribution of u and $1-u$ are the same.

Examples of Exponential Distributions

1. The decay time t of a radioactive atom is given by

   $$t = -\tau \ln(u) \ ,$$

   where τ is the expectation value of the decay time. The random numbers u and $1-u$ have the same probability density function.

2. The distance made by a decaying particle, whose momentum is p and decay time is τ, may be generated with the equations:

   $$l = tv = t\frac{p}{m_0} \qquad \text{and}$$

   $$l = -\frac{\tau p}{m_0} \ln(u) \ .$$

3. The decay time t of a radioactive particle in the subinterval (a,b) of the whole $(0,\infty)$ interval one may generate with the equations

   $$u = \frac{\frac{1}{\tau}\int_a^t e^{-\frac{t}{\tau}}}{\frac{1}{\tau}\int_a^b e^{-\frac{t}{\tau}}},$$

   $$\left[-e^{-\frac{t}{\tau}}\right]_a^t = \left[e^{-\frac{a}{\tau}} - e^{-\frac{b}{\tau}}\right]u \qquad \text{and so}$$

   $$t = -\tau \ln\left[e^{-\frac{a}{\tau}} + \left(e^{-\frac{b}{\tau}} - e^{-\frac{a}{\tau}}\right)u\right]$$

6.3.11 Gamma Distribution

The $\xi_1, \xi_2, ..., \xi_n$ independent random variables are distributed according to the same exponential distribution whose parameter is λ. The random variable

$$\eta = \sum_{i=1}^{n} \xi_i$$

is distributed according to the gamma distribution with n degrees of freedom (see Section 2.3.6).

- If $n = 1$, then the distribution is an exponential one (see Section 6.3.10).
- If n is a small integer, repeat the $n = 1$ case n times and add the results.

6.3.12 Gaussian Distribution

Take the two-dimensional Gaussian distribution, whose parameters are $\mu = 0$ and $\sigma = 1$. The probability on the area element $dS = rdrd\varphi$ is given by

$$f(r,\varphi)dS = \frac{e^{-\frac{r^2}{2}}}{2\pi} dS = \frac{e^{-\frac{r^2}{2}}}{2\pi} r dr d\varphi \ .$$

The two random variables r and φ are independent so

$$f(r,\varphi) = f_r(r) f_\varphi(\varphi) = \frac{1}{2\pi} r e^{-\frac{r^2}{2}} \ .$$

The value of the random variable r is generated from a uniform random number u as follows:

$$u = \int_0^r r e^{-\frac{r^2}{2}} dr = \left[-e^{-\frac{r^2}{2}} \right]_0^r = -e^{-\frac{r^2}{2}} + 1$$

$$r = \sqrt{-2 \ln u}, \qquad \text{because} \qquad u = 1 - u \ .$$

The value of the random variable φ is generated from a uniform random number u as follows:

$$u = \frac{1}{2\pi} \int_0^\varphi d\varphi = \frac{\varphi}{2\pi} \qquad \text{from this} \qquad \varphi = 2\pi u \ .$$

The two normal random number in the rectangular coordinate system are given by

$$x = r \cos\varphi \qquad \text{and} \qquad y = r \sin\varphi \ .$$

There are many faster variants of this basic algorithm. One of them is described here.

1. Generate random numbers u_1 and u_2. They are uniform on $(0,1)$. Then $v_1 = 2u_1 - 1$ and $v_2 = 2u_2 - 1$ are uniform on $(-1,1)$.

2. If $r^2 = v_1^2 + v_2^2 > 1$ go to 1. Otherwise calculate

$$z_1 = v_1\sqrt{\frac{-2lnr^2}{r^2}} \qquad \text{and} \qquad z_2 = v_2\sqrt{\frac{-2lnr^2}{r^2}}$$

which are independent random numbers from a normal distribution with expectation value 0 and variance 1.

The random variable γ is distributed according to the $N(0,1)$ Gaussian distribution. We prove that the random variable $\eta = \mu + \sigma\gamma$ is distributed according to a Gaussian distributions whose parameters are $E(\eta) = \mu$ and $D^2(\eta) = \sigma^2$. The probability density function of γ is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} .$$

The probability that γ is in the (x_1, x_2) interval is given by

$$P(x_1 < \gamma < x_2) = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx .$$

Now we express γ with η:

$$P(x_1 < \gamma < x_2) = P\Big(x_1 < \frac{\eta - \mu}{\sigma} < x_2\Big) .$$

For the random variable η:

$$P\Big(x_1 < \frac{\eta - \mu}{\sigma} < x_2\Big) = P\Big(\mu + x_1\sigma < \eta < \mu + x_2\sigma\Big) = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx .$$

Now we introduce the new variable $y = \mu + x\sigma$ and so

$$P(y_1 < \eta < y_2) = \frac{1}{\sqrt{2\pi}\sigma} \int_{y_1}^{y_2} e^{-\frac{y-\mu^2}{2\sigma^2}} dy .$$

Gaussian Distribution with Center Limit Theorem The indepenent random variables $\xi_1, \xi_2, ..., \xi_n$ are uniform on $(0,1)$. The mathematical expectation values and the variances are given by

$$E(\xi_1) = E(\xi_2) = ... = E(\xi_n) = \frac{1}{2} \qquad \text{and}$$

$$D^2(\xi_1) = D^2(\xi_2) = ... = D^2(\xi_n) = \frac{1}{12} .$$

The mathematical expectation value and the variance of the sum of these random variables $\eta_n = \xi_1 + \xi_2 + ... + \xi_n$ are given by

$$E(\eta_n) = \frac{n}{2} \qquad \text{and}$$

$$D^2(\eta_n) = \frac{n}{12} .$$

The random variable

$$\gamma = \frac{\eta_n - nm}{\sigma\sqrt{n}}$$

has and $N(0,1)$ distribution, if n is a big number. Even in case $n = 12$, in good approximation, one has Gaussian distribution.

6.3.13 χ^2 Distribution

The $z_1, z_2, ..., z_n$ random variables are distributed according to $N(0,1)$ Gaussian distribution. The random variable

$$\chi^2 = \sum_{i=1}^{n} z_i^2$$

is distributed according to the $\chi^2(n_D)$, where $n_D = n$ is the number of degrees of freedom. A faster algorithm is given below:

- For n_D even, generate $n_D/2$ uniform numbers u_i; then

$$y = -2\ln\Big(\prod_{i=1}^{n_D/2} u_i\Big) \qquad \text{is} \qquad \chi^2(n_D)\,.$$

- For odd n_D, generate $(n_D-1)/2$ uniform random numbers u_i and one Gaussian z (see Section 6.3.12); then

$$y = -2\ln\Big(\prod_{i=1}^{(n_D-1)/2} u_i\Big) + z^2 \qquad \text{is} \qquad \chi^2(n_D)\,.$$

6.3.14 Student's t Distribution

The $\xi_1, \xi_2, ..., \xi_n$ and ξ are independent Gaussian distributed random variables with expectation value $\mu = 0$ and variance $\sigma^2 = 1$. We construct the random variables

$$z = \sum_{i=1}^{n} \xi_i^2 \qquad \text{and}$$

$$t = \xi\sqrt{\frac{n}{z}}\,.$$

The variable z belongs to a $\chi^2(n)$ distribution. The random variable t is distributed according to the Student's t distribution with n degrees of freedom.

For $n_D > 0$ degrees of freedom (n_D not necessarily integer):

1. Generate x from a Gaussian distribution with expectation value $\mu = 0$ and variance $\sigma^2 = 1$ according to the method described in Section 6.3.12.
2. Next generate z, obeying $\chi^2(n_D)$ as it is described in Section 6.3.11.
3. Then $t = x\sqrt{n_D/z}$ is distributed as a Student's t with n_D degrees of freedom.

For special case $n = 1$, the Breit-Wigner distribution:

1. Generate random numbers uniformly distributed on $(0,1)$ u_1 and u_2. Calculate $v_1 = 2u_1 - 1$ and $v_2 = 2u_2 - 1$ which are distributed uniformly on $(-1,1)$.

2. If $v_1^2+v_2^2 \leq 1$ accept $z = v_1/v_2$ as a random number distributed according to a Breit-Wigner distribution with unit area, center at 0.0 and FWHM 2.0. Otherwise go to 1.

For center M_0 and FWHM Γ use $W = M_0 + z\Gamma/2$.

7 Monte-Carlo Integration

Let us consider a function $f(x)$. The values of the independent variable are from the (a, b) interval. We want to calculate the integral

$$I = \int_a^b f(x)dx \ .$$

Let us chose a random variable ξ whose probability density function is $p(x)$. The random variable ξ may have any value from the (a, b) interval. Let us calculate the random variable η which is given by

$$\eta = f(x)/p(x) \ .$$

The mathematical expectation value of η is given by

$$E(\eta) = \int_a^b [f(x)/p(x)]p(x)dx = I \ . \tag{41}$$

Let us take N independent values $x_1, x_2, ..., x_N$ of the random variable η. According to the center limit theorem (see Section 2.4):

$$P\Big(\Big|\frac{1}{N}\sum_{i=1}^{N} x_i - I\Big|\Big) < 3\sqrt{\frac{D^2(\eta)}{N}} \sim 0.997 \ .$$

The error of the estimation is sensitive to the selected random variable:

$$D^2(\eta) = M(\eta^2) - I^2 = \int_a^b [f^2(x)/p(x)]dx - I^2 \ .$$

One may show that the error has minimum when $p(x)$ is proportional to $|f(x)|$. Of course, if we chose a complicated $p(x)$ probability density function, then the calculations take a lot of time.

In one dimension there are more efficient numerical methods, but in several dimensions in many cases the Monte-Carlo method is the most usual one.

7.1 Crude Monte-Carlo

The estimated value of the integral is given by the mathematical expectation value of the random variable $\eta = f(x)/p(x)$:

$$\int_a^b f(x)dx \sim \frac{1}{N}\sum_{i=1}^{N} \eta_i \ .$$

In the most simple case we may chose a uniform probability density function:

$$p(x) = \begin{cases} 1/(b-a)\ , & \text{if } a < x < b; \\ 0\ , & \text{otherwise.} \end{cases} .$$

In this case

$$E(\eta) = \frac{(b-a)}{N} \sum_{i=1}^{N} f(x_i)\ .$$

7.2 Acceptance-Rejection Method

We suppose that in the (a,b) interval any given value of x the function $f(x)$ can be computed and the function has an upper limit

$$f(x) \leq M\ ,$$

so one may enclose the function in a rectangle. One may generate points $Q(x,y)$, which are distributed uniformly on the rectangle with the following algorithm (Fig. 14a):

1. Generate two random numbers u_1 and u_2 distributed uniformly on (0,1). Calculate the coordinates which are distributed uniformly on the rectangle:
$$x = a + (b-a)u_1 \qquad \text{and}$$
$$y = Mu_2\ .$$
2. If $y \leq f(x)$, then it is a 'fortunate' event.

Repeat these procedure N times. The area of the rectangle is $M(b-a)$. The ratio of the area below the function $f(x)$ and the whole area $M(b-a)$ is approximately the ratio of the 'fortunate' events k and whole number N of the events:

$$\frac{\int_a^b f(x)dx}{M(b-a)} \sim \frac{k}{N} \qquad \text{and so}$$

$$\int_a^b f(x)dx \sim \frac{k}{N} M(b-a)\ .$$

7.3 Control Variate Monte-Carlo

The $f(x)$ function is constructed as the sum of two functions (Fig. 14b):

$$I = \int_a^b f(x)dx = \int_a^b G(x)dx + \int_a^b (f(x) - G(x))dx\ . \tag{42}$$

The integral of the first function $G(x)$ is known and the error of the other function $(f(x) - G(x))$ is less than the error of the original function $f(x)$.

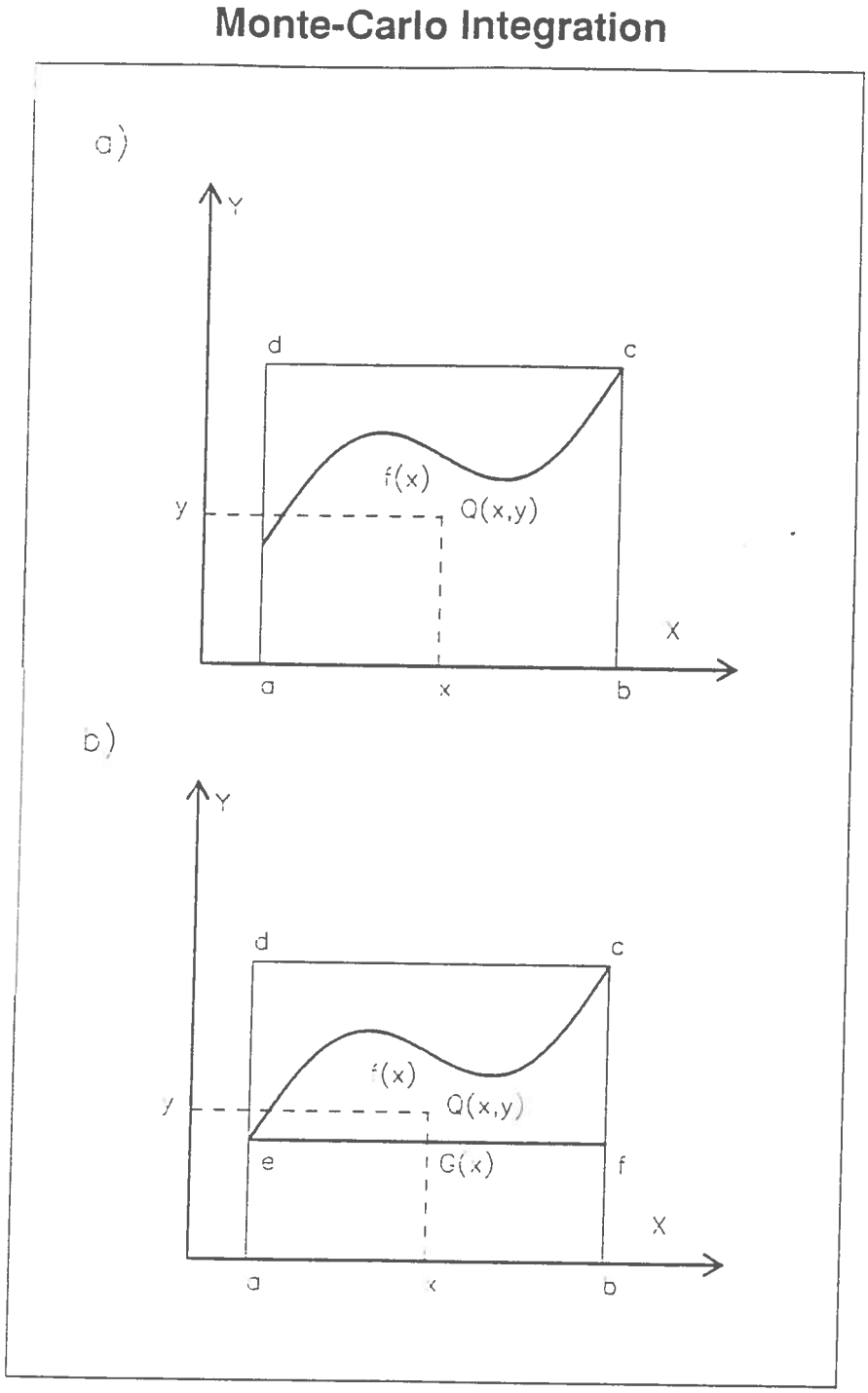

Figure 14: Control Variate Monte-Carlo

7.4 Stratisfied Sampling

There are cases where it is useful to divide into subintervals the whole (a, b) interval. The errors of the subitervals are less than the error of the whole interval (Figs. 15a and 15b):

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \ . \tag{43}$$

7.5 Antithetic Variate

In order to reduce the error, one may construct more uniform function, if the original one has symmetry property as it is show in Figs. 16a and 16b.

- $$\frac{1}{2N}\sum_{i=1}^{N}(f(u) + f(0.5 + u)) \ ,$$
- $$\frac{1}{2N}\sum_{i=1}^{N}(f(u) + f(1.0 - u)) \ .$$

8 Simulation Examples

8.1 Radioactive Decay

There are N_t radioactive atoms at time t. The number of decayed atoms dN during the interval $(t, t+dt)$ is directly proportional to the number of the radioactive atoms N_t at time t and the length of the interval dt:

$$dN = -\lambda N_t dt \ .$$

The minus sign is because the number of radioactive atoms decreases. The constant λ depends on the properties of the radioactive atom. Due to the experiments the constant λ is independent of any physical or chemical effect. Its dimension is s^{-1}. If at time $t = 0$ the number of radioactive atoms is N_0, then at time t the number of radioactive atoms is

$$N_t = N_0 e^{-\lambda t} \ .$$

The number of radioactive atoms decreases according to a negative exponential law. The half life time T of the radioactive atoms is given by

$$\frac{N_0}{2} = N_0 e^{-\lambda T} \qquad \text{and so}$$

$$T = \frac{\ln 2}{\lambda} = \frac{0.69315}{\lambda} \ .$$

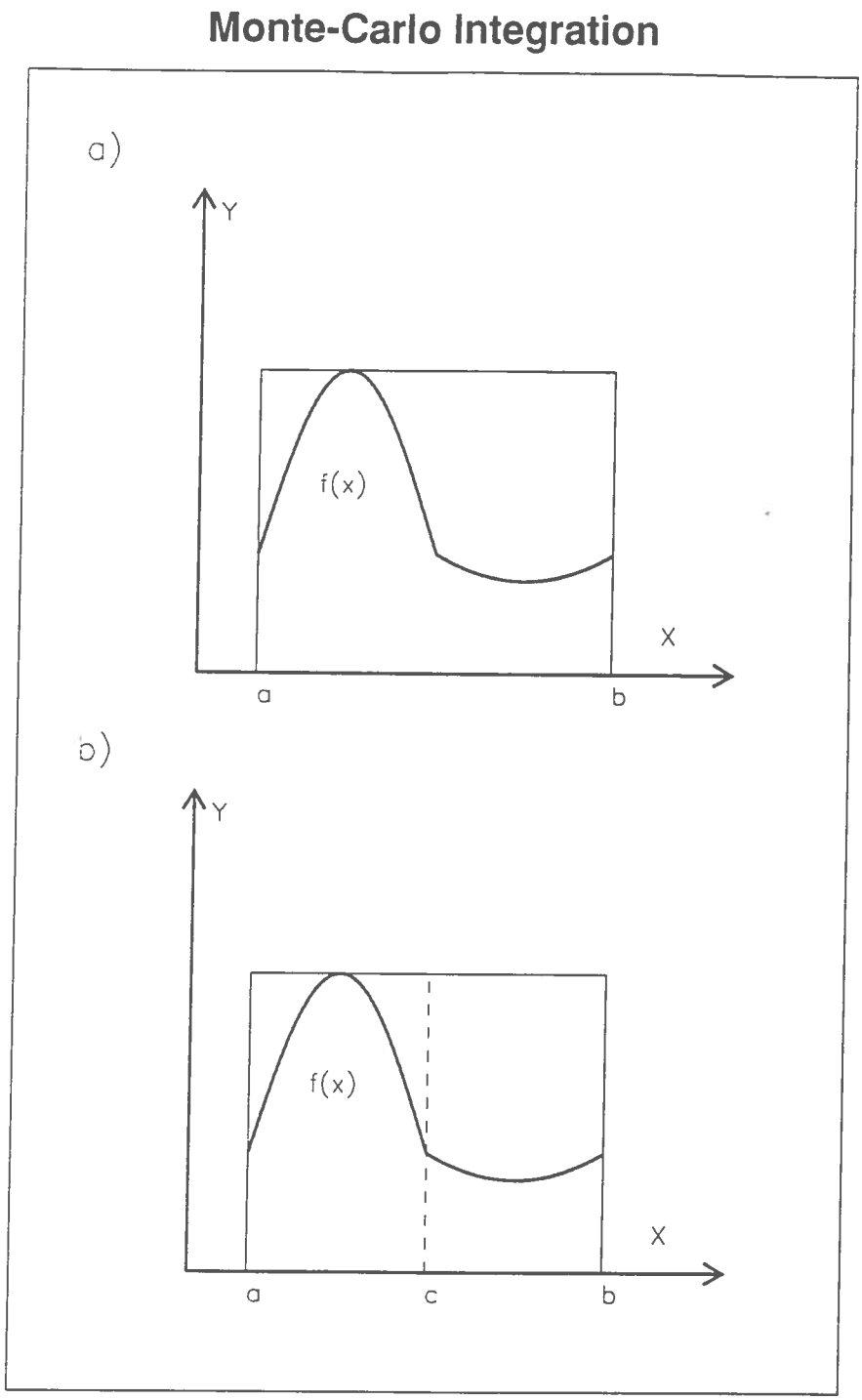

Figure 15: Stratisfied Sampling

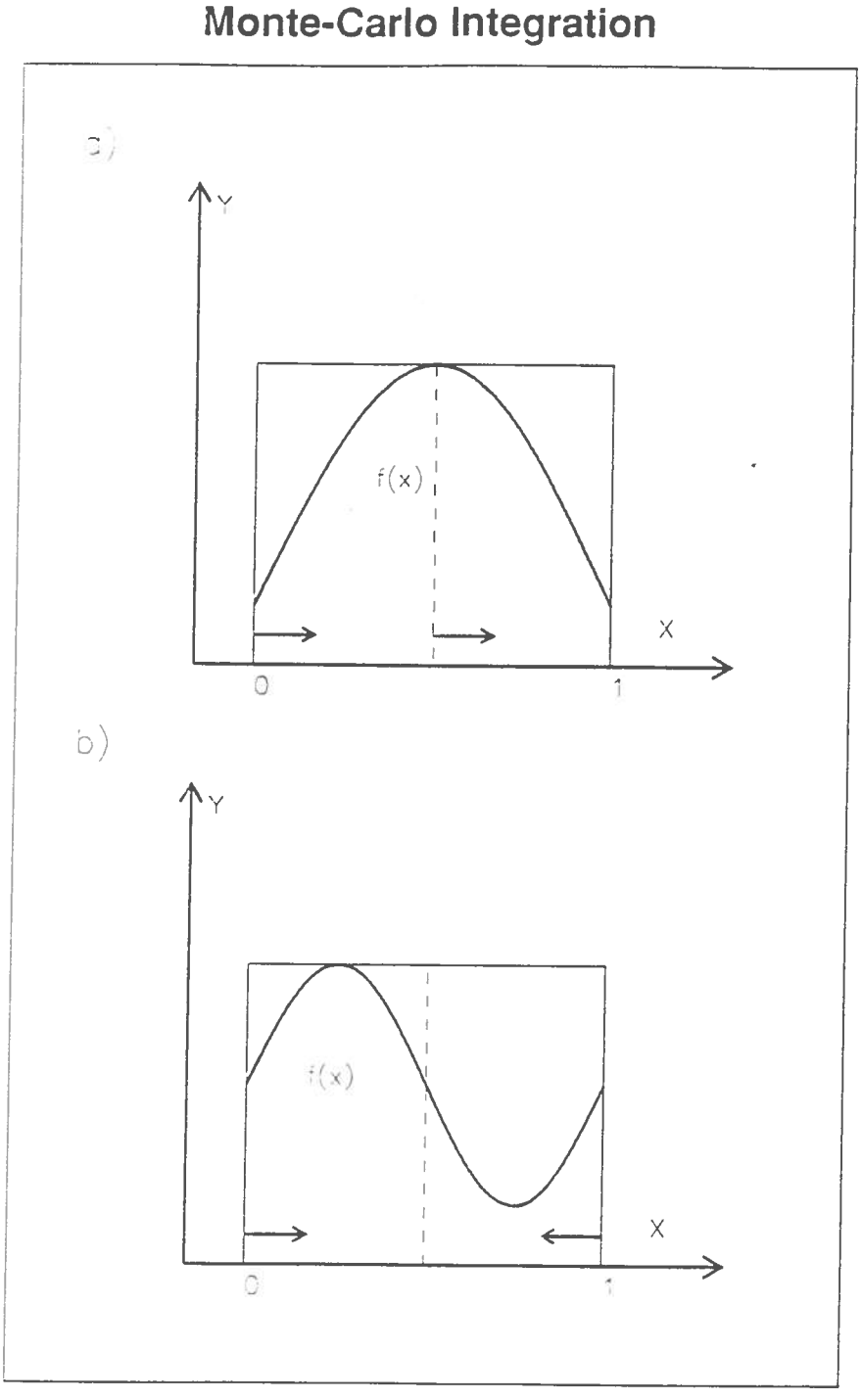

Figure 16: Antithetic Variate

The number of atoms which decay between t and $t + dt$ is given by

$$\lambda N_t dt = \lambda N_0 e^{-\lambda t} dt \ .$$

The sum of the life times is

$$T_\Sigma = \int_0^\infty t\lambda N_0 e^{-\lambda t} dt = \frac{N_0}{\lambda} \ .$$

The mean life time τ of the radioactive atoms is given by

$$\tau = \frac{T_\Sigma}{N_0} = \frac{1}{\lambda} = \frac{T}{0.69315} = 1.4427 \ T \ .$$

The probability that the radioactive atom decays in the interval $(t, t + dt)$ is given by

$$f(t)dt = \frac{1}{\tau} e^{-\frac{t}{\tau}} dt \ .$$

The exponential distribution is shown in Fig. 5. The values of a random variable t, which has exponential distribution, may be generated with the equation

$$u = \frac{1}{\tau} \int_0^t e^{-\frac{t}{\tau}} dt = \left[-e^{-\frac{t}{\tau}} \right]_0^t = -e^{-\frac{t}{\tau}} + 1 \ ,$$

where u is a uniform random number on $(0, 1)$. The value of the random variable t is given by

$$t = -\tau \ln(u) \ . \tag{44}$$

The random numbers u and $1 - u$ have the same probability density function.

8.2 Propagation of Neutrons

Neutrons are launched with E energy against h thick homogeneous plate. The neutron interact with the atoms of the plate. For simplicity we suppose that the energy of the neutron does not change in the interactions and it is scattered in each direction with the same probability. (The last assumption is true if the plate consists from heavy atoms.) In Fig. 17 there are shown the three different events with may happen to the neutrons:

- a) The neutron goes through the plate (Fig. 17a).
- b) The neutron is absorbed in the plate (Fig. 17b).
- c) The neutron is reflected back by the plate (Fig. 17c).

We want to determine the probabilities of the transmission (p_t), the absorption (p_a) and the reflection (p_r). The interactions of the neutron in the medium are characterized by two constants which are called the crossection of the capture (σ_c) and the crossection of the scattering (σ_s). The total crossection is given by

$$\sigma = \sigma_c + \sigma_s \ . \tag{45}$$

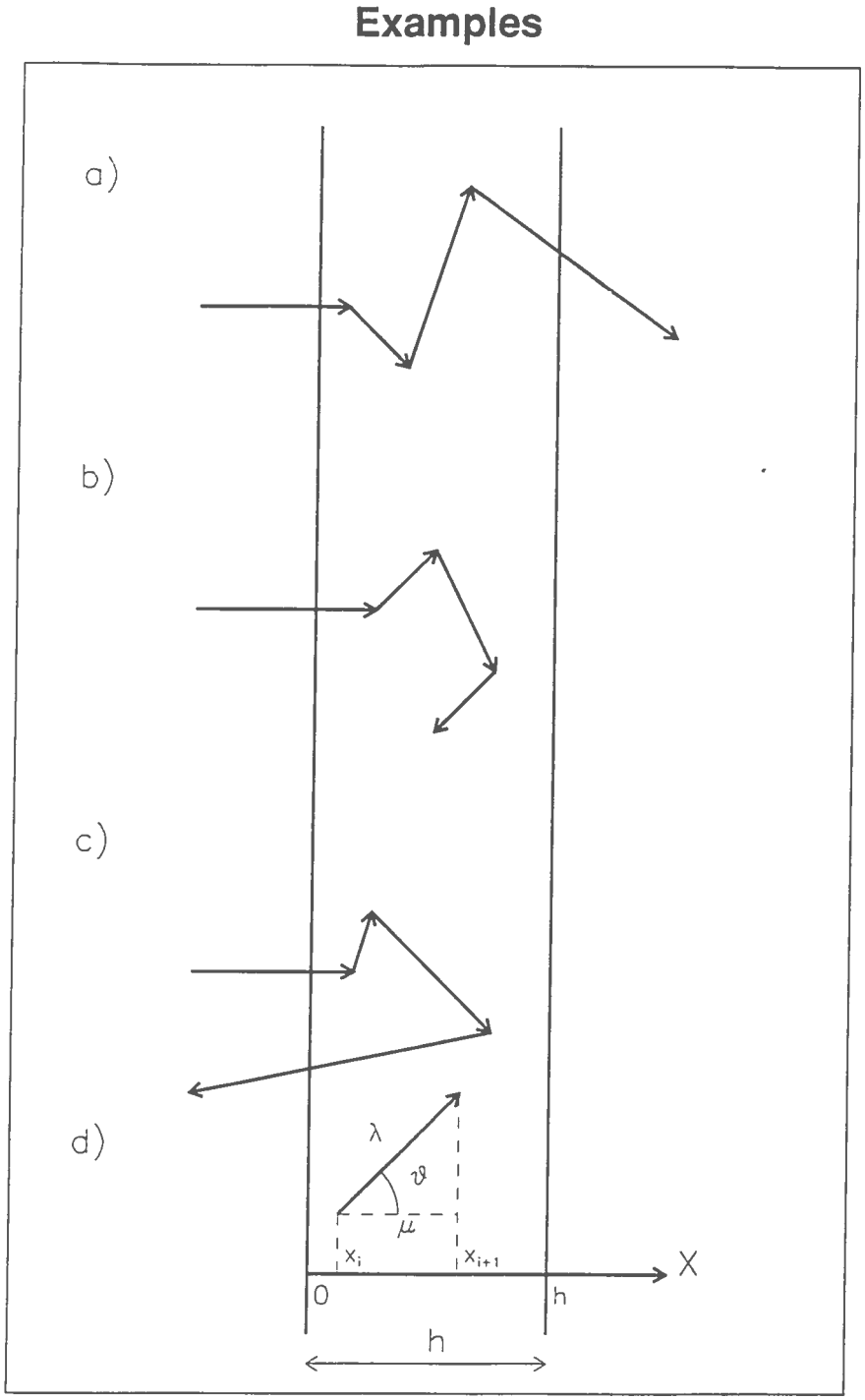

Figure 17: Propagation of Neutrons

1. *Type of interaction.* In each encounter the probability of scattering is σ_s/σ and that of capture is σ_c/σ. One has generate a random number u distributed uniformly on $(0,1)$. If $u < \sigma_c/\sigma$, then the neutron is captured, else it is scattered.

2. *Path length.* The probability density distribution of the path length λ is given by

$$f(\lambda) = \sigma e^{-\sigma\lambda} .$$

The mean free path length is

$$E(\lambda) = \int_0^\infty \lambda\sigma e^{-\sigma\lambda} d\lambda = \frac{1}{\sigma} . \tag{46}$$

The value of λ is generated with the equation (see Section 6.3.10)

$$\int_0^\lambda x\sigma e^{-\sigma x} dx = u ,$$

where u is a random number on $(0,1)$ and so

$$\lambda = -\frac{1}{\sigma} \ln u . \tag{47}$$

3. *Fly direction.* The cosine of the angle ϑ between the fly direction and the x-axis, in the case of isotrop scattering (see Section 6.3.7), is uniformly distributed on $(-1,1)$:

$$\cos\vartheta = 2u - 1 ,$$

where u is a random number distributed uniformly on $(0,1)$. The projection of the path length λ on the x-axis is given by

$$\mu = \lambda \cos\vartheta$$

(see Fig. 17d). The abscissa of the neutron position in the n-th interaction is given by

$$x_n = x_{n-1} + \lambda \cos\vartheta_n .$$

- $x_n < 0$, then it is reflected back.
- $x_n > h$, then it crosses the plate.
- $0 < x_n < h$, then the neutron is in the plate. One should generate the next interaction (go to 1).

One simulates N neutrons and counts the neutrons which are reflected N_r, captured N_c or go through the plate N_t. The probabilities are given by

$$p_r \sim \frac{N_r}{N}, \qquad p_c \sim \frac{N_c}{N} \qquad \text{and} \qquad p_t \sim \frac{N_t}{N} . \tag{48}$$

8.3 Simulation of $K_S^0 \to \mu^+\mu^-$ Decay

Different decay modes of the Kaons were studied with the *BIS* spectrometer which is shown in Fig. 18. The spectrometer, which was one of the first electronic detectors at the Serpukhov accelerator, is a wire chamber spectrometer. It consists of 18 wire chambers (at the beginning spark chambers, later multiwire proportional counters). Each wire chamber measures the x and y coordinates of the traversing charged particles. The charged particles are deflected by the magnet M, so one may determine their momentum. The different decay modes are selected by the A_R, A_L and A anticounters and by the F_1, F_2, G_1 and G_2 scintillator hodoscopes. The electrons and the muons are identified by means of the electron detector eD and the muon detector μD.

The main steps of the simulation of the $K_S^0 \to \mu^+\mu^-$ decay are explained below.

1. Birth point $\vec{p}_l$ of the K_S^0 in the target. The coordinates of the birth point are given by

$$x_l = L_x(2u_1 - 1.) \ , \qquad y_l = L_y(2u_2 - 1.) \qquad \text{and} \qquad z_l = -970. \ ,$$

   where u_1 and u_2 are uniform random numbers (see Section 6.3.1). The L_x and L_y are the extensions of the tharget in the x and y directions. The z-coordinate of the target is $-970.$ cm in the spectrometer's coordinate system. The thickness of the target is negligible.

2. Absolute value of the momentum $|\vec{p}_K|$ of the K_S^0 meson is generated by:

$$|\vec{p}_K| = 15. + 40.u \ , \tag{49}$$

   where u is a uniform random number. The absolute value $|\vec{p}_K|$ of the momentum is generated uniformly between 15 GeV/c and 55 GeV/c (see Section 6.3.1).

3. The fly direction of the K_S^0 meson (see Section 6.3.7) is given by

$$\varphi = 2\pi u_1 \ ,$$

$$\vartheta = 0.06u_2 \ ,$$

$$v_x = \sin\vartheta\cos\varphi \ , \qquad v_y = \sin\vartheta\sin\varphi \qquad \text{and} \qquad v_z = \cos\vartheta \ . \tag{50}$$

   The z-axis of the spectrometer coincides with the z-axis of the coordinate system. The maximum value of ϑ is limited by the acceptance of the detector.

4. Fly distance d of the K_S^0 to its decay point. The mean life time of the K_S^0 meson is $\tau = 0.8922 \times 10^{-10}$ s in the Center of Mass System (CM System). In the Laboratory System (L System) the mean life time is given by

$$\tau_L = \gamma\tau \ , \tag{51}$$

   where $\gamma = E^L/m^{CM}$ and $E^2 = p^2 + m^2$ (the speed of light is $c = 1$). $E^2 \sim p^2$ because $m^2 << p^2$. The mean fly distance is given by

$$\lambda \sim c\tau_L \tag{52}$$

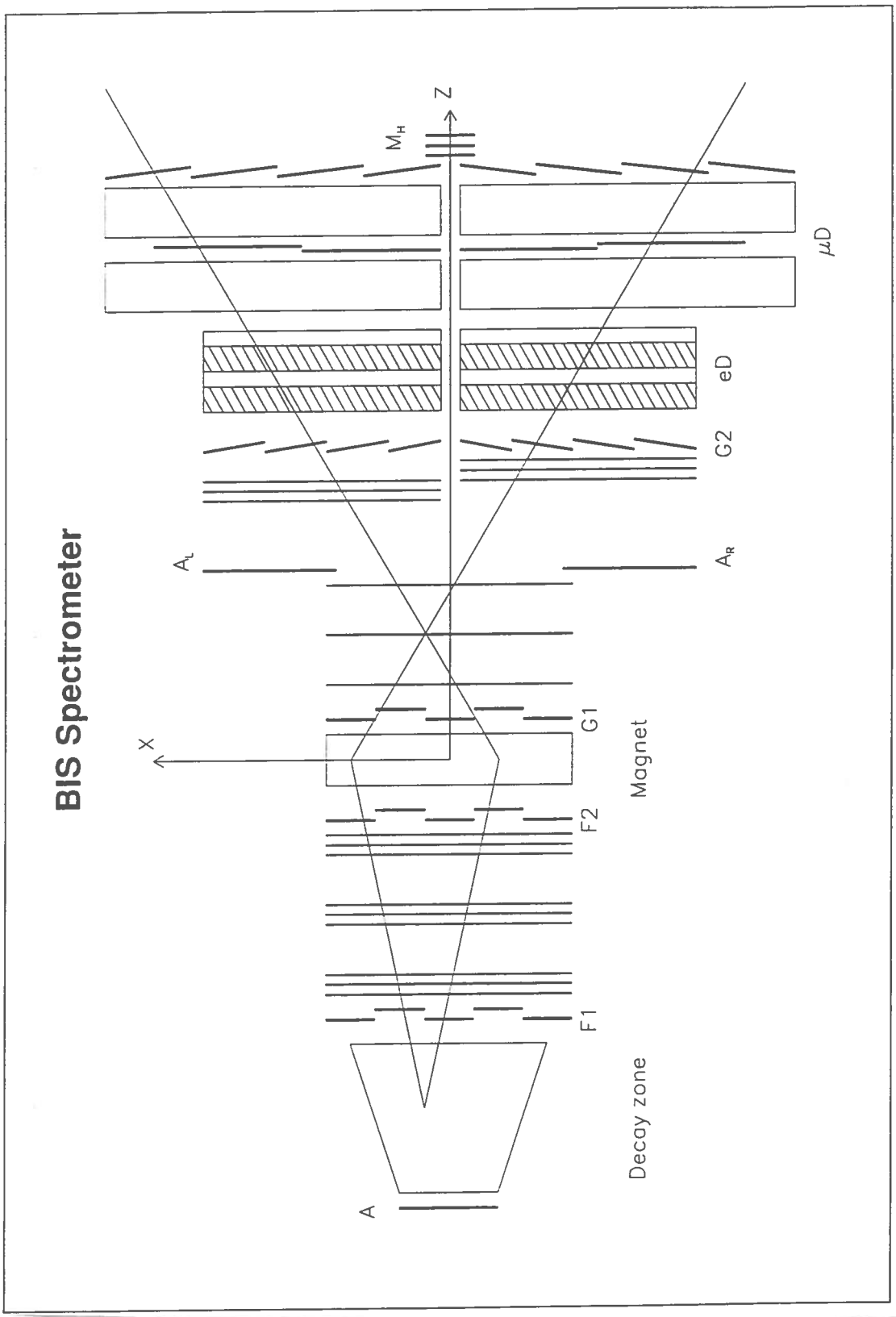

Figure 18: Spectrometer to Study the Kaon Decays

because the K_S^0 meson flys nearly with the speed of light. So the fly distance of the K_S^0 meson is given by

$$d = -\lambda \ln(1. - (1. - e^{-1./\lambda})u) \; . \tag{53}$$

The fly distance is generated from 0 to one mean fly distance (see Section 6.3.10). The space vector of the decay point is given by

$$\vec{p}_d = \vec{p}_\iota + d\vec{v} \; . \tag{54}$$

5. Decay of the K_S^0 meson in the CM System. The mass of μ^- is equal to the mass of μ^+. Due to the momentum conservation the absolute value of the momentum of μ^- is equal to the absolute value of the momentum of μ^+:

$$|\vec{p}_{\mu^-}^{\,CM}| = |\vec{p}_{\mu^+}^{\,CM}| \; . \tag{55}$$

From the energy conservation we have

$$E_{\mu^-}^{CM} = E_{\mu^+}^{CM} = \frac{1}{2} m_{K_S^0}^{CM} \; . \tag{56}$$

The sum of momenta in the CM System is zero, so

$$\vec{p}_{\mu^-}^{\,CM} = -\vec{p}_{\mu^+}^{\,CM} \; . \tag{57}$$

The absolute value of the momentum of the muons is given by

$$|\vec{p}_{\mu^-}^{\,CM}| = |\vec{p}_{\mu^+}^{\,CM}| = \sqrt{(E_\mu^{CM})^2 - (m_\mu^{CM})^2} \; .$$

The decay products (μ^- and μ^+) are emitted isotropically in the CM System (see Section 6.3.7):

$$\varphi = 2\pi u_1 \; ,$$

$$\cos\vartheta = 2u_2 - 1. \; ,$$

$$v_x = \sin\vartheta\cos\varphi \; , \qquad v_y = \sin\vartheta\sin\varphi \qquad \text{and} \qquad v_z = \cos\vartheta \; . \tag{58}$$

The momentum of the muons in the CM System are given by

$$\vec{p}_{\mu^-}^{\,CM} = |\vec{p}_\mu^{\,CM}|\vec{v} \qquad \text{and} \qquad \vec{p}_{\mu^+}^{\,CM} = -\vec{p}_{\mu^-}^{\,CM} \; .$$

6. The transformation of momentum of the muons into the L System. The Lorentz transformation is given by

$$\begin{pmatrix} E^{\,L} \\ p_\parallel^{\,L} \\ p_\perp^{\,L} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 \\ -\gamma\beta & \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E^{\,CM} \\ p_\parallel^{\,CM} \\ p_\perp^{\,CM} \end{pmatrix} \; , \tag{59}$$

where $\gamma = E^L/m^{CM}$ and $\gamma\beta = |p^L|/m^{CM}$. The CM System of the muons is moving with the velocity of the K_S^0 meson and so

$$\gamma = \frac{E_K^L}{m_K^{CM}} \qquad \text{and} \qquad \gamma\beta = \frac{|p_K^L|}{m_K^{CM}} \; , \tag{60}$$

where $(E_K^L)^2 = (p_K^L)^2 + (m_K^L)^2$ (the units are $\hbar = c = 1$). The parallel components of the muon μ (μ^- or μ^+) is:

$$p_{\parallel\,\mu}^{CM} = \frac{\vec{p}_\mu^{\,CM}\vec{p}_K^{\,L}}{|\vec{p}_K^{\,L}|} ,$$

$$P_{Boost} = (\gamma - 1)p_{\parallel\,\mu}^{CM} + \gamma\beta E_\mu^{CM} \qquad \text{and finally}$$

$$\vec{p}_\mu^{\,L} = \vec{p}_\mu^{\,CM} + P_{Boost}\frac{\vec{p}_K^{\,L}}{|\vec{p}_K^{\,L}|} . \tag{61}$$

7. The charged particles are deflected by the Lorentz force in the magnet M. The Lorentz force is given by

$$\vec{F} = q(\vec{v} \times \vec{B}) . \tag{62}$$

8. Intersection point of the trajectory of a charged particle and a wire chamber. The trajectory of the charged particle is given by the straight line

$$\vec{r} = \vec{r}_0 + t\vec{v} , \tag{63}$$

where $\vec{r}_0$ is a point on the staight line and $\vec{v}$ is the fly direction of the particle along the line. The intersection point is determined by the equation:

$$(\vec{r} - \vec{q})\vec{u} = 0 , \tag{64}$$

where $\vec{q}$ is a point on the chamber's plane and $\vec{u}$ is the unit normal vector of the chamber's plane.

In Fig. 19 one may see a simulated event. Analysing the simulated events one may determine the acceptance of the detector, etc.

8.4 Quality of Equipments

Let us consider an equipment which consists of several component parts. The quality of the equipment is described with a function which depends on the parameters of the component parts. For example an electronic equipment contains resistors, capacitors etc. It is described by a function which depends on the parameters of the resistors, capacitors etc. For example this function f may be the voltage U measured at a given point of the equipment:

$$U = f(R_1, R_2, ..., C_1, C_2, ...) . \tag{65}$$

The real value of the resistor, capacitor etc. is in an interval around the nominal value. There are cases in which one may calculate the extreme values of U using the extreme values of the parameters. In complicated cases these calculations are not simple and on the other hand the probability is small that each parameter has extreme value. If we study a lot of component parts, then we may determine the distribution of each parameter and its mean value. So it seems more resonable to calculate with the mean values:

$$E(U) = f(E(R_1), E(R_2), ..., E(C_1), E(C_2), ...) . \tag{66}$$

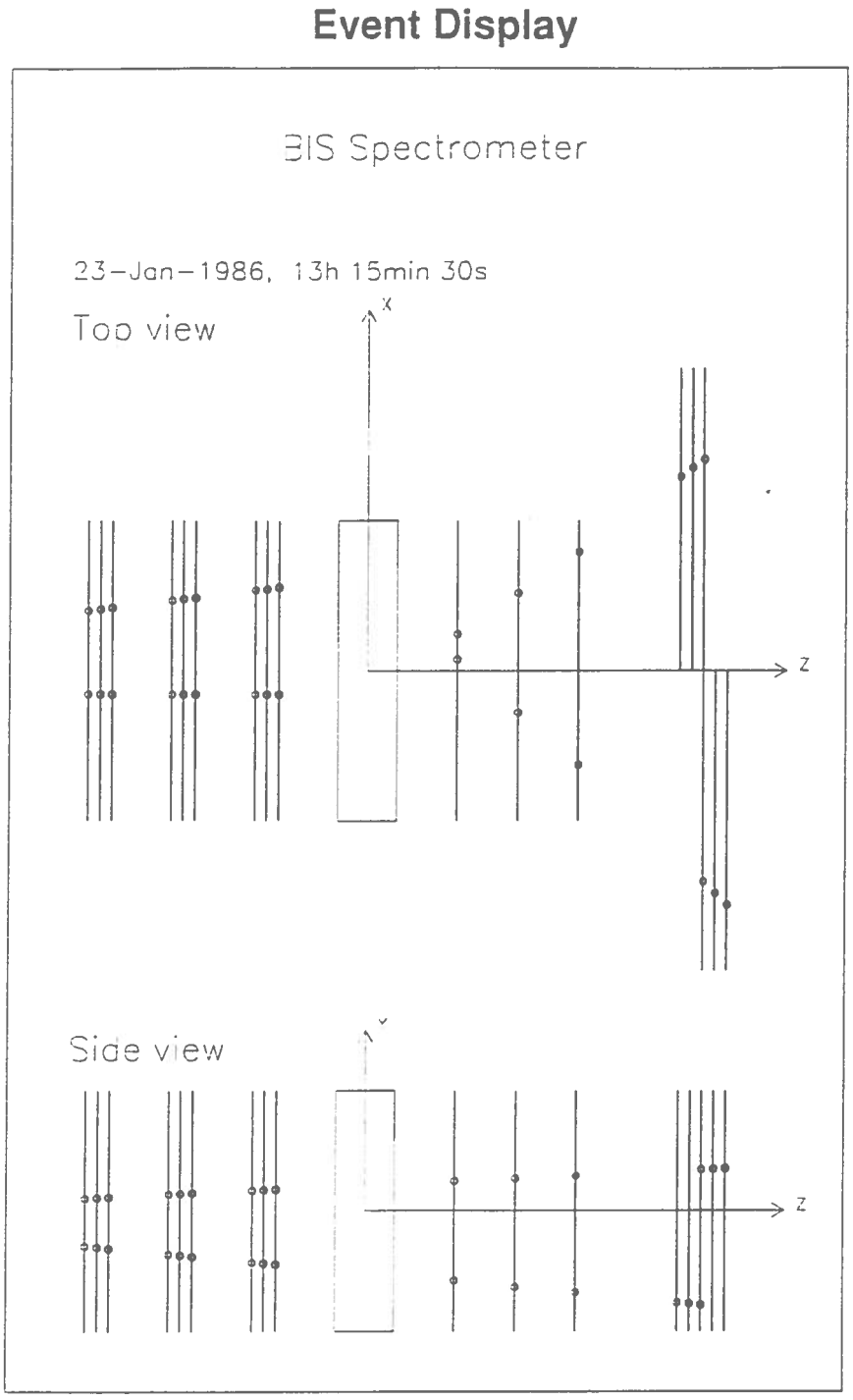

Figure 19: Simulated Event

In most of the cases it is difficult to give the analytical form of the function f. One may use the Monte-Carlo method in this case too. Knowing the distribution of the parameters one may generate the parameter values. With the generated parameter values one may calculate the mean value and the standard deviation:

$$E(U) \sim \frac{1}{N}\sum_{i=1}^{N} U_i \qquad \text{and}$$

$$D^2(U) \sim \frac{1}{N-1}\sum_{i=1}^{N}(U_i - E(U))^2 = \frac{1}{N-1}\Big(\sum_{i=1}^{N}(U_i)^2 - \frac{1}{N}\Big(\sum_{i=1}^{N} U_i\Big)^2\Big)$$

8.5 Reliability of Equipments

We would like to estimate the life time of an equipment. Let us suppose that we know the probability of the break down of each component part. The break down of one component part means the breakdown of the whole equipment. In case of four components the life time of the equipment is given by

$$t = \min(t_1, t_2, t_3, t_4)\ . \tag{67}$$

If we duplicate one component part, then in the case of the break down of one of the two parts the whole equipment does not break down. There are five component parts in the equipment. The third and the fourth may substitute each other. In this case the life time of the equipment is given by

$$t = \min(t_1, t_2, \max(t_3, t_4), t_5)\ . \tag{68}$$

One may generate the time values of the component parts with Monte-Carlo method and calculate the life time of the equipment. Repeating N times the calculations one may calculate the mean life time and the standard deviation

$$E(t) \sim \frac{1}{N}\sum_{i=1}^{N} t_i \qquad \text{and}$$

$$D^2(t) \sim \frac{1}{N-1}\sum_{i=1}^{N}(t_i - E(t))^2\ .$$

One may develop the method taking into account that the whole equipment break down only if some combinations of the component parts do break down and so on.

References

[1] L.Jánossy, Theory and Practice of the Evaluation of Measurements, Oxford University Press, (1965).

[2] D.E.Knuth, The Art of Computer Programming, Addison-Wesley Publishing Company, London, (1969).

[3] G.E.Forsyte, M.A.Malcolm, C.B.Moler, Computer Methods for Mathematical Computations, Prince-Hall, Inc. Englewood Cliffs, N.J. 07632, (1977).

[4] S.Bonetti et al.,The Monte-Carlo code for the Borexino detector at LNGS, Nuclear instruments and Methods in Physics Research A329 (1993) **314-328**, North-Holland.

Appendices

Appendix A: Gamma Function

The gamma function $\Gamma(p)$ is defined for $p > 0$ with the integral:

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

The properties of the gamma function are

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

- $$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty x^{\frac{1}{2}-1} e^{-x} dx = \int_0^\infty \frac{1}{\sqrt{x}} e^{-x} dx = 2\int_0^\infty e^{-u^2} du = \sqrt{\pi} \ .$$

- $$\Gamma(p+1) = \int_0^\infty x^p e^{-x} dx = \left[-x^p e^{-x}\right]_0^\infty + p\int_0^\infty x^{p-1} e^{-x} dx = p\Gamma(p) \ .$$

- In the case of positive integer numbers

$$\Gamma(n+1) = n! \ .$$

Appendix B: Binary Search

There are n values $x_1, x_2, ..., x_n$. To find the value of these n values which is the closest to a given x in average one has to make $n/2$ comparisons in the sequential search. With the binary search one has to make only $1 + [\log_2 n]$ comparisons. In the case of 512 elements the numbers are 256 and 9 respectively. In the binary search one has to arrange the n values into increasing order. One may find the serial number of the smallest element which is greater than a given value x with the algorithm as follows:

1. $I = 1$ and $J = n + 1$.
2. $K = [(I + J)/2]$.
3. If $x < x_K$, then $J = K$ and go to 5.
4. IF $x \geq x_K$, then $I = K$ and go to 5.
5. If $J > I + 1$, then go to 2.
6. The serial number of the smallest element, which is greater than or equal to x, is in the variable I.

Appendix C: The Function URAND

Here is presented a uniform random number generator. The program is written in FORTRAN, so it is rather independent of the computers. The URAND (Universal RANDom number generator or Uniform RANDom number generator) uses the formula described in Section 3.2.3. The parameter IY of the function URAND may be optional integer number at the first call. One must not change its value further on. The values of the constants of the multiplicative congruent method (a, b and c) are calculated at the first call.

```
      REAL FUNCTION URAND(IY)
      INTEGER IY
C
C     The function URAND generates uniform random numbers
C     on the (0,1) interval. The value of the variable IY
C     determines the first value of the generator. In the
C     successive call the value of IY must not change.
C
      INTEGER IA,IC,ITWO,M2,M,MIC
      DOUBLE PRECISION HALFM
      REAL S
      DOUBLE PRECISION DATAN,DSQRT
      DATA M2/0/,ITWO/2/
      IF (M2.NE.0) GO TO 20
C
C     The length of the computer word (Integer number).
C
      M=1
 10   M2=M
      M=ITWO*M2
      IF (M.GT.M2) GO TO 10
      HALFM=M2
C
C     The coefficient and the constant of the linear congruent method.
C
      IA=8*IDINT(HALFM*DATAN(1.D0)/8.D0)+5
      IC=2*IDINT(HALFM*(0.5D0-DSQRT(3.D0)/6.D0))+1
      MIC=(M2-IC)+M2
C
C     S is the coefficient of the floating point transformation.
C
      S=0.5D0/HALFM
C
      The next random number
C
 20   IY=IY*IA
C
```

```
C     These instructions are needed for the computers where the
C     overflow is not allowed for integer numbers.
C
      IF (IY.GT.MIC) IY=(IY-M2)-M2
      IY=IY+IC
C
      These instructions are needed for the computers whose word
C     length is longer in the case of addition than in the case
C     of multiplication.
C
      IF (IY/2.GT.M2) IY=(IY-M2)-M2
C
C     These instruction are for computers where the overflow
C     changes the sign.
C
      IF (IY.LT.0) IY=(IY+M2)+M2
      URAND=FLOAT(IY)*S
      RETURN
      END
```

List of Figures