

# Monte Carlo Simulation of Stochastic Processes

## MONTÉ CARLO METHOD

- **Monte Carlo (MC) method:** A computational method that utilizes random numbers.
- Two major applications of the MC method:
  1. Multidimensional integrations (e.g., statistical mechanics in physics);
  2. Simulation of stochastic natural phenomena (e.g., stock price).

In this lecture, we discuss the MC method used to simulate stochastic natural and artificial processes.

## §1 Probability—Foundation of the MC Method

### ELEMENTS OF PROBABILITY THEORY<sup>1</sup>

- **Random variable:** Arises when repeated attempts under apparently identical conditions to measure a quantity produce different results. For example, the reading from tossing a dice,  $x \in \{1, 2, 3, 4, 5, 6\}$ , is a random variable.
- **Probability density function:** Defines the probability,  $p(x)$ , when the variable is sampled, that the value  $x$  will result. For a large number of samples,

$$p(x) = (\text{Number of samples with result } x) / (\text{Total number of samples})$$

For example, the probability that the reading of dice is 1,  $p(1)$ , is  $1/6$ .

- **Normalization condition:** Satisfied by the probability density function simply states that the number of samples for all the readings add up to the total number of samples.

$$\sum_k p(x_k) = 1,$$

i.e.,  $\sum_x (\text{Number of samples with result } x) / (\text{Total number of samples}) = 1$ . For example, for a dice,  $p(1) + p(2) + \dots + p(6) = 1$ .

- **Expectation (mean):** of a random variable  $x$  is defined as

$$E[x] = \langle x \rangle = \sum_k x_k p(x_k).$$

For example, the expected reading of a dice is,

$$1 \times \frac{1}{6} + \dots + 6 \times \frac{1}{6} = \frac{6 \times (1+6)}{2} \times \frac{1}{6} = 3.5.$$

### CONTINUOUS RANDOM VARIABLES

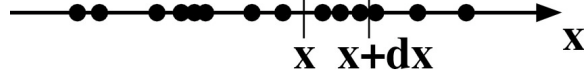
Random variables can be continuous, e.g., the position of an atom along  $x$ -axis in a box. The probability density function now defines the probability that, when the variable is sampled, a value lying in the range  $x$  to  $x + dx$  will result; this probability is

---

<sup>1</sup> M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods, Vol. I: Basics* (John Wiley & Sons, New York, 1986)—a good introduction to the general MC method; P. K. MacKeown, *Stochastic Simulation in Physics* (Springer, Singapore, 1997)—contains a practical introduction to probability with many applications in physics; W. Feller, *An Introduction to Probability Theory and Its Application, Vol. I, 3rd Ed.* (John Wiley & Sons, New York, 1968)—a classic of probability theory with rigorous treatment of Markov chains.

$$p(x)dx.$$

This is realized in the limit of very large number of sample points such that  $p(x)dx = (\text{Number of samples in the range } [x, x+dx]) / (\text{Total number of samples})$ .



The normalization condition now involves an integral,

$$\int_{-\infty}^{\infty} dx p(x) = 1.$$

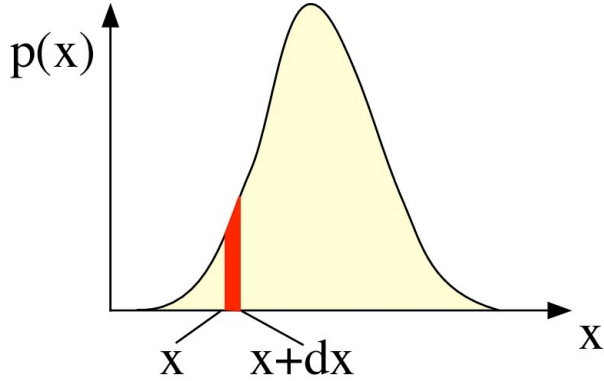
and the mean is calculated as

$$E[x] = \langle x \rangle = \int_{-\infty}^{\infty} dx x p(x).$$

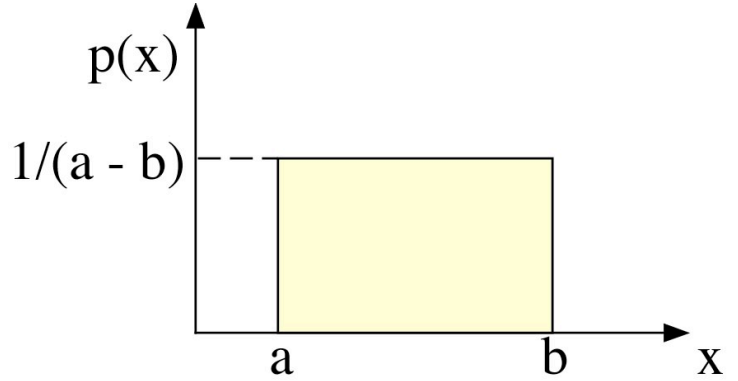
We can also define the expectation of a function value,

$$E[f(x)] = \langle f(x) \rangle = \int_{-\infty}^{\infty} dx f(x) p(x),$$

i.e., if we repeatedly generate random numbers,  $x$ , and sample the function value,  $f(x)$ , then  $\langle f(x) \rangle$  is the average value over the entire samples.



**Figure:** A continuous probability density function.



**Figure:** Uniform probability in the range  $[a, b]$ .

For example, if a random variable,  $x$ , is uniformly distributed in the range,  $[a, b]$ , then the probability density function is,

$$p(x) = 1/(b - a).$$

The mean of the random variable  $x$  is

$$\langle x \rangle = \int_a^b dx x \frac{1}{b-a} = \frac{1}{b-a} \left[ \frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}.$$

For a function,  $f(x) = x^2$ , the expected value is given by

$$\langle x^2 \rangle = \int_a^b dx x^2 \frac{1}{b-a} = \frac{1}{b-a} \left[ \frac{x^3}{3} \right]_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}.$$

## VARIANCE AND STANDARD DEVIATION

- **Variance:** of a random variable  $x$  is defined as

$$\text{Var}[x] = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2.$$

For the above uniform distribution, the variance is

$$\text{Var}[x] = \frac{b^2 + ba + a^2}{3} - \left(\frac{b+a}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

- **Standard deviation (standard error):** of a random variable  $x$  is the square root of its variance,

$$\text{Std}[x] = \sqrt{\text{Var}[x]},$$

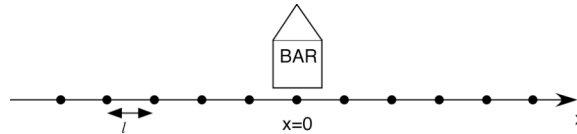
and it measures how broadly the probability density spreads around its mean. In the above example, the standard deviation is  $|b-a|/2\sqrt{3}$ .

## §2 Random Walks<sup>2</sup>

We consider the simplest but most fundamental stochastic process, i.e., random walks in one dimension.

### DRUNKARD'S WALK PROBLEM

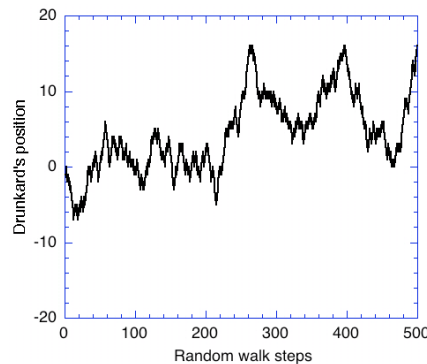
Consider a drunkard on a street in front of a bar, who starts walking at time  $t = 0$ . At every time interval  $\tau$  (say 1 second) the drunkard moves randomly either to the right or to the left by a step of  $l$  (say 1 meter). The position of the drunkard  $x$  along the street is a random variable.



A MC simulation of the drunkard is implemented according to the following pseudocode.

- Program `diffuse.c`

```
Initialize a random number sequence
for walker = 1 to N_walker
  position = 0
  for step = 1 to Max_step
    if rand()3 > RAND_MAX/2 then
      Increment position by 1
    else
      Decrement position by 1
    endif
  endfor step
endfor walker
```



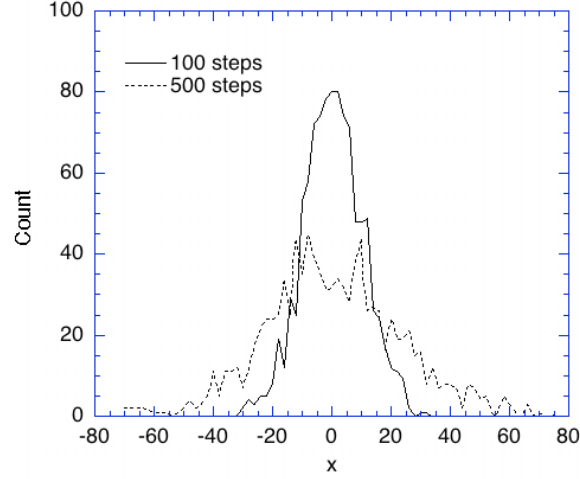
**Figure:** An MC simulation result of a walking drunkard's position for 500 steps.

<sup>2</sup> H. Gould and J. Tobochnik, *An Introduction to Computer Simulation Methods, Part 2* (Addison-Wesley, Reading, MA, 1988).

<sup>3</sup> The function `rand()` returns a random integer with uniform distribution in the range between 0 and `RAND_MAX`.

## PROBABILITY DISTRIBUTION

The drunkard's position,  $x(t)$ , at time  $t$  is a random variable, which follows the probability density function,  $P(x, t)$ . By generating many drunkards (with different random-number seeds), we can have a MC estimate of  $P(x, t)$ . The following graph shows a histogram of the drunkard's position over 1,000 samples at 100 and 500 steps. Note that the initial probability density is  $P(x, 0) = \delta_{x,0}$ , meaning that the drunkard is at the origin with probability 1. As time progresses, the probability distribution becomes broader.



**Figure.** A histogram of the drunkard's position for 1,000 random drunkards.

Let's analyze the probability density of the drunkard's position. First consider the probability,  $P_n(x)$ , that the drunkard is at position  $x$  at time  $n\tau$ . Suppose that the drunkard has walked to the right  $n_{\rightarrow}$  times to the right and  $n_{\leftarrow} = n - n_{\rightarrow}$  times to the left. Then the drunkard's position  $x$  is  $(n_{\rightarrow} - n_{\leftarrow})l$ . There are many ways that the drunkard can reach the same position at the same time; the number of possible combinations is

$$\frac{n!}{n_{\rightarrow}!n_{\leftarrow}!},$$

where  $n!$  is the factorial of  $n$ . (There are  $n!$  combinations to arrange  $n$  distinct objects in a list. However  $n_{\rightarrow}$  objects are indistinguishable and therefore the number of combinations is reduced by a factor of  $n_{\rightarrow}!$ . Due to a similar reason, the number must be further divided by  $n_{\leftarrow}!$ .) Let's assume that the drunkard walks to the right with probability,  $p$ , and to the left with probability,  $q = 1 - p$ . Then each of the above path occurs with probability,  $p^{n_{\rightarrow}}(1 - p)^{n_{\leftarrow}}$ . Consequently the probability that the drunkard is at position  $x = (n_{\rightarrow} - n_{\leftarrow})l$  at time  $n\tau$  is given by

$$P_n(x = (n_{\rightarrow} - n_{\leftarrow})l) = \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} (1 - p)^{n_{\leftarrow}}.$$

The mean value of  $x$  at time  $n\tau$  is thus

$$\begin{aligned} \langle x_n \rangle &= \sum_{n_{\rightarrow}=0}^n P_n(x = (n_{\rightarrow} - n_{\leftarrow})l) (n_{\rightarrow} - n_{\leftarrow})l \\ &= \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} (1 - p)^{n_{\leftarrow}} (n_{\rightarrow} - n_{\leftarrow})l. \end{aligned}$$

Now recall the binomial identity, which we will use as a **generating function** for the binomial series,

$$\sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} q^{n_{\leftarrow}} = (p+q)^n.$$

By differentiating both sides of the above identity by  $p$ , we obtain

$$\begin{aligned} \frac{\partial}{\partial p} \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} q^{n_{\leftarrow}} &= \sum_{n_{\rightarrow}=1}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} p^{n_{\rightarrow}-1} q^{n_{\leftarrow}} \\ &= \frac{\partial}{\partial p} (p+q)^n = n(p+q)^{n-1} \end{aligned}$$

By multiplying both sides by  $p$ , and noting the term  $n_{\rightarrow} = 0$  is zero,

$$\sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} p^{n_{\rightarrow}} q^{n_{\leftarrow}} = np(p+q)^{n-1}.$$

For  $q = 1 - p$ , we get

$$\sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} p^{n_{\rightarrow}} (1-p)^{n_{\leftarrow}} = np.$$

Similarly,

$$\sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\leftarrow} p^{n_{\rightarrow}} (1-p)^{n_{\leftarrow}} = nq.$$

Using the above two equalities in the expression for  $x_n$ ,

$$\langle x_n \rangle = \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} (1-p)^{n_{\leftarrow}} (n_{\rightarrow} - n_{\leftarrow}) l = n(p-q)l.$$

If the drunkard walks to the right and left with equal probability, then  $p = q = 1/2$  and  $x_n = 0$  as can be easily expected.

Now let's consider the variance of  $x_n$ .

$$\begin{aligned} \langle x_n^2 \rangle &= \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} (1-p)^{n_{\leftarrow}} ((n_{\rightarrow} - n_{\leftarrow})l)^2 \\ &= \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} (1-p)^{n_{\leftarrow}} (n_{\rightarrow}^2 - 2n_{\rightarrow}n_{\leftarrow} + n_{\leftarrow}^2) l^2. \end{aligned}$$

We can again make use of the binomial relation. By differentiating both sides by  $p$  twice,

$$\sum_{n_{\rightarrow}=2}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow}^2 p^{n_{\rightarrow}-2} q^{n_{\leftarrow}} = n(n-1)(p+q)^{n-2} + \sum_{n_{\rightarrow}=2}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} p^{n_{\rightarrow}-2} q^{n_{\leftarrow}}.$$

Multiplying both sides by  $p^2$ ,

$$\sum_{n_{\rightarrow}=2}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow}^2 p^{n_{\rightarrow}} q^{n_{\leftarrow}} = n(n-1)p^2(p+q)^{n-2} + \sum_{n_{\rightarrow}=2}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} p^{n_{\rightarrow}} q^{n_{\leftarrow}}$$

$$\begin{aligned} \therefore \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow}^2 p^{n_{\rightarrow}} q^{n_{\leftarrow}} - npq^{n-1} &= n(n-1)p^2(p+q)^{n-2} + \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} p^{n_{\rightarrow}} q^{n_{\leftarrow}} - npq^{n-1} \\ \therefore \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow}^2 p^{n_{\rightarrow}} q^{n_{\leftarrow}} &= n(n-1)p^2(p+q)^{n-2} + np(p+q)^{n-1} \end{aligned}$$

For  $p+q=1$ ,

$$\sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow}^2 p^{n_{\rightarrow}} q^{n_{\leftarrow}} = n(n-1)p^2 + np.$$

Similarly,

$$\sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\leftarrow}^2 p^{n_{\rightarrow}} q^{n_{\leftarrow}} = n(n-1)q^2 + nq.$$

Now differentiate both sides of the binomial relation with respect to  $p$  and then by  $q$ ,

$$\begin{aligned} \sum_{n_{\rightarrow}=1}^{n-1} \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} n_{\leftarrow} p^{n_{\rightarrow}-1} q^{n_{\leftarrow}-1} &= n(n-1)(p+q)^{n-2} \\ \therefore \sum_{n_{\rightarrow}=1}^{n-1} \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} n_{\leftarrow} p^{n_{\rightarrow}} q^{n_{\leftarrow}} &= n(n-1)pq(p+q)^{n-2} \\ \therefore \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} n_{\leftarrow} p^{n_{\rightarrow}} q^{n_{\leftarrow}} &= n(n-1)pq(p+q)^{n-2} \end{aligned}$$

For  $p+q=1$ ,

$$\sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} n_{\rightarrow} n_{\leftarrow} p^{n_{\rightarrow}} q^{n_{\leftarrow}} = n(n-1)pq.$$

By combining the above results,

$$\begin{aligned} \langle x_n^2 \rangle &= \sum_{n_{\rightarrow}=0}^n \frac{n!}{n_{\rightarrow}!n_{\leftarrow}!} p^{n_{\rightarrow}} q^{n_{\leftarrow}} (n_{\rightarrow}^2 - 2n_{\rightarrow}n_{\leftarrow} + n_{\leftarrow}^2) l^2 \\ &= [n(n-1)p^2 + np - 2n(n-1)pq + n(n-1)q^2 + nq] l^2 \\ &= [n(n-1)(p-q)^2 + n] l^2 \end{aligned}$$

The variance is obtained as

$$\begin{aligned} \langle x_n^2 \rangle - \langle x_n \rangle^2 &= [n(n-1)(p-q)^2 + n] l^2 - [n(p-q)l]^2 \\ &= [1 - (p-q)^2] nl^2 \\ &= [(p+q)^2 - (p-q)^2] nl^2 \\ &= 4pqnl^2. \end{aligned}$$

For  $p=q=1/2$ ,

$$\text{Var}[x_n] = nl^2.$$

## DIFFUSION LAW

The main result of the above analysis is the linear relation between the steps and the variance of the random walk (the latter is also called the **mean square displacement**). The following graph confirms this linear relation. This relation means that a drunkard cannot go far. If he walks straight to the right, he can reach to the distance,  $nl$ , in  $n$  steps. On the other hand, the drunkard can reach only,  $Std[x_n] = \sqrt{n}l$ , on average.

The time evolution of  $P(x, t)$  for the drunkard's walk problem is typical of the so call **diffusion** processes. Diffusion is characterized by a linear relation between the mean square displacement and time,

$$\langle \Delta R(t)^2 \rangle = 2Dt.$$

The above drunkard follows this general relation, since

$$\langle x(t = n\tau)^2 \rangle = nl^2 = 2\left(\frac{l^2}{2\tau}\right)t.$$

The “diffusion constant” in this example is  $D = l^2/2\tau$ .

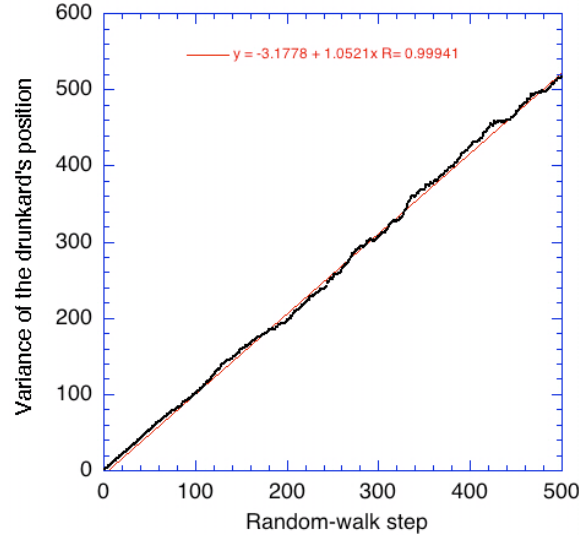


Figure. Variance of the drunkard's position (1,000 samples.)

## CONTINUUM LIMIT—DIFFUSION EQUATION

Diffusion is central to many stochastic processes. The probability density function,  $P(x, t)$ , is often analyzed by partial differential equations, which is derived as follows. We start from a recursive relation,

$$P(x, t) = \frac{1}{2}P(x-l, t-\tau) + \frac{1}{2}P(x+l, t-\tau),$$

*i.e.*, the probability density is obtained by adding two conditionally probabilities that he was: i) at one step left at the previous time and walked to the right with probability 1/2; and ii) at one step right at the previous time and walked to left with probability 1/2. By subtracting  $P(x, t-\tau)$  from both sides and dividing them by  $\tau$ ,

$$\frac{P(x, t) - P(x, t-\tau)}{\tau} = \frac{l^2}{2\tau} \frac{P(x-l, t-\tau) - 2P(x, t-\tau) + P(x+l, t-\tau)}{l^2}.$$

Let's take the limit that  $\tau \rightarrow 0$  and  $l \rightarrow 0$  with  $l^2/2\tau = D$  is finite. The above equation then becomes

$$\frac{\partial}{\partial t} P(x, t) = D \frac{\partial^2}{\partial x^2} P(x, t).$$

This parabolic equation is known as the diffusion equation.

### CENTRAL-LIMIT THEOREM

Now we will see a manifestation of a very important theorem in probability theory, namely the central-limit theorem. Consider a sequence of random numbers,  $\{y_n \mid n = 1, 2, \dots, N\}$ , which may follow an arbitrary probability density. The sum of all the random variables,  $Y = (y_1 + y_2 + \dots + y_N)$ , itself is a random variable. The central-limit theorem states that this sum follows the normal (Gaussian) distribution for a large  $N$ .

The drunkard's position is a special example of this theorem. Let's rewrite the binomial distribution as

$$P_N(x) = \frac{N!}{\left(\frac{N+x}{2}\right)! \left(\frac{N-x}{2}\right)!} p^{(N+x)/2} q^{(N-x)/2},$$

where  $x = (n_{\rightarrow} - n_{\leftarrow})$ . For  $p = q = 1/2$ ,

$$P_N(x) = \frac{N!}{\left(\frac{N+x}{2}\right)! \left(\frac{N-x}{2}\right)!} \left(\frac{1}{2}\right)^N.$$

For  $N \rightarrow \infty$ , this distribution reduces to the normal distribution,

$$\lim_{N \rightarrow \infty} P_N(x) = P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (1)$$

where  $\sigma = \sqrt{N}$ .

### STIRLING'S FORMULA

The proof of the above limiting behavior requires the knowledge about the asymptotic behavior of factorials. This is answered by the Stirling's theorem,<sup>4</sup>

$$N! = \sqrt{2\pi} N^{N+1/2} e^{-N} \left(1 + \frac{1}{12N} + \dots\right).$$

(Factorial is an extremely fast growing function of its argument!) The proof of Stirling's theorem exemplifies an interesting observation: Integer problems are hard, but approximate solutions to them are often easily obtained by expanding the solution space to the real or sometimes even to the complex numbers. This is particularly true for asymptotic behaviors, since  $N$  is so large that the discrete unit,  $1 \ll N$ , is negligible.

$\therefore$  Let's first define the **gamma function**,

$$\Gamma(z) \equiv \int_0^\infty e^{-t} t^{z-1} dt \quad (z \in \mathbb{C}; \text{Re } z > 0).$$

The factorial  $n!$  is a special case of the gamma function where  $z$  is an integer. To prove this, let's recall a recursive relation for the gamma function,

<sup>4</sup> G. Arfken, *Mathematical Methods for Physicists*, 3rd Ed. (Academic Press, San Diego, 1985).



$$\Gamma(z+1) = z \Gamma(z),$$

which is easily proven by integrating by part,

$$[f(x)g(x)]_a^b = \int_a^b \frac{d}{dx} [f(x)g(x)] dx = \int_a^b f(x)g'(x) dx + \int_a^b f'(x)g(x) dx,$$

where  $f'(x) = df(x)/dx$ :

$$\Gamma(z+1) = \int_0^\infty e^{-t} t^z dt = \left[ -e^{-t} t^z \right]_0^\infty - \int_0^\infty (-e^{-t}) z t^{z-1} dt = z \Gamma(z).$$

Also note that

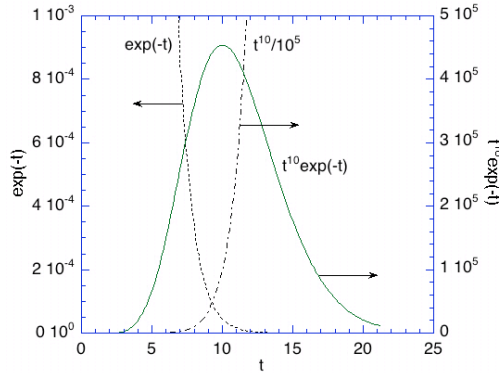
$$\Gamma(0) = \int_0^\infty e^{-t} dt = \left[ -e^{-t} \right]_0^\infty = 1.$$

Therefore,  $\Gamma(N+1) = N\Gamma(N) = N(N-1)\Gamma(N-1) = \dots = N!$ .

Now let's perform an asymptotic expansion of

$$\Gamma(z+1) \equiv \int_0^\infty e^{-t} t^z dt.$$

To get a handle on this, you should first plot the integrand,  $f(t) = e^{-t} t^z$ , for a large  $z$ .

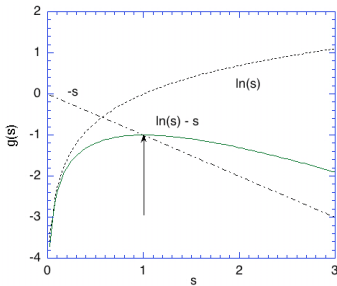


**Figure:** Integrand of the gamma function  $\Gamma(z)$  for  $z = 10$ .

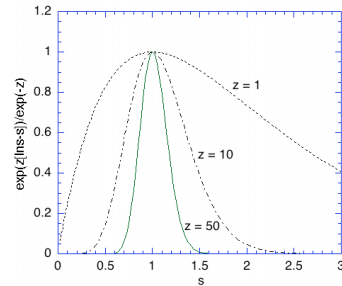
Note that the most significant contribution to the integral comes from the maximum of  $f(t)$ , which is located by  $df/dt = e^{-t} t^{z-1} (-t + z) = 0$ , as  $t = z$ . As we increase  $z$ , you will notice that the distribution of this function becomes sharper and sharper around its peak. Our strategy is thus to expand the integrand around its maximum. Since everything occurs near  $t \sim z$  (very big), let's scale the integration variable as  $t = zs$ , so that the main contribution to the integral comes from  $s \sim 1$ .

$$\Gamma(z+1) = \int_0^\infty e^{-zs} (zs)^z z ds = z^{z+1} \int_0^\infty e^{-zs} \exp(z \ln s) ds = z^{z+1} \int_0^\infty \exp(z(\ln s - s)) ds.$$

Now the function,  $g(s) = \ln s - s$ , is peaked at  $s = 1$  ( $dg/ds = 1/s - 1 = 0$  at  $s = 1$ ).



**Figure:** Function  $g(s) = \ln s - s$ .



**Figure.** Function  $\exp(z(\ln s - s)) / \exp(-z)$ .

Note that the exponential function with a large prefactor in its argument acts as a discriminator. It emphasizes the maximum value and makes the other regions less and less significant for larger prefactors (see the Figure above right).

Since the most significant contribution comes from a very narrow range near  $s = 1$  for a large  $z$ , let's expand  $g(s)$  around  $s = 1$ . Note that  $g'(s) = 1/s - 1$ ,  $g''(s) = -1/s^2$ , ..., so that  $g(1) = -1$ ,  $g'(1) = 0$ ,  $g''(1) = -1$ , ... The Taylor expansion of  $g(s)$  around  $s = 1$  is thus,

$$\begin{aligned} g(s) &= g(1) + g'(1)(s-1) + \frac{1}{2}g''(1)(s-1)^2 + \dots \\ &= -1 - \frac{1}{2}(s-1)^2 + \dots \end{aligned}$$

Substituting this expansion in the integrand, we obtain

$$\begin{aligned} \Gamma(z+1) &= z^{z+1} \int_0^\infty ds \exp\left(z\left[-1 - \frac{1}{2}(s-1)^2 + \dots\right]\right) \\ &= z^{z+1} e^{-z} \int_0^\infty ds \exp\left(-\frac{z}{2}(s-1)^2 + \dots\right) \\ &\approx z^{z+1} e^{-z} \int_{-\infty}^\infty ds \exp\left(-\frac{z}{2}(s-1)^2\right) \\ &= z^{z+1} e^{-z} \sqrt{\frac{2}{z}} \int_{-\infty}^\infty du \exp(-u^2) \\ &= \sqrt{2\pi} z^{z+1/2} e^{-z} \end{aligned}$$

Here we have changed the variable to  $u = \sqrt{z/2}(s-1)$ , and used the fact that the function is so concentrated around  $s = 1$  that changing the lower limit of the integration range from  $-1$  to  $-\infty$  does not affect the result. (We have only derived the leading term in the Stirling's formula. The other terms can be obtained by keeping subsequent terms in the above Taylor expansion.) //

---

## PROOF OF EQUATION 1

By substituting the leading term of the Stirling's expansion into the binomial probability density,

$$\begin{aligned} \frac{N!}{\left(\frac{N+x}{2}\right)! \left(\frac{N-x}{2}\right)!} \left(\frac{1}{2}\right)^N &= \frac{1}{\sqrt{2\pi}} \frac{N^{N+1/2}}{\left(\frac{N+x}{2}\right)^{\frac{N+x}{2} + \frac{1}{2}} \left(\frac{N-x}{2}\right)^{\frac{N-x}{2} + \frac{1}{2}}} \exp\left(-N + \frac{N+x}{2} + \frac{N-x}{2}\right) \left(\frac{1}{2}\right)^N \\ &= \frac{1}{\sqrt{2\pi}} \frac{\left(\frac{N}{2}\right)^N}{\left(\frac{N+x}{2}\right)^{\frac{N+x}{2}} \left(\frac{N-x}{2}\right)^{\frac{N-x}{2}}} \left(\frac{4N}{(N+x)(N-x)}\right)^{1/2} \\ &= \sqrt{\frac{2}{\pi N}} \frac{1}{\left(1 + \frac{x}{N}\right)^{\frac{N+x}{2}} \left(1 - \frac{x}{N}\right)^{\frac{N-x}{2}}} \end{aligned}$$

Consider

---

---


$$\ln \left[ \left(1 + \frac{x}{N}\right)^{\frac{N+x}{2}} \left(1 - \frac{x}{N}\right)^{\frac{N-x}{2}} \right] = \frac{N}{2} \left(1 + \frac{x}{N}\right) \ln \left(1 + \frac{x}{N}\right) + \frac{N}{2} \left(1 - \frac{x}{N}\right) \ln \left(1 - \frac{x}{N}\right).$$

We know that the standard deviation of this distribution is  $\sqrt{N}$ , so that  $x \ll N$  in the range where  $P_N(x)$  has any significant value. By expanding the above expression in  $x/N$  and retaining only the leading term, we get

$$\begin{aligned} \ln \left[ \left(1 + \frac{x}{N}\right)^{\frac{N+x}{2}} \left(1 - \frac{x}{N}\right)^{\frac{N-x}{2}} \right] &= \frac{N}{2} \left(1 + \frac{x}{N}\right) \left( \frac{x}{N} - \frac{1}{2} \left(\frac{x}{N}\right)^2 \right) + \frac{N}{2} \left(1 - \frac{x}{N}\right) \left( -\frac{x}{N} - \frac{1}{2} \left(\frac{x}{N}\right)^2 \right) \\ &= \frac{x}{2} \left(1 + \frac{x}{N}\right) \left(1 - \frac{1}{2} \frac{x}{N}\right) - \frac{x}{2} \left(1 - \frac{x}{N}\right) \left(1 + \frac{1}{2} \frac{x}{N}\right) \\ &= \frac{x}{2} \left(1 + \frac{1}{2} \frac{x}{N}\right) - \frac{x}{2} \left(1 - \frac{1}{2} \frac{x}{N}\right) = \frac{x^2}{2N} \end{aligned}$$

where we have used the expansion,

$$\ln(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \dots$$

Therefore

$$\begin{aligned} \frac{N!}{\left(\frac{N+x}{2}\right)! \left(\frac{N-x}{2}\right)!} \left(\frac{1}{2}\right)^N &= \sqrt{\frac{2}{\pi N}} \frac{1}{\left(1 + \frac{x}{N}\right)^{\frac{N+x}{2}} \left(1 - \frac{x}{N}\right)^{\frac{N-x}{2}}} \\ &= \sqrt{\frac{2}{\pi N}} \exp\left(-\frac{x^2}{2N}\right) \\ &= \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \end{aligned}$$

where  $\sigma = \sqrt{N}$ . //

---

(Normalization)

Note that the binomial distribution function satisfies the following normalization relation,

$$\sum_{n_{\rightarrow}=0}^N P_N(x) = \frac{N!}{\left(\frac{N+x}{2}\right)! \left(\frac{N-x}{2}\right)!} \left(\frac{1}{2}\right)^N = 1.$$

For  $n_{\rightarrow} = 0, 1, \dots$ ,  $x = n_{\rightarrow} - n_{\leftarrow} = 2n_{\rightarrow} - N = -N, -N+2, \dots$ . Therefore  $x$  values are distributed uniformly with stride 2. Now let's define a continuous probability density function,  $P(x)$ , such that the number of sample points generated by  $N_{\text{try}}$  trials in the range  $[x, x+\Delta x]$  is  $N_{\text{try}} P(x) \Delta x$ .

$$N_{\text{try}} \Delta x P(x) = N_{\text{try}} \frac{\Delta x}{2} \frac{N!}{\left(\frac{N+x}{2}\right)! \left(\frac{N-x}{2}\right)!} \left(\frac{1}{2}\right)^N.$$

The factor  $\Delta x/2$  is the number of possible  $x$  values in the range. Therefore,

$$P(x) = \lim_{N \rightarrow \infty} \frac{1}{2} \frac{N!}{\left(\frac{N+x}{2}\right)! \left(\frac{N-x}{2}\right)!} \left(\frac{1}{2}\right)^N = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right),$$

where  $\sigma = \sqrt{N}$ .

### §3 Random Walks in Finance<sup>5</sup>

#### GEOMETRIC BROWNIAN MOTION

Stock price,  $S(t)$ , as a function of time  $t$ , is a random variable. Time evolution of a stock price is often idealized as a **diffusion process**,

$$dS = \mu S dt + \sigma S \varepsilon \sqrt{dt},$$

where  $\mu$  is the **drift** term (or the expected rate of return on the stock),  $\sigma$  is the **volatility** of the stock price, and  $\varepsilon$  is a random variable following the normal distribution with unit variance.

Suppose the second, stochastic term is zero, then the solution to the above differential equation is

$$S(t) = S_0 \exp(\mu t).$$

(Confirm that the above solution satisfies  $dS/dt = \mu S$ .) Therefore the first term in right-hand side of the differential equation describes the stock-price growth at a compounded rate of  $\mu$  per unit time.

Suppose, on the other hand, the first term is zero (no growth). Let's define  $U = \ln S$  so that  $dU = dS/S$ . Then the above differential equation leads to

$$dU = \sigma \varepsilon \sqrt{dt}.$$

Or

$$U(t) - U(0) = \sigma \sqrt{\Delta t} \sum_{i=1}^N \varepsilon_i.$$

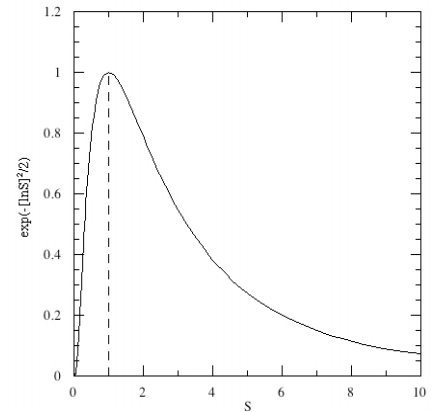
According to the central-limit theorem, the sum over  $N$  random variables,  $E = \sum \varepsilon_i$ , follows the normal distribution with variance  $N$ . By defining  $t = N\Delta t$ ,

$$U(t) - U(0) = \sigma \sqrt{t} \varepsilon.$$

Namely the logarithm of  $U(t)$  is a diffusion process whose variance scales as  $t$ . ( $\sigma$  is the diffusion constant.)  $S(t)$ , whose logarithm follows the normal distribution, is said to follow the **log-normal distribution**.

#### MC SIMULATION OF STOCK PRICE

An MC simulation of a stock price is performed by interpreting the time-evolution equation to be discrete ( $dt$  is small but finite). At each MC step, stock-price increment relative to its current price,  $dS/S$ , follows a normal distribution which has a mean value,  $\mu dt$ , and standard deviation  $\sigma \sqrt{dt}$  (or variance  $\sigma^2 dt$ ). Or you can generate the increment as



**Figure:** Log-normal distribution.

<sup>5</sup> J. C. Hull, *Options, Futures, and Other Derivatives*, 3rd Ed. (Prentice Hall, 1997); E. Barucci, L. Landi, and U. Cherubini, *IEEE Comput. Sci. & Eng.*, Spring 1996, p. 66.

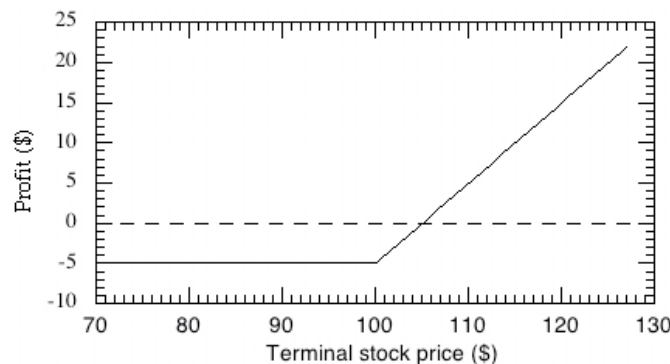
$$\frac{dS}{S} = \mu dt + \sigma \sqrt{dt} \xi,$$

where  $\xi$  is a random number which follows a normal distribution with variance 1 (you can use the Box-Muller algorithm to generate  $\xi$ ).

## BLACK-SCHOLES ANALYSIS<sup>6</sup>

We will not get into the details of the Black-Scholes analysis of an option price. However, let's look briefly at what it does. It determines the price of options.

A (European<sup>7</sup>) **call<sup>8</sup> option** gives its holder the right to buy the underlying asset at a certain date (called the **expiration date** or maturity) for a certain price (called the **strike price**). Note that an option gives the holder the right to do something but that the holder does not have to exercise this right. Consider an investor who buys an European call option on IBM stock with a strike price of \$100. Suppose that the current stock price is \$98, the expiration date is in two months, and the option price is \$5. If the stock price on the expiration day is less than \$100, he or she will clearly not exercise. (There is no point in buying for \$100 a stock that has a market value of less than \$100.) In this circumstance the investor loses the entire initial investment of \$5. Suppose, for example, the stock price is \$115. By exercising the option, the investor buys a stock for \$100. If the share is sold immediately, he or she makes a gain of \$15. The net profit is \$10 by subtracting the initial investment from the gain.



**Figure:** Profit from buying a call option: option price is \$5, strike price is \$100.

The Black-Scholes analysis determines the price of an option based on the assumptions:

- i) The underlying stock price follows the simple diffusive equation in the previous page;
- ii) In a competitive market, there are no risk-less arbitrage<sup>9</sup> opportunities;
- iii) The risk-free rate of interest,  $r$ , is constant and the same for all risk-free investments.

The main observation is that the option price, which is a function of the underlying stock price, itself a stochastic process, which depends on the same random variable  $\varepsilon$ . By constructing a portfolio that contains a right combination of the option and the stock, we can eliminate the random contribution to the growth rate of the portfolio. From the no arbitrage principle above, the growth rate of such a risk-less portfolio must be  $r$ . The resulting equation gives a partial differential equation that must be followed by the price,  $f$ , of the call option,

<sup>6</sup> F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy* **81**, 637 (1973). See also <http://www.nobel.se/economics/laureates/1997>, 1997 Nobel Economy Prize homepage.

<sup>7</sup> An option that can be exercised only at the expiration date. In contrast, an American option can be exercised at any time up to the expiration date.

<sup>8</sup> Put option, on the other hand, is the right to sell.

<sup>9</sup> Buying/selling portfolios of financial assets in such a way as to make a profit in a risk-free manner.

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial^2 S} = rf. {}^{10}$$

## §4 Random Number Generator

### LINEAR CONGRUENTIAL GENERATOR<sup>11</sup>

- **Uniform random number generator:** A routine to generate a sequence of random numbers within a specified range (typically 0 to 1).

The simplest uniform random number generator uses a sequence of large positive integers  $X_i$ , each generated from the previous one by multiplication conducted in finite modulus arithmetic:

$$X_{i+1} = aX_i \bmod M,$$

where  $a$  and  $M$  are large positive integers and  $\bmod$  is the modulo operator. Each integer in the sequence lies in the range  $[0, M-1]$ . Random numbers  $r_i$  in the range  $[0, 1]$  is obtained by

$$r_i = X_i/M.$$

Since  $X_i$  can never exceed  $M-1$  the sequence must repeat after at most  $M-1$  iterations. If  $M$  is chosen to be prime and  $a$  is carefully chosen, the maximum possible period of  $M-1$  is achieved.

The most famous generator is  $(M, a) = (2^{31}-1 = 2147483647, 7^5 = 16807)$ , in which  $M$  is the largest prime that can be represented by a signed 32-bit integer. This cannot be implemented using the 32-bit integer arithmetic since the multiplication can easily overflow.

The following algorithm (due to Schrage) solves this difficulty by using a factorization of  $m$ ,

$$m = aq + r, \text{ i.e., } q = [m/a], r = m \bmod a$$

with square brackets denoting integer part. If  $r < q$ , then we can show, for an arbitrary  $z$  in the range  $(0, m-1)$ , that both  $a(z \bmod q)$  and  $r[z/q]$  lie in the range  $[0, m-1]$  and

$$az \bmod m = \begin{cases} a(z \bmod q) - r[z/q] & \text{if } \geq 0 \\ a(z \bmod q) - r[z/q] + m & \text{otherwise} \end{cases}.$$

---

$\therefore$  Let factorize  $z = xq + y$  ( $y < q$ : remainder) so that  $x = [z/q]$  and  $y = z \bmod q$ . Then

$$\begin{aligned} az &= a(xq + y) \\ &= x(m - r) + ay \\ &= xm + (ay - xr). \end{aligned}$$

This suggests that  $ay - xr$  is equal to  $az \bmod m$  in finite modulus arithmetic. Now note that

$$0 < ay < aq = m - r \leq m, \text{ therefore } 0 < ay < m$$

$$\begin{aligned} 0 < xr &\leq ar \text{ (} z < m \text{ and } x \text{ \& } a \text{ are their respective quotients)} < aq \text{ (} r \text{ is a remainder with divider } q) \\ &= m - r \leq m, \text{ therefore } 0 < xr < m \end{aligned}$$

---

Therefore  $-m < ay - xr < m$ , and it is equal to either  $(az \bmod m)$  itself or  $(az \bmod m - m)$ . //

---

<sup>10</sup> This equation is worth a Nobel prize!

<sup>11</sup> M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Univ. Press, Oxford, 1987) Appendix G; W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C, 2nd Ed.* (Cambridge Univ. Press, Cambridge, 1992) p. 274.

The following random number generator uses this algorithm with  $q = [2147483647/16807] = 127773$  and  $r = 2147483647 - 16807 \cdot 127773 = 2836$ .

- The cycle of this random number generator is  $m - 1 = 2^{31} - 2 \sim 2.1 \times 10^9$ : You can obtain up to 2 billion independent random numbers.
- Program ran0.c

```
#define IA 16807
#define IM 2147483647
#define AM (1.0/IM)
#define IQ 127773
#define IR 2836
float ran0(int *idum) {
/*
Random number generator. Returns a uniform random number between 0.0 and 1.0.
Set *idum (to be modified) to any nonzero integer value to initialize the sequence.
*/
    long k;
    float ans;

    k = (*idum)/IQ;
    *idum = IA*(*idum - k*IQ) - IR*k;
    if (*idum < 0) *idum += IM;
    ans = AM*(*idum);
    return ans;
}
```

The “perfect random-number generator” in *Numerical Recipes* by W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (Cambridge Univ. Press) is an extension of this routine with “shuffle” and “mix” operations.

## NONUNIFORM RANDOM NUMBER GENERATION: TRANSFORMATION METHOD

- **Maxwell-Boltzmann distribution:** Let’s consider a particle with mass  $m$ . If this particle is in equilibrium at temperature  $T$ , the equipartition principle in the statistical mechanics state its velocity  $v$  in any one of the  $x$ ,  $y$ , and  $z$  directions is a random variable with the probability density,

$$\rho(v) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left(-\frac{mv^2}{2k_B T}\right).$$

This is in fact the “**normal**” distribution with standard deviation  $\sigma = \sqrt{k_B T/m}$ .

- **Normalization:**

$$\begin{aligned} \int_{-\infty}^{\infty} dv \rho(v) &= \sqrt{\frac{m}{2\pi k_B T}} \int_{-\infty}^{\infty} \sqrt{\frac{2k_B T}{m}} ds \exp(-s^2) \quad (s = \sqrt{\frac{m}{2k_B T}} v) \\ &= \sqrt{\frac{1}{\pi}} \int_{-\infty}^{\infty} ds \exp(-s^2) = 1. \end{aligned}$$

(See Appendix A for the last equality.)

- **Box-Muller algorithm:** Generates a random velocity  $v$ , which follows the above normal distribution. Let  $v = \sigma \zeta$  so that  $\zeta$  is the normally distributed random number with unit variance,

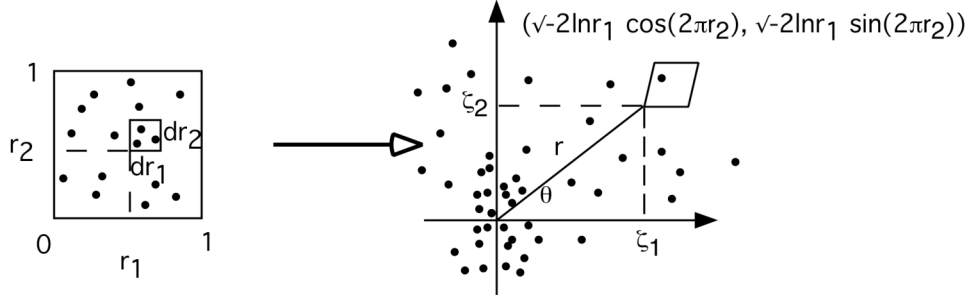
$$\rho(\xi) = \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{\xi^2}{2}\right).$$

- (a) Generate uniform random numbers  $r_1$  and  $r_2$  in the range  $(0, 1)$ ;
- (b) Calculate  $\xi_1 = (-2\ln r_1)^{1/2} \cos(2\pi r_2)$  and  $\xi_2 = (-2\ln r_1)^{1/2} \sin(2\pi r_2)$ .

Both  $\xi_1$  and  $\xi_2$  are the desired normally distributed random number.

- **Coordinate transformation:** An arbitrary pair of velocities  $(\xi_1, \xi_2)$  can be specified using an alternative coordinate system, the polar coordinates,  $(r, \theta)$  ( $0 \leq r < \infty$ ;  $0 \leq \theta < 2\pi$ ), where  $\xi_1 = r \cos \theta$  and  $\xi_2 = r \sin \theta$ . Furthermore the entire 2D Euclidean space  $(-\infty, \infty)^2$  can be mapped back to a unit square by the following transformation:

$$r = \sqrt{-2\ln r_1} \quad \theta = 2\pi r_2 \quad (0 \leq r_1, r_2 < 1).$$



Now the question is: “If we uniformly generate  $N$  random points  $(r_1, r_2)$  in the unit square, what will be the distribution of  $(\xi_1, \xi_2)$  in the infinite 2D space according to the above coordinate transformation?” In order to answer this question, let’s count the number of sample points in the small area at  $(r_1, r_2)$  and with edges  $(dr_1, dr_2)$ , which should be approximately  $Ndr_1dr_2$ . These sample points, once transformed to the  $(\xi_1, \xi_2)$  space, occupy a skewed rectangular whose area is given by (see Appendix B)

$$\left| \frac{\partial(\xi_1, \xi_2)}{\partial(r_1, r_2)} \right| dr_1 dr_2,$$

where  $|\partial(\xi_1, \xi_2)/\partial(r_1, r_2)| = |\partial\xi_1/\partial r_1 \partial\xi_2/\partial r_2 - \partial\xi_1/\partial r_2 \partial\xi_2/\partial r_1|$  is called the Jacobian of the coordinate transformation. Consider a (not skewed) small rectangle around  $(\xi_1, \xi_2)$  of size  $(d\xi_1, d\xi_2)$ . The number of sample points in this area is given by

$$Ndr_1dr_2 \times \frac{d\xi_1 d\xi_2}{|\partial(\xi_1, \xi_2)/\partial(r_1, r_2)| dr_1 dr_2} = N \left| \frac{\partial(r_1, r_2)}{\partial(\xi_1, \xi_2)} \right| d\xi_1 d\xi_2.$$

Here, we have used the relation

$$\left| \frac{\partial(r_1, r_2)}{\partial(\xi_1, \xi_2)} \right| = \frac{1}{\left| \frac{\partial(\xi_1, \xi_2)}{\partial(r_1, r_2)} \right|}.$$

The definition of the probability density function  $(\xi_1, \xi_2)$  is that this number be equal to

$$Np(\xi_1, \xi_2) d\xi_1 d\xi_2,$$

so that

$$p(\xi_1, \xi_2) = \left| \frac{\partial(r_1, r_2)}{\partial(\xi_1, \xi_2)} \right|.$$



Note that

$$r_1 = \exp\left[-\frac{\xi_1^2 + \xi_2^2}{2}\right]$$

$$r_2 = \frac{1}{2\pi} \arctan \frac{\xi_2}{\xi_1}$$

so that

$$\frac{\partial r_1}{\partial \xi_1} = -\xi_1 \exp\left[-\frac{\xi_1^2 + \xi_2^2}{2}\right]$$

$$\frac{\partial r_1}{\partial \xi_2} = -\xi_2 \exp\left[-\frac{\xi_1^2 + \xi_2^2}{2}\right]$$

$$\frac{\partial r_2}{\partial \xi_1} = -\frac{\xi_2 \cos^2(2\pi r_2)}{2\pi \xi_1^2} = -\frac{\xi_2}{2\pi \xi_1^2 (1 + \tan^2(2\pi r_2))} = -\frac{\xi_2}{2\pi (\xi_1^2 + \xi_2^2)}$$

$$\frac{\partial r_2}{\partial \xi_2} = \frac{\cos^2(2\pi r_2)}{2\pi \xi_1} = \frac{1}{2\pi \xi_1 (1 + \tan^2(2\pi r_2))} = \frac{\xi_1}{2\pi (\xi_1^2 + \xi_2^2)}$$

and the Jacobian is

$$\left| -\xi_1 \exp\left[-\frac{\xi_1^2 + \xi_2^2}{2}\right] \frac{\xi_1}{2\pi (\xi_1^2 + \xi_2^2)} - \xi_2 \exp\left[-\frac{\xi_1^2 + \xi_2^2}{2}\right] \frac{\xi_2}{2\pi (\xi_1^2 + \xi_2^2)} \right|$$

$$= \frac{1}{2\pi} \exp\left[-\frac{\xi_1^2 + \xi_2^2}{2}\right] = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\xi_1^2}{2}\right] \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\xi_2^2}{2}\right]$$

i.e., the random sample points in the  $(\xi_1, \xi_2)$  space generated in this way follows the product of two 1D normal distributions.

## APPENDIX A—Gaussian Integral

Consider the Gaussian integral,

$$I = \int_{-\infty}^{\infty} dx \exp(-x^2).$$

A common trick is instead to evaluate

$$I^2 = \int_{-\infty}^{\infty} dx \exp(-x^2) \int_{-\infty}^{\infty} dy \exp(-y^2)$$

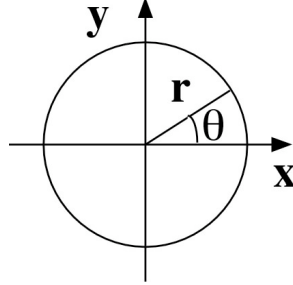
$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp(-(x^2 + y^2)) ,$$

and to introduce a polar coordinate,  $(x = r \cos \theta, y = r \sin \theta)$ . Note that

$$\begin{aligned}\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy &= \int_0^{\infty} \int_0^{2\pi} dr d\theta \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \int_0^{\infty} \int_0^{2\pi} dr d\theta \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = \int_0^{\infty} \int_0^{2\pi} dr d\theta |r \cos^2 \theta + r \sin^2 \theta| \\ &= \int_0^{\infty} \int_0^{2\pi} r dr d\theta\end{aligned}$$

(2D integration is a summation of function values multiplied by a small area, see Appendix B). Using this transformation,

$$I^2 = \int_0^{\infty} dr \int_0^{2\pi} d\theta r \exp(-r^2) = 2\pi \int_0^{\infty} dr r \exp(-r^2).$$



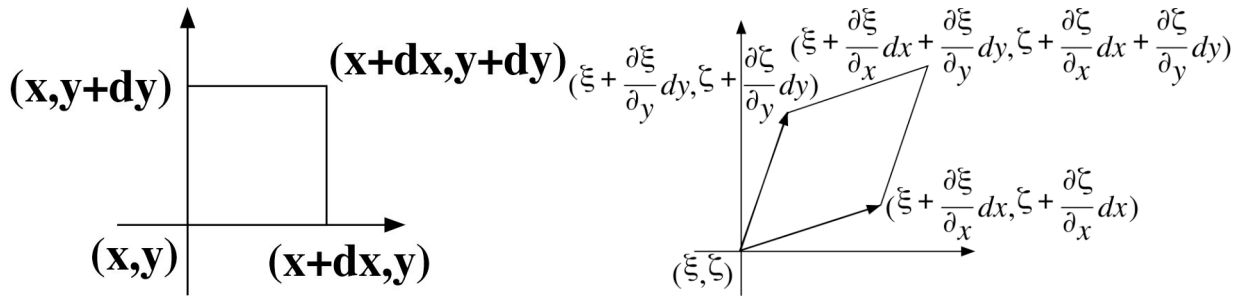
Now a further coordination transformation,  $x = r^2$  and  $dx = (dx/dr)dr = 2rdr$ , makes

$$I^2 = \pi \int_0^{\infty} dx \exp(-x) = \pi [-\exp(-x)]_0^{\infty} = \pi.$$

Therefore,  $I = \int_0^{\infty} dx \exp(-x^2) = \sqrt{\pi}.$

## APPENDIX B—TRANSFORMATION OF AREAL ELEMENT

Q. Consider a variable transformation from  $(x, y)$  to  $(\xi, \zeta) = (\xi(x, y), \zeta(x, y))$  and a tiny rectangle in the  $xy$ -space with area  $dx dy$ , which will be transformed to a parallel-piped in the  $\xi\zeta$ -space shown below. What is the area of the transformed parallel-piped?



A.  $S = \begin{vmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \zeta}{\partial x} & \frac{\partial \zeta}{\partial y} \end{vmatrix} dx dy = \left( \frac{\partial \xi}{\partial x} \frac{\partial \zeta}{\partial y} - \frac{\partial \xi}{\partial y} \frac{\partial \zeta}{\partial x} \right) dx dy$ , where  $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = |ad - bc|$  is the determinant of matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ .

(Proof)

Consider a parallel-piped formed by two 2D vectors  $\vec{k} = (a, b)$  and  $\vec{l} = (c, d)$  and the angle  $\theta$  between them. Note that

$$\cos \theta = \frac{\vec{k} \cdot \vec{l}}{\|\vec{k}\| \|\vec{l}\|} = \frac{ac + bd}{\sqrt{(a^2 + b^2)(c^2 + d^2)}}$$

$$|\sin \theta| = \sqrt{1 - \cos^2 \theta} = \frac{\sqrt{(a^2 + b^2)(c^2 + d^2) - (ac + bd)^2}}{\sqrt{(a^2 + b^2)(c^2 + d^2)}} = \frac{|ad - bc|}{\sqrt{(a^2 + b^2)(c^2 + d^2)}}$$

The area of the parallel-piped is

$$S = \|\vec{k}\| \|\vec{l}\| \sin \theta = \sqrt{a^2 + b^2} \sqrt{c^2 + d^2} \sin \theta = |ad - bc|$$

The parallel-piped under question is instead formed by two vectors,

$$\vec{k} = \left( \frac{\partial \xi}{\partial x} dx, \frac{\partial \xi}{\partial x} dx \right),$$

$$\vec{l} = \left( \frac{\partial \xi}{\partial y} dy, \frac{\partial \xi}{\partial y} dy \right).$$

Here, we have used the linear approximation,

$$(\xi(x + dx, y + dy), \zeta(x + dx, y + dy)) \approx (\xi(x, y) + \frac{\partial \xi}{\partial x} dx + \frac{\partial \xi}{\partial y} dy, \zeta(x, y) + \frac{\partial \zeta}{\partial x} dx + \frac{\partial \zeta}{\partial y} dy)$$

With this substitution, we get

$$S = \left| \frac{\partial \xi}{\partial x} dx \frac{\partial \zeta}{\partial y} dy - \frac{\partial \zeta}{\partial x} dx \frac{\partial \xi}{\partial y} dy \right| = \left| \frac{\partial \xi}{\partial x} \frac{\partial \zeta}{\partial y} - \frac{\partial \zeta}{\partial x} \frac{\partial \xi}{\partial y} \right| dx dy = \begin{vmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \zeta}{\partial x} & \frac{\partial \zeta}{\partial y} \end{vmatrix} dx dy.$$

## PROBLEM—NORMAL DISTRIBUTION

What is the standard deviation of the random number,  $\zeta$ , that follow the normal probability density,

$$P(\zeta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\zeta^2}{2\sigma^2}\right)?$$

(Answer)

Note that

$$I(\sigma) = \int_{-\infty}^{\infty} d\zeta \exp\left(-\frac{\zeta^2}{2\sigma^2}\right) = \int_{-\infty}^{\infty} \sqrt{2}\sigma ds \exp(-s^2) = \sqrt{2\pi}\sigma,$$

where we have introduced a new variable,  $s$ , through  $\zeta = \sqrt{2}\sigma s$ . By differentiate both sides by  $\sigma$ ,

$$\frac{dI}{d\sigma} = \int_{-\infty}^{\infty} d\zeta \frac{\zeta^2}{\sigma^3} \exp\left(-\frac{\zeta^2}{2\sigma^2}\right) = \sqrt{2\pi}.$$

Or

$$\int_{-\infty}^{\infty} \frac{d\xi}{\sqrt{2\pi}\sigma} \xi^2 \exp\left(-\frac{\xi^2}{2\sigma^2}\right) = \langle \xi^2 \rangle = \sigma^2.$$

From the symmetry, the average value,  $\langle \xi \rangle = 0$ , and therefore the variance of  $\xi$  is  $\langle \xi^2 \rangle - \langle \xi \rangle^2 = \sigma^2$  and the standard deviation is  $\sigma$ .

## APPENDIX C—DERIVATION OF THE BLACK-SCHOLES EQUATION

Let us assume that the stock price  $S$  is the geometric diffusion process as described in the lecture,

$$dS = \mu S dt + \sigma S \varepsilon \sqrt{dt}, \quad (C1)$$

Suppose that  $f$  is the price of a call option contingent on  $S$ . Ito's lemma (K. Ito, 1951) states that the time change in  $f$  during  $dt$  is given by

$$df = \left( \frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S \varepsilon \sqrt{dt}.^{12} \quad (C2)$$

---

$\therefore$  Equation (2) is understood as the Taylor expansion as follows:

$$\begin{aligned} & f(S + dS, t + dt) - f(S, t) \\ &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S} dS + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (dS)^2 + \dots \\ &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S} (\mu S dt + \sigma S \varepsilon \sqrt{dt}) + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (\mu S dt + \sigma S \varepsilon \sqrt{dt})^2 + \dots \\ &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial S} (\mu S dt + \sigma S \varepsilon \sqrt{dt}) + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (\mu^2 S^2 dt^2 + 2\mu\sigma S^2 \varepsilon dt \sqrt{dt} + \sigma^2 S^2 \varepsilon^2 dt) + \dots \\ &= \left( \frac{\partial f}{\partial S} \sigma S \varepsilon \right) (dt)^{1/2} + \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial S} \mu S + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \varepsilon^2 \right) dt + O((dt)^{3/2}) \end{aligned} \quad (C3)$$

In the above derivation, we have arranged the terms in increasing power of  $(dt)^{1/2}$  and retained the terms up to  $dt$ . Now consider the random variable,

$$\frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \varepsilon^2 dt, \quad (C4)$$

which appears in the last term of the order  $dt$ . Since  $\varepsilon$  follows the normal distribution of variance 1 (i.e.,  $\langle \varepsilon^2 \rangle = 1$ ), its expectation value is

$$\frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 dt. \quad (C5)$$

Though the random variable (C4) fluctuates around the mean value (C5), the effect of the fluctuation on the growth of  $f$  is higher order in  $dt$  and can be neglected. Consequently, we can regard this term as deterministic and hence

---

<sup>12</sup> Ito received the first Gauss prize from the International Mathematical Union in 2006 for his work on stochastic differential equations including this equation.

---


$$f(S + dS, t + dt) - f(S, t) = \left( \frac{\partial f}{\partial S} \sigma S \varepsilon \right) (dt)^{1/2} + \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial S} \mu S + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + O((dt)^{3/2}), \quad (C6)$$


---

which proves Eq. (C2). //

---

Now let us define a portfolio, which is a linear combination of  $S$  and  $f$  as follows:

$$\Pi = -f + \frac{\partial f}{\partial S} S. \quad (C7)$$

Then its time change during  $dt$  is given by

$$\begin{aligned} d\Pi &= -df + \frac{\partial f}{\partial S} dS \\ &= -\left( \frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt - \frac{\partial f}{\partial S} \sigma S \varepsilon \sqrt{dt} + \frac{\partial f}{\partial S} (\mu S dt + \sigma S \varepsilon \sqrt{dt}) \\ &= -\left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt \end{aligned} \quad (8)$$

Because the random terms arising from  $df$  and  $dS$  cancel each other, the change  $d\Pi$  is deterministic (i.e., risk-free). From the assumption, the growth rate of such a portfolio is equal to the risk-free interest rate,  $r$ . Therefore,

$$d\Pi = -\left( \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt = r\Pi dt = r\left( f - \frac{\partial f}{\partial S} S \right) dt. \quad (C9)$$

Dividing Eq. (C9) by  $dt$  and rearranging the terms, we obtain the Black-Scholes equation,

$$\frac{\partial f}{\partial t} + r \frac{\partial f}{\partial S} S + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 = rf. \quad (C10)$$