

# Modelling of Complex Systems

## Probability distributions

### Outline:

- Binomial, Poisson, and Gaussian distributions
- Central limit theorem
- Sampling random numbers

## 1. Gambling with a slot machine at a casino.

There are three windows and the combination 777 gives you a prize.  $M$  numbers (or symbols) appear at random in three windows. Each symbol appears with probability  $1/M$ .

### Problem:

You have  $N$  coins. What is a probability to win  $n$  times?

The probability to get a combination 777 (**to win**) is

$$p = \frac{1}{M^3}$$

The probability to get another combination (**to lose**) is

$$1 - p$$




We represent  $N$  attempts as a sequence of white and black windows.

**A white window - you lose; a black window - you win.**

**We define  $B(n, N)$  as a probability to win  $n$  times after  $N$  attempts.**

$n = 0$   probability  $B(0, N) = (1 - p)^N$

$n = 1$   probability  $B(1, N) = Np(1 - p)^{N-1}$

$n = 2$   probability  $B(2, N) = \frac{N(N-1)}{2} p^2 (1 - p)^{N-2}$

**Probability to win  $n$  times ( $n$  black windows) is**

$$B(n, N) = C_n^N p^n (1 - p)^{N-n}$$

where the binomial coefficient is

$$C_n^N = \frac{N!}{(N - n)! n!}$$

$C_n^N$  gives us the number of ways to distribute  $n$  black windows among  $N$  possible positions.  **$B(n, N)$  is called “the binomial distribution”.**

## 2. Russian roulette

**Russian roulette** is a deadly game in which a player places a single *bullet* in a *revolver*, spins the *cylinder*, and pulls the *trigger*.

For a six-shot revolver the probability to kill himself is

$$p = \frac{1}{6},$$

the probability to survive is

$$1 - p.$$

**Rules of the game:**  $N$  people spin the cylinder one by one. A new bullet replaces the used one.

Using the same approach as for gambling, we find that the probability that  $n$  people ( $n$  *black windows*) will die is  $B(n, N)$ .



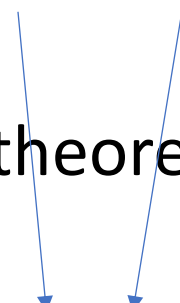
## Properties of the binomial distribution function

$$B(n, N) = C_n^N p^n (1 - p)^{N-n}$$

Normalization:

$$\sum_{n=0}^N B(n, N) = \sum_{n=0}^N C_n^N p^n (1 - p)^{N-n} = (p + (1 - p))^N = 1,$$

because of Newton's binomial theorem,

$$\sum_{n=0}^N C_n^N a^n b^{N-n} = (a + b)^N$$


## Properties of the binomial distribution function

$$B(n, N) = C_n^N p^n (1 - p)^{N-n}$$

The mean number of wins (deaths in the Russian roulette):

$$\begin{aligned} \sum_{n=0}^N n B(n, N) &= \sum_{n=0}^N n C_n^N p^n (1 - p)^{N-n} = \sum_{n=0}^N n \frac{N!}{(N-n)! n!} p^n (1 - p)^{N-n} \\ &= \sum_{n=1}^N N p \frac{(N-1)!}{(N-n)! (n-1)!} p^{n-1} (1 - p)^{N-n} \\ &= N p \sum_{n'=0}^{N'} \frac{N'!}{(N' - n')! n'!} p^{n'} (1 - p)^{N' - n'} = N p \end{aligned}$$

## Properties of the binomial distribution function

$$B(n, N) = C_n^N p^n (1 - p)^{N-n}$$

### Limiting cases:

The probability to lose all of the  $N$  attempts is

$$B(n = 0, N) = (1 - p)^N = e^{N \ln(1-p)} \approx e^{-Np} \ll 1$$

The probability to win all of the  $N$  attempts is

$$B(n = N, N) = p^N = e^{-N |\ln p|} \ll 1$$

Unfortunately, it is difficult to make analytical calculations by use of  $B(n, N)$ . It is useful to find an approximate formula in the case

$$p \ll 1, N \gg 1, \text{ and } n \ll N.$$

We use the **Stirling's approximation**:

$$N! \approx \sqrt{2\pi N} N^N e^{-N},$$

The Stirling's formula is amazingly good!

$$\text{Ratio } \sqrt{2\pi N} N^N e^{-N} / N!$$

$N = 1$	0.92
$N = 2$	0.96
$N = 3$	0.97
...	...
$N = 10$	0.99



In the case when

$$N \gg 1, n \ll N, \text{ and } p \ll 1$$

using the Stirling's formula  $N! \approx \sqrt{2\pi N} N^N e^{-N}$ , we find an approximation

$$C_n^N = \frac{N!}{n! (N-n)!} \approx \frac{e^{-n}}{n!} \frac{N^n}{\left(1 - \frac{n}{N}\right)^{N-n+\frac{1}{2}}}$$

Therefore,

$$B(n, N) = C_n^N p^n (1-p)^{N-n} \approx \frac{e^{-n}}{n!} \frac{N^n p^n (1-p)^{N-n}}{\left(1 - \frac{n}{N}\right)^{N-n+\frac{1}{2}}}$$
$$= \frac{e^{-n}}{n!} (Np)^n \exp\left[(N-n) \ln(1-p) - \left(N-n+\frac{1}{2}\right) \ln\left(1 - \frac{n}{N}\right)\right]$$

Using the Taylor expansion,  $\ln(1-p) \approx -p$  and  $\ln(1 - n/N) \approx -\frac{n}{N}$ , we get

$$\exp[\dots] \approx \exp[n - Np]$$

Therefore, in the cases when  $p \ll 1$ ,  $N \gg 1$ , and  $n \ll N$ ,

$$B(n, N) = C_n^N p^n (1 - p)^{N-n} \approx \frac{(Np)^n e^{-Np}}{n!} = P_n(Np)$$

Where  $P_n(Np)$  is the **Poisson distribution**

$$P_n(c) = \frac{c^n e^{-c}}{n!}$$

**Properties of  $P_n(c)$ :**


Normalization:

$$\begin{aligned} \sum_{n=0}^{\infty} P_n(c) &= \sum_{n=0}^{\infty} \frac{c^n e^{-c}}{n!} \\ &= e^{-c} \left( \sum_{n=0}^{\infty} \frac{c^n}{n!} \right) = e^{-c} e^c = 1 \end{aligned}$$

## Average value (or expectation):

$$\begin{aligned}\langle n \rangle &= \sum_{n=0}^{\infty} n P_n(c) = \sum_{n=0}^{\infty} n \frac{c^n e^{-c}}{n!} \\ &= e^{-c} c \sum_{n=1}^{\infty} \frac{c^{n-1}}{(n-1)!} = e^{-c} c \sum_{n'=0}^{\infty} \frac{c^{n'}}{n'!} = c e^{-c} e^c = c\end{aligned}$$

## Variance:

$$\begin{aligned}\langle (n - \langle n \rangle)^2 \rangle &= \sum_{n=0}^{\infty} (n - \langle n \rangle)^2 P_n(c) = \sum_{n=0}^{\infty} [n(n-1) + n - 2nc + c^2] P_n(c) \\ &= c^2 + c - 2c^2 + c^2 = c\end{aligned}$$


Thus, the Poisson function  $P_n(c)$  is a probability distribution function of random integer numbers  $n$  with equal mean and variance:

$$\frac{\langle (n - \langle n \rangle)^2 \rangle}{\langle n \rangle} = 1$$

We can use this property to check that a given series of integer random numbers  $a_i$  is generated by the Poisson process.

It is necessary to calculate

$$\frac{\sum_i (a_i - \langle a \rangle)^2}{\sum_i a_i} = 1?$$

In 1837, Simeón-Denis Poisson used this distribution to describe the number wrongful convictions in criminal courts. There are many examples of Poisson distributed random variables, from the number of stars found in a unit of space, to the number of soldiers killed accidentally by horse kicks.

## The Poisson distribution function

$$P_n(c) = \frac{c^n e^{-c}}{n!}$$

is still a complicated function for analytical calculations due to the factorial  $n!$ .

Let us use again the Stirling's formula  $n! \approx \sqrt{2\pi n} n^n e^{-n}$

$$P_n(c) = \frac{c^n e^{-c}}{n!} \approx \frac{c^n e^{-c}}{\sqrt{2\pi n} n^n e^{-n}} = \frac{e^{-c}}{\sqrt{2\pi}} \exp \left[ n \ln c - \left( n + \frac{1}{2} \right) \ln n + n \right].$$

We introduce a function

$$f(n) = n \ln c - \left( n + \frac{1}{2} \right) \ln n + n$$

This function has a maximum at a point where  $\frac{\partial f(n)}{\partial n} = 0$ . Thus,

$$\ln c - \frac{\left( n + \frac{1}{2} \right)}{n} - \ln n + 1 = \ln \frac{c}{n} - \frac{1}{2n} = 0$$

At  $c \gg 1$  we get a solution  $n_0 \approx c$ .

Taylor expansion of  $f(n)$  at the point  $n = n_0$

$$\begin{aligned} f(n) = f(n_0 + (n - n_0)) &\approx f(n_0) + \overbrace{f'(n_0)}^0 (n - n_0) + \frac{1}{2} f''(n_0) (n - n_0)^2 \\ &= f(n_0) + \frac{1}{2} f''(n_0) (n - n_0)^2 \end{aligned}$$

The second derivative

$$f''(n_0) = \frac{\partial}{\partial n} \left( \ln c - \ln n - \frac{1}{2n} \right) = -\frac{1}{n_0} + \frac{1}{n_0^2} \approx -\frac{1}{c}$$

We have  $f(n_0) \approx c$ . Therefore,

$$f(n) \approx c - \frac{1}{2c} (n - c)^2$$

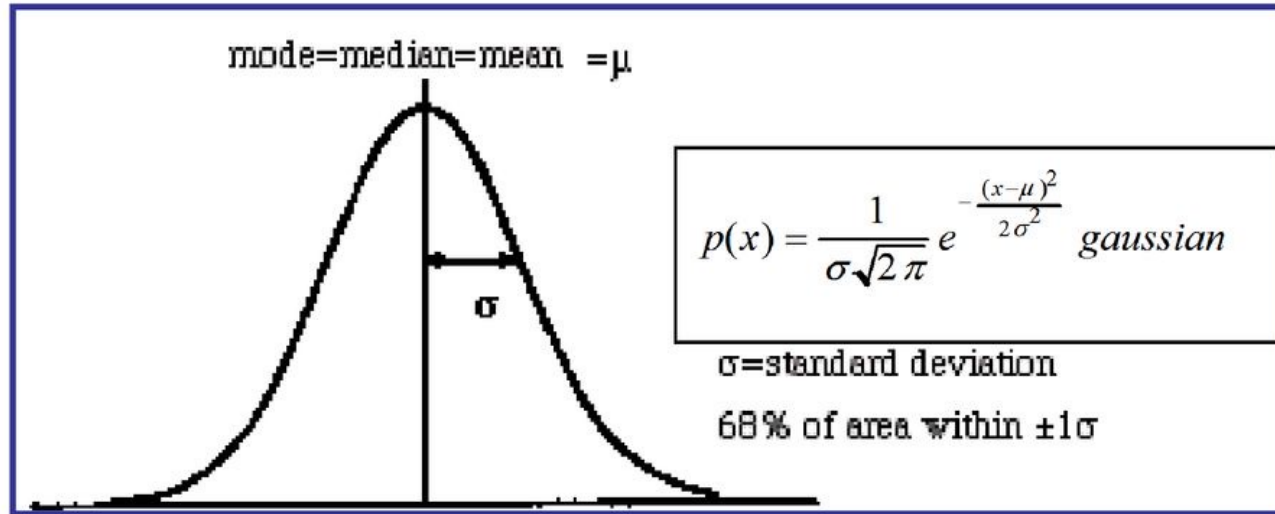
Then we get

$$P_n(c) = \frac{e^{-c}}{\sqrt{2\pi}} \exp[f(n)] \approx \frac{1}{\sqrt{2\pi c}} \exp \left[ -\frac{1}{2c} (n - c)^2 \right],$$

which is the **Gaussian distribution** (or normal distribution) with mean and variance equal to  $c$ .

# Gaussian (Normal) Distribution

- The *Gaussian Distribution* is one of the most used distributions in all of science. It is also called the “bell curve” or the *Normal Distribution*.



In our case, the Gaussian distribution corresponds to the mean value  $\mu = c$  and the standard deviation  $\sigma = \sqrt{c}$ .

## Summary

Binomial distribution  $B(n, N) = C_n^N p^n (1 - p)^{N-n}$

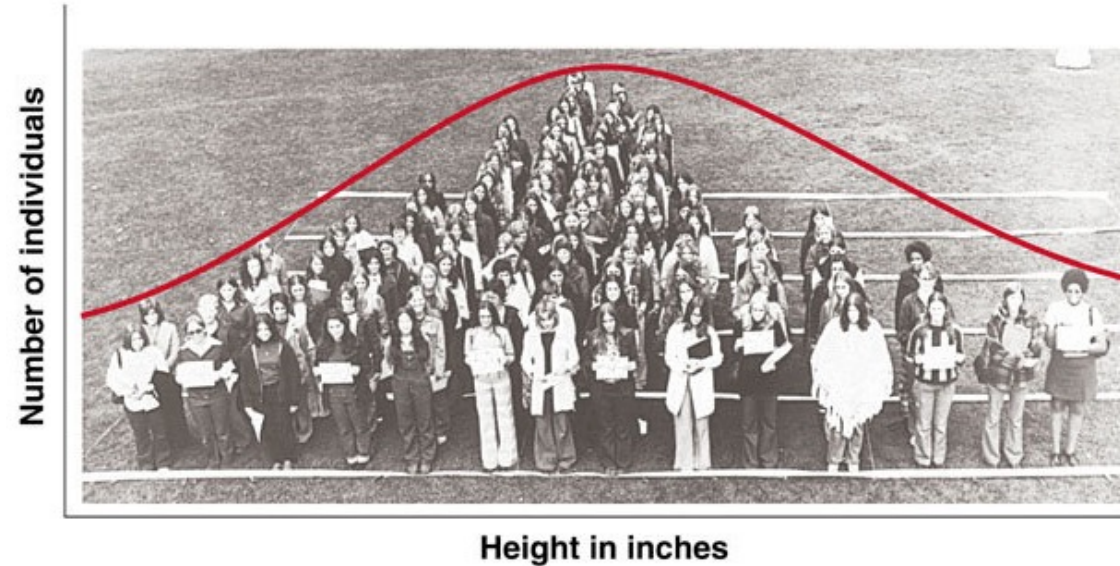
Poisson distribution  $\approx P_n(c) = \frac{c^n e^{-c}}{n!}$   $(p \ll 1 \text{ and } N \gg 1)$

Gaussian distribution  $\approx G(n) = \frac{1}{\sqrt{2\pi c}} \exp \left[ -\frac{1}{2c} (n - c)^2 \right]$   $(c \gg 1)$

where  $c = Np$ .



The **Gaussian distribution** appears everywhere in the real-world. The classical textbook example of the distribution of human height of female/male adults:



variations

# When will we see people of negative height?

The heights of human beings follow the normal distribution. Or do they? If they do, say **Patrik Perlman** and colleagues, why don't we see human beings with negative heights?

# The Central Limit Theorem

The CLT may help explain how the height distribution approaches a Gaussian distribution, if we assume the height is the sum of many random variables.

History of CLT: Moivre (1733), Laplace (1812), Lyapunov (1901), ....

Recall the example of how many cars pass on our street a given time interval. Let us measure the (random) number of cars  $N$  times per day, during  $M$  days. We get a set of sequences of random numbers:

first day	$a_1^{(1)}, a_2^{(1)}, a_3^{(1)}, a_4^{(1)}, \dots, a_N^{(1)}$
second day	$a_1^{(2)}, a_2^{(2)}, a_3^{(2)}, a_4^{(2)}, \dots, a_N^{(2)}$
.....	
$M$ -th day	$a_1^{(M)}, a_2^{(M)}, a_3^{(M)}, a_4^{(M)}, \dots, a_N^{(M)}$

The total number of measurements is  $N \times M$ .

At the day with the index  $\alpha$ , the average value is

$$X_{\alpha} = \frac{1}{N} \sum_{i=1}^N a_i^{(\alpha)}$$

This defines a sequence of random numbers  $X_{\alpha}: X_1, X_2, X_3, \dots, X_M$ .

It represents daily averaged values of the random numbers  $a_i^{(\alpha)}$ .

We are interesting in how strong are the fluctuation of  $X_{\alpha}$ .

For this purpose let us first find the mean value and variance of the random numbers  $a_i^{(\alpha)}$ :

$$\langle a \rangle = \frac{1}{MN} \sum_{i=1}^N \sum_{\alpha=1}^M a_i^{(\alpha)}$$

$$\sigma^2 = \langle (a - \langle a \rangle)^2 \rangle = \frac{1}{MN} \sum_{i=1}^N \sum_{\alpha=1}^M (a_i^{(\alpha)} - \langle a \rangle)^2$$

Let us compare the mean value and variance of the random number  $a$  with **the mean value and variance** of the random numbers  $X_\alpha$ :

$$\langle X \rangle = \frac{1}{M} \sum_{\alpha=1}^M X_\alpha = \frac{1}{MN} \sum_{i=1}^N \sum_{\alpha=1}^M a_i^{(\alpha)} = \langle a \rangle$$

The variance is

$$\begin{aligned} \Lambda^2 &= \langle (\delta X)^2 \rangle = \frac{1}{M} \sum_{\alpha=1}^M (X_\alpha - \langle X \rangle)^2 = \frac{1}{M} \sum_{\alpha=1}^M \left( \frac{1}{N} \sum_{i=1}^N \delta a_i^{(\alpha)} \right)^2 \\ &= \frac{1}{MN^2} \sum_{\alpha=1}^M \sum_{i=1}^N \sum_{j=1}^N \delta a_i^{(\alpha)} \delta a_j^{(\alpha)} = \frac{1}{MN^2} \sum_{\alpha=1}^M \sum_{i=1}^N \left( \delta a_i^{(\alpha)} \right)^2 = \frac{\sigma^2}{N} \end{aligned}$$

Here we used the fact that fluctuations are uncorrelated:

$$\frac{1}{M} \sum_{\alpha=1}^M \delta a_i^{(\alpha)} \delta a_j^{(\alpha)} = 0 \text{ for } i \neq j \text{ (i.e. variables } a_i \text{ and } a_j \text{ are independent).}$$

Most remarkably, the **Central Limit Theorem** states that the distribution of a random variable  $X$  defined as the sum of  $N$  **independent** variables  $a_i$ ,

$$X = \frac{1}{N} \sum_i a_i,$$

tends to a **Gaussian distribution** with **mean** equal to the sum of the  $\langle a_i \rangle$ ,

$$\langle X \rangle = \frac{1}{N} \sum_i \langle a_i \rangle,$$

and **variance**,

$$\begin{aligned} \Lambda^2 &= \langle (\delta X)^2 \rangle = \left\langle \left( \frac{1}{N} \sum_i \delta a_i \right)^2 \right\rangle = \frac{1}{N^2} \left\langle \left( \sum_i \delta a_i \right) \left( \sum_j \delta a_j \right) \right\rangle \\ &= \frac{1}{N^2} \sum_i \sum_j \langle \delta a_i \delta a_j \rangle = \frac{1}{N^2} \sum_i \langle \delta a_i^2 \rangle = \frac{1}{N^2} \sum_i \sigma_i^2, \end{aligned}$$

when the **number of variables  $N \rightarrow \infty$** .

(The covariances  $\langle \delta a_i \delta a_j \rangle = 0$  for  $i \neq j$  because variables are independent.)

The **Central Limit Theorem** explains why the Normal distribution is so ubiquitous in science and nature:

- The CTL holds for arbitrary distributions of the  $a_i$ 's, including when these  $a_i$ 's have different distributions among them.
- Therefore, *any* variable which is determined by the sum of many random variables, whatever their individual distributions are, will generally display a Bell-curve distribution.

In the case that all the  $a_i$ 's have the same distribution with mean  $\langle a \rangle$  and variance  $\sigma^2$ , we have:

$$\langle X \rangle = \langle a \rangle \text{ and } \Lambda^2 = \frac{\sigma^2}{N}.$$

## Implications of CTL

**Temperature and pressure fluctuations**, which are caused by fluctuations of the number and energy of the molecules that hit our body, are very weak. It makes our life comfortable.

The averaged kinetic energy of molecules at time  $t$  is

$$E(t) = \frac{1}{N} \sum_{i=1}^N \frac{mv_i^2(t)}{2}$$

where  $N \gg 1$ . Averaging over time  $t$ , we get the mean value  $\langle E \rangle = \frac{3}{2} k_B T$  and the variance  $\langle (E - \langle E \rangle)^2 \rangle = \frac{1}{N} \langle (\delta e)^2 \rangle \sim T^2 / N$ .

The pressure is proportional to the kinetic energy, and so are their fluctuations.

## Inverse transform sampling

If *all* pseudo-random numbers generators are uniform, **how to generate numbers from a Gaussian distribution?**

To sum a *large number* of uniformly distributed numbers is very unpractical!  
And it will not work for other probability density distributions.

Fortunately, it is possible to convert a uniformly distributed random number to another random number that follows the different **desired distribution**.

Let us write the probability that a number is smaller or equal to  $x$ , call it the **cumulative distribution function**

$$P_{cum}(x) = \int_{-\infty}^x P(x') dx' .$$

*To switch distributions we simply have to match their cumulatives.*



## Inverse transform sampling

**Start with a uniform random number  $x$** , distributed according to the function  $\mathcal{U}(x) = 1$  for  $0 \leq x < 1$ , and 0 elsewhere.

To generate a random number  $y$  from **any desired distribution  $f(y)$** , we take the cumulative  $F(y) = \int_{-\infty}^y f(y') dy'$  and match it with the cumulative distribution of  $x$ , which is simply given by

$$\int_0^x \mathcal{U}(x') dx' = x \text{ (for } 0 \leq x < 1 \text{):}$$

$$F(y) = \int_0^x \mathcal{U}(x') dx' = x$$

➡  $y = F^{-1}(x)$

*Just plug  $x$  into  $F^{-1}(x)$  to get  $y$ .*

