A.V. Goltsev (1)

# Sistemas Complexos e Desordenados

## Lecture 2.

## Discrete probability distributions

Gambling:
Binomial and Poisson distributions

Gambling with a fruit machine
There are three windos

[ ][ ][ ] ; [7][7][7] - prize combination

M numbers (or fruits or symbols) appears at random in the windows.
You have N coins. What is the probability to win n times?
The probability to have combination 777 is

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

Probability to have another combination is

$$1-p$$

The probability to win $n$ times using $N$ attempts is

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

Here $p^n$ is the probability to win $n$ times, while other $N-n$ attempts are unsuccessful

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

is binomial coefficient. It gives us the number of ways to win $n$ times and lose $N-n$ times. For example, first $n$ attempts can be successful while other $N-n$ attempt will be unsuccessful.

Normalization:

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

$$= (p+1-p)^N = 1$$

In avarage, you can win

$$C = Np$$

times. Indead,

$$C \equiv \langle n \rangle \equiv \sum_{n=0}^{N} B(n,N)\, n =$$

$$= \sum_{n=0}^{N} n \, \frac{N! \, p^n (1-p)^{N-n}}{n!\,(N-n)!}$$

We replace $n = n'+1$, i.e. $n' = 0, \dots N-1$

Then

$$C = \sum_{n'=0}^{N-1} \frac{N! \, p^{n'+1} (1-p)^{N-1-n'}}{n'!\,(N-1-n')!} =$$

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

$$= pN$$

Good news is that

Probability to lose $N$ times decreases exponentialy

$$B(0,N) = (1-p)^N = e^{-N|\ln(1-p)|}$$

$$\simeq e^{-Np}, \text{ if } p \ll 1$$

Unfortunately, probability to win $N$ times is small

$$B(N,N) = p^N = \exp(-N|\ln p|)$$

Unfortunately, it is very difficult to work with the function $B(n, N)$.

Let us find approximate distribution in the case

$$p \ll 1, \quad N \gg 1$$

$$N - n \gg 1$$

we use ~~the~~ Stirling formula (De Moivre, 1733)

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

This approximation is amazingly good

| | Ratio $\dfrac{\sqrt{2\pi n}\, n!\, e^{-n}}{n!}$ |
|---|---|
| $n = 1$ | 0.92 |
| $n = 2$ | 0.96 |
| $n = 3$ | 0.97 |
| $n = 10$ | 0.99 |

We get

$$C_n^N = \frac{N!}{(N-n)!\,n!} = \frac{1}{n!}\,\frac{\sqrt{2\pi N}\,N^N e^{-N}}{\sqrt{2\pi(N-n)}\,(N-n)^{N-n}\,e^{-N+n}}$$

$$= \frac{1}{n!}\,\frac{\cancel{\sqrt{N}}\,\cancel{N^N}\,\cancel{e^{-N}}}{\cancel{\sqrt{N}}\left(1-\frac{n}{N}\right)^{1/2} N^{\cancel{N}-n}\left(1-\frac{n}{N}\right)^{N-n} e^{-\cancel{N}+n}}$$

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

The probability

$$B(n,N) \approx \frac{e^{-n}}{n!}\,\frac{N^n p^n (1-p)^{N-n}}{\left(1-\frac{n}{N}\right)^{N-n+1/2}}$$

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

In the leading order, at $p \ll 1$ and $\frac{n}{N} \ll 1$, we get

$$\exp[\;] \approx \exp\Big[-(N-n)p + \left(N-n+\tfrac{1}{2}\right)\frac{n}{N}\Big]$$

$$= \exp\Big[-Np + np + n - \frac{\left(n-\frac{1}{2}\right)n}{N}\Big]$$

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

Substituting this result into $B(n,N)$, we obtain

$$B(n,N) \cong \frac{(Np)^n \cancel{e^{-n}}}{n!} \cdot e^{-Np \cancel{+n}}$$

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

<u>Poisson distribution</u>

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

where

$$c = Np$$

Properties of $P_n(c)$

1) Normalization

$$\sum_{n=0}^{\infty} P_n(c) = \sum_{n=0}^{\infty} \frac{c^n}{n!} e^{-c}$$

$$= e^{-c} \sum_{n=0}^{\infty} \frac{c^n}{n!} = e^{-c} e^{c} = 1$$

2) Mean value

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

$$= \sum_{n=1}^{\infty} \frac{c^n e^{-c}}{(n-1)!} = \sum_{n'=0}^{\infty} \frac{c^{n'+1} e^{-c}}{n'!} = c$$

$$= c \sum_{n'=0}^{\infty} \frac{c^{n'} e^{-c}}{n'!} = c \sum_{n'=0}^{\infty} P_{n'}(c) = c$$

Thus

$$\langle n \rangle = c$$

3) Variance

$$\langle (n - \langle n \rangle)^2 \rangle = \sum_{n=0}^{\infty} (n-c)^2 P_n(c)$$

$$= \sum_{n=0}^{\infty} (n^2 - 2nc + c^2) P_n(c) =$$

$$= \sum_{n=0}^{\infty} [n(n-1) + n - 2nc + c^2] P_n(c)$$

$$= \sum_{n=0}^{\infty} n(n-1) P_n(c) + \underbrace{\sum_{n=0}^{\infty} n P_n(c)}_{= \langle n \rangle = c}$$

$$- 2c \underbrace{\sum_{n=0}^{\infty} n P_n(c)}_{= \langle n \rangle = c} + c^2 \underbrace{\sum_{n=0}^{\infty} P_n(c)}_{=1}$$

$$= \sum_{n=0}^{\infty} n(n-1) P_n(c) + c - c^2$$

We have

$$\sum_{n=0}^{\infty} n(n-1) P_n(c) = \sum n(n-1) \frac{c^n e^{-c}}{n!} =$$

$$= \sum_{n=2}^{\infty} \frac{c^n e^{-c}}{(n-2)!} = \sum_{n'=0}^{\infty} \frac{c^{n'+2} e^{-c}}{n'!} = c^2 \underbrace{\sum_{n'=0}^{\infty} P_{n'}(c)}_{=1}$$

$$V = c^2$$

We arrive at

$$\langle (n-\langle n\rangle)^2\rangle = c^2 + c - c^2 = c$$

$\Rightarrow$ That is the variance is $c$!

$$\boxed{\sum_{n=0}^{\infty}(n-\langle n\rangle)^2 P_n(c) = c}$$

Relation for discrete Poisson distribution

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

We can use this relation to check that a random number $a$ is generated according to Poisson distribution. We should ~~calculate~~ check

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

Dependence of $P_n(c)$ on $n$.

We use Stirling formular

$$P_n(c) = \frac{c^n e^{-c}}{n!} \approx \frac{c^n e^{-c}}{\sqrt{2\pi n}\, n^n e^{-n}}$$

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

We introduce a function

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

This function has a maximum at a point

$$\frac{d f(n)}{dn} = 0 \Rightarrow 1 + \ln c - \ln n - \frac{n+1/2}{n} = 0$$

$$\Rightarrow \cancel{f(n)+} \ln c - \ln n - \frac{1}{2n} = 0$$

In the case $c \gg 1$, we obtain a maximum at

$$n_0 = c$$

Teylor expansion of $f(n)$ near $n_0$

$$f(n) = f(n_0 + (n-n_0)) =$$

$$= f(n_0) + \underbrace{f'(n_0)}_{=0}(n-n_0) + \frac{1}{2} f''(n_0)(n-n_0)^2 + \ldots$$

$$= f(n_0) + \frac{1}{2} f''(n_0)(n-n_0)^2$$

$$f(n_0) = n_0 + n_0 \ln c - \left(n_0 + \frac{1}{2}\right) \ln n_0$$
$$= c + c\cancel{\ln c} - \cancel{c \ln c} - \frac{1}{2} \ln c$$

$$f''(n_0) = \frac{d}{dn}\left( \ln c - \ln n - \frac{1}{2n}\right) =$$

$$= -\frac{1}{n_0} + \frac{1}{2n_0^2} \approx -\frac{1}{c}$$

Therefore

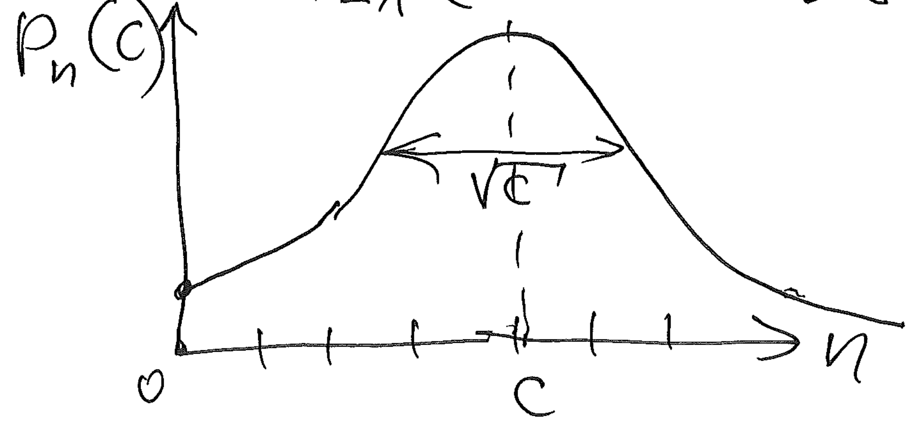
$$f(n) \cong c - \frac{(n-c)^2}{2c}$$

We obtain

$$P_n(c) = \frac{c^n e^{-c}}{n!} \approx \frac{e^{-c} e^{f(n)}}{\sqrt{2\pi}} =$$

$$= \frac{e^{\cancel{-c}\cancel{+c} - \frac{1}{2c}(n-c)^2}}{\sqrt{2\pi c}}$$

$$P_n(c) \approx \frac{1}{\sqrt{2\pi c}} e^{-\frac{(n-c)^2}{2c}}$$ – Gaussian distribution

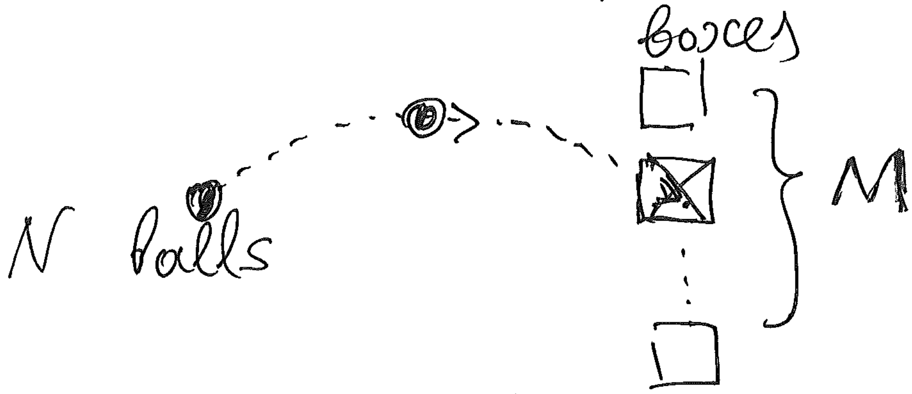


## Ball in boxes

Another simple example is throwing ball in boxes.

We have $M$ boxes and $N$ balls; we assume that the probability that a ball will be in a given box is

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

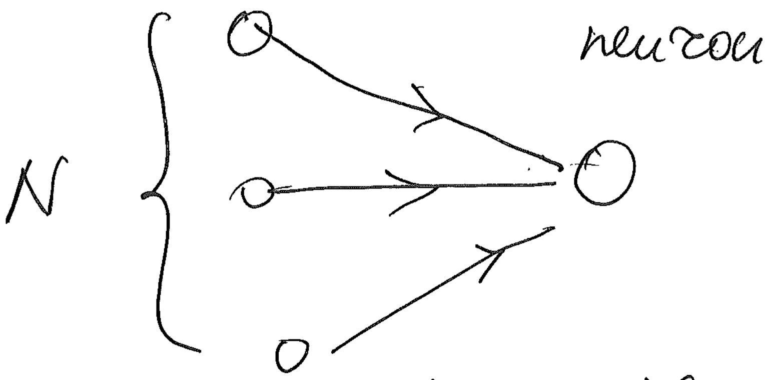



The probability to find $n$ balls in a given box is again given by the binomial distribution

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

$$\approx P_n(Np)$$

# Neurons in the brain




An neuron has $N$ neighborin neurons
Probability to receive a spike from a neighbor during time $t$ is

$$p = t\nu$$

where is $\nu$ is the mean frequence of spikes.

Probability to receive $n$ spikes from $N$ neighbor during time $t$ is

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

$$\approx P_n(Np) = P_n(Nt\nu)$$

Mean number of spikes

$$\langle n \rangle = \sum_n P_n(Nt\nu) = Nt\nu$$

Variance

$$\langle (n - \langle n \rangle)^2 \rangle = \langle n \rangle = Nt\nu$$

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