Probability distributions

Part 3

Normal distribution

The *Normal Distribution*, also known as *Gaussian Distribution*, plays a very important role in natural and social sciences.

Modelling a certain phenomenon with a Normal distribution means assuming this hypothesis:

There is a "true" or "normal" value for this phenomenon, driven by the main factor, the "cause".

But real events are affected by others factors, each one introducing a noise, which makes the observed value different from the "normal" expected value.

These noise sources are assumed having these features:

- There are a lot of noise factors.
- No one dominates the others.
- Each one varies random around 0 with no precise tendency.

The "normal" value is commonly denoted with the greek letter μ.

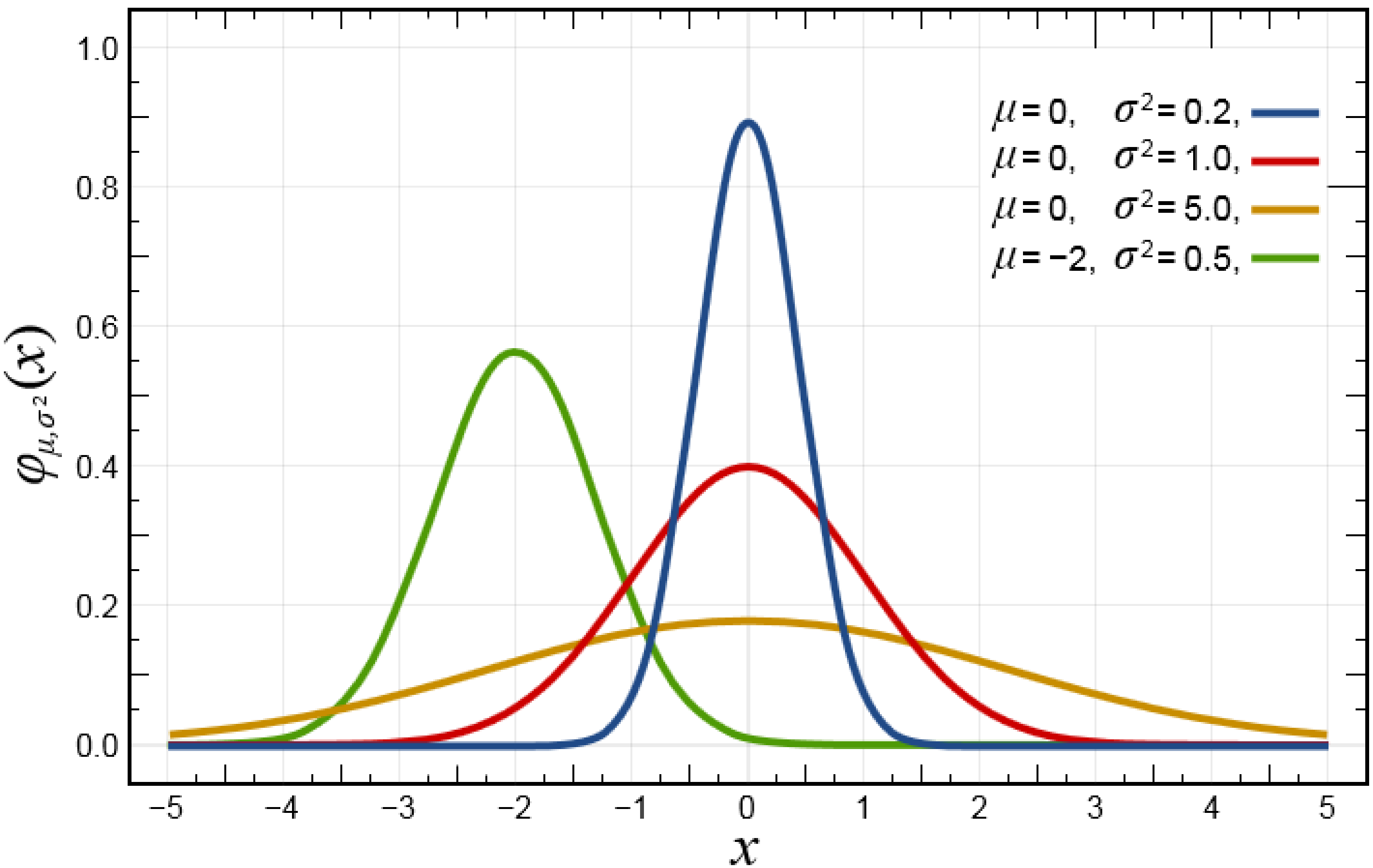
It is the *expected value*, the rigorous formulation of the concept of *mean*.

The expected value μ is the "most likely to be observed" value.

The closer a value is to μ, the more likely to be observed it is.

The *variance* of the distribution is denoted with σ^2.

Mean and variance are the two parameters which completely describe a normal distribution.



$\mu = 0, \quad \sigma^2 = 0.2,$
$\mu = 0, \quad \sigma^2 = 1.0,$
$\mu = 0, \quad \sigma^2 = 5.0,$
$\mu = -2, \quad \sigma^2 = 0.5,$
$\varphi_{\mu,\sigma^2}(x)$
x
1.0
0.8
0.6
0.4
0.2
0.0
−5
−4
−3
−2
−1
0
1
2
3
4
5

Given μ and σ^2, we can know how much likely it is to observe a realization of the phenomenon modelled with the distribution falling inside a certain range.

For example, we know that in the 95.44% of cases, the observation will be in the range from $\mu - 2\sigma$ to $\mu + 2\sigma$.

Note we are using σ, the *standard deviation*, which is the square root of the variance.

Normal Distribution