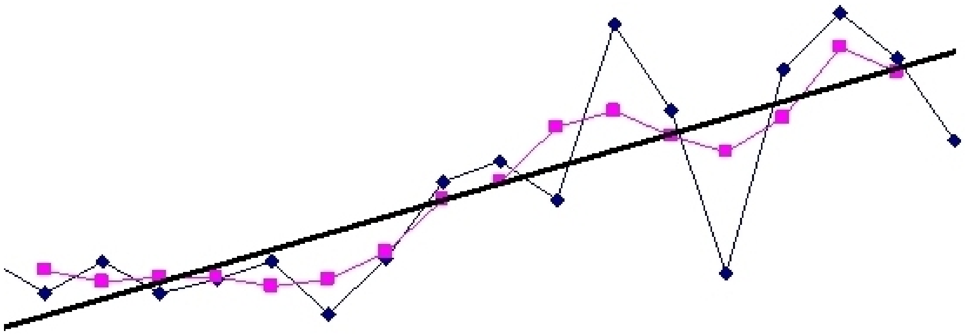
Mathematical and Statistical Modelling in Medicine

Author: **Tibor Nyári** PhD

University of Szeged
Department of Medical Physics and Informatics

www.model.u-szeged.hu
www.szote.u-szeged.hu/dmi

Probability distributions



Random experiment

- The outcome is not determined uniquely by the considered conditions.
- For example, tossing a coin, rolling a dice, measuring the concentration of a solution, measuring the body weight of an animal, etc. are experiments.
- Every experiment has more, sometimes infinitely large outcomes

Event: **the result (or outcome) of an experiment**

- elementary events: the possible outcomes of an experiment.
- composite event: it can be divided into sub-events.
- Example. The experiment is rolling a dice.
  - Elementary events are 1,2,3,4,5,6.
  - Composite events:
    - E1={1,3,5} (the result is an odd number).
    - E2={2,4,6} (the result is an even number).
    - E3={5,6} (the result is greater than 4).
    - Ω={1,2,3,4,5,6} (the result is the certain event).

Example: the tossing of a (fair) coin

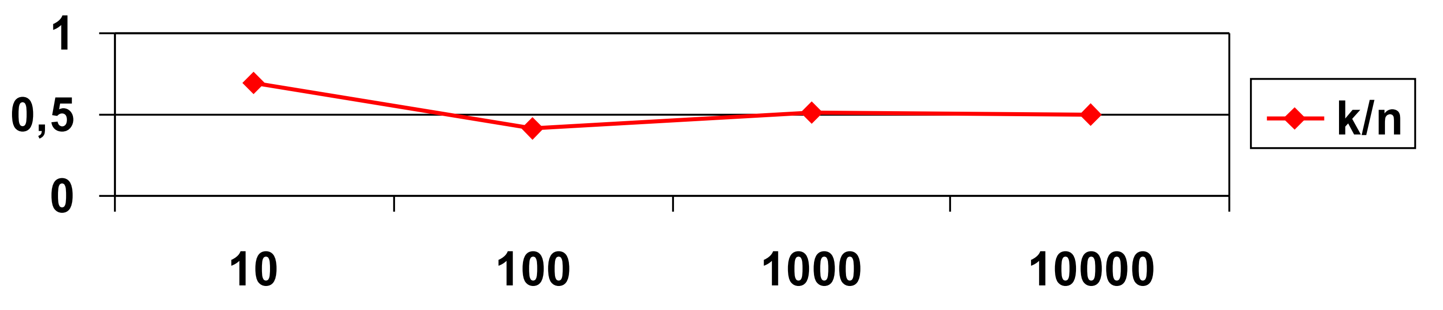
k: number of „head"s

n= 10 100 1000 10000 100000

k= 7 42 510 5005 49998

k/n= 0.7 0.42 0.51 0.5005 0.49998

P(„head")=0.5



The concept of probability

- Lets repeat an experiment n times under the same conditions. In a large number of n experiments the event A is observed to occur k times ($0 \leq k \leq n$).
- ***k*** : **frequency** of the occurrence of the event A.
- ***k/n*** : **relative frequency** of the occurrence of the event A.

$$0 \leq k/n \leq 1$$

If n is large, k/n will approximate a given number. This number is called the probability of the occurrence of the event A and it is denoted by P(A).

$$0 \leq P(A) \leq 1$$

Probability facts

- Any probability is a number between 0 and 1.
- All possible outcomes together must have probability 1.
- The probability of the complementary event of A is 1-P(A).

Rules of probability calculus

- Assumption: all elementary events are equally probable

$$P(A) = \frac{F}{T} = \frac{\text{number of favorite outcomes}}{\text{total number of outcomes}}$$

Examples:

- Rolling a dice. What is the probability that the dice shows 5?
  - If we let X represent the value of the outcome, then P(X=5)=1/6.
- What is the probability that the dice shows an odd number?
  - P(odd)=1/2. Here F=3, T=6, so F/T=3/6=1/2.

Discrete (categorical) random variable

- A discrete random variable *X* has finite number of possible values

- **The probability distribution of *X* lists the values and their probabilities:**

Value of X:	x_1 x_2 x_3 … x_n
Probability:	p_1 p_2 p_3 … p_n

$$p_i \geq 0, \;\; p_1 + p_2 + p_3 \ldots + p_n = 1$$

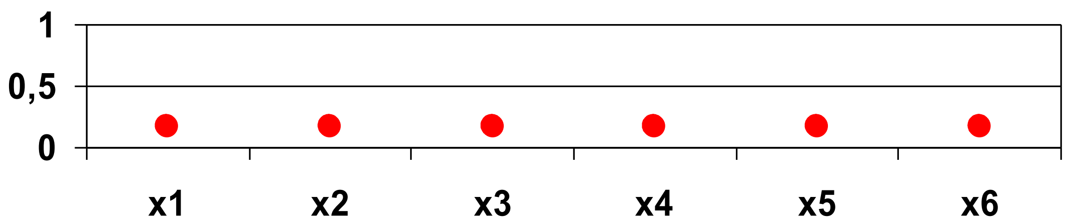
Examples

- The experiment is tossing a coin.

  $p_1 = 0.5,\ p_2 = 0.5$

- The experiment is rolling a dice.

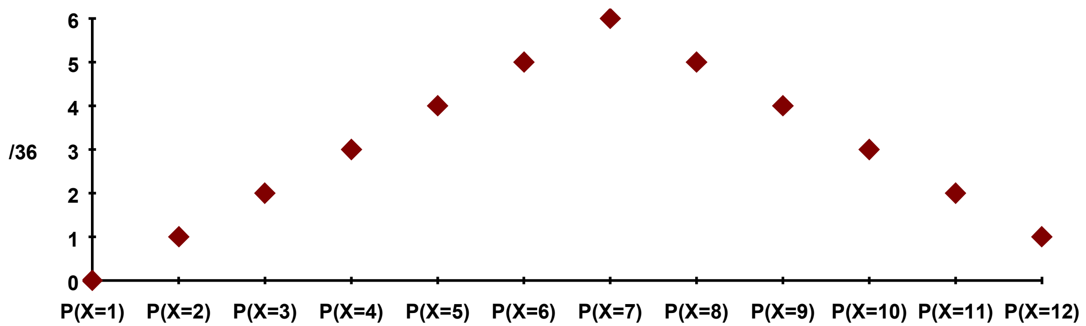
  $p_1 = 1/6,\ p_2 = 1/6, \ldots,\ p_6 = 1/6$



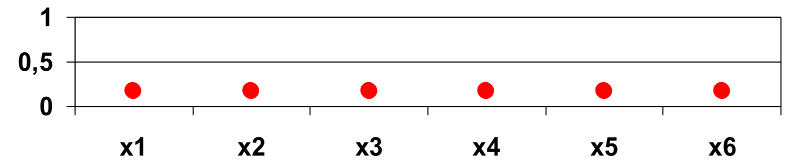
Example: rolling two dices

- Let random variable X be the sum of the two numbers shown on the two dices.
- P(X=1)=0, (X=1 is impossible)
- P(X=2)=1/36 (the only favourable event is (1,1), and the number of all possible event is 36.)
-

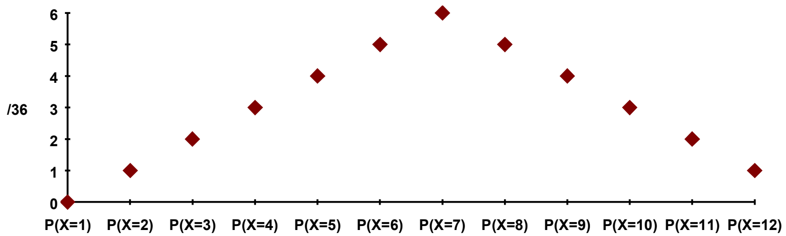
	j=1	j=2	j=3	j=4	j=5	j=6
i=1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
X	**2**	**3**	**4**	**5**	**6**	**7**
i=2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
X	**3**	**4**	**5**	**6**	**7**	**8**
i=3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
X	**4**	**5**	**6**	**7**	**8**	**9**
i=4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
X	**5**	**6**	**7**	**8**	**9**	**10**
i=5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
X	**6**	**7**	**8**	**9**	**10**	**11**
i=6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)
X	**7**	**8**	**9**	**10**	**11**	**12**



Uniform discrete distributions: all p_i-s are equal



Not uniform

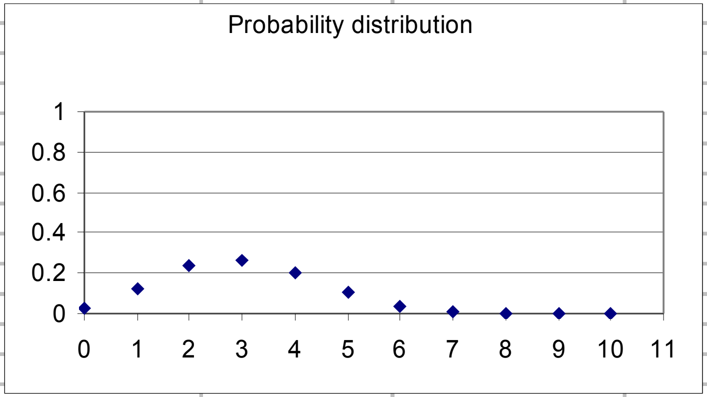


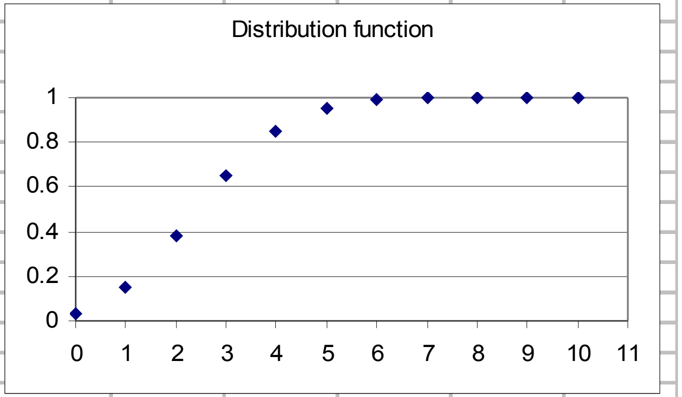
The binomial distribution

- Let's consider an experiment ***A*** that may have only two possible mutually exclusive outcomes (success, failure)
- Let P(***A***)=*p*
- We now repeat the experiment *n* times, let X denote the absolute frequency of the event ***A***.
- The probability that ***X*** will assume any given possible value *k* is expressed by the binomial formula

$$P_k = P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0,1,\ldots,n$$

Number of successful cases	Probability distribution	Distribution function	Probability of "success"
0	0.028247525	0.028247525	0.3
1	0.121060821	0.149308346	
2	0.233474441	0.382782786	
3	0.266827932	0.649610718	
4	0.200120949	0.849731667	
5	0.102919345	0.952651013	
6	0.036756909	0.989407922	
7	0.009001692	0.998409614	
8	0.001446701	0.999856314	
9	0.000137781	0.999994095	
10	5.9049E-06	1	
Összesen	1		





Binomial distribution n=10, input p, k=0,1,…,10

Poisson eloszlás – „ritka” események száma

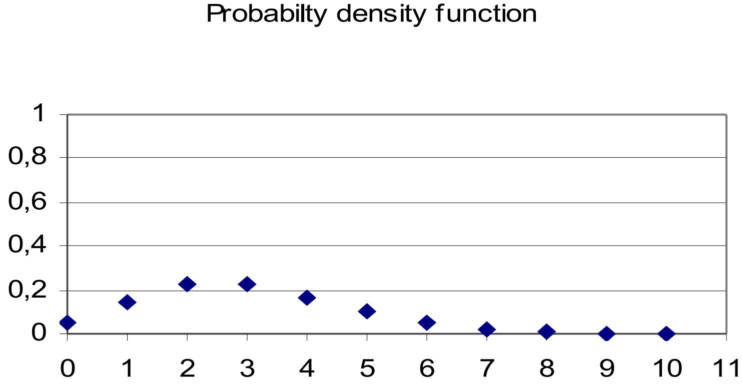
- Given a Poisson process, the probability of obtaining exactly   successes in   trials is given by the limit of a binomial distribution
- If n is huge and np=λ constant.

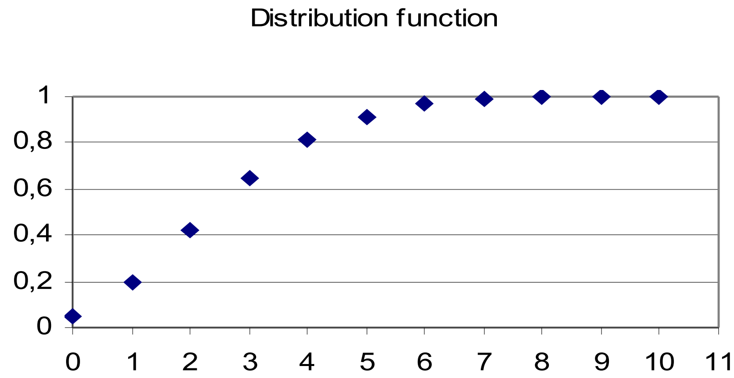
$$P(X = k) = f(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

- Viewing the distribution as a function of the expected number of successes
- λ is both the expected value and variance.

- Eg:. Number of new cases diagnosed with a certain disease is 3 (in average). If the number of new cases follow Poisson distribution what is the probability
  - There is no new case in the next month (0.0498)
  - There are 2 new cases in the next month (0.224)

No. of new cases	Probability density function	Distribution function	Expected value
0	0,049787068	0,049787068	3
1	0,149361205	0,199148273	
2	0,224041808	0,423190081	
3	0,224041808	0,647231889	
4	0,168031356	0,815263245	
5	0,100818813	0,916082058	
6	0,050409407	0,966491465	
7	0,021604031	0,988095496	
8	0,008101512	0,996197008	
9	0,002700504	0,998897512	
10	0,000810151	0,999707663	
Total	0,999707663		



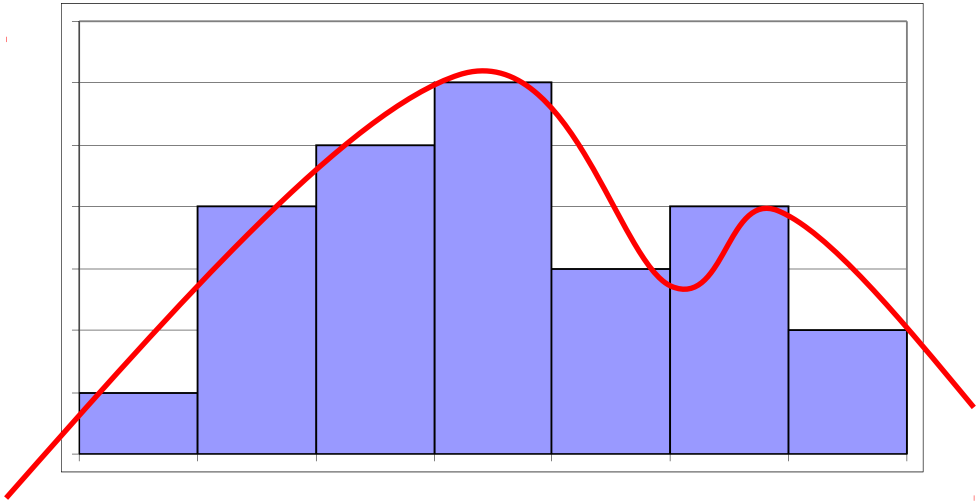


Continuous random variable

- A continuous random variable *X* has takes all values in an interval of numbers.
- **The probability distribution of *X*** is described by a **density curve.**
- The density curve
  - is on the above the horizontal axis, and
  - has area exactly 1 underneath it.
- The probability of any event is the area under the density curve and above the values of X that make up the event.

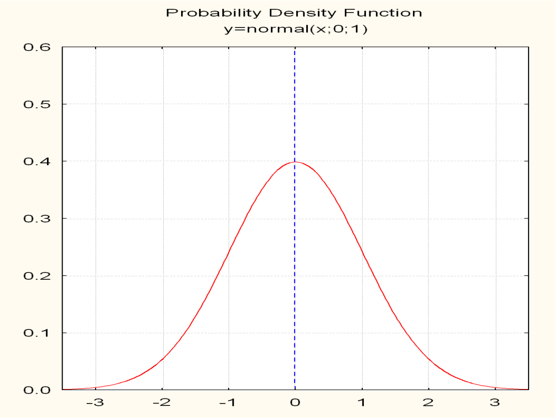
The density curve

- The density curve is an idealized description of the overall pattern of a distribution that smooths out the irregularities in the actual data.
- The density curve
  - is on the above the horizontal axis, and
  - has area exactly 1 underneath it.
- The area under the curve and above any range of values is the **proportion** of all observations that fall in that range.

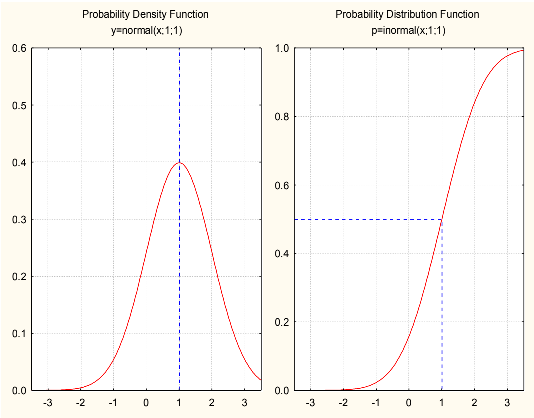


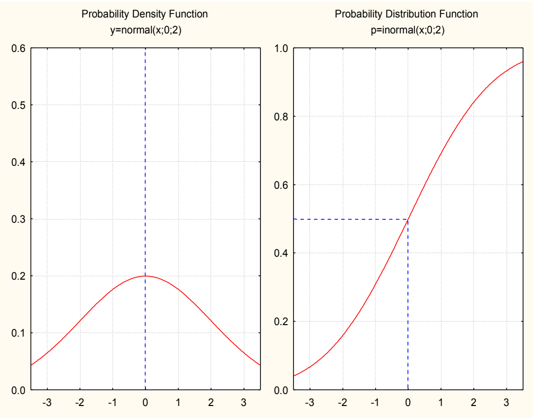
Normal distributions N(μ, σ)

N(0,1)



N(1,1)



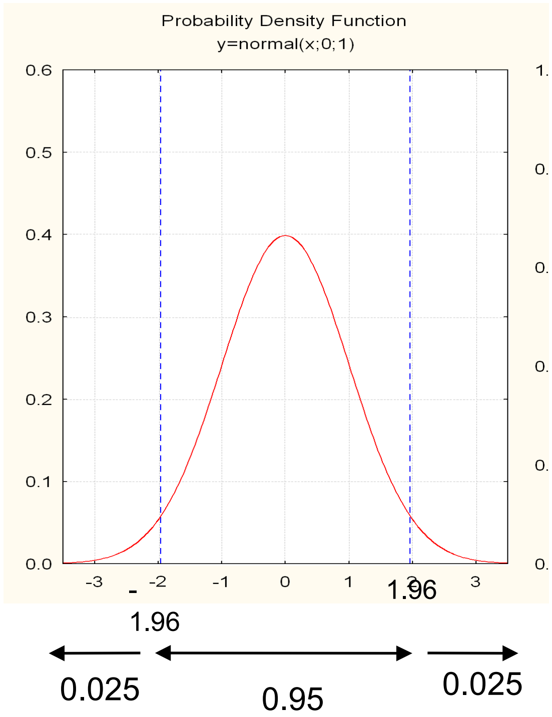


N(0,2)

μ, σ : parameters
(a **parameter** is a number that describes the distribution)

Standard normal probabilities

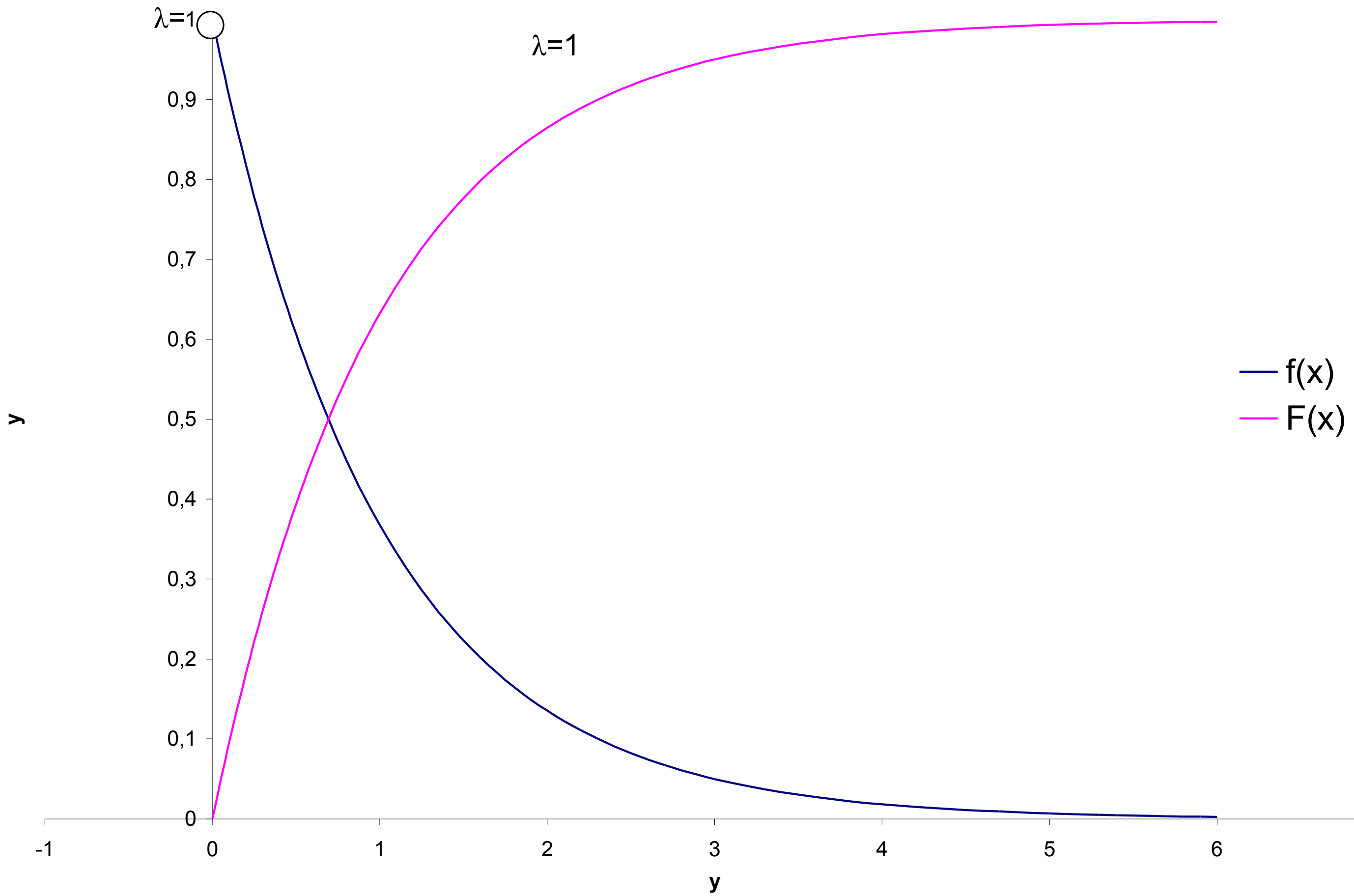
x	(x): proportion of area to the left of x
-4	0.0003
-3	0.0013
-2.58	0.0049
-2.33	0.0099
-2	0.0228
-1.96	0.0250
-1.65	0.0495
-1	0.1587
0	0.5
1	0.8413
1.65	0.9505
1.96	0.975
2	0.9772
2.33	0.9901
2.58	0.9951
3	0.9987
4	0.99997



Other continous distributions

- Exponential distribution
- Density function $f(x)=\lambda e^{-\lambda x}$, if $x > 0$, otherwise 0.
- λ is a constant parameter
- Distribution function $F(x)=1- e^{-\lambda x}$, if $x > 0$.

The exponential density and distribution function



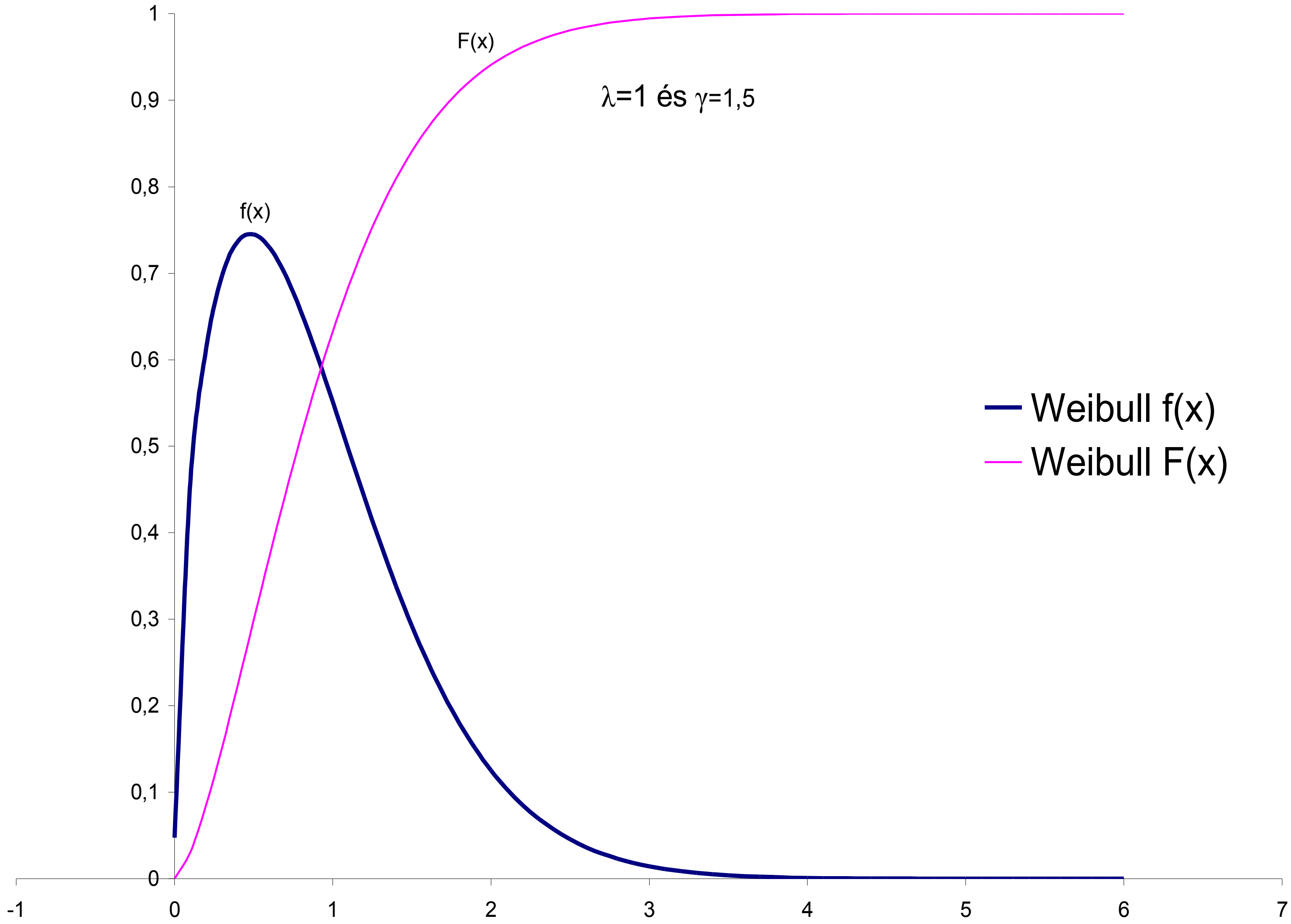
Weibull-distribution

- Generalize the exponential (γ =1) distribution results in **Weibull–distribution**, density function ($\lambda > 0$ and $\gamma > 0$ are constants)

$$f(x) = \begin{cases} \lambda \gamma (x^{\gamma - 1}) e^{-\lambda x^{\gamma}} & ha \quad x \geq 0 \\ 0 & ha \quad x < 0 \end{cases}$$

Weibull distribution function

$$F(x) = \begin{cases} 1 - e^{-\lambda x^{\gamma}} & ha \quad x \geq 0 \\ 0 & ha \quad x < 0 \end{cases}$$



F(x)
λ=1 és γ=1,5
f(x)
Weibull f(x)
Weibull F(x)
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-1
0
1
2
3
4
5
6
7

The f(x) is a probability density function (PDF) as f(x)≥0 and integrate using (u=λx^{γ} ; du=$\lambda\gamma$ $x^{\gamma-1}$ dx)
]-∞; ∞[(improprius : β→∞) equals to 1

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} \lambda\,\gamma\,\mathrm{x}^{\gamma-1}\mathrm{e}^{-\lambda \mathrm{x}^{\gamma}} dx = \lim_{\beta \to \infty} \left[-e^{-\lambda x^{\gamma}} \right]_{0}^{\beta} = 0 + 1 = 1.$$

- **A Weibull–distribution in pharmacocinetics**
- Example:
  - Calculate the area under Weibull PDF(λ=0.5 & γ=2) on] –∞ és 1] intervall (Note: this equivalent using [0 ; 1] intervall since the area under curve is 0 on]–∞ – 0[intervall)

Solution

$$\int_{-\infty}^{1} f(x)dx = \int_{0}^{1} \lambda\gamma x^{\gamma-1} e^{-\lambda x^{\gamma}} dx = \left[-e^{-\lambda x^{\gamma}} \right]_{0}^{1} = -0,6065 + 1 = 0,3934.$$

The λ=0.5 and γ=2 Weibull-probability density function