

# Probability Calculus

**Anna Janicka**

lecture XV, 22.01.2019

**MARKOV CHAINS**

**SOME IMPORTANT DISTRIBUTIONS**

# Plan for Today

---

- Markov chains
- Some important distributions we haven't talked about yet
- A short summary



## Markov chains – definitions (reminder)

---

A Markov chain is **irreducible**, if for any  $i, j \in E$  there exists an  $n > 0$ , such that  $p_{ij}(n) > 0$ ; in other words, it is possible to go from any state to any state (not necessarily in one step) – each two states **communicate**.

A state  $i$  has a **period** equal to  $k$ , if  $o(i) = \text{GCD}(n : p_{ii}(n) > 0) = k$ . A state is **aperiodic** if  $o(i) = 1$ , and **periodic** if  $o(i) > 1$ .



## Markov chains: – definitions (reminder 2)

---

**Theorem:** *If a Markov chain is irreducible, all chains have the same period.*

### More definitions:

*An irreducible Markov chain is **periodic**, if all the states are periodic; it is **aperiodic**, if all the states are aperiodic.*

*A distribution (vector)  $\pi$  is a **stationary distribution** (or state) of a Markov chain of a transition matrix  $P$ , if  $\pi \cdot P = \pi$ .*



## Markov chains: Ergodic Theorem (reminder 3)

---

*Let  $(X_n)_{n \geq 0}$  be an aperiodic irreducible Markov chain over a finite set of states. Then, this Markov chain has a single stationary distribution  $\pi$ , which also satisfies the following property:  
for any  $i, j \in E$ , we have  $\lim_{n \rightarrow \infty} p_{ij}(n) = \pi_j > 0$ .*

### Consequences:

- limit distribution does not depend on initial state
- stationary state describes behavior in the far future



## Markov chains: still more definitions

---

*The mean first passage time from state  $i$  to  $j$  for an irreducible Markov chain is the expected number of steps to reach state  $j$  from  $i$  for the first time, denoted by  $m_{ij}$ .*

*The mean recurrence time for state  $i$  for an irreducible Markov chain is the expected number of steps to return to state  $i$  for the first time, denoted by  $m_i$ .*

Calculation: systems of equations:

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj} \qquad m_i = 1 + \sum_k p_{ik} m_{ki}$$

— By convention,  $m_{ii} = 0$  —



## Markov chains: more properties

---

### Theorem (extension of ergodic theorem):

*Let  $(X_n)_{n \geq 0}$  be an aperiodic irreducible Markov chain over a finite set of states. Then, we have that the stationary distribution satisfies  $\pi_j = \frac{1}{m_j}$ , where  $m_j$  is the mean recurrence time for state  $j$ .*



## Markov chains: more properties (2)

---

### Definition:

*A state  $i$  is **absorbing** if it is impossible to leave the state; in other words,  $p_{ii} = 1$  (while  $p_{ij} = 0$  for  $j \neq i$ ).*

### Typical problems:

- calculate the probability of reaching an absorbing state
  - calculate the average time until reaching an absorbing state
- 



# Some useful distributions

---

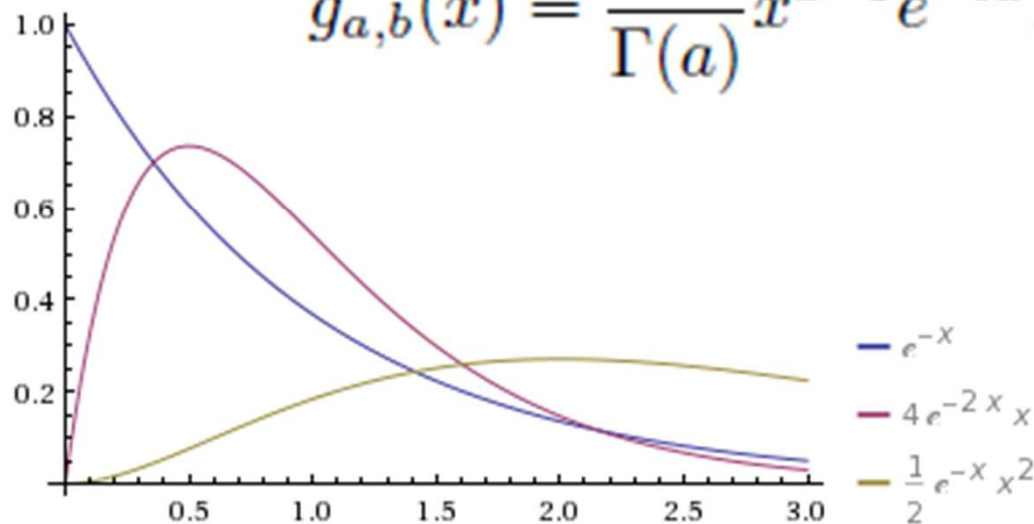
used in statistics and economics when modeling or as a result of operations performed on other distributions



# Gamma Distribution $\Gamma(a,b)$ , $a,b>0$

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt \quad \Gamma(n) = (n-1)! \text{ dla } n = 1, 2, \dots$$

$$g_{a,b}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} 1_{(0,\infty)}(x)$$



$$\mathbb{E}X = a/b$$

$$\text{Var}X = a/b^2$$



## Gamma Distribution properties

---

- for  $a=1$  – exponential distribution  $\exp(b)$
- for integer  $a$  – Erlang Distribution
- $\Gamma(n/2, 1/2)$  – **chi-squared distribution**  $\chi^2(n)$

Theorem:

A sum of independent random variables from gamma distributions  $\Gamma(a_i, b)$  has a  $\Gamma(\sum a_i, b)$  distribution

$$g_{a,b}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} 1_{(0,\infty)}(x)$$

# How the Gamma Distribution is used

---

In econometrics: duration of events

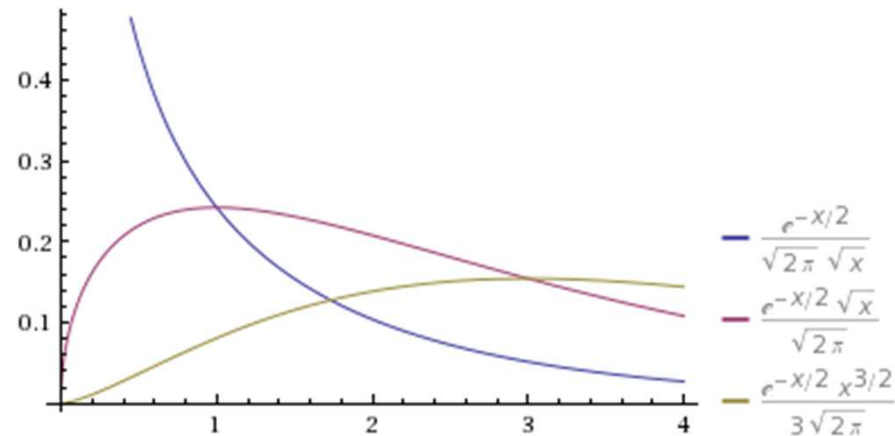
In statistics: bayesian statistics.

....the chi-squared distribution



# Chi-squared Distribution $\chi^2(n)$

---



Theorem: sum of squares of  $n$  IIN random variables (independent identically  $N(0,1)$  distributed) has a  $\chi^2(n)$  distribution

$$\mathbb{E}X = n, \quad \text{Var}X = 2n$$

for large  $n$ : like the normal distr.

# The distribution of the sample mean and variance

---

## Theorem:


*Let  $X_1, X_2, \dots, X_n$  be independent identically distributed random variables from a standard normal distribution and let  $\bar{X}, s^2$  be the sample mean and variance, respectively:*

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$s^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

*Then  $\sqrt{n}\bar{X}$  has a  $\mathcal{N}(0, 1)$ , distribution,*

*—  $ns^2$  has a  $\chi_{n-1}^2$  distribution*

 *and  $\bar{X}, s^2$  are independent.*

# Chi-squared distribution uses

---

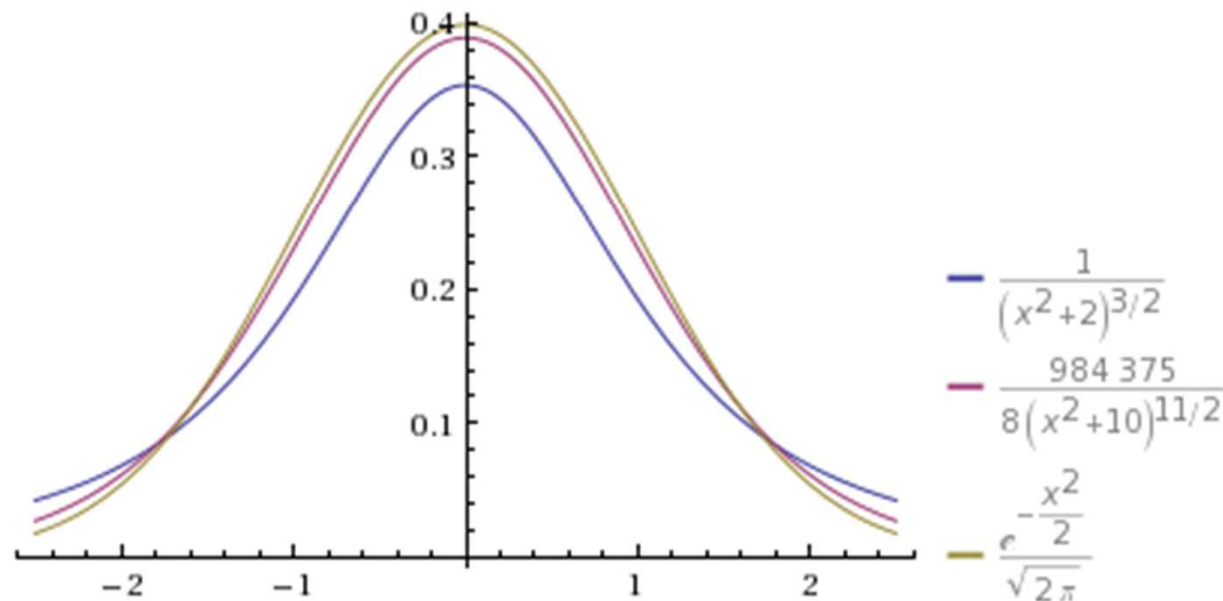
Used very frequently in statistics: appears in „standard” models when testing hypotheses, constructing confidence intervals, as a „component” of other distributions.



# ***t*-Student Distribution $t(n)$ , $n=1,2,\dots$**

$n^{1/2}X/Y^{1/2}$  for independent  $X$  and  $Y$ ,  $X \sim N(0,1)$ ,  $Y \sim \chi^2(n)$

$$g_n(x) = \frac{1}{\sqrt{n\pi}} \cdot \frac{\Gamma(\frac{1}{2}(n+1))}{\Gamma(\frac{1}{2}n)} \left(1 + \frac{x^2}{n}\right)^{-\frac{1}{2}(n+1)}, \quad n = 1, 2, \dots$$



$$\mathbb{E}X = 0, \quad n > 1$$

$$\text{Var}X = n/(n-2), \quad n > 2$$

## ***t*-Student distribution uses**

---

Statistics: regression analysis, hypothesis testing, construction of confidence intervals

Econometrics: as an alternative for the normal distribution in cases where fat tails are needed

for large  $n$ : almost  $N(0, 1)$ .



## ***F*-Snedecor $F(d_1, d_2)$ , $d_1, d_2 = 1, 2, \dots$**

---

$X$  has a  $F(d_1, d_2)$  distribution, if  $X = (Y_1/d_1)/(Y_2/d_2)$ ,  
where  $Y_i$  are independent,  $\chi^2(d_i)$

$$g_{d_1, d_2}(x) = \frac{\left(\frac{d_1 x}{d_1 x + d_2}\right)^{d_1/2} \left(1 - \frac{d_1 x}{d_1 x + d_2}\right)^{d_2/2}}{x B(d_1/2, d_2/2)} 1_{(0, \infty)}(x)$$

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad a, b > 0$$

$$\mathbb{E}X = d_2 / (d_2 - 2) \quad d_2 > 2$$

$$\text{Var}X = \frac{2 d_2^2 (d_1 + d_2 - 2)}{d_1 (d_2 - 2)^2 (d_2 - 4)} \quad d_2 > 4$$



# ***F* distribution uses**

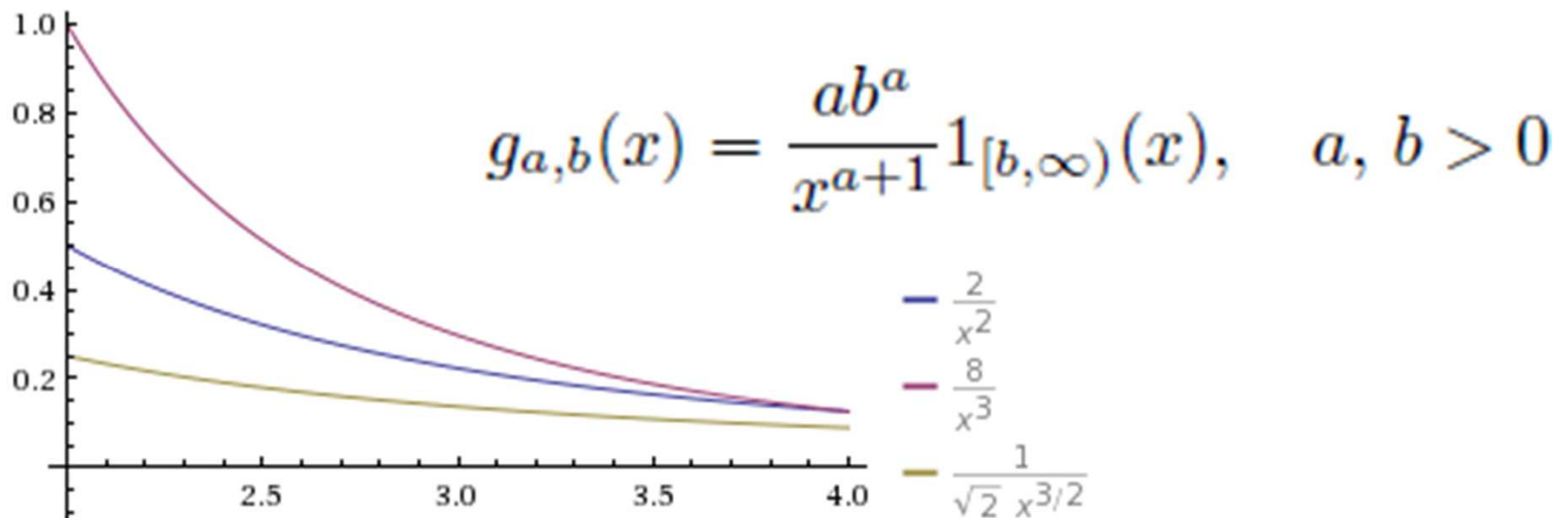
---

Statistics: testing hypotheses



# Pareto distribution

---



$$\mathbb{E}X = ab/(a - 1) \text{ dla } a > 1$$

$$\text{Var}X = ab^2 / [(a - 1)^2(a - 2)] \text{ dla } a > 2.$$



## Pareto distribution uses

---

Pareto described the distribution of wealth/income among various groups in a population

(property: larger part of wealth accumulated by a smaller part of the population

→ Pareto 80-20 rule, which corresponds to a value  $\alpha > 1$  )

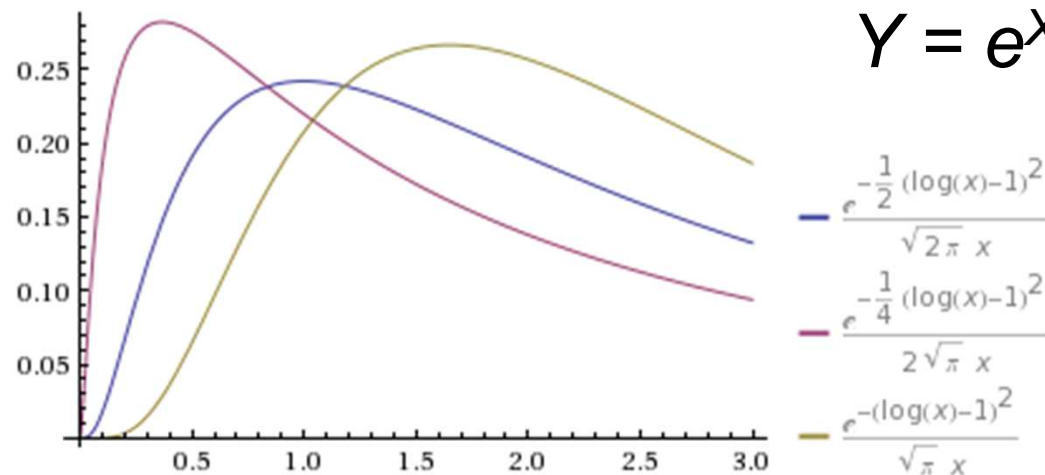
Used not only for income and wealth, but also in finance, insurance, actuarial

---



# Lognormal distribution $L(m, \sigma^2)$ , where $m \in \mathbb{R}$ , $\sigma > 0$

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln x - m)^2}{2\sigma^2}\right) 1_{(0, \infty)}(x)$$



$$Y = e^X, \text{ where } X \sim N(m, \sigma^2)$$

$$\mathbb{E}Y = \exp\left(m + \frac{1}{2}\sigma^2\right)$$

$$\text{Var}Y = (\exp(\sigma^2) - 1) \cdot \exp(2m + \sigma^2)$$

Theorem: A product of lognormal variables has a lognormal distribution



## Lognormal distribution uses

---

In most cases where we model positive random variables of multiplicative character (the CLT may be used for logarithms)

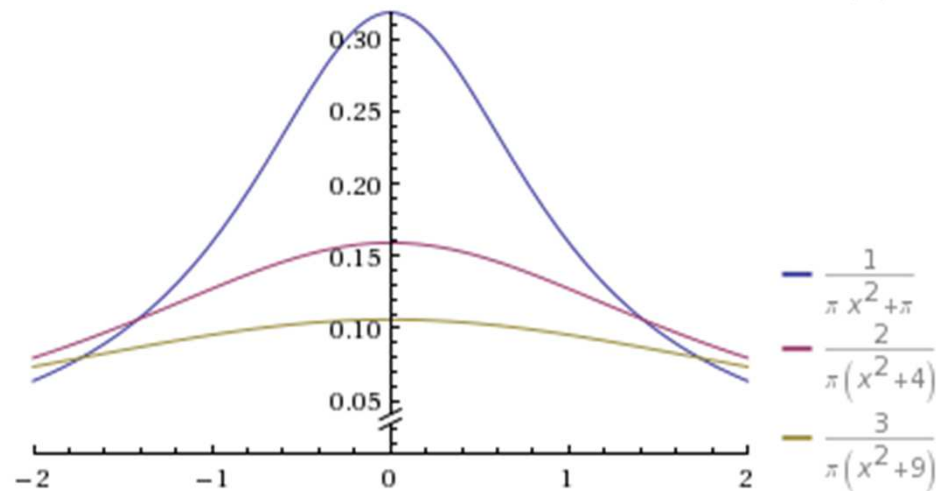
For modeling income (apart from the most rich).  
In finance (eg. Black-Scholes model)



# Cauchy distribution $\text{Cau}(a, m)$ , where $a > 0, m \in \mathfrak{R}$

---

$$g(x) = \frac{a}{\pi((x - m)^2 + a^2)}$$



Does not have an expected value nor higher moments.

The LLN nor the CLT may not be applied!

---



# Cauchy distribution properties

---

Theorem: an average of identical independent Cauchy random variables is a Cauchy variable  
→ conclusions based on the mean are worthless

A ratio of two IIN random variables has a Cauchy distribution

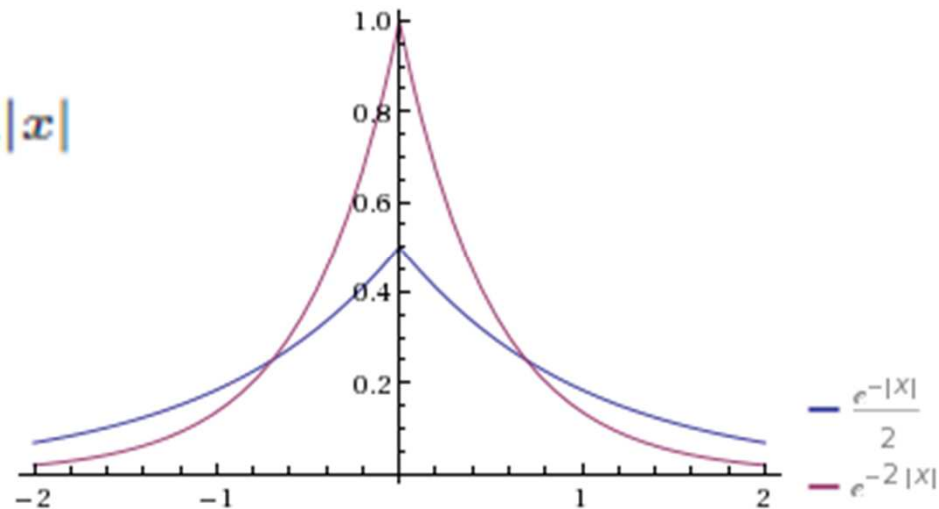
$\text{Cau}(1,0)$  is equivalent to the t-Student distribution with one degree of freedom (1).



# Two-sided exponential distribution (Laplace distribution) with parameter $\lambda > 0$

---

$$g(x) = \frac{1}{2} \lambda e^{-\lambda|x|}$$



$$\mathbb{E}X = 0, \text{Var}X = 2/\lambda^2$$



# Weibull distribution

---

A different generalization of the exp. distr.

$$g(x) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^\alpha}1_{(0,\infty)}(x), \quad \alpha, \beta > 0.$$

$$\mathbb{E}X = \beta\Gamma\left(1 + \frac{1}{\alpha}\right)$$

$$\text{Var}X = \beta^2\left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \left(\Gamma\left(1 + \frac{1}{\alpha}\right)\right)^2\right]$$

For modeling duration; depending on the value of  $\alpha$ : decreasing, constant, increasing incidence. Actuarial sciences.

---





WARSAW UNIVERSITY  
**Faculty of Economic Sciences**