NMAI059 Probability and statistics 1
Class 7

Robert Šámal

Overview

Continuous random variables

Particular continuous distributions and their parameters

General and continuous random variable – what we have learned

- R.v. is a mapping $X : \Omega \to \mathbb{R}$, that for every $x \in \mathbb{R}$ satisfies $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$.
- Discrete r.v. is a r.v.
- CDF of a r.v. X is a function $F_X(x) := P(X \leq x)$.
- CDF F_X is nondecreasing right-continuous function with limits in ± 1 equal to 0/1.
- A continuous r.v. has a PDF $f_X \geq 0$ such that $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt$.
- $P(a \leq X \leq b) = \int_a^b f_X(t)dt$ for every $a, b \in \mathbb{R}$.
- $P(X \in A) = \int_A f_X(t)dt$ for a “reasonable set A”.

Expectation of a continuous r.v.

Definition

Consider a continuous r.v. X with PDF f_X. Then its expectation (expected value, mean) is denoted by $\mathbb{E}(X)$ and defined by

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \, f_X(x) dx,$$

whenever the integral is defined; that is unless it is of type $\infty - \infty$.

- An analogy with computing a center of mass of a pole from a formula for its density.
- Discretization.

Properties of expectation

Theorem (LOTUS)

Consider a continuous r.v. X with density f_X and a real function g. Then we have

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx,$$

whenever the integral is defined.

(We skip the proof.)

Theorem (Linearity of expectation)

For $X_1, \ldots, X_n$ discrete or continuous random variables we have

$$\mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n).$$

(Proof later.)

Variance of a continuous r.v.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \; f_X(x) dx$$

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \; f_X(x) dx$$

Writing $\mu = \mathbb{E}(X)$, we have

$$var(X) := \mathbb{E}((X - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 \; f_X(x) dx.$$

Theorem

For continuous random variables we have the same formula as for discrete ones, $var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$*.*

(Proof is the same as for discrete r.v.)

Variance of a sum

Theorem (Variance of a sum)

For $X_1, \dots, X_n$ independent discrete or continuous r.v. we have

$$var(X_1 + \cdots + X_n) = var(X_1) + \cdots + var(X_n).$$

Overview

Continuous random variables

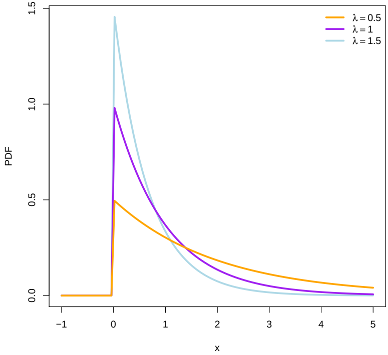
Particular continuous distributions and their parameters

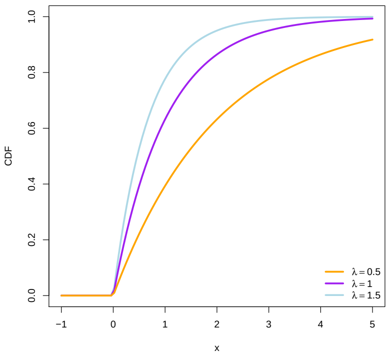
Uniform distribution

- R.v. X has a uniform distribution on $[a, b]$, we write $X \sim U(a, b)$, if $f_X(x) = 1/(b - a)$ for $x \in [a, b]$ and $f_X(x) = 0$ otherwise.

Exponential distribution $Exp(\lambda)$ with rate λ

$$F_X(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 1 - e^{-\lambda x} & \text{for } x \geq 0 \end{cases}$$

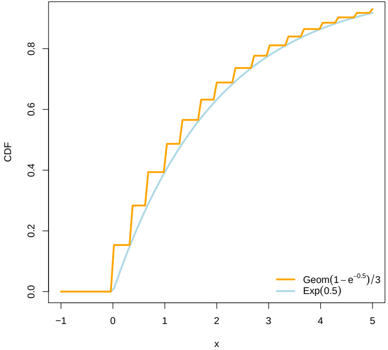




- X models time before next phone call in a call-center / web-server response / time till another lightning in a storm / ...

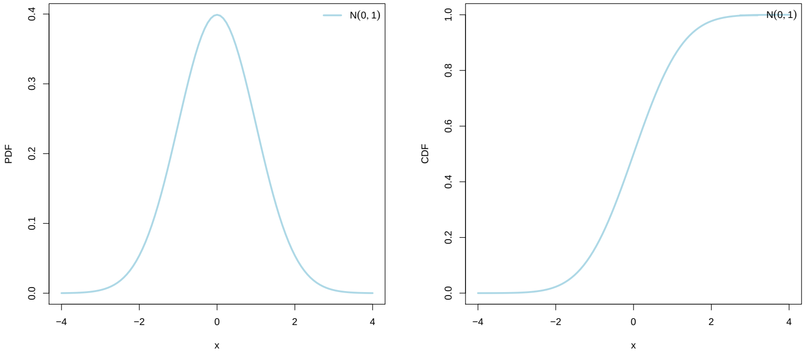
Relating $Exp(\lambda)$ and $Geom(p)$

- $P(X > x) = e^{-\lambda x}$ for $x > 0$
- $P(Y > n) = (1 - p)^n$ for $n \in \mathbb{N}$



Standard normal distribution

- $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
- $\Phi(x)$ – antiderivative of φ
- *Standard normal distribution* $N(0,1)$ has PDF φ and CDF Φ.
- If $Z \sim N(0,1)$, then $\mathbb{E}(Z) = 0$, $var(Z) = 1$.



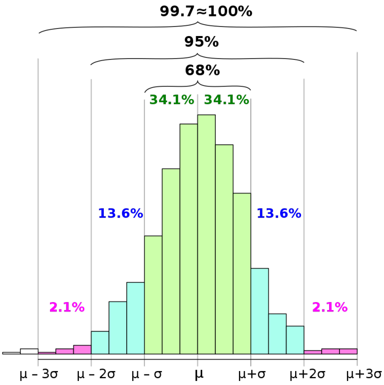
General normal distribution

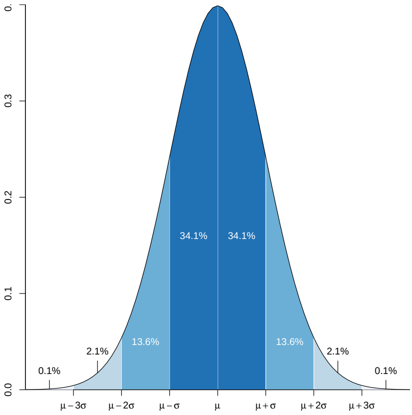
- For $\mu, \sigma \in \mathbb{R}$, $\sigma > 0$ we put $X = \mu + \sigma \cdot Z$, where $Z \sim N(0, 1)$.
- We write $X \sim N(\mu, \sigma^2)$ – general normal distribution
- Normal distribution $N(\mu, \sigma^2)$ has density $\frac{1}{\sigma}\varphi(\frac{x-\mu}{\sigma})$.

Resistance to a sum

Normal distribution – key properties

- 68–95–99.7 rule (3σ rule)
- Central limit theorem

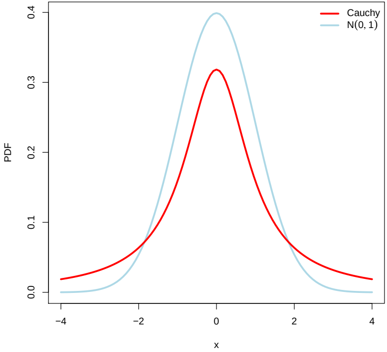




(Image on the left is from Wikipedia, author Melikamp.)

Cauchy distribution

- *Cauchy distribution:* PDF $f(x) = \frac{1}{\pi(1+x^2)}$
- does not have an expectation!



Gamma distribution

- $Gamma(w, \lambda)$, *gamma distribution with parameters* $w > 0$ *and* $\lambda > 0$ has PDF

$$f(x) = \begin{cases} 0 & \text{pro } x \leq 0 \\ \frac{1}{\Gamma(w)} \lambda^w x^{w-1} e^{-\lambda x} & \text{pro } x \geq 0 \end{cases}$$

  where $\Gamma(w) = (w-1)! = \int_0^\infty x^{w-1} e^{-x} dx$.
- For $w = 1$ we get exponential distribution again.
- If $X_1, \ldots, X_n$ are i.i.d with distribution $Exp(\lambda)$, then $X_1 + \cdots + X_n \sim Gamma(n, \lambda)$.
- Models lifetime of an electronic component, total of rainfall in a year, web-server latency.

A many others

- $Beta(s, t)$ – beta distribution
- χ^2 distribution with k degrees of freedom = chi-square (χ^2_k) is an alternative name for $Gamma(\frac{1}{2}k, \frac{1}{2})$. It is the distribution $Z_1^2 + \cdots + Z_k^2$, where $Z_i \sim N(0, 1)$ are i.i.d.
- Student t-distribution
- etc. etc.

Uniform distribution

- R.v. X has a uniform distribution on $[a, b]$, we write $X \sim U(a, b)$, if $f_X(x) = 1/(b-a)$ for $x \in [a, b]$ and $f_X(x) = 0$ otherwise.

Universality of uniform

Theorem

Let X be a r.v. with CDF $F_X = F$, let F be continuous and increasing. Then $F(X) \sim U(0,1)$.

Theorem

Let F be a function "of CDF-type": nondecreasing right-continuous function with $\lim_{x\to-\infty} F(x) = 0$ a $\lim_{x\to+\infty} F(x) = 1$. Let Q be the corresponding quantile function.
Let $U \sim U(0,1)$ and $X = Q(U)$. Then X has CDF F.