NMAI059 Probability and statistics 1
Class 6

Robert Šámal

Overview

Random vectors

Conditional distribution

Continuous random variables

Particular continuous distributions and their parameters

What have we learned

- Joint PMF: $p_{X,Y}(x,y) = P(X = x \,\&\, Y = Y)$
- Example: multinomial distribution
- Marginal PMF: $p_X(x) = \sum_{y \in Im(y)} p_{X,Y}(x,y)$
- Example: coupling
- X, Y are independent iff
  $P(X = x \,\&\, Y = y) = P(X = x)P(Y = y)$
  That is, iff $p_{X,Y}(x,y) = p_X(x)p_Y(y)$.
- If X, Y are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.
- $\mathbb{E}(g(X,Y)) = \sum_{x \in ImX} \sum_{y \in ImY} g(x,y)P(X = x, Y = y)$
- **Linearity of expectation** For any r.v.s X, Y and $a, b \in \mathbb{R}$ we have $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.
- **convolution formula**
  $P(X + Y = n) = \sum_{k \in Im(X)} P(X = k, Y = n - k)$

Overview

Random vectors

Conditional distribution

Continuous random variables

Particular continuous distributions and their parameters

Conditional PMF

X, Y – discrete random variables on $(\Omega, \mathcal{F}, P)$, $A \in \mathcal{F}$

- $p_{X|A}(x) := P(X = x \mid A)$
  example: X is outcome of a roll of a die, $A =$ we got an even number
- $p_{X|Y}(x|y) = P(X = x \mid Y = y)$ example: X, Z is an outcome of two independent die rolls, $Y = X + Z$.

$$p_{X|Y}(6|10) =$$

- $p_{X|Y}$ from $p_{X,Y}$:

Joint vs. conditional PMF

$p_{X,Y}$	...	10	11	12
1				
2				
3				
4				
5				
6				

$p_{X\|Y}$	...	10	11	12
1				
2				
3				
4				
5				
6				

Overview

Random vectors

Conditional distribution

Continuous random variables

Particular continuous distributions and their parameters

General random variable

Definition

Random variable on $(\Omega, \mathcal{F}, P)$ *is a mapping* $X : \Omega \to \mathbb{R}$*, such that for each* $x \in \mathbb{R}$

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}.$$

- discrete r.v. is a r.v.

CDF

Definition

Cumulative distribution function, CDF of a r.v. X is a function

$$F_X(x) := P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x).$$

- F_X is a nondecreasing function
- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to +\infty} F_X(x) = 1$
- F_X is right-continuous

CDF examples

Quantile function

For a r.v. X we define its *quantile function* $Q_X : [0, 1] \to \mathbb{R}$ by

$$Q_X(p) := \min\{x \in \mathbb{R} : p \leq F_X(x)\}$$

- If F_X is continuous and increasing then $Q_X = F_X^{-1}$.
- $Q_X(1/2) =$ median (watch out if F_X is not strictly increasing!)
- $Q_X(10/100) =$ tenth percentile, etc.

Continuous random variable

Definition

R.v. X is called continuous, if there is nonnegative real function f_X such that

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt.$$

(Sometimes such X is said to be absolutely continuous.)
Function f_X is called the probability density function, PDF of X.

- Alternatively: we pick a point from the probability space corresponding to the area under graph of f – nonnegative function with $\int_{-\infty}^{\infty} f = 1$.
- Let (X, Y) denote the coordinates of the point.
- Then X is a random variable with PDF f.

Using density

Theorem

Let X be a continuous r.v. with PDf f_X. Then

1. $P(X = x) = 0$ *for every* $x \in \mathbb{R}$.
2. $P(a \leq X \leq b) = \int_a^b f_X(t)dt$ *for every* $a, b \in \mathbb{R}$.

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 a type $\infty - \infty$. TODO EXPLAIN?

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

Continuous LOTUS

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.)

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

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

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 a uniform distribution

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.

1. *Let $U \sim U(0,1)$ and $X = Q(U)$. Then X has CDF F.*
2. *Let X be a r.v. with CDF $F_X = F$, suppose F is increasing. Then $F(X) \sim U(0,1)$.*

Overview

Random vectors

Conditional distribution

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.

Exponencial distribution

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

Relating Exp and $Geom$

Normal distribution

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