# NMAI059 Probability and statistics 1
# Class 4

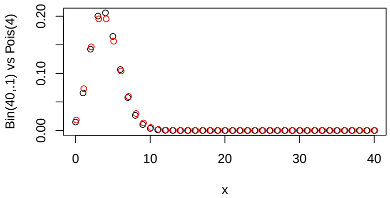
Robert Šámal

# Overview

# What we have learned

- What is a discrete r.v.
- How to describe it using a PMF and/or CDF.
- Examples of distributions: Bernoulli, binomial, hypergeometric, Poisson, geometric.
- Expectation: two possible definitions
- $\mathbb{E}(X) = \sum_{x \in Im(X)} x \cdot P(X = x)$
- $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\})$
- $\mathbb{E}(g(X)) = \sum_{x \in Im(X)} g(x) P(X = x)$ (LOTUS)
- "How much we expect to get on average, when we repeat independent experiments with result given by $X$" … we will discuss later as the law of large numbers.

# Comparing binomial and Poisson distribution: PMF




**Generated by the following code in R**

```
x = 0:40
bin = dbinom(x,40,0.1)
pois = dpois(x,4)
plot(x,bin, ylab="Bin(40,.1)␣vs␣Pois(4)")
points(x+.1,pois,col="red")
```

# Properties of $\mathbb{E}$

## Theorem

*Suppose $X, Y$ are discrete r.v. and $a, b \in \mathbb{R}$.*

1. *If $P(X \geq 0) = 1$ and $\mathbb{E}(X) = 0$, then $P(X = 0) = 1$.*
2. *If $\mathbb{E}(X) \geq 0$ then $P(X \geq 0) > 0$.*
3. $\mathbb{E}(a \cdot X + b) = a \cdot \mathbb{E}(X) + b$.
4. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

# Another formula for expectation

## Theorem

*Let $X$ be a discrete r.v. such that $Im(X) \subseteq \mathbb{N}_0 = \{0, 1, 2, \dots\}$. Then we have*

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} P(X > n).$$

# Variance

## Definition

*Variance of a r.v. $X$ is the number $\mathbb{E}((X - \mathbb{E}X)^2)$. It is denoted by $var(X)$.*

## Theorem

$$var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

# Conditional expectation

## Definition

*Let $X$ be a discrete r.v. and $P(B) > 0$. Conditional expectation of $X$ given $B$ is*

$$\mathbb{E}(X \mid B) = \sum_{x \in Im(X)} x \cdot P(X = x \mid B),$$

*whenever the sum is defined.*

# Law Of Total Expectation

### Theorem

*Suppose* $B_1, B_2, \ldots$ *is a partition of* $\Omega$ *and* $A \in \mathcal{F}$*. Then*

$$\mathbb{E}(X) = \sum_i P(B_i)\mathbb{E}(X \mid B_i),$$

*whenever the sum is defined. (Terms with* $P(B_i) = 0$ *are counted as 0.)*

# Law Of Total Expectation

# Overview

Discrete r.v. – expectation and variance

Parameters of discrete distributions

Random vectors

# Distribution parameters – Bernoulli

Pro $X \sim Bern(p)$ je

- $\mathbb{E}(X) = p$
- $var(X) = p - p^2$

# Distribution parameters – binomial

Pro $X \sim Bin(n,p)$ je

- $\mathbb{E}(X) = np$
- $var(X) = np(1-p)$

- First way: $X = \sum_{i=1}^{n} X_i$, where $X_i =$
- $\mathbb{E}(X_i) = P(X_i = 1) =$

- Second way:
  $\mathbb{E}(X) = \sum_{k=0}^{n} k \cdot P(X = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$

# Distribution parameters – hypergeometric

Pro $X \sim Hyper(N, K, n)$

- $\mathbb{E}(X) = n\frac{K}{N}$
- $var(X) = n\frac{K}{N}(1 - \frac{K}{N})\frac{N-n}{N-1}$

- First way: $X = \sum_{i=1}^{n} X_i$, where $X_i =$
- $\mathbb{E}(X_i) = P(X_i = 1) =$

- Second way: $X = \sum_{j=1}^{K} Y_j$, where $Y_j =$
- $\mathbb{E}(Y_j) = P(Y_j = 1) =$

# Distribution parameters – geometric

For $X \sim Geom(p)$ we have

- $\mathbb{E}(X) = 1/p$
- $var(X) = \frac{1-p}{p^2}$

# Distribution parameters – Poisson

Pro $X \sim Pois(\lambda)$ je

- $\mathbb{E}(X) = \lambda$
- $var(X) = \lambda$

# Overview

Discrete r.v. – expectation and variance

Parameters of discrete distributions

Random vectors

# Basic description of random vectors

- $X, Y$ – random variables on the same probability space $(\Omega, \mathcal{F}, P)$.
- We wish to treat $(X, Y)$ as one object – a random vector.
- How to do that?
- Example: we roll twice a 4-sided dice, $X =$ first outcome, $Y =$ second one.

# Joint distribution

## Definition

*For a discrete r.v.* $X, Y$ *on a probability space* $(\Omega, \mathcal{F}, P)$ *we define their joint PMF* $p_{X,Y} : \mathbb{R}^2 \to [0,1]$ *by a formula*

$$p_{X,Y}(x,y) = P(\{\omega \in \Omega : X(\omega) = x \& Y(\omega) = y).$$

- We can define it also for more than two r.v.'s $p_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$.

# Marginal distribution

- Given $p_{X,Y}$, how to find the distribution of each of the coordinates, that is $p_X$ and $p_Y$?

# Independence of r.v.'s

## Definition

*Discrete r.v.'s $X, Y$ are independent if for every $x, y \in \mathbb{R}$ the events $\{X = x\}$ a $\{Y = y\}$ are independent. That happens if and only if*

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

# Product of independent r.v.'s

## Theorem

*For independent discrete r.v.'s $X, Y$ we have*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

# Function of a random vector

### Theorem

*Suppose $X, Y$ are r.v.'s on $(\Omega, \mathcal{F}, P)$, let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function.*

- *Then $Z = g(X, Y)$ is a r.v. on $(\Omega, \mathcal{F}, P)$*
- *and it satisfies*

$$\mathbb{E}(g(X,Y)) = \sum_{x \in ImX} \sum_{y \in ImY} g(x,y) P(X = x, Y = y),$$

*whenever the sum is defined.*

### Theorem

*For $X, Y$ r.v.'s and $a, b \in \mathbb{R}$ we have*

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

# Proof of the theorem about variance

# Sum of independent r.v.'s

- Given $p_{X,Y}$, how to find the distribution of the sum, $Z = X + Y$?

# Sum of independent r.v.'s – convolution

## Theorem

*Let $X$, $Y$ be discrete random variables. Then their sum $Z = X + Y$ has PMF given by*

$$P(Z = z) = \sum_{x \in Im(X)} P(X = x, Y = z - x).$$

*If we further assume that $X$, $Y$ are independent, then*

$$P(Z = z) = \sum_{x \in Im(X)} P(X = x)P(Y = z - x).$$