# Probability Theory & Statistics

Innopolis University, BS-I,II

Spring Semester 2016-17

Lecturer: Nikolay Shilov

Part I

# DISCRETE RANDOM VARIABLES

# Concept

- *Discrete random variable* is any (total) real function on finite domain $X:\Omega\rightarrow R$.
- Notation convention: If X, Y. Z are random variables then x, y, and z are reserved for values of these functions.

# From Random Variable to …

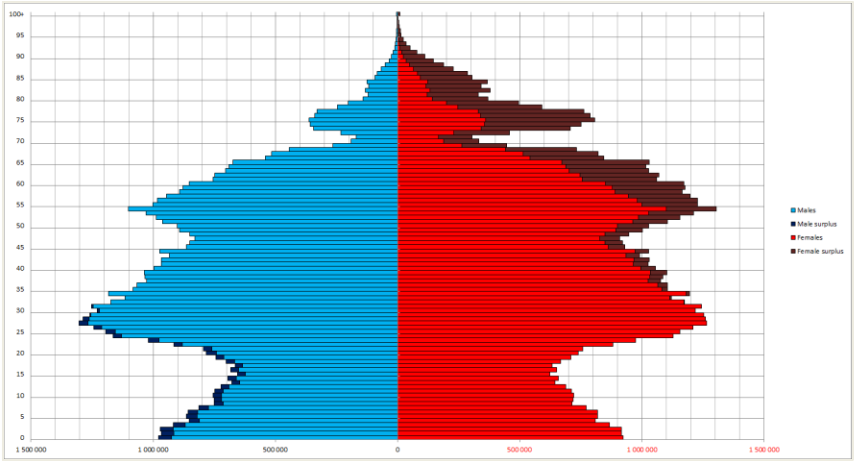
- Discrete random variable $X:\Omega\rightarrow R$ is a table that assigns an individual value to each outcome in $\Omega$.
- Example: X is a table of the current age (in years as on February 20-25, 2017) to each RF citizen, i.e. Nikolay Shilov (with INN=… ) is 55, etc.

# … to Distribution Function …

- *(Frequency) distribution* is a table/function that assigns number (i.e. non-normalized frequency) $|X^{-}(x)|$ of the corresponding outcomes to each value x of the variable X.
- Here
  - $X^{-}$ is the inverse of X as a function,
  - $|\ldots|$ is the number of elements of a set.

# … Distribution Function (cont) …

Example: a table how many people (RF citizens) are/were 1, 2, 3, … 55, … years old at some date (represented in the form of the population pyramid).




(https://en.wikipedia.org/wiki/Demographics_of_Russia#/media/File:Russia_Sex_by_Age_20150101.png)

# … and to Probability Distribution Function

- *(Discrete) probability distribution* (or *probability mass function*) is a table/function that assigns probability (i.e. *the normalized frequency*)

$$P_X(x) = P(X=x) = \frac{|X^{-}(x)|}{|\Omega|}$$

of the corresponding outcomes to each value x in the range of the variable X.

# Example: RF Population Probability Mass Function

- Probability distribution for random variable RFPop.date (Russian Federation Population on concrete date) is a function that assigns
  - to each age x
  - the ratio

$$\frac{\text{RF citizens at age x on this date}}{\text{total number of RF citizens on this date}}$$

Part II

# SIMPLE EXAMPLES AND DISCUSSION

# Domino Example

Exercise: build distributions for random variable that assigns the sum of spots (pips, nips, or dobs) to each piece (i.e. tile)

# Pair of Dices Example

- Exercise: build distributions for random variable that assigns the sum of pips for tossed pair of idealized dices.

# Lottery Example

- Some lottery has probability distribution function defined by the following table:

| Prize | 1000 | 100 | 1 | 0 |
|---|---|---|---|---|
| Probability | 0.0001 | 0.001 | 0.01 | ? |

- Fill in the gap and suggest a discrete random variable with this discrete probability distribution.

# Lottery Example (cont)

- What is the smallest size n of the set of the outcomes of a random variable that has this distribution?
- How many exist different discrete random variables with a fixed set of outcomes with this size n?

# What is the probability space?

- Let $X:\Omega \rightarrow R$ be a (discrete) random variable and $P_X$ be its pobability mass function

$$P_X(x) = P(X=x) = |X^{-}(x)|/|\Omega|.$$

- What is the probability space where $P_X$ serves the probability function role?

# Probability Space affiliated with Random Variable

- The sample space $\Omega' = \{ S \subseteq \Omega : S = X^{-}(x), x \in R \}$;
- The event space $\mathcal{F} = 2^{\Omega'}$;
- The probability function $P: \mathcal{F} \rightarrow [0,1]$ is the additive continuation on $\mathcal{F}$ of a function defined on samples as $P(X^{-}(x)) = x$ for any $x \in R$.

# Exercise about affiliated space

- Are outcomes affiliated with a random variable same as samples of the probability space affiliated with the random variable?
- Is the definition of the affiliated probability space correct, i.e.
  - does it define unique space,
  - and it defines really a probability space?

# Cumulative Distribution

- Let $X:\Omega \to R$ be a (discrete) random variable is any (total) real function on finite domain.
- It defines probability mass function $P_X$ and *cumulative distribution function*

$$F_X(x) = P_X(X \leq x) = \Sigma_{y \leq x} P_X(y).$$

- You used cumulative distribution when you calculated probability for Nikolay Shilov to survive next 10 years in lab classes on week 4.

# Functions of a Random Variable

- Let $X:\Omega\to R$ be a random variable and $g:R\to R$ be function (that is defined on the range of X at least).
- This function g defines a function of this random variable X with the following probability distribution $P_{Y=g(X)}(y) = \Sigma_{g(x)=y}P_X(x)$.

# Type it!

- Question: Consider a total function $g:R \to R$ as a function on probability distributions and type it!

# Example of a random variable function

Distribution of X:

| | |
|---|---|
| $P_X(1)$ | 1/4 |
| $P_X(-1)$ | 1/8 |
| $P_X(2)$ | 5/8 |

$Y=X^2$

Distribution of $Y=X^2$:

| | |
|---|---|
| $P_Y(1)$ | 3/8 |
| $P_Y(4)$ | 5/8 |