Probability Theory & Statistics

Innopolis University, BS-I,II

Spring Semester 2016-17

Lecturer: Nikolay Shilov

Part I

LAW(S) OF LARGE NUMBERS

Bernoulli' law: statement

- Assume we are given an infinite series of trials X_1, … X_n, … where X_n= binomial(n, p) and $0<p<1$.
- Then for every $\varepsilon>0$ the probability

 that normalized frequency of success

 equals p with accuracy ε

 converges to 1:

$$P\left(\left|\frac{m}{n}-p\right| \leq \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 1$$

Bernoulli' law: proof sketch

$$-n\varepsilon \le m - np \le n\varepsilon$$

$$x = \frac{m-np}{\sqrt{npq}}$$

$$x_1 = -\varepsilon\sqrt{\frac{n}{pq}} \le x \le \varepsilon\sqrt{\frac{n}{pq}} = x_2$$

$$P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) = P(x_1 \le x \le x_2) \approx$$

$$\approx \Phi_0(x_2) - \Phi_0(x_1) = 2\Phi_0(x_2) = 2\Phi_0\left(\varepsilon\sqrt{\frac{n}{pq}}\right) \underset{n\to\infty}{\longrightarrow} 1$$

$$P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 1$$

Chebyshev's inequality: statement

- Let X be a random variable with a finite expectation μ and finite non-zero deviation σ.
- Then for any real number $k>0$

$$P(|X - \mu| \geq k\sigma) \leq 1/k^2.$$

Chebyshev's inequality: proof

- $P(|X - \mu| \geq k\,\sigma) = M(I_{|X - \mu| \geq k\,\sigma}) =$

$$= M(I_{|X - \mu|/(k\sigma) \geq 1}) \leq M((X - \mu)^2/\,(k\,\sigma)^2) =$$

$$= M((X - \mu)^2)/(k^2\,\sigma^2) = 1/k^2.$$

Khintchin's (weak) law: statement

- Let $X_1, X_2, \ldots$ be an infinite sequence of *independent and identically distributed* (i.i.d. or IID) random variables with same sets of outcomes, finite expectation μ and finite non-zero deviation σ.
- Let $\underline{\mathbf{X}}_{\mathbf{n}}$ be $(X_1 + \ldots + X_n)/n$
- Then for every $\varepsilon>0$ the probability

$$P(|\underline{\mathbf{X}}_{\mathbf{n}} - \mu|>\varepsilon)$$

converges to 0 (as $n\to\infty$)

Khintchin’s (weak) law: proof

- $M(\underline{\mathbf{X}}_n) = \mu$ and $D(\underline{\mathbf{X}}_n) = \sigma^2/n$;
- Using Chebyshev’ inequality

$$P(|X - \mu| \geq k\sigma) \leq 1/k^2$$

for $\underline{\mathbf{X}}_n$: $P(|\underline{\mathbf{X}}_n - \mu| \geq \varepsilon) \leq \sigma^2/(n\varepsilon^2)$.

Part II

A NEED OF CONTINUOUS SAMPLE/OUTCOME SPACES

Finite case doesn't work any more

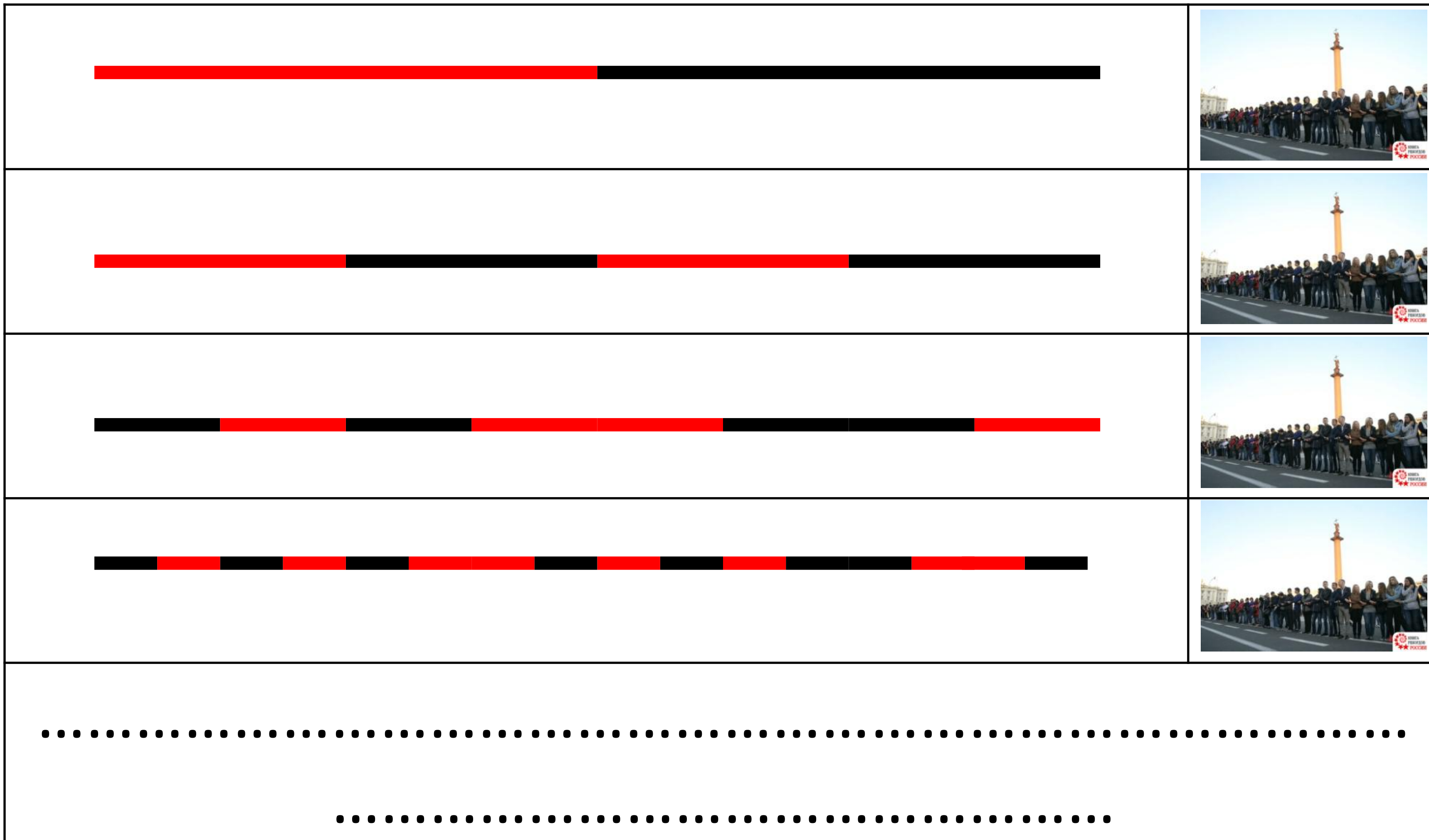
- If X is a random variable with finite set of outcomes than the set of all random variables Y that are IID with X is *finite*.
- For example: there is no any random variable Y that is IID with
  - tossing idealized coin,
  - rolling idealized dice,
  - a random variable X with n values and outcomes.

Tossing coin as a random variables with infinite outcomes

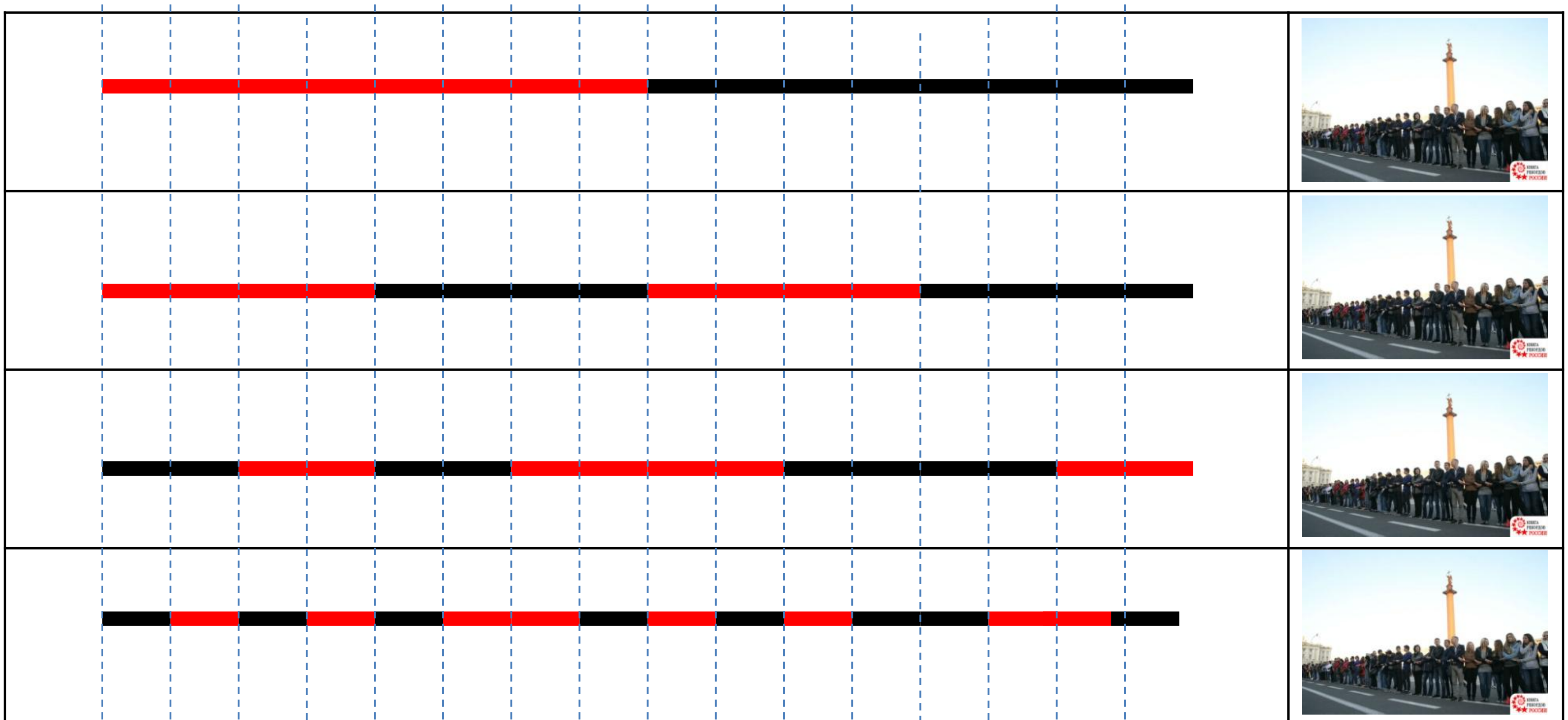
- There is an infinite chain of people but with a finite *measure*,
- they individually flip coins and half of them get *head* another half – *tails*.

Самая длинная цепочка рукопожатий (Книга Рекордов России, http://knigarekordovrossii.ru/index.php/rekordy/kategorii/massovye-meropriyatiya/1035-samaya-dlinnaya-tsepochka-rukopozhatij.html#!prettyPhoto[galleryaf22339285]/2/)

An infinite sequence of IID random variables



How to understand Law of Large Numbers



X_4	1/2	3/4	3/4	1	1/4	1/2	3/4	1/2	3/4	1/2	1/2	1/4	0	1/4	1/2	1/4

Part III

PROBABILITY SPACES (CONTINUES EUCLIDEAN CASE)

Probability Space Definition (compare with week 2)

- A *probability space* is a sample space Ω together with a set of events $\mathcal{F} \subseteq 2^{\Omega}$ (that must be a σ-algebra) and a non-negative additive probability function P to all events and satisfying normalization condition $P(\Omega)=1$.

In other words…

- A probability space is a triple

$$(\Omega, \mathcal{F}, \mathrm{P})$$

where
  - Ω is a finite event/sample space,
  - $\mathcal{F} \subseteq 2^{\Omega}$ is the set of events,
  - and P: $\mathcal{F} \to [0,1]$ a (total) probability function satisfying *axioms*.

Probability Axioms

- Non-negativity: $0 \leq P(A)$ for every event;
- Normalization: $P(\Omega)=1$;
- Countable additivity: $P(\cup_{k \in N} A_k) = \Sigma_{k \in N} P(A_k)$ assuming that all events are pair-wise exclusive.

Simple Properties

- Boundness: $P(A) \leq 1$ for every event.
- Impossibility: $P(\varnothing)=0$.
- Additivity (finite case): for any finite collection of (pair-wise) mutually exclusive events

$$P(\cup_{1 \leq j \leq n} A_j) = \Sigma_{1 \leq j \leq n} P(A_j).$$

Further Properties

- Complimentarity: $P(A^c) = 1 - P(A)$ for every event.
- Difference: $P(A\backslash B) = P(A) - P(A\cap B)$ for all events.
- Monotonicity: if an event B implies A then A is more probable than B, i.e.

$$B\subseteq A \text{ implies } P(B)\leq P(A).$$

Variants in Euclidean Spaces

- Probability within
  - an interval of length L: for any set of subintervals of total length l let probability be l/L;
  - a figure with area S: for any subfigure with area s let probability be s/S;
  - a 3D-body with volume V: for any sub-body with volume v let probability be v/V.

Example

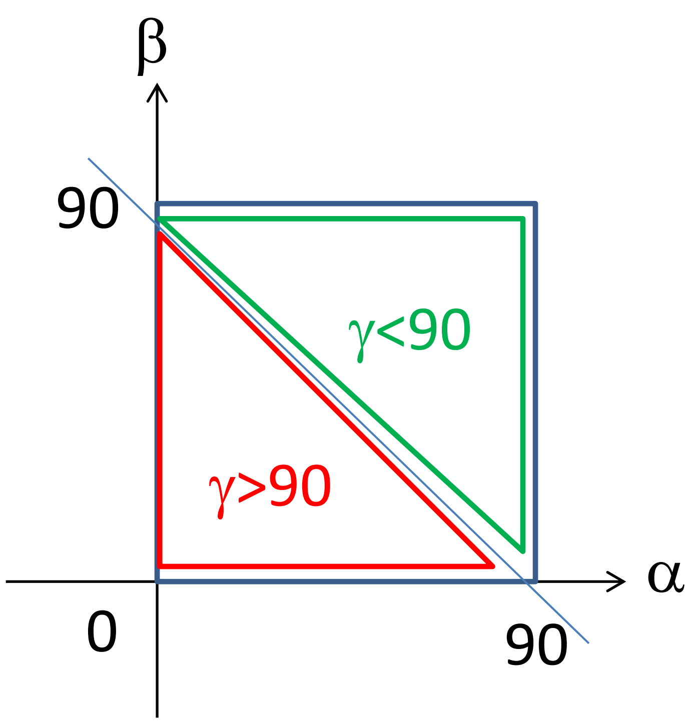
- What is probability that a randomly drown triangle be
  - an acute triangle;
  - a right triangle;
  - an obtuse triangle?

An intuitive approach

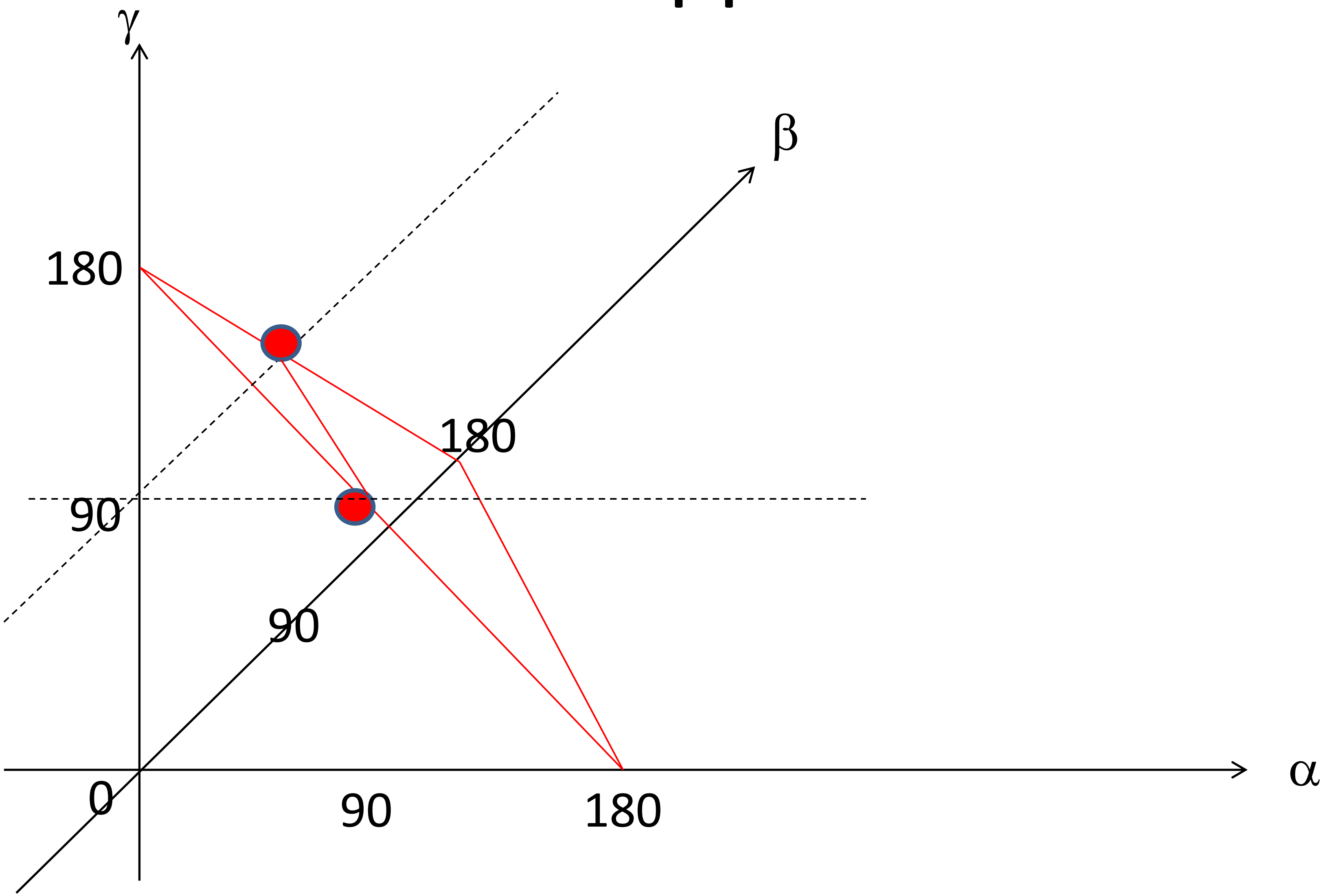
- Consider the angle for some fixed vertex of a “randomly drown” triangle: it ranges from 0 to 180 degrees.
- The probability that the angle
  - ranges in [0,90) is 0.5,
  - is 90 exactly is 0,
  - ranges in (90,180] is 0.5.

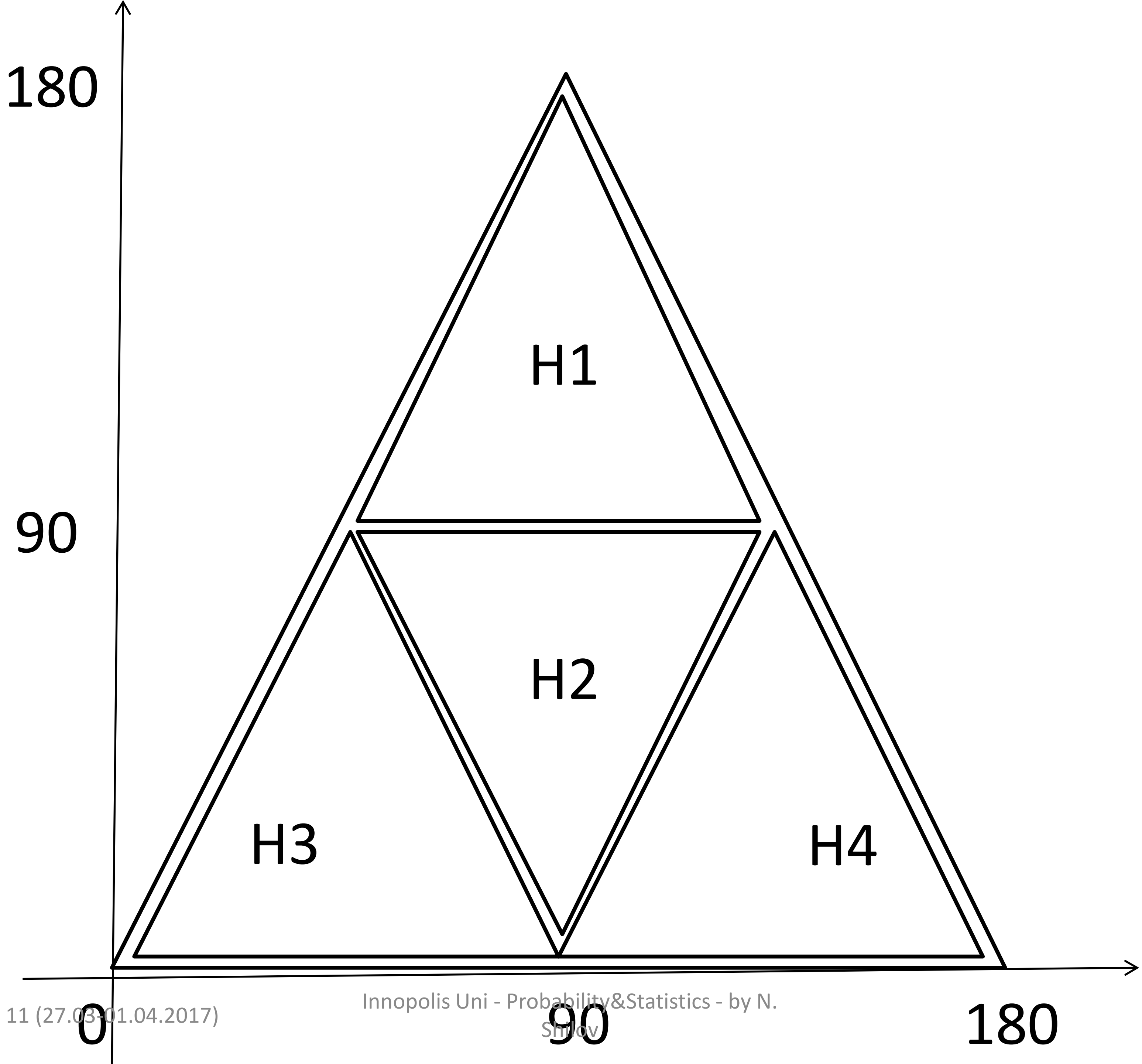
Another approach to the problem

- Two angles α and β are independently ranging in [0, 90], the third angle γ is $(180 - \alpha - \beta)$.
- The probability that γ
  - ranges in [0,90) is 0.5,
  - is 90 exactly is 0,
  - ranges in (90,180] is 0.5.



One more approach…




180
90
H1
H2
H3
H4
0
90
180

Results of the study…

- The probability that
  - all angles range in [0,90) is 0.25,
  - any angle is 90 exactly is 0,
  - any angle ranges in (90,180] is 0.75.