Probability Theory & Statistics

Innopolis University, BS-I,II

Spring Semester 2016-17

Lecturer: Nikolay Shilov

Part I

EVENT SPACES (FINITE CASE)

Just a Set?

- Let Ω be a finite set … Sometimes it makes sense to call/to think about its elements in other way than just elements.
- In particular, in the probability theory elements are *elementary events*, *samples*, *choices*, *options*, etc. In this cases a set becomes a *space.*
- Consider/discuss examples: dice, coin, cards, elephants in Innopolis, etc…

Subset or Events?

- If a set is a space then subsets are called *events*.
- The set of all events (in finite case!) is the power-set denoted as P(Ω) or 2^{Ω}.

Operations on Events

- Evens inherited the standard set-theoretic operations (sometimes with a special terminology):
  - *union* or *sum* ($\cup$);
  - *intersection* or *product* ($\cap$);
  - *compliment* ($\_^{c}$),
  - *difference* ($\_\backslash\_$)…

Impossible, Certain and Disjoint Events

- If a set is a space then
  - the empty set $\varnothing$ is called *the impossible event*;
  - the space Ω is called *the certain/sure event*;
  - disjoint events (i.e. $A \cap B = \varnothing$) are called (*mutually*) *exclusive*.
- Sometimes there may be impossible events other than $\varnothing$, certain events other than Ω…

Probability vs. Non-Determinism

- In non-determinism elementary events have no numeric measure. (In other words: there is no any reasonable way to assign numeric values to elementary events.)
- Probability theory assumes that there is a reasonable way to assign *measures* (*non-negative* numeric values) to all elementary events (by statistics for instance) in a *sample space* Ω.

Expanding a Measure on Events

- If $V:\Omega \rightarrow R$ is an assignment of *non-negative* numeric values to all elementary events
- then a numeric value $V(A)$ may/can be assigned to every event $A \in P(\Omega)$ in *additive manner*:

$$V(A) = \Sigma_{a \in A} V(a) = V(a_1) + \ldots + V(a_n)$$

where $a_1, \ldots a_n$ is an explicit enumeration of elementary events in A.

Part II

PROBABILITY SPACES (FINITE CASE)

Probability Definition

- Let Ω be a (finite) space with a measure V for all elementary events, such that the certain event Ω *is not an impossible* (i.e. $V(\Omega)\neq 0$).
- Then a probability of an event $A\subseteq\Omega$ is

$$P(A)= V(A)/V(\Omega).$$

Probability Properties

- Non-negativity: $0 \leq P(A)$ for every event;
- Normalization: $P(\Omega)=1$;
- Additivity: $P(A \cup B)= P(A) + P(B)$ for all exclusive events.

(finite) Probability Space Definition

- A (finite) *probability space* is a (finite) event/sample space Ω together with the set of events 2^{Ω} and a non-negative additive probability function P to all events and satisfying normalization condition $P(\Omega)=1$.

In other words…

- A (finite) probability space is a triple

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

where

– Ω is a finite event/sample space,

– $\mathcal{F} = 2^{\Omega}$ is the set of events,

– and P: $\mathcal{F} \to [0,1]$ a (total) probability function satisfying *axioms* on slide about Probability Properties.

Does It Makes Sense…

- Philosophically? Mathematically?
- “In random” usually means a “flat” probability assignment: $P(\omega) = 1/|\Omega|$ for every elementary event $\omega \in \Omega$.
- Example: Course has 90 enrolled students, a project group must consists of 3 students, groups will be formatted in random. What does it means?

Part III

PROBABILITY PROPERTIES CALCULUS

Simple Properties

- Boundness: $P(A) \leq 1$ for every event.
- Impossibility: $P(\varnothing)=1$.
- 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: $P(A \setminus B) = P(A) - P(A \cap B)$ for all events.
- Monotonicity: if an event B implies A then A is more probable than B, i.e.

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

Inclusion-Exclusion Principle

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
- $P(A \cup B \cup C) = P(A) + P(B) + P(C) -$

$$- P(A \cap B) - P(A \cap C) - P(B \cap C) +$$

$$+ P(A \cap B \cap C);$$

- Example: There are 1000 smart students; 750 of these students have iPads, 450 owned individual cars, 350 – both. How many smart students have either a car or iPad?

General Inclusion-Exclusion Principle

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

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

  $- \Sigma_{1\leq j<k\leq n}P(A_j \cap A_k) +$

  $+ \Sigma_{1\leq j<k<m\leq n}P(A_j \cap A_k \cap A_m) -$

  ....

  $(-1)^{n-1}P(A_1 \cap A_2 \cap \ldots \cap A_n)$

Probability of a dearangemnet

- There are n people numbered 1, 2, ..., n and the hats also numbered 1, 2, ..., n.
- People pick up hats in random. Assuming $n=4$ what is a probability that nobody picks up his/her hat?

Counting Derangements

- Let us assume that the first person takes hat k. There are (n – 1) options for this choice.
- Then there are two alternatives:
  – Person k does not take the hat 1; in this case each of the remaining (n-1) people has precisely 1 forbidden choice in (n-1) hats.
  – Person k takes the hat 1; the problem reduces to (n-2) people and (n-2) hats.

Subfactorial

- The arguments on slide Counting Derangements lead to the following recurrence formula for the number of derangements (also known as the *subfactorial*)

$$!n=(n-1)(!(n-1)+!(n-2))$$

with initial values $!0= 1$ and $!1 = 0$.

What a Surprise!

- Factorial has the same recurrence formula

$$n!=(n-1)((n-1)!+(n-2)!)$$

but with other initial values $0! = 1! = 1$.