Probability Theory & Statistics

Innopolis University, BS-I,II

Spring Semester 2016-17

Lecturer: Nikolay Shilov

Part I

PROBLEMS ABOUT PROBABILITY SPACE

Problem 1: statement

- All seats on flight Novosibirsk-Kazan are booked and all passengers are queuing in line for boarding.
- But the first passenger, a crazy Professor, rushes to occupy a random seat in the cabin.

Problem 1: statement (cont)

- All other passengers are polite and each obeys to the following discipline: upon boarding passenger checks its boarding pass and whether assigned seat is free or occupied; in the former case the passenger occupies the assigned seat, in the late case – any free seat.
- What is probability that the last passenger occupies a seat according to boarding pass?

Problem 1: solution

- Let us assume that there are passengers and enumerate passengers in line by numbers [1. . n] and enumerate cabin seats according to passenger number in line. (That is: if a seat is assigned to a passenger with number k in [1. .n] in the queue then we refer this seat by this number k.)

Problem 1: solution (cont.)

- Then the probability space is collection of events (1,m_1,m_2,… m_p) where
  - $0 \le p \le n$ is a number of passengers that occupy non-assigned seats,
  - $1 < m_1 < m_2 < \ldots < m_p$ is a increasing sequence in [1. . n];

Problem 1: solution (cont.)

- The event $(1,m_1,m_2,\ldots m_p)$ means that
  - the passenger 1 occupies the seat of the passenger m_1,
  - the passenger m_1 occupies the seat of the passenger m_2,
  - …
  - but the passenger m_p occupies the seat of the first passenger.

Problem 1: solution (cont.)

Since

- every event $(1, m_1, m_2, \dots m_p)$ where $m_p \neq n$ (i.e. the last passenger occupies the seat according to boarding pass)

corresponds in 1-1 manner

- to event $(1, m_1, m_2, \dots m_p, n)$ when the last passenger occupies not-assigned seat,

the desired probability is exactly 0.5.

Problem 2: statement

- There are 3 parties A,B and C. Party A has 50% support, party B – 30% support and C – 20% support nation-wide.
- There are 100 voting districts to elect a member of Parliament with 100 seats. All voting districts are equal in number of voters.

Problem 2: statement (cont)

- A party candidate wins in a district if he/she gets at least 50% support in the district.
- What are maximal and minimal numbers of seats that parties can get after elections?
- What are probabilities of the best and the worst outcome for each party?

Part II

ELEMENTS OF ENUMERATIVE COMBINATORICS

350 years of Combinatorics

- Combinatorics studies arrangements, partitions and other ways of organization/structuring finite sets.
- The term was introduced by G.W. Leibniz in 1666

Combinatorics Main Rules - I

Pigeonhole principle (or Dirichlet's box/drawer principle): if n items are put into m containers, with n>m, then at least one container must contain more than one item.

Combinatorics Main Rules - III

- Rule of sum (or addition principle): the size of the union of a finite collection of pairwise disjoint sets is the sum of the sizes of these sets.
  - Example:…

Combinatorics Main Rules - II

- Rule of product (or multiplication principle): the size of the Cartesian product of a finite collection of finite sets is the product of the sizes of these sets.
  - Example:…

Arrangements

- *Arrangement* of k elements from a set of n elements is any sequence (vector) of k different elements from the set.
- The number of arrangements A_n^k is

$$n*(n-1)*\ldots*(n-k) = n!/(n-k)!$$

Permutations

- *Permutation* is an “arrangement of n form n”.
- The number of permutations P_n is n!
- Explain why $P_0=0!=1$.

Combinations

- k-combination of a set that has n elements is a subset of k distinct elements of the set.
- The number of combinations C_n^k also denoted as $\binom{n}{k}$ is

$$\frac{n!}{k!(n-k)!}$$

Combinations

- Explain (not prove!) why
  - C_n^k equals to the appropriate binomial coefficient;
  - $C_n^k = C_{(n-1)}^{(k-1)} + C_{(n-1)}^k$;
  - $\Sigma_{0 \leq k \leq n} C_n^k = 2^n$.

Part III

1,000,000$ FOR ENUMERATE COMBINATIONS?

A Hard Problem

- Suppose that you are organizing housing accommodations for a group of 400 university students. Space is limited and only 100 of the students will receive places in the dormitory. To complicate matters, the Dean has provided you with a list of pairs of incompatible students, and requested that no pair from this list appear in your final choice.

(http://www.claymath.org/millennium-problems/p-vs-np-problem)

A Hard Problem (cont.)

- Indeed, the total number of ways of choosing 100 students from the 400 applicants is greater than the number of atoms in the known universe! Thus no future civilization could ever hope to build a supercomputer capable of solving the problem by brute force; that is, by checking every possible combination of 100 students.

(http://www.claymath.org/millennium-problems/p-vs-np-problem)

Exercises for you

- Estimate (without computer aid) the number of digits in the decimal representation of the mentioned number of combinations.
- Check precision (accuracy) of your estimation with aid of computer (but without symbolic computations of factorials).

A Hard Problem (cont.)

- However, this apparent difficulty may only reflect the lack of ingenuity of your programmer. In fact, one of the outstanding problems in computer science is determining whether questions exist whose answer can be quickly checked, but which require an impossibly long time to solve by any direct procedure.

(http://www.claymath.org/millennium-problems/p-vs-np-problem)

A Hard Problem (cont.)

- Problems like the one listed above certainly seem to be of this kind, but so far no one has managed to prove that any of them really are so hard as they appear, i.e., that there really is no feasible way to generate an answer with the help of a computer.

(http://www.claymath.org/millennium-problems/p-vs-np-problem)

Why Millennium Prize

- In order to celebrate mathematics in the new millennium, The Clay Mathematics Institute of Cambridge, Massachusetts (CMI) has named seven "Millennium Prize Problems." The Scientific Advisory Board of CMI selected these problems, focusing on important classic questions that have resisted solution over the years. The Board of Directors of CMI have designated a $7 million prize fund for the solution to these problems, with $1 million allocated to each.

(http://www.claymath.org/millennium-problems/p-vs-np-problem)