MVE055 2017 Lecture 1

Marco Longfils

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg

Monday 28th August, 2017

Introduction to Statistics

- Statistics deals with the collection, analysis, interpretation, and visualization of data.
- In Statistics we try to build models which takes into account the uncertainty to describe the real world.
- Statistical (or stochastic) models:
  - try to describe a real phenomenon;
  - are used to predict new observations of the phenomenon;
  - make assumptions and often are simplification of what they try to model;
- Statistical models are used in everyday life to predict events characterized by uncertainty. An example is weather forecasting, where we try to predict how likely is that tomorrow will be rainy.

Probability theory

- The mathematics on which Statistics is built upon is called probability theory.
- There are different schools when it comes to the interpetation of predictions from statistical models. Example:
  1. The probability that this geological formation contains oil is 34%.
  2. The probability of a successful heart surgery is about 96%.
- In a subjective or Bayesian model one thinks the statistical model as a model for an individual knowledge of a certain phenomenon. Therefore two persons with the same informations will obtain the same probabilities. We can talk about the probability of events which cannot be repeated, as (1) above.

Probability theory

- In a frequentists approach we can assign a probability only to events that can be repeated. The probability of an event is defined as the frequency of the event when the number of repetition approach infinity. In (2) 96% can be obtain as a frequency.

Some definitions

Outcome (or sample point) is a particular physical outcome of experiment under investigation.

Sample space is the collection of all possible outcomes.

Event is a subset of the sample space.

Axioms of Probability theory

Axiom

Let S be the sample space. A probability distribution is a function which associate to all (measurable) events A in S a value $Pr[A]$ such that

- $\Pr[S] = 1$
- $Pr[A] \geq 0$
- *Let $A_1, \ldots, A_k, \ldots$ be non-overlapping events, i.e. $A_i \cap A_j = \emptyset$ if $i \neq j$. Then*

$$\Pr[A_1 \cup A_2 \cup \cdots \cup A_k \cup \cdots] = Pr[A_1] + Pr[A_2] + \cdots + \Pr[A_k] + \cdots$$

Some consequences of the axioms

- $\Pr[\emptyset] = 0$;
- $\Pr[A^c] = 1 - \Pr[A]$, where A^c is the complement of A in S (sometimes we write $S \setminus A$);
- If A_1 and A_2 are two events, then

$$\Pr[A_1 \cup A_2] = \Pr[A_1] + \Pr[A_2] - \Pr[A_1 \cap A_2].$$

Conditional probability

- The conditional probability of an event A_2 given an event A_1, with $\Pr[A_1] > 0$, is defined as

$$\Pr[A_2|A_1] = \frac{\Pr[A_1 \cap A_2]}{\Pr[A_1]}$$

- Interpretation: if we know (or given that) that A_1 has happened, what is the probability of A_2?
- $\Pr[A_2|A_1] \neq \Pr[A_1|A_2]$
- By definition of conditional probability we have

$$\Pr[A_1 \cap A_2] = \Pr[A_2|A_1] \cdot \Pr[A_1]$$

Important rules

- If $0 < \Pr[A_1] < 1$ then

$$\begin{aligned}\Pr[A_2] &= \Pr[(A_2 \cap A_1) \cup (A_2 \cap A_1^c)] = \Pr[A_2 \cap A_1] + \Pr[A_2 \cap A_1^c] \\ &= \Pr[A_2|A_1]\Pr[A_1] + \Pr[A_2|A_1^c]\Pr[A_1^c]\end{aligned}$$

- More generally, if $S = B_1 \cup \cdots \cup B_k$, $B_i \cap B_j = \emptyset$ if $i \neq j$ and $\Pr[B_i] > 0$ we have

$$\begin{aligned}\Pr[A] &= \Pr[(A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k)] \\ &= \Pr[A \cap B_1] + \Pr[A \cap B_2] + \cdots + \Pr[A \cap B_k] \\ &= \Pr[A|B_1]\Pr[B_1] + \Pr[A|B_2]\Pr[B_2] + \cdots + \Pr[A|B_k]\Pr[B_k]\end{aligned}$$

Important rules (continued)

- $$\Pr[A|B\cap C] = \frac{\Pr[A\cap B\cap C]}{\Pr[B\cap C]} = \frac{\Pr[A\cap B|C]}{\Pr[B|C]}$$

- $$\Pr[A\cap B|C] = \Pr[A|B\cap C]\,\Pr[B|C]$$

- $$\begin{aligned}\Pr[A\cap B\cap C] &= \Pr[A\cap B|C]\,\Pr[C]\\ &= \Pr[A|B\cap C]\,\Pr[B|C]\,\Pr[C]\end{aligned}$$

- and many others

Independence

Definition (Independent events)

Two events A_1 and A_2 are independent if and only if

$$\Pr[A_1 \cap A_2] = \Pr[A_1] \Pr[A_2]$$

Theorem

Two events A_1 *and* A_2 *with* $\Pr[A_1] > 0$ *are independent if and only if*

$$\Pr[A_2 | A_1] = \Pr[A_2].$$

Definition

$A_1, A_2, \ldots, A_n$ are independent events if and only if, given any subcollection $A_{(1)}, \ldots, A_{(k)}$

$$\Pr[A_{(1)} \cap A_{(2)} \cap \cdots \cap A_{(k)}] = \Pr[A_{(1)}] \Pr[A_{(2)}] \cdots \Pr[A_{(k)}]$$

Bayes Theorem

For events A and B with $\Pr[A] > 0, \Pr[B] > 0$ it holds

$$\Pr[A|B] = \frac{\Pr[B|A]\Pr[A]}{\Pr[B]} = \frac{\Pr[B|A]\Pr[A]}{\Pr[B|A]\Pr[A] + \Pr[B|A^c]\Pr[A^c]}$$

In general,

Theorem (Bayes Theorem)

Let $A_1, \ldots, A_n$ be non-overlapping events such that $S = A_1 \cup A_2 \cdots \cup A_n$ and $\Pr[A_j] > 0, \Pr[B] > 0$. Then

$$\Pr[A_j|B] = \frac{\Pr[B|A_j]\Pr[A_j]}{\sum_{i=1}^{n}\Pr[B|A_i]\Pr[A_i]}$$

Counting

Definition (Permutation)

A **permutation** is an arrangement of objects in a definite order. The number of ways we can choose r objects among n distinct objects is

$$n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}$$

Definition (Combination)

A **combination** is a selection of objects disregarding the order. The number of combinations of r objects extracted from n distinct objects is given by

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$