Parametric Statistics

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

October 4, 2022

Lecture Summary

2.1 The Definition of Conditional Probability

2.2 Independent Events

2.3 Bayes' Theorem

- SKIP: 2.4 The Gambler's Ruin Problem

3.1 Random Variables and Discrete Distributions.

Reminder

- Conditional Probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ for} P(B) > 0$$

- Independence:

$$P(A \cap B) = P(A)P(B)$$

$$P(A|B) = P(A)$$

- Bayes Rule

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Law of Total Probability

Partition

Let Ω be a sample space. If $A_1, A_2, A_3, \ldots$ are disjoint and $\bigcup_i A_i = \Omega$ then the collection $A_1, A_2, A_3, \ldots$ is called a partition of Ω.

Law of total probability

If events $B_1, \ldots B_k$ form a partition of hte sample space Ω and $P(B_j > 0$ for all j, then for every event A in Ω

$$P(A) = \sum_{i=1}^{k} P(A \cap B_j) = \sum_{i=1}^{k} P(A|B_j)P(B_j)$$

Bayes Theorem revisited

- Vacc: Yes if vaccinated, zero otherwise
- Hosp: Yes if hospitalized, zero otherwise.
- $P(Hosp|Vacc) = 0.01$
- $P(Hosp|\neg Vacc) = 0.2$
- Three different possibilities: $P(Vacc) = 0.8, 0.5, 0.99$

Let' s use Bayes rule to compute $P(Vacc|Hosp)$ for all three cases.

Independence

Definition

Independence of multiple events A set of events $\{A_i : i \in I\}$ are independent if

$$\mathbb{P}(\cap_{i \in J} A_i) = \prod \mathbb{P}(A_i)$$

for every finite subset J of I.

- If $P(A_i \cup A_j) = P(A_i)P(A_j)$ for every pair $(i,j), i \neq j$ then the events are called pairwisely independent.
- Pairwise Independence does not imply joint independence.
- Example: Tossing a fair coin twice.
  - Event A: Heads on the first toss.
  - Event B: Heads on the second toss.
  - Event C: The two tosses are the same.

Conditional Independence

- A and B are independent given C:

$$P(A \cap B|C) = P(A|C)P(B|C)$$

$$P(A|B, C) = P(A|C)$$

- Conditional Independence does not imply independence.
- Independence does not imply conditional independence.

Random Variables

Random Variable

A random variable is a mapping $X : \Omega \rightarrow \mathbb{R}$ that assigns a real number $X(\omega)$ to each outcome ω.

Example

Consider the experiment of flipping a coin 10 times, and let $X(\omega)$ denote the number of heads in the outcome ω. For example, if $\omega = HHHTHHHTTHT$, $X(\omega) = 6$.

Random Variables

- Why do we need random variables? (easier to work with than original sample space)
- A random variable is NOT a variable (in the algebraic sense).
- A random variable takes a specific value AFTER the experiment is conducted.

Notation

- Letter near the end of the alphabet since it is a variable in the context of the experiment.
- Capital letter to distinguish from algebraic variable.
- Lower case denotes a specific value of the random variable.

Random Variables

- An assignment of a value (number) to every possible outcome.
- Strictly speaking: A function from the sample space Ω to the real numbers.
- It can take discrete or continuous values.
- We will start with discrete values to build our intuition, and then proceed to continuous.

Probability mass function

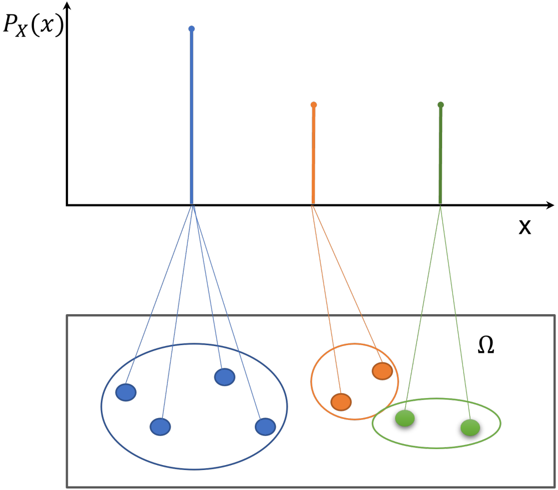
Definition (Probability (mass) function)

If a random variable X has a discrete distribution, the probability function of X is defined as the function p_X such that

$$P_X(x) = \mathbb{P}(X = x) = P(\omega \in \Omega \quad s.t. \quad X(\omega) = x).$$

The closure of the set $x : P_X(x) > 0$ is called the support of the (distribution of) X.

Computing Probability mass functions



Bernoulli distribution

Some distributions come up so often, that they have a name.

Definition

A random variable X that only takes two values 0 (failure) and 1 (success) with $P(X = 1) = p$ has the Bernoulli distribution with parameter p.

Example: You flip a coin one time. X is the result of the coin flip. You can define "Heads" as success and "Tails" as failure (or the other way around, it does not matter) Then X follows a Bernoulli distribution with parameter p.

$$P_X(x) = p^x(1-p)^{1-x}$$

Binomial Distribution

Now imagine you toss the fair coin 5 times. Some possible outcomes are 00000, 1000, 01100 etc. Let X be the number of successes (heads) after the 5 times. Then X follows a Binomial distribution with parameters (5, 0.5).

Definition

The binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent experiments, each with a binary outcome: success (with probability p) or failure (with probability $q = 1 - p$). The pmf of the binomial distribution is

$$P_X(x) = \binom{n}{x} p^x (1-p)^{1-x} \text{ for } x = 0, \ldots, n,$$

Geometric distribution

Now imagine you toss the fair coin until you get heads. times. Some possible outcomes are 1, 01, 0001, $0\ldots1$, etc. Let X be the trial of the first success (heads). Then X follows a Geometric distribution parameter 0.5.

Definition