Probability

Salvatore Andaloro

March 2023

Contents

Chapter 1

Combinatorics

1.1 Permutations

We have n objects, in how many ways we can pick/order them (order matters):

$$n! = n * (n-1) * (n-2) ... * 1$$

We have n objects, but some objects are equal:

$$\frac{n!}{r_1!r_2!...}$$

where $r_1, r_2 ...$ are the number of repetitions of same objects (look at anagram example).

1.2 Dispositions

From a group n, pick k objects (order matters, we can choose each object once):

$$\frac{n!}{(n-k)!} = n * (n-1) * ... * (k+1)$$

Example: 10 objects, 3 slots: $\#dispositions = 10 * 9 * 8 = \frac{10!}{7!}$
From a group n, pick k objects (order matters, we can choose each object multiple times):

$$n^k = n * n * ... * n \text{ (repeated k times)}$$

1.3 Combinations

From a group n, pick k objects (order doesn't matter, we can pick each object once):

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

From a group n, pick k objects (order doesn't matter, we can choose each object multiple times):

$$\frac{(n+k-1)!}{k!(n-1)!}$$

1.4 Examples

Multiply possibilities of first case with those of second case etc.

Example. Un dipartimento di statistica decide di assegnare ai propri 25 laureati tre premi di diversa tipologia. Se ciascuno dei laureati potesse ricevere al massimo un premio, quante assegnazioni sarebbero possibili?
$\#E = 25 * 24 * 23$

First permutate outer group, then inner group.

Example. Il Signor Amadori deve sistemare 10 libri in un ripiano della scaffalatura. Quattro libri sono di matematica, tre sono di chimica, due sono di storia e uno è di grammatica. Amadori, che è un tipo ordinato, vuole fare in modo che i libri sullo stesso argomento siano vicini in libreria. In quanti modi ciò si può realizzare?
$\#E = 4! * 4! * 3! * 2!$

Anagrams: $\#E = \frac{\text{(num. of letters)!}}{\text{(num. of repeted letter A)!}*\text{(num. of repeted letter B)!}}$

Example. Quanti sono gli anagrammi di PEPPER?
$\#E = \frac{6!}{3!*2!}$

Pick k elements in n. Order doesn't matter.

Example. Dieci ragazzi devono formare 2 squadre A e B di 5 membri ciascuna. Quante sono lo suddivisioni possibili?
$\#E = \binom{10}{5} = \frac{10!}{5!5!}$

Chapter 2

First steps into probability

Classical definition of probability:

$$P = \frac{\#\text{possible cases}}{\#\text{total cases}}$$

Only works with finite cases and equally probable cases.

Definition 2.0.1. An experiment is **random** if the outcome, given the initial configuration is uncertain.

Definition 2.0.2. Ω - **sample space**: the results, pairwise incompatible, of a random experiment.

Definition 2.0.3. $\mathcal{P}(\Omega)$ - **power set** of Ω: set of all subsets of Ω. Cardinality is $2^{\#\Omega}$ (same cardinality of when we have n bits in binary (ex. if we have 4 bits then we have 2^4 possible numbers), since we can think that 0 represents that we don't take the element, 1 if we take it).

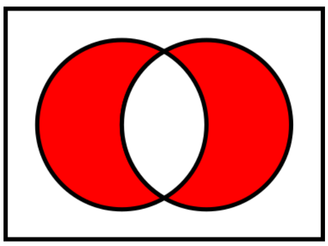
2.1 Algebras and probability spaces

Definition 2.1.1. $\mathcal{F}$ - **algebra**: family of subsets where the following holds:

1. $\Omega \in \mathcal{F}$
2. if $A \in \mathcal{F}$ then $A^C \in \mathcal{F}$
3. (finite case) if $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$

Properties of $\mathcal{F}$:

1. $\emptyset \in \mathcal{F}$
2. if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$



$A \Delta B$

3. if $\{A_i\}_{i=1}^{n} \subseteq \mathcal{F}$, then $\cap_{i=1}^{n} A_i \in \mathcal{F}$
4. if $A, B \in \mathcal{F}$, then $A \setminus B \in \mathcal{F}$
5. if $A, B \in \mathcal{F}$, then $A \Delta B \in \mathcal{F}$ (Δ = elements present only in A or B, equivalent is $(A \cup B) \cap (A^C \cup B^C)$)

Definition 2.1.2. σ-algebra: same algebra as before, but also infinite unions are defined. Properties of σ-algebras:

1. $\Omega \in \mathcal{F}$
2. if $A \in \mathcal{F}$ then $A^C \in \mathcal{F}$
3. for every countable family $\{A_i\}_{i=1}^{+\infty}$ of subsets of Ω, if all the sets A_i are in $\mathcal{F}$, then $\cup_{i=1}^{+\infty} A_i \in \mathcal{F}$

Since σ-algebras accept also finite unions, all algebras are σ-algebras.

Example. Difference between normal algebras and σ-algebras.
$\Omega = \mathbb{N}$ $\mathcal{A}$=$\{A \subseteq \mathbb{N}$: A is finite or A^C is finite$\}$
Let $A \in \mathcal{A}$ and $B \in \mathcal{A}$.
If both finite, also union is finite => element of $\mathcal{A}$. Now, let's look at numbers 2n. For any n, $2n \in \mathcal{A}$ as it is finite. But $\cup_{i=1}^{+\infty} 2n \notin \mathcal{A}$ => infinite union not contained in $\mathcal{A}$=> not a σ-algebra.

Definition 2.1.3. E - event: every element $E \in \mathcal{F}$ ($\mathcal{F}$ is a σ-algebraon Ω). Singletons are elementary or atomic events.

Example. Let $\Omega = \{a, b, c\}$.
Then we can define our σ-algebra as $\mathcal{F} = \{\emptyset, \{a\}, \{b, c\}, \{a, b, c\}\}$
$\{a\}$ is an atomic event
$\{b\}$ is not an event
$\{a, b, c\}$ is an event, but not atomic
Notice that we have checked all 3 properties of a σ-algebra: we have Ω and all complements.

Definition 2.1.4. Given a set Ω and a σ-algebra $\mathcal{F}$ on Ω, the pair $(\Omega, \mathcal{F})$ is a **measurable space** or **Borel space**.

Definition 2.1.5. Given a measurable space $(\Omega, \mathcal{F})$, a function $P : \mathcal{F} \to \mathbb{R}$ is a probability measure or probability function if it satisfies the following properties (Kolmogorov's axioms):

1. For every event E, $P(E) \geq 0$ (non negativity)
2. $P(\Omega) = 1$ (normalization or total mass)
3. given a countable family $\{E_i\}_{i=1}^{+\infty}$ of pairwise disjoint events, $P(\cup_{i=1}^{+\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$ (σ-additivity)

The probability of E is the value P(E).

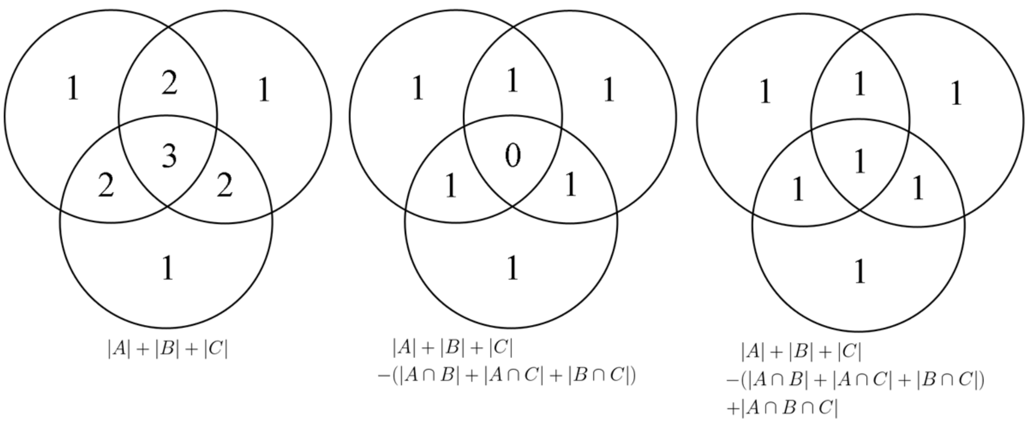
Definition 2.1.6. Let Ω be a set, $\mathcal{F}$ a σ-algebra on Ω, P a probability function on $\mathcal{F}$. The triple $(\Omega, \mathcal{F}, P)$ is a **probability space**.

Properties of probability measures:

1. $P(\emptyset) = 0$
2. if $E \in \mathcal{F} \Rightarrow P(E^C) = 1 - P(E)$
3. Let E, F events s.t. $E \subseteq F$. Then $P(E) \leq P(F)$
4. Image of any probability function is in unit interval $[0, 1]$
5. $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
6. $P(E \cup F) \leq P(E) + P(F)$

2.1.1 Inclusion-exclusion principle

We can extend the notion of point 5 to the union of any number of sets. This is known as the inclusion-exclusion principle. The idea is that we have to remove all elements that we have counted twice, add elements that we have removed three times etc.



Counting elements using the inclusion-exclusion principle with 3 sets

Proposition 2.1.1. Let $\{E_i\}_{i=1}^n \subseteq \mathcal{F}$ a finite family of sets. Then

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i) - \sum_{i<j} P(E_i \cap E_j)^1 + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) + \cdots + (-1)^{n+1} P\left(\bigcap_{i=1}^n E_i\right)$$

Remark. We can estimate from above (stopping at odd intersections) or below (stopping at even intersections). These are called Bonferroni bounds.

[1] $\sum_{i<j} = \sum_{i=1}^n \left(\sum_{j=i+1}^n P(E_i \cap E_j)\right)$

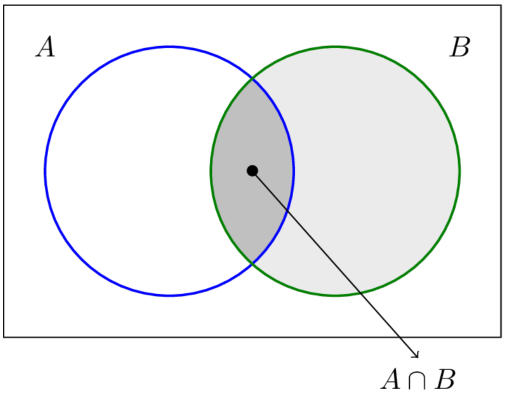
2.2 Conditional probability

It is possible to extend the notion of probability spaces by adding conditions to our events.

Definition 2.2.1. Given a probability space $(\Omega, \mathcal{F}, P)$ and two events E, F in $\mathcal{F}$ with $P(F) \neq 0$, we define the probability of E conditional to F ("E given F" per gli amici) as

$$P(E|F) := \frac{P(E \cap F)}{P(F)}$$

WARNING! $P(E|F)$ is not the same thing as $P(E \cap F)$! $P(E|F)$ denotes the probability of the intersection ONLY on the F set, while $P(E \cap F)$ denotes the probability on the whole Ω!



Remark. $P_F(\cdot) = P(\cdot|F)$ is a probability function, since it satisfies Kolmogorov axioms. Therefore also P_F is a probability measure on the Borel space $(\Omega, \mathcal{F})$, that is in general different from P.

Remark. Product rule

$$\begin{aligned} P(E|F) &= \frac{P(E \cap F)}{P(F)} \\ P(E \cap F) &= P(E|F)P(F) \\ &= P(F|E)P(E) \end{aligned}$$

2.2.1 Independence

Let E be the event that it rains tomorrow, and suppose that $P(E) = \frac{1}{3}$. Also suppose that I toss a fair coin; let F be the event that it lands heads up. We have $P(F) = \frac{1}{2}$. Now I ask you, what is $P(E|F)$? What is your guess? You probably guessed that $P(E|F) = P(E) = \frac{1}{3}$. You are right! The result of my coin toss does not have anything to do with tomorrow's weather. Thus, no matter if F happens or not, the probability of E should not change. This is an example of two independent events. Two events are independent if one does not convey any information about the other.

Definition 2.2.2. In a probability space $(\Omega, \mathcal{F}, P)$, two events E, F in $\mathcal{F}$ are **independent** (with respect to P) if the following holds: $P(E \cap F) = P(E) \cdot P(F)$. Sometimes the notation $E \perp F$ is used in this case.

Now, let's first reconcile this definition with what we mentioned earlier, $P(E|F) = P(E)$. If two events are independent, then $P(E \cap F) = P(E)P(F)$, so

$$\begin{aligned} P(E|F) &= \frac{P(E \cap B)}{P(F)} \\ &= \frac{P(E)P(F)}{P(F)} \\ &= P(E). \end{aligned}$$

An intuitive question we can ask ourselves is: is the probability of E happening the same as E happening after F? If that is the case, then the events are independent. Going back to the rain and coin case, the probability of getting heads is the same as the probability of getting heads after raining.

Example. We have a $d6$. $E = \{2, 4, 6\}$ (getting an even number), $F = \{3, 6\}$ (getting a multiple of 3). Are these events independent?

$$\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6} = P(E)P(F) = P(E \cap F) = \frac{1}{6}$$

Yes, they are. If we ask our "intuitive question", the probability of F happening after E is $\frac{1}{2}$, but also the probability of E happening after F is $\frac{1}{2}$, therefore the two events are independent.

Now we can of course (to be read with a with a pedantic British accent) extend this notion to more than to sets. For example, three events A, B and C are independent if all of the following conditions hold:

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \\ P(A \cap B \cap C) &= P(A)P(B)P(C) \end{aligned}$$

Now we can apply what we have seen with 3 sets to any number of sets.

Definition 2.2.3. In a probability space $(\Omega, \mathcal{F}, P)$, the events $E1, ..., E_n$ are independent (with respect to P) if for any choice of indices (without repetition) $i_1, ..., i_m$ in $\{1, ..., n\}$ (with $m \leq n$) it holds

$$P\left(\bigcap_{j=1}^{m} E_{i_j}\right) = \prod_{j=1}^{m} P(E_{i_j}).$$

2.2.2 Law of total probability/factorisation formula

Theorem 2.2.1. Let $(\Omega, \mathcal{F}, P)$, $\{E_i\}_{i=1}^n$ disjoint, $P(E_i) > 0\ \forall i$, $\bigcup_{i=1}^n E_i = \Omega$

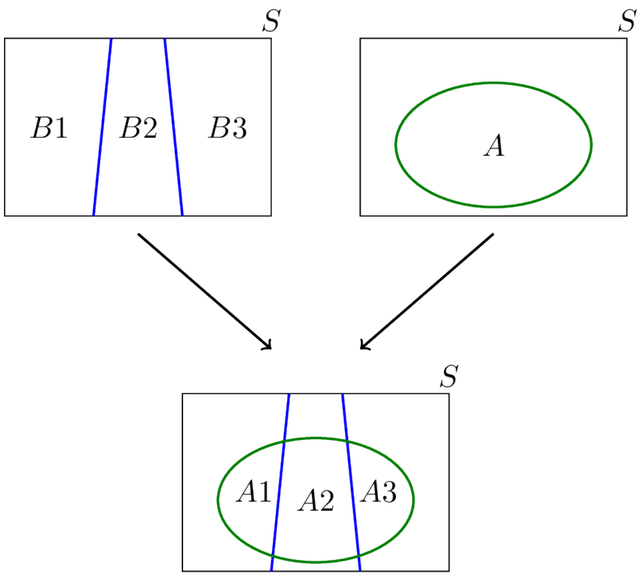
$$\forall E \in \mathcal{F} \quad P(E) = \sum_{i=1}^n P(E \cap E_i) = \sum_{i=1}^n P(E|E_i)P(E_i)$$

Using a Venn diagram, we can pictorially see the idea behind the law of total probability. In the next figure, we have

$$A_1 = A \cap B_1$$
$$A_2 = A \cap B_2$$
$$A_3 = A \cap B_3$$

As it can be seen from the figure, A_1, A_2, and A_3 form a partition of the set A, and thus by the third axiom of probability

$$P(A) = P(A_1) + P(A_2) + P(A_3).$$



Law of total probability

2.2.3 Bayes theorem

Theorem 2.2.2. Let $(\Omega, \mathcal{F}, P)$ be a probability space and E, F two events, both with non-zero probability. Then

$$P(E|F) = \frac{P(F|E)}{P(F)} \cdot P(E).$$

Proof:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E \cap F)}{P(E)} \cdot \frac{P(E)}{P(F)} = \frac{P(F|E) \cdot P(E)}{P(F)}$$

Chapter 3

Random variables

A random variable is a propriety of the outcome of the experiment