Graph - Lezione 1 - 28/10/2019

sabato 2 novembre 2019 12:02

Introduction

Some Element of theory and combinatory

Discrete time Markov chain --> discrete in time and space and useful to model movements on a graph.

Markov chain to analyse random graph to represent for example social networks.

Probability space

Probability space

Let Ω be any set and let Σ be some "appropriate" class of subsets of Ω.

Elements of Σ are called events.

For $A \subseteq \Omega$ we write A^C for the complement of A in Ω, i.e.

$$A^C = \{s \in \Omega : s \notin A\}.$$

The randomness of the experiment is summerize by Omega. Omega is defined a family of class of appropriate subset of Omega. Many cases sigma is all subset of Omega. In general is important that the family (sigma) not only includes the event but also the complements of the event A^C.

Definition

A probability measure on Ω is a function $P : \Sigma \to [0,1]$, satisfying

1. $P(\emptyset) = 0$.
2. $P(A^C) = 1 - P(A)$ for any event A.
3. If A and B are disjoint events (that is if $A \cap B = \emptyset$), then $P(A \cup B) = P(A) + P(B)$. More generally, if $A_1, A_2, \dots$ is a countable sequence of disjoint events ($A_i \cap A_j = \emptyset$, for any $i \neq j$), then $P(\bigcup_{t=1}^{\infty} A_t) = \sum_{t=1}^{\infty} P(A_t)$.

Note that the first two conditions imply that $P(\Omega) = 1$.

The triple (Ω, Σ, P) is called probability space.

Is a function from the probability function from 0 to 1. Probability of empty set should be zero --> nothing is happening.
P(complement of event) = 1 - P (event).
If A and B are disjoint event the probe of the union is the sum of the prob of the single event. The sum should be good even for infinity disjoint events.

Conditional Probability

Conditional probability

Definition

If A and B are events, and $P(B) > 0$, we define the conditional probability of A given B as

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

Interpretation: $P(A|B)$ is how likely we consider that A happens, knowing that B happened

How likely we expect realization of A knowing B.

Example

$A =$ Tomorrow here will rain
$B =$ Today a storm occurred 100 Km on the west of my position

If I don't know anything about weather forecast or conditions in the surrounding (and I don't know if B occurred) I can only guess that $P(A) = P(\text{tomorrow here will rain}) = \frac{1}{2}$.

But if I know that B happened, it becomes more likely that tomorrow here will rain, thus $P(A|B) > \frac{1}{2}$.

Independence

Definition

Two events A and B are said to be independent if

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

More in general

Definition

The events $A_1, \dots, A_k$ are said to be independent if for any $l \leq k$ and any $i_1, \dots, i_l \in \{1, \dots, k\}$ with $i_1 < i_2 < \dots < i_l$ we have

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_l}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_l}).$$

The probability of the intersecion of the two events is the product of the probability.

Why is related to the conditional probability? Since we know the prob of A|B so P(A|B) = P(A)

Note that if A and B are independent, then, since $P(A|B) = \frac{P(A \cap B)}{P(B)}$, we have

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

Then

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

Example.
$A =$ Tomorrow here will rain
$B =$ Today I make a cake
A is not influenced by B and viceversa, thus they are independent and $P(A|B) = P(A)$.

If A and B are independent the probability of intersection is the probability of the two intersection. If A and B are independent this does not modify the probability of A.

In practice, in particular if the space Ω is finite, we compute the probability of an event A as

$$P(A) = \frac{\#\ \text{cases in favor of } A}{\#\ \text{possible cases}}$$

The correct counting of cases is the subject of combinatorics.

Combinatorics: counting problems

[C.M.Grinstead, J.L.Snell, Introduction to Probability, AMS publisher, 1997 - Chapter 3]

Consider an experiment that takes place in several stages and is such that the number of outcomes m at the nth stage is independent of the outcomes of the previous stages.
The number m may be different for different stages.

We want to count the number of ways that the entire experiment can be carried out.

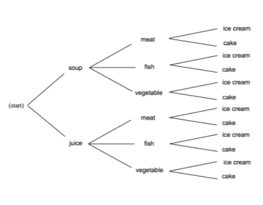
An experiment in several stages. The outcomes of m is indepedent of the outcomes of the previos stage. We wanto to count the number of way the number of experiments can be go on.

Example 1. You are eating at Emile's restaurant and the waiter informs you that you have

(a) two choices for appetizers: soup or juice;
(b) three for the main course: a meat, fish, or vegetable dish;
(c) two for dessert: ice cream or cake.

How many possible choices do you have for your complete meal?

We can represent this concept with a three.



Your menu is decided in three stage. At each stage the number of possible choices does not depend on what is chosen in the previous stages: two choices at the first stage, three at the second, and two at the third.
From the tree diagram we see that the total number of choices is the product of the number of choices at each stage.
In this example we have $2 \cdot 3 \cdot 2 = 12$ possible menus.

At each stage you have different number of possible choises. So, number m is different from different stages. Counting the number of leaves you have all the possible menus that you can compose. If you want to count them you have to multiply 2 appetizer * 3 main course * 2 desert = 12. In general it's a good procedure and we can generalize this rules.

Our menu example is an example of the following general counting technique:

Counting technique. A task is to be carried out in a sequence of r stages. There are n_1 ways to carry out the first stage; for each of these n_1 ways, there are n_2 ways to carry out the second stage; for each of these n_2 ways, there are n_3 ways to carry out the third stage, and so forth. Then the total number of ways in which the entire task can be accomplished is given by the product

$$N = n_1 \cdot n_2 \cdot \dots \cdot n_r$$

If the stage are indepedent the total number is given by the product of the number of ways in each step.

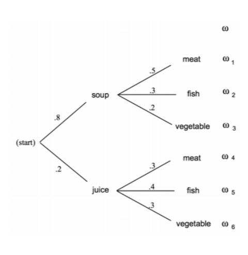
Tree diagrams

It will often be useful to use a tree diagram when studying probabilities of events relating to experiments that take place in stages and for which we are given the probabilities for the outcomes at each stage.

Example 1: consider only appetizers and main course, and assume that the owner of Emile's restaurant has observed that 80% of his customers choose the soup for an appetizer and 20% choose juice.
Of those who choose soup, 50% choose meat, 30% choose fish, and 20% choose the vegetable dish.
Of those who choose juice for an appetizer, 30% choose meat, 40% choose fish, and 30% choose the vegetable dish.

We represent these probabilities on the tree diagram.

Why can be useful to introduce these three diagrams? Using the example you can imaging that the owner of the restaurant want toforget about the dessert. The owner observe that the 30% chosse appetizer and the other 30% choose soup. We can represent this example with a three diagram.

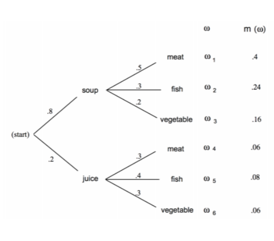


We choose for our sample space the set $\Omega = \{\omega_1, \dots, \omega_k\}$ of all possible paths along the tree.

Question: what is the probability that a customer chooses first soup and then meat?

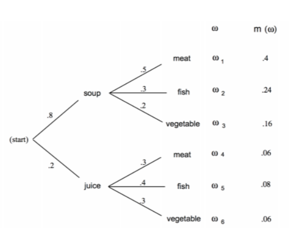
These means that we have 6 possible compositions of our meal. 6 possible outcomes that we label with the symbol w1, ..,wk.

We have to moltiply 0.8 * 0.5 to get soup and meat.



Answer: $P([1st\ soup]) \cdot P([2nd\ meat|1st\ soup]) = 0.8 \cdot 0.5 = 0.4$

This suggests choosing our probability distribution for each path through the tree to be the product of the probabilities at each of the stages along the path.



Note that $\sum_{i=1}^{6} m(\omega) = 1$. And if we want to know the probability that a customer eats meat, whatever appetizer he/she gets, we sum the probabilities of having meat:

$$P(meat) = 0.4 + 0.06 = 0.46$$

So we have conditional probability

SECOND Part

If we wanto only to count our possible outcome we can introduce an intstrument for the counting which is a lot automatical. So first of all, image that we have a finite set A with a finite number of elements an we want to compute the different reordering of the elements of A. Which mean that for example our set of three elements, we want to count the element reorder in different sequences. We are counting the permutations of the elements of A. If A scompose by 3 elements the possible permutation are 6.

PERMUTATIONS

DEFINITION. Let A be a finite set, composed by n distinct elements.
Suppose to order the elements of A.
A permutation is a reordering of the elements of A.

Example: $A = \{a, b, c\}$

permutations:
abc
acb
bca
bac
cab
cba

6 permutations are possible

If you want to compute a set of n element we can count n! permutation. Positioning an element we have to subtract it, infact we got then n-1, n-2 ecc till 1.

In general, if A is composed by n objects, how many permutations can we form with its elements?

elements: 1 2 3 --- n-1 n

n. of ways in which we can fill the position: n (n-1) (n-2) --- 2 1

Thus the number of permutations of n different objects is

n·(n-1)·(n-2)·...·2·1 = n!

n factorial = product of the first n integers

In our case now we want to compute only a subset of k of them.

Assume now to have n different objects but you want to compute the permutations only of a subset of k of them.

Example: in genetics, DNA is composed by a sequence of the 4 bases A, G, C, T. Triplets, that is sequences of 3 bases, are responsible of the synthesis of amino acids. How many different triplets can be formed?

triplet: ☐ ☐ ☐

ways to be filled: 4 3 2

N. of triplets $= 4 \cdot 3 \cdot 2 = 24$

In general: if we have n different elements and we want to fill k (<n) positions, we can fill them in

position: 1 2 3 --- k-1 k

ways to fill the positions: $n \cdot (n-1) \cdot (n-2) \cdots (n-(k-2)) \cdot (n-(k-1))$

Thus the number of permutations of k objects chosen among n is

$$n(n-1)(n-2) \cdots (n-k+2)(n-k+1) = P_{n,k}$$

that is the product of k decreasing integers starting from n

Pn,k can be written in another way that is more convenient to remember.

Note that we can rewrite $P_{n,k}$ as follows:

$$P_{n,k} = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1) = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1) \cdot (n-k) \cdot (n-k-1) \cdot \dots \cdot 2 \cdot 1}{(n-k)(n-k-1) \cdot \dots \cdot 2 \cdot 1} = \frac{n!}{(n-k)!}$$

$$= \frac{\text{n. of permutations if we had } n \text{ "boxes to fill"}}{\text{n. of permutations in the "missing boxes"}}$$

If we consider an example with some elements are repeted.

Suppose finally that we want to count the permutations of the elements of a set A, which contains repeated elements

Example: $A = \{a, a, b, c\}$

The sequences

baca
baca

are the same, since we only switched the position of two repeated elements.
Then the number of different permutations of the 4 elements of A here is

$$\frac{\text{N. of permutations of all elements of A}}{\text{N. of permutations of the repeated elements}} = \frac{4!}{2!}$$

General rule:

If there are n objects of which n_1 are equal, n_2 are equal, ... , n_k are equal, the number of different permutations of the n objects is

$$\frac{n!}{n_1!\, n_2! \dots n_k!}$$

Example: the number of anagrams of the word

MISSISSIPPI

is

$$\frac{11!}{\underbrace{4!}_{I}\,\underbrace{4!}_{S}\,\underbrace{2!}_{P}} = 34650$$

Number of I S P

How would you compute this number with a pocket calculator?

$$\frac{11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{4 \cdot 3 \cdot 2 \cdot 1 \quad 4 \cdot 3 \cdot 2 \cdot 1 \quad 4 \cdot 3 \cdot 2 \cdot 1}$$

WE PUT SOME PARENTESYS AND THEN WE AVOID OVERFLOW

Combination --> let A be a set of composed different object
Object of A are not repeted but they are different. This regarding the order. Combination of k object chosen amoung n.

COMBINATIONS

Let A be a set composed by n different objects. Assume that we want to count how many sequences of k objects can be composed by the elements of A, but now two sequences are different only if they contain different elements, and not if they contain the same elements with a different order.
We call these sequences combinations of k objects chosen among n: $C_{n,k}$

EXAMPLE.
In a class there are 12 students and the school has a laboratory which contains only 6 students at a time. Thus the teacher has to organize two working groups to attend the laboratory, Group A and Group B.

In how many different ways Group A can be formed?
(Note that Group B is automatically formed once you define the composition of Group A)

In this way, we have that two composition of the two group are different if they are composed by different student.
They way we obtain this numbers is to compute the number of permutation divided by the number of object disregarding the order.

Clearly two compositions of Group A are different only if they are formed by different students, and not by the way the teacher lists their names. Thus we are counting the combinations of 6 elements among 12

$$C_{12,6}$$

In order to obtain this number we can compute the number of permutations of 6 objects chosen among 12, and then divide by the number of sequences of 6 objects which differ only for the order of their elements, that is 6!

$$C_{12,6} = \frac{P_{12,6}}{6!} = \frac{12!}{(12-6)!\,6!} = 924$$

So we can get a general rule.

General rule:

The number of combinations of k objects chosen among n is given by

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

Binomial coefficient

What is an efficient way to compute $C_{n,k}$ with a pocket calculator or a computer?

The same strategy that we already introduce. Compute the ratio 2 by 2.

Esercizi

URN WITH 4 WHITE BALLS AND 3 BLACK BALLS
WITHDRAW WITH REPLACEMENT 2 BALLS

1) PROB THAT 2 BALLS HAVE THE SAME COLOUR
2) PROB AT LEAST 1 OF TWO BALL IS BLACK

P(2 SAME COLOUR) WE HAVE TO COUNT EVENTS IN FAVOUR

$A = \{w_1, w_2, w_3, w_4, b_1, b_2, b_3\}$

$\Omega = \{(B_1, B_2),\ B_1, B_2 \in A\}$ SET OF POSSIBLE OUTCOMES

N. OF POSSIBLE CASES ⇒ NUMBER OF ELEMENTS OF THIS SET

N. OF POSSIBLE CASES = N. ELEMENTS OF $\Omega = 7 \cdot 7 = 49$

$(\ , \ )$
$\uparrow \ \uparrow$
7 7

POSSIBLE WAYS

a) WE WANT TO COUNT THE FAVORABLE CASES OF "TWO BALLS OF THE SAME COLOR"

(w, w) (b, b)
4 4 3 3
(REPLACEMENT)

COUPLE $(w,w) = 16$ COUPLE $(b,b) = 9$

$16 + 9 = 25$

$P(\text{2 BALL OF SAME COLOUR}) = \frac{25}{49}$

b) FAVORABLE CASES WITH "AT LEAST ONE BALL IS BLACK"

(b,b) (w,b) (b,w)

$3 \cdot 3 + 4 \cdot 3 + 3 \cdot 4 = \frac{33}{49}$

ANOTHER SOLUTION
CONSIDER PROBABILITY OF THE COMPLEMENTS
P(AT LEAST ONE BALL IS BLACK) = 1 - P(BOTH BALLS WHITE)

(w, w)
$4 \cdot 4$

$$P \Rightarrow 1 - \frac{16}{49} = \frac{49-16}{49} = \frac{33}{49}$$

ESERCIZIO 2

→ BALL WITHOUT REPLACEMENT

a) (w, w) (b, b)
$4 \cdot 3 + 3 \cdot 2$

POSS. CASES $= 7 \cdot 6 = 42$

$P(a) = \frac{18}{42}$

b) (w, w) → $\frac{12}{42}$
$4 \cdot 3$

$P(b) = 1 - P(w,w) = 1 - \frac{12}{42} = \frac{30}{42}$

ESERCIZIO 3

5 BUTTON 1, 2, 3, 4, 5
COMBINATION CHANGE EVERY WEEK

a) HOW MANY COMBINATIONS OF THE LOCK CAN BE FORM PRESSING EACH BUTTON ONLY ONES

b) PRESS TWO DIFFERENT BUTTONS AT THE TIME PLUS 3 SINGLE BUTTONS.
HOW MANY MORE COMBINATIONS

a) PERMUTATION OF THE SET $\{1, \dots, 5\} = 5!$

$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

b)

FIRST TWO BUTTON CAN BE CHOSEN IN $C_{5,2} = \binom{5}{2} = \frac{5!}{3!\,2!} = \frac{5 \cdot 4}{2} = 10$

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

LAST 3 IS BUTTONS CAN BE CHOSEN IN 3!

$10 \cdot 3! = 10 \cdot 6 = \boxed{60}$

c) 3 BUTTON CAN BE THE SAME

FIRST TWO $C_{5,2} = 10$

LAST 3 $= 5^3$

$5^3 \cdot 10$