Ideas in probability, counting and Bayes theorem

Iván Andrés Trujilllo Abella

ivantrujillo1229@gmail.com

Probability

The sample space (Ω) represent the total possibles outcomes in a experiment. An **event** (A_i) is a subset of Ω $(A \subseteq \Omega)$ and the number of favorable cases to the event is (a_i).

Rolling dice - Probability of get an even number

$$\begin{aligned} \Omega &= \{1, 2, 3, 4, 5, 6\} \\ A_i &= \{2, 4, 6\} \end{aligned} \tag{1}$$

Probability

The sample space (Ω) represent the total possibles outcomes in a experiment. An **event** (A_i) is a subset of Ω $(A \subseteq \Omega)$ and the number of favorable cases to the event is (a_i).

Rolling dice - Probability of get an even number

$$
\begin{aligned}
\Omega &= \{1, 2, 3, 4, 5, 6\} \\
A_i &= \{2, 4, 6\}
\end{aligned}
\tag{1}
$$

...

$$
P(A_i) = \frac{n(A_i)}{n(\Omega)} = \frac{3}{6} \tag{2}
$$

Probability

Naive Definition

The probability that occur the event A_i is equal to $P(A_i) = \frac{a_i}{\Omega}$ we can see that $\Omega = n(\Omega)$ is the number of favorable and unfavorable cases.

Note that we refer to the number of elements, specifically to the cardinally of the sets, then $a_i = n(A_i)$.

Probability

To compute

- What is the probability of get 12 in two dice rolling
- What is the probability of get two tails flipping two times a coin

Those above listed problems are counting problems!

Multipliation rule

...

If one action can be make it of p ways and later other can be make it of q ways then the total numbers of ways that the actions can make it jointly is p.q

There are "q" ways to carry out the second action, by to "p" ways of make the first action, therefore there are $p.q$ total ways of carry out both actions.

Paiting doors

if you have 4 doors and 3 different colors of paint, how many ways you can painted the doors?

Paiting doors

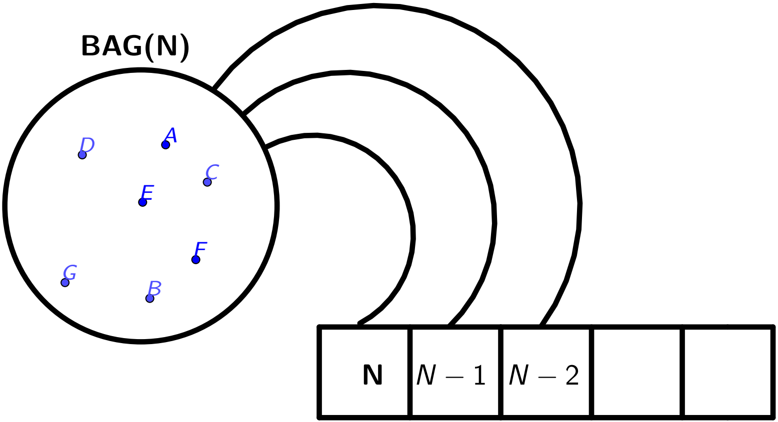
if you have 4 doors and 3 different colors of paint, how many ways you can painted the doors?

Bag model

for n marbles and drawing k of them one by one, how many arragments we can get?... the first time we can get n marbles, after $(n-1)$, the third time $(n-2)$ thus the $k-time$ we have $(n-k+1)$ ways of drawing one.

$$p(n,k) = \frac{n!}{(n-k)!} = n(n-1)(n-2)...(n-k+1) \tag{3}$$

Bag model ilustration



...

How many possibles arrangements are result of drawing two marbles in a bag with letters (a,b,c).

first time we can get three ways of drawing a marble, there are two ways of draw the second marble (don't insert the marble again - without replacement) multiplication rule total of possibles results are six:

...

How many possibles arrangements are result of drawing two marbles in a bag with letters (a,b,c).

first time we can get three ways of drawing a marble, there are two ways of draw the second marble (don't insert the marble again - without replacement) multiplication rule total of possibles results are six:

Arragements	(a,b)	(a,c)	(b,a)	(b,c)	(c,a)	(c,b)

Permutations

...

$$
\begin{aligned}
n! &= n(n-1)! \\
\frac{n!}{n-1!} &= n \\
\frac{n}{n-2!} &= n(n-1) \\
\frac{n}{n-3!} &= n(n-1)(n-2)
\end{aligned} \tag{4}
$$

The generalization:

Formula

$$p(n,k) = \frac{n!}{(n-k)!} = n(n-1)(n-2)...(n-k+1) \tag{5}$$

Combinations

in a permutation there is not replacement and order matter, However for each arrangement of k longitude have $k!$ ways of ordered therefore the number of possible ways of different total elements it is a combination.

Formula

$$C(n, k) = \frac{P(n, k)}{k!} = \frac{n!}{(n - k!)k!} \tag{6}$$

Binomial coefficient

Note that combinations are also noted as follow:

$$C(n, k) = \binom{n}{k} \tag{7}$$

Summary

We need consider the following in counting problems:

...

- Repetition
- order

	order	not-order
Replacement	n^k	$\binom{n+k-1}{n-1}$
Without replacement	$\dfrac{n!}{(n-k)!}$	$\binom{n}{k}$

Combinatorial identity

Using the following binomial theorem:

$$(a+b)^n = \sum_{i=0}^{n} a^{n-i} b^i C(n,i) \tag{8}$$

$$(1+1)^n = \sum_{r=0}^{n} C(n,r) = 2^n - C(n,0) = \sum_{r=1}^{n} C(n,r) \tag{9}$$

...

$$2^n - 1 = \sum_{r=1}^{n} C(n,r) \tag{10}$$

What is the practical importance of this identity?

Philosophy!

...

The problem with **naive definition;** is that require that all events are equally likely...

No naive defintion

probability space, a new concept, that include a function f such that:

$$x \subseteq \Omega; \quad f(x) \in [0, 1] \tag{11}$$

Assumptions

According to the literature $f() = P()$.

...

- $P(A_i) \in [0,1] \quad \forall A_i$.
- $P(\emptyset) = 0$ the probability that something that never will happened is equal to zero.
- $P(\Omega) = 1$ the probability of something that always occur is equal to one.
- $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$ if and only if $\bigcap_{i=1}^{n} A_i = \emptyset$.

Kolgomorov approach

The certainty, it is one (1), for instance; the following statement *A person is live or dead it is certainty*

$$P(A \cup A^c) = 1 \tag{12}$$

Could be a natural property that two two two events mutually exclusive could be sum up their probabilities.

$$P(A \cup B) = P(A) + P(B) \tag{13}$$

You can extend this properties to n events.

Probability

there are two main theorems that we must consider and allow us calculate probabilities.

Theorem (Multiplication rule)

if the events $A_1, A_2, A_3, ..., A_n$ are independent each one then the probability that they occur is $\prod_{i=1}^{n} P(A_i)$

If there are $a_1, a_2, a_3, ..., a_n$ favorable cases and $\Omega_1, \Omega_2, \Omega_3, ...\Omega_n$ possible outcomes to the events $A_1, A_2, A_3, ..., A_n$ respectively then by **??** the are $\Omega_1\Omega_2\Omega_3...\Omega_n$ in total possible outcomes, and there are $a_1 a_2 a_3 ... a_n$ of get a favorable case therefore by naive definition we have

$$P = \frac{a_1 a_2 a_3 ... a_n}{\Omega_1\Omega_2\Omega_3...\Omega_n} = P(A_1)P(A_2)P(A_3)...P(A_n)$$

Independece

Independence not is equal not occurrence, the independence is related with the change of the probability that occur a event given that another occur.

$$P(A \cap B) = P(A)P(B) \tag{14}$$

Tossing coins it is a bernoulli trial, and the occurrence of a event not affect the another event.

Birthday paradox

Programming and mathematics are practical. Statistics it is a field in which we can interact with real data, the **birthday paradox** it is a practical example of this.

If we have k persons then the probability that two of them born in the same day is; if the year have 360 days:

$$1-\frac{360.359.358...(360-k+1)}{360^k}$$

In the case of :

$$1-\left(\prod_{i=311}^{360} i(360^{-50})\right)\approx 0,97 \tag{15}$$

is 97% is a high probability that at least two person born the same day.

The following Figure 1 show the relation among the probability of two coincidences among two person in the sample.

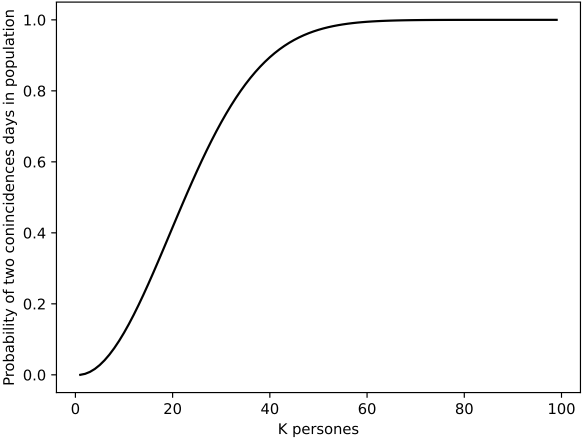


Figure: Birthday paradox with 360 days

Contingency table

		Diagnose	
		Disease	No-Disease
Risk Factor	Smoke	a	b
	Not Smoke	c	d

What it is $P(Disase \mid Smoke) = \frac{a}{a+b}$ note that marginal distribution. Note that $P(Disease \cap Smoke) = \frac{a}{(a+b+c+d)}$. Also note that $P(Smoke) = \frac{a+b}{(a+b+c+d)}$. Note the result of divide the last two probabilities.

Conditional Probability

The probability of event given a "information". *Probability of A occur given B occurs.*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{16}$$

Note that $P(A \cap B)$ it is equal to $P(B \cap A)$.

$$P(A \mid B)P(B) = P(B \mid A)P(A) \tag{17}$$

Conditional probability

We can make the following lecture of conditional probability

$$P(A \mid B) \tag{18}$$

What is the probability of A occur given that B already happened.

Permutation and combinations

For what is used statistics?

Bayes Theorem

Uses

- You could uses to filter spam!
- Predict if a person pay or not a credit
- an a lot of examples...

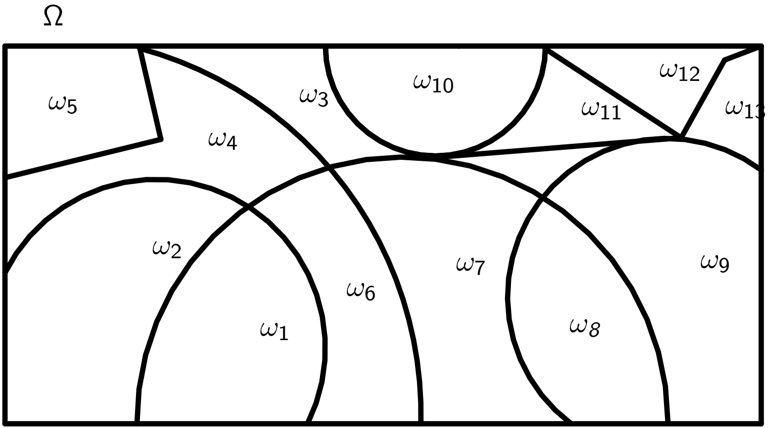
We need?

- Conditional probability
- Law of total probability

Law of total probability

The sample space defined as Ω if we split omega in ω k subsets in order that each *subset* no overlap with others. $\bigcup_{i=1}^{k} \omega_i = \Omega$ y $\bigcap_{i=1}^{k} \omega_i = \emptyset$ For instance the sample space defined as $\Omega = \{a, b, c, d, e, f\}$ $\omega_1 = \{a, f\}$ $\omega_2 = \{b, c, d\}$ $\omega_3 = \{e\}$.

Split Ω



Ω
ω_5
ω_4
ω_3
ω_{10}
ω_{11}
ω_{12}
ω_{13}
ω_2
ω_1
ω_6
ω_7
ω_8
ω_9

$P(A)$

total law probability

We need remember by set theory that a event A could be rewrite as $A = (A \cap B) \cup (A \cap C)$. if $(B \cup C) = \Omega$ for this case we can rewrite

$$\begin{aligned} A &= (A \cap \omega_1) \cup (A \cap \omega_2) ... (A \cap \omega_k) \\ P(A) &= P(A \cap \omega_1) + ... + P(A \cap \omega_k) \\ P(A) &= P(A \mid \omega_1) P(\omega_1) + P(A \mid \omega_2) P(\omega_2) + ... + P(A \mid \omega_k) P(\omega_k) \end{aligned} \tag{19}$$

note that by total law $P(A)$

Bayes Theorem

Missused

$$P(A \mid B) \neq P(B \mid A) \tag{20}$$

A daily practical context

Medical diagnose

Clinical and epidemiological research

give us information about patterns of diseases, when you are analyzing the data of medical history record you can compute:

$$P(Symptom \mid Disease) \tag{21}$$

Consider, that an a symptom could be associated with several diseases and in a practical sense you need compute:

$$P(Disease \mid Symptom) \tag{22}$$

Medical diagnose

An a test have a sensitivity $P(\hat{+} \mid +)$ but in a practical sense the really question is:

...

$$P(+ \mid \hat{+}) \tag{23}$$

In simple words: **The probability that the person really present a condition given that the test is positive.**

A daily practical context!

Bayes gives us a way of compute it!.

...

$$P(+ \mid \hat{+}) = \frac{P(\hat{+} \mid +)P(+)}{P(\hat{+})} \tag{24}$$

Now split our population in Disease(D) and Not disease(ND) persons and using **Total probability law** we have:

...

$$P(+ \mid \hat{+}) = \frac{P(\hat{+} \mid +)P(+)}{P(\hat{+} \mid D)P(D) + P(\hat{+} \mid ND)P(ND)} \tag{25}$$

Diagnose zika

Activity

given that $P(\hat{+} \mid Disease) = 0.98$, but in the year is reported that 3% of the population is infected, y is known that 80% of time the test is positive independent if have or not the disease, find:

$$P(Disease \mid \hat{+}) \tag{26}$$