Gregorii Bezdudnyi

Probability Theory Lecture Notes

Contents

Chapter 1

Lecture 1

1.1 The Subject of Probability Theory

The main concept of theory is probability.

Definition 1.1. *Probability is a quantitative measure of randomness.*

The theory is a collection of theorems, rules, formulas that allow one, through the probability of others, to find the probability of certain events in the presence of certain relationships between events.

The principle is applied widely in practice: in physics, chemistry, economics, psychology, medicine and all the disciplines of STEM [1]. For example, probabilistic methods are applied for numerical integration (e.g., Monte Carlo method) of multiple integrals with high multiplicity (e.g. 10, 20, 30), which are hard to compute analytically. Mathematical statistics which is commonly applied for data analysis is founded upon the probability theory.

1.1.1 Basic notions

The basic concepts are experiment, or trial. These words are synonyms.

Definition 1.2 The actual implementation of experiment is called *trial outcome.*

Definition 1.3 The trial outcome, which is bound to occur with any implementation, is called a *sure event.*

Definition 1.4 A trial outcome which never happens is called *impossible event.*

Other trial outcomes are called *random events.*

Example 1.1. *Without aiming a thin uniform coin is thrown onto the table. Thin coin means that it can not fall on the edge. Uniform coin means that the*

[1] Science, technology, engineering, and mathematics

heads and tales are equally likely.
The actual implementation of experiment is a throw of coin.
The coin that fell on the table is a sure event.
A coin that hung in the air is an impossible event.
Random events are heads or tails.

Example 1.2 *Without aiming a uniform dice is thrown onto the table. We exclude the following cases: the dice falls on the edge or vertex (which is theoretically possible, but practically unlikely).*
The actual implementation of experiment is a throw of dice.
Examples of random events are 2 ,3, 4, 5 and 6 points; an even/odd number of points.

1.1.2 Stability of relative frequencies

Let A be a random event that may or may not occur as a result of the experiment. We carry out n independent repetitions of our experiment. By independent repetitions we mean that no single outcome will affect the other implementations. For example, imagine that we toss a dice (or a coin) n times independently.
Let m denote the number of experiments in which the event A will happen, where m is unknown in advance. Then the ratio $\frac{m}{n}$ randomly takes values from 0 to 1.

Definition 1.6. *For n independent experiment repetitions, the ratio $\frac{m}{n}$ is called the relative frequency of event A occurrence.*

As it was noticed in practice, when the number of repetitions n increases (tends to ∞), the relative frequency $\frac{m}{n}$ (though random) tends to a certain fixed number $P(A)$:

$$\lim_{n\to\infty} \frac{m}{n} = P(A) \tag{1.1}$$

Definition 1.7 *This phenomenon is known as stability of relative frequencies.*

Properties

1. $0 \leq P(A) \leq 1$ (obviously).

2. If A is a sure event, then $P(A) = 1$;
   If A is an impossible event, then $P(A) = 0$.
   Intuitively, it seems reasonable to think of the probability of event A in individual experiment as a limit number $P(A)$. We give a rigorous definition in the next section.

Table 1.1: Let us consider the summary of n independent coin tosses, where heads randomly occur m times:

n	m	Relative frequency
4040	2048	$\frac{2048}{4040} = 0.5080$
12000	6019	$\frac{6019}{12000} = 0.5080$
24000	12012	$\frac{12012}{24000} = 0.5005$

The summary above shows that the relative frequencies of heads tend to 0.5.

1.1.3 Space of elementary events

Consider the experiment and it's many possible outcomes. Of this set of outcomes, we consider a subset Ω of outcomes ω fulfilling the following requirements:

1. Each implementation of experiment will always give one of the ω from Ω.
2. Two different ω_1 and ω_2 from Ω can not result from the same implementation of experiment.

Definition 1.8 *This ω is called an elementary event (or elementary outcome).*

Definition 1.9 *This Ω is called a space of elementary events.*

Definition 1.10 *The space Ω is called discrete, if it is finite or countable* [2]

Example 1.4 *Without aiming a thin uniform coin is flipped once. From the subset of outcomes we consider two events: heads (H) or tails (T). We consider them as the elementary events:*

$$\Omega = \{H, T\}$$

Example 1.5 *Without aiming a thin uniform coin is flipped twice. We can build the space of elementary events as follows:*

$$\Omega = \{HH, TT, TH, HT\}$$

Example 1.6 *Without aiming a thin uniform coin is flipped three times. We can build the space of elementary events as follows:*

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

Example 1.7 *Without aiming a uniform dice is thrown on the table once. We can distinguish six different outcomes:*

[2] Countable set is a finite or infinite set which can be mapped to the set of natural numbers ($\mathbb{N} = \{1, 2, \ldots\}$).

$$\omega_i - i \text{ points, where } \quad 1 \leq i \leq 6 \tag{1.2}$$

These events form the space of elementary outcomes:

$$\Omega = \{\omega_i\}$$

Example 1.8 *Independently and without aiming a uniform dice is thrown twice. This experiment can be also interpreted as two uniform dices thrown at once. Then the first dice stands for the first throw, and the second dice - for the second. Let us select a pair (i, j) as an elementary outcome, where i and j denote a number of points at the first and second throws, respectively, $i, j \in [1, \ldots, 6]$. The overall number of such pairs equals 36. Altogether the set of such pairs form a space Ω of elementary events.*

Example 1.9 *Independently and without aiming a uniform dice is thrown on the table three times. The set of elementary events is*

$$\Omega = \{(i, j, k) | i, j, k \in [1, \ldots, 6]\} \tag{1.3}$$

Altogether Ω contains 216 of such a triples.

Example 1.10 *From a deck containing 36 cards we pick randomly 6 cards. In this case, an elementary event - is a set containing 6 different cards. The number of such sets is C_{36}^{6}, where the number of k-combinations is calculated by the formula:*

$$C_n^k = \frac{n!}{k!(n-k)!}.$$

1.2 Algebra of Events

Let us consider only the experiments with discrete set of elementary events Ω.

Definition 1.11. By *random event* we call any arbitrary subset of Ω. Let us further denote them by $A, B, C, D, \ldots$, sometimes supplied with the indexes.

Definition 1.12. Let us call all the elementary events that fell into set A favourable for event A.

Definition 1.13. Let us call Ω a sure event, while the empty set $\emptyset$ will stand for an impossible event.

Definition 1.14. By *sum of events* A and B ($A \cup B$, $A + B$) we mean the event which results from union of the events A and B. It contains all the elementary events that are favourable for either A, B, or both of the events A and B.

Definition 1.15. By product of events A and B (AB) we mean the set which results from intersection of A and B. In other words, by the product of two events we understand the event which implies simultaneous occurrence of the two events.

Definition 1.16. By *difference of events* A and B ($A \backslash B$) we mean the event which contains all the elementary events, that fell into A and not into B.

Definition 1.17. By the event *inverse* to A we understand $\overline{A} = \Omega \backslash A$. Logically, $\overline{\overline{A}} = A$, $\overline{\Omega} = \emptyset$, $\overline{\emptyset} = \Omega$.

Definition 1.18. Two events A and B are called *incompatible* (*mutually exclusive*), if AB is impossible event.

Definition 1.19. Let us call $A_1, A_2, \ldots, A_n$ a *full group of events*, if the following properties hold:

1. A_i are pairwise mutually exclusive;

2. $A_1 + A_2 + \ldots + A_n = \Omega$

In this case A_i are referred to as *hypothesis*.

Example 1.13. A and $\overline{A}$ form full group of events.

De Morgan's laws:

$$\overline{A + B} = \overline{A}\overline{B} \tag{1.4}$$

$$\overline{AB} = \overline{A} + \overline{B} \tag{1.5}$$

This formulas hold for any number of events.

1.3 Classical definition of probability

Let us consider only the experiments for which the space of elementary events Ω fulfils the following two requirements:

1. $|\Omega| \leq \infty$ (the number of elementary events is finite)

2. All the elementary events ω_i contained in Ω are equally likely.

Definition 1.20. If the above two conditions hold, by probability of event A we understand a number:

$$P(A) = \frac{|A|}{|\Omega|}, \tag{1.6}$$

where $|A|$ stands for the number of elementary events favourable for A, i.e., the cardinality of set A.

Properties

1. $0 \leq P(A) \leq 1$;

2. $P(A) = 1$ if and only if $A = \Omega$ (A is sure event);

3. $P(A) = 0$ if and only if $A = \emptyset$ (A is impossible event);

4. $P(A) + P(\overline{A}) = 1$ *Proof:*

$$P(\overline{A}) = \frac{|\Omega \backslash A|}{|\Omega|} = \frac{|\Omega| - |A|}{|\Omega|} = 1 - \frac{|A|}{|\Omega|}, P(A) + P(\overline{A}) = 1$$

Example 1.14 A uniform coin is flipped without aiming. Then

$$\Omega = \{H, T\},$$

$|\Omega| = 2 < \infty$. Are those outcomes equally likely? Yes, thus we can use the above-formulated definition of probability 1.6, A - the coin landed on heads, $|A| = 1, P(A) = \frac{1}{2}$, $\overline{A}$ - the coin landed on tails, $P(\overline{A}) = 1 - \frac{1}{2} = \frac{1}{2}$.

Example 1.15 A uniform coin is flipped two times without aiming. Find the probability of:

1. A) getting head twice;
2. B) getting head once;
3. C) getting head at least once.

$$\Omega = \{HH, TT, TH, HT\}.$$

All the outcomes are equally likely, hence we apply the classical definition of probability:

1. $|A| = 1, P(A) = \frac{1}{4}$;
2. $|B| = 2, P(B) = \frac{1}{2}$;
3. $|C| = 3, P(C) = \frac{3}{4}$.

Example 1.16 A uniform coin is flipped three times without aiming. Find the probability of:

1. A) getting head three times;
2. B) getting head twice;
3. B) getting head once;
4. D) getting head at least once.

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

1. $|A| = 1, P(A) = \frac{1}{8}$;
2. $|B| = 3, P(B) = \frac{3}{8}$;
3. $|C| = 3, P(C) = \frac{3}{8}$;
4. $|D| = 7, P(D) = \frac{7}{8}$;

Example 1.17. A dice is thrown one time without aiming. Find the probability that

1. A) the number of points is even;

2. B) the number of points is multiple of three;
3. C) the number of points < 5;
4. D) the number of points > 2.

$$\Omega = \{\omega_i\}_{i=1}^{6}, \omega_i \text{ - the number of points.}$$

1. $|A| = 3, P(A) = \frac{1}{2}$;
2. $|B| = 2, P(B) = \frac{1}{3}$;
3. $|C| = 4, P(C) = \frac{2}{3}$;
4. $|D| = 4, P(D) = \frac{2}{3}$;

Exercise Two dices are thrown at once without aiming. Find the probability of

1. A) getting the same number of points on both dice;
2. B) getting the sum of points which equals 8;
3. C) getting the sum of points not exceeding 4;

Chapter 2

Lecture 2

2.1 Geometric Probability

The application of classical definition is limited, since it requires $|\Omega| < \infty$ and equally probable elementary outcomes. In many real applications these conditions do not hold. Other probability concepts do exist, which have no such limitations. One of such examples is the concept of geometric probability.

Definition 2.1 Consider $D \subset G \subset \mathbb{R}^n$ - Lebesgue measurable sets. Experiment consists of random sampling of a point from the set G. We are interested in probability that the sampled point will fell into set D. The classical definition is not applicable, since the set of all points is infinite. In this case, the probability is defined as a ratio of areas corresponding to sets D and G:

$$P(D) = \frac{|D|}{|G|}. \tag{2.1}$$

Properties

1. $0 \leq P(D) \leq 1$;
2. If $D = G$, then D is sure event;
3. If $D = \emptyset$, then D is impossible event.
4. If D is sure event $\implies P(D) = 1$ (but the reverse statement does not hold!)

   Proof. Let us consider $D = G \backslash \{x_0\}$. Then $|D| = |G| \implies P(|G|) = 1$. The last statement holds, since Lebesgue measure of the set consisting of any countable number one points is zero ($|\{x_0\}| = 0$).

5. D - is impossible event ($D = \emptyset$) $\implies P(D) = 0$ (the reverse statement does not hold!)

Proof. Let us consider the set consisting of a discrete point $D = \{x_0\}$. Then D is not empty, but $|D| = 0$.

6. $P(D) + P(\overline{D}) = 1$.

Example 2.2. *"Meeting Problem".* Two friends agreed to meet at a specific place during one hour. Each of two persons arrives independently of each other at any random moment. At the same time, the one who came earlier is waiting for another no more than 20 minutes. If the second person does not show up, the first person will leave. What is the probability that the meeting will take place?

Let us denote by X the moment of arrival of the first person ($0 \leq X \leq 1$); by Y we denote the moment of arrival of the second person ($0 \leq Y \leq 1$). X and Y are independent of one another. The experiment consists of two independent arrivals. Elementary event is defined as an ordered pair of numbers (X, Y). This is a point on the plane. A set of all such pairs comprises a unit square $[0, 1] \times [0, 1]$.

Since they arrive independently of one another at any random moment, the experiment can be interpreted as sampling a random point from unit square G. Obviously, $|G| = 1$. The meeting will happen if and only if $|x - y| \leq \frac{1}{3}$. Let us denote a set of points defined by the last inequality by D.

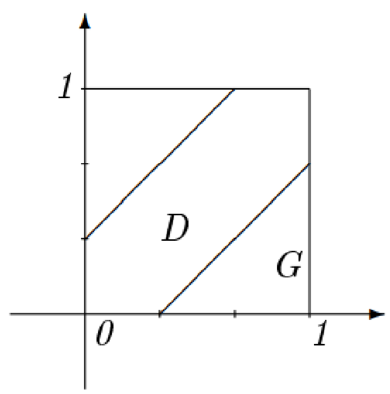
$|x - y| = \frac{1}{3} \implies y = x \pm \frac{1}{3}$;

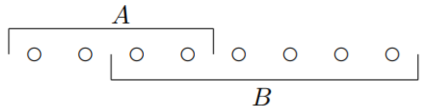
$|D| = 1 - \dfrac{4}{9} = \dfrac{5}{9}$, where $\dfrac{2}{9}$ is the square area of each of the triangles.

$$P(D) = \frac{|D|}{|G|} = \frac{\frac{5}{9}}{1} = \frac{5}{9}$$

.

Example 2.3. $\bigoplus$ Check the Buffon's needle problem.





2.1.1 Event Union Probability

Theorem 1. For any events A, B, such that $AB \neq \emptyset$

$$P(A+B) = P(A) + P(B) - P(AB) \tag{2.2}$$

If two events A and B are not mutually exclusive ($AB \neq \emptyset$), then the probability of their union (the event that at least one of the events A and B will occur) is equal to the sum of their probabilities minus the sum of their intersection.

Corollary 1 If the events A and B are mutually exclusive (i.e., $AB = \emptyset$), then $P(A+B) = P(A) + P(B)$.

Theorem 2. For any given number of random events $A_1, A_2, \ldots A_n$

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \tag{2.3}$$
$$\sum_{1\leq i\leq j\leq n} P(A_iA_j) + \sum_{1\leq i\leq j\leq k\leq n} P(A_iA_jA_k) - \ldots + (-1)^{n-1}P(A_1A_2\ldots A_n)$$

Corollary 2 If A_i are mutually independent events, than

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i). \tag{2.4}$$

Using De Morgan's law, we obtain:

$$\overline{\sum_{i=1}^{n} A_i} = \prod_{i=1}^{n} \overline{A}_i$$

$$P\left(\sum_{i=1}^{n} A_i\right) = 1 - P\left(\prod_{i=1}^{n} \overline{A}_i\right)$$

Example 2.3. We randomly select one card from a deck of 36 cards. What is the probability to take out a picture (jack, queen, king) or a spade?

A - drawn card is a picture; B - card taken has a spade suit. Elementary event is drawn out card. All the cards are equally possible. Ω is a deck of cards.

$$|\Omega| = 36; |A| = 12; |B| = 9; |AB| = 3$$

$$P(AB) = \frac{12}{36}; \quad P(B) = \frac{9}{36}; \quad p(AB) = \frac{3}{36}$$

$$P(A+B) = \frac{12}{36} + \frac{9}{36} - \frac{3}{36} = \frac{1}{2}$$

2.1.2 Event Product and Conditional Probability

Definition 2.2 Let us consider events A and B such, that $P(B) \neq 0$. We define *conditional* probability of event A given that event B has occurred as

$$P(A|B) = \frac{P(AB)}{P(B)}. \tag{2.5}$$

The above expression is read as "P of A after B".

Definition 2.3. $P(A)$ is called *unconditional* probability of event A.

Example 2.5. The urn contains 3 white and 7 black balls. We randomly select one ball twice, without returning. Let us consider events A - the second ball is white; and B - the first ball is black. What is $P(A|B)$?

By experiment we mean subsequent random selection of two balls. Imagine that all the balls are numbered. Then elementary event - an ordered pair of balls.
$|\Omega| = A_{10}^2 = 90; \quad |B| = 7 \cdot 9 = 63; \quad |AB| = 7 \cdot 3 = 21$

$$P(B) = \frac{63}{90} = \frac{7}{10}; \quad P(AB) = \frac{21}{90} = \frac{7}{30}; \quad P(A|B) = \frac{P(AB)}{P(B)} = \frac{1}{3}$$

Theorem 3. For any pair of events A and B. From definition of conditional probability 2.5, we obtain

$$P(AB) = P(A)P(B|A) = P(B)P(A|B). \tag{2.6}$$

Definition 2.4. Two events A and B are called *independent* events if $P(AB) = P(A)P(B)$. The literal meaning of independent events is the events which occur freely of each other.

Note that in this case $P(B|A) = P(B|\bar{A}) = P(B)$ in 2.6.

Lemma 1. Independence of any pair of events from $AB, \overline{A}B, A\overline{B}, \bar{A}\bar{B}$ implies the independence of the remaining 3 pairs.

Definition 2.6 Events $A_1, A_2, \ldots, A_n$ are called *pairwise independent* if for any i, j A_i and A_j are independent.

Definition 2.6 Events $A_1, A_2, \ldots, A_n$ are called *mutually (jointly) independent* if for any $A_{i_1}, A_{i_2}, \ldots, A_{i_k}, k \in [2, n]$

$$P(A_{i_1}, A_{i_2}, \ldots, A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$$

.

Example 2.6. A box contains 4 details: one piece has type A defect, the second one features type B defect, the third piece has type C defect. The fourth one has all the three defects. For luck, one piece is taken out of the box.

Let us consider three events: that randomly taken detail has defect of type A, B or C, respectively.

$$|\Omega| = 4; \quad |A| = |B| = |C| = 2 \implies P(A) = P(B) = P(C) = \frac{1}{2},$$

$$|AB| = |AC| = |BC| = 1 \implies P(AB) = P(A)P(B) = \frac{1}{4}$$

Similarly, $P(AC) = P(A)P(C); P(BC) = P(B)P(C)$

$$|ABC| = 1 \implies P(ABC) = \frac{1}{4} \neq P(A)P(B)P(C)$$

These events are pairwise independent, but not mutually independent.
Theorem 4. Probability of event product. For any events $A_1, A_2, \ldots, A_n$

$$P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_2A_1) \cdots P(A_n|A_1 \cdots A_{n-1}) \tag{2.7}$$

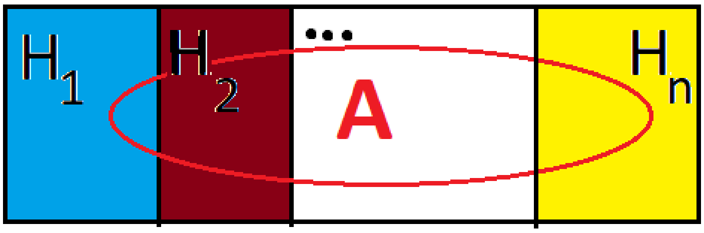
Note that the above formula is an extension of 2.6.

If the events $A_1, A_2, \ldots, A_n$ are *mutually independent*, then

$$P\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i) \tag{2.8}$$

2.1.3 Law of total probability

Let us consider event A, full group of events (hypothesis) $H_1, H_2, \ldots H_n$.
Event A can occur together with any of the hypothesis H_i, $i = 1, \ldots, n$. The probabilities $P(H_i)$ are known (we call them *a prior probabilities of hypotheses*). Also $P(A|H_i)$ are known. The the formula of total probability is:



$$P(A) = \sum_{i=1}^{n} P(H_i)P(A|H_i) = P(H_1)P(A|H_1) + \ldots P(H_n)P(A|H_n) \qquad (2.9)$$

The above formula is base on the representation of event A as a union of pairwise independent events:

$$A = \sum_{i=1}^{n} H_i A,$$

then, by 2.4 from Collary 2 we obtain:

$$P(A) = \sum_{i=1}^{n} P(H_i A)$$

.

By definition of conditional probability 2.5:

$$P(A) = \sum_{i=1}^{n} P(H_i A) = \sum_{i=1}^{n} P(H_i)P(A|H_i)$$

.

Example 3.2. The junk folder contains 10 messages, three of which contain the word 'award'. The probability of spam message given word 'award' in the caption is $\frac{9}{10}$, from a normal message the chance is $\frac{4}{5}$. These is known by statistics. For luck, a message is open from the folder. What is the probability of receiving spam?

Event A - a spam message read, H_1 - randomly selected message contains word 'award', H_2 - word 'award' is missing from the message. H_1 and H_2 form a full group of events (they are mutually exclusive and in any implementation on of these hypotheses for sure will hold). Event A can occur with either H_1 or H_2.

$$P(H_1) = \frac{3}{10}, \quad P(H_2) = \frac{7}{10}.$$
$$P(A|H_1) = \frac{9}{10}, \quad P(A|H_2) = \frac{4}{5}.$$

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) = 0.83$$

2.2 Bayes theorem

We preserve notations from the full probability law. Additionally it becomes known that event *occured* as an outcome of experiment. But it's not known which of the hypotheses happened together with A. New information allows to reassess the probability of hypotheses:

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)}, \quad i \in [1, n] \tag{2.10}$$

Proof From $AH_i = H_iA$ and by Theorem 3:

$$P(AH_i) = P(A)P(H_i|A)$$

$$P(H_iA) = P(H_i)P(A|H_i)$$

$$P(A)P(H_i|A) = P(H_i)P(A|H_i)$$

Example 3.3. Let us consider the previous example, given that selected message was spam. What is the chance that the message contains the word "award"?

$P(A) = 0.83$ (from the previous example).

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)} = \frac{0.3 \cdot 0.9}{0.83} = \frac{27}{83}$$

.

Definition 3.1. $P(H_i|A)$ - is a posterior probability of hypothesis H_i.

Definition 3.1. $P(H_i)$ - is a prior probability of hypothesis H_i.

2.3 Discrete random variable

Example Let us consider experiment with two possible outcomes: A - success and $\bar{A}$ - failure.

$$P(A) = p, \quad P(\bar{A}) = 1 - p = q$$

Consider a complex experiment which consist of successive repetitions of experiment until the first successful event.

Elementary events: A,$\{\bar{A}, A\}$, $\{\bar{A}, \bar{A}, A\}$, etc.

$$\omega_n = \underbrace{\bar{A}\bar{A}, \dots \bar{A}A}_{\text{n letters}}$$

$\Omega = \{\omega_n\}$ - is an infinite countable set. Ω - is discrete space of elementary events. Let's denote all the possible subsets of Ω by $\mathbb{F}$. Then the probability measure can be defined as follows: $P(\omega_n) = q^{n-1}, \quad n = 1, 2, \dots$.

$$\forall B \in \mathbb{F} P(B) = \sum_{n:\omega_n \in B} P(\omega_n)$$

.

Let us define a discrete random variable on space Ω: $\xi(\omega_n) = n, \quad \xi : \Omega \to \mathbf{R}^1$.
$P(\xi = n) = P(\omega_n) = q^{n-1}p, \quad n = 1, 2, \dots$.
Probability can be given as a table.

ξ	1	2	3	...	n	$\dots$
p	p	qp	q^2p	$\dots$	$q^{n-1}p$	$\dots$

The sum of binomial coefficients in the second row equals to one. This table defines the geometric distribution with parameter p.

Alternatively, if we consider all possible n repetitions of experiment A, with possible successes and failures, but the order does not matter (for example, $\underbrace{AA, \dots A}_{\text{n letters}}$ and $\underbrace{\bar{A}\bar{A}, \dots \bar{A}}_{\text{n letters}}$) . Then $|\Omega| = 2^n$. Number of outcomes is finite, same as the set of all the subsets of Ω, $\mathbb{F}$:

$$|\mathbb{F}\| = C^0_{2^n} + C^1_{2^n} + C^2_{2^n} + \dots + C^{2^n}_{2^n} = 2^{2^n}$$

.

Let the elementary outcome ω contain k letters A and $n - k$ letters $\bar{A}$. $(0 \leq k \leq n)$. Order of the latter does not matter Obviously, $P(\omega) = p^k q^{n-k}$. If we extend this probability to all the subsets in $\mathbb{F}$, we can define the probability of k latter A occurring in the sequence, this there exist C^k_n of ω, such that A occurs k times in different positions. Let's define a random variable $\xi : \Omega \to [1, \dots, n]$. ξ is a random number of successful outcomes in Bernoulli scheme with n repetitions:

$$P(\xi = k) = C^k_n p^k q^{n-k}, \quad 0 \leq k \leq n. \tag{2.11}$$

ξ is discrete random variable with binomial distribution and parameters n and p.

2.4 Mean and variance, covariance and correlation