N.I. Lobachevsky State University
of Nizhni Novgorod

Probability theory and mathematical statistics:

Independent trials

Associate Professor
A.V. Zorine

Suppose $\Omega = \{\omega_1, \omega_2, \ldots\}$ is a finite or denumerable set. Assign a number $p(\omega) \geqslant 0$ to each outcome $\omega \in \Omega$ such that

$$p(\omega_1) + p(\omega_2) + \ldots = 1.$$

Then $\mathrm{P}(A) = \sum_{\omega \in A} p(\omega)$, $A \subset \Omega$, is a probability.

Suppose $\Omega = \{\omega_1, \omega_2, \ldots\}$ is a finite or denumerable set. Assign a number $p(\omega) \geqslant 0$ to each outcome $\omega \in \Omega$ such that

$$p(\omega_1) + p(\omega_2) + \ldots = 1.$$

Then $\mathrm{P}(A) = \sum_{\omega \in A} p(\omega)$, $A \subset \Omega$, is a probability.
1. $\mathrm{P}(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$.

Suppose $\Omega = \{\omega_1, \omega_2, \ldots\}$ is a finite or denumerable set. Assign a number $p(\omega) \geqslant 0$ to each outcome $\omega \in \Omega$ such that

$$p(\omega_1) + p(\omega_2) + \ldots = 1.$$

Then $\mathrm{P}(A) = \sum_{\omega \in A} p(\omega)$, $A \subset \Omega$, is a probability.
1. $\mathrm{P}(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$.
2. $\mathrm{P}(A) = \sum_{\omega \in A} p(\omega) \geqslant 0$.

Suppose $\Omega = \{\omega_1, \omega_2, \ldots\}$ is a finite or denumerable set. Assign a number $p(\omega) \geqslant 0$ to each outcome $\omega \in \Omega$ such that

$$p(\omega_1) + p(\omega_2) + \ldots = 1.$$

Then $\mathrm{P}(A) = \sum_{\omega \in A} p(\omega)$, $A \subset \Omega$, is a probability.
1. $\mathrm{P}(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$.
2. $\mathrm{P}(A) = \sum_{\omega \in A} p(\omega) \geqslant 0$.
3. Let $A_1, A_2, \ldots$ be mutually exclusive events, then

$$\begin{aligned} \mathrm{P}(A_1 \cup A_2 \cup \ldots) &= \sum_{\omega \in A_1 \cup A_2 \cup \ldots} p(\omega) \\ &= \sum_{\omega \in A_1} p(\omega) + \sum_{\omega \in A_2} p(\omega) + \ldots \\ &= \mathrm{P}(A_1) + \mathrm{P}(A_2) + \ldots \end{aligned}$$

People frequently keep independence in mind when assigning $p(\omega)$ to an outcome ω.
Consider a sequence of n independent trials. In each trial one of mutually exclusive events E_1, E_2 …, E_r occurs. Put $p_k = \mathrm{P}(E_k)$, $k = 1, 2, \ldots, r$.
Let's construct a probability model for this situation. Put

$$\Omega = \{(x_1, x_2, \ldots, x_n) \colon x_1 \in \{1, 2, \ldots, r\}, \ldots, x_n \in \{1, 2, \ldots, r\}\}$$

Here $x_1 = k$ if event E_k was observed in kth trial. To introduce independence to the model, put $p(\omega) = p_{x_1} p_{x_2} \cdots p_{x_n}$ when $\omega = (x_1, x_2, \ldots, x_n)$.

Bernoulli trials

Let $r = 2$, i.e. an event A occurs ($E_1 = A$) or does not occur ($E_2 = \bar{A}$) in a single trial. It is useful here to introduce binary vectors $\omega = (x_1, x_2, \dots, x_n)$, $x_i \in \{0, 1\}$ to describe outcomes. By p denote the probability of A in a single trial and define $p(\omega) = p^{x_1+\dots+x_n}(1-p)^{n-(x_1+\dots+x_n)}$. For example,

$$p((0, 0, \dots, 0)) = (1-p)^n,$$

$$p((1, 1, \dots, 1)) = p^n,$$

$$p((1, \underbrace{0, \dots, 0}_{n-1 \text{ zeros}})) = p(1-p)^{n-1},$$

$$p((0, 1, \underbrace{0, \dots, 0}_{n-2 \text{ zeros}})) = p(1-p)^{n-1},$$

Examples of events in Bernoulli trials and their probabilities:

- What's the probability that A never happens?

$$\mathrm{P}(\text{"}A \text{ never happens"}) = \mathrm{P}(\{(0,0,\ldots,)\}) = (1-p)^n$$

- What's the probability that A happens at least once?

$$\mathrm{P}(\text{"}A \text{ happens at least once"}) = 1-\mathrm{P}(\text{"}A \text{ never happens"}) = 1-(1-p)^n$$

- What's the probability A happens exactly once?

$$\mathrm{P}(\{(1,\underbrace{0,\ldots,0}_{n-1 \text{ zeros}}),(0,1,\underbrace{0,\ldots,0}_{n-2 \text{ zeros}}),\ldots,(\underbrace{0,\ldots,0}_{n-1 \text{ zeros}},1)\}) = np(1-p)^{n-1}$$

Event A is usually referred to as a "success", event $\bar{A}$ as "failure".

We have to prove now that our models does not lead to contradictions with initial problem statement.

1) The probability of A in a single trial should be equal to p. Let's compute this probability for the first trial:

$$\mathrm{P}(\{(1, x_2, \ldots, x_n) \colon x_2 \in \{0,1\}, \ldots, x_n \in \{0,1\}\}) =$$

$$= \sum_{x_2=0}^{1} \cdots \sum_{x_{n-1}=0}^{1} \sum_{x_n=0}^{1} p^{1+x_2+\ldots+x_{n-1}+x_n}(1-p)^{n-(1+x_2+\ldots+x_{n-1}+x_n)}$$

$$= p \sum_{x_2=0}^{1} \cdots \sum_{x_{n-1}=0}^{1} \sum_{x_n=0}^{1} p^{x_2+\ldots+x_{n-1}}(1-p)^{n-2-(x_2+\ldots+x_{n-1})} p^{x_n}(1-p)^{1-x_n}$$

$$= p \sum_{x_2=0}^{1} \cdots \sum_{x_{n-1}=0}^{1} p^{x_2+\ldots+x_{n-1}}(1-p)^{n-2-(x_2+\ldots+x_{n-1})}(1-p+p)$$

$$= \ldots = p$$

2) Using the same technique we can show that

$$\mathrm{P}(\{(1,0,x_3,\dots,x_n)\colon x_3\in\{0,1\},\dots,x_n\in\{0,1\}\})=p(1-p)$$

This means that occurrences of A in the first and the second trials are independent:

$$\begin{aligned}&\mathrm{P}(\{(1,0,x_3,\dots,x_n)\colon x_3\in\{0,1\},\dots,x_n\in\{0,1\}\})=\\&\quad=\mathrm{P}(\{(1,x_2,\dots,x_n)\colon x_3\in\{0,1\},\dots,x_n\in\{0,1\}\})\\&\qquad\times\mathrm{P}(\{(x_1,0,x_3,\dots,x_n)\colon x_3\in\{0,1\},\dots,x_n\in\{0,1\}\}).\end{aligned}$$

Similar results can be obtained for any on n trials.

Formula of Jacob Bernoulli

Theorem

The probability to observe event A exactly k times in n repeated (independent) trials equals

$$b(k;n,p) = C_n^k p^k (1-p)^{n-k}$$

Numbers $\{b(k;n,p) : k = 0, 1, \ldots, n\}$ are called *binomial probabilities*.

Formula of Jacob Bernoulli

Theorem

The probability to observe event A exactly k times in n repeated (independent) trials equals

$$b(k;n,p) = C_n^k p^k (1-p)^{n-k}$$

Numbers $\{b(k;n,p) \colon k = 0, 1, \ldots, n\}$ are called *binomial probabilities*.

Proof. The number of favorable outcomes equals to the number of ways to place exactly k 1's into n places, which is C_n^k. Each favorable outcome ω' is assigned the number $p(\omega') = p^k(1-p)^{n-k}$. So

$$\sum_{\omega' \colon x_1 + \ldots + x_n = k} p(\omega') = C_n^k p^k (1-p)^{n-k}.$$

Usage of Bernoulli's formula

Often this formula is used without constructing specific probability model!

Example. A die is rolled 5 times. What is the probability to have 6 points twice?

Solution. The probability to have 6 points in one roll is $\frac{1}{6}$ and rolls are made independently. Thus the probability is $b(2; 5, \frac{1}{6})$ which equals

$$C_5^2 \left(\frac{1}{6}\right)^2 \left(1 - \frac{1}{6}\right)^{5-2} = \frac{3125}{23328} \approx 0{,}133959106721536\ldots$$

Problem. How many throws of two dice is it enough for a pair of ⚅ appears at least once with probability $\geqslant \frac{1}{2}$?
Solution. Here $p = (1/36)^2$, n is unknown. The probability to have two ⚅'s in n throws is

$$1 - \left(\frac{35}{36}\right)^n$$

Find the smallest n such that

$$1 - \left(\frac{35}{36}\right)^n \geqslant \frac{1}{2}.$$

By trial-and-error (or taking logarithms), $n = 25$.

A power supply problem

Suppose that $n = 10$ workers are to use electric power intermittently, and we are interested in estimating the total load to be expected. Assume that each worker has the same probability p of requiring a unit of power. If they work independently, the probability of exactly k workers requiring power at the same time should be $b(k; n, p)$.
If, on the average, a worker uses power for 12 minutes per hour, we should put $p = \frac{1}{5}$. The probability of seven or more workers requiring current at the same time is then

$$b(7; 10, 0{,}2) + \ldots + b(10; 10, 0{,}2) = 0{,}00086\ldots.$$

If the supply is adjusted to six power units, an overload has probability $0{,}00086\ldots$ and should be expected about one minute in twenty hours.

Problem. The probability that an article is defective equals 0,005. What is the probability that among 10000 randomly sampled articles a) exactly 40 articles are defective? b) at most 70 articles are defective?

Solution. Here $n = 10000$, $p = 0{,}005$.
a) $b(40; 10000, 0{,}005) = C_{100000}^{40} \cdot 0{,}005^{40} \cdot 0{,}995^{9960} = 0{,}021435\ldots$
b) There could be 0, 1, ..., 70 defective articles:

$$Q(70;10000, 0{,}005) = \sum_{k=0}^{70} b(k; n, p) =$$

$$= \sum_{k=0}^{70} C_{100000}^{k} \cdot 0{,}005^{k} \cdot 0{,}995^{10000-k} = 0{,}997098\ldots$$

Let's study $b(k;n,p)$ as a function of k. Since

$$R(k;n,p)=\frac{b(k;n,p)}{b(k-1;n,p)}=\frac{C_n^k p^k(1-p)^{n-k}}{C_n^{k-1}p^{k-1}(1-)p^{n-k+1}}=$$
$$=\frac{p}{1-p}\cdot\frac{n!(k-1)!(n-k+1)!}{k!(n-k)!n!}=\frac{p}{1-p}\cdot\frac{n-k+1}{k}=$$
$$=\frac{p}{1-p}\left(\frac{n+1}{k}-1\right).$$

Obviously $R(k;n,p)$ is decreasing as k increases.

Solving inequation

$$\frac{p}{1-p} \cdot \frac{n-k+1}{k} > 1$$

for k obtain

$$k < (n+1)p$$

So $R(k;n,p) < 1$ when $\frac{k}{n+1} < p$ and $R(k;n,p) < 1$ when $\frac{k}{n+1} > p$. It means that probabilities $b(k;n,p)$ increase until $k \geqslant p(n+1)$ and then decrease.

Consider an example with $p = \frac{1}{6}$:

k	0	1	2	3	4	5	6
$b(k; 4, \frac{1}{6})$	0,4823	0,3858	0,1157	0,0154	0,0008		
$b(k; 5, \frac{1}{6})$	0,4019	0,4019	0,1608	0,0322	0,0032	0,0001	
$b(k; 10, \frac{1}{10})$	0,1615	0,3230	0,2907	0,1550	0,0543	0,0130	0,0022

For $n = 4$ we have $p(n+1) < 1$ and all probabilities are decreasing. For $n = 5$ $p(n+1)$ is a whole number so there are two values $k = 0$ and $k = 1$ with the greatest probability.

Computation using spreadsheets

Numbers $R(k; n, p)$ give the means to compute binomial probabilities without using factorials.

	A	B	C
1	p=	0,4	
2	n=	10	
3	k	R(k;n,p)	b(k;n,p)
4	0		=POWER(1-B$1;B$2)
5	1	=B$1*(B$2-A5+1)/(1-B$1)/A5	=C4*B5
6	2	=B$1*(B$2-A6+1)/(1-B$1)/A6	=C5*B6
7	3	=B$1*(B$2-A7+1)/(1-B$1)/A7	=C6*B7
8	4	=B$1*(B$2-A8+1)/(1-B$1)/A8	=C7*B8
9	5	=B$1*(B$2-A9+1)/(1-B$1)/A9	=C8*B9
10	6	=B$1*(B$2-A10+1)/(1-B$1)/A10	=C9*B10
11	7	=B$1*(B$2-A11+1)/(1-B$1)/A11	=C10*B11
12	8	=B$1*(B$2-A12+1)/(1-B$1)/A12	=C11*B12
13	9	=B$1*(B$2-A13+1)/(1-B$1)/A13	=C12*B13
14	10	=B$1*(B$2-A14+1)/(1-B$1)/A14	=C13*B14
15		Total	=SUM(C4:C14)
16			
17			

Computation using spreadsheets

Numbers $R(k;n,p)$ give the means to compute binomial probabilities without using factorials.

	A	B	C	D
1	p=	0,4		
2	n=	10		
3	k	R(k;n,p)	b(k;n,p)	
4	0		**0,00605**	
5	1	6,66667	**0,04031**	
6	2	3,00000	**0,12093**	
7	3	1,77778	**0,21499**	
8	4	1,16667	**0,25082**	
9	5	0,80000	**0,20066**	
10	6	0,55556	**0,11148**	
11	7	0,38095	**0,04247**	
12	8	0,25000	**0,01062**	
13	9	0,14815	**0,00157**	
14	10	0,06667	**0,00010**	
15		Total	1,00000	
16				
17				

Now we can estimate sums

$$Q(k;n,r) = \sum_{j=0}^{k} b(j;n,p)$$

Here $Q(k;n,r)$ equals to the probability to have at most k successes (occurrences of event A). First notice that

$$b(k-1;n,p) = \frac{b(k;n,p)}{R(k;n,p)},$$

$$b(k-2;n,p) = \frac{b(k-1;n,p)}{R(k-1;n,p)} = \frac{b(k;n,p)}{R(k;n,p)R(k-1;n,p)},$$

and so on.

Thus

$$\begin{aligned} Q(k;n,p) &= b(k;n,p)\left(1 + \frac{1}{R(k;n,p)} + \frac{1}{R(k;n,p)R(k-1;n,p)} + \dots\right) \\ &\leqslant b(k;n,p)\left(1 + \frac{1}{R(k;n,p)} + \frac{1}{R(k;n,p)R(k;n,p)} + \dots\right) \\ &\leqslant b(k;n,p)\frac{R(k;n,p)}{R(k;n,p)-1} = b(k;n,p)\frac{(n-k+1)p}{(n+1)p-k}. \end{aligned}$$

This estimate is quite accurate when k and n are large and $k/(pr)$ is not close to 1.

For example, let $n = 30$, $p = 0{,}7$, $k = 16$.

$$\begin{aligned} Q(k;n,p) &= 0{,}040052547682131 34, \\ b(k;n,p) &= 0{,}023115236189581 77, \\ \frac{(n+1-k)p}{(n+1)p-k} &= 1{,}842105263157895 \\ b(k;n,p)\frac{(n+1-k)p}{(n+1)p-k} &= 0{,}042580698243966 42 \end{aligned}$$

with 15 decimal places.

Multinomial trials

In case of $r \geqslant 2$ mutually exclusive outcomes $E_1, E_2, \ldots, E_r$,
$p_k = \mathrm{P}(E_k)$, $p_1 + p_2 + \ldots + p_r = 1$,

$$\Omega = \{(x_1, x_2, \ldots, x_n) \colon x_1 \in \{1, 2, \ldots, r\}, \ldots, x_n \in \{1, 2, \ldots, r\}\},$$

$$p(\omega) = p_{x_1} p_{x_2} \cdots p_{x_n}, \quad \omega = (x_1, x_2, \ldots, x_n)$$

Similar to the Bernoulli trials case it can be proved, that
1) The probability of event E_k in a single trial equals still p_k,
2) Outcomes of different trials are independent.

Theorem

The probability of n_1 occurrences of E_1, n_2 occurrences of E_2, $\ldots$, n_r occurrences of E_r in n independent trials, $n_1 + n_2 + \ldots + n_r = n$, equals

$$\mathrm{multi}(n_1, \ldots, n_r; n, p_1, \ldots, p_r) = \frac{n!}{n_1! n_2! \cdots n_r!} p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

Probabilities $\mathrm{multi}(n_1, \ldots, n_r; n, p_1, \ldots, p_r)$ are called *multinomial probabilities*.

Example. Three dice are rolled. What is the probability that the product of numbers shown on each die is 20?

Example. Three dice are rolled. What is the probability that the product of numbers shown on each die is 20?
Solution. We have 3 independent trials here with outcomes ⚀, ⚁, ⚂, ⚃, ⚄, ⚅. Here $n = 3$, $p_1 = p_2 = \ldots = p_6 = \frac{1}{6}$. Number 20 can be factored in two ways with factors 1, 2, 3, 4, 5, 6:

$$20 = 2 \cdot 2 \cdot 5 = 1 \cdot 4 \cdot 5$$

Now, for $n_2 = 2$, $n_5 = 1$,

$$\text{multi}\left(0, 2, 0, 0, 1, 0; 3, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) = \frac{3!}{2!1!(0!)^4} \cdot \frac{1}{6^3} = \frac{3}{216},$$

and for $n_1 = n_4 = n_5 = 1$

$$\text{multi}\left(1, 0, 0, 1, 1, 0; 3, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) = \frac{3!}{(1!)^3(0!)^3} \cdot \frac{1}{6^3} = \frac{6}{216}.$$

The probability in question is $\frac{3}{216} + \frac{6}{216} = \frac{1}{24}$.

Many interesting problems involve infinite series of independent trials, for example ruin of gamesters, problem of runs, etc. We'll get focused on Bernoulli trials here. Denote by p_1 the probability of success, by $p_0 = 1 - p_1$ the probability of failure. Put

$$\Omega = \{(x_1, x_2, \dots, x_n, \dots) \colon x_k \in \{0, 1\},\ k = 1, 2, \dots\}$$

Keeping independence in mind, we have to assign zero probability to any elementary outcome:

$$\mathrm{P}(\{(0, 0, \dots, 0, \dots)\}) = p_0 p_0 \cdots p_0 \cdots = 0.$$

$$\mathrm{P}(\{(x_1, x_2, \dots, x_n, \dots)\}) = p_{x_1} p_{x_2} \cdots p_{x_n} \cdots = 0.$$

Consider events

$$A^{i_1,i_2,\ldots,i_n}_{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n} = \{(x_1,x_2,\ldots,x_n,\ldots): x_{i_1} = \varepsilon_1, x_{i_2} = \varepsilon_2,\ldots,x_{i_n} = \varepsilon_n\}. \tag{1}$$

Event $A^{i_1,i_2,\ldots,i_n}_{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n}$ is called cylindrical. For independent trials,

$$\mathrm{P}(A^{i_1,i_2,\ldots,i_n}_{\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n}) = p_{\varepsilon_1}p_{\varepsilon_2}\ldots p_{\varepsilon_n}. \tag{2}$$

Formula (2) defines a unique probability for any event which can be expressed using finite or denumerable number of intersections, unions, etc of events (1).

Geometric interpretation of infinite series of trials

As it was noticed in 1920's, any infinite sequence of successes and failures can be represented with a point in a unite segment $[0, 1]$. In fact, any point $x \in [0, 1]$ has its binary expansion

$$x = 0{,}x_1x_2 \dots x_n \dots, x_k \in \{0, 1\},\ k = 1, 2, \dots$$

In this way the probability of an event is a sort of a measure of a point set in the unit segment. It became possible to treat different question in real analysis in a probabilistic way, using notion of independence in particular.

Uniform distribution and symmetric Bernoulli trials

Suppose a random point x is selected in a unit segment $[0, 1]$. The probability $\mathrm{P}(x \in [\alpha, \beta])$ that the point belongs to a segment $[\alpha, \beta]$, $0 \leqslant \alpha \leqslant \beta \leqslant 1$ equals $\lambda([\alpha, \beta]) = \beta - \alpha$, the length of the segment $[\alpha, \beta]$.
Consider binary expansion

$$x = 0{,}x_1x_2 \ldots x_n \ldots, \; x_k \in \{0, 1\}, \; k = 1, 2, \ldots$$

One has

$$\mathrm{P}(x_1 = 0) = \mathrm{P}(0 \leqslant x \leqslant 0{,}5) = \lambda([0,\ 0{,}5]) = 0{,}5,$$

and the same is true for any digit x_k: $\mathrm{P}(x_k = 0) = 0{,}5$. Moreover, any digits $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ are independent.