Problems in the Theory of Probability

B. Sevastyanov, V. Chistyakov, A. Zubkov

B. A. SEVASTYANOV, V. P. CHISTYAKOV,
AND A. M. ZUBKOV

Problems in the Theory of Probability

Translated from the Russian
by
Irene Aleksanova

MIR PUBLISHERS MOSCOW

First published 1985
Revised from the 1980 Russian edition

На английском языке

Preface

This problem book is intended for students of universities and technical colleges as a study aid in the elementary course of the theory of probability. Most of the problems are suitable for mathematics students at technical colleges, but problems 1.19-1.22, 2.18-2.20, 3.114-3.132, 3.147-3.160, 4.9-4.22, 4.41-4.55, 4.89-4.94, 4.101-4.115, 4.135-4.140 are meant primarily for university students.

Command of the basic concepts and methods of the theory of probability is necessary not only for mathematicians but also for applied scientists. The solution of practical problems depends on the choice of a correct stochastic model, which both reflects the essential features of the event under investigation and is easy enough to investigate. One cannot select a model and estimate its accessibility without sufficient knowledge of the probability theory and its methods.

Comparatively few problems in the book are formulated in applied terms since it is primarily intended for general courses in probability theory rather than for various specialized courses at technical colleges. The authors' wish was also to give as much information on the methods of the probability theory as possible in so small a book. Solutions of applied problems are presented in two stages: (1) mathematical formulation of an applied problem and (2) solution of the subsequent mathematical problem. The authors believe that their principal aim is to acquaint students with standard models of the theory of probability and to teach them to cope with the difficulties of the stage of problem solving; the problems were chosen precisely with this idea in mind; problems that differ from ordinary mathematical problems only in "applied" terminology are no substitute for instruction in the techniques of constructing mathematical models of real phenomena. Obviously, models of real phenomena that are actually employed in the fields of science and technology relevant to each technical college will be the most useful.

Only a few of the problems included here involve simple "number crunching". Instead of choosing problems in which students merely substitute numbers in formulas, the authors have tried to choose problems that will acquaint the students with the main concepts and methods of the probability theory as well as illustrate the relationship between the concepts and enable the students to estimate the possibilities of the methods.

In this connection, a number of problems are theoretical, requiring proof of an assertion or investigation of a particular problem. Problems of this kind are usually included as part of a series; their solution in succession should not present difficulties. It is much easier to solve the problems with the aid of the instructions provided in Part 2. The problems marked with asterisks are supplied with solutions.

Theoretical problems often contain material that is significant in principle but is almost ignored in standard courses in the theory of probability. This is especially true of such techniques of problem solving as the representation of the required random variable as a sum of indicators, the use of the linearity of mathematical expectation, the introduction of generating the characteristic functions, the method of moments, a consideration of random variables which are similar to the given quantities but easier to investigate, and so on.

The authors

PART 1

Problems

CHAPTER 1

The Simplest Probability Schemes

Mathematical models of random events considered in the theory of probability are based on the concept of a *sample space* or *probability space*, i.e. a triple $(\Omega, \mathcal{A}, \mathbf{P})$, where $\Omega = \{\omega\}$ is a nonempty set, whose elements ω are interpreted as mutually exclusive outcomes of the random event in question; $\mathcal{A}$ is a collection of subsets of the set Ω called *events* (the set $\mathcal{A}$ is assumed to contain Ω and to be closed with respect to an opposite event or a sum of events in not more than a countable number, i.e. $\mathcal{A}$ is a σ-algebra); the probabibity $\mathbf{P}$ is a function defined on the events $A \in \mathcal{A}$ and satisfying the following conditions:

1. $\mathbf{P}(A) \geqslant 0$ for any $A \in \mathcal{A}$;
2. $\mathbf{P}(\Omega) = 1$; (1.1)
3. $\mathbf{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)$ if $A_i A_j = \varnothing$ for any $i \neq j$.

The sign $\varnothing$ denotes here an empty set (or an impossible event).

The definition of operations on events and that of the algebra and σ-algebra of events can be found in courses in probability theory (see, for instance, [6], [2], [3]). In this chapter we consider two of the simplest classes of sample spaces.

Suppose $\Omega = \{\omega_1, \omega_2, \ldots, \omega_s\}$. The σ-algebra of events $\mathcal{A}$ includes all 2^s subsets $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ of the set Ω. In the classical definition of probability, it is assumed that all $\mathbf{P}(\omega_i) = 1/s$ and, therefore, the probability $\mathbf{P}(A)$ of the event $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ is equal to the ratio of the number of the simple events* ω_i in A to the whole

* Here and in what follows the number of elements of any finite set M will be designated as $|M|$.

number of simple events in Ω:

$$\mathsf{P}(A) = \frac{|A|}{|\Omega|} = \frac{k}{s}. \tag{1.2}$$

The classical definition of probability is a good mathematical model of random events for which the outcomes of the experiment are symmetric in some sense, and it is, therefore, logical to assume that they are equiprobable.

Here is a description of two of the most frequently encountered probability schemes in which the general classical definition is presented in detail. Let us designate as $\mathcal{N}$ a set of N numbers: $\mathcal{N} = \{1, 2, \ldots, N\}$; suppose $\omega = (i_1, i_2, \ldots, i_n)$ is an ordered collection of n elements of the set $\mathcal{N}$. The probability scheme in which

$$\Omega = \{\omega = (i_1, i_2, \ldots, i_n) : i_k \in \mathcal{N}, \quad k = 1, 2, \ldots, n\} \tag{1.3}$$

and all elementary events in ω are equiprobable is called *sampling with replacement*.

Sampling without replacement is such a probability scheme in which

$$\Omega = \{\omega = (i_1, i_2, \ldots, i_n) : i_k \in \mathcal{N}, \quad k = 1, 2, \ldots, n, \text{ there are no identical elements among } i_1, \ldots, i_n\} \tag{1.4}$$

and the elementary events in ω are equiprobable.

Various combinatorial formulas prove useful in calculating probabilities by formula (1.2). Here are the main formulas. Suppose we are given a set $\mathcal{N}$ of N elements: $\mathcal{N} = \{a_1, a_2, \ldots, a_N\}$. The subsets of the set $\mathcal{N}$ are called *combinations*. The number of combinations that can be formed from N elements of $\mathcal{N}$, using various methods to choose subsets with n elements, is designated as C_N^n or $\binom{N}{n}$. The following formulas hold true:

$$C_N^n = \frac{N^{[n]}}{n!}, \quad C_N^n = \frac{N!}{n!\,(N-n)!}, \quad C_N^n = C_N^{N-n},$$

where $n! = 1 \cdot 2 \cdot \ldots \cdot n$ and

$$N^{[n]} = N(N-1)\ldots(N-n+1). \tag{1.5}$$

The ordered chains $a_{i_1} a_{i_2} \ldots a_{i_n}$, formed from various elements of $\mathcal{N}$, are called *arrangements*. The number of arrangements formed by selecting various ordered chains of length n from N elements of $\mathcal{N}$ is designated as A_N^n. For A_N^n we have a formula $A_N^n = N^{[n]}$. A special case of arrangements for $n = N$ is called a *permutation*. The number of different permutations formed from N elements is equal to $N!$

The following classical formula known as refined Stirling's formula (see [5], formula (9.8)) is useful in many cases:

$$n! = \sqrt{2\pi n}\, n^n e^{-n} e^{\frac{\theta n}{12n}}, \quad \frac{12n}{12n+1} < \theta_n < 1. \qquad (1.6)$$

The formulations of some problems include the expression 'the integer a is comparable with the integer b modulo m' (m is an integer), or, in symbolic notation,

$$a \equiv b \pmod{m}. \qquad (1.7)$$

Comparison (1.7) is equivalent to the following statement: *there is an integer t such that* $a - b = tm$ (i.e. a and b, when divided by m, leave the same remainder). In particular, the notation $a \equiv 0 \pmod{m}$ means that a is exactly divisible by m.

We shall designate the integral part of the real number x (the largest integer not exceeding x) as $[x]$ (not to be confused with $a^{(x)}$, where x is an integer, see (1.5)).

Let us consider the second class of the general sample spaces. Suppose Ω is a bounded set of an n-dimensional Euclidean space. We shall assume that Ω has a volume. Consider a system $\mathcal{A}$ of the subsets of Ω which have volume. For any $A \in \mathcal{A}$ wet put

$$\mathsf{P}(A) = \frac{\mu(A)}{\mu(\Omega)}, \qquad (1.8)$$

where $\mu(C)$ is the volume of the set C. If by the volume of a set we mean its Lebesgue measure, then the system $\mathcal{A}$ is the σ-algebra of the Lebesgue measurable sets, and then the function $\mathsf{P}(A)$, defined by formula (1.8), is a probability. Note that in special cases, the system $\mathcal{A}$ contains all the Jordan measurable subsets of Ω, i.e ordinary squarable or cubable figures studied in every course of mathematical analysis. Most problems in this book are concerned with this particular case. The definition of probability (1.8) is known as the *geometrical definition of probability.*

Here are some of the formulas most often employed in problem solving. For any events $A_1, A_2, \ldots$, we have

$$\overline{\bigcup_{n=1}^{\infty} A_n} = \bigcap_{n=1}^{\infty} \overline{A_n}, \quad \overline{\bigcap_{n=1}^{\infty} A_n} = \bigcup_{n=1}^{\infty} \overline{A_n}, \qquad (1.9)$$

(Here and in what follows a bar over a letter indicates an opposite event.) The following formula is valid for any A and B:

$$\mathbf{P}(A+B)=\mathbf{P}(A)+\mathbf{P}(B)-\mathbf{P}(AB). \tag{1.10}$$

For $AB=\varnothing$, in particular, we have

$$\mathbf{P}(A+B)=\mathbf{P}(A)+\mathbf{P}(B). \tag{1.11}$$

The probability of the sum of n arbitrary events can be found by the formula

$$\begin{aligned}\mathbf{P}(A_1+A_2+\ldots+A_n)&=\sum_{k=1}^{n}\mathbf{P}(A_k)-\sum_{1\leqslant k_1<k_2\leqslant n}\mathbf{P}(A_{k_1}A_{k_2})\\&+\sum_{1\leqslant k_1<k_2<k_3\leqslant n}\mathbf{P}(A_{k_1}A_{k_2}A_{k_3})\\&-\ldots+(-1)^{n-1}\mathbf{P}(A_1A_2\ldots A_n)\\&=\sum_{l=1}^{n}(-1)^{l-1}\sum_{1\leqslant k_1<\ldots<k_l\leqslant n}\mathbf{P}(A_{k_1}\ldots A_{k_l}).\end{aligned} \tag{1.12}$$

In all the problems given in Sec. 1.1 it is assumed that the simple events are equiprobable; the terms 'accidentally', 'chosen at random' refer to the equiprobability of simple events. The expression 'the point is uniformly distributed on the set Ω' in Sec. 1.2 means that the probabilities must be calculated by formula (1.8).

1. Classical Definition of Probability

1.1. A box contains three tickets numbered 1, 2, and 3. The tickets are drawn from the box one at a time. Assume that all sequences of the numbers of the tickets are equiprobable. Find the probability that the ordinal number of at least one ticket coincides with its own number.

1.2. A deck of 36 playing cards is well shuffled (that is, all possible arrangements of cards are equiprobable). Find the probabilities of the following events:

A = {four aces are dealt in succession},
B = {the places, occupied by the aces form an arithmetic progression with a common difference 7}.

1.3. There is a three-volume dictionary among 40 books arranged on a shelf in random order. Find the probability of these volumes standing in increasing order from left to right (the volumes are not necessarily side-by-side).

1.4. Three coins are tossed. Assuming the simple events to be equiprobable, find the probabilities of the following events:

A = {the first coin comes up heads},
B = {exactly two heads have occurred},
C = {not more than two heads have occurred}.

1.5. One sequence is chosen at random from the set of all sequences of length consisting of the numbers 0, 1, 2. Find the probabilities of the following events:

A = {the sequence begins with 0},
B = {the sequence contains exactly $m + 2$ zeros, two of them being at the end-points of the sequence},
C = {the sequence contains exactly m unities},
D = {the sequence contains exactly m_0 zeros, m_1 unities, and m_2 twos}.

1.6. Two domino pieces are chosen at random from 28 pieces. Find the probability P_2 that a chain can be formed from them in accordance with the rules of the game.

1.7. The last three digits of a telephone number beginning 135-3.-.. have been erased. Assuming that all combinations of the last three digits are equiprobable, find the probabilities of the following events:

A = {distinct digits, different from 1, 3, 5, have been erased},
B = {identical digits have been erased},
C = {two of the missing digits coincide}.

1.8. What are the chances that a four-digit number on the licence plate of a car chosen at random in a city: (a) consists of different digits? (b) includes only two identical digits? (c) includes two pairs of identical digits? (d) contains only three identical digits? (e) consists of identical digits?

1.9. Find the probability p_N that a natural number chosen at random from the set $\{1, 2, \ldots, N\}$ is divisible by a fixed natural number k. Find $\lim_{N\to\infty} p_N$.

1.10. A number a is chosen at random from the numbers $\{1, 2, \ldots, N\}$. Find the probability p_N that: (a) the number a is not divisible either by a_1 or by a_2, where a_1 and a_2 are fixed natural coprime numbers; (b) the number a is not divisible by either of the numbers $a_1, a_2, \ldots, a_k$, where the numbers a_i are natural pairwise coprime numbers. Find $\lim\limits_{N\to\infty} p_N$ in cases (a) and (b).

1.11. A number a is chosen at random from the set $\{1, 2, \ldots, N\}$. Find $\lim\limits_{N\to\infty} p_N$, where p_N is the probability that $a^2 - 1$ is divisible by 10.

1.12. A number a is chosen at random from the set $\{1, 2, \ldots, N\}$. Find the probability p_N that when divided by the integer $r \geqslant 1$, a will leave a remainder q. Find $\lim\limits_{N\to\infty} p_N$.

1.13. An integer ξ is chosen at random from the set $\{0, 1, 2, \ldots, 10^n - 1\}$. Find the probability that in decimal notation this number is a k-digit number, i.e. it can be represented as $\xi = \xi_k \cdot 10^{k-1} + \xi_{k-1} \cdot 10^{k-2} + \ldots + \xi_2 \cdot 10 + \xi_1$, where $0 \leqslant \xi_i \leqslant 9$ for all $i = 1, \ldots, k$ and $\xi_k > 0$ $(k \geqslant 1)$.

1.14. The numbers ξ and η are chosen at random from the set of natural numbers $\{1, 2, \ldots, N\}$ with replacement. Find the probability q_N that ξ and η are coprime numbers. Find $\lim\limits_{N\to\infty} q_N$ using the familiar equation $\sum\limits_{n=1}^{\infty} \frac{1}{n_2} = \frac{\pi^2}{6}$.

1.15. The numbers ξ and η are chosen at random from the set of integers $\{1, 2, \ldots, N\}$ with replacement. If the probability of the event $\xi^2 + \eta^2 \leqslant N^2$ is p_N, find $\lim\limits_{N\to\infty} p_N$.

1.16. The numbers ξ and η are chosen at random from the set of integers $\{0, 1, 2, \ldots, 10^n - 1\}$ with replacement. Let p_m be the probability that the sum $\xi + \eta$ is an m-digit natural number in decimal notation. Find the probabilities p_{n-k+1}, $k = 0, 1, \ldots, n$, and $q_k = \lim\limits_{N\to\infty} p_{n-k+1}$, $k = 0, 1, 2, \ldots$.

1.17. The numbers ξ and η are chosen at random from the set of integers $\{0, 1, 2, \ldots, 10^n - 1\}$ with replacement. Let p_m be the probability that the product $\xi\eta$ is an m-digit natural number in decimal notation. Find $q_k = \lim\limits_{N\to\infty} p_{2n-k}$, $k = 0, 1, 2, \ldots$.

1.18. Show that in problems 1.14-1.17 the limit probabilities will remain the same if the numbers ξ and η are chosen at random from a single set without replacement.

1.19*. The number X and Y are chosen at random from the set of natural numbers $\{1, 2, \ldots, N\}$, $N \geqslant 3$ with replacement. Find which is greater:

$$P_2 = \mathbf{P}\,\{X^2 - Y^2 \text{ is divisible by } 2\}$$

or

$$P_3 = \mathbf{P}\,\{X^2 - Y^2 \text{ is divisible by } 3\}.$$

1.20. The numbers X and Y are chosen at random from the set of natural numbers $\{1, 2, \ldots, N\}$, $N \geqslant 6$ with replacement. Show that

$$\mathbf{P}\,\{X^4 - Y^4 \equiv 0 \pmod 2\} < \mathbf{P}\,\{X^4 - Y^4 \equiv 0 \pmod 3\} < \mathbf{P}\,\{X^4 - Y^4 \equiv 0 \pmod 5\}.$$

1.21. The numbers X and Y are chosen at random from the set $\{1, 2, \ldots, N\}$ with replacement. Using Fermat's *little* theorem (*if p is a prime number and the integer a is not divisible by p, then* $a^{p-1} \equiv 1 \pmod p$), find the probability $Q_N(p)$ that the number $X^{p-1} - Y^{p-1}$ is divisible by the prime number p. Find $\lim\limits_{N\to\infty} Q_N(p) = Q(p)$, $\lim\limits_{P,N\to\infty} Q_N(p) = Q$.

1.22*. The numbers X and Y are chosen at random from the set $\{1, 2, \ldots, N\}$ with replacement. Show that when $N \geqslant 4$,

$$\mathbf{P}\,\{X^3 + Y^3 \equiv 0 \pmod 3\} < \mathbf{P}\,\{X^3 + Y^3 \equiv 0 \pmod 7\}.$$

1.23. The sets A_1 and A_2 are chosen with replacement from the collection of all subsets of the set $S = \{1, 2, \ldots, N\}$. Find the probability that $A_1 \cap A_2 = \varnothing$.

1.24. The subsets $A_1, A_2, \ldots, A_r$ are chosen with replacement from all subsets of the set $S = \{1, 2, \ldots, N\}$. Find the probability that the sets $A_1, A_2, \ldots, A_r$ do not intersect pair-by-pair.

1.25*. An urn contains $(2n + 1)^2$ cards labelled with an ordered pair of integers (x, y) (x and y assume values from $-n$ to n and each pair of numbers is written on exactly one card). Three cards (ξ_1, η_1), (ξ_2, η_2), (ξ_3, η_3) are drawn from the urn without replacement. Let us consider these pairs as the coordinates of the random points Ξ_1, Ξ_2, Ξ_3 of a plane in a Cartesian system of coordinates. Find the probability p_n that Ξ_1 is symmetric with respect to Ξ_2 about Ξ_3.

1.26. Ten dice are tossed. Assume that all combinations of spots are equiprobable. Find the probability of the following events:

(a) not a single 6 is obtained;
(b) exactly three 6 are obtained;
(c) at least one 6 is obtained.

1.27. Some people consider a six-digit number on a tram or bus ticket to be lucky if the sum of its first three digits is equal to the sum of the last three digits. Find the probability of obtaining a "lucky" ticket.

1.28 (see problem 1.27). Calculate the probability of obtaining at least one lucky ticket when n successive tickets are bought, $1 < n < 9$.

1.29. Ten different numbers are selected at random from 30 numbers (1, 2, . . ., 29, 30). Find the probabilities of the following events:

A = {all the numbers are odd},
B = {exactly 5 numbers are divisible by 3},
C = {5 numbers are even and 5 numbers are odd, exactly one number being divisible by 10}.

1.30. From an urn containing M_1 balls labelled with the number 1, M_2 balls with number 2, . . ., and M_N balls with the number N, we draw n balls at random without replacement. Find the probabilities of the following events:

(a) we draw m_1 balls labelled with the number 1, m_2 balls with the number 2, . . ., and m_N balls with the number N;
(b) we draw each of the N numbers at least once.

1.31. The numbers ξ_1 and ξ_2 are chosen from the set of numbers $\{1, 2, \ldots, N\}$ without replacement. Find $\mathsf{P}\{\xi_2 > \xi_1\}$. When selecting three numbers, find the probability that the second number lies between the first and the third.

1.32. From the set of numbers $\{1, 2, \ldots, N\}$ we chose n different numbers without replacement. Arrange them in increasing order: $z_{(1)} < z_{(2)} < \ldots < z_{(n)}$. Find the probability that $z_{(m)} \leqslant M < z_{(m+1)}$; calculate its limit for N $M \to \infty$, $M/N \to \alpha \in [0, 1]$.

1.33. From the set $\{1, 2, \ldots, N\}$ we chose at random $k+1$ numbers $x_1, x_2, \ldots, x_{k+1}$ without replacement. The first k numbers, arranged in increasing order, are designated $x_{(1)} < x_{(2)} < \ldots < x_{(k)}$. Find

$$\mathsf{P}\{x_{(l)} < x_{k+1} < x_{(l+1)}\}.$$

1.34. Ten manuscripts are arranged in 30 files (3 files for one manuscript). Find the probability that no 6 files selected at random contain an entire manuscript.

1.35. The number of guests seated in random order at a round table is $2n$. What is the probability that the guests can be divided into n nonintersecting pairs so that each pair consists of a man and a woman sittig side-by-side?

1.36. Every ticket of the "Sports lotto" lottery contains 49 different numbers. The ticket holder marks 6 numbers on each ticket. If his guess is correct, he wins the grand prize. If he guesses only 3 winning numbers out of 49, he gets a consolation prize. A ticket holder marks numbers 4, 12, 20, 31, 32, 33 on the first ticket and numbers 4, 12, 20, 41, 42, 43 on the second ticket. What are the chances that the ticket holder wins exactly two consolation prizes?

1.37. In *an equiprobable scheme for arranging particles in cells*, the numbers of cells consecutively occupied by the particles are obtained by random choice with replacement.

Let us designate as $\mu_r = \mu_r(n, N)$ the number of cells containing exactly r particles each after n particles have been arranged in N cells. Find the probabilities of the following events:

(a) $\mu_0(n, N) > 0$ (for $n = N$);
(b) $\mu_0(n, N) = 0$ (for $n = N + 1$);
(c) $\mu_0(n, N) = 1$ (for $n = N + 1$);
(d) there is a cell containing at least two particles (for any ratio of n and N).

1.38 (see problem 1.37.) Find $\mathsf{P}\{\mu_0(n, N) = 0\}$ for arbitrary n, N.

1.39. We arrange at random n mutually indistinguishable particles in N mutually distinguishable cells. (The simple events are ordered samples of numbers $(r_1, r_2, \ldots, r_N)$, where r_k is the number of particles in the kth cell, $k = 1, 2, \ldots, N$.) Find the probabilities of the following events:

(a) $\mu_0(n, N) > 0$;
(b) $\mu_0(n, N) = 1$.

1.40. There are n people in the first row of a theatre. The row contains N seats. Assuming that all possible ar-