

MATHEMATICAL METHODS OF STATISTICS

By

HARALD CRAMÉR

PROFESSOR IN THE UNIVERSITY
OF STOCKHOLM

1946

PREFACE.

During the last 25 years, statistical science has made great progress, thanks to the brilliant schools of British and American statisticians, among whom the name of Professor R. A. Fisher should be mentioned in the foremost place. During the same time, largely owing to the work of French and Russian mathematicians, the classical calculus of probability has developed into a purely mathematical theory satisfying modern standards with respect to rigour.

The purpose of the present work is to join these two lines of development in an exposition of the mathematical theory of modern statistical methods, in so far as these are based on the concept of probability. A full understanding of the theory of these methods requires a fairly advanced knowledge of pure mathematics. In this respect, I have tried to make the book self-contained from the point of view of a reader possessing a good working knowledge of the elements of the differential and integral calculus, algebra, and analytic geometry.

In the first part of the book, which serves as a mathematical introduction, the requisite mathematics not assumed to be previously known to the reader are developed. Particular stress has been laid on the fundamental concepts of a distribution, and of the integration with respect to a distribution. As a preliminary to the introduction of these concepts, the theory of Lebesgue measure and integration has been briefly developed in Chapters 4—5, and the fundamental concepts are then introduced by straightforward generalization in Chapters 6—7.

The second part of the book contains the general theory of random variables and probability distributions, while the third part is devoted to the theory of sampling distributions, statistical estimation, and tests of significance. The selection of the questions treated in the last part is necessarily somewhat arbitrary, but I have tried to concentrate in the first hand on points of general importance. When these are fully mastered, the reader will be able to work out applications to particular problems for himself. In order to keep the volume

of the book within reasonable limits, it has been necessary to exclude certain topics of great interest, which I had originally intended to treat, such as the theory of random processes, statistical time series and periodograms.

The theory of the statistical tests is illustrated by numerical examples borrowed from various fields of application. Owing to considerations of space, it has been necessary to reduce the number of these examples rather severely. It has also been necessary to restrain from every discussion of questions concerning the practical arrangement of numerical calculations.

It is not necessary to go through the first part completely before studying the rest of the book. A reader who is anxious to find himself *in medias res* may content himself with making some slight acquaintance with the fundamental concepts referred to above. For this purpose, it will be advisable to read Chapters 1—3, and the paragraphs 4.1—4.2, 5.1—5.3, 6.1—6.2, 6.4—6.6, 7.1—7.2, 7.4—7.5 and 8.1—8.4. The reader may then proceed to Chapter 13, and look up the references to the first part as they occur.

The book is founded on my University lectures since about 1930, and has been written mainly during the years 1942—1944. Owing to war conditions, foreign scientific literature was during these years only very incompletely and with considerable delay available in Sweden, and this must serve as an excuse for the possible absence of quotations which would otherwise have been appropriate.

The printing of the Scandinavian edition of the book has been made possible by grants from the Royal Swedish Academy of Science, and from Stiftelsen Lars Hiertas Minne. I express my gratitude towards these institutions.

My thanks are also due to the Editors of the Princeton Mathematical Series for their kind offer to include the book in the Series, and for their permission to print a separate Scandinavian edition.

I am further indebted to Professor R. A. Fisher and to Messrs Oliver and Boyd for permission to reprint tables of the t - and χ^2 distributions from »Statistical methods for research workers».

A number of friends have rendered valuable help during the preparation of the book. Professors Harald Bohr and Ernst Jacobsthal, taking refuge in Sweden from the hardships of the times, have read parts of the work in manuscript and in proof, and have given stimulating criticism and advice. Professor Herman Wold has made a very careful scrutiny of the whole work in proof, and I have greatly profited

from his valuable remarks. Gösta Almqvist, Jan Jung, Sven G. Lindblom and Bertil Matérn have assisted in the numerical calculations, the revision of the manuscript, and the reading of the proofs. To all these I wish to express my sincere thanks.

Department of Mathematical Statistics
University of Stockholm

May 1945

H. C.

TABLE OF CONTENTS.

First Part.

MATHEMATICAL INTRODUCTION.

CHAPTERS 1—3. SETS OF POINTS.

	Page
Chapter 1. General properties of sets	3
1. Sets. — 2. Subsets, space. — 3. Operations on sets. — 4. Sequences of sets. — 5. Monotone sequences. — 6. Additive classes of sets.	
Chapter 2. Linear point sets	10
1. Intervals. — 2. Various properties of sets in R_1 . — 3. Borel sets.	
Chapter 3. Point sets in n dimensions	15
1. Intervals. — 2. Various properties of sets in R_n . — 3. Borel sets. — 4. Linear sets. — 5. Subspace, product space.	
References to chapters 1—3	18

CHAPTERS 4—7. THEORY OF MEASURE AND INTEGRATION IN R_1 .

Chapter 4. The Lebesgue measure of a linear point set	19
1. Length of an interval. — 2. Generalization. — 3. The measure of a sum of intervals. — 4. Outer and inner measure of a bounded set. — 5. Measurable sets and Lebesgue measure. — 6. The class of measurable sets. — 7. Measurable sets and Borel sets.	
Chapter 5. The Lebesgue integral for functions of one variable.	33
1. The integral of a bounded function over a set of finite measure. — 2. B -measurable functions. — 3. Properties of the integral. — 4. The integral of an unbounded function over a set of finite measure. — 5. The integral over a set of infinite measure. — 6. The Lebesgue integral as an additive set function.	
Chapter 6. Non-negative additive set functions in R_1	48
1. Generalization of the Lebesgue measure and the Lebesgue integral. — 2. Set functions and point functions. — 3. Construction of a set function. — 4. P -measure. — 5. Bounded set functions. — 6. Distributions. — 7. Sequences of distributions. — 8. A convergence theorem.	

	Page
Chapter 7. The Lebesgue-Stieltjes integral for functions of one variable	62
1. The integral of a bounded function over a set of finite P -measure. — 2. Unbounded functions and sets of infinite P -measure. — 3. Lebesgue-Stieltjes integrals with a parameter. — 4. Lebesgue-Stieltjes integrals with respect to a distribution. — 5. The Riemann-Stieltjes integral.	
References to chapters 4—7	75

CHAPTERS 8—9. THEORY OF MEASURE AND INTEGRATION IN R_n .

Chapter 8. Lebesgue measure and other additive set functions in R_n	76
1. Lebesgue measure in R_n . — 2. Non-negative additive set functions in R_n . — 3. Bounded set functions. — 4. Distributions. — 5. Sequences of distributions. — 6. Distributions in a product space.	
Chapter 9. The Lebesgue-Stieltjes integral for functions of n variables	85
1. The Lebesgue-Stieltjes integral. — 2. Lebesgue-Stieltjes integrals with respect to a distribution. — 3. A theorem on repeated integrals. — 4. The Riemann-Stieltjes integral. — 5. The Schwarz inequality.	

CHAPTERS 10—12. VARIOUS QUESTIONS.

Chapter 10. Fourier integrals	89
1. The characteristic function of a distribution in R_1 . — 2. Some auxiliary functions. — 3. Uniqueness theorem for characteristic functions in R_1 . — 4. Continuity theorem for characteristic functions in R_1 . — 5. Some particular integrals. — 6. The characteristic function of a distribution in R_n . — 7. Continuity theorem for characteristic functions in R_n .	
Chapter 11. Matrices, determinants and quadratic forms	103
1. Matrices. — 2. Vectors. — 3. Matrix notation for linear transformations. — 4. Matrix notation for bilinear and quadratic forms. — 5. Determinants. — 6. Rank. — 7. Adjugate and reciprocal matrices. — 8. Linear equations. — 9. Orthogonal matrices. Characteristic numbers. — 10. Non-negative quadratic forms. — 11. Decomposition of Σx_i^2 . — 12. Some integral formulae.	
Chapter 12. Miscellaneous complements	122
1. The symbols O , o and ∞ . — 2. The Euler-MacLaurin sum formula. — 3. The Gamma function. — 4. The Beta function. — 5. Stirling's formula. — 6. Orthogonal polynomials.	

Second Part.

RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS.

CHAPTERS 13—14. FOUNDATIONS.

	Page
Chapter 13. Statistics and probability	137
1. Random experiments. — 2. Examples. — 3. Statistical regularity. — 4. Object of a mathematical theory. — 5. Mathematical probability.	
Chapter 14. Fundamental definitions and axioms	151
1. Random variables. (Axioms 1—2.) — 2. Combined variables. (Axiom 3.) — 3. Conditional distributions. — 4. Independent variables. — 5. Functions of random variables. — 6. Conclusion.	

CHAPTERS 15—20. VARIABLES AND DISTRIBUTIONS IN R_1 .

Chapter 15. General properties	166
1. Distribution function and frequency function. — 2. Two simple types of distributions. — 3. Mean values. — 4. Moments. — 5. Measures of location. — 6. Measures of dispersion. — 7. Tchebycheff's theorem. — 8. Measures of skewness and excess. — 9. Characteristic functions. — 10. Semi-invariants. — 11. Independent variables. — 12. Addition of independent variables.	
Chapter 16. Various discrete distributions	192
1. The function $\epsilon(x)$. — 2. The binomial distribution. — 3. Bernoulli's theorem. — 4. De Moivre's theorem. — 5. The Poisson distribution. — 6. The generalized binomial distribution of Poisson.	
Chapter 17. The normal distribution	208
1. The normal functions. — 2. The normal distribution. — 3. Addition of independent normal variables. — 4. The central limit theorem. — 5. Complementary remarks to the central limit theorem. — 6. Orthogonal expansion derived from the normal distribution. — 7. Asymptotic expansion derived from the normal distribution. — 8. The rôle of the normal distribution in statistics.	
Chapter 18. Various distributions related to the normal	233
1. The χ^2 -distribution. — 2. Student's distribution. — 3. Fisher's z -distribution. — 4. The Beta distribution.	
Chapter 19. Further continuous distributions	244
1. The rectangular distribution. — 2. Cauchy's and Laplace's distributions. — 3. Truncated distributions. — 4. The Pearson system.	

	Page
Chapter 20. Some convergence theorems.....	250
1. Convergence of distributions and variables. — 2. Convergence of certain distributions to the normal. — 3. Convergence in probability. — 4. Tchebycheff's theorem. — 5. Khintchine's theorem. — 6. A convergence theorem.	
Exercises to chapters 15—20	255

CHAPTERS 21—24. VARIABLES AND DISTRIBUTIONS IN R_n .

Chapter 21. The two-dimensional case.....	260
1. Two simple types of distributions. — 2. Mean values, moments. — 3. Characteristic functions. — 4. Conditional distributions. — 5. Regression, I. — 6. Regression, II. — 7. The correlation coefficient. — 8. Linear transformation of variables. — 9. The correlation ratio and the mean square contingency. — 10. The ellipse of concentration. — 11. Addition of independent variables. — 12. The normal distribution.	
Chapter 22. General properties of distributions in R_n	291
1. Two simple types of distributions. Conditional distributions. — 2. Change of variables in a continuous distribution. — 3. Mean values, moments. — 4. Characteristic functions. — 5. Rank of a distribution. — 6. Linear transformation of variables. — 7. The ellipsoid of concentration.	
Chapter 23. Regression and correlation in n variables.....	301
1. Regression surfaces. — 2. Linear mean square regression. — 3. Residuals. — 4. Partial correlation. — 5. The multiple correlation coefficient. — 6. Orthogonal mean square regression.	
Chapter 24. The normal distribution	310
1. The characteristic function. — 2. The non-singular normal distribution. — 3. The singular normal distribution. — 4. Linear transformation of normally distributed variables. — 5. Distribution of a sum of squares. — 6. Conditional distributions. — 7. Addition of independent variables. The central limit theorem.	
Exercises to chapters 21—24	317

Third Part.

STATISTICAL INFERENCE.

CHAPTERS 25—26. GENERALITIES.

Chapter 25. Preliminary notions on sampling	323
1. Introductory remarks. — 2. Simple random sampling. — 3. The distribution of the sample. — 4. The sample values as random variables. Sampling	

distributions. — 5. Statistical image of a distribution. — 6. Biased sampling. Random sampling numbers. — 7. Sampling without replacement. The representative method.

Chapter 26. Statistical inference 332

1. Introductory remarks. — 2. Agreement between theory and facts. Tests of significance. — 3. Description. — 4. Analysis. — 5. Prediction.

CHAPTERS 27—29. SAMPLING DISTRIBUTIONS.

Chapter 27. Characteristics of sampling distributions 341

1. Notations. — 2. The sample mean \bar{x} . — 3. The moments a_r . — 4. The variance m_2 . — 5. Higher central moments and semi-invariants. — 6. Unbiased estimates. — 7. Functions of moments. — 8. Characteristics of multi-dimensional distributions. — 9. Corrections for grouping.

Chapter 28. Asymptotic properties of sampling distributions . . 363

1. Introductory remarks. — 2. The moments. — 3. The central moments. — 4. Functions of moments. — 5. The quantiles. — 6. The extreme values and the range.

Chapter 29. Exact sampling distributions 378

1. The problem. — 2. Fisher's lemma. Degrees of freedom. — 3. The joint distribution of \bar{x} and s^2 in samples from a normal distribution. — 4. Student's ratio. — 5. A lemma. — 6. Sampling from a two-dimensional normal distribution. — 7. The correlation coefficient. — 8. The regression coefficients. — 9. Sampling from a k -dimensional normal distribution. — 10. The generalized variance. — 11. The generalized Student ratio. — 12. Regression coefficients. — 13. Partial and multiple correlation coefficients.

CHAPTERS 30—31. TESTS OF SIGNIFICANCE, I.

Chapter 30. Tests of goodness of fit and allied tests 416

1. The χ^2 test in the case of a completely specified hypothetical distribution. — 2. Examples. — 3. The χ^2 test when certain parameters are estimated from the sample. — 4. Examples. — 5. Contingency tables. — 6. χ^2 as a test of homogeneity. — 7. Criterion of differential death rates. — 8. Further tests of goodness of fit.

Chapter 31. Tests of significance for parameters 452

1. Tests based on standard errors. — 2. Tests based on exact distributions. — 3. Examples.

CHAPTERS 32—34. THEORY OF ESTIMATION.

Chapter 32. Classification of estimates 473

1. The problem. — 2. Two lemmas. — 3. Minimum variance of an estimate.

Efficient estimates. — 4. Sufficient estimates. — 5. Asymptotically efficient estimates. — 6. The case of two unknown parameters. — 7. Several unknown parameters. — 8. Generalization.

Chapter 33. Methods of estimation	497
1. The method of moments. — 2. The method of maximum likelihood. — 3. Asymptotic properties of maximum likelihood estimates. — 4. The χ^2 minimum method.	
Chapter 34. Confidence regions	507
1. Introductory remarks. — 2. A single unknown parameter. — 3. The general case. — 4. Examples.	

CHAPTERS 35—37. TESTS OF SIGNIFICANCE, II.

Chapter 35. General theory of testing statistical hypotheses	525
1. The choice of a test of significance. — 2. Simple and composite hypotheses. — 3. Tests of simple hypotheses. Most powerful tests. — 4. Unbiased tests. — 5. Tests of composite hypotheses.	
Chapter 36. Analysis of variance	536
1. Variability of mean values. — 2. Simple grouping of variables. — 3. Generalization. — 4. Randomized blocks. — 5. Latin squares.	
Chapter 37. Some regression problems	548
1. Problems involving non-random variables. — 2. Simple regression. — 3. Multiple regression. — 4. Further regression problems.	

TABLES 1—2. THE NORMAL DISTRIBUTION	557
TABLE 3. THE χ^2-DISTRIBUTION	559
TABLE 4. THE t-DISTRIBUTION	560
LIST OF REFERENCES	561
INDEX	571

FIRST PART

MATHEMATICAL INTRODUCTION

CHAPTERS 1-3. SETS OF POINTS.

CHAPTER 1.

GENERAL PROPERTIES OF SETS.

1.1. Sets. — In pure and applied mathematics, situations often occur where we have to consider the collection of all possible objects having certain specified properties. Any collection of objects defined in this way will be called a *set*, and each object belonging to such a set will be called an *element of the set*.

The elements of a set may be objects of any kind: points, numbers, functions, things, persons etc. Thus we may consider e.g. 1) the set of all positive integral numbers, 2) the set of all points on a given straight line, 3) the set of all rational functions of two variables, 4) the set of all persons born in a given country and alive at the end of the year 1940. In the first part of this book we shall mainly deal with cases where the elements are points or numbers, but in this introductory chapter we shall give some considerations which apply to the general case when the elements may be of any kind.

In the example 4) given above, our set contains a finite, though possibly unknown, number of elements, whereas in the three first examples we obviously have to do with sets where the number of elements is not finite. We thus have to distinguish between *finite* and *infinite* sets.

An infinite set is called *enumerable* if its elements may be arranged in a *sequence*: $x_1, x_2, \dots, x_n, \dots$, such that a) every x_n is an element of the set, and b) every element of the set appears at a definite place in the sequence. By such an arrangement we establish a *one-to-one correspondence* between the elements of the given set and those of the set containing all positive integral numbers $1, 2, \dots, n, \dots$, which forms the simplest example of an enumerable set.

We shall see later that there exist also infinite sets which are *non-enumerable*. If, from such a set, we choose any sequence of elements x_1, x_2, \dots , there will always be elements left in the set which do not appear in the sequence, so that a non-enumerable set may be

said to represent a higher order of infinity than an enumerable set. It will be shown later (cf 4. 3) that the set of all points on a given straight line affords an example of a non-enumerable set.

1.2. Subsets, space. — If two sets S and S_1 are such that every element of S_1 also belongs to S , we shall call S_1 a *subset* of S , and write

$$S_1 < S \quad \text{or} \quad S > S_1.$$

We shall sometimes express this also by saying that S_1 is *contained in* S or *belongs to* S . — When S_1 consists of one single element x , we use the same notation $x < S$ to express that x *belongs to* S .

In the particular case when both the relations $S_1 < S$ and $S < S_1$ hold, the sets are called *equal*, and we write

$$S = S_1.$$

It is sometimes convenient to consider a set S which does not contain any element at all. This we call the *empty set*, and write $S = 0$. The empty set is a subset of any set. If we regard the empty set as a particular case of a finite set, it is seen that *every subset of a finite set is itself finite, while every subset of an enumerable set is finite or enumerable*. Thus the set of all integers between 20 and 30 is a finite subset of the set 1, 2, 3, . . . , while the set of all odd integers 1, 3, 5, . . . is an enumerable subset of the same set.

In many investigations we shall be concerned with the properties and the mutual relations of various subsets of a given set S . The set S , which thus contains the totality of all elements that may appear in the investigation, will then be called the *space* of the investigation. If, e. g., we consider various sets of points on a given straight line, we may choose as our space the set S of all points on the line. Any subset S of the space S will be called briefly a *set in* S .

1.3. Operations on sets. — Suppose now that a space S is given, and let us consider various sets in S . We shall first define the operations of *addition*, *multiplication* and *subtraction* for sets.

The *sum* of two sets S_1 and S_2 ,

$$S' = S_1 + S_2,$$

is the set S' of all elements *belonging to at least one* of the sets S_1 and S_2 . — The *product*

$$S'' = S_1 S_2$$

is the *common part* of the sets, or the set S'' of all elements *belonging to both* S_1 and S_2 . — Finally, the *difference*

$$S''' = S_1 - S_2$$

will be defined only in the case when S_2 is a subset of S_1 , and is then the set S''' of all elements *belonging to* S_1 *but not to* S_2 .

Thus if S_1 and S_2 consist of all points inside the curves C_1 and C_2 respectively (cf Fig. 1), $S_1 + S_2$ will be the set of all points inside at least one of the two curves, while $S_1 S_2$ will be the set of all points common to both domains.

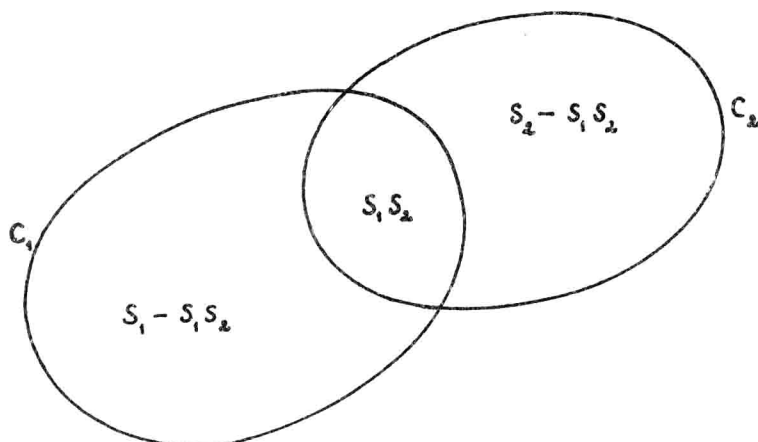


Fig. 1. Simple operations on sets.

The product $S_1 S_2$ is evidently a subset of both S_1 and S_2 . The difference $S_n - S_1 S_2$, where n may denote 1 or 2, is the set of all points of S_n which do not belong to $S_1 S_2$.

In the particular case when S_1 and S_2 have no common elements, the *product* is empty, so that we have $S_1 S_2 = 0$. On the other hand, if $S_1 = S_2$ the *difference* is empty, and we have $S_1 - S_2 = 0$.

In the particular case when S_2 is a subset of S_1 we have $S_1 + S_2 = S_1$ and $S_1 S_2 = S_2$.

It follows from the symmetrical character of our definitions of the sum and the product that the operations of addition and multiplication are *commutative*, i. e. that we have

$$S_1 + S_2 = S_2 + S_1 \quad \text{and} \quad S_1 S_2 = S_2 S_1.$$

1.3

Further, a moment's reflection will show that these operations are also *associative* and *distributive*, like the corresponding arithmetic operations. We thus have

$$\begin{aligned}(S_1 + S_2) + S_3 &= S_1 + (S_2 + S_3), \\ (S_1 S_2) S_3 &= S_1 (S_2 S_3), \\ S_1 (S_2 + S_3) &= S_1 S_2 + S_1 S_3.\end{aligned}$$

It follows that we may without ambiguity talk of the sum or product of any finite number of sets:

$$S_1 + S_2 + \cdots + S_n \quad \text{and} \quad S_1 S_2 \cdots S_n,$$

where the order of terms and factors is arbitrary.

We may even extend the definition of these two operations to an enumerable sequence of terms or factors. Thus, given a sequence S_1, S_2, \dots of sets in S , we define the sum

$$\sum_1^{\infty} S_v = S_1 + S_2 + \cdots$$

as the set of all elements *belonging to at least one* of the sets S_v , while the product

$$\prod_1^{\infty} S_v = S_1 S_2 \cdots$$

is the set of all elements *belonging to all* S_v . — We then have, e.g., $S(S_1 + S_2 + \cdots) = SS_1 + SS_2 + \cdots$.

Thus if S_v denotes the set of all real numbers x such that $\frac{1}{v+1} \leq x \leq \frac{1}{v}$, we find that $\sum_1^{\infty} S_v$ will be the set of all x such that $0 < x \leq 1$, while the product set will be empty, $\prod_1^{\infty} S_v = 0$. — On the other hand, if S_v denotes the set of all x such that $0 \leq x \leq \frac{1}{v}$, the sum $\sum_1^{\infty} S_v$ will coincide with S_1 , while the product $\prod_1^{\infty} S_v$ will be a set containing one single element, viz. the number $x = 0$.

For the operation of subtraction, an important particular case arises when S_1 coincides with the whole space S . The difference

$$S^* = S - S$$

is the set of all elements of our space which do not belong to S , and will be called the *complementary set* or simply the *complement* of S . We obviously have $S + S^* = S$, $SS^* = 0$, and $(S^*)^* = S$.

It is important to observe that the complement of a given set S is relative to the space S in which S is considered. If our space is the set of all points on a given straight line L , and if S is the set of all points situated on the positive side of an origin O on this line, the complement S^* will consist of O itself and all points on the negative side of O . If, on the other hand, our space consists of all points in a certain plane P containing L , the complement S^* of the same set S will also include all points of P not belonging to L . — In all cases where there might be a risk of a mistake, we shall use the expression: S^* is the complement of S with respect to S .

The operations of addition and multiplication may be brought into relation with one another by means of the concept of complementary sets. We have, in fact, for any finite or enumerable sequence S_1, S_2, \dots the relations

$$(1.3.1) \quad \begin{aligned} (S_1 + S_2 + \dots)^* &= S_1^* S_2^* \dots, \\ (S_1 S_2 \dots)^* &= S_1^* + S_2^* + \dots. \end{aligned}$$

The first relation expresses that *the complementary set of a sum is the product of the complements of the terms*. This is a direct consequence of the definitions. As a matter of fact, the complement $(S_1 + \dots)^*$ is the set of all elements x of the space, of which it is not true that they occur in at least one S_v . This is, however, the same thing as the set of all elements x which are absent from every S_v , or the set of all x which belong to every complement S_v^* , i.e. the product $S_1^* S_2^* \dots$. The second relation is obtained from the first by substituting S_v^* for S_v . — For the operation of subtraction, we obtain by a similar argument the relation

$$(1.3.2) \quad S_1 - S_2 = S_1 S_2^*.$$

The reader will find that the understanding of relations such as (1.3.1) and (1.3.2) is materially simplified by the use of figures of the same type as Fig. 1.

1.4. Sequences of sets. — When we use the word *sequence* without further specification, it will be understood that we mean a *finite or*