Undergraduate Texts in Mathematics

Kai Lai Chung, Farid AitSahlia

Elementary Probability Theory

With Stochastic Processes and an Introduction to Mathematical Finance

Fourth Edition

初等概率论 第4版

Springer

世界图书出版公司
www.wpcbj.com.cn

图书在版编目（CIP）数据

初等概率论：第4版：英文/（美）钟开莱著.
—影印本.—北京：世界图书出版公司北京公司，2010.2
书名原文：Elementary Probability Theory 4th ed
ISBN 978-7-5100-0462-9

Ⅰ.①初… Ⅱ.①钟… Ⅲ.①概率论—高等学校—教材—英文
Ⅳ.①O211

中国版本图书馆CIP数据核字（2010）第010557号

书　　名：	Elementary Probability Theory 4th ed.
作　　者：	Kai Lai Chung，Farid AitSahlia
中 译 名：	初等概率论 第4版
责任编辑：	高蓉　刘慧
出 版 者：	世界图书出版公司北京公司
印 刷 者：	三河国英印务有限公司
发　　行：	世界图书出版公司北京公司（北京朝内大街137号　100010）
联系电话：	010-64021602，010-64015659
电子信箱：	kjb@wpcbj.com.cn
开　　本：	24开
印　　张：	17.5
版　　次：	2010年01月
版权登记：	图字：01-2009-1077
书　　号：	978-7-5100-0462-9/O·677　　**定　　价：** 45.00元

Preface to the Fourth Edition

In this edition two new chapters, 9 and 10, on mathematical finance are added. They are written by Dr. Farid AitSahlia, *ancien élève*, who has taught such a course and worked on the research staff of several industrial and financial institutions.

The new text begins with a meticulous account of the uncommon vocabulary and syntax of the financial world; its manifold options and actions, with consequent expectations and variations, in the marketplace. These are then expounded in clear, precise mathematical terms and treated by the methods of probability developed in the earlier chapters. Numerous graded and motivated examples and exercises are supplied to illustrate the applicability of the fundamental concepts and techniques to concrete financial problems. For the reader whose main interest is in finance, only a portion of the first eight chapters is a "prerequisite" for the study of the last two chapters. Further specific references may be scanned from the topics listed in the Index, then pursued in more detail.

I have taken this opportunity to fill a gap in Section 8.1 and to expand Appendix 3 to include a useful proposition on martingale stopped at an optional time. The latter notion plays a basic role in more advanced financial and other disciplines. However, the level of our compendium remains *elementary*, as befitting the title and scheme of this textbook. We have also included some up-to-date financial episodes to enliven, for the beginners, the stratified atmosphere of "strictly business". We are indebted to Ruth Williams, who read a draft of the new chapters with valuable suggestions for improvement; to Bernard Bru and Marc Barbut for information on the Pareto-Lévy laws originally designed for income distributions. It is hoped that a readable summary of this renowned work may be found in the new Appendix 4.

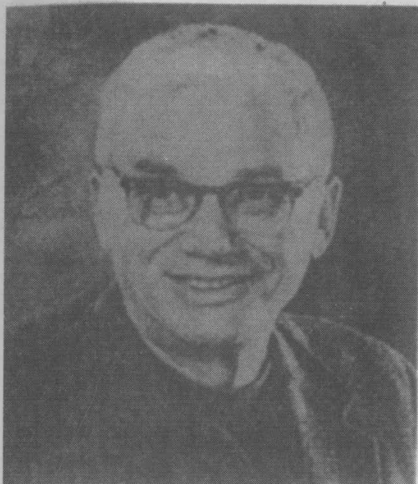
Kai Lai Chung
August 3, 2002

Prologue to Introduction to Mathematical Finance

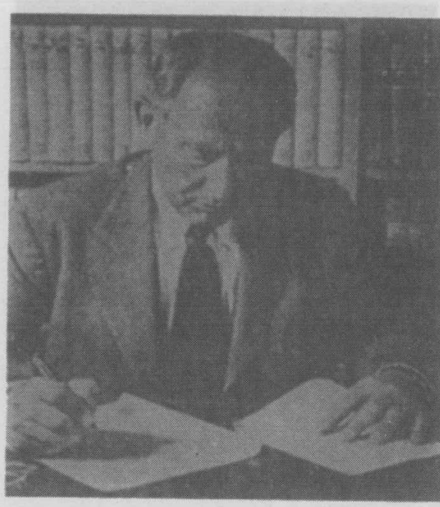
The two new chapters are self-contained introductions to the topics of mean-variance optimization and option pricing theory. The former covers a subject that is sometimes labeled "modern portfolio theory" and that is widely used by money managers employed by large financial institutions. To read this chapter, one only needs an elementary knowledge of probability concepts and a modest familiarity with calculus. Also included is an introductory discussion on stable laws in an applied context, an often neglected topic in elementary probability and finance texts. The latter chapter lays the foundations for option pricing theory, a subject that has fueled the development of finance into an advanced mathematical discipline as attested by the many recently published books on the subject. It is an initiation to martingale pricing theory, the mathematical expression of the so-called "arbitrage pricing theory", in the context of the binomial random walk. Despite its simplicity, this model captures the flavors of many advanced theoretical issues. It is often used in practice as a benchmark for the approximate pricing of complex financial instruments.

I would like to thank Professor Kai Lai Chung for inviting me to write the new material for the fourth edition. I would also like to thank my wife Unnur for her support during this rewarding experience.

Farid AitSahlia
November 1, 2002



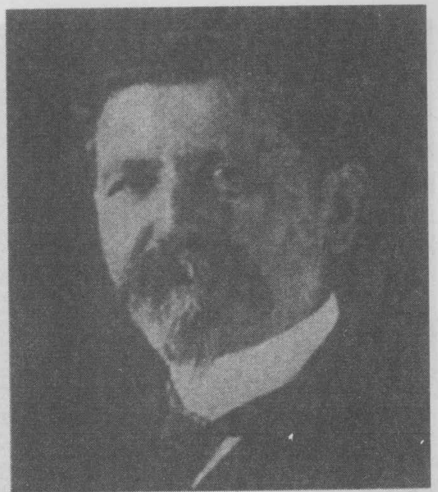
Doob



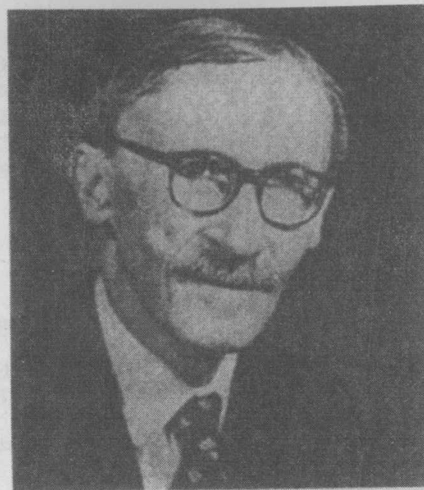
Polya

Kolmogorov

Cramer

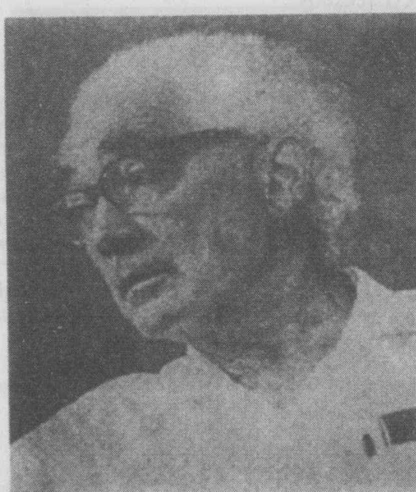


Borel



Levy

Keynes



Feller

Contents

PREFACE TO THE FOURTH EDITION xi

PROLOGUE TO INTRODUCTION TO MATHEMATICAL FINANCE xiii

1 SET 1

1.1 Sample sets 1
1.2 Operations with sets 3
1.3 Various relations 7
1.4 Indicator 13
Exercises 17

2 PROBABILITY 20

2.1 Examples of probability 20
2.2 Definition and illustrations 24
2.3 Deductions from the axioms 31
2.4 Independent events 35
2.5 Arithmetical density 39
Exercises 42

3 COUNTING 46

3.1 Fundamental rule 46
3.2 Diverse ways of sampling 49
3.3 Allocation models; binomial coefficients 55
3.4 How to solve it 62
Exercises 70

4 RANDOM VARIABLES 74
4.1 What is a random variable? 74
4.2 How do random variables come about? 78
4.3 Distribution and expectation 84
4.4 Integer-valued random variables 90
4.5 Random variables with densities 95
4.6 General case 105
Exercises 109

APPENDIX 1: BOREL FIELDS AND GENERAL RANDOM VARIABLES 115

5 CONDITIONING AND INDEPENDENCE 117
5.1 Examples of conditioning 117
5.2 Basic formulas 122
5.3 Sequential sampling 131
5.4 Pólya's urn scheme 136
5.5 Independence and relevance 141
5.6 Genetical models 152
Exercises 157

6 MEAN, VARIANCE, AND TRANSFORMS 164
6.1 Basic properties of expectation 164
6.2 The density case 169
6.3 Multiplication theorem; variance and covariance 173
6.4 Multinomial distribution 180
6.5 Generating function and the like 187
Exercises 195

7 POISSON AND NORMAL DISTRIBUTIONS 203
7.1 Models for Poisson distribution 203
7.2 Poisson process 211
7.3 From binomial to normal 222
7.4 Normal distribution 229
7.5 Central limit theorem 233
7.6 Law of large numbers 239
Exercises 246

APPENDIX 2: STIRLING'S FORMULA AND DE MOIVRE–LAPLACE'S THEOREM 251

8 FROM RANDOM WALKS TO MARKOV CHAINS 254

8.1 Problems of the wanderer or gambler 254
8.2 Limiting schemes 261
8.3 Transition probabilities 266
8.4 Basic structure of Markov chains 275
8.5 Further developments 284
8.6 Steady state 291
8.7 Winding up (or down?) 303
Exercises 314

APPENDIX 3: MARTINGALE 325

9 MEAN-VARIANCE PRICING MODEL 329

9.1 An investments primer 329
9.2 Asset return and risk 331
9.3 Portfolio allocation 335
9.4 Diversification 336
9.5 Mean-variance optimization 337
9.6 Asset return distributions 346
9.7 Stable probability distributions 348
Exercises 351

APPENDIX 4: PARETO AND STABLE LAWS 355

10 OPTION PRICING THEORY 359

10.1 Options basics 359
10.2 Arbitrage-free pricing: 1-period model 366
10.3 Arbitrage-free pricing: N-period model 372
10.4 Fundamental asset pricing theorems 376
Exercises 377

GENERAL REFERENCES 379

ANSWERS TO PROBLEMS 381

VALUES OF THE STANDARD NORMAL DISTRIBUTION FUNCTION 393

INDEX 397

1

Set

1.1. Sample sets

These days schoolchildren are taught about sets. A second grader* was asked to name "the set of girls in his class." This can be done by a complete list such as:

"Nancy, Florence, Sally, Judy, Ann, Barbara, ... "

A problem arises when there are duplicates. To distinguish between two Barbaras one must indicate their family names or call them B_1 and B_2. The same member cannot be counted twice in a set.

The notion of a set is common in all mathematics. For instance, in geometry one talks about "the set of points which are equidistant from a given point." This is called a circle. In algebra one talks about "the set of integers which have no other divisors except 1 and itself." This is called the set of prime numbers. In calculus the domain of definition of a function is a set of numbers, e.g., the interval (a, b); so is the range of a function if you remember what it means.

In probability theory the notion of a set plays a more fundamental role. Furthermore we are interested in very general kinds of sets as well as specific concrete ones. To begin with the latter kind, consider the following examples:

(a) a bushel of apples;
(b) fifty-five cancer patients under a certain medical treatment;

*My son Daniel.

(c) all the students in a college;
(d) all the oxygen molecules in a given container;
(e) all possible outcomes when six dice are rolled;
(f) all points on a target board.

Let us consider at the same time the following "smaller" sets:

(a′) the rotten apples in that bushel;
(b′) those patients who respond positively to the treatment;
(c′) the mathematics majors of that college;
(d′) those molecules that are traveling upwards;
(e′) those cases when the six dice show different faces;
(f′) the points in a little area called the "bull's-eye" on the board.

We shall set up a mathematical model for these and many more such examples that may come to mind, namely we shall abstract and generalize our intuitive notion of "a bunch of things." First we call the things points, then we call the bunch a space; we prefix them by the word "sample" to distinguish these terms from other usages, and also to allude to their statistical origin. Thus a *sample point* is the abstraction of an apple, a cancer patient, a student, a molecule, a possible chance outcome, or an ordinary geometrical point. The *sample space* consists of a number of sample points and is just a name for the totality or aggregate of them all. Any one of the examples (a)–(f) above can be taken to be a sample space, but so also may any one of the smaller sets in (a′)–(f′). What we choose to call a space [a *universe*] is a relative matter.

Let us then fix a sample space to be denoted by Ω, the capital Greek letter *omega*. It may contain any number of points, possibly infinite but at least one. (As you have probably found out before, mathematics can be very pedantic!) Any of these points may be denoted by ω, the small Greek letter omega, to be distinguished from one another by various devices such as adding subscripts or dashes (as in the case of the two Barbaras if we do not know their family names), thus $\omega_1, \omega_2, \omega', \dots$. Any partial collection of the points is a *subset* of Ω, and since we have fixed Ω we will just call it a set. In extreme cases a set may be Ω itself or the *empty set*, which has no point in it. You may be surprised to hear that the empty set is an important entity and is given a special symbol $\emptyset$. The number of points in a set S will be called its *size* and denoted by $|S|$; thus it is a nonnegative integer or ∞. In particular $|\emptyset| = 0$.

A particular set S is well defined if it is possible to tell whether any given point *belongs to* it or not. These two cases are denoted respectively by

$$\omega \in S; \quad \omega \notin S.$$

Thus a set is determined by a specified rule of membership. For instance, the sets in (a′)–(f′) are well defined up to the limitations of verbal descriptions. One can always quibble about the meaning of words such as "a rotten apple," or attempt to be funny by observing, for instance, that when dice are rolled on a pavement some of them may disappear into the sewer. Some people of a pseudo-philosophical turn of mind get a lot of mileage out of such *caveats*, but we will not indulge in them here. Now, one sure way of specifying a rule to determine a set is to enumerate all its members, namely to make a complete list as the second grader did. But this may be tedious if not impossible. For example, it will be shown in §3.1 that the size of the set in (e) is equal to $6^6 = 46656$. Can you give a quick guess as to how many pages of a book like this will be needed just to record all these possibilities of a mere throw of six dice? On the other hand, it can be described in a systematic and unmistakable way as the set of all ordered 6-tuples of the form below:

$$(s_1, s_2, s_3, s_4, s_5, s_6)$$

where each of the symbols s_j, $1 \le j \le 6$, may be any of the numbers 1, 2, 3, 4, 5, 6. This is a good illustration of mathematics being economy of thought (and printing space).

If every point of A belongs to B, then A is *contained* or *included* in B and is a *subset* of B, while B is a *superset* of A. We write this in one of the two ways below:

$$A \subset B, \quad B \supset A.$$

Two sets are *identical* if they contain exactly the same points, and then we write

$$A = B.$$

Another way to say this is: $A = B$ if and only if $A \subset B$ and $B \subset A$. This may sound unnecessarily roundabout to you, but is often the only way to check that two given sets are really identical. It is not always easy to identify two sets defined in different ways. Do you know for example that the set of even integers is identical with the set of all solutions x of the equation $\sin(\pi x/2) = 0$? We shall soon give some examples of showing the identity of sets by the roundabout method.

1.2. Operations with sets

We learn about sets by operating on them, just as we learn about numbers by operating on them. In the latter case we also say that we compute with numbers: add, subtract, multiply, and so on. These operations performed on given numbers produce other numbers, which are called their sum, difference, product, etc. In the same way, operations performed on sets produce other sets with new names. We are now going to discuss some of these and the laws governing them.

Complement. The complement of a set A is denoted by A^c and is the set of points that do not belong to A. Remember we are talking only about points in a fixed Ω! We write this symbolically as follows:

$$A^c = \{\omega \mid \omega \notin A\},$$

which reads: "A^c is the set of ω that does not belong to A." In particular $\Omega^c = \emptyset$ and $\emptyset^c = \Omega$. The operation has the property that if it is performed twice in succession on A, we get A back:

$$(A^c)^c = A. \tag{1.2.1}$$

Union. The union $A \cup B$ of two sets A and B is the set of points that belong to at least one of them. In symbols:

$$A \cup B = \{\omega \mid \omega \in A \text{ or } \omega \in B\}$$

where "or" means "and/or" in pedantic [legal] style and will always be used in this sense.

Intersection. The intersection $A \cap B$ of two sets A and B is the set of points that belong to both of them. In symbols:

$$A \cap B = \{\omega \mid \omega \in A \text{ and } \omega \in B\}.$$

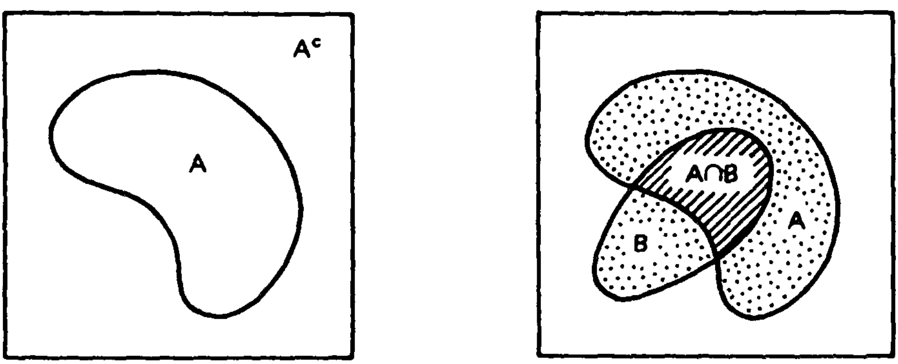


Figure 1

We hold the truth of the following laws as self-evident:

Commutative Law. $A \cup B = B \cup A$, $A \cap B = B \cap A$.

Associative Law. $(A \cup B) \cup C = A \cup (B \cup C)$,

$$(A \cap B) \cap C = A \cap (B \cap C).$$

But observe that these relations are instances of identity of sets mentioned above, and are subject to proof. They should be compared, but not confused, with analogous laws for sum and product of numbers:

$$a + b = b + a, \quad a \times b = b \times a$$

$$(a + b) + c = a + (b + c), \quad (a \times b) \times c = a \times (b \times c).$$

Brackets are needed to indicate the order in which the operations are to be performed. Because of the associative laws, however, we can write

$$A \cup B \cup C, \quad A \cap B \cap C \cap D$$

without brackets. But a string of symbols like $A \cup B \cap C$ is ambiguous, therefore not defined; indeed $(A \cup B) \cap C$ is not identical with $A \cup (B \cap C)$. You should be able to settle this easily by a picture.

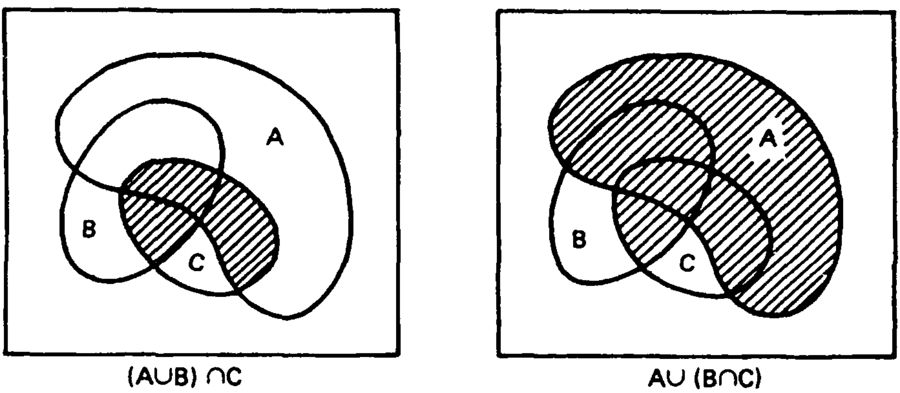
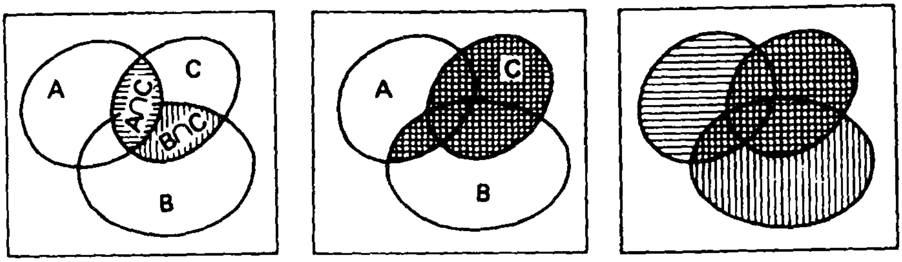


Figure 2

The next pair of *distributive laws* connects the two operations as follows:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C); \qquad (D_1)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C). \qquad (D_2)$$



(A∪B)∩C = (A∩C)∪(B∩C) (A∩B)∪C = (A∪C)∩(B∪C)

Figure 3

Several remarks are in order. First, the analogy with arithmetic carries over to (D_1):

$$(a + b) \times c = (a \times c) + (b \times c);$$

but breaks down in (D_2):

$$(a \times b) + c \neq (a + c) \times (b + c).$$

Of course, the alert reader will have observed that the analogy breaks down already at an earlier stage, for

$$A = A \cup A = A \cap A;$$

but the only number a satisfying the relation $a + a = a$ is 0; while there are exactly two numbers satisfying $a \times a = a$, namely 0 and 1.

Second, you have probably already discovered the use of diagrams to prove or disprove assertions about sets. It is also a good practice to see the truth of such formulas as (D_1) and (D_2) by well-chosen examples. Suppose then that

$$A = \text{inexpensive things}, \ B = \text{really good things},$$
$$C = \text{food [edible things]}.$$

Then $(A \cup B) \cap C$ means "(inexpensive or really good) food," while $(A \cap C) \cup (B \cap C)$ means "(inexpensive food) or (really good food)." So they are the same thing all right. This does not amount to a proof, as one swallow does not make a summer, but if one is convinced that whatever logical structure or thinking process involved above in no way depends on the precise nature of the three things A, B, and C, so much so that they can be *anything*, then one has in fact landed a general proof. Now it is interesting that the same example applied to (D_2) somehow does not make it equally obvious (at least to the author). Why? Perhaps because some patterns of logic are in more common use in our everyday experience than others.

This last remark becomes more significant if one notices an obvious duality between the two distributive laws. Each can be obtained from the other by switching the two symbols $\cup$ and $\cap$. Indeed each can be deduced from the other by making use of this duality (Exercise 11).

Finally, since (D_2) comes less naturally to the intuitive mind, we will avail ourselves of this opportunity to demonstrate the roundabout method of identifying sets mentioned above by giving a rigorous proof of the formula. According to this method, we must show: (i) each point on the left side of (D_2) belongs to the right side; (ii) each point on the right side of (D_2) belongs to the left side.

(i) Suppose ω belongs to the left side of (D_2), then it belongs either to $A \cap B$ or to C. If $\omega \in A \cap B$, then $\omega \in A$, hence $\omega \in A \cup C$; similarly $\omega \in B \cup C$. Therefore ω belongs to the right side of (D_2). On the other hand, if $\omega \in C$, then $\omega \in A \cup C$ and $\omega \in B \cup C$ and we finish as before.

(ii) Suppose ω belongs to the right side of (D_2), then ω may or may not belong to C, and the trick is to consider these two alternatives. If $\omega \in C$, then it certainly belongs to the left side of (D_2). On the other hand, if $\omega \notin C$, then since it belongs to $A \cup C$, it must belong to A; similarly it must belong to B. Hence it belongs to $A \cap B$, and so to the left side of (D_2). Q.E.D.

1.3. Various relations

The three operations so far defined: complement, union, and intersection obey two more laws called *De Morgan's laws:*

$$(A \cup B)^c = A^c \cap B^c; \tag{C$_1$}$$

$$(A \cap B)^c = A^c \cup B^c. \tag{C$_2$}$$

They are dual in the same sense as (D_1) and (D_2) are. Let us check these by our previous example. If A = inexpensive, and B = really good, then clearly $(A \cup B)^c$ = not inexpensive nor really good, namely high-priced junk, which is the same as $A^c \cap B^c$ = inexpensive and not really good. Similarly we can check (C_2).

Logically, we can deduce either (C_1) or (C_2) from the other; let us show it one way. Suppose then (C_1) is true, then since A and B are arbitrary sets we can substitute their complements and get

$$(A^c \cup B^c)^c = (A^c)^c \cap (B^c)^c = A \cap B \tag{1.3.1}$$

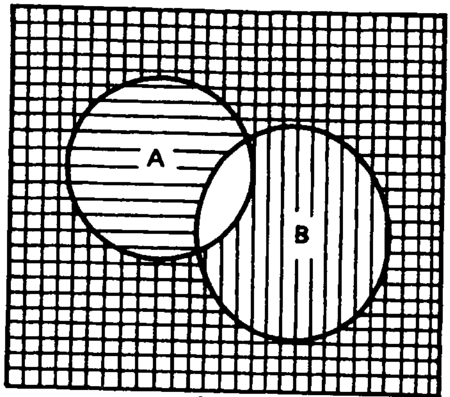


Figure 4

where we have also used (1.2.1) for the second equation. Now taking the complements of the first and third sets in (1.3.1) and using (1.2.1) again we get

$$A^c \cup B^c = (A \cap B)^c.$$

This is (C_2). Q.E.D.

It follows from De Morgan's laws that if we have complementation, then either union or intersection can be expressed in terms of the other. Thus we have

$$A \cap B = (A^c \cup B^c)^c,$$
$$A \cup B = (A^c \cap B^c)^c;$$

and so there is redundancy among the three operations. On the other hand, it is impossible to express complementation by means of the other two although there is a magic symbol from which all three can be derived (Exercise 14). It is convenient to define some other operations, as we now do.

Difference. The set $A \setminus B$ is the set of points that belong to A and (but) not to B. In symbols:

$$A \setminus B = A \cap B^c = \{\omega \mid \omega \in A \text{ and } \omega \notin B\}.$$

This operation is neither commutative nor associative. Let us find a *counterexample* to the associative law, namely, to find some A, B, C for which

$$(A \setminus B) \setminus C \neq A \setminus (B \setminus C). \tag{1.3.2}$$