

Zhengyan Lin  
Zhidong Bai

# Probability Inequalities

 Science Press  
Beijing

 Springer

Zhengyan Lin  
Zhidong Bai

## **Probability Inequalities**

Zhengyan Lin  
Zhidong Bai

# Probability Inequalities

 Science Press  
Beijing

 Springer

*Authors*

Zhengyan Lin  
Department of Mathematics  
Zhejiang University, China  
zlin@zju.edu.cn

Zhidong Bai  
School of Mathematics and Statistics  
Northeast Normal University, China  
baizd@nenu.edu.cn  
Department of Statistics and Applied Probability  
National University of Singapore, Singapore  
stabaizd@leonis.nus.edu.sg

ISBN 978-7-03-025562-4  
Science Press Beijing

ISBN 978-3-642-05260-6                      e-ISBN 978-3-642-05261-3  
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009938102

© Science Press Beijing and Springer-Verlag Berlin Heidelberg 2010

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

*Cover design:* Frido Steinen-Broo, EStudio Calamar, Spain

Printed on acid-free paper

Springer is a part of Springer Science+Business Media (www.springer.com)

# Preface

In almost every branch of quantitative sciences, inequalities play an important role in its development and are regarded to be even more important than equalities. This is indeed the case in probability and statistics. For example, the Chebyshev, Schwarz and Jensen inequalities are frequently used in probability theory, the Cramer-Rao inequality plays a fundamental role in mathematical statistics. Choosing or establishing an appropriate inequality is usually a key breakthrough in the solution of a problem, e.g. the Berry-Esseen inequality opens a way to evaluate the convergence rate of the normal approximation.

Research beginners usually face two difficulties when they start researching—they choose an appropriate inequality and/or cite an exact reference. In literature, almost no authors give references for frequently used inequalities, such as the Jensen inequality, Schwarz inequality, Fatou Lemma, etc. Another annoyance for beginners is that an inequality may have many different names and reference sources. For example, the Schwarz inequality is also called the Cauchy, Cauchy-Schwarz or Minkovski-Bnyakovski inequality. Bennet, Hoeffding and Bernstein inequalities have a very close relationship and format, and in literature some authors cross-cite in their use of the inequalities. This may be due to one author using an inequality and subsequent authors just simply copying the inequality's format and its reference without checking the original reference. All this may distress beginners very much.

The aim of this book is to help beginners with these problems. We provide a place to find the most frequently used inequalities, their proofs (if not too lengthy) and some references. Of course, for some of the more popularly known inequalities, such as Jensen and Schwarz, there is no necessity to give a reference and we will not do so.

The wording “frequently used” is based on our own understanding. It can be expected that many important probability inequalities are not

collected in this work. Any comments and suggestions will be appreciated.

The writing of the book is supported partly by the National Science Foundation of China.

The authors would like to express their thanks to Ron Lim Beng Seng for improving our English in this book.

Zhengyan Lin

May, 2009

# Contents

<b>Chapter 1 Elementary Inequalities of Probabilities of Events</b> .....	1
1.1 Inclusion-exclusion Formula .....	1
1.2 Corollaries of the Inclusion-exclusion Formula .....	2
1.3 Further Consequences of the Inclusion-exclusion Formula .....	2
1.4 Inequalities Related to Symmetric Difference .....	6
1.5 Inequalities Related to Independent Events .....	6
1.6 Lower Bound for Union (Chung-Erdős) .....	8
References .....	8
<b>Chapter 2 Inequalities Related to Commonly Used Distributions</b> .....	9
2.1 Inequalities Related to the Normal d.f. ....	9
2.2 Slepian Type Inequalities .....	12
2.3 Anderson Type Inequalities .....	17
2.4 Khatri-Šidák Type Inequalities .....	18
2.5 Corner Probability of Normal Vector .....	19
2.6 Normal Approximations of Binomial and Poisson Distributions .....	20
References .....	22
<b>Chapter 3 Inequalities Related to Characteristic Functions</b> .....	23
3.1 Inequalities Related Only with c.f. ....	23
3.2 Inequalities Related to c.f. and d.f. ....	26
3.3 Normality Approximations of c.f. of Independent Sums .....	27
References .....	28
<b>Chapter 4 Estimates of the Difference of Two Distribution Functions</b> .....	29
4.1 Fourier Transformation .....	29
4.2 Stein-Chen Method .....	33
4.3 Stieltjes Transformation .....	34

References .....	36
<b>Chapter 5 Probability Inequalities of Random Variables</b> .....	<b>37</b>
5.1 Inequalities Related to Two r.v.'s .....	37
5.2 Perturbation Inequality .....	39
5.3 Symmetrization Inequalities .....	40
5.4 Lévy Inequality .....	41
5.5 Bickel Inequality .....	42
5.6 Upper Bounds of Tail Probabilities of Partial Sums .....	44
5.7 Lower Bounds of Tail Probabilities of Partial Sums .....	44
5.8 Tail Probabilities for Maximum Partial Sums .....	45
5.9 Tail Probabilities for Maximum Partial Sums (Continuation) .....	46
5.10 Reflection Inequality of Tail Probability (Hoffmann-Jørgensen) .....	47
5.11 Probability of Maximal Increment (Shao) .....	48
5.12 Mogulskii Minimal Inequality .....	49
5.13 Wilks Inequality .....	50
References .....	50
<b>Chapter 6 Bounds of Probabilities in Terms of Moments</b> .....	<b>51</b>
6.1 Chebyshev-Markov Type Inequalities .....	51
6.2 Lower Bounds .....	52
6.3 Series of Tail Probabilities .....	53
6.4 Kolmogorov Type Inequalities .....	54
6.5 Generalization of Kolmogorov Inequality for a Submartingale .....	56
6.6 Rényi-Hájek Type Inequalities .....	57
6.7 Chernoff Inequality .....	60
6.8 Fuk and Nagaev Inequality .....	62
6.9 Burkholder Inequality .....	63
6.10 Complete Convergence of Partial Sums .....	65
References .....	65
<b>Chapter 7 Exponential Type Estimates of Probabilities</b> .....	<b>67</b>
7.1 Equivalence of Exponential Estimates .....	67
7.2 Petrov Exponential Inequalities .....	68



7.3	Hoeffding Inequality	70
7.4	Bennett Inequality	72
7.5	Bernstein Inequality	73
7.6	Exponential Bounds for Sums of Bounded Variables	74
7.7	Kolmogorov Inequalities	75
7.8	Prokhorov Inequality	79
7.9	Exponential Inequalities by Censoring	80
7.10	Tail Probability of Weighted Sums	82
	References	83
<b>Chapter 8 Moment Inequalities Related to One or Two</b>		
	<b>Variables</b>	84
8.1	Moments of Truncation	84
8.2	Exponential Moment of Bounded Variables	84
8.3	Hölder Type Inequalities	85
8.4	Jensen Type Inequalities	86
8.5	Dispersion Inequality of Censored Variables	87
8.6	Monotonicity of Moments of Sums	88
8.7	Symmetrization Moment Inequalities	88
8.8	Kimball Inequality	89
8.9	Exponential Moment of Normal Variable	90
8.10	Inequalities of Nonnegative Variable	93
8.11	Freedman Inequality	94
8.12	Exponential Moment of Upper Truncated Variables	95
	References	96
<b>Chapter 9 Moment Estimates of (Maximum of) Sums of</b>		
	<b>Random Variables</b>	97
9.1	Elementary Inequalities	97
9.2	Minkowski Type Inequalities	99
9.3	The Case $1 \leq r \leq 2$	100
9.4	The Case $r \geq 2$	101
9.5	Jack-knifed Variance	105
9.6	Khinchine Inequality	106
9.7	Marcinkiewicz-Zygmund-Burkholder Type Inequalities	108
9.8	Skorokhod Inequalities	111
9.9	Moments of Weighted Sums	112
9.10	Doob Crossing Inequalities	113

9.11	Moments of Maximal Partial Sums	114
9.12	Doob Inequalities	115
9.13	Equivalence Conditions for Moments	117
9.14	Serfling Inequalities	123
9.15	Average Fill Rate	127
	References	128
<b>Chapter 10 Inequalities Related to Mixing Sequences</b>		130
10.1	Covariance Estimates for Mixing Sequences	131
10.2	Tail Probability on $\alpha$ -mixing Sequence	134
10.3	Estimates of 4-th Moment on $\rho$ -mixing Sequence	135
10.4	Estimates of Variances of Increments of $\rho$ -mixing Sequence	136
10.5	Bounds of $2 + \delta$ -th Moments of Increments of $\rho$ -mixing Sequence	138
10.6	Tail Probability on $\varphi$ -mixing Sequence	140
10.7	Bounds of $2 + \delta$ -th Moment of Increments of $\varphi$ -mixing Sequence	142
10.8	Exponential Estimates of Probability on $\varphi$ -mixing Sequence	143
	References	148
<b>Chapter 11 Inequalities Related to Associative Variables</b>		149
11.1	Covariance of PQD Variables	149
11.2	Probability of Quadrant on PA (NA) Sequence	150
11.3	Estimates of c.f.'s on LPQD (LNQD) Sequence	151
11.4	Maximal Partial Sums of PA Sequence	152
11.5	Variance of Increment of LPQD Sequence	153
11.6	Expectation of Convex Function of Sum of NA Sequence	154
11.7	Marcinkiewicz-Zygmund-Burkholder Inequality for NA Sequence	156
	References	157
<b>Chapter 12 Inequalities about Stochastic Processes and Banach Space Valued Random Variables</b>		158
12.1	Probability Estimates of Supremums of a Wiener Process	158

12.2	Probability Estimate of Supremum of a Poisson Process .....	162
12.3	Fernique Inequality .....	164
12.4	Borell Inequality .....	167
12.5	Tail Probability of Gaussian Process .....	168
12.6	Tail Probability of Randomly Signed Independent Processes .....	168
12.7	Tail Probability of Adaptive Process .....	170
12.8	Tail Probability on Submartingale .....	172
12.9	Tail Probability of Independent Sum in B-Space .....	173
12.10	Isoperimetric Inequalities .....	173
12.11	Ehrhard Inequality .....	174
12.12	Tail Probability of Normal Variable in B-Space .....	175
12.13	Gaussian Measure on Symmetric Convex Sets .....	175
12.14	Equivalence of Moments of B-Gaussian Variables .....	176
12.15	Contraction Principle .....	176
12.16	Symmetrization Inequalities in B-Space .....	178
12.17	Decoupling Inequality .....	178
	References .....	180

# Chapter 1

## Elementary Inequalities of Probabilities of Events

In this Chapter, we shall introduce some basic inequalities which can be found in many basic textbooks on probability theory, such as Feller (1968), Loève (1977), etc.

We shall use the following popularly used notations. Let  $\Omega$  be a space of elementary events,  $\mathcal{F}$  be a  $\sigma$ -algebra of subsets of  $\Omega$ ,  $P$  be a probability measure defined on the events in  $\mathcal{F}$ .  $(\Omega, \mathcal{F}, P)$  is so called a probability space. The events in  $\mathcal{F}$  will be denoted by  $A_1, A_2, \dots$  or  $A, B, \dots$  etc.  $A \cup B, AB$  (or  $A \cap B$ ),  $A - B$  and  $A \Delta B$  denote the union, intersection, difference and symmetric difference of  $A$  and  $B$  respectively.  $A^c$  denotes the complement of  $A$  and  $\emptyset$  denotes the empty set.

### 1.1 Inclusion-exclusion Formula

Let  $A_1, A_2, \dots, A_n$  be  $n$  events, then we have

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \dots \\ &\quad + (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k}) \\ &\quad + \dots + (-1)^{n-1} P(A_1 \dots A_n). \end{aligned}$$

**Proof.** When  $n = 2$ , it is trivially known that

$$\begin{aligned} P\left(A_1 \bigcup A_2\right) &= P(A_1) + P(A_2 - A_1 A_2) \\ &= P(A_1) + P(A_2) - P(A_1 A_2). \end{aligned} \tag{1}$$

We show the formula by induction. Assume that the formula holds for  $n$ . We will show that it holds also for  $n + 1$ . In fact, by (1) and the induction hypothesis,

$$\begin{aligned}
 P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P\left(\bigcup_{i=1}^n A_i A_{n+1}\right) \\
 &= \sum_{i=1}^{n+1} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \cdots + (-1)^{n-1} P(A_1 \cdots A_n) \\
 &\quad - \left\{ \sum_{i=1}^n P(A_i A_{n+1}) - \sum_{1 \leq i < j \leq n} P(A_i A_j A_{n+1}) \right. \\
 &\quad \left. + \cdots + (-1)^{n-1} P(A_1 \cdots A_n A_{n+1}) \right\} \\
 &= \sum_{i=1}^{n+1} P(A_i) - \sum_{1 \leq i < j \leq n+1} P(A_i A_j) + \cdots + (-1)^n P(A_1 \cdots A_{n+1}).
 \end{aligned}$$

## 1.2 Corollaries of the Inclusion-exclusion Formula

From the inclusion-exclusion formula, it is easy to deduce the following two conclusions.

**1.2.a.** When  $A_1, \dots, A_n$  are exchangeable, we have

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} P(A_1, \dots, A_i).$$

**Remark.** A set of events  $\{A_1, \dots, A_n\}$  is said to be exchangeable if the probability of the intersection of any subset depends only on the size of the subset, that is, for any integers  $1 \leq i_1 < \cdots < i_j \leq n$  and  $1 \leq j \leq n$ ,  $P(A_{i_1} A_{i_2} \cdots A_{i_j}) = p_j$ .

**1.2.b.** When  $A_1, \dots, A_n$  are independent and  $p = P(A_i)$ , we have

$$p \left( \bigcup_{i=1}^n A_i \right) = \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} p^i.$$

## 1.3 Further Consequences of the Inclusion-exclusion Formula

The following inequalities are also consequences of the inclusion-exclusion formula.

$$1.3.a. \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) \leq P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i).$$

**Remark.** The right hand side (RHS) of the above inequality can be improved to

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i) - \sum_{i=2}^n P(A_1 A_i).$$

**Proof.** When  $n = 2$ , the inequality with the improved RHS reduces to the inclusion-exclusion formula. Now, by induction we have

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P\left(\left(\bigcup_{i=1}^n A_i\right) \cap A_{n+1}\right) \\ &\leq \sum_{i=1}^n P(A_i) - \sum_{i=2}^n P(A_1 A_i) + P(A_{n+1}) - P(A_1 A_{n+1}). \end{aligned}$$

This proves the improved right hand side. Similarly, by induction, we have

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P\left(\left(\bigcup_{i=1}^n A_i\right) \cap A_{n+1}\right) \\ &\geq \sum_{i=1}^n P(A_i) - \sum_{i < j \leq n} P(A_i A_j) + P(A_{n+1}) - \sum_{i=1}^n P(A_i A_{n+1}). \end{aligned}$$

This proves the left hand side (LHS) of the inequality.

$$1.3.b. |P(AB) - P(A)P(B)| \leq \frac{1}{4}.$$

**Proof.**

$$\begin{aligned} |P(AB) - P(A)P(B)| &= |P(A)P(AB) + P(A^c)P(AB) \\ &\quad - P(A)P(AB) - P(A)P(A^c B)| \\ &= |P(A^c)P(AB) - P(A)P(A^c B)|. \end{aligned}$$

Since  $A^c$  and  $AB$  are disjoint,  $P(A^c)P(AB) \leq 1/4$  (by noticing that  $\max_{0 < p < 1} p(1-p) = 1/4$ ). Similarly,  $P(A)P(A^c B) \leq 1/4$  as desired.

**Remark.** The difference  $P(AB) - P(A)P(B)$  can be regarded as the covariance of the indicators  $I_A$  and  $I_B$ . The inequality 1.3.b can be easily proved by using the Cauchy-Schwarz inequality. Here, we proved the inequality by deliberately avoiding the use of moments.

$$1.3.c. |P(A) - P(B)| \leq P(A \Delta B), \quad (A \Delta B = (A - B) \cup (B - A))$$

**Proof.** By 1.3.a,

$$P(A\Delta B) \geq P(A - B) = P(A) - P(AB) \geq P(A) - P(B).$$

By the symmetry of  $A$  and  $B$ , 1.3.c is proved.

**1.3.d.** (Boole inequality).  $P(AB) \geq 1 - P(A^c) - P(B^c)$ .

**Proof.**  $P(AB) + P(B^c) \geq P(A) = 1 - P(A^c)$ .

**1.3.e.** Let  $\limsup_{n \rightarrow \infty} A_n = \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A_n$ ,  $\liminf_{n \rightarrow \infty} A_n = \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} A_n$ . Then

$$\begin{aligned} P(\liminf_{n \rightarrow \infty} A_n) &\leq \liminf_{n \rightarrow \infty} P(A_n) \leq \limsup_{n \rightarrow \infty} P(A_n) \\ &\leq P(\limsup_{n \rightarrow \infty} A_n) \leq \lim_{N \rightarrow \infty} \sum_{n=N}^{\infty} P(A_n). \end{aligned}$$

**Proof.** For any positive integer  $N$ , we have

$$\bigcap_{n=N}^{\infty} A_n \subset A_N \subset \bigcup_{n=N}^{\infty} A_n,$$

which simply implies

$$P\left(\bigcap_{n=N}^{\infty} A_n\right) \leq P(A_N) \leq P\left(\bigcup_{n=N}^{\infty} A_n\right) \leq \sum_{n=N}^{\infty} P(A_n).$$

Letting  $N \rightarrow \infty$ , we obtain the desired inequalities.

**1.3.f.** If  $P(A) \geq 1 - \varepsilon$ ,  $P(B) \geq 1 - \varepsilon$  for some  $0 < \varepsilon < \frac{1}{2}$ , then  $P(AB) \geq 1 - 2\varepsilon$ .

**Proof.**  $P(AB) = P(A) + P(B) - P(A \cup B) \geq 1 - 2\varepsilon$ .

**1.3.g** (Bonferroni inequality). Let  $P_{[m]}(P_m)$  be the probability that exactly (at least, correspondingly)  $m$  events among  $A_1, \dots, A_n$  occur simultaneously. Putting

$$S_m = \sum_{1 \leq i_1 < \dots < i_m \leq n} P(A_{i_1} \cdots A_{i_m}).$$

Then

$$S_m - (m+1)S_{m+1} \leq P_{[m]} \leq S_m, \quad S_m - mS_{m+1} \leq P_m \leq S_m.$$

**Proof.** Let  $A_{[m]}$  ( $A_{(m)}$ ) denote the event that exactly (at least, correspondingly)  $m$  events among  $A_1, \dots, A_n$  happen simultaneously. Then we have

$$A_{[m]} \subset A_{(m)} = \bigcup_{1 \leq i_1 < \dots < i_m \leq n} A_{i_1} \cdots A_{i_m}.$$

The RHS of the above inequalities follows by using the semi-additivity of probability measure.

On the other hand, we have

$$A_{[m]} \supset \bigcup_{1 \leq i_1 < \dots < i_m \leq n} A_{i_1} \cdots A_{i_m} - \bigcup_{1 \leq i_1 < \dots < i_{m+1} \leq n} A_{i_1} \cdots A_{i_{m+1}}.$$

This implies  $p_{[m]} \geq p_m - S_{m+1}$ . Also, for each  $i_1 < \dots < i_{m+1}$ , the probability  $P(A_{i_1} \cdots A_{i_{m+1}})$  is included at most in each of  $P(A_{i'_1} \cdots A_{i'_m})$  in  $S_m$ , where  $(i'_1, \dots, i'_m)$  is a subset of  $(i_1, \dots, i_{m+1})$ . Among the  $m+1$ , one needs to contribute to  $p_m$ . Therefore,

$$S_m - p_m \geq mS_{m+1}.$$

The LHS of 1.3.g then follow.

**Remark.** In fact, the inequalities 1.3.g can be proved from the following identities:

$$\begin{aligned} P_{[m]} &= S_m - \binom{m+1}{m} S_{m+1} + \binom{m+2}{m} S_{m+2} + \dots + (-1)^{n-m} \binom{n}{m} S_n, \\ P_m &= S_m - \binom{m}{m-1} S_{m+1} + \binom{m+1}{m-1} S_{m+2} + \dots + (-1)^{n-m} \binom{n-1}{m-1} S_n, \\ S_m &= \sum_{i=m}^n \binom{i}{m} P_{[i]} \quad \text{and} \quad S_m = \sum_{i=m}^n \binom{i-1}{m-1} P_i. \end{aligned}$$

By definition, we have

$$P_{[i]} = \sum_{F \in \mathcal{F}_i} P \left( \bigcap_{j \in F} A_j \bigcap_{\ell \in F^c} A_\ell^c \right),$$

where  $\mathcal{F}_i$  is the collection of all subsets of size  $i$  of the set  $\{1, 2, \dots, n\}$ . Note that for each  $\tilde{F} \in \mathcal{F}_m$ , the set  $\bigcap_{t \in \tilde{F}} A_t$  can be written as the union of

disjoint subsets  $\bigcap_{j \in F} A_j \bigcap_{\ell \in F^c} A_\ell^c$ , for all  $i \geq m$  where  $F \in \mathcal{F}_i$  and  $F \subset \tilde{F}$ .

This proves that

$$S_m = \sum_{i=m}^n \binom{i}{m} P_{[i]},$$

which implies  $S_m = \sum_{i=m}^n \binom{i-1}{m-1} P_i$  by noticing that  $P_{[i]} = P_i - P_{i+1}$ .



Substituting the expression of  $S_m$  in terms of  $P_{[i]}$  into the RHS of the first identity, we obtain

$$\begin{aligned} & S_m - \binom{m+1}{m} S_{m+1} + \binom{m+2}{m} S_{m+2} + \cdots + (-1)^{n-m} \binom{n}{m} S_n \\ &= \sum_{j=m}^n (-1)^{j-m} \binom{j}{m} \sum_{i=j}^n \binom{i}{j} P_{[i]} \\ &= \sum_{i=m}^n \binom{i}{m} P_{[i]} \sum_{j=m}^i \binom{i-m}{j-m} (-1)^{j-m} = P_{[m]}. \end{aligned}$$

By the same approach, one can prove the second identity.

## 1.4 Inequalities Related to Symmetric Difference

$$1.4.a. \quad P\left\{\left(\bigcup_n A_n\right) \Delta \left(\bigcup_n B_n\right)\right\} \leq P\left\{\bigcup_n (A_n \Delta B_n)\right\} \leq \sum_n P(A_n \Delta B_n).$$

**Proof.** The left inequality follows from  $\left(\bigcup_n A_n\right) \Delta \left(\bigcup_n B_n\right) \subset \bigcup_n (A_n \Delta B_n)$  by the definition of the symmetric difference. The right one follows from 1.3.a.

$$1.4.b. \quad P\{(A_1 - A_2) \Delta (B_1 - B_2)\} \leq P(A_1 \Delta B_1) + P(A_2 \Delta B_2).$$

**Proof.** The inequality follows from the observation

$$(A_1 - A_2) \Delta (B_1 - B_2) \subset (A_1 \Delta B_1) \cup (A_2 \Delta B_2).$$

## 1.5 Inequalities Related to Independent Events

1.5.a. Let  $\{A_n\}$  be a sequence of mutually independent events. Then

$$\begin{aligned} 1 - P\left(\bigcup_{k=1}^n A_k\right) &\leq \exp\left\{-\sum_{k=1}^n P(A_k)\right\}, \\ 1 - P\left(\bigcup_{k=1}^{\infty} A_k\right) &\leq \lim_{n \rightarrow \infty} \exp\left\{-\sum_{k=1}^n P(A_k)\right\}. \end{aligned}$$

**Remark.** The inequalities are useful in the proof of the Borel-Cantelli lemma.

**Proof.** The conclusions follow from an application of the inequality  $1 - x \leq e^{-x}$  for real  $x$  to the RHS of the identity

$$1 - P\left(\bigcup_{k=1}^n A_k\right) = P\left(\bigcap_{k=1}^n A_k^c\right) = \prod_{k=1}^n (1 - P(A_k)).$$

**1.5.b.** Suppose that  $A$  and  $B$  are independent,  $AB \subset D$  and  $A^c B^c \subset D^c$ . Then  $P(AD) \geq P(A)P(D)$ .

**Proof.**

$$\begin{aligned} P(AD) &= P(ADB) + P(ADB^c) = P(AB) + P(AB^c) - P(AD^c B^c) \\ &= P(A)P(B) + P(AB^c) - P(D^c B^c) + P(A^c D^c B^c) \\ &= P(A)P(B) + P(AB^c) - P(D^c B^c) + P(A^c B^c) \\ &= P(A)P(B) + P(B^c) - P(D^c B^c) \\ &\geq P(A)P(BD) + P(A)P(B^c D) = P(A)P(D). \end{aligned}$$

**1.5.c** (Feller-Chung). Let  $A_0 = \emptyset$ ,  $\{A_n\}$  and  $\{B_n\}$  be two sequences of events. Suppose that either

(i)  $B_n$  is independent of  $A_n A_{n-1}^c \cdots A_0^c$  for all  $n \geq 1$ , or

(ii)  $B_n$  is independent of  $\{A_n, A_n A_{n+1}^c, A_n A_{n+1}^c A_{n+2}^c, \dots\}$  for all  $n \geq 1$ .

1. Then

$$P\left(\bigcup_{n=1}^{\infty} A_n B_n\right) \geq \inf_{n \geq 1} P(B_n) P\left(\bigcup_{n=1}^{\infty} A_n\right).$$

**Proof.** In case (i),

$$\begin{aligned} P\left\{\bigcup_{n=1}^{\infty} A_n B_n\right\} &= P\left\{\bigcup_{n=1}^{\infty} B_n A_n \bigcap_{j=0}^{n-1} (B_j A_j)^c\right\} = \sum_{n=1}^{\infty} P\left\{B_n A_n \bigcap_{j=0}^{n-1} (B_j A_j)^c\right\} \\ &\geq \sum_{n=1}^{\infty} P\left\{B_n A_n \bigcap_{j=0}^{n-1} A_j^c\right\} = \sum_{n=1}^{\infty} P(B_n) P\left\{A_n \bigcap_{j=0}^{n-1} A_j^c\right\} \\ &\geq \inf_{n \geq 1} P(B_n) P\left(\bigcup_{n=1}^{\infty} A_n\right), \end{aligned}$$

and in case (ii),

$$\begin{aligned} P\left\{\bigcup_{j=1}^n A_j B_j\right\} &= \sum_{j=1}^n P\left\{A_j B_j \bigcap_{i=j+1}^n (A_i B_i)^c\right\} \geq \sum_{j=1}^n P\left\{A_j B_j \bigcap_{i=j+1}^n A_i^c\right\} \\ &= \sum_{j=1}^n P(B_j) P\left\{A_j \bigcap_{i=j+1}^n A_i^c\right\} \geq \inf_{1 \leq j \leq n} P(B_j) P\left\{\bigcup_{j=1}^n A_j\right\}. \end{aligned}$$

## 1.6 Lower Bound for Union (Chung-Erdős)

$$P\left(\bigcup_{i=1}^n A_i\right) \geq \left(\sum_{i=1}^n P(A_i)\right)^2 / \left(\sum_{i=1}^n P(A_i) + 2 \sum_{1 \leq i < j \leq n} P(A_i A_j)\right).$$

**Proof.** Define random variables  $X_k(\omega), \omega \in \Omega$ , by

$$X_i(\omega) = \begin{cases} 0, & \text{if } \omega \notin A_i, \\ 1, & \text{if } \omega \in A_i. \end{cases}$$

Then

$$2 \sum_{1 \leq i < j \leq n} P(A_i A_j) = E(X_1 + \cdots + X_n)^2 - E(X_1^2 + \cdots + X_n^2).$$

By the Cauchy-Schwarz inequality (see 8.4.b), we have

$$(E(X_1 + \cdots + X_n))^2 \leq P(X_1 + \cdots + X_n > 0)E(X_1 + \cdots + X_n)^2.$$

Note that  $EX_i = EX_i^2 = P(A_i), P(X_1 + \cdots + X_n > 0) = P\left(\bigcup_{i=1}^n A_i\right)$  by definition. Combining the above two relations yields the desired inequality.

## References

- Chow Y S, Teicher H. 1997. Probability Theory: Independence, Interchangeability, Martingales. 2nd ed.. New York: Springer-Verlag.
- Chung K L. 1974. A Course in Probability Theory. 2nd ed.. New York: Academic Press.
- Feller W. 1968. An Introduction to Probability Theory and Its Applications. 3rd ed.. New York: Wiley.
- Hardy G H, Littlewood J E, Polya G. 1952. Inequalities. Cambridge: Cambridge University Press.
- Loève M. 1977. Probability Theory. 4th ed.. New York: Springer-Verlag.
- Petrov V V. 1995. Limit Theorems of Probability Theory: Sequences of Independent Random Variables. Oxford & New York: Clarendon Press & Oxford University Press.
- Shorack G R, Wellner J A. 1986. Empirical Processes with Applications to Statistics. New York: Wiley.

## Chapter 2

# Inequalities Related to Commonly Used Distributions

Commonly used distributions play an important role in applied statistics, statistical computing and applied probability. So, inequalities related to these distributions are of great interest in these areas.

Let  $\xi$  be a random variable (r.v.). Then its distribution function (d.f.) is defined by  $F(x) = P(\xi < x)$  and its probability density function (pdf.)  $p(x)$  (if it exists) is defined to be a measurable function such that  $F(x) = \int_{-\infty}^x p(y)dy$ . Write

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \quad \text{and} \quad \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

for the standard normal d.f. and pdf. respectively,

$$b(k; n, p) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \dots, n, \quad 0 < p < 1, \quad q = 1 - p$$

for the binomial distribution with parameters  $n$  and  $p$ ,

$$p(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, \dots, \quad \lambda > 0$$

for the Poisson distribution with parameter  $\lambda$ .

### 2.1 Inequalities Related to the Normal d.f.

**2.1.a.**  $\frac{1}{\sqrt{2\pi}}(b-a) \exp\{-(a^2 \vee b^2)/2\} \leq \Phi(b) - \Phi(a) \leq \frac{1}{\sqrt{2\pi}}(b-a), -\infty < a < b < \infty$ .

**Proof.** It follows from the fact that  $e^{-x^2/2}$  on  $[a, b]$  is between  $\exp\{-(a^2 \vee b^2)/2\}$  and 1.