独立随机变量之和的极限定理

〔苏〕B.B. 佩特罗夫 著

苏 淳 黄可明 译

中国科学技术大学出版社

1991·合肥

内 容 简 介

本书是关于随机变量之和的极限的理论的一本系统性专著，收入了直到80年代后期的许多新结论、新成果，论证了一系列关于任意数目的独立随机变量之和的概率不等式，以及一些强大数律和重对数律方面的定理，还特别注重了中心极限定理中的估计问题。

本书可作为概率统计专业研究生的教材，也可供有关教学、研究人员及专家参考。

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ
ДЛЯ СУММ НЕЗАВИСИММ
НЕЗАВИСИМЫХ СЛУЧАЙНЫХ
ВЕЛИЧИН

В. В. ПЕТРОВ

Москов Наука 1987

独立随机变量之和的极限定理

〔苏〕B. B. 佩特罗夫 著

苏 淳 黄可明 译

•

中国科学技术大学出版社出版

（安徽省合肥市金寨路96号，邮政编码：230026）

合肥炮兵学院印刷厂印刷

安徽省新华书店发行

•

开本：850×1168/32 印张：11.125 字数：287千

1991年5月第1版 1991年5月第1次印刷

印数 1—2500 册

ISBN7-312-00226-9/O·86 定价5.70元

译者的话

本书是B.B.佩特罗夫教授的新作，成稿于1986年6月，由苏联科学出版社1987年出版．佩特罗夫教授在序言中介绍了他写作这本书新著的原因及其与1972年俄文版《独立随机变量的和》的关系（1975年由Springer-Verlag出版社出版的英文版《Sums of Independent Random Variables》即是译自1972年的俄文版）．

B.B．佩特罗夫的1975年英文版久已蜚声海内外概率统计界．用陈希孺教授的话来说，就是“这些年来有关概率统计的书出了不少，我在国外时，唯见佩特罗夫的书在同行们当中人手一册”．在我国同行们中间，佩特罗夫教授的书也早已引起高度重视，并已广为流传．魏宗舒教授说过：“这是一本概率统计界人士人人应读的好书”．

B.B.佩特罗夫的新著出版的当年就传到了我国，并立即引起了广泛的注意．中国科技大学当年就采用它作为研究生教材，在为1987和1988级概率统计专业硕士研究生开设的“极限理论”课上使用；杭州大学、安徽大学等一些高校也已先后采用它作为研究生教材．北京、上海、福州、杭州、南京和合肥的许多概率统计界和出版界人士都为新著的翻译和正式出版作了许多努力，尤其是江苏教育出版社的喻纬同志和中国科技大学出版社的一些同志更为其努力奔波，积极地想办法．此次的公开出版发行，得到了中国科技大学（合肥）研究生院的资助．译者谨在此致以深深的谢意．

1989年12月31日

序　　言

自从我的著作《独立随机变量的和》1972年出版以来，在独立随机变量之和的理论方面，又经历了一个急剧发展的时期．为了能在一定的程度上反映出这些进步，使我产生了撰写现在这本著作的愿望．

尽管相对来说，本书的篇幅不算太大，但它却包含了许多国内外书籍文献中所未曾收入的资料．书中特意收入了不久前刚获得的一系列关于任意数目的独立随机变量之和的概率不等式，以及一些强大数律和重对数律方面的定理，我们还特别注重了中心极限定理中的估计问题．为了不增加书本的篇幅，书中未再收入《独立随机变量的和》中有关局部极限定理和大偏差概率的章节，并删去了中心极限定理渐近展开方面的内容．这是因为若要收入这些内容，我们就不能不拿出大量篇幅，以便陈述和证明大偏差概率和渐近展开方面的数目众多的新成果．而另一方面，最近15年来，在局部极限定理方面的进展却并不很大．

本书在写作上，既兼顾了正在攻读独立随机变量之和的理论的本科生和研究生们的需要，也顾及了概率理论及其应用方面的专家们的兴趣．

每一章中的基本内容，都可供初次学习概率论极限定理的人们阅读，还可选为相应的讲课材料．而在每一章之后的附录里，则收入了许多不久前才在杂志上发表的新结果，以使本书对概率论方面的专家也将有所裨益．

对于本书的阅读，只需要具有概率论的基础知识，本书的第一章还梗概地介绍了一些必要的预备知识．

应当深深感谢 А. И. Мартикайнен, С. М. Ананьевский, В. М. Круглов 和 В. Б. Невзоров 提出的许多建议，它们使得本书大为增色.

我尤其要诚挚地感谢 Ю. В. Прохоров，是他提议我筹备写作这本书，并一贯关注和支持着它的全部写作过程.

B.B.佩特罗夫

目　录

译者的话……………………………………………………(i)

序言……………………………………………………(iii)

第一章　概率分布和特征函数……………………………(1)

§ 1　随机变量和概率分布……………………………(1)

§ 2　随机变量的矩和其它数字特征……………………(6)

§ 3　特征函数……………………………………………(12)

§ 4　逆转公式……………………………………………(18)

§ 5　分布序列和特征函数序列的收敛性………………(20)

§ 6　附录…………………………………………………(27)

第二章　无穷可分分布……………………………………(36)

§ 1　无穷可分分布的定义及最简单的性质……………(36)

§ 2　无穷可分特征函数的典则表示……………………(38)

§ 3　一个辅助命题………………………………………(44)

§ 4　附录…………………………………………………(48)

第三章　关于独立随机变量之和的分布的一些不等式………(54)

§ 1　浓度函数……………………………………………(54)

§ 2　关于独立随机变量之和的浓度函数的一些不等式…(63)

§ 3　关于独立随机变量之和的最大值分布的一些不等式(78)

§ 4　关于独立随机变量之和的分布的指数估计…………(83)

§ 5　关于独立随机变量之和的矩的一些不等式…………(87)

§ 6　附录…………………………………………………(92)

第四章　无穷可分分布的收敛定理和中心极限定理………(109)

§ 1　无穷可分分布——独立随机变量和的分布的极限(109)

§ 2　收敛于给定的无穷可分分布的条件………………(122)

§ 3　L族极限分布与稳定分布……………………………（126）
§ 4　中心极限定理…………………………………………（137）
§ 5　附录……………………………………………………（154）
第五章　中心极限定理中的估计………………………（160）
§ 1　通过相应的 Fourier—Stieltjes 变换间的距离来估计有界变差函数间的接近程度……………………（160）
§ 2　Esseen 不等式和 Berry-Esseen不等式…………（166）
§ 3　Esseen 不等式的推广…………………………………（170）
§ 4　具有相同阶的上下界估计……………………………（179）
§ 5　非一致估计……………………………………………（188）
§ 6　中心极限定理中的渐近展开公式的结构…………（194）
§ 7　关于独立同分布随机变量和的中心极限定理的渐近展开………………………………………………（199）
§ 8　附录……………………………………………………（203）
第六章　大数定律………………………………………（227）
§ 1　弱大数律………………………………………………（227）
§ 2　独立随机变量级数的收敛性…………………………（235）
§ 3　强大数律………………………………………………（243）
§ 4　通过矩的和来估计部分和 S_n 的上升的阶………（257）
§ 5　附录……………………………………………………（264）
第七章　重对数律………………………………………（282）
§ 1　КОЛМОГОРОВ 定理…………………………………（282）
§ 2　Hartman-Wintner 定理……………………………（294）
§ 3　广义重对数律…………………………………………（297）
§ 4　附录……………………………………………………（303）
名词索引…………………………………………………（315）
苏联期刊、文集及出版社名称索引……………………（319）
参考文献…………………………………………………（321）

第一章　概率分布和特征函数

我们在本章中列出了一系列后面章节中将要用到的概率论中的基本概念和定理．其中大部分命题的证明均予略去，它们可以在一些常见的参考书，例如 А. А. Боровков〔18〕，Б. В. Гнеденко〔27〕，А. Н. Ширяев〔194〕以及Loev〔84〕，Feller〔183〕和 Lukacs〔85〕等人的著作中找到．对于某些较为专门的论断则附出了证明．

§1　随机变量和概率分布

设 Ω 是某个非空集合．其中的元素称为点或基本事件，用带有下标或不带下标的字母 ω 表示它们．集合 Ω 称为基本事件空间（或必然事件）．

设$\mathfrak{A}$是某个由基本事件空间 Ω 的子集形成的非空的集族，具有如下性质：1)如果 $A\in\mathfrak{A}$，则 $\Omega\setminus A\in\mathfrak{A}$；2)如果$A_1,A_2,\cdots$是属于$\mathfrak{A}$的有限或无限的集合序列，则 $\bigcup\limits_n A_n\in\mathfrak{A}$．集族$\mathfrak{A}$称为事件的 σ 代数，或称为事件的 Borel 域，它的元素都是事件．

如果$\mathfrak{A}$是事件的 σ 代数，则容易看出，有 $\Omega\in\mathfrak{A}$；并且，空集 ϕ（称为不可能事件）以及$\mathfrak{A}$中的有限或可列个事件集合的交集也属于$\mathfrak{A}$．

对于任何事件 $A\in\mathfrak{A}$ 都有定义的非负的可列可加函数 $P(A)$，满足正则化条件$P(\Omega)=1$，称为概率测度．$P(A)$之值称为事件 A 的概率．三元体（Ω，$\mathfrak{A}$，P）称为概率空间．

任何定义在 Ω 上的实函数 $X=X(\omega)$，都将基本事件空间 Ω 映射到实直线$\mathbf{R}$上．设 B是实直线上的某个点集，记

$X^{-1}(B)=\{\omega: X(\omega)\in B\}$. 集合 $X^{-1}(B)$ 是基本事件空间 Ω 的子集，称为集合 B 的逆象. 如果对于任何实直线上点的Borel 集合 B*，都有 $X^{-1}(B)\in\mathfrak{A}$，则函数 $X(\omega)$ 称为可测的.

有限实值可测函数 $X(\omega)$（$\omega\in\Omega$）称为随机变量. 对实直线上点的任何 Borel 集合 B 都有定义的函数 $P_X(B)=P(\{\omega: X(\omega)\in B\})$ 叫做随机变量 X 的概率函数. 今后，我们常用较为简单的记号 $P(X\in B)$ 来代替 $P(\{\omega: X(\omega)\in B\})$.

设 $(\Omega, \mathfrak{A}, P)$ 是概率空间，在它上面定义了随机变量 X. 随机变量 X 导出了一个新的概率空间（$\mathbf{R}$, $\mathscr{B}$, P_X），其中 $\mathscr{B}$ 是实直线 $\mathbf{R}$ 上的 Borel 集合的 σ 代数.

我们来对 B 是区间 $(-\infty, x)$，亦即实直线上满足不等式 $y<x$ 的点 y 的集合的场合来考察概率 $P(X\in B)$. 记 $F(x)=P(X<x)$. 函数 $F(x)$ 对任何实数 x 都有定义，它称为随机变量 X 的分布函数（d. f. ）.

分布函数 $F(x)$ 具有如下性质：1）$F(x)$ 非降，左连续；2）$\lim\limits_{x\to-\infty}F(x)=0$；3）$\lim\limits_{x\to+\infty}F(x)=1$. 反之亦真：即任何满足这三个条件的函数 $F(x)$ 都是定义在某个概率空间上的某个随机变量的分布函数.

我们在说到随机变量 X 的概率分布，或随机变量 X 的分布时，既可以是指该变量的概率函数 $P_X(B)$，又可以是指它的分布函数 $F(x)$.

随机变量 X 的分布称作离散的，如果存在实直线上点的有限或可数集合 B，使得 $P(X\in B)=1$. 如果 X 是具有离散分布的随机变量，且 $P(X=x)>0$，则数 x 称为随机变量 X 的可能值. 随机变量 X 具有格子点分布，如果它以概率 1 取形如 $a+kh$（$k=0, \pm1, \pm2, \cdots$）的值，其中 a 和 $h>0$ 为定数. 数 h 称为分布的步长. 如果对任何 a_1 和 $h_1>h$，随机变量 X 以概

---

*实直线上的 Borel 集合类定义为包含着所有区间在内的集合的最小 σ 代数.

率1所取的值都不能表示为 a_1+kh_1 ($k=0, \pm 1, \pm 2, \cdots$) 的形式，则步长 h 称为最大的.

随机变量 X 的分布称作连续的，如果对实直线上点的任何有限或可数集合 B，都有 $P(X\in B)=0$. 随机变量 X 的分布称作绝对连续的，如果对任何 Lebesgue 测度为零的 Borel 集合 B，都有 $P(X\in B)=0$. 随机变量 X 的分布称作奇异的，如果它是连续的且存在某个 Lebesgue 测度为零的 Borel 集合 B，使得 $P(X\in B)=1$.

使得随机变量 X 的分布为离散的充要条件是它的分布函数 $F(x)$ 是纯间断的. 使得随机变量 X 的分布为连续的充要条件是它的分布函数处处连续. 分布 F 为绝对连续的，当且仅当对任何 x 有

$$F(x)=\int_{-\infty}^{x} p(t)\,\mathrm{d}t,$$

其中 $p(x)$ 为实直线上可积的非负函数，称为分布密度.（此处的积分应理解为 Lebesgue 积分. ）我们仅能在分布为绝对连续的场合下谈及分布密度.

根据 Lebesgue 的分解定理. 分布函数 $F(x)$ 可以唯一地表示为如下的和的形式：

$$F(x)=c_1F_1(x)+c_2F_2(x)+c_3F_3(x) \tag{1.1}$$

这里 $c_k\geqslant 0$ ($k=1, 2, 3$), $c_1+c_2+c_3=1$，而 $F_1(x)$, $F_2(x)$ 和 $F_3(x)$ 分别为离散的、绝对连续的和奇异的分布函数.

点 x 称为分布函数 $F(x)$ 的增长点，如果对任何 $\varepsilon>0$. 都有 $F(x+\varepsilon)-F(x-\varepsilon)>0$. 分布 F 的所有增长点所组成的集合称作分布 F 的谱.

起着特别重要作用的分布有退化分布、二项分布和 Poisson 分布等三种离散分布，以及正态分布这样一种绝对连续分布. 随机变量 X 具有退化分布，如果存在实数 c，使得 $P(X=c)$

$=1$. 它的分布函数 $F(x)$ 为：当 $x\leqslant c$ 时 $F(x)=0$，而当 $x>c$ 时 $F(x)=1$.

设 n 是正整数，$0<p<1$. 随机变量具有参数为 (n, p) 的二项分布，如果对 $m=0, 1, \cdots, n$，有 $P(X=m)=C_n^m p^m(1-p)^{n-m}$.

设 λ 是正数，a 和 $b\neq 0$ 是实数. 随机变量 X 具有参数为 (a, b, λ) 的 Poisson 分布，如果对于任何非负整数 m 有

$$P(X=a+bm)=\frac{\lambda^m}{m!}e^{-\lambda}.$$

这种定义较之通常取 $a=0$ 和 $b=1$ 的定义更为广义.

设 $a\in\mathbf{R}$，$\sigma>0$. 随机变量 X 具有参数为 (a, σ) 的正态分布，或称具有正态 (a, σ) 分布，如果它具有分布密度

$$P(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}.$$

我们将正态 $(0, 1)$ 分布函数称为标准正态分布函数，今后恒用 $\Phi(x)$ 表示. 于是就有

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

如果 $X_1=X_1(\omega), \cdots, X_n=X_n(\omega)$ 是定义在同一个概率空间 $(\Omega, \mathfrak{A}, P)$ 上的随机变量，则向量 $X=(X_1, \cdots, X_n)$ 称为随机向量，或 n 维随机向量. 随机向量 X 的值域是 n 维欧氏空间 $\mathbf{R}^n$. 对于空间 $\mathbf{R}^n$ 中的任何 Borel 集合 B 都有定义的概率

$$P(X\in B)=P(\{\omega: (X_1(\omega), \cdots, X_n(\omega))\in B\})$$

称为随机向量 X 的概率函数. 特别地，对于任何实数 $x_1, \cdots, x_n$ 都有定义的函数

$$F(x_1, \cdots, x_n)=P\left(\bigcap_{k=1}^{n}\{\omega: X_k(\omega)<x_k\}\right)^{*}$$

* 通常将右端的概率写成如下形式：

$$P(X_1<x_1, \cdots, X_n<x_n).$$

叫做随机向量 $X=(X_1,\cdots,X_n)$ 的分布函数.

设 $(\Omega,\mathfrak{A},P)$ 是概率空间 而 $A_k\in\mathfrak{A}$ $(k=1,\cdots,n)$. 事件 $A_1,\cdots,A_n$ 称为是相互独立的，如果对任何整数 k $(2\leqslant k\leqslant n)$ 和任何满足条件 $1\leqslant i_1<\cdots<i_k\leqslant n$ 的整数 $i_1,\cdots,i_k$，都有

$$P\Big(\bigcap_{s=1}^{k}A_{i_s}\Big)=\prod_{s=1}^{k}P(A_{i_s}).$$

设 $X_1,\cdots,X_n$ 是定义在同一个概率空间 $(\Omega,\mathfrak{A},P)$ 上的随机变量. 称这些随机变量是相互独立的，或者简单地称为独立的，如果对于实直线上的任何 Borel 集合 $B_1,\cdots,B_n$，事件 $\{\omega:X_k(\omega)\in B_k\}$ $(k=1,\cdots,n)$ 相互独立. 随机变量 $X_1,\cdots,X_n$ 相互独立，当且仅当对任何实数 $x_1,\cdots,x_n$ 有

$$F(x_1,\cdots,x_n)=\prod_{k=1}^{n}F_k(x_k),$$

其中 $F(x_1,\cdots,x_n)=P(X_1<x_1,\cdots,X_n<x_n)$，$F_k(x)=P(X_k<x_k)$. 具有离散分布的随机变量 $X_1,\cdots,X_n$，如果相应的可能值集合分别为 $\{x_k^{(1)}\},\cdots,\{x_k^{(n)}\}$ $(k=1,2,\cdots)$，则它们的独立性等价于对任何整数 $k_1,\cdots,k_n$ 成立等式

$$P(X_1=x_{k_1}^{(1)},\cdots,X_n=x_{k_n}^{(n)})=\prod_{m=1}^{n}P(X_m=x_{k_m}^{(m)}).$$

如果随机变量 X_1 与 X_2 独立，且分别具有分布函数 $F_1(x)$ 和 $F_2(x)$，则和 X_1+X_2 具有分布函数

$$F(x)=\int_{-\infty}^{\infty}F_1(x-y)\,\mathrm{d}F_2(y).\qquad(1.2)$$

如上的积分称为分布 F_1 和 F_2 的卷积或褶合，记作 F_1*F_2. 对 $\mathbf{R}$ 上的有界变差函数 F_1 和 F_2 也可研究其卷积，它们不一定是分布函数. 此时我们仍按照 (1.2) 式来定义卷积 $F=F_1*F_2$.

对有界变差函数 $F(x)$ 的 n 重卷积，我们用 F^{*n} 来表示.

定义在同一个概率空间上的随机变量序列 X_1，X_2，…称为独立随机变量序列，如果对任何 n，随机变量 X_1，…，X_n 都相互独立. 对任何分布函数序列 F_1，F_2，…，都存在概率空间 $(\Omega, \mathfrak{A}, P)$ 和定义于其上的独立随机变量序列 X_1，X_2，…，使得对任何 n，随机变量 X_n 的分布函数是 F_n.

§2 随机变量的矩和其它数字特征

设 $(\Omega, \mathfrak{A}, P)$ 为概率空间，$X = X(\omega)$ $(\omega \in \Omega)$ 为随机变量. 既然概率空间是赋有测度的可测空间，所以能引入积分概念. 如果 $\int_{\Omega} |X| \mathrm{d}P < \infty$，则称随机变量 X 的数学期望存在，记作 EX，并将其定义为

$$EX = \int_{\Omega} X \mathrm{d}P.$$

显然有等式

$$EX = \int_{-\infty}^{\infty} x \mathrm{d}F(x)$$

成立，该式右端的积分是 Stieltjes 积分，其中 $F(x)$ 为随机变量 X 的分布函数.

设随机变量 X 具有分布函数 $F(x)$，而 $g(x)$ 是 Borel 函数*. 如果如下二条件之一成立：1）存在数学期望 $Eg(X)$，2）$\int_{-\infty}^{\infty} |g(x)| \mathrm{d}F(x) < \infty$，则另一条件也必成立，且有等式

*定义在 $\mathbf{R}$ 上的实函数 $g(x)$，如果对任何 $a \in \mathbf{R}$，集合 $\{x: g(x) < a\}$ 都是 Borel 集合，则称其为 Borel 函数.

$$Eg(X)=\int_{-\infty}^{\infty}g(x)\,\mathrm{d}F(x).$$

设k是正数．如果随机变量X^k的数学期望存在，就称它为随机变量X的k阶原点矩，记作α_k．于是就有

$$\alpha_k=EX^k=\int_{-\infty}^{\infty}x^k\mathrm{d}F(x),$$

其中$F(x)$是随机变量X的分布函数．如果原点矩α_k存在，则k阶绝对原点矩有限，我们用β_k表示它并定义为

$$\beta_k=E|X|^k=\int_{-\infty}^{\infty}|x|^k\mathrm{d}F(x).$$

如果对某给定的k，有α_k存在，则显然对所有正数$m\leqslant k$，都有α_m和β_m存在．k阶中心矩和k阶绝对中心矩分别由如下等式定义

$$\mu_k=E(X-EX)^k=\int_{-\infty}^{\infty}(x-\alpha_1)^k\mathrm{d}F(x),$$

$$\nu_k=E|X-EX|^k=\int_{-\infty}^{\infty}|x-\alpha_1|^k\mathrm{d}F(x).$$

2阶绝对中心矩ν_2称为方差．我们也可以用DX来表示随机变量X的方差．

如果$X_1,\cdots,X_n$是具有数学期望的独立随机变量，则有$E(X_1\cdots X_n)=EX_1\cdots EX_n$．如果$X_1,\cdots,X_n$是具有方差的两两独立的随机变量，则有$D(X_1+\cdots+X_n)=DX_1+\cdots+DX_n$.

如果X是具有数学期望的非负随机变量，则对任何$t>0$都有$P(X\geqslant t)\leqslant EX/t$．由此可以推知，对于任何具有2阶矩的随机变量$X$，都对任何$\varepsilon>0$成立如下的不等式

$$P(|X|\geqslant\varepsilon)\leqslant\varepsilon^{-2}EX^2 \tag{2.1}$$

$$P(|X-EX|\geqslant\varepsilon)\leqslant\varepsilon^{-2}DX,$$

（Чебышев不等式）.

如果 X 和 Y 是随机变量，$r>1$，$\frac{1}{r}+\frac{1}{s}=1$，则有

$$E|XY|\leqslant(E|X|^r)^{1/r}(E|Y|^s)^{1/s} \tag{2.2}$$

（Hölder 不等式）.

如果 $r\geqslant 1$，则有

$$(E|X+Y|^r)^{1/r}\leqslant(E|X|^r)^{1/r}+(E|Y|^r)^{1/r} \tag{2.3}$$

（Minkowski 不等式）.

Hölder 不等式的特例是 Cauchy–Буняковский 不等式

$$E|XY|\leqslant(EX^2)^{1/2}(EY^2)^{1/2}. \tag{2.4}$$

设 I 是实直线 $\mathbf{R}$ 上的有限或无限的开区间．设 $g(x)$ 是在 I 上连续的凸的实函数．又设 X 是以概率 1 取值于区间 I 的随机变量．如果存在着数学期望 EX 和 $Eg(X)$，则有

$$g(EX)\leqslant Eg(X) \tag{2.5}$$

（Jensen 不等式）.

为证(2.5)，我们注意到 $g(x)$ 的凸性，知对任何 $y\in I$，存在数 c，使得对一切 $X\in I$，都有 $g(x)\geqslant g(y)+c(x-y)$．在其中令 $x=X$，$y=EX$，再对所得的不等式取数学期望，即得(2.5)式.

如下的定理给出了另一个 Чебышев 不等式.

定理 1 设(a, b)是 $\mathbf{R}$ 上的有限或无限的区间(可以重合于 $\mathbf{R}$)．设 $u(x)$ 与 $v(x)$ 是(a, b)上的两个同为非增(或同为非降)的函数．又设 X 是以概率 1 取值于(a, b)的随机变量．则当相应的数学期望存在时，有

$$Eu(X)Ev(X)\leqslant E(u(X)v(X)). \tag{2.6}$$

证明 对任何 x，$y\in(a, b)$，都有如下不等式

$$(u(x)-u(y))(v(x)-v(y))\geqslant 0$$

成立．所以对于任何取值于$(a,\ b)$的随机变量X_1和X_2，我们都有

$$E[(u(X_1)-u(X_2))(v(X_1)-v(X_2))]\geqslant 0,$$

亦即

$$Eu(X_1)v(X_1)-Eu(X_2)v(X_1)-Eu(X_1)v(X_2)+Eu(X_2)v(X_2)\geqslant 0$$

（假定上述各数学期望都存在）．现设X_1和X_2是可交换的随机变量，亦即随机向量$(X_1,\ X_2)$与$(X_2,\ X_1)$具有相同的分布，就有

$$Eu(X_1)v(X_1)-2Eu(X_1)v(X_2)+Eu(X_1)v(X_1)\geqslant 0,$$

也就是

$$Eu(X_1)v(X_1)\geqslant Eu(X_1)v(X_2).$$

现在再设X_1与X_2为独立随机变量，且都与X同分布，我们就得到了$Eu(X)v(X)\geqslant Eu(X)Ev(X)$.

如果$u(x)$非增，而$v(x)$非降，则不等式（2.6）中的$\leqslant$号要改为$\geqslant$．只需在定理的证明中做某些必要的修改，就可证得这个结论．

如果随机变量X的k阶矩α_k存在，那么对于任何正数$m\leqslant k$就都有$\beta_m^{1/m}\leqslant\beta_k^{1/k}$和$\nu_m^{1/m}\leqslant\nu_k^{1/k}$．由此可知，对于任何$l$和$m$，都有$\beta_l\beta_m\leqslant\beta_{l+m}$和$\nu_l\nu_m\leqslant\nu_{l+m}$．这些不等式可由如下命题推出．

定理 2　如果X是随机变量，$\beta_r=E|X|^r$且$0<r<s$，则有

$$\beta_r^{1/r}\leqslant\gamma^{1/r-1/s}\beta_s^{1/s},\tag{2.7}$$

其中$\gamma=P(X\neq 0)$.

在$\gamma<1$，亦即$P(X=0)>0$的场合，不等式(2.7)是Ляпунов不等式的加强．(2.7)式中的等号可在随机变量X仅取0和1两个值，且相应的概率为$1-p$和p时达到，这里$0<p<1$.

只需对$\gamma>0$的场合证明(2.7)式，因为在$\gamma=0$，亦即$p(X=0)=1$时，它是显然的．不难验证$\ln\beta_r$是r的凸函数．事实上，只需在Cauchy–Буняковский不等式(2.4)中将X换成$|X|^{\frac{r-t}{2}}$，将Y换成$|X|^{\frac{r+t}{2}}$，其中$0<t\leqslant r$，我们即可得到

$$E|X|^r\leqslant(E|X|^{r-t}E|X|^{r+t})^{1/2},$$

也就是

$$\ln\beta_r\leqslant\frac{1}{2}\ln\beta_{r-t}+\frac{1}{2}\ln\beta_{r+t}.$$

于是当$\gamma>0$时，函数$\ln(\beta_r/\gamma)$也是凸的．曲线$y=\ln(\beta_r/\gamma)$，$r\geqslant0$在$(r,\ y)$平面上通过点(0，0)．故知，经过点(0，0)和点$(r,\ \ln(\beta_r/\gamma))$的直线的斜率$\frac{1}{r}\ln(\beta_r/\gamma)$是$r$的非降函数．因而$(\beta_r/\gamma)^{1/r}$是自变量$r$的非降函数。由此得到(2.7)式.

由(2.7)可以推知，对任何l和m有$\beta_l\beta_m\leqslant\gamma\beta_{l+m}$.

我们定义随机变量X的矩母函数$M(t)$为$M(t)=Ee^{tX}$，该数学期望对于$t=0$总是存在的，但却不一定总能在非退化的区间中存在．如果在区间$|t|<a$中有矩母函数存在，则在该区中有

$$M(t)=1+\sum_{k=1}^{\infty}\alpha_k\frac{t^k}{k!}.$$

如果存在常数C，使得$P(|X|\leqslant C)=1$，则随机变量X的矩母函数对于任何$t\in\mathbf{R}$都存在.

如果随机变量的矩母函数在某个以零为中心的非退化区间上存在，则称该随机变量满足Cramer条件.

如果存在非负常数b^2，使得随机变量X的矩母函数对任何$t\in\mathbf{R}$满足条件

$$M(t)\leqslant\exp\{b^2t^2/2\},$$