概率论极限理论引论

陆传荣　林正炎
陆传赉　编著

高等教育出版社

概率论极限理论引论

陆传荣　林正炎
陆传赉
编　著

高等教育出版社

内 容 提 要

本书是概率论极限理论的一本入门教材，内容分为：第一章预备知识；第二章叙述分析概率论的主要结果；第三章综述大数律及重对数律等较完整的结果；第四章论述概率测度弱收敛的基础理论，给出了泛函极限定理；第五章讨论强不变原理基本结果；第六章讨论鞅的弱收敛与强不变原理．本书可作为数学、概率论数理统计等专业的选修课与硕士研究生学位课程的教材，也可供概率统计专业的工作者参考．

概率论极限理论引论

陆传荣　林正炎　陆传赉

*

高等教育出版社出版

新华书店北京发行所发行

人民教育出版社印刷厂印装

*

开本 850×1168　1/32　印张 10.875　字数 260 000

1989 年 4月第 1 版　1989 年 4月第 1 次印刷

印数 0 001— 2,320

ISBN7-04-000595-6/O·225

定价 2.50 元

序 言

苏联著名概率论学者柯尔莫哥洛夫和格涅坚科在评论概率论极限理论时曾说过:“概率论的认识论的价值只有通过极限定理才能被揭示，没有极限定理就不可能去理解概率论的基本概念的真正含义.”概率论极限理论是概率论的主要分支之一，也是概率论的其它分支和数理统计的重要基础，极限理论的基本内容是每一概率统计工作者必须掌握的常识与工具. 十九世纪二十年代以前,中心极限定理是概率论研究的中心课题,经典极限理论是概率论发展史上的重要成果. 近代极限理论的研究至今方兴未艾，它不仅深化了经典理论的许多基本结果，也极大地拓展了自己的研究领域. 这些都是和概率论其他分支以及数理统计的最新发展相联系的.

基于包含这一学科的主要方面的基本内容、且能部分地反映近年来的一些新进展的教材或参考书至今还不多见，在概率统计界的前辈和朋友们的支持下我们编写了这本材料. 本书内容包括:关于独立随机变量和的经典的分析概率论；数理统计及随机过程理论中十分有用的泛函极限定理；近十几年中发展起来的强不变原理的基本结果和鞅的极限理论初步. 全书分六章，第一章是有关的概率论基本知识的简要回顾和必要补充；第二章叙述了分析概率论的主要结果；第三章综述了大数定律、重对数律和随机变量级数的收敛性；第四章论述了概率测度弱收敛——弱不变原理的基本理论；第五章介绍了强逼近理论的基本结果；第六章给出了鞅的弱与强不变原理的若干结果.

本书是概率论极限理论的一本入门教材，可作为概率统计(数理统计)专业研究生的学位课程教材，也可作为概率统计专业和数学专业的大学生的选修课教材，并可供需要这方面的知识的读者自学，作者也期望本书对于概率统计的专业工作者有参考价值.

本书主要由林正炎、陆传荣编写，陆传赉参加了第一章及第二章一部分的起草工作.

汪嘉冈教授、程士宏、刘秀芳副教授详细地审阅了原稿，提供了许多修改的具体建议，使本书增色不少. 严士健教授、陈希孺教授、江泽培教授始终关心和支持本书的出版. 高等教育出版社，特别是高尚华同志为本书的出版做了大量的工作，作者谨向他们表示衷心的感谢.

限于作者水平，不妥或谬误之处在所难免，恳请同行专家和广大读者不吝赐教.

作　　者

1987年12月17日

于杭州大学

缩写及记号

r.v. 随机变量

d.f. 分布函数

c.f. 特征函数

a.s. 几乎必然地

i.d. 无穷可分

i.i.d. 相互独立同分布

A_n i.o. 无穷多个 A_n 发生

$(\Omega, \mathscr{A}, P)$ 概率空间

A^c A 的余集(A 的逆事件)

$A \triangle B$ A、B 的对称差,即$(A \cap B^c) \cup (A^c \cap B)$

∂A A 的边界

mX 随机变量X的中位数

EX 随机变量X的数学期望

Var$X$ 随机变量X的方差

Cov(X,Y) 随机变量X与Y的协方差

$X_n \to X$(a.s.) 随机变量列$\{X_n\}$几乎必然地收敛于随机变量X

$X_n \to X$(P) 随机变量列$\{X_n\}$依概率收敛于随机变量X

$\xrightarrow{\mathrm{d}}$ 依分布收敛(或弱收敛)

$\xrightarrow{L_p}$ p 阶平均收敛

$\mu_n \Rightarrow \mu$ 测度序列$\{\mu_n\}$弱收敛于测度 μ

$\sigma(X_1, \cdots, X_n)$ 由随机变量 $X_1, \cdots, X_n$生成的σ域

$\sigma(\mathscr{G})$ 由集类$\mathscr{G}$生成的σ域

$I(A)$ 集A的示性函数

X^+ $\max(X, 0)$

X^- $-\min(X, 0)$

$\boldsymbol{R}^k$　k维欧氏空间

$\mathscr{B}^k$　k维 Borel 集全体

$\boldsymbol{R}^\infty$　无穷维欧氏空间

$\mathscr{B}^\infty$　无穷维 Borel 集

$w_x(\delta)$　函数 $x(t)$ $(0\leqslant t\leqslant 1)$ 的连续模

$[a]$　实数 a 的整数部分

$a_n=o(b_n)$　$\lim a_n/b_n=0$

$a_n=O(b_n)$　$\limsup a_n/b_n<\infty$

$\pi_{t_1,\cdots,t_k}$　在 $(t_1,\cdots,t_k)$ 上的投影映射

$X\circ\Phi$　复合函数 $X(\Phi(\cdot))$

目　录

序言 …… i
缩写及记号 …… iii
第一章　准备知识 …… 1
§1　随机变量与概率分布 …… 1
§2　数学期望及其性质 …… 5
§3　特征函数及其性质 …… 11
§4　分布函数列与特征函数列的收敛性 …… 19
§5　随机变量序列的收敛性 …… 22
§6　距离空间上的概率测度 …… 32
§7　鞅的基本概念 …… 39
习题 …… 43
第二章　无穷可分分布与普遍极限问题 …… 54
§1　无穷可分分布函数 …… 55
§2　独立随机变量和的极限分布 …… 67
§3　L族和稳定分布族 …… 82
§4　中心极限定理 …… 89
§5　中心极限定理的收敛速度，Esseen与Berry-Esseen不等式 …… 97
§6　非一致估计，分布函数的渐近展开 …… 106
§7　大偏差概率的估计 …… 111
习题 …… 118
第三章　大数定律 …… 128
§1　弱大数定律 …… 128
§2　独立随机变量和的收敛性 …… 138
§3　强大数定律 …… 148
§4　重对数律 …… 156

习题……171
第四章 概率测度的弱收敛……176
§1 几个常见的距离空间上概率测度的弱收敛性……177
§2 随机元序列的收敛性……183
§3 胎紧性(tightness),Прохоров 定理……190
§4 $C[0,1]$中概率测度弱收敛, Donsker 定理……196
§5 $D[0,1]$空间, Скороход 拓扑……205
§6 $D[0,1]$中概率测度弱收敛, Donsker 定理的一般化……214
§7 随机指标泛函中心极限定理……225
§8 经验过程的弱收敛性……233
§9 附录……241
习题……248
第五章 强不变原理……253
§1 Wiener过程及其基本性质……254
§2 Wiener过程的增量有多大……262
§3 Wiener过程的重对数律……269
§4 Скороход 嵌入定律……276
§5 强不变原理……283
习题……286
第六章 鞅的极限定理……288
§1 鞅收敛定理……288
§2 关于鞅的中心极限定理……295
§3 鞅的弱不变原理……304
§4 鞅的强不变原理……314
习题……330
参考书目……332
索引……334

第一章 准备知识

§1 随机变量与概率分布

设Ω是由一些元素组成的非空集，其元素(常记作ω)叫做**点**或**基本事件**. 通常称Ω为**基本事件空间**. 记$\mathscr{A}$为由Ω的某些子集组成的集类，如果它具有性质:

(i) $\Omega\in\mathscr{A}$,

(ii) 若$A\in\mathscr{A}$,则$A^c=\Omega\setminus A\in\mathscr{A}$,

(iii) 若$A_i\in\mathscr{A}$, $i=1,2,\cdots$, 则$\bigcup_i A_i\in\mathscr{A}$,

那么称$\mathscr{A}$为**事件的σ域**,并称$\mathscr{A}$中的元素为**事件**.

设$P(A)$ $(A\in\mathscr{A})$是定义在σ域$\mathscr{A}$上的实值集函数. 若它满足条件:

(i) 对每一$A\in\mathscr{A}$,有$0\leqslant P(A)\leqslant 1$,

(ii) $P(\Omega)=1$,

(iii) 对任意$A_n\in\mathscr{A}$ $(n=1,2,\cdots)$, $A_n\cap A_m=\varnothing$ $(n\neq m)$,

有

$$P\left\{\bigcup_{n=1}^{\infty}A_n\right\}=\sum_{n=1}^{\infty}P(A_n),$$

则称$P(\cdot)$是$\mathscr{A}$上的**概率测度**，简称**概率**. 称值$P(A)$为事件A的**概率**. 三元总体$(\Omega,\mathscr{A},P)$称为**概率空间**.

设$X=X(\omega)$是定义在Ω上的有限实值函数，如果它关于$\mathscr{A}$是可测的，就称X为一**随机变量**(简记为r. v.). 在以后我们也允

许 r.v. 在概率为零的集合上取无穷值. 记$\mathscr{B}$为 $R^1=(-\infty,\infty)$ 中所有Borel 集组成的集类(它是一个 σ 域). 我们称定义在$\mathscr{B}$上的集函数 $P_X(B)=P\{\omega: X(\omega)\in B\}$ $(B\in\mathscr{B})$ 为 r.v. X的**概率分布**. 这样, 由概率空间$(\Omega,\mathscr{A},P)$上的一个 r.v. X可以诱导出一个新的概率空间$(R^1,\mathscr{B},P_X)$.

记 $F(x)=P_X\{(-\infty,x)\}=P(X<x)$ $(x\in R^1)$, 称它是 r.v. X 的**分布函数**(简记为 d.f.). 易知它具有性质:

(i) $F(x)$是非负、不减、左连续的;

(ii) $\lim\limits_{x\to-\infty}F(x)=0, \lim\limits_{x\to\infty}F(x)=1.$

反之,可以证明任一具有性质(i)、(ii)的函数 $F(x)$必为某个概率空间上某一 r.v. X的 d.f.

如果在 R^1 上存在一个有限或可列的点集 B, 使得 $P(X\in B)=1$, 则称 r.v. X 具有**离散分布**. 此时称使得 $P(X=x)>0$ 的 x 为 r.v. X 的**可能值**.

如果 r.v. X 对 R^1上任一有限或可列的点集 B, 有$P(X\in B)=0$, 就称 r.v. X 的分布是**连续的**. 若对任何 L测度 (Lebesgue 测度)为 0 的 Borel 集 B, 有 $P(X\in B)=0$, 则称 r.v. X 具有**绝对连续的分布**. 若X的分布是连续的, 且存在 L 测度为 0 的 Borel 集 B, 使 $P(X\in B)=1$, 则称 r.v. X 的分布是**奇异的**.

r.v. X 的分布为离散的充要条件是它的 d.f. $F(x)$ 是离散型的. X 的分布为连续的充要条件是它的 d.f. $F(x)$是处处连续的. X 的分布是绝对连续的充要条件是它的 d.f. $F(x)$是绝对连续的, 即对每一 x 有

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

其中 $p(x)$ 是非负可积函数, 称为 r.v. X 的**概率(分布)密度**或**密度函数**.

若对每一$\varepsilon>0$，有$F(x+\varepsilon)-F(x-\varepsilon)>0$，则称$x$是d.f. $F(x)$的一个支撑点. 支撑点的全体组成的集合称为F的**支撑**. d.f. 的不连续点是它的支撑点.

设$(\Omega,\mathscr{A},P)$是一概率空间，如果$A\in\mathscr{A}$，$P(A)>0$，且对任何$B\subset A$，或者$P(B)=0$，或者$P(B)=P(A)$，则称A是P的一个**原子**.

由Lebesgue分解定理和单调函数的性质可知，任一d.f. $F(x)$可唯一地分解成如下形式：

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

其中$c_i\geqslant 0\ (i=1,2,3)$，$\sum_{i=1}^{3}c_i=1$，$F_1(x)$、$F_2(x)$、$F_3(x)$分别是离散的、绝对连续的和奇异的d.f.

设X是r.v.，当X与$-X$有相同的分布时，我们称X是**对称的r.v.**

设$X_1,X_2,\cdots,X_n$是定义在同一概率空间上的r.v.，则称$\boldsymbol{X}=(X_1,\cdots,X_n)$为$n$**维随机变量或随机向量**. 对任一$n$维欧氏空间$R^n$中的Borel集$B$，记

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

称$P_{\boldsymbol{X}}$是随机向量$\boldsymbol{X}$的**概率分布**. 记

$$F(x_1,\cdots,x_n)=P\{X_1<x_1,\cdots,X_n<x_n\},\qquad (x_1,\cdots,x_n)\in R^n,$$

称它为n维随机变量$(X_1,\cdots,X_n)$的n**元分布函数**或随机向量$\boldsymbol{X}$**的分布函数**. 易知它具有性质：

(i) 对每一$x_i\ (i=1,\cdots,n)$，$F(x_1,\cdots,x_n)$是非负不减左连续的.

(ii) $\lim\limits_{x_i\to-\infty}F(x_1,\cdots,x_n)=0\quad (i=1,\cdots,n)$,

$$\lim_{x_1\to\infty,\cdots,x_n\to\infty}F(x_1,\cdots,x_n)=1.$$

(iii) 对 R^n 中的任一矩形集 $A=\{x=(x_1,\cdots,x_n):\ a_i\leqslant x_i<b_i,1\leqslant i\leqslant n\}$，有

$$P(X\in A)=P\{a_i\leqslant x_i<b_i;1\leqslant i\leqslant n\}$$

$$=F(b_1,\cdots,b_n)-\sum_{i=1}^{n}F(b_1,\cdots,a_i,\cdots,b_n)+\cdots$$

$$+(-1)^nF(a_1,\cdots,a_n).$$

称

$$F_j(y)=\lim_{\substack{y_i\to\infty\\ i\neq j}}F(y_1,\cdots,y_{j-1},y,y_{j+1},\cdots,y_n)$$

为随机向量 X 的**一元边际分布**.

设 $X_1,\cdots,X_n$ 是 n 个 r.v.，称它们是**相互独立的**，若对 R^1 中任意 n 个 Borel 集 $B_1,\cdots,B_n$，事件$\{\omega: X_i(\omega)\in B_i\}$ $(1\leqslant i\leqslant n)$是相互独立的. 易知 r.v. $X_1,\cdots,X_n$ 相互独立当且仅当对任何实数 $x_1,\cdots,x_n$ 有

$$F(x_1,\cdots,x_n)=\prod_{i=1}^{n}F_i(x_i),$$

其中 $F_i(x)=P(X_i<x)$.

设 r.v. X_1 和 X_2 独立且它们的 d.f. 分别为 $F_1(x)$ 和 $F_2(x)$，那么 r.v. X_1+X_2 的 d.f. 为

$$F_1*F_2(x)=\int_{-\infty}^{\infty}F_1(x-u)dF_2(u),$$

F_1*F_2 称为 F_1 与 F_2 的**卷积**.

若 $X_1,X_2,\cdots$是一列 r.v.，且对每一n，$X_1,\cdots,X_n$ 是相互独立的，则称此 r.v. 序列是**相互独立的**. 又若 $X_1,\cdots,X_{n+m}$ 是相互独立的 r.v.，而 f 和 g 分别是定义在 R^m 和 R^n 上、取值于 R^1 中的 Borel 函数，那么 $f(X_1,\cdots,X_m)$和$g(X_{m+1},\cdots,X_{n+m})$是独立的 r.v.

§2 数学期望及其性质

设 $X=X(\cdot)$ 是定义在概率空间 $(\Omega, \mathscr{A}, P)$ 上的 r.v.，如果 $\int_{\Omega}|X|dP<\infty$，就称r.v. X **的数学期望**或**均值**存在(或称 r.v. X 是可积的)，记作 EX，它由下式定义：

$$EX=\int_{\Omega} XdP. \tag{2.1}$$

利用积分变换，也可写 $EX=\int_{-\infty}^{\infty}xdF(x)$，此处 $F(x)$ 是 r.v. X 的 d.f.

设 $g(x)$ 是 R^1 上的 Borel 可测函数，如果 r.v. $g(X)$ 的数学期望存在，即 $E|g(X)|<\infty$，则由积分变换可知

$$Eg(X)=\int_{\Omega} g(X)dP=\int_{-\infty}^{\infty}g(x)dF(x). \tag{2.2}$$

设 k 是正整数，若 r.v. X^k 的数学期望存在，就称它为 X 的 k **阶原点矩**，记为 α_k. 由(2.2)知

$$\alpha_k=EX^k=\int_{-\infty}^{\infty}x^k dF(x). \tag{2.3}$$

设 k 是正实数，$|X|^k$ 的数学期望存在，就称它为 X **的 k 阶绝对矩**，记为 β_k. 由(2.2)知

$$\beta_k=E|X|^k=\int_{-\infty}^{\infty}|x|^k dF(x). \tag{2.4}$$

类似地，X 的 k **阶中心矩** μ_k 和 k **阶绝对中心矩** ν_k 分别由下列等式定义：

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

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

我们称二阶中心矩为方差，记作 $\mathrm{Var}X$ 或 DX. 显然有

$$\mathrm{Var}X=\mu_2=\nu_2=\alpha_2-\alpha_1^2.$$

容易验证，若 α_k 存在，则对一切正整数 $m\leqslant k$, α_m 存在；若 β_k 存在，则对一切正数 $m\leqslant k$, β_m 存在且有 $\beta_m^{1/m}\leqslant\beta_k^{1/k}$. 从而，对每一 l 和 m 有 $\beta_m\beta_l\leqslant\beta_{m+l}$. 对 ν_k 也有类似的结论.

关于数学期望，容易验证下列性质成立：

1) 若 r. v. X,Y 的期望 EX 和 EY 存在，则对任意实数 α、β, $E(\alpha X+\beta Y)$ 也存在，且

$$E(\alpha X+\beta Y)=\alpha EX+\beta EY.$$

2) 设 $A\in\mathscr{A}$, 用 I_A 表示集 A 的示性函数. 若 EX 存在，则 $E(XI_A)$ 也存在，且

$$E(XI_A)=\int_A XdP.$$

此外，$E(|X|I_A)=0$ 的充要条件是 $P(A)=0$ 或者在 A 中 a. s. 成立着 $X=0$(a. s. 是“几乎必然”的简写).

3) 若 EX 存在，则下面三命题等价：

(i) $X=0$ a. s.,

(ii) 对一切 $A\in\mathscr{A}$, $E(XI_A)=0$,

(iii) $E|X|=0$.

4) 若 $\{A_k\}$ 是 Ω 的一个分划，即 $A_k\cap A_m=\varnothing\ (k\neq m)$ 且 $\Omega=\bigcup_k A_k$, 则

$$EX=\int_\Omega XdP=\sum_k\int_{A_k}XdP.$$

5) 设 X 是 r. v., $A\in\mathscr{A}$. 若有 α、$\beta\in R^1$ 使得在 A 上 a. s. 成立着 $\alpha\leqslant X\leqslant\beta$, 则 $\int_A XdP$ 存在且有

$$\alpha P(A)\leqslant\int_A XdP\leqslant\beta P(A).$$

特别地，若 X 是在 Ω 上 a. s. 有界的，则 EX 必存在. 进一步，若有 r. v. X_1 和 X_2，EX_1 和 EX_2 都存在，且在 A 上 a. s. 成立着 $X_1\leqslant X\leqslant X_2$，则 EX 存在并且

$$\int_A X_1 dP\leqslant\int_A X dP\leqslant\int_A X_2 dP.$$

6) 设 EX 存在，并由关系式

$$G_X(A)=\int_A X dP,\quad A\in\mathscr{A}$$

定义 $\mathscr{A}$ 上的集函数 $G_X(\cdot)$，那么它是可列可加的. 特别地，当 $X\geqslant 0$ a. s. 时，G_X 是 $\mathscr{A}$ 上的一个有限测度.

下面叙述关于矩的若干不等式，它们是十分有用的.

1) Чебышёв **不等式** 对任一 $\varepsilon>0$，

$$P(|X|\geqslant\varepsilon)\leqslant EX^2/\varepsilon^2. \tag{2.5}$$

2) 设 $g(x)$ 是 R^1 上的非负 Borel 可测函数，且 $Eg(X)<\infty$. 若 $g(x)$ 是偶的，且在 $[0,\infty)$ 上是不减的，那么对任一 $\varepsilon>0$，有*

$$\frac{Eg(X)-g(\varepsilon)}{\text{a. s. sup}\,g(X)}\leqslant P(|X|\geqslant\varepsilon)\leqslant\frac{Eg(X)}{g(\varepsilon)}. \tag{2.6}$$

证 记 $A=\{|X|\geqslant\varepsilon\}$. 由

$$\begin{aligned}Eg(X)&=\int_A g(X)dP+\int_{A^c}g(X)dP\\&\leqslant\text{a. s. sup}\,g(X)P(A)+g(\varepsilon)\end{aligned}$$

即可推出(2.6) 式的第一个不等式. 由条件不难验证第二个不等式也成立.

注 若令 $g(x)=|x|^r$，$r>0$，则有

$$\frac{E|X|^r-\varepsilon^r}{\text{a. s. sup}|X|^r}\leqslant P(|X|\geqslant\varepsilon)\leqslant\frac{E|X|^r}{\varepsilon^r}. \tag{2.7}$$

后一不等式常称为 Марков **不等式**.

若令 $g(x)=\dfrac{|x|^r}{1+|x|^r}$，$r>0$，则易知 a. s. sup$\dfrac{|X|^r}{1+|X|^r}\leqslant 1$. 故

---

* 定义 a. s. sup$|X|=\inf\{c: 0\leqslant c\leqslant\infty, P(|X|>c)=0\}$.

$$E\frac{|X|^r}{1+|X|^r}-\frac{\varepsilon^r}{1+\varepsilon^r}\leqslant P(|X|\geqslant\varepsilon)\leqslant\frac{1+\varepsilon^r}{\varepsilon^r}E\frac{|X|^r}{1+|X|^r}. \tag{2.8}$$

3) c_r-**不等式** 设 X、Y 是任两个 r. v.，那么

$$E|X+Y|^r\leqslant c_r(E|X|^r+E|Y|^r)\ (r>0), \tag{2.9}$$

其中 $c_r=1$，当 $0<r\leqslant 1$；$c_r=2^{r-1}$，当 $r\geqslant 1$.

4) Hölder **不等式** 设实数 p、q 都大于 1 且 $\frac{1}{p}+\frac{1}{q}=1$，那么就有

$$E|XY|\leqslant(E|X|^p)^{\frac{1}{p}}(E|Y|^q)^{\frac{1}{q}}. \tag{2.10}$$

5) Schwarz **不等式** 当 4) 中 $p=q=2$ 时，有

$$E|XY|\leqslant(EX^2)^{\frac{1}{2}}(EY^2)^{\frac{1}{2}}. \tag{2.11}$$

6) Minkowski **不等式** 对任意 r. v. X 和 Y，实数 $p\geqslant 1$，

$$(E|X+Y|^p)^{\frac{1}{p}}\leqslant(E|X|^p)^{\frac{1}{p}}+(E|Y|^p)^{\frac{1}{p}}. \tag{2.12}$$

7) Jensen **不等式** 设 $g(x)$ 是定义在 R^1 上的实函数，若对任意的实数 x_1、x_2 及 $\lambda\in[0,1]$ 有

$$g(\lambda x_1+(1-\lambda)x_2)\leqslant\lambda g(x_1)+(1-\lambda)g(x_2), \tag{2.13}$$

则称 $g(x)$ 是 R^1 上的**凸函数**.

若 $g(x)$ 是凸函数，r. v. X 的数学期望 EX 存在，则成立下列 Jensen 不等式：

$$g(EX)\leqslant Eg(X). \tag{2.14}$$

证明如下：由凸函数的定义易知

$$g(x)\geqslant g(x_0)+(x-x_0)g'_-(x_0),$$

其中 g'_- 表示 g 的左导数. 取 $x=X$，$x_0=EX$，得

$$g(X)\geqslant g(EX)+(X-EX)g'_-(EX).$$

两边同时取期望即得(2.14)式.

除去上述不等式外，关于矩的存在性可写出如下的必要条件