顺序极值引论

谢盛荣

西南师范学院数学系

一九八四年五月

目　录

序

关于序列极值的问题已有几十年的研究历史．这份教材是为高等院校数学系四年级学生编写的选修课教材，也可以作为青年科研工作者的读物．内容是引论性的，证明力求严密而初等，只设想读者具有较好的数学分析基础及概率论基础．编者期待读者通过这门课程的学习进一步加深和巩固过去所学过的基础知识，并期待某些优秀生能由此而迅速通往科学研究的前沿．

编者在东北师大学习期间，得到苏明礼老师的指导，近几年仍不断得到苏明老师的指引和支持，感到受益不浅．现在又在我系领导、办公室及资料室的支持和协助下写成此稿，借此一并表示感谢．

由于编者水平有限，错误难免，恳请专家和读者提出宝贵意见，以便迅速改进．

编者 1984年4月

第一章 顺序统计量与极值

§1 背景

设$\xi_1, \cdots \xi_n$是n个随机变量，若$(x_1, \cdots, x_n)$是$(\xi_1, \cdots, \xi_n)$的一组取值，将这组值由大到小排列，得到

$$x_1^* \geq x_2^* \geq \cdots \geq x_n^*.$$

我们定义$M_n^{(r)} = x_r^*$，于是$M_n^{(r)}$是向量$(\xi_1, \cdots, \xi_n)$的函数，且显然有$M_n^{(n)} \leq M_n^{(n-1)} \leq \cdots \leq M_n^{(1)}$及$M_n^{(n)} = \min\limits_{1\leq i\leq n} \xi_i$，$M_n^{(1)} = \max\limits_{1\leq i\leq n} \xi_i$。

称$M_n^{(1)}, \cdots, M_n^{(n)}$为$(\xi_1, \cdots, \xi_n)$的一组顺序统计量，而$M_n^{(r)}$称为第$r$个上极值，$M_n^{(1)}$常简称为最大值。由于$-M_n^{(n-r+1)}$也即是$(-\xi_1, \cdots -\xi_n)$的第$r$个上极值，所以下极值$M_n^{(n-r+1)}$方面的研究可转化为上极值方面的研究。为节省篇幅，我们只研究上极值。

极值问题曾引起不少学者的注意。一方面是因为这个问题具有现实意义。例如，洪水、暴雨、高温、台风、地震等灾害性问题都可以建立极值模型来研究，以便为工程设计提供理论依据(本章§3将介绍一些有关的应用)。另一方面，极端值的研究往往比较具有理论价值。从1922年以来就有人研究这个问题，早期有许多知名学者，像Fréchet, von Mises, Гнеденко, Смирнов, Cramer等都曾对之作过贡献，他们主要研究了独立同分布序列的极值的弱收敛问题。近十多年来，在平稳序列及一般相依序列的极值的研究方面也取得了许多成果。作为大学生的选修课教材，我们将着重介绍一些基本的易于接受的成果。

§2 iid序列的极值的分布

设$\xi_1, \xi_2, \cdots \xi_n$是相互独立且同分布的随机变序列，常简称

这类序列为iid.序列．设其公共分布函数为$F(x)$，并以$\hat{M}_n^{(r)}$记此iid.序列的第r个上极值．

注意事件"$\hat{M}_n^{(r)}\le x$"等价于"n个随机变另的取值超过x者至多有$r-1$个"，因此可得$\hat{M}_n^{(r)}$的分布函数为

$$F_r(x)=P(\hat{M}_n^{(r)}\le x)=\sum_{i=0}^{r-1}\binom{n}{i}(1-F(x))^i\,F^{n-i}(x) \tag{1}$$

特别是有最大值及最小值的分布函数分别为

$$F_1(x)=F^n(x) \text{ 及 } F_n(x)=1-[1-F(x)]^n \tag{2}$$

(1)式也可写成

$$F_r(x)=\frac{n!}{(r-1)!(n-r)!}\int_0^{F(x)}x^{n-r}(1-x)^{r-1}\,dx \tag{3}$$

事实上，(3)右端是所谓不完全β函数，经重复使用P分积分法即可转化为(1)式的右端．又若密度函数$f(x)=F'(x)$存在，则得$\hat{M}_n^{(r)}$的密度函数为

$$f_r(x)=F_r'(x)=\frac{n!}{(r-1)!(n-r)!}F^{n-r}(x)(1-F(x))^{r-1}f(x) \tag{4}$$

例1 设iid.序列$X_1,\cdots,X_n$服从参数为λ的指数分布．其密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}\qquad(\lambda>0).$$

由(4)得$M_n^{(1)}$，$M_n^{(n)}$的密度函数分别为

$$f_1(x)=\begin{cases}n\lambda e^{-\lambda x}[1-e^{-\lambda x}]^{n-1}, & x\geqslant 0\\ 0, & x<0.\end{cases}\qquad(\lambda>0)$$

$$f_2(x)=\begin{cases}n\lambda e^{-n\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}\qquad(\lambda>0).$$

下面求最大值与最小值的联合密度.

设 $y>x$, $\hat{M}_n^{(1)}$ 与 $\hat{M}_n^{(n)}$ 分别落在 $(y, y+dy)$ 及 $(x, x+dx)$ 内相当于 $\xi_1, \cdots, \xi_n$ 中各有一个落在 $(y, y+dy)$ 及 $(x, x+dx)$ 之中, 而其余都落在 $(x+dx, y)$ 中, 此事件发生的概率为

$$n(n-1)f(x)f(y)(F(y)-F(x))^{n-2}\,dx\,dy$$

故 $(\hat{M}_n^{(1)}, \hat{M}_n^{(n)})$ 的联合密度当 $y \geq x$ 时为

$$n(n-1)f(x)f(y)(F(y)-F(x))^{n-2} \tag{5}$$

而 $y<x$ 时为 0.

用类似的方法可得 $(\hat{M}_n^{(1)}, \cdots, \hat{M}_n^{(n)})$ 的联合密度函数当 $x_1 \geq x_2 \geq \cdots \geq x_n$ 时为

$$n!\,f(x_1)\cdots f(x_n) \tag{6}$$

而其它情形下则为 0.

例 2 设 iid. 序列 $\xi_1, \cdots \xi_5$ 在 $[0.1]$ 上服从均匀分布, 则 (6) 表明 $(\hat{M}_5^{(1)}, \cdots, \hat{M}_5^{(5)})$ 的联合密度函数当 $1 \geq x_1 \geq \cdots \geq x_5 \geq 0$ 时为 $5!$, 此外则为 0. 于是 $(\hat{M}_n^{(1)}, \hat{M}_n^{(n)})$ 的联合密度当 $x_1 \geq x_5$ 时用求边缘分布的方法可得

$$5!\int_{x_5}^{x_1} dx_4 \int_{x_4}^{x_1} dx_3 \int_{x_3}^{x_1} dx_2 = \frac{5!}{3!}(x_1-x_5)^3 = 20(x_1-x_5)^3$$

而当 $x_1 < x_5$ 时为 0. 显然上式右端也可直接由 (5) 得到.

关于极值的函数, 在这里我们仅涉及中程与极差, 其定义分别为

$$M = \frac{1}{2}(\hat{M}_n^{(1)} + \hat{M}_n^{(n)}), \qquad R = \hat{M}_n^{(1)} - \hat{M}_n^{(n)} \tag{7}$$

其逆变换的 Jacobi 行列式之值易知为 1, 于是由 (5) 可得 (M, R) 的联合密度为 $y \geq 0$ 时为

$$n(n-1)f\left(\frac{2x+y}{2}\right)f\left(\frac{2x-y}{2}\right)\left(F\left(\frac{2x+y}{2}\right)-F\left(\frac{2x-y}{2}\right)\right)^{n-2} \quad (8)$$

而当 $y<0$ 时为 0 .

§3 复合极值的分布与应用

设 $\{X_i\}$ 是 iid. 序列，Y 是某一随机变易，N 是独立于 $\{X_i\}$ 及 Y 的非负整值随机变易。称随机变易

$$Z=\begin{cases} Y & N=0 \\ \max\limits_{1\le i\le n}\{X_i\} & N\geqslant 1 \end{cases} \quad (9)$$

为复合最大值。关于复合极值的分布有如下定理：

定理 1 iid 序列 $\{X_i\}$ 具有公共分布函数 $G(x)$，随机变易 Y 具有分布函数 $Q(x)$，则 (9) 式中复合极值 Z 的分布函数为

$$F(x)=\sum_{k=0}^{\infty} p_k[G(x)]^k-p_0[1-Q(x)]$$

其中 $P(N=k)=p_k \qquad k=0,1,\cdots$

证 $F(x)=P(Z<x)=P(Z<x,N=0)+\sum\limits_{k=1}^{\infty}P(Z<x,N=k)$

$$=P(Z<x\mid N=0)P(N=0)+\sum_{k=1}^{\infty}P(Z<x\mid N=k)P(N=k)$$

$$=P(Y<x)p_0+\sum_{k=1}^{\infty}P(\max_{1\le i\le k}\{X_i\}<x)\,p_k$$

$$=\sum_{k=0}^{\infty}p_k[G(x)]^k-p_0[1-Q(x)]$$

常记 $F_0(x)=\sum\limits_{k=0}^{\infty}p_k[G(x)]^k$，故当 $p_0=0$ 或 $Q(x)=1$ 时有 $F_0(x)=F(x)$.

例 1. N 服从单点分布，$P(N=k_0)=1$ 而 $k_0\neq 0$. 则得

$$F(x)=\sum_{k=0}^{\infty}p_k[G(x)]^k=[G(x)]^{k_0}=P(M_{k_0}^{(1)}\leq x)$$

例2．N服从二项分布 $p_k=\binom{n}{k}p^k(1-p)^{n-k}$ 则得

$$F_0(x)=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}[G(x)]^k=(pG(x)+1-p)^n$$

例3．N服从 Poisson 分布，$p_k=\frac{\lambda^k}{k!}e^{-\lambda}$ 则得

$$F_0(x)=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!}[G(x)]^k=e^{-\lambda[1-G(x)]}$$

复合极值的分布有许多实际应用．下面引进几个实例

1．用于研究最大震级 建设规划中关心的是本地T年内可能发生的最大地震是几级，为此需研究T年内最大震级ξ^*的分布．

我们作下列假定：

1）承认李希特与古登堡关于地震次数与震级之间的经验公式

$$n(x)=10^{a-bx}\qquad(x>0)$$

其中x代表震级，$n(x)$是x附近单位震级范围内的地震次数．通过对该地区的历史资料的统计，可以近似估计a与b之值．令$\alpha=a\ln 10$，$\beta=b\ln 10$，上式可写成

$$n(x)=e^{\alpha-\beta x}$$

由频率近似代替概率的思想可得地震级ξ的分布函数为

$$P(\xi\leq x)=\frac{\int_0^x n(y)dy}{\int_0^{\infty}n(y)dy}=\frac{\frac{e^{\alpha}}{\beta}(1-e^{-\beta x})}{\frac{e^{\alpha}}{\beta}}=1-e^{-\beta x}\qquad(x\geq 0)$$

其中$\beta>0$是常数．当$x<0$时，分布函数为0．

2）一年内该地区发震次数是一Poisson流，具有Poisson分布$P(\lambda)$．从而T年内发生的地震总次数N服从分布$P(T\lambda)$，

显然 $\lambda = \dfrac{EN}{T}$.

3). ξ_i 表示第 i 次地震的震级，$\{\xi_i\}$ 是 iid. 序列具有 1) 中所述的指数分布．且认为 $N, \xi_1, \xi_2, \cdots$ 独立．

4) η 记"无震"时的震级．另取值 0．得其分布函数为

$$Q(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

在上述有关假定下，由 (9) 及上面的例 2 得

$$P(\xi^* \leq x) = F(x) = \begin{cases} e^{-\lambda T e^{-\beta x}}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Ⅱ．用于研究暴雨　历史资料表明，長江流域一带年暴雨次数 N 服从某 Poisson 分布，参数 λ 随各地区有所不同．重庆、北碚地区资料表明，暴雨量 ξ 服从某指数分布

$$F(x) = \begin{cases} 1 - e^{-\theta(x-5)}, & x \geq 5 \\ 0, & x < 5 \end{cases} \quad (\theta > 0 \text{ 是常数})$$

其中雨量在 5 以上才记为暴雨量．以 η 记 T 年中无暴雨时的最大雨量，可以认为 $P(\eta \leq 5) = 1$．与讨论地震情形类似，可得北碚地区 T 年内最大暴雨量 ξ_T^* 的分布函数为

$$P(\xi_T^* \leq x) = \begin{cases} e^{-\lambda T e^{-\theta(x-5)}}, & x \geq 5 \\ 0, & x < 5 \end{cases}.$$

Ⅳ．研究台风影响　若频次看成服从 Poisson 分布 $P(\lambda)$，而台风影响下形成的波高经 χ^2 检验与标准 Gumbel 分布 $G(x) = e^{-e^{-x}}$ 相符合．在设计中主要是要解决对已给设计值 R 求 x_R 使下式成立

$$\sum_{k=0}^{\infty} p_k [G(x_R)]^k = R$$

用例 3 即 x_R 满足下式

$$e^{-\lambda[1-G(x_R)]}=R$$

所以得
$$e^{-e^{-x_R}}=1+\frac{1}{\lambda}\ln R$$

解之得
$$x_R=-\ln(\ln(1+\frac{1}{\lambda}\ln R))$$

此即台风影响下多年一遇设计波高的一种形式。

§4 指数分布与极值的期望

如众所周知，指数分布

$$F(x)=\begin{cases}1-e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}\quad(常数\lambda>0)\tag{10}$$

具有重要的实际应用。它的重要性还在于它有"无记忆性"。事实上，设随机变量X服从指数分布(10)，则对任意的$x\geq 0$，$y\geq 0$有

$$P(X>x+y|X>x)=\frac{P(X>x+y)}{P(X>x)}=\frac{e^{-\lambda(x+y)}}{e^{-\lambda x}}=e^{-\lambda y}\tag{11}$$

即 $P(X>x+y|X>x)=P(X>y)$，此等价于

$$P(X>x+y)=P(X>x)P(X>y),\quad x\geq 0且y\geq 0.\tag{12}$$

定理1 指数分布是唯一具有性质(12)的连续分布函数。

证 下面只证充分性即可。

若(12)成立，令$G(x)=P(X>x)$，得

$$G(x+y)=G(x)G(y),\quad x\geq 0且y\geq 0.\tag{13}$$

于是$G(nx)=[G(x)]^n$，从而有 $G(1)=[G(\frac{1}{n})]^n$，记$a=G(1)$得$G(\frac{m}{n})=a^{\frac{m}{n}}$。

设$x>0$是一无理数，存在正整数i_n使得$i_n-1\leq nx<i_n$，即$x\in[\frac{i_n-1}{n},\frac{i_n}{n})$。由于$G(x)$单调下降，得

$$a^{\frac{i_n-1}{n}}=G\left(\frac{i_n-1}{n}\right)\geqslant G(x)\geqslant G\left(\frac{i_n}{n}\right)=a^{\frac{i_n}{n}}$$

上式中令 $n\uparrow\infty$ 得 $G(x)=a^x$. 由于 $1-G(x)$ 是分布函数，可知必有 $0<a<1$. 记 $a=e^{-\lambda}$，此 $\lambda>0$. 又因 X 的分布函数是连续型的，故其分布函数如 (10) 所示.

指数分布的特征还可以由其最小值的期望的特点来确定.

由 (2) 可得对 $x>0$ 及 $n\geqslant 1$ 有

$$P(\hat{M}_n^{(n)}\geqslant x)=G^n(x)=(1-F(x))^n$$

而 (13) 表明 $G^n(x)=G(nx)$. 令 $x=\frac{y}{n}$，上式可写为

$$P(n\hat{M}_n^{(n)}\geqslant y)=P(X_1>y),\qquad y>0. \tag{14}$$

可见无记忆性蕴含着 $n\hat{M}_n^{(n)}$ 与 X_1 的分布相同，从而 $n\hat{M}_n^{(n)}$ 与 X_1 有相同的期望. 反之，可以证明极值期望具有这一特点的连续分布函数必是指数分布.

为此，要证明一个有关极值期望的结果.

定理 2　设 $F(x)$ 与 $G(x)$ 分别是 i.i.d. 序列 $\{X_i\}$ 与 $\{Y_i\}$ 的公共分布函数. 且 EX_1 及 EY_1 均存在. 则对 $n\geqslant 1$ 有 $E(\max\limits_{1\leqslant j\leqslant n}X_j)=E(\max\limits_{1\leqslant j\leqslant n}Y_j)$ 的充要条件是 $F(x)\equiv G(x)$.

证　由定义知

$$E(\max_{1\leqslant j\leqslant n}X_j=n\int_{-\infty}^{\infty}x(F(x))^{n-1}\,dF(x)=n\int_0^1F^{-1}(u)\,u^{n-1}du$$

其中 $F^{-1}(u)=\inf\{x\setminus F(x)\geqslant u\}$. 已知 EX_1 存在，得 $\left|E(\max\limits_{1\leqslant j\leqslant n}X_j)\right|\leqslant n\int_0^1\left|F^{-1}(u)\right|du=nE|X_1|$，即对 $n\geqslant 2$ 上面的积分存在. 类似 $E(\max\limits_{1\leqslant j\leqslant n}Y_j)=n\int_0^1G^{-1}(u)\,u^{n-1}du$ 存在，　$n\geqslant 2$.

充分性是显然成立的，下面证明定理的必然性方面. 因为对每个实数 t 及每个正整数 k 有

$$\left|F^{-1}(u\sum_{r=0}^{k}\frac{(iut)^r}{r!}\right|\leqslant\sum_{r=0}^{\infty}\left|F^{-1}(u)\frac{(ut)^r}{r!}\right|\leqslant e^{|t|}\left|F^{-1}(u)\right|$$

且$\int_0^1 e^{|t|}\left|F^{-1}(u)\right|du = e^{|t|}E\left|X\right| < \infty$，用 Lebesgue 控制收敛定理得

$$\sum_{r=0}^{\infty}\frac{(it)^r}{r!}\int_0^1 F^{-1}(u)\cdot u^r du = \int_0^1 F^{-1}(u)e^{iut}du$$

类似的有

$$\sum_{r=0}^{\infty}\frac{(it)^r}{r!}\int_0^1 G^{-1}(u)\cdot u^r du = \int_0^1 G^{-1}(u)e^{iut}du$$

由条件知对$n\geq 1$有 $\int_0^1 F^{-1}(u)u^{n-1}du = \int_0^1 G^{-1}(u)u^{n-1}du$，于是对各个实数$t$有

$$\int_0^1 F^{-1}(u)e^{iut}du = \int_0^1 G^{-1}(u)e^{iut}du$$

由Fourier 变换的唯一性得$F^{-1}(u) = G^{-1}(u)$, a.e.。另一方面因$F(x)$与$G(x)$是非降且右连续的函数，所以$F^{-1}(u)$与$G^{-1}(u)$是非降且右连续的，从而对每个$0\leq u\leq 1$有$F^{-1}(u) = G^{-1}(u)$，故$F(x)\equiv G(x)$。

推论1　在定理2的条件中，最大值的期望换成最小值的期望结论仍成立。

下面继续讨论指数分布的特征，即证明

定理3　设$\{X_i\}$是具有公共分布函数$F(x)$的iid·序列，则$F(x)$是指数分布的充要条件是对一切$n\geq 1$有

$$nE(\hat{M}_n^{(n)}) = EX_1 \qquad (15)$$

且EX_1存在为正数。

证　(10)及(14)表明必要性是显然的。

今若(15)成立且$0<EX_1<+\infty$，得一指数分布函数

$$G(x) = \begin{cases} 1-e^{-x/EX_1}, & x\geq 0 \\ 0, & x<0 \end{cases} \qquad (16)$$

设iid·序列$\{Y_i\}$具有公共分布函数$G(x)$，于是本定理的必要性表明对一切$n\geq 1$有

$$nE(\min_{1\leq i\leq n} Y_i) = EY_1$$

但(16)表明 $E\eta_1=E\xi_1$，从而结合(15)得对一切 $n\geq 1$ 有

$$E(\min_{1\leq j\leq n}\xi_j)=E(\min_{1\leq j\leq n}\eta_j)$$

用定理2的推论1即得 $F(x)\equiv G(x)$ 为一指数分布．

类似地可以证明，若将定理3中的条件(15)改成

$$(n+1)E(\hat{M}_n^{(n)})=E\xi_1,\quad n\geq 1$$

则可推得 $F(x)$ 是 $[0,\alpha]$ 上的均匀分布，此 α 是某个正数．

此外，定理3中的条件 $E\xi_1$ 为正数是可以略去不写的．事实上只要 $F(x)$ 不是在零处退化的分布，则由(15)及 $E\xi_1$ 存在可推知 $E\xi_1>0$．

最后，我们向读者介绍两篇70年代的短文献~~[8]~~及[7]，[9]给出的结果比本节定理2的结果更一般化，它表明了本节定理2中最大值的期望相等的条件可以换成第 k 个极值的期望相等的条件．~~短文[9]则给出了分布函数列的收敛性与相应的极值的期望之收敛性之间的关系~~．有余力的读者不妨试读一下这~~两~~篇短文．

第二章　iid 序列的最大值

§1　引言

设$\{\xi_i\}$是 iid 序列，具有公共分布函数 $F(x)$，我们知道 $P(\hat{M}_n^{(1)}\le x)=F^n(x)$，如果进一步讨论极限问题，上式取极限显然是没有多大意义的。但人们发现 $P(\frac{\hat{M}_n^{(1)}-a_n}{b_n}\le x)=F^n(a_n+b_n x)$ 的极限问题很值得研究。其中 a_n 及 $b_n>0$ 是数列。此即所谓规范化最大值的分布之弱收敛问题。如果有 $\lim P(\frac{\hat{M}_n^{(1)}-a_n}{b_n})\le x)=H(x)$ 存在，于是可得当 n 充分大时

$$P(\hat{M}_n^{(1)}\le x)=P(\frac{\hat{M}_n^{(1)}-a_n}{b_n}\le\frac{x-a_n}{b_n}) \text{ 接近于 } H(\frac{x-a_n}{b_n})$$

1943年 Гнеденко [4] 对规范化最大值的分布之弱收敛问题给出了全面而完整的解决，并大体上研究了此弱收敛与原始分布之间的密切关系。这一经典结果至今仍是一般极值之分布的弱收敛理论的基石。1971年 Haan [10] 进一步研究了 Гнеденко 遗留的一个抽象结果，使问题解决得更加完整和彻底。本章及下一章将着重讲述这个方面上的有关理论。

以下各定理中所述的分布列的收敛性，一般都指弱收敛，为简化叙述，常略而未提。

§2　几个引理

引理1　设 $F_n(x)$ 是一序列分布函数，C_n 与 $d_n>0$ 是实数列使 $\lim\limits_{n\uparrow\infty} F_n(C_n+d_n x)=G(x)$ ……对 $G(x)$ 的一切连续点成立。若数列 $\overset{*}{C}_n$ 与 $\overset{*}{d}_n>0$ 满足当 $n\uparrow\infty$ 有

$$\frac{c_n - c_n^*}{d_n} \longrightarrow 0 \quad 及 \quad \frac{d_n}{d_n^*} \longrightarrow 1 \tag{1}$$

则对 $G(x)$ 的一切连续点有

$$\lim_{n\to\infty} F_n(c_n^* + d_n^* x) = G(x) \tag{2}$$

证　选 x_1, x 及 x_2 ($x_1 < x < x_2$) 是 $G(x)$ 的连续点，由条件(1)知当 n 充分大有

$$x_1 < \frac{d_n^*}{d_n} x + \frac{c_n^* - c_n}{d_n} < x_2$$

因为 $c_n^* + d_n^* x = c_n + d_n\left(\frac{d_n^*}{d_n} x + \frac{c_n^* - c_n}{d_n}\right)$，所以当 n 充分大有

$c_n + d_n x_1 < c_n^* + d_n^* x < c_n + d_n x_2$，于是

$F_n(c_n + d_n x_1) \le F_n(c_n^* + d_n^* x) \le F_n(c_n + d_n x_2)$

由已知，对上式取极限得

$$G(x_1) \le \varliminf_{n\to\infty} F_n(c_n^* + d_n^* x) \le \varlimsup_{n\to\infty} F_n(c_n^* + d_n^* x) \le G(x_2)$$

因为 x 是 $G(x)$ 的连续点，取 x_1 及 x_2 从左右两方逼近 x 可得(2)成立

引理 2　设 $F_n(x)$ 是一序列分布函数，而 c_n, $d_n > 0$, f_n 及 $\tau_n > 0$ 是数列，满足

$$\lim_{n\to\infty} F_n(c_n + d_n x) = G(x) \quad 及 \quad \lim_{n\to\infty} F_n(f_n + \tau_n x) = T(x) \tag{3}$$

对极限函数的一切连续点 x 成立，此 $G(x)$ 与 $T(x)$ 是非退化的分布函数，则

$$\lim_{n\to\infty} \frac{\tau_n}{d_n} = B > 0, \qquad \lim_{n\to\infty} \frac{f_n - c_n}{d_n} = A \tag{4}$$

是有限的，且

$$T(x) = G(A + Bx) \tag{5}$$

证　分别选择 $G(x)$ 与 $T(x)$ 的两个连续点 x_1，x_2 及 y_1，y_2 使

$$0<G(x_1)\le G(x_2)<1 \quad 及\ T(y_1)<G(x_1),\ T(y_2)>G(x_2),$$

则当 n 充分大，由(3)可得

$$\rho_n+\tau_n y_1\le c_n+d_n x_1\le c_n+d_n x_2\le \rho_n+\tau_n y_2 \qquad (6)$$

而中间两项及外边两项取差可得 $d_n(x_2-x_1)\le \tau_n(y_2-y_1)$，即

$$\frac{d_n}{\tau_n}\le\frac{y_2-y_1}{x_2-x_1}$$

说明 d_n/τ_n 是上有介的．在证明开始若交换 $G(x)$ 与 $T(x)$，类似地可以得 τ_n/d_n 是上有介的．

另一方面，(6)的前一不等式即

$$\frac{\rho_n-c_n}{d_n}\le x_1-\frac{\tau_n}{d_n}y_1$$

说明 $\dfrac{\rho_n-c_n}{d_n}$ 是上有介的．(6)的后一不等式类似地可推出 $\dfrac{c_n-\rho_n}{d_n}$ 是上有介的．

于是可设存在子列 n_i，有

$$\lim_{i\uparrow\infty}\frac{\tau_{n_i}}{d_{n_i}}=B>0 \qquad 及\ \lim_{i\uparrow\infty}\frac{\rho_{n_i}-c_{n_i}}{d_{n_i}}=A$$

此 A 及 B 是有限数．

任给 $\varepsilon>0$，当 n_i 充分大时有 $(B-\varepsilon)d_{n_i}\le\tau_{n_i}\le d_{n_i}(B+\varepsilon)$，且 $c_{n_i}+d_{n_i}(A-\varepsilon)\le\rho_{n_i}\le c_{n_i}+D_{n_i}(A+\varepsilon)$．因此，对 $x\geqslant 0$ 有

$$F_{n_i}(c_{n_i}+Ad_{n_i}+Bd_{n_i}x-\varepsilon(x+1)d_{n_i})\le F_{n_i}(\rho_{n_i}+\tau_{n_i}x)\le F_{n_i}(c_{n_i}+d_{n_i}(A+Bx+\varepsilon(x+1)))$$

于是如果 x 和 ε 使 $A+Bx-\varepsilon(x+1)$ 及 $A+Bx+\varepsilon(x+1)$ 是 $G(x)$ 的连续点．由(3)得当 $n_i\uparrow\infty$，有

$$G(A+Bx-\varepsilon(x+1))\le\varliminf_{i\uparrow\infty}F_{n_i}(\rho_{n_i}+\tau_{n_i}x)\le\varlimsup_{i\uparrow\infty}F_{n_i}(\rho_{n_i}+\tau_{n_i}x)\le G(A+Bx+\varepsilon(x+1))$$