《概率论》补充材料
— 50个反例

复旦大学管理学院统计学系

2009年5月30日

1　事件之间的关系

(1) 从$A-B=C$推不出$A=B\cup C$

当$A-B=C$时，只能推出$A\subset B\cup C$. 事实上，令

$$A=\{1,2,3,4,5\},\quad B=\{1,3,5,7\},$$

于是

$$C=A-B=\{2,4\}.$$

而

$$B\cup C=\{1,2,3,4,5,7\},$$

从而$B\cup C\supset A$，但是$A\neq B\cup C$.

仅当$A\supset B$时，方能得出$A=B\cup C$.

(2) 从$A=B\cup C$推不出$A-B=C$

当$A=B\cup C$成立时，只能推出$A-B\subset C$. 当$B\subset A$，$C\subset A$　且　$B\cap C\neq\emptyset$时，可以得出$A-B=C$. 例如，令

$$A=\{1,2,3,4,5,6\},\quad B=\{1,2,3\},\quad C=\{2,4,5,6\}.$$

则有

$$B\cup C=\{1,2,3,4,5,6\}=A.$$

但

$$A-B=\{4,5,6\}.$$

从而$A-B\neq C$.

(3) $\bigcup_k A_k-\bigcup_k B_k\neq\bigcup_k(A_k-B_k)$

对于一般情况的A，B，C，有

$$\bigcup_k A_k-\bigcup_k B_k\subset\bigcup_k(A_k-B_k).$$

例如，令$k=2$，取

$$A_1=A_2=\{1,2,3,4,5,6\},\quad B_1=\{1,2\},\quad B_2=\{5,6\}.$$

则有

$$(A_1\cup A_2)-(B_1\cup B_2)=\{3,4\},\quad (A_1-B_1)\cup(A_2-B_2)=\{1,2,3,4,5,6\}.$$

从而$(A_1\cup A_2)-(B_1\cup B_2)\neq(A_1-B_1)\cup(A_2-B_2)$.

2 从概率关系推不出事件关系

若两个事件A，B之间有关系$A \subset B$，则其对应的概率关系如下：$P(A) \leq P(B)$，反之不然.

例如，设

$$P(A) = 0.3, \quad P(B) = 0.35, \quad P(A \cup B) = 0.35.$$

这时，$A \subset B$不成立. 事实上，由

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

及假设可得$P(A \cap B) = 0.3$.

于是，$A - (B \cap A)$的概率

$$P(A - B \cap A) = P(A) - P(B \cap A) = 0,$$

但这不意味着$A - B \cap A = \emptyset$.

我们改变一下上面所给的条件，就可以用来说明另一种情形. 设

$$P(A) = 0.3, \quad P(B) = 0.05, \quad P(A \cup B) = 0.35.$$

此时可得$P(A \cap B) = 0$，但这不能说明$A \cap B = \emptyset$，因而不能得出A，B互斥的结论.

通过这两个例子可见，不能由概率关系推出事件关系.

3 概率为零的事件未必是不可能事件

不可能事件的概率必为零. 那么，概率为零的事件是否为不可能事件？回答一般是否定的.

当考虑的对象为古典概率模型，概率按照古典概率定义时，概率为零的事件一定是不可能事件.

但是，当考虑的对象为几何概率模型，概率按几何概率定义时，概率为零的事件未必是一个不可能事件. 例如，设$\Omega = \{(x, y), 0 \leq x, y \leq 1\}$，$A = \{x = y, 0 \leq x, y \leq 1\}$，显然$P(A) = 0$. 但$A$是可能发生的. 另外对于连续性随机变量，它在一点处取值的概率为零，但它不是不可能发生.

4　非离散型又非连续型的分布函数

例　设
$$F(x)=\begin{cases}0, & x\le 0,\\ \frac{1+x}{2}, & 0<x\le 1,\\ 1, & x>1.\end{cases}$$
则$F(x)$是一个分布函数. 但是，$F(x)$显然不是离散型的，也非连续型的.

5　有限可加而非可列可加的概率测度

设Ω是$[0,1]$中所有有理数组成的集合，$\mathcal{F}_1$表示由形式为$[a,b],(a,b],[a,b),(a,b)$所组成的Ω的子集类，这里a,b都为有理数. 令$\mathcal{F}_2$表示由$\mathcal{F}_1$中所有不相交的集合的有限并组成的集合. 则$\mathcal{F}_2$是一个域. 我们在这个域上定义概率测度:

$$P(A)=b-a,\quad \text{如果}A\in\mathcal{F}_1,$$

$$P(B)=\sum_{i=1}^{n}P(A_i),\quad \text{如果 } B\in\mathcal{F}_2.$$
这里$B\in\mathcal{F}_2$表示$B=\sum_{i=1}^{n}A_i$，$A_i\in\mathcal{F}_1$.

考虑$\mathcal{F}_2$中两个不相交的集合B，B'，即

$$B=\sum_{i=1}^{n}A_i,\quad B'=\sum_{j=1}^{m}A'_j,$$

其中A_i，$A_j\in\mathcal{F}_1$，且A_i，A_j都互不相交. 则$B+B'=\sum_{k=1}^{m+n}C_k$，这里$C_k=A_i$或者$C_k=A'_j$. 接下来，

$$\begin{aligned}P(B+B')&=P(\sum_k C_k)=\sum_k P(C_k)=\sum_{i,j}(P(A_i)+P(A'_j))\\&=\sum_i P(A_i)+\sum_j P(A'_j)=P(B)+P(B').\end{aligned}$$

易知，P满足有限可加性.

对于每个单点集$\{r\}\in\mathcal{F}_2$，$P(\{r\})=0$. 由于集合Ω是可数集，即$\Omega=\sum_{i=1}^{\infty}\{r_i\}$，则
$$P(\Omega)=1\neq 0=\sum_{i=1}^{\infty}P(\{r_i\}),$$
即P非可列可加.

6 边际分布与联合分布可以不是同类型分布

我们知道，正态分布的边际分布仍为正态分布，多项分布的边际分布亦为多项分布. 那么是否联合分布与边际分布皆为同一种类的分布呢？一般回答是否定的，这种问题的例子很多.

例如，随机变量(ξ_1, ξ_2)有联合密度函数

$$f(x,y)=\begin{cases}\frac{1}{4}\left[1+xy(x^2+y^2)\right], & |x|\le 1, |y|\le 1,\\ 0, & \text{反之}.\end{cases}$$

则ξ_1，ξ_2的密度函数分别为$f_{\xi_1}(x)=1/2$，$|x|\le 1$，$f_{\xi_2}(y)=1/2$，$|y|\le 1$. 显然，三者不是同一类型的分布.

7 由边际分布无法求出联合分布

由两个随机变量ξ，η的联合分布$f(x,y)$可以很容易地计算出它们各自的边际分布$f_\xi(x)$和$f_\eta(y)$. 但是，若仅知道ξ, η各自的边际分布，却可能求不出它们的联合分布.

例如，把三个球放在三个盒中. 这时样本空间有27个点，令N表示的是3个球随机的放入三个盒中被装进球的盒子的个数，X_i表示第i个盒子中球的个数$(i=1,2,3)$. 对于每个样本点赋以概率$\frac{1}{27}=q$. 此时，形式上考虑(N, X_1)的联合分布由下表给出.

N \ X_1	0	1	2	3	N的分布
1	2q	0	0	q	3q=$\frac{1}{9}$
2	6q	6q	6q	0	18q=$\frac{2}{3}$
3	0	6q	0	0	6q=$\frac{2}{9}$
X_1的分布	8q	12q	6q	q	1

显然无法由X_1，N的边际分布得出(X_1, N)的联合分布. 造成这种情形的原因在于X_1与N不独立.

8 相同边际分布但是联合分布不同-1

若随机向量$(X_1,..,X_n)$的分布函数为$F(x_1,..,x_n)$，则边际分布$F_k(x_k), k=1,...,n$被唯一确定，但反之不然.

例 令$p=\{p_{ij},i,j=1,2,...\}$是一个二维离散分布. 选择两个点(x_1,y_1)和(x_2,y_2)，满足每个点上都有正的概率且$x_1\neq x_2$，$y_1\neq y_2$. 取ε使得$0<\varepsilon\leq p_{11}$，$0<\varepsilon\leq p_{22}$. 考虑$q=\{q_{ij},i,j=1,2,...\}$，定义如下：

$$q_{11}=p_{11}-\varepsilon,\quad q_{12}=p_{12}+\varepsilon,\quad q_{21}=p_{21}+\varepsilon,\quad q_{22}=p_{22}-\varepsilon,$$

对于其余的$i,\ j\neq 1,2$，令$q_{ij}=p_{ij}$. 易得q也是一个二维分布，且和p有着同样的边际分布，虽然$p\neq q$.

9 相同边际分布但是联合分布不同-2

例 假定F_1和F_2的密度函数分别为f_1和f_2. 考虑函数

$$f(x_1,x_2)=f_1(x_1)f_2(x_2)[1+\varepsilon(2F_1(x_1)-1)(2F_2(x_2)-1)],\quad (x_1,x_2)\in\mathbb{R}^2,$$

其中ε为任意实数，且满足$|\varepsilon|\leq 1$. 可以看出f是一个密度函数，且它的边际密度函数分别为f_1和f_2，与ε无关，故确定了边际分布也无法确定联合分布.

10 相同边际分布，但是联合分布不同-3

虽然，边际分布函数由联合分布函数唯一决定，但反之却不成立. 也就是说，不相同的分布函数却可以有相同的边际分布函数. 下面举出一例.

设有两个二元分布函数为$F(x,y)$及$G(x,y)$，分别有密度函数为：

$$f(x,y)=\begin{cases} x+y & \text{若}0\leq x\leq 1,\ 0\leq x\leq 1,\\ 0 & \text{其他,}\end{cases}$$

$$g(x,y)=\begin{cases} (0.5+x)(0.5+y) & \text{若}0\leq x\leq 1,\ 0\leq x\leq 1,\\ 0 & \text{其他.}\end{cases}$$

易知$F,\ G$不恒等. 然而，两对边际分布函数却相等，因为他们的两对密度函数相等. 事实上

$$\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y=\int_{-\infty}^{\infty}g(x,y)\mathrm{d}y=0.5+x,$$

$$\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x=\int_{-\infty}^{\infty}g(x,y)\mathrm{d}x=0.5+y.$$

11 二维概率密度函数连续，边际密度函数不一定连续

例 令

$$f(x,y)=(2\sqrt{2\pi})^{-1}|x|\exp(-|x|-\frac{1}{2}x^2y^2),\quad (x,y)\in\mathbb{R}^2. \tag{1}$$

容易验证f是一个概率密度函数. 对于第一个边际密度函数

$$f_1(x)=\begin{cases}0, & x=0,\\ \frac{1}{2}\exp(-|x|), & x\neq 0.\end{cases} \tag{2}$$

易知，虽然f连续，但f_1在$x=0$处不连续.

注意到函数f只有一点不连续. 现在我们根据f构造一个新的连续的密度函数，使它的边际密度函数有无穷多个不连续点.

令$\{r_k,\, k\geq 1\}$为一组已排序的有理数，令

$$g(x,y)=\sum_{n=1}^{\infty}2^{-n}f(x-r_n,y). \tag{3}$$

易知（1）中f在$\mathbb{R}^2$中有界，（3）式中右端的级数在$\mathbb{R}^2$中一致收敛. 另外，g是一个处处连续的概率密度函数，它的边际密度函数为

$$g_1(x)=\sum_{n=1}^{\infty}2^{-n}f_1(x-r_n). \tag{4}$$

同理易知（4）式右端的级数一致收敛，但是在有理数点r_1，r_2，... 上g_1都不连续. 虽然它在其他无理数点都是连续的.

12 数学期望不存在的离散型随机变量

在离散型随机变量的数学期望定义中（见《概率论基础》，p.172），要求级数$\sum_{i=1}^{\infty}x_kp_k$绝对收敛. 易知，若绝对收敛，则级数$\sum_{i=1}^{\infty}x_kp_k$收敛，反之不然.

例 设随机变量X取值为

$$x_k=(-1)^k\frac{2^k}{k},\ k=1,2,...$$

相应的概率为

$$p_k=\frac{1}{2^k},\ k=1,2,...$$

这是一个离散型的随机变量. 由于

$$\sum_{k=1}^{\infty}|x_k|p_k=\sum_{k=1}^{\infty}\frac{1}{k}=\infty,$$

从而EX不存在. 然而

$$\sum_{k=1}^{\infty}x_kp_k=\sum_{k=1}^{\infty}(-1)^k\frac{1}{k}=-\ln 2.$$

13　数学期望不存在的连续型随机变量

在数学期望的定义中，要求积分绝对收敛. 我们知道，若一个积分绝对收敛则该积分一定收敛，反之则不一定成立.

例　设随机变量X的密度函数为

$$f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2},\ x\in\mathbb{R}.$$

由于$f(x)\geq 0$，且

$$\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1.$$

故$f(x)$确实是一个密度函数. 但是，因为

$$\int_{-a}^{a}|x|\frac{1}{\pi}\cdot\frac{1}{1+x^2}\mathrm{d}x=\frac{1}{\pi}\int_{0}^{a}\frac{\mathrm{d}(1+x^2)}{1+x^2}=\frac{1}{\pi}\ln(1+a^2),$$

当$a\to\infty$时，$\frac{1}{\pi}\ln(1+a^2)\to\infty$. 故$EX$不存在.

14　数学期望存在但方差不存在的随机变量

密度函数为

$$f(x)=\frac{\Gamma(\frac{3}{2})}{\sqrt{\pi}\Gamma(\frac{1}{2})}\cdot\frac{1}{(1+x^2)^{3/2}}$$

的随机变量X，其数学期望为0，方差不存在.

15　ξ与η不独立但$E(\xi\eta)=E\xi\cdot E\eta$

设ξ，η为两个随机变量. 若ξ和η独立，且各自数学期望存在，则

$$E(\xi\eta)=E\xi\cdot E\eta.$$

反之则不然.

例 取$\Omega=[0,1]$，S为Ω的博雷尔集，P为通常的勒贝格测度. 考虑如下两个随机变量

$$\xi(x)=\sin 2\pi x,\quad \eta(x)=\cos 2\pi x.$$

不难得到

$$E(\xi\eta)=E\xi=E\eta=0.$$

取ε足够，使得

$$A_\xi=\{x:\ |\sin 2\pi x-1|<\varepsilon\},\quad A_\eta=\{x:\ |\cos 2\pi x-1|<\varepsilon\}$$

不相交，即$P(A_\xi A_\eta)=0$. 另一方面,

$$P(A_\xi)\neq 0\neq P(A_\eta).$$

故ξ，η不独立.

16 各阶矩存在也不足以确定分布律

从随机变量的分布函数，可以确定它的各阶矩，但反之不然. 实际上，存在着不同的分布函数，其各阶矩都是一样的.

例 设随机变量ξ，η分别具有如下密度函数$f_\xi(x)$和$f_\eta(x)$：

$$f_\xi(x)=\begin{cases} C\exp(-x^u\cos u\pi), & x>0,\\ 0, & x\leq 0,\end{cases}$$

$$f_\eta(x)=\begin{cases} C[1+\sin(x^u\cos u\pi)]\exp(-x^u\cos u\pi), & x>0,\\ 0, & x\leq 0,\end{cases}$$

其中$0<u<\frac{1}{2}$，$C=\frac{u(\cos u\pi)^{1/u}}{\Gamma(1/u)}$.

显然，$f_\xi(x)\neq f_\eta(x)$，故易证两随机变量分布函数不同. 但是它们却有相同的各阶矩：

$$\int_0^\infty x^n f_\xi(x)\mathrm{d}x=\frac{\Gamma(\frac{n+1}{u})}{\Gamma(\frac{1}{u})}(\cos u\pi)^{-n/u}=\int_0^\infty x^n f_\eta(x)\mathrm{d}x.$$

17 两两独立但不相互独立-1

所谓事件A_k，$k=1,2,...,n$ 两两独立，是指其中任意两个A_i，A_j之间都有关系式

$$P(A_iA_j)=P(A_i)P(A_j),\quad (i\neq j,\ i,j=1,2,...,n)$$

成立. 而相互独立的定义见《概率论基础》，p143.

我们知道，若事件A_k，$k=1,2,...,n$ 相互独立，则他们一定两两独立，但反之不然.

例 有四张卡片，各有数字112，121，222，211. 随机变量ξ_1，ξ_2，ξ_3分别表示随机取得的某张卡片上的第一、第二、第三位数字. 取四张卡片的概率相等. 由于

$$P(\xi_i=1)=0.5,\quad (i=1,2,3)$$

$$P((\xi_i=1)\cap(\xi_j=1))=0.25.\quad (i\neq j,\ i,j=1,2,3)$$

所以ξ_1,ξ_2,ξ_3两两独立. 但由于

$$P((\xi_1=1)\cap(\xi_2=1)\cap(\xi_3=1))=0,$$

$$P(\xi_1=1)\cdot P(\xi_2=1)\cdot P(\xi_3=1)=\frac{1}{8}\neq 0.$$

故ξ_1，ξ_2，ξ_3不相互独立.

18 两两独立但不相互独立-2

设三维随机向量(X,Y,Z)的联合密度函数为

$$f(x,y,z)=\begin{cases}\frac{1}{8\pi^3}(1-\sin x\sin y\sin z), & 0<x,y,z<2\pi,\\ 0, & \text{其他}.\end{cases}$$

从中可以求出X，Y，Z各自的边际分布，从而可以得出X，Y，Z两两独立但不相互独立.

19 两两独立不符合传递律

设三个事件A，B，C. 若A与B独立，且与独立，则有A与C独立，我们就说A，B，C的独立关系符合传递律.

两两独立不符合传递律. 考虑有两个孩子的家庭全体，假定生男孩与生女孩是等可能的. 因而样本空间

$$\Omega = \{(b,b),(b,g),(g,b),(g,g)\},$$

其中$b=$男孩，$g=$女孩. 每一对里的次序是指出生的次序，四点中每一点具有概率$\frac{1}{4}$. 现在随机的选择这样一个家庭，并考虑下面三个事件：

$$\begin{aligned} A &= \text{“第一个孩子是男孩”}, \\ B &= \text{“两个孩子不同性别”}, \\ C &= \text{“第一个孩子是女孩”}. \end{aligned}$$

则有

$$AB = \{(b,g)\}, \quad BC = \{(g,b)\}, \quad AC = \emptyset.$$

经简单计算可得

$$P(AB) = P(A)P(B) = \frac{1}{4}, \quad P(BC) = P(B)P(C) = \frac{1}{4}.$$

即A与B独立，C与B独立. 但是,

$$P(AC) = 0 \neq \frac{1}{4} = P(A)P(C).$$

因此事件A与C不独立. 这样得出，A，B，C的独立关系不符合传递律.

20 随机变量不独立而其函数独立

众所周知，正态分布有一个特性：任何$n(n>1)$维的正态分布的随机变量，可以由坐标轴的旋转变为一组n个独立的正态分布的随机变量. 这说明了，对$n=2$，即使ξ，η不独立，但当(ξ,η)服从正态分布时，随机向量(ξ',η'):

$$\xi' = \xi\cos\alpha + \eta\sin\alpha, \quad \eta' = -\xi\sin\alpha + \eta\cos\alpha$$

服从正态分布：

$$f(x',y') = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\Big\{-\frac{1}{2(1-r^2)}(Ax'^2 - 2Bx'y' + Cy'^2)\Big\}.$$

只要适当的选择α：

$$\tan 2\alpha = \frac{2r\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2},$$

则$B=0$. 此时ξ'，η'独立.

21 P(ABC)=P(A)P(B)P(C)不一定推出A,B,C相互独立

令$\Omega = \{1,2,3,4,5,6,7,8\}$，且在每一点的概率都为1/8. 考虑事件$B_1 = \{1,2,3,4\}$, $B_2 = B_3 = \{1,5,6,7\}$. 则$P(B_1) = P(B_2) = P(B_3) = 1/2$，$B_1B_2B_3 = \{1\}$，$P(B_1B_2B_3) = \frac{1}{8} = P(B_1)P(B_2)P(B_3)$. 然而$B_2$与$B_3$之间显然不独立.

22 X^2与Y^2独立但X与Y不独立-1

如果随机变量X与Y独立，那么X^2与Y^2一定独立. 反之不然.

考虑一个二维随机向量(X,Y)，其概率分布为

$$p_{i,j} := P(X = i, Y = j), \quad i,j = -1,0,1,$$

其中$p_{1,1} = p_{-1,1} = 1/32$，$p_{-1,-1} = p_{1,-1} = p_{1,0} = p_{0,1} = 3/32$，$p_{-1,0} = p_{0,-1} = 5/32$，$p_{0,0} = 8/32$. 容易验证$X^2, Y^2$独立，但$X, Y$不独立.

23 X^2与Y^2独立但X与Y不独立-2

设随机向量(X,Y)的联合密度函数为

$$f(x,y) = \begin{cases} \frac{1}{4}(1+xy), & |x| < 1，|y| < 1, \\ 0, & \text{其他}. \end{cases}$$

由于$f_X(x) = f_Y(y) = \frac{1}{2}$，所以$f(x,y) \neq f_X(x)f_Y(y)$. 可见$X$与$Y$不相互独立. 但是

$$f_{X^2}(x) = (4x)^{-1/2}, \quad f_{Y^2}(y) = (4y)^{-1/2},$$

而且

$$f_{X^2,Y^2}(x,y) = (4\sqrt{xy})^{-1}.$$

可见对一切x，y有

$$f_{X^2}(x) \cdot f_{Y^2}(y) = f_{(X^2,Y^2)}(x,y).$$

故X^2与Y^2独立.

24　不相关也不独立的随机变量-1

随机变量相互独立则它们必然不相关，反之则不然.

例　设随机变量X与Y的联合分布为

X \ Y	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

容易验证X与Y不相关($\mathrm{cov}(X,Y)=0$)，但也不独立.

25　不相关也不独立的随机变量-2

设随机变量X的概率密度函数是偶函数，且$EX^2<\infty$. 则X与$|X|$不相关，而且不独立.

26　随机变量ξ_1，ξ_2独立但其函数η_1，η_2未必独立

设随机变量ξ_1与ξ_2相互独立，η_1与η_2为它们的函数. 由ξ_1与ξ_2独立性未必能得出η_1与η_2的独立性. 下面分两部分谈这一问题.

(1) ξ_1，ξ_2独立同分布，其函数η_1，η_2独立

例　当ξ_1，ξ_2独立且具有相同的分布$N(0,1)$时，令

$$\eta_1=\xi_1+\xi_2,\quad \eta_2=\xi_1-\xi_2.$$

易得η_1，η_2均服从分布$N(0,2)$，且有

$$f_{\eta_1,\eta_2}(y_1,y_2)=f_{\eta_1}(y_1)\cdot f_{\eta_2}(y_2),$$

其中$f_{\eta_1,\eta_2}(y_1,y_2)$，$f_{\eta_1}(y_1),f_{\eta_2}(y_2)$分别为$\eta_1$与$\eta_2$的联合分布密度函数及边际分布密度. 从而$\eta_1$与$\eta_2$独立.

(2)ξ_1，ξ_2独立同分布，其函数η_1，η_2不独立

若ξ_1与ξ_2独立同分布:

$$\begin{pmatrix} 1, & 2, & 3, & 4, & 5, & 6 \\ \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6}, & \frac{1}{6} \end{pmatrix},$$

即取为两颗骰子的点数. 令

$$\eta_1 = \xi_1 + \xi_2, \quad \eta_2 = \xi_1 - \xi_2.$$

此时η_1与η_2或者同为偶数或者同为奇数. 所以η_1与η_2不相互独立.

但是，由于

$$E(\eta_1\eta_2) = E(\xi_1^2) - E(\xi_2^2) = 0,$$

故

$$\operatorname{cov}(\eta_1, \eta_2) = E(\eta_1\eta_2) - E(\eta_1)E(\eta_2) = 0,$$

从而η_1与η_2不相关.

27 分布相同但是随机变量不相同

考虑定义在同一个概率空间上的两个随机变量X，Y. 假设它们是相等的，即$P\{\omega : X(\omega) \neq Y(\omega)\} = 0$. 因此

$$F_X(x) = P\{\omega : X(\omega) \leq x\} = P\{\omega : Y(\omega) \leq x\} = F_Y(x), \quad \forall x \in \mathbb{R}.$$

因此X和Y分布相同, 这时我们记作$X \stackrel{d}{=} Y$或者$X \stackrel{L}{=} Y$.

但是反之不然. 设随机变量X服从分布$N(0,1)$. 令$Y = -X$，则有对称性易得$F_X = F_Y$. 但是

$$P\{\omega : X(\omega) = Y(\omega)\} = P\{\omega : X(\omega) = -X(\omega)\} = P\{\omega : X(\omega) = 0\} = 0,$$

即X，Y不相等.

28 $X \stackrel{d}{=} Y$不一定能推出$XZ \stackrel{d}{=} YZ$

令X，Y为两个随机变量，且$X \stackrel{d}{=} Y$. 设函数$g(x)$，$x \in \mathbb{R}$是一个$\mathcal{B}^1$可测函数. 则有$g(X)$，$g(Y)$也是随机变量且$g(X) \stackrel{d}{=} g(Y)$. 这使得我们容易产生下面这个推测：若$X$，$Y$，$Z$定义在同一个概率空间内，则

$$X \stackrel{d}{=} Y \Rightarrow XZ \stackrel{d}{=} YZ, \quad \text{对任意随机变量} Z.$$

容易证明这是错的. 令X为对称的随机变量，$Y = -X$. 则$X \stackrel{d}{=} Y$. 再令$Z = Y$，则易知$XZ \stackrel{d}{=} YZ$不可能成立.

29　随机变量ξ，ζ有函数关系且相互独立

考虑这样三个随机变量ξ，η和ζ，ξ与η具有相同的分布：

$$\begin{pmatrix} 0, & 1 \\ p, & q \end{pmatrix}, \quad p+q=1, p>0, q>0.$$

令ζ

$$\zeta = \begin{cases} 0, & \xi+\eta = \text{偶数}, \\ 1, & \xi+\eta = \text{奇数}. \end{cases}$$

并假定ξ与η相互独立.

由于ξ与η相互独立，所以可得ξ与η的联合概率分布为：

ξ \ η	0	1
0	$(1-p)^2$	$p(1-p)$
1	$(1-p)p$	p^2

根据假设，ζ的概率分布为

$$\begin{pmatrix} 0, & 1 \\ (1-p)^2+p^2, & 2p(1-p) \end{pmatrix}.$$

而ζ与ξ的联合概率分布为

ξ \ ζ	0	1
0	$(1-p)^2$	$p(1-p)$
1	p^2	$p(1-p)$

假定ζ与ξ独立，即p应满足

$$\begin{gathered} (1-p)(2p^2-2p+1) = (1-p)^2, \\ (1-p)[2p(1-p)] = p(1-p), \\ p(2p^2-2p+1) = p^2, \\ p \cdot 2p(1-p) = p(1-p). \end{gathered}$$

解得$p=\frac{1}{2}$. 于是当$p=\frac{1}{2}$时，ζ与ξ相互独立（同样可证得ζ与η相互独立，但ξ，η，ζ却不相互独立）.

这个例子指明了，虽然ζ与ξ之间存在着函数关系，但它们可以相互独立.

30 ξ与ξ^2不相关

设ξ各以$\frac{1}{4}$的概率取值±1，±2. 由于$E\xi = 0$，且$E(\xi^3) = 0$. 所以$\mathrm{cov}(\xi, \xi^2) = 0$，即$\xi$与$\xi^2$不相关. 但是$\xi$与$\xi^2$之间有着非线性关系.

可见，相关系数并不是相依性的一般度量，仅仅指出两个随机变量间的线性相依性.

31 一阶矩存在的一个必要但不充分条件

设随机变量X的分布函数为F. 易知若EX存在，则有$\lim_{x\to\infty} x(1-F(x)) = 0$. 但如果$\lim_{x\to\infty} x(1-F(x)) = 0$存在，不一定能得到数学期望$EX$存在.

例 假定分布函数为

$$F(x) = \begin{cases} 0, & x \leq 1 \\ 1 - 1/(kx), & e^{k-1} < x \leq e^k, k = 1, 2, ... \end{cases}$$

直接可得$\lim_{x\to\infty} x(1-F(x)) = 0$，$\int_0^\infty (1-F(x))\mathrm{d}x = \infty$，故$EX = \int_0^\infty (1-F(x))\mathrm{d}x$不存在.

32 在有限区间内特征函数的值不足以唯一决定分布函数

从唯一性定理（《概率论基础》，p229）知，分布函数由特征函数唯一决定. 如果特征函数的定义域为一有限区间，则唯一性定理不成立.

例 令

$$\varphi_\xi(t) = \begin{cases} 1 - |t|, & \text{当}|t| \leq 1, \\ 0, & \text{当}|t| > 1. \end{cases}$$

因为$\varphi_\xi(t)$在$\mathbb{R}_1$上可积，故对应的分布密度为

$$f_\xi(x) = \frac{1-\cos x}{\pi x^2}.$$

另外，考虑一离散型分布$f_\eta(k)$:

$$P(\eta = 0) = \frac{1}{2},$$
$$P(\eta = (2k-1)\pi) = \frac{2}{(2k-1)^2\pi^2}, \quad k = 0, \pm1, ...$$

其特征函数为

$$\varphi_\eta = \frac{1}{2} + \frac{4}{\pi^2}\sum_{k=1}^{\infty}\frac{\cos(2k-1)\pi t}{(2k-1)^2}.$$

当$|t| \le 1$时，有$\varphi_\xi(t) = \varphi_\eta(t)$. 事实上，将$g(t) = |t|$在区间$|t| \le 1$上展开成傅立叶级数:

$$g(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos n\pi t,$$

其中

$$a_0 = 1, \quad a_n = \frac{2(\cos n\pi - 1)}{n^2\pi^2},$$

因而$a_{2k} = 0, (k \ne 0)$，$a_{2k-1} = -\frac{4}{(2k-1)^2\pi^2}$，于是

$$g(t) = |t| = \frac{1}{2} - \frac{4}{\pi^2}\sum_{k=1}^{\infty}\frac{\cos(2k-1)\pi t}{(2k-1)^2}.$$

将其带入$\varphi_\xi(t)$，并将结果于$\varphi_\eta(t)$比较，即得

$$\varphi_\xi(t) = \varphi_\eta(t), \quad |t| \le 1.$$

当$t > 1$时，显然有

$$\varphi_\xi(t) \ne \varphi_\eta(t), \quad |t| > 1.$$

由此可知，$\varphi_\xi(t)$与$\varphi_\eta(t)$是两个不同分布的特征函数，故由唯一性定理知$\varphi_\xi(t)$与$\varphi_\eta(t)$是两个不同的特征函数.

所以，一特征函数在有限区间内的值不足以决定此特征函数，从而也不足以唯一决定分布函数.

33 正态随机变量的联合分布非正态-1

正态分布具有许多好的性质，其中之一是：二维正态分布的边缘分布仍是正态分布. 反之，两边缘分布都是正态分布，其联合分布未必是正态分布.

例 取两个独立的随机变量ξ_1，ξ_2，两者都服从分布$N(0,1)$. 考虑如下二维分布：

$$(X_1, X_2) = \begin{cases} (\xi_1, |\xi_2|), & \xi_1 \ge 0 \\ (\xi_1, -|\xi_2|), & \xi_1 < 0. \end{cases}$$

显然(X_1, X_2)不服从正态分布，但是边际分布X_1，X_2都服从正态分布.

34 正态随机变量的联合分布非正态-2

例 二维分布函数$F(x,y)$的密度函数如下

$$f(x,y)=\frac{1}{2\pi}\exp\big(-\frac{1}{2}(x^2+y^2)\big)\cdot(1+\sin x\sin y),\ -\infty<x,y<\infty.$$

它不是正态分布，但其边际密度函数分别为

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y=\frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2}),\quad -\infty<x<\infty,$$

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\exp(-\frac{y^2}{2}),\quad -\infty<y<\infty.$$

即$X\sim N(0,1)$，$Y\sim N(0,1)$.

35 若$(X_1,...,X_n)$服从正态分布，则$X_1,...,X_n$的任一线性组合都服从正态分布，但当只是$(X_1,...,X_n)$的部分线性组合服从正态分布时，$(X_1,...,X_n)$不一定服从正态分布

定义密度函数如下：

$$f_\varepsilon(x_1,...,x_n)=(2\pi)^{-n/2}\prod_{k=1}^{n}\varphi_0(x_k)\big[1+\varepsilon(x_1^2-x_2^2)\prod_{k=1}^{n}x_kI_{(-1,1)}(x_k)\exp(\frac{1}{2}\sum_{j=1}^{n}x_j^2)\big],$$

其中$(x_1,..,x_n)\in\mathbb{R}^n$，$\varphi_0=\exp(-\frac{1}{2}x^2),I_{(-1,1)}$是区间$(-1,1)$上的示性函数. 选择合适的常数$\varepsilon$使得

$$\Big|\varepsilon(x_1^2-x_2^2)\prod_{k=1}^{n}x_kI_{(-1,1)}(x_k)\exp(\frac{1}{2}\sum_{j=1}^{n}x_j^2)\Big|\leq 1.$$

可以验证，f_ε是一个n维密度函数，且它显然不是一个正态密度函数.

接下来，我们求$(X_1,...,X_n)$的边际分布. 我们先求出f_ε的特征函数ϕ，如下

$$\phi(t_1,...,t_n)=\prod_{k=1}^{n}\varphi_0(t_k)+\varepsilon\Big[\big(\psi(t_1)\tilde{\psi}(t_2)-\tilde{\psi}(t_1)\psi(t_2)\big)\prod_{k=3}^{n}\psi(t_k)\Big],$$

其中

$$\psi(t)=\begin{cases}(2i/t^2)(\sin t-t\cos t), & t\neq 0,\\ 0, & t=0,\end{cases}$$

$$\tilde{\psi}(t)=\begin{cases}(2i/t)+(6/t^2)\psi(t), & t\neq 0,\\ 0, & t=0.\end{cases}$$

由上式我们可以得到以下结论：

1. $X_1, ..., X_n$每个都服从分布$N(0,1)$.

2. $\forall k < n$，向量$(X_{i_1}, ..., X_{i_k})$服从正态分布.

3. 令$U = X_1 \pm X_2$，V为$X_3, ..., X_n$的任一线性组合，则$U + V$服从正态分布.

4. 设$a_1, ..., a_n$为非零的实数，且$|a_1| \neq |a_2|$. 则$\sum_{k=1}^n a_k X_k$非正态.

5. $E\big(\prod_{k=1}^n X_k\big) = 0$.

36　依分布收敛但不依概率收敛-1

随机变量序列$\{\xi_n\}$，若依概率收敛于ξ，则一定依分布收敛. 反之不然.

例　设$\Omega = \{\omega_1, \omega_2\}$，$P(\omega_1) = P(\omega_2) = \frac{1}{2}$. 定义随机变量$\xi$：$\xi(\omega_1) = -1$，$\xi(\omega_2) = 1$. 则$\xi$的分布为：

$$\begin{pmatrix} -1, & 1 \\ \frac{1}{2}, & \frac{1}{2} \end{pmatrix} \qquad (*).$$

令$\xi_n(\omega) = -\xi(\omega)$. 显然，$\xi_n(\omega)$的分布也是(*). 因此，

$$\xi_n(\omega) \xrightarrow{L} \xi(\omega).$$

但是，对任意的$0 < \varepsilon < 2$,

$$P\{|\xi_n(\omega) - \xi(\omega)| > \varepsilon\} = P(\Omega) = 1,$$

即$\xi_n(\omega) \overset{P}{\nrightarrow} \xi(\omega)$.

37　依分布收敛但不依概率收敛-2

例　设$\{\xi_n, n \geq 1\}$为独立同分布非退化随机变量序列，且$F_n(x) \equiv F(x)$，则

$$F_n(x) \xrightarrow{W} F(x).$$

由于独立性，(ξ_n, ξ_m)的联合密度函数

$$F_{\xi_n,\xi_m}(x,y) = F_{\xi_n}(x)F_{\xi_m}(y) = F(x)F(y).$$

因而

$$\begin{aligned} P(|\xi_n - \xi_m| \leq \varepsilon) &= \int\int_{|x-y|\leq\varepsilon} \mathrm{d}F_{\xi_n,\xi_m}(x,y) \\ &= \int_{-\infty}^{\infty} [F(y+\varepsilon) - F(y-\varepsilon)]\mathrm{d}F(x). \end{aligned}$$

要使$P(|\xi_n-\xi_m|\le\varepsilon)\to 1$，必须对一切$y$有

$$F(y+\varepsilon)-F(y-\varepsilon)\to 1.$$

但是，它不成立. 这说明了由$\xi_n\xrightarrow{L}\xi$推不出$\xi_n\xrightarrow{P}\xi$.

如果将$\{\xi_n\}$局限为独立随机变量序列，且有$\xi_n\xrightarrow{L}C$，C为常数，则$\xi_n\xrightarrow{P}C$.

38 依概率收敛但不几乎处处收敛

随机变量序列$\{\xi_n,\,n\ge 1\}$，若几乎处处收敛于ξ，则一定依概率收敛. 反之不然.

例 取$\Omega=(0,1]$，F为$(0,1]$中博雷尔点集全体组成的σ代数全体，P为勒贝格测度. 令

$$\eta_{k_i}(\omega)=\begin{cases}1, & \omega\in(\frac{i-1}{k},\frac{i}{k}],\\ 0, & \omega\notin(\frac{i-1}{k},\frac{i}{k}].\end{cases}\quad(i=1,2,...,k,\ k=1,2,...)$$

定义

$$\xi_n(\omega)=\eta_{k_i}(\omega),\quad n=i+\frac{k(k-1)}{2}.$$

于是$\{\xi_n,\,n\ge 1\}$是一个随机变量序列.

(1) $\xi_n\xrightarrow{P}\xi$：对于任意$\varepsilon>0$，由于

$$P\{|\eta_{k_i}(\omega)|\ge\varepsilon\}\le\frac{1}{k},$$

故

$$P\{|\xi_n(\omega)|\ge\varepsilon\}\to 0,\quad(n\to\infty).$$

即$\xi_n\xrightarrow{P}\xi$.

(2) $\xi_n\overset{a.s.}{\nrightarrow}\xi$：由于对任意一个固定的$\omega\in\Omega$，任一正整数$k$，恰有一个$i$，使$\eta_{k_i}(\omega)=1$. 而对其余的$i$有$\eta_{k_i}(\omega)=0$. 由此知，$\{\xi_n(\omega)\}$对每一个$\omega\in\Omega$都不收敛.

39 依概率收敛但不r阶矩收敛-1

令$\{X_n,n\ge 1\}$是一列随机变量序列，满足

$$P(X_n=e^n)=1/n,\quad P(X_n=0)=1-1/n,\quad n\ge 1.$$

则，$\forall \varepsilon > 0$有

$$P(|X_n| < \varepsilon) = P(X_n = 0) = 1 - 1/n \to 1, \quad n \to \infty,$$

因此$X_n \xrightarrow{P} 0$，当$n \to \infty$. 然而，对$r > 0$，

$$E(X_n^r) = e^{rn}\frac{1}{n} \to \infty, \quad n \to \infty.$$

因此$X_n \stackrel{r}{\nrightarrow} 0$.

40 依概率收敛但不r阶矩收敛-2

定义随机变量序列$\{Y_n, n \geq 2\}$和$\{Z_n, n \geq 1\}$:

$$P(Y_n = 1) = 1/\log n = 1 - P(Y_n = 0),$$

$$P(Z_n = 0) = 1 - n^{-\alpha}, \quad P(Z_n = \pm n) = 1/(2n^{\alpha}), \quad 0 < \alpha \leq 2.$$

则当$n \to \infty$时，$\forall r > 0$，有$Y_n \xrightarrow{P} 0$，$Y_n \stackrel{r}{\nrightarrow} 0$，$Z_n \xrightarrow{P} 0$，$Z_n \stackrel{r}{\nrightarrow} 0$.

41 r阶矩收敛但不几乎处处收敛

令$\Omega = [0,1]$，$\mathcal{F} = \mathcal{B}_{[01]}$，$P$为勒贝格测度. 对每个自然数$n$，有且仅有一对整数$(m,k)$满足$m \geq 0$和$0 \leq k \leq 2^m - 1$，使得$n = 2^m + k$. 定义事件列

$$A_n = [k2^{-m}, (k+1)2^{-m}).$$

令随机变量序列$X_n = X_n(\omega) = 1_{A_n}(\omega), n \geq 1$. 则$E(|X_n|) = E(X_n) = 2^{-m} \to 0$，当$n \to \infty$. 因此$(r = 2)$

$$X_n \xrightarrow{r} 0, \quad n \to \infty.$$

另外，易证$X_n \stackrel{a.s.}{\nrightarrow} 0$，当$n \to \infty$.

42 几乎处处收敛也无法推出r阶矩收敛

定义随机变量序列$\{X_n, n \geq 1\}$如下:

$$P(X_n = 0) = 1 - 1/n^{\alpha}, \quad P(X_n = n) = P(X_n = -n) = 1/(2n)^{\alpha}, \quad \alpha > 0.$$

由$E(|X_n|^{1/2}) = 1/n^{\alpha-1/2}$，得到$\sum_{n=1}^{\infty} E(|X_n|^{1/2}) < \infty$，若$\alpha > 3/2$. 由Markov不等式，知$P(|X_n| > \varepsilon) \leq \varepsilon^{-1/2}E(|X_n|^{1/2})$，因此$\sum_{n=1}^{\infty} P(|X_n| > \varepsilon) < \infty, \forall\varepsilon > 0$. 再利用Borel-Cantelli引理，可推出$X_n \stackrel{a.s.}{\nrightarrow} 0$，当$n \to \infty$.

然而，$E(|X_n|^2) = 1/n^{\alpha-2}$，因此，对于$\alpha \leq 2$，$X_n \stackrel{2}{\nrightarrow} 0$.

所以，若$\alpha \in [3/2, 2]$，则$X_n \stackrel{a.s.}{\to} 0$，但$X_n \stackrel{2}{\nrightarrow} 0$.

43 $X_n \stackrel{L}{\to} X$和$Y_n \stackrel{L}{\to} Y$不一定推出$X_n + Y_n \stackrel{L}{\to} X + Y$

当$\{X_n, n \geq 1\}$和$\{Y_n, n \geq 1\}$独立时，$X_n \stackrel{L}{\to} X$，$Y_n \stackrel{L}{\to} Y$能够推出$X_n + Y_n \stackrel{L}{\to} X + Y$. 但有些情况，我们推不出$X_n + Y_n \stackrel{L}{\to} X + Y$.

例 令$\{X_n, n \geq 1\}$为独立同分布随机变量序列，且$P(X_n = 1) = P(X_n = 0) = 1/2$. 令$Y_n = 1 - X_n$，则$X_n \stackrel{L}{\to} X$和$Y_n \stackrel{L}{\to} Y$成立，其中$P(X = 1) = P(X = 0) = P(Y = 1) = P(Y = 0) = 1/2$，且$X$和$Y$独立. 由于$X_n + Y_n = 1$，而$P(X + Y = 1) = 2P(X + Y = 0) = 2P(X + Y = 2) = 1/2$，故不能得到$X_n + Y_n \stackrel{L}{\to} X + Y$.

44 对于任意函数g，$X_n \stackrel{P}{\to} X$不能推出$g(X_n) \stackrel{P}{\to} g(X)$

随机变量X，X_n，$n \geq 1$满足$X_n \stackrel{P}{\to} X$，当$n \to \infty$. 函数$g(x)$，$x \in \mathbb{R}$是一个连续函数，则$g(X_n) \stackrel{P}{\to} g(X), n \to \infty$.

若$g(x)$不连续，则上式不一定成立. 例如，定义函数

$$g(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0. \end{cases}$$

令$P(X_n = 1) = n^{-1}$，$P(X_n = n^{-1}) = 1 - n^{-1}$，则$X_n \stackrel{P}{\to} X, X = 0, a.s.$. 但有$g(X_n) = 1$和$g(X) = 0$，因此$g(X_n) \stackrel{P}{\nrightarrow} g(X)$.

45 随机变量序列收敛，其期望不一定收敛

例 定义随机变量序列$\{X_n, n \geq 1\}$如下：

$$P(X_n = -n - 4) = 1/(n + 4), \quad P(X_n = -1) = 1 - 4/(n + 4),$$

$$P(X_n = n + 4) = 3/(n + 4).$$

$\forall \varepsilon > 0$，我们有

$$P(|X_n - (-1)| > \varepsilon) = 4/(n+4).$$

因此，$X_n \xrightarrow{P} -1$，当$n \to \infty$. 另一方面

$$EX_n = 1 + 4/(n+1), \quad \lim_{n\to\infty} EX_n = 1.$$

因此，依概率收敛不能保证期望收敛.

46 分布函数列的极限未必是分布函数

设$\{D(x+n)\}$为一分布函数列，其中

$$D(x) = \begin{cases} 1, & x > 0, \\ 0, & x \le 0. \end{cases}$$

这个序列的极限函数就不是一个分布函数. 另外，分布函数序列$\{D(x-n)\}$的极限函数也不是一个分布函数.

47 分布函数列未必弱收敛于分布函数

从上面知，分布函数列的极限函数未必是分布函数. 若将收敛性改为弱收敛，仍未必能得出其极限函数是一个分布函数. 《概率论基础》270页已经给出一个例子，这里给出另外一个例子.

定义分布函数列$\{F_n\}$:

$$F_n(x) = \begin{cases} 0, & x \le 1 - \frac{1}{n}, \\ 1, & x > 1 - \frac{1}{n}. \end{cases} \quad (n = 1, 2, ...)$$

及

$$F(x) = \begin{cases} 0, & x \le 1, \\ 1, & x > 1. \end{cases}$$

则当$x > 1$时，$F_n(x) = F(x) = 1$；当$x < 0$时，$F_n(x) = F(x) = 0$；当$0 < x < 1$时，必存在正整数m，使得

$$1 - \frac{1}{m} \le x < 1 - \frac{1}{m+1}.$$

于是，当$n \ge m+1$时，$F_n(x) = 0$. 此时仍有

$$F_n(x) \to F(x).$$

但$x = 1$是$F(x)$的不连续点. 故除了$x = 1$之外的一切x均有

$$F_n(x) \to F(x).$$

48 弱收敛的极限函数不唯一

考虑正态分布函数序列

$$F_n(x) = \frac{n}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-\frac{n^2y^2}{2})\mathrm{d}y, \quad n = 1, 2, ...$$

另外，定义

$$F(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 1, \end{cases} \qquad G(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{2}, & x = 0, \\ 1 & x > 0. \end{cases}$$

则$F_n(x) \xrightarrow{W} F(x)$，$F_n(x) \xrightarrow{W} G(x)$，但是$F(x) \neq G(x)$.

49 一列离散的独立随机变量序列满足弱大数律不满足强大数律

令$\{X_n, n \geq 2\}$为一列独立的随机变量序列，满足

$$P(X_n = \pm n) = 1/(2n\log n), \quad P(X_n = 0) = 1 - 1/(n\log n), \quad n = 2, 3, ...$$

令$A_n = \{|X_n| \geq n\}$，则

$$P(A_n) = 1/(n\log n) \Rightarrow \sum_{n=2}^{\infty} P(A_n) = \infty.$$

再由Borel-Cantelli引理，得

$$P(A_n, \mathrm{i.o.}) = P(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} A_n) = 1 \Rightarrow P(\lim_{n\to\infty} S_n/n \neq 0) = 1,$$

即序列$\{X_n, n \geq 2\}$不满足强大数律，这里$S_n = X_1 + \cdots + X_n$.

下面我们证明$\{X_n, n \geq 2\}$满足弱大数律. 显然$DX_k = k/\log k$. 由于函数$x/\log x$在$x = e$处取得局部最小值，且$\sum_{k=3}^{n} k/\log k \leq \int_3^{n+1}(x/\log x)\mathrm{d}x$，则我们易得

$$\frac{1}{n^2}\sum_{k=2}^{n} DX_k \leq \Big(\frac{2}{\log 2} + \int_3^{n+1}(x\log x)\mathrm{d}x\Big) \leq \frac{2}{n^2\log n} + \frac{(n-2)(n+1)}{n^2\log n} \to 0,\ n \to \infty,$$

即$\{X_n\}$满足Markov条件，故$\{X_n, n \geq 2\}$服从弱大数律.

50 一列绝对连续的独立随机变量序列满足弱大数律但不满足强大数律

令$\{X_n, n\geq 1\}$为一列独立的随机变量序列，密度函数为

$$f_n(x) = (\sqrt{2}\sigma_n)^{-1}\exp(-\sqrt{2}|x|/\sigma_n), \quad x\in\mathbb{R}.$$

计算可得$DX_n = \sigma_n^2$. 定义$\sigma_n^2 = 2n^2/(\log n)^2$，$n\geq 2$.

首先，我们证明$\{X_n\}$不满足强大数律. 事实上，令$A_n = \{|X_n|\geq n\}$，则

$$P(A_n) = 2(\sqrt{2}\sigma_n)^{-1}\int_n^{\infty}\exp(-\sqrt{2}x/\sigma_n)\mathrm{d}x = \exp\big(-\frac{1}{2}\sqrt{2}(\log n)^2/n\big).$$

由于$(\log n)^2/n\to 0$，故$\sum_{n=2}^{\infty}P(A_n) = \infty$，即$\{X_n,\, n\geq 1\}$不满足强大数律.

接下来，我们证明$\{X_n\}$满足弱大数律. 首先

$$P\big(\frac{1}{n}|X_k|>\varepsilon\big) = \exp(-n\varepsilon\frac{\log k}{k}), \quad k=2,3,...,n.$$

任意取$c>0, \varepsilon>0$两数，定义新的随机变量$\tilde{X}_k$

$$\tilde{X}_k = \begin{cases} X_k, & |X_k|\leq c, \\ c, & |X_k|>c. \end{cases}$$

显然，

$$\sum_{k=2}^{n}P\big(\frac{1}{n}|\tilde{X}_k|>\varepsilon\big)\to 0, n\to\infty.$$

与离散的情形类似，可以得到

$$\frac{1}{n^2}\sum_{k=2}^{n}D\tilde{X}_k\to 0, \quad n\to\infty.$$

利用Feller定理，$\{X_n,\, n\geq 1\}$服从弱大数律.

参考文献

[1] Stoyanov, J.M. (1997). Counterexamples in probability, second edition. *John Wiley & Sons.*

[2] 李贤平，概率论基础（第二版），高等教育出版社，1997.

[3] 张朝金，概率论中的反例，陕西人民出版社，1984.