大样本理论

司继春

[1] 上海对外经贸大学

2017年11月

概览

收敛的概念

收敛的定义

若$\{a_n, n = 1, 2, ...\}$为实数序列，如果对于任意的$\epsilon > 0$，存在$n_0 = n_0(\epsilon)$使得：

$$|a_n - a| < \epsilon, \forall n > n_0$$

那么我们称数列$\{a_n\}$的极限为a，或者$\{a_n\}$收敛到（converges to）a，记为

$$\lim_{n \to \infty} a_n = a$$

或者

$$a_n \to a \text{ } as \text{ } n \to \infty$$

收敛的概念

有界的定义

若$\{a_n, n = 1, 2, ...\}$为实数序列，如果存在常数$b < \infty$，使得

$$|a_n| < b$$

那么我们称数列$\{a_n\}$为有界的（bounded），否则称之为无界的（unbounded）。

小o符号

小o符号的定义

对于两个序列$\{a_n\},\{b_n\}$，如果随着$n\to\infty$，有：

$$\frac{a_n}{b_n}\to 0$$

那么我们记为$a_n=o(b_n)$。特别的，如果令$b_n=1$，那么$a_n=o(1)$等价为$a_n\to 0$。

小o符号

小o实例

如果$a_n = \frac{1}{n^2}, b_n = \frac{1}{n}$，那么:

$$\frac{a_n}{b_n} = \frac{\frac{1}{n^2}}{\frac{1}{n}} = \frac{1}{n} \to 0$$

因而$\frac{1}{n^2} = o\left(\frac{1}{n}\right)$，即$\frac{1}{n^2}$以更快的速度收敛到0。如果两个序列$a_n \to 0, b_n \to 0$，且$a_n = o\left(b_n\right)$，那么我们称$a_n$为比$b_n$高阶的无穷小量。

小o符号

小o符号经常用来对一个复杂的式子进行化简，通过将无穷小量舍掉从而减少了计算量。比如：

小o的应用

假设有两个数列，$a_n = \frac{1}{n} + \frac{6}{n^2} - \frac{8}{n^3}$而另外一个序列：$b_n = \frac{1}{n}$如果定义$R_n = \frac{6}{n^2} - \frac{8}{n^3}$，显然$R_n = o\left(\frac{1}{n}\right)$，因而$a_n = b_n + o\left(\frac{1}{n}\right)$，即：

$$\frac{a_n}{b_n} = \frac{b_n + o\left(\frac{1}{n}\right)}{b_n} \to 1$$

因而尽管两个序列a_n和b_n并不相等，但是当$n \to \infty$时，两者误差趋向于0，因而我们可以舍去无穷小量R_n，使用更简单的序列b_n去逼近a_n。

小o符号的应用：泰勒展开

当$x \to a$时，$(x-a) = o(1)$，同时我们有$(x-a)^{k+1} = o\left((x-a)^k\right)$，即当$x \to a$时，$(x-a)$的高阶幂是低阶幂的无穷小量。对于一个单变量实值函数$f(x)$且k阶可微，那么有：

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + o\left(|x-a|^k\right)$$

因而对于一个难以计算的函数f，我们经常使用其前k阶泰勒多项式对其进行逼近。

泰勒展开

泰勒展开

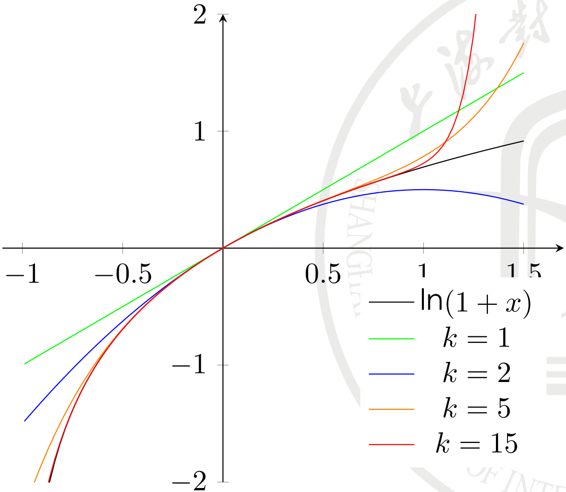
函数$f(x)=\ln(1+x)$在$x=0$处的泰勒展开为:

$$f(x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots$$

因而当x充分靠近0时，我们可以使用前k阶泰勒展开对其进行逼近。特别的，如果令

$$k=1,\quad \ln(1+x)=x+o(x)\approx x$$

泰勒展开



多元函数的泰勒展开

更一般的，如果$f:\mathbb{R}^n\to\mathbb{R}$为多元实值函数，那么其泰勒级数为：

$$f(x)=f(a)+\frac{\partial f}{\partial x'}(a)(x-a)+\frac{1}{2!}(x-a)'\frac{\partial^2 f}{\partial x\partial x'}(a)(x-a)\\+o\left(\|x-a\|^2\right)$$

其中x和a为$n\times 1$向量。

多元函数的泰勒展开

多元函数泰勒展开示例

令$f(x) = e^{x_1}\ln(1+x_2)$，其中$x = (x_1, x_2)'$。那么：

$$\frac{\partial f}{\partial x} = \begin{bmatrix} e^{x_1}\ln(1+x_2) \\ \frac{e^{x_1}}{1+x_2} \end{bmatrix}, \frac{\partial^2 f}{\partial x \partial x'} = \begin{bmatrix} e^{x_1}\ln(1+x_2) & \frac{e^{x_1}}{1+x_2} \\ \frac{e^{x_1}}{1+x_2} & -\frac{e^{x_1}}{(1+x_2)^2} \end{bmatrix}$$

因而其在$a = (0,0)'$处的二阶泰勒展开：

$$\begin{aligned} p_2(x) &= [0,1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{1}{2}[x_1, x_2]\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= x_2 + \frac{1}{2}\left(2x_1x_2 - x_2^2\right) \end{aligned}$$

小o符号

小o的性质

1. 若$a_n = o(b_n), b_n = o(c_n)$，那么$a_n = o(c_n)$
2. 对于任意的常数$c \neq 0$，及$a_n = o(b_n)$，有$ca_n = o(b_n)$
3. 对于任意的数列$c_n \neq 0$，及$a_n = o(b_n)$，有$c_n a_n = o(c_n b_n)$
4. 如果$d_n = o(b_n), e_n = o(c_n)$，那么$d_n e_n = o(b_n c_n)$
5. 如果$a_n, b_n > 0, c_n, d_n > 0$，$a_n = o(b_n), c_n = o(d_n)$，那么$a_n + c_n = o(b_n + d_n)$。

大O符号

大O符号的定义

对于两个序列$\{a_n\},\{b_n\}$，如果随着$n \to \infty$，$\left|\frac{a_n}{b_n}\right|$是有界的，即存在一个M使得：

$$\left|\frac{a_n}{b_n}\right| < M$$

那么我们记为$a_n = O(b_n)$。特别的，如果令$b_n = 1$，那么$a_n = O(1)$等价为a_n是有界的。

同阶的定义

对于两个序列$\{a_n\},\{b_n\}$，如果$a_n = O(b_n)$，且同时$b_n = O(a_n)$，那么我们称两个序列是同阶的，简记为$a_n \asymp b_n$。

大O符号

大O符号示例

对于序列$a_n = \frac{1}{n} + \frac{b}{n\sqrt{n}} + \frac{c}{n^2} + \frac{d}{n^2\sqrt{n}}$，同时定义$R_n = \frac{b}{n\sqrt{n}} + \frac{c}{n^2} + \frac{d}{n^2\sqrt{n}}$那么：

1. $a_n \sim \frac{1}{n}$
2. 若$b = 0$，$R_n = O\left(\frac{1}{n^2}\right)$
3. 若$b = 0$，$R_n \asymp \frac{1}{n^2}$
4. 若$b \neq 0$，$R_n \sim \frac{b}{n\sqrt{n}}$
5. 若$b = c = 0$，$R_n = o\left(\frac{1}{n^2}\right)$

大O符号

大O符号的性质

1. 若$a_n = O(b_n), b_n = O(c_n)$，那么$a_n = O(c_n)$
2. 对于任意的常数$c \neq 0$，及$a_n = O(b_n)$，有$ca_n = O(b_n)$
3. 对于任意的数列$c_n \neq 0$，及$a_n = O(b_n)$，有$c_n a_n = O(c_n b_n)$
4. 如果$d_n = O(b_n), e_n = O(c_n)$，那么$d_n e_n = O(b_n c_n)$
5. 如果$a_n = o(b_n)$，$c_n = O(b_n)$，那么$a_n c_n = o(b_n)$
6. 如果$a_n = o(b_n)$，$c_n = O(b_n)$，那么$a_n + c_n = O(b_n)$

几乎必然收敛

几乎必然收敛

如果概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的一系列随机变量$\{X_n\}$满足:

$$\mathscr{P}\left(\left\{\lim_{n\to\infty} X_n(\omega) = X\right\}\right) = 1$$

那么我们称X_n几乎必然收敛于X，记为$X_n \xrightarrow{\text{a.s.}} X$，或者$X_n \to X\ a.s.$。

依概率收敛

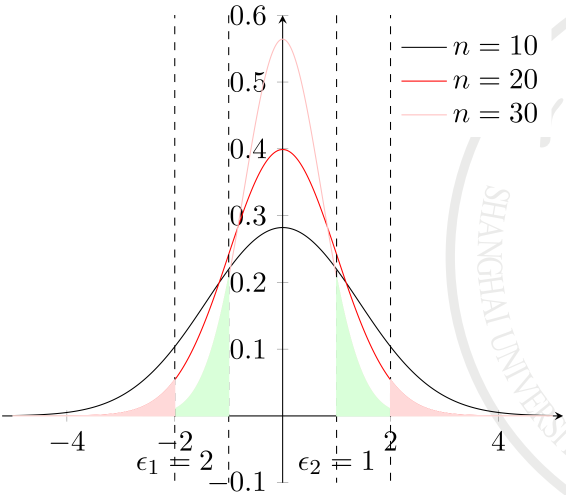
依概率收敛

如果对于任意的$\epsilon > 0$，概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的一系列随机变量序列$\{X_n\}$满足：

$$\mathscr{P}\left(|X_n - X| > \epsilon\right) \to 0$$

那么我们称X_n依概率收敛于X，记为$X_n \xrightarrow{p} X$，或$\mathrm{plim} X_n = X$。

依概率收敛



o_p符号

o_p符号

$\{X_n\}$与$\{Y_n\}$为定义在概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的两个随机变量序列，如果

$$\frac{X_n}{Y_n} \xrightarrow{p} 0$$

那么我们记为$X_n = o_p(Y_n)$。特别的，当$Y_n = 1$时，即$X_n = o_p(1)$，等价于$X_n \xrightarrow{p} 0$。

O_p符号

O_p符号

$\{X_n\}$与$\{Y_n\}$为定义在概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的两个随机变量序列，如果对于任意的$\epsilon > 0$，存在一个$C_\epsilon$使得:

$$\sup_n \mathscr{P}\left(|X_n| \geq C_\epsilon |Y_n|\right) < \epsilon$$

那么我们记$X_n = O_p(Y_n)$。特别的，当$Y_n = 1$时，我们称X_n依概率有界（bounded in probability）。

O_p符号

O_p符号是对O符号在概率意义上的扩展。注意如果$\text{Var}\left(X_n\right) < M$，即随机变量序列$\{X_n\}$的方差有界，那么对于任意的$\epsilon > 0$，取$C_\epsilon = \sqrt{M/\epsilon + 1}$，那么：

$$\mathscr{P}\left(|X_n| \geq C_\epsilon\right) \leq \frac{\mathbb{E}\left(X_n^2\right)}{C_\epsilon^2} = \frac{\mathbb{E}\left(X_n^2\right)}{M/\epsilon + 1} < \epsilon$$

因而$X_n = O_p\left(1\right)$。因而如果X_n的方差有界，那么必然是$O_p\left(1\right)$。

O_p和o_p符号的性质

O_p和o_p的性质

如果$X_n = o_p(1), Y_n = o_p(1), Z_n = O_p(1), W_n = O_p(1)$，那么：

1. $X_n + Y_n = o_p(1)$
2. $X_n + Z_n = O_p(1)$
3. $Z_n + W_n = O_p(1)$
4. $X_n Y_n = o_p(1)$
5. $X_n Z_n = o_p(1)$
6. $Z_n W_n = O_p(1)$

均方收敛

均方收敛的定义

如果概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的一系列随机变量序列$\{X_n\}$随着$n \to \infty$满足：

$$\mathbb{E}\left(|X_n - X|^2\right) \to 0$$

那么我们称X_n均方收敛于X，记为$X_n \xrightarrow{L^2} X$。

均方收敛

均方收敛与依概率收敛

如果随机变量序列$X_n \xrightarrow{L^2} X$，那么$X_n \xrightarrow{p} X$。

Proof.

据切比雪夫不等式，对于任意的$\epsilon > 0$，有：

$$\mathscr{P}\left(|X_n - X| > \epsilon\right) \leq \frac{\mathbb{E}\left(|X_n - X|^2\right)}{\epsilon^2} \to 0$$

$\square$

依分布收敛

依分布收敛的定义

令F_n,F为分布函数，如果对于每一个$F(x)$连续的点x，有：

$$\lim_{n\to\infty} F_n(x) = F(x)$$

那么我们称$F_n(x)$弱收敛于$F(x)$，记为$F_n \xrightarrow{w} F$。

如果一系列随机变量$\{X_n\}$的分布函数$F_{X_n}(x) \xrightarrow{w} F_X$，我们称$X_n$依分布收敛于$X$，记为$X_n \xrightarrow{D} X$或者：$X_n \overset{a}{\sim} F$，其中$a$代表渐进的（asymptotically），即$X_n$渐进服从分布函数为$F$的分布。

依分布收敛

依分布收敛与O_p的关系

1. 如果$X_n \neq O_p(1)$，那么分布函数极限$\lim_{n\to\infty} F_n(x) = F(x)$不是一个分布函数。
2. 如果X_n依分布收敛，那么$X_n = O_p(1)$。

因而当我们讨论依分布收敛时，一定要保证我们讨论的$X_n = O_p(1)$。

大数定律

大数定律（Law of Large Numbers, LLN）讨论样本均值的极限，即在何种条件下，以下结论：

$$\frac{S_N - \mathbb{E}\left(S_N\right)}{N} = \bar{x} - \mathbb{E}\left(\bar{x}\right) \xrightarrow{p} 0$$

成立，其中：

$$S_N = \sum_{i=1}^{N} x_i$$

大数定律

实际上，$S_N - \mathbb{E}(S_N) = \sum_{i=1}^{N} [x_i - \mathbb{E}(x_i)]$，因而：

$$
\begin{aligned}
\mathbb{E}[S_N - \mathbb{E}(S_N)]^2 &= \mathbb{E}\left(\sum_{i=1}^{N} [x_i - \mathbb{E}(x_i)]\right)^2 \\
&= \mathbb{E}\left(\sum_{i=1}^{N} [x_i - \mathbb{E}(x_i)]^2\right) \\
&\quad + \mathbb{E}\left(2 \sum_{1\le j<i\le N} [x_i - \mathbb{E}(x_i)][x_j - \mathbb{E}(x_j)]\right) \\
&= \sum_{i=1}^{N} \mathrm{Var}(x_i) + 2 \sum_{1\le j<i\le N} \mathrm{Cov}(x_i, x_j)
\end{aligned}
$$

大数定律

$$\mathbb{E}\left[S_N - \mathbb{E}\left(S_N\right)\right]^2 = \sum_{i=1}^{N} \mathsf{Var}\left(x_i\right) + 2 \sum_{1\le j<i\le N} \mathsf{Cov}\left(x_i, x_j\right)$$

如果我们假设$\mathsf{Var}\left(x_i\right) < M$，那么：

1. $\sum_{i=1}^{N} \mathsf{Var}\left(x_i\right) < NM = O\left(N\right)$
2. 而根据Cauthy-Schwartz不等式，$\mathsf{Cov}\left(x_i, x_j\right) \le \sqrt{\mathsf{Var}\left(x_i\right)\mathsf{Var}\left(x_j\right)} \le M$，而上式中有$\frac{N(N-1)}{2}$个协方差，所以

$$\sum_{1\le j<i\le N}^{n} \mathsf{Cov}\left(x_i, x_j\right) = O\left(N^2\right)$$

大数定律

$$\mathbb{E}\left[S_N-\mathbb{E}\left(S_N\right)\right]^2=O\left(N\right)+O\left(N^2\right)$$

从而：

$$\mathbb{E}\left[\frac{S_N-\mathbb{E}\left(S_N\right)}{N}\right]^2=\frac{1}{N^2}O\left(N\right)+\frac{1}{N^2}O\left(N^2\right)=o\left(1\right)+O\left(1\right)$$

根据均方收敛定义，如果希望$\frac{S_N}{N}-\frac{\mathbb{E}(S_N)}{N}\xrightarrow{L^2}0$，必须

$$\mathbb{E}\left[\frac{S_N-\mathbb{E}\left(S_N\right)}{N}\right]^2=o\left(1\right)$$

大数定律

$$\mathbb{E}\left[\frac{S_N-\mathbb{E}\left(S_N\right)}{N}\right]^2=o\left(1\right)+O\left(1\right)$$

其中的$O\left(1\right)$来源于$\frac{N(N-1)}{2}$个协方差，如果$\mathrm{Cov}\left(x_i,x_j\right)=0$，那么自然有：

$$\mathbb{E}\left[\frac{S_N-\mathbb{E}\left(S_N\right)}{N}\right]^2=o\left(1\right)$$

从而

$$\frac{S_N}{N}-\frac{\mathbb{E}\left(S_N\right)}{N}\xrightarrow{L^2}0\Rightarrow\frac{S_N}{N}-\frac{\mathbb{E}\left(S_N\right)}{N}\xrightarrow{p}0$$

大数定律

大数定律1

如果概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的一个随机变量序列$\{x_i\}$两两不相关，且存在一个M使得对于所有的$i = 1, 2, \ldots$，都有$\mathrm{Var}\left(x_i\right) < M$，那么：

$$\bar{x} - \mathbb{E}\left(\bar{x}\right) = \frac{S_N - \mathbb{E}\left(S_N\right)}{N} \xrightarrow{L^2} 0$$

从而

$$\bar{x} - \mathbb{E}\left(\bar{x}\right) = \frac{S_N - \mathbb{E}\left(S_N\right)}{N} \xrightarrow{p} 0$$

如果额外假设$\{x_i\}$是同分布的，那么$\mathbb{E}\left(\bar{x}\right) = \mathbb{E}\left(x_i\right) = \mu$，从而：

$$\bar{x} \xrightarrow{p} \mu$$

大数定律

以下的定理放松了二阶矩有限的假定以及独立的假定，保留了独立同分布的假定：

大数定律2

令$\{x_i\}$为概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的两两独立且同分布的随机变量序列，若

$$\mathbb{E}\left|x_i\right| < \infty$$

那么

$$S_N/N = \bar{x} \xrightarrow{p} \mu$$

其中$\mu = \mathbb{E}\left(x_i\right)$。

大数定律

以下的定理则同时放宽了同分布的假定以及二阶矩的假定。

大数定律3

令$\{x_i\}$为概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上的相互独立的随机变量序列，如果存在一个常数$p \in [1,2]$，随着$N \to \infty$，使得：

$$\frac{1}{N^p} \sum_{i=1}^{N} \mathbb{E}\left|x_i\right|^p \to 0$$

那么$S_N/N = \bar{x} \xrightarrow{p} \mathbb{E}\left(\bar{x}\right)$

大数定律

大数定律的应用

如果令$\{x_i\}$为一系列i.i.d的随机变量，且$x_i \sim Ber\left(p\right)$，那么$\mathbb{E}\left(x_i\right) = p, \mathsf{Var}\left(x_i\right) = p\left(1-p\right) < \infty$，定义：

$$\hat{p} = \frac{\sum_{i=1}^{N} x_i}{N}$$

即成功的比例，那么根据上例，可以得到

$$\hat{p} \xrightarrow{p} p$$

中心极限定理

中心极限定理（Central Limit Theorem, CLT）讨论样本均值的极限分布，即在大样本条件下，样本均值的分布情况，通常中心极限定理可以得到如下结论：

$$\sqrt{N}\,(\bar{x}-\mu)=\frac{S_N-\mu}{\sqrt{N}} \overset{a}{\sim} N\left(0,\mathsf{Var}\left(x_i\right)\right)$$

其中：

$$S_N=\sum_{i=1}^{N} x_i$$

即样本均值的极限分布为正态分布。

中心极限定理

前面我们提到，如果希望讨论$\sqrt{N}\left(\bar{x}-\mu\right)$的极限分布，需要保证$\sqrt{N}\left(\bar{x}-\mu\right)=O_p\left(1\right)$，实际上：如果假设$\{x_i\}$之间两两不相关，那么：

$$
\begin{aligned}
\mathsf{Var}\left(S_N\right) &= \mathbb{E}\left[S_N-\mathbb{E}\left(S_N\right)\right]^2 \\
&= \mathbb{E}\left[\sum_{i=1}^{n}\left(X_i-\mathbb{E}\left(X_i\right)\right)\right]^2 \\
&= \sum_{i=1}^{n}\mathsf{Var}\left(X_i\right) \\
&= O\left(N\right)
\end{aligned}
$$

从而$\mathsf{Var}\left(\sqrt{N}\frac{S_N}{N}\right)=O\left(1\right)$，或者其方差有界，因而$\sqrt{N}\frac{S_N}{N}=\sqrt{N}\bar{x}=O_p\left(1\right)$。

中心极限定理

中心极限定理（标量）

令$\{x_i\}$为概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上$i.i.d$的随机变量序列，且$\mathbb{E}\left(x_i\right)=\mu, \mathsf{Var}\left(x_i\right)=\sigma^2$，那么：

$$\sqrt{N}\left(\bar{x}_N-\mu\right) \xrightarrow{D} N\left(0, \sigma^2\right)$$

或：

$$\sqrt{N}\left(\frac{\bar{x}_N-\mu}{\sigma}\right) \xrightarrow{D} N(0,1)$$

中心极限定理

中心极限定理示例

如果$\{x_i\}$为$i.i.d$的随机变量，且$x_i \sim Ber(p)$，令$\hat{p}_N$如前定义，那么：

$$\sqrt{N}\left(\hat{p}_N - p\right) \xrightarrow{D} N\left(0, p\left(1-p\right)\right)$$

如果$\{x_i\}$为$i.i.d$的随机变量，且$x_i \sim N(0,1)$，那么可知$\mathbb{E}\left(x_i^2\right) = 1, \mathbb{E}\left(x_i^4\right) = 3$，因而：

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} x_i^2 - 1\right) \xrightarrow{D} N\left(0, 2\right)$$

中心极限定理

中心极限定理（向量）

令$\{x_i\}$为概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上$i.i.d$的随机向量序列，且$\mathbb{E}(x_i) = \mu, \mathsf{Var}(x_i) = \Sigma$，那么：

$$\sqrt{N}(\bar{x}_N - \mu) \xrightarrow{D} N(0, \Sigma)$$

或：

$$\sqrt{N}\Sigma^{-\frac{1}{2}}(\bar{x}_N - \mu) \xrightarrow{D} N(0, I)$$

随机变量连续函数的收敛

中心极限定理

令$\{x_i\}$为概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上$i.i.d$的随机变量序列，且$\mathbb{E}(x_i) = \mu, \mathrm{Var}(x_i) = \sigma^2$，那么：

$$\sqrt{N}(\bar{x}_N - \mu) \xrightarrow{D} N(0, \sigma^2)$$

或：

$$\sqrt{N}\left(\frac{\bar{x}_N - \mu}{\sigma}\right) \xrightarrow{D} N(0, 1)$$

随机变量连续函数的收敛

样本相关系数的极限

对于二维随机向量(x_i, y_i)，令(x_i, y_i)为i.i.d的样本，那么在可积性条件下，

$$\begin{cases} \frac{1}{N}\sum_{i=1}^{N} x_i \xrightarrow{p} \mathbb{E}(x_i) & \frac{1}{N}\sum_{i=1}^{N} y_i \xrightarrow{p} \mathbb{E}(y_i) \\ \frac{1}{N}\sum_{i=1}^{N} x_i y_i \xrightarrow{p} \mathbb{E}(x_i y_i) & \\ \frac{1}{N}\sum_{i=1}^{N} x_i^2 \xrightarrow{p} \mathbb{E}(x_i^2) & \frac{1}{N}\sum_{i=1}^{N} y_i^2 \xrightarrow{p} \mathbb{E}(y_i^2) \end{cases}$$

从而

$$\frac{\frac{1}{N}\sum_{i=1}^{N} x_i y_i - \frac{1}{N}\sum_{i=1}^{N} x_i \frac{1}{N}\sum_{i=1}^{N} y_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N} y_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} y_i\right)^2}} \xrightarrow{p} \mathsf{Corr}(x_i, y_i)$$

Slutsky定理

Slutsky定理

如果随机变量$X_n \xrightarrow{D} X$，$R_n = o_p(1)$，那么

$$X_n + R_n \xrightarrow{D} X$$

同时如果$Y_n \xrightarrow{p} a \neq 0$，那么

$$\frac{X_n}{Y_n} \xrightarrow{D} \frac{X}{a}$$

如果$Y_n \xrightarrow{p} a$，那么

$$X_n Y_n \xrightarrow{D} aX$$

Slutsky定理

t统计量的大样本分布

之前曾讨论过，如果$x_i \sim N\left(\mu, \sigma^2\right)$ $i.i.d$，那么

$$\frac{\sqrt{N}\left(\bar{x}-\mu\right)}{\sqrt{\frac{\sum_{i=1}^{N}\left(x_i-\bar{x}\right)^2}{N-1}}} \sim t\left(N-1\right)$$

现在我们不假设x_i服从正态分布，而是假设其独立同分布且具有有限的二阶矩，那么我们有$\bar{x} \xrightarrow{p} \mathbb{E}\left(x_i\right), \frac{1}{N}\sum_{i=1}^{n} x_i^2 \xrightarrow{p} \mathbb{E}\left(x_i^2\right)$

Slutsky定理

t统计量的大样本分布

因而：

$$
\begin{aligned}
\frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N-1} &= \frac{\sum_{i=1}^{N}x_i^2}{N-1}-\frac{N}{N-1}\bar{x}^2 \\
&= \frac{\sum_{i=1}^{N}x_i^2-N\bar{x}^2}{N-1} \\
&\xrightarrow{p} \mathbb{E}\left(x_i^2\right)-\left[\mathbb{E}\left(x_i\right)\right]^2=\mathsf{Var}\left(x_i\right)
\end{aligned}
$$

进而：

$$
\frac{\sqrt{N}\left(\bar{x}-\mu\right)}{\sqrt{\frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N-1}}} \xrightarrow{p} \frac{\sqrt{N}\left(\bar{x}-\mu\right)}{\sqrt{\mathsf{Var}\left(x_i\right)}}=\sqrt{N}\left(\frac{\bar{x}-\mu}{\sqrt{\mathsf{Var}\left(x_i\right)}}\right)\xrightarrow{D} N\left(0,1\right)
$$

delta方法

delta方法示例

令$\{x_i\}$为概率空间$(\Omega, \mathscr{F}, \mathscr{P})$上$i.i.d$的随机变量序列，且$\mathbb{E}\left(x_i\right) = \mu, \mathsf{Var}\left(x_i\right) = \sigma^2$，那么根据中心极限定理：

$$\sqrt{N}\left(\bar{x} - \mu\right) \xrightarrow{D} N\left(0, \sigma^2\right)$$

delta方法

delta方法示例

如果我们关心$Y=\exp\left(\bar{X}_n\right)$的分布，那么可以对其进行泰勒展开：

$$\begin{aligned}
\sqrt{N}Y-\sqrt{N}\exp\left(\mu\right) =& \sqrt{N}\exp\left(\bar{X}_n\right)-\sqrt{N}\exp\left(\mu\right)\\
=& \sqrt{N}\exp\left(\mu\right)\left(\bar{X}_n-\mu\right)+\frac{1}{2}\sqrt{N}\exp\left(\mu\right)\left(\bar{X}_n-\mu\right)^2+\cdots\\
=& \sqrt{N}\exp\left(\mu\right)\left(\bar{X}_n-\mu\right)+\frac{1}{2}\exp\left(\mu\right)O_p\left(1\right)o_p\left(1\right)\\
=& \sqrt{N}\exp\left(\mu\right)\left(\bar{X}_n-\mu\right)+o_p\left(1\right)\\
\xrightarrow{D}& \sqrt{N}\exp\left(\mu\right)\left(\bar{X}_n-\mu\right)\\
\overset{a}{\sim}& N\left(0,\exp\left(2\mu\right)\sigma^2\right)
\end{aligned}$$

因而$Y\overset{a}{\sim}N\left(e^{\mu},\frac{\exp(2\mu)\sigma^2}{N}\right)$。