



Fluctuations of Maxima

Modelling Extremal Events: Chap 3

庄源

2024 年 5 月 7 日

南开大学精算学系

1. 章节框架介绍与示例
2. 3.1 Limit Probabilities for Maxima
3. 3.2 Weak Convergence of Maxima Under Affine Transformations
4. 3.3 Maximum Domains of Attraction and Norming Constants
5. 3.3.1 Fréchet 分布的最大吸引域

6. 3.3.2 Weibull 分布的最大吸引域

7. 3.3.3 Gumbel 分布的最大吸引域

8. 3.4 The Generalised Extreme Value Distribution and the Generalised Pareto Distribution

9. 3.5 Almost Sure Behaviour of Maxima

章节框架介绍与示例

章节结构

- 主题：研究最大值的极限分布
- 3.1：累积分布函数 F 在满足什么条件时，能使这个极限分布存在？退一步讲，能否使中心化和标准化后的 M_n 依分布收敛于某个分布？
- 3.2：如果真的能收敛于一个简单的分布（不是 0 或 1 那样的常数），那么这个分布是什么？
- 3.3：我怎样知道，对于给定的 F ，它会收敛于极值分布中的哪一个？与中心极限定理相似，仿射变换的参数序列 c_n 和 d_n 到底是什么？
- 3.4：能否把三种极值分布归为一类？这些分布又跟广义帕累托分布 (GPD) 有什么关系？
- 3.5：有没有比依分布收敛更强的结论？（如几乎处处收敛）

定义、命题、定理、引理或推论

公式、定理编号均与原书相同。

图片的源代码均可在[个人主页](#)上找到。

评论

对应原书 Remark 部分。

例题

本章例题较多。

3.1 Limit Probabilities for Maxima

最大值随机变量的累积分布函数

- X, X_1, X_2, \dots 为 iid、非退化的随机变量, 累积分布函数为 F ;
- 定义样本最大值 M_n :

$$M_1 = X_1, \quad M_n = \max(X_1, \dots, X_n), \quad n \geq 2$$

- 对于 iid 的 X_1, X_2, \dots , 有:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x)$$

$$= P(X_1 \leq x) \times \dots \times P(X_n \leq x) = F^n(x), \quad x \in \mathbb{R}, \quad n \in \mathbb{N}$$

最大值随机变量的收敛性

- 定义 x_F 为 F 的右端点, 即 $x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$;
- 若 $x < x_F$:

$$P(M_n \leq x) = F^n(x) \rightarrow 0, \quad n \rightarrow \infty$$

- 若 $x_F \leq \infty, x \geq x_F$:

$$P(M_n \leq x) = F^n(x) = 1$$

- M_n 是 n 的单调不减函数, 则可以印证一个直觉性的结论:

$$M_n \xrightarrow{\text{a.s.}} x_F, \quad n \rightarrow \infty \tag{3.1}$$

寻找仿射变换下的收敛

- 是否存在两列数 $c_n > 0$, $d_n \in \mathbb{R}$, 使得

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H, \quad n \rightarrow \infty \quad (3.2)$$

- 下式中, $u_n = c_n x + d_n$:

$$\begin{aligned} H(x) &= P(c_n^{-1} (M_n - d_n) \leq x) \\ &= P(M_n \leq c_n x + d_n) = P(M_n \leq u_n) \end{aligned} \quad (3.3)$$

- F 满足什么性质时, $P(M_n \leq u_n)$ 能够有收敛性, 并不要收敛于一个简单的数?

Proposition 3.1.1 (Poisson Approximation)

对于给定 $\tau \in [0, \infty]$ 和一系列实数 (u_n) , 下面两式等价:

$$n\bar{F}(u_n) \rightarrow \tau, \quad (3.4)$$

$$P(M_n \leq u_n) \rightarrow e^{-\tau}. \quad (3.5)$$

泊松近似 (cont'd)

Remark: 泊松近似的来源

在 $0 \leq \tau < \infty$ 时, 令

$$B_n = \sum_{i=1}^n I_{\{X_i > u_n\}} \sim B(n, \bar{F}(u_n))$$

则

$$B_n \xrightarrow{d} \text{Poi}(\tau) \iff E(B_n) = n\bar{F}(u_n) \rightarrow \tau$$

于是有:

$$P(M_n \leq u_n) = P(B_n = 0) \rightarrow \exp\{-\tau\}$$

在何时, $P(M_n \leq u_n)$ 会收敛于简单的 0 或 1

Corollary 3.1.2

假设 $x_F < \infty$ 且

$$\bar{F}(x_F-) = F(x_F) - F(x_F-) > 0.$$

那么对于任意的序列 (u_n) 都有:

$$P(M_n \leq u_n) \rightarrow \rho,$$

要么 $\rho = 0$, 要么 $\rho = 1$ 。

在何时, $P(M_n \leq u_n)$ 会收敛于简单的 0 或 1 (cont'd)

Corollary 3.1.2

如果一个分布在很接近右端点 (右端点有限大) 处仍有跳, M_n 的非退化极限分布不存在。

Theorem 3.1.3

令 F 为具有右端点 $x_F \leq \infty$ 的累积分布函数, 再令 $\tau \in (0, \infty)$ 。存在满足 $n\bar{F}(u_n) \rightarrow \tau$ 的序列 (u_n) , 当且仅当

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{F}(x-)} = 1 \quad (3.6)$$

当 F 为离散型分布时, $\bar{F}(x-) = \bar{F}(x-1)$ 。

几个无穷大右端点下离散分布的例子

Example 3.1.4–3.1.6

对于 Poisson 分布、几何分布和负二项分布，证明式 3.6 并不成立：

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{F}(x-)} = 1 \quad (3.6)$$

3.2 Weak Convergence of Maxima Under Affine Transformations

收敛到的分布是什么？

Definition 3.2.1 最大稳定分布

如果对于 iid 的 X, X_1, \dots, X_n , 有合适的序列 $c_n > 0, d_n \in \mathbb{R}$ 和每个 $n > 2$, 都有:

$$\max(X_1, \dots, X_n) \stackrel{d}{=} c_n X + d_n \quad (3.7)$$

则称一个非退化的随机变量 X (对应的分布或累积分布函数) 是最大稳定的。则有:

$$c_n^{-1} (M_n - d_n) \stackrel{d}{=} X \quad (3.8)$$

最大稳定分布和最大值极限分布的关系

Theorem 3.2.2 最大稳定分布和最大值极限分布的关系

最大稳定分布刚好就是任意可能的独立同分布随机变量最大值（经过中心化标准化）的极限分布。

证明方式：两种分布是“同一种”、“同一类”。

同分布 (same distribution) 和同类 (same type)

定义

同分布 (same distribution) 为:

$$X \stackrel{d}{=} Y$$

同类 (same type) 为, 存在 $a \in \mathbb{R}$ 和 $b > 0$, 有:

$$X \stackrel{d}{=} bY + a$$

向某类型收敛 (Convergence to Types)

Theorem A1.5: Convergence to Types Theorem

令 A, B, A_1, A_2, \dots 为随机变量, $b_n > 0, \beta_n > 0, a_n, \alpha_n \in \mathbb{R}$ 为常数, 假设 $b_n^{-1}(A_n - a_n) \xrightarrow{d} A$, 那么下面的关系:

$$\beta_n^{-1}(A_n - \alpha_n) \xrightarrow{d} B \quad (\text{A.2})$$

成立, 当且仅当

$$\lim_{n \rightarrow \infty} b_n / \beta_n = b \in [0, \infty), \quad \lim_{n \rightarrow \infty} (a_n - \alpha_n) / \beta_n = a \in \mathbb{R}. \quad (\text{A.3})$$

如果 (A.2) 成立, 那么 $B \stackrel{d}{=} bA + a$, 且 a, b 是唯一能让关系成立的常数。

向某类型收敛 (Convergence to Types, cont'd)

Theorem A1.5: Convergence to Types Theorem

令 A, B, A_1, A_2, \dots 为随机变量, $b_n > 0, \beta_n > 0, a_n, \alpha_n \in \mathbb{R}$ 为常数, 假设 $b_n^{-1}(A_n - a_n) \xrightarrow{d} A$, 那么下面的关系:

$$\beta_n^{-1}(A_n - \alpha_n) \xrightarrow{d} B \quad (\text{A.2})$$

成立, 当且仅当

$$\lim_{n \rightarrow \infty} b_n / \beta_n = b \in [0, \infty), \quad \lim_{n \rightarrow \infty} (a_n - \alpha_n) / \beta_n = a \in \mathbb{R}. \quad (\text{A.3})$$

当 (A.2) 成立, A 非退化当且仅当 $b > 0$, A 和 B 属于同一类。

非高斯领域的中心极限定理: Fisher-Tippett theorem

Theorem 3.2.3: Fisher-Tippett theorem

让 (X_n) 是独立同分布随机变量序列。如果存在规范化参数 $c_n > 0$, $d_n \in \mathbb{R}$ 和一些非退化的累积分布函数 H 使得

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H, \quad (3.9)$$

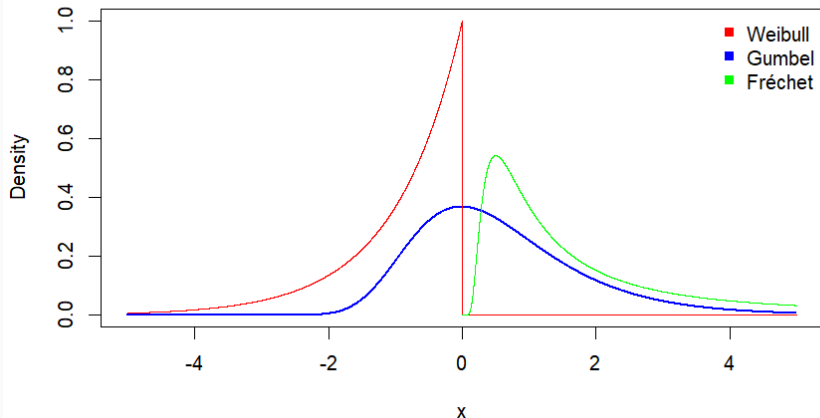
那么 H 属于以下三种分布的其中一种:

$$\text{Fréchet: } \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp \{-x^{-\alpha}\}, & x > 0 \end{cases} \quad \alpha > 0.$$

$$\text{Weibull: } \Psi_\alpha(x) = \begin{cases} \exp \{-(-x)^\alpha\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0.$$

$$\text{Gumbel: } \Lambda(x) = \exp \{-e^{-x}\}, \quad x \in \mathbb{R}.$$

Fisher-Tippett 定理中的三种分布



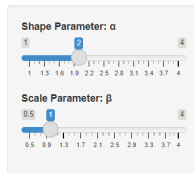
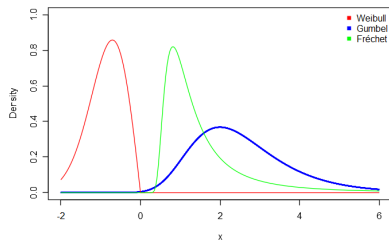
Fisher-Tippett 定理的可视化

相关 R 代码

```
library(shiny)
```

```
runUrl("https://yuanzhuang.xyz/uploads/EVT/FTG.zip")
```

Fisher-Tippett-Gnedenko Visualization



Reference: Embrechts P., Klüppelberg C., Mikosch T. Modelling Extremal Events[M]. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997.
Notice: The Parametrization may be different.

Fisher-Tippett theorem 的证明思路

对于所有 $t > 0$, 有:

$$F^{[nt]}(c_{[nt]}x + d_{[nt]}) \rightarrow H(x), \quad x \in \mathbb{R}$$

也有:

$$F^{[nt]}(c_n x + d_n) = (F^n(c_n x + d_n))^{[nt]/n} \rightarrow H^t(x)$$

由 Covergence to types theorem, 存在函数 $\gamma(t) > 0, \delta(t) \in \mathbb{R}$ 满足

$$\lim_{n \rightarrow \infty} \frac{c_n}{c_{[nt]}} = \gamma(t), \quad \lim_{n \rightarrow \infty} \frac{d_n - d_{[nt]}}{c_{[nt]}} = \delta(t), \quad t > 0,$$

且

$$H^t(x) = H(\gamma(t)x + \delta(t)) \tag{3.10}$$

Fisher-Tippett theorem 的证明思路 (cont'd)

由

$$H^t(x) = H(\gamma(t)x + \delta(t)) \quad (3.10)$$

有

$$\gamma(st) = \gamma(s)\gamma(t), \quad \delta(st) = \gamma(t)\delta(s) + \delta(t) \quad (3.11)$$

上述函数方程有三类解。

关于 Fisher-Tippett theorem 的几点注意事项：仿射变换

- 极限定理在仿射变换下也成立，只不过需要更换规范化参数。如果：

$$\lim_{n \rightarrow \infty} P \left(c_n^{-1} (M_n - d_n) \leq x \right) = H(cx + d)$$

那么，对于 $\tilde{c}_n = c_n/c$ 和 $\tilde{d}_n = d_n - dc_n/c$ ，有

$$\lim_{n \rightarrow \infty} P \left(\tilde{c}_n^{-1} \left(M_n - \tilde{d}_n \right) \leq x \right) = H(x)$$

关于 Fisher-Tippett theorem 的几点注意事项: Weibull 分布的定义

- 我们学过的 Weibull 分布定义域一般在 $[0, +\infty)$ 上:

$$F_{\alpha}(x) = 1 - e^{-x^{\alpha}}, \quad x \geq 0$$

- 书中的 Weibull 分布和我们学过的 Weibull 分布可以这样转换:

$$\Psi_{\alpha}(x) = 1 - F_{\alpha}(-x), \quad x < 0$$

- 部分精算教材中将 $\Psi_{\alpha}(x)$ 称为 “Weibull EV”, 但在本书中, 为了传统, 还是使用 $\Psi_{\alpha}(x)$ 作为 Weibull 分布。

关于 Fisher-Tippett theorem 的几点注意事项：极值分布间的关系

- 虽然三种分布看起来十分不同，但它们有这样的转换关系：

$$X \text{ has df } \Phi_\alpha \iff \ln X^\alpha \text{ has df } \Lambda \iff -X^{-1} \text{ has df } \Psi_\alpha$$

Denition 3.2.6

在定理 3.2.3 中列出的 Φ_α , Ψ_α 和 Λ 叫作标准极值分布，相应的随机变量叫标准极值随机变量。 Φ_α , Ψ_α 和 Λ 这一类的分布叫作极值分布，相应的随机变量叫极值随机变量。

标准极值分布的最大稳定特性

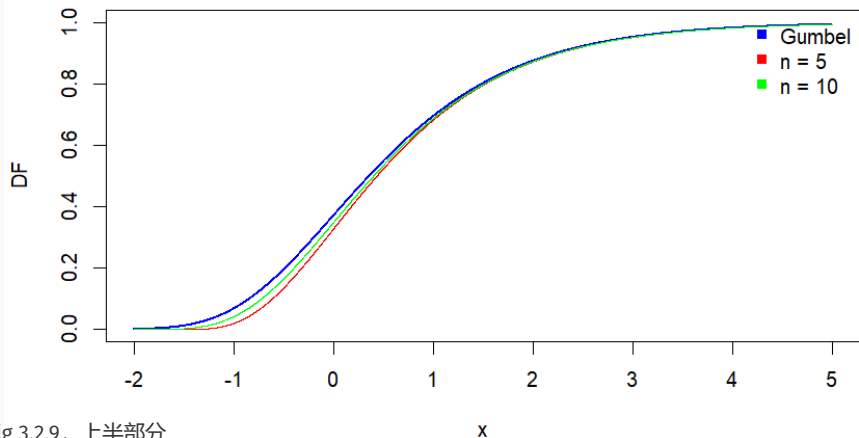
- Fréchet: $M_n \stackrel{d}{=} n^{1/\alpha} X$
- Weibull: $M_n \stackrel{d}{=} n^{-1/\alpha} X$
- Gumbel: $M_n \stackrel{d}{=} X + \ln n$

在不同的 F 下, 各自最大值的极限分布是什么?

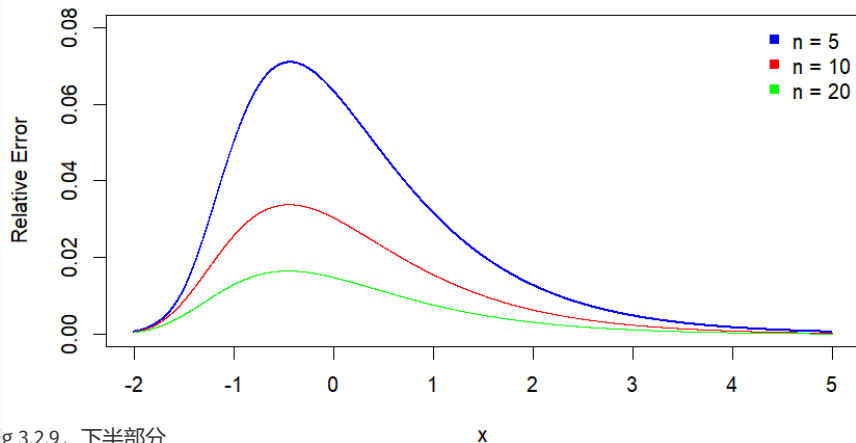
Example 3.2.7 指数分布最大值的极限分布

$$\begin{aligned} P(M_n - \ln n \leq x) &= (P(X \leq x + \ln n))^n \\ &= (1 - n^{-1}e^{-x})^n \\ &\rightarrow \exp\{-e^{-x}\} = \Lambda(x), \quad x \in \mathbb{R} \end{aligned}$$

从图像查看最大值的依分布收敛特性：指数 VS Gumbel



使用 Gumbel 分布近似多个指数变量最大值的相对误差



在不同的 F 下，各自最大值的极限分布是什么？(cont'd)

Example 3.2.8 柯西分布最大值的极限分布

柯西分布的 df 为：

$$f(x) = (\pi (1 + x^2))^{-1}, \quad x \in \mathbb{R}$$

所以有：

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x)}{(\pi x)^{-1}} = \lim_{x \rightarrow \infty} \frac{f(x)}{\pi^{-1} x^{-2}} = \lim_{x \rightarrow \infty} \frac{\pi x^2}{\pi (1 + x^2)} = 1$$

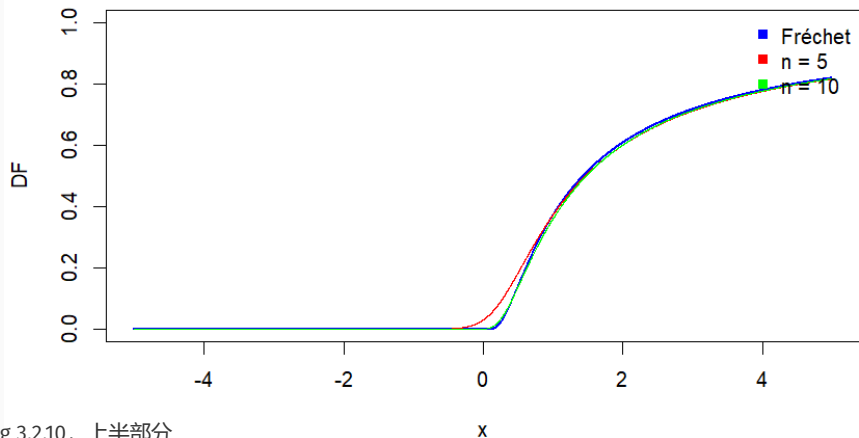
也就是 $\bar{F}(x) \sim (\pi x)^{-1}$ 。

在不同的 F 下, 各自最大值的极限分布是什么? (cont'd)

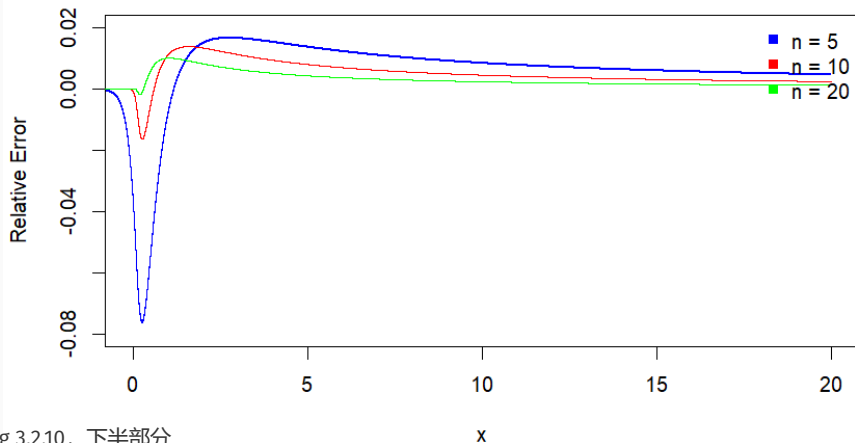
Example 3.2.8 柯西分布最大值的极限分布

$$\begin{aligned} P\left(M_n \leq \frac{nx}{\pi}\right) &= \left(1 - \bar{F}\left(\frac{nx}{\pi}\right)\right)^n \\ &= \left(1 - \frac{1}{nx} + o(1)\right)^n \\ &\rightarrow \exp\{-x^{-1}\} = \Phi_1(x), \quad x > 0 \end{aligned}$$

从图像查看最大值的依分布收敛特性：柯西 VS Fréchet



使用 Fréchet 分布近似多个柯西变量最大值的相对误差



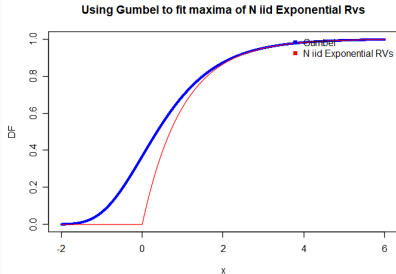
不同 N 下，最大值确切分布与极值分布的相似程度

相关 R 代码

```
library(shiny)
```

```
runUrl("https://yuanzhuang.xyz/uploads/EVT/FTGApproximation.zip")
```

Sample Maxima Approximation Using Extreme Value Distributions



Number of iid Exponential Rvs:



Number of iid Cauchy Rvs:



Number of iid Normal Rvs:



3.3 Maximum Domains of Attraction and Norming Constants

- c_n 和 d_n 应该怎样选取？（其实在上一节中已经告诉了我们）
- 如何很快就知道，分布 F 下最大值的极限分布是 Fréchet、Weibull 和 Gumbel 中的哪一个？（最大吸引域和冯·米塞斯条件）

Definition 3.3.1

如果存在常数 $c_n > 0, d_n \in \mathbb{R}$ 满足 $c_n^{-1}(M_n - d_n) \xrightarrow{d} H$, 就说随机变量 X (X 的累积分布函数 F , X 的分布) 属于极值分布 H 的最大吸引域。该关系可写为 $X \in \text{MDA}(H)$ ($F \in \text{MDA}(H)$)。

上述条件也相当于在说：

$$\lim_{n \rightarrow \infty} P(M_n \leq c_n x + d_n) = \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x), \quad x \in \mathbb{R}$$

最大吸引域的特征（后续证明最大吸引域的方式）

Proposition 3.3.2

F 属于极值分布 H 的最大吸引域，具有规范化参数 $c_n > 0, d_n \in \mathbb{R}$ 当且仅当

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) = -\ln H(x), \quad x \in \mathbb{R}.$$

$H(x) = 0$ 时，极限视作 ∞ 。

- 正则变化相关知识 (附录 A3.1)
- 尾等价 (Definition 3.3.3)

Definition 3.3.5: 分位数函数

$$F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad 0 < t < 1$$

定义了 F 的 t 分位数。

3.3.1 Fréchet 分布的最大吸引域

c_n 的选取: F 的 $1 - 1/n$ 分位数

$$\begin{aligned} c_n &= F^{\leftarrow}(1 - n^{-1}) = \inf \{x \in \mathbb{R} : F(x) \geq 1 - n^{-1}\} \\ &= \inf \{x \in \mathbb{R} : (1/\bar{F})(x) \geq n\} \\ &= (1/\bar{F})^{\leftarrow}(n). \end{aligned} \tag{3.13}$$

d_n 为 0, 也就是不需要中心化。

Φ_α 最大吸引域的充要条件: F 尾部正则变化, 参数为 $-\alpha$

Theorem 3.3.7: Φ_α 的最大吸引域

F 属于 $\Phi_\alpha, \alpha > 0$ 的最大吸引域, 当且仅当 $\bar{F}(x) = x^{-\alpha}L(x)$ (也就是 $\bar{F} \in \mathcal{R}_{-\alpha}$), 其中 L 是某种慢变函数。

如果 $F \in \text{MDA}(\Phi_\alpha)$, 那么

$$c_n^{-1}M_n \xrightarrow{d} \Phi_\alpha, \quad (3.14)$$

其中 c_n 是 F 的 $1 - 1/n$ 分位数 (即式 3.13)。

是否存在更简单、更易辨识的 $F \in \text{MDA}(\Phi_\alpha)$ 条件?

Corollary 3.3.8: von Mises condition

令 F 绝对连续, 具有密度函数 f , 其满足

$$\lim_{x \rightarrow \infty} \frac{x f(x)}{\bar{F}(x)} = \alpha > 0, \quad (3.16)$$

那么 $F \in \text{MDA}(\Phi_\alpha)$ 。

这是一个充分条件, 而不是充要条件。

最大吸引域中是否还存在其它分布？

Proposition 3.3.9: Closure property of $\text{MDA}(\Phi_\alpha)$

令 F 和 G 是累积分布函数, 假设 $F \in \text{MDA}(\Phi_\alpha)$, 其正规化参数为 $c_n > 0$, 也就是说

$$\lim_{n \rightarrow \infty} F^n(c_n x) = \Phi_\alpha(x), \quad x > 0. \quad (3.17)$$

那么

$$\lim_{n \rightarrow \infty} G^n(c_n x) = \Phi_\alpha(x), \quad x > 0,$$

当且仅当 F 和 G 尾等价:

$$\lim_{x \rightarrow \infty} \bar{F}(x)/\bar{G}(x) = c^\alpha.$$

Fréchet 分布最大吸引域中的分布是什么？

- 是满足 von Mises 条件的分布和与它们尾等价的分布；
- 是 \bar{F} 以 $-\alpha$ 为参数正则变化的分布；
- 具体来说，是：
 1. Cauchy (Example 3.2.8)
 2. Pareto
 3. Burr
 4. Loggamma (Example 3.3.11, 接下来有案例)
 5. $\alpha < 2$ 的稳定分布

Example 3.3.11: Loggamma distribution

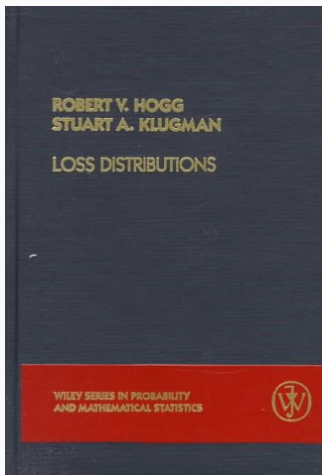
Loggamma 分布有尾:

$$\bar{F}(x) \sim \frac{\alpha^{\beta-1}}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{-\alpha}, \quad x \rightarrow \infty, \quad \alpha, \beta > 0. \quad (3.18)$$

代入 c_n 是 F 的 $1 - 1/n$ 分位数, 解得:

$$c_n \sim \left((\Gamma(\beta))^{-1} (\ln n)^{\beta-1} n \right)^{1/\alpha}$$

Loggamma 分布的尾部行为来自于其密度函数



VI. LOGGAMMA DISTRIBUTION

Support: $x > 1$

Parameters: $\alpha > 0, \lambda > 0$

D.f.: $F(x) = \Gamma(\alpha; \lambda \ln x)$

P.d.f.: $f(x) = \frac{\lambda^\alpha (\ln x)^{\alpha-1}}{x^{\lambda+1} \Gamma(\alpha)}$

Moments: $E[X^n] = \left(1 - \frac{n}{\lambda}\right)^{-\alpha}, \lambda > n$

Mode: $1, \alpha \leq 1; \quad \exp\left(\frac{\alpha-1}{\lambda+1}\right), \alpha > 1$

3.3.2 Weibull 分布的最大吸引域

Recall: Weibull 和 Fréchet 分布的关系

- X has df $\Phi_\alpha \iff -X^{-1}$ has df Ψ_α
- 可以预料到 Weibull 最大吸引域的条件和 Fréchet 非常相似，可能只需要做变量变化
- 不同之处：在 Weibull 最大吸引域中的分布 F 都有有限的右端点 x_F

c_n 和 d_n 的选取

- $c_n = x_F - F^{\leftarrow}(1 - n^{-1})$
- $d_n = x_F$
- c_n 即是 F 的 $1 - 1/n$ 分位数到右端点的距离
- d_n 即是右端点 x_F

Ψ_α 最大吸引域的充要条件: $F(x_F - x^{-1})$ 尾部正则变化, 参数为 $-\alpha$

Theorem 3.3.12: Ψ_α 的最大吸引域

F 属于 $\Psi_\alpha, \alpha > 0$ 的最大吸引域, 当且仅当 $x_F < \infty$ 且 $\bar{F}(x_F - x^{-1}) = x^{-\alpha}L(x)$ (也就是 $\bar{F}(x_F - x^{-1}) \in \mathcal{R}_{-\alpha}$), 其中 L 是某种慢变函数。

如果 $F \in \text{MDA}(\Psi_\alpha)$, 那么

$$c_n^{-1}(M_n - x_F) \xrightarrow{d} \Phi_\alpha, \quad (3.20)$$

其中 c_n 、 d_n 分别为: F 的 $1 - 1/n$ 分位数到右端点的距离和右端点 x_F 。

Weibull 最大吸引域的 von Mises 条件以及 $\text{MDA}(\Psi_\alpha)$ 的封闭性

- Weibull 的 von Mises 条件仅与 Frechet 的略有不同，不同之处来源于变量变换 (Corollary 3.3.13):

$$\lim_{x \uparrow x_F} \frac{(\mathbf{x}_F - \mathbf{x}) f(x)}{\bar{F}(x)} = \alpha > 0 \quad (3.23)$$

- $\text{MDA}(\Psi_\alpha)$ 在尾等价下封闭 (Proposition 3.3.14)。也即: $\text{MDA}(\Psi_\alpha)$ 中包含满足 von Mises 条件的分布，还有与这些分布尾等价的分布。

Weibull 最大吸引域中的具体分布

分布名称	密度函数或生存函数	c_n	d_n
均匀分布	$f(x) = 1, 0 < x < 1$	n^{-1}	1
Power law at x_F	$\bar{F}(x) = K (x_F - x)^\alpha$	$(Kn)^{-1/\alpha}$	x_F
Beta 分布	$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$	$\left(n \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b+1)} \right)^{-1/b}$	1

三个分布分别对应 Example 3.3.15、Example 3.3.16、Example 3.3.17。

Karamata Theorem: 对慢变函数的积分

Theorem A3.6: Karamata 定理

令 $L \in \mathcal{R}_0$ 在 $[x_0, \infty)$ 上局部有界, $x_0 \geq 0$ 。那么:

(a) 对于 $\alpha > -1$,

$$\int_{x_0}^x t^\alpha L(t) dt \sim (\alpha + 1)^{-1} x^{\alpha+1} L(x), \quad x \rightarrow \infty,$$

(b) 对于 $\alpha < -1$,

$$\int_x^\infty t^\alpha L(t) dt \sim -(\alpha + 1)^{-1} x^{\alpha+1} L(x), \quad x \rightarrow \infty.$$

Karamata Theorem: 积分项中含正则变化项的积分

Theorem 12 (Karamata's theorem, direct part). *Let $f \in \mathcal{RV}_\rho$ be locally bounded on $[a, \infty)$. Then*

(i) *for $\sigma \geq -(\rho + 1)$*

$$\frac{x^{\sigma+1}f(x)}{\int_a^x t^\sigma f(t)dt} \rightarrow \sigma + \rho + 1;$$

(ii) *for $\sigma < -(\rho + 1)$*

$$\frac{x^{\sigma+1}f(x)}{\int_x^\infty t^\sigma f(t)dt} \rightarrow -(\sigma + \rho + 1).$$

(The latter also holds for $\sigma = -(\rho + 1)$ if the integral is finite.)

It turns out that this behavior also characterizes regular variation.

Beta 在 Weibull 分布最大吸引域中

Example 3.3.17: Beta distribution

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}, \quad 0 < x < 1, \quad a, b > 0$$

3.3.3 Gumbel 分布的最大吸引域

Gumbel 最大吸引域与 Fréchet 和 Weibull 分布的不同

- Gumbel 最大吸引域和正则变化没有直接联系
- 很多重尾和轻尾分布都在 Gumbel 分布的最大吸引域中
- Gumbel 分布最大吸引域中的分布有些有无限的右端点 $x_F = \infty$, 有些只有有限的右端点 ($x_F < \infty$)

Definition 3.3.18: von Mises function

令 F 为累积分布函数, 右端点为 $x_F \leq \infty$ 。假设存在 $z < x_F$ 使 F 有如下表示:

$$\bar{F}(x) = c \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\}, \quad z < x < x_F, \quad (3.24)$$

其中 c 是某个正数, $a(\cdot)$ 是正且一致连续的函数, 导数 a' 存在, 且 $\lim_{x \uparrow x_F} a'(x) = 0$ 。那么 F 叫作 von Mises 函数, 函数 $a(\cdot)$ 被称为 F 的辅助函数。

很多常见分布函数的生存函数都是 von Mises 函数

- 指数分布 (Example 3.3.19)
- Weibull 分布 (Example 3.3.20)
- Erlang 分布 (Example 3.3.21)
- 有限右端点下的指数行为 (Example 3.3.22)
- 右端点下的可导性 (Example 3.3.23)

Proposition 3.3.24: Properties of von Mises functions

每个 von Mises 函数 F 在 (z, x_F) 上都绝对连续, 密度函数 f 为正。辅助函数可以选为 $a(x) = \bar{F}(x)/f(x)$ 。另外, 下面的特点都成立:

(a) 如果 $x_F = \infty$, 那么 $\bar{F} \in \mathcal{R}_{-\infty}$ 且

$$\lim_{x \rightarrow \infty} \frac{x f(x)}{\bar{F}(x)} = \infty. \quad (3.26)$$

(b) 如果 $x_F < \infty$, 那么 $\bar{F}(x_F - x^{-1}) \in \mathcal{R}_{-\infty}$ 且

$$\lim_{x \uparrow x_F} \frac{(x_F - x) f(x)}{\bar{F}(x)} = \infty. \quad (3.27)$$

Definition A3.11: Rapid Variation

一个在 $(0, \infty)$ 上正的, 勒贝格可测的函数 h 是快速变化的, 参数为 $-\infty$ (写为 $h \in \mathcal{R}_{-\infty}$), 如果:

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = \begin{cases} 0 & \text{if } t > 1, \\ \infty & \text{if } 0 < t < 1. \end{cases}$$

正则变化函数与快变函数的表示定理

Theorem A3.12(b): Properties of functions of rapid variation

如果 $h \in \mathcal{R}_{-\infty}$, 那么存在函数 c 和 δ , 满足 $c(x) \rightarrow c_0 \in (0, \infty)$; $\delta(x) \rightarrow -\infty, x \rightarrow \infty$ 。且对于某些 $z > 0$,

$$h(x) = c(x) \exp \left\{ \int_z^x \frac{\delta(u)}{u} du \right\}, \quad x \geq z. \quad (\text{A.14})$$

反之亦然。

正则变化函数与快变函数的表示定理 (cont'd)

Theorem A3.3: Representation theorem for regularly varying functions

如果 $h \in \mathcal{R}_{-\infty}$, 那么存在函数 c 和 δ , 满足 $c(x) \rightarrow c_0 \in (0, \infty)$; $\delta(x) \rightarrow \alpha, x \rightarrow \infty$ 。且对于某些 $z > 0$,

$$h(x) = c(x) \exp \left\{ \int_z^x \frac{\delta(u)}{u} du \right\}, \quad x \geq z. \quad (\text{A.12})$$

反之亦然。

Gumbel 最大吸引域的 von Mises 条件

Proposition 3.3.29: von Mises functions and $\text{MDA}(\Lambda)$

假设 F 是 von Mises 函数, 那么 $F \in \text{MDA}(\Lambda)$ 。正规化参数为:

$$d_n = F^{\leftarrow}(1 - n^{-1}) \quad \text{and} \quad c_n = a(d_n),$$

其中 a 是 F 的辅助函数。

Theorem 3.3.26: Characterisation I of $\text{MDA}(\Lambda)$

F 属于 Λ 的最大吸引域, 当且仅当有 $z < x_F$ 使 F 有表示:

$$\bar{F}(x) = c(x) \exp \left\{ - \int_z^x \frac{g(t)}{a(t)} dt \right\}, \quad z < x < x_F, \quad (3.33)$$

c 和 g 是可测函数, 满足 $c(x) \rightarrow c > 0, g(x) \rightarrow 1, x \uparrow x_F$, $a(x)$ 是正的、绝对连续的函数。其密度 $a'(x)$ 有 $\lim_{x \uparrow x_F} a'(x) = 0$ 。

Remark: Gumbel 最大吸引域的充要条件

- 定理 3.3.26 中, $d_n = F^{\leftarrow}(1 - n^{-1})$, $c_n = a(d_n)$;
- $a(x)$ 的其中的一个选取为:

$$\int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt \quad (3.24)$$

这其实就是 $E(X - x \mid X > x)$ 。(mean excess)

- 很多种表示其实都可接受, 如: $\bar{F}(x) = c(x) \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\}$, 这本质来源于 c 和 g 之间的权衡。

Gumbel 最大吸引域的另一充要条件

Theor 3.3.27: Characterisation II of $\text{MDA}(\Lambda)$

累积分布函数 F 属于 A 的最大吸引域, 当且仅当存在正函数 \tilde{a} 使下列等式满足:

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x + t\tilde{a}(x))}{\bar{F}(x)} = e^{-t} \quad (3.36)$$

可能的 $\tilde{a} = a$ 如式 (3.34) 所示。

Gumbel 分布最大吸引域中的分布是什么？

- $\text{MDA}(\Lambda_\alpha)$ 在尾等价下封闭 (Proposition 3.3.28)。
- 是生存函数为 von Mises 函数的分布和与它们尾等价的分布；
- 具体来说，是：
 1. 指数分布
 2. Gamma 分布
 3. 正态分布
 4. 对数正态分布
 5. 尾部的指数行为 ($\bar{F}(x) = K \exp \left\{ -\frac{\alpha}{x_F - x} \right\}$, $x < x_F$, $\alpha, K > 0$)
 6. Benktander-type-I 和 Benktander-type-II 分布

使用 Gumbel 近似 iid 正态分布的最大值

Example 3.3.29 Normal distribution

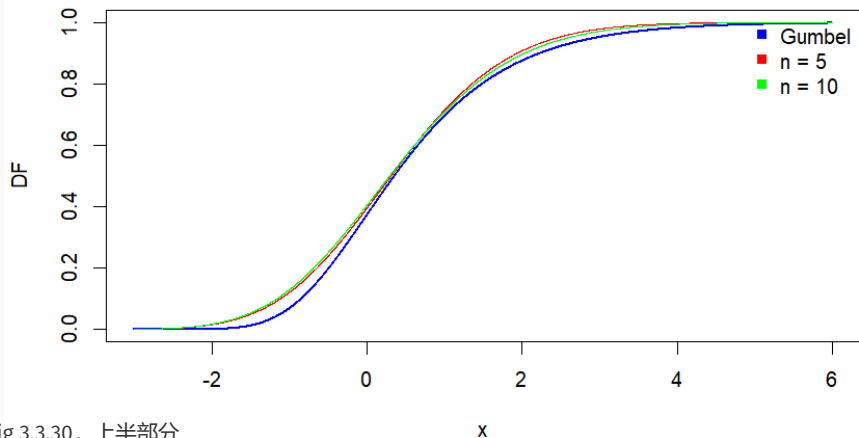
$$d_n = (2 \ln n)^{1/2} - \frac{\ln \ln n + \ln 4\pi}{2(2 \ln n)^{1/2}} + o((\ln n)^{-1/2})$$

$$c_n = a(d_n) \sim (2 \ln n)^{-1/2}$$

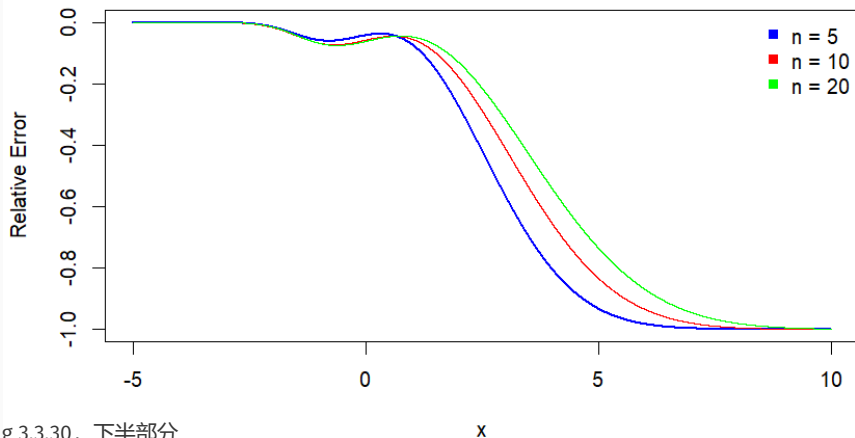
有:

$$\sqrt{2 \ln n} \left(M_n - \sqrt{2 \ln n} + \frac{\ln \ln n + \ln 4\pi}{2(2 \ln n)^{1/2}} \right) \xrightarrow{d} \Lambda \quad (3.40)$$

从图像查看最大值的依分布收敛特性：正态 VS Gumbel



使用 Gumbel 分布近似多个正态变量最大值的相对误差



Corollary 3.3.32: Existence of moments

X 有 $F \in \text{MDA}(\Lambda)$, 具有有限右端点。那么 $\bar{F} \in \mathcal{R}_{-\infty}$ 。特别地, $E(X^+)^\alpha < \infty$ 存在。 $\alpha > 0$, 其中 $X^+ = \max(0, X)$ 。

3.4 The Generalised Extreme Value Distribution and the Generalised Pareto Distribution

- 将三种极值分布归为一类 (GEV)
- 寻找 mean excess 的极限分布 (GPD)

极值分布的 Jenkinson-von Mises 表示

Definition 3.4.1: The generalised extreme value distribution (GEV)

定义如下累积分布函数:

$$H_{\xi}(x) = \begin{cases} \exp \{-(1 + \xi x)^{-1/\xi}\} & \text{if } \xi \neq 0 \\ \exp\{-\exp\{-x\}\} & \text{if } \xi = 0 \end{cases}$$

其中 $1 + \xi x > 0$ 。

Remark: GEV

- 定义域:

$$x > -\xi^{-1} \quad \text{for} \quad \xi = \alpha^{-1} > 0 \dots \text{Fréchet}$$

$$x < -\xi^{-1} \quad \text{for} \quad \xi = -\alpha^{-1} < 0 \dots \text{Weibull}$$

$$x \in \mathbb{R} \quad \text{for} \quad \xi = 0 \dots \text{Gumbel}$$

- 也把经过平移或放缩的后的随机变量 $H_{\xi;\mu,\psi}$ 称为 GEV:

$$(x - \mu)/\psi, \mu \in \mathbb{R}, \psi > 0$$

Theorem 3.4.5: Characterisation of MDA(H_ξ)

对于 $\xi \in \mathbb{R}$, 下列命题等价: (a) $F \in \text{MDA}(H_\xi)$

(b) 存在一个正的、可测函数 $a(\cdot)$, 对于 $1 + \xi x > 0$, 有:

$$\lim_{u \uparrow x_F} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{if } \xi \neq 0, \\ e^{-x} & \text{if } \xi = 0. \end{cases} \quad (3.42)$$

(c) 对于 $x, y > 0$, $y \neq 1$, 令 $U(t) = F^{\leftarrow}(1 - t^{-1})$, $t > 0$, 有:

$$\lim_{s \rightarrow \infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^\xi - 1}{y^\xi - 1} & \text{if } \xi \neq 0 \\ \frac{\ln x}{\ln y} & \text{if } \xi = 0 \end{cases} \quad (3.43)$$

Remarks: $\text{MDA}(H_\xi)$ 的表征

- 式 (3.42) 还可被表示为如下 mean excess 的形式:

$$\lim_{u \uparrow x_F} P\left(\frac{X-u}{a(u)} > x \mid X > u\right) = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{if } \xi \neq 0 \\ e^{-x} & \text{if } \xi = 0 \end{cases} \quad (3.45)$$

Definition 3.4.6: Excess distribution function, mean excess function

令 X 为随机变量, 累积分布函数为 F , 右端点为 x_F 。对于 $u < x_F$,

$$F_u(x) = P(X - u \leq x \mid X > u), \quad x \geq 0, \quad (3.47)$$

为超过限额 u 的条件分布函数。Mean excess function 则为:

$$e(u) = E(X - u \mid X > u)$$

Mean Excess 的极限：广义帕累托分布

Definition 3.4.9: The generalised Pareto distribution (GPD)

定义累积分布函数 G_ξ :

$$G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - e^{-x} & \text{if } \xi = 0, \end{cases} \quad (3.50)$$

其中,

$$\begin{aligned} x &\geq 0 && \text{if } \xi \geq 0, \\ 0 \leq x &\leq -1/\xi && \text{if } \xi < 0. \end{aligned}$$

G_ξ 被称为标准广义帕累托分布 (GPD)。

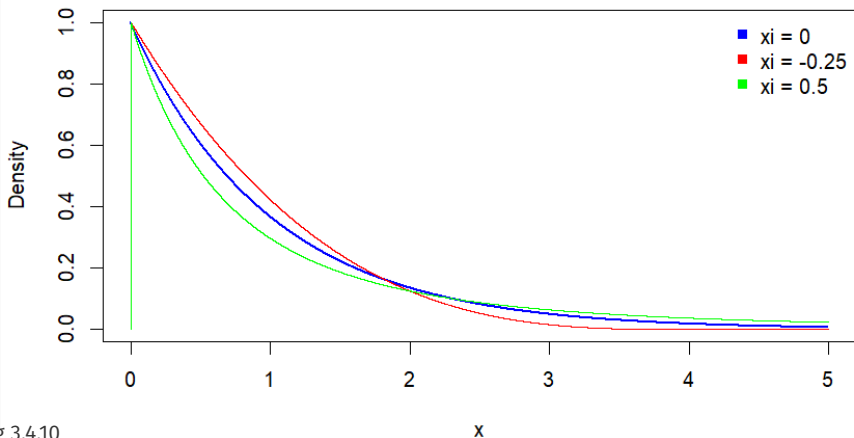
Remarks: 广义帕累托分布

- $G_{\xi;\nu,\beta}$, 也就是 $(x - \nu)/\beta$, 是标准广义帕累托分布平移、放缩后的结果, 同样也被称为 GPD
- $G_{\xi;0,\beta}$, 即不平移, 只放缩的分布十分重要, 即:

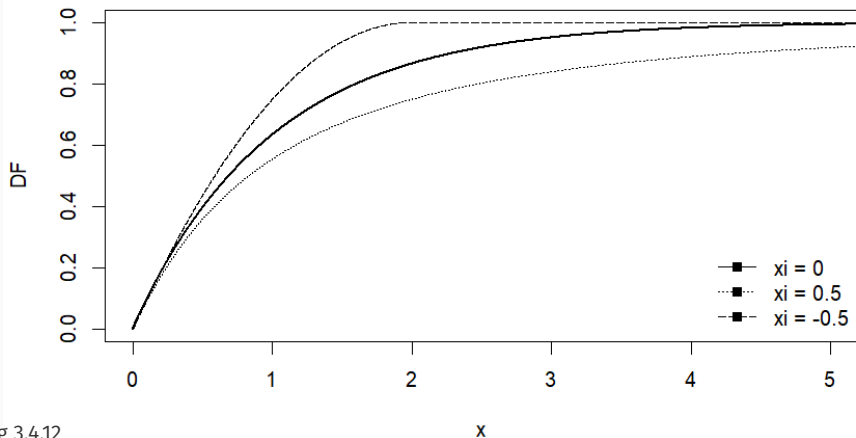
$$G_{\xi,\beta}(x) = 1 - \left(1 + \xi \frac{x}{\beta}\right)^{-1/\xi}, \quad x \in D(\xi, \beta) \quad (3.51)$$

$$x \in D(\xi, \beta) = \begin{cases} [0, \infty) & \text{if } \xi \geq 0 \\ [0, -\beta/\xi] & \text{if } \xi < 0 \end{cases}$$

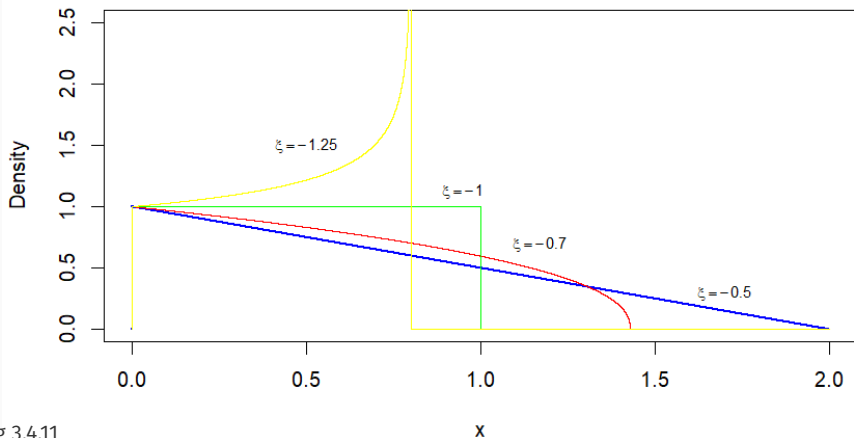
不同 ξ 下 GPD 的密度函数 (ξ 越大, 尾越厚)



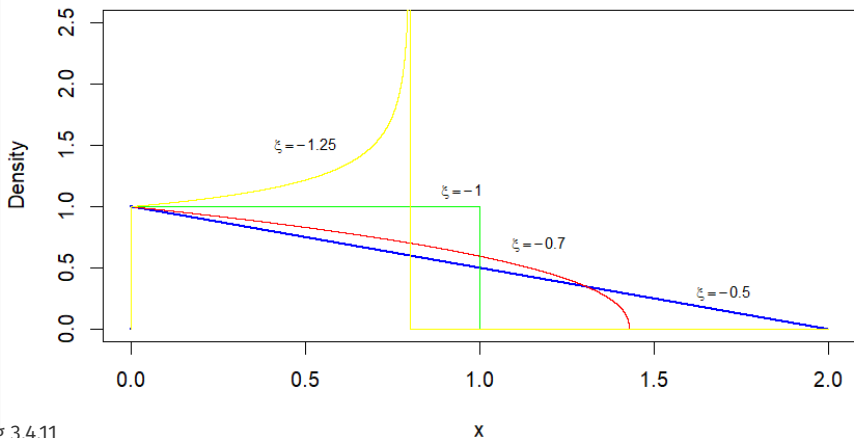
不同 ξ 下 GPD 的密度函数 (ξ 越大, 尾越厚, cont'd)



不同 ξ 下 GPD 的密度函数



不同 ξ 下 GPD 的密度函数 (负数, 即 Weibull)



关于 GPD 的一些结论

- $EX < \infty$, 如果 $\xi < 1$
- 假设 $x_i \in D(\xi, \beta), i = 1, 2$, 那么

$$\frac{\bar{G}_{\xi, \beta}(x_1 + x_2)}{\bar{G}_{\xi, \beta}(x_1)} = \bar{G}_{\xi, \beta + \xi x_1}(x_2) \quad (3.53)$$

Theorem 3.4.13: Properties of GPD

假设 N 服从 $\text{Poi}(\lambda)$, 与 iid 序列 (X_n) 相独立。 (X_n) 服从 GPD 分布, 参数为 ξ and β 。令 $M_N = \max(X_1, \dots, X_N)$ 。那么

$$P(M_N \leq x) = \exp \left\{ -\lambda \left(1 + \xi \frac{x}{\beta} \right)^{-1/\xi} \right\} = H_{\xi; \mu, \psi}(x),$$

其中 $\mu = \beta \xi^{-1} (\lambda^\xi - 1)$, 其中 $\psi = \beta \lambda^\xi$ 。

3.5 Almost Sure Behaviour of Maxima

- 并没有像强大数定律那样的统一的定理
- 但有一些 a.s. 的结论

部分最大值的最大 a.s. 增长

Theorem 3.5.1 Characterisation of the maximal a.s. growth of partial maxima

假设 (u_n) 非减, 那么

$$P(M_n > u_n \text{ i.o.}) = P(X_n > u_n \text{ i.o.}) \quad (3.54)$$

特别地,

$$P(M_n > u_n \text{ i.o.}) = 0 \text{ 或 } = 1$$

决定于

$$\sum_{n=1}^{\infty} P(X > u_n) < \infty \text{ 或 } = \infty. \quad (3.55)$$

部分最大值的最小 a.s. 增长

Theorem 3.5.2 Characterisation of the minimal a.s. growth of partial maxima

假设 (u_n) 非减, 且下列条件满足:

$$\bar{F}(u_n) \rightarrow 0, \quad (3.57)$$

$$n\bar{F}(u_n) \rightarrow \infty. \quad (3.58)$$

那么

$$P(M_n \leq u_n \text{ i.o.}) = 0 \text{ 或 } = 1$$

决定于

$$\sum_{n=1}^{\infty} \bar{F}(u_n) \exp \{-n\bar{F}(u_n)\} < \infty \text{ 或 } = \infty. \quad (3.59)$$

证明几乎处处收敛：上确界和下确界

Corollary 3.5.3: Characterisation of the upper and lower limits of maxima

(a) 假设序列 $u_n(\epsilon) = c_n(1 + \epsilon) + d_n, n \in \mathbb{N}$ 对于所有 $\epsilon \in (-\epsilon_0, \epsilon_0)$ 非减。那么关系

$$\sum_{n=1}^{\infty} \bar{F}(u_n(\epsilon)) < \infty \text{ 或 } = \infty$$

决定于 $\epsilon \in (0, \epsilon_0)$ 或 $\epsilon \in (-\epsilon_0, 0)$ 。这表示：

$$\limsup_{n \rightarrow \infty} c_n^{-1} (M_n - d_n) = 1 \quad \text{a.s.}$$

证明几乎处处收敛：上确界和下确界 (cont'd)

Corollary 3.5.3: Characterisation of the upper and lower limits of maxima

(b) 假设序列 $u_n(\epsilon) = c_n(1 + \epsilon) + d_n, n \in \mathbb{N}$ 是非减序列且对于所有 $\epsilon \in (-\epsilon_0, \epsilon_0)$ 满足 (3.57), (3.58)。那么下列关系

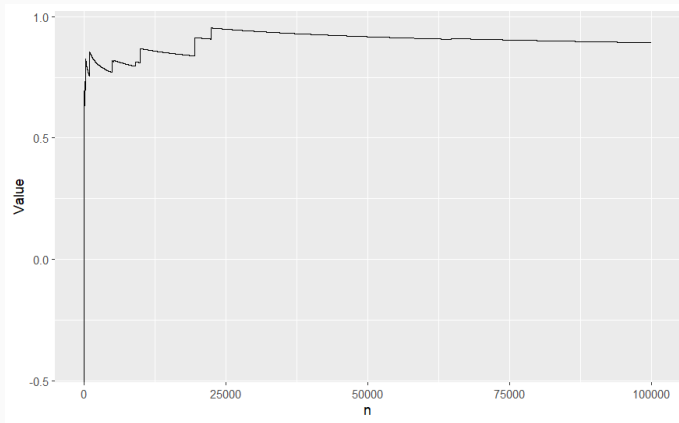
$$\sum_{n=1}^{\infty} \bar{F}(u_n(\epsilon)) \exp \{-n \bar{F}(u_n(\epsilon))\} < \infty \text{ 或 } = \infty$$

决定于 $\epsilon \in (-\epsilon_0, 0)$ 或 $\epsilon \in (0, \epsilon_0)$ 。这表示：

$$\liminf_{n \rightarrow \infty} c_n^{-1} (M_n - d_n) = 1 \quad \text{a.s.}$$

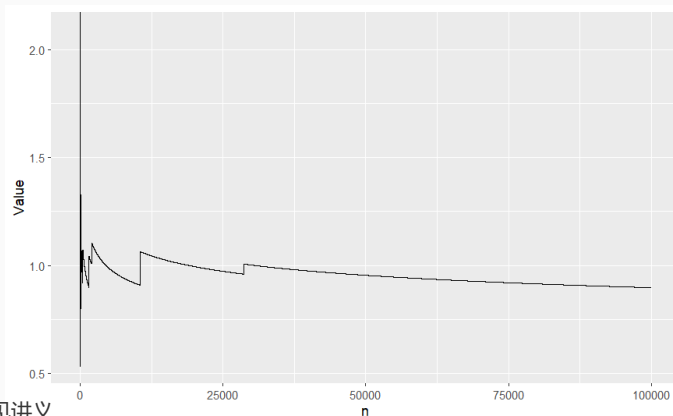
正态分布最大值的几乎处处收敛特性

- $\lim_{n \rightarrow \infty} \frac{M_n}{\sqrt{2 \ln n}} = 1 \quad \text{a.s.}$



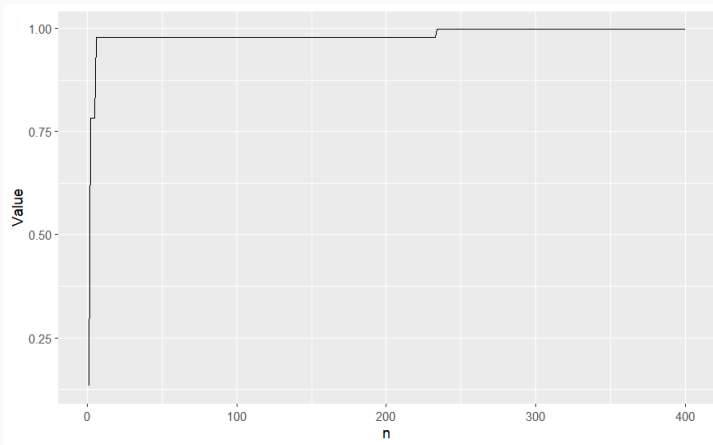
指数尾分布最大值的几乎处处收敛特性

- $\bar{F}(x) \sim K e^{-ax}, \quad x \rightarrow \infty$
- $\lim_{n \rightarrow \infty} \frac{M_n}{\ln n} = \frac{1}{a} \quad \text{a.s.}$



均匀分布最大值的几乎处处收敛特性

- $M_n \xrightarrow{\text{a.s.}} 1$



Thanks!