K15 Extremwerte

- Leistungen im Sport
  — Bob Beamon, Sven Hannawald

- Wetterdaten — Wasserpegel, Schnee, Temperatur

- Bruchfestigkeit

- Finanzmärkte — Risiken, Versicherungen, Aktien

Wir interessieren uns nicht für Maxima/Minima von einer bestimmten Anzahl von Fällen, sondern für Werte, die „extrem“sind. Trotzdem wird das Konzept vom Maximum von unendlich vielen u.i.v. Werten benutzt, um Extremwertverteilungen herzuleiten.

15.1 Die Form einer Extremwertverteilung

Für eine bestimmte Anzahl, n, sei

$$M_n = \mathrm{max}\{X_1, X_2, \ldots, X_n\}$$

$$F_{M_n}(x) = F(x)^n \quad \text{für} \quad \{X_i\} \quad \text{u.i.v.}$$

Daraus entsteht eine degenerierte Grenzverteilung

$$\lim_{n\to\infty} F_{M_n}(x) = 1 \quad \text{für} \quad F(x) = 1 \quad \text{und sonst} \quad 0$$

Fréchet hat in 1927 einen Ausweg vorgeschlagen. Wenn es eine Grenzverteilung, $G(x)$, gibt, sollte diese dieselbe Form für die Maxima der N Untergruppen der Größe n als auch für das Maximum von allen Nn Werten haben. Man muss aber die Lage- und Skalierungs-faktoren berücksichtigen.

Betrachten wir das Maximum aus den Nn Zufallsvariablen $X_1, X_2, \ldots, X_{Nn}$. Das wird auch das Maximum von den folgenden N Zufallsvariablen sein

$$\mathrm{max}(X_{(j-1)n+1}, \ldots, X_{jn}) \quad j = 1, 2, \ldots, N$$

Daraus folgt, dass

$$G(x)^N = G(a_N x + b_N)$$

Wenn wir annehmen, dass $a_N = 1$, finden wir, dass

$$\begin{aligned} G(x)^N &= G(x + b_N) \\ G(x)^{Nr} &= G(x + b_N)^r \\ &= G(x + b_N + b_r) \\ &= G(x + b_{Nr}) \end{aligned}$$

so dass

$$b_N + b_r = b_{Nr}$$

$$\Rightarrow b_N = c \log N \qquad c \text{ konstant}$$

$$G(x) < 1 \Rightarrow G(x)^N < G(y) \qquad \forall y > x$$

$$\Rightarrow c < 0$$

Sei $c = -\theta$ und $\theta > 0$

$$G(x)^N = G(x - \theta \log N)$$

$$N \log G(x) = \log G(x - \theta \log N)$$

$$\log N + \log(-\log G(x)) = \log(-\log G(x - \theta \log N))$$

$$(-\log \text{ weil } G(x) \leq 1)$$

Setzen wir $h(x) = \log(-\log G(x))$

$$h(x) - h(x - \theta \log N) = -\log N$$

$$\Rightarrow h(x) = h(0) - \frac{x}{\theta}$$

$$\Rightarrow -\log G(x) = \exp\left(-\frac{x - \theta h(0)}{\theta}\right)$$

und daher

$$G(x) = \exp\left[-e^{-\frac{x-\xi}{\theta}}\right]$$

die Gumbel-Extremwertverteilung.

$$E[X] = \xi + \gamma\theta$$

wo γ ist der Euler Konstant:

$$\gamma = \lim_{n\to\infty}(\sum_{i=1}^{n} \frac{1}{n} - \log n) = 0.577216...$$

$$V[X] = \frac{\pi^2}{6}\theta^2$$

Für $\xi = 0$ und $\theta = 1$ hat die Gumbelverteilung die standardisierte Form

$G(x) = \exp\left(-e^{-x}\right)$ mit $E[X] = 0.57722, V[X] = 1.645$

15.2 Extremwertverteilungen im allgemeinen

Es gibt 3 Typen

Typ 1 Gumbel (wie gerade besprochen)

$$G(x) = \exp\left[-e^{-\frac{x-\xi}{\theta}}\right] \quad -\infty < x < \infty$$

Typ 2 Fréchet

$$\begin{aligned} G(x) &= 0 & x < \xi \\ &= \exp\left[-\left(\frac{x-\xi}{\theta}\right)^{-\alpha}\right] & x \geq \xi \quad (\alpha > 0) \end{aligned}$$

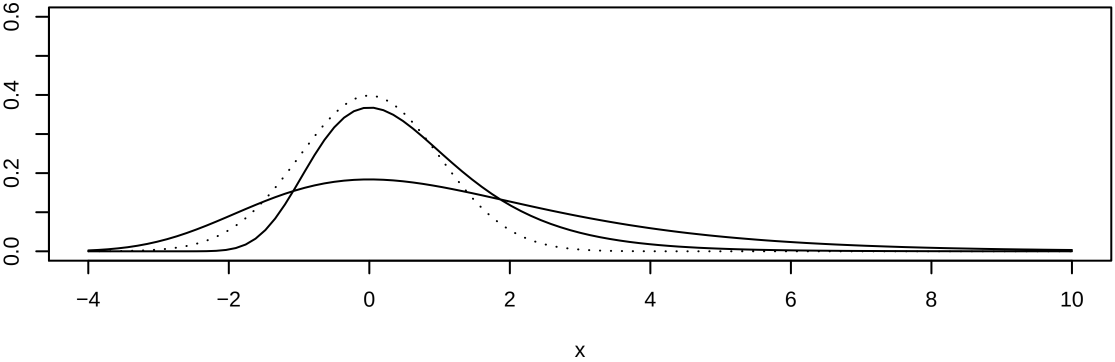
Typ 3 „Weibull"

$$\begin{aligned} G(x) &= \exp\left[-\left(\frac{\xi - x}{\theta}\right)^{\alpha}\right] & x \leq \xi \quad (\alpha > 0) \\ &= 1 & x > \xi \end{aligned}$$

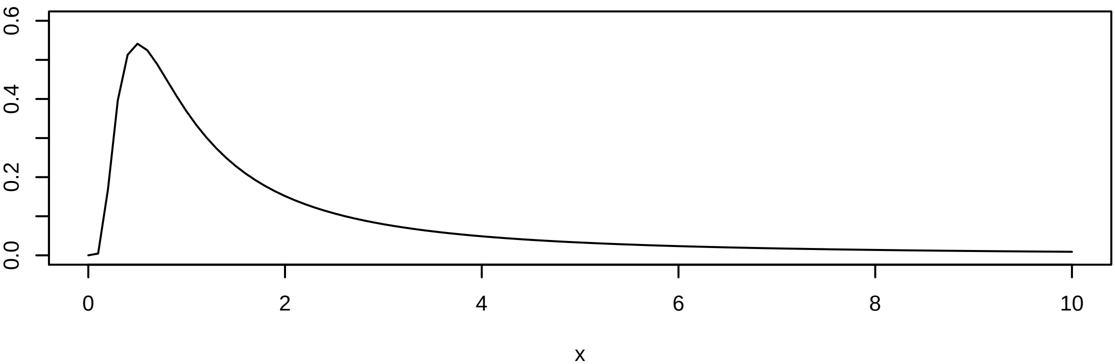
(Eigentlich hat $-X$ die Weibullverteilung.)

Alle Typen sind engverwandt. Die Fréchetverteilung kann durch $Z = \log(X - \xi)$ in eine Gumbelverteilung transformiert werden und die Weibullverteilung in eine Gumbelverteilung durch $Z = -\log(\xi - X)$.

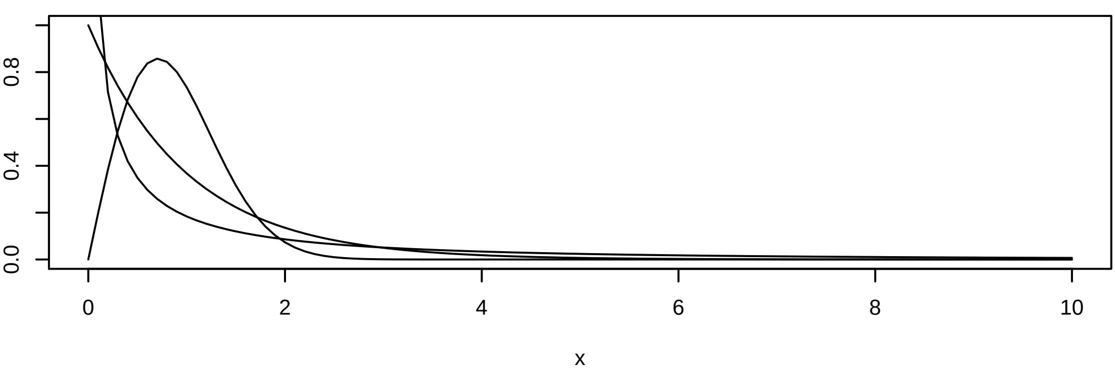
Extremwertverteilungsdichten: Gumbel(scale=1,2) und eine N(0,1)



Extremwertverteilungsdichten: Frechet



Extremwertverteilungsdichten: Weibull (shape=0.5,1,2)



15.3 Die allgemeine Extremwertverteilung

Alle drei Typen sind Mitglieder der allgemeinn Familie:

$$G(x) = \exp\left[-\left(1 + \frac{1}{\nu}\left(\frac{x-\xi}{\theta}\right)\right)^{-\nu}\right]$$

für $1 + \frac{1}{\nu}\left(\frac{x-\xi}{\theta}\right) > 0$

$$\nu \to \infty \Rightarrow \text{ Typ 1} \qquad -\infty < x < \infty$$

$$\nu > 0 \Rightarrow \text{ Typ 2} \qquad \xi - \theta\nu \le x < \infty$$

$$\nu < 0 \Rightarrow \text{ Typ 3} \qquad -\infty < x \le \xi - \theta\nu$$

Typ 2 und Typ 3 kann man aus der Gleichung

$$G(x)^N = G(a_N x + b_N)$$

gewinnen, wenn man für $a_N \neq 1$

$$x = a_N x + b_N$$

betrachtet. Dann muss entweder

$$G(b_N(1-a_N)^{-1}) = 1 \quad \text{Typ 2}$$

oder

$$G(b_N(1-a_N)^{-1}) = 0 \quad \text{Typ 3}$$

15.4 Resultate von Gnedenko

Die Form der Grenzverteilung hängt von der Form der Basisverteilung ab. Für Typ 1 muss gelten, wo

$$F(X_\alpha) = \alpha$$

$$\lim_{n\to\infty} n[(1-F(X_{1-n^{-1}}+y(X_{1-(ne)^{-1}}-X_{1-n^{-1}}))] = e^{-y}$$

Das gilt für die Normal- und Exponential-verteilungen.

Für Typ 2 muss gelten

$$\lim_{x\to\infty} \frac{1-F(x)}{1-F(cx)} = c^k \quad c > 0 \quad k > 0$$

Das gilt für die Cauchyverteilung.