Distribuição BGEV

Seja $Y \sim GEV(\xi, \mu, \sigma)$. Assim, a função de distribuição acumulada de Y é dada por

$$F(y;\xi,\mu,\sigma) = \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, & \xi \neq 0 \\ \exp\left\{-\exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}, & \xi = 0, \end{cases}$$

$\xi \in \mathbb{R}$
$\mu \in \mathbb{R}$
$\sigma > 0$

desde que para $\xi \neq 0$ é necessário ter $1 + \xi\left(\frac{y-\mu}{\sigma}\right) > 0$, e para $\xi = 0$ tem-se $y \in \mathbb{R}$.

Para definir a distribuição GEV bimodal (BGEV), considere a seguinte transformação na variável Y:

$$T_{\mu,\delta}(Y) = (Y-\mu)|Y-\mu|^{\delta}, \quad \delta > -1, \ \mu \in \mathbb{R}.$$

Daí, a distribuição BGEV é obtida a partir da avaliação da função de distribuição da $GEV(\xi, 0, \sigma)$ no ponto $T_{\mu,\delta}(y)$, isto é,

$$F(T_{\mu,\delta}(y);\xi,0,\sigma,\delta) =$$

(2)

$$= \begin{cases} \exp\left\{-\left[1+\xi\left(\frac{T_{\mu,\delta}(y)}{\sigma}\right)\right]^{-1/\xi}\right\}, & \xi \neq 0 \\ \exp\left\{-\exp\left[-\left(\frac{T_{\mu,\delta}(y)}{\sigma}\right)\right]\right\}, & \xi = 0, \end{cases}$$

notação $Y \sim BGEV(\xi, \mu, \sigma, \delta)$

desde que

$$\begin{cases} y \in \left[\mu - \left(\frac{\sigma}{\xi}\right)^{\frac{1}{\delta+1}}, \infty\right), & \text{se } \xi > 0, \\ y \in \left(-\infty, \mu + \left|\frac{\sigma}{\xi}\right|^{\frac{1}{\delta+1}}\right], & \text{se } \xi < 0, \\ y \in \mathbb{R}, & \text{se } \xi = 0. \end{cases}$$

Mostra-se que a função quantílica de $Y \sim BGEV(\xi, \mu, \sigma, \delta)$ é dada por

$$Q(\tau; \xi, \mu, \sigma, \delta) = \text{sinal}(-\ln(-\ln(\tau))) \cdot \left|-\sigma \ln(-\ln(\tau))\right|^{\frac{1}{\delta+1}} + \mu,$$

em que $0 < \tau < 1$.

A partir da função quantílica, obtemos que

$$\mu = q_\tau - \text{sinal}(-\ln(-\ln(\tau)))\left|-\sigma \ln(-\ln(\tau))\right|^{\frac{1}{\delta+1}},$$

em que q_τ denota o τ-ésimo quantil da dis-
-tribuição BGEV.

notação

Se $\tau = 0,5$, então $m = q_{0,5}$ é a mediana da
distribuição BGEV. Lembre-se que μ é apenas um parâmetro
de localização, e portanto não representa a média ou
a mediana da distribuição. Nesse sentido, para
obtermos a modelagem de dados com uma interpre-
-tação direta se torna interessante reparametrizar
a distribuição BGEV em termos de sua mediana.

Veja que se $\tau = 0,5$, então

$$\mu = m - \text{sinal}\Big(\overbrace{-\ln(-\ln(0,5))}^{0,3665>0}\Big)\Big|\overbrace{-\sigma \ln(-\ln(0,5))}^{\sigma * 0,3665>0}\Big|^{\frac{1}{\delta+1}}$$

$$\mu = m - \left(-\sigma \ln(-\ln(0,5))\right)^{\frac{1}{\delta+1}},$$

$\delta+1 > -1+1 \Rightarrow \delta+1 > 0.$

com $\sigma > 0$ e $\delta > -1$. Daí, tem-se que

mediana depende de μ, σ e δ.

$$m = \mu + \Big(\overbrace{-\sigma \ln(-\ln(0,5))}^{>0}\Big)^{\frac{1}{\delta+1}}, \quad m \in \mathbb{R}.$$

Logo, quanto maior for μ maior será a mediana m.

A função de distribuição da BGEV reparametri-
-zada fica dada por

$$F(T_{m,\delta}(y);\xi,0,\sigma,\delta)=\begin{cases}\exp\left\{-\left[1+\xi\left(\frac{T_{m,\delta}(y)}{\sigma}\right)\right]^{-1/\xi}\right\}, & \xi\neq 0\\[2ex] \exp\left\{-\exp\left[-\left(\frac{T_{m,\delta}(y)}{\sigma}\right)\right]\right\}, & \xi=0,\end{cases}$$

com o suporte tendo a mesma expressão, porém substi-
-tuindo μ por m. Denotaremos por
$Y\sim BGEV(\xi, m, \sigma, \delta)$.

Regressão

Considere n realizações independentes de tal forma que
$Y_i \overset{ind}{\sim} BGEV(\xi, m_i, \sigma, \delta)$, $i=1,2,\dots,n$, e
suponha a seguinte estrutura de regressão para
a mediana:

$$g(m_i)=\eta_i=\sum_{j=1}^{p} x_{ij}\beta_j=\underset{\sim}{x}_i^{T}\underset{\sim}{\beta},$$

em que $\underset{\sim}{\beta}=(\beta_1\ \beta_2 \dots \beta_k)^T\in\mathbb{R}^p$ é um vetor de
parâmetros desconhecidos associado à mediana m,

$\underset{\sim}{X}_i = (x_{i1} \; x_{i2} \cdots x_{ip})^T \in \mathbb{R}^p$ é o vetor de valores conhecidos das p variáveis explicativas para a i-ésima observação, e $g(\cdot)$ é uma função de ligação contínua, estritamente monó-tona e duas vezes diferenciável.

Neste cenário, as funções de ligação mais uti-lizadas são

1) Identidade: $g(z) = z$, $z \in \mathbb{R}$.

2) logarítmica: $g(z) = \underset{\ln(z)}{\log(z)}$, $z > 0$.

Estimação

O vetor da regressão $\underset{\sim}{\beta} = (\beta_1 \ldots \beta_p)$ e os parâmetros ξ, σ e δ são desconhecidos e devem ser estimados. Para tanto, pode-se utilizar o método de máxima verossimilhança.

O logaritmo da função de verossimilhança para o vetor de parâmetros $\underset{\sim}{\theta} = (\underset{\sim}{\beta}^T, \xi, \sigma, \delta) \in \mathbb{R}^{p+3}$ é dado por

$l_i(\underset{\sim}{\theta})$ → logaritmo da densidade BGEV(ξ, m_i, σ, δ)

$$l(\underset{\sim}{\theta}) = \sum_{i=1}^{n} \overbrace{\log\left[f(T_{m_i,\delta}(y_i); \xi, 0, \delta)\right]}$$

em que $f(T_{mi,\delta}(y_i); \xi, 0, \delta)$ é ⑥

- para $\xi \neq 0$:

$$\frac{1}{\sigma}\left[1+\xi\left(\frac{T_{mi,\delta}(y_i)}{\sigma}\right)\right]^{-\frac{1}{\xi}-1} \exp\left\{-\left[1+\xi\left(\frac{T_{mi,\delta}(y_i)}{\sigma}\right)\right]^{-1/\xi}\right\} T'_{mi,\delta}(y_i)$$

- para $\xi = 0$:

$$\frac{1}{\sigma}\exp\left\{-\frac{T_{mi,\delta}(y_i)}{\sigma}\right\} \exp\left\{-\exp\left\{-\frac{T_{mi,\delta}(y_i)}{\sigma}\right\}\right\} T'_{mi,\delta}(y_i)$$

Daí, segue que $l_i(\underset{\sim}{\theta})$ é dada por

- $\xi \neq 0$:

$$-\log(\sigma) - \left(\frac{1}{\xi}+1\right)\log\left[1+\xi\left(\frac{T_{mi,\delta}(y_i)}{\sigma}\right)\right]$$
$$-\left[1+\xi\left(\frac{T_{mi,\delta}(y_i)}{\sigma}\right)\right]^{-1/\xi} + \log\left[T'_{mi,\delta}(y_i)\right].$$

- $\xi = 0$:

$$-\log(\sigma) - \frac{T_{mi,\delta}(y_i)}{\sigma} - \exp\left\{-\frac{T_{mi,\delta}(y_i)}{\sigma}\right\}$$
$$+\log\left(T'_{mi,\delta}\right).$$

Assim, o logaritmo da função de verossimilhança para $\underset{\sim}{\theta}$ é

7

- $\xi \neq 0$:

$$\ell(\underset{\sim}{\theta}) = -n \log(\sigma) - \left(\frac{1}{\xi}+1\right)\sum_{i=1}^{n} \log\left[1+\xi\left(\frac{T_{m_i,\delta}(y_i)}{\sigma}\right)\right]$$

$$+ \sum_{i=1}^{n}\left\{ \log\left[T'_{m_i,\delta}(y_i)\right] - \left[1+\xi\left(\frac{T_{m_i,\delta}(y_i)}{\sigma}\right)\right]^{-1/\xi}\right.$$

- $\xi = 0$:

$$\ell(\underset{\sim}{\theta}) = -n \log(\sigma) + \sum_{i=1}^{n}\left\{\log\left[T'_{m_i,\delta}(y_i)\right] - \frac{T_{m_i,\delta}(y_i)}{\sigma} - \exp\left\{-\frac{T_{m_i,\delta}(y_i)}{\sigma}\right\}\right\}.$$

Próximo passo é obter o vetor escore: