

## APLICAÇÃO DA METODOLOGIA *PEAKS OVER THRESHOLD* AO SALTO EM COMPRIMENTO DO ATLETISMO

Domingos SILVA<sup>1</sup>  
Frederico CAEIRO<sup>2</sup>  
Manuela OLIVEIRA<sup>1</sup>

- RESUMO: O método *Peaks Over Threshold* (POT) consiste no ajuste da distribuição generalizada de Pareto aos excessos aleatórios independentes e identicamente distribuídos acima de um limiar elevado. Neste estudo, usamos a metodologia POT para estimar os parâmetros da distribuição dos excessos na aplicação ao salto em comprimento do atletismo, no período de 1935 a 2016 nos homens e de 1964 a 2016 nas mulheres. Determinado o limiar  $u$ , adequado, são apresentadas as estimativas de alguns parâmetros de interesse, como a probabilidade de excedência, os quantis extremos, o limite superior do suporte e os níveis e períodos de retorno. Nos homens, os resultados apontam para  $u=8,45\text{m}$  com  $\hat{\gamma} > 0$  e nas mulheres  $u = 7,06\text{m}$  com  $\hat{\gamma} < 0$ . Prevê-se que o recorde mundial seja batido, em média, acima de 585 recordes individuais nos homens, e acima de 427 recordes individuais nas mulheres.
- PALAVRAS-CHAVE: Teoria de valores extremos; probabilidade de excedência; quantis extremos; limite superior do suporte; níveis de retorno.

### 1 Introdução

O salto em comprimento é uma especialidade do atletismo, cujo objetivo é saltar o mais longe possível a partir de um ponto determinado (tábua de chamada). Na antiguidade, o salto em comprimento surgiu como prática de preparação para a guerra, uma vez que os soldados eram muitas vezes obrigados a ter que ultrapassar obstáculos como os riachos e ravinas. Enquanto especialidade do atletismo, foi introduzida nos primeiros Jogos Olímpicos da Era Moderna em 1896 (Atenas) para os homens e em 1948 (Londres) para as mulheres. São fatores determinantes para um bom salto em comprimento, a velocidade de deslocamento, a força explosiva, a resistência, a coordenação e a agilidade específicas (SCHMOLINSKY, 1982).

Desde os primeiros Jogos Olímpicos da Era Moderna em Atenas, 1896, o recorde mundial dos homens aumentou cerca de 29%, e o das mulheres cerca de 24% (Tabela 1). Porém, há mais de 25 anos que o recorde mundial estabilizou. De uma forma geral, as melhores marcas de sempre ocorreram nas décadas de 80 e 90 do Séc. XX. No Séc. XXI, a

---

<sup>1</sup>Universidade de Évora, Escola de Ciências e Tecnologia, Departamento de Matemática, Portugal. E-mail: [domingosjsilva@gmail.com](mailto:domingosjsilva@gmail.com); [mno@uevora.pt](mailto:mno@uevora.pt)

<sup>2</sup>Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia, Departamento de Matemática, Portugal. E-mail: [fac@fct.unl.pt](mailto:fac@fct.unl.pt)

melhor marca dos homens (8,74m em 2009) surge em 7º lugar das melhores de sempre, enquanto que nas mulheres (7,42m em 2002) surge em 5º lugar. Se considerarmos apenas os últimos 5 anos, a melhor marca dos homens surge em 13º lugar (8,58m em 2016) e das mulheres em 11º lugar (7,31m em 2016).

Tabela 1- Primeiro recorde mundial e atuais recordes mundial e olímpico do salto em comprimento masculino e feminino

Atleta	Marca	Ano	Obs.
Ellery Clark (EUA)	6,35	1896	1º recorde mundial (masc)
Mike Powell (EUA)	8,95	1991	Atual recorde mundial (masc)
Bob Beamon (EUA)	8,90	1968	Atual recorde olímpico (masc)
Olga Gyarmati (HUN)	5,695	1948	1º recorde mundial (fem)
Galina Chistyakova (URSS)	7,52	1988	Atual recorde mundial (fem)
Jackie Joyner-Kersey (EUA)	7,40	1988	Atual recorde olímpico (fem)

Fonte: Wikipedia

Quando em 1988, Galina Chistyakova (URSS), e em 1991, Mike Powell (EUA) bateram o recorde mundial, apesar do notável feito, estávamos longe de imaginar que tal proeza iria manter-se por praticamente 30 anos, e que parece querer perpetuar-se tais têm sido as marcas obtidas nas principais competições internacionais (Campeonatos do Mundo e Jogos Olímpicos) dos últimos anos (Tabela 2). Ora, isto faz exatamente jus à definição da Teoria de Valores Extremos (EVT) enquanto resultado de extremos de grande magnitude, grande impacto e de muito baixa probabilidade de ocorrência.

Tabela 2 - Marca obtida na obtenção da medalha de ouro no salto em comprimento (masculino e feminino) nos últimos Campeonatos do Mundo e Jogos Olímpicos

Campeonatos do Mundo			Jogos Olímpicos			
	Homens	Mulheres		Homens	Mulheres	
	2015	8,41	7,14	2016	8,38	7,17
	2013	8,56	7,01	2012	8,31	7,12
	2011	8,45	6,82	2008	8,34	7,04
	2009	8,54	7,10	2004	8,59	7,07
	2007	8,57	7,03	2000	8,55	6,99
	2005	8,60	6,89	1996	8,50	7,12
	2003	8,32	6,99	1992	8,67	7,08
	2001	8,40	7,02	1988	8,72	7,40

Fonte: Wikipedia: URL: [https://en.wikipedia.org/wiki/Long\\_jump](https://en.wikipedia.org/wiki/Long_jump)

A teoria de valores extremos (“*Extreme Value Theory*”) é um ramo da teoria da probabilidade e da estatística matemática cujo foco é a análise e inferência de eventos extremos, i.e., eventos com muito baixa probabilidade de ocorrência, como por exemplo o novo recorde mundial do salto em comprimento. Para ajudar a perceber este assunto, dentro da EVT, recorreremos à distribuição generalizada de Pareto (GPD, “*generalized Pareto distribution*”) e à metodologia *Peaks Over Threshold* (POT), a qual, apoiada pelo Teorema de Pickands-Balkema-de Haan, é adequada para ajustar os excessos aleatórios

independentes e identicamente distribuídos (i.i.d.'s) (ou relativamente estacionários) acima de um limiar  $u$  suficientemente alto.

A seleção do limiar, bem como o seu procedimento não são temas consensuais. As dificuldades existem e a problemática está em aberto. De qualquer forma, procura-se sempre o melhor compromisso entre a redução do viés e a menor variância dos estimadores (EMBRECHTS *et al.*, 1997; COLES, 2001; BEIRLANT *et al.*, 2004). Existem vários métodos para a determinação do limiar. Alguns dos mais utilizados são o gráfico da vida média residual e o gráfico de estabilidade dos parâmetros (HALL e WELLNER, 1981; DAVISON e SMITH, 1990; EMBRECHTS *et al.*, 1997; COLES, 2001; BEIRLANT *et al.*, 2004; GHOSH e RESNICK, 2010; GHOSH e RESNICK, 2011).

Com este estudo pretendemos utilizar a GPD e o método POT com aplicação às melhores performances de sempre no salto em comprimento, para modelar os excessos acima de um nível elevado e apresentar as estimativas de alguns parâmetros de interesse, designadamente a probabilidade de excedência de um nível  $x$  elevado, alguns quantis extremos, o limite superior do suporte ( $\gamma < 0$ ) e os níveis de retorno.

## 2 Metodologia

Assumindo que existem suficientes observações acima de um certo limiar  $u$ , procuramos uma distribuição condicional apropriada para estas excedências. Seja  $X = (X_1, X_2, \dots, X_n)$  uma sequência de variáveis aleatórias (v.a.'s) i.i.d.'s com função de distribuição (f.d.)  $F$  comum, desconhecida, sendo  $(X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n})$  as estatísticas ordinais (e.o.'s) correspondentes, ordenadas crescentemente, onde  $X_{1:n}$  denota o valor mínimo e  $X_{n:n}$  o valor máximo, sendo  $u$  um limiar alto próximo de  $X_{n:n}$ . O interesse recai unicamente na amostra dos excessos, dada pela f.d.  $(X - u)$  condicional em  $X > u$ ,

$$Y_j = X_i - u \mid X_i > u, \quad j = 1, 2, \dots, N_u,$$

em que  $N_u$  denota o número de observações da amostra original que excedem o limiar  $u$ ,  $N_u = \{i : X_i > u, i = 1, 2, \dots, n\}$ , com  $N_u$  suficientemente grande. A esta nova amostra aleatória poderá ser ajustada a GPD,  $H_{\gamma, \sigma_u}(y)$ .

### 2.1 Distribuição condicional dos excessos

Seja  $(X_1, X_2, \dots, X_n)$  uma sequência de v.a.'s, i.i.d.'s, com f.d.  $F$  desconhecida, a função de distribuição condicional dos excessos acima do nível  $u$ , representada por  $F_u$ , é definida como,

$$F_u(y) = P[X - u \leq y \mid X > u] = \frac{P(u < X \leq y + u)}{P(X > u)} \quad (1)$$

em que  $y = x - u$  são os excessos, com  $0 \leq y \leq x^F - u$ , em que  $x^F := \sup\{x : F(x) < 1\}$  é o limite superior do suporte de  $F$ . Verificamos que  $F_u$  pode ser escrita em termos de  $F$ , da seguinte forma,

$$F_u(y) = \frac{F(y+u) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)}. \quad (2)$$

Isolando  $F(x)$  em [2], tem-se

$$F(x) = [1 - F(u)]F_u(y) + F(u). \quad (3)$$

Os resultados em (1), (2) e (3) dependem de uma f.d. que é desconhecida, pelo que têm pouco interesse. Assim, na procura de um modelo apropriado para a cauda direita sob a f.d.  $F$ , Balkema e de Haan (1974) e Pickands (1975) provaram que acima um limiar  $u$  suficientemente alto, dado que  $u$  é excedido, e sob condições adequadas,  $F_u(y)$  e  $H_{\gamma, \sigma_u}(y)$  são assintoticamente equivalentes, à medida que  $u$  se aproxima do  $x^F$ ,

$$F_u(y) \approx H_{\gamma, \sigma_u}(y), \quad u \rightarrow x^F \quad (4)$$

para algum  $\gamma$  e  $\sigma_u$ , com  $y \in [0, x^F - u]$  se  $\gamma \geq 0$ , ou  $y \in [0, -\frac{\sigma_u}{\gamma}]$  se  $\gamma < 0$ .

A aproximação (4) é enunciada no Teorema de Pickands-Balkema-de Haan:

$F \in D(G_\gamma)$ ,  $\gamma \in \mathfrak{R}$  se e só se,

$$\lim_{u \rightarrow x^F} \sup_{0 \leq y < x^F - u} |F_u(y) - H_{\gamma, \sigma_u}(y)| = 0, \quad (5)$$

para alguma distribuição da família generalizada de Pareto com parâmetros de forma  $\gamma$  e escala  $\sigma_u > 0$ .

Se (5) for assegurado, então dizemos que a f.d.  $F$  pertence ao domínio de atração POT-GPD. Informações detalhadas sobre domínios de atração constam em Embrechts *et al.* (1997) e Beirlant *et al.* (2004). Balkema e de Haan (1974) e Pickands (1975) provaram, assim, que a GPD é a distribuição limite para os excessos acima de um limiar  $u$ .

## 2.2 Distribuição generalizada de Pareto

A f.d. da GPD tem a seguinte forma funcional,

$$H_{\gamma,u,\sigma_u}(x) := \begin{cases} 1 - \left(1 + \gamma \left(\frac{x-u}{\sigma_u}\right)\right)_+^{-\frac{1}{\gamma}}, & x \geq 0, & \text{se } \gamma > 0 \\ 1 - \exp\left(-\frac{x-u}{\sigma_u}\right), & x \geq 0, & \text{se } \gamma = 0 \\ 1 - \left(1 + \gamma \left(\frac{x-u}{\sigma_u}\right)\right)_+^{-\frac{1}{\gamma}}, & 0 \leq x \leq -\frac{\sigma_u}{\gamma}, & \text{se } \gamma < 0 \end{cases} \quad (6)$$

em que  $u$  é o limiar (*threshold*),  $x > u$ ,  $1 + \gamma(x-u)/\sigma_u \geq 0$ ,  $\sigma_u$  é o parâmetro escala ( $\sigma_u > 0$ ) correspondente ao nível  $u$  selecionado,  $\gamma$  é o parâmetro forma ( $\gamma \in \mathfrak{R}$ ) e  $a_+ = \max(a, 0)$ .

Tal como na distribuição generalizada de valores extremos, também na GPD o parâmetro de forma  $\gamma$  ou índice de valor extremo (EVI, “*extreme value index*”) está diretamente relacionado com o “peso” da cauda direita subjacente à f.d.  $F$ . Os mais importantes submodelos são:

- (i)  $\gamma = 0$  ( $\gamma \rightarrow 0$ ), a GPD reduz-se à f.d. Exponencial, pelo que  $x^F < \infty$ .
- (ii)  $\gamma > 0$ , a GPD reduz-se à f.d. Pareto (versão reparametrizada), pelo que  $x^F = \infty$
- (iii)  $\gamma < 0$ , a GPD reduz-se à f.d. Beta, pelo que  $x^F = u - \frac{\sigma_u}{\gamma}$

A média e a variância da GPD são dadas por,

$$E[X] = \begin{cases} u + \frac{\sigma_u}{1-\gamma} & (\gamma < 1) \\ \infty & (\gamma \geq 1) \end{cases} \quad (7)$$

$$\text{Var}[X] = \begin{cases} \frac{\sigma_u^2}{(1-\gamma)^2(1-2\gamma)} & (\gamma < 0.5) \\ \infty & (\gamma \geq 0.5) \end{cases} \quad (8)$$

Quando  $\gamma < 1$ , um aumento de  $\sigma_u$ , desde que não se altere o valor de  $\gamma$ , traduz-se num aumento do valor esperado; quando  $\gamma < 0$ ,  $E[X] < \sigma_u$ ; quando  $0 < \sigma_u < 1$ ,  $E[X] > \sigma_u$  desde que  $0 < \gamma < 1$ .

### 2.3. Estimação dos parâmetros

Os parâmetros da GPD (6) podem ser estimados por diferentes métodos. No ajuste da GPD para os excessos acima do nível  $u$ , Smith (1984), Davison (1984), Hosking e Wallis (1987), Embrechts *et al.* (1997) e Beirlant *et al.* (2004) referem que as propriedades estandardizadas pelo método de estimação por máxima verossimilhança (MLE, “*maximum likelihood estimation*”) podem ser aplicadas quando  $\gamma > -0.5$ . Portanto, a função log-verossimilhança para a amostra  $y_1, \dots, y_{N_u}$  de v.a.’s i.i.d.’s com GPD, conforme o comportamento de  $\gamma$ , é dada por,

(i) Para  $\gamma = 0$ ,

$$\begin{aligned} \ell(0, \sigma_u | y_1, \dots, y_{N_u}) &= \log L(0, \sigma_u | y_1, \dots, y_{N_u}) = \\ &= -N_u \log \sigma_u - \frac{1}{\sigma_u} \sum_{i=1}^{N_u} y_i \end{aligned} \quad (9)$$

(ii) Para  $\gamma \neq 0$ ,

$$\begin{aligned} \ell(\gamma, \sigma_u | y_1, \dots, y_{N_u}) &= \log L(\gamma, \sigma_u | y_1, \dots, y_{N_u}) = \\ &= -N_u \log \sigma_u - \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^{N_u} \log \left(1 + \frac{\gamma y_i}{\sigma_u}\right), \end{aligned} \quad (10)$$

para  $1 + \frac{\gamma y_i}{\sigma_u} > 0$ ,  $i = 1, \dots, N_u$ ; caso contrário  $\ell(\gamma, \sigma_u) = -\infty$ .

Uma vez que para as equações (9) e (10) não existe solução analítica, a maximização pode ser feita recorrendo a algoritmos numéricos. Neste caso, os métodos de otimização de Nelder-Mead (NELDER e MEAD, 1965) ou BFGS (LEWIS e OVERTON, 2013), ambos incorporado no pacote *ismev* (HEFFERNAN e STEPHENSON, 2016) do R (R CORE TEAM, 2017), podem ser boas opções. Os erros-padrão (*se*, “*standard error*”) podem ser obtidos pelo método delta e os intervalos de confiança (IC) para  $(\gamma, \sigma_u)$  por aproximação à Normal (COLES, 2001).

### 2.4 Outros parâmetros de interesse

Para além de  $(\hat{\gamma}, \hat{\sigma}_u)$ , importa também estimar outros parâmetros de eventos extremos, nomeadamente, os quantis extremos, a probabilidade de excedência, o limite superior do suporte ( $\hat{\gamma} < 0$ ) e os valores de retorno.

Seja  $F$  uma f.d. contínua, denotando-se por  $F^{\leftarrow}(\cdot)$  a respectiva inversa generalizada, que se define por,

$$F^{\leftarrow}(y) = \inf\{x: F(x) \geq y\} \quad (11)$$

e  $U(\cdot)$  a função quantil de cauda empírica da v.a.  $X$ , definida por,

$$U(t) := F^{\leftarrow}(1-1/t) = \inf\{x: F(x) \geq 1-1/t\}, \quad t \in [1, \infty[. \quad (12)$$

(i) Quantis extremais

As estimativas dos quantis extremais com probabilidade de excedência  $p$ , com  $p$  suficientemente pequeno, obtêm-se invertendo a f.d. da GPD dada em (6), ou seja,

$$\hat{U}_{H_\gamma}\left(\frac{1}{p}\right) = \begin{cases} \frac{\hat{\sigma}_u}{\hat{\gamma}}(p^{-\hat{\gamma}} - 1), & (\hat{\gamma} \neq 0) \\ -\hat{\sigma}_u \log p, & (\hat{\gamma} = 0) \end{cases} \quad (13)$$

Alternativamente, e que é o que realmente interessa, podemos usar a função quantil de cauda definida em (12), com  $t = 1/p$ , para estimar quantis elevados de  $F$ ,  $U(1/p) = F^{\leftarrow}(1-p)$ , dados por,

$$\hat{U}_{H_\gamma}\left(\frac{1}{p}\right) = \hat{F}^{\leftarrow}(1-p) = \begin{cases} u + \frac{\hat{\sigma}_u}{\hat{\gamma}} \left[ \left( \frac{np}{N_u} \right)^{-\hat{\gamma}} - 1 \right], & \hat{\gamma} \neq 0 \\ u - \hat{\sigma}_u \log \left( \frac{np}{N_u} \right), & \hat{\gamma} = 0 \end{cases} \quad (14)$$

(ii) Probabilidades de excedência

Dada a complexidade, as probabilidades de excedência de um nível  $x$  elevado,  $\hat{F}(x)$ , não podem ser obtidas diretamente da cauda da GPD com os parâmetros estimados, pelo que terão que ser ponderadas pela proporção de excessos de nível na amostra original,  $N_u/n$ , e posteriormente substituir os parâmetros  $(\gamma, \sigma_u)$  pelas respectivas estimativas de máxima verossimilhança  $(\hat{\gamma}, \hat{\sigma}_u)$ , obtendo-se assim um estimador de probabilidade para a cauda direita da distribuição,

$$\hat{\hat{F}}(x) := P(X > x) = \frac{N_u}{n} \left( 1 + \hat{\gamma} \left( \frac{x-u}{\hat{\sigma}_u} \right) \right)^{-\frac{1}{\hat{\gamma}}}, \quad \text{com } x > u \quad (15)$$

(iii) Limite superior do suporte

Quando  $\hat{\gamma} < 0$ , o limite superior do suporte de  $F$  é dado por,

$$\hat{x}^F = \hat{U}(\infty) = u - \frac{\hat{\sigma}_u}{\hat{\gamma}}. \quad (16)$$

(iv) Valores de retorno

Os valores de retorno compreendem os níveis de retorno e o período de retorno. Supondo que dispomos de  $t$  recortes individuais independentes,  $F(U(t)) = 1 - 1/t$  significa que o nível  $U(t)$  (nível de retorno) é excedido em média uma vez em cada  $t$ -recortes individuais (período de retorno), o qual é a solução de,

$$\hat{\lambda}_u \left[ 1 + \hat{\gamma} \left( \frac{\hat{x}_t - u}{\hat{\sigma}_u} \right) \right]_+^{\frac{1}{\hat{\gamma}}} = \frac{1}{t} \quad (17)$$

em que  $\hat{\lambda}_u = N_u / n$  é o rácio empírico de excedência.

Rearranjando (17), o nível de retorno para  $t$ -recortes individuais, com  $t$  suficientemente grande para assegurar que  $\hat{U}(t) > u$ , é dado por,

$$\hat{U}(t) = \begin{cases} u + \frac{\hat{\sigma}_u}{\hat{\gamma}} \left[ (t\hat{\lambda}_u)^{\hat{\gamma}} - 1 \right], & \hat{\gamma} \neq 0 \\ u + \hat{\sigma}_u \ln(t\hat{\lambda}_u), & \hat{\gamma} = 0 \end{cases} \quad (18)$$

Os intervalos de confiança foram construídos pela *profile log-likelihood function*, conforme recomendação de Coles (2001) e Beirlant *et al.* (2004),

$$IC_{100(1-\alpha)\%}(\hat{U}(t)) = \left\{ \hat{U}(t) : \log L_p(U(t)) \geq \log L_p(\hat{U}(t)) - \frac{\chi_{1,1-\alpha}^2}{2} \right\}, \quad (19)$$

com  $\log L_p(U(t)) = \max L(\hat{U}(t))$ . Deve-se notar que geralmente os intervalos de confiança são assimétricos, i.e., a distância entre a estimativa pontual por máxima verossimilhança e os seus limites são diferentes.

No caso de se pretender estimar o período de retorno, o seu estimador é dado por,

$$\hat{T}_x \approx \frac{1}{\hat{F}(x)}. \quad (20)$$

Ainda que não tenha sido alvo de abordagem no presente estudo, decorrente da infrequência do número de observações nos anos de recolha, mas quando os dados estão dispostos por  $t$ -anos, com  $n_y$  observações por ano, a estimativa do nível de retorno para o período de retorno de  $t$ -anos,  $\hat{U}(t)$ , é definida como,

$$\hat{U}(t) = \begin{cases} u + \frac{\hat{\sigma}_u}{\hat{\gamma}} \left[ (t.n_y.\hat{\lambda}_u)^{\hat{\gamma}} - 1 \right], & \hat{\gamma} \neq 0 \\ u + \hat{\sigma}_u \log(t.n_y.\hat{\lambda}_u), & \hat{\gamma} = 0 \end{cases} \quad (21)$$



## 2.5. Seleção do limiar

Analogamente ao problema da seleção do tamanho dos blocos no método dos blocos máximos, um aspecto fundamental no processo de ajuste da GPD é a escolha de  $u$ , que por ser especificado antes de ajustar a GPD (por ser esta a distribuição para os excessos acima de  $u$ ) não pode ser considerado um parâmetro. Um limiar demasiado alto resulta em poucas observações na cauda da distribuição ( $N_u$  pequeno) e conseqüentemente grande variância assintótica na estimação de  $\gamma$  e  $\sigma_u$ ; opostamente, um limiar demasiado baixo ainda que resulte na diminuição da variância, faz com que a inclusão de observações da região central da distribuição origine um grande viés na estimação de  $\gamma$  e  $\sigma_u$ , além de invalidar o Teorema de Pickands-Balkema-de Haan. Portanto, na adequada seleção de  $u$  deverá existir um compromisso entre valores elevados de  $u$  onde o viés dos estimadores é menor e valores pequenos de  $u$  onde a variância é menor.

Existem diversos métodos para determinar  $u$ . Neste estudo, utilizamos o gráfico da função de excesso médio (MEF, “*mean excess function*”), também conhecida como função de vida média residual (*mean residual life function*) e o gráfico de estabilidade dos parâmetros (“*parameter stability plot*”). A MEF é uma ferramenta analítica muito usada quer na ajuda à seleção do nível  $u$  quer na determinação da adequação prática do modelo GPD. Começou por ser estudada por Benktander e Segedahl (1960). Posteriormente, Hall e Wellner (1981) discutiram as suas propriedades (em 2012, estes autores apresentaram uma revisão do seu trabalho). Mais recentemente, Northrop *et al.* (2017) introduziram a seleção do limiar em modelos não-estacionários. Mas foram Davison e Smith (1990) a proporcionarem as bases metodológicas para a escolha conveniente de  $u$ .

Definição: a função de excesso médio,  $e(u)$ , da v.a.  $X$  i.i.d., representa a média condicional dos excessos acima de um limiar  $u$ , dada por,

$$e(u) := E[X - u | X > u], \quad (21)$$

assumindo que para o modelo proposto,  $E[X]_+ < \infty$ .

A MEF é linear com declive positivo se  $0 < \gamma < 1$ , declive negativo se  $\gamma < 0$ , ou declive nulo se  $\gamma = 0$ . Quando  $\gamma \geq 1$ , a média teórica é infinita. Mais precisamente, conforme Coles (2001) e Ghosh e Resnick (2010, 2011), para uma v.a.  $X \sim H_{\gamma, \sigma_u}$ , o  $E(X) < \infty$  se e só se  $\gamma < 1$ , a MEF de  $X$  é linear em  $u_0$  (com declive  $-\gamma/(1+\gamma)$  e ordenada na origem  $\sigma_u/(1+\gamma)$ ), daí que para  $u > u_0$

$$e(u) := E[X - u | X > u] = \frac{\sigma_u}{1-\gamma} = \frac{\sigma_{u_0} + \gamma u}{1-\gamma}, \quad \sigma_{u_0} + \gamma u > 0 \quad (22)$$

com  $0 \leq u < \infty$  se  $0 \leq \gamma < 1$  e  $0 \leq u \leq -\sigma_u/\gamma$  se  $\gamma < 0$ .

De outra maneira pode dizer que a linearidade da MEF caracteriza o modelo GPD. Na prática, a MEF é estimada pela correspondente função empírica  $\hat{e}_{N_u}$ . Explicitamente, a MEF é definida pelo conjunto de pontos,

$$\left\{ \left( \frac{\sum_{i=1}^{N_u} (x_i - u) I_{[X_i > u]}}{\sum_{i=1}^{N_u} I_{[X_i > u]}} = \frac{\sum_{i=1}^{N_u} x_i I_{(u, \infty)}(x_i)}{\sum_{i=1}^{N_u} I_{(u, \infty)}(x_i)} - u : u < x_{\max} \right) \right\}, \quad (23)$$

em que  $x_1, \dots, x_{N_u}$  são as  $N_u$  observações que excedem  $u$ , e  $x_{\max}$  é o valor mais elevado da

amostra observada e  $I_{(u, \infty)}(x_i) = \begin{cases} 1 & \text{se } x_i > u \\ 0 & \text{caso contrário} \end{cases}$ .

Alternativamente,  $\hat{e}_{N_u}$  é representada nos valores  $u = x_{n-k:n}$ ,  $k = 1, \dots, n-1$  (as  $k+1$  maiores observações), pelo que

$$\sum_{i=1}^{N_u} x_i I_{(u, \infty)}(x_i) = \sum_{j=1}^k x_{n-j+1:n}, \text{ com } k \equiv \# x_i : x_i > u.$$

Neste caso, a estimativa da MEF é dada por,

$$\hat{e}_{k,n} := \hat{e}_n(x_{n-k:n}) = \frac{1}{k} \sum_{j=1}^k (x_{n-j+1:n} - x_{n-k:n}). \quad (24)$$

A MEF apresentada em (22) para a GPD é linear em  $u$ , pelo que o gráfico correspondente é obtido através de

$$\left\{ \hat{e}_{n-k,n}, \hat{e}_n(x_{n-k:n}), 1 \leq k \leq n-1 \right\},$$

esperando-se que haja linearidade à direita de  $x_{n-k:n}$ .

Fica assim claro que a MEF é o quociente da soma dos excessos acima de  $u$  pelo número de observações que excedem esse limiar, determinando a média aritmética de valores que excedem  $u$ . A estimativa empírica da MEF é geralmente observada pelo gráfico da função de excesso médio, também conhecido como gráfico de vida média residual (MRL plot, “mean residual life plot”), que é uma representação gráfica de pontos, parecendo-se com uma serra, cujos dentes crescem para a direita. O MRL plot é o gráfico de pontos  $\{(x_{N_u}, \hat{e}_{N_u}) : 1 < N_u \leq n\}$  que permite visualizar a partir de que ponto (limiar) se observa a linearidade, bem como a inclinação do declive (positivo, negativo ou nulo). Se os dados suportam o modelo GPD acima de um limiar alto, então é esperado que o gráfico se torne linear a partir desse ponto. Portanto, o limiar  $u$  deverá ser escolhido, precisamente no ponto a partir do qual se verifica uma certa linearidade à sua direita, podendo esse ponto estar ou não presente nos dados disponíveis. A este propósito, Gosh e Resnick (2010) referem que

a linearidade observada em (23) para as e.o.'s extremas indica que não existe evidência contra a hipótese de considerar a boa adequação do modelo GPD para as excedências. O MRL *plot* tem boa capacidade de análise quando  $-\infty < \gamma < 1$ , e inconsistente quando  $\gamma \geq 1$  (GHOSH e RESNICK, 2010, 2011; DAS e GHOSH, 2013). Pode-se observar que na parte final deste gráfico (cauda direita) não se verifica a linearidade, uma vez que estão a ser calculadas médias de um número reduzido de excessos elevados, pelo que esses pontos deverão ser ignorados (e.g., Figura 2).

De fato, os resultados de Davison e Smith (1990) e de Ghosh e Resnick (2010, 2011) mostram que se a distribuição subjacente aos dados está no domínio de atração de uma GPD, então os pontos no MRL *plot* devem variar de modo linear enquanto função de  $u$ . Porém, na prática, nem sempre é fácil a determinação de  $u$  e a interpretação muitas vezes requer uma boa dose de julgamento subjetivo.

Outra metodologia usada neste estudo no apoio à seleção do limiar mais adequado inclui o gráfico *parameter stability plot* (e.g., Figuras 3 e 4). Esta abordagem, baseia-se na estimativa de máxima verossimilhança para  $\gamma$  e para  $\sigma^*$ , os quais são representados graficamente contra todos os candidatos a limiar, definida por,

$$\{(u, \sigma^*) : u < x_{\max}\} \text{ e } \{(u, \gamma_u) : u < x_{\max}\}.$$

As estimativas dos parâmetros deverão ser estáveis (i.e., constantes) acima do limiar para o qual o modelo GPD se torna válido. Portanto, se os valores acima de  $u_0$  seguem uma GPD, então o mesmo ocorre para outro qualquer limiar  $u$  ( $u > u_0$ ). Além disso,  $\gamma$  permanece inalterável para todo o  $u$ , i.e.,

$$GPD_u(\gamma, \sigma_u) = GPD(\gamma, \sigma_u - \gamma u),$$

e  $\sigma_u$  é uma função linear de  $u$  (COLES, 2001). Isto significa que caso a distribuição dos excessos tenha uma GPD exata (em vez de aproximada), então o aumento de  $u_0$  para  $u$  resultará noutra GPD com o mesmo parâmetro de forma  $\gamma$ , mas com parâmetro de escala  $\sigma_u$  ajustado (HOSKING e WALLIS, 1987; DARGAHI-NOUBARY, 1989; EMBRECHTS *et al.*, 1997; COLES, 2001). Portanto, se um novo limiar for encontrado, então o resultado de Pickands-Balkema-de Haan poderá ser aplicado acima desse limiar. Ou seja, se a GPD com os parâmetros de forma  $\gamma$  e de escala  $\sigma_u$  é o modelo correto para os excessos acima de  $u_0$ , para qualquer limiar  $u > u_0$ , os excessos serão também uma GPD com o mesmo parâmetro de forma  $\gamma$ , mas com o parâmetro de escala definido como,

$$\sigma_u = \sigma_{u_0} + \gamma(u - u_0), \quad \gamma \neq 0, \quad (25)$$

que variam com  $u$ , exceto se  $\gamma = 0$ . A dificuldade desta solução passa por reparametrizar o parâmetro de escala  $\sigma_u$ , originando um parâmetro de escala modificado  $\sigma^*$ ,

$$\sigma^* = \sigma_u - \gamma u, \quad (26)$$

que ao ser constante deixa de depender de  $u$ , em virtude de (25). Por consequência, também  $\sigma^*$  é independente de  $u$ . Por isso, as estimativas de  $\sigma^*$  e  $\gamma$  serão também constantes para todo o  $u > u_0$ , desde que  $u_0$  seja um limiar válido para os excessos que seguem a GPD.

Na prática nem sempre podemos esperar que o ajuste dos parâmetros permaneça constante, tendo em vista que estamos lidando com uma amostra. Para limitar este problema, uma sugestão poderá ser plotar graficamente  $\hat{\sigma}^*$  e  $\hat{\gamma}$  com os seus intervalos de confiança versus  $u$ , e escolher  $u_0$  como sendo o valor mais baixo de  $u$  para o qual as estimativas permanecem relativamente constantes. Para obter os intervalos de confiança para  $\gamma$  pode-se usar a matriz de variâncias e covariâncias. Contudo, os intervalos de confiança para  $\sigma^*$  dependem de  $\sigma_u$  e de  $\gamma$ . A variância de  $\sigma^*$  é calculada pelo método delta, utilizando os valores “plug-in”,  $Var(\hat{\sigma}^*) \approx \nabla \hat{\sigma}^{*T} V \nabla \hat{\sigma}^*$ , em que

$$\nabla_{\hat{\sigma}^*}^T = \left[ \frac{\partial \hat{\sigma}^*}{\partial \hat{\sigma}_u}, \frac{\partial \hat{\sigma}^*}{\partial \hat{\gamma}} \right] = [1, -u],$$

e  $V$  é a matriz de variâncias e covariâncias para  $\hat{\sigma}^*$  (COLES, 2001).

Existem também metodologias multivariadas para determinação do limiar. Por exemplo, Northrop e Coleman (2014) numa tentativa de escolher o limiar mais adequado apresentam uma metodologia baseada na seleção de múltiplos limiares, e Rootzén *et al.* (2017) apresentam uma abordagem baseada em modelos multivariados dos excessos acima de um limiar.

## 2.6. Qualidade do ajuste

A qualidade do ajuste do modelo foi verificada pelos gráficos de probabilidade (PP *plot*) e quantil-quantil (QQ *plot*).

Seja  $y_i (i = 1, \dots, n)$  uma amostra ordenada das excedências,  $y_1 \leq y_2 \leq \dots \leq y_{N_u}$ , estimadas de um modelo  $\hat{H}$  (i.e., a GPD com os parâmetros estimados). Para  $\hat{\gamma} \neq 0$ , o PP *plot* é um gráfico com os pontos,

$$\left\{ \left[ \hat{H}(y_i), \left( \frac{i}{N_u + 1} \right) \right] : i = 1, \dots, N_u \right\}. \quad (27)$$

A quantidade  $i/(N_u + 1)$  corresponde à f.d. empírica dos excessos avaliada em  $y_i$ , e  $\hat{H}(y)$  é a GPD dada em [6], conforme  $\hat{\gamma} > 0$ ,  $\hat{\gamma} < 0$  ou  $\hat{\gamma} = 0$ .

Assumindo  $\hat{\gamma} \neq 0$ , o QQ *plot* é um gráfico com os pontos,

$$\left\{ \left[ \hat{H}^{-1} \left( \frac{i}{N_u + 1} \right), y_i \right] : i = 1, \dots, N_u \right\}, \quad (28)$$

em que  $\hat{H}^{-1}(y) = u + \frac{\hat{\sigma}_u}{\hat{\gamma}} [y^{-\hat{\gamma}} - 1]$ ; a quantidade  $\hat{H}^{-1}(i/(N_u + 1))$  proporciona a estimativa de um modelo do quantil  $i/(N_u + 1)$  da distribuição  $\hat{H}$ ; os valores de  $y_i$  proporcionam uma estimativa empírica deste quantil.

Se  $\hat{H}$  é um modelo adequado para modelar os excessos de  $u$ , então os pontos dos gráficos *PP plot* e *QQ plot* deverão ter uma configuração aproximadamente linear, situando-se na proximidade da reta de diagonal  $y=x$ .

## 2.7 Aleatoriedade

Com os dados estruturados em função do tempo em que ocorrem, a aleatoriedade pode ser verificada pelo *Runs Test* (WALD e WOLFOWITZ, 1940) e/ou *Turning Point Test* (MOORE e WALLIS, 1943).

Para  $n_1 \geq 20$  e  $n_2 > 20$ , a estatística de teste do *Runs Test*, é dada por,

$$z = \frac{R - \mu_R}{\sigma_R^2} \stackrel{assint}{\sim} N(0,1), \quad (29)$$

em que  $R$  é o número total de sequências, e  $\mu_R$  é a média de  $R$ , dada por ,

$$\mu_R = \frac{2n_1n_2}{n_1 + n_2} + 1, \text{ em que } n_1 = 1, 2, \dots, n_1 \text{ e } n_2 = 1, 2, \dots, n_2 \quad (30)$$

e  $\sigma_R^2$  é a variância de  $R$ , dada por,

$$\sigma_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}. \quad (31)$$

A decisão de rejeitar “ $H_0$ : a sequência é aleatória”, ocorre se  $|z| \leq \Phi_{\alpha/2}$  ou  $|z| \geq \Phi_{1-\alpha/2}$ , onde  $\Phi_{\alpha}$  é o quantil de probabilidade  $\alpha$  da distribuição Normal padrão.

Seja  $X = \{X_1, \dots, X_n\}$ , uma sequência aleatória. Diz-se que existe “ponto de viragem” (*turning point test*) no tempo  $i$ ,  $1 < i < n$ , se  $X_{i-1} < X_i$  e  $X_i > X_{i+1}$  ou se  $X_{i-1} > X_i$  e  $X_i < X_{i+1}$ . Sendo  $T$  o número de “pontos de viragem” da sequência  $X_i$ ,  $i = 1, \dots, n$ . O valor esperado de  $T$  é dado por,

$$\mu_T := E(T) = \frac{2n-4}{3} \quad (32)$$

e variância,

$$\sigma_T^2 := Var(T) = \frac{16n-29}{90}. \quad (33)$$

Brockwell e Davis (2002) mostraram que para uma amostra aleatória  $\{X_1, \dots, X_n\}$ , com  $n$  grande,  $T \sim N(\mu_T, \sigma_T^2)$ . A decisão de rejeitar “ $H_0$ : a série é aleatória”, a um certo nível  $\alpha$ , ocorre se  $z = |T - \mu_T| / \sigma_T > \Phi_{1-\alpha/2}$ .

## 2.8 Estacionariedade

Tal como na aleatoriedade, também na verificação da estacionariedade os dados são estruturados em função do momento temporal em que ocorrem. Um teste usual é o Augmented Dickey-Fuller Test (ADF) (DICKEY e FULLER, 1979), dado por,

$$\Delta z_t = \alpha_0 + \beta t + \theta z_{t-1} + \omega_1 \Delta z_{t-1} + \omega_2 \Delta z_{t-2} + \dots + \omega_{p-1} \Delta z_{t-p+1} + \varepsilon_t, \quad (34)$$

em que  $\alpha$  é a constante,  $\beta$  é o coeficiente de tendência temporal (ou determinística),  $p$  representa o número de defasagens ou *lags* da variável dependente em primeira diferença, incluídas no modelo para remover a autocorrelação serial dos resíduos, assegurando que  $\varepsilon_t$  seja um processo de ruído branco (processo estacionário), ou seja, uma sequência de v.a.'s i.i.d.'s,  $\varepsilon_t \sim N(0, \sigma^2)$ . As restrições  $\alpha=0$  e  $\beta=0$  correspondem ao passeio aleatório, e usando apenas  $\beta=0$  tem-se um modelo sem tendência.

A hipótese  $H_0 : \theta = 0$  indica que a série é não estacionária, por oposição à hipótese  $H_1 : \theta < 0$

## 3 ANÁLISE DE VALORES EXTREMOS NO SALTO EM COMPRIMENTO

### 3.1 Coleta de dados

Os dados utilizados neste estudo foram obtidos em 31-12-2016 no *Website* <http://www.all-athletics.com/en-us/all-time-lists>. Foram solicitadas as melhores performances de sempre no salto em comprimento para homens e mulheres, onde cada atleta “colaborou” apenas com um registo, a sua melhor performance. Nesta base de dados apenas constam as marcas homologadas pela IAAF (*International Association of Athletics Federations*), não considerando os resultados obtidos de forma anómala (e.g., uso de doping e/ou substâncias ilícitas, vento favorável, +2m/s, uso de equipamento não-homologado, desnível do terreno, tábua de chamada irregular, competição em praça de cidade). Nos homens foram disponibilizados os registos de 1935 a 2016 e nas mulheres de 1964 a 2016. O número de recordes individuais por ano consta na Tabela 3.

Tabela 3 - Número de recordes individuais por ano, para homens (H) e mulheres (M)

Ano	H	M	Ano	H	M	Ano	H	M	Ano	H	M	Ano	H	M
1935	2	--	1969	3	--	1981	12	2	1993	16	2	2005	12	4
1947	1	--	1970	6	1	1982	16	2	1994	25	5	2006	18	7
1952	1	--	1971	2	1	1983	13	9	1995	23	4	2007	14	12
1955	2	--	1972	7	1	1984	24	15	1996	24	13	2008	19	7
1957	1	--	1973	3	--	1985	19	7	1997	14	6	2009	18	6
1960	3	--	1974	8	--	1986	12	5	1998	20	9	2010	15	7
1963	2	--	1975	5	--	1987	26	8	1999	16	8	2011	26	9
1964	4	1	1976	12	3	1988	34	13	2000	27	10	2012	28	13
1965	2	--	1977	4	2	1989	18	5	2001	9	2	2013	17	5
1966	1	--	1978	3	1	1990	18	6	2002	14	6	2014	21	4
1967	2	--	1979	11	1	1991	22	4	2003	21	3	2015	26	15
1968	10	1	1980	11	8	1992	24	5	2004	21	9	2016	41	17

Dado que as performances foram registradas em metros e centímetros, em certos casos ocorreu uma grande concentração de marcas com o mesmo registo. Assim, para não causar problemas na estimação, estas marcas foram “suavizadas” antes de operar a análise das mesmas. Uma metodologia semelhante à proporcionada por Einmahl e Magnus (2008), no âmbito da corrida de velocidade, foi aplicada neste estudo, a qual consiste em criar um critério de desempate, sem afetar as marcas pessoais. Ou seja, assumindo que existem  $m$  atletas com a mesma melhor marca pessoal, por exemplo,  $d=8.26$  metros, então os  $m$  pontos serão “suavizados” no intervalo (8.255; 8.265) da seguinte forma:

$$d_j = 8.26 - 0.005 + 0.01 \times \frac{2j-1}{2m}, \quad j = 1, 2, \dots, m.$$

### 3.2 Resultados e discussão

#### 3.2.1 Aleatoriedade e estacionariedade

Antes de iniciar o processo de análise de valores extremos, i.e., usando os dados organizados em função do tempo em que ocorreram, realizamos a verificação da aleatoriedade e da estacionariedade. Pela Tabela 4, observa-se quer com a aplicação do *run test* com valor  $p$  exato (WALD e WOLFOWITZ, 1940) quer do *turning point test* (MOORE e WALLIS, 1943), para uma hipótese bilateral, que os valores de  $p > \alpha$  (5%), indicando assim a aleatoriedade nas amostras, não-rejeitando a hipótese nula de aleatoriedade. Ou seja, as observações são i.i.d.'s. A estacionariedade foi verificada pelo Augmented Dickey-Fuller test (DICKEY e FULLER, 1979). Nas duas amostras  $\theta < 0$ . Paralelamente, a um nível de significância de 5% existem evidências estatisticamente significativas para rejeitar a hipótese nula da não-estacionariedade, pelo que a série é estacionária.

Tabela 4 – Verificação da aleatoriedade pelo *run test* e *turning point test*, e da estacionariedade pelo *augmented Dickey-Fuller test*, para os dados relativos ao salto em comprimento de homens e mulheres

	<i>Run test</i>		<i>Turning Point test</i>		ADF		
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	$\theta$	Lag order	<i>p</i>
S. Comp. homens	-1,1376	0,270	-0,2514	0,802	-8,7130	9	0,01
S. Comp. mulheres	-0,8978	0,402	0,4740	0,636	-5,6144	6	0,01

### 3.2.2 Análise exploratória de dados

Observa-se na Tabela 5 mostra as estatísticas descritivas básicas das amostras. O valor máximo corresponde ao atual recorde mundial. A diferença entre a média e a mediana é, nas duas amostras de 4cm, indicando a grande semelhança entre estas duas estatísticas. *Igualmente, o coeficiente de variação (CV) revela fraca dispersão nas duas amostras (Homens: CV~1.8%; Mulheres: CV~2.3%).*

Tabela 5 – Caracterização das amostras: estatísticas descritivas de base

	Homens	Mulheres
<i>n</i>	829	284
$\xi$	8,16	6,91
<i>s</i>	0,146	0,156
Mín.	8,00	6,75
Máx.	8,95	7,52
$Q_1$	8,05	6,80
$Q_2$	8,12	6,87
$Q_3$	8,22	6,99
$IC_{95\%}(\mu)$	(8,15, 8,17)	(6,90, 6,93)
Assimetria	1,640	1,465
Curtose	3,862	2,113

O histograma do salto em comprimento com a curva de densidade (Figura 1) mostra, como seria de esperar, que nas duas amostras existe uma forte assimetria à direita, que é suportada pelas estimativas dos coeficientes de assimetria e curtose apresentados na Tabela 5.



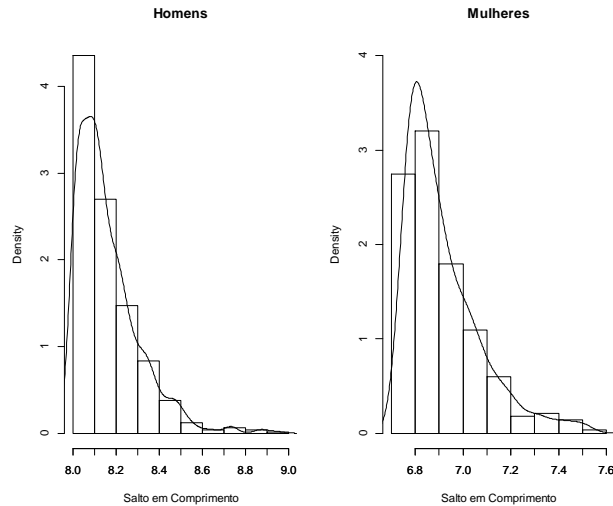


Figura 1 – Histograma com a curva de densidade dos dados referentes ao salto em comprimento de homens e mulheres.

### 3.2.3 Determinação do limiar

#### 3.2.3.1. Quantis empíricos

Pode-se observar na Tabela 6 os quantis empíricos  $q(0,75)$ ,  $q(0,80)$ ,  $q(0,85)$ ,  $q(0,90)$ ,  $q(0,95)$  e  $q(0,99)$ , bem como o número de observações ( $N_u$ ) que estão acima de cada um dos quantis. Obviamente, que numa lógica de valores extremos interessam os quantis mais afastados, desde que se disponha de número suficiente de observações para estimação, não afetando o viés e a variância dos estimadores. DuMouchel (1983) sugere que o valor limiar se situe na proximidade do quantil empírico de probabilidade  $q(0,90)$ . Gong (2012) sugere que varie entre  $q(0,90)$  e  $q(0,97)$ .

Tabela 6 – Quantis empíricos:  $q(0,75)$ ,  $q(0,80)$ ,  $q(0,85)$ ,  $q(0,90)$ ,  $q(0,95)$  e  $q(0,99)$ , para os dados referentes ao salto em comprimento, homens e mulheres

	$q(0,75)$	$q(0,80)$	$q(0,85)$	$q(0,90)$	$q(0,95)$	$q(0,99)$
Homens	8,22	8,25	8,30	8,35	8,45	8,70
	( $N_u=212$ )	( $N_u=172$ )	( $N_u=126$ )	( $N_u=84$ )	( $N_u=42$ )	( $N_u=9$ )
Mulheres	6,99	7,02	7,06	7,12	7,21	7,44
	( $N_u=73$ )	( $N_u=56$ )	( $N_u=43$ )	( $N_u=30$ )	( $N_u=15$ )	( $N_u=3$ )

#### 3.2.3.2 Gráfico de vida média residual

Para ajustar a GPD aos excessos acima de um certo limiar  $u$ , recorreremos à informação gráfica obtida pelo gráfico de vida média residual, i.e., MRL *plot*

(Figura 2). Aqui, o limiar suficientemente alto é selecionado no ponto do gráfico a partir do qual se visualiza uma tendência de linearização. Nos homens, o gráfico mostra que o intervalo  $u = (8,45; 8,50)$  e nas mulheres o intervalo  $u = (7,06; 7,11)$ , a MEF apresenta-se bastante linear. A partir de 8.50m nos homens começa a verificar-se uma trajetória alternadamente ascendente-descendente (com vários pontos de inflexão), e a partir de 7,11m nas mulheres a trajetória é predominantemente descendente.

A ausência de uma solução geral para a escolha do limiar induz uma incerteza latente acerca do próprio limiar. Por isso, testamos vários modelos, considerando diferentes níveis de  $u$ . E, apoiados nas Figuras 2, 3 e 4, bem como nos gráficos de diagnóstico de cada um dos modelos testados, selecionamos o modelo que apresentou o melhor compromisso entre a redução do viés e a menor variância na estimação de máxima verossimilhança dos parâmetros  $(\gamma, \sigma_u)$ . Assim, para os homens definimos  $u = 8,45$ m e para as mulheres definimos  $u = 7,06$ m, que correspondem aos quantis 0,95 e 0,85, respectivamente, não muito distantes do percentil 90 sugerido por Du Mouchel (1983), ou do intervalo  $q(0,90, 0,97)$  proposto por Gong (2012).

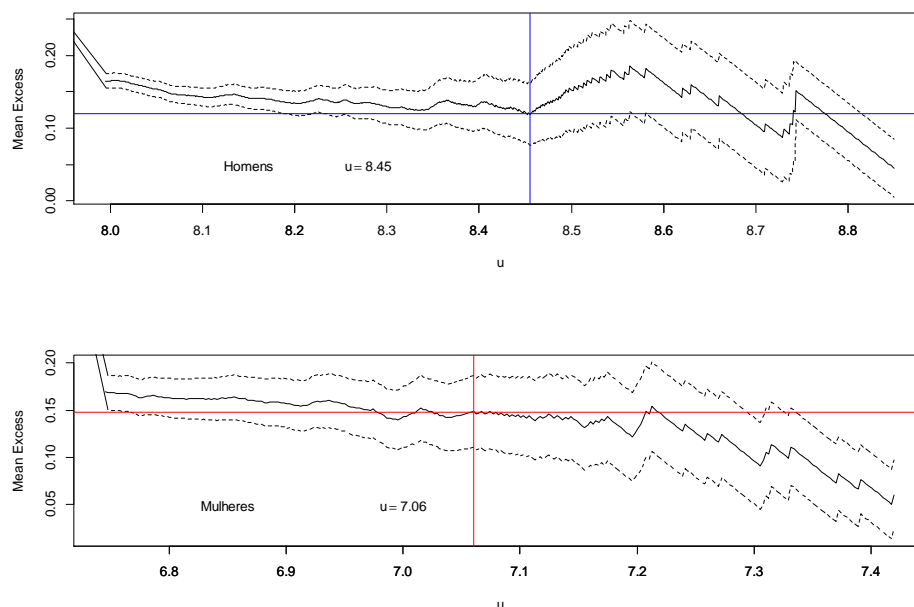


Figura 2 – Gráfico de vida média residual para a performance no salto em comprimento: homens (em cima) e mulheres (em baixo).

Para ajudar na seleção do limiar mais adequado também recorreremos ao *parameter stability plot* (Figura 3 e Figura 4), onde se projetam graficamente as

estimativas dos parâmetros da GPD contra uma gama de valores candidatos a limiar. As estimativas obtidas por MLE dos parâmetros  $(\sigma_u, \gamma)$  deverão ser estáveis (i.e., relativamente constantes) acima do limiar no qual o modelo GPD se torna válido. As Figuras 3 e 4, apresentam o *parameter stability plot*, para homens e mulheres, respetivamente, mostrando para cada candidato a limiar os valores reparametrizados de  $\hat{\sigma}_u^*$  (*modified scale*) e  $\hat{\gamma}$  (*shape*), com os respectivos intervalos de confiança ( $IC_{95\%}$ ). Esta reparametrização parece confirmar a informação recolhida na Figura 2. Ou seja, a combinação entre MRL plot e *parameter stability plot*, dá-nos  $u = 8,45$  nos e homens ( $N_u = 42$ ) e  $u = 7,06$  nas mulheres ( $N_u = 43$ ) como escolhas razoáveis.

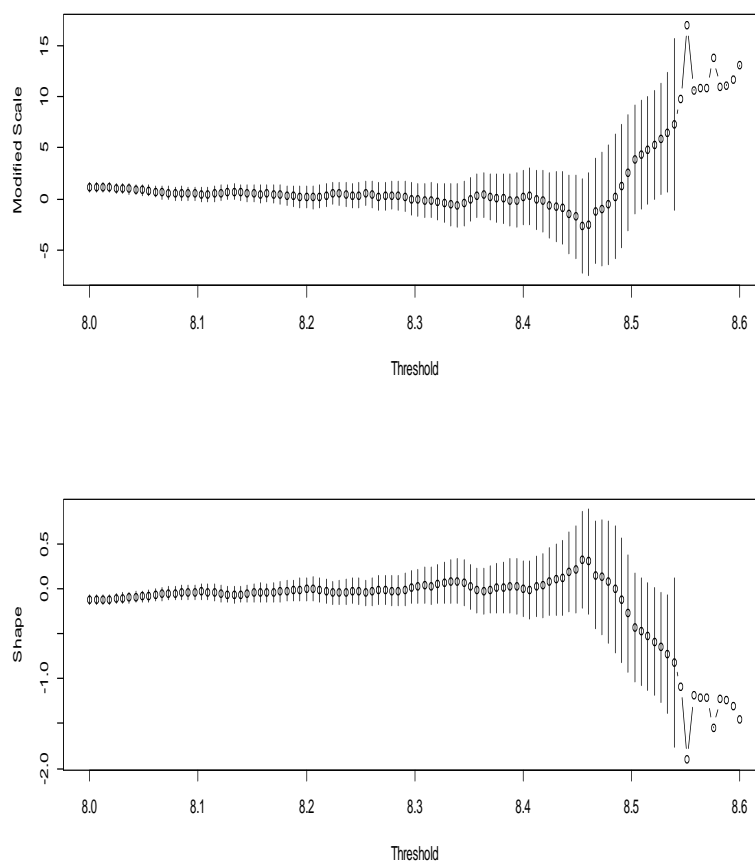


Figura 3 – Gráfico de estabilidade dos parâmetros de escala modificado e forma, por máxima verossimilhança, com os respetivos intervalos com 95% de confiança baseados na função *profile log-likelihood*, para os dados referentes ao salto em comprimento masculino.

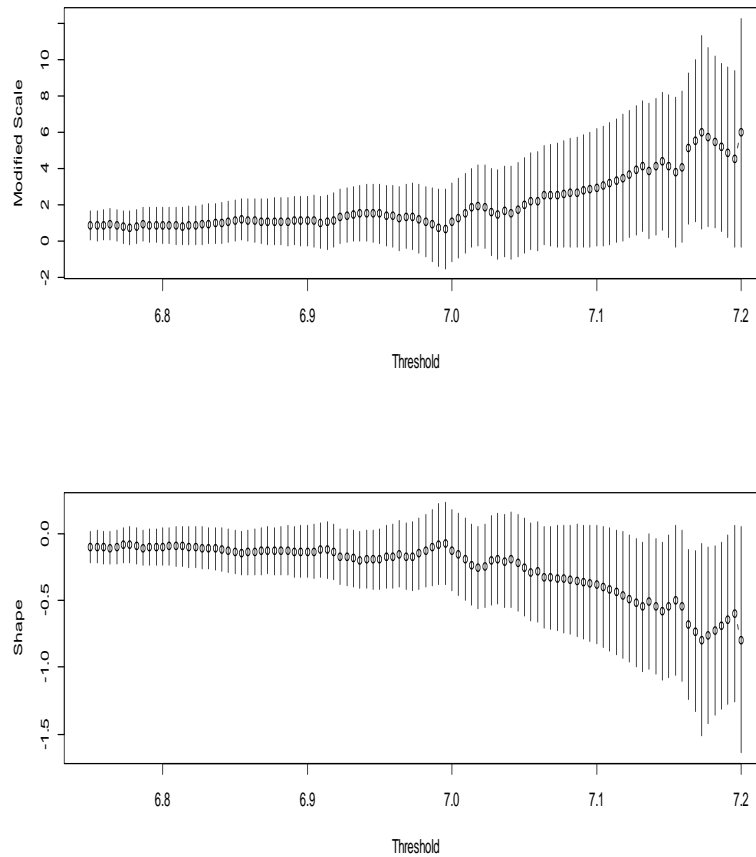


Figura 4 – Gráfico de estabilidade dos parâmetros de escala modificado e forma, por máxima verossimilhança, com os respectivos intervalos com 95% de confiança baseados na função *profile log-likelihood*, para os dados referentes ao salto em comprimento feminino.

### 3.2.4. Estimação dos parâmetros de interesse

Em função do limiar  $u$  selecionado e consequentes excedências  $N_u$ , segundo o método da máxima verossimilhança para a GPD e o algoritmo de Nelder-Mead, apresenta-se na Tabela 7 a estimativa de alguns parâmetros de interesse, nomeadamente, os parâmetros de escala  $\hat{\sigma}_u$  e forma  $\hat{\gamma}$  com os respectivos erros-padrão e  $IC_{95\%}$ , a probabilidade de excedência de um nível  $x$  elevado e o quantil extremal para uma probabilidade de 0,1%.

Hosking e Wallis (1987) demonstraram que quando  $\gamma > -0.5$ , verificam-se as condições de regularidade da máxima verossimilhança e a estimativa de máxima verossimilhança de  $(\hat{\gamma}_{N_u}, \hat{\sigma}_{N_u})$  baseadas na amostra dos excessos  $N_u$  tem distribuição Normal assintótica. Nos homens,  $\hat{\gamma} = 0.25428 (se = 0.253)$ , aponta para um modelo Pareto ( $\gamma > 0$ ) para as excedências. Nas mulheres,  $\hat{\gamma} = -0.31877 (se = 0.189)$ , sugere um modelo Beta ( $\gamma < 0$ ). Contudo, a zona de estimação nas duas amostras tem variância elevada, uma vez que os intervalos de confiança para o parâmetro de forma apresentam amplitudes elevadas. Esta constatação baseia-se no fato da estimativa do  $IC_{95\%}(\gamma)$  não estar inteiramente contido na região  $]0; +\infty[$  no caso dos homens e na região  $] -\infty; 0[$  no caso das mulheres. No caso concreto do salto em comprimento, o estudo de Einmhal e Magnus (2008) é concordante com o presente estudo no caso dos homens,  $\hat{\gamma} = 0.06$ , e relativamente concordante no caso das mulheres  $\hat{\gamma} = -0.07$ , e envolvendo apenas amostras de homens e com resultados obtidos em pista coberta, Gomes (2008), Gomes *et al.* (2009) e Rodrigues *et al.* (2011) registaram estimativas do parâmetro de forma diferentes do presente estudo,  $\hat{\gamma} = -0.26$ ,  $\hat{\gamma} = -0.058$  e  $\hat{\gamma} = -0.23$ , respectivamente.

A probabilidade de exceder o limiar, dada pela frequência relativa  $N_u/n$  das observações que excedem  $u$  na amostra original, é de  $N_u/n = 0,0507 (se = 0,00762)$  nos homens e de  $N_u/n = 0,1514 (se = 0,02127)$  nas mulheres.

Sabendo que o máximo da amostra  $X_i$  corresponde ao atual recorde mundial de 8,95m de Mike Powell (EUA) em 1991 e 7,52m de Galina Chistyakova (URSS) em 1988, será interessante conhecer, dado o atual estado de arte, a probabilidade de destas marcas serem ultrapassadas. Assim, pretendemos saber a probabilidade de excedência de um nível  $x$  elevado, ou seja,  $P[X > 8,95]$  e  $P[X > 7,52]$ . Pela expressão (15) e apresentado na Tabela 7, observamos que a probabilidade do recorde mundial ser batido é de, aproximadamente, 0,171% nos homens e 0,234% nas mulheres.

Na amostra de mulheres, uma vez que  $\hat{\gamma} < 0$ , o limite superior do suporte, dado em [16], é de  $\hat{x}^F = u - \hat{\sigma}_u / \hat{\gamma} \sim 7.69m$ .

Tabela 7 – Limiar ( $u$ ), excedências ( $N_u$ ), estimativa, erro-padrão e  $IC_{95\%}$  para os parâmetros de forma e escala, probabilidade de excedência de um nível  $x$  elevado, com base na GPD, segundo o método de estimação por máxima verossimilhança

	$u$	$N_u$	$\hat{\sigma}_u$	$\hat{\gamma}$	$P[X > x]$
Homens	8,45	42	0,09305 (0,027) (0,039; 0,147)	0,25428 (0,253) (-0,242; 0,751)	$P[X > 8,95] =$ $= 0,00171231$
Mulheres	7,06	43	0,1994 (0,047) (0,106; 0,292)	-0,31877 (0,189) (-0,689; 0,051)	$P[X > 7,52] =$ $= 0,00233829$

Outra questão de interesse é saber os valores que são excedidos com baixa probabilidade – os quantis extremos – e que associado a estas quantidades podemos obter os níveis de retorno. Usando a expressão apresentada em (14), e considerando o quantil extremal  $\chi_{0,999}(U(1000))$ , sabemos que existe aproximadamente 0.1% de probabilidade que um saltador em comprimento consiga o seu melhor registo pessoal acima de 9.08m se for homem e acima de 7.56m se for mulher. Como a escolha de  $u$  não é fácil nem consensual, decidimos verificar o impacto que diferentes limiares têm no quantil extremal  $\chi_{0,999}$ , os quais estão apresentados na Tabela 8. Verifica-se que na amostra de mulheres com um quantil extremal de probabilidade de excedência 0,1%, as performances previsíveis de ser realizadas mantêm-se relativamente estáveis, independentemente de  $u$ ; na amostra de homens, observa-se a partir de  $u=8,46$  uma tendência decrescente, sendo que acima de  $u=8,48$  as marcas previsíveis são inferiores ao atual recorde mundial.

Tabela 8 – Estimativa do quantil extremal  $\hat{\chi}_{0,999}$ , com base na GPD, segundo o método de estimação por máxima verossimilhança, para diferentes limiares

	$u=8,45$	$u=8,46$	$u=8,47$	$u=8,48$	$u=8,49$	$u=8,50$
Homens	9,076909	9,095718	9,027628	8,973121	8,942373	8,918035
	$u=7,06$	$u=7,07$	$u=7,08$	$u=7,09$	$u=7,10$	$u=7,11$
Mulheres	7,55916	7,54992	7,548959	7,547069	7,54409	7,539976

A Tabela 9 apresenta as estimativas pontuais por máxima verossimilhança, segundo a expressão (18), os intervalos de confiança calculados pela *profile log-likelihood function* dado na expressão (19) para os níveis de retorno para diferentes períodos de retorno (i.e., recordes individuais),  $U(100)$ ,  $U(200)$ , ...,  $U(700)$ . Paralelamente, com base nas expressões (15) e (20), verificamos que para o atual recorde mundial ser ultrapassado são necessários cerca de 585 recordes individuais nos homens e cerca de 427 recordes individuais nas mulheres. Contudo, nas duas amostras, duas situações poderão colocar em causa esta previsão: (i) o  $IC_{95\%}$  para o nível de retorno  $U(200)$  e  $U(100)$ , para homens e mulheres respectivamente, já

contém um novo recorde mundial, (ii) o limite inferior do  $IC_{95\%}$  para todas as estimativas pontuais dos níveis de retorno é inferior ao atual recorde mundial. Ou seja, se por um lado é possível que um novo recorde mundial ocorra antes dos 585 e 427 recordes individuais de homens e mulheres, também pode acontecer que um novo máximo mundial não se verifique com 700 recordes individuais. Deve-se notar que a interpretação dos valores de retorno se deve ao fato dos “blocos” corresponderem a atletas de salto em comprimento que não estão definidos por qualquer período temporal, pelo que a forma como analisamos esta informação pode ser um pouco mais exagerada, deverá ser lidada com extremo cuidado, mas ainda assim interessante. Por isso, à semelhança de Caeiro *et al.* (2015) que alertam para o cuidado a ter com a extrapolação para longos períodos de retorno, também julgamos que o mesmo cuidado deverá estar presente na extrapolação para grande número de recordes individuais.

Tabela 9 – Estimativas pontuais por máxima verossimilhança e intervalos de confiança pela *profile log-likelihood function*, do modelo GPD dos níveis de retorno para 100, 200, 300, 400, 500, 600 e 700 recordes individuais

	<i>U</i> (100)	<i>U</i> (200)	<i>U</i> (300)	<i>U</i> (400)	<i>U</i> (500)	<i>U</i> (600)	<i>U</i> (700)
Homens	8,64 (8,58; 8,73)	8,74 (8,65; 8,99)	8,82 (8,70; 9,23)	8,87 (8,37; 9,47)	8,92 (8,75; 9,69)	8,96 (8,77; 9,90)	8,99 (8,79; 10,11)
Mulheres	7,42 (7,36; 7,55)	7,47 (7,41; 7,69)	7,50 (7,43; 7,78)	7,52 (7,45; 7,84)	7,53 (7,46; 7,89)	7,54 (7,46; 7,94)	7,54 (7,47; 7,97)

### 3.2.5 Ajuste do modelo

Na Figura 5 (homens) e na Figura 6 (mulheres) apresenta-se o ajuste do modelo pelos gráficos de probabilidade e quantil-quantil. Nas duas amostras, os gráficos de probabilidade e de quantis empíricos,  $y_i : N_u$  vs quantis teóricos,  $\left(\frac{(1-p_i)^{-\gamma}-1}{\gamma}, y_i : N_u\right)$ ,  $p_i := i/(N_u + 1), i = 1, \dots, N_u$ , i.e., a GPD, exibem uma tendência quase linear, embora menos evidente nos homens. Todos os pontos estão contidos no  $IC_{95\%}$ . De uma forma geral, podemos concluir que os gráficos de diagnóstico, suportados pela GPD, apresentam um ajuste global razoável, mas existem algumas evidências que indicam, que sobretudo na amostra de homens, que o modelo subestima os dados nos níveis mais extremos, e isto é um pouco preocupante uma vez que estamos interessados em extrapolar para níveis ainda mais extremos. Neste, e como recomendação para trabalho futuro, podemos considerar o processo de modelação pela distribuição generalizada de valores extremos (von MISES, 1936, 1954, JENKINSON, 1955) e o método dos máximos de blocos (GUMBEL, 1935, 1958).

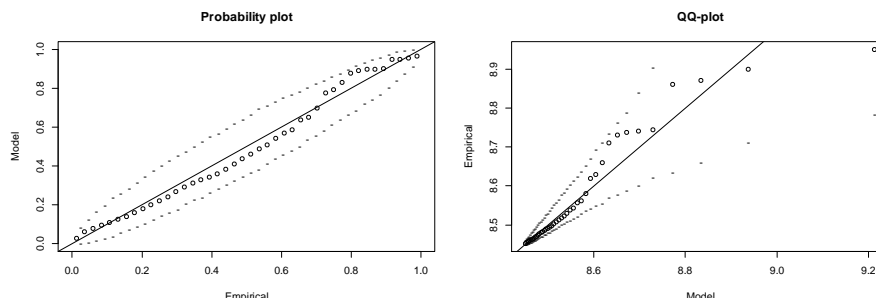


Figura 5 – Verificação da qualidade do ajuste do modelo obtido para o salto em comprimento: homens.

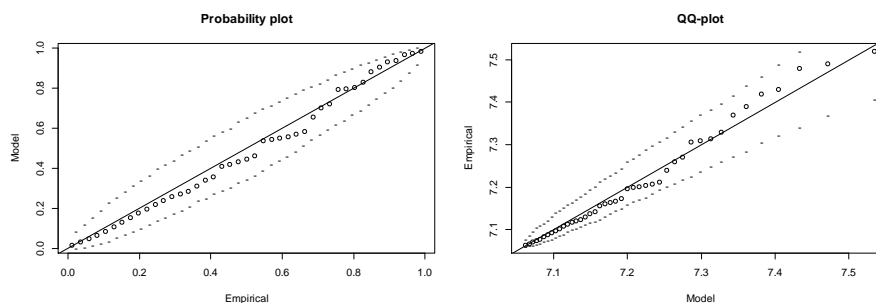


Figura 6 – Verificação da qualidade do ajuste do modelo obtido para o salto em comprimento: mulheres.

#### 4 Conclusões

Neste estudo utilizamos a distribuição generalizada de Pareto para ajustar os excessos acima de um limiar. A metodologia POT mostrou-se adequada para modelar o salto em comprimento dos melhores atletas do mundo ao longo dos tempos, homens e mulheres. Por intermédio do método da máxima verossimilhança e o algoritmo de Nelder-Mead obtivemos as estimativas dos parâmetros de forma e escala. Com base no gráfico de vida média residual, selecionamos o limiar. Antes, observamos o comportamento do parâmetro de forma para vários candidatos a limiar. Entre vários modelos testados, optamos por aquele que apresentou o melhor compromisso entre a menor variância e o menor viés dos estimadores. Os resultados apontam para  $u=8,45\text{m}$  com  $\hat{\gamma} > 0$  nos homens e  $u=7,06\text{m}$  com  $\hat{\gamma} < 0$  nas mulheres. Prevê-se que o recorde mundial seja batido acima de 585 recordes individuais nos homens, e acima de 427 recordes individuais nas mulheres. Os resultados deste estudo podem servir de guia para todos os que se interessam pela problemática dos extremos no desporto, especialmente no salto em comprimento do atletismo.



## Agradecimentos

Esta pesquisa foi parcialmente suportada pela Fundação para a Ciência e a Tecnologia através do projeto UID/MAT/04674/2013 (CIMA) (Centro de Investigação em Matemática e Aplicações, Universidade de Évora) e pelo projeto UID/MAT/00297/2013 (Centro de Matemática e Aplicações). Os autores agradecem aos revisores e editores da Revista Brasileira de Biometria pelas importantes e pertinentes sugestões para melhoria da qualidade do artigo.

SILVA, D.; CAEIRO, F.; OLIVEIRA, M. Application of the peaks over threshold methodology to the long jump of athletics. *Rev. Bras. Biom.* Lavras, v.36, n.4, p.753-779, 2018.

- **ABSTRACT:** *The Peaks Over Threshold (POT) method consists in the fit of a generalized Pareto distribution to a sequence of independent and identically distributed random excesses over a high threshold. In this study, we use the POT method to estimate the distributional parameters of excesses in the application of track and field long jump, between 1935 to 2016 in males, and 1964 to 2016 in females. After determining the appropriate threshold  $u$ , we present the estimates of some relevant parameters, like the exceedance probability, the extremal quantile, the right endpoint (only when  $\hat{\gamma} < 0$ ), and the return levels and return period. In males, the results they point to  $u=8.45m$  with  $\hat{\gamma} > 0$ , and in females  $u=7.06m$  with  $\hat{\gamma} < 0$ . It is expected that the world record will be beaten above 585 individual records in men, and above 427 individual records in women.*
- **KEYWORDS:** *Extreme value theory; exceedance probability; extremal quantile; right endpoint; return levels.*

## Referências

- BALKEMA, A. A.; HAAN, L. Residual life time at great age. *Annals of Probability*, v.2, n.5, p.792-804, 1974
- BEIRLANT, J.; GOEGEBEUR, Y.; SEGERS, J.; TEUGELS, J. *Statistics of extremes: Theory and applications*. Chichester: John Wiley & Sons, 2004. 490p.
- BENKTANDER, G.; SEGERDAHL, C. (1960). On the analytical representation of claim distributions with special reference to excess of loss reinsurance. In: *Transactions 16-th International Congress of Actuaries*, Brussels, 1960. p.626-646.
- BROCKWELL, P. J.; DAVIS, R. A. *Introduction to time series and forecasting*. New York: Springer Verlag, 2002. 434p.
- CAEIRO, F.; MATEUS, A.; RAMOS, L. Extreme value analysis of the sea levels in Venice. *AIP Conference Proceedings*, 1648, p.540006-1540006-4, 2015
- COLES, S. *An introduction to statistical modeling of extreme values*. London: Springer-Verlag, 2001. 208p.
- DARGAHI-NOUBARY, G. R. On tail estimation: an improved method. *Mathematical Geology*, v.21, n.8, p.829-842, 1989.

- DAS, B.; GHOSH, S. Weak limits for exploratory plots in the analysis of extremes. *Bernoulli*, v.19, n.1, p.308-343, 2013.
- DAVISON, A. Modeling excesses over high threshold with an application. In: TIAGO DE OLIVEIRA, J. (Ed.). *Statistical Extremes and Applications*. Lisbon: Springer, 1984. p.461-482.
- DAVISON, A. C.; SMITH, R. L. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, Series B (Methodological)*, v.52, n.3, p.393-442, 1990.
- DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, v.74, n.366, p.427-431, 1979.
- Du MOUCHEL, W. Estimating the stable index  $\alpha$  in order to measure tail thickness: a critique. *Annals of Statistics*, v.11, n.4, p.1019-1036, 1983.
- EINMAHL, J. H. J.; MAGNUS, J. R. Records in athletics through extreme-value theory. *Journal of the American Statistical Association*, v.103, n.484, p.1382-1391, 2008.
- EMBRECHTS, P.; KLUPPELBERG, C.; MIKOSCH, T. *Modelling extremal events. Applications in Mathematics*, v.33, New York: Springer-Verlag, 1997. 644p.
- GHOSH, S.; RESNICK, S. A discussion on mean excess plots. *Stochastic Processes and their Applications*, v.120, n.8, p.1492-1517, 2010.
- GHOSH, S.; RESNICK, S. When does the mean excess plot look linear? *Stochastic Models*, v.27, n.4, p.705-722, 2011.
- GOMES, M. I. *Estatística de extremos e desporto - como estimar alguns parâmetros úteis*. Seminário de Probabilidades e Estatística da UTL, FMH, 2008.
- GOMES, M. I.; PESTANA, D.; RODRIGUES, L. H. *Athletic events and statistics of extremes: estimation of useful parameters*. Technical Report CEAUL 06/09, 2009.
- GONG, S. *Estimation of hot and cold spells with extreme value theory*. Project Report 19, Uppsala University, 2012. 55p.
- GUMBEL, E.J. Les valeurs extrêmes des distributions statistiques. *Annals of Institute Henri Poincaré*, v.5, n.2, p.115-158, 1935.
- GUMBEL, E.J. *Statistics of extremes*. New York: Columbia University Press, 1958. 375p.
- HALL, W. J.; WELLNER, J. Mean residual life. Statistics and related topics (Ottawa, Ont., 1980). North-Holland, Amsterdam, 1981. p.169-184, 1981.
- HEFFERNAN, J.E.; STEPHENSON, A.G. *ismev: An introduction to statistical modeling of extreme values*. R package version 1.41. 2016.
- HOSKING, J. R. M.; WALLIS, J. R. Parameter and quantile estimation for the Generalized Pareto Distribution. *Technometrics*, v.29, n.3, p.339-349, 1987.
- JENKINSON, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, v.81, n.348, p.158-171, 1955.
- LEWIS, A. S.; OVERTON, M. Nonsmooth optimization via quasi-Newton BFGS. *Mathematical Programming*, v.141, n.1, p.135-163, 2013.

- MOORE, G. H.; WALLIS, W. A. Time series significance tests based on signs of differences. *Journal of the American Statistical Association*, v.38, n.222, p.153-164, 1943.
- NELDER, J. A.; MEAD, R. A simplex method for function minimization. *Computer Journal*, v.7, n.4, p.308-313, 1965.
- NORTHROP, P.J.; COLEMAN, C. Improved threshold diagnostic plots for extreme value analyses. *Extremes*, v.17, n.2, p.289-303, 2014.
- NORTHROP, P. J.; JONATHAN, P.; RANDELL, D. Threshold modeling of nonstationary extremes. In: DEY, D. K.; YAN, J. (Ed.). *Extreme value modeling and risk analysis: methods and applications*. Boca Raton, FL: Chapman and Hall/CRC Press, p.87-108, 2017.
- PICKANDS, J. Statistical inference using extreme order statistics. *Annals of Statistics*, v.3, n.1, p.119-131, 1975.
- R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2017.
- RODRIGUES, L. H.; GOMES, M. I.; PESTANA, D. Statistics of extremes in athletics. *REVSTAT - Statistical Journal*, v.9, n.2, p.127-153, 2011.
- ROOTZÉN, H.; SEGERS, J.; WADSWORTH, J. L. Multivariate peaks over threshold models. *Extremes*, DOI 10.1007/s10687-017-0294-4, 2017.
- SCHMOLINSKY, G. *Atletismo*. 2.ed. Lisboa: Editorial Estampa, 1982. 508p.
- SMITH, R. L. Threshold methods for sample extremes. In: TIAGO DE OLIVEIRA, J. (Ed.). *Statistical Extremes and Applications*. Lisbon: Springer, 1984. p 623-640.
- von MISES, R. La distribution de la plus grande de n valeurs. *Revue Mathématique de l'Union Interbalcanique*, v.1, p.141-160, 1936.
- von MISES, R. La distribution de la plus grande de n valeurs. *Revue Mathématique de l'Union Interbalcanique*, 1, 141-160. Reprinted in Selected Papers of Richard von Mises, *American Mathematical Society*, v.2, p.271-294, 1954.
- WALD, A.; WOLFOWITZ, J. On a test whether two samples are from the same population. *Annals of Mathematical Statistics*, v.11, n.2, p.147-162, 1940.

Recebido em 31.03.2017

Aprovado após revisão em 31.10.2017