

XIX CONGRESSO DE PÓS-GRADUAÇÃO DA UFPA
27 de setembro a 01 de outubro de 2010

**ANÁLISE DOS MÉTODOS DE ESTIMAÇÃO PARA OS PARÂMETROS DA
DISTRIBUIÇÃO GUMBEL NA PRECIPITAÇÃO DE CHUVAS MÁXIMAS PARA A CIDADE
DE PIRACICABA-SP**

FABRÍCIA RESENDE FREIRE¹, LUIZ ALBERTO BEIJO².

Resumo:

As obras de construções civis como barragens, torres de alta tensão, entre outras, são projetadas para suportar o limite máximo (ou mínimo) conhecido dos eventos meteorológicos de uma região. A idéia é que se determinada estrutura suporta um valor extremo de um evento ela está em segurança para os valores correntes. Uma das distribuições de probabilidade que se trata de valores extremos é a distribuição Gumbel, ou distribuição tipo I de Fisher-Tippet.. As estimativas dos parâmetros α e β para essa distribuição podem ser obtidas pelos métodos dos momentos, da regressão, de Lieblein e da máxima verossimilhança. Neste trabalho os métodos de estimação foram avaliados quanto ao ajuste da distribuição Gumbel aos dados agrupados em períodos mensal e anual, e na estimação da precipitação pluvial diária máxima provável, para os tempos de retorno de 5, 25, 50, 100 e 200 anos, para o município de Piracicaba no estado de São Paulo. Verificou-se que para o período de retorno de 5, 25, 50 anos não houve diferença significativa de precipitação diária máxima provável entre os métodos analisados. Para o período de 100 anos o método da máxima verossimilhança apresentou maior precipitação diária máxima provável no mês de janeiro e para o período de 200 anos nos meses de janeiro e outubro. No entanto, o método da máxima verossimilhança apresentou-se mais preciso e menos tendencioso em relação aos demais métodos analisados.

Palavras-chave: Teoria dos Valores Extremos, Métodos de Estimação, Precipitação diária máxima.

INTRODUÇÃO

A teoria de valores extremos em especial a distribuição Gumbel, também conhecida como distribuição assintótica dos extremos do tipo I de Fisher-Tippet ou dupla exponencial desenvolvida por (GUMBEL, 1958), tem apresentado um papel de grande importância em vários campos de pesquisa, principalmente na estatística aplicada à análise de fenômenos meteorológicos, entre os quais as precipitações pluviométricas máximas. Em situações climáticas que ocorrem eventos extremos, a distribuição Gumbel tem se mostrado adequada para descrever os eventos e tem sido aplicada com grande frequência na análise estatística de dados correspondentes a medições de grandezas físicas.

Existem diversas formas de fazer inferências sobre parâmetros da distribuição Gumbel. Haan (1977), propõe o uso dos estimadores do método dos momentos e do método da máxima verossimilhança para a estimação dos parâmetros da distribuição Gumbel. Sendo que o método da verossimilhança apresenta algumas características desejáveis para uma estimativa, entre elas: eficiência com menor variância do que quaisquer outros estimadores; consistência, isto é, assintoticamente não-tendenciosa, com variância tendendo a zero, (SILVA et al., 2000). Assis et al. (1996), sugerem ainda, o método de regressão e o método de Lieblein, sendo que este último necessita de valores tabelados para a obtenção das estimativas.

1 Discente de Graduação em Matemática – Universidade Federal de Alfenas (UNIFAL-MG).
fabriar.freire@yahoo.com.br

2 Prof. Adjunto II – Universidade Federal de Alfenas (UNIFAL-MG).
luizbeijo@yahoo.com.br

XIX CONGRESSO DE PÓS-GRADUAÇÃO DA UFLA 27 de setembro a 01 de outubro de 2010

Este trabalho teve como objetivo comparar, via simulados, os estimadores fornecidos pelos métodos dos momentos, da regressão, de Lieblein e da máxima verossimilhança quanto ao comportamento com diferentes valores paramétricos e tamanhos de amostra; avaliar ambos os métodos no ajuste da distribuição Gumbel aos dados agrupados em períodos anual e mensal, e na estimação da precipitação pluviométrica diária máxima provável, para os tempos de retorno de 5, 25, 50, 100 e 200 anos, para o município de Piracicaba no estado de São Paulo.

MATERIAL E MÉTODOS

Material

Foram analisadas as séries históricas dos valores máximos diários de precipitação (em mm), registradas na Estação Agrometeorológica da ESALQ, USP, em Piracicaba, SP (latitude 22°43' S, longitude 47°30' W e altura 545m). O clima da região é tropical úmido com chuvas de verão e seca no inverno.

A série histórica de precipitação abrange o período de janeiro de 1917 a dezembro de 2008 num total de 92 anos. Ocorreu uma variação no tamanho da amostra em alguns períodos, isso aconteceu em virtude de que em alguns anos verificados, não havia dados para precipitação. Os dados foram agrupados em períodos de um mês e um ano, extraindo-se a máxima precipitação diária observada de cada série.

Métodos

Para a análise de dados reais observou-se a precipitação máxima nos meses de outubro, novembro, dezembro, janeiro, fevereiro e a precipitação máxima anual em Piracicaba - SP, durante o período de 1917 a 2008. Foi criado um algoritmo no programa R (R Development Core Team, 2010) para a obtenção de estimativas para os parâmetros α e β pelos métodos dos momentos, da regressão, de Lieblein e da máxima verossimilhança. Estudou-se o ajuste do modelo de distribuição Gumbel aos dados de precipitação pluviométrica máxima. E determinou-se a influência dos diferentes métodos de estimação na obtenção de precipitações máximas para períodos de retorno de 5, 25, 50, 100 e 200 anos.

Distribuição de Gumbel

A função de densidade de probabilidades da distribuição Gumbel, da variável aleatória X , associada a valores máximos, denotada por $f(x; \alpha, \beta)$, é dada por:

$$f(x; \alpha, \beta) = \frac{1}{\beta} \exp \left[- \left(\frac{x - \alpha}{\beta} \right) - \exp \left(- \left(\frac{x - \alpha}{\beta} \right) \right) \right] \quad (1)$$

dados que x é a variável aleatória associada a valores máximos do período e $-\infty < x < \infty$, α denominado parâmetro de posição e $-\infty < \alpha < \infty$, β parâmetro de escala e $\beta > 0$.

A probabilidade P de que ocorra uma precipitação pluviométrica máxima maior que um certo valor X é estimada utilizando-se a complementar da função de distribuição acumulada da distribuição Gumbel, que é dada por:

$$P(X > x) = 1 - \exp \left[- \exp \left(- \left(\frac{x - \hat{\alpha}}{\hat{\beta}} \right) \right) \right] \quad (2)$$

XIX CONGRESSO DE PÓS-GRADUAÇÃO DA UFLA
27 de setembro a 01 de outubro de 2010

em que x é a precipitação máxima do período e varia de $0 < x < \infty$.

Segundo Hann (1977), o período de retorno é definido como o inverso da probabilidade de frequência do evento a saber, portanto:

$$T = 1 / P \quad (3)$$

Método dos momentos

De acordo com Haan (1977), os parâmetros α e β da distribuição Gumbel são estimados pelo método dos momentos respectivamente pelas seguintes expressões,

$$\hat{\alpha} = \bar{X} - 0,5772 \beta \quad (4)$$

$$\hat{\beta} = \frac{\sqrt{6}}{\pi} s \quad (5)$$

onde \bar{X} e s , são respectivamente as estimativas da média e do desvio padrão da amostra.

Método da regressão

Tomamos $Ln [Ln (n / (N + 1))] = Y$, $\alpha / \beta = a$ e $1 / \beta = b$, ela toma a forma $Y = a + bX$, que é a equação da reta. Desse modo os parâmetros a e b podem ser estimados por:

$$a = \bar{Y} - b\bar{X} \quad (6)$$

$$b = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}} \quad (7)$$

Portanto,

$$\hat{\beta} = -1 / b \quad (8)$$

$$\hat{\alpha} = a \cdot \hat{\beta} \quad (9)$$

Os valores de $n / (N + 1)$, de $Ln [n / (N + 1)]$ e de X dos dados são organizados em forma crescente.

Método de Lieblein

De acordo com Assis (1996) para a aplicação do método de Lieblein os dados devem ser organizados em ordem cronológica e subdivididos em grupos, com no máximo seis observações para se efetuarem os cálculos. Dentro de cada grupo os dados devem ser ordenados em ordem crescente e ponderados de acordo com os pesos estatísticos para estimativa dos parâmetros da distribuição de valores extremos.

XIX CONGRESSO DE PÓS-GRADUAÇÃO DA UFLA
27 de setembro a 01 de outubro de 2010

As estimativas de α e β são dadas por:

$$\hat{\alpha} = \frac{a_i \sum_{i=1}^n X_i}{K} \quad (10)$$

$$\hat{\beta} = \frac{b_i \sum_{i=1}^n X_i}{K} \quad (11)$$

onde K é o número de grupos formados.

Método da máxima verossimilhança

Segundo Lawless (1982) os estimadores da máxima verossimilhança (E.M.V.) para os parâmetros α e β , obtidos a partir da f.d.p conjunta da distribuição Gumbel são respectivamente,

$$\hat{\alpha} = -\hat{\beta} \log \left[\frac{1}{n} \sum_{i=1}^n \exp \left(\frac{-x_i}{\hat{\beta}} \right) \right] \quad (12)$$

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^n x_i - \frac{\sum_{i=1}^n x_i \exp \left(\frac{-x_i}{\hat{\beta}} \right)}{\sum_{i=1}^n \exp \left(\frac{-x_i}{\hat{\beta}} \right)} \quad (13)$$

Teste de Kolmogorov-Smirnov (K-S)

De acordo com campos (1979), trata-se de um teste utilizado para verificar o ajuste de uma distribuição de probabilidade aos dados originais. É aplicado em distribuições populacionais contínuas e completamente especificadas. Tem por base a análise da proximidade ou ajuste entre a função de distribuição amostral $S(x)$ e a função de distribuição populacional sob a hipótese nula, $F_0(x)$.

RESULTADOS E DISCUSSÃO

Estimativas de Parâmetros e Resultados do Teste de Kolmogorov-Smirnov (K-S)

Na Tabela 1, pode-se notar que a distribuição Gumbel ajustou-se aos dados de precipitação máxima diária, nos meses de outubro, novembro, dezembro, janeiro, fevereiro e para o ano na cidade de Piracicaba (SP). O ajuste ocorreu em todos os métodos analisados.

XIX CONGRESSO DE PÓS-GRADUAÇÃO DA UFPA
27 de setembro a 01 de outubro de 2010

Tabela 1: Resultados do teste Kolmogorov-Smirnov (5%) referente à distribuição Gumbel, com estimativas de parâmetros pelos métodos dos momentos (Mom), da regressão (Regres), de Lieblein e da máxima verossimilhança (Máx. Ver.).

Métodos	Período	p-valor	Parâmetros		Período	P-valor	Parâmetros	
			α	β			α	β
Mom	Out	0,72	26,922	10,736	Jan	0,73	40,664	13,181
Regres		0,75	26,691	11,503		0,57	40,375	14,133
Lieblein		0,68	26,403	12,254		0,64	40,768	13,565
Máx. Ver.		0,67	26,533	12,363		0,37	40,252	15,337
Mom	Nov	0,69	30,342	14,373	Fev	0,70	35,443	18,140
Regres		0,61	30,064	15,349		0,55	35,090	19,370
Lieblein		0,44	31,064	13,738		0,67	35,682	17,720
Máx. Ver.		0,73	30,270	14,557		0,75	35,495	17,799
Mom	Dez	0,52	40,461	15,704	Ano	0,41	63,289	14,986
Regres		0,51	40,143	16,796		0,45	63,000	15,991
Lieblein		0,50	40,422	16,071		0,31	63,838	14,000
Máx. Ver.		0,49	40,263	16,770		0,43	63,280	14,580

Tempo de Retorno

Na Tabela 2, pode-se observar que para o período de retorno de 5, 25, 50 anos não houve diferença significativa de precipitação diária máxima provável entre os métodos analisados. Para o período de 100 anos, o método da máxima verossimilhança apresentou maior precipitação diária máxima provável no mês de janeiro em relação aos demais métodos, e nos meses de outubro e dezembro apresentou uma precipitação máxima equivalente aos métodos de Lieblein e da regressão respectivamente. Para o período de 200 anos, o método da máxima verossimilhança apresentou maior precipitação diária máxima provável nos meses de outubro e janeiro, e nos meses de dezembro e fevereiro apresentou precipitação máxima equivalente aos métodos da regressão e Lieblein respectivamente.

Tabela 2: Estimativas das precipitações máximas (mm) nos meses mais chuvosos: outubro, novembro, dezembro, janeiro, fevereiro e no ano, para os tempos de retorno 5, 25, 50, 100 e 200 anos obtidas pelos métodos dos momentos (Mom), da regressão (Regres), de Lieblein e da máxima verossimilhança (Máx. Ver.).

Métodos	Período	Tempo de retorno					Período	Tempo de retorno				
		5	25	50	100	200		5	25	50	100	200
Mom	Out	43	61	69	76	84	Jan	60	83	92	101	110
Regres		44	63	72	80	88		62	86	96	105	115
Lieblein		45	66	74	83	91		61	84	94	103	113
Máx.Ver.		45	66	75	83	92		63	89	100	111	121
Mom	Nov	52	76	86	96	106	Fev	63	93	106	119	132
Regres		53	79	90	101	111		64	97	111	124	138
Lieblein		52	75	85	94	104		62	92	105	117	130
Máx.Ver.		52	77	87	97	107		62	92	105	117	130
Mom	Dez	64	91	102	113	124	Ano	86	111	122	132	143
Regres		65	94	106	117	129		87	114	125	137	148
Lieblein		65	92	103	114	126		85	109	118	128	138
Máx.Ver.		65	94	106	117	129		85	110	120	130	140

XIX CONGRESSO DE PÓS-GRADUAÇÃO DA UFLA
27 de setembro a 01 de outubro de 2010

CONCLUSÕES

Para o período de retorno de 5, 25, 50 anos não houve diferença significativa de precipitação diária máxima provável entre os métodos analisados, para o período de 100 anos o método da máxima verossimilhança apresentou maior precipitação diária máxima provável no mês de janeiro e para o período de 200 anos nos meses de janeiro e outubro.

No entanto, o método da máxima verossimilhança forneceu estimativas de parâmetros para a distribuição Gumbel mais precisos e menos tendenciosos, seguido respectivamente do método de Lieblein, dos momentos e da regressão.

REFERÊNCIAS

ASSIS, F. N.; ARRUDA, H. V.; PEREIRA, A.R. **Aplicações de Estatística à Climatologia**. Pelotas: Ed. Universitária / UFPel, 1996. 161 p.

CAMPOS, H., de. **Estatística experimental não-paramétrica**. 3. ed. Piracicaba: ESALQ, 1979. 343p.

GUMBEL, E.J. **Statistical of Extremes**. New York: Columbia University Press, 1958. 375 p.

HAAN, C.T. **Statistical methods in hidrology**. Ames, Iowa: The Iowa State University Prees, 1977. 377 p.

LAWLESS, J.F. **Statistical models and methods for lifetime data** . New York : Wiley,1982. 580p.

R DEVELOPMENT CORE TEAM, **An Introduction to R:Version: 2.10.1** (2010). In: <
<http://www.r-project.org> >. Acesso em 4 de jan de 2010.

SILVA, M. M.; GONÇALVES A. R. T.; MAKINO, M. Distribuição de probabilidades de valores extremos da precipitação máxima de 24 horas de Belém do Pará. In: XI CONGRESSO BRASILEIRO DE METEOROLOGIA, 2000, Rio de Janeiro. Disponível em:
<<http://www.criatividadecoletiva.net/cbm-files/12-f6a674c4841fc07e50585d4218bf803a.PDF>>.
Acesso em: 05 Mai. 2009.