

AJUSTE DE SEIS DISTRIBUIÇÕES DENSIDADE DE PROBABILIDADE A SÉRIES HISTÓRICAS DE TEMPERATURA MÉDIA MENSAL, EM MOSSORÓ/RN.

Janilson Pinheiro de Assis¹, Ben Dêvide de Oliveira Batista², José Espínola Sobrinho³

¹ Departamento de Ciências Vegetais, Universidade Federal Rural do Semi-árido, Av. Francisco Mota, s/n, km 47 da Br 110, Bairro Presidente Costa e Silva, Cep. 59500-900, Mossoró, Rio Grande do Norte, Brasil. E-mail: janilson@ufersa.edu.br.

² Estudante de Graduação em Engenharia Agrônoma da UFERSA, Av. Francisco Mota, s/n, km 47 da Br 110, Bairro Presidente Costa e Silva, Cep. 59500-900, Mossoró, Rio Grande do Norte, Brasil. E-mail: ben.deivide@gmail.com.

³ Departamento de Ciências Ambientais, Universidade Federal Rural do Semi-árido, Av. Francisco Mota, s/n, km 47 da Br 110, Bairro Presidente Costa e Silva, Cep. 59500-900, Mossoró, Rio Grande do Norte, Brasil. E-mail: jespínola@ufersa.edu.br.

Apresentado no XVI Congresso Brasileiro de Agrometeorologia – 22 a 25 de Setembro de 2009 – GranDarrell Minas Hotel, Eventos e Convenções – Belo Horizonte – MG.

Resumo: No presente trabalho foram ajustadas sete distribuições densidades de probabilidade, Normal, Log-Normal, Beta, Gama, Log-Pearson (Tipo III), Gumbel, Weibull. Os critérios de ajuste utilizados para verificar a aderência das distribuições empíricas à distribuição teórica sob hipótese, e adotando um nível de significância de 10% de probabilidade, foram os testes de Kolmogorov-Smirnov, Qui-quadrado, Cramer Von-Mises, Anderson Darling, Kuiper, e o Logaritmo da Máxima Verossimilhança, totalizando assim, 504 testes. A série de temperatura foi constituída de 1970 a 2007. Verificou-se ajustamento para todos os modelos estudados, sendo as distribuições Normal e Log-Normal aquelas que melhor representam a série de temperatura, devido a parcimônia desses modelos. Deve-se salientar que o teste de Kolmogorov Smirnov apresenta um nível de aprovação de uma distribuição sob teste muito elevado, o que gera uma certa insegurança em relação aos critérios do teste, mas neste estudo de dados aproximadamente simétricos e séries pequenas, ele é o mais recomendado.

Palavras-chave: Temperatura, probabilidade, ajustamento.

Introdução

O estudo de distribuições de probabilidade de variáveis climáticas ao longo do tempo, como forma de compreender os fenômenos meteorológicos para determinar seus padrões de ocorrência e previsões probabilísticas, permitem ao pesquisador uma previsibilidade razoável do comportamento climático de uma região, sendo este fato, uma ferramenta de grande importância para o planejamento e gestão de diversas atividades agrícolas e humanas. O objetivo deste trabalho foi verificar o ajuste e a qualidade deste à série de dados de temperatura média mensal no município de Mossoró/RN, às distribuições densidade de probabilidade Normal, Log-Normal, Beta, Gama, Log-Pearson (Tipo III), Gumbel e Weibull, através dos critérios dos testes Kolmogorov-Smirnov, Qui-quadrado, Cramer Von-Mises, Anderson Darling, Küiper, e do Logaritmo da Máxima Verossimilhança.

Material e Métodos

Foram ajustados os seguintes modelos de funções densidade de probabilidade: Normal, Log-normal, Beta, Gama, Gumbel, Weibull e Log Pearson tipo III (CAMPOS, 1983; COOKE, 1993).

Segundo Morettin & Bussab (2003), a Normal é uma das mais importantes variáveis aleatórias contínuas, cuja distribuição é chamada distribuição normal ou Gaussiana, a qual serve como modelo de distribuição para muitos problemas da vida real, mas também aparece em muitas investigações teóricas, pois sua importância em análise matemática resulta do fato de que muitas técnicas estatísticas, como análise de variância e de regressão além de alguns testes de hipóteses assumem ou exigem a normalidade dos dados.

Na Distribuição log-normal os logaritmos das variáveis aleatórias são normalmente distribuídos. Nesse trabalho foi usada a distribuição a três parâmetros.

A função Gama de probabilidade possui dois parâmetros, o de forma (α) e o de escala (β). Thom (1958), citado por Miller & Weaver (1968), afirmou que para valores de α maiores ou igual a 100, a distribuição gama se aproxima da distribuição normal. O parâmetro de escala (β) indica o grau de dispersão entre os dados de uma série estudada.

A distribuição Weibull é utilizada em análise hidrológica para eventos extremos; contudo, sua utilização em séries históricas de variáveis climáticas e biológicas ainda é pouco conhecida (CATALUNHA et al., 2002). A distribuição densidade de probabilidade de Gumbel é outro modelo bastante utilizado na representação de dados climáticos, como por exemplo, para radiação solar e temperatura (ASSIS et. al., 2004)

A distribuição densidade de probabilidade Log-Pearson Tipo III, é usada para representar variáveis médias e extremas (SANSIGOLO, 2008).

Ao se ajustar uma série de dados a uma distribuição densidade de probabilidade, trabalha-se com a hipótese de que a distribuição pode representar adequadamente aquele conjunto de dados. Um critério de comprovar esta hipótese é através de alguns testes não paramétricos, sendo neste trabalho utilizado os testes de qui-quadrado (χ^2), Kolmogorov Smirnov, Cramer Von-Mises, Kuiper, Anderson-Darling, além do Logaritmo da Máxima Verossimilhança, como critério para indicar o grau de ajustamento (WORLEY et al., 1990; COOKE, 1993). Essa estatística possui poder para discriminar ajustes em que os demais testes fracassam em apontar (SHAPIRO e BRAIN, 1981).

Resultados e Discussão

Foram ajustados seis modelos de distribuição de probabilidade Normal, Log-Normal, Beta, Gama, Log-Pearson (Tipo III), Gumbel e Weibull, através dos critérios dos testes Kolmogorov-Smirnov, Qui-quadrado, Cramer Von-Mises, Anderson Darling, Küiper, e do Logaritmo da Máxima Verossimilhança Campos (1983) Cooke, 1993, cujos resultados indicaram um bom ajuste a essas funções de distribuição de probabilidade, sendo a maioria testadas com o valor-p $\geq 0,10$. Considerando um nível de significância dos testes igual a 0,05, ou seja, com probabilidade de um erro tipo I de 0,05, ou número de série de dados que não rejeitar H_0 , isto é, os dados que se ajustam à distribuição em teste, foi bastante pequena. Já através do Logaritmo da Máxima Verossimilhança, verificou-se ainda que nenhuma série de dados se ajustou ou aderiram aos modelos propostos. Isto se deve provavelmente ao número reduzido de dados da série histórica, pois quanto maior dessa estatística, ou seja, quanto maior é o valor do logaritmo da função de máxima verossimilhança, melhor é o ajuste obtido (SHAPIRO e BRAIN, 1981; COOKE, 1993; SANSIGOLO, 2008).

No entanto conforme Campos (1983) o teste de Kolmogorov-Smirnov, é mais indicado para grandes séries de valores, além do mais nesse teste não há necessidade de agrupamento dos dados em intervalos de classe como é o caso para a aplicação do teste de qui-quadrado, sendo assim nesse trabalho o teste que se mostrou mais conveniente é o de Kolmogorv-Smirnov. Sendo assim, pode-se inferir que entre as sete distribuições testadas, os parâmetros de qualquer uma delas poderiam ser utilizados para representar o comportamento

da temperatura média mensal. Porém, a obtenção da estimativa dos parâmetros dessas distribuições e a estimativa das probabilidades diferem quanto ao grau de dificuldade. Então convém verificar qual das funções estudadas tem o melhor ajuste, podem a mesma, coincidir com uma distribuição que apresente menor dificuldade de obtenção dos parâmetros e ainda facilidade nas estimativas de probabilidades (CARGNELUTTI FILHO et. al., 2004).

Entre as sete distribuições de probabilidade testadas na classe de um valor-p $\geq 0,10$, a distribuição Normal apresentou o maior número de aderências, revelada de forma mais freqüente através do teste de aderência de Kolmogorov-Smirnov, bastando estimar os parâmetros desta distribuição (média e desvio-padrão) para os cálculos de probabilidade dentro de limites de intervalos desejados (SANSIGOLO, 2008).

Conclusões

Os dados de temperatura média mensal se ajustam as distribuições de probabilidade Normal, Log-Normal, Beta, Gama, Log-Pearson (Tipo III), Gumbel e Weibull e apresentam melhor ajuste, a distribuição de probabilidade Normal. Os critérios de ajuste Kolmogorov-Smirnov, Qui-quadrado, Cramer Von-Mises, Anderson Darling, Küiper, foram semelhantes, podendo neste caso, serem utilizados como elemento indicador do melhor modelo. No entanto, o Logaritmo da Máxima Verossimilhança não se prestou como elemento indicador do ajuste.

Tabela 1. Freqüência do número de ajuste de séries históricas de temperatura média mensal a sete modelos de distribuição de densidade de probabilidade, Mossoró, RN, 2009.

MÊS	DISTRIBUIÇÕES DE PROBABILIDADE													
	Normal		Log-Normal		Beta		Gama		Log-Pearson III		Gumbel		Weibull	
Jan	5	11,11%	5	11,11%	4	10,00%	4	9,09%	5	10,87%	4	8,70%	5	11,36%
Fev	4	8,89%	4	8,89%	4	10,00%	4	9,09%	5	10,87%	4	8,70%	4	9,09%
Mar	4	8,89%	4	8,89%	4	10,00%	4	9,09%	4	8,70%	4	8,70%	3	6,82%
Abr	3	6,67%	3	6,67%	3	7,50%	3	6,82%	4	8,70%	3	6,52%	3	6,82%
Mai	3	6,67%	3	6,67%	3	7,50%	3	6,82%	4	8,70%	4	8,70%	4	9,09%
Jun	3	6,67%	3	6,67%	3	7,50%	3	6,82%	3	6,52%	3	6,52%	3	6,82%
Jul	5	11,11%	5	11,11%	3	7,50%	5	11,36%	4	8,70%	5	10,87%	5	11,36%
Ago	4	8,89%	5	11,11%	4	10,00%	5	11,36%	5	10,87%	5	10,87%	4	9,09%
Set	4	8,89%	4	8,89%	3	7,50%	4	9,09%	3	6,52%	4	8,70%	3	6,82%
Out	3	6,67%	3	6,67%	3	7,50%	3	6,82%	3	6,52%	3	6,52%	3	6,82%
Nov	4	8,89%	3	6,67%	3	7,50%	3	6,82%	3	6,52%	4	8,70%	4	9,09%
Dez	3	6,67%	3	6,67%	3	7,50%	3	6,82%	3	6,52%	3	6,52%	3	6,82%
TOTAL	45	100%	45	100%	40	100%	44	100%	46	100%	46	100%	44	100%

Tabela 2. Freqüência do número de ajuste de séries históricas de temperatura média mensal a sete modelos de distribuição de densidade de probabilidade, Mossoró, RN, 2009.

Distribuição de probabilidade	MÊS											
	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Normal	5	4	4	3	3	3	5	4	4	3	4	3
Log-Normal	5	4	4	3	3	3	5	5	4	3	3	3
Beta	4	4	4	3	3	3	3	4	3	3	3	3
Gama	4	4	4	3	3	3	5	5	4	3	3	3
Log-Pearson III	5	5	4	4	4	3	4	5	3	3	3	3
Gumbel	4	4	4	3	4	3	5	5	4	3	4	3
Weibull	5	4	3	3	4	3	5	4	3	3	4	3
TOTAL	32	29	27	22	24	21	32	32	25	21	24	21

Tabela 3. Valor-p dos testes de Qui-quadrado e Kolmogorov-Smirnov dos ajustes de sete distribuições densidade de probabilidade à séries de temperatura média mensal.

DATA DA SEMEADURA	NORMAL		LOG NORMAL		BETA		GAMA		LOG-PEARSON TIPO III		GUMBEL		WEIBULL	
	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2
JANEIRO	0,176	0,107	0,193	0,107	0,183	0,107	0,171	0,107	0,164	0,107	0,184	0,044	0,202	0,491
FEVEREIRO	0,143	0,864	0,136	0,864	0,264	0,077	0,136	0,864	0,154	0,864	0,173	0,024	0,172	0,024
MARÇO	0,118	0,888	0,114	0,888	0,225	0,000	0,107	0,888	0,110	0,888	0,149	0,021	0,231	0,013
ABRIL	0,244	0,000	0,223	0,000	0,415	0,000	0,234	0,000	0,307	0,000	0,235	0,000	0,271	0,000
MAIO	0,207	0,038	0,197	0,038	0,219	0,003	0,203	0,038	0,154	0,038	0,148	0,038	0,287	0,001
JUNHO	0,226	0,005	0,242	0,005	0,368	0,005	0,223	0,005	0,177	0,005	0,176	0,005	0,265	0,000
JULHO	0,154	0,185	0,166	0,185	0,166	0,005	0,149	0,185	0,181	0,185	0,157	0,185	0,142	0,185
AGOSTO	0,135	0,060	0,132	0,602	0,184	0,241	0,125	0,602	0,151	0,660	0,171	0,491	0,136	0,060
SETEMBRO	0,215	0,269	0,216	0,269	0,314	0,001	0,205	0,269	0,263	0,001	0,226	0,269	0,229	0,000
OUTUBRO	0,255	0,000	0,279	0,000	0,381	0,000	0,256	0,000	0,217	0,000	0,224	0,000	0,369	0,000
NOVEMBRO	0,160	0,080	0,176	0,004	0,203	0,075	0,164	0,075	0,208	0,075	0,193	0,075	0,195	0,075
DEZEMBRO	0,228	0,001	0,220	0,005	0,223	0,001	0,222	0,005	0,231	0,005	0,259	0,005	0,264	0,005

Tabela 3. Natureza dos ajustes em função dos testes de Qui-quadrado e Kolmogorov-Smirnov através de sete modelos de distribuição densidade de probabilidade à séries de temperatura média mensal.

DATA DA SEMEADURA	NORMAL		LOG NORMAL		BETA		GAMA		LOG-PEARSON TIPO III		GUMBEL		WEIBULL	
	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2	P-VALOR KS	P-VALOR χ^2
JANEIRO	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	-	AJ	AJ
FEVEREIRO	AJ	AJ	AJ	AJ	AJ	-	AJ	AJ	AJ	AJ	AJ	-	AJ	-
MARÇO	AJ	AJ	AJ	AJ	AJ	-	AJ	AJ	AJ	AJ	AJ	-	AJ	-
ABRIL	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-
MAIO	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-
JUNHO	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-
JULHO	AJ	AJ	AJ	AJ	AJ	-	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ
AGOSTO	AJ	-	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	AJ	-
SETEMBRO	AJ	AJ	AJ	AJ	AJ	-	AJ	AJ	AJ	-	AJ	AJ	AJ	-
OUTUBRO	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-
NOVEMBRO	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-
DEZEMBRO	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-	AJ	-

(1) AJ = Distribuição ajustada, - = Distribuição não ajustada (significativo a 10% de probabilidade).

Referências Bibliográficas

- ASSIS, J. P.; NETO, D. D.; MANFRON, P. A.; MARTIN, T. N.; SPAROVEK, G.; TIMM, L. C. Ajuste de series históricas de temperatura e radiação solar global diária às funções densidade de probabilidade normal e log-normal, em Piracicaba, SP. *Revista Brasileira de Agrometeorologia*, Santa Maria, v. 12, n.1, p. 113-121, 2004.
- CAMPOS, H. **Estatística não paramétrica**, Piracicaba: ESALQ/USP, 1983. 349 p. 4ª ed.
- CARGNELUTTI FILHO, Alberto; MATZENAUER, R.; TRINDADE, J. K. da. Ajustes de funções de distribuição de probabilidade à radiação solar global no Estado do rio Grande do Sul. *Pesq. Agropec. Brás.*, Brasília, v. 39, n. 12, p. 1157-1166, dezembro, 2004.
- CATALUNHA, M.J.; SEDIYAMA, G.C.; LEAL, B.G.; SOARES, C.P.B.; RIBEIRO, A. Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, v.10, n.1, p.153-162, 2002.
- COOKE, R.A. **VTFIT: A routine for fitting homogenous Probability density functions**. User documentation. Blacksburg: Department of agricultural Engineering. Virginia, Polytecnic Institute, 1993. 21p.
- COOKE, R.A.; MOSTAGHIMI, S.; WOESTE, F.E. **VTFIT: A microcomputer – based routine for fitting probability distribution functions to data**. American Society of Agricultural Engineers. St. Joseph, v.9, n.4, p.401-408, 1993.
- MILLER, M. E.; WEAVER, C. R. **Monthly and annual precipitation probabilities for climatic divisions in Ohio**. Research bulletin, n. 1005. 11p., 1968.
- MORETTIN, P.A.; BUSSAB, W.O. **Estatística básica**. , São Paulo: Saraiva, 2003. 526p.
- SANSIGOLO, C. A. Distribuição de extremos de precipitação diária, temperatura máxima e mínima e velocidade do vento em Piracicaba, SP (1917-2006). *Revista Brasileira de Meteorologia*, v. 23, n. 3, 341-346, 2008.
- SHAPIRO, S. S.; BRAIN, C. W. A review of distributional testing procedure and development of a censored sample distributional test. *In: Statistical distribution in scientific work*, v. 5. Dordrecht: D. Reidel Publishing Co. 1981.
- SHAPIRO, S. S.; BRAIN, C. W. A review of distributional testing procedure and development of a censored sample distributional test. *In: Statistical distribution in scientific work*, v.5, Dordrecht: D. Reidel Publishing Co. 1981.
- THOM, H.C.S. Some methods of climatological analysis. Roma: FAO, 1966. 50p. (Technical Notes, 81).
- WORLEY, J. W. *et al.* Graphic distribution analysis (GDA). **American Society Agricultural Engenharing**. , St. Joseph, v.6, n.3, p.367-371, 1990.
- WORLEY, J. W. *et. al.* Graphic distribution analysis (GDA). *Am. Soc. Agric. Eng.*, St. Joseph, v. 6, n. 3, p. 367-371, 1990.