

AJUSTE DE FUNÇÕES DE DISTRIBUIÇÃO DE PROBABILIDADE À TEMPERATURA MÁXIMA NO ESTADO DO RIO GRANDE DO SUL

Alberto Cargnelutti Filho¹, Ronaldo Matzenauer², Júlio Kuhn da Trindade³

ABSTRACT - The objective of this work was to verify the adjustment of data series for maximum temperature average decennial to the probability distribution functions normal, log-normal, gamma, gumbel and weibull. Data were collected from twenty three locations in Rio Grande do Sul State, Brazil. The Kolmogorov-Smirnov test was applied in the 828 series of data (23 localities x 36 periods of ten days) of maximum temperature average decennial to verify the adjustment of the data to the probability distributions normal, log-normal, gamma, gumbel and weibull, totalizing 4,140 tests. The data maximum temperature average decennial adjusted to the probability distributions functions normal, log-normal, gamma, gumbel and weibull, and present a better adjustment to the normal probability function.

INTRODUÇÃO

Na tomada de decisões relacionadas às atividades agropecuárias, turísticas e esportivas, é importante conhecer o comportamento de variáveis climáticas como a precipitação pluvial, temperatura e umidade relativa do ar, evaporação, direção e velocidade do vento, radiação solar global, ocorrência de orvalho, nevoeiro, granizo, geada e neve. Entre essas variáveis climáticas, a temperatura do ar, é fundamental, principalmente em relação às atividades agropecuárias.

O ajuste de uma função de distribuição de probabilidade da temperatura máxima é importante no cálculo de probabilidade de ocorrência de temperaturas extremas prejudiciais às culturas, principalmente nos períodos críticos. Com isto, o efeito negativo causado por temperaturas extremas do ar pode ser minimizado por meio da semeadura em época na qual o período crítico coincida com a menor probabilidade de ocorrência de temperaturas prejudiciais à cultura.

A simples visualização dos dados amostrais de uma variável em um histograma de frequência é insuficiente para inferir, entre as diversas funções de distribuição de probabilidade conhecidas, a que melhor se ajusta aos dados em estudo. Portanto, faz-se necessário o uso de testes de aderência, como o Qui-quadrado, Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk e Cramer-von Mises (Campos, 1983; Assis et al., 1996; Morettin & Bussab, 2004), para verificar se a distribuição de probabilidade dos dados de uma variável em análise pode ser estudada por uma função de distribuição de probabilidade conhecida.

O objetivo deste trabalho foi verificar o ajuste das séries de dados de temperatura máxima média decennial, de vinte e três municípios do Estado do Rio Grande do Sul às funções de distribuições de probabilidade normal, log-normal, gama, gumbel e weibull.

MATERIAL E MÉTODOS

Os dados diários de temperatura máxima (°C) das 23 estações meteorológicas, localizadas em vinte e três municípios do Estado do Rio Grande do Sul, foram obtidos no Banco de Dados do Laboratório de Agrometeorologia, da Fundação Estadual de Pesquisa Agropecuária – Fepagro/SCT-RS (Tabela 1).

O período considerado foi o de 1954 a 2004, totalizando 51 anos de observações. Em cada município, com os dados diários de temperatura máxima, estimou-se a média decennial dos 36 decênios do ano (1º decênio de janeiro até 3º decênio de dezembro), formando 828 séries temporais (23 municípios x 36 decênios), com número diferenciado de anos de observações em cada série em função da disponibilidade de dados meteorológicos (Tabela 1).

Tabela 1. Estações meteorológicas localizadas em 23 municípios do Estado do Rio Grande do Sul, período de coleta dos dados de temperatura máxima média decennial e número de observações (mínimo, máximo e média) entre as 36 séries (decênios) de dados em cada município.

Município	Período	Nº observações		
		Mn	Mx	Md
Alegrete	1968-2001	18	22	19
Cachoeirinha	1975-2004	29	30	29
Caxias do Sul	1986-2004	18	19	19
Cruz Alta	1973-2004	28	31	30
Encruzilhada do Sul	1958-2003	38	43	40
Erechim	1966-2004	36	39	38
Farroupilha	1963-2004	33	42	41
Ijuí	1963-2004	36	41	40
Júlio de Castilhos	1956-2004	45	47	46
Maquiné	1956-2004	47	49	48
Passo Fundo	1961-2004	43	44	44
Quaraí	1966-2004	35	39	37
Rio Grande	1954-2004	43	46	44
Santa Maria	1963-2004	38	42	41
Santa Rosa	1975-2004	29	30	30
Santana do Livramento	1965-2004	16	21	18
Santo Augusto	1966-1986	17	19	17
São Borja	1956-2004	43	45	44
São Gabriel	1963-2004	38	41	40
Taquari	1963-2004	40	42	42
Uruguaiana	1963-2004	33	39	36
Vacaria	1966-1994	22	25	24
Veranópolis	1956-2004	45	48	47
Todos os municípios		16	49	35

Aplicou-se o teste de aderência de Kolmogorov-Smirnov (Campos, 1983; Assis et al., 1996) nas 828 séries de dados de temperatura máxima média decennial, para verificar o ajuste dos dados às distribuições normal, log-normal, gama, gumbel e weibull (Assis et al., 1996; Morettin & Bussab, 2004; Spiegel et al., 2004), totalizando 4.140 testes (828 séries de dados x 5 distribuições).

¹ Eng. Agr., Dr., Pesquisador em Estatística/Experimentação Agropecuária da Fundação Estadual de Pesquisa Agropecuária (FEPAGRO), Rua Gonçalves Dias, 570, 90130-060, Porto Alegre, RS. E-mail: alberto-cargnelutti@fepagro.rs.gov.br

² Eng. Agr., Dr., Pesquisador em Agrometeorologia – FEPAGRO. E-mail: ronaldo-matzenauer@fepagro.rs.gov.br

³ Acadêmico de Agronomia da UFRGS - Estagiário do Laboratório de Agrometeorologia da FEPAGRO

RESULTADOS E DISCUSSÃO

O valor da estatística (D máximo) do teste de aderência de Kolmogorov-Smirnov (Campos, 1983; Assis et al., 1996; Morettin & Bussab, 2004) informa a máxima distância entre as probabilidades empíricas e as teóricas obtidas sob a função de distribuição de probabilidade em teste. Assim, menores valores da estatística fornecem maiores valores de p-valor e, conseqüentemente, maior evidência de não - rejeição da hipótese nula (H_0), ou seja, maior aderência dos dados a distribuição em teste.

Independentemente da função de distribuição de probabilidade (normal, log-normal, gama, gumbel e weibull), 3.851 (93,02%), dos 4.140 casos analisados (cinco distribuições x 23 municípios x 36 decêndios), se ajustaram às funções de distribuições de probabilidades testadas com p-valor $\geq 0,20$, ou seja, com boa aderência (Tabela 2). No nível de significância do teste de 0,05, ou seja, com probabilidade de erro tipo I de 0,05, o número de séries de dados que não rejeitaram H_0 , isto é, os dados que se ajustam à distribuição em teste, elevou-se para 3.919 (94,66%). Portanto, apenas 5,34% dos casos não aderiram a nenhuma das distribuições neste nível de significância. Assim, pode-se inferir que entre as cinco distribuições testadas, os parâmetros de qualquer uma delas poderiam ser utilizados para representar o comportamento da temperatura máxima média decendial. Porém, a obtenção da estimativa dos parâmetros dessas distribuições e o cálculo das probabilidades diferem quanto ao grau de dificuldade. Então, convém verificar qual entre as funções estudadas tem o melhor ajuste, podendo a mesma coincidir com uma distribuição que apresente menor dificuldade de obtenção dos parâmetros e ainda, facilidade nas estimativas de probabilidades.

Tabela 2. Frequência absoluta (f_i), frequência acumulada (F_i), frequência relativa absoluta (fr_i) e frequência relativa acumulada (Fr_i) de ajustes das 828 séries de dados de temperatura máxima média decendial às funções de distribuições de probabilidade normal, log-normal, gama, gumbel e weibull em vinte e três municípios do Estado do Rio Grande do Sul, em cada classe de p-valor, pelo teste de aderência de Kolmogorov-Smirnov.

p-valor	f_i	F_i	fr_i (%)	Fr_i (%)
p<0,01	216	216	5,22	5,22
p<0,05	5	221	0,12	5,34
p<0,10	12	233	0,29	5,63
p<0,15	14	247	0,34	5,97
p<0,20	42	289	1,01	6,98
p \geq 0,20	3.851	4.140	93,02	100,00

Entre as cinco funções de distribuições de probabilidades testadas na classe de p-valor $\geq 0,20$, as distribuições normal, log-normal e gama apresentaram os maiores números de aderências (826 séries de dados), ou seja, apenas duas séries de dados (0,24%) não se ajustaram a essas distribuições, em nível de significância de 0,20, evidenciando boa aderência dos dados a estas distribuições. Em outro extremo, a distribuição weibull apresentou os menores índices de aderência. Esta distribuição apontou 220 dos 221 casos

com p-valor<0,05. A distribuição gumbel está em situação intermediária (Tabela 3).

Tabela 3. Frequência absoluta (f_i) e frequência relativa absoluta (fr_i) de casos em que os dados de 828 séries de temperatura máxima média decendial se ajustaram às funções de distribuições de probabilidade normal, log-normal, gama, gumbel e weibull em vinte e três municípios do Estado do Rio Grande do Sul em cada classe de p-valor pelo teste de aderência de Kolmogorov-Smirnov.

p-valor	Normal		Log-normal		Gama		Gumbel		Weibull	
	f_i	fr_i	f_i	fr_i	f_i	fr_i	f_i	fr_i	f_i	fr_i
p<0,01	0	0,00	0	0,00	0	0,00	0	0,00	216	26,09
p<0,05	0	0,00	0	0,00	0	0,00	1	0,12	4	0,48
p<0,10	0	0,00	0	0,00	0	0,00	3	0,36	9	1,09
p<0,15	1	0,12	0	0,00	0	0,00	6	0,72	7	0,85
p<0,20	1	0,12	2	0,24	2	0,24	13	1,57	24	2,90
p \geq 0,20	826	99,76	826	99,76	826	99,76	805	97,22	568	68,60
Total	828	100,00	828	100,00	828	100,00	828	100,00	828	100,00

O teste de aderência de Kolmogorov-Smirnov revelou as distribuições normal, log-normal e gama adequadas ao estudo para representar temperatura máxima média decendial, bastando estimar os parâmetros destas distribuições, para os cálculos de probabilidade dentro de limites de intervalo desejados. No entanto, pela facilidade de obtenção das estimativas dos parâmetros da distribuição normal (média e desvio - padrão) esta torna-se mais adequada, para estimar as probabilidades acima ou abaixo de qualquer valor de temperatura máxima média decendial, visando minimização de riscos.

Então, os dados de temperatura máxima média decendial se ajustam às funções de distribuições de probabilidade normal, log-normal, gama, gumbel e weibull e apresentam melhor ajuste à função de distribuição de probabilidade normal.

REFERÊNCIAS

- ASSIS, F.N.de; ARRUDA, H.V.de; PEREIRA, A.R. **Aplicações de estatística à climatologia: teoria e prática**. Pelotas: UFPEL, 1996. 161 p.
- CAMPOS, H. de **Estatística experimental não-paramétrica**. 4.ed. Piracicaba: Departamento de Matemática e Estatística - ESALQ, 1983. 349 p.
- MORETTIN, P.A.; BUSSAB, W.O. **Estatística básica**. 5 ed. São Paulo: Saraiva, 2004. 526 p.
- SPIEGEL, R.A.; SCHILLER, J.; SRINIVASAN, R.A. **Probabilidade e estatística**. 2 ed. Porto Alegre: Bookman, 2004. 398 p.