

Aplicação de seis distribuições de probabilidade a séries de temperatura máxima em Iguatu - CE¹

Evaluation of six distributions of probability at maximum temperature series in Iguatu - CE, Brazil

Eliakim Martins Araújo^{2*}, Italo Nunes Silva³, Joaquim Branco de Oliveira⁴, Edmilson Gomes Cavalcante Junior⁵ e Bruno Marçal de Almeida⁶

Resumo - Esse estudo teve como objetivo avaliar a distribuição de probabilidade que melhor se ajusta a séries diárias de temperatura máxima da cidade de Iguatu no Ceará, situada no sertão nordestino, com clima, segundo Köppen, do tipo BSw'h'. Foram utilizadas as funções distribuição de probabilidade Gumbel I, Gama, Log-normal, Normal, Weibull e Beta. A série estudada abrangeu um período de 45 anos (1961 a 2005), obtidos do 3º Distrito do Instituto Nacional de Meteorologia (INMET). O ajuste das distribuições à série de dados foi realizado através dos testes de aderência Qui-quadrado e Lilliefors ambos ao nível de significância de 5%. A distribuição que melhor se ajustou à série de dados de temperatura máxima de Iguatu - CE foi a Normal e a de pior desempenho foi a distribuição Gumbel tipo I.

Palavras-chave - Análise de séries temporais. Distribuição de frequência. Teste de aderência.

Abstract - This study had as objective to evaluate the distribution of probability that better adapts to daily series of maximum temperature in Iguatu city, Ceará, located in the Northeastern Brazil, with climate, according to Köppen, of the type BSw'h'. Were used the distribution of probability functions Gumbel I, Gama, Log-normal, Normal, Weibull and Beta. The studied series included a period of 45 years (1961 to 2005), obtained from the 3rd District of the Instituto Nacional de Meteorologia (INMET). The adherence of the distributions was tested by the tests Chi-square and Lilliefors both to the level of significance of 5%. The distribution that best fit the data series of maximum temperature for Iguatu - CE was the Normal distribution and the worse was the distribution of Gumbel type I.

Key words - Time series. Distributions of frequency. Adherence tests.

* Autor para correspondência

¹Recebido para publicação em 13/04/2009; aprovado em 05/01/2010

Trabalho de pesquisa acadêmica desenvolvido no laboratório de Topografia e Geoprocessamento do IFCE - Campus Iguatu

²Instituto Federal de Educação Ciências e Tecnologia do Ceará/IFCE - Campus Iguatu, Rodovia Iguatu, Varzea Alegre, Km 05, Sítio Cajazeiras-CE, Brasil, 63500-000, eliakim.araujo@bol.com.br

³Instituto Federal de Educação Ciências e Tecnologia do Ceará/IFCE - Campus Iguatu, Sítio Cajazeiras-CE, Brasil, italobolinha@gmail.com

⁴Instituto Federal de Educação Ciências e Tecnologia do Ceará/IFCE - Campus Iguatu, Sítio Cajazeiras-CE, Brasil, joaquimbrancodeoliveira@gmail.com

⁵Instituto Federal de Educação Ciências e Tecnologia do Ceará/IFCE - Campus Iguatu, Sítio Cajazeiras-CE, Brasil, ediguatu@yahoo.com.br

⁶Instituto Federal de Educação Ciências e Tecnologia do Ceará/IFCE - Campus Iguatu, Sítio Cajazeiras-CE, Brasil, br255@hotmail.com

Introdução

O clima vem recebendo atenção especial do Homem ao longo do tempo; a temperatura, um dos principais fatores de caracterização climática, exerce diversos efeitos na prática agrícola, por isso é um dos elementos meteorológicos mais comumente medidos.

Estefanel et al. (1994), afirmam que este elemento exerce dois efeitos bioclimáticos ao longo do ano: nos meses frios tem ação desvernalizante sobre as criófitas e nos meses quentes, pela ação deprimente, tanto no crescimento como no desenvolvimento das plantas.

Segundo Costa e Sediya (1999), os efeitos da temperatura máxima do ar nas culturas podem ser: influência na duração da fase de enchimento de grãos, na taxa de respiração e na taxa de expansão foliar. Todos esses efeitos podem provocar redução na produtividade das culturas da ordem de 5 a 40%.

O estudo das distribuições de variáveis, como um meio de compreender os fenômenos meteorológicos, para determinar seus padrões de ocorrência e permitir uma previsibilidade razoável do comportamento climático de uma região, é uma ferramenta de grande valor para o planejamento e gestão de inúmeras atividades agropecuárias e humanas. As previsões probabilísticas auxiliam no planejamento e condução das atividades agropecuárias, ao racionalizar os procedimentos e evitar ou minimizar os possíveis prejuízos causados pela ação das intempéries (ASSIS et al., 2004).

Para Catalunha et al. (2002), o uso de funções densidade de probabilidade está diretamente ligado à natureza dos dados a que elas se relacionam. Algumas têm boa capacidade de estimação para pequeno número de dados, outras requerem grande série de observações. Desde que respeitado o aspecto da representatividade dos dados, as estimativas dos seus parâmetros para uma determinada região podem ser estabelecidas como de uso geral, sem prejuízo da precisão na estimação da probabilidade.

As distribuições de probabilidade contínuas são bastante utilizadas em diversos trabalhos probabilísticos (ASSIS et al., 2004; CATALUNHA et al., 2002; JUNQUEIRA JÚNIOR et al., 2007; LYRA et al., 2006), devido o ajustamento de suas variáveis, que pode até não ser perfeito, mas descrevem bem uma situação real, fornecendo respostas às hipóteses que podem ter sido levantadas na pesquisa. Segundo Ferreira (2005), as variáveis aleatórias das distribuições contínuas são aquelas que assumem seus valores em uma escala real, modeladas por uma função de densidade $f(x)$ com as seguintes propriedades:

a) o valor de $f(x)$ é sempre igual ou superior a zero;

b) a área sob a curva estabelecida pela densidade e delimitada pelo eixo da abscissa é igual à unidade, se for considerado o domínio da variável X.

A utilização de funções de distribuição de probabilidade requer o uso de testes para provar a adaptação dos dados ou da série de dados às funções. Esses testes são conhecidos como testes de aderência e sua real função é verificar a forma de uma distribuição, através da análise da adequação dos dados à curva de um modelo de distribuição hipotética.

De acordo com Assis et al. (1996), os testes de aderência Qui-quadrado, Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Cramer-von Mises, servem para comparar as probabilidades empíricas de uma variável com as probabilidades teóricas estimadas pela função de distribuição em teste, verificando se os valores da amostra podem ser provenientes de uma população com aquela distribuição teórica.

O objetivo deste estudo foi analisar qual, entre seis distribuições de probabilidade, melhor se ajusta a uma série histórica de dados diários de temperatura máxima do ar da cidade de Iguatu, no Ceará.

Material e métodos

Foram utilizados dados diários de temperatura máxima da cidade de Iguatu (Figura 1), obtidos junto ao 3º Distrito do Instituto Nacional de Meteorologia (INMET) dos anos de 1961 a 2005, compreendendo assim, uma série de 45 anos, cuja localização é 6,22° latitude Sul e 39,18° longitude Oeste (Figura 2), situada no estado do Ceará que possui características climáticas tipicamente tropicais (SILVA, 1987).

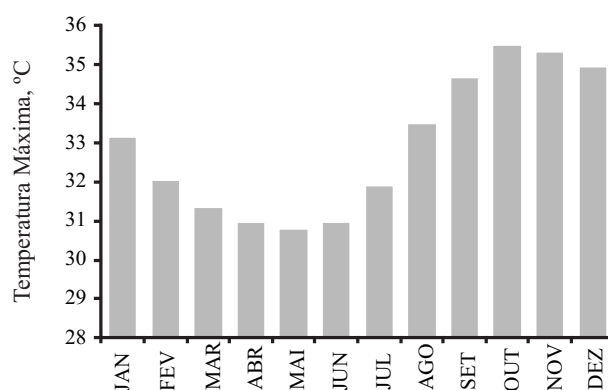


Figura 1 - Médias mensais da temperatura máxima para a cidade de Iguatu

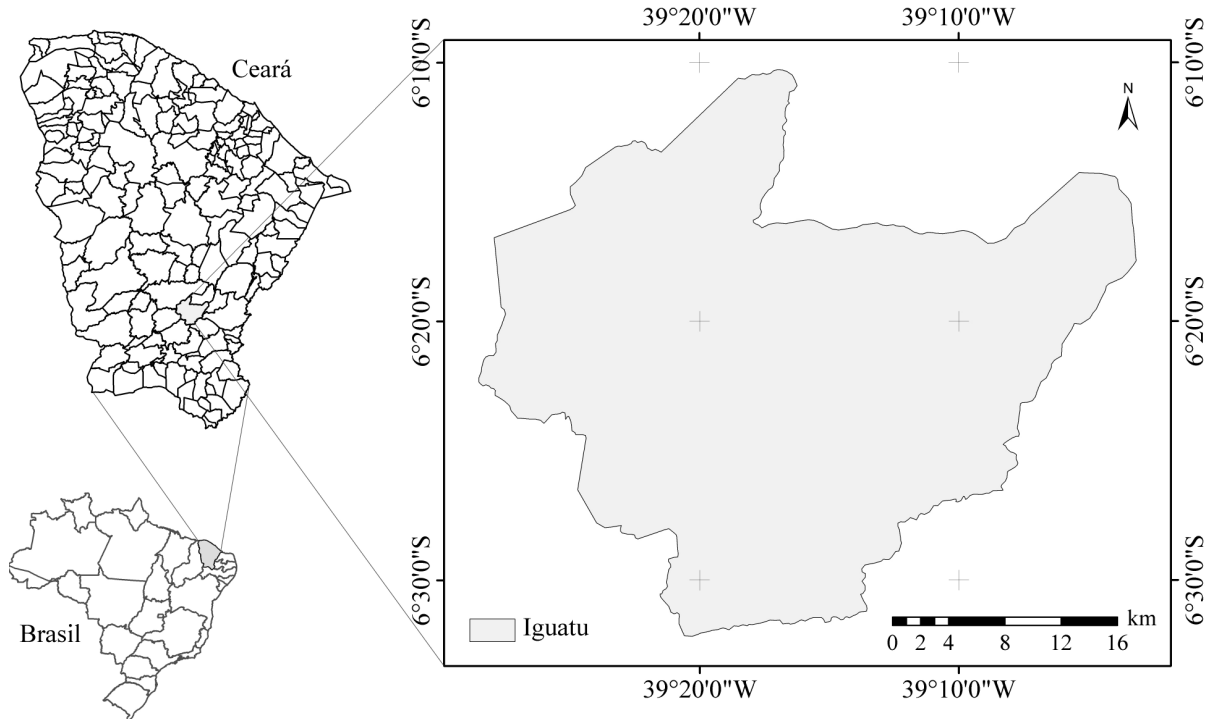


Figura 2 – Localização da cidade em estudo

A cidade de Iguatu acha-se inserida na região fisiográfica do sertão nordestino, compreendendo, segundo a classificação de Köppen, o tipo de clima: BSw'h', clima quente e semi-árido, caracterizado pela insuficiência das chuvas, com temperaturas elevadas acarretando numa forte evaporação, e tendo apenas duas estações climáticas bem definidas durante o ano, chuvosa e seca (SUDENE, 1973).

A homogeneidade dos dados foi testada segundo a metodologia proposta por Thom (1966), sendo constatado que a série é homogênea. Em toda a série não foram encontradas falhas, fato esse que não exigiu nenhum tipo de tratamento especial aos dados.

Os dados diários foram separados por mês e agrupados em classes conforme os procedimentos adotados por Doane (1976) e demonstrados pelas Equações 1 a 4:

$$k = 1 + \log_2 \left(1 + \frac{\sqrt{b_1}}{S_{\sqrt{b_1}}} \right) + \log_2 (n) \quad (1)$$

onde,

n - número total de observações

$S_{\sqrt{b_1}}$ - erro padrão do coeficiente de assimetria, $\sqrt{b_1}$ os quais são dados pelas equações 2 e 3.

$$\sqrt{b_1} = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n \left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \right]^{3/2}} \quad (2)$$

$$S_{\sqrt{b_1}} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}} \quad (3)$$

O intervalo de classe (c), ou seja, o comprimento das classes foi determinado da seguinte forma:

$$c = \left[\frac{(X_x + 0,01) - (X_n - 0,01)}{k} \right] \quad (4)$$

onde,

X_x - maior valor da série,

X_n - menor valor da série.

Foram aplicadas as distribuições de probabilidade: Gumbel I, Gama, Log-normal, Normal, Weibull e Beta.

Após a aplicação das distribuições de probabilidade, suas respectivas aderências foram testadas utilizando os testes Qui-quadrado e Lilliefors ambos ao nível de significância de 5%.

a) Distribuição Gumbel I - tem como uma das formas de apresentar sua função densidade de probabilidade (BEIJO et al., 2005; NAGHETTINI; PINTO, 2007):

$$F(x) = \frac{1}{\beta} e^{-\left(\frac{x-\alpha}{\beta}\right)} - e^{-\left(\frac{x-\alpha}{\beta}\right)} \quad (5)$$

onde,

α - parâmetro de posição,

β - é o parâmetro de escala.

Sua função de distribuição acumulada é dada pela equação:

$$F(x) = e^{-e^{\pm\left(\frac{x-\alpha}{\beta}\right)}} \quad (6)$$

onde,

\pm - no segundo expoente, refere-se aos valores extremos máximos (sinal negativo) e mínimos (sinal positivo).

b) Distribuição Gama - se x for uma variável aleatória contínua, tal que ($0 < x < \infty$), com distribuição Gama de parâmetros $\alpha > 0$ e $\beta > 0$, então a sua função densidade de probabilidade é definida como (BOTELHO; MORAIS, 1999; JUNQUEIRA JÚNIOR et al., 2007):

$$F(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}; \text{ para } 0 < x < \infty \quad (7)$$

Algumas formas de estimar os parâmetros da distribuição Gama foram desenvolvidas, contribuindo, junto com a sua flexibilidade de formas, para sua utilização em diversas áreas (HAAN, 1977; apud CATALUNHA et al., 2002). O principal método para estimar seus parâmetros é o método de máxima verossimilhança; que para satisfazer a condição $\alpha > 0$ (por definição) e após algumas considerações matemáticas na equação original, Equação 8, a solução que interessa será:

$$\hat{\alpha} = \frac{1 + \sqrt{1 + \frac{4[\ln(x) - x_g]}{3}}}{4[\ln(\bar{x}) - x_g]} \quad (8)$$

O estimador do parâmetro β poderá ser obtido por:

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \quad (9)$$

onde,

\bar{x} - média aritmética,

x_g - média geométrica das observações.

Sendo $F(x)$ a probabilidade de ocorrência de um evento menor ou igual a x , pode-se escrever que a função de distribuição acumulada de probabilidade é representada pela função Gama incompleta, segundo Thom (1958) apud Catalunha et al. (2002):

$$F(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x u^{\alpha-1} e^{-\frac{u}{\beta}} du \quad (10)$$

onde,

$F(x)$ - probabilidade de ocorrer um valor menor ou igual a x ,

x - variável aleatória contínua,

$\Gamma(\alpha)$ - função Gama do parâmetro alfa,

α - parâmetro de forma da variável aleatória x ,

β - parâmetro de escala da variável aleatória x ,

e - base do logaritmo neperiano (2,718281828...),

u - variável aparente utilizada para integração.

A função de distribuição acumulada da distribuição de probabilidade Gama possui uma integral, que pode ser resolvida por métodos numéricos ou pelo desenvolvimento em série de uma expressão exponencial. Para isso, utilizando-se de transformação de variáveis, tal que $t = u/\beta \rightarrow du = \beta dt$, não mudando o limite inferior zero e substituindo x por t , e após as simplificações adequadas, tem-se, da Equação 8:

$$F(t) = \frac{1}{\Gamma(\alpha)} \int_0^t t^{\alpha-1} e^{-t} dt \quad (11)$$

Considerando o desenvolvimento em série de e^{-t} , multiplicando-se e dividindo-se a equação anterior por e^{-t} , reunindo, do produto resultante, os termos em t , t^2 , t^3 e os demais termos correspondentes aos diversos expoentes, têm-se, segundo Assis et al. (1996):

$$F(t) = \frac{t^\alpha}{\alpha \Gamma(\alpha)} \left[1 + \frac{t}{\alpha + 1} + \frac{t^2}{(\alpha + 1)(\alpha + 2)} + \frac{t^3}{(\alpha + 1)(\alpha + 2)(\alpha + 3)} + \dots \right] \quad (12)$$

onde,

t - x/β , uma aproximação da distribuição Gama por expansão em série.

c) Distribuição Log-normal - conforme Ribeiro et al. (2007), a função densidade da distribuição Log-normal a dois parâmetros e a três parâmetros são representadas pela seguinte equação:

$$F(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} e^{-\frac{[\ln(x-a)-\mu]^2}{2\sigma^2}} \quad (13)$$

onde,

F(x) - função densidade de probabilidade da variável,

e - base do logaritmo neperiano,

x - valor da variável aleatória,

μ - média dos logaritmos da variável x,

σ - desvio-padrão dos logaritmos da variável x,

a - limite inferior da amostra.

Para encontrar a probabilidade de que uma variável aleatória x tendo distribuição Log-normal, assuma valores entre a e b ($a \leq x \leq b$), tem-se:

$$F(a \leq x \leq b) = \int_a^b \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(x-a)-\mu]^2}{2\sigma^2}\right) dx \quad (14)$$

O valor de “a” pode ser zero, quando se considera a distribuição Log-normal a dois parâmetros, ou um valor mínimo da série, quando se considera Log-normal a três parâmetros. O valor de “b” pode ser o da variável aleatória, quando se considera a probabilidade cumulativa de ocorrência daquele valor (HASTINGS; PEACOCK, 1975; apud CATALUNHA et al., 2002).

d) Distribuição Normal – tem sua função densidade de probabilidade da seguinte forma (ABRAMOVITZ; STEGUN, 1972; apud ASTOLPHO et al., 2005; KORTUM, 2002; apud MARTIN et al., 2007):

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}; \text{ para } -\infty < x < +\infty \quad (15)$$

onde,

μ - média,

σ - desvio padrão da variável aleatória.

A probabilidade de que uma variável x assuma valores menores ou iguais a x quando ela tem $N(\mu, \sigma^2)$, distribuição Normal com média μ e variância σ^2 , é estimada por:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (16)$$

Essa equação não pode ser resolvida analiticamente sem o uso de métodos de integração aproximada. Por esse motivo, usa-se a transformação $Z = (x - \mu)/\sigma$, a variável Z tem distribuição Normal com média zero e variância um $[N(0,1)]$. A variável Z é chamada variável reduzida e a Equação 12 pode ser reescrita na seguinte forma:

$$F(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Z e^{-\frac{Z^2}{2}} dz; \text{ para } -\infty < Z < +\infty \quad (17)$$

e) Distribuição Weibull - tem sua função de densidade de probabilidade mais comumente apresentada da seguinte forma (CATALUNHA et al., 2002):

$$F(x) = \frac{\gamma}{\beta} \left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{x-\alpha}{\beta}\right)^\gamma\right]; \text{ para } x \geq \alpha \quad (18)$$

onde,

F(x) = 0 para outros intervalos,

x - variável aleatória,

$\alpha \geq 0, \beta > 0, \gamma > 0$ os parâmetros da distribuição.

f) Distribuição Beta - é uma distribuição definida no intervalo (0,1) com a seguinte função densidade de probabilidade (ASSIS et al., 1996):

$$F(x) = \frac{x^{(\alpha-1)}(1-x)^{\beta-1}}{\beta^{(\alpha,\beta)}} \quad (19)$$

onde,

$0 < X < 1$ e $\alpha, \beta > 0$.

B(α, β) - função beta; definida por:

$$B(\alpha, \beta) = \int_0^1 X^{\alpha-1}(1-X)^{\beta-1} dX \quad (20)$$

Os testes de aderência usados no estudo são assim definidos:

Teste Qui-quadrado (χ^2) - a hipótese de nulidade admite que a distribuição seja a especificada (Gama, Weibull etc...), com os seus parâmetros estimados com base nos dados amostrais. A hipótese é testada fazendo-se a comparação entre as frequências observadas e as frequências teóricas, em cada classe de frequência da amostra, com a variável aleatória do χ^2 , dada por Campos (1979) citado por Catalunha et al. (2002):

$$\chi^2 = \sum_{i=1}^k \left(\frac{(Fo_i - Fe_i)^2}{Fe_i} \right) \quad (21)$$

onde,

k - número de classes,

Fo_i - frequência observada,

Fe_i - frequência esperada, de acordo com a distribuição que está sendo testada.

Os valores críticos do χ^2 para alguns níveis de significância são descritos por tabelas próprias.

Teste de Lilliefors - é um teste amplo, uma adaptação do teste de Kolmogorov-Smirnov, usado para verificação de normalidade de um conjunto de dados, que tem sua função distribuição empírica para a hipótese nula (composta) de normalidade. A estatística de teste é diferença máxima absoluta entre a função distribuição acumulada hipotética e empírica. Ele pode ser calculado como (WILKS, 2006):

$$D = \max(D^+, D^-) \quad (22)$$

onde,

$$D^+ = \max_{i=1, \dots, n} \left(\frac{i}{n} - p_{(i)} \right) \quad (23)$$

$$D^- = \max_{i=1, \dots, n} \left(p_{(i)} - \frac{i-1}{n} \right) \quad (24)$$

$$p_{(i)} = \Phi \left(X_{(i)} - \frac{\bar{X}}{S} \right) \quad (25)$$

Φ - função distribuição acumulada da distribuição normal padrão;

\bar{X} - média dos valores;

S - desvio padrão dos valores

Resultados e discussão

Através da análise de distribuição de frequência foi possível ajustar as seis funções distribuição de probabilidade (Figura 3) para os dados diários de temperatura máxima, separados por mês, para todos os meses do ano. Observando-se a Figura 3, é possível notar que as temperaturas máximas diárias da cidade de Iguatu têm sua maior frequência no intervalo entre 32 °C e 35 °C e que a partir das curvas teóricas de cada distribuição contínua existem distribuições (como é o caso da Normal e da Gama) que podem representar a série melhor que outras (Gumbel).

A aderência das distribuições para análise dos dados segundo o teste χ^2 (Tabela 1) demonstra a rigorosidade desse teste, já que poucos foram os ajustamentos aos dados. O uso desse teste pode ser considerado limitado. De acordo com Ferreira (2005), o teste é sensível à presença de classes com pequenos valores das frequências esperadas. A solução alternativa para solucionar o problema que esse teste tem de trabalhar com classes com pequenos valores é agrupar classes adjacentes cujas frequências, de ambas ou de uma delas apenas, sejam inferiores a esses limites.

A Tabela 2 mostra a aderência das distribuições segundo o teste de Lilliefors. Este teste, assim como o Kolmogorov-Smirnov (K-S), é baseado no módulo da maior diferença entre a probabilidade observada e a estimada, não ocorrendo o aspecto cumulativo dos erros.

Segundo Catalunha et al. (2002), o nível de aprovação do K-S (base do teste de Lilliefors) numa distribuição de probabilidade sob teste é elevado, o que gera uma insegurança em relação aos critérios do teste, no entanto, isto é bem provável quando se ajustam distribuições assimétricas as quais apresentam maiores valores nas classes iniciais e menores nas finais. Assim, quando se aplicam esses valores à equação de definição do teste de χ^2 , tem-se um somatório dos erros absolutos. Por outro lado, aplicando-se a equação de definição do teste de Lilliefors (a mesma do K-S), tem-se um único valor, o módulo da diferença. Isto mostra que os erros, no teste do χ^2 , são considerados de forma cumulativa e em todas as classes e que, no teste de Lilliefors, eles são considerados somente na classe em que foi maior.

Os valores calculados do χ^2 são comparados com valores críticos ou tabelados referenciados pelo nível de significância e pelo grau de liberdade. Já o Lilliefors tem os valores calculados comparados a tabelas próprias (baseadas nas tabelas do K-S) pelo nível de significância e pelo número de observações.

Um fator que diferencia bem o teste de Lilliefors do χ^2 é que o valor crítico ou tabelado depende, unicamente, do número de observações algo que não varia de distribuição para distribuição, estando em

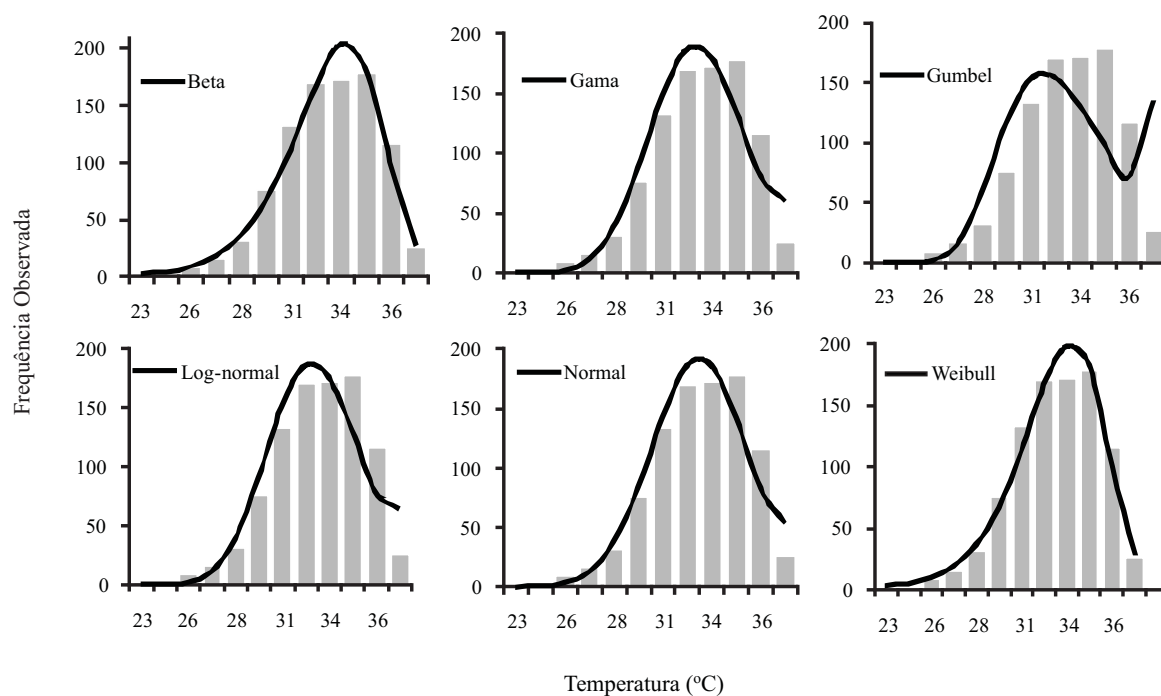


Figura 3 - Histograma das temperaturas máximas diárias e respectivas curvas teóricas propostas

Tabela 1 - Aderência das funções distribuição de probabilidade para o teste Qui-quadrado. Onde (*) - significativo e (ns) - não significativo, ao nível de 5% de probabilidade

Função distribuição de probabilidade	Aderência para o qui-quadrado											
	J	F	M	A	M	J	J	A	S	O	N	D
Gumbel I	*	*	*	*	*	*	*	*	*	*	*	*
Gama	*	*	*	*	*	ns	*	*	*	*	*	*
Log-normal	*	*	*	*	*	ns	*	*	*	*	*	*
Normal	*	*	*	*	*	ns	*	*	*	*	*	*
Weibull	*	*	*	*	*	*	*	*	*	*	*	*
Beta	*	*	*	*	*	*	*	*	*	*	*	*

Tabela 2 - Aderência das funções de distribuição de probabilidade para o teste de Lilliefors. Onde (*) - significativo e (ns) - não significativo, ao nível de 5% de probabilidade

Função distribuição de probabilidade	Aderência para o Lilliefors											
	J	F	M	A	M	J	J	A	S	O	N	D
Gumbel I	ns	ns	*	*	*	*	*	*	*	*	*	*
Gama	ns	ns	ns	ns	ns	ns	*	ns	ns	ns	*	ns
Log-normal	ns	ns	ns	ns	ns	ns	*	*	ns	ns	*	ns
Normal	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Weibull	ns	ns	*	ns	*	*	ns	*	*	*	*	*
Beta	ns	*	*	*	*	ns	ns	ns	ns	*	*	*

função apenas da série estudada. Já no teste de χ^2 , verifica-se que o grau de liberdade depende do número de parâmetros da distribuição, no caso igual a dois, e do número de classes.

Neste estudo, a ocorrência de agrupamentos de classes para evitar o uso de frequências menores que três ou cinco foi um fator bastante limitante para o χ^2 , pois em todos os meses foi utilizada essa restrição de uso, analisando-se que o grau liberdade fica menor, reduzindo o valor tabelado para o χ^2 . Isto faz com que seu nível de restrição seja ainda maior e aconteçam poucas aderências aos dados. Essa falta de aderência faz com que se torne impossível determinar qual distribuição de probabilidade é a mais indicada para determinar a probabilidade de ocorrência de temperaturas máxima na cidade de Iguatu. Por este motivo, o teste de Lilliefors foi utilizado para determinar qual distribuição de probabilidade obteve o melhor desempenho no ajuste aos dados em estudo.

Pode ser observado na Tabela 2, que segundo o teste de Lilliefors, a distribuição de probabilidade que mais se aderiu aos dados diários de temperatura do ar máxima em Iguatu foi a distribuição Normal e que a que menos se aderiu foi a Gumbel I.

A distribuição Gumbel I bastante utilizada na estimação de eventos climatológicos extremos, servindo de subsídio para projetos de engenharia agrícola (ASSIS et al., 1996), obteve em vários estudos (BEIJO et al., 2005; LYRA et al., 2006) bons ajustes na estimativa da probabilidade tanto aos dados de temperatura quanto de precipitação e até mesmo para dados de velocidade do vento. Diferente de diversos trabalhos que podem ser mencionados, neste estudo, a distribuição Gumbel I não obteve a aderência exigida pelo teste de Lilliefors, obtendo valores de Dobs bem maiores do que os de Dtab. Seu ajuste foi observado apenas para o meses de janeiro

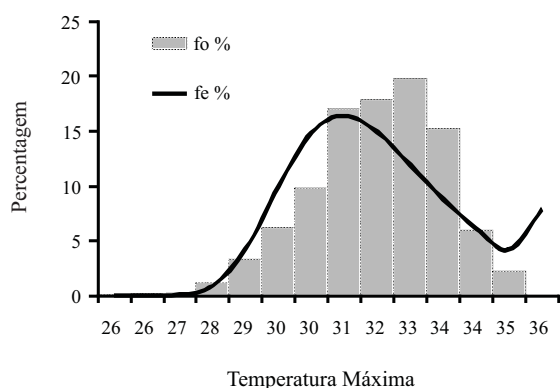


Figura 4 - Distribuição de frequência observada (fo) e estimada (fe) pela distribuição Gumbel I para a temperatura máxima diária em julho para Iguatu, Ceará

e fevereiro (Tabela 2), no estudo de Sansigolo (2008) a distribuição Gumbel I se sobressaiu no ajustamento de dados de temperatura máxima absoluta e precipitação diária, já no ajustamento dessa distribuição aos dados de velocidade do vento vale-se citar os trabalhos de Bautista et al. (2004). A pouca aderência da distribuição Gumbel I pode ser observada na Figura 4:

A distribuição Normal, em diversos trabalhos (ASTOLPHO, et al., 2005; SILVA et al., 2005) foi considerada a distribuição de melhor ajustamento, indiferente da homogeneidade dos dados em estudo. Neste estudo, assim como no trabalho de Assis et al. (2004), a distribuição Normal obteve um bom ajustamento para dados diários de temperatura do ar máxima segundo o teste de Lilliefors. Foi observado que essa distribuição obteve ajuste para todos os meses do ano (Tabela 2), mostrando assim que essa distribuição pode ser utilizada em diferentes épocas sem que haja uma restrição para específicas épocas do ano. A boa aderência da distribuição Normal pode ser observada na Figura 5:

Outra distribuição que obteve destaque neste e em diversos estudos, é a Gama. Nos resultados observados por Catalunha et al. (2002), para precipitação pluvial, a distribuição Gama obteve grande aderência pelo teste K-S e reprovação pelo χ^2 . Um ponto que pode ser observado segundo a Tabela 2, é que a distribuição Gama, obteve quase a mesma quantidade de ajustes que a Normal. Isso pode ser explicado pelo fato de que essa distribuição possui uma grande diversidade de forma, todas influenciadas pelo valor de seu parâmetro η , podendo assim, algumas vezes, se aproximar da forma da distribuição Normal.

A aproximação da forma da distribuição Gama a distribuição Normal pode ser observada na Figura 6:

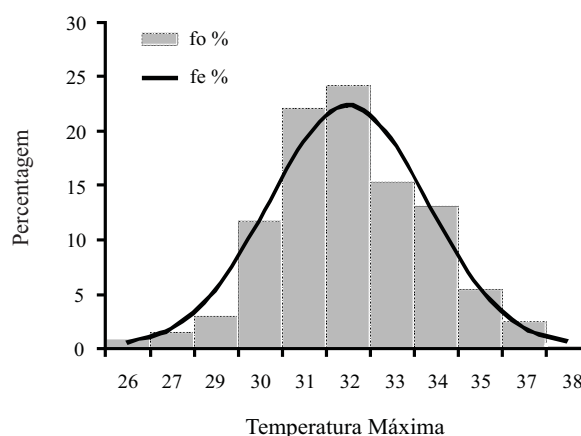


Figura 5 - Distribuição de frequência observada (fo) e estimada (fe) pela distribuição Normal para a temperatura máxima diária em fevereiro para Iguatu, Ceará

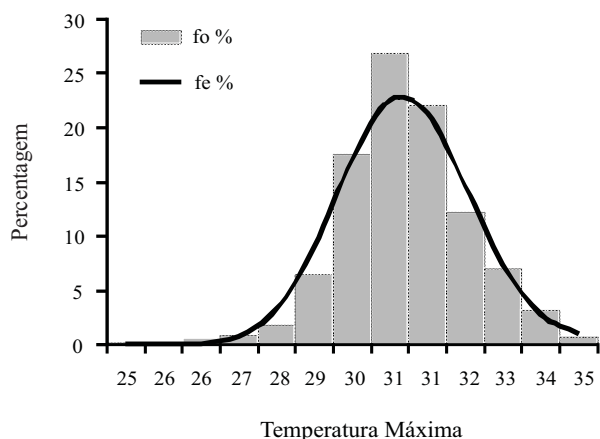


Figura 6 - Distribuição de frequência observada (fo) e estimada (fe) pela distribuição Gama para a temperatura máxima diária em abril para Iguatu, Ceará

Conclusões

Conforme os resultados obtidos, concluiu-se, para as condições estudadas, que:

1. Para as análises diárias de probabilidade de temperatura máxima para a cidade de Iguatu, pelo teste de Lilliefors, a função que melhor se ajustou aos dados de temperatura do ar diária máxima foi a distribuição Normal;
2. A função densidade de probabilidade que apresentou menor ajuste foi a função Gumbel I;
3. A distribuição de probabilidade Gama teve sua forma aproximada a da distribuição Normal, obtendo ambas as distribuições bons resultados, segundo o teste de Lilliefors.

Referências

ASSIS, F. N.; ARRUDA, H. V.; PEREIRA, A. R. **Aplicações de estatística à climatologia**. 1. ed. Pelotas: Universidade Federal de Pelotas, 1996. 161 p.

ASSIS, J. P. *et al.* Ajuste de séries históricas de temperatura e radiação solar global diária às funções densidade de probabilidade normal e log-normal, em Piracicaba, SP. **Revista Brasileira de Agrometeorologia**, v. 12, n. 01, p. 113-121, 2004.

ASTOLPHO, F. *et al.* Regionalização de riscos de ocorrência de temperaturas mínimas absolutas anuais para o estado de São Paulo com base em modelos probabilísticos e digitais de elevação. **Bragantia**, v. 64, n. 01, p. 139-148, 2005.

BAUTISTA, E. A. L.; ZOCCHI, S. S.; ANGELOCCI, L. R. A distribuição generalizada de valores extremos aplicada ao ajuste dos dados de velocidade do vento em Piracicaba São Paulo, Brasil. **Revista Matemática Estatística**, v. 22, n. 01, p. 95-111, 2004.

BEIJO, L. B.; MUNIZ, J. A.; CASTRONETO, P. Tempo de retorno das precipitações máximas em Lavras (MG) pela distribuição de valores extremos do tipo I. **Ciência e Agrotecnologia**, v. 29, n. 03, p. 657-667, 2005.

BOTELHO, V. A. V. A.; MORAIS, A. R. Estimativas dos parâmetros da distribuição Gama de dados pluviométricos do município de Lavras, estado de Minas Gerais. **Ciência e Agrotecnologia**, v. 23, n. 03, p. 697-706, 1999.

CAMPOS, H. **Estatística experimental não-paramétrica**. 4. ed. Piracicaba: ESALQ, 1983. 349 p.

CATALUNHA, M. J. *et al.* Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, v. 10, n. 01, p. 153-162, 2002.

COSTA, L. C.; SEDIYAMA, G. C. Elementos climáticos e produtividade agrícola. **Ação Ambiental**, v. 2, n. 07, p. 24-28, 1999.

DOANE, D. P. Aesthetic frequency classifications. **The American Statistician**, v. 30, n. 04, p. 181-183, 1976.

ESTEFANEL, V.; SCHNEIDER, F. M.; BURIOL, G. A. Probabilidade de ocorrência de temperaturas máximas do ar prejudiciais aos cultivos agrícolas em Santa Maria, RS. **Revista Brasileira de Agrometeorologia**, v. 02, n. 01, p. 57-63, 1994.

FERREIRA, F. F. **Estatística Básica**. 1. ed. Lavras: Editora UFLA, 2005. 664 p.

JUNQUEIRA JÚNIOR, J. A. *et al.* Precipitação provável para a região de Madre de Deus, Alto Rio Grande: modelos de probabilidades e valores característicos. **Ciência e Agrotecnologia**, v. 31, n. 03, p. 842-850, 2007.

LYRA, G. B. *et al.* Regiões homogêneas e funções de distribuição de probabilidade da precipitação pluvial no Estado de Táchira, Venezuela. **Pesquisa Agropecuária Brasileira**, v. 41, n. 02, p. 205-215, 2006.

MARTIN, T. N.; STORCK, L.; DOURADO NETO, D. Simulação estocástica da radiação fotossinteticamente ativa e da temperatura do ar por diferentes métodos. **Pesquisa Agropecuária Brasileira**, v. 42, n. 09, p. 1211-1219, 2007.

NAGHETTINI, M.; PINTO, E. J. A. **Hidrologia estatística**. Belo Horizonte: CPRM, 2007. 552 p.

RIBEIRO, B. T. *et al.* Comparação de distribuições de probabilidade e estimativa da precipitação provável para região de Barbacena, MG. **Ciência e Agrotecnologia**, v. 31, n. 05, p. 1297-1302, 2007.

SANSIGOLO, C. A. Distribuições de extremos de precipitação diária, temperatura máxima e mínima e velocidade do vento em

Piracicaba, SP. **Revista Brasileira de Meteorologia**, v. 23, n. 03, p.341-346, 2008.

SILVA, Z. M. Climas do Estado do Ceará. **Ciência Agrônômica**, v. 18, n. 02, p. 89-95, 1987.

SILVA, J. B. da; LLOPART, M. P.; BOIASKI, N. Temperatura máxima do ar em Pelotas, RS - Tabelas de probabilidades em escalas de tempo pentadal. **Revista Brasileira de Meteorologia**, v.20, n.02, 267-276, 2005.

SUPERINTENDÊNCIA DO DESENVOLVIMENTO DO NORDESTE (SUDENE). **Levantamento Exploratório – Reconhecimento de Solos do Estado do Ceará – Volume I**. Recife: SUDENE, 1973. 301 p.

THOM, H. C. S. **Some methods of climatological analysis**. Geneva, 1966. 53 p. (Technical note n 81).

WILKS, D. S. **Statistical methods in the atmospheric sciences**. 2 ed. San Diego: Elsevier, 2006. 630 p.