

# FUNÇÃO HYPER: PROPOSTA DE UM NOVO MODELO DE DISTRIBUIÇÃO PROBABILÍSTICA

Daniel Pereira Guimarães<sup>1</sup>

## 1. Introdução

O emprego de funções densidade de probabilidade tem sido cada vez mais frequente. Dentre os fatores que contribuíram para a expansão do uso dessas funções, tem-se as facilidades de estimativa dos parâmetros dos modelos não lineares, decorrentes dos avanços da informática, os conceitos de Modelos Lineares Generalizados definidos por NELDER & WEDDERBURN (1972), onde modelos de regressão se baseiam em distribuições de erros não normais e as técnicas da modelagem de processos. Descrições detalhadas e os principais métodos de solução para a maioria das funções probabilísticas contínuas são encontrados em JOHNSON et al. (1994) e EVANS et al. (2000). Embora exista na literatura um grande número de funções probabilísticas contínuas e univariadas, em muitos casos a mesma função recebe diferentes denominações.

As funções com capacidade de descrever distribuições com diferentes curvaturas tem sido as mais empregadas na área de climatologia.

A função Gama tem sido usada com frequência em trabalhos relativos a estudos meteorológicos, os quais tiveram impulso a partir dos resultados de VIVALDI (1973). RIBEIRO & LUNARDI (1996) usaram-na para determinar as probabilidades mensais de chuva em Londrina, PR. MACENA et al. (1998) aplicaram essa distribuição para estimar e espacializar a disponibilidade hídrica no Estado de Goiás. SAAD & FRIZZONE (1998) propuseram o dimensionamento de equipamentos de irrigação, através do ajuste dessa função às frequências das intensidades de precipitação na região de Piracicaba, SP.

QUINE (2000), empregou a função Weibull para estimar a probabilidade de riscos de vendavais na Inglaterra, SANTOS & GUERRINI (1998), para avaliar o fluxo horizontal da água no solo e TATE & FREEMAN (2000), para caracterizar os períodos de estiagem no sudeste da África.

Outras funções comumente usadas são: Beta, logística, Richards, Log-normal e Gompertz e a curva normal.

## 2. Material e métodos

De maneira análoga à função tangente, a qual é resultante da razão entre as funções seno e cosseno, a função tangente hiperbólica, denominada de tanh, é definida pela razão entre as funções seno hiperbólico (senh) e cosseno hiperbólico (cosh).

A função tangente hiperbólica descreve uma distribuição probabilística de curvatura rígida passando pela origem. Substituindo-se a variável aleatória (X) por uma função de potência e acrescentando-se um parâmetro assintótico ( $\omega$ ) e

outro de escala ( $\alpha$ ), tem-se o modelo denominado de função Hyper, cuja forma integral é dada por:

$$Y = \omega \cdot \tanh(\beta \cdot (X - \alpha)^\gamma),$$

sua forma derivativa é:

$$f(x) = \omega \cdot \beta \cdot \gamma \cdot X^{(\gamma-1)} \cdot (1 - \tanh(\beta \cdot X^\gamma))^2,$$

com o seguinte ponto de inflexão:

$$I = \exp(-\ln(\beta)/\gamma) + \alpha$$

Para a validação do modelo, testou-se sua aplicabilidade em descrever fenômenos com diferentes tendências de variação tendo como base comparativa os resultados obtidos pelo ajuste da função Weibull.

## 3. Resultados e discussão

O modelo proposto gera distribuições com pontos de inflexão variando desde zero ( $\gamma = 1$ ) até o limite superior definido por  $I = \tanh(1) = 0.76$ , o que confere a esta função uma flexibilidade superior à obtida pela função Weibull, cujos pontos de inflexão variam entre zero e o limite de  $I = (1-1/e) = 0.63$  e também superior à função Gama onde esses limites variam entre 0 e 0,5. Curvaturas descritas pela função Hyper em sua forma integral são mostradas na Figura 1 e na forma derivativa na Figura 2.

**Teste de validação 1.** Ajuste de distribuições com assimetria positiva. Utilizou-se do exemplo empregado pelo software "TableCurve" para determinar a distribuição de gotas de óleo em suspensão ([http://www.spssscience.com/White\\_Papers/oil\\_droplet\\_si ze.cfm](http://www.spssscience.com/White_Papers/oil_droplet_si ze.cfm)), onde a função Weibull forneceu o melhor ajuste dentre as 3665 funções disponíveis no programa.

O ajuste das funções Hyper e Weibull forneceu os seguintes resultados:

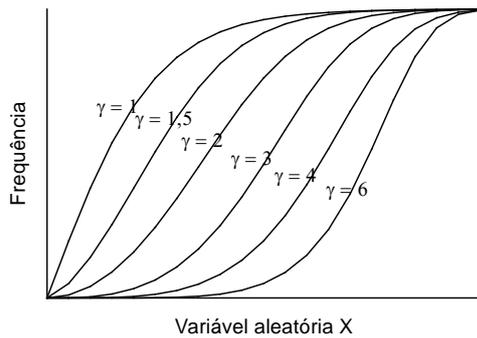
Função	$\omega$	$\beta$	$\gamma$	SQR	R2
<b>Hyper</b>	167,2	0.2989	1.0971	35.61	0.9874
<b>Weibull</b>	160,8	0.3082	1.3801	37.29	0.9868

Observam-se resultados similares entre as funções para o ajuste de distribuições com assimetria positiva.

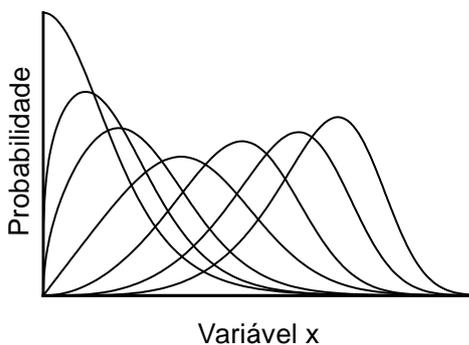
**Teste de validação 2.** Ajuste de distribuições com assimetria negativa. Os dados empregados referem-se aos valores disponibilizados pelo CPTEC/INPE no site [www.cptec.inpe.br/products/queimadas/queimabrasilacu.htm](http://www.cptec.inpe.br/products/queimadas/queimabrasilacu.htm) e representam o número de focos de calor mensalmente captados através de imagens de satélite. Os resultados obtidos são apresentados na Tabela 1 e ilustrados na Figura 3. Os resultados dos ajustamentos demonstram a superioridade da função Hyper quando os dados tendem a apresentar assimetria negativa, sempre

<sup>1</sup> Dr. Pesquisador da Embrapa Milho e Sorgo, Cx. P. 151, 35701-970 Sete Lagoas, MG. e-mail: daniel@cnpms.embrapa.br

que o ponto de inflexão for maior que o limite de inflexão imposto para a função Weibull ( $I = 0,63$ ).



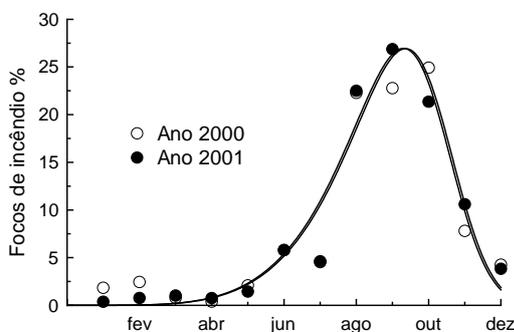
**Figura 1.** Distribuições probabilísticas determinadas pela função Hyper em sua forma derivativa.



**Figura 2.** Distribuições probabilísticas determinadas pela função Hyper em sua forma derivativa.

**Tabela 1.** Estimativas dos parâmetros e precisão das estimativas das funções Hyper e Weibull para explicar a frequência de focos de incêndio no Brasil nos anos de 2000 e 2001. ([www.inpe.gov.br](http://www.inpe.gov.br)).

	ano de 2000		ano de 2001	
	Hyper	Weibull	Hyper	Weibull
$\omega$	100 Fixo	100 Fixo	100 Fixo	100 Fixo
$\beta$	2.11E-06	1.45E-06	3.29E-06	1.07E-06
$\gamma$	5.6757	5.9515	5.49212	6.09376
SQR	85.96	110.33	64.79	77.51
$R^2$	0.9097	0.8842	0.9376	0.9254



**Figura 2.** Frequência de ocorrência de focos de incêndio no Brasil e tendências determinadas pela função Hyper.

#### 4. Conclusão

A função de distribuição Hyper descreve curvas com diferentes assimetrias, incluindo desde a tradicional distribuição do J invertido (sem ponto de inflexão), curvas de assimetria positiva ( $I < 0,5$ ), distribuição simétrica ( $I = 0,5$ ) e distribuições com assimetria negativa ( $I > 0,5$  até  $I = 0,76$ ). Esses atributos colocam essa função em posição vantajosa em relação aos modelos tradicionalmente usados, uma vez que a função Gama descreve apenas curvas com pontos de inflexão  $I < 0,5$  e a função Weibull descreve curvas com pontos de inflexão variando até o limite de  $I = 0,63$ .

#### 5. Referências bibliográficas

- EVANS, M.; HASTINGS, N.; PEACOCK, B. **Statistical Distributions**, 3rd. Ed., John Wiley and Sons. 2000, 221 p.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions**, 2nd. Ed., John Wiley and Sons. 1994, 784 p.
- MACENA, F. A.; ASSAD, E. D.; MATTOS, H.; LUIZ, A. J. B. Variação espaço-temporal da disponibilidade hídrica climática no Estado de Goiás, **Revista Agropecuária Brasileira**, v. 33 no. 5, p. 1998.
- NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society**, v. 135 Série A, p. 370-384. 1972.
- QUINE, C.P. Estimation of mean wind climate and probability of strong winds for wind risk assessment. **Forestry**, Vol. 73 No. 3, p. 247-258. 2000.
- RIBEIRO, A.M.; LUNARDI, D. M. C. A precipitação mensal provável para Londrina - PR, através da função Gama. **Energia na Agricultura**, v. 12 no. 4, p. 37-44. 1996.
- SAAD, J. C. C.; FRIZZONE, J, A. Estudo da distribuição de frequência da precipitação pluvial, visando o dimensionamento de sistemas de irrigação. **Revista Irriga**, v. 3 no. 1, p. 117-126. 1998.
- SANTOS, E; GUERRINI, I. A; Determinação experimental da dimensão fractal do movimento da água no solo em fluxos horizontais. **Irriga**, 3: 2, p. 96-107. 1998.
- TATE, E.L; FREEMAN S.N. Three modelling approaches for seasonal streamflow droughts in southern Africa: the use of censored data. **Hydrological Sciences Journal**, Vol. 45 No. 1, p. 27-42. 2000.
- VIVALDI, L.J. **Utilização da distribuição gama em dados pluviométricos**. Piracicaba, 1973. 73p. Dissertação de Mestrado em Agrometeorologia - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.