

MODELANDO COM ALGUMAS DISTRIBUIÇÕES DISCRETAS AS SEQUÊNCIAS DE DIAS CHUVOSOS

Marta Eliane Echeverria Borges¹, Josmar Mazucheli², Diogo Francisco Rossoni³

1. Introdução

O conhecimento dos fatores climáticos é primordial para direcionar decisões socioeconômicas de um local ou região, especialmente, a informação da precipitação pluviométrica, um fator determinante para estimar o abastecimento de água de uso doméstico, industrial, agropecuário e na geração de energia.

Os estudos relacionados às secas e aos impactos das chuvas têm recebido atenção dos pesquisadores, sendo analisados em diversas áreas do conhecimento, como a geografia, a meteorologia, a sociologia, a engenharia sanitária, estatística dentre outras áreas (MONTEIRO et al., 2012).

Um modelo é a representação simplificada de algum problema ou situação da vida real destinado a ilustrar certos aspectos do problema. E mais de um modelo pode descrever um mesmo fenômeno, haja vista que cada pesquisador tem a liberdade de modelar o fenômeno seguindo a metodologia que julgar mais adequada (EMILIANO, 2013).

Segundo WMO (2009), os modelos probabilísticos são usados em uma ampla variedade de estudos hidrológicos, incluindo estudos de fluxos, secas, volumes de reservatórios e quantidades de chuva. A escolha e o ajuste da distribuição de probabilidade que melhor represente o comportamento da variável em estudo é a parte primordial da análise. Sendo assim, inúmeras distribuições de probabilidade têm sido consideradas na literatura. Entre elas, as mais utilizadas são a distribuição gama, Weibull, log-normal, log-logística e Gumbel.

As variáveis originais podem ser contínuas na natureza mas discretas pela observação e portanto é razoável e conveniente modelar a situação por uma distribuição discreta apropriada gerada dos modelos contínuos subjacentes que preservam uma ou mais características importantes da distribuição contínua (CHAKRABORTY, 2015).

Para ZEKAI (2015), informações completas sobre o número de períodos de chuva e seca têm se tornado cada vez mais relevantes em muitos campos, como em projetos de gestão de água, planejamento agrícola, estudos de enchentes e em várias atividades industriais.

Segundo TOLIKA e MAHERAS (2005) e MANDAPAKA (2016), o período chuvoso é definido como um período consecutivo de dias com chuva está acima de um limite específico. Vários estudos sobre períodos de chuva e seca mostraram que a seleção do limiar a fim de definir um dia chuvoso ou seco não é fixo, mas depende em determinadas características, o clima e as necessidades de cada área de estudo em particular.

Deste modo, este trabalho tem como objetivo empregar as distribuições gama discreta, Weibull discreta, binomial negativa, binomial negativa truncada em zero, geométrica e a geométrica truncada em zero para modelar a sequência de dias chuvosos da estação meteorológica de Calçoene - AP, visto ser a cidade que mais chove no Brasil.

¹ Mestranda em Pós graduação em Bioestatística - UEM. email: marta_est@yahoo.com.br

² Departamento de Estatística - UEM. email: jmazucheli@gmail.com

³ Departamento de Estatística - UEM. email: dfrossoni@uem.br

2. Metodologia

O banco de dados utilizado para estudo foi obtido a partir dos registros disponíveis no portal de Rede Hidrometeorológica Nacional (RHN) coordenado pela Agência Nacional de Águas (ANA). Para o estudo foi utilizado a série histórica do período de 1 de janeiro de 1978 a 30 de novembro de 2018 da estação 08250002 em Calçoene - AP.

Para encontrar modelos probabilísticos que caracterizam adequadamente as sequências de dias chuvosos, algumas distribuições como gama discreta, Weibull discreta, binomial negativa, geométrica, binomial negativa truncada em zero e a geométrica truncada em zero foram avaliadas. Neste estudo o limiar para considerar um período chuvoso utilizou-se 1 mm de precipitação por dia.

Para verificar se a distribuição de probabilidade utilizada ajusta-se a série de dados estudada a qualidade do ajuste foi verificada por meio do teste qui-quadrado, considerando nível de significância de 5%. Os parâmetros das distribuições estudadas foram estimados por meio do método da máxima verossimilhança.

A seleção do modelo mais adequado foi realizada por meio dos critérios de informação de Akaike (AIC) e Bayesiano (BIC) para avaliar a qualidade de um modelo ajustado. Todas as análises foram conduzidas no software R (IHAKA e GENTLEMAN 1996).

3. Resultados e Discussão

O período da série histórica estudado mostrou que a precipitação pluviométrica média anual em Calçoene foi de 4350.2 mm ($s = 799.7$, $cv = 18.4\%$). Calçoene é caracterizada por um clima tropical (Am) de acordo com a classificação climática de Köppen-Geiger.

As observações diárias de precipitação em relação ao meses indicaram uma grande variabilidade dos dados, logo os quartis são as melhores medidas para representar os dados, deste modo pode-se notar que maio e abril foram os meses que mais choveram durante os últimos 40 anos, pois 50% dos dias de maio choveram no mínimo 12 mm e em abril 50% dos dias de choveram mais de 11.2 mm. Enquanto que outubro caracterizou como o mês mais seco, uma vez que 75% dos dias de outubro não houve chuva, o mesmo ocorreu em setembro, porém teve o maior índice de precipitação diária entre todos os meses (Tab. 1).

Tabela 1: Análise descritiva da precipitação (mm) de 1978 a 2018 da estação de Calçoene - AP.

Medidas	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Q ₁	2.2	3.0	1.2	2.6	3.5	0.3	0.0	0.0	0.0	0.0	0.0	0.0
Md	9.2	10.4	8.5	11.2	12.0	6.3	1.3	0.0	0.0	0.0	0.0	3.4
\bar{x}	18.9	21.4	19.7	21.8	20.9	12.4	6.6	2.7	1.1	1.3	3.2	10.5
Q ₃	25.5	28.2	26.5	29.2	28.4	17.2	8.5	1.8	0.0	0.0	1.5	12.3
Máximo	338.9	220.4	261.3	175.5	207.9	158.9	124.2	74.3	340.6	112.2	121.6	177.1
s	27.1	28.2	28.1	27.7	25.9	17.0	12.2	6.8	10.4	6.4	9.5	18.9
cv (%)	142.9	132.0	142.8	127.2	124.0	137.5	184.6	249.2	923.7	498.2	295.5	179.9

As distribuições observadas e ajustadas das sequências de chuvas da estação de Calçoene em relação aos meses estão apresentadas na Figura 1 e pode-se observar que nos meses de janeiro a junho houve pelo menos um ano que choveu durante o mês todo, enquanto que de agosto a outubro as sequências de dias chuvosos não ultrapassaram 9 dias.

Os períodos chuvosos observados na estação de Calçoene são estatisticamente mais adequados pela distribuição binomial negativa truncada em zero (BNTZ), com um nível de confiança de 95% usando a estatística de teste do Qui-quadrado. Verificou-se também que no mês de outubro apenas a distribuição de geométrica truncada em zero (GTZ) se ajustou adequadamente para descrever o comportamento das sequências de dias chuvosos (Quadro 1).

Deve-se notar que as distribuições com um valor χ^2 mais baixo e maior p valor produzem um melhor ajuste. Por outro lado, se nenhuma das distribuições se encaixa na distribuição dos períodos chuvosos e secos em uma determinada estação pluviométrica, a estatística mínima de χ^2 será considerada para a seleção da distribuição mais apropriada (DENI et al. 2008).

Os critérios de informações AIC e BIC também sinalizam que o modelo binomial negativa truncada em zero é mais indicado dentre as distribuições de probabilidade avaliadas exceto nos meses de outubro e setembro que a distribuição de geométrica truncada em zero apresentou os menores valores (Quadro 2).

4. Considerações Finais

Com auxílio do teste qui-quadrado de aderência constatou-se que a distribuição binomial negativa truncada em zero foi mais adequada para os meses mais chuvosos, enquanto que a distribuição geométrica truncada em zero foi mais indicada para meses mais secos como ocorreu no mês de outubro.

Além disso, os resultados indicaram que os meses de maio e abril foram os mais chuvosos da série histórica estudada. Aliás, os meses de janeiro a junho apontaram as maiores sequências de dias chuvosos da estação meteorológica de Calçoene.

5. Referencias Bibliográficas

- CHAKRABORTY, S. Generating discrete analogues of continuous probability distributions-A survey of methods and constructions. *J. Stat. Distrib. App.*, n.2, v.6, 30 p., 2015.
- DENI S.M.; JEMAIN A.A.; IBRAHIM K. The spatial distribution of wet and dry spells over peninsular Malaysia. *Theor. Appl. Climatol.*, v.94, p. 163-173, 2008.
- EMILIANO, P. C. Critérios de Informação: Como Eles Se Comportam Em Diferentes Modelos? Tese - Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, Universidade Federal de Lavras, Lavras - MG, 2013.
- IHAKA, R. e GENTLEMAN, R. R: A language for lata analysis and graphics. *J. Comput. Graph. Statist.*, v.5, n.3, p. 299-314, 1996.
- MANDAPAKA, P.V.; QIN, X.; LO E.Y.M. Seasonal and interannual variability of wet and dry spells over two urban regions in the western maritime continent. *J. Hydrometeorol*, v.17, p. 1579-1600, 2016.
- MONTEIRO, J.B.; ROCHA, A. B.; ZANELLA, M.E. Técnica dos quantis para caracterização de anos secos e chuvosos (1980-2009): baixo curso do Apodi-Mossoró-RN. *R. Depto. Geogr.*, v. 23, p. 232-249, 2012.

TOLIKA K, MAHERAS P. Spatial and temporal characteristics of wet spells in Greece. *Theor. App. Climatol.*, v.81, p. 71-85, 2005.

WMO. World Meteorological Organization. Guide to Hydrological Practices, Volume II. [S.l.: s.n.]. Management of Water Resources and Application of Hidrological Practices. 2009.

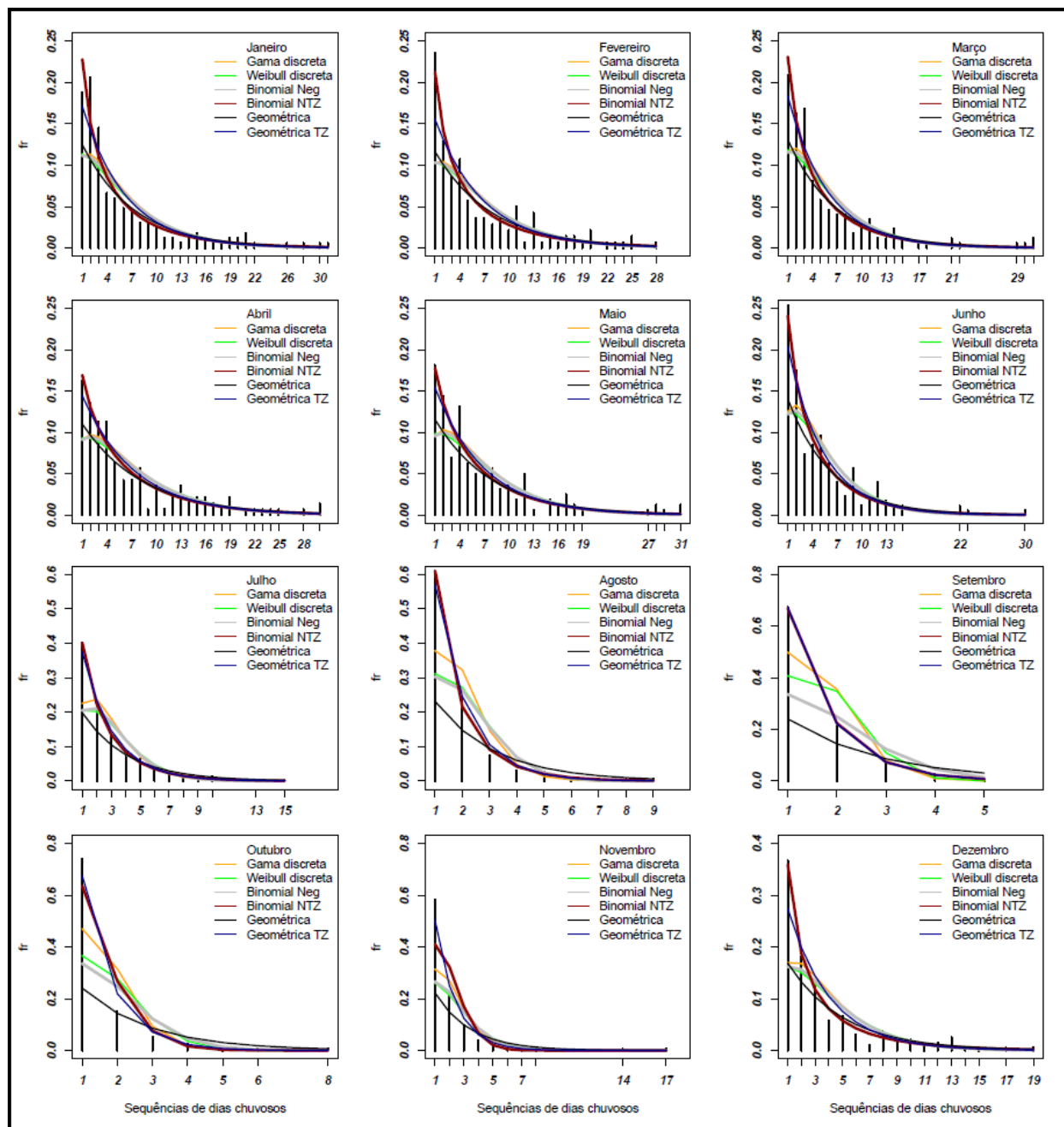


Figura1: Ajuste das sequências de dias chuvosos da série histórica de Calçoene (AP) de 1978 a 2018.

Quadro 1: Teste qui-quadrado de aderência para os modelos ajustados da série histórica de Calçoene (AP) de 1978 a 2018.

Mês	Estatística do teste	Distribuição					
		GD	WD	BN	BNTZ	G	GTZ
Jan	χ^2	29.09	29.93	29.13	9.12	28.84	16.89
	p valor	< 0.0001	< 0.0001	< 0.0001	0.1044	< 0.0001	0.0097
Fev	χ^2	12.60	10.27	10.39	4.40	5.26	13.91
	p valor	0.0498	0.1136	0.1092	0.6231	0.6281	0.0528
Mar	χ^2	13.33	15.28	13.95	5.56	19.38	7.29
	p valor	0.0204	0.0092	0.0159	0.3510	0.0036	0.2952
Abr	χ^2	11.01	11.42	10.87	4.39	12.18	6.07
	p valor	0.0880	0.0761	0.0925	0.6237	0.0947	0.5322
Mai	χ^2	7.63	7.04	7.03	1.04	8.13	2.93
	p valor	0.1779	0.2179	0.2184	0.9593	0.2285	0.8181
Jun	χ^2	11.53	9.29	9.81	6.48	12.04	8.37
	p valor	0.0733	0.1578	0.1327	0.3711	0.0992	0.3012
Jul	χ^2	22.69	16.30	16.27	1.19	18.86	2.83
	p valor	0.0001	0.0026	0.0027	0.8805	0.0020	0.7269
Ago	χ^2	18.60	15.87	14.56	1.34	20.02	2.11
	p valor	< 0.0001	< 0.0001	0.0001	0.2465	< 0.0001	0.3487
Set	χ^2	9.44	9.57	5.77	0.13	8.22	0.09
	p valor	0.0021	0.0020	0.0163	0.7156	0.0041	0.7647
Out	χ^2	15.38	15.77	13.97	7.71	7.59	3.07
	p valor	<0.0001	<0.0001	0.0002	0.0055	0.0059	0.0797
Nov	χ^2	14.86	12.28	12.24	21.79	11.39	4.13
	p valor	0.0001	0.0005	0.0005	< 0.0001	0.0034	0.1267
Dez	χ^2	30.29	22.77	23.20	2.78	8.45	20.32
	p valor	< 0.0001	0.0001	0.0001	0.5957	0.1331	0.0011

Legenda: gama discreta (GD), Weibull discreta (WD), binomial negativa (BN), binomial negativa truncada em zero (BNTZ), geométrica (G) e a geométrica truncada em zero (GTZ).

Quadro 2: Critérios de informação para os modelos ajustados da série histórica de Calçoene (AP) de 1978 a 2018.

Mês	Critérios de informação	Distribuição					
		GD	WD	BN	BNTZ	G	GTZ
Jan	AIC	930.4	937.5	935.2	879.2	941.3	884.5
	BIC	936.6	943.7	941.4	885.4	944.4	887.6
Fev	AIC	816.7	820.2	819.8	775.8	823.7	780.0
	BIC	822.6	826.1	825.7	781.7	826.7	782.9
Mar	AIC	942.7	952.1	948.7	892.4	958.9	895.4
	BIC	949.0	958.4	954.9	898.7	962.1	898.6
Abr	AIC	838.78	843.32	842.18	809.33	849.58	808.78
	BIC	844.68	849.22	848.08	815.23	852.53	811.73
Mai	AIC	929.28	935.61	933.55	895.28	944.13	894.66
	BIC	935.43	941.76	939.70	901.43	947.21	897.73
Jun	AIC	928.93	938.05	935.83	880.09	953.48	880.95
	BIC	935.28	944.41	942.18	886.45	956.66	884.13
Jul	AIC	891.33	920.39	916.06	813.52	988.13	813.11
	BIC	898.20	927.25	922.93	820.38	991.56	816.55
Ago	AIC	537.20	582.28	584.12	454.98	684.50	455.86
	BIC	543.68	588.77	590.60	461.46	687.74	459.10
Set	AIC	229.71	253.32	278.29	200.37	350.94	198.38
	BIC	235.00	258.61	283.58	205.66	353.59	201.03
Out	AIC	272.80	307.71	313.64	230.83	376.83	211.43
	BIC	278.23	313.14	319.07	236.27	379.55	214.14
Nov	AIC	512.86	549.35	542.28	507.69	588.51	426.19
	BIC	518.93	555.43	548.35	513.77	591.55	429.23
Dez	AIC	903.74	917.51	915.18	818.06	936.16	828.98
	BIC	910.26	924.03	921.71	824.58	939.43	832.24

Legenda: gama discreta (GD), Weibull discreta (WD), binomial negativa (BN), binomial negativa truncada em zero (BNTZ), geométrica (G) e a geométrica truncada em zero (GTZ).