Modelagem do tempo de vida dos papas via análise de sobrevivência

Daniel Valentins de Lima[1], Fabio Douglas Soares Bezerra[2], Evandro Mariano Barros da Silva[3], Marcelino Alves Rosa de Pascoa [4], José Nilton da Cruz [5], Graziela Dutra Rocha Gouvêa [6]

Introduçao

Este presente trabalho apresenta uma análise via análise de sobrevivência em relação ao tempo de vida até um papa morrer ou sair do cargo. É de conhecimento geral que o tempo de pontificado de um papa é por período vitalício, ou seja, em casos que não há a existência de renúncia ou deposição, o papa exerce o cargo até o fim de sua vida.

Não é sempre que ocorre o cenário do exercício de cargo por período vitalício, no decorrer da história houveram papas que renunciaram, e também papas que foram depostos de seus cargos. Visto isso, há o cenário do tempo de ocorrência de um evento específico, e também a ausência desse evento por parte de algumas observações, o que nos permite uma implementação das técnicas de análise de sobrevivência.

Até o presente momento, houveram 266 papas, incluindo o atual. No que se refere a renúncias e deposições, há o total de 13 ocorrências. Ou seja, há uma censura em 4,89% dos dados.

Distribuição Weibull

A distribuição de Weibull tem uma gama de aplicabilidades. Criada em 1939, pelo engenheiro e matemático sueco Ernst Hjalmar Waloddi Weibull, a distribuição de Weibull é uma das mais conhecidas e utilizadas distribuições na modelagem de tempo de falhas de equipamentos e outros fenômenos.

Esta ampla utilização se deve basicamente a facilidade de interpretá-la e a sua flexibilidade, sendo passível de utilização na modelagem de dados simétricos, com assimetria à direita ou à esquerda e dados com censura. A expressão para as funções densidade de probabilidade e sobrevivência são dadas, respectivamente, por:

$$f(t) = \frac{\gamma}{\lambda^\gamma} t^{\gamma-1} exp\left\{-\left(\frac{t}{\lambda}\right)^\gamma\right\}, t \geq 0,$$

$$S(t) = exp\left\{-\left(\frac{t}{\lambda}\right)^\gamma\right\},$$

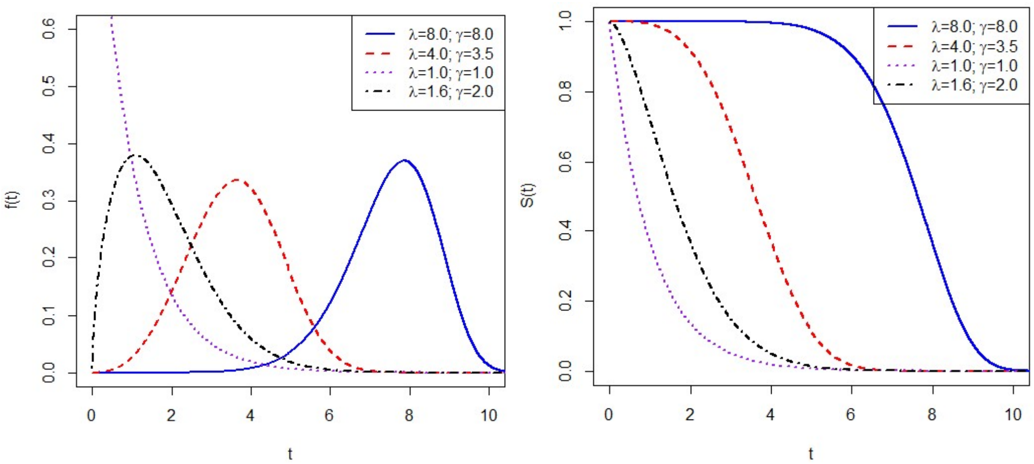
Sendo $\gamma > 0$ e $\lambda > 0$, respectivamente, o parâmetro de forma e de escala.

[1] Departamento de Estatística - UFMT. email: dvalentins@outlook.com
[2] Departamento de Estatística - UFMT. email: fabiodouglas.soares@gmail.com
[3] Departamento de Estatística - UFMT. email: evandromariano_barros@live.com
[4] Departamento de Estatística - UFMT. email: marcelino.pascoa@gmail.com
[5] Departamento de Estatística - UFMT. email: niltonn.cruz@gmail.com
[6] Universidade Federal de Ouro Preto. email: gragouvea@gmail.com

Algumas formas das funções densidade de probabillidade e Sobrevivência para a distribuição Weibull são apresentadas na Figura 1, no qual podemos verificar a versatilidade da distribuição.

Figura 1: Formas das funções de densidade probabilidade e de sobrevivência da distribuição

Weibull para alguns valores dos parâmetros.



Distribuição Odd Log-logística Weibull (OLLW)

Com o objetivo de buscar distribuições mais flexíveis que melhor se ajustem a dados em que as distribuições mais usuais como Exponencial, Weibull, Gama, não conseguem ou oferecem um mal ajuste, ? propuseram uma extensão da distribuição Weibull. Com base na família "Odd Log-logística", propuseram a distribuição Odd Log-logística Weibull (OLLW), produzindo maior flexibilidade a distribuição Weibull por meio do acréscimo de um parâmetro de forma $\alpha > 0$.

Seja uma variável aleatória T com distribuição OLLW, com parâmetros de escala $\lambda > 0$ e de forma $\gamma > 0$ e $\alpha > 0$. Assim sendo, a função densidade de probabilidade de T é dada pela seguinte expressão

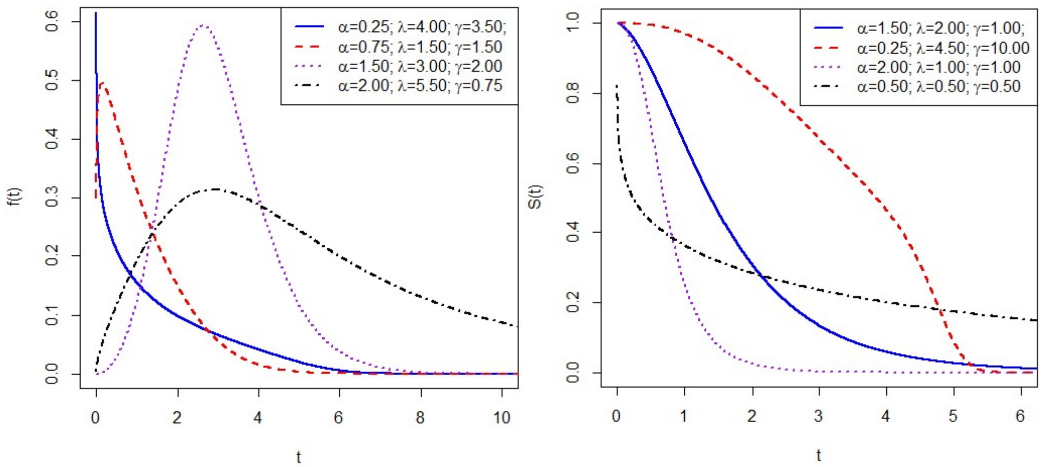
$$f(t) = \frac{\alpha\gamma t^{\gamma-1}\left\{exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha}\left\{1-exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha-1}}{\lambda^{\gamma}\left\{\left\{1-exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha}+\left\{exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha}\right\}^{2}}$$

A função de Sobrevivência de T é dada por

$$S(t;\alpha,\gamma,\lambda) = 1 - \frac{\left\{1-exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha}}{\left\{1-exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha}+\left\{exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha}}$$

Algumas formas das funções densidade de probabilidade e de sobrevivência da OLLW podem ser vistas na Figura 2.

Figura 2: Formas das funções de densidade probabilidade e de sobrevivência da distribuição OLLW para alguns valores dos parâmetros.



Objetivo

Este trabalho tem como objetivo, através do método do estimar de máxima verossimilhança, estimar os parâmetros dos modelos de distribuição Weibull e odd log-logístico Weibull (OLLW), utilizando um banco de dados com a data de início e término do pontificado dos papas, e fazer um comparativo entre o ajuste e o estimador de Kaplan-Meier.

Metodologia

Os dados foram obtidos virtualmente, visto que é informação pública. Foram analisados 266 registros de exercício de cargo, o período de tempo começa no século I, no ano 30, e se estende até o presente momento. As técnicas a serem utilizadas são provenientes do ramo aplicado da estatística denominado de análise de sobrevivência.

O comportamento da função risco pode ser determinado de maneira empírica, através do método da construção do gráfico do tempo total em teste (curva TTT), proposto por Aarset (1987). A curva TTT pode ser obtida pela seguinte expressão

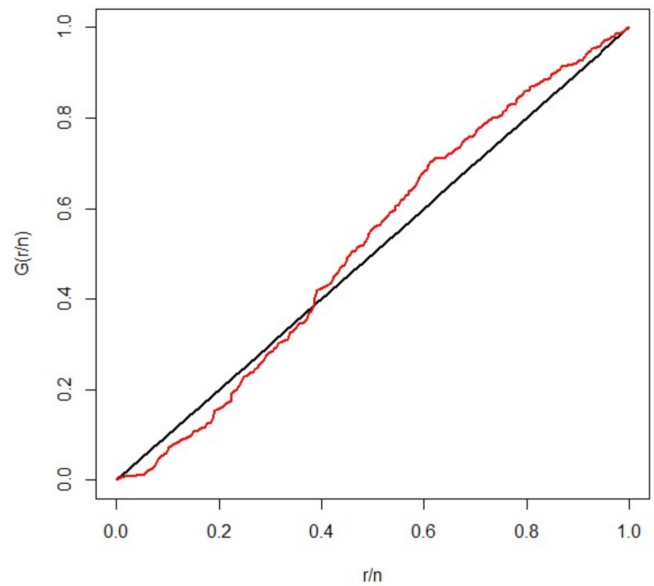
$$G\left(\frac{r}{n}\right) = \frac{\sum_{i=1}^{r} T_{i:n} + (n-r)T_{r:n}}{\sum_{i=1}^{n} T_{i:n}} \quad \text{por} \quad \frac{r}{n},$$

em que n é o tamanho da amostra, $r = 1, \ldots, n$ e $T_{i:n}$, $i = 1, \ldots, n$ são estatísticas de ordem da amostra.

Resultados e discussões

A Curva TTT para o conjunto de dados na Figura 3 mostra uma função risco com uma curvatura primeiramente convexa e depois côncava e indica um risco em forma de "U".

Figura 3: Curva TTT



Na Tabela 1, podem ser vistos as estimativas de máxima verossimilhança (e os correspondentes erros-padrão que estão entre parênteses) dos parâmetros e os valores das estatísticas dos modelos Weibull e OLLW. Os resultados indicam que o modelo OLLW tem os menores valores das estatísticas AIC (Critério de Informação de Akaike), BIC (Critério de Informação Bayesiano) e CAIC (Critério de Informação Akaike Consistente) entre os modelos ajustados, portanto, o modelo OLLW é o mais adequado para os dados. O teste da razão de verossimilhança é apresentado na Tabela 2. Os resultados nessa tabela sugerem que o modelo OLLW produz um ajuste mais adequado a esses dados quando comparado com a distribuição exponencial.

Tabela 1: Ajuste final dos modelos comparados

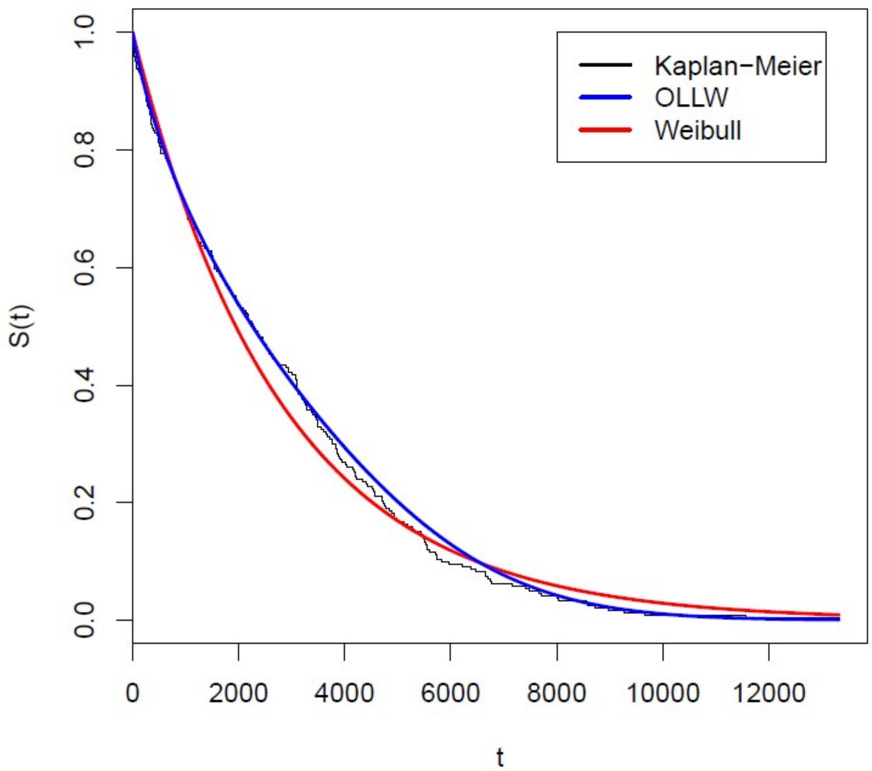
Modelo	λ	γ	α	AIC	CAIC	BIC
Weibull	2807,22	0,9972		4528,1	4528,1	4535,1
	(183,72)	(0,05139)				
OLLW	2807,17	1,6836	0,5358	4514,4	4514,5	4525,2
	(179,70)	(0,1808)	(0,07169)			

Tabela 2: Teste da razão de verossimilhança

Modelo	Estatística do Teste	Valor *p*
Weibull *vs* OLLW	15,7	<0,001

A Figura 4 apresenta a comparação das estimativas da função de sobrevivência segundo Kaplan-Meier e segundo os modelos Weibull e OLLW para os dados. Observa-se pela figura que a distribuição OLLW nos fornece um ajuste satisfatório para os dados em estudo. Logo, conclui-se que o tempo médio em que um papa fica no cargo é de aproximadamente 6,3 anos. Verifica-se também que mais de 80% dos papas não ficam mais que 12 anos no cargo.

Figura 4: Estimativas da função de sobrevivência segundo Kaplan-Meier e segundo os modelos Weibull e OLLW



Conclusão

A distribuição OLLW proposta por apresentou melhor ajuste para os dados em estudo, segundo o teste da razão de verossimilhança e as estatísticas AIC, BIC e CAIC, quando comparada com a distribuição Weibull, se mostrando mais flexível.

Referencias Bibliográficas

AARSET, M. V. How to identify bathtub hazard rate. **Transaction son Reliability**. v. 36, p. 106-108, 1987.

COLOSIMO, E. A.; GIOLO, S.R., **Analise de Sobrevivência Aplicada.**Projeto Fisher – ABE. São Paulo: Edgard Blucher Ltda., 1º Edição 89 p., 2006.

CRUZ, José Nilton da. **A nova família de distribuições odd log-logística: teoria e aplicações.** 2015. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2016.

PASCOA, Marcelino Alves Rosa de. Extensões da distribuição gama generalizada: propriedades e aplicações . 2012. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2012.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013. ISBN 3-900051-07-0, URL http://www.R-project.org/.