

Estimador de Kaplan-Meier na avaliação do tempo de conclusão do Profmat/UFSJ

Kaplan-Meier estimator in the evaluation of the completion time of Profmat/UFSJ

Roberta Cristina Ferreira Arvelos

Universidade Federal de São João del-Rei (UFSJ), Departamento de Matemática, Programa de Mestrado Profissional em Matemática em Rede Nacional, São João del-Rei, MG, Brasil

<http://orcid.org/0000-0002-6823-9822>, betaf85@yahoo.com.br

Carla Regina Guimarães Brighenti

Universidade Federal de São João del-Rei (UFSJ), Departamento de Zootecnia, Programa de Mestrado Profissional em Matemática em Rede Nacional, São João del-Rei, MG, Brasil

<http://orcid.org/0000-0002-7822-3744>, carlabrighenti@ufsj.edu.br

Resumo

A presença de dados incompletos ou censuras no conjunto de dados requer técnicas estatísticas especializadas para acomodar a informação contida nessas observações, tais como a Análise de Sobrevivência. O objetivo deste estudo é a análise dos dados referentes ao tempo de conclusão do curso de Mestrado Profissional em Matemática – Profmat/UFSJ. Foram realizados levantamentos quantitativos para registros das datas de conclusão de curso dos discentes entre os anos de 2011 e 2018. Realizou-se inicialmente uma análise estatística descritiva apresentando a relação de matriculados e desligados, estatísticas do tempo médio de conclusão em função do ano e do Exame Nacional de Qualificação (ENQ) de aprovação, sendo o evento de interesse o tempo até a defesa do TCC. Utilizou-se a Análise de Sobrevivência com a aplicação de técnicas não-paramétricas de Kaplan-Meier e teste Log Rank com o intuito de comparar a data de ingresso no curso e aprovação no ENQ com a data de defesa do TCC pelo discente. Verificou-se que 81,76% não concluem o curso no tempo regular de 24 meses. Desconsiderando as censuras, que correspondem a discentes desligados ou evadidos, o tempo médio foi de 798 dias. Contabilizando as mesmas, este valor aumenta para 1232 dias. Ao comparar as curvas pelo teste de Log-Rank, obtém-se diferença significativa entre o tempo de conclusão para discentes com aprovação na primeira aplicação do ENQ e na segunda, sendo superior para esses últimos.

Informações do Artigo



Histórico do Artigo

Submissão: 08 de maio de 2019.

Aceite: 27 de maio de 2019.

Palavras-chave

Profmat/UFSJ

Análise de Sobrevivência

Técnicas Não-paramétricas

Kaplan-Meier

Log-Rank

Abstract

The presence of incomplete data or censors in the sample dataset requires specialized statistical techniques to accommodate the information contained in such observations, such as Survival Analysis. The objective of this study is to analyze the data related to the time of completion of the Professional Master's Degree in Mathematics – Profmat/UFSJ. Quantitative surveys were carried out to record student completion dates between 2011 and 2018. A descriptive statistical analysis was performed, showing disconnection ratio, statistics of the average completion time were calculated based on the year and the National Qualification Examination (ENQ) in which approval was given, and the event of interest was the time until defense. Survival Analysis was also used with non-parametric techniques such as Kaplan-Meier and Log Rank test. It was concluded that 81.76% did not complete the course in the regular time of 24 months. Disregarding the censors, which correspond to students turned off or evaded, the average time was 798 days. By posting them, this value increases to 1232 days. When comparing the survival curves by the Log-Rank test, we obtain a significant difference between the conclusion time for students with approval in the first application of the ENQ and in the second, being superior for the latter.

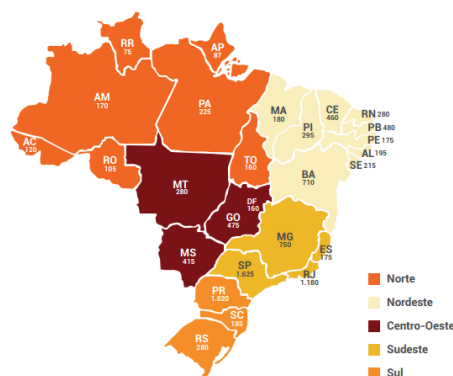
Keywords

Profmat/UFSJ
Survival Analysis
Non-parametric Techniques
Kaplan-Meier
Log-Rank

1 Introdução

O Mestrado Profissional em Matemática em Rede Nacional (Profmat) é um programa de mestrado semipresencial na área de Matemática, formado por uma rede de instituições de Ensino Superior, no contexto da Universidade Aberta do Brasil/Coordenação e Aperfeiçoamento Pessoal de Nível Superior (CAPES), e coordenado pela Sociedade Brasileira de Matemática (SBM), com apoio do Instituto Nacional de Matemática Pura e Aplicada (IMPA), (CALDATTO *et al.*, 2016). Surgiu em 2011, como um programa integrado por 48 instituições, com oferta de 1192 vagas em 54 *campi*. Nos primeiros 7 anos de funcionamento, o programa ofertou mais de 10000 vagas espalhadas por todo o território brasileiro no total (Figura 1).

Figura 1 – Número total de vagas ofertadas pelo Profmat (2011-2017).



Fonte: SBM (2017, p. 4).

A admissão de discentes no Profmat é realizada exclusivamente por meio do Exame Nacional de Acesso (ENA), o qual é regulamentado por edital elaborado pela Coordenação Acadêmica Nacional. No ano de implantação, o Profmat atendia preferencialmente os professores das redes públicas de educação básica. Estes eram beneficiados com a concessão de bolsas, conforme quantitativo disponibilizado pela CAPES (BRASIL, 2013). Atualmente, a cota de bolsas está condicionada à disponibilidade orçamentária da CAPES e é vinculada a alguns critérios particulares. Para a conclusão do Profmat, antes da elaboração da dissertação propriamente dita, o aluno deve ser aprovado no ENQ que consiste em uma avaliação escrita, ofertada duas vezes por ano contendo questões discursivas envolvendo os conteúdos de Números e Funções Reais, Matemática Discreta, Geometria e Aritmética. Excepcionalmente, no ano de 2012, o ENQ foi ofertado três vezes devido à solicitação dos discentes à Comissão Acadêmica Nacional mediante alto índice de reprovação ocorrido no ENQ 2012-2.

O prazo máximo para a conclusão do Profmat pelo discente é definido pela Comissão Acadêmica Institucional, respeitadas suas normas e as da CAPES. Na Universidade Federal de São João del-Rei (UFSJ), o Profmat tem duração mínima de 12 meses e máxima de 24 meses, contados a partir da data da matrícula, podendo o discente solicitar 6 meses de prorrogação. Além disso, há possibilidade de renovar a prorrogação, totalizando 36 meses de curso (CALDATTO *et al.*, 2016).

Apesar do prejuízo, tanto profissional quanto acadêmico, o pedido de prorrogação é muito usual no Profmat, devido ao acúmulo de atividades exercidas pelos discentes, principalmente assumidas após aprovação no ENQ. No entanto, não se tem inferência acerca do tempo de conclusão de curso. Há, também, alunos que são desligados do curso pelos mais variados motivos. Desta forma, o uso da estatística para avaliar o tempo de conclusão precisa considerar tais fatos.

Assim, o objetivo deste trabalho foi avaliar a relação entre o tempo de conclusão do Profmat e a aprovação no ENQ em relação à data de matrícula, utilizando a Análise de Sobrevivência não-paramétrica de Kaplan-Meier.

2 Referencial teórico

Nas pesquisas de longa duração é possível que alguns dos participantes não completem o evento de interesse. Na análise estatística clássica, os dados de tais participantes seriam descartados já que gerariam dados incompletos por não concluir o evento sob estudo durante o período

de acompanhamento. No entanto, uma alternativa possível é o uso desses dados incompletos, por meio da Análise de Sobrevida considerando-os como dados censurados.

Nos estudos de Análise de Sobrevida, os dados incompletos são utilizados na análise final. Nos dados de tempo de conclusão do Profmat, tal fato pode ocorrer porque discentes abandonaram o estudo ou não concluíram o curso antes que a análise fosse realizada. Quando isso acontece, designa-se o indivíduo por censurado. Segundo Costa *et al.* (2015), censurado é o dado correspondente ao indivíduo que não desenvolveu o evento de interesse do estudo, independentemente do motivo ou do que lhe aconteceu posteriormente.

A grande vantagem da Análise de Sobrevida é permitir utilizar a informação de todos os participantes até ao momento em que completam o evento ou são censurados. Assim, a Análise de Sobrevida é a técnica ideal para analisar respostas binárias (ter ou não ter um evento) em estudos longitudinais que se caracterizam por tempo de seguimento diferente entre os indivíduos e perdas de acompanhamento. Portanto, na Análise de Sobrevida, compara-se a rapidez com que os participantes desenvolvem determinado evento, ao contrário de comparar as percentagens de observados que desenvolvem o evento, ao fim de determinado período de tempo (BOTELHO *et al.*, 2009).

A função sobrevivência $S(t)$ fornece a probabilidade de um evento não falhar até o período especificado t , $p(T \geq t)$ e, segundo Penido e Machado (2002), pode ser estimada seguindo modelos não-paramétricos, tais como estimador limite-produto de Kaplan-Meier ou ainda utilizar modelos paramétricos ou semi-paramétricos como modelo de Cox.

O estimador limite-produto (ou estimador de Kaplan-Meier) é caracterizado pela não exigência de uma especificação de função de risco e parte do pressuposto de que as observações consistem em tempos de sobrevivência $(t_1, t_2, t_3, \dots, t_n)$ independentes e distribuídos de forma idêntica (COLOSIMO; GIOLO, 2006). Define-se como:

$$\hat{S}(t) = \prod_j^{t_j < t} \left(1 - \frac{d_j}{n_j}\right) \quad (1)$$

onde

- $t_1 < t_2 < \dots < t_k$, são os k tempos distintos e ordenados de falha,
- d_j é o número de falhas em t_j , $j = 1, \dots, k$, e
- n_j é o número de indivíduos sob risco em t_j , ou seja, os indivíduos que não falharam e não foram censurados até o instante imediatamente anterior a t_j .

$S(t)$ tem uma representação gráfica de uma curva como função escada com degraus nos tempos observados de falha, no tamanho $\frac{1}{n}$, em que n é o tamanho da amostra. Se existirem empates em certo tempo t_j , o tamanho do degrau fica multiplicado pelo número de empates. O estimador de Kaplan-Meier, na sua construção, considera tantos intervalos de tempo quantos forem os números de ocorrências distintas (ROCHA, 2013).

A utilização direta da curva de Kaplan-Meier informa a probabilidade estimada de sobrevivência para um determinado tempo. Se o valor do tempo de interesse estiver ao longo de um degrau da curva de Kaplan-Meier, pode-se também utilizar uma interpolação linear, conforme (2).

$$\frac{t_j - t_{j-1}}{\widehat{S}(t_j) - \widehat{S}(t_{j-1})} = \frac{TM - t_{j-1}}{0,5 - \widehat{S}(t_{j-1})} \quad (2)$$

Outra quantidade que pode ser de interesse é o tempo médio de ocorrência de um evento. Uma estimativa para o tempo médio é então obtida calculando-se a área sob a curva de Kaplan-Meier estimada. Como esta curva é uma função escada, esta integral é simplesmente a soma das áreas de retângulos, isto é,

$$\widehat{t}_m = t_1 + \sum_{j=1}^{k-1} \widehat{S}(t_j)(t_{j+1} - t_j) \quad (3)$$

em que $t_1 < t_2 < \dots < t_k$ são os k tempos distintos e ordenados de falha.

Outra quantidade possivelmente de interesse é o tempo médio restante para aquele indivíduo em que o evento ainda não ocorreu. Este tempo é estimado pela área sob a curva de sobrevivência à direita de t dividido por $\widehat{S}(t)$, isto é,

$$\widehat{vmr}(t) = \frac{\text{área sob a curva } \widehat{S}(t) \text{ à direita de } t}{\widehat{S}(t)} \quad (4)$$

3 Metodologia

Foram realizados levantamentos quantitativos fundamentados em pesquisas no site da Sociedade Brasileira de Matemática, com intuito de registrar as datas de apresentação do trabalho de conclusão de curso de cada discente do Profmat na Universidade Federal de São João del-Rei (UFSJ), que ingressaram no programa entre 2011 e 2016. Complementação dos dados foi

obtida de informações geradas por meio de levantamento documental oriundos da secretaria do Profmat/UFSJ.

Inicialmente, realizou-se uma análise estatística descritiva dos resultados apresentando a relação de matriculados e desligados do Profmat, assim como gráficos de números de defesas ocorridas. Além disso, buscaram-se relações estatísticas entre o tempo médio de conclusão do curso em função do ano e da etapa ENQ de aprovação, se na primeira oportunidade ou segunda.

O foco de interesse da discussão consiste no tempo decorrido de ingresso no programa até a ocorrência do evento “defesa do trabalho de conclusão”, usando a Análise de Sobrevivência.

Foram considerados dados censurados aqueles em que a defesa do trabalho de conclusão de curso não ocorreu até os 36 meses permitidos legalmente, incluindo-se duas prorrogações de seis meses cada, ou o aluno se desligou do curso.

Uma vez reconhecido o evento de interesse e os dados censurados, estimou-se a função de sobrevivência e risco por meio da equação (1) que forneceu a probabilidade $S(t)$ para cada um dos tempos até o período de $t = 36$. A partir da curva de Kaplan-Meier, também foi possível obter estimativas de estatísticas tais como tempo mediano de conclusão de curso TM , (equação 2), tempo médio de conclusão de curso (equação 3) e tempo médio restante para aqueles que ainda não apresentaram sua defesa (equação 4).

Utilizou-se, ainda, como covariável, a aprovação no ENQ nas etapas 1 ou 2, constituindo-se duas curvas de sobrevivência separadamente.

Para avaliar se duas ou mais curvas são estatisticamente equivalentes utilizou-se o teste de Log-Rank, que compara a distribuição da ocorrência dos eventos observados em cada grupo com a distribuição que seria esperada se a incidência fosse igual em todos os grupos (LOPES, 2017). Deve-se observar que o teste Log-Rank é similar ao teste Mantel-Hanzel. Então, com o teste de Log-Rank é possível avaliar a relação entre o tempo de conclusão de curso e o fato de o discente ter sido aprovado no ENQ 1 ou ENQ 2 (incluindo-se o ENQ 3).

Assim, comparam-se as curvas de sobrevivência em sua globalidade e se indica se há diferenças significativas entre as mesmas.

O teste Log-Rank é baseado nas mesmas suposições da curva de sobrevivência de Kaplan-Meier, ou seja, que se a censura não está relacionada ao prognóstico e as probabilidades de sobre-

vivência são as mesmas para os sujeitos recrutados no início ou durante o estudo (BLAND; ALTMAN, 2004).

A estatística de comparação Log-Rank envolve o cálculo das matrizes de variâncias e covariâncias das diferenças entre os valores observados e esperados para cada um dos grupos e segue uma distribuição Qui-quadrado segundo H_0 , onde H_0 indica que as curvas são semelhantes e ordena os tempos de falhas em dois ou mais grupos atribuindo postos.

Sua estatística envolve a diferença entre valores observados e esperados para todos os grupos, mas é necessário que se calcule a matriz de variâncias e covariâncias dessas diferenças (SANTOS *et.al*, 2011).

De acordo com Fidalgo e Scalon (2012), o teste de Log-Rank é baseado na Distribuição Hipergeométrica e é similar ao teste de Mantel-Hanzel. O procedimento Mantel-Hanzel é realizado ao se dispor as contagens dos discentes em relação à etapa de aprovação no ENQ, em K tabelas de contingência de 2×2 , onde K é o número de tempos distintos de defesa. Assim, para cada tempo t , teremos uma tabela de contingência 2×2 , com os membros do grupo (ENQ 1/ENQ 2) em uma das entradas e a resposta ao item defesa (sim/não) na outra.

Então, a particularidade do Log-Rank é particionar a tabela de contingência “final” entre defesas e não-defesas ocorridas, ao longo dos 38 tempos distintos, considerando o número de alunos sob risco, em cada tempo, utilizando desta para formar os dados censurados.

4 Resultados e discussão

Os resultados obtidos foram subdivididos em análise descritiva dos dados, Sobrevivência utilizando o Estimador de Kaplan-Meier e comparação do tempo de conclusão em função do ENQ de Aprovação utilizando o teste de Log-Rank.

4.1 Análise descritiva dos dados Profmat/UFSJ

No período de 2011 a 2016 foram ofertadas 150 vagas na Universidade Federal de São João del-Rei (UFSJ), das quais 147 foram preenchidas. As vagas foram distribuídas entre os dois *campi*: Santo Antônio (CSA) e Alto Paraopeba (CAP). Dentre os ingressantes, há um número considerável

de desligamentos do curso. A Tabela 1 mostra a relação entre o número de alunos inscritos e desligados em cada ano e a relação percentual destes com os dados do Profmat Nacional.

Tabela 1 – Relação entre alunos matriculados e desligados do Profmat/UFSJ.

Ano/Situação	2011	2012	2013	2014	2015	2016
Matriculados	15	29	30	29	29	15
Desligados	1	8	10	11	7	3
% Desligados	6,7	27,6	33,3	37,9	24,2	20
*% Desligados - Profmat Nacional	–	–	20,6	19,3	29,6	14,6

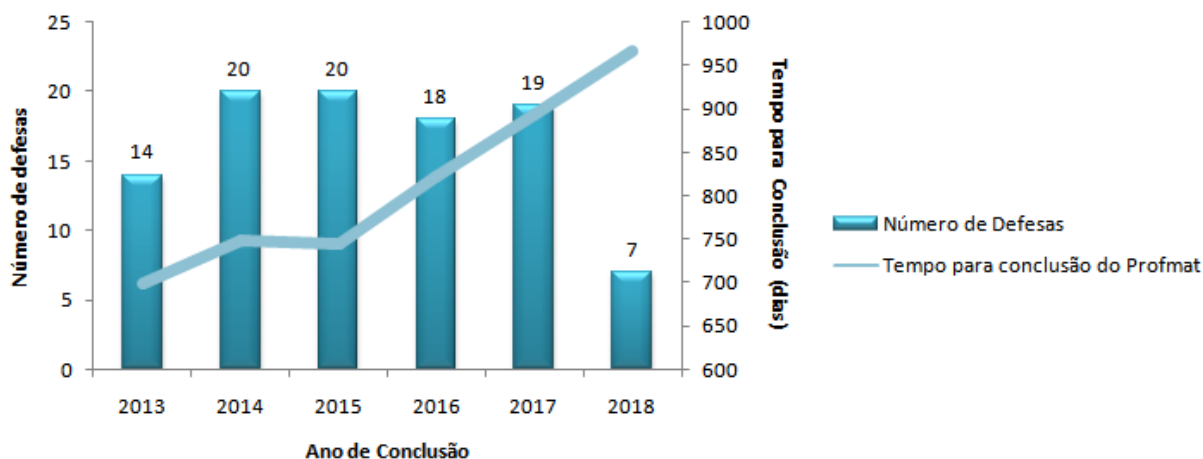
Fonte: Secretaria do Profmat/UFSJ. *SBM (2017).

O percentual de desligamentos do Profmat/UFSJ entre 2011 e 2016 apresentou uma média de 24,95% ao ano, índice um pouco acima dos dados fornecidos pela análise feita pelo Profmat Nacional, no período de 2013 a 2016, que é de 21% ao ano, segundo dados da SBM (2017, p. 25). Já no ano de 2014, observou-se que o número de discentes desligados do programa foi de 37,9%, estando muito acima da média nacional, que foi de 19,3%. Por outro lado, no ano de 2015, a média de desligados no Profmat Nacional (29,6%) é superior se comparada à média do Profmat/UFSJ (24,2%).

É importante ressaltar que, no universo de alunos desligados, há cinco alunos que reingressaram no Programa em outro processo seletivo da UFSJ após não obterem êxito em duas provas consecutivas no ENQ.

Quanto ao número de defesas, os registros aparecem a partir de 2013, ano no qual os discentes completam os dois anos mínimos para a conclusão (Figura 2).

Figura 2 – Número de defesas ocorridas no Profmat/UFSJ entre os anos 2013 e 2017.



Fonte: Secretaria do Profmat/UFSJ.

O número de titulados no Profmat/UFSJ nos anos de 2013 a 2018 foi de 98 discentes, o que corresponde a 66,7% do total de matriculados no Profmat/UFSJ. Percebe-se na Figura 2, que o número de titulados não corresponde ao número esperado da diferença entre matriculados e desligados em cada ano, fato este que está relacionado às prorrogações do tempo de defesa, que, em média, percebe-se um aumento significativo, a partir do ano de 2015, em que esta estatística supera os 800 dias para ocorrência da conclusão do curso. Este fato pode estar relacionado com a aprovação no ENQ realizado na primeira ou segunda aplicação.

Assim, faz-se necessário, para uma análise mais aprofundada, avaliar o tempo de conclusão, em dias, até a data de defesa em cada ano, relacionando-o com a aprovação nas diferentes etapas do ENQ. Os dados referentes a estes registros encontram-se na Tabela 2, que analisa o tempo médio de defesa (em dias).

Tabela 2 – Estatística descritiva do tempo de conclusão do curso Profmat/UFSJ considerando o ano de realização do ENQ.

Ano	ENQ	N	Tempo Médio	Desvio Padrão	Mínimo	1º Quartil	Mediana	3º Quartil	Máximo
2012	1	7	700	3	699	699,0	699	699,0	706
	2	4	701	4	699	699,0	699	700,8	706
	3	3	694	4	692	692,0	692	695,5	699
2013	1	16	748	18	721	741,0	750	752,0	814
	2	4	747	4	741	746,2	748	748,8	751
2014	1	19	745	4	737	740,5	748	748,0	748
	2	2	984	339	744	864,0	984	1104,0	1224
2015	1	6	809	76	710	758,0	814	849,5	916
	2	12	891	98	808	818,0	822	960,8	1035
2016	1	10	873	132	704	760,0	844	994,5	1087
	2	12	882	190	529	746,0	963	1020,8	1068
2017	1	0	0	-	-	-	-	-	-
	2	3	823	0	823	823,0	823	823,0	823
Geral		98	798	119	529	740,0	748	823,0	1224

Fonte: Dados da pesquisa.

Além dos dados citados, existem, ainda, seis discentes aprovados no ENQ em 2014 e 2015, que foram desligados do programa por não apresentarem a defesa do TCC em tempo hábil estipulado pelo regimento do curso. Cabe ressaltar também que nove discentes matriculados no ano de 2016, aprovados no ENQ em 2017, ainda não tinham defendido o TCC até a data final da análise, dados estes que não aparecem na Tabela 2.

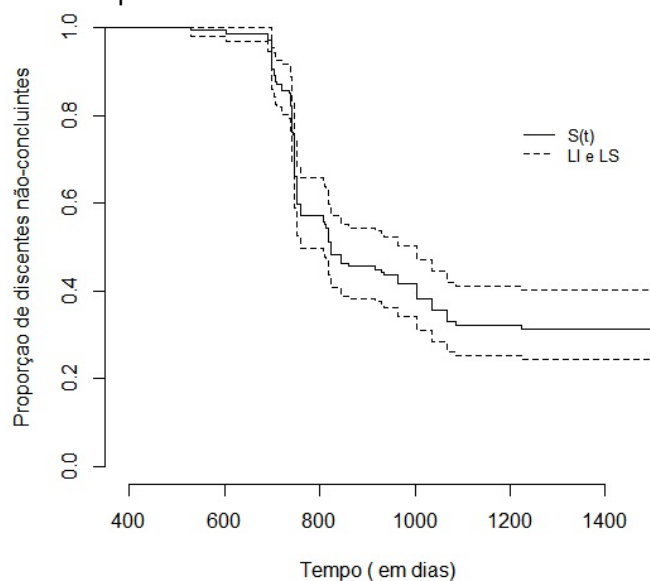
Sendo assim, a partir dos tempos (em dias) registrados, dos 147 alunos matriculados no Profmat/UFSJ, apenas 113 estariam aptos a defender o TCC. No entanto, até o registro final de defesas, em 31 de janeiro de 2019, apenas 98 alunos haviam defendido.

4.2 Estimador de Kaplan-Meier na descrição do tempo de conclusão do Profmat/UFSJ

Para os dados do Profmat/UFSJ, as estimativas da função de sobrevivência $S(t)$, utilizando a equação (1) e seus respectivos intervalos de confiança a 95%, foram obtidos utilizando-se o estimador de Kaplan-Meier e encontram-se na Tabela 3, em que t_j é o tempo decorrido entre o ingresso e a defesa, e $j = 1, \dots, k$, n_j é o número de discentes sob risco, d_j é o número de eventos ocorridos e k corresponde aos 38 tempos diferentes obtidos a partir dos 98 concluintes.

A primeira defesa ocorre nos 529 dias de curso, sendo a estimativa de $\hat{S}(t)$ igual a 0,993. Pode-se observar também que o número máximo de defesas em um mesmo período ocorre aos 748 dias, atingidas por 12 alunos. A representação gráfica de $\hat{S}(t)$, com os respectivos intervalos de 95% de confiança para todo t tal que $400 \leq t \leq 1400$ é mostrado na Figura 3.

Figura 3 – Sobrevivência e respectivos intervalos de 95% de confiança estimados a partir do estimador de Kaplan-Meier para os dados do Profmat/UFSJ.



É possível notar que não há uma “tendência” a zero, indicando uma não expectativa que ocorra conclusão do TCC entre todos os alunos no período até 1400 dias. Este fato está diretamente relacionado aos desligamentos ocorridos no Profmat, e estimado em 29,7%.

A partir da Tabela 3 e da Figura 3, pode-se verificar que os intervalos de confiança obtidos para $S(t)$ são amplos em alguns casos, como, por exemplo, para os desligados. Neste caso, o intervalo corresponde à proporção de $[0,226; 0,389]$.

É importante notar que, apesar de amplo, este intervalo não contém a média nacional, que é de 20,1%, segundo a SBM (2017, p. 25). Isto indica que o tempo de conclusão na UFSJ tende a ser maior que a média nacional. Neste caso, a curva de Kaplan-Meier não atinge o valor zero e o valor do tempo médio de conclusão de curso fica subestimado. Uma alternativa é utilizar a mediana em vez de tempo médio de conclusão.

Pela Tabela 3 pode-se obter o tempo mediano TM , ou seja, tempo necessário para que metade dos alunos concluam o curso. Tal cálculo é feito por meio de interpolação linear, conforme equação (2), da seguinte maneira:

$$\frac{823 - 818}{0,490 - 0,510} = \frac{TM - 818}{0,50 - 0,510} \Rightarrow TM = 820,5$$

Assim, 820,5 dias é uma estimativa do tempo em que 50% dos discentes permanecem sem concluir o mestrado. É importante notar que tal mediana considera os dados censurados. Do contrário, essa estimativa seria igual a 798 dias, conforme calculado na Tabela 2, ou seja, subestimada.

Considerando o tempo regular para o final do curso, segundo Portaria Normativa número 17 (BRASIL, 2009), que é de 2 anos, ou seja, de 730 dias, tem-se, entre os discentes do Profmat/UFSJ, um percentual P de não-concluintes, dado por:

$$\frac{737 - 721}{0,850 - 0,857} = \frac{730 - 721}{P - 0,857} \Rightarrow P = 0,8176$$

Assim, cerca de 81,76% dos discentes não concluem o curso no tempo regular. Da mesma forma, considerando o período de prorrogação de 180 dias, teríamos, então, 910 dias e um percentual D de não-concluintes igual a:

$$\frac{916 - 860}{0,449 - 0,456} = \frac{910 - 860}{D - 0,456} \Rightarrow D = 0,4498$$

Ou seja, quase metade dos alunos do Profmat/UFSJ, no período de 2011 a 2018 recorreu à segunda prorrogação. Estes resultados refletem uma realidade problemática ao redor da estrutura do Profmat que, por muitas vezes, cria uma expectativa de conclusão em dois anos, a qual percebe-se não ser uma situação plausível. Por outro lado, tal ocorrência possibilita perceber uma flexibilidade no curso do Profmat, muitas vezes não permitida em cursos presenciais de mestrados acadêmicos, mostrando, assim, a fidelidade do curso ao perfil profissional do discente para o qual foi destinado, desde sua criação.

No caso do tempo médio de defesa estimado, tem-se que, pela equação(3), o mesmo é de $\widehat{t}_m = 1232,28$ dias.

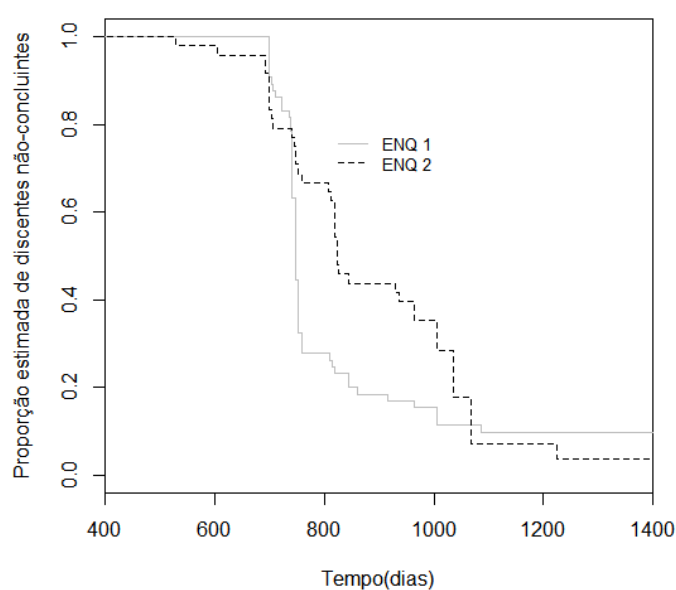
Outra análise se refere ao tempo médio restante para a defesa, que, conforme equação (4), será para discentes que não defenderam até, por exemplo, 916 dias, estimada em $\widehat{vmr} = 56,85$ dias.

É importante observar que, no tempo $t = 737$ tem-se $\widehat{vmr} = 16,13$ dias e no tempo $t = 916$ dias tem-se $\widehat{vmr} = 56,85$ dias. Este fato parece, inicialmente, incoerente, pois o tempo restante esperado de defesa é menor no tempo $t = 737$ do que em $t = 916$ dias. No entanto, tal informação está de acordo com os dados que indicam que se tem uma expectativa de tempo menor até a primeira prorrogação de defesa do que quando o discente já fez uso da segunda prorrogação.

4.3 Comparação do tempo de conclusão em função do ENQ de aprovação

Quando se quer comparar a evolução de grupos segundo algum fator de interesse, podem ser construídas curvas separadas para cada subgrupo e compará-las por um teste estatístico, como, por exemplo, o teste de Log-Rank. Para verificar se há diferença entre o tempo de conclusão do curso para o alunos com aprovação nas diferentes etapas do ENQ, foi considerada a construção de curvas de sobrevivência incluindo a covariável ENQ (Figura 4).

Figura 4 – Tempo de conclusão do curso para alunos com aprovação nas diferentes etapas do ENQ.



Fonte: Dados da pesquisa.

A partir das curvas de sobrevivência geradas pelos tempos, a partir da Tabela 3, em cada caso, pode-se estimar que o tempo mediano de conclusão para o discente com a aprovação no ENQ 1 é de 747,65 dias e no ENQ 2 é de 821,5 dias.

É possível observar que há uma redução no tempo de conclusão para a aprovados no ENQ 1. Este fato pode ser comprovado pelo teste de Log-Rank. Log-Rank é um teste de hipóteses usado para comparar as distribuições de sobrevivência de duas ou mais amostras. É um teste não-paramétrico e apropriado para ser usado quando os dados estão desalinhados e censurados.

Neste caso, foram aprovados, no ENQ 1, um total de 65 alunos e no ENQ 2 (incluindo o ENQ 3, no ano de 2012), 48 alunos.

O valor calculado do teste de Log-Rank é dado por $\chi^2 = 4,2$ obtendo-se o valor p de 0,04 indicando uma significância estatística entre o tempo de conclusão de curso entre os discentes aprovados no ENQ 1 ou posterior.

Tabela 3 – Estimativa de Kaplan-Meier para dados do Profmat/UFSJ com intervalo de confiança de 95%.

t_j	n_j	d_j	$\hat{S}(t)$	Erro Padrão	<i>L.I.</i>	<i>L.S.</i>
529	147	1	0,993	0,00678	0,980	1,000
605	146	1	0,986	0,00955	0,968	1,000
692	145	2	0,973	0,01342	0,947	0,999
699	143	10	0,905	0,02421	0,859	0,953
704	133	2	0,891	0,02569	0,842	0,943
706	131	2	0,878	0,02704	0,826	0,932
710	129	1	0,871	0,02767	0,818	0,927
721	128	2	0,857	0,02886	0,802	0,916
737	126	1	0,850	0,02942	0,795	0,910
738	125	1	0,844	0,02996	0,787	0,904
739	124	1	0,837	0,03048	0,779	0,899
740	123	2	0,823	0,03147	0,764	0,887
741	121	9	0,762	0,03513	0,696	0,834
744	112	1	0,755	0,03547	0,689	0,828
747	111	2	0,741	0,03611	0,674	0,816
748	109	12	0,660	0,03907	0,588	0,741
751	97	5	0,626	0,03991	0,552	0,709
752	92	4	0,599	0,04043	0,524	0,683
760	88	4	0,571	0,04082	0,497	0,657
808	84	1	0,565	0,04089	0,490	0,651
809	83	1	0,558	0,04096	0,483	0,644
811	82	1	0,551	0,04102	0,476	0,638
814	81	1	0,544	0,04108	0,469	0,631
818	80	5	0,510	0,04123	0,435	0,598
823	75	3	0,490	0,04123	0,415	0,578
825	72	1	0,483	0,04122	0,409	0,571

t_j	n_j	d_j	$\hat{S}(t)$	Erro Padrão	$L.I.$	$L.S.$
844	71	3	0,463	0,04112	0,389	0,551
860	68	1	0,456	0,04108	0,382	0,544
916	67	1	0,449	0,04102	0,375	0,537
930	66	1	0,442	0,04096	0,369	0,530
936	65	1	0,435	0,04089	0,362	0,523
963	64	3	0,415	0,04064	0,342	0,503
1005	49	4	0,381	0,04070	0,309	0,470
1035	45	3	0,356	0,04054	0,284	0,445
1068	42	3	0,330	0,04021	0,260	0,419
1087	39	1	0,322	0,04006	0,252	0,411
1224	38	1	0,313	0,03989	0,244	0,402
2168	19	1	0,297	0,04106	0,226	0,389

Fonte: Dados da pesquisa.

5 Considerações finais

A utilização da estatística não-paramétrica de Kaplan-Meier, com inclusão dos dados censurados, permitiu melhorar as estimativas do tempo de conclusão do Profmat.

A análise dos dados do tempo de conclusão do Profmat com a inclusão do tempo de prorrogação permite estimar melhor o tempo esperado para término de curso. Esta estimativa pode auxiliar órgãos de pesquisa e fomento para realizar ações envolvendo discentes e docentes do Profmat, tais como promoção de workshop para divulgação dos trabalhos em datas definidas e planejamento das datas da emissão dos selos nacionais de conclusão emitidos pela SBM.

No período avaliado e instituição onde foi feita a coleta de dados, percebeu-se um aumento significativo no tempo de conclusão a partir de 2015. Deve-se ressaltar que este período coincide com a diminuição do oferecimento de bolsas aos integrantes do Profmat.

O tempo médio de conclusão para discentes aprovados na primeira aplicação do Exame Nacional de Qualificação é de 747,65 dias, sendo este tempo estatisticamente inferior ($P < 0,05$) ao dos alunos que são aprovados no segundo exame, que é de 821,5 dias. Desta forma, pode-se dizer que a aprovação no primeiro exame de qualificação oferecida permite redução no tempo de conclusão do Profmat.

A avaliação do tempo de conclusão do curso de Mestrado Profissional em Matemática na instituição UFSJ, mostrou que, o curso de modalidade semi-presencial, pode necessitar de um acom-

panhamento mais próximo dos discentes por parte dos docentes, para que o tempo de término não seja superior ao desejado para o mestrado.

Referências

- BLAND, J. M.; ALTMAN, D. G. The logrank test. **British Medical Journal**, v. 328, p. 1073, 2004.
- BOTELHO, F.; SILVA, C.; CRUZ, F. Epidemiologia Explicada: análise de sobrevivência. **Acta Urológica**, v. 26, n. 4, p. 33-38, 2009.
- BRASIL. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). **Avaliação Suplementar Externa do Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT)**. Brasília, DF, 2013.
- BRASIL. **Portaria Normativa Nº 17, de 28 de dezembro de 2009**. Diário Oficial da União Nº 248. Disponível em: https://www.capes.gov.br/images/stories/download/legislacao/PortariaNormativa_17MP.pdf. Acesso em: 30 jan. 2019.
- CALDATTO, M. E.; PAVANELLO, R. M.; FIORENTINI, D. O PROFMAT e a Formação do Professor de Matemática: uma análise curricular a partir de uma perspectiva processual e descentralizadora. **Bolema**, v. 30, n. 56, p. 906, 2016.
- COLOSIMO E. A.; GIOLO S. R. **Análise de Sobrevivência Aplicada**. v. 1. São Paulo: Edgar Blücher, 2006.
- COSTA, E. O.; SILVA, C. S.; SOARES, M. E. S. M.; SILVA, R. G.; AMARAL, P. B. Análise do tempo de internação de crianças com pneumonia em Hospital Público de João Pessoa-PB. **Revista Brasileira de Ciências da Saúde**, v. 18, n. 2, p. 147-150, 2015.
- FIDALGO, A. M.; SCALON, J. D. Using Mantel-Haenszel methods for detecting differential item functioning. **Psicologia: Reflexão e Crítica**, v. 25, n. 1, p. 60-68, 2012.
- LOPES, M. T. **Modelos estatísticos para suporte a avaliação cirúrgica em crianças portadoras de cardiopatias congênitas**. 2017. 120f. Dissertação (Mestrado em Modelos de Decisão e Saúde) – Universidade Federal da Paraíba, João Pessoa, 2017.
- PENIDO, M.; MACHADO, A. F. Desemprego: evidências da duração no Brasil metropolitano. **Anais do XIII Encontro Nacional de Estudos Populacionais**, p. 1-22, 2002. Disponível em: <http://www.abep.org.br/publicacoes/index.php/anais/article/view/1215/1179>. Acesso em: 28 jun. 2019.
- ROCHA, T. S. **Modelos de regressão discretos para dados grupados: uma aplicação em avaliação de risco em produto de crédito parcelado**. 2013. 31 f. Monografia (Graduação em Estatística) – Departamento de Estatística, Universidade de Brasília, Brasília, 2013.
- SANTOS, L. M.; SILVA, G. M.; NEVES, J. A. B. Risco de sobrevivência de micro e pequenas empresas comerciais. **Revista de Contabilidade e Organizações**, v. 5, n. 11, p. 107-124, 2011.

SBM. Sociedade Brasileira de Matemática. **Profmat: uma reflexão e alguns resultados**. 2017.

Disponível em:

<http://www.profmat-sbm.org.br/2017/06/20/profmat-uma-reflexao-e-alguns-resultados>.

Acesso em: 30 jan. 2019.