

**UNIVERSIDADE FEDERAL FLUMINENSE
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS GRADUAÇÃO *STRICTO SENSU* EM ENGENHARIA DE
PRODUÇÃO
CURSO DE MESTRADO ACADÊMICO EM ENGENHARIA DE PRODUÇÃO**

LUIZ FREDERICO HORÁCIO DE SOUZA DE BARROS TEIXEIRA

**ANÁLISE DOS TESTES DE ADERÊNCIA EM TÁBUAS ATUARIAIS: UMA
CONTRIBUIÇÃO PARA O SISTEMA DE PROTEÇÃO SOCIAL DOS MILITARES
DAS FORÇAS ARMADAS**

**NITERÓI, RJ
2020**

LUIZ FREDERICO HORÁCIO DE SOUZA DE BARROS TEIXEIRA

**ANÁLISE DOS TESTES DE ADERÊNCIA EM TÁBUAS ATUARIAIS: UMA
CONTRIBUIÇÃO PARA O SISTEMA DE PROTEÇÃO SOCIAL DOS MILITARES
DAS FORÇAS ARMADAS**

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Engenharia de Produção da Universidade Federal Fluminense como requisito parcial para obtenção do Grau de Mestre em Engenharia de Produção

Orientador:

Prof. Dr. CARLOS FRANCISCO SIMÕES GOMES

Coorientador:

Prof. Dr. MARCOS DOS SANTOS

Niterói, RJ
2020

Ficha catalográfica automática - SDC/BEE

T266a Teixeira, Luiz Frederico Horácio de Souza de Barros
Análise dos testes de aderência em tábuas atuariais : uma contribuição para o sistema de proteção social dos militares das forças armadas. / Luiz Frederico Horácio de Souza de Barros Teixeira ; Carlos Francisco Simões Gomes, orientador ; Marcos dos Santos, coorientador. Niterói, 2020.
116 f. : il.

Dissertação (mestrado)-Universidade Federal Fluminense, Niterói, 2020.

DOI: <http://dx.doi.org/10.22409/TPP.2020.m.10479532770>

1. Pesquisa operacional. 2. Apoio à decisão. 3. Sistemas. 4. Logística. 5. Produção intelectual. I. Gomes, Carlos Francisco Simões, orientador. II. Santos, Marcos dos, coorientador. III. Universidade Federal Fluminense. Escola de Engenharia. IV. Título.

CDD -

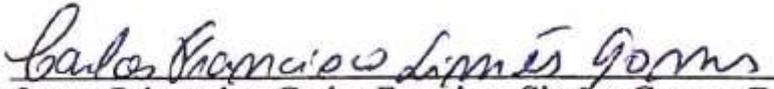
LUIZ FREDERICO HORÁCIO DE SOUZA DE BARROS TEIXEIRA

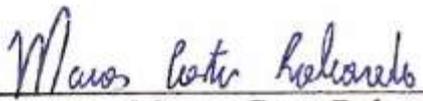
ANÁLISE DOS TESTES DE ADERÊNCIA EM TÁBUAS ATUARIAIS: UMA
CONTRIBUIÇÃO PARA O SISTEMA DE PROTEÇÃO SOCIAL DOS MILITARES DAS
FORÇAS ARMADAS

Dissertação apresentada ao Programa de Pós-
Graduação *Stricto Sensu* em Engenharia de
Produção da Universidade Federal Fluminense
como requisito parcial para obtenção do Grau de
Mestre em Engenharia de Produção

Aprovada em 06 de fevereiro de 2020.

BANCA EXAMINADORA


Professor Orientador: Carlos Francisco Simões Gomes, D.Sc.
Universidade Federal Fluminense


Marcos Costa Roboredo, D.Sc.
Universidade Federal Fluminense


Professor Coorientador: Marcos dos Santos, D.Sc.
Instituto Militar de Engenharia

Niterói, RJ
2020

Dedico este trabalho à minha família. Aos meus pais, Luiz Frederico e Ester, que sempre me apoiaram e incentivaram a estudar; aos meus queridos filhos Frederico e Alice, que me proporcionaram momentos de ternura e descontração em meio à seriedade do estudo; e, em especial, à minha amada esposa Bruna, pessoa com quem amo partilhar a vida. Obrigado pela paciência, pelo incentivo, pela força e, principalmente, por ser meu porto seguro e me trazer paz em meio à correria de cada semestre. Grato pela sua compreensão com as minhas horas de ausência e por diversas vezes assumir o "papel de pai". Eu não teria conseguido sem você. Te amo.

AGRADECIMENTOS

A Deus, pelo dom da vida e por me permitir alcançar mais esta conquista.

Ao meu orientador, Prof. Dr. Carlos Francisco Simões Gomes, pelos diversos ensinamentos profissionais e pessoais, pela paciência e dedicação durante este período de orientação. Em especial, agradeço pelo voto de confiança em aceitar-me no seu seletivo grupo de pesquisa e pelos diversos momentos agradáveis vividos em congressos e simpósios durante o curso.

Ao meu coorientador, Capitão de Fragata Marcos dos Santos, Prof. Dr., por ter compartilhado toda a sua pesquisa, pelo companheirismo e pelas dicas valorosas que facilitaram que eu trilhasse este caminho.

Aos demais professores do programa de pós-graduação em Engenharia de Produção da Universidade Federal Fluminense, pela excelência em sala de aula, abrilhantando este curso. Em especial, à Prof^a Dr^a. Priscilla Cristina Cabral Ribeiro, membra da Banca do Seminário de Qualificação e ao Prof. Dr. Marcos Costa Roboredo, membro da Banca Examinadora desta dissertação, pelas críticas e comentários que enriqueceram este trabalho.

À Marinha do Brasil e ao Centro de Análises de Sistemas Navais (CASNAV), particularmente na pessoa do Diretor, Capitão de Mar e Guerra Mauricio Pires Malburg da Silveira, pelo apoio institucional para que eu pudesse me dedicar ao curso e participar de eventos acadêmicos que contribuíssem com a minha formação.

Aos amigos do Laboratório de Estudos de Governança, Gestão e Otimização (LEGGO), em especial aos pesquisadores Capitão de Mar e Guerra Ernesto Rademaker Martins, Prof. Msc., e Prof. Msc. Fabrício Maione Tenório, pelo apoio e convívio cordial e agradável.

Aos amigos de turma, contemporâneos na Escola Naval e companheiros de praça d'armas de navios da Marinha que, assim como eu, optaram pela carreira acadêmica, Thiago Muzi Machado, Antônio Luiz Carlucio Doneda, Diego Fróes e Coelho da Silva, Márcio Selemen Coelho, Diego Felipe Gimenez de Andrade e Felipe Araújo Marins, pela troca de conhecimentos, experiências e “bizus” nesta nossa nova jornada.

Aos amigos Ronaldo Cesar Evangelista dos Santos e Prof. Msc. Thierry Faria da Silva Gregorio, do CASNAV e da Arcadia Sistemas, por compartilharem suas experiências e conhecimentos adquiridos como integrantes do projeto “Estudo Atuarial das Pensões dos Militares” e pela licença de utilização do *software* Aderência 4D.

RESUMO

Os militares das Forças Armadas não estão submetidos a nenhum regime de previdência, sendo amparados por um sistema de proteção social específico. Este sistema não é baseado no acúmulo de capital por parte dos beneficiários, sendo sustentado pela previsão orçamentária da União. Para a correta mensuração de custos com os militares veteranos e pensionistas, é necessário, portanto, o contínuo estudo e acompanhamento das suas previsões atuariais. Estas, por sua vez, requerem a adequada seleção das tábuas atuariais, de modo a representar o mais fidedignamente possível a amostra da população em análise. O objetivo deste trabalho é determinar, dentre os principais testes de aderência difundidos no meio acadêmico, qual o teste de aderência que melhor se aplica para a seleção das tábuas atuariais para uso nos dados dos militares das Forças Armadas Brasileiras. Desta forma, conseqüentemente, possibilitando a seleção das tábuas atuariais que apresentam melhor aderência para aquela amostra populacional. Para tal, foi realizada uma revisão bibliométrica sobre o tema, que deslumbrou um relativo ineditismo nesta pesquisa aplicada aos militares das Forças Armadas, comprovando a sua contribuição para a Academia. Adicionalmente, considerando que um militar durante a sua carreira pode se encontrar em diferentes estados biométricos e sociais, este trabalho apresenta as probabilidades de estado estável, a partir da aplicação da teoria de Cadeia de Markov. A contribuição para a sociedade pode ser comprovada pela abrangência do estudo, que afeta diretamente um sistema de proteção social com mais de um milhão de beneficiários. O presente estudo pode ser considerado uma pesquisa exploratória em continuação ao trabalho de Santos (2018), contribuindo para verificar a robustez do modelo proposto por aquele autor, o que denota a relevância deste trabalho para o estado da arte. Os resultados apresentam que o teste de Kolmogorov-Smirnov foi o mais adequado para a amostra populacional das Forças Armadas analisada.

Palavras-chave: Testes de Aderência; Forças Armadas Brasileiras; Cadeia de Markov; Cálculo Atuarial.

ABSTRACT

The military of the Armed Forces is not subject to any social security system and is supported by a specific social protection system. This system is not based on the accumulation of capital by the beneficiaries, but is supported by the budget forecast of the Union. For the correct measurement of costs with the military veterans and pensioners, therefore, the continuous study and monitoring of their actuarial forecasts is necessary. These, in turn, require the proper selection of actuarial tables, so as to represent the sample of the population under analysis as reliably as possible. The objective of this paper is to determine, among the main goodness-of-fit test disseminated in the academic environment, which goodness-of-fit test best applies to the selection of actuarial tables for the military of the Brazilian Armed Forces. Thus, allowing the selection of actuarial tables that present better adherence to that population sample. To this end, a bibliometric review was conducted on the subject, which denoted relative unprecedentedness in this research applied to the military of the Armed Forces, proving their contribution to the Academy. Additionally, considering that a military man during his career may be in different biometric and social states, this paper presents the steady state probabilities, based on the application of Markov Chain theory. The contribution to society can be evidenced by the scope of the study, which directly affects a social protection system with more than one million potential beneficiaries. This study can be considered an exploratory research in continuation to the work of Santos (2018), contributing to verify the robustness of the model proposed by that author, which denotes the relevance of this work to the state of the art. The results show that the Kolmogorov-Smirnov test was the most appropriate for the Armed Forces population sample analyzed.

Keywords: Goodness-of-fit tests; Brazilian Armed Forces; Markov chain; Actuarial calculation.

SUMÁRIO

1. INTRODUÇÃO	17
1.1. CONTEXTUALIZAÇÃO	17
1.2. SITUAÇÃO PROBLEMA	18
1.3. PERGUNTAS DA PESQUISA	19
1.4. OBJETIVO GERAL E OBJETIVOS SECUNDÁRIOS	20
1.5. JUSTIFICATIVA E RELEVÂNCIA DA PESQUISA	21
1.6. DELIMITAÇÃO DA PESQUISA	22
1.7. ESTRUTURA DO TRABALHO	23
2. REVISÃO DA LITERATURA	23
2.1. BASE CONCEITUAL	25
2.1.1. TESTE DE ADERÊNCIA (GOODNESS-OF-FIT – GOF - TEST)	26
2.1.2. PRINCIPAIS TESTES DE ADERÊNCIA	29
2.1.2.1. TESTE DO QUI-QUADRADO	29
2.1.2.2. TESTE DE DESVIO QUADRÁTICO MÉDIO (ROOT MEAN SQUARE DEVIATION – RMSD)	30
2.1.2.3. TESTE BINOMINAL	31
2.1.2.4. TESTE DE KOLMOGOROV-SMIRNOV	32
2.1.2.5. TESTE DE LILLIEFORS	33
2.1.2.6. TESTE DE SHAPIRO-WILK	34
2.2. MAPEAMENTO DO ESTADO DA ARTE	35
2.2.1. PESQUISA NAS BASES SCOPUS E WEB OF SCIENCE	35
2.1.2. ESTATÍSTICA DOS ARTIGOS SELECIONADOS	44
3. METODOLOGIA DA PESQUISA	48
4. MODELAGEM DOS TESTES DE ADERÊNCIA	51
4.1 O SOFTWARE ATUARIAL DAS FORÇAS ARMADAS (SAFA)	51
4.2 A SELEÇÃO DE TÁBUAS ATUARIAIS MAIS ADERENTES	53
4.2.1 – TÁBUAS ATUARIAIS	53
4.2.1.1 – TÁBUAS DE MORTALIDADE	53
4.2.1.2 – TÁBUAS DE ENTRADA EM INVALIDEZ	54
4.2.1.3 – TÁBUAS DE MORTALIDADE DE INVÁLIDOS	54
4.2.2 – A SELEÇÃO DAS TÁBUAS ATUARIAIS E O "HIPERCUBO DE ADERÊNCIA"	55
4.3 APLICAÇÃO DOS TESTES DE ADERÊNCIA	58
4.3.1 RESULTADO DE MORTALIDADE DE ATIVOS, INATIVOS E PENSIONISTAS DAS FORÇAS ARMADAS	59
4.3.1.1 APLICAÇÃO DOS TESTES COM CÁLCULO DA IDADE EXATA	59
4.3.1.2 APLICAÇÃO DOS TESTES COM CÁLCULO POR CLASSES DE IDADES ..	63

4.3.1.3 ANÁLISE DOS RESULTADOS PARA MORTALIDADE DE ATIVOS, INATIVOS E PENSIONISTAS	68
4.3.2 RESULTADO DE MORTALIDADE DE INVÁLIDOS DAS FORÇAS ARMADAS	70
4.3.2.1 APLICAÇÃO DOS TESTES COM CÁLCULO DA IDADE EXATA	70
4.3.2.2 APLICAÇÃO DOS TESTES COM CÁLCULO POR CLASSES DE IDADES..	73
4.3.2.3 ANÁLISE DOS RESULTADOS PARA MORTALIDADE DE INVÁLIDOS	77
4.3.3 RESULTADO DE ENTRADA EM INVALIDEZ DAS FORÇAS ARMADAS	80
4.3.3.1 APLICAÇÃO DOS TESTES COM CÁLCULO DA IDADE EXATA	80
4.3.3.2 APLICAÇÃO DOS TESTES COM CÁLCULO POR CLASSES DE IDADES..	84
4.3.3.3 ANÁLISE DOS RESULTADOS PARA ENTRADA EM INVALIDEZ	89
4.3.4 CONSIDERAÇÕES FINAIS SOBRE A ANÁLISE DE RESULTADOS.....	92
5. ESTUDO MARKOVIANO SOBRE ESTADOS BIOMÉTRICOS	93
6. CONCLUSÃO E RECOMENDAÇÕES PARA TRABALHOS FUTUROS	107
6.1 CONCLUSÃO.....	107
6.2. RECOMENDAÇÕES PARA TRABALHOS FUTUROS	109
REFERÊNCIAS BIBLIOGRÁFICAS	110

LISTA DE EQUAÇÕES

Equação 1: Teste do Qui-quadrado.	29
Equação 2: Teste do Desvio Quadrado Médio.	30
Equação 3: Teste Binominal	31
Equação 4: Valor Supremo de Teste de Kolmogorov-Smirnov	32
Equação 5: Diferença entre as funções de Distribuição de Kolmogorov-Smirnov	32
Equação 6: Estatística do Teste de Kolmogorov-Smirnov	32
Equação 7: Hipótese Nula vs Hipótese Alternativa	32
Equação 8: Coeficientes do Teste de Lilliefors	34
Equação 9: Valor Crítico para $n \leq 100$	34
Equação 10: Valor Crítico para $n > 100$	34
Equação 11: P-valor para a correção do Lilliefors	34
Equação 12: Estatística do Teste de Shapiro-Wilk	35
Equação 13: Estatística do Teste de Anderson-Darling	47
Equação 14: Função Cumulativa Teórica.....	47

LISTA DE GRÁFICOS

Gráfico 1: Tábua de melhor aderência Teste KS para todas as idades.....	60
Gráfico 2: Tábua de melhor aderência Teste DQM para todas as idades.....	60
Gráfico 3: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 20 a 90 anos....	61
Gráfico 4: Tábua de melhor aderência Teste KS para intervalo de 20 a 90 anos.....	61
Gráfico 5: Tábua de melhor aderência Teste DQM para intervalo de 20 a 90 anos	62
Gráfico 6: Tábua de melhor aderência Teste KS para todas as idades	63
Gráfico 7: Tábua de melhor aderência Teste DQM para todas as idades.....	64
Gráfico 8: Tábua de melhor aderência Teste KS para intervalo de 20 a 90 anos.....	64
Gráfico 9: Tábua de melhor aderência Teste DQM para intervalo de 20 a 90 anos	65
Gráfico 10: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 22 a 87 anos..	66
Gráfico 11: Tábua de melhor aderência Teste KS para intervalo de 22 a 87 anos.....	66
Gráfico 12: Tábua de melhor aderência Teste DQM para intervalo de 22 a 87 anos	67
Gráfico 13: Tábua de melhor aderência Teste KS para todas as idades.....	70
Gráfico 14: Tábua de melhor aderência Teste DQM para todas as idades	70
Gráfico 15: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 21 a 93 anos..	71
Gráfico 16: Tábua de melhor aderência Teste KS para intervalo de 21 a 93 anos.....	71
Gráfico 17: Tábua de melhor aderência Teste DQM para intervalo de 21 a 93 anos	72
Gráfico 18: Tábua de melhor aderência Teste KS para todas as idades.....	73
Gráfico 19: Tábua de melhor aderência Teste DQM para todas as idades.....	73
Gráfico 20: Tábua de melhor aderência Teste KS para intervalo de 21 a 90 anos.....	74
Gráfico 21: Tábua de melhor aderência Teste DQM para intervalo de 21 a 90 anos	74
Gráfico 22: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 74 anos .	75
Gráfico 23: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 74 anos .	76
Gráfico 24: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 74 anos .	76
Gráfico 25: Tábua de melhor aderência Teste KS para todas as idades	80
Gráfico 26: Tábua de melhor aderência Teste DQM para todas as idades.....	81
Gráfico 27: Tábua de melhor aderência Teste KS para intervalo de 19 a 90 anos.....	81
Gráfico 28: Tábua de melhor aderência Teste DQM para intervalo de 19 a 90 anos	82
Gráfico 29: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 41 anos .	82
Gráfico 30: Tábua de melhor aderência Teste KS para intervalo de 25 a 41 anos	83
Gráfico 31: Tábua de melhor aderência Teste DQM para intervalo de 25 a 41 anos	83

Gráfico 32: Tábua de melhor aderência Teste KS para todas as idades.....	85
Gráfico 33: Tábua de melhor aderência Teste DQM para todas as	85
Gráfico 34: Tábua de melhor aderência Teste KS para intervalo de 19 a 90 anos.....	86
Gráfico 35: Tábua de melhor aderência Teste DQM para intervalo de 19 a 90 anos	86
Gráfico 36: Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 48 anos..	87
Gráfico 37: Tábua de melhor aderência Teste KS para intervalo de 25 a 48 anos.....	88
Gráfico 38: Tábua de melhor aderência Teste DQM para intervalo de 25 a 48 anos	88

LISTA DE ILUSTRAÇÕES

Figura 1: Organização do Capítulo 2 – Revisão da Literatura.	25
Figura 2: Organização do item 2.1 Base Conceitual.	26
Figura 3: Distribuição aleatória de amostras com similaridade crescente à curva normal.....	27
Figura 4: Interpretação gráfica da estatística do teste Kolmogorov-Smirnov	33
Figura 5: Seleção de Artigos Relevantes	38
Figura 6: Metodologia da pesquisa.....	48
Figura 7: Aplicação dos Testes de Aderência.	50
Figura 8: Cubo de Aderência.....	55
Figura 9: Grafo de um processo de Markov de dois estados.....	93
Figura 10: Matriz de transição em n etapas da Cadeia de Markov.....	94
Figura 11: Grafo de doze estados biométricos e sociais.....	96
Figura 12: Matriz de Transição de doze estados biométricos e sociais.....	97
Figura 13: Grafo de seis estados biométricos e sociais..	98
Figura 14: Matriz de Transição de seis estados biométricos e sociais..	99
Figura 15: Matriz de transição de seis estados com valores numéricos normalizados.....	106
Figura 16: Matriz de probabilidade estável..	106

LISTA DE QUADROS

Quadro 1: Lista de artigos selecionados para a pesquisa	39
Quadro 2: Tábuas de mortalidade de válidos utilizadas.	57
Quadro 3: Tábuas de entrada em invalidez e mortalidade de inválidos utilizadas	57
Quadro 4: Tábuas de Mortalidade de Ativos, Inativos e Pensionistas Melhor Aderidas para Cada Teste.	69
Quadro 5: Tábuas de Mortalidade de Inválidos Melhor Aderidas para Cada Teste.....	79
Quadro 6: Tábuas de Entrada em Invalidez Melhor Aderidas para Cada Teste.....	91

LISTA DE TABELAS

Tabela 1: Palavras-chave da pesquisa inicial .	36
Tabela 2: Palavras-chave da pesquisa refinada.	37
Tabela 3: Distribuição dos artigos selecionados por ano de publicação.	45
Tabela 4: Distribuição dos artigos selecionados por periódico.	45
Tabela 5: Distribuição dos artigos selecionados por autor.	46
Tabela 6: Distribuição dos artigos selecionados por número de autores.	47
Tabela 7: Quantitativo de militares ativos/inativos e pensionistas em 2015.	59
Tabela 8: Quantidade de Tábuas de Mortalidade Geral aderidas a partir de cada teste .	62
Tabela 9: Tábuas de Mortalidade Geral com melhor aderência a partir de cada teste.	63
Tabela 10: Quantidade de Tábuas de Mortalidade Geral aderidas a partir de cada teste.	67
Tabela 11: Tábuas de Mortalidade Geral com melhor aderência a partir de cada teste.	67
Tabela 12: Quantidade de Tábuas Mortalidade de Inválidos aderidas a partir de cada teste ...	72
Tabela 13: Tábuas Mortalidade de Inválidos com melhor aderência a partir de cada teste ...	72
Tabela 14: Quantidade Tábuas de Mortalidade de Inválidos aderidas a partir de cada teste ...	77
Tabela 15: Tábuas de Mortalidade de Inválidos com melhor aderência para cada teste	77
Tabela 16: Quantidade de Tábuas de Entrada em Invalidez aderidas a partir de cada teste. ...	84
Tabela 17: Tábuas de Entrada em Invalidez com melhor aderência a partir de cada teste	84
Tabela 18: Quantidade Tábuas de Mortalidade de Inválidos aderidas a partir de cada teste. ...	89
Tabela 19: Tábuas de Mortalidade de Inválidos com melhor aderência para cada teste.....	89
Tabela 20: Probabilidades dos Estados Biométricos.....	100
Tabela 21: Razões de probabilidades não normalizadas.	105
Tabela 22: Tipos de Normalização.....	105
Tabela 23: Probabilidades de estado estável.	106

LISTA DE ABREVIATURAS E SIGLAS

AD – Teste de Aderência de Anderson-Darling
ANEMAR - Anuário Estatístico da Marinha do Brasil
BD – Banco de Dados
BIEG – Banco de Informações Estratégicas e Gerenciais
CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CASNAV – Centro de Análises de Sistemas Navais
CDF – *Cumulative Distribution Function*
CF – Constituição Federal
CRFB - Constituição da República Federativa do Brasil
CVM – Teste de Cramér-von Mises
DAP – Teste de D’Agostino e Pearson
DQM – Desvio Quadrático Médio
EB – Exército Brasileiro
EC – Emenda Constitucional
ECDF – *Empirical Cumulative Distribution Function*
EDF – *Empirical Distribution Function*
FAB – Força Aérea Brasileira
FFAA – Forças Armadas Brasileiras
GOF – *Goodness-of-Fit*
GPS – *Global Positioning System*
IBA – Instituto Brasileiro de Atuária
IBGE - Instituto Brasileiro de Geografia e Estatística
JB – Teste de Jarque-Bara
JCR – *Journal Citation Report*
KS – Teste de Aderência de Kolmogorov-Smirnov
MAD – Teste Modificado de Anderson-Darling
MB – Marinha do Brasil
MD – Ministério da Defesa
PO – Pesquisa Operacional
RGPS – Regime Geral de Previdência Social
RMSD – *Root Mean Square Deviation*
RPPS – Regime Próprio de Previdência Social

SAFA – Software Atuarial das Forças Armadas

SF – Teste de Shapiro-Francia

SOA – *Society of Actuaries*

SPSM – Sistema de Proteção Social dos Militares

SW – Teste de Aderência de Shapiro-Wilk

TI – Tecnologia da Informação

TCU – Tribunal de Contas da União

VLBI – *Very Long Baseline Interferometry*

1. INTRODUÇÃO

1.1. CONTEXTUALIZAÇÃO

A continua discussão da Reforma da Previdência, retomada com mais ênfase a partir de 2017, trouxe à tona o debate sobre os gastos da União com os militares e seus pensionistas. Tecnicamente, a Constituição da República Federativa do Brasil (CRFB) de 1988 não enquadrava aquela categoria em nenhum regime de previdência.

O sistema previdenciário brasileiro é estruturado, basicamente, em três regimes, que são: o Regime Geral da Previdência Social (RGPS); o Regime Próprio da Previdência Social (RPPS), públicos e obrigatórios; e o Regime de Previdência Complementar (RPC), de caráter facultativo, privado e autônomo (LOBATO JUNIOR, 2018).

Cardoso (2007) destaca que a Constituição Federal (CF) não estabeleceu um único conjunto de regras previdenciárias para toda a população. A previdência social dos servidores públicos civis mereceu um tratamento diferenciado por parte do legislador constituinte, de modo a inserir princípios fundamentais à saúde econômico-financeira deste sistema previdenciário imposto ao servidor público.

A Emenda Constitucional (EC) 18/98, seguindo o mandamento constitucional vigente à época, afastou quaisquer possibilidades de enquadramento dos militares das Forças Armadas Brasileiras (FFAA) em algum dos regimes de previdência social. Ainda, a mesma EC define o sistema no qual os militares estão vinculados como sendo um “Regime Constitucional” com características próprias. Este regime chamado de Sistema de Proteção Social dos Militares (SPSM) é um contrato social compensatório fundamental para a existência das FFAA (LEAL *et al.*, 2019).

Ferreira (2017) diferencia o SPSM das FFAA, ou seja, Marinha do Brasil (MB), Exército Brasileiro (EB) e Força Aérea Brasileira (FAB), dos regimes de previdência social. Para aquele autor, o SPSM visa a assegurar dignidade aos militares e dependentes, em razão das peculiaridades da profissão militar. Santos (2018) aduz que o SPSM das FFAA é constituído por um conjunto integrado de instrumentos legais permanentes e interativos, em seus aspectos de remuneração, assistência social e saúde, que visam garantir o amparo e a dignidade aos militares das FFAA e seus dependentes, haja vista as peculiaridades da carreira militar.

As questões que envolvem o SPSM das FFAA não são simples e ultrapassam as esferas sociais e econômicas. Este sistema deve se preocupar em garantir uma relativa

tranquilidade para que os militares possam assumir o risco inerente às suas profissões, sabendo que suas famílias estão amparadas, o que inclui, especialmente uma aposentadoria digna e um sistema de pensões eficaz. Por outro lado, o sistema deve ser equilibrado de modo a fazer bom uso dos recursos públicos. Segundo Conti e Carvalho (2011), na literatura contabilista, uma das principais formas de controle na atividade pública é o controle interno. Assim, a estruturação do sistema de controle definirá a qualidade de seu gasto público e possui função essencial para uma gestão fiscal responsável no Brasil.

Visando o aumento da governança pública, o Tribunal de Contas da União (TCU) instou o Ministério da Defesa (MD), a partir de 2010, a realizar anualmente uma projeção dos custos atuariais com os pensionistas das FFAA, considerando um horizonte temporal de 75 anos. Tal tarefa mostrou-se um grande desafio organizacional, devido ao grande volume de dados, esforço computacional exigido pelos referidos cálculos matemáticos e incertezas envolvidas (GOMES, SANTOS e MARTINS, 2017).

Gomes *et al.* (2017) destacam que projeções atuariais de receitas e despesas são importantes para garantir o equilíbrio atuarial e diminuir o risco de falta de liquidez, possibilitando que os órgãos públicos provisionem recursos para os anos subsequentes. O objetivo destas projeções para uma entidade de previdência e pensões, portanto, seria quantificar os custos futuros estimados com pagamento de benefícios e as receitas futuras estimadas de contribuições de participantes.

1.2. SITUAÇÃO PROBLEMA

Com o intuito de gerar informações cada vez mais confiáveis em seu benefício, é crescente a necessidade das instituições se organizarem e tratarem seus dados. Neste contexto, a Engenharia de Produção é uma ferramenta fundamental no assessoramento dos gestores na tomada de decisão correta (SANTOS et al., 2019). Dentro da área da Engenharia de Produção, por sua vez, a Pesquisa Operacional (PO) é uma subárea que envolve conceitos multidisciplinares para a resolução de questões complexas.

Segundo Santos (2018), os problemas de Pesquisa Operacional possuem metodologia razoavelmente simples e que respeita um encadeamento lógico, iniciando com uma situação problemática, isto é, uma assimetria entre o que determinada pessoa, grupo ou organização deseja e o que efetivamente obtém. Desta forma, só é possível obter a solução de um problema, a partir do perfeito entendimento do mesmo e das suas condições de contorno.

No Brasil, os militares das Forças Armadas não estão submetidos às mesmas regras dos regimes de previdência dos cidadãos civis, possuindo um sistema próprio. De acordo com Gomes, Santos e Martins (2017), naqueles países em que se preza pela solvência das entidades de seguridade social, os sistemas de previdência e pensões necessitam realizar a projeção atuarial de despesas e receitas. Esta prática consiste em mensurar os pagamentos e contribuições futuros de participantes ativos, inativos e pensionistas. Assim sendo, a gestão do Sistema de Proteção Social dos Militares exige a realização de cálculos atuariais.

Desta forma, faz-se necessária a determinação das tábuas atuariais de mortalidade e invalidez que melhor representam as expectativas de decrementos, ou seja, morte e invalidez da população estudada. Analisando-se a situação de um indivíduo ativo deste grupo, verifica-se que o mesmo está sujeito a múltiplos decrementos probabilísticos, chamados também de “multidecrementos”; isto é, probabilidades conjuntas de morrer ou se invalidar (SANTOS, 2018).

Para Gomes, Santos e Martins (2017), uma forma de avaliar se duas distribuições de frequências são consideradas heterogêneas ou aproximadamente idênticas é por meio da realização dos testes de aderência. Dentro deste contexto, nos estudos atuariais do SPSM, faz-se necessária a utilização de testes de aderência para a seleção de tábuas atuariais. Em Santos (2018), os testes de aderência consistiram em analisar se a distribuição de decrementos (falecimentos e invalidez) dados pelas tábuas atuariais analisadas representavam ou não a distribuição de decrementos das bases de dados das FFAA. Naquele trabalho, dentre os diversos tipos de teste de aderência existentes na academia, Santos utilizou como teste de aderência apenas o teste do Qui-quadrado, pela sua predominância no mercado previdenciário.

1.3. PERGUNTAS DA PESQUISA

Com isso, encontra-se uma problemática motivante para este trabalho: na academia existem diferentes modelos de testes de aderência para verificar a normalidade na distribuição de frequências. Embora o teste de aderência do Qui-quadrado seja o mais empregado no mercado previdenciário, não necessariamente seja sempre o mais adequado. Neste contexto, surge a seguinte questão central da pesquisa: **dentre os principais testes de aderência difundidos no meio acadêmico, qual o teste de aderência que melhor se aplica para a seleção das tábuas atuariais para os militares das Forças Armadas Brasileiras?** Ainda neste trabalho, outras questões secundárias serão respondidas, a saber:

- (a) Dentre os principais testes de aderência, quais os métodos de cálculo da estatística de Teste?
- (b) Quais os parâmetros que melhor formam as dimensões a serem consideradas por ocasião de um teste de aderência?
- (c) Qual teste de aderência se mostrou mais adequado para o cálculo de Mortalidade de Ativos Inativos e Pensionistas das Forças Armadas?
- (d) Qual teste de aderência se mostrou mais adequado para o cálculo de Mortalidade de Inválidos das Forças Armadas?
- (e) Qual teste de aderência se mostrou mais adequado para o cálculo de Entrada em Invalidez das Forças Armadas?
- (f) As tábuas atuariais existentes hoje no mercado previdenciário podem ser empregadas para a população das Forças Armadas?
- (g) Quais tábuas atuariais apresentam melhor aderência para a amostra populacional das forças armadas para Mortalidade de Ativos Inativos e Pensionistas?
- (h) Quais tábuas atuariais apresentam melhor aderência para a amostra populacional das forças armadas para Mortalidade de Inválidos?
- (i) Quais tábuas atuariais apresentam melhor aderência para a amostra populacional das forças armadas para Entrada em Invalidez?
- (j) Os decrementos probabilísticos a que um indivíduo ativo está sujeito tendem a algum estado de convergência?
- (k) Caso haja um estado de convergência, qual a probabilidade de estado estável para os estados biométricos e sociais nos quais um indivíduo ativo pode se encontrar?

1.4. OBJETIVO GERAL E OBJETIVOS SECUNDÁRIOS

Na busca de uma resposta para a questão central abordada anteriormente, este estudo tem como objetivo geral identificar o teste de aderência mais adequado para seleção das tábuas atuariais referentes aos militares das Forças Armadas Brasileiras. Para complementar o atendimento do objetivo geral, foram estabelecidos como objetivos secundários:

- (a) Identificar, dentre os principais testes de aderência, quais os métodos de cálculo da estatística de Teste;
- (b) Identificar quais os parâmetros que melhor formam as dimensões a serem consideradas por ocasião de um teste de aderência;

- (c) Identificar o teste de aderência que se mostra mais adequado para o cálculo de Mortalidade de Ativos Inativos e Pensionistas das Forças Armadas;
- (d) Identificar o teste de aderência que se mostra mais adequado para o cálculo de Mortalidade de Inválidos das Forças Armadas;
- (e) Identificar o teste de aderência que se mostra mais adequado para o cálculo de Entrada em Invalidez das Forças Armadas;
- (f) Verificar se tábuas atuariais existentes hoje no mercado previdenciário podem ser empregadas para a população das Forças Armadas.
- (g) Identificar a tábua(s) atuarial (is) que apresenta(m) melhor aderência para a amostra populacional das forças armadas para Mortalidade de Ativos Inativos e Pensionistas;
- (h) Identificar a tábua(s) atuarial (is) que apresenta(m) melhor aderência para a amostra populacional das forças armadas para Mortalidade de Inválidos;
- (i) Identificar a tábua(s) atuarial (is) que apresenta(m) melhor aderência para a amostra populacional das forças armadas para Entrada em Invalidez;
- (j) Verificar se os decrementos probabilísticos a que um indivíduo ativo está sujeito tendem a algum estado de convergência; e
- (k) Calcular, se houver, qual a probabilidade de estado estável para os estados biométricos e sociais em que um indivíduo ativo pode se encontrar.

1.5. JUSTIFICATIVA E RELEVÂNCIA DA PESQUISA

O teste de aderência Qui-quadrado já está consagrado pelo uso na literatura acadêmica. Assim, a pesquisa pretende verificar se tal método, em relação aos demais, é o que melhor atende às especificidades das FFAA. Caso este teste seja identificado como o mais adequado, confirmar-se-á que o seu emprego está sendo coerente; porém, caso contrário, abrirá um leque para que outros testes de aderência venham a ser usados em outros estudos que utilizaram o Qui-quadrado.

Pretende-se verificar a aplicabilidade das Cadeias de Markov para determinar, a partir do modelo de estados biométricos e sociais dos militares das FFAA, se os decrementos probabilísticos convergem para um estado funcional de estabilidade.

Espera-se, por meio desta pesquisa, portanto, verificar a robustez do modelo proposto por Santos (2018), a fim de contribuir com o Ministério da Defesa e com a União; assim, permitir que possa ser disponibilizado um conjunto de informações relativas ao custo

constitucional das pensões dos militares para o correto provisionamento de recursos financeiros a médio e longo prazos, indo ao encontro do preconizado pelo TCU no que diz respeito à manutenção e melhoria da governança pública.

Destaca-se, ainda, que se trata uma pesquisa aplicada a um problema real, referente às FFAA, que afeta diretamente mais de um milhão de cidadãos brasileiros.

1.6. DELIMITAÇÃO DA PESQUISA

O presente trabalho é uma continuação da pesquisa de Santos (2018), a partir de sua tese “Proposta de Modelagem Atuarial aplicada ao setor militar, considerando influências econômicas e biométricas”, que restringiu-se às pensões militares definidas pela Lei nº 3.765, de 4 de maio de 1960; portanto, sem incluir as pensões indenizatórias ou pensões especiais do grupo de ex-combatentes e anistiados políticos. Naquele trabalho, Santos desenvolveu uma metodologia para um rápido cálculo das projeções atuariais dos custos com pensionistas das FFAA, a partir de tábuas de mortalidade selecionadas pelo teste do Qui-quadrado. Com o intuito de ampliar o escopo daquele trabalho, Santos identificou as seguintes possibilidades de pesquisas futuras:

- (1) Ampliar a quantidade de tábuas atuariais analisadas;
- (2) Realizar outros testes de aderência, diferentes do teste Qui-quadrado, para a seleção das tábuas atuariais;
- (3) Aplicar Cadeias de Markov a fim de verificar a convergência dos decrementos probabilísticos; e
- (4) Avaliar a possibilidade do desenvolvimento de tábuas dinâmicas, com o propósito de antecipar os acréscimos de longevidade que se tem observado ao longo do tempo.

Esta dissertação se propôs estender a tese de doutorado de Santos (2018), desenvolvendo uma pesquisa relacionada aos itens (1), (2) e (3) supracitados.

No âmbito da revisão da literatura serão pesquisados os testes de aderência relatados em artigos indexados nas bases SCOPUS e ISI *Web of Science*, acessados a partir do portal de periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

No presente trabalho, focou-se nos testes de aderência para verificação de ajuste à distribuição normal ou para calcular a probabilidade de uma variável aleatória estar normalmente distribuída. A pesquisa se propõe a identificar o teste de aderência mais adequado entre os citados por Santos (2018) como principais, a saber: Qui-quadrado, Desvio

Quadrático Médio, Kolmogorov-Smirnov e Binomial, além dos testes de Shapiro-Wilk e Lilliefors.

No presente trabalho, não serão realizadas projeções atuarias, focando-se em expandir a premissa assumida por Santos (2018) de utilização somente do teste de aderência do Qui-quadrado para a seleção de tábuas atuarias e na análise dos estados biométricos e sociais propostos por aquele autor.

No estudo markoviano dos Estados Biométricos e Sociais a que um militar possa estar submetido, foi empregado como base de dados o Anuário Estatístico da Marinha (ANEMAR), volume 1, de 2018. Utilizou-se somente de dados públicos, não havendo a disponibilidade de dados consolidados de probabilidade de assunção dos diferentes estados. Portanto, as probabilidades foram estimadas mediante a razão entre o número de militares entre diferentes estados biométricos e sociais, conforme descrito no Capítulo 5. Para aplicação da Cadeia de Markov, foi adotado um modelo simplificado de grafo dos estados biométricos e sociais.

1.7. ESTRUTURA DO TRABALHO

Esta dissertação está estruturada em seis capítulos.

O primeiro capítulo apresenta uma breve introdução. O segundo capítulo revela o referencial teórico que fundamenta este estudo a partir de um estudo bibliométrico dos testes de aderência. O terceiro capítulo descreve a metodologia utilizada para obtenção das respostas aos questionamentos da pesquisa. O quarto capítulo apresenta a implementação dos testes de aderência e analisa os resultados obtidos pelos modelos descritos, relacionando os resultados com a questão central da pesquisa e algumas das perguntas secundárias. O quinto capítulo apresenta um estudo markoviano dos estados biométricos e sociais, fazendo uma correspondência entre os resultados alcançados e os demais questionamentos secundários da pesquisa. O sexto capítulo apresenta as conclusões do trabalho e as recomendações para estudos posteriores.

2. REVISÃO DA LITERATURA

O termo Pesquisa Operacional deriva das expressões em inglês *Operational Research* (empregada no Reino Unido) ou *Operations Research* (empregada nos Estados Unidos da América).

O início da atividade moderna que recebeu a denominação de PO remete às ações militares dos comandos britânicos e americanos durante a Segunda Guerra Mundial (1939-1945). Frente a escassez de recursos e a conseqüente necessidade de melhor aloca-los, cientistas foram chamados a realizar pesquisas sobre operações militares a fim de proporem melhores soluções para problemas estratégicos e táticos (HILLIER E LIEBERMAN, 2013).

Sir Robert Watson-Watt, um dos cientistas quem reivindicou a autoria do termo *Operational Research*, a definiu como a atividade de examinar quantitativamente se uma organização está obtendo com a operação de seus equipamentos a melhor contribuição possível para o seu objetivo geral; quais são os fatores predominantes afetando os resultados obtidos; quais mudanças em equipamentos ou métodos podem gerar razoável expectativa de melhoria destes resultados ao menor custo de esforço e tempo e o grau em que as variações nos objetivos táticos são susceptíveis a contribuir para uma forma mais econômica e expedita de alcançar o objetivo estratégico global (NATURE, 1947).

Para Philip McCord Morse, considerado nos Estados Unidos como “*the father of Operations Research*” a pesquisa operacional é, em uma definição mais simples, um método científico de prover aos departamentos executivos uma base quantitativa para as decisões relativas às operações sob seu controle (MORSE e KIMBALL, 1951; LITTLE, 2003).

Harper (1975) considera que no período pós-guerra, o desenvolvimento da PO se intensificou pois as empresas perceberam que poderiam melhorar sua performance por meio de um estudo mais cuidadoso de como geriam suas operações. Para aquele pesquisador, a PO seria o conjunto de técnicas que provê as respostas aos questionamentos das empresas em relação aos seus procedimentos: ou seja, segundo a qual os gerentes conseguem obter as informações para verificar as conseqüências dos métodos de operação existente e alternativos. Para Zachariah *et al.* (2009), uma definição de PO mais moderna e que estende seu uso para áreas das ciências humanas é que se trata da pesquisa visando conhecimento sobre intervenções, estratégias ou ferramentas que melhorarem a qualidade, a eficácia ou a cobertura de programas em que a pesquisa está sendo realizada. Kruger *et al.* (2019) ressaltam a importância da modelagem matemática para a PO, porém defendem agregar as contribuições conceituais e práticas da teoria da complexidade. Para estes autores, novas interações prospectivas e transdisciplinares para a PO, de forma mais sustentável e ética, poderiam ser exploradas.

Drobne, Dumičić e Stirn (2018) ressaltam que a PO teve um impacto importante na melhoria da eficiência das organizações e contribuiu para aumentar a produtividade e o bem-estar social. Sua aplicação é ampla, envolvendo problemas reais em diversos campos

distintos. O caráter abrangente da PO está intimamente ligado à sua interdisciplinaridade. Para Corbett e Wassenhove (1993), o mote interdisciplinar da PO se justifica em suas próprias origens, pois quando começou a ser utilizada durante a guerra, não havia profissionais habilitados neste novo campo da ciência. Assim, foi necessária a combinação de pesquisadores que dominavam diferentes áreas do conhecimento. Hillier e Lieberman (2013), por exemplo, afirmam que um estudo de PO pode requerer conhecimentos de matemática, estatística e teoria da probabilidade, administração, economia, informática, engenharia, física e ciências comportamentais, além de experiência e habilidades específicas relacionadas ao problema analisado. Santos *et al.* (2015) concordam que a PO apresenta uma alta multidisciplinaridade, geralmente fazendo uso de modelos matemáticos e / ou lógicos. Estes autores aduzem que uma característica da maioria das técnicas de PO é que as soluções obtidas não traduzem meramente a aplicação de uma fórmula, mas são encontradas apenas por meio de alguns algoritmos.

No presente trabalho, dentro do contexto de multidisciplinariedade da PO, serão utilizadas técnicas matemáticas, estatísticas e de teoria da probabilidade. Este capítulo descreve os principais conceitos relacionados ao estudo de aderência, apresentando a revisão da literatura desenvolvida a partir de pesquisa bibliométrica, considerando o modelo descrito por Costa (2010). Este capítulo está dividido em duas seções: Base Conceitual e Mapeamento do Estado da Arte, conforme ilustra a Figura 1.

Figura 1 – Organização do Capítulo 2 – Revisão da Literatura

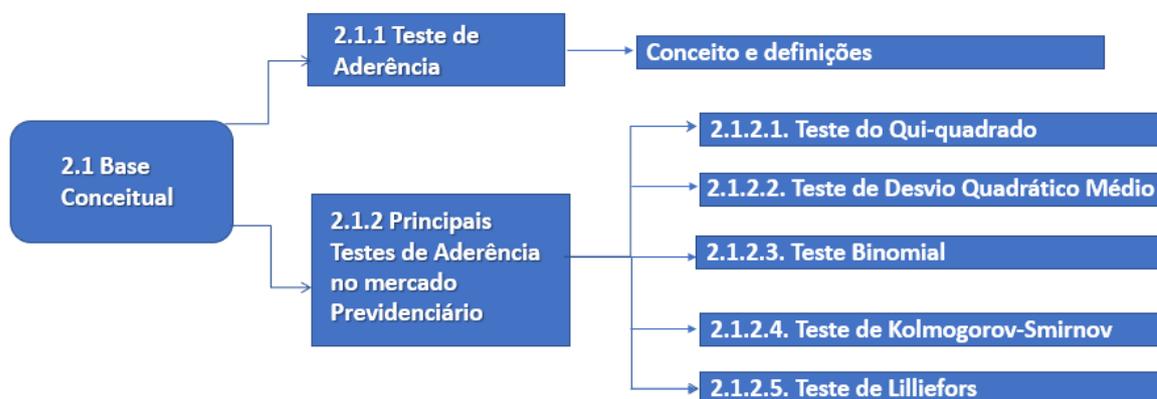


Fonte: Elaboração Própria (2020)

2.1. Base conceitual

Nesta seção são abordados os conceitos e fundamentos dos aspectos abrangidos pelo tema central, criando o alicerce teórico para o posterior desenvolvimento da pesquisa. Desta sorte, a base conceitual foi estruturada como ilustrado na Figura 2.

Figura 2 – Organização do item 2.1 Base Conceitual



Fonte: Elaboração Própria (2020)

2.1.1. Teste de Aderência (*Goodness-of-Fit* – GOF – Test)

A validação de um modelo é um passo necessário para a sua aceitação. Nenhuma combinação única de testes de validação será aplicável em diversos modelos e seus usos. A escolha da técnica é importante, pois alguns contêm problemas e inconsistências. A eficiência da modelagem é proposta como a melhor medida global de concordância entre os valores observados e simulados (MAYER e BUTLER, 1993).

Para D’Agostino e Stephens (1986), testes de aderência são métodos estatísticos que permitem examinar o quão bem uma amostra de dados concorda com uma dada distribuição como sua população. O objetivo da realização dos testes de aderência, portanto é avaliar se duas distribuições de frequências são aproximadamente idênticas ou se são consideradas heterogêneas.

Em que pese não haja uma única definição padronizada internacionalmente para os Testes de Aderência, Assis, Souza e Dias (2019), em seu Glossário de Estatística, apresentam um conceito geral sobre o termo em inglês *Goodness-of-Fit*:

Teste estatístico de modelo de distribuição de probabilidades, no qual as proporções observadas se ajustam às proporções esperadas, deduzidas matematicamente ou estabelecidas de acordo com alguma teoria. É também denominado de teste de aderência [ASSIS, SOUZA e DIAS, 2019].

Os testes de aderência (também chamados de testes de qualidade de ajuste ou testes de bondade de ajuste, em tradução livre do termo em inglês *Goodness-of-Fit*), portanto, consistem em verificar a adequabilidade de um modelo probabilístico a um conjunto de dados. Nos testes de aderência, tem-se uma hipótese nula H_0 que X , uma variável aleatória,

segue uma lei de probabilidade declarada $F(x)$. As técnicas destes testes consistem em modelos matemáticos para medir a conformidade dos dados de uma amostra, isto é, conjunto de valores de x com a distribuição hipotética; ou, de forma equivalente, com sua discrepância em relação a ela. Ou seja, o conceito básico é que, dada uma amostra aleatória de tamanho n , observada de uma variável aleatória X , deseja-se testar a hipótese nula H_0 que a amostra segue uma determinada função de distribuição $F(x)$, confrontando-a com a hipótese alternativa H_1 de que aquela amostra não siga a função de distribuição $F(x)$:

$$H_0: X \text{ tem distribuição } F(x) \quad \text{vs} \quad H_1: X \text{ não tem distribuição } F(x).$$

Na estrutura formal do teste de aderência, a hipótese nula H_0 pode ser uma hipótese simples, quando $F(x)$ é especificado completamente ou H_0 pode dar uma especificação incompleta e então será uma hipótese composta.

Para McAssey (2013), é necessário que um Teste de Aderência seja teoricamente sólido, simples de implementar em aplicações científicas, adaptável a qualquer distribuição de qualquer dimensão e ter robustez suficiente. Santos (2018) aponta como principais testes de aderência presentes na academia os testes: Qui-quadrado, Desvio Quadrático Médio, Kolmogorov-Smirnov e Binomial. Nesta pesquisa, foram incluídos na revisão da literatura, ainda, o teste de Lilliefors, que é uma evolução do teste de Kolmogorov-Smirnov e o teste de Shapiro-Wilk, que possui grande relevância acadêmica.

De acordo com Maydeu-Olivares e Forero (2010), a avaliação de um teste de aderência envolve necessariamente um julgamento subjetivo. Esses autores afirmam que os modelos são apenas aproximações de fenômenos da vida real e, conseqüentemente, qualquer modelo será rejeitado se a amostra tiver tamanho suficientemente grande. Contudo, aqueles autores alertam que isso não deve ser levado a implicar que o teste de aderência não tenha sentido. Pelo contrário, eles opinam que os pesquisadores devem analisar a aderência de seus modelos usando um teste estatístico, para avaliar a magnitude da discrepância entre os dados e os modelos, considerando a variabilidade da amostragem.

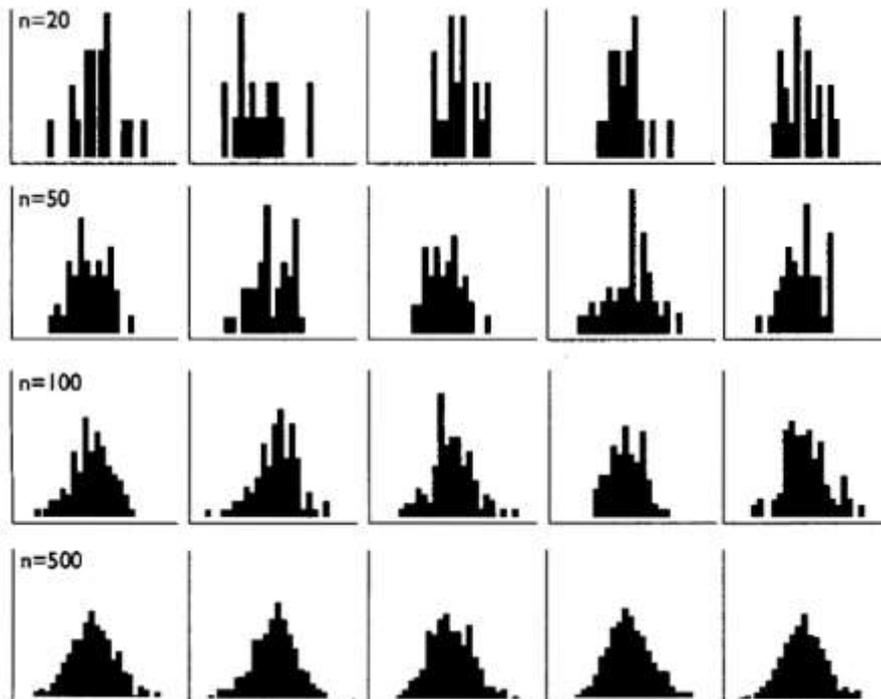
Nos estudos de estatística, comumente utiliza-se um índice de significância marginal arbitrado, chamado de *p-value* (ou valor-p), que representa a probabilidade da ocorrência de um dado evento. O valor-p é usado como uma alternativa aos pontos de rejeição para fornecer o menor nível de significância no qual a hipótese nula seria rejeitada (VICKERS, 2007; DALLAL, 2012). O *p-value* é, portanto, a probabilidade de se obter um efeito igual ou maior que o observado considerando que a hipótese nula seja verdadeira (GLANTZ, 2002; BIAU *et al.*, 2015).

Em outras palavras, o valor-p é o limite assumido pelo pesquisador de que a diferença entre a amostra esperada e observada invalide a hipótese admitida no estudo científico como sendo verdadeira, ainda que, estatisticamente, demonstrada.

Segundo Dallal (2012), o nível padrão de significância normalmente utilizado para justificar uma reivindicação de um efeito estatisticamente significativo é de 0,05, isto é, $p\text{-value} \leq 0,05$. Desta forma, no presente trabalho também será assumido este valor-p como referência.

Quando o tamanho de uma amostra é relativamente grande, pode-se aplicar o teorema do limite central e inferir a normalidade de sua distribuição; isto é, conforme o tamanho da amostra aumenta, a distribuição amostral da sua média aproxima-se cada vez mais de uma distribuição normal (ALTMAN e BLAND, 1995). A Figura 3 ilustra distribuições de 5 amostras de tamanhos diferentes tiradas aleatoriamente de distribuições normais, com tamanhos crescentes $n=20$, $n=50$, $n=100$ e $n=500$. Evidencia-se que poucas das amostras pequenas parecem uma distribuição normal, mas a similaridade aumenta à medida que o tamanho da amostra cresce. No presente trabalho, conforme item 1.6, limitou-se ao estudo dos principais testes de aderência para verificação de ajuste à distribuição normal ou para calcular a probabilidade de uma variável aleatória estar normalmente distribuída.

Figura 3 – Distribuição aleatória de amostras com similaridade crescente à curva normal



Fonte: Altman e Bland (1995)

2.1.2. Principais Testes de Aderência

2.1.2.1. Teste do Qui-quadrado

O objetivo do método do teste do Qui-quadrado (ou Chi-quadrado) é comparar as divergências entre as frequências esperadas (E) e as observadas (O). Ao ser aplicada para testes de aderência, portanto, a estatística chi-quadrado (χ^2) consiste em comparar os desvios entre o número de eventos observados na população estudada e aqueles estimados de acordo com determinado padrão teórico, os quais são avaliados quanto à sua significância, mediante os valores tabelados da distribuição chi-quadrado (GOMES, SANTOS e MARTINS, 2017).

Para testar se as discrepâncias calculadas possuem significância estatística, calcula-se o índice χ^2 e compara-se com o mesmo fator (χ^2 crítico) obtido da tabela de distribuição chi-quadrado (SANTOS, 2018).

Segundo Ross (2014), o índice χ^2 é calculado pela Equação 1:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (\text{Equação 1})$$

Onde,

χ^2 = Estatística do Teste Qui-quadrado

O = Frequência Observada

E = Frequência Esperada

O resultado de χ^2 exprime a divergência entre a frequência observada e a frequência esperada e, quanto menor este valor, maior a probabilidade de não se rejeitar a hipótese de aderência. Uma vez calculada a variável χ^2 , compara-se com $\chi^2_{\text{crítico}}$, considerando-se o nível de significância arbitrado e os graus de liberdade considerados no teste (BARCELÓ, 2018).

Se o número esperado de observações em qualquer categoria for muito pequeno, isto é, tamanho total da amostra for menor que 1000, o teste do Qui-quadrado pode fornecer resultados imprecisos (MCDONALD, 2014).

Para fins desta pesquisa, considera-se que cada faixa de idades de indivíduos expostos ao risco é representada como uma observação independente da amostra. Desta forma, o número de graus de liberdade do teste Qui-quadrado é representado pelo número de faixas etárias analisadas, subtraídas de um, devido à característica intrínseca do modelo de teste utilizado. Assim sendo, entende-se que a distribuição de decrementos gerada por uma tábua

atuarial determinada será idêntica à distribuição de decrementos da massa populacional quando a estatística Qui-Quadrado calculada for inferior à estatística Qui-Quadrado tabelada, baseada no nível de significância e nos graus de liberdade do teste.

2.1.2.2. Teste de Desvio Quadrático Médio (*Root Mean Square Deviation – RMSD*)

Segundo Armstrong e Collopy (1992), dada uma série de medidas de uma grandeza, as diferenças entre os valores medidos e o valor mais provável, \bar{x} , são chamados desvios ou resíduos. Uma relação que poderia ser utilizada para a medida da dispersão poderia ser o desvio quadrático médio, definida pela Equação 2:

$$S_{rms} = \sqrt{\frac{\sum_{i=1}^n (\delta x_i)^2}{n}} \quad (\text{Equação 2})$$

Onde,

- S_{rms} é o Desvio Quadrático Médio na ocorrência do evento
- δx_i é a diferença entre as probabilidades esperadas e observadas do evento

O Desvio Quadrático Médio é sempre não negativo e um valor de 0 indicaria um ajuste perfeito aos dados. Esta medida indica simplesmente como um conjunto de n valores desvia-se de sua média. É uma medida frequentemente usada das diferenças entre valores amostrais (ou populacionais) previstos por um modelo e os valores observados. Representa a raiz quadrada da média dos erros quadrados, isto é, a média quadrática das diferenças entre os valores previstos e os valores observados. É uma medida de precisão para comparar erros de previsão de modelos diferentes para um conjunto de dados específico e não entre conjuntos de dados, pois é dependente da escala (WILLMOTT *et al.*, 2006 ; HYNDMAN e KOEHLER, 2006).

O teste do Desvio Quadrático Médio (DQM) ou Desvio Quadrado Médio visa analisar o quão discrepantes são duas distribuições específicas. Entretanto, distingue-se de outros testes de aderência pois não indica isoladamente se uma determinada tábua atuarial é aderente à massa de dados populacionais. Por este método são calculadas as estatísticas DQM de diversas tábuas, indicando que aquela que possui o menor valor DQM é a que possuiu

melhor aderência. Aquela tábua atuarial com o segundo menor valor da estatística DQM seria a segunda tábua com melhor aderência e assim sucessivamente.

Para emprego deste teste, faz-se necessário que, a partir da base de dados histórica da população, sejam calculados os desvios entre os decrementos observados e os decrementos esperados. Estes, por sua vez, são calculados mediante a multiplicação da quantidade de indivíduos expostos ao risco em determinada faixa etária pela probabilidade da tábua atuarial testada. Considera-se, então, que o teste do Desvio Quadrático Médio fornece uma ordenação das tábuas testadas, das menos discrepantes aos dados observados da população analisada para aquelas que mais se diferem da amostra. Na prática, este teste é especialmente importante quando não é encontrada aderência estatística a nenhuma tábua atuarial a partir dos métodos Qui-Quadrado e/ou Kolmogorov-Smirnov.

2.1.2.3. Teste Binominal

O teste binomial é um teste de aderência que deve ser utilizado para comparar a distribuição observada com a distribuição esperada quando há apenas duas categorias, situação na qual o seria mais preciso do que o Qui-quadrado (HOWELL, 2007).

Segundo McDonald (2014), o teste binomial é empregado quando há uma variável nominal com apenas dois valores (como masculino ou feminino, esquerdo ou direito, verde ou amarelo), comparando os dados observados com os dados esperados, que são algum tipo de expectativa teórica, não sendo comumente utilizado se o número de observações for muito alto.

O valor estatístico do teste binomial, $P(B)$, está indicado na Equação 3:

$$P_k (B=k) = \binom{n}{k} \times p^k \times (1-p)^{n-k} \quad (\text{Equação 3})$$

Onde,

- n é o tamanho da amostra (número de tentativas);
- k é o número de sucessos;
- p é a probabilidade de sucesso de um único estudo ou a proporção de população (hipótese); e
- $\binom{n}{k}$ é uma equação de números fatoriais, que também pode ser descrita por $\frac{n!}{k! (n-k)!}$.

2.1.2.4. Teste de Kolmogorov-Smirnov

Segundo Young (1977), o teste de aderência de Kolmogorov-Smirnov (KS) é um teste estatístico não-paramétrico realizado em dados experimentais com dois propósitos distintos: testar se um histograma condiz com uma distribuição de referência; ou testar se dois histogramas são amostras da mesma função de densidade de probabilidade contínua. Para fins deste estudo, foca-se no primeiro caso, em que o teste de aderência de KS se baseia na suposição paramétrica de que o erro de amostragem tem uma distribuição normal.

A estatística de teste de Kolmogorov-Smirnov (KOLMOGOROV, 1933) é calculada a partir do valor supremo, isto é, o máximo dos valores em um determinado intervalo (Equações 4 e 5).

$$D_n = \sup_x |F_n(x) - F_0(x)| \quad (\text{Equação 4})$$

Onde $F_n(x)$ e $F_0(x)$ representam as funções de distribuição acumulada dos óbitos esperados e observados, respectivamente. A estatística do teste D_n será a maior distância vertical entre $F_n(x)$ e $F_0(x)$, isto é:

$$D_n = \max \{ |D_n^-|, |D_n^+| \}$$

$$\text{Onde } D_n^- = F(x_{k-1}) - F_0(x_k) \text{ e } D_n^+ = F(x_k) - F_0(x_k) \quad (\text{Equação 5})$$

A estatística KS_z pode ser dada por: $KS_z = \sqrt{n}D_n$ (Equação 6)

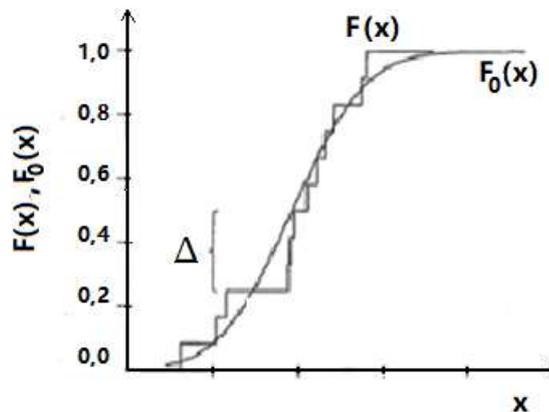
Para Maroco (2010), testar a normalidade por meio do KS consiste em mensurar a distância entre as funções distribuições empíricas para as duas amostras, que são padronizadas e comparadas com uma distribuição normal. Isto equivale a tornar a média e a variância da distribuição de referência iguais aos estimados das amostras (STEPHENS, 1974). Isto é, sejam os parâmetros média da população “ μ ” e desvio padrão da população “ δ ” quaisquer, pretende-se testar a hipótese nula H_0 que a amostra “X” observa uma determinada distribuição “N” contra a hipótese alternativa H_1 de que a amostra não observe aquela distribuição (Equação 7):

$$H_0: X \sim N(\mu, \delta) \text{ vs } H_1: X \neq N(\mu, \delta) \quad (\text{Equação 7})$$

Monish e Rehana (2020) reiteram que o teste KS mede a diferença máxima entre a distribuição cumulativa teórica e distribuição cumulativa empírica de pontos das amostras. A estatística do teste KS será baseada na maior diferença vertical (Δ) entre as funções de distribuição cumulativa (*cumulative distribution function* - CDF) teóricos e empíricos da

amostra testada. Segundo Steel e Torrie (1980), inicia-se o teste ordenando as observações da variável X por ordem crescente e, em seguida, calculando-se a frequência acumulada de cada observação. Deve-se, então, calcular a diferença entre este resultado e a frequência acumulada esperada para a mesma observação, caso tivesse uma distribuição de probabilidade normal, bem como a mesma diferença relativamente à observação anterior. A estatística do teste é então dada pela maior diferença destas duas diferenças, cuja interpretação gráfica é apresentada na Figura 4.

Figura 4 – Interpretação gráfica da estatística do teste Kolmogorov-Smirnov



Fonte: Maroco (2010)

Compara-se o valor crítico da distribuição estatística de KS com o tabelado e rejeita-se H_0 se $D > D_{\text{tabelado}}(\alpha)$. A maioria dos softwares estatísticos calculam o p -value, isto é, o menor valor de “ α ” a partir do qual $D > D_{\text{tabelado}}$. Assim para uma determinada probabilidade de erro de tipo I $\alpha \times 100\%$ rejeita-se H_0 se $p < \alpha$ (SMIRNOV, 1948; MAROCO, 2010).

2.1.2.5. Teste de Lilliefors

No teste de Kolmogorov-Smirnov (KS), assume-se que os parâmetros populacionais “ μ ” e “ δ ” são conhecidos, o que nem sempre é possível. Na prática, em grande parte dos casos, obtém-se uma amostra representativa da população a partir da qual pode-se inferir sobre os verdadeiros valores de “ μ ” e “ δ ”. Assim, a estatística D não poderia ser aplicada a partir apenas estimativas amostrais, em vez dos verdadeiros valores de “ μ ” e “ δ ” (MAROCO, 2010). Ou seja, as tabelas padrão usadas para o teste de KS são válidas ao testar se um conjunto de observações é de uma distribuição contínua completamente especificada. Se um ou mais parâmetros precisarem ser estimados a partir da amostra, as tabelas não serão mais válidas (LILLIEFORS, 1967).

Lilliefors (1967) propôs uma correção às tabelas com os valores críticos da distribuição de KS (Correção de Lilliefors) que deve ser utilizada quando se compara a distribuição de frequências acumuladas e das observações da variável com uma distribuição teórica cujos parâmetros foram estimados a partir da amostra. O *p-value* pode ser calculado usando-se a aproximação analítica da estatística de teste de Lilliefors proposta por Dallal e Wilkinson (1986), rejeitando-se H_0 se $p\text{-value} < \alpha$. (LILLIEFORS, 1967; CONOVER, 1999; MAROCO, 2010).

Dallal e Wilkinson (1986) propõem que se encontre os valores críticos tabelados correspondentes a D_{20} ($\alpha = 0,20$) e D_{15} ($\alpha = 0,15$), referentes a “n”, utilizando-se de interpolação linear quando necessário. Então, deve-se encontrar o valor de a_1 , b_1 , c_1 , a_2 , b_2 e c_2 , conforme descrito a seguir nas Equações 8, 9, 10 e 11:

$$a_1 = -7,01256(n + 2,78019)$$

$$b_1 = 2,99587 \sqrt{n + 2,78019}$$

$$c_1 = 2,1804661 + 0,974598 / \sqrt{n} + 1,67997/n$$

$$a_2 = -7,90289126054 n^{0,98}$$

$$b_2 = 3,180370175721 n^{0,49}$$

$$c_2 = 2,2947256 \quad (\text{Equação 8})$$

Em seguida, deve-se calcular o valor crítico D_{10} ($\alpha = 0,15$)

$$\text{Se } n \leq 100, \quad a = a_1, \quad b=b_1, \quad c=c_1, \quad \text{tem-se que: } D_{10} = \frac{-b_1 - \sqrt{(b_1)^2 - 4a_1c_1}}{2a_1} \quad (\text{Equação 9})$$

$$\text{Se } n > 100, \quad a = a_2, \quad b=b_2, \quad c=c_2, \quad \text{tem-se que: } D_{10} = \frac{-b_2 - \sqrt{(b_2)^2 - 4a_2c_2}}{2a_2} \quad (\text{Equação 10})$$

Finalmente, calcula-se o *p-value* de acordo com a seguinte fórmula:

$$\begin{array}{ll}
 0,10 & \text{se } D_n = D_{10} \\
 \exp (aD_n^2 + bD_n + c - 2,3025851) & \text{se } D_n > D_{10} \\
 p\text{-value} = & 0,15 + (D_n - D_{15}) [(0,10 - 0,15) / (D_{10} - D_{15})] \quad \text{se } D_n \geq D_{15} \\
 & 0,20 + (D_n - D_{10}) [(0,15 - 0,20) / (D_{15} - D_{20})] \quad \text{se } D_n \geq D_{20} \\
 & p\text{-value} > 0,20 \quad \text{se } D_n < D_{10} \quad (\text{Equação 11})
 \end{array}$$

2.1.2.6. Teste de Shapiro-Wilk

O teste de Shapiro-Wilk é um teste de aderência difundido no meio acadêmico que não utiliza procedimentos gráficos para testar a adequação à distribuição normal, podendo ser obtido facilmente nos principais pacotes estatísticos, como R, SAS e SPSS.

Dada uma amostra aleatória ordenada, $y_1 < y_2 < \dots < y_n$, a estatística original do teste Shapiro-Wilk (SHAPIRO e WILK, 1965) é definido como

$$W = \frac{(\sum_{i=1}^n a_i y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (\text{Equação 12})$$

Onde,

- y_i é a estatística de i -ésima ordem;
- \bar{y} é a média da amostra,
- $a_i = (a_1, a_2, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}$
- $m = (m_1, m_2, \dots, m_n)$ são os valores esperados das estatísticas de variáveis aleatórias independentes e distribuídas de forma idêntica de amostras a partir da distribuição normal padrão; e
- V é a covariância.

O valor de W situa-se entre zero e um. Pequenos valores de W levam à rejeição da normalidade, enquanto o valor de um indica aderência à normalidade dos dados (RAZALI e WAH, 2011). Segundo Maroco (2010), este teste é particularmente apropriado para amostras de pequena dimensão. Whetherill (1986), por exemplo, já considerava que o teste de Shapiro-Wilks seria poderoso para amostras menores ou iguais a 50.

2.2. MAPEAMENTO DO ESTADO DA ARTE

Com o objetivo de identificar um referencial bibliográfico atualizado foi realizado um estudo bibliométrico nas Bases de Dados SCOPUS e *Web of Science*. Primeiramente, a pesquisa foi consolidada em março de 2019; posteriormente, foi atualizada em dezembro de 2019. Sobre o tema de testes de aderência, buscaram-se contribuições para a discussão sobre o emprego dos diferentes testes.

2.2.1. Pesquisa nas Bases SCOPUS e Web of Science

As buscas nas bases de dados SCOPUS e *Web of Science* foram realizadas por meio do Portal de Periódicos da Capes, acessadas entre os meses de fevereiro e maio de 2019, e atualizadas em dezembro de 2019. Inicialmente, foram utilizadas, de forma isolada, as seguintes palavras-chave acerca do tema: “*goodness-of-fit test*”; “*adherence test*”; “*normal distribution*”. O termo “*goodness-of-fit*”, cuja tradução literal seria teste de bondade de ajuste,

é a expressão mais utilizada no idioma inglês para se referir aos testes de aderência; entretanto, por excesso de zelo, também foi incluída na busca a versão literal de teste de aderência para o inglês “*adherence test*”. O objetivo inicial da referida busca, foi identificar e investigar o quantitativo geral de publicações nas respectivas bases de dados a serem pesquisadas, sem restrições quanto aos parâmetros de refinamento de pesquisa. A seguir são apresentados os resultados na Tabela 1. Ao lado das palavras-chaves, encontra-se entre parêntesis a tradução livre para os termos utilizados. A pesquisa consistiu na entrada dos termos especificados no campo de busca das bases de dados, digitados entre aspas para garantir que as palavras não fossem procuradas separadamente. Para a base SCOPUS foi utilizado o filtro de seleção de pesquisa para os campos “*Title*” (título), “*Abstract*” (Resumo) e “*Keywords*” (Palavras-chaves), enquanto na base *Web of Science*, por falta desta opção de filtro, a pesquisa foi referente à todos os campos. Vale ressaltar que, mesmo com o filtro utilizado, o número de publicações na base SCOPUS foi superior nesta primeira pesquisa.

Tabela 1: Palavras-chave da pesquisa inicial.

TERMOS DA PESQUISA	NÚMERO DE REGISTROS ENCONTRADOS		
	SCOPUS	WEB OF SCIENCE	TOTAL
<i>Goodness-of-fit test</i> (Teste de bondade de ajuste)	5.818	2.919	23.780
<i>Adherence test</i> (Teste de Aderência)	438	191	629
<i>Normal distribution</i> (Distribuição normal)	55.408	22.141	77.549
SubTotal	61.664	25.251	101.958

Fonte: Elaboração Própria (2020)

A verificação de registros encontrados nas bases de dados SCOPUS e *Web of Science* com as três palavras-chave pesquisadas isoladamente, evidenciou variados temas acerca do assunto. Em virtude da alta dispersão e variedade de registros encontrados em relação ao contexto de testes de aderência, fez-se necessário um refinamento da pesquisa com a combinação de algumas palavras-chave, com o objetivo de focar no tema central do trabalho em questão. Os termos utilizados na pesquisa anterior foram combinados com os nomes dos testes de aderência já citados, na sua versão em inglês: “*Chi-Square*”, “*Root Mean Square Deviation*”, “*Binomial*”, “*Kolmogorov-Smirnov*”, “*Shapiro-Wilk*” e “*Lilliefors*”. A pesquisa consistiu na entrada dos termos especificados, dois a dois, no campo de busca das bases de dados, com as expressões digitadas entre aspas e com filtro para busca em todos os campos. Estas duas etapas ocorreram no período de 8 de fevereiro a 18 de março de 2019, sendo atualizadas de 27 a 31 de dezembro de 2019 (Tabela 2).

Tabela 2: Palavras-chave da pesquisa refinada.

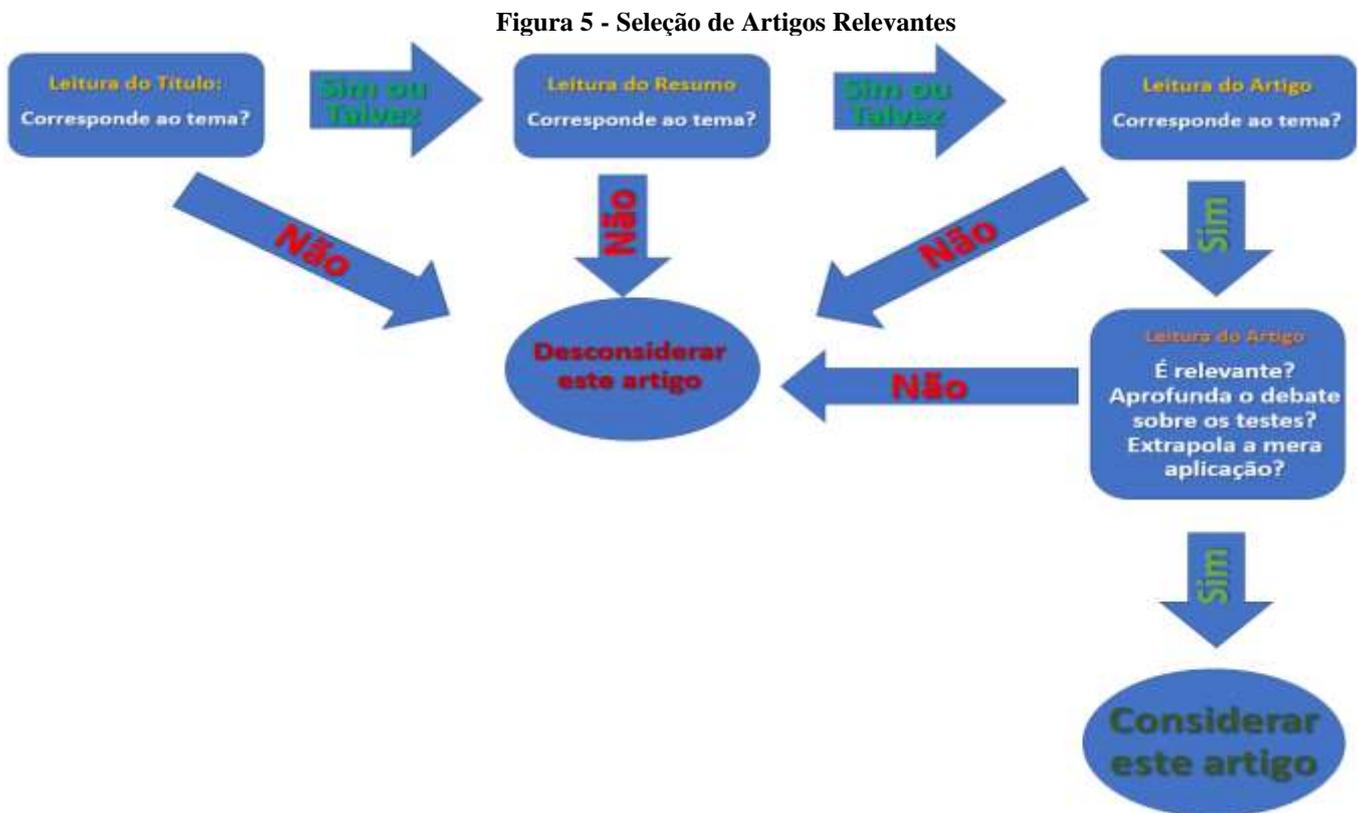
TERMOS DA PESQUISA	NÚMERO DE REGISTROS ENCONTRADOS		
	SCOPUS	WEB OF SCIENCE	TOTAL
“Goodness-of-fit test” AND “Chi-Square”	1.190	384	1.574
“Goodness-of-fit test” AND “Root Mean Square Deviation”	3	1	4
“Goodness-of-fit test” AND “Binomial”	882	59	941
“Goodness-of-fit test” AND “Kolmogorov-Smirnov”	1.915	302	2.217
“Goodness-of-fit test” AND “Shapiro – Wilk”	268	25	293
“Goodness-of-fit test” AND “Lilliefors”	343	13	356
“Adherence test” AND “Chi-Square”	14	6	20
“Adherence test” AND “Root Mean Square Deviation”	0	0	0
“Adherence test” AND “Binomial”	2	0	2
“Adherence test” AND “Kolmogorov-Smirnov”	11	6	17
“Adherence test” AND “Shapiro – Wilk”	2	1	3
“Adherence test” AND “Lilliefors”.	2	0	2
“Normal distribution” AND “Chi-Square”	1.895	289	2.184
“Normal distribution” AND “Root Mean Square Deviation”	25	4	29
“Normal distribution” AND “Binomial”	2.681	274	2.955
“Normal distribution” AND “Kolmogorov-Smirnov”	1.531	349	1.880
“Normal distribution” AND “Shapiro – Wilk”	559	200	759
“Normal distribution” AND “Lilliefors”.	459	36	495
Sub-Total	11.782	1.949	13.731

Fonte: Elaboração Própria (2020)

Cabe ressaltar que, como se pode perceber, a pesquisa nas bases de dados utilizando tão apenas as citações dos nomes dos testes de aderência resulta numa grande quantidade de registros. Entretanto, na maioria das vezes, estes autores estão apenas aplicando os testes como ferramentas estatísticas e os referidos trabalhos não desenvolvem a teoria envolvida por trás dos testes ou sequer realizam comparação entre diferentes testes.

Em que pese uma redução significativa na quantidade de documentos em relação à primeira pesquisa, o número total de trabalhos continuou muito elevado. Entre 19 de março e 24 de maio foi realizada nova fase de pesquisa. Decidiu-se descartar o termo de busca “*Adherence Test*” e realizar busca dos termos “*Goodness-of-fit test*” e “*Normal distribution*”, fazendo a interseção com cada um dos nomes dos testes separadamente. Os resultados desta nova etapa foram atualizados em 31 de dezembro de 2019. Cada busca, portanto, continha três expressões como entrada. Com o fito de ater a pesquisa às publicações mais recentes sobre o tema, para refinar ainda mais a pesquisa do estado da arte nas duas bases, foi restrito então o período a partir do ano de 2013 até dezembro de 2019, considerando apenas os artigos em periódicos.

Após a realização dessas buscas, os registros em duplicidade encontrados nas bases SCOPUS e Web of Science foram devidamente descartados, permanecendo um total de 94 artigos a serem analisados. Em seguida, a partir da análise dos resumos e dos títulos, foram descartados outros 40 artigos que certamente não correspondiam ao objetivo do estudo. Finalmente, buscando-se encontrar aqueles com maior relevância teórica ao tema da pesquisa, foi realizada a leitura completa dos demais 54 artigos. Novamente, observou-se que alguns artigos não se referiam aos testes de aderência à curva normal e a maioria dos trabalhos aplicavam os testes de aderência sem se aprofundar neste tópico. Desta forma, foram selecionados 11 artigos para estudo mais detalhado. A Figura 5 representa como foi realizada a seleção dos artigos mais relevantes.



Fonte: Elaboração Própria (2020)

Os resultados dos 11 artigos selecionados como fonte para a pesquisa, estão registrados no Quadro 1. Adicionalmente, optou-se por incluir no referido quadro o trabalho de Santos (2018), devido à sua relevância para a presente pesquisa, uma vez que esta dissertação é uma extensão daquela tese de doutorado, conforme explicitado no item 1.6. O quadro 1 apresenta, ainda, quais foram os testes de aderência utilizados em cada trabalho.

Quadro 1: Lista de trabalhos selecionados para a pesquisa.

Continua

AUTOR (ES)	ARTIGO SELECIONADO PARA A PESQUISA	PUBLICAÇÃO / PERIÓDICO	ANO	TESTES
Marcos dos Santos	Proposta de Modelagem Atuarial aplicada ao setor militar, considerando influências econômicas e biométricas	Tese de Doutorado apresentada ao Programa de Pós-Graduação Stricto Sensu em Engenharia de Produção da Universidade Federal Fluminense	2018	Qui-quadrado
Steffen Betsch; Bruno Ebner	Testing normality via a distributional fixed point property in the Stein characterization	Test	2019	Kolmogorov-Smirnov / Shapiro-Wilk / Shapiro – Francia / Anderson–Darling / Baringhaus–Henze–Epps–Pulley / Barrio–Cuesta–Albertos–Mátran–Rodríguez–Rodríguez / Henze–Jiménez-Gamero / Cràmer-von Mises
Jean-François Quessy; Louis-Paul Rivest; Toupin, Marie-Hélène	Goodness-of-fit tests for the family of multivariate chi-square copulas	Computational Statistics & Data Analysis	2019	Qui-quadrado
Piotr Sulewski	Modified Lilliefors goodness-of-fit test for normality.	Communications In Statistics-Simulation And Computation	2019	Kolmogorov-Smirnov / Lilliefors
Hina Khan; Muhammad Aslam; Masood Amjad Khan.	Properties of Exponential Ratio Type Estimators in Equal Probability Sampling: A Simulation Study	Communications in Mathematics and Statistics	2018	Qui-quadrado / Kolmogorov-Smirnov / Desvio Quadrático Médio / Anderson–Darling
Samiran Das	Goodness-of-Fit Tests for Generalized Normal Distribution for Use in Hydrological Frequency Analysis	Pure and Applied Geophysics	2018	Kolmogorov-Smirnov / Cràmer-von Mises / Anderson–Darling
Metin Öner e İpek Deveci Kocakoç	A Compilation of Some Popular Goodness of Fit Tests for Normal Distribution: Their Algorithms and MATLAB Codes	Journal of Modern Applied Statistical Methods	2017	Kolmogorov-Smirnov / Cràmer-von Mises / Shapiro-Wilk / Shapiro – Francia / Anderson–Darling / Jarque-Bera / D’Agostino / Pearson

Continuação

AUTOR (ES)	ARTIGO SELECIONADO PARA A PESQUISA	PUBLICAÇÃO / PERIÓDICO	ANO	TESTES
Avijit Hazra e Nithya Gogtay	Module on Biostatistics and Research Methodology for the Dermatologist	Indian Journal of Dermatology	2016	Kolmogorov-Smirnov / Lilliefors / Shapiro-Wilk
Almir Sedic; Danijel Pavkovic; Mihajlo Firak.	A methodology for normal distribution-based statistical characterization of long-term insolation by means of historical data	Solar Energy	2015	Qui-quadrado / Kolmogorov-Smirnov / Lilliefors / Shapiro-Wilk
Pedrosa, Ignacio; Juarros-Basterretxea, Joel ; Robles-Fernandez, Adan; Basteiro, Julia; Garcia-Cueto, Eduardo.	Goodness of Fit Tests for Symmetric Distributions, which Statistical Should I Use?	Universitas Psychologica	2015	Kolmogorov-Smirnov / Lilliefors / Shapiro-Wilk / Anderson-Darling / Jarque-Bera
Emine Tanır Kayıkçı; Eyüp Sopacı	Testing the normality of the residuals of surface temperature data at VLBI/GPS co-located sites by goodness of fit tests	Arabian Journal of Geosciences	2015	Kolmogorov-Smirnov / Shapiro-Wilk / Shapiro – Francia / Anderson-Darling / Jarque-Bera / D'Agostino / Pearson
Michelli Barros; Victor Leiva; Raydonal Ospina; Aline Tsuyuguchi	Goodness-of-Fit Tests for the Birnbaum-Saunders Distribution With Censored Reliability Data	IEEE Transactions On Reliability	2014	Qui-quadrado / Kolmogorov-Smirnov / Crâmer-von Mises / Anderson-Darling

Fonte: Elaboração Própria (2020)

O trabalho de Santos (2018) apresentou uma metodologia capaz de calcular, rapidamente e com a melhor acurácia possível, as projeções Atuariais dos custos com pensionistas das Forças Armadas Brasileiras num horizonte temporal de 75 anos. Para a realização de tais projeções, por sua vez, fez-se necessária a seleção das tábuas atuariais que melhor representam aquela parcela da população estudada. Esta seleção ocorreu a partir da utilização do teste de aderência Qui-quadrado, com 99 agravamentos e 99 desagravamentos para cada tábua atuarial. Como banco de dados, empregou-se o Banco de Informações Estratégicas e Gerenciais (BIEG), administrado pelo MD. A metodologia apresentada por

aquele autor trouxe uma significativa contribuição para a sociedade brasileira, uma vez que reduz a possibilidade de haver déficits ou superávits financeiros nas projeções atuariais, contribuindo para o equilíbrio das contas públicas. Além disso, também trouxe uma relevante contribuição para a Academia, dado o ineditismo e a aplicabilidade do estudo realizado.

Betsch e Ebner (2019) propõem duas famílias de testes para o problema clássico do teste de aderência à curva normal. Esses autores fazem um estudo comparativo de poder dos novos procedimentos numa amostra finita, mostrando a competitividade aos procedimentos clássicos. Para aqueles autores, testar a normalidade é a forma mais utilizada e estudada de emprego dos testes de aderência, o que é justificado pelo pressuposto de adequação à normal de modelos clássicos. Dentre os testes de aderência citados que coincidem com os pesquisados para a presente dissertação, os autores destacam que, enquanto o teste de Kolmogorov-Smirnov aborda o problema medindo a distância da função de distribuição empírica ao representante estimado da amostra N , outros testes empregam a técnica de correlação e regressão, como o teste Shapiro-Wilk.

Quessy, Rivest e Toupin (2019) desenvolvem testes de aderência para a família Qui-quadrado multivariada. Embora não seja exatamente o escopo do presente estudo, este artigo foi incluído pela descrição do método Qui-quadrado e discussão acerca de testes de aderência que permitam assimetrias de cauda. Os testes propostos são baseados em dois coeficientes de correlação de classificação e as estatísticas de teste são calculadas a partir de estimativas que derivam de suas distribuições. Para aqueles autores, uma vantagem importante do uso dessas medidas é que elas podem ser explicitamente calculadas a partir de resultados conhecidos nos momentos da distribuição normal bivariada, bem como facilmente estimadas por versões naturais com base em coeficientes de correlação empírica.

Sulewski (2019) propõe uma correção para o teste de aderência de Lilliefors para verificar a normalidade. Neste trabalho, o autor apresenta uma variação na fórmula de cálculo da função de distribuição empírica (*empirical distribution function* - EDF). Para ele, os valores das constantes na fórmula de cálculo do EDF influenciam diretamente o poder do teste de Lilliefors. O artigo visa, ainda, propor uma medida de similaridade entre a distribuição normal e uma distribuição alternativa. Este trabalho foi incluído por apresentar explicações sobre os métodos Kolmogorov-Smirnov e Lilliefors, além de calcular o poder dos testes para novas distribuições alternativas.

O objetivo do trabalho de Khan, Aslam & Khan (2018) foi estabelecer as propriedades do estimador tipo razão exponencial, consistentes com a média populacional, a partir do trabalho prévio de Khan *et al.* (2014). As propriedades de alguns estimadores

existentes baseados no estudo de simulação também foram discutidas. Neste contexto, portanto, os testes de aderência foram utilizados apenas como teste de qualidade de ajuste no estimador. No entanto, este trabalho vale ser mencionado porque a distribuição normal revelou-se como a curva melhor ajustada no estimador e foi possível constatar o emprego dos testes de aderência de Kolmogorov-Smirnov, Anderson-Darling, Desvio Quadrático Médio e Qui-quadrado.

O estudo de Das (2018), baseado na função empírica de distribuição, visava obter valores críticos de testes de aderência amplamente utilizados sob a hipótese nula de que as observações surgem de uma distribuição normal. O foco é o caso mais relevante para o hidrólogo nos quais os valores dos parâmetros não são identificados e passam a ser estimados de uma amostra usando o método de momentos. Os testes considerados no estudo são Kolmogorov – Smirnov, Cramér-von Mises (CVM) e Anderson–Darling (AD), além de uma versão modificada do AD, nomeada de teste Modificado de Anderson-Darling (MAD). Os valores críticos dos testes estatísticos propostos são aproximados usando técnicas de Monte Carlo para mostrar a dependência do parâmetro de forma e tamanho da amostra. O desempenho e os resultados obtidos a partir do estudo sugerem que o teste DA e uma variante do MAD (MAD-L) são os testes mais poderosos. Os valores dos testes CVM, AD e MAD mostram dependência com o parâmetro de forma, enquanto que o teste de KS mostra dependência fraca em comparação com outros testes considerados. Em termos de dependência do tamanho da amostra, os testes se comportam de forma oposta: os valores críticos do teste KS mostraram forte dependência, enquanto os demais testes mostraram pouca dependência.

O trabalho de Öner e Kocakoç (2017) revisou os algoritmos de cálculo para alguns dos testes não-paramétricos mais comuns para verificar adequação à curva normal, compilando-os como função MATLAB. Segundo esses autores, para diversas análises estatísticas, um teste de normalidade é um dos pontos mais importantes e necessários. Assim, os autores citam os testes de aderência mais usuais, classificando-os em: testes da Função de Distribuição Empírica (EDF), como Teste de Kolmogorov-Smirnov e outros que derivaram a partir deste, Teste de Anderson-Darling e Teste de Cramér-von Mises; Testes baseados em regressão e correlação, como os Teste de Shapiro-Wilk e Teste de Shapiro-Francia (SF); e Testes baseados no Momento, como o Teste Jarque-Bera (JB) e o Teste de D'Agostino e Pearson (DAP).

Hazra e Gogtay (2016) apresentam os testes de aderência como alternativas à análise visual da plotagem gráfica para fins de adequação à curva normal. Esses autores decorrem sobre diferentes testes de aderência, destacando dentre aqueles para determinar a normalidade

com base no princípio da hipótese nula, os testes de Kolmogorov-Smirnov, Shapiro-Wilk e Lilliefors. Para aqueles autores, não há resposta satisfatória para a pergunta sobre qual seria o melhor teste de aderência em uma dada situação. Porém, eles ressaltam que o teste de Kolmogorov-Smirnov é o mais antigo desta família de testes, sendo amplamente utilizado no meio acadêmico e tolerando mais desvios da normalidade estrita.

Sedic, Pavkovic e Firak (2015) propõe uma abordagem estatística para a caracterização da disponibilidade de energia solar a longo prazo com base na análise de dados históricos de insolação, aumentada por um procedimento de normalização implementando transformações simples de conjuntos de dados. A abordagem proposta produziu um conjunto normalizado de distribuições de probabilidade para os dados de energia solar coletados para a maior área metropolitana de Zagreb (República da Croácia) em um período de 10 anos, que pode ser usado para a avaliação da disponibilidade futura de energia solar. O interessante deste trabalho é que a metodologia proposta foi validada por meio de critérios de normalidade estatística, mediante o uso dos testes de aderência. Neste trabalho, portanto, foi possível ver a diferença na aplicação dos testes de Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors e Qui-quadrado.

Segundo Pedrosa *et al.* (2015), o uso de testes de aderência que não são apropriados para a amostra avaliada são aspectos comuns. Em muitos casos, isso implica o uso de testes estatísticos não ajustados para a distribuição real de dados e, conseqüentemente, o estabelecimento de conclusões incorretas. Neste trabalho, aqueles autores analisaram o poder de detecção de cinco testes de ajuste de aderência: Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Anderson-Darling e Jarque-Bera. Os testes foram aplicados em distribuições simétricas e analisados em seis tamanhos de amostra entre 30 e 1000 participantes gerados pela simulação de Monte Carlo. Os resultados mostraram uma tendência conservadora acentuada à medida que o tamanho da amostra se tornasse maior. Em relação aos tamanhos amostrais para detectar não-normalidade: analisando as amostras pequenas, os melhores resultados foram fornecidos pelos testes de Lilliefors e Anderson-Darling; para a amostra de tamanho médio estudada (200 participantes), destacou-se o teste de Kolmogorov-Smirnov; e quando as amostras terminaram em 500 participantes, ressaltou-se o teste de Shapiro-Wilk.

Kayıkçı e Sopacı (2015) estudam dados de temperatura da superfície coletados a partir da Interferometria de Linha de Base Muito Longa (*Very Long Baseline Interferometry* - VLBI) e do sistema de posicionamento global (*Global Positioning System* - GPS), localizados em duas regiões diferentes na Europa. Esses autores utilizam métodos gráficos e estatísticos de testes de aderência para verificar a normalidade dos resíduos de uma série temporal da

temperatura recebida dos GPS / VLBI. Os testes estatísticos utilizados para testar a normalidade foram: Pearson χ^2 , Kolmogorov-Smirnov e Anderson-Darling (como testes de frequência), Shapiro-Wilk e Shapiro-Francia (como testes de correlação - regressão) e D'Agostino, Jarque-Bera e teste de assimetria e curtose (como testes baseados em momentos). Os desempenhos dos testes de aderência para normalidade foram avaliados usando dois diferentes conjuntos de dados com os mesmos tamanhos de amostra. Os testes SW/SF, de Curtose e JB se mostraram mais poderosos, detectando desvios da suposição de normalidade, enquanto os testes Pearson χ^2 , KS e de assimetria foram considerados os menos poderosos. Observou-se que o teste de Anderson-Darling para a distribuição estudada provê maior atenção às caudas. Aqueles autores acrescentam que os métodos gráficos fornecem alguns dados sobre a forma de uma distribuição, mas não fornecem informações exatas sobre a diferença entre as distribuições da amostra e o esperado numa curva normal. Assim, para aqueles autores na maioria dos casos, testes de aderência para normalidade são necessários para confirmar os resultados dos métodos gráficos.

No trabalho de Barros *et al.* (2014), os autores propõem testes de aderência para a distribuição do tipo Birnbaum-Saunders, considerando as estatísticas de teste mais usadas, baseadas na função empírica de distribuição cumulativa (*empirical cumulative distribution function* - ECDF), como Anderson-Darling, Cramér-von Mises e Kolmogorov-Smirnov. Embora não seja o mesmo escopo desta dissertação, vale menção desse artigo a classificação dos tipos de testes de aderência. Segundo aqueles autores, os métodos de teste de aderência podem ser classificados como técnicas gráficas, de momento ou de transformação; ou eles podem ser do tipo qui-quadrado, ou formulado por meio de estatística com base na função de distribuição cumulativa empírica, função característica empírica, ou correlação e regressão. Os autores reforçam que estes testes têm sido propostos para avaliar as hipóteses de adequação distributiva ou sua discrepância, onde o vetor de parâmetros que indexa a distribuição pode ser conhecido ou desconhecido.

2.1.2. Estatística dos artigos selecionados

Nesta seção são apresentadas estatísticas dos artigos selecionados para a pesquisa referentes a ano de publicação, distribuição por periódico, autor, filiação e idioma. Decidiu-se por não incluir nas estatísticas dos artigos a tese de Santos (2018).

A Tabela 3 exhibe a distribuição dos artigos por ano de publicação. Observa-se que, à exceção dos anos de 2015 e 2019, em que foram selecionados 3 artigos nas bases, nos outros anos foi selecionado apenas um artigo. Pode-se dizer que os artigos selecionados foram

distribuídos uniformemente no período. Por se tratar de um espaço amostral muito pequeno, não é possível extrair outras conclusões, mas vale ressaltar que no período estudado não houve um ano sequer em que não tenha sido publicado pelo menos 1 artigo sobre o tema.

Tabela 3: Distribuição dos artigos selecionados por ano de publicação.

ANO	2014	2015	2016	2017	2018	2019
Quantidade de Artigos	1	3	1	1	1	3
Percentual de Artigos	9%	27%	9%	9%	9%	27%

Fonte: Elaboração Própria (2020)

De acordo com o Quadro 1, observa-se que os 11 artigos estavam uniformemente distribuídos em 11 diferentes periódicos, não havendo nenhuma publicação que se destacasse no assunto. Isto é, cada artigo selecionado foi publicado em um periódico diferente, que, por sua vez, contou com apenas um artigo na lista, o que representou uma participação de 9% do total de artigos para cada *Journal*. A Tabela 4 destaca as áreas de atuação de cada periódico, bem como seu fator de impacto (*Journal Citation Report – JCR*) e índice H referentes ao ano de 2018. Os dados dos periódicos foram obtidos em 28 de Fevereiro de 2020, a partir do sítio do Portal de Periódicos da CAPES na internet e/ou dos sítios das bases SCOPUS e *Web of Science*.

Tabela 4 : Distribuição dos artigos selecionados por periódico.

PERIÓDICOS	ÁREA DO CONHECIMENTO	FATOR DE IMPACTO / JCR 2018 (WEB OF SCIENCE)	ÍNDICE H 2018 (SCOPUS)
Computational Statistics & Data Analysis	Ciência da Computação / Matemática	1,323	93
Communications In Statistics-Simulation And Computation	Matemática	0,49	39
Test	Estatística e Probabilidade	1,423	35
Communications in Mathematics and Statistics	Matemática	1,18	7
Pure and Applied Geophysics	Ciências Exatas e da Terra / Ciências Biológicas	1,466	78
Journal of Modern Applied Statistical Methods	Matemática / Ciências da Decisão	0,401	21
Indian Journal of Dermatology	Ciências da Saúde / Medicina	1,411	27
Solar Energy	Ciências Exatas e da Terra / Engenharias	4,674	151
Universitas Psychologica	Ciências Humanas / Psicologia	0,331	18
Arabian Journal of Geosciences	Ciências Exatas e da Terra / Geociências	1,141	33
IEEE Transactions On Reliability	Ciências Exatas e da Terra / Ciências da Computação / Engenharias	2,888	86

Fonte: Elaboração Própria (2020)

Da mesma forma, em relação a autoria dos trabalhos, a partir do Quadro 1 é possível constatar que, dentre os 11 artigos selecionados, não houve repetição de autores. Cada autor participou de um artigo, o que representa uma contribuição de 3,7% em relação ao total de autores de todos os artigos listados. A Tabela 5 apresenta individualmente cada autor, ilustrando o país da pesquisa, a organização de filiação e Fator-H de cada pesquisador. Os dados dos autores foram obtidos em 29 de Fevereiro de 2020 dos sítios das bases SCOPUS e *Web of Science* na internet, acessados a partir do Portal de Periódicos da CAPES.

Tabela 5 : Distribuição dos artigos selecionados por autor.

AUTORES	PAÍS	FILIAÇÃO	FATOR H	
			WEB OF SCIENCE	SCOPUS
Aslam, Muhammad	Arábia Saudita	King Abdulaziz University	26	26
Barros, Michelli	Brasil	Universidade Federal de Campina Grande	11	11
Basteiro, Julia	Espanha	Universidad de Oviedo	1	0
Betsch, Steffen	Alemanha	Karlsruhe Institut für Technologie	2	1
Das, Samiran	China	Nanjing University of Information Science and Technology	4	5
Ebner, Bruno	Alemanha	Karlsruhe Institut für Technologie	4	4
Firak, Mihajlo	Croácia	University of Zagreb	5	5
Garcia-Cueto, Eduardo	Espanha	Universidad de Oviedo	13	14
Hazra, Avijit	Índia	Institute of Post Graduate Medical Education & Research	18	18
Juarros-Basterretxea, Joel	Chile	Universidad de O'Higgins	4	3
Kayıkcı, Emine Tanır	Turquia	Karadeniz Teknik Üniversitesi	3	1
Khan, Hina	Paquistão	Government College University of Lahore	6	3
Khan, Masood Amjad	Paquistão	Government College University of Lahore	8	1
Kocakoç, İpek Deveci	Turquia	Dokuz Eylül Üniversitesi	6	5
Leiva, Victor	Chile	Pontificia Universidad Católica de Valparaíso	30	31
Öner , Metin	Turquia	Celal Bayar Üniversitesi	1	2
Ospina, Raydonal	Brasil	Universidade Federal de Pernambuco	6	7
Pavkovic, Danijel	Croácia	University of Zagreb	13	14
Pedrosa, Ignacio	Espanha	Universidad de Oviedo	7	8
Quessy, Jean-François	Canadá	Université du Québec à Trois-Rivières	12	13
Rivest, Louis-Paul	Canadá	Université Laval	24	22
Robles-Fernandez, Adan	Espanha	Universidad de Oviedo	1	1
Sedic, Almir	Croácia	University of Zagreb	4	4
Sopacı, Eyüp Sopacı	Turquia	Orta Doğu Teknik Üniversitesi	1	0
Sulewski, Piotr	Polônia	Szkola Główna Gospodarstwa Wiejskiego	3	2
Toupin, Marie-Hélène	Canadá	Université Laval	4	4
Tsuyuguchi, Aline	Brasil	Universidade de São Paulo	4	1

Fonte: Elaboração Própria (2020)

A Tabela 5 contém 27 autores, ou seja, fornece a informação de que há um número maior de autores do que de artigos. Isto se deve ao fato de alguns artigos possuírem mais do que um autor. A Tabela 6 apresenta a relação entre o número de autores por artigo. Observa-se que a maioria dos artigos selecionados apresenta entre 2 e 3 autores (63,6%).

Tabela 6: Distribuição dos artigos selecionados por número de autores.

Nº DE AUTORES	1	2	3	4	5	≥ 6
Quantidade de Artigos	2	4	3	1	1	0
Percentual de Artigos	18,2%	36,4%	27,3%	9%	9%	0%

Fonte: Elaboração Própria (2020)

Como consequência desta pesquisa nos artigos recentemente publicados, pôde-se observar que não foram encontrados artigos utilizando a técnica do teste de aderência Binomial para verificação de adequação à curva normal. Por outro lado, verificou-se o frequente uso do teste de Anderson-Darling (AD). Este teste é uma modificação de Kolmogorov-Smirnov, sendo mais sensível a desvios nas caudas de uma distribuição. Como o KS, este teste informa quando é improvável que se tenha uma distribuição normal. A estatística do teste AD transforma-se na seguinte fórmula (Equação 13):

$$A^2 = -n - (1/n) \sum_{i=1}^n (2i-1) [\ln(u_i) + \ln(1-u_{n-i+1})] \quad (\text{Equação 13})$$

Tomando a função de peso para ser $\psi(x) = [F(x)(1 - F(x)) - 1]$, onde $F(x)$ é o acumulado teórico subjacente distribuição (Anderson e Darling, 1952, 1954). Onde $u_i = \Phi(z_i) = \Phi((y_i - \bar{x})/\hat{\sigma})$ (Equação 14)

é a probabilidade cumulativa do padrão distribuição normal. Média e variância são desconhecidas e estimadas a partir da amostra.

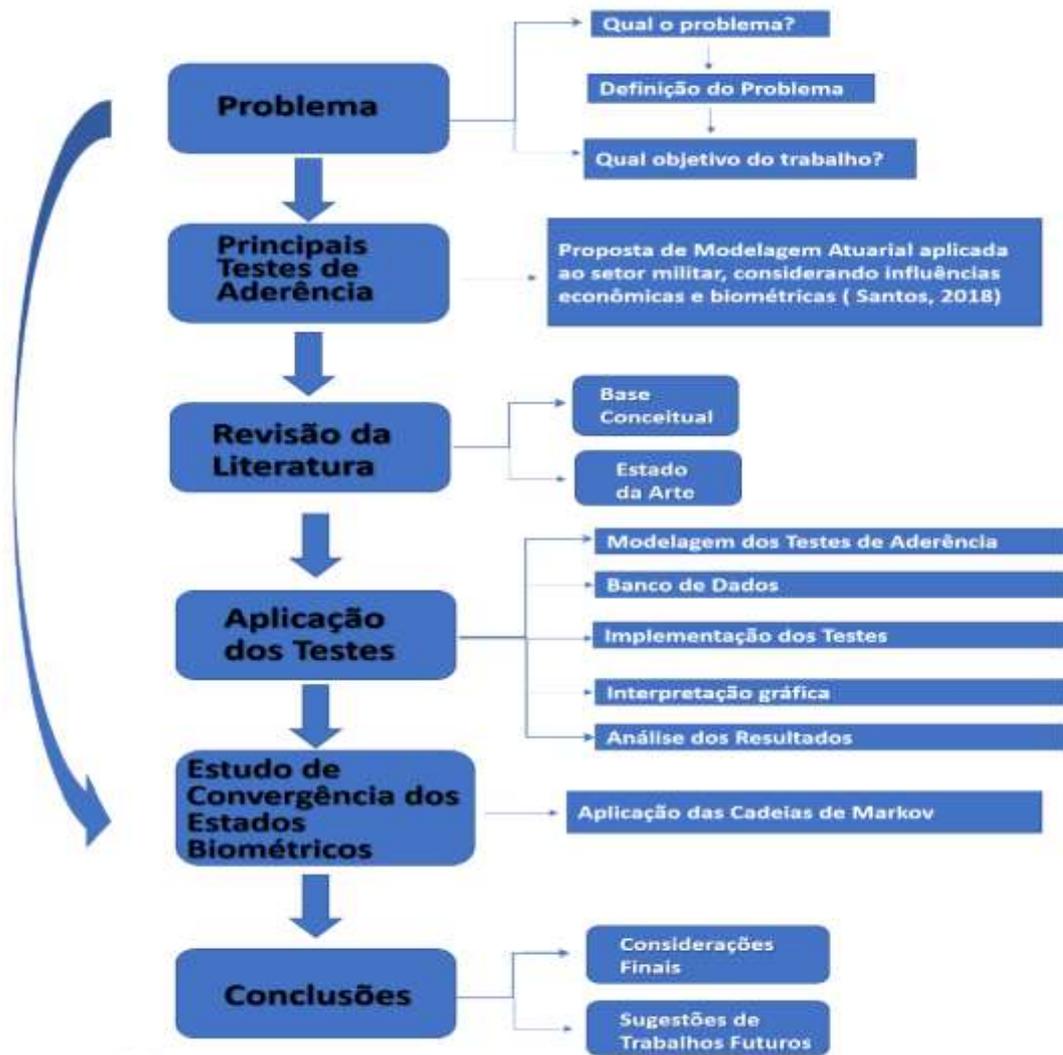
Cabe ressaltar, no entanto, que no presente trabalho, os dados populacionais históricos são conhecidos e devido às características da amostra populacional que se deseja testar, não foram utilizados o teste de Anderson-Darling ou Lilliefors para teste da aderência às tabuas atuariais. Da mesma forma, verificou-se pela revisão da literatura que o teste Binomial também não seria aplicável a este trabalho. Outra conclusão da pesquisa do estado da arte foi que, conforme já se esperava pelo referencial teórico, o teste de Shapiro-Wilk vem sendo utilizado particularmente em amostras pequenas, não sendo, portanto, compatível ao tamanho do banco de dados (BD) desta pesquisa. Assim sendo, a revisão da literatura

permitiu que o autor concentrasse a continuidade da pesquisa no estudo e aplicação dos testes de aderência Qui-quadrado, Desvio Quadrático Médio e Kolmogorov-Smirnov.

3. METODOLOGIA DA PESQUISA

As metodologias utilizadas num trabalho acadêmico estão diretamente ligadas às respostas que o pesquisador deseja obter. Este capítulo descreve a metodologia utilizada nesta pesquisa e o respectivo processo de obtenção de resposta para os problemas expressos nos seguintes questionamentos, detalhados no item 1.3. As diferentes etapas deste trabalho foram planejadas para desenvolver e atingir os objetivos propostos no item 1.4. Para melhor entendimento do desenvolvimento da pesquisa, as referidas etapas são detalhadas a seguir e ilustradas na Figura 6.

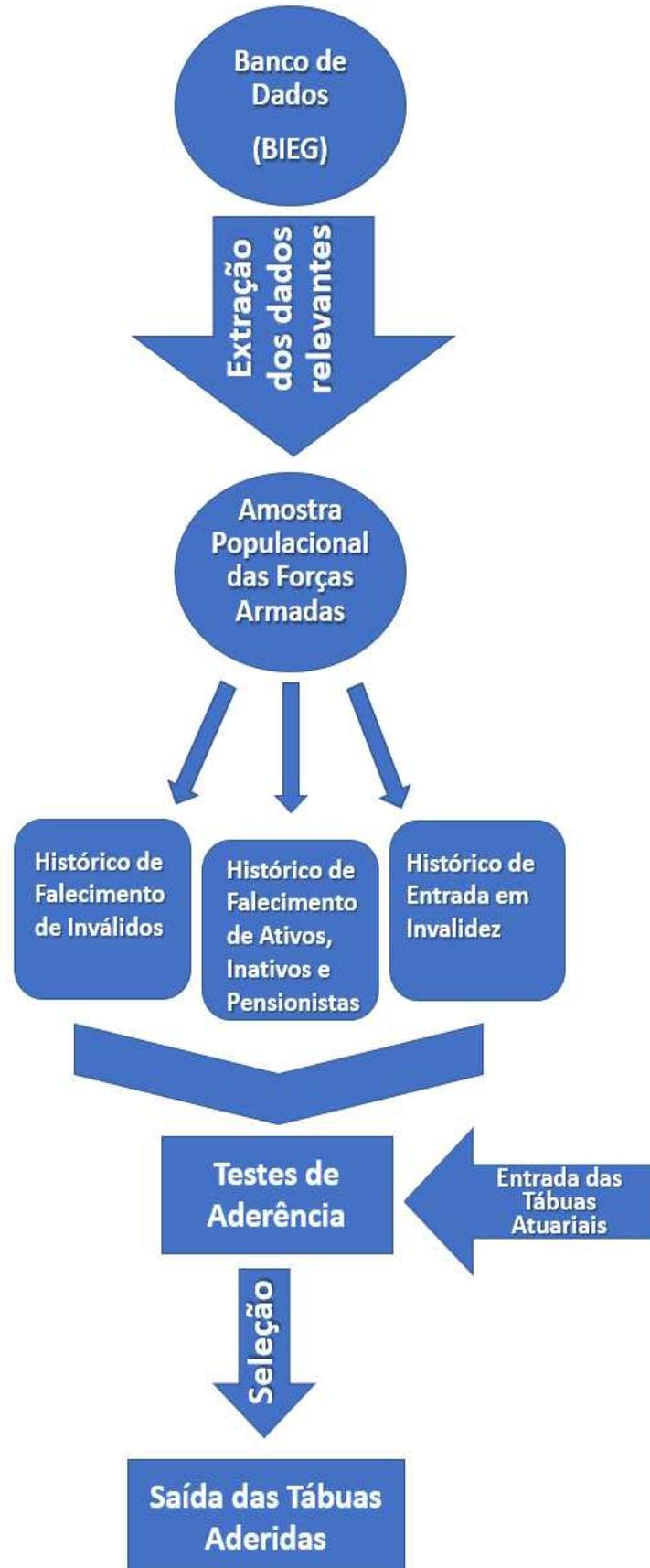
Figura 6 - Metodologia da pesquisa



Fonte: Elaboração Própria (2020)

- 1) O trabalho de pesquisa se iniciou com a definição clara do problema e das questões a serem respondidas.
- 2) Em seguida, foi feito o levantamento dos principais Testes de Aderência. Assumiu-se que os testes destacados por Santos (2018) seriam verdadeiramente os principais testes de aderência utilizados no meio acadêmico. Foi realizado um estudo da base conceitual destes testes, como alicerce da pesquisa, inclusive analisando aqueles mais citados pelos autores. Verificou-se então a relevância de acrescentar-se o teste de Shapiro-Wilk e Lilliefors à revisão bibliográfica do trabalho. A partir deste estudo os principais testes de aderência relacionados com o tema são obtidos, totalizando-se seis diferentes testes a serem pesquisados.
- 3) Para o desenvolvimento desta pesquisa realizou-se uma revisão de literatura sistematizada e estruturada nas bases Scopus e ISI *Web of Science*, conforme descrito no Capítulo 2, com o objetivo de fornecer suporte teórico ao estudo. Este estudo bibliométrico foi consolidado em março de 2019 e atualizado em dezembro do mesmo ano, obtendo 54 trabalhos com maior aproximação ao tema e que foram analisados detalhadamente. Destes, onze trabalhos foram selecionados, conforme detalhado no item 2.2.
- 4) Desenvolvimento e Aplicação do instrumento de pesquisa, isto é, realização dos testes de aderência. Esta etapa incluiu ainda os procedimentos de coleta de dados e análise dos resultados. A Figura 7 ilustra o passo a passo até a aplicação dos testes de aderência.
- 5) Estudo de Convergência dos Estados Biométricos, que consistiu, basicamente, na etapa de modelagem estatística para aplicar a metodologia de Cadeia de Markov de forma adequada aos dados, obter e analisar os resultados. Estes resultados devem ser capazes de dirimir as perguntas da pesquisa.
- 6) Consolidação do Trabalho, que contempla as conclusões gerais da pesquisa e sugestões de pesquisa futuras.

Figura 7 - Aplicação dos Testes de Aderência



Fonte: Elaboração Própria (2020)

4. MODELAGEM DOS TESTES DE ADERÊNCIA

4.1 O *Software* Atuarial das Forças Armadas (SAFA)

Os primeiros estudos de cálculo atuarial registrados foram realizados por companhias de seguro de vida no início século XVIII para fornecer uma base científica para a gestão dos ativos e passivos das empresas. O passivo dependia do número de mortes ocorridas entre as vidas seguradas a cada ano. A modelagem da mortalidade tornou-se um tópico de interesse científico, comercial e geral. Desta forma, a atuária atraiu muitos matemáticos e outros cientistas, de modo que muito do trabalho inicial no campo da probabilidade estava intimamente ligado ao desenvolvimento de soluções para problemas atuariais (DICKSON, HARDY e WATERS, 2009).

O Cálculo Atuarial aplica princípios científicos e técnicas de uma série de outras disciplinas, entre as quais matemática, teoria da probabilidade e estatística, a problemas envolvendo risco, incerteza e finanças (OGBORN, 1956; BÜHLMANN, 1997). Para D'Arcy (1989), as ciências atuariais se aproximam do estudo de economia financeira, pois ambas têm forte inclinação para matemática, abordam questões monetárias e incorporam risco em seus cálculos.

Uma das principais evoluções nas ciências atuariais veio pelo desenvolvimento de tabelas de mortalidade e demais tábuas atuariais (FREES 1990; HICKMAN, 2004). Segundo Gomes, Santos e Martins (2017) diversos cálculos econômicos e sociológicos derivam do uso das tábuas atuariais, como os relacionados à expectativa de vida, invalidez, população economicamente ativa, densidade populacional, entre outros. Para estes autores, no contexto das entidades de previdência, as tábuas atuariais mais relevantes são: tábua de mortalidade de válidos, tábua de entrada em invalidez e tábua de mortalidade de inválidos.

Segundo Gomes e Ribeiro (2014), originalmente a computação surge como um mecanismo que permite automatizar determinadas tarefas específicas nas organizações e, com o passar dos anos e evolução da tecnologia da informação (TI) se torna um fator estratégico para as mesmas. De acordo com Santos (2018), devido à necessidade de um aplicativo que permitisse a entrada dos dados dos militares das FFAA, levando em consideração as especificidades desta parcela da população, e gerasse os resultados das projeções atuariais em tempo satisfatório, em 2018 foi desenvolvido o *Software* Atuarial das Forças Armadas (SAFA). O esforço computacional para realizar esta projeção atuarial dos pensionistas das

FFAA com a melhor acurácia possível demandou a utilização de recursos de computação paralela.

O *Software* Atuarial das FFAA possibilita que os dados do BIEG sejam inseridos, calculando, em poucos minutos, a tábua atuarial que proporciona a maior precisão dos eventos decrementais mortalidade e invalidez da população militar brasileira. Ademais, o SAFA também gera como resultado as projeções atuariais a partir das premissas atuariais adotadas. O objetivo do SAFA é, portanto, apoiar o alto comando das forças armadas por meio de resultados precisos gerados a tempo de auxiliar a decisão (SANTOS, 2018).

Faz-se mister destacar o incremento da capacidade técnica que o *Software* Atuarial das Forças Armadas representou para o estudo de cálculo atuarial nas FFAA. Uma forma prática e simples de medir tal evolução se dá pela comparação entre o tempo demandado para o processamento dos dados pela versão do programa anteriormente utilizada e pelo SAFA, utilizando computação paralela. O SAFA possibilita uma diminuição de aproximadamente 99% do tempo de processamento, reduzindo de cerca de 43 horas para 20 minutos. Entretanto, cabe ressaltar que, por ocasião da seleção da melhor tábua atuarial para aquela amostra de dados inseridos, o programa utiliza como teste de aderência o Qui-quadrado, escolhido pela sua presença maciça no meio previdenciário. Vale mencionar, que o *Software* Atuarial das Forças Armadas, até o presente momento, se encontra em uso exclusivamente interno pelo Centro de Análises de Sistemas Navais (CASNAV) da Marinha do Brasil para os problemas atuariais referentes ao MD.

Atualmente, o SAFA é capaz de realizar apenas o teste de aderência do Qui-quadrado. No presente estudo, foram verificadas outras opções de teste de aderência para obter qual apresentava melhor adequação às amostras populacionais. Para a aplicação dos Testes de Aderência foi utilizado um software comercial denominado “Aderência 4D”, desenvolvido pela empresa Arcadia Sistemas, a partir de pesquisas conjuntas com o Corpo Técnico do CASNAV. O software Aderência 4D também utiliza as mais modernas técnicas de Engenharia de Software e, mediante os recursos da Computação Paralela, apresenta os resultados dos testes de aderência em tempo computacional ínfimo. Cabe destacar, portanto, que a aplicação dos outros testes de aderência neste estudo atuarial representou um acréscimo no esforço computacional insignificante (inferior a 1%) em relação à utilização unicamente do teste do Qui-quadrado.

4.2 A seleção de Tábuas Atuariais mais aderentes

4.2.1 – Tábuas Atuariais

As Tábuas Atuariais são tabelas contendo características biométricas e sociais de uma determinada população para fins de análises de risco e expectativas no campo das Ciências Atuariais. As tábuas selecionadas num estudo atuarial devem representar de forma efetiva os eventos biométricos (morte, invalidez, doença etc.) aos quais a população analisada está submetida, devendo ser escolhidas com base na experiência histórica e nas perspectivas da massa. O uso de tábuas atuariais descasadas da realidade pode resultar em perdas ou ganhos atuariais cumulativos ao longo do tempo, criando desequilíbrios estruturais ao sistema.

De acordo com Castro (1997), a tábua atuarial é a ferramenta básica para análise de mudança populacional, sendo o modelo demográfico mais antigo em uso, tendo sido empregadas, historicamente, para medir a longevidade de uma população. As primeiras tábuas conhecidas datam do século III, porém as tábuas atuariais modernas foram desenvolvidas a partir do século XVII. Desde então, as suas aplicações práticas foram diversificadas, com novas relações desenvolvidas e funções aprimoradas.

Segundo Gomes, Santos e Martins (2017), as tábuas atuariais são fundamentais para a realização de diversos cálculos econômicos e sociológicos, como cálculo de densidade populacional, mensuração de população economicamente ativa, estimativa de expectativa de vida de ativos, inativos e inválidos, quantidade esperada de pessoas que virão a se invalidar, entre outros. Para Santos (2018), as Tábuas Atuariais são o alicerce para qualquer produto na área de previdência, podendo ser oriundas de dados de censos demográficos ou da experiência de seguradoras.

As tábuas atuariais mais importantes para as entidades de previdência são: tábua de mortalidade de válidos, tábua de entrada em invalidez e tábua de mortalidade de inválido, as quais são detalhadas ao longo deste item.

4.2.1.1 – Tábuas de Mortalidade

As Tábua de Mortalidade, Tábua de Vida, Tabela de vida ou Tabela de Sobrevivência são as nomenclaturas mais comuns para este tipo de tábua atuarial que indica a expectativa de vida e probabilidade de morte em função de características como idade, sexo e/ou ocupação.

Segundo Nogueira (2015), as tábuas de mortalidade podem ser calculadas por idade simples, grupo etário e por sexo, designando um conjunto de medidas construídas para determinada coorte, ou segundo determinadas forças específicas de mortalidade aplicadas em uma coorte hipotética.

A premissa de mortalidade é a mensuração, mediante o uso das tábuas de mortalidade, do evento morte em uma população. Trata-se da probabilidade de falecimento e sobrevivência de uma população a cada idade seguinte, desde a raiz da tábua (idade inicial) até a idade inalcançável (idade final) (CORRÊA, 2018).

4.2.1.2 – Tábuas de Entrada em Invalidez

As Tábuas de Entrada em Invalidez são tábuas atuariais que expressam por meio de tabelas a probabilidade de um indivíduo se tornar inválido, em função da idade de uma determinada população.

Estas tábuas proporcionam, ademais das probabilidades de entrada em invalidez, a quantidade esperada de pessoas que teoricamente irão se invalidar por idade em uma determinada população fechada, além de outras informações estatísticas que podem ser extraídas mediante operações matemáticas (GOMES, SANTOS e MARTINS, 2017).

4.2.1.3 – Tábuas de Mortalidade de Inválidos

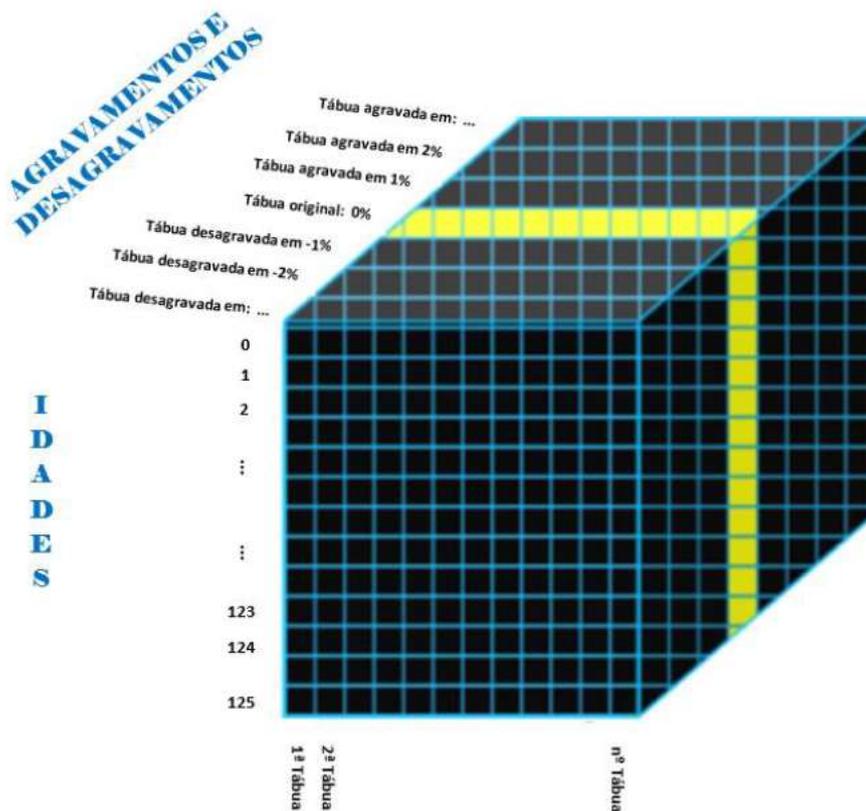
As Tábuas de Mortalidade de Inválidos são tábuas atuariais que expressam por meio de tabelas a probabilidade de falecimento de um indivíduo inválido, em função da idade de uma determinada população.

As tábuas de mortalidade de inválidos se aplicam exclusivamente aos indivíduos que deixaram de ser ativos, isto é, contraíram alguma incapacidade que lhes impede de exercer a sua atividade laboral. Estas tábuas proporcionam, ademais das probabilidades de falecimento de inválidos, a quantidade esperada de falecidos e sobreviventes por idade em uma determinada população fechada, a expectativa de vida esperada de um participante incapaz, além de informações adicionais que podem ser obtidas mediante cálculos algébricos diversos (GOMES, SANTOS e MARTINS, 2017).

4.2.2 – A Seleção das Tábuas Atuariais e o “Hiper-cubo de Aderência”

A metodologia de seleção de Tábuas Atuariais proposta por Santos (2018) consiste na aplicação do teste de aderência do Qui-Quadrado para cada tábua atuarial e suas respectivas 198 variações, que representam 99 agravamentos e 99 desagravamentos percentuais distintos. Denominou-se “Cubo de Aderência” um modelo conceitual representado por uma matriz de três dimensões que contém as probabilidades de falecimento/invalidez originais, agravadas e desagravadas. A primeira dimensão do referido cubo indica as tábuas atuariais de falecimento ou invalidez, a segunda dimensão corresponde as idades dos participantes e a terceira dimensão representa o quanto à tábua original sofreu de agravamentos e desagravamentos. A Figura 8 ilustra o Cubo de Aderência.

Figura 8 – Cubo de Aderência.



Fonte: Santos (2018)

Naquele trabalho, foram avaliados agravamentos percentuais tradicionais, que variam desde 1% até 99%, os quais consistiam em incrementar as probabilidades de falecimento/invalidez de todas as idades desde 1% ($[qx$ ou $ix] * 1,01$) até 99% ($[qx$ ou $ix] *$

1,99) respectivamente. Analogamente, os desagravamentos testados consistiam em reduzir as probabilidades de todas as idades desde 1% ($[qx$ ou $ix]^* 0,99$) até 99% ($[qx$ ou $ix]^* 0,01$).

Os agravamentos e desagravamentos, portanto, representam uma adequação matemática na função de probabilidade de morte ou entrada em invalidez a fim de tornar a tábua atuarial analisada mais aderente à população estudada. Para Caldart *et al.* (2014), de maneira simplificada, o desagravamento de uma tábua atuarial consistiria, portanto, em reduzir as taxas de mortalidade e, conseqüentemente, aumentar as expectativas de vida.

Entretanto, além de ser possível ajustar as tábuas atuariais para que estas sejam mais aderentes à amostra populacional numa grandeza diferente da observada na tábua original mediante o aumento ou redução na função de probabilidade, este ajuste também pode ser feito pelo deslocamento em anos na escala de idades da população.

No presente trabalho, optou-se, então, por incluir uma quarta dimensão no cubo de aderência descrito anteriormente, criando-se o “Hipercubo de Aderência”. Além dos agravamentos e desagravamentos percentuais já explicados, foram inseridos agravamentos e desagravamentos em anos de vida, nas idades das tábuas atuariais. Empiricamente, foram determinados os limites de 20 anos de idade para os ajustes por idades.

Os agravamentos e desagravamentos por idade testados neste trabalho consistem em deslocar as probabilidades de falecimento/invalidez de todas as idades desde 1 até 20 anos. Por exemplo, seja uma tábua atuarial qualquer denominada TABATU. Quando esta tábua é agravada em 5 anos (TABATU + 5 anos), a probabilidade da tábua original TABATU relativa à idade 5 é assumida como sendo a probabilidade da idade 0 na tábua agravada. A probabilidade da tábua original TABATU relativa à idade 6 é assumida como sendo a probabilidade da idade 1 na tábua agravada e assim por diante.

Nesta pesquisa, visando obter as tábuas atuariais que possuíssem a melhor aderência aos dados advindos do BIEG, realizaram-se testes para três grupos de indivíduos distintos, a saber: mortalidade de válidos, mortalidade de inválidos e entrada em invalidez.

As tábuas atuariais de mortalidade e entrada em invalidez utilizadas neste trabalho são relacionadas nos Quadros 2 e 3. As especificidades das tábuas atuariais utilizadas nessa pesquisa podem ser consultadas no site do Instituto Brasileiro de Atuária (IBA) e no site da Society of Actuaries (SOA).

Quadro 2 - Tábuas de mortalidade de válidos utilizadas.

ALLG-72	AMERICAN EXPERIENCE	AT-49	AT-50	AT-55
AT-71	AT-83	AT-83 masculina	AT-2000	AT-2000 F
BR-EMSsb-v.2010-m	BR-EMSsb-v.2010-f	BR-EMSmt-v.2010-m	BR-EMSmt-v.2010-f	BR-EMSsb-2015-m
BR-EMSsb – 2015 – f	BR-EMSmt-2015-m	BR-EMSmt-2015-f	CSO2001MALE	CSO2001FEMALE
CSO-41	CSO-58	CSO-80	CSG-60	GAM-1971
GAM 1994 Masculina	GKM-70	GKM-80	GRM-80	GRF-80
GRM-95	GRF-95	GRUPAL AMERICANA	HUNTER SEMITROPICAL	IAPC
IBGE-2011-M	IBGE-2011-F	IBGE-2011	IBGE-2012-M	IBGE-2012-F
IBGE-2012	PRUDENTIAL 1950	SGB-51	RENTIERS FRANCAIS	RP-2000 - 1992 Base-Male Aggregate
SGB-71	SGB-75	USTP-61	UP-84	UP94Homens
UP94Mulheres	UP-94 MT-M-ANB	X-17	-----	----- --

Fonte: Elaboração Própria (2020)

Quadro 3 - Tábuas de entrada em invalidez e mortalidade de inválidos utilizadas

ALLG_72	IAPB-57	TASA-1927
ÁLVARO VINDAS	IAPB-57 FRACA	USTP61
BENTZIEN	IAPB-57 FORTE	X17
GRUPAL AMERICANA	IAPB-57 (AJUST/ITAU)	WINKLEVOSS
HUNTER'S	MULLER	ZIMMERMANN
IAPC	PRUDENTIAL (FERR. APOSENT.)	ZIMMERMANN (EMPR. ESCRIT)
IBA (FERROVIARIOS)	RENTIERS_FRANCAIS	ZIMMERMANN (FERR. ALEMAES)

Fonte: Elaboração Própria (2020)

Cabe destacar que nesta pesquisa, no que tange à seleção da Tábua Atuarial, o presente trabalho vem repetindo a prática tradicional de que o agravamento ou desagravamento seja realizado de forma uniforme ao longo de todas as idades para adoção da Tábua Biométrica.

4.3 Aplicação dos Testes de Aderência

A aplicação dos Testes de Aderência pode ser realizada de duas formas: a partir do Cálculo da Idade Exata ou para Classes de Idades. No primeiro, o teste de aderência é realizado para toda a amostra populacional, considerando todas as idades da população. No teste de aderência por Classe de Idades, a amostra é dividida por classes de faixas etárias arbitradas e cada classe representa uma observação independente da amostra. Neste segundo caso, o teste é realizado para cada classe e integrado para apresentar o resultado que melhor se adequou àquela população. Em que pese na academia seja consagrado o uso dos testes de aderência, durante o presente estudo, verificou-se que normalmente os testes de aderência são aplicados para o cálculo de idade exata. Porém, nesta pesquisa, observou-se que no mercado previdenciário utiliza-se também o teste de aderência por classes de idades.

Cabe destacar que como as bases de dados das FFAA são bem grandes, historicamente, utiliza-se apenas o teste de aderência por Idade Exata e não por Classes de Idades. Durante a presente pesquisa, não foram encontrados registros de realização de testes de aderência por Classes de Idades com a amostra populacional das FFAA, o que denota um provável ineditismo nesta aplicação. Desta forma, nesta dissertação, decidiu-se realizar os testes de aderência por Idade Exata e, também, por Classes de Idades.

No presente estudo, foram utilizados os dados do BIEG consolidados em 2017 relativos ao ano de 2015, referente à MB, EB e FAB. Assim sendo, conforme ilustrado na Tabela 7, mais de 1,5 milhão de registros históricos foram analisados. O referido Banco de Dados (BD) engloba informações dos militares quanto às datas de nascimento, ingresso nas FFAA, invalidez, falecimento e outros, cujos quantitativos estão subdivididos por militares ativos e inativos, pensionistas e beneficiários.

Cabe ressaltar que, dentro do objetivo do presente trabalho, uma amostra de dados relativo ao ano de 2015, ainda que não seja a mais recente, atente às necessidades da pesquisa. Este volume de dados foi escolhido por ter sido tratado no Centro de Análises de Sistemas Navais e considerado mais robusto, consistente e coerente que outras amostras; além disso,

por ter sido exatamente a mesma massa de dados utilizada na tese de Santos (2018), a qual esta pesquisa se propõe a dar continuidade.

Tabela 7: Quantitativo de militares ativos/inativos e pensionistas em 2015.

	MB	EB	FAB	FFAA
Militares ativos/inativos	212.580	951.941	229.172	1.393.693
Pensionistas	47.985	101.569	34.652	184.206
Total	260.565	1.053.510	263.824	1.577.899

Fonte: Ministério da Defesa (2017).

Além de utilizar toda a faixa etária do banco de dados da população disponível no BIEG, isto é, entre zero e cento e vinte e cinco anos, como primeira entrada para os testes de aderência, outros intervalos de idades ou de classes de idades foram estabelecidos como entradas alternativas. Estas outras entradas foram estabelecidas empiricamente, para efeitos de comparação neste estudo, pois a primeira entrada não resultou em tábuas aderidas para todos os tipos de teste executados. Para a aplicação dos Testes de Aderência foi empregado o Software Aderência 4D.

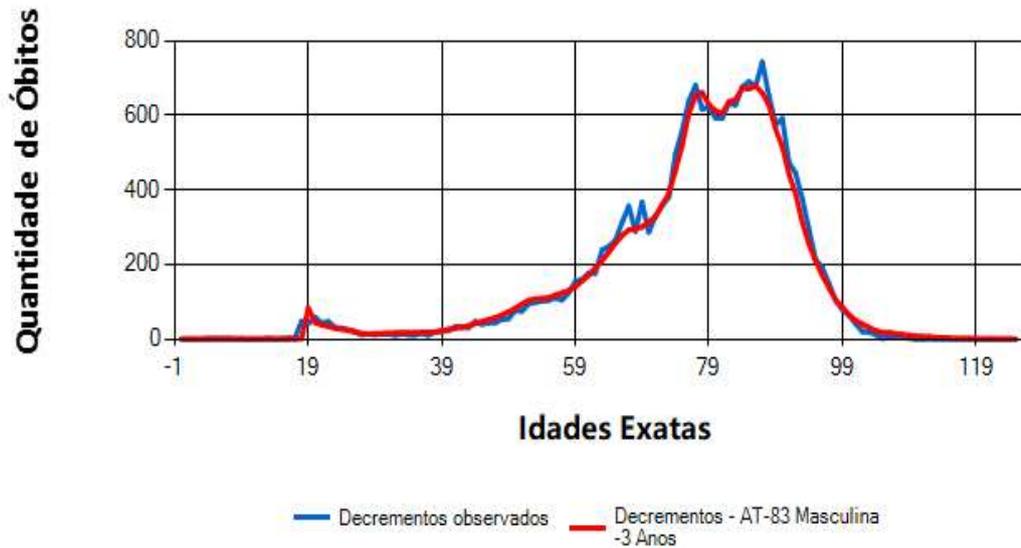
4.3.1 Resultado de Mortalidade de Ativos, Inativos e Pensionistas das Forças Armadas

4.3.1.1 Aplicação dos Testes com Cálculo da Idade Exata

Para a população de ativos, inativos e pensionistas das Forças Armadas foram realizados os testes de aderência para duas entradas possíveis. A primeira entrada de dados de mortalidade considerando toda a população, isto é, todas as idades; enquanto a segunda entrada de dados considerando uma amostra representativa desta população, com o intervalo de idades entre 20 e 90 anos, que engloba 89% dos dados populacionais.

Considerando a primeira entrada, com toda a população, não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de mortalidade, a partir do teste do Qui-Quadrado. Com o teste de Kolmogorov-Smirnov, houve aderência de 481 tábuas. A tábua de melhor aderência foi a tábua AT-83 Masculina desagravada por idade em 3 anos, conforme Gráfico 1.

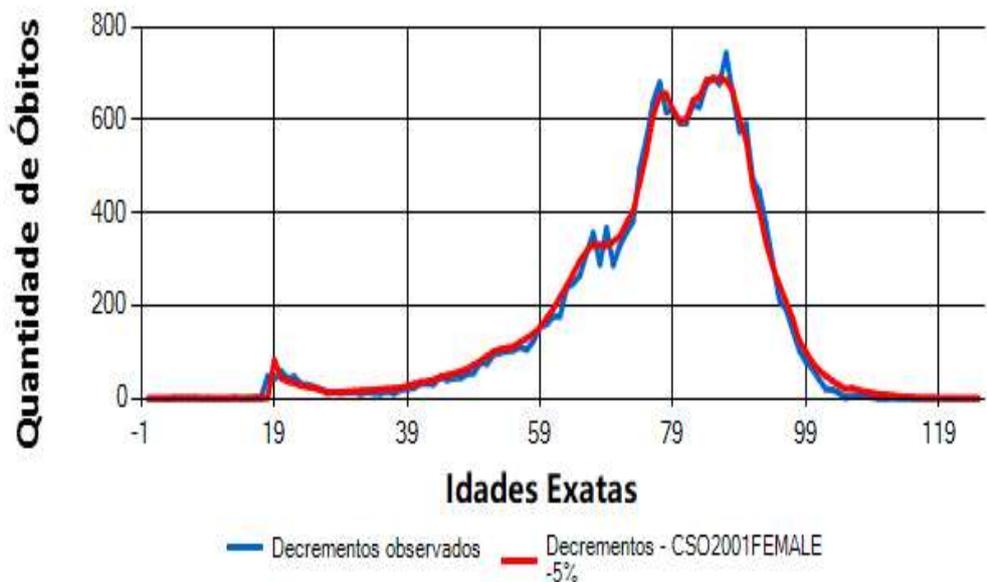
Gráfico 1 – Tábua de melhor aderência Teste KS para todas as idades



Fonte: Elaboração Própria (2020)

A partir do teste do Desvio Quadrático Médio, a tábua de melhor aderência foi a tábua CSO2001FEMALE desagradada em 5%, conforme Gráfico 2.

Gráfico 2 – Tábua de melhor aderência Teste DQM para todas as idades



Fonte: Elaboração Própria (2020)

Para a segunda entrada, com o intervalo de idades entre 20 e 90 anos, uma única tábua aderiu à mortalidade observada a partir do teste do Qui-Quadrado e foi a GKM-70 desagradada em 61%, conforme apresentado no Gráfico 3.

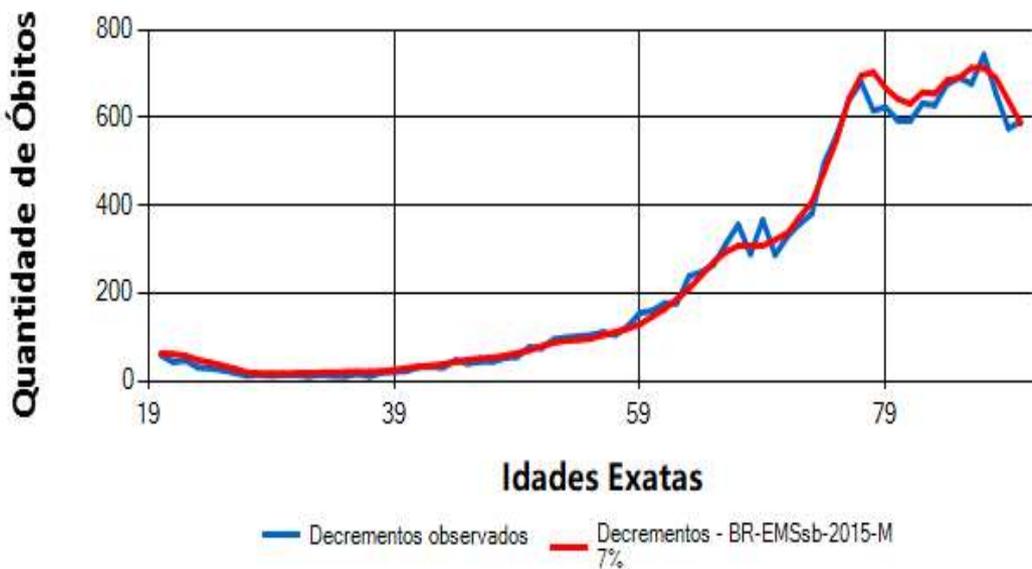
Gráfico 3 – Tábua de melhor aderência Teste Qui-quadrado para intervalo de 20 a 90 anos



Fonte: Elaboração Própria (2020)

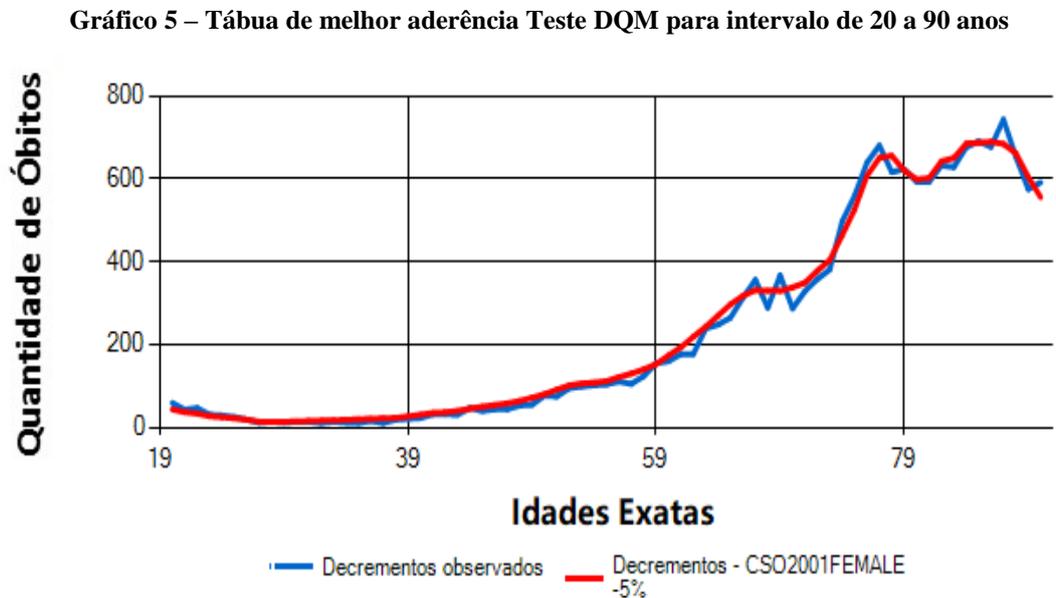
Com o teste de Kolmogorov-Smirnov, houve aderência de 1.071 tábuas, sendo a tábua de melhor aderência foi a tábua BR-EMSsb-2015-M agravada em 7%, ilustrada no Gráfico 4.

Gráfico 4 – Tábua de melhor aderência Teste KS para intervalo de 20 a 90 anos



Fonte: Elaboração Própria (2020)

A partir do teste do Desvio Quadrático Médio, a tábua de melhor aderência para a segunda entrada com idades entre 20 e 90 anos foi a tábua CSO20001FEMALE desagravada em 5%, conforme Gráfico 5.



Fonte: Elaboração Própria (2020)

A Tabela 8 resume a comparação da quantidade de tábuas aderidas em cada um do teste, enquanto a Tabela 9 destaca as tábuas que apresentaram melhor aderência no cálculo para idade exata.

Tabela 8 – Quantidade de Tábuas de Mortalidade Geral aderidas a partir de cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	481	12.667
Amostra com intervalo de idade entre 20 e 90 anos	1	1.071	12.667

Fonte: Elaboração Própria (2020)

Tabela 9 – Tábuas de Mortalidade Geral com melhor aderência a partir de cada teste

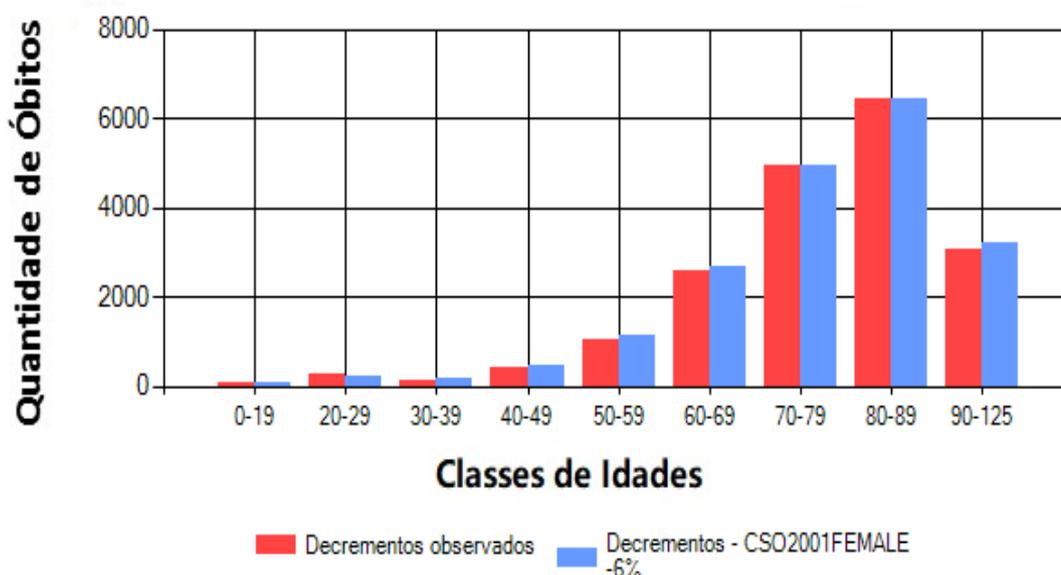
Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	AT-83 Masculina desagravada por idade em 3 anos	CSO 2001FEMALE desagravada de 5%
Amostra com intervalo de idade entre 20 e 90 anos	GKM-70 desagravada 61%	BR-EMSSb-2015-M agravada em 7%,	CSO 2001FEMALE desagravada de 5%

Fonte: Elaboração Própria (2020)

4.3.1.2 Aplicação dos Testes com Cálculo por Classes de Idades

Considerando a primeira entrada, com toda a população, foi estabelecida como classe inferior a amostra compreendida no intervalo de idade inferior 0 (zero) anos e idade superior 19 (dezenove) anos. Foram estabelecidas sete classes intermediárias com amplitude de dez anos. A classe superior foi estabelecida com idade inferior da classe de 90 (noventa) anos e idade superior da classe 125 (cento e vinte e cinco) anos. Para esta entrada, não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de mortalidade a partir do teste do Qui-Quadrado. Com o teste de KS, houve aderência de 1.203 tábuas. A tábua de melhor aderência foi CSO 2001FEMALE desagravada de 6%, conforme Gráfico 6.

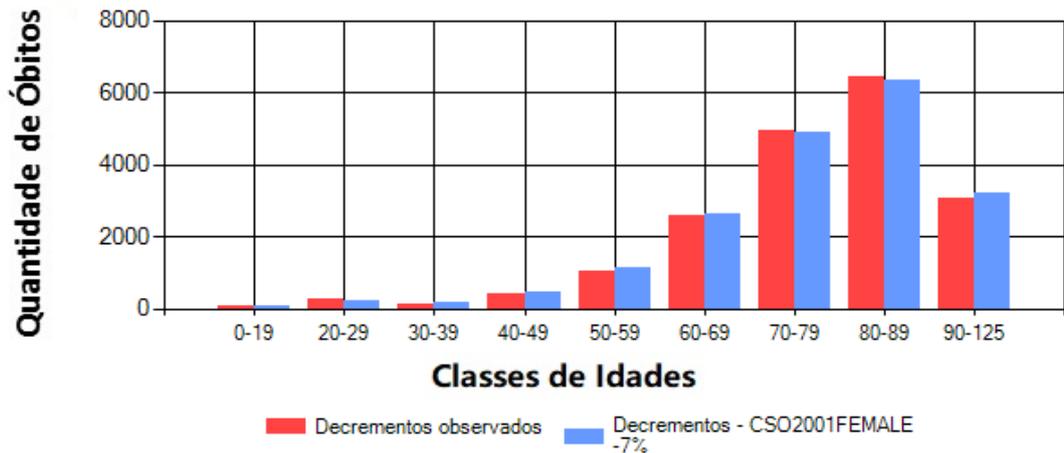
Gráfico 6 – Tábua de melhor aderência Teste KS para todas as idades



Fonte: Elaboração Própria (2020)

Para esta mesma entrada, considerando toda a população, com a mesma divisão de classe de idades, mediante o teste do Desvio Quadrático Médio, a tábua de melhor aderência foi a CSO 2001 FEMALE desagravada em 7%, conforme ilustrada no Gráfico 7.

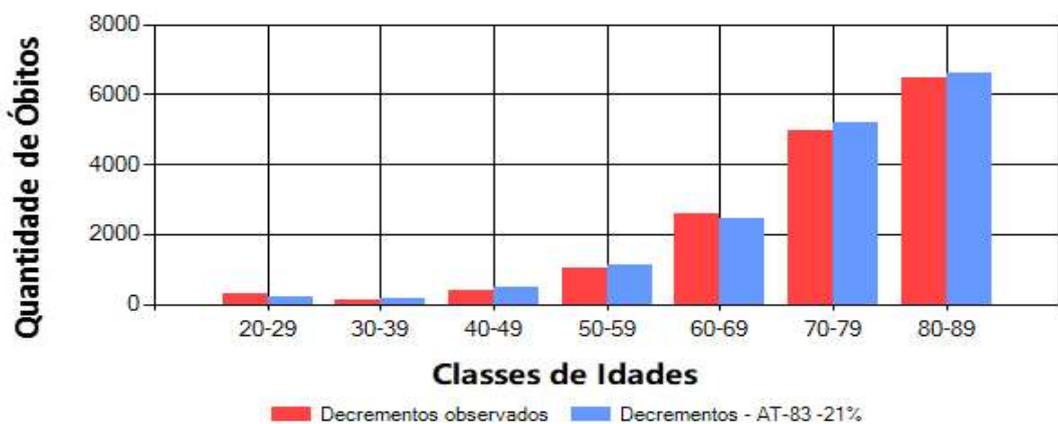
Gráfico 7 – Tábua de melhor aderência Teste DQM para todas as idades



Fonte: Elaboração Própria (2020)

Para a segunda entrada, com o intervalo de idades entre 20 e 90 anos, mediante o teste do Qui-Quadrado, ainda não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de mortalidade. Utilizando-se o teste de Kolmogorov-Smirnov, houve aderência de 1.936 tábuas. A classe inferior da amostra foi considerada compreendida no intervalo de idade inferior vinte anos e idade superior trinta anos. Foram estabelecidas cinco classes intermediárias com amplitude de dez anos. A classe superior foi estabelecida com idade inferior da classe de oitenta anos e idade superior da classe oitenta e nove anos. Para esta divisão de classe de idades, a tábua de melhor aderência foi a AT-83 desagravada em 21%, conforme ilustrada no Gráfico 8.

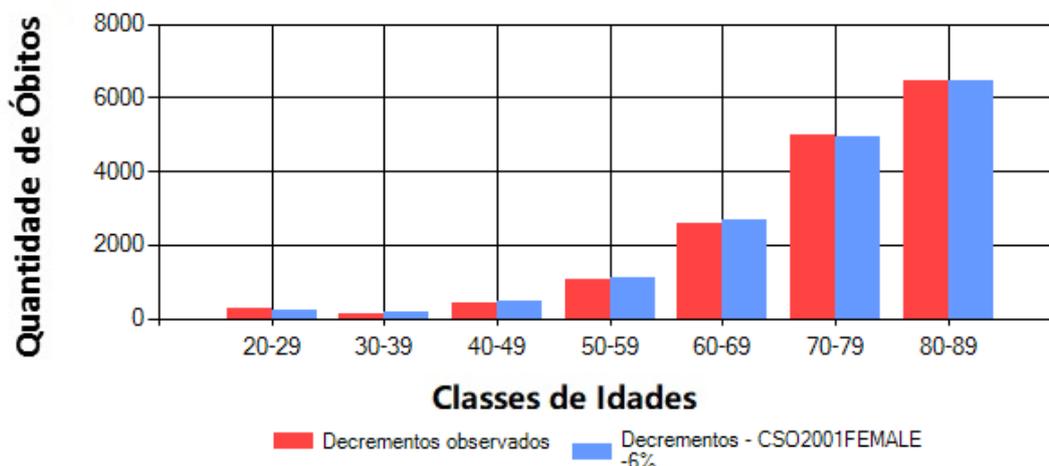
Gráfico 8 – Tábua de melhor aderência Teste KS para intervalo de 20 a 90 anos



Fonte: Elaboração Própria (2020)

Para esta mesma entrada, com amostra entre 20 e 90 anos e com a mesma divisão de classe de idades, mediante o teste do Desvio Quadrático Médio, a tábua de melhor aderência foi a CSO 2001 FEMALE desagravada em 6%, conforme ilustrada no Gráfico 9.

Gráfico 9 – Tábua de melhor aderência Teste DQM para intervalo de 20 a 90 anos

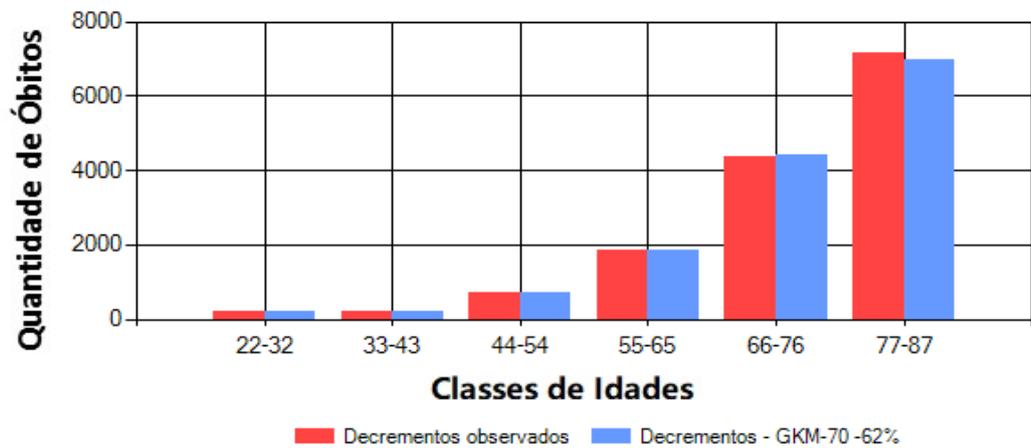


Fonte: Elaboração Própria (2020)

Como não foi possível encontrar uma tábua aderente pelo Cálculo por Classes de Idades a partir do Teste do Qui-quadrado para as duas entradas propostas no Cálculo por Idade Exata, com o único objetivo de dar prosseguimento com a pesquisa forçando a aderência, decidiu-se experimentalmente testar outros intervalos de idade. A partir de diversos testes, obteve-se, empiricamente, que o intervalo de idade entre 22 e 87 anos, correspondente a cerca de 79% dos dados populacionais, possibilita que uma tábua seja aderente. Considerando, portanto, como classe inferior a amostra compreendida no intervalo de idade inferior vinte e dois anos e idade superior trinta e dois anos, foram estabelecidas quatro classes intermediárias com amplitude de onze anos. A classe superior foi estabelecida com idade inferior da classe de setenta e sete anos e idade superior da classe oitenta e sete anos. Para esta entrada específica, a única tábua aderente mediante o teste do Qui-quadrado foi a GKM-70, desagravada de 62%, conforme Gráfico 10.

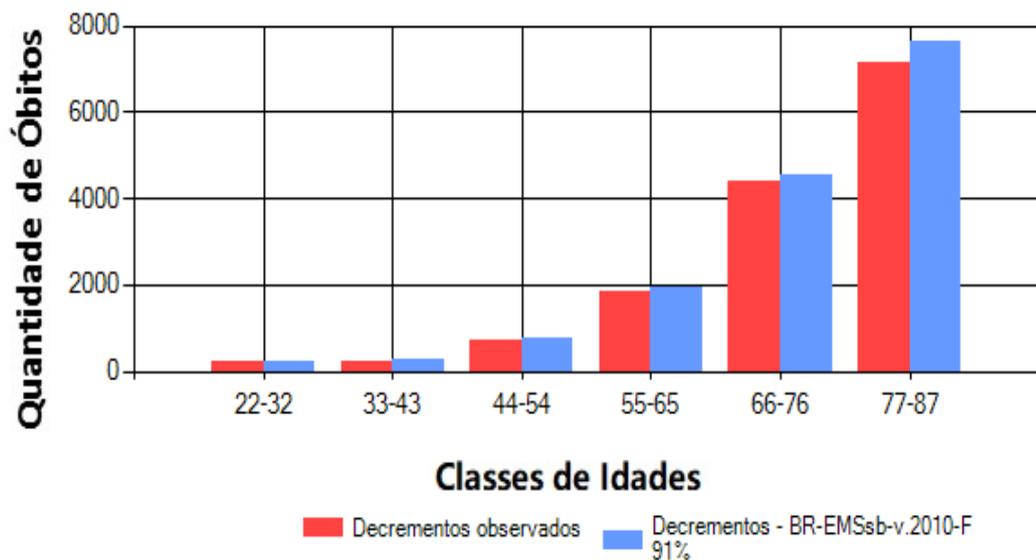
Posteriormente, para fins de comparação, com a mesma entrada de intervalo entre 22 e 87 anos e mesma divisão de classes anterior foi realizado o teste de Kolmogorov-Smirnov. Então, foram obtidas 2.559 tábuas aderidas. A tábua com melhor aderência foi a tábua BR – SEMsb-v.2010F agravada em 91%, conforme ilustrado no gráfico 11.

Gráfico 10 – Tábua de melhor aderência Teste Qui-quadrado para intervalo de 22 a 87 anos



Fonte: Elaboração Própria (2020)

Gráfico 11 – Tábua de melhor aderência Teste KS para intervalo de 22 a 87 anos

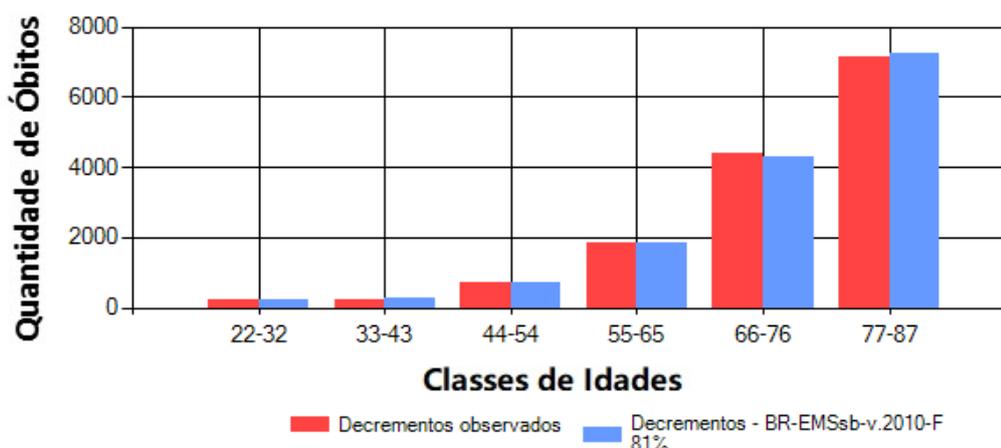


Fonte: Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada de intervalo entre 22 e 87 anos e mesma divisão de classes anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua BR – SEMsb-v.2010F agravada em 81%, conforme ilustrado no Gráfico 12.

A tabela 10 resume a comparação da quantidade de tábuas aderidas em cada um do teste, enquanto a Tabela 11 destaca as tábuas que apresentaram melhor aderência no cálculo por classe de idades.

Gráfico 12 – Tábua de melhor aderência Teste DQM para intervalo de 22 a 87 anos



Fonte: Elaboração Própria (2020)

Tabela 10 – Quantidade de Tábuas de Mortalidade Geral aderidas a partir de cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	1.203	12.667
Amostra com intervalo de idade entre 20 e 90 anos	Não houve aderência	1.936	12.667
Amostra com intervalo de idade entre 22 e 87 anos	01	2.559	12.667

Fonte: Elaboração Própria (2020)

Tabela 11 – Tábuas de Mortalidade Geral com melhor aderência a partir de cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	CSO 2001FEMALE desagravada de 6%	CSO 2001FEMALE desagravada de 7%,
Amostra com intervalo de idade entre 20 e 90 anos	Não houve aderência	AT-83 desagravada em 21%	CSO 2001FEMALE desagravada de 6%
Amostra com intervalo de idade entre 22 e 87 anos	GKM-70 desagravada de 62%	BR – SEMsb-v.2010F agravada em 91%	BR – SEMsb-v.2010F agravada em 81%

Fonte: Elaboração Própria (2020)

4.3.1.3 Análise dos Resultados para Mortalidade de Ativos, Inativos e Pensionistas

Os testes de aderência para Mortalidade de Ativos, Inativos e Pensionistas demonstraram que o teste Qui-quadrado não é o mais adequado para esta amostra populacional. Para que alguma tábua fosse aderente a partir deste teste, fez-se necessário reduzir o espaço amostral e, ainda assim, o número de tábuas aderidas foi mínimo. Por outro lado, o teste de Kolmogorov-Smirnov demonstrou-se aplicável para toda a massa de dados. Assim como o Desvio Quadrático Médio.

Devido às diferentes estatísticas de teste de cada tipo de teste de aderência, torna-se difícil comparar os resultados das tábuas melhor aderidas em cada teste aplicado de forma distinta. No teste do Qui-quadrado, calcula-se a estatística X^2 e o melhor resultado do teste é aquele com o menor valor desta estatística, dentre aqueles cujo valor seja inferior ao valor crítico oriundo da tabela de distribuição Qui-Quadrado com nível de significância de 5%. No teste KS, por sua vez, calcula-se a estatística p-valor e o melhor resultado do teste é aquele com o menor valor desta estatística, dentre aqueles cujo valor seja inferior ao valor crítico oriundo da tabela de distribuição KS com nível de significância de 5%. E no teste do Desvio Quadrático Médio, o melhor resultado do teste é o resultado da menor estatística DQM. Portanto, tratam-se de cálculos algébricos e grandezas matemáticas distintas.

Durante esta pesquisa, não foi encontrada na literatura ou no mercado previdenciário uma maneira específica de realizar tal comparação. Assim sendo, optou-se por uma análise qualitativa dos dados encontrados. O Quadro 4 resume a relação entre as tábuas melhores aderidas em cada teste e o seu resultado nos outros testes. Nesta tabela, a legenda “S” significa “sim”, isto é, aquela determinada tábua foi aderida no teste em questão. A legenda “N”, significa “não”, isto é, aquela determinada tábua não foi aderida naquele teste. A legenda “S*” indica que aquela tábua foi a de melhor aderência no referido teste.

Após a análise qualitativa do Quadro 4, verifica-se que para este banco de dados de entrada de Mortalidade de Ativos, Inativos e Pensionistas no ano de 2015, a tábua atuarial obtida com melhor aderência foi a tábua AT-83 Masculina desagravada por idade em 3 anos, a partir do teste de Kolmogorov-Smirnov.

Quadro 4 – Tábuas de Mortalidade de Ativos, Inativos e Pensionistas Melhor Aderidas para Cada Teste

Teste	Qui-Quadrado					Kolmogorov-Smirnov					Desvio Quadrático Médio				
	Idade Exata		Classes de Idades			Idade Exata		Classes de Idades			Idade Exata		Classes de Idades		
Tábua \ Amostra	Toda População	20 a 90 anos	Toda População	20 a 90 anos	22 a 87 anos	Toda População	20 a 90 anos	Toda População	20 a 90 anos	22 a 87 anos	Toda População	20 a 90 anos	Toda População	20 a 90 anos	22 a 87 anos
AT-83 Masculina – 3 anos	N	N	N	N	N	S*	S	S	S	S	S	S	S	S	S
AT-83 – 21%	N	N	N	N	N	N	N	S	S*		S	S	S	S	S
BR – SEMsb- v.2010F + 81%	N	N	N	N	N	N	N	N	N		S	S	S	S	S*
BR – SEMsb- v.2010F + 91%	N	N	N	N	N	N	N	N	N	S*	S	S	S	S	S
BR – EMSsb-2015- M + 7%	N	N	N	N	N	N	S*	S	N		S	S	S	S	S
CSO 2001 FEMALE – 5%	N	N	N	N	N	N	N	S	N		S*	S*	S	S	S
CSO 2001 FEMALE – 6%	N	N	N	N	N	N	N	S*	N		S	S	S	S*	S
CSO 2001 FEMALE – 7%	N	N	N	N	N	N	N	S	N		S	S	S*		S
GKM-70 – 61%	S*	N	N	N	N	N	N	N	N	N	S	S	S	S	S
GKM-70 – 62%	N	N	N	N	S*	S*	N	N	N	N	S	S	S	S	S

Fonte: Elaboração Própria (2020)

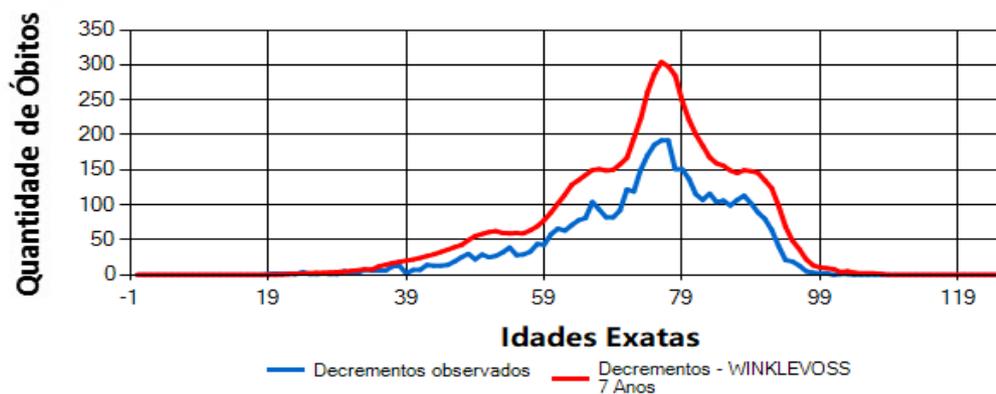
4.3.2 Resultado de Mortalidade de Inválidos das Forças Armadas

4.3.2.1 Aplicação dos Testes com Cálculo da Idade Exata

Para a população de inválidos das Forças Armadas foram realizados os testes de aderência para duas entradas possíveis. A primeira entrada de dados de mortalidade considerando toda a população, isto é, todas as idades; enquanto a segunda entrada de dados considerando uma amostra representativa desta população, com o intervalo de idades entre 21 e 93 anos, que engloba 94% dos dados populacionais.

Considerando a primeira entrada, com toda a população, não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de mortalidade, a partir do teste do Qui-Quadrado. Com o teste de KS, houve aderência de 71 tábuas. A tábua de melhor aderência foi a tábua Winklevoss +7 anos, conforme Gráfico 13.

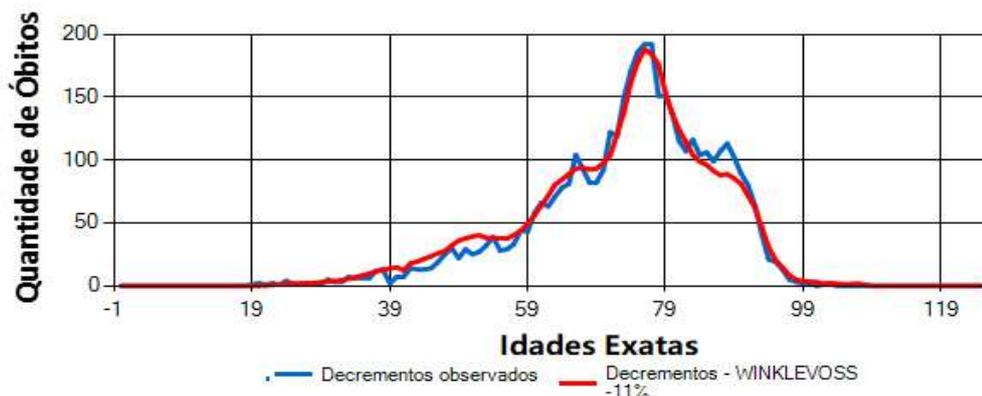
Gráfico 13 – Tábua de melhor aderência Teste KS para todas as idades



Fonte: Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Winklevoss desagravada em 11%, conforme ilustrado no Gráfico 14.

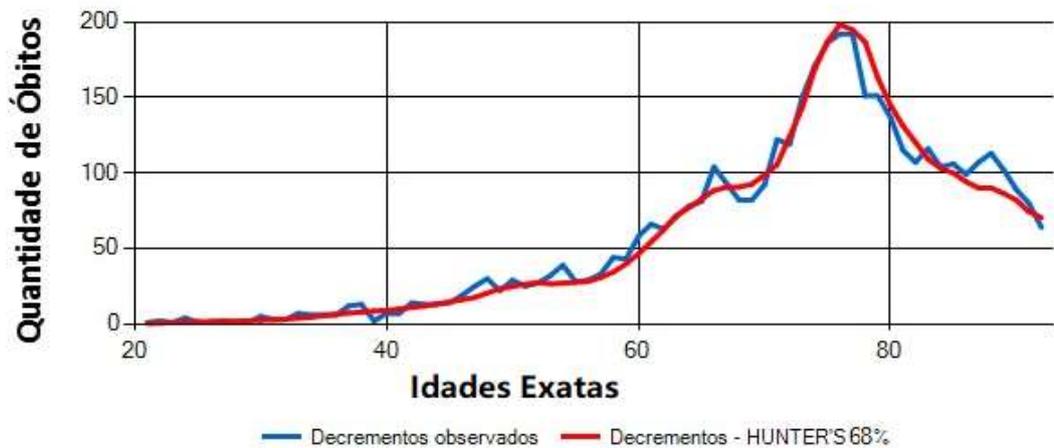
Gráfico 14 – Tábua de melhor aderência Teste DQM para todas as idades



Fonte: Elaboração Própria (2020)

Para a segunda entrada, com o intervalo de idades entre 21 e 93 anos, dezessete tábuas aderiram à ocorrência de falecimento de inválidos, observadas a partir do teste do Qui-Quadrado. As tábuas que aderiram à mortalidade de inválidos foram as tábuas Hunter's para todos os agravamentos entre 60% e 76%. A tábua de melhor aderência, isto é, menor estatística X^2 , foi a tábua Hunter's agravada em 68%, conforme apresentado no Gráfico 15.

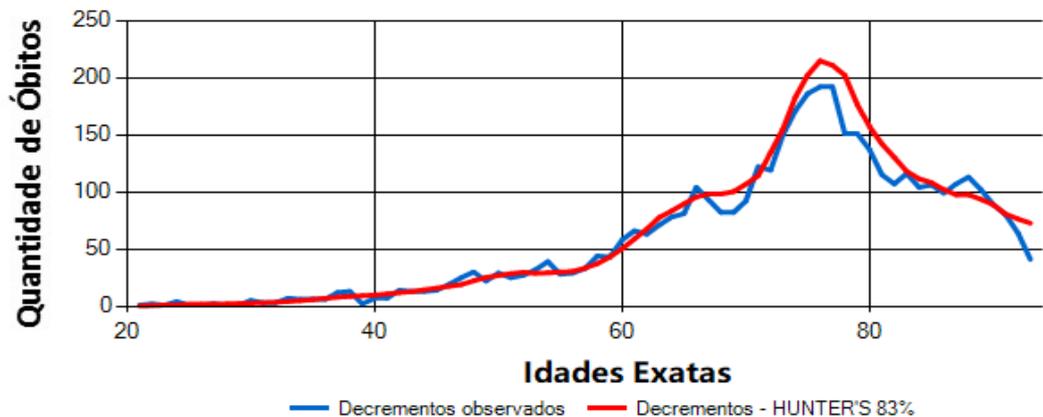
Gráfico 15 – Tábua de melhor aderência Teste Qui-quadrado para intervalo de 21 a 93 anos



Fonte: Elaboração Própria (2020)

Com o teste de KS, houve aderência de 258 tábuas, sendo a tábua de melhor aderência foi a Hunter's agravada em 83%, ilustrada no Gráfico 16.

Gráfico 16 – Tábua de melhor aderência Teste KS para intervalo de 21 a 93 anos



Fonte: Elaboração Própria

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Winklevoss desagravada em 10%, conforme ilustrado no Gráfico 17.

Gráfico 17 – Tábua de melhor aderência Teste DQM para intervalo de 21 a 93 anos



Fonte: Elaboração Própria

A tabela 12 resume a comparação da quantidade de tábuas aderidas em cada um do teste, enquanto a Tabela 13 destaca as tábuas que apresentaram melhor aderência no cálculo por classe de idades.

Tabela 12 – Quantidade de Tábuas de Mortalidade de Inválidos aderidas a partir de cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	71	12.667
Amostra com intervalo de idade entre 21 e 93 anos	17	258	12.667

Fonte: Elaboração Própria (2020)

Tabela 13 – Tábuas de Mortalidade de Inválidos com melhor aderência a partir de cada teste

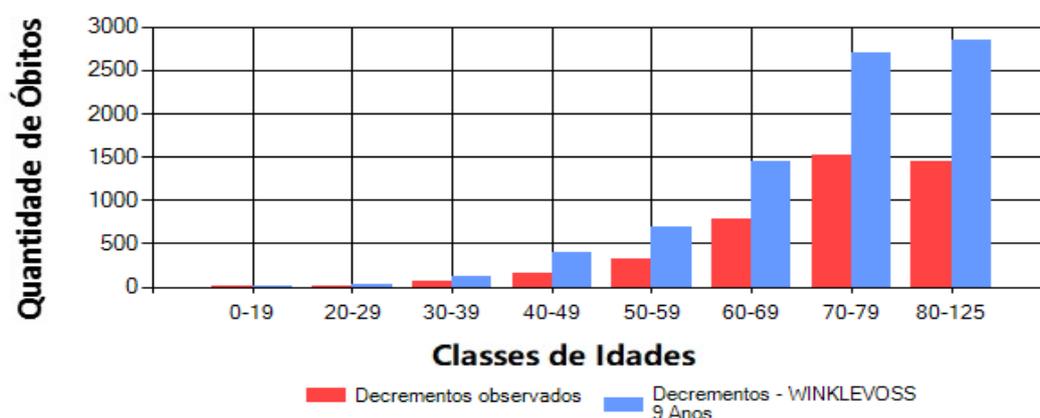
Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	Winklevoss +7 anos	Winklevoss desagravada em 11%
Amostra com intervalo de idade entre 21 e 93 anos	Hunter's agravada em 68%	Hunter's agravada em 83%,	Winklevoss desagravada em 10%

Fonte: Elaboração Própria (2020)

4.3.2.2 Aplicação dos Testes com Cálculo por Classes de Idades

Considerando a primeira entrada, com toda a população, foi estabelecida como classe inferior a amostra compreendida no intervalo de idade inferior 0 (zero) anos e idade superior 19 (dezenove) anos. Foram estabelecidas seis classes intermediárias com amplitude de dez anos. A classe superior foi estabelecida com idade inferior da classe de 80 (oitenta) anos e idade superior da classe 125 (cento e vinte e cinco) anos. Para esta entrada, não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de mortalidade a partir do teste do Qui-Quadrado. Com o teste de KS, houve aderência de 99 tábuas. A tábua de melhor aderência foi a tábua Winklevoss + 9 anos, conforme Gráfico 18.

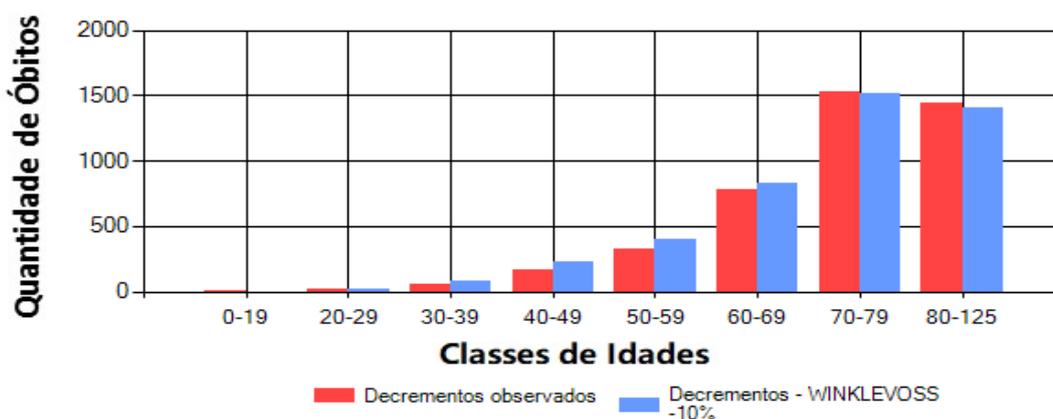
Gráfico 18 – Tábua de melhor aderência Teste KS para todas as idades



Fonte: Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Winklevoss desagravada em 10%, conforme ilustrado no Gráfico 19.

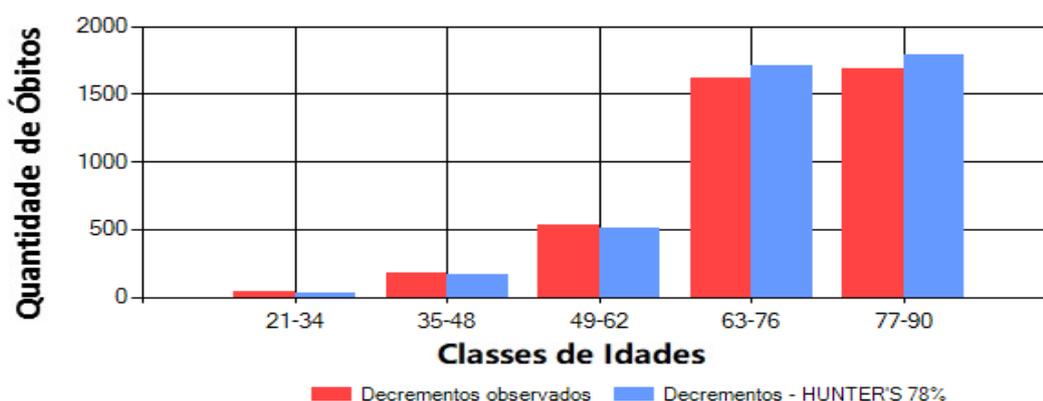
Gráfico 19 – Tábua de melhor aderência Teste DQM para todas as idades



Fonte: Elaboração Própria (2020)

Para a segunda entrada, com o intuito de se aproximar ao intervalo da entrada do Cálculo por Idade Exata, optou-se por testar a amostra de idades entre 21 e 90 anos. Esta segunda entrada corresponde a aproximadamente 90% dos dados populacionais. A classe inferior da amostra foi considerada compreendida no intervalo de idade inferior vinte e um anos e idade superior trinta e quatro anos. Foram estabelecidas três classes intermediárias com amplitude de quatorze anos. A classe superior foi estabelecida com idade inferior da classe de setenta e sete anos e idade superior da classe noventa anos. Porém, mediante o teste do Qui-Quadrado, ainda não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de mortalidade de inválidos. Utilizando-se o teste de Kolmogorov-Smirnov, houve aderência de 294 tábuas. Para esta divisão de classe de idades, a tábua de melhor aderência foi a Hunter's agravada em 78%, conforme ilustrada no Gráfico 20.

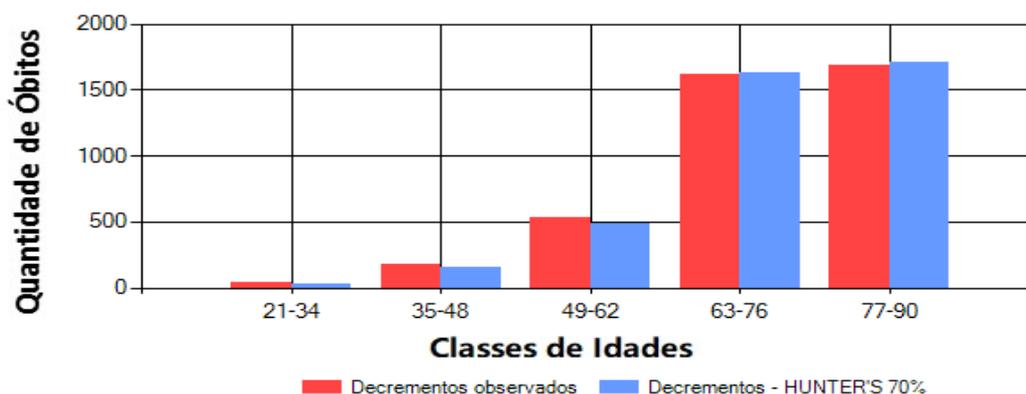
Gráfico 20 – Tábua de melhor aderência Teste KS para intervalo de 21 a 90 anos



Fonte: Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Hunter's agravada em 70%, conforme ilustrado no Gráfico 21.

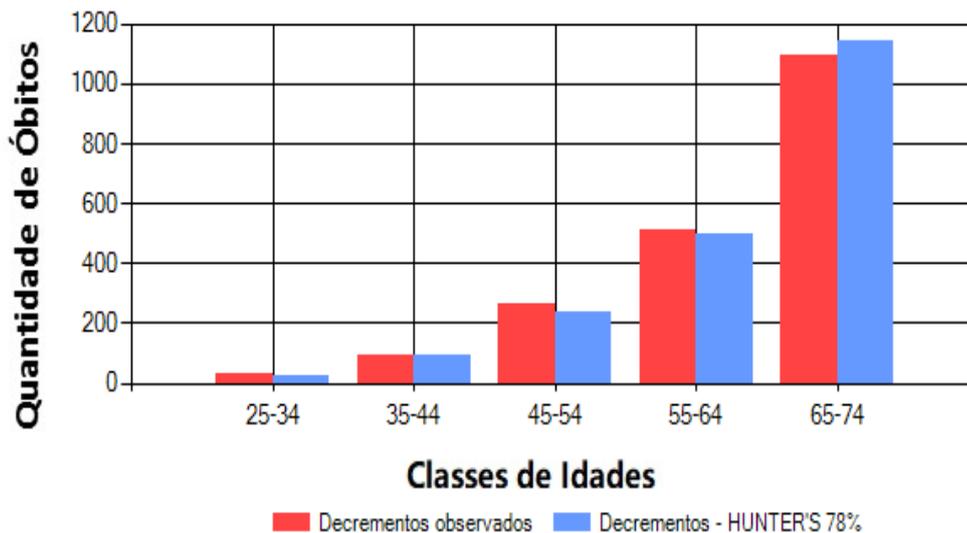
Gráfico 21 – Tábua de melhor aderência Teste DQM para intervalo de 21 a 90 anos



Fonte Elaboração Própria (2020)

Como não foi possível encontrar uma tábua aderente pelo Cálculo por Classes de Idades a partir do Teste do Qui-quadrado para as duas entradas propostas anteriormente, com o único objetivo de dar prosseguimento com a pesquisa forçando a aderência, decidiu-se experimentalmente testar outros intervalos de idade. A partir de diversos testes, obteve-se, empiricamente, que o intervalo de idade entre 25 e 74 anos, correspondente a 74% dos dados populacionais, possibilita que uma tábua seja aderente. Considerando, portanto, como classe inferior a amostra compreendida no intervalo de idade inferior vinte e cinco anos e idade superior trinta e quatro anos, foram estabelecidas três classes intermediárias com amplitude de dez anos. A classe superior foi estabelecida com idade inferior da classe de sessenta e cinco anos e idade superior da classe setenta e quatro anos. Para esta entrada específica, a única tábua aderente mediante o teste do Qui-quadrado foi a Hunter's, agravada de 78%, conforme Gráfico 22.

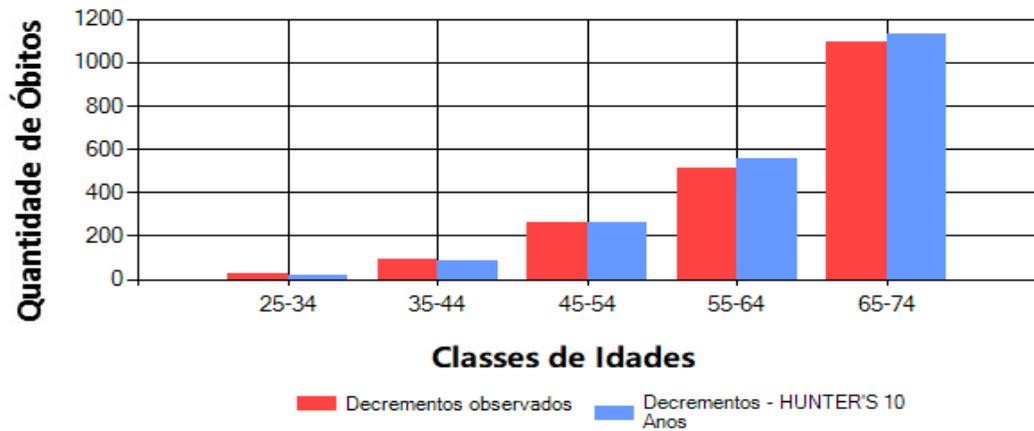
Gráfico 22 – Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 74 anos



Fonte: Elaboração Própria (2020)

Posteriormente, para fins de comparação, com a mesma entrada de intervalo entre 25 e 74 anos e mesma divisão de classes anterior foi realizado o teste de Kolmogorov-Smirnov. Então, foram obtidas 601 tábuas aderidas. A tábua com melhor aderência foi a tábua Hunter's 10 anos, conforme ilustrado no Gráfico 23.

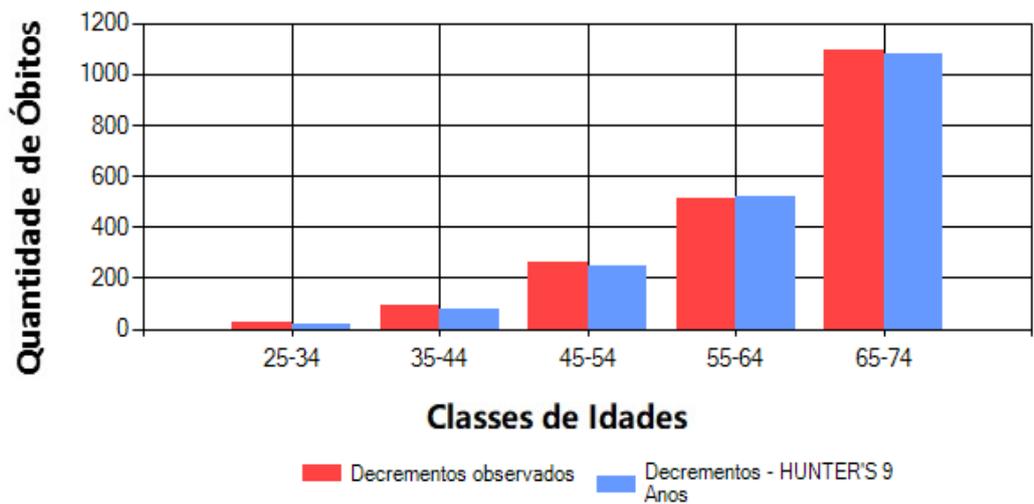
Gráfico 23 – Tábua de melhor aderência Teste KS para intervalo de 25 a 74 anos



Fonte: Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Hunter's agravada em 9 anos, conforme ilustrado no Gráfico 24.

Gráfico 24 – Tábua de melhor aderência Teste DQM para intervalo de 25 a 74 anos



Fonte Elaboração Própria (2020)

A tabela 14 resume a comparação da quantidade de tábuas aderidas em cada um dos testes, enquanto a Tabela 15 destaca as tábuas que apresentaram melhor aderência no cálculo por classe de idades.

Tabela 14 – Quantidade de Tábuas de Mortalidade de Inválidos aderidas a partir de cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	99	5.019
Amostra com intervalo de idade entre 21 e 90 anos	Não houve aderência	294	5.019
Amostra com intervalo de idade entre 25 e 74 anos	01	601	5.019

Fonte: Elaboração Própria (2020)

Tabela 15 – Tábuas de Mortalidade de Inválidos com melhor aderência para cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	Winklevoss + 9 anos	Winklevoss desagravada em 10%
Amostra com intervalo de idade entre 21 e 90 anos	Não houve aderência	Hunter's agravada em 78%	Hunter's agravada em 70%,
Amostra com intervalo de idade entre 25 e 74 anos	Hunter's, agravada de 78%	Hunter's 10 anos	Hunter's agravada em 9 anos

Fonte: Elaboração Própria (2020)

4.3.2.3 Análise dos Resultados para Mortalidade de Inválidos

Os testes de aderência para Mortalidade de Inválidos demonstraram que o teste Qui-quadrado não é o mais adequado para esta amostra populacional. Por outro lado, o teste de Kolmogorov-Smirnov demonstrou-se aplicável para toda a massa de dados. Assim como o Desvio Quadrático Médio.

Conforme explicado no item 4.3.1.3, devido às diferentes estatísticas de teste de cada tipo de teste de aderência, torna-se difícil comparar os resultados das tábuas melhor aderidas em cada teste aplicado de forma distinta. Durante esta pesquisa, não foi encontrada na literatura ou no mercado previdenciário uma maneira específica de realizar tal comparação.

Assim sendo, optou-se por uma análise qualitativa dos dados encontrados. O Quadro 5 resume a relação entre as tábuas melhores aderidas em cada teste e o seu resultado nos outros testes. Nesta tabela, a legenda “S” significa “sim”, isto é, aquela determinada tábua foi aderida no teste em questão. A legenda “N”, significa “não”, isto é, aquela determinada tábua não foi aderida naquele teste. A legenda “S*” indica que aquela tábua foi a de melhor aderência no referido teste.

Após a análise qualitativa do Quadro 5, verifica-se que para este banco de dados de mortalidade de inválidos no ano de 2015, a tábua atuarial que obteve a melhor aderência foi a tábua Hunter’s agravada em 78%. O teste de Kolmogorov-Smirnov foi o teste com mais robustez, sendo aplicável a qualquer amostra populacional.

Quadro 5 – Tábuas de Mortalidade de Inválidos Melhor Aderidas para Cada Teste

Teste	Qui-Quadrado					Kolmogorov-Smirnov					Desvio Quadrático Médio				
	Idade Exata		Classes de Idades			Idade Exata		Classes de Idades			Idade Exata		Classes de Idades		
Tábua \ Amostra	Toda População	21 a 93 anos	Toda População	21 a 90 anos	25 a 74 anos	Toda População	21 a 93 anos	Toda População	21 a 90 anos	25 a 74 anos	Toda População	21 a 93 anos	Toda População	21 a 90 anos	25 a 74 anos
Hunter's + 9 anos	N	N	N	N	N	N	N	N	N	N	S	S	S	S	S
Hunter's + 10 anos	N	N	N	N	N	N	N	N	N	N	S	S	S	S	S
Hunter's + 68%	N	S*	N	N	N	N	N	N	N	N	S	S	S		
Hunter's + 70%	N	S	N	N	N	N	N	N	N	N	S	S	S	S	S
Hunter's + 78%	N	N	N	N	S*	S	S	N	S*	N	S	S	S	S	S
Hunter's + 83%	N	N	N	N	N	N	S*	N	N	N	S	S	S	S	S
Winklevoss + 7 anos	N	N	N	N	N	S*	N	N	N	N	S	S	S	S	S
Winklevoss + 9 anos	N	N	N	N	N	N	N	S*	N	N	S	S	S	S	S
Winklevoss – 10%	N	N	N	N	N	N	N	N	N	N	S	S*	S	S	S
Winklevoss – 11%	N	N	N	N	N	N	N	N	N	N	S*	S	S	S	S

Fonte: Elaboração Própria (2020)

4.3.3 Resultado de Entrada em Invalidez das Forças Armadas

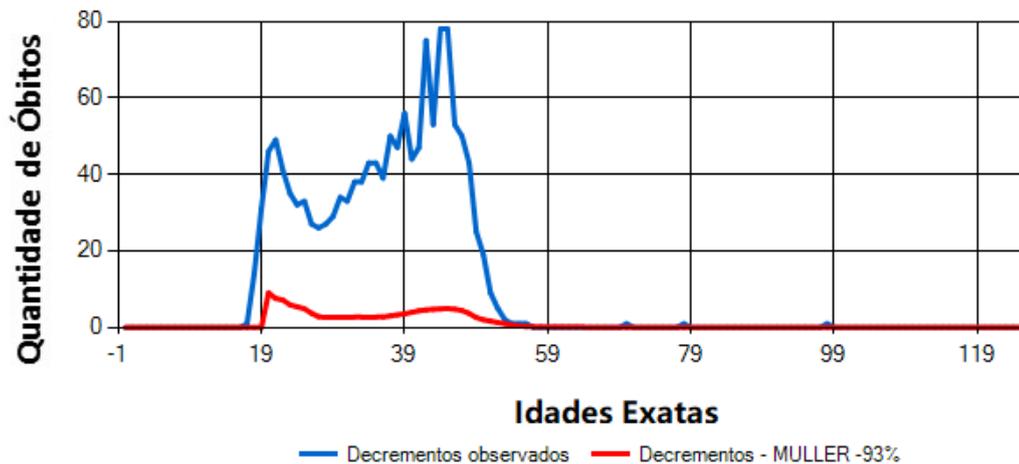
4.3.3.1 Aplicação dos Testes com Cálculo da Idade Exata

Para o cálculo da entrada em invalidez nas Forças Armadas foram realizados os testes de aderência para três entradas possíveis. A primeira entrada de dados de entrada em invalidez considerando toda a população, isto é, todas as idades entre zero e cento e vinte e cinco anos. A segunda entrada representa 99% dos dados da população, considerando o intervalo entre 19 e 90 anos de idade; enquanto a terceira entrada de dados considera uma amostra representativa desta população, com o intervalo de idades entre 25 e 41 anos, que engloba apenas 35% dos dados populacionais.

Considerando a primeira entrada, com toda a população, não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de entrada em invalidez, a partir do teste do Qui-Quadrado.

Com o teste de KS, houve aderência de 24 tábuas. A tábua de melhor aderência foi a tábua Muller desagravada em 93%, conforme Gráfico 25.

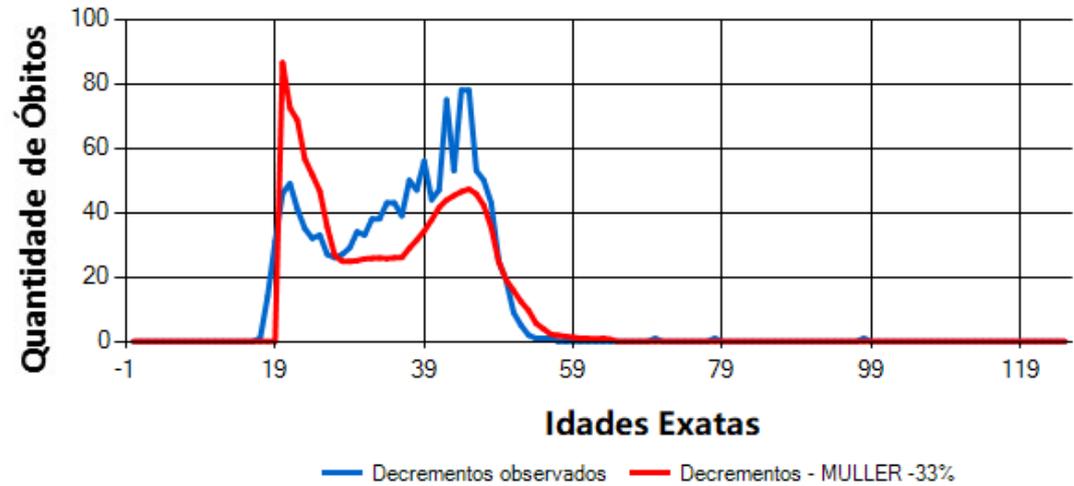
Gráfico 25 – Tábua de melhor aderência Teste KS para todas as idades



Fonte Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Muller desagravada em 33%, conforme ilustrado no Gráfico 26.

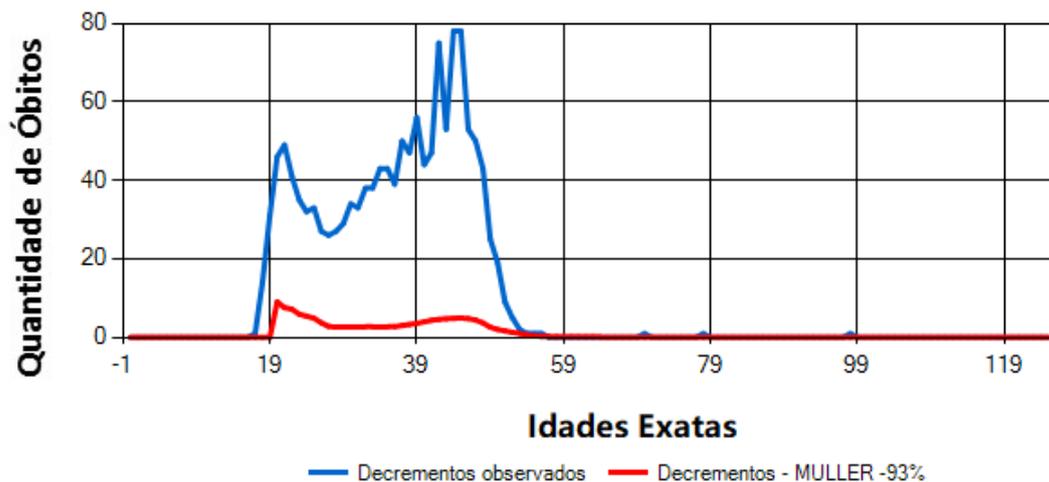
Gráfico 26 – Tábua de melhor aderência Teste DQM para todas as idades



Fonte Elaboração Própria (2020)

Para a segunda entrada, com o intervalo de idades entre 19 e 90 anos, nenhuma tábua aderiu à curva de entrada em invalidez observada a partir do teste do Qui-Quadrado. Com o teste de KS, houve aderência de 23 tábuas e a tábua de melhor aderência novamente foi a tábua Muller desagravada em 93%, conforme Gráfico 27.

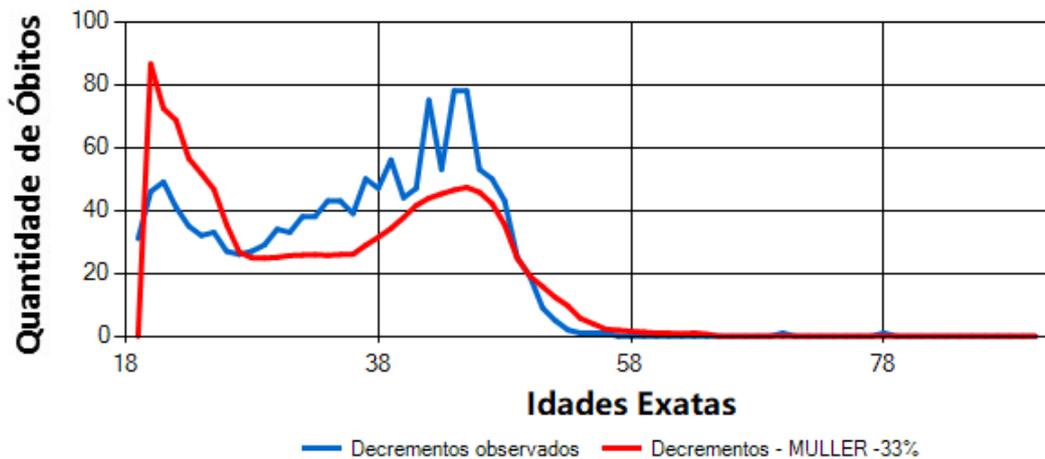
Gráfico 27 – Tábua de melhor aderência Teste KS intervalo de 19 a 90 anos



Fonte Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Muller desagravada em 33%, conforme ilustrado no Gráfico 28.

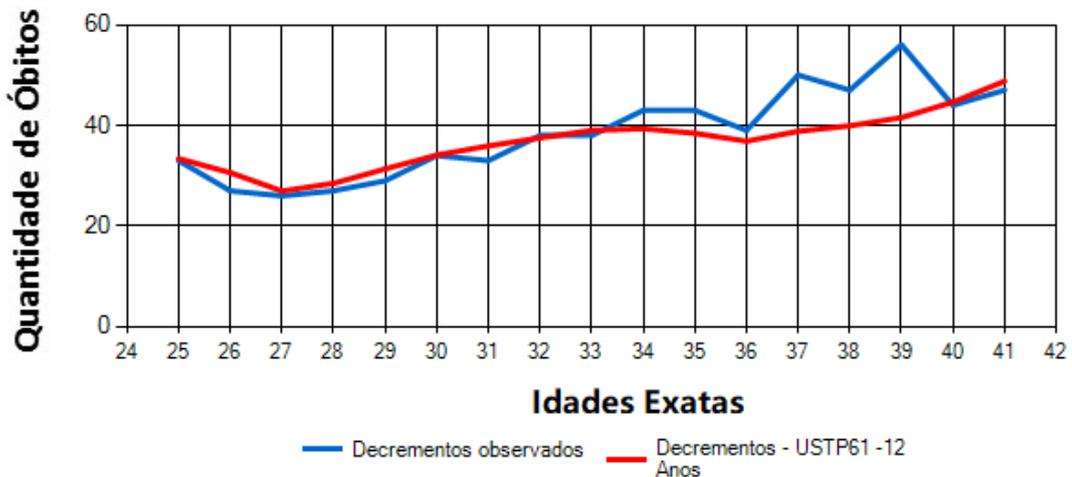
Gráfico 28 – Tábua de melhor aderência Teste DQM para intervalo de 19 a 90 anos



Fonte Elaboração Própria (2020)

Considerando a terceira entrada, com o intervalo de idades entre 25 e 41 anos, 34 tábuas aderiram à curva de entrada em invalidez observada a partir do teste do Qui-Quadrado. A tábua de melhor aderência foi USTP61 desagravada em 12 anos, conforme Gráfico 29.

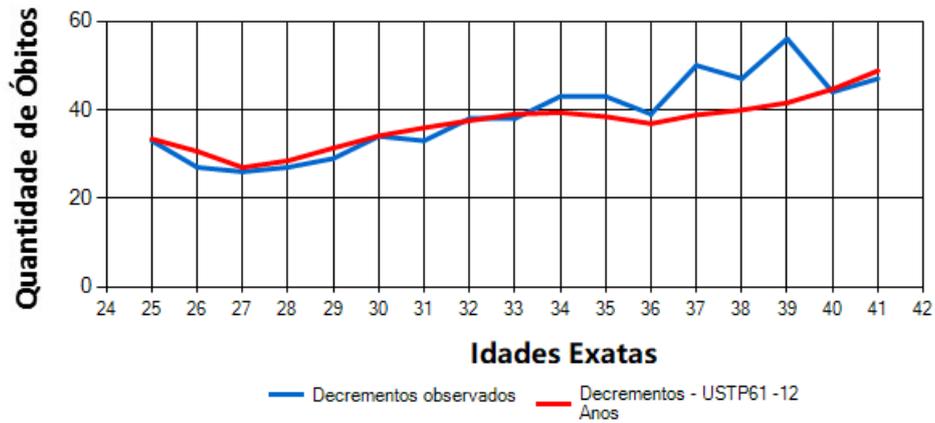
Gráfico 29 – Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 41 anos



Fonte Elaboração Própria (2020)

Com o teste de KS, para esta terceira entrada, houve aderência de 839 tábuas, sendo a tábua de melhor aderência foi a tábua USTP61 desagravada em idade por 12 anos, conforme Gráfico 30.

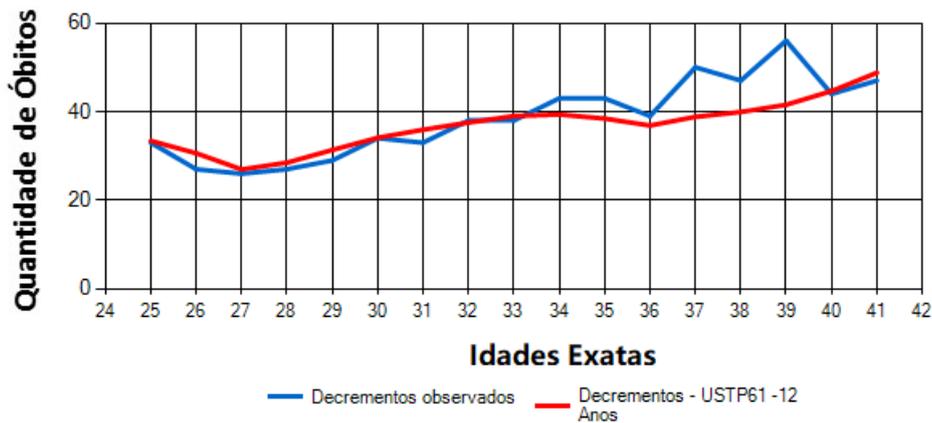
Gráfico 30 – Tábua de melhor aderência Teste KS para intervalo de 25 a 41 anos



Fonte Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência também foi a tábua USTP61 desagravada em 12 anos, conforme ilustrado no Gráfico 31.

Gráfico 31 – Tábua de melhor aderência Teste DQM para intervalo de 25 a 41 anos



Fonte Elaboração Própria (2020)

A tabela 16 resume a comparação da quantidade de tábuas aderidas em cada um do teste, enquanto a Tabela 17 destaca as tábuas que apresentaram melhor aderência no cálculo para idade exata.

Tabela 16 – Quantidade de Tábuas de Entrada em Invalidez aderidas a partir de cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	24	5.019
Amostra com intervalo de idade entre 19 e 90 anos	Não houve aderência	23	5.019
Amostra com intervalo de idade entre 25 e 41 anos	34	839	5.019

Fonte: Elaboração Própria (2020)

Tabela 17 –Tábuas de Entrada em Invalidez com melhor aderência a partir de cada teste

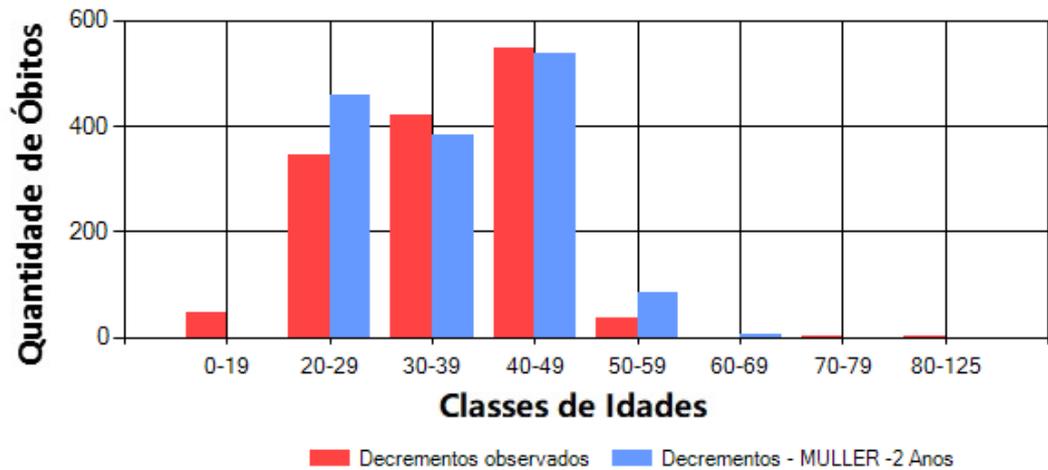
Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	Muller desagravada em 93%	Muller desagravada em 33%
Amostra com intervalo de idade entre 19 e 90 anos	Não houve aderência	Muller desagravada em 93%	Muller desagravada em 33%
Amostra com intervalo de idade entre 25 e 41 anos	USTP61 desagravada em idade por 12 anos	USTP61 desagravada em idade por 12 anos	USTP61 desagravada em idade por 12 anos

Fonte: Elaboração Própria (2020)

4.3.3.2 Aplicação dos Testes com Cálculo por Classes de Idades

Considerando a primeira entrada, com toda a população, foi estabelecida como classe inferior a amostra compreendida no intervalo de idade inferior 0 (zero) anos e idade superior 19 (dezenove) anos. Foram estabelecidas seis classes intermediárias com amplitude de dez anos. A classe superior foi estabelecida com idade inferior da classe de 80 (oitenta) anos e idade superior da classe 125 (cento e vinte e cinco) anos. Para esta entrada, não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de mortalidade a partir do teste do Qui-Quadrado. Com o teste de KS, houve aderência de 48 tábuas. A tábua de melhor aderência foi a tábua Muller desagravada em idade de dois anos, conforme Gráfico 32.

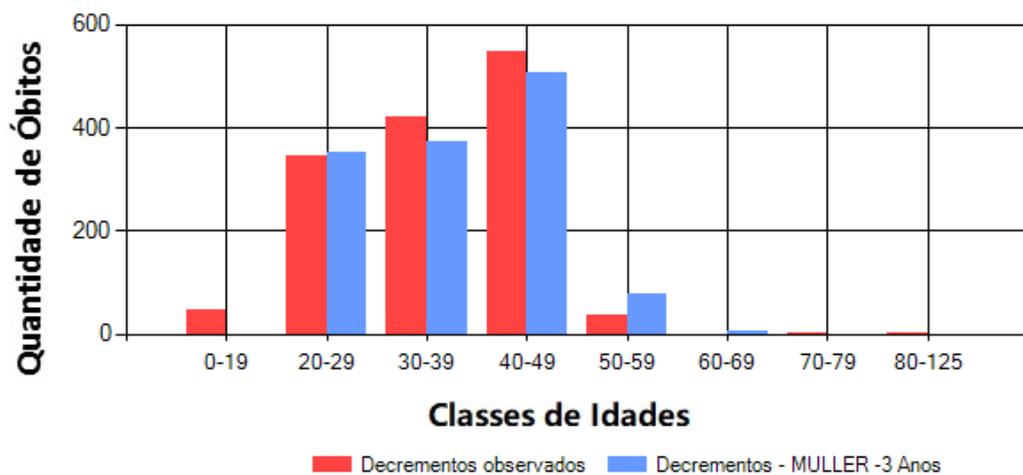
Gráfico 32 – Tábua de melhor aderência Teste KS para todas as idades



Fonte Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua Muller desagradada em idade de três anos, conforme ilustrado no Gráfico 33.

Gráfico 33 – Tábua de melhor aderência Teste DQM para todas as

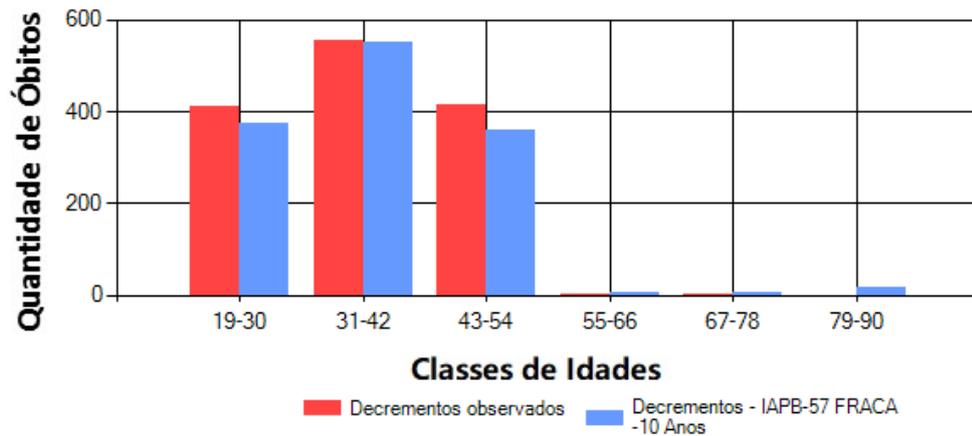


Fonte Elaboração Própria (2020)

Para a segunda entrada, optou-se por testar a amostra de idades igual ao intervalo da segunda entrada do Cálculo por Idade Exata, isto é, entre 19 e 90 anos. Esta segunda entrada corresponde a 99% dos dados populacionais. A classe inferior da amostra foi considerada compreendida no intervalo de idade inferior dezoito e idade superior trinta anos. Foram estabelecidas seis classes intermediárias com amplitude de doze anos. A classe superior foi

estabelecida com idade inferior da classe de setenta e nove anos e idade superior da classe noventa anos. Porém, mediante o teste do Qui-Quadrado, ainda não houve nenhuma tábua atuarial que tivesse aderência aos dados observados de entrada em invalidez. Utilizando-se o teste de Kolmogorov-Smirnov, houve aderência de 54 tábuas. Para esta divisão de classe de idades, a tábua de melhor aderência foi a IAPB-57 Fraca desagravada por idade em 10 anos, conforme ilustrada no Gráfico 34.

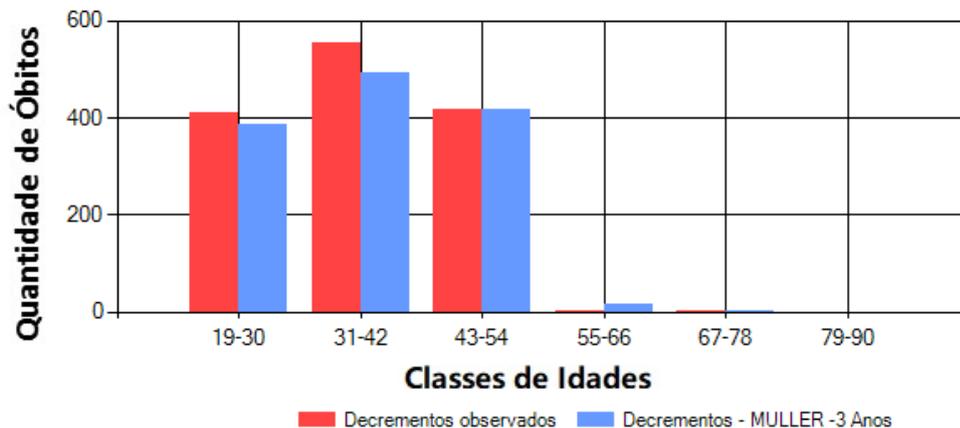
Gráfico 34 - Tábua de melhor aderência Teste KS para intervalo de 19 a 90 anos



Fonte Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi, novamente, a tábua Muller desagravada em idade de três anos conforme ilustrado no Gráfico 35.

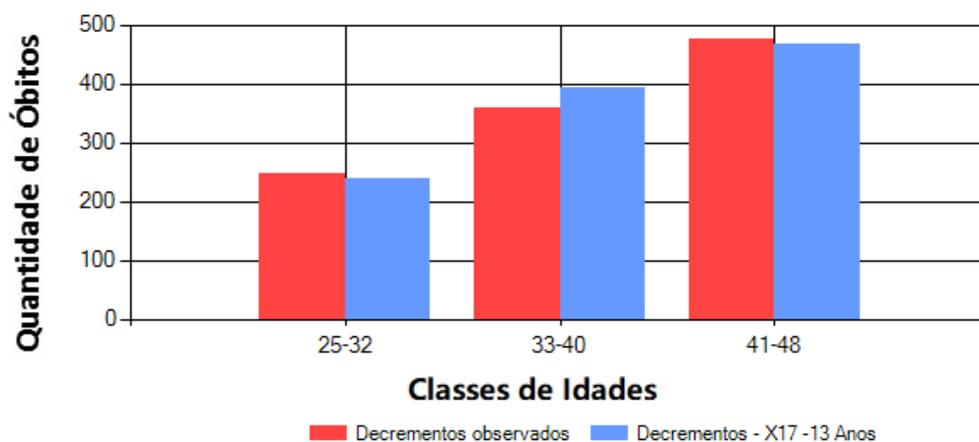
Gráfico 35 - Tábua de melhor aderência Teste DQM para intervalo de 19 a 90 anos



Fonte Elaboração Própria (2020)

Como não foi possível encontrar uma tábua aderente pelo Cálculo por Classes de Idades a partir do Teste do Qui-quadrado para as duas entradas propostas anteriormente, com o único objetivo de dar prosseguimento com a pesquisa forçando a aderência, decidiu-se experimentalmente testar outros intervalos de idade. Visando se aproximar ao espaço etário da terceira entrada do Cálculo por Idade Exata, optou-se por testar amostras de idades nas faixas próximas ao intervalo de 25 a 41 anos. A partir de diversos testes, obteve-se, empiricamente, que o intervalo de idade entre 25 e 48 anos, correspondente a aproximadamente 47% dos dados populacionais, possibilita que uma única tábua seja aderente pelo método Qui-quadrado. Considerando, portanto, como classe inferior a amostra compreendida no intervalo de idade inferior vinte e cinco anos e idade superior trinta e dois anos, foram estabelecidas três classes intermediárias com amplitude de oito anos. A classe superior foi estabelecida com idade inferior da classe de quarenta e um anos e idade superior da classe de quarenta e oito anos. Para esta entrada específica, a única tábua aderente mediante o teste do Qui-quadrado foi a tábua X17 desagravada em idade de 13 anos, conforme o Gráfico 36.

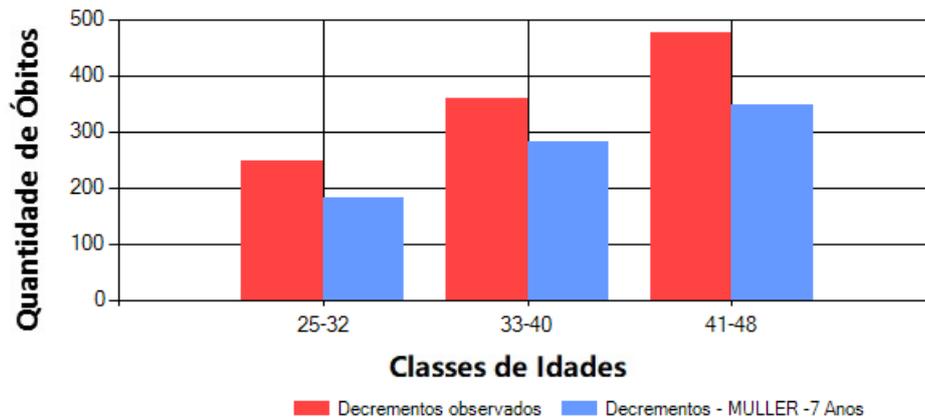
Gráfico 36 - Tábua de melhor aderência Teste Qui-quadrado para intervalo de 25 a 48 anos



Fonte Elaboração Própria (2020)

Posteriormente, para fins de comparação, com a mesma entrada de intervalo entre 25 e 48 anos e mesma divisão de classes anterior foi realizado o teste de Kolmogorov-Smirnov. Então, foram obtidas 563 tábuas aderidas. A tábua com melhor aderência foi a tábua Muller desagravada por idade em 7 anos, conforme ilustrado no Gráfico 37.

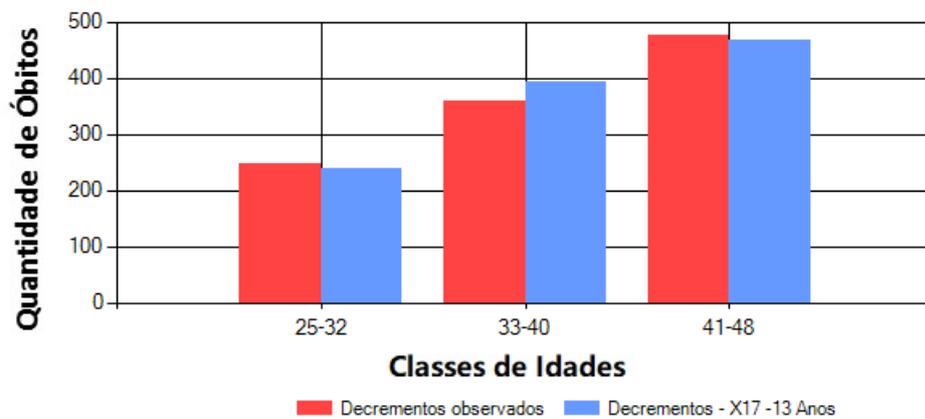
Gráfico 37 - Tábua de melhor aderência Teste KS para intervalo de 25 a 48 anos



Fonte Elaboração Própria (2020)

Foi realizado, ainda, a partir da mesma entrada anterior, o teste do desvio quadrático médio. A tábua com melhor aderência foi a tábua X17 desagradada em 13 anos, conforme ilustrado no Gráfico 38.

Gráfico 38 - Tábua de melhor aderência Teste DQM para intervalo de 25 a 48 anos



Fonte Elaboração Própria (2020)

A tabela 18 resume a comparação da quantidade de tábuas aderidas em cada um dos testes, enquanto a Tabela 19 destaca as tábuas que apresentaram melhor aderência no cálculo por classe de idades.

Tabela 18 – Quantidade de Tábuas de Mortalidade de Inválidos aderidas a partir de cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrado Médio
Toda a População	Não houve aderência	48	5.019
Amostra com intervalo de idade entre 19 e 90 anos	Não houve aderência	54	5.019
Amostra com intervalo de idade entre 25 e 48 anos	01	563	5.019

Fonte: Elaboração Própria (2020)

Tabela 19 – Tábuas de Mortalidade de Inválidos com melhor aderência para cada teste

Dados \ Teste	Qui-quadrado	Kolmogorov-Smirnov	Desvio Quadrático Médio
Toda a População	Não houve aderência	Muller desagravada em 2 anos	Muller desagravada em 3 anos
Amostra com intervalo de idade entre 19 e 90 anos	Não houve aderência	IAPB-57 Fraca desagravada em 10 anos	Muller desagravada em 3 anos
Amostra com intervalo de idade entre 25 e 48 anos	X17 desagravada em 13 anos	Muller desagravada em 7 anos	X17 desagravada em idade de 13 anos

Fonte: Elaboração Própria (2020)

4.3.3.3 Análise dos Resultados para Entrada em Invalidez

Os testes de aderência para Entrada e Invalidez demonstraram que o teste Qui-quadrado não é o mais adequado para esta amostra populacional. Para que alguma tábua fosse aderente a partir deste teste, fez-se necessário reduzir o espaço amostral a menos de 50% da população total. Desta forma, é possível que a grande perda da massa de dados gere imprecisão nos resultados. Por outro lado, o teste de Kolmogorov-Smirnov demonstrou-se aplicável para toda a massa de dados. Assim como o Desvio Quadrático Médio.

Conforme explicado no item 4.3.1.3, devido às diferentes estatísticas de teste de cada tipo de teste de aderência, torna-se difícil comparar os resultados das tábuas melhor aderidas

em cada teste aplicado de forma distinta. Durante esta pesquisa, não foi encontrada na literatura ou no mercado previdenciário uma maneira específica de realizar tal comparação. Assim sendo, optou-se por uma análise qualitativa dos dados encontrados. O Quadro 6 resume a relação entre as tábuas melhores aderidas em cada teste e o seu resultado nos outros testes. Nesta tabela, a legenda “S” significa “sim”, isto é, aquela determinada tábua foi aderida no teste em questão. A legenda “N”, significa “não”, isto é, aquela determinada tábua não foi aderida naquele teste. A legenda “S*” indica que aquela tábua foi a de melhor aderência no referido teste.

Após a análise qualitativa do Quadros 6, verifica-se que para este banco de dados de entrada em invalidez no ano de 2015, a melhor tábua obtida foi a tábua Muller, desagravado em 93%, a partir do teste de Kolmogorov-Smirnov.

Quadro 6 – Tábuas de Entrada em Invalidez Melhor Aderidas para Cada Teste

Teste	Qui-Quadrado						Kolmogorov-Smirnov						Desvio Quadrático Médio					
Tipo	Idade Exata			Classes de Idades			Idade Exata			Classes de Idades			Idade Exata			Classes de Idades		
Tábua \ Amostra	Toda População	19 a 90 anos	25 a 41 anos	Toda População	19 a 90 anos	25 a 48 anos	Toda População	19 a 90 anos	25 a 41 anos	Toda População	19 a 90 anos	25 a 48 anos	Toda População	19 a 90 anos	25 a 41 anos	Toda População	19 a 90 anos	25 a 48 anos
Muller -2 anos	N	N	N	N	N	N	N	N	N	S*	S	N	S	S	S	S	S	S
Muller -3 anos	N	N	N	N	N	N	N	N	N	S	S	N	S	S	S	S*	S*	S
Muller -7 anos	N	N	N	N	N	N	N	N	N	N	N	S*	S	S	S	S	S	S
Muller -33% 33%	N	N	N	N	N	N	N	N	N	N	N	S	S*	S*	S	S	S	S
Muller -93% 93%	N	N	N	N	N	N	S*	S*	S	S	S	S	S	S	S	S	S	S
IAPB-57 Fraca – 10 anos	N	N	N	N	N	N	N	N	N	S	S*	N	S	S	S	S	S	S
USTP61 – 12 anos	N	N	S*	N	N	N	N	N	S*	N	N	S	S	S	S*	S	S	S
X17 – 13 anos	N	N	N	N	N	S*	N	N	N	N	N	S	S	S	S	S	S	S*

Fonte: Elaboração Própria (2020)

4.3.4 Considerações finais sobre a Análise de Resultados

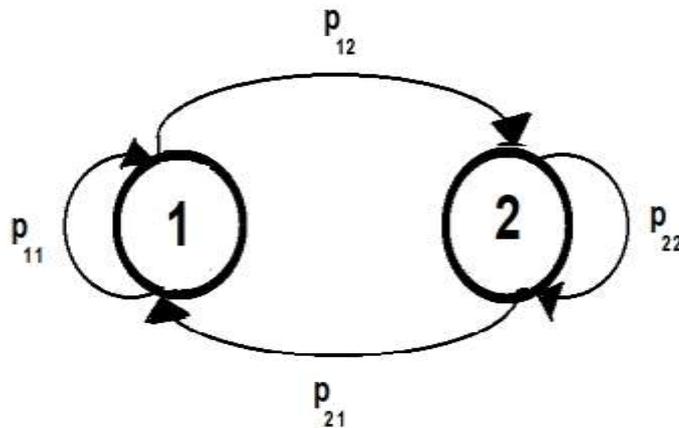
Durante a referida pesquisa, não foi encontrada na literatura ou no mercado previdenciário um método quantitativo consagrado de comparação das diferentes estatísticas de teste dos diversos testes de aderência. Assim sendo foi realizada uma análise qualitativa que concluiu que o teste de aderência Qui-quadrado não seria o mais adequado para o banco de dados utilizado referente à amostra populacional FFAA. Esse teste restringe as tábuas atuariais satisfatórias de tal modo que, em alguns dos testes, para que alguma tábua atuarial fosse aderente, era preciso reduzir o espaço amostral para apenas 1/3 da população em estudo. Já o teste do Desvio Quadrático Médio possibilita uma ordenação das tábuas testadas, o que não necessariamente reflete que a tábua mais aderente seja adequada ao propósito em questão. O teste de Kolmogorov-Smirnov, por sua vez, demonstrou robustez, uma vez que um grande número de tábuas atuariais apresentou aderência, considerando um espaço amostral mais amplo da população das FFAA e com o resultado da estatística de teste satisfatório.

5. ESTUDO MARKOVIANO SOBRE ESTADOS BIOMÉTRICOS

Um processo de Markov é definido como um processo estocástico no qual apenas o valor atual da variável aleatória é relevante na previsão futura. Os valores anteriores, além do mais recente, não afetam, portanto, os valores subsequentes (SPITZER, 1970; D'ARCY, 1989). Processos estocásticos, por sua vez, são modelos matemáticos de fenômenos aleatórios que evoluem de acordo com a dinâmica prescrita (SERFOZO, 2009) O pressuposto das Cadeias de Markov é, portanto, a independência do presente em relação ao passado. Desta forma, o conhecimento do estado em qualquer instante seria o suficiente para prever o futuro (GOMES e GOMES, 2019).

Segundo Gomes e Ribeiro (2014), um grafo é um elemento de abstração que representa um conjunto não vazio e finito de elementos, chamados nós, e suas relações de interdependência, denominadas arestas. Um processo markoviano pode ser representado por meio de grafos, onde os nós representam os estados e as arestas indicam as probabilidades de transição entre esses estados (RESTUM, 1999). A Figura 9 ilustra um processo de Markov de dois estados.

Figura 9 – Grafo de um processo de Markov de dois estados.



Fonte: Elaboração Própria (2020)

Para Seneta (1996), outra forma de representar as probabilidades de transição em n etapas numa Cadeia de Markov é por meio de uma matriz de transição, onde ocorre a passagem do estado representado na coluna lateral para o estado representado pela linha superior. A Figura 10 representa a probabilidade de se mover do estado “ i ” para o estado “ j ” em uma determinada etapa “ n ”.

Seja P uma matriz $k \times k$ com elementos $\{P_{ij} : i, j = 1, \dots, k\}$. Um processo aleatório (X_0, X_1, \dots) com espaço de estado finito $S = \{s_1, s_2, \dots, s_k\}$ é dito ser uma Cadeia de Markov com matriz de transição P , se $\forall n$, todo $i, j = \{1, \dots, k\}$ e todo $i_0, \dots, i_{n+1} = \{1, \dots, k\}$, tem-se:

$$P(X_{n+1} = s_j | X_0 = s_{i_0}, X_1 = s_{i_1}, \dots, X_{n-1} = s_{i_{n-1}}, X_n = s_i) = P(X_{n+1} = s_j | X_n = s_i) = P_{ij}$$

Figura 10 – Matriz de transição em n etapas da Cadeia de Markov.

$$\begin{array}{c}
 \text{Estado} \\
 1 \\
 2 \\
 \vdots \\
 i
 \end{array}
 P^n =
 \begin{bmatrix}
 \begin{array}{c}
 1 \\
 (n) \\
 p_{11}
 \end{array}
 &
 \begin{array}{c}
 2 \\
 (n) \\
 p_{12}
 \end{array}
 &
 \dots
 &
 \begin{array}{c}
 J \\
 (n) \\
 p_{1j}
 \end{array}
 &
 \dots \\
 \begin{array}{c}
 p_{21} \\
 p_{22} \\
 \vdots \\
 p_{i1}
 \end{array}
 &
 \begin{array}{c}
 p_{22} \\
 p_{22} \\
 \vdots \\
 p_{i2}
 \end{array}
 &
 \dots
 &
 \begin{array}{c}
 p_{2j} \\
 p_{2j} \\
 \vdots \\
 p_{ij}
 \end{array}
 &
 \dots \\
 \vdots & \vdots & \ddots & \vdots & \ddots
 \end{bmatrix}$$

Fonte: Seneta (1996)

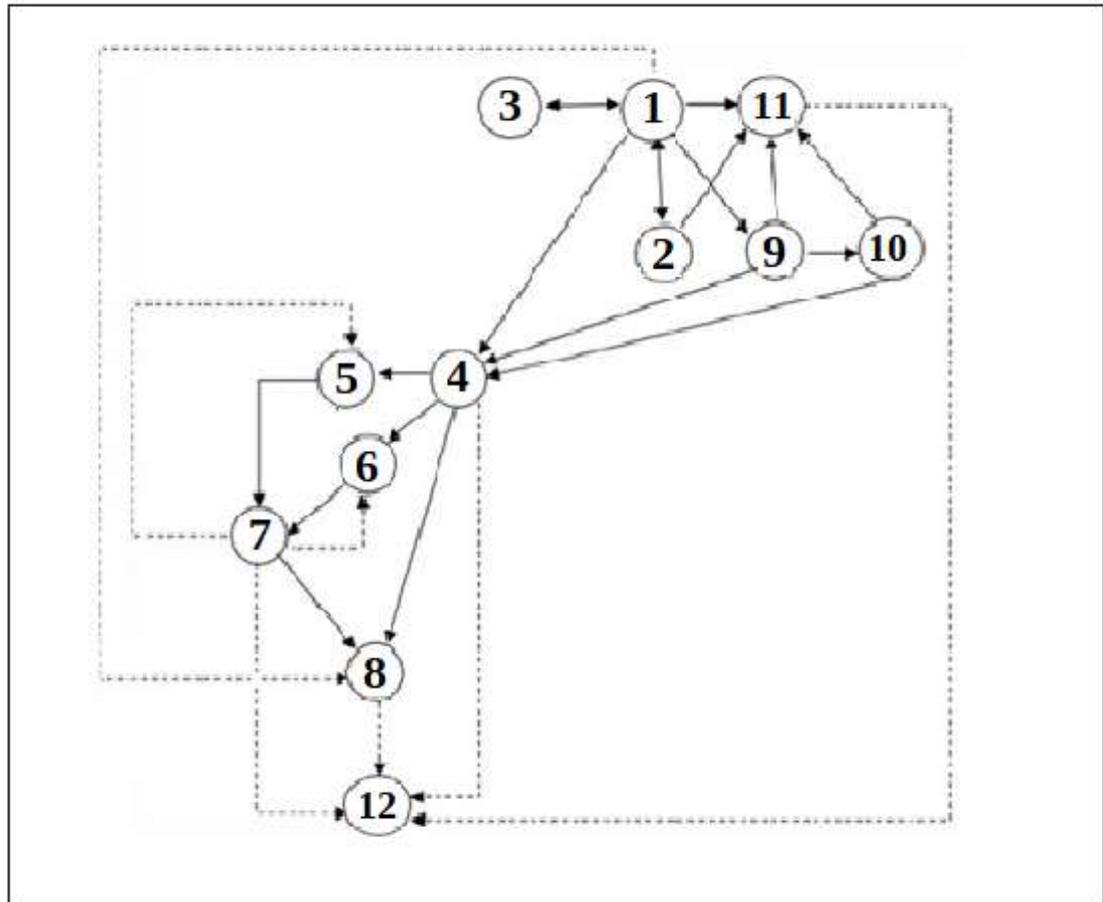
Santos (2018) apresenta 12 estados biométricos no fluxo de carreira de um militar que possa vir a gerar pensão, são eles:

- 1) Militar Ativo
- 2) Invalidez / Reforma (temporária ou não)
- 3) Demissão do Serviço Ativo (saída do sistema)
- 4) Casamento
- 5) Divórcio
- 6) Viuvez
- 7) Novo Casamento
- 8) Paternidade
- 9) Reserva Remunerada
- 10) Reserva por Idade
- 11) Falecimento
- 12) Geração de Pensão

O grafo apresentado na Figura 11 mostra que um indivíduo ativo está sujeito a probabilidades conjuntas de morrer ou se invalidar. Cada vértice do grafo estabelece um estado no qual o militar pode se encontrar (ativo, reserva, casado, reformado, com filhos etc.). Costuma-se dizer, com isso, que esse indivíduo está submetido a multidecrementos probabilísticos. Assim, a Figura 11 mostra que:

- ⇒ Um militar ativo (estado 1) pode invalidar-se (estado 2), casar-se (estado 4), ir para a reserva (estado 9), pedir demissão (estado 3), ser pai (estado 8) ou falecer (estado 11).
- ⇒ Um militar inválido (estado 2) pode voltar para a ativa (estado 1), casar-se (estado 2) ou falecer (estado 11).
- ⇒ Caso o militar peça demissão (estado 3), ele sai do sistema e não gera pensão.
- ⇒ Um militar casado (estado 4) pode divorciar-se (estado 5), tornar-se viúvo (estado 6) ou tornar-se pai (estado 8). Uma vez casado, a linha pontilhada mostra que o militar pode gerar uma pensão (estado 12).
- ⇒ Um militar divorciado (estado 5) pode contrair novo matrimônio (estado 7).
- ⇒ Um militar viúvo (estado 6) pode contrair um novo matrimônio (estado 7).
- ⇒ Um militar que contraiu um novo matrimônio (estado 7) pode divorciar-se (estado 5), tornar-se viúvo (estado 6) ou tornar-se pai (estado 8).
- ⇒ Um militar que se tornou pai (estado 8) pode gerar uma pensão (estado 12), conforme indica a linha pontilhada.
- ⇒ Um militar na Reserva Remunerada (estado 9) pode contrair matrimônio (estado 4), ir para a reserva por idade (estado 10) ou falecer (estado 11).
- ⇒ Um militar na reserva por idade (estado 10) pode casar-se (estado 4) ou falecer (estado 11).
- ⇒ A geração de pensão (estado 12) pode ser gerada pelo casamento (estado 4), por um novo casamento (estado 7) e pela paternidade (estado 8), uma vez que o militar tenha falecido (estado 12).

Figura 11 – Grafo de doze estados biométricos e sociais.



Fonte: Adaptado de Santos (2018)

De acordo com Hillier e Lieberman (2013), as propriedades duradouras de um processo markoviano dependem das características de seus estados e de sua matriz de transição. Assim, os estados da Cadeia de Markov poderiam ser classificados em:

- Acessível, a partir de outro estado “y”, quando for possível eventualmente para o sistema entrar naquele estado quando se inicia no estado “y”. Se um estado “x” é acessível a partir do estado “y” e vice-versa, diz-se que os estados “x” e “y” se comunicam. Dois estados que se comunicam são ditos como de mesma classe. Quando todos os estados de uma determinada Cadeia de Markov se comunicam esta é dita irredutível;
- Transiente, quando após atingir este estado existir a possibilidade de o processo nunca retornar a este estado novamente;
- Recorrente, quando após atingir este estado existir a certeza de o processo retornar a este estado novamente;

- Absorvente, quando após atingir este estado existir a certeza de o processo nunca deixar este estado novamente;
- Aperiódico, quando, havendo dois números consecutivos “ t ” e “ $t+1$ ”, o processo possa se encontrar no estado “ x ” nos instantes “ t ” e “ $t+1$ ”; e
- Ergódico, quando o estado for simultaneamente recorrente e aperiódico, num conjunto de estados finitos. Uma cadeia de Markov é dita ergódica se todos os seus estados forem estados ergódicos.

Segundo as equações de Chapman-Kolmogorov, a matriz de probabilidade de transição em “ n ” etapas pode ser obtida pelo cálculo da n -ésima potência da matriz de transição em uma determinada etapa. Quando uma Cadeia de Markov é ergódica e irredutível, existe uma probabilidade limite, após um número suficientemente grande de transições entre os estados, de que o sistema se encontrará em um determinado estado “ x ”, não dependendo mais do estado inicial do mesmo. Chama-se de probabilidade de estado estável esta condição na qual uma Cadeia de Markov atinge uma matriz de transição inalterada. Cabe ressaltar que a probabilidade de estado estável não significa que um determinado sistema não realize mais transições entre os seus estados, mas sim que as probabilidades envolvidas no processo não sofrem mais alterações.

A partir dos Estados Biométricos da Figura 11, obtêm-se a matriz de transição que representa a relação entre as alterações dos estados, onde $p_{i,j}$ indica a probabilidade de mudança do estado “ i ” para o “ j ”. Quando um determinado estado “ j ” não é acessível a partir do estado i , têm-se $p_{i,j} = 0$ (zero). A Figura 12 ilustra a referida matriz de transição com 12 estados biométricos.

Figura 12 – Matriz de Transição de doze estados biométricos e sociais.

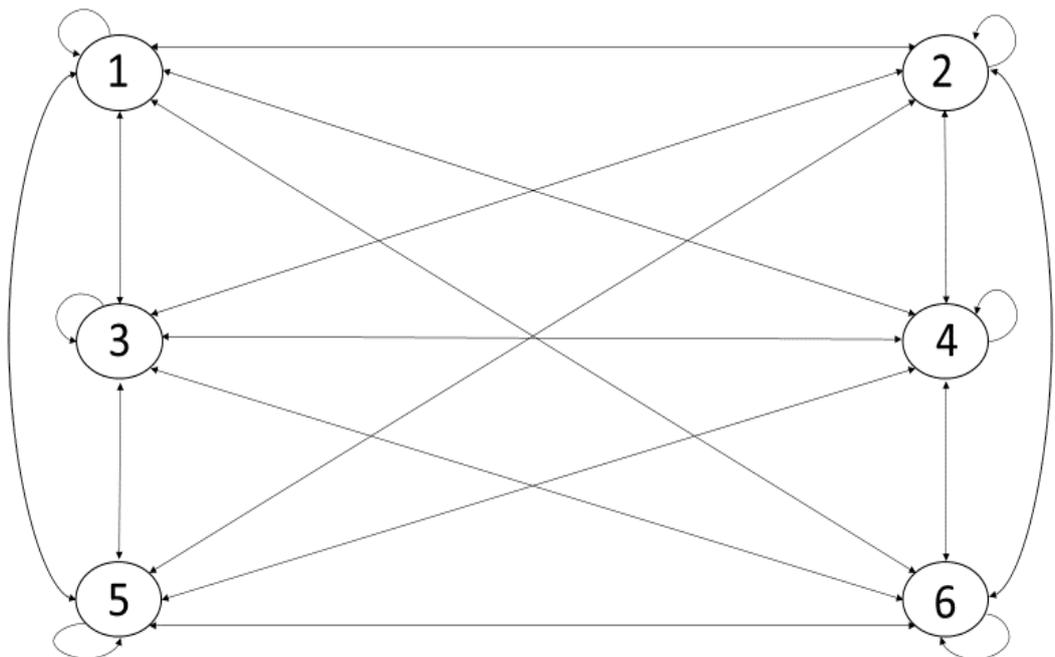
Estados Biométricos	1	2	3	4	5	6	7	8	9	10	11	12
1	0	$p_{1.2}$	$p_{1.3}$	$p_{1.4}$	0	0	0	$p_{1.8}$	$p_{1.9}$	0	$p_{1.10}$	0
2	$p_{2.1}$	0	0	0	0	0	0	0	0	0	$p_{2.11}$	0
3	$p_{3.1}$	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	$p_{4.5}$	$p_{4.6}$	0	$p_{4.8}$	0	0	0	$p_{4.12}$
5	0	0	0	0	0	0	$p_{5.7}$	0	0	0	0	0
6	0	0	0	0	0	0	$p_{6.7}$	0	0	0	0	0
7	0	0	0	0	$p_{7.5}$	$p_{7.6}$	0	$p_{7.8}$	0	0	0	$p_{7.12}$
8	0	0	0	0	0	0	0	0	0	0	0	$p_{8.12}$
9	0	0	0	$p_{9.4}$	0	0	0	0	0	$p_{9.10}$	0	0
10	0	0	0	$p_{10.4}$	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	$p_{11.12}$
12	0	0	0	0	0	0	0	0	0	0	0	0

Fonte: Elaboração Própria (2020)

Observa-se que a matriz representada na Figura 12 não é ergódica, pois não satisfaz as condições de recorrência e periodicidade. Desta sorte, não é possível, por meio da aplicação da Cadeia de Markov, obter-se a probabilidade de estado estável. Assim, concluir-se-ia, em relação às questões de pesquisa (i) e (j) do item 1.3 que os decrementos probabilísticos a que um indivíduo ativo está sujeito não tendem a nenhum estado de convergência. Consequentemente, foram tentadas novas modelagens dos estados biométricos e sociais a que um militar possa estar sujeito. Contudo, verificou-se que mesmo em diferentes modelos, alguns dos estados tendem a serem absorventes. Por exemplo, o estado “Falecimento” é um estado no qual, para fins deste estudo, sempre existe a certeza de que, uma vez que o militar o atinja, nunca deixará este estado novamente.

Com o objetivo de dar continuidade a pesquisa, a partir de uma nova análise dos estados biométricos, decidiu-se por realizar um novo arranjo de decrementos de estados, desconsiderando os estados absorventes e obtendo-se um modelo simplificado. Neste novo grafo, privilegiou-se a interação entre os estados biométricos e sociais, com foco nos dependentes do militar e na sua entrada ou saída do serviço ativo. Nesta nova abordagem, também passou a ser considerada a hipótese de um militar não alterar seu estado biométrico e social entre um instante “ t ” e o seguinte “ $t+ 1$ ”. Desta forma, foram integrados 6 estados biométricos e sociais, conforme Figura 13:

Figura 13– Grafo de seis estados biométricos e sociais.



Fonte: Elaboração Própria (2020)

Onde os números que se encontram nos nós do grafo indicam os seguintes estados biométricos / sociais:

- 1) Militar ativo sem dependentes
- 2) Casamento
- 3) Paternidade
- 4) Reserva
- 5) Reforma / Invalidez (temporária ou não)
- 6) Demissão

O grafo representado pela Figura 13 também pode ser representado por uma matriz de transição, conforme ilustrado na Figura 14.

Figura 14 – Matriz de Transição de seis estados biométricos e sociais.

	1	2	3	4	5	6
1	p1.1	p1.2	p1.3	p1.4	p1.5	p1.6
2	p2.1	p2.2	p2.3	p2.4	p2.5	p2.6
3	p3.1	p3.2	p3.3	p3.4	p3.5	p3.6
4	p4.1	p4.2	p4.3	p4.4	p4.5	p4.6
5	p5.1	p5.2	p5.3	p5.4	p5.5	p5.6
6	p6.1	p6.2	p6.3	p6.4	p6.5	p6.6

Fonte: Elaboração Própria (2020)

Observa-se que a matriz representada na Figura 14 é ergódica, pois satisfaz as condições de recorrência e periodicidade. Ou seja, neste novo grafo simplificado é possível, por meio da aplicação da Cadeia de Markov, obter-se a probabilidade de estado estável.

Com base nos dados públicos do Anuário Estatístico da Marinha do Brasil de 2018 (ANEMAR 2018), foi possível estimar as probabilidades de transição entre os estados biométricos. Os valores foram estimados pelas razões entre os quantitativos de militares da MB em cada estado no ano de 2018.

O Anuário Estatístico da Marinha é uma publicação anual expedida pela Diretoria de Administração da Marinha, que permite ter uma visão geral da Marinha do Brasil, com informações estatísticas das atividades operativas e administrativas referentes ao ano em questão. O ANEMAR também apresenta séries históricas, que permitem a comparabilidade de informações e o acompanhamento da evolução, retratando variadas épocas da História Naval. Com a identificação de comportamentos e padrões dos dados apresentados no Anuário Estatístico da Marinha, a Alta Administração Naval dispõe de uma ferramenta para subsidiar processos de tomada de decisão. As informações do Volume I (Ostensivo) estão disponíveis aos servidores da instituição na intranet em www.dadm.mb, no ícone ANEMAR.

A Tabela 20 resume os critérios de cálculo para a estimativa de probabilidades.

Tabela 20: Probabilidades dos Estados Biométricos.

Estado Anterior	Estado Seguinte	Índice	Probabilidade	Critério de Cálculo Estimado
Militar ativo sem dependentes	Militar ativo sem dependentes	p_{1.1}	32.294/ 32.294 = 1	Razão entre o número total de militares ativos sem dependentes na MB no ano de 2018 e o número total de militares ativos sem dependentes na MB no ano de 2018.
Militar ativo sem dependentes	Casamento	p_{1.2}	44.005/32.294 = 0,1364	Razão entre o número de militares casados na MB no ano de 2018 e o número total de militares ativos sem dependentes na MB no ano de 2018.
Militar ativo sem dependentes	Paternidade	p_{1.3}	44.391/32.294 = 1,3745	Razão entre o número de militares com filhos na MB no ano de 2018 e o número total de militares ativos sem dependentes na MB no ano de 2018.
Militar ativo sem dependentes	Reserva	p_{1.4}	17.168 /32.294 = 0,5316	Razão entre o número de militares da MB na reserva remunerada no ano de 2018 e o número total de militares ativos sem dependentes na MB no ano de 2018.
Militar ativo sem dependentes	Reforma / Invalidez	p_{1.5}	34.774/32.294 = 1,076	Razão entre o número de militares da MB na situação de reformados / invalidez no ano de 2018 e o número

dependentes				total de militares ativos sem dependentes na MB no ano de 2018.
Militar ativo sem dependentes	Demissão	p1.6	$156 / 32.294 = 0,0048$	Razão entre o número de militares da MB que foram demitidos do serviço ativo no ano de 2018 e o número total de militares ativos sem dependentes na MB no ano de 2018.
Casamento	Militar ativo sem dependentes	p2.1	$32.294 / 44.005 = 0,7338$	Razão entre o número total de militares ativos sem dependentes na MB no ano de 2018 e o número de militares casados na MB no ano de 2018.
Casamento	Casamento	p2.2	$44.005 / 44.005 = 1$	Razão entre o número de militares casados na MB no ano de 2018 e o número de militares casados na MB no ano de 2018.
Casamento	Paternidade	p2.3	$44.391 / 44.005 = 1,0088$	Razão entre o número de militares com filhos na MB no ano de 2018 e o número de militares casados na MB no ano de 2018.
Casamento	Reserva	p2.4	$17.168 / 44.005 = 0,3901$	Razão entre o número de militares da MB na reserva remunerada no ano de 2018 e o número de militares casados na MB no ano de 2018.
Casamento	Reforma / Invalidez	p2.5	$34.774 / 44.005 = 0,7902$	Razão entre o número de militares da MB na situação de reformados / invalidez no ano de 2018 e o número de militares casados na MB no ano de 2018.
Casamento	Demissão	p2.6	$156 / 44.005 = 0,0035$	Razão entre o número de militares da MB que foram demitidos do serviço ativo no ano de 2018 e o número de militares casados na MB no ano de 2018.
Paternidade	Militar ativo sem	p3.1	$32.294 / 44.391 = 0,7275$	Razão entre o número total de militares ativos sem dependentes na

	dependentes			MB no ano de 2018 e o número de militares com filhos na MB no ano de 2018.
Paternidade	Casamento	p3.2	44.005/ 44.391= 0,9913	Razão entre o número de militares casados na MB no ano de 2018 e o número de militares com filhos na MB no ano de 2018.
Paternidade	Paternidade	p3.3	44.391/ 44.391= 1	Razão entre o número de militares com filhos na MB no ano de 2018 e o número de militares com filhos na MB no ano de 2018.
Paternidade	Reserva	p3.4	17.168 / 44.391= 0,3867	Razão entre o número de militares da MB na reserva remunerada no ano de 2018 e o número de militares com filhos na MB no ano de 2018.
Paternidade	Reforma / Invalidez	p3.5	34.774/ 44.391= 0,7834	Razão entre o número de militares da MB na situação de reformados / invalidez no ano de 2018 e o número de militares com filhos na MB no ano de 2018.
Paternidade	Demissão	p3.6	156 / 44.391= 0,0035	Razão entre o número de militares da MB que foram demitidos do serviço ativo no ano de 2018 e o número de militares com filhos na MB no ano de 2018.
Reserva	Militar ativo sem dependentes	p4.1	32.294 / 17.168 = 1,8811	Razão entre o número total de militares ativos sem dependentes na MB no ano de 2018 e o número de militares da MB na reserva remunerada no ano de 2018.
Reserva	Casamento	p4.2	44.005 / 17.168 = 2,5632	Razão entre o número de militares casados na MB no ano de 2018 e o número de militares da MB na reserva remunerada no ano de 2018.
Reserva	Paternidade	p4.3	44.391 / 17.168 =	Razão entre o número de militares com filhos na MB no ano de 2018 e o

			2,5857	número de militares da MB na reserva remunerada no ano de 2018.
Reserva	Reserva	p4.4	$\frac{17.168}{17.168} = 1$	Razão entre o número de militares da MB na reserva remunerada no ano de 2018 e o número de militares da MB na reserva remunerada no ano de 2018.
Reserva	Reforma / Invalidez	p4.5	$\frac{34.774}{17.168} = 2,0255$	Razão entre o número de militares da MB na situação de reformados / invalidez no ano de 2018 e o número de militares da MB na reserva remunerada no ano de 2018.
Reserva	Demissão	p4.6	$\frac{156}{17.168} = 0,0091$	Razão entre o número de militares da MB que foram demitidos do serviço ativo no ano de 2018 e o número de militares da MB na reserva remunerada no ano de 2018.
Reforma / Invalidez	Militar ativo sem dependentes	p5.1	$\frac{32.294}{34.774} = 0,9286$	Razão entre o número total de militares ativos sem dependentes na MB no ano de 2018 e o número de militares da MB na situação de reformados / invalidez no ano de 2018.
Reforma / Invalidez	Casamento	p5.2	$\frac{44.005}{34.774} = 1,2654$	Razão entre o número de militares casados na MB no ano de 2018 e o número de militares da MB na situação de reformados / invalidez no ano de 2018.
Reforma / Invalidez	Paternidade	p5.3	$\frac{44.391}{34.774} = 1,2766$	Razão entre o número de militares com filhos na MB no ano de 2018 e o número de militares da MB na situação de reformados / invalidez no ano de 2018.
Reforma / Invalidez	Reserva	p5.4	$\frac{17.168}{34.774} = 0,4937$	Razão entre o número de militares da MB na reserva remunerada no ano de 2018 e o número de militares da MB

				na situação de reformados / invalidez no ano de 2018.
Reforma / Invalidez	Reforma / Invalidez	p5.5	$34.774 / 34.774 = 1$	Razão entre o número de militares da MB na situação de reformados / invalidez no ano de 2018 e o número de militares da MB na situação de reformados / invalidez no ano de 2018.
Reforma / Invalidez	Demissão	p5.6	$156 / 34.774 = 0,0045$	Razão entre o número de militares da MB que foram demitidos do serviço ativo no ano de 2018 e o número de militares da MB na situação de reformados / invalidez no ano de 2018.
Demissão	Militar ativo sem dependentes	p6.1	$32.294 / 156 = 207,0128$	Razão entre o número total de militares ativos sem dependentes na MB no ano de 2018 e o número de militares da MB que foram demitidos do serviço ativo no ano de 2018.
Demissão	Casamento	p6.2	$44.005 / 156 = 282,0833$	Razão entre o número de militares casados na MB no ano de 2018 e o número de militares da MB que foram demitidos do serviço ativo no ano de 2018.
Demissão	Paternidade	p6.3	$44.391 / 156 = 284,5577$	Razão entre o número de militares com filhos na MB no ano de 2018 e o número de militares da MB que foram demitidos do serviço ativo no ano de 2018.
Demissão	Reserva	p6.4	$17.168 / 156 = 110,05112$	Razão entre o número de militares da MB na reserva remunerada no ano de 2018 e o número de militares da MB que foram demitidos do serviço ativo no ano de 2018.
Demissão	Reforma / Invalidez	p6.5	$34.774 / 156 = 222,9103$	Razão entre o número de militares da MB na situação de reformados /

				invalidez no ano de 2018 e o número de militares da MB que foram demitidos do serviço ativo no ano de 2018.
Demissão	Demissão	p_{6.6}	$156 / 156 = 1$	Razão entre o número de militares da MB que foram demitidos do serviço ativo no ano de 2018 e o número de militares da MB que foram demitidos do serviço ativo no ano de 2018.

Fonte: Elaboração própria (2020)

A relação entre as probabilidades baseada nas razões estabelecidas na tabela acima pode ser ilustrada conforme a Tabela 21.

Tabela 21 – Tabela de razões de probabilidades não normalizadas

		Ativo sem dependentes	Casamento	Paternidade	Reserva	Invalidez / Reforma	Demissão
		1	2	3	4	5	6
Ativo sem dependentes	1	1	0,1364	1,3745	0,532	1,0768	0,0048
Casamento	2	0,7338	1	1,0088	0,39	0,7902	0,0035
Paternidade	3	0,7275	0,9913	1	0,387	0,7834	0,0035
Reserva	4	1,8811	2,5632	2,5857	1	2,0255	0,0091
Invalidez / Reforma	5	0,9286	1,2654	1,2766	0,494	1	0,0045
Demissão	6	207,0128	282,08	284,5577	110,1	222,91	1

Fonte: Elaboração própria.

Os valores da tabela acima foram normalizados, de modo que fossem enquadrados em uma faixa de valores entre 0 e 1. Segundo Gomes e Gomes (2019), há 4 formas distintas de se realizar uma normalização, conforme ilustrado na Tabela 22.

Tabela 22 – Tipos de Normalização

Procedimento	Fórmula	Observação sobre o Resultado
1	$A_{ij}/\max a_{ij}$	Gera pesos negativos; atribui peso zero ao critério que está no meio das preferências
2	$(a_{ij} - \min a_{ij})/(\max a_{ij} - \min a_{ij})$	Gera pesos positivos; atribuindo peso zero ao critério menos preferido
3	$a_{ij} / \sum a_{ij}$	O denominador assume valor zero
4	$A_{ij} / \sqrt{(\sum a_{ij}^2)}$	Gera valores negativos para pesos e valor zero para o critério no meio do conjunto de preferências

Fonte: Adaptado de Gomes e Gomes (2019)

Como todos os valores a serem normalizados são positivos e a soma das normalizações deve ser igual a 1, optou-se pela normalização 3. Esta normalização consiste em realizar a divisão do número a ser normalizado pelo somatório dos números do subconjunto. Desta forma, garante-se que todos os valores são não nulos e encontram-se dentro do intervalo entre 0 e 1. A Figura 15 representa a matriz de transição com os valores já normalizados.

Figura 15 – matriz de transição de seis estados com valores numéricos normalizados.

0,242477	0,0331	0,333285	0,129	0,2611	0,0012
0,186889	0,2547	0,256927	0,099	0,20125	0,0009
0,186903	0,2547	0,256911	0,099	0,20126	0,0009
0,186903	0,2547	0,25691	0,099	0,20125	0,0009
0,186886	0,2547	0,256923	0,099	0,20126	0,0009
0,1869	0,2547	0,25691	0,099	0,20125	0,0009

Fonte: Elaboração Própria (2020)

Utilizando-se a equação de Chapman-Kolmogorov e elevando-se a matriz de transição da Cadeia de Markov de 6 estados à 54ª potência, foi possível calcular a probabilidade de estado estável. A Figura 16 ilustra a matriz de probabilidades após a 54ª aplicação da Cadeia de Markov.

Figura 16 – Matriz de probabilidade estável.

0.197891869	0.21082041	0.272027876	0.105199722	0.213097157	0.00095182258
0.197891869	0.21082041	0.272027876	0.105199722	0.213097157	0.00095182258
0.197891869	0.21082041	0.272027876	0.105199722	0.213097157	0.00095182258
0.197891869	0.21082041	0.272027876	0.105199722	0.213097157	0.00095182258
0.197891869	0.21082041	0.272027876	0.105199722	0.213097157	0.00095182258
0.197891869	0.21082041	0.272027876	0.105199722	0.213097157	0.00095182258

Fonte: Elaboração Própria (2020)

A Tabela 23 resume as probabilidades de estado estável para cada um dos estados separadamente.

Tabela 23 - Probabilidades de estado estável

Estado Biométrico e Social	Probabilidade de Estado Estável
Ativo sem dependentes	0.1978918690
Casamento	0.2108204100
Paternidade	0.2720278760
Reserva	0.1051997220
Invalidez / Reforma	0.2130971570
Demissão	0.00095182258

Fonte: Elaboração Própria (2020)

6. CONCLUSÃO E RECOMENDAÇÕES PARA TRABALHOS FUTUROS

6.1 CONCLUSÃO

Esta dissertação realizou uma pesquisa exploratória em continuação ao trabalho de Santos (2018), apresentando novas contribuições àquela linha de pesquisa. Santos (2018) teve o propósito de apresentar uma metodologia capaz de calcular rapidamente as projeções Atuariais dos custos com pensionistas das Forças Armadas Brasileiras num horizonte temporal de 75 anos, com a melhor acurácia possível. Naquela tese, aplicou-se somente o teste de aderência Qui-quadrado com 99 agravamentos e 99 desagravamentos para cada tábua atuarial, exclusivamente pelo método estatístico do cálculo por idade exata.

Durante esta pesquisa, no entanto, verificou-se que há dois métodos de cálculo da estatística de teste de aderência, que foram aplicados no presente estudo, a saber: cálculo por idade exata e cálculo por faixa de idades. Cabe ressaltar que não foram encontrados registros que o segundo método já houvesse sido aplicado para a população das forças armadas brasileiras. Foram analisados, ainda, quais os parâmetros que melhor formariam as dimensões a serem consideradas por ocasião de um teste de aderência, incluindo-se uma quarta dimensão ao Cubo de Aderência de Santos (2018), criando-se o “Hiper-cubo de Aderência”. Ou seja, além dos tradicionais parâmetros de idade, tábua atuarial propriamente dita e agravamentos e desagravamentos percentuais da probabilidade, foi incluído um quarto parâmetro de agravamento e desagravamento em anos no perfil da idade.

Adicionalmente, esta dissertação ampliou a quantidade tábuas atuariais analisadas: em Santos (2018) foram avaliadas 68 tábuas atuariais, enquanto o presente trabalho utilizou 75 tábuas. Desta sorte, se segundo Santos (2018) o Cubo de Aderência permitiu a realização de 9.353 testes de mortalidade de válidos e 4.179 testes de entrada em invalidez e mortalidade de inválidos em um único espaço amostral pré-definido pra cada teste, no presente trabalho, através do Hiper-cubo de aderência, foram realizados 76.002 testes de mortalidade de válidos e 30.114 testes de entrada em invalidez e mortalidade de inválidos para cada um dos diversos espaços amostrais verificados.

Esta pesquisa não encontrou na literatura ou no mercado previdenciário uma maneira reconhecidamente consagrada de comparar quantitativamente os resultados de aderência de tábuas atuarias em diferentes testes de aderência. Isto se deve ao fato de que os diferentes testes empregam estatísticas de testes distintas, até mesmo em termos de grandezas

matemáticas. Desta forma, optou-se por uma análise qualitativa dos dados encontrados, que foi considerada satisfatória para o objetivo deste estudo específico. Vale mencionar, no entanto, que os resultados apresentados nessa pesquisa são sensíveis a variações de hipóteses, da base normativa e da base cadastral.

Conforme apresentado nas análises de resultados, não há uma necessidade premente do cálculo de uma tábua atuarial específica para a população das Forças Armadas, dada a aderência daquela amostra às tábuas comerciais. Entretanto, a comparação entre o resultado de uma tábua especificamente desenvolvida para a população das FFAA e daquelas tábuas comerciais já analisadas pode apresentar alguma melhoria na estatística do teste.

O presente trabalho comprovou que, embora o teste de aderência Qui-quadrado já esteja consagrado pelo uso na literatura acadêmica, não é o que melhor atende às especificidades das FFAA. Além de apresentar menor quantidade de tábuas aderentes, em alguns casos, para que alguma tábua atuarial apresentasse aderência, fazia-se necessário reduzir o espaço amostral para apenas 1/3 da população em estudo. Por outro lado, o teste do Desvio Quadrático Médio sempre apresenta resultado satisfatório para toda população e todas as tábuas analisadas. Isto porque, pela natureza do teste, fornece uma ordenação das tábuas testadas, das menos discrepantes aos dados observados da população analisada para aquelas que mais se diferem da amostra. Entretanto, conforme verificado na revisão da literatura, isto não significa uma maior robustez. Pelo contrário, este teste é particularmente recomendado quando não é encontrada aderência estatística a nenhuma tábua atuarial a partir dos métodos Qui-Quadrado ou Kolmogorov-Smirnov.

Concluiu-se que, dentre os principais testes de aderência difundidos no meio acadêmico, aquele que melhor se aplica para a seleção das tábuas atuariais para os militares das Forças Armadas Brasileiras é o teste de Kolmogorov-Smirnov. O teste de Kolmogorov-Smirnov demonstrou-se robusto, possibilitando a aderência de diversas tábuas atuariais, incluindo um espaço amostral da população das FFAA mais abrangente e com o resultado da estatística de teste satisfatório.

O presente estudo também comprovou a inaplicabilidade das Cadeias de Markov para determinar se os decrementos probabilísticos originalmente propostos por Santos (2018) convergiriam para um estado funcional de estabilidade. Entretanto, para fins acadêmicos, nesta dissertação, foram aceitas simplificações do grafo de estados biométricos e sociais, de modo a permitir que os decrementos probabilísticos a que um indivíduo ativo está sujeito tendessem a algum estado de convergência. Assim, foi calculada a matriz de probabilidade de estado estável para os estados biométricos e sociais.

Diante de todo o exposto, verifica-se que este trabalho atendeu a seus objetivos principal e secundários. Insta salientar, ainda, que a presente dissertação tratou de um problema atual e real, referente às FFAA, que afeta diretamente mais de um milhão de cidadãos brasileiros.

6.2. RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Com o objetivo de ampliar a abrangência desta dissertação e em consonâncias com as recomendações de Santos (2018) para pesquisas futuras, foram identificadas as seguintes possibilidades de trabalhos futuros:

- Realização de novas projeções atuariais para os militares das FFAA a partir das tábuas atuariais de melhor aderência encontradas mediante o teste de Kolmogorov-Smirnov;
- Desenvolvimento de uma tábua atuarial específica para a população das Forças Armadas e comparação entre o resultado dos testes de aderência à uma tábua especificamente desenvolvida para a população das FFAA e daquelas tábuas comerciais já analisadas; e
- Avaliar a possibilidade do desenvolvimento de tábuas dinâmicas, por meio de Simulação, com o propósito de antecipar os acréscimos de longevidade que se tem observado ao longo do tempo.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALTMAN, D. G.; BLAND, J. M. **Statistics notes: the normal distribution**. *BMJ*, v.310(6975), p.298, 1995. DOI: <https://doi.org/10.1136/bmj.310.6975.298>
Disponível em <<https://www.bmj.com/content/310/6975/298.short>> . Acesso em 13/08/2019
- ANDERSON, T. W; DARLING, D. A. Asymptotic theory of certain goodness of fit criteria based on stochastic processes. **The Annals of Mathematical Statistics**, v.23(2), p.193-212. DOI: 10.1214/aoms/1177729437
- ANDERSON, T.W.; DARLING, D.A. A Test of Goodness of Fit. **Journal of The American Statistical Association**, v. 49 (268), p. 765-769, 1954.
- ARMSTRONG, J. Scott; COLLOPY, Fred. Error Measures For Generalizing About Forecasting Methods: Empirical Comparisons. **International Journal of Forecasting**, v. 8, p.69-80, 1992.
- ASSIS, Janilson Pinheiro de; SOUZA, Roberto Pequeno de; DIAS, Carlos Tadeu dos Santos. **Glossário de estatística**, Mossoró / RN, EdUFERSA, 901f., 2019. ISBN: 978-85-5757-104-4
- BARCELÓ, Juan A. Chi-Square Analysis. **The Encyclopedia of Archaeological Sciences**, 2018. DOI: 10.1002/9781119188230.saseas0090
- BARROS, Michelli; LEIVA, Victor; OSPINA, Raydonal; TSUYUGUCHI, Aline. Goodness-of-Fit Tests for the Birnbaum-Saunders Distribution With Censored Reliability Data. **IEEE Transactions On Reliability**, v. 63 (2), p.543-554, 2014. DOI: 10.1109/TR.2014.2313707
- BETSCH , Steffen; EBNER, Bruno. Testing normality via a distributional fixed point property in the Stein characterization. **Test**, p. 1-34, 2019. DOI: 10.1007/s11749-019-00630-0
- BIAU, D.J.; JOLLES, B.M.; PORCHER, R. **P value and the theory of hypothesis testing: an explanation for new researchers**. *Clin Orthop Relat Res.*, v. 463 (3), p. 885–892. DOI:10.1007/s11999-009-1164-4. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2816758/>> Acesso em: 30/07/2019
- BÜHLMANN, Hans. The Actuary: the Role and Limitations of the Profession Since the Mid-19th Century. **ASTIN Bulletin**, v.27 (2), p.165–171, 1997. DOI: 10.2143/ast.27.2.542046. ISSN 0515-0361. Disponível em: <<https://www.casact.org/library/astin/vol27no2/165.pdf>> Acesso em: 19/05/2019
- BRASIL, Marinha do Brasil, Secretaria Geral da Marinha, Diretoria de Administração da Marinha. **Anuário Estatístico da Marinha**, Volume 1, 45ª Edição, 2018
- CALDART, Paulo Roberto; MOTTA, Silvia Teixeira da; CAETANO, Marcelo Abi-Ramia; BONATTO, Taís Vieira. **Adequação das Hipóteses Atuariais e Modelo Alternativo de Capitalização para o Regime Básico do RPPS: o Caso do Rio Grande do Sul**. *Revista Contabilidade e Finanças*, v..25, n.66, São Paulo, Sept./Dec, 2014. DOI: <http://dx.doi.org/10.1590/1808-057x201412200>

CARDOSO, Rodrigo Felix Sarruf. A solidariedade social e a contributividade como alicerces da Previdência Social dos servidores públicos civis. **Âmbito Jurídico**, Rio Grande, IX, nº 36, jan, 2007. Disponível em: <http://www.ambitojuridico.com.br/site/index.php?n_link=revista_artigos_leitura&artigo_id=1629>. Acesso em: 13/03/2019.

CASTRO, Marcia Caldas de. **Entradas e Saídas no Sistema Previdenciário Brasileiro: Uma Aplicação de Tábuas de Mortalidade**. Dissertação (Mestrado) apresentada ao Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR) da Faculdade de Ciências Econômicas (FACE) da Universidade Federal de Minas Gerais (UFMG), como requisito parcial para a obtenção do grau de Mestre em Demografia. Orientadora: Prof^a. Diana Oya Sawyer. 229f. 1997

CONOVER, W.J. **Practical nonparametric statistics**, 3rd ed, Wiley, New York, 1999.

CONTI, José Maurício. CARVALHO, André Castro. O Controle Interno na Administração Pública Brasileira: Qualidade do Gasto Público e Responsabilidade Fiscal. **Revista da Defensoria Pública da União**, nº 37, Jan-Fev, Seção Especial de Teorias e Estudos Científicos, p. 201 – 220, 2011. DOI: 10.11117//22361766.37.01.08

CORBETT, Charles J.; WASSENHOVE, Luk N. Van. The Natural Drift: What Happened to Operations Research? **Operations Research**, v. 41, nº. 4, July-August, p. 625-640, 1993 DOI: 10.1287/opre.41.4.625

CORRÊA, Raphael Baseggio. **Modelo De Simulação De Governança De Passivo Atuarial De Um Fundo De Pensão Brasileiro**. 86f. Dissertação de Mestrado Acadêmico apresentada ao Programa de Pós-Graduação em Administração da Universidade Federal do Rio Grande do Sul. 2018

COSTA, H.G. Modelo para webibliomining: proposta e caso de aplicação. **Revista da FAE**, v. 13, n. 1, p. 115-126, 2010.

D'AGOSTINO, Ralph B; STEPHENS, Michael A. **Goodness-of-fit Techniques**, v.68, 1st Edition, 1986. Disponível em <<https://books.google.com.br/books?id=1BSEaGVBj5QC&printsec=frontcover&hl=pt-BR#v=onepage&q&f=false>>. Acessado em 15/03/2019.

DALLAL, Gerard E. An analytic approximation to the distribution of Lilliefors's test statistic for normality. **The American Statistician**, v. 40, p. 294-296, 1986.

DALLAL, Gerard E. (2012). **The Little Handbook of Statistical Practice**. Disponível em: <<http://www.jerrydallal.com/LHSP/LHSP.htm>> Acesso em: 30/07/2019

D'ARCY, Stephen P. On Becoming An Actuary of the Third Kind. **Proceedings of the Casualty Actuarial Society**, v. LXXVI (145), p. 45–76, May, 1989. Disponível em: <<https://www.casact.org/pubs/proceed/proceed89/89045.pdf>> Acesso em: 22/05/2019

DAS, Samiran. Goodness-of-Fit Tests for Generalized Normal Distribution for Use in Hydrological Frequency Analysis. **Pure and Applied Geophysics**, v.175 (10), p.3605-3617, 2018. DOI: 10.1007/s00024-018-1877-y

DICKSON, David C. M.; HARDY, Mary R.; WATERS, Howard R. **Actuarial Mathematics for Life Contingent Risks**, Cambridge University Press, 2009. ISBN-13 978-0-521-11825-5

DOS SANTOS, Ronaldo César Evangelista; GREGÓRIO, Thierry Faria da Silva. **ADERÊNCIA 4D**, Arcádia Sistemas, Versão 1, 17 jul. 2018, Disponível em: <http://aderencia4d.arcadiasistemas.com.br/> Acesso em: 20 set.2019

DROBNE, Samo; DUMICIC, Ksenija; STIRN, Lidija Zadnik. Novel Solutions or Novel Approaches in Operational Research. **Business Systems Research**, v. 9, nº. 2, 2018.

FERREIRA, E. B. L. Secretaria Geral da Marinha. Sem privilégios - há muito as Forças Armadas contribuem para o equilíbrio do orçamento. 23 fev. 2017. Disponível em: <https://www.marinha.mil.br/spsm/node/52>>. Acesso em 17 nov. 2019.

FREES, Edward W. Stochastic Life Contingencies with Solvency Considerations. **Transactions of the Society of Actuaries**. v. XLII, p. 91–148, January, 1990. Disponível em: <https://web.archive.org/web/20060103161211/http://library.soa.org/library/tsa/1990-95/TSA90V427.pdf>> Acesso em: 22/05/2019

GLANTZ, Stanton A. **Primer in Biostatistics**, 5th ed. New York: McGraw-Hill; 2002.

GOMES, Carlos Francisco Simões; RIBEIRO, Priscilla Cristina Cabral. **Gestão da Cadeia de Suprimentos integrada à Tecnologia da Informação**. 2ª Edição. Editora SENAC. Rio de Janeiro, RJ.2014

GOMES, Carlos Francisco Simões. SANTOS, Marcos dos. MARTINS, Ernesto Rademaker. Seleção de Tábuas Atuariais para os Militares das Forças Armadas Brasileiras: uma abordagem prospectiva. *In: Anais do VII ENFEPro - ENCONTRO FLUMINENSE DE ENGENHARIA DE PRODUÇÃO*, editora SFEPro - Sociedade Fluminense de Engenharia de Produção, CEFET/RJ - Campus Nova Iguaçu, 22 e 23 de Novembro de 2017. ISSN 2178-3272.

GOMES, Carlos Francisco Simões. SANTOS, Marcos dos. MARTINS, Ernesto Rademaker. GREGÓRIO, Thierry Faria da Silva. SANTOS, Ronaldo Cesar Evangelista. Redução do tempo de processamento do Cálculo Atuarial das Forças Armadas: uma aplicação da Computação Paralela. *In: Anais do XIV SEGeT – SIMPÓSIO DE EXCELÊNCIA EM GESTÃO E TECNOLOGIA*, Resende, RJ, 26 e 27 de Outubro de 2017.

GOMES, Luiz Flavio Autran; GOMES, Carlos Francisco Simões. **Princípios e métodos para a tomada de decisão: Enfoque multicritério** (6a ed.). São Paulo, Atlas, 2019.

GOUVEIA, André Luiz Lemos Andrade; SOUZA, Filipe Costa de; RÊGO, Leandro Chaves. Justiça atuarial nos cálculos previdenciários: aplicação de um modelo multidecremental para comparação da regra do fator previdenciário e da idade mínima. **Revista Contabilidade &**

Finanças. vol.29 no.78 São Paulo Sept./Dec. 2018. On-line version ISSN 1808-057X. DOI: <http://dx.doi.org/10.1590/1808-057x201805740>

HAZRA, Avijit; GOGTAY, Nithya. Module on Biostatistics and Research Methodology for the Dermatologist. **Indian Journal of Dermatology**, v.62(2), p.123-129, 2016.

HARPER, W. M. **Operational Research – M & E Handbooks**. MacDonal & Evans Ltd. Londres, Reino Unido, 1975. ISBN: 0 7121 1514 5

HELLER, Robert. **Making Decisions – Essential Managers**. Dorling Kindersley Ltd. Londres, Reino Unido, 1998. ISBN: 0 7513 0631 2

HILLIER, Frederick S. LIEBERMAN, Gerald J. **Introdução à Pesquisa Operacional**, 9ª Edição, 2013.

HICKMAN, James History of Actuarial Profession. **Encyclopedia of Actuarial Science**, 2004. Disponível em: <https://web.archive.org/web/20060724173339/http://www.wiley.co.uk/eoas/pdfs/TAH012-.pdf> Acesso em: 22/05/2019

HOWELL, David C. **Statistical methods for psychology**, 6. ed., Belmont, Calif.: Thomson, 2007. ISBN 978-0495012870.

HYNDMAN, Rob J.; KOEHLER, Anne B. Another look at measures of forecast accuracy. **International Journal of Forecasting**, v.22 (4), p.679–688, 2006. DOI:10.1016/j.ijforecast.2006.03.001.

KAYIKÇI, Emine Tanır; SOPACI, Eyüp. Testing the normality of the residuals of surface temperature data at VLBI/GPS co-located sites by goodness of fit tests. **Saudi Society for Geosciences**, v.8 (11), p. 10119-10134, 2015. DOI 10.1007/s12517-015-1911-7

KHAN, Hina; Aslam, Muhammad; Khan, Masood Amjad. Properties of Exponential Ratio Type Estimators in Equal Probability Sampling: A Simulation Study. **Communications in Mathematics and Statistics**, v. 6 (1), p. 91–118, 2018.

KOLMOGOROV, A. Sulla determinazione empirica di una legge di distribuzione. **Giornale dell’Istituto Italiano degli Attuari**, v. 4, p. 83-91, 1933.

KRUGER, Hennie; VERHOEF, Anne; PREISER, Rika. The Epistemological Implications of Critical Complexity Thinking for Operational Research. **Systems**, v. 7 (1), 2019.

LEAL, Carlos Ivan Simonsen; CUNHA, Armando Santos Moreira; LYRA, Irapoan Cavalcanti de; MOTTA, Paulo Roberto de Mendonça. **As Forças Armadas e a PEC da Previdência**. Fundação Getúlio Vargas, Rio de Janeiro, RJ, 2019. Disponível em: <https://www.marinha.mil.br/spsm/content/fgv-for%20as-armadas-e-pec-da-previd%20ancia-2> Acesso em: 01/05/2019.

LILLIEFORS, Hubert W. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. **Journal of the American Statistical Association**, v. 62 (318), p. 399-402, 1967. DOI: 10.1080/01621459.1967.10482916

LITTLE, John D.C. IFORS' Operational Research Hall of Fame: Philip McCord Morse. **International Transactions in Operational Research**, v. 10, p. 307–309, 2003.

LOBATO JÚNIOR, José Maria dos Santos. **Provisões da Previdência Social: Uma Abordagem Motivacional ao Ensino de Matemática Financeira na Educação Básica**. 2018. 64 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional) – Universidade Federal do Pará, Belém, 2018.

MAROCO, João. **Análise Estatística - com utilização do SPSS**. 3ª Edição. Edições Sílabo, Lisboa, Portugal, 2010. ISBN: 978-972-618-452-2

MAYDEU-OLIVARES, Alberto; FORERO, Carlos G. Goodness-of-Fit Testing. **International Encyclopedia of Education**, v. 7, p. 190-196, 2010. DOI: 10.1016/B978-0-08-044894-7.01333-6

MAYER, D.G.; BUTLER, D.G. Statistical validation. **Ecological Modelling**, v. 68 (1–2), p.21-32, July, 1993. DOI: 10.1016/0304-3800(93)90105-2

McASSEY, Michael P. An empirical goodness-of-fit test for multivariate distributions. **Journal of Applied Statistics**, v. 40 (5), p. 1120-1131, 2013.

McDONALD, John H. G–test of goodness-of-fit. **Handbook of Biological Statistics**, 3rd ed., Sparky House Publishing, Baltimore, Maryland, p. 53–58, 2014.

MONISH, N.T., REHANA, S. Suitability of distributions for standard precipitation and evapotranspiration index over meteorologically homogeneous zones of India. **Journal of Earth System Science**, 129, 25 (2020) doi:10.1007/s12040-019-1271-x

MORSE, Philip McCord; KIMBALL, George Elbert. **Methods of Operations Research**. The Technology Press and John Wiley & Sons Inc. New York. v. 55 (492), p. 809-810, December, 1951.

NATURE, v. 160, p. 660 – 662, 15 de Novembro de 1947.

NOGUEIRA, Adriano Roner Melo. **Envelhecimento populacional brasileiro: idade prospectiva e retrospectiva**. Rio de Janeiro, 2015. 97 f. Dissertação (Mestrado em Estudos Populacionais e Pesquisas Sociais) – Escola Nacional de Ciências Estatísticas.

OGBORN, M.E. The Professional Name of Actuary. **Journal of the Institute of Actuaries**, Faculty and Institute of Actuaries, v.82, p. 233–246, December, 1956.

ÖNER, Metin; KOCAKOÇ, İpek Deveci. A Compilation of Some Popular Goodness of Fit Tests for Normal Distribution: Their Algorithms and MATLAB Codes. **Journal of Modern Applied Statistical Methods**, v.16 (2), p. 547-575, 2017.

PEDROSA, Ignacio; JUARROS-BASTERRETxea, Joel ; ROBLES-FERNANDEZ, Adan; BASTEIRO, Julia; GARCIA-CUETO, Eduardo. Goodness of Fit Tests for Symmetric Distributions, which Statistical Should I Use? **Universitas Psychologica**, v.14, n°.1, Jan./Mar, Bogotá, 2015. DOI:10.11144/Javeriana.upsy13-5.pbad

QUESSY, Jean-François; RIVEST, Louis-Paul; TOUPIN, Marie-Hélène. Goodness-of-fit tests for the family of multivariate chi-square copulas. **Computational Statistics & Data Analysis**. Volume 140, December 2019, Pages 21-40, 2019.

RAZALI, Nornadiah; WAH, Yap Bee (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. **Journal of Statistical Modeling and Analytics**. 2 (1): 21–33. 2011

ROSS, Sheldon M. **Introduction to Probability Models**, 11th Edition, Academic Press, Elsevier, 2014. ISBN-10: 0124079482

SEDIC, Almir ; PAVKOVIC, Danijel ; FIRAK, Mihajlo. A methodology for normal distribution-based statistical characterization of long-term insolation by means of historical data. **Solar Energy**, v. 122, p. 440-454, 2015. DOI: 10.1016/j.solener.2015.09.014

SANTOS, Marcos dos; MARTHA, Leonardo da Costa; REIS, Marcone Freitas dos; TEIXEIRA, Luiz Frederico Horácio de Souza de Barros. Utilização do algoritmo Branch and Bound na otimização da produção de uma indústria de produtos plásticos. **Revista de Trabalhos Acadêmicos Lusófona**, v. 2, p. 217-237, 2019.

SANTOS, Marcos dos. **Proposta de Modelagem Atuarial aplicada ao setor militar considerando influências econômicas e biométricas**. Orientador: Carlos Francisco Simões Gomes. 2018. 237f. Tese (Doutorado em Engenharia de Produção) – Escola de Engenharia, Universidade Federal Fluminense, Niterói, RJ, 2018. DOI: 10.22409/TPP.2018.d.05163096761

SANTOS, Marcos dos; QUINTAL, Renato Santiago; PAIXÃO, Alexandre Camacho da; GOMES, Carlos Francisco Simões. Simulation of Operation of An Integrated Information for Emergency Pre-Hospital Care In Rio de Janeiro Municipality. **Procedia Computer Science**, v. 55, p.931 – 938, 2015. DOI: 10.1016/j.procs.2015.07.111

SENETA, E. Markov and the Birth of Chain Dependence Theory. **International Statistical Review / Revue Internationale De Statistique**, vol. 64, no. 3, 1996, pp. 255–263. JSTOR, Disponível em: <www.jstor.org/stable/1403785>. Acesso em 02/08/2019.

SERFOZO, Richard. **Basics of Applied Stochastic Processes**. Springer Science & Business Media. p. 2, 24 January 2009. ISBN 978-3-540-89332-5.

SHAPIRO, S. S; WILK, M. B. An analysis of variance test for Normality. **Biometrika**, v. 52(3-4), p. 591-611, 1965. DOI: 10.1093/biomet/52.3-4.591

SMIRNOV, N. Table for Estimating the Goodness of Fit of Empirical Distributions. **Ann. Math. Statist.** v.19, n°. 2, p.279-281, 1948. DOI:10.1214/aoms/1177730256. Disponível em: <<https://projecteuclid.org/euclid.aoms/1177730256>> Acesso em: 23/06/2019

SPITZER, Frank (1970). Interaction of Markov Processes. **Advances in Mathematics**, v.5 (2), p. 246–290. DOI:10.1016/0001-8708(70)90034-4

STEEL, Robert G. D.; TORRIE, James H. **Principles and procedures of statistics a biometrical approach**. McGraw-Hill International, 2nd ed, 1980. ISBN 0-07-060926-8

STEPHENS, M. A. EDF Statistics for Goodness of Fit and Some Comparisons. **Journal of the American Statistical Association**, v.69 (347), p. 730–737, Setembro de 1974. ISSN 0162-1459. DOI:10.2307/2286009

SULEWSKI, Piotr. Modified Lilliefors goodness-of-fit test for normality. **Communications In Statistics-Simulation And Computation**, 2019. DOI: 10.1080/03610918.2019.1664580

VICKERS, Andrew J. **To P or Not to P: Why Use a P Value, Anyway?** Medscape Business of Medicine, v.7(1), 2006. Disponível em: < <https://www.medscape.com/viewarticle/524206>> Acesso em: 30/07/2019

WETHERILL, G. Barrie. **Regression Analysis with Applications**. Monographs in Statistics and Applied Probability. Ed. Chapman and Hall, 1986. Disponível em: <https://books.google.com.br/books/about/Regression_analysis_with_applications.html?id=N BvvAAAAMAAJ&redir_esc=y> Acesso em: 20/07/2019

WILLMOTT, Cort; MATSUURA, Kenji. On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. **International Journal of Geographic Information Science**, v.20, p. 89–102, 2006. DOI:10.1080/13658810500286976.

YOUNG, Ian T. Proof without prejudice: use of the Kolmogorov-Smirnov Test for the analysis of Histograms from Flow Systems. **The Histochemical Society, Inc.**, v. 25, nº 7, p. 935-941, 1977.

ZACHARIAH, R; Harries, A D; Ishikawa, N; Rieder, H L; Bissell, K; Laserson, K; Massaquoi, M; Van Herp, M; Reid, T. Operational research in low-income countries: what, why, and how? **Journal Lancet Infectious Diseases**, v.9 (11), 2009. DOI: 10.1016/S1473-3099(09)70229-4