

Ítalo de Moraes Dolores

**Teoria Assintótica: uma Abordagem em  
Testes de Aderência**

Niterói - RJ, Brasil

21 de julho de 2023

Ítalo de Moraes Dolores

**Teoria Assintótica: uma Abordagem  
em Testes de Aderência**

**Trabalho de Conclusão de Curso**

Monografia apresentada para obtenção do grau de Bacharel em  
Estatística pela Universidade Federal Fluminense.

Orientador(a): Prof. Dr. Valentin Sisko

Co-Orientador(a): Ma. Daiane de Souza Oliveira

Niterói - RJ, Brasil

21 de julho de 2023

**Ítalo de Moraes Dolores**

**Teoria Assintótica: uma Abordagem em  
Testes de Aderência**

Monografia de Projeto Final de Graduação sob o título “*Teoria Assintótica: uma Abordagem em Testes de Aderência*”, defendida por Ítalo de Moraes Dolores e aprovada em 21 de julho de 2023, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

---

**Prof. Dr. Valentin Sisko**  
Departamento de Estatística - UFF

---

**Prof. Dr. Marco Aurélio Santos Sanfins**  
Departamento de Estatística - UFF

---

**Profa. Dra. Mariana Albi de Oliveira Souza**  
Departamento de Estatística – UFF

---

**Ma. Daiane de Souza Oliveira**

Niterói, 21 de julho de 2023

Ficha catalográfica automática - SDC/BIME  
Gerada com informações fornecidas pelo autor

D659t Dolores, Ítalo de Moraes  
Teoria Assintótica: uma Abordagem em Testes de Aderência /  
Ítalo de Moraes Dolores. - 2023.  
51 f.: il.

Orientador: Valentin Sisko.  
Coorientador: Daiane de Souza Oliveira.  
Trabalho de Conclusão de Curso (graduação)-Universidade  
Federal Fluminense, Instituto de Matemática e Estatística,  
Niterói, 2023.

1. Teste de Aderência de Qui-quadrado de Pearson. 2. Teoria  
Assintótica. 3. Revisão Bibliográfica. 4. Simulação  
computacional. 5. Produção intelectual. I. Sisko, Valentin,  
orientador. II. Oliveira, Daiane de Souza, coorientadora. III.  
Universidade Federal Fluminense. Instituto de Matemática e  
Estatística. IV. Título.

CDD - XXX

# Resumo

Os testes de hipóteses são amplamente utilizados para realizar inferências na população em estudo. Em particular, os Testes de Aderência constituem parte fundamental de orientação para demais investigações sobre as distribuições de probabilidade, por isso, o conhecimento metodológico é fundamental para análises fidedignas. Neste trabalho, foi estudado o Teste de Aderência de Qui-quadrado de Pearson tanto no caso de parâmetros conhecidos quanto no caso de parâmetros estimados. Esta pesquisa, resume-se como um trabalho investigativo para fornecer informações sobre o Teste de Aderência de Qui-quadrado de Pearson, ilustrar as conclusões dos teoremas e realizar revisão literária. Após extensivas revisões literárias, notou-se a importância da diferença entre os Testes de Aderência com e sem estimação de parâmetros, por isso, realizou-se simulações computacionais para ilustrar esta diferença. Nas simulações realizadas, fica evidente a diferença entre os números de graus de liberdade das distribuições Qui-quadrado para quais as estatísticas de testes nos casos sem estimação e com estimação de parâmetros convergem assintoticamente. Pode-se concluir que os nossos resultados das simulações computacionais coincidiram com os resultados teóricos descritos na literatura.

Palavras-chave: Teste de Aderência de Qui-quadrado de Pearson. Teoria Assintótica. Revisão Bibliográfica. Simulação computacional.

# Dedicatória

Dedico esta obra a todos que contribuíram com minha formação, de alguma forma, em especial, ao meu Orientador Valentin Sisko.

# Agradecimentos

Agradeço ao meu orientador a paciência e zelo que o faz um excelente acadêmico e pesquisador, agradeço a minha coorientadora pelas contribuições valorosas, posteriormente, não menos importante, ao professor Marco Aurélio que sempre esteve disposto à contribuir para seus alunos, seja trocando experiências de vida como trazendo conhecimentos acadêmicos valorosos. Aos meus amigos mais próximos de faculdade que sempre estiveram comigo em diversos momentos bons e ruins no decorrer deste jornada, Gabriel Oliveira, Bruno Hauret, Brunno Costa, Thamires Lousada, Thaís Machado, Eliane Carnaúba, Rodrigo Mourão e João Pedro Fernandes. Aos meus amigos de longa data Breno Negrão, Guilherme Rodrigues, Sávio da Cunha e David Machado que sempre estiveram ao meu lado em diversas fases da minha vida e que nenhuma barreira pode destruir este laço de amizade. Aos meus familiares que sempre me desejaram sortes de bênçãos no decorrer da minha vida. A minha amada namorada Marlene que sempre esteve ao meu lado e que sempre me inspirou admiração pelos seus esforços. Em especial, a minha avó Elizabeth que sempre esteve presente e que nunca mediu esforços para demonstrar o quanto me ama. Aos meus amigos Gabriela, Douglas Pinheiro, Maurício e Thiago que contribuíram com inúmeras histórias e emoções ao decorrer destes anos. Aos meus amigos da Mackenzie que levei para toda a vida Pedro Henrique, Bruno Borges, Mauro Borges, Stephanny e Arthur, não menos importante deste, ao Professor Paulo Santos que descansa em paz. Agradeço a Deus, por ter me guardado e protegido e até aqui ter me sustentado. Deixo espaço para mais agradecimentos em trabalhos futuros, na pós-graduação. Para todos aqueles que estiveram comigo e sei que sempre estarão dispostos a contribuir para o meu crescimento, saibam que se precisarem, ADSUMUS!

“A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original”.

Albert Einstein

# Sumário

Lista de Figuras

Lista de Tabelas

Lista de Códigos	13
<b>1 Introdução</b>	<b>14</b>
1.1 Revisão da Literatura . . . . .	15
1.2 Objetivos . . . . .	18
1.3 Organização . . . . .	19
<b>2 Materiais e Métodos</b>	<b>20</b>
2.1 Estrutura Geral de Testes Assintóticos . . . . .	20
2.2 Teste para média com $\sigma$ conhecido . . . . .	21
2.3 Teste para média com $\sigma$ estimado . . . . .	22
2.4 Teste de aderência $\chi^2$ sem estimação de parâmetros . . . . .	23
2.4.1 Formulação do Teorema de Pearson . . . . .	23
2.4.2 Formulação do Teorema 2.4 . . . . .	24
2.4.3 Formulação do Teorema 2.5 . . . . .	25
2.5 Teste de aderência $\chi^2$ com estimação de parâmetros . . . . .	25
<b>3 Análise de Resultados</b>	<b>28</b>
3.1 Simulações Computacionais do Teste para média com $\sigma$ conhecido . . . . .	29
3.2 Simulações Computacionais do Teste para média com $\sigma$ estimado . . . . .	29
3.3 Comparação das f.d.a.s de estatísticas dos testes . . . . .	31
3.4 Simulações computacionais do Teste de Aderência de $\chi^2$ sem estimação de parâmetros . . . . .	34
3.5 Simulações computacionais do Teste de Aderência de $\chi^2$ com estimação de parâmetros . . . . .	34
3.6 Comparação das f.d.a.s de estatísticas dos testes . . . . .	39
3.7 Ilustração da convergência (2.14) feita de jeito errado . . . . .	39

3.8	Comparação das f.d.a.s de estatísticas dos testes . . . . .	45
3.9	Código para as tabelas . . . . .	45
<b>4</b>	<b>Conclusões</b>	<b>47</b>
	<b>Referências</b>	<b>49</b>
	<b>Apêndice A - Definições e Teoremas</b>	<b>51</b>

# Lista de Figuras

1	Ilustra a conclusão do Teorema 2.1, isto é, a convergência (2.3): em verde, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 1000$ amostras de tamanho $n = 1000$ . . . . .	30
2	Ilustra a conclusão do Teorema 2.1, isto é, a convergência (2.3): em verde, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 10000$ amostras de tamanho $n = 10000$ . . . . .	30
3	Ilustra a conclusão do Teorema 2.2, isto é, a convergência (2.5): em azul, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 1000$ amostras de tamanho $n = 1000$ . . . . .	32
4	Ilustra a conclusão do Teorema 2.2, isto é, a convergência (2.5): em azul, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 10000$ amostras de tamanho $n = 10000$ . . . . .	32
5	Combinação das Figuras 1 e 3 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes coincidem. . . . .	33
6	Combinação das Figuras 2 e 4 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes coincidem. . . . .	35
7	Ilustra a conclusão do Teorema 2.5, isto é, a convergência (2.11): em verde, fica a f.d.a. da Qui-quadrado com $k$ graus de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 1000$ amostras de tamanho $n = 1000$ . . . . .	35
8	Ilustra a conclusão do Teorema 2.5, isto é, a convergência (2.11): em verde, fica a f.d.a. da Qui-quadrado com $k$ graus de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 10000$ amostras de tamanho $n = 10000$ . . . . .	36

9	Ilustra a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com $k - r$ graus de liberdade, e em vermelho, a f.d.a. empírica da estatística de teste, baseada em $m = 1000$ amostras de tamanho $n = 1000$ . . . . .	38
10	Ilustra a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com $k - r$ graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 10000$ amostras de tamanho $n = 10000$ . . . . .	38
11	Combinação das Figuras 7 e 9 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes não coincidem. . . . .	41
12	Combinação das Figuras 8 e 10 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes não coincidem. . . . .	41
13	Ilustra do jeito errado a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com $k - r$ graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 1000$ amostras de tamanho $n = 1000$ . . . . .	43
14	Ilustra do jeito errado a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com com $k - r$ graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em $m = 10000$ amostras de tamanho $n = 10000$ . . . . .	43
15	Combinação das Figuras 7 e 13 para ilustrar que a f.d.a. da distribuição limite está entre a f.d.a. da distribuição $\chi_{k-r}^2$ e a f.d.a. da distribuição $\chi_k^2$ . . . . .	46
16	Combinação das Figuras 8 e 14 para ilustrar que a f.d.a. da distribuição limite está entre a f.d.a. da distribuição $\chi_{k-r}^2$ e a f.d.a. da distribuição $\chi_k^2$ . . . . .	46

# Lista de Tabelas

1	Estatísticas do Teste de Aderência de Pearson sem estimação de parâmetros do artigo de Pearson (1900) e calculados no R. . . . .	16
2	Estatísticas do Teste de Aderência de Pearson com estimação de parâmetros do artigo de Pearson (1900) e calculados no R. . . . .	16

# Lista de Códigos

1	Ajuda ilustrar a conclusão do Teorema 2.1, isto é, a convergência (2.3); gera a Figura 2. . . . .	31
2	Ajuda ilustrar a conclusão do Teorema 2.2, isto é, a convergência (2.5); gera a Figura 4. . . . .	33
3	Ajuda ilustrar a conclusão do Teorema 2.5, isto é, a convergência (2.11); gera a Figura 8. . . . .	36
4	Ajuda ilustrar a conclusão do Teorema 2.6, isto é, a convergência (2.14); gera a Figura 10. . . . .	40
5	Ajuda ilustrar do jeito errado a conclusão do Teorema 2.6, isto é, a convergência (2.14); gera a Figura 14. . . . .	44

# 1 Introdução

No decorrer dos anos, observamos à importância dos conceitos de inferência assintótica na estatística moderna, porém sua natureza abstrata dificulta a compreensão dos alunos. Para superar esse obstáculo, utilizamos o software R como uma ferramenta de simulação computacional, enriquecendo a teoria (R Core Team, 2021). Atualmente, o software R é uma das principais ferramentas computacionais para análise estatística e de dados, capaz de lidar com diversas funções, como manipulação de dados, modelagem estatística e métodos avançados. Assim, alinhados com a recente tendência de utilizar o software R para cálculos estatísticos e análise de dados, empregamos extensivamente essa ferramenta em nossa pesquisa para ilustrar as conclusões dos teoremas do Teste de Aderência de Qui-quadrado de Pearson.

Ao longo da última década, obteve-se uma percepção acentuada de que o volume massivo de dados produzidos poderia ser compreendido e estudado através de análises estatísticas, resultando na aplicabilidade de teoremas em problemas cotidianos e que, outrora, utilizavam-se em pesquisas puramente acadêmicas.

Atualmente, a Estatística é utilizada na inteligência de negócios, instrumento para o tomador de decisão em temas que versam sobre saúde pública, segurança pública, mercado financeiro, sociologia, análises espaciais e análise de risco, e conseqüentemente, é necessária a adequada utilização de métodos e procedimentos estatísticos para obter resultados fidedignos das análises nas mais diversas áreas do conhecimento.

Para a modelagem de dados nas mais diversas áreas do conhecimento é necessário que seja feita a detecção das distribuições de probabilidade que estão relacionadas à pesquisa, deste modo, o usuário das ferramentas estatísticas necessita de recursos bibliográficos e computacionais suficientes para compreender os métodos empregados no dia-a-dia da análise de dados.

Testes de aderência avaliam se a distribuição dos dados coincide com uma distribuição de referência. Há diversos testes de aderência, em particular, este presente trabalho aborda

o Teste de Aderência de Qui-quadrado de Pearson.

Uns testes de aderência bastante populares são os testes de Kolmogorov-Smirnov, Kupier, Cramér-von Mises e Anderson-Darling. Para cada um desses testes, na construção da estatística de teste, é utilizada função de distribuição acumulada empírica (FDAE) da amostra. Então esses testes são frequentemente chamados de testes baseados em FDAE, veja (DASGUPTA, 2008).

Os Testes de Aderência Qui-quadrado são concorrentes bem conhecidos dos testes baseados em FDAE. Eles discretizam a distribuição nula de algum jeito e avaliam a concordância entre as contagens observadas e as contagens postuladas, o que implica em uma perda de informação e, conseqüentemente, uma perda de poder de teste. No entanto, eles são versáteis. Ao contrário dos testes baseados em FDAE, um Teste Qui-quadrado pode ser usado tanto para dados contínuos quanto discretos, em uma dimensão ou várias dimensões. Assim, uma perda de informação está sendo trocada pela versatilidade do princípio e facilidade de cálculo.

O desenvolvimento e a construção dos teoremas formulados em (CASELLA; BERGER, 1990), (LEHMANN; ROMANO, 2005) e (MOOD, 1950) trouxeram a contribuição necessária para a compreensão das diferenças dos testes. Nos Teoremas 2.1 e 2.2 que correspondem aos Testes para a Média com  $\sigma$  conhecida e estimada, respectivamente, apesar de que as estatísticas de teste sejam diferentes já que no primeiro é utilizada a variância populacional e no segundo a variância amostral, ambas convergem assintoticamente para a mesma distribuição de probabilidade, como observado na Figura 6. Enquanto nos Teoremas 2.5 e 2.6 que correspondem aos Testes de Aderência de Qui-quadrado, fica evidente a diferença entre os números de graus de liberdade das distribuições Qui-quadrado para quais as estatísticas de testes nos casos sem estimação e com estimação de parâmetros assintoticamente convergem, como observado na Figura 12. Os Testes para a Média com  $\sigma$  conhecida e estimada estão aqui só para fazer contraste com os Testes de Aderência de Qui-quadrado sem e com estimação de parâmetros.

## 1.1 Revisão da Literatura

O Teste Qui-quadrado de Pearson tem sido amplamente celebrado como uma grande conquista na metodologia estatística. Mas também, como Stigler (2008, p. 266) observou, “praticamente nenhuma menção histórica do artigo é feita por estatísticos sem acrescentar críticas ou reprovações no sentido de que Pearson errou ao afirmar que nenhuma correção

	1	2	3
k+1	13	13	13
$\chi^2$	43,87241	17,775	172,43
P	0,000016	0,1227	14,5/10 <sup>30</sup>
$\chi_{k-1}^2$	42,03072	16,52597	164,7010
$\chi_k^2$	43,88116	17,77484	167,6400
$\chi_{k+1}^2$	45,69750	19,01304	170,5159

Tabela 1: Estatísticas do Teste de Aderência de Pearson sem estimação de parâmetros do artigo de Pearson (1900) e calculados no R.

	4	5	6	7	8
k+1	7	17	21	11	20
$\chi^2$	4,885528	23,50	36,2872	45,811	22,0422
P	0,5586	0,101	0,01423	0,00000155	0,2817
$\chi_{k-1}^2$	3,93646	22,26667	34,93263	43,77919	20,95479
$\chi_k^2$	4,885349	23,50033	36,28659	45,81385	22,05151
$\chi_{k+1}^2$	5,839587	24,72654	37,63163	47,79967	23,14564

Tabela 2: Estatísticas do Teste de Aderência de Pearson com estimação de parâmetros do artigo de Pearson (1900) e calculados no R.

em graus de liberdade precisa ser feita quando os parâmetros são estimados sob a hipótese nula.”

De acordo com Cochran (1952), na primeira parte do artigo de Pearson (1900), chega-se no resultado de que a distribuição limite da estatística de teste (quando o tamanho de amostra  $n$  tende para o infinito), com  $k + 1$  células, segue a distribuição  $\chi^2$  com  $k$  graus de liberdade.

Os Testes de Aderência de Qui-quadrado podem ser classificados em duas situações, quando os parâmetros são conhecidos e quando são estimados. Na primeira situação, quando os parâmetros são conhecidos, o trabalho desenvolvido por Pearson (1900) encontra-se mais delineado, porém, na segunda situação, quando procura-se estimar tais parâmetros, Pearson (1900) desconhece a aplicação do teste de aderência, como destacado por Cochran (1952, p. 319) que disse o seguinte: “Este assunto causou alguma confusão e controvérsia em aplicações práticas e não foi resolvido por 20 anos”. A modificação com parâmetros estimados foi desenvolvida por Fisher (1924), que claramente mostrou que a estimação de parâmetros altera a distribuição limite.

Quando Pearson (1900) escreveu seu artigo, ele deu oito exemplos do uso do Teste de

Aderência separando em duas seções, sendo três na primeira seção e os cinco exemplos restantes na segunda seção, nomeando-as em “Frequência conhecida, ou supostamente conhecida a priori” e “Frequência da população geral não conhecida a priori”, notadamente, quando o parâmetro é conhecido e estimado, respectivamente. Temos as Tabelas 1 e 2, que correspondem as essas seções, respectivamente. Os nomes das linhas das tabelas têm o seguinte significado:  $k + 1$  é o número de classes de frequência,  $\chi^2$  e  $P$  são, respectivamente, a estatística de teste e o p-valor, calculados por Pearson (1900),  $\chi_{k-1}^2$ ,  $\chi_k^2$  e  $\chi_{k+1}^2$  são os quantis superiores de distribuição Qui-quadrado que correspondem ao p-valor de Pearson para graus de liberdade,  $k - 1$ ,  $k$  e  $k + 1$ , calculados pelo R, respectivamente. Com isso, podemos observar quais graus de liberdade Pearson utilizou em cada exemplo para calcular o p-valor dele a partir da estatística de teste dele, já que isso não é colocado explicitamente nos exemplos.

Note que nos primeiros três exemplos, os números de graus de liberdade estão como no Teorema 2.5, ou seja, para o caso em que os parâmetros são conhecidos, ele está usando  $k$  graus de liberdade para  $k + 1$  classes de frequência.

Observe que a afirmação do parágrafo anterior não é completamente correta, pois ela não é válida para o terceiro exemplo, mas a razão para isso é que neste exemplo o p-valor que corresponde a estatística de teste é muito pequeno e o cálculo fica muito menos preciso do que nos outros exemplos, mas neste texto este assunto não vai ser abordado com mais detalhes.

Esta Tabela 2 merece atenção do leitor em diversos aspectos. Primeiramente, observe que a tabela extraiu os resultados do artigo de Pearson (1900) nas duas primeiras linhas. Na sequência, ele colocou esses exemplos na seção de “Frequência da população geral não conhecida a priori.”, ou seja, a tentativa era clara de ilustrar casos quando havia estimação de parâmetros. Note que nos últimos cinco exemplos dele, os números de graus de liberdade não estão como no Teorema 2.6 e isso não é apropriado, mas corrigido anos depois por Fisher (1924).

Rao e Robson (1974), apresentaram um aperfeiçoamento para o Teste de Aderência de Qui-quadrado no caso da família exponencial.

Uma questão prática chave na implementação do Teste de Aderência de Qui-quadrado, é a escolha dos intervalos de classe na reta real. É um problema difícil e, apesar de existir uma enorme quantidade de literatura sobre o assunto, não há solução clara. Algumas referências importantes sobre este difícil problema são Mann e Wald (1942), Oosterhoff (1985) e Stuart e Ord (1991).

Podemos sintetizar a cronologia do Teste de Aderência de Pearson com suas principais evoluções do seguinte modo:

- Pearson (1900) introduz seu Teste de Aderência;
- A modificação com parâmetros estimados, usando as frequências de classes, foi considerada por Fisher (1924), que corrigiu a crença errônea de que estimar os parâmetros não altera a distribuição limite da estatística de teste;
- Mann e Wald (1942), desenvolvem um método otimizado de selecionar o número de intervalos de classes nos testes de aderência;
- O artigo de Cochran (1952) contém uma discussão expositiva sobre o Teste de Aderência de Qui-quadrado, destinado aos estudantes e aos que necessitam de informações para aplicação dos testes apresentados e não para familiarizados com o rigor da teoria estatística; e
- Rao e Robson (1974) aperfeiçoaram o Teste de Aderência de Qui-quadrado no caso da família exponencial.

Como pode ser visto no decorrer da Revisão da Literatura, apesar de o Teste de Aderência de Qui-quadrado de Pearson ser amplamente conhecido no meio científico por mais de cem anos, houveram momentos da história que grandes cientistas não sabiam distinguir as diferenças da estatística de teste quando os parâmetros eram conhecidos ou estimados. Apesar de tal problemática ter sido solucionada vinte e quatro anos depois da publicação do artigo de Pearson (1900), ainda assim, cento e vinte e três anos após da publicação, podemos nos deparar com situações onde o usuário da estatística aplicada pode não saber diferenciar estas situações. Portanto, esta seção destacou os principais autores e temas ligados ao desenvolvimento do Teste de Aderência de Qui-quadrado de Pearson.

## 1.2 **Objetivos**

O objetivo geral deste trabalho é investigar a utilização do Teste de Aderência de Qui-quadrado de Pearson, e os objetivos específicos são:

- fornecer informações sobre o Teste de Aderência de Qui-quadrado de Pearson com e sem estimação de parâmetros;

- ilustrar as conclusões dos Teoremas, por meio de recursos computacionais; e
- revisar a literatura da teoria do Teste de Aderência de Qui-quadrado de Pearson.

## 1.3 Organização

Este trabalho está organizado em quatro capítulos e um apêndice. No Capítulo 1, são introduzidos o tema abordado, a revisão da literatura e objetivos gerais e específicos da pesquisa. No Capítulo 2, são apresentados os materiais utilizados e os métodos aplicados para compreender e diferenciar os Testes de Aderência de Qui-quadrado de Pearson. Na sequência, o Capítulo 3 expõe e discute os gráficos gerados a partir de simulações computacionais e os códigos das simulações. No Capítulo 4, fala-se das conclusões do trabalho realizado e suas contribuições para a comunidade acadêmica. Para fins de consulta, criou-se o apêndice de definições e teoremas utilizados nesta pesquisa.

## 2 Materiais e Métodos

Na Seção 2.1 apresentamos uma estrutura geral de Testes Assintóticos. Essa estrutura pode ser vista em (LEHMANN; ROMANO, 2005, Seção 15.6), também uma estrutura parecida está considerada em (CASELLA; BERGER, 1990, Seção 10.3.2).

### 2.1 Estrutura Geral de Testes Assintóticos

Vamos supor que os dados  $X_1, X_2, \dots, X_n$  sejam gerados a partir de alguma lei desconhecida  $P$ . A hipótese nula  $H_0$  afirma que  $P$  pertence a uma certa família de distribuições  $\mathbf{P}_0$ , enquanto a hipótese alternativa  $H_1$  conjunto vazio, e a lei desconhecida  $P$  pertence a  $\mathbf{P}$ , a união de  $\mathbf{P}_0$  e  $\mathbf{P}_1$ . Em outras palavras,  $\mathbf{P}_1$  é o complementar de  $\mathbf{P}_0$  em relação a  $\mathbf{P}$ .

Consideramos a estatística  $T_n = T_n(X_1, X_2, \dots, X_n)$  que ajuda na construção da região crítica do teste. Sob a hipótese nula, a distribuição da estatística  $T_n$ , é aproximada por uma distribuição Qui-quadrado (pelo menos nos testes apresentados aqui a distribuição é essa, em alguns outros testes assintóticos, por exemplo no teste de Kolmogorov-Smirnov, a distribuição pode ser diferente). Grandes valores de  $T_n$  rejeitam a hipótese nula. Assim, tendo escolhido uma estatística de teste adequada  $T_n$ , nosso objetivo é obter um valor crítico, digamos  $c_\alpha$ , onde  $c_\alpha$  é o  $\alpha$ -quantil superior da distribuição Qui-quadrado, de modo que o teste rejeita  $H_0$ , se e somente se,  $T_n$  excede  $c_\alpha$ . O teste assintótico baseado na distribuição Qui-quadrado pode ser justificado de seguinte jeito. Primeiro, precisamos verificar que temos

$$P \{T_n(X_1, X_2, \dots, X_n) > c_\alpha\} \rightarrow \alpha \text{ quando } n \rightarrow \infty \quad (2.1)$$

quando  $P \in \mathbf{P}_0$ . Assim obtemos a base para o teste assintótico.

Além disso, precisamos verificar alguma afirmação do tipo

$$P \{T_n(X_1, X_2, \dots, X_n) > c_\alpha\} \rightarrow 1 \text{ quando } n \rightarrow \infty \quad (2.2)$$

quando  $P \in \mathbf{P}_1$ , para que o teste apresente propriedades de poder assintótico razoáveis.

Nas Seções 2.2, 2.3, 2.4.2, 2.4.3 e 2.5 são apresentados os teoremas que demonstram as afirmações análogas a aquela que envolve (2.1). Para ver isso basta, respectivamente, comparar (2.3), (2.5), (2.10), (2.11), (2.14) com (2.1), lembrando a definição de convergência em distribuição (veja a Definição A.3) junto com o fato que se  $c_\alpha$  é  $\alpha$ -quantil superior para a distribuição de uma variável aleatória  $Y$ , então

$$1 - F_Y(c_\alpha) = P(Y > c_\alpha) = \alpha.$$

Neste trabalho não são apresentados os teoremas que demonstram as afirmações análogas a aquela que envolve (2.2). Note que precisa tomar cuidado quando está fazendo isso no caso dos Testes de Aderência de Qui-Quadrado correspondentes as Seções 2.4.3 e 2.5, já que os testes discretizam a distribuição nula e avaliam a concordância entre as contagens observadas e as contagens postuladas, o que implica em uma perda de informação e, conseqüentemente, uma perda de poder de teste.

Os Teoremas 2.1 e 2.2 que correspondem aos Testes para a Média com  $\sigma$  conhecida e estimada estão aqui para fazer contraste com os Teoremas 2.5 e 2.6 que correspondem aos Testes de Aderência de Qui-quadrado sem e com estimação de parâmetros.

## 2.2 Teste para média com $\sigma$ conhecido

Sejam

- $\mu_0$  um número real,
- $\sigma$  número real maior do que zero,
- $\mathbf{P}$  uma família de distribuições  $Q$  tais que  $\text{Var}(Y) = \sigma^2$ , sendo  $Y$  uma variável aleatória com distribuição  $Q$ , e
- $\mathbf{P}_0$  uma família de distribuições  $Q \in \mathbf{P}$ , tais que  $E(Y) = \mu_0$ , sendo  $Y$  uma variável aleatória com distribuição  $Q$ .

**Teorema 2.1.** *Sejam  $X_1, X_2, \dots$  uma seqüência de variáveis aleatórias i.i.d. cada uma com distribuição  $P \in \mathbf{P}$ . Definimos  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .*

Se  $P \in \mathbf{P}_0$ , então

$$\left( \sqrt{n} \frac{\bar{X}_n - \mu_0}{\sigma} \right)^2 \xrightarrow{d} \chi_1^2 \quad (2.3)$$

quando  $n \rightarrow \infty$ .

*Observação 2.1.* A convergência em (2.3) é convergência em distribuição, veja a Definição A.3.

*Demonstração.* Já que  $P \in \mathbf{P}_0$ , em particular,  $E(X_1) = \mu_0$ , então, pelo Teorema do Limite Central (Teorema A.5),

$$\sqrt{n} \frac{\bar{X}_n - \mu_0}{\sigma} \xrightarrow{d} N(0, 1) \quad (2.4)$$

quando  $n \rightarrow \infty$ .

De (2.4), pelo Teorema do Mapeamento Contínuo (Teorema A.1), com função  $g$  de elevar ao quadrado, temos

$$\left( \sqrt{n} \frac{\bar{X}_n - \mu_0}{\sigma} \right)^2 \xrightarrow{d} \chi_1^2 \quad \square$$

quando  $n \rightarrow \infty$ .

## 2.3 Teste para média com $\sigma$ estimado

Sejam

- $\mu_0$  um número real,
- $\mathbf{P}$  uma família de distribuições  $Q$  tais que  $0 < \text{Var}(Y) < \infty$  e  $0 < \theta_4 < \infty$ , sendo  $Y$  uma variável aleatória com distribuição  $Q$  enquanto  $\theta_4 = E(Y - \mu)^4$  e  $\mu = E(Y)$ , e
- $\mathbf{P}_0$  uma família de distribuições  $Q \in \mathbf{P}$  tais que  $E(Y) = \mu_0$ , sendo  $Y$  uma variável aleatória com distribuição  $Q$ .

**Teorema 2.2.** *Sejam  $X_1, X_2, \dots$  uma sequência de variáveis aleatórias i.i.d. cada uma com distribuição  $P \in \mathbf{P}$ .*

Seja  $S_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2}$  e  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

Se  $P \in \mathbf{P}_0$ , então

$$\left( \sqrt{n} \frac{\bar{X}_n - \mu_0}{S_n} \right)^2 \xrightarrow{d} \chi_1^2 \quad (2.5)$$

quando  $n \rightarrow \infty$ .

*Demonstração.* Vamos verificar que,  $\frac{\sigma}{S_n}$  converge para 1, em probabilidade.

Note que,  $\text{Var}(S_n^2) \rightarrow 0$  quando  $n \rightarrow \infty$ , já que pelo Exercício 5.8 de (CASELLA; BERGER, 1990), temos  $\text{Var} S_n^2 = \frac{1}{n}(\theta_4 - \frac{n-3}{n-1}\sigma^2)$ .

Usando a Desigualdade de Chebyshev, temos

$$P(|S_n^2 - \sigma^2| \geq \varepsilon) \leq \frac{E(S_n^2 - \sigma^2)^2}{\varepsilon^2} = \frac{\text{Var}(S_n^2)}{\varepsilon^2}.$$

E, portanto,  $S_n^2$  converge em probabilidade para  $\sigma^2$ .

Usando o Teorema do Mapeamento Contínuo (Teorema A.1) com função  $h(x) = \frac{\sqrt{x}}$ , do fato que  $S_n^2 \xrightarrow{P} \sigma^2$  segue que

$$\frac{S_n}{\sigma} \xrightarrow{P} 1. \quad (2.6)$$

Note que a convergência em (2.6) é convergência em probabilidade, veja a Definição A.1. Pelo Teorema do Limite Central (Teorema A.5), temos que

$$\sqrt{n} \frac{\bar{X}_n - \mu_0}{\sigma} \xrightarrow{d} N(0, 1). \quad (2.7)$$

De (2.7) e (2.6), usando o item (b) do Teorema de Slutsky (Teorema A.6), segue que

$$\sqrt{n} \frac{\bar{X}_n - \mu_0}{S_n} = \frac{\sqrt{n}(\bar{X}_n - \mu_0)/\sigma}{S_n/\sigma} \xrightarrow{d} N(0, 1). \quad (2.8)$$

De (2.8), pelo Teorema do Mapeamento Contínuo (Teorema A.1), com função  $h$  de elevar ao quadrado, temos

$$\left( \sqrt{n} \frac{\bar{X}_n - \mu_0}{S_n} \right)^2 \xrightarrow{d} \chi_1^2$$

quando  $n \rightarrow \infty$ . □

Uma demonstração de uma versão do Teorema 2.2 pode ser vista na Seção 10.3.2 de (CASELLA; BERGER, 1990).

## 2.4 Teste de aderência $\chi^2$ sem estimação de parâmetros

O Teorema 2.5 é o principal nesta seção. O Teorema 2.4 é usado na demonstração do Teorema 2.5. O Teorema 2.3 é usado na demonstração do Teorema 2.4.

### 2.4.1 Formulação do Teorema de Pearson

#### Teorema 2.3.

- *Sejam  $n$  e  $r$  números naturais, tais que  $n > 1$  e  $k > 1$ .*

- Sejam  $p_j, j = 1, \dots, k$ , números reais tais que  $0 < p_j < 1$  e  $\sum_{j=1}^k p_j = 1$ .
- Seja  $\nu = (\nu_1, \dots, \nu_k)$  um vetor aleatório.
- $\nu \sim \text{Multinomial}(n; p)$ , onde  $p := (p_1, \dots, p_k)$ .

Então

$$T_n = \sum_{j=1}^k \frac{(\nu_j - np_j)^2}{np_j} \xrightarrow{d} \chi_{k-1}^2, \quad (2.9)$$

quando  $n \rightarrow \infty$ .

*Observação 2.2.* Para a demonstração veja o item (i) do Teorema 14.3.1 de (LEHMANN; ROMANO, 2005).

## 2.4.2 Formulação do Teorema 2.4

Seja

$$\Omega = \{(\pi_1, \dots, \pi_r) \in \mathbb{R}^r : \pi_i \geq 0, \sum_{i=1}^r \pi_i \leq 1\}.$$

Assumiremos que  $(\pi_1, \dots, \pi_r)$  é um ponto interior de  $\Omega$ .

Seja  $\pi_{r+1} = 1 - \sum_{i=1}^r \pi_i$ .

Sejam

- $B_1, \dots, B_{r+1}$  números reais,
- $\mathbf{P}$  uma família de distribuições  $Q$  tais que os valores  $B_1, \dots, B_{r+1}$  são assumidos com alguns probabilidades  $p_1, \dots, p_{r+1}$ , correspondentemente, sendo

$$\sum_{i=1}^{r+1} p_i = 1,$$

- $\mathbf{P}_0$  uma família de distribuições que consiste só de uma distribuição  $Q \in \mathbf{P}$  tal que  $p_i = \pi_i$ , para  $i = 1, \dots, r + 1$ .

**Teorema 2.4.** *Sejam  $X_1, \dots, X_n$  variáveis aleatórias i.i.d. com distribuição  $P \in \mathbf{P}$ .*

- Sejam  $\nu_i = \#\{j : X_j = B_i\}$ , onde  $i = 1, \dots, r + 1$ .
- Seja

$$Q_n = \sum_{j=1}^{r+1} \frac{(\nu_j - n\pi_j)^2}{n\pi_j}.$$

Se  $P \in \mathbf{P}_0$ , então

$$Q_n \xrightarrow{d} \chi_r^2 \quad (2.10)$$

quando  $n \rightarrow \infty$ .

*Demonstração.* Observe que sob  $H_0$ , temos que  $\nu = (\nu_1, \dots, \nu_{r+1}) \sim \text{Multinomial}(n; \pi)$ , onde  $\pi = (\pi_1, \dots, \pi_{r+1})$ . Então, pelo Teorema 2.3, temos (2.10).  $\square$

### 2.4.3 Formulação do Teorema 2.5

Sejam

- $F_0$  uma função de distribuição acumulada,
- $\mathbf{P}$  uma família de distribuições que consiste de todas distribuições, e
- $\mathbf{P}_0$  uma família de distribuições que consiste de um elemento só, que é a distribuição  $Q$ , tal que, ela tem f.d.a.  $F_0$ .

**Teorema 2.5.** *Vamos supor  $X_1, \dots, X_n$  são i.i.d. variáveis aleatórias com distribuição  $P \in \mathbf{P}$ .*

- *Seja  $(v_0, \dots, v_{k+1})$  uma sequência crescente finita de números reais, onde  $v_0 = -\infty$  e  $v_{k+1} = +\infty$ .*
- *Seja  $Y_j = \#\{i | X_i \in [v_{j-1}, v_j)\}$ ,  $j = 1, \dots, k + 1$*

Se  $P \in \mathbf{P}_0$ , então

$$\sum_{j=1}^{k+1} \frac{(Y_j - (F_0(v_j) - F_0(v_{j-1}))n)^2}{(F_0(v_j) - F_0(v_{j-1}))n} \xrightarrow{d} \chi_k^2 \quad (2.11)$$

quando  $n \rightarrow \infty$ .

*Demonstração.*  $X_i \in [v_{j-1}, v_j)$  com probabilidade  $F_0(v_j) - F_0(v_{j-1})$ ,  $j = 1, \dots, k + 1$ . Então, pelo Teorema 2.4, temos (2.11).  $\square$

## 2.5 Teste de aderência $\chi^2$ com estimação de parâmetros

Na seção anterior, como pode ser visto, tratou-se de situação em que a hipótese de teste é simples, agora, consideramos a situação que surge quando a hipótese de teste é composta.

Suponha que agora  $F_0$  é especificado quanto à sua forma, mas alguns (ou talvez todos) parâmetros são deixados sem especificação. No Teorema 2.6, o novo elemento é que as probabilidades teóricas  $F_0(v_j) - F_0(v_{j-1}), j = 1, \dots, k + 1$ , não são agora calculáveis imediatamente, pois são funções de  $s$  parâmetros não especificados  $\theta_1, \theta_2, \dots, \theta_r$  ( $r < k - 1$ ). Assim devemos escreve-los

$$F_0(\theta_1, \dots, \theta_r, v_j) - F_0(\theta_1, \dots, \theta_r, v_{j-1}), \quad j = 1, \dots, k + 1. \quad (2.12)$$

Para progredir, devemos estimar  $\theta_1, \dots, \theta_r$  por algum vetor de estimadores  $\hat{\Theta}_1, \dots, \hat{\Theta}_r$ , com notação  $\hat{F}_0 = F_0(\hat{\Theta}_1, \dots, \hat{\Theta}_r)$ , temos que (2.12) fica

$$\hat{F}_0(v_j) - \hat{F}_0(v_{j-1}), \quad j = 1, \dots, k + 1. \quad (2.13)$$

Então parte esquerda de (2.11) vira parte esquerda de (2.14). Isso pode mudar a parte direita, pois agora (2.13) são variáveis aleatórias, e não é óbvio que a distribuição assintótica da parte esquerda de (2.14) será a mesma e até que terá a mesma forma que no caso de  $H_0$  simples. Sejam

- $F_0$  uma família de funções de distribuição acumuladas com  $r$  parâmetros deixados sem especificação.
- $\mathbf{P}$  uma família de distribuições que consiste de todas distribuições, e
- $\mathbf{P}_0$  uma família de distribuições  $Q \in \mathbf{P}$ , tais que, existem  $\theta_1, \dots, \theta_r$ , tais que,  $Q$  tem f.d.a.  $F_0(\theta_1, \dots, \theta_r)$ .

**Teorema 2.6.** *Vamos supor  $X_1, \dots, X_n$  são i.i.d. variáveis aleatórias com distribuição  $P \in \mathbf{P}$ .*

- *Seja  $(v_0, \dots, v_{k+1})$  uma sequência crescente finita de números reais, onde  $v_0 = -\infty$  e  $v_{k+1} = +\infty$ .*
- *Seja  $Y_j = \#\{i | X_i \in [v_{j-1}, v_j)\}$ ,  $j = 1, \dots, k + 1$ .*
- *$\hat{\Theta}_1, \dots, \hat{\Theta}_r$  são estimadores de máxima verossimilhança baseados no  $Y_1, \dots, Y_{k+1}$ .*
- *Vamos supor que  $F_0(\theta_1, \dots, \theta_r)$  satisfaz condições de regularidade gerais.*

Se  $P \in \mathbf{P}_0$ , então

$$\sum_{j=1}^{k+1} \frac{(Y_j - (\hat{F}_0(v_j) - \hat{F}_0(v_{j-1})))^2}{(\hat{F}_0(v_j) - \hat{F}_0(v_{j-1})))n} \xrightarrow{d} \chi_{k-r}^2 \quad (2.14)$$

quando  $n \rightarrow \infty$ .

*Observação 2.3.* Como pode ser visto, a explícita formulação das condições de regularidade gerais foi deixada fora do escopo deste texto. Também não verificamos se a definição de  $\mathbf{P}$  está completamente correta, isto é, possivelmente as condições de regularidade gerais devem ser usados para a família  $\mathbf{P}$  ficar mais restrita para que  $\mathbf{P}_1$  fique mais restrita também.

*Observação 2.4.* A demonstração do Teorema 2.6 pode ser vista em (KENDALL; STUART, 1961, Cap. 30).

*Observação 2.5.* No Teorema 2.6, exigimos que  $\hat{\Theta}_1, \dots, \hat{\Theta}_r$  sejam estimadores de máxima verossimilhança, contudo, pode-se flexibilizar esta exigência para que seja melhor estimador assintoticamente normal, veja a definição de BAN em (MOOD, 1950, pág. 296). Tradução de BAN (Best asymptotically normal) para o português foi retirada de (SOCIEDADE PORTUGUESA DE ESTATÍSTICA, SPE, 2022).

### 3 Análise de Resultados

O objetivo deste capítulo consiste, principalmente, em ilustrar as conclusões dos Teoremas 2.1, 2.2, 2.5 e 2.6 utilizando simulações em R. Além disso, nas Seções 3.7 e 3.8 é apresentado um erro típico que pode acontecer quando está tentando fazer a simulação para ilustrar a conclusão do Teorema 2.6.

Geralmente, cada seção deste capítulo corresponde a um teorema. Para facilitar o entendimento das etapas necessárias para ilustrar a conclusão do teorema, apresenta-se, primeiramente, um passo a passo e, na sequência, a implementação do passo a passo na linguagem R. Esta estrutura de apresentação pode ser vista também no livro (DESHMUKH; KULKARNI, 2021).

Para ilustrar as conclusões dos teoremas acima, nas simulações, geram-se  $m$  amostras de tamanho  $n$ . Geralmente, as figuras correspondentes às simulações são divididas em pares. Em cada par, ambas as figuras ilustram a conclusão do mesmo teorema sendo uma com  $m = 1000$  e  $n = 1000$  e a outra com  $m = 10000$  e  $n = 10000$ . O objetivo principal da segunda figura é ilustrar a conclusão do teorema, enquanto, o objetivo principal da primeira é servir para o leitor entender melhor os gráficos de quais funções estão presentes nas figuras, já que na segunda figura alguns dos gráficos frequentemente coincidem.

Uma recomendação prática, influenciada pelo resultado de Mann e Wald (1942), é

$$k + 1 = 2n^{\frac{2}{5}}.$$

A recomendação parece produzir valores de  $k + 1$  que concordam bem com escolhas de  $k + 1$ .

DasGupta (2008, Seção 27.5) faz as seguintes observações. Esses valores parecem estar próximos da prática comum. O ponto importante é que  $k + 1$  deve ser maior quando  $n$  é grande. Mas não é recomendado que use-se um valor muito grande para  $k + 1$ , e uma escolha na faixa de 5 a 15 parece correta.

Por isso, e também para as figuras das Seções 3.7 e 3.8 ilustrar melhor o erro cometido,

foi escolhido  $k + 1 = 6$  para todos os tamanhos das amostras.

### 3.1 Simulações Computacionais do Teste para média com $\sigma$ conhecido

O passo a passo para simular a conclusão do Teorema 2.1 pode ser visto a seguir:

1. geram-se  $m$  amostras de tamanho  $n$  de distribuição Gama com parâmetros  $\alpha = 4$  e  $\beta = 2$ ;
2. a partir de cada amostra, obtemos a estatística de teste, isto é, a parte esquerda de (2.3);
3. usando a amostra de tamanho  $m$  para a estatística de teste obtida no item anterior, construímos a função de distribuição acumulada empírica da estatística de teste;
4. é feito o gráfico da f.d.a. empírica;
5. é feito o gráfico da f.d.a. de Qui-quadrado com um grau de liberdade, isto é, a parte direita de (2.3).

Nas Figuras 1 e 2 pode ser visto, em verde, f.d.a. da Qui-quadrado com um grau de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste. Na Figura 2 pode ser visto para onde convergiu a distribuição da estatística de teste, quando o tamanho da amostra e o número de amostras aumentam. Essa distribuição convergiu para a distribuição Qui-quadrado com um grau de liberdade, já que na Figura 2 pode observar que a f.d.a. empírica da estatística de teste coincide com a f.d.a. teórica da distribuição Qui-quadrado com um grau de liberdade.

O passo a passo acima é implementado em R como o Código 1.

### 3.2 Simulações Computacionais do Teste para média com $\sigma$ estimado

O passo a passo para simular a conclusão do Teorema 2.2 pode ser visto a seguir:

1. geram-se  $m$  amostras de tamanho  $n$  de distribuição Gama com parâmetros  $\alpha = 4$  e  $\beta = 2$ ;

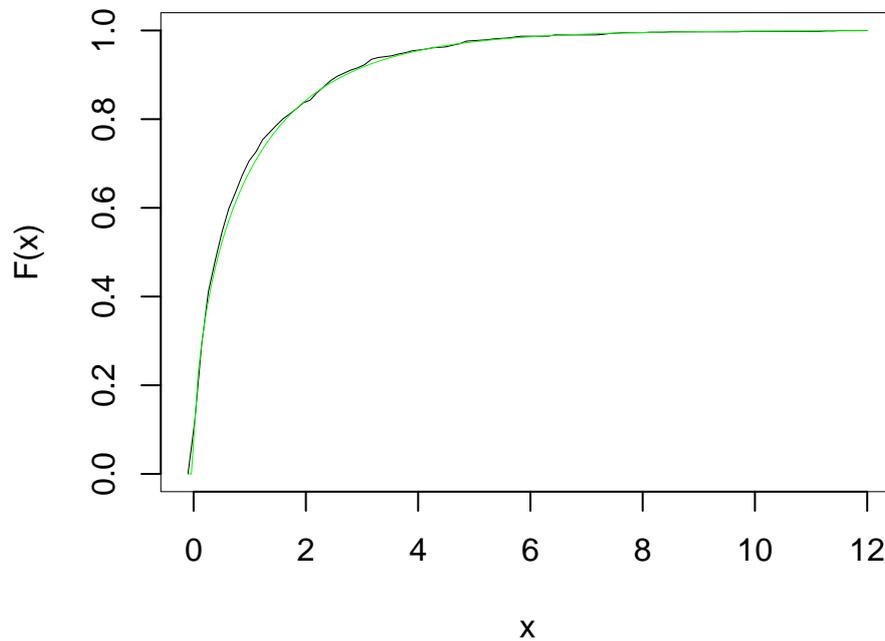


Figura 1: Ilustra a conclusão do Teorema 2.1, isto é, a convergência (2.3): em verde, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 1000$  amostras de tamanho  $n = 1000$ .

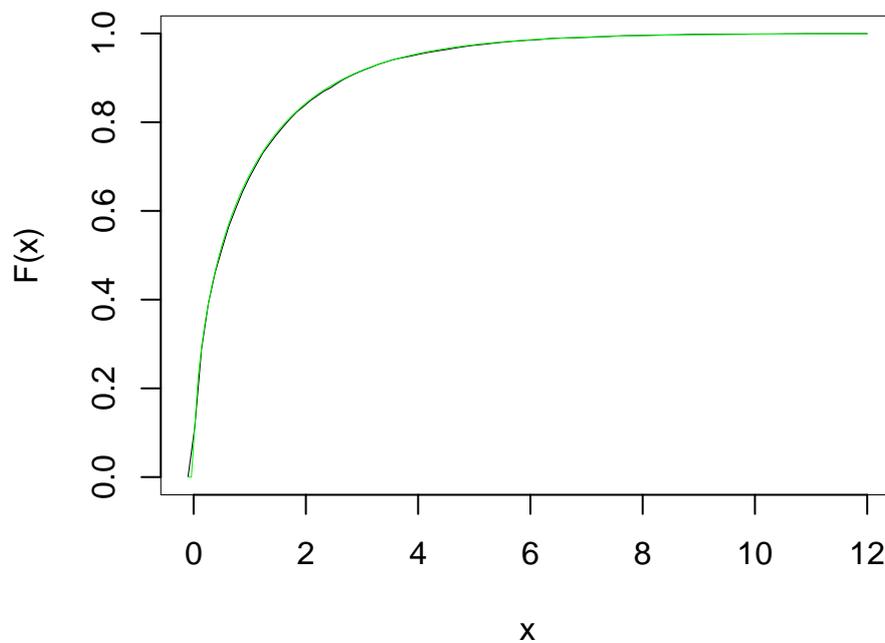


Figura 2: Ilustra a conclusão do Teorema 2.1, isto é, a convergência (2.3): em verde, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 10000$  amostras de tamanho  $n = 10000$ .

```

n <- 10000
m <- 10000
fb <- function(x) {
  y <- rgamma(n,4,2)
  X <- (sqrt(n))*((mean(y)-2)/sqrt(4/(2**2)))*2
  return(X)
}
b <- sapply(1:m, fb)
fdaeb <- ecdf(b)
curve(fdaeb,main="",ylab="F(x)", xlim=c(-0.1,12))
fc1 <- function(x) pchisq(x,df=1)
curve(fc1, add=T, n=200,col="green")

```

Código 1: Ajuda ilustrar a conclusão do Teorema 2.1, isto é, a convergência (2.3); gera a Figura 2.

2. a partir de cada amostra, obtemos a estatística de teste, isto é, a parte esquerda de (2.5);
3. usando a amostra de tamanho  $m$  para a estatística de teste obtida no item anterior, construímos a função de distribuição acumulada empírica da estatística de teste;
4. é feito o gráfico da f.d.a. empírica; e
5. é feito o gráfico da f.d.a. de Qui-quadrado com um grau de liberdade, isto é, a parte direita de (2.5).

Nas Figuras 3 e 4 pode ser visto, em azul, f.d.a. da Qui-quadrado com um grau de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste.

Na Figura 4 pode ser visto para onde convergiu a distribuição da estatística de teste, quando o tamanho da amostra e o número de amostras aumentam. Essa distribuição convergiu para a distribuição Qui-quadrado com um grau de liberdade, já que na Figura 4 pode observar que a f.d.a. empírica da estatística de teste coincide com a f.d.a. teórica da distribuição Qui-quadrado com um grau de liberdade.

O passo a passo acima é implementado em R como o Código 2.

### 3.3 Comparação das f.d.a.s de estatísticas dos testes

Na Figura 5 juntamos as Figuras 1 e 3. Na Figura 6 juntamos as Figuras 2 e 4. A figura ilustra que as distribuições limites de estatísticas de ambos os testes coincidem.

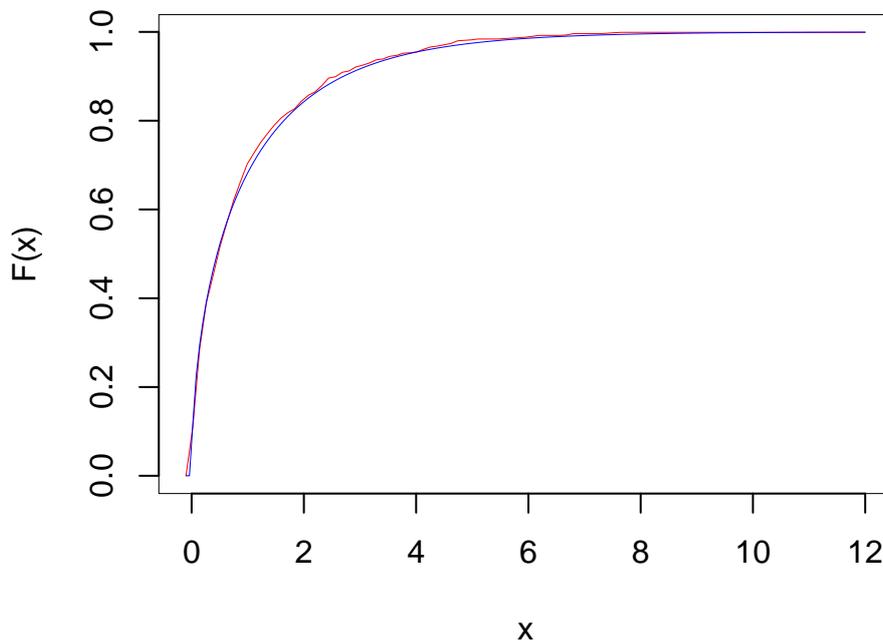


Figura 3: Ilustra a conclusão do Teorema 2.2, isto é, a convergência (2.5): em azul, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 1000$  amostras de tamanho  $n = 1000$ .

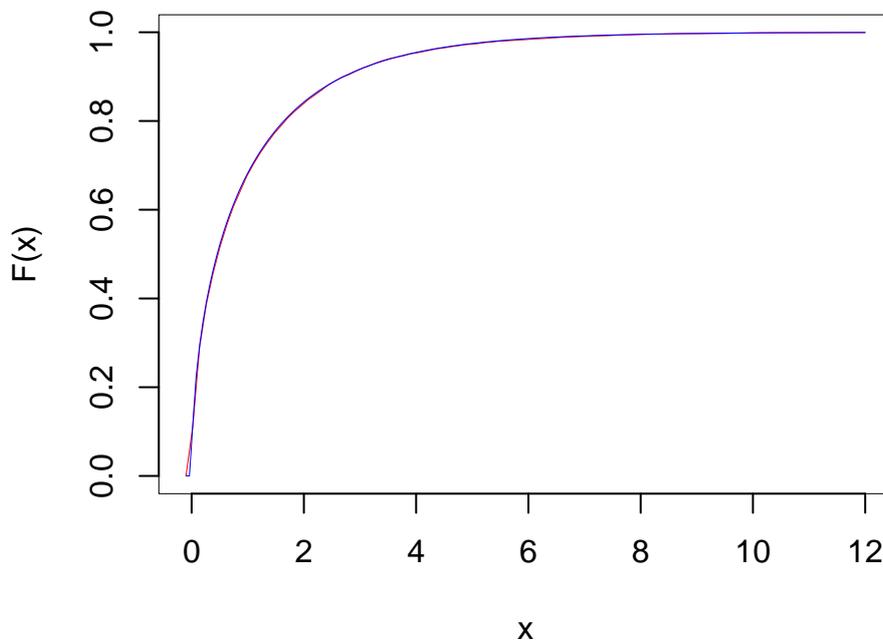


Figura 4: Ilustra a conclusão do Teorema 2.2, isto é, a convergência (2.5): em azul, fica a f.d.a. da Qui-quadrado com um grau de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 10000$  amostras de tamanho  $n = 10000$ .

```

n <- 10000
m <- 10000
fa <- function(x) {
  y <- rgamma(n,4,2)
  X <- (sqrt(n)*((mean(y)-2)/sd(y)))**2
  return(X)
}
a <- sapply(1:m, fa)
fdaea <- ecdf(a)
curve(fdaea, main="", col="red", ylab="F(x)", xlim=c(-0.1,12))
fc1 <- function(x) pchisq(x,df=1)
curve(fc1, add=T, n=200,col="blue")

```

Código 2: Ajuda ilustrar a conclusão do Teorema 2.2, isto é, a convergência (2.5); gera a Figura 4.

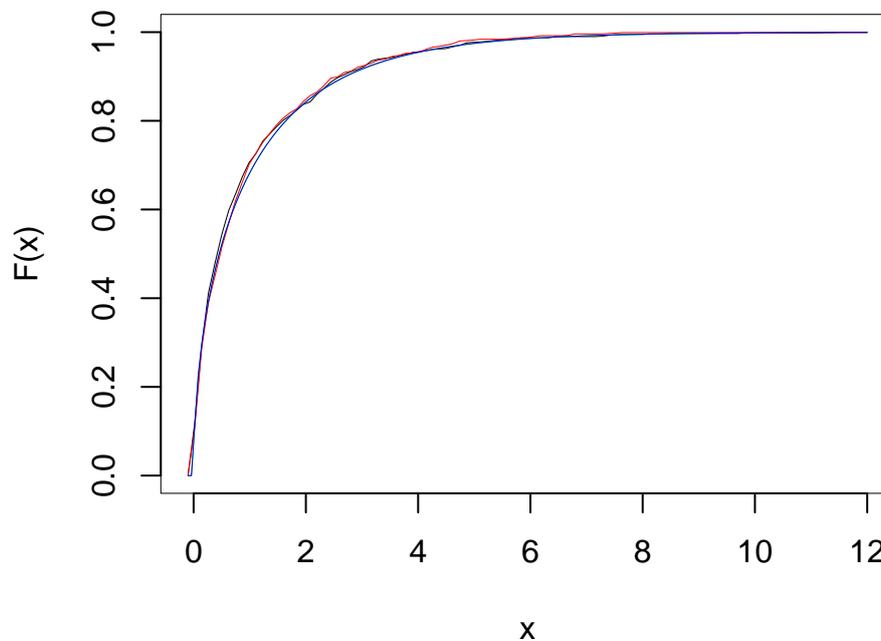


Figura 5: Combinação das Figuras 1 e 3 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes coincidem.

### 3.4 Simulações computacionais do Teste de Aderência de $\chi^2$ sem estimação de parâmetros

O passo a passo para simular a conclusão do Teorema 2.5 pode ser visto a seguir:

1. geram-se  $m$  amostras de tamanho  $n$  de distribuição Normal Padrão;
2. a reta é dividida nos seguintes intervalos:  $(-\infty; -3)$ ,  $[-3; -1,5)$ ,  $[-1,5; 0)$ ,  $[0; 1,5)$ ,  $[1,5; 3)$ ,  $[3; +\infty)$ ;
3. para cada amostra, contamos o número de observações em cada intervalo;
4. a partir desses números, obtemos a estatística de teste, isto é, a parte esquerda de (2.11), onde  $F_0$  é a f.d.a. da distribuição Normal Padrão;
5. usando a amostra de tamanho  $m$  para a estatística de teste obtida no item anterior, construímos a função de distribuição acumulada empírica da estatística de teste;
6. é feito o gráfico da f.d.a. empírica; e
7. é feito o gráfico da f.d.a. de Qui-quadrado com  $k$  graus de liberdade, isto é, a parte direita de (2.11).

Nas Figuras 7 e 8 pode ser visto, em verde, f.d.a. da Qui-quadrado com  $k$  graus de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste. Na Figura 7 pode ser visto para onde convergiu a distribuição da estatística de teste, quando o tamanho da amostra e o número de amostras aumentam. Essa distribuição convergiu para a distribuição Qui-quadrado com  $k$  graus de liberdade, já que na Figura 8 pode observar que a f.d.a. empírica da estatística de teste coincide com a f.d.a. teórica da distribuição Qui-quadrado com  $k$  graus de liberdade.

O passo a passo acima é implementado em R como o Código 3.

### 3.5 Simulações computacionais do Teste de Aderência de $\chi^2$ com estimação de parâmetros

Supondo que  $Y_1 = n_1, \dots, Y_{k+1} = n_{k+1}$ , na simulação calculamos  $\hat{\Theta}_1, \dots, \hat{\Theta}_r$ , que são estimadores de máxima verossimilhança baseados no  $Y_1, \dots, Y_{k+1}$ . Para fazer isso, usaremos a seguinte função de verossimilhança

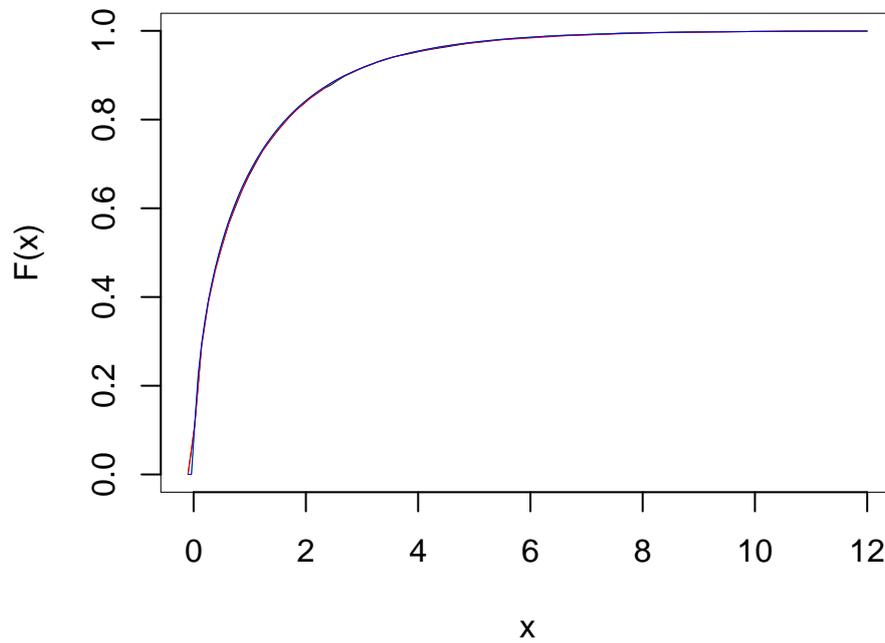


Figura 6: Combinação das Figuras 2 e 4 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes coincidem.

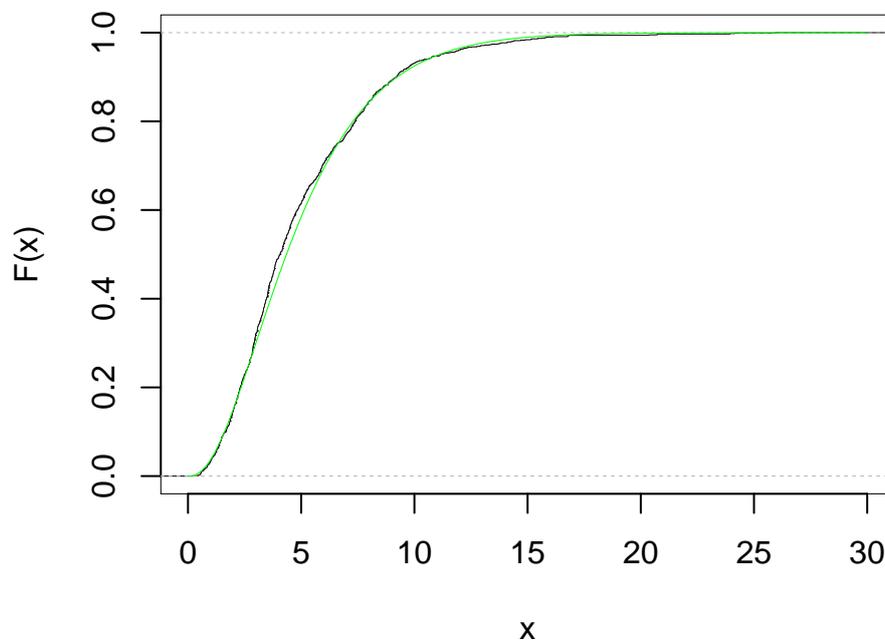


Figura 7: Ilustra a conclusão do Teorema 2.5, isto é, a convergência (2.11): em verde, fica a f.d.a. da Qui-quadrado com  $k$  graus de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 1000$  amostras de tamanho  $n = 1000$ .

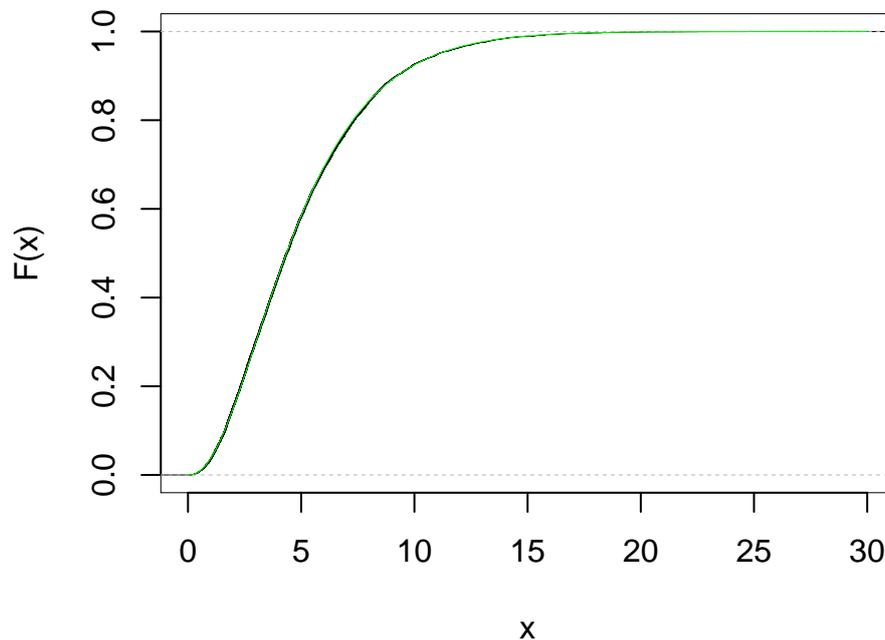


Figura 8: Ilustra a conclusão do Teorema 2.5, isto é, a convergência (2.11): em verde, fica a f.d.a. da Qui-quadrado com  $k$  graus de liberdade, em preto, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 10000$  amostras de tamanho  $n = 10000$ .

```
n <- 10000
m <- 10000
v <- c(-Inf, -3, -1.5, 0, 1.5, 3, Inf)
E <- diff(pnorm(v))*n
func <- function(x) {
  O <- as.vector(table(cut(x, breaks = v)))
  sum((O-E)^2/E)
}
fa <- function(x) {
  y <- rnorm(n)
  X <- func(y)
  return(X)
}
a <- sapply(1:m, fa)
fdaea <- ecdf(a)
plot(fdaea,main="",ylab = "F(x)", xlim=c(0,30), do.points=F)
fc <- function(x) pchisq(x, df=length(v)-2)
curve(fc, add=T, n=200,col="green")
```

Código 3: Ajuda ilustrar a conclusão do Teorema 2.5, isto é, a convergência (2.11); gera a Figura 8.

$$L = L(\theta_1, \dots, \theta_r | n_1, \dots, n_{k+1}) = \frac{n!}{n_1! n_2! \dots n_{k+1}!} p_1^{n_1} p_2^{n_2} \dots p_{k+1}^{n_{k+1}}, \quad (3.1)$$

sendo

$$p_i(\theta_1, \dots, \theta_r) = F_0(\theta_1, \dots, \theta_r, v_j) - F_0(\theta_1, \dots, \theta_r, v_{j-1}), \quad i = 1, \dots, k+1. \quad (3.2)$$

Em princípio, da (3.1) pegando o logaritmo e derivando, podemos obter

$$\frac{\partial \log L}{\partial \theta_j} = \sum_{i=1}^k n_i \frac{\partial p_i}{\partial \theta_j} \frac{1}{p_i}, \quad j = 1, 2, \dots, r. \quad (3.3)$$

E os estimadores Máxima Verossimilhança neste caso regular são as raízes das  $r$  equações obtidas igualando (3.3) a zero para cada  $j$ .

Mas na nossa simulação, seguimos um outro caminho, encontramos os estimadores ML a partir da (3.1) por meio de otimização numérica, da função `optim` como visto no código abaixo.

O passo a passo para simular a conclusão do Teorema 2.6 pode ser visto a seguir:

1. geram-se  $m$  amostras de tamanho  $n$  de distribuição Normal Padrão;
2. a reta é dividida nos seguintes intervalos:  $(-\infty; -3)$ ,  $[-3; -1,5)$ ,  $[-1,5; 0)$ ,  $[0; 1,5)$ ,  $[1,5; 3)$ ,  $[3; +\infty)$ ;
3. para cada amostra, contamos o número de observações em cada intervalo;
4. a partir desses números, obtemos a estatística de teste, isto é, a parte esquerda de (2.14), onde  $\hat{F}_0$  é a f.d.a. da distribuição Normal com os parâmetros obtidos do seguinte jeito: eles são os argumentos que maximizam a função  $L$  de (3.1);
5. usando a amostra de tamanho  $m$  para a estatística de teste obtida no item anterior, construímos a função de distribuição acumulada empírica da estatística de teste;
6. é feito o gráfico da f.d.a. empírica; e
7. é feito o gráfico da f.d.a. de Qui-quadrado com  $k - r$  graus de liberdade, isto é, a parte direita de (2.14).

Nas Figuras 9 e 10 pode ser visto, em azul, f.d.a. da Qui-quadrado com  $k - r$  graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste.

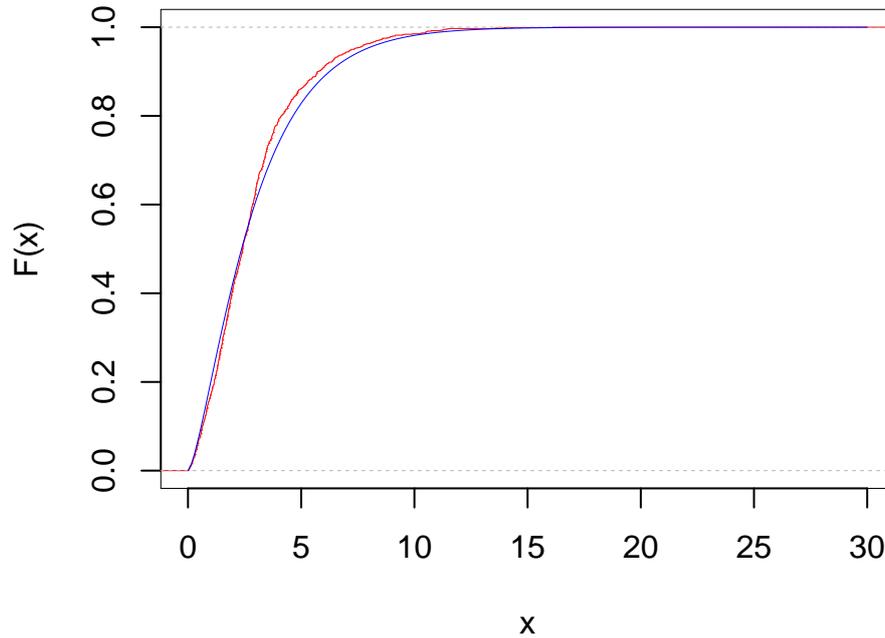


Figura 9: Ilustra a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com  $k - r$  graus de liberdade, e em vermelho, a f.d.a. empírica da estatística de teste, baseada em  $m = 1000$  amostras de tamanho  $n = 1000$ .

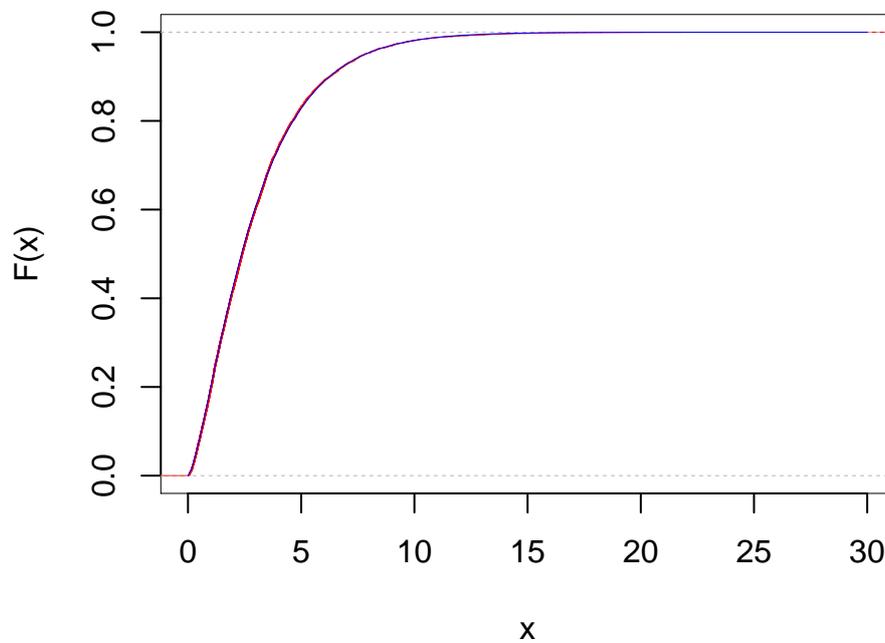


Figura 10: Ilustra a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com  $k - r$  graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 10000$  amostras de tamanho  $n = 10000$ .

Na Figura 10 pode ser visto para onde convergiu a distribuição da estatística de teste, quando o tamanho da amostra e o número de amostras aumentam. Essa distribuição convergiu para a distribuição Qui-quadrado com  $k - r$  graus de liberdade, já que na Figura 10 pode observar que a f.d.a. empírica da estatística de teste coincide com a f.d.a. teórica da distribuição Qui-quadrado com  $k - r$  graus de liberdade.

O passo a passo acima é implementado em R como o Código 4. Note que para implementar o item 4 do passo a passo a função `optim` é utilizada. Define-se no código a função a ser maximizada com sinal negativo, já que `optim`, por padrão, minimiza a função. Também, não procuramos os argumentos que maximizam a função  $L$ , mas procuramos os argumentos que maximizam logaritmo dela, o que é equivalente. Além disso, para o cálculo do logaritmo da parte direita de (3.1) a função `dmultinom` é usada com o argumento `log=T`. Finalmente, uma verificação é feita se o algoritmo usado na função `optim` convergiu. Nas simulações feitas aqui ele sempre estava convergindo.

### 3.6 Comparação das f.d.a.s de estatísticas dos testes

Na Figura 11 juntamos as Figuras 7 e 9. Na Figura 12 juntamos as Figuras 8 e 10. A figura ilustra que as f.d.a's das distribuições limites das estatísticas dos testes não coincidem.

### 3.7 Ilustração da convergência (2.14) feita de jeito errado

Na seção anterior, foram estimados  $\hat{\Theta}_1, \dots, \hat{\Theta}_r$  usando as estatísticas  $Y_1, \dots, Y_{k+1}$  no lugar de  $X_1, \dots, X_n$ . Na prática, frequentemente os valores de  $X_i$ 's não são gravados, e então  $Y_1, \dots, Y_{k+1}$  são a única informação que pode ser utilizada.

Por outro lado, se as observações  $X_1, \dots, X_n$  são gravadas, então podemos estimar  $\theta_i$ ,  $i = 1, \dots, r$ , mais eficientemente usando Estimadores de Máxima Verossimilhança baseados em  $X_1, \dots, X_n$ . Segundo Kendall e Stuart (1961, Seção 30.19), quando tais estimadores são usados, a distribuição limite da parte esquerda de (2.14) não é mais  $\chi_{k-r}^2$ , invés disso, a f.d.a. dela está entre a f.d.a. da distribuição  $\chi_{k-r}^2$  e a f.d.a. da distribuição  $\chi_k^2$ . Conforme  $k + 1$  se torna grande, elas estão tão próximos que a diferença pode ser ignorada. Mas para  $k + 1$  pequeno, o efeito de usar a distribuição  $\chi_{k-r}^2$  para fins de teste pode levar a erro grave: pois a probabilidade de exceder qualquer valor dado

```

n <- 10000
m <- 10000
v <- c(-Inf, -3, -1.5, 0, 1.5, 3, Inf)
E <- diff(pnorm(v))*n
funce <- function(x) {
  O <- as.vector(table(cut(x, breaks = v)))
  L <- function(theta) {
    mu <- theta[1]
    sigmasq <- theta[2]
    p <- diff(pnorm(v,mean=mu,sd=sqrt(sigmasq)))
    -dmultinom(O, prob=p, log = T)
  }
  qaux <- optim(c(mean(x),var(x)), L)
  q <- qaux$par
  if(qaux$convergence!=0) stop("nao convergiu")
  E <- diff(pnorm(v,mean=q[1],sd=sqrt(q[2])))*n
  sum((O-E)^2/E)
}
fb <- function(x) {
  y <- rnorm(n)
  X <- funce(y)
  return(X)
}
l <- numeric(m)
k <- 1
b <- sapply(1:m, fb)
fdaeb <- ecdf(b)
plot(fdaeb, main="", col="red", ylab = "F(x)", xlim=c(0,30), do.points=F)
fd <- function(x) pchisq(x, df=(length(v)-2)-2)
curve(fd, add=T, n=200,col="blue")

```

Código 4: Ajuda ilustrar a conclusão do Teorema 2.6, isto é, a convergência (2.14); gera a Figura 10.

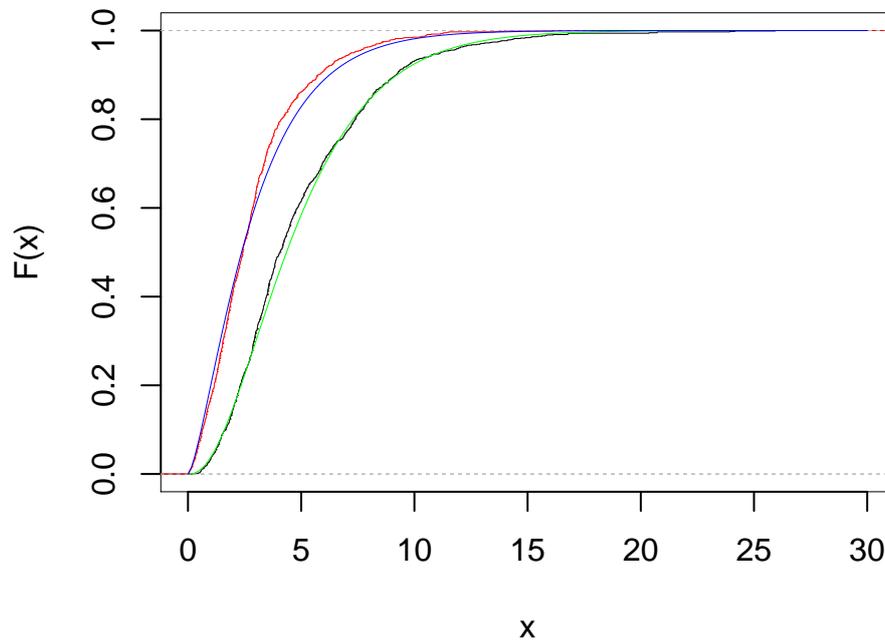


Figura 11: Combinação das Figuras 7 e 9 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes não coincidem.

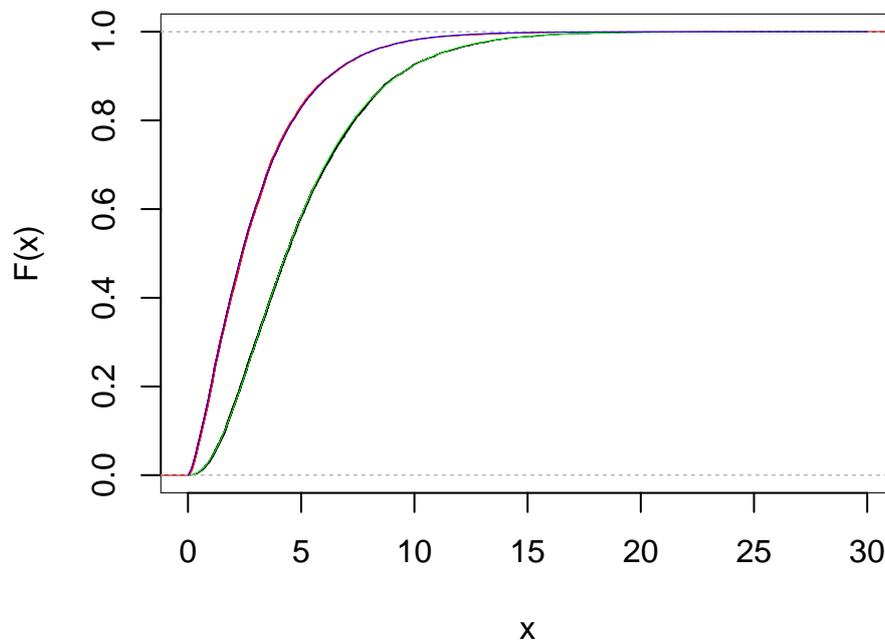


Figura 12: Combinação das Figuras 8 e 10 para ilustrar que as f.d.a's das distribuições limites das estatísticas dos testes não coincidem.

será maior do que supomos. Ele também faz uma sugestão como prosseguir, na prática, quando nos deparamos com uma situação dessa. Voinov, Nikulin e Balakrishnan (2013) observaram que “O uso errôneo do teste de Pearson sob essa situação é reproduzido mesmo em alguns livros recentes” e deram como exemplos de livros com este tipo de erro dois livros: (CLARK, 1997, pág. 273) e (WEIERS, 1991, pág. 602).

O passo a passo para simular a conclusão do Teorema 2.6 do jeito errado pode ser visto a seguir:

1. geram-se  $m$  amostras de tamanho  $n$  de distribuição Normal Padrão;
2. a reta é dividida nos seguintes intervalos:  $(-\infty; -3)$ ,  $[-3; -1,5)$ ,  $[-1,5; 0)$ ,  $[0; 1,5)$ ,  $[1,5; 3)$ ,  $[3; +\infty)$ ;
3. para cada amostra, contamos o número de observações em cada intervalo;
4. a partir desses números, obtemos a estatística de teste, isto é, a parte esquerda de (2.14), observe que nessa fórmula usamos estimadores de parâmetros obtidos pelo método de máxima verossimilhança diretamente a partir das observações geradas no primeiro item, isto é, usamos média amostral e variância amostral com  $n$  e não  $(n - 1)$ ;
5. usando a amostra de tamanho  $m$  para a estatística de teste obtida no item anterior, construímos a função de distribuição acumulada empírica da estatística de teste;
6. é feito o gráfico da f.d.a. empírica; e
7. é feito o gráfico da f.d.a. de Qui-quadrado com  $k - r$  graus de liberdade, isto é, a parte direita de (2.14).

Nas Figuras 13 e 14 pode ser visto, em azul, f.d.a. da Qui-quadrado com  $k - r$  graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste.

Na Figura 14 pode ser visto para onde convergiu a distribuição da estatística de teste, quando o tamanho da amostra e o número de amostras aumentam. Essa distribuição não convergiu para a distribuição Qui-quadrado com  $k - r$  graus de liberdade, já que na Figura 14 pode observar que a f.d.a. empírica da estatística de teste não coincide com a f.d.a. teórica da distribuição Qui-quadrado com  $k - r$  graus de liberdade.

O passo a passo acima é implementado em R como o Código 5.

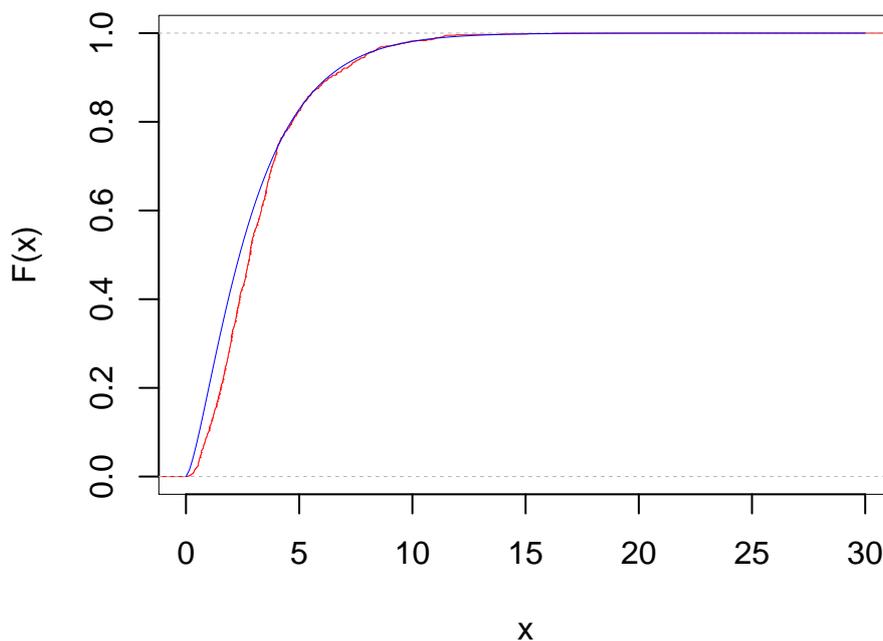


Figura 13: Ilustra do jeito errado a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com  $k-r$  graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 1000$  amostras de tamanho  $n = 1000$ .

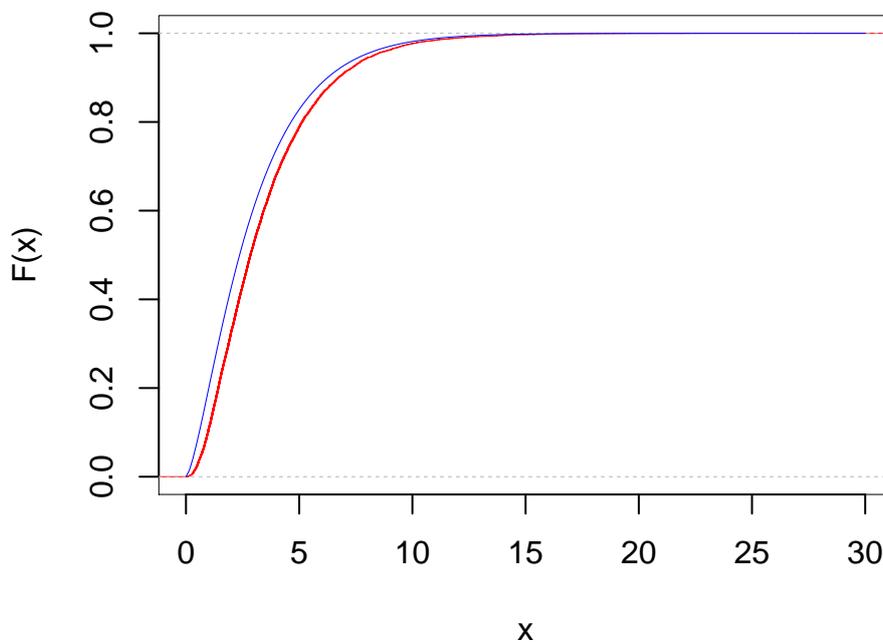


Figura 14: Ilustra do jeito errado a conclusão do Teorema 2.6, isto é, a convergência (2.14): em azul, fica a f.d.a. da Qui-quadrado com com  $k-r$  graus de liberdade, em vermelho, encontra-se a f.d.a. empírica da estatística de teste, baseada em  $m = 10000$  amostras de tamanho  $n = 10000$ .

```

n <- 10000
m <- 10000
v <- c(-Inf, -3, -1.5, 0, 1.5, 3, Inf)
E <- diff(pnorm(v))*n
funce <- function(x) {
  O <- as.vector(table(cut(x, breaks = v)))
  E <- diff(pnorm(v, mean=mean(x), sd=sqrt((n-1)*var(x)/n)))*n
  sum((O-E)^2/E)
}
fb <- function(x) {
  y <- rnorm(n)
  X <- funce(y)
  return(X)
}
l <- numeric(m)
k <- 1
b <- sapply(1:m, fb)
fdaeb <- ecdf(b)
plot(fdaeb, main="", col="red", ylab = "F(x)", xlim=c(0,30), do.points = F)
fd <- function(x) pchisq(x, df=(length(v)-2)-2)
curve(fd, add=T, n=200, col="blue")

```

Código 5: Ajuda ilustrar do jeito errado a conclusão do Teorema 2.6, isto é, a convergência (2.14); gera a Figura 14.

## 3.8 Comparação das f.d.a.s de estatísticas dos testes

Na Figura 15 juntamos as Figuras 7 e 13. Na Figura 16 juntamos as Figuras 8 e 14. A Figura 16 ilustra a observação de Kendall e Stuart (1961, Seção 30.19) que a f.d.a. da distribuição limite está entre a f.d.a. da distribuição  $\chi_{k-r}^2$  e a f.d.a. da distribuição  $\chi_k^2$ .

## 3.9 Código para as tabelas

Nesta seção, é apresentado o código em R que ajuda a preencher as linhas com os nomes  $\chi_{k-1}^2$ ,  $\chi_k^2$  e  $\chi_{k+1}^2$  das Tabelas 1 e 2.

A explicação é feita usando um exemplo. Neste exemplo, é calculado o número que fica na Tabela 1, na linha  $\chi_{k-1}^2$  e na coluna que corresponde ao segundo exemplo de Pearson.

Depois de executar o comando `qchisq(0.1227,df=11, lower.tail = F)` em R, o resultado 16,52597 é obtido.

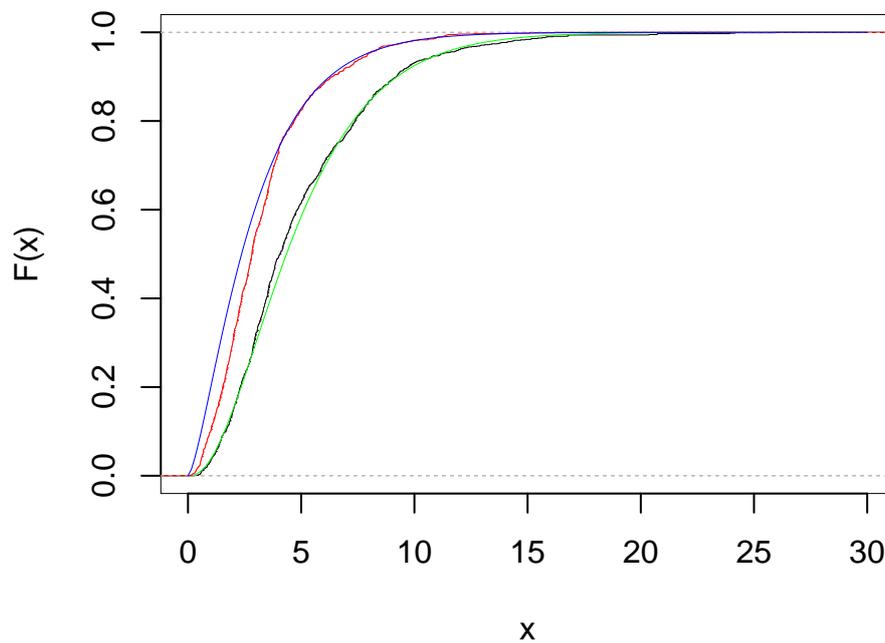


Figura 15: Combinação das Figuras 7 e 13 para ilustrar que a f.d.a. da distribuição limite está entre a f.d.a. da distribuição  $\chi_{k-r}^2$  e a f.d.a. da distribuição  $\chi_k^2$ .

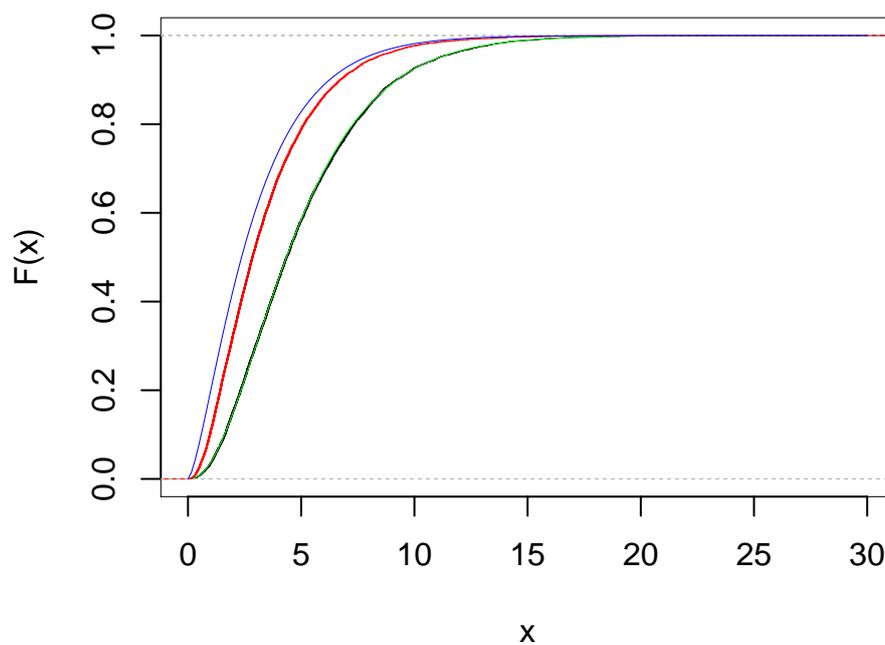


Figura 16: Combinação das Figuras 8 e 14 para ilustrar que a f.d.a. da distribuição limite está entre a f.d.a. da distribuição  $\chi_{k-r}^2$  e a f.d.a. da distribuição  $\chi_k^2$ .

## 4 Conclusões

Este trabalho apresentou um estudo teórico e simulado dos Testes de Aderência de Qui-quadrado, quando há estimação dos parâmetros e quando não há estimação. Neste texto, foram abordados, inicialmente, os Testes para a Média com  $\sigma$  conhecida e estimada só para fazer contraste com os Testes de Aderência de Qui-quadrado, sugerindo uma razão porque as pessoas podem errar. Pois, nos Testes para a Média, que as pessoas conhecem melhor, apesar de que as estatísticas de teste sejam diferentes, já que no primeiro é utilizada a variância populacional e no segundo a variância amostral, ambas convergem assintoticamente para a mesma distribuição de probabilidade. Enquanto nos Testes de Aderência de Qui-quadrado existe a diferença entre os números de graus de liberdade das distribuições Qui-quadrado para quais as estatísticas de testes nos casos sem estimação e com estimação de parâmetros assintoticamente convergem.

Ao longo da Revisão Literária, ao se tratar do Teste de Aderência de Qui-quadrado de Pearson, observamos que não foi tão trivial para grandes pesquisadores na área da estatística, que haviam diferenças entre os casos em que os parâmetros eram estimados ou conhecidos. Por conta disto, no decorrer desta parte do trabalho, foram realizadas diversas menções de autores renomados no tocante à importância desta diferenciação, como também ilustramos estas diferenças. Podemos perceber, cronologicamente, que o Teste de Aderência de Qui-quadrado de Pearson sofreu críticas e aperfeiçoamento por décadas. No capítulo “Materiais e Métodos”, apresentamos os teoremas, para as conclusões de quais as simulações computacionais descritas no capítulo “Análise de resultados” são feitas, e, com o uso das simulações computacionais, conseguimos distinguir, visualmente, que as distribuições são distintas quando o Teste de Aderência de Qui-quadrado tem ou não estimação de parâmetros.

Este trabalho se propôs a investigar à utilização do Teste de Aderência de Qui-quadrado de Pearson, no qual obteve êxito ao se basear nas revisões literárias, material teórico e computacional para trazer clareza ao leitor.

No decorrer das investigações, como resposta à principal abordagem desta pesquisa, verificou-se graficamente para onde convergiu a estatística de teste para o Teste de Aderência de Qui-quadrado de Pearson com estimação de parâmetros, com o intuito do leitor perceber, visualmente, as diferenças não realizadas por Pearson (1900), então, pode-se observar que há diferença nos graus de liberdade quando é necessário estimar parâmetros no Teste de Aderência de Qui-quadrado de Pearson.

Dito isto, conclui-se de forma sumarizada, que este trabalho fornece ao leitor as conclusões necessárias para refutar qualquer tentativa, desapercibida, de algum usuário do Teste de Aderência de Qui-quadrado de Pearson utilizar o teste supondo que não faz diferença nos graus de liberdade ao se deparar com a necessidade de estimar os parâmetros.

Como sugestão para trabalhos futuros, destacamos: um estudo onde se aprofunda a discussão, simulação e apresentação dos resultados abordados por Pearson (1900) com apresentações intuitivas e linguagem contemporânea; a ampliação do escopo de estudo para outros testes de aderência utilizando referenciais teóricos e simulações computacionais; e, posteriormente, comparar os Testes de Aderência e verificar se alguma estatística de teste converge, em lei, mais rápido que as demais.

# Referências

- CASELLA, G.; BERGER, R. L. *Statistical inference*. [S.l.]: Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1990. xviii+650 p. (The Wadsworth & Brooks/Cole Statistics/Probability Series).
- CLARK, J. *Business Statistics*. [S.l.]: Barron's Educational Series, Inc., 1997.
- COCHRAN, W. G. The  $\chi^2$  test of goodness of fit. *Ann. Math. Statistics*, v. 23, p. 315–345, 1952.
- DASGUPTA, A. *Asymptotic theory of statistics and probability*. [S.l.]: Springer, New York, 2008. xxviii+722 p. (Springer Texts in Statistics).
- DESHMUKH, S.; KULKARNI, M. *Asymptotic statistical inference—a basic course using R*. [S.l.]: Springer, Singapore, 2021. xviii+529 p.
- FISHER, R. A. The conditions under which  $\chi^2$  measures the discrepancy between observations and hypothesis. *Journal Royal Statist.*, v. 87, p. 442–450, 1924.
- KENDALL, M. G.; STUART, A. *The advanced theory of statistics. Vol. 2*. First. [S.l.]: Hafner Publishing Co., New York, 1961. ix+676 p. Inference and relationship.
- LEHMANN, E. L.; ROMANO, J. P. *Testing statistical hypotheses*. Third. [S.l.]: Springer, New York, 2005. xiv+784 p. (Springer Texts in Statistics).
- MANN, H. B.; WALD, A. On the choice of the number of class intervals in the application of the chi square test. *Ann. Math. Statistics*, v. 13, p. 306–317, 1942.
- MOOD, A. M. *Introduction to the Theory of Statistics*. [S.l.]: McGraw-Hill Book Co., Inc., New York-Toronto-London, 1950. xiv+433 p.
- OOSTERHOFF, J. The choice of cells in chi-square tests. *Statist. Neerlandica*, v. 39, n. 2, p. 115–128, 1985.
- PEARSON, K. On the criterion, that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Phil. Mag. (5)*, v. 50, p. 157–175, 1900.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <https://www.R-project.org/>.
- RAO, K. C.; ROBSON, D. S. A chi-square statistic for goodness-of-fit tests within the exponential family. *Comm. Statist.*, v. 3, p. 1139–1153, 1974.
- SOCIEDADE PORTUGUESA DE ESTATÍSTICA, SPE. *Sociedade Portuguesa de Estatística — Associação Brasileira de Estatística. Glossário de Estatística: Inglês-Português*. Lisboa, PT, 2022. Disponível em: <https://www.spestatistica.pt/glossario/>.

STIGLER, S. M. Karl Pearson's theoretical errors and the advances they inspired. *Statist. Sci.*, v. 23, n. 2, p. 261–271, 2008.

STUART, A.; ORD, J. K. *Kendall's advanced theory of statistics. Vol. 2.* Fifth. [S.l.]: The Clarendon Press, Oxford University Press, New York, 1991. i–xii and 605–1323 p. Classical inference and relationship.

VOINOV, V.; NIKULIN, M. S.; BALAKRISHNAN, N. *Chi-squared goodness of fit tests with applications.* [S.l.]: Amsterdam: Elsevier/Academic Press, 2013. xi+229 p.

WEIERS, R. M. *Introduction to Business Statistics.* [S.l.]: The Dryden Press, Chicago, Illinois, 1991.

## APÊNDICE A – Definições e Teoremas

**Definição A.1.** (Convergência em Probabilidade) Uma sequência de variáveis aleatórias,  $X_1, X_2, \dots$ , converge em probabilidade para a uma variável aleatória  $X$  se, para cada  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P(|X_n - X| \geq \varepsilon) = 0 \quad \text{ou, de modo equivalente,} \quad \lim_{n \rightarrow \infty} P(|X_n - X| < \varepsilon) = 1.$$

Notação:  $X_n \xrightarrow{P} X$ .

**Definição A.2.** (Convergência Quase Certa) Uma sequência de variáveis aleatórias,  $X_1, X_2, \dots$ , converge quase certamente para uma variável aleatória  $X$  se, para  $\varepsilon > 0$ ,

$$P(\lim_{n \rightarrow \infty} |X_n - X| < \varepsilon) = 1.$$

**Definição A.3.** (Convergência em Distribuição) Uma sequência de variáveis aleatórias,  $X_1, X_2, \dots$ , converge em distribuição para uma variável aleatória  $X$  se

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$$

em todos os pontos  $x$  onde  $F_X(x)$  é contínua. Notação:  $X_n \xrightarrow{d} X$ .

**Teorema A.1.** (Teorema do Mapeamento Contínuo) Suponha que  $X_n$  convirja para  $X$  em probabilidade, quase certamente, ou em lei. Seja  $g(\cdot)$  uma função contínua. Então  $g(X_n)$  converge para  $g(X)$  em probabilidade, quase certamente ou em lei, respectivamente.

**Teorema A.2.** (Lei Forte dos Grandes Números) Sejam  $X_1, X_2, \dots$  variáveis aleatórias i.i.d, com  $E(X_i) = \mu$  e  $\text{Var}(X_i) = \sigma^2 < \infty$ , e definimos  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ . Então, para cada  $\varepsilon > 0$ ,

$$P(\lim_{n \rightarrow \infty} |\bar{X}_n - \mu| < \varepsilon) = 1;$$

isto é,  $\bar{X}_n$  converge quase certamente para  $\mu$ .

**Teorema A.3.** *Se a sequência de variáveis aleatórias,  $X_1, X_2, \dots$ , converge em probabilidade para uma variável aleatória  $X$ , a sequência também converge em distribuição para  $X$ .*

**Teorema A.4** (Convergência em Probabilidade para uma constante). *A sequência de variáveis aleatórias,  $X_1, X_2, \dots$ , converge em probabilidade para uma constante  $\mu$  se, e somente se, a sequência também convergir em distribuição para  $\mu$ , isto é, a declaração*

$$P(|X_n - \mu| > \varepsilon) \rightarrow 0 \quad \text{para todo } \varepsilon > 0.$$

é equivalente a

$$P(X_n \leq x) \rightarrow \begin{cases} 0 & \text{se } x < \mu, \\ 1 & \text{se } x \geq \mu. \end{cases}$$

**Teorema A.5** (Versão mais forte do Teorema do Limite Central). *Sejam  $X_1, X_2, \dots$  uma sequência de variáveis aleatórias i.i.d. com  $E(X_i) = \mu$  e  $0 < \text{Var}(X_i) = \sigma^2 < \infty$ . Definimos  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ . Seja  $G_n(x)$  denotando a f.d.a. de  $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$ . Então, para qualquer  $x$ ,  $-\infty < x < \infty$ ,*

$$\lim_{n \rightarrow \infty} G_n(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy;$$

isto é,

$$\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \xrightarrow{d} N(0, 1) \tag{A.1}$$

quando  $n \rightarrow \infty$ .

**Teorema A.6** (Teorema de Slutsky).

(a) *Se  $X_n \xrightarrow{d} X$  e  $Y_n \xrightarrow{P} c$ , então  $X_n \cdot Y_n \xrightarrow{d} cX$ .*

(b) *Se  $X_n \xrightarrow{d} X$  e  $Y_n \xrightarrow{P} c \neq 0$ , então  $\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}$ .*

(c) *Se  $X_n \xrightarrow{d} X$  e  $Y_n \xrightarrow{P} c$ , então  $X_n + Y_n \xrightarrow{d} X + c$ .*

*Observação A.1.* As Definições A.1, A.2, A.3, são as Definições 5.5.4, 5.5.9, 5.5.12 de Casella e Berger (1990), e os Teoremas A.2, A.3, A.4 e A.5 são os Teoremas 5.5.9, 5.5.12, 5.5.13 e 5.5.15 de Casella e Berger (1990), respectivamente. Os Teoremas A.1 e A.6 foram retirados do livro de DasGupta (2008).