



Universidade de Brasília
IE- Departamento de Estatística
Métodos Estatísticos 2

Juliana Magalhães Rosa

Testes de Aderência

Brasília

2020

Sumário

1 Introdução 2

2 Distribuições das Estatísticas do Tipo Kolmogorov 3

 2.1 Estatística do Teste de Kolmogorov-Smirnov para Uma Amostra 3

 2.2 Estatística do Teste de Lilliefors para Normalidade 7

3 Testes para Normalidade 11

 3.1 Teste de Shapiro-Wilk 11

 3.2 Teste de Anderson-Darling 14

4 Considerações Finais 17

1 Introdução

Testes de Aderência (ou Ajustamento) são testes de hipóteses que estudam as distribuições de probabilidade das variáveis. Geralmente indicam se uma variável segue ou não certa distribuição conhecida, a partir da análise de amostras. É comum existirem vários testes com o mesmo propósito, mas com métodos diferentes.

Neste trabalho, serão analisadas as distribuições das estatísticas de teste tipo Kolmogorov, tanto para o teste de Kolmogorov-Smirnov para uma amostra, como para o teste de Lilliefors para normalidade. Isso será feito a partir de mil amostras aleatórias de tamanho 15 de uma variável normalmente distribuída com média de 116 (escolhida aleatoriamente) e desvio padrão de 11,6.

Além disso, serão escolhidas cinco amostras, entre essas mil, e nelas serão realizados os testes de Shapiro-Wilk e Anderson-Darling, também para testar normalidade.

2 Distribuições das Estatísticas do Tipo Kolmogorov

2.1 Estatística do Teste de Kolmogorov-Smirnov para Uma Amostra

Tabela 1: Distribuição de Frequências das Estatísticas de Kolmogorov em Mil Amostras Aleatórias, Software R, 2020.

Classes de Valores das Estatísticas de Kolmogorov		Frequência Absoluta
(0.08, 0.116]		27
(0.116, 0.152]		160
(0.152, 0.188]		257
(0.188, 0.224]		203
(0.224, 0.26]		145
(0.26, 0.296]		102
(0.296, 0.332]		54
(0.332, 0.368]		32
(0.368, 0.404]		15
(0.404, 0.44]		3
(0.44, 0.476]		2
Fonte:	Função	norm,
		Software
		R.

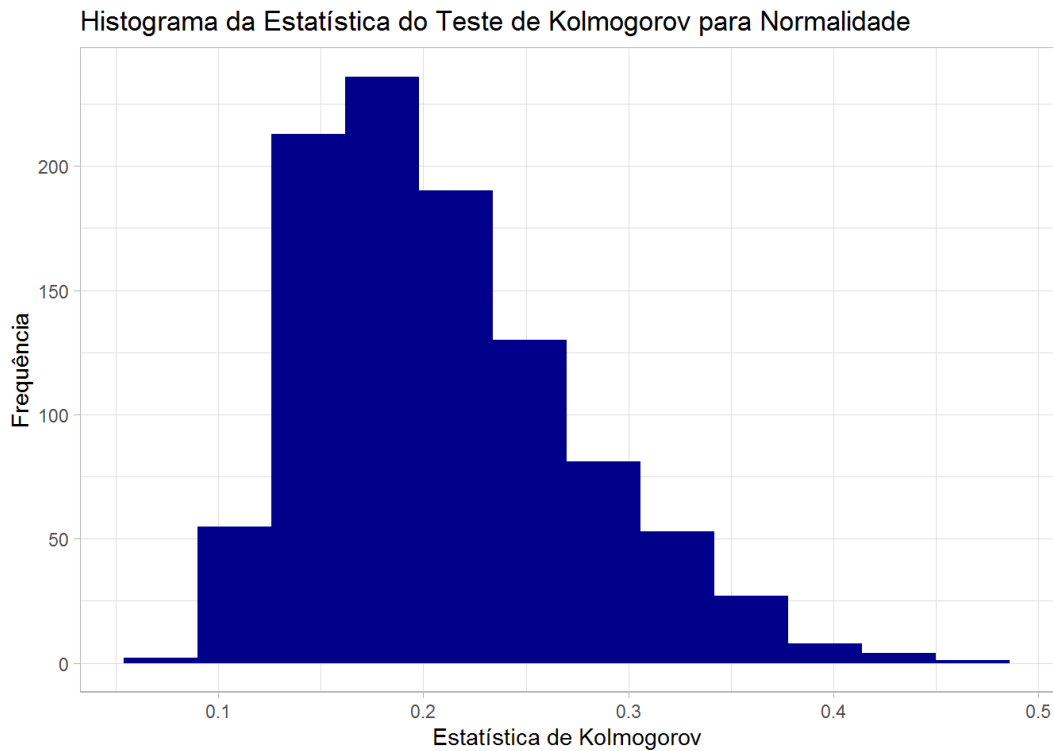


Tabela 2: Valores de Medidas Descritivas para as Estatísticas de Kolmogorov em Mil Amostras Aleatórias, Software R, 2020.

Medidas Descritivas	Valores
Média	0.209
Primeiro Quartil	0.160
Mediana/ Segundo Quartil	0.197
Terceiro Quartil	0.245
Quarto Quartil	0.475
Amplitude	0.388
Intervalo Interquartil	0.086
Variância	0.004
Desvio Padrão	0.065
Coeficiente de Variação	0.310
Coeficiente de Assimetria de Pearson	0.528
Coeficiente Quartil de Assimetria	0.121
Coeficiente de Achatamento	0.260

Fonte: Função summarize, pacote dplyr, Software R.

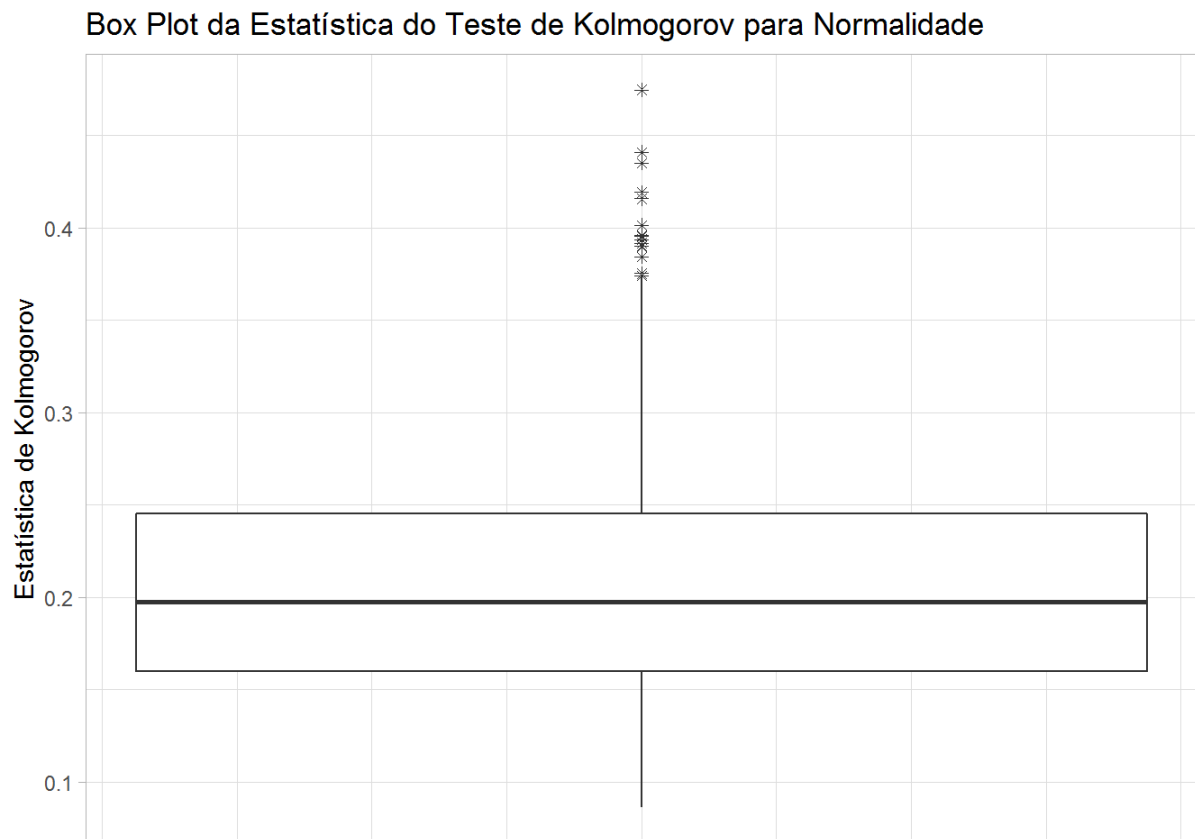


Tabela 3: Quantis da Distribuição das Estatísticas de Kolmogorov em Mil Amostras Aleatórias, Software R, 2020.

Quantis	Valores
80%	0.263
85%	0.277
90%	0.300
95%	0.333
97.5%	0.363
99%	0.390

Fonte: Função summarize, pacote dplyr, Software R.

O teste de Kolmogorov-Smirnov para uma amostra foi feito para checar se as mil amostras seguiam distribuições normais com parâmetros especificados (média igual a 116 e desvio padrão igual a 11,6). As mil estatísticas de teste resultantes formam sua própria distribuição, a partir da qual foram feitos os gráficos e as tabelas acima.

Pela tabela de frequências por classes, vê-se que todos os valores calculados da estatística estão entre 0,08 e 0,476. As frequências estão mais concentradas nas primeiras classes, indicando que talvez haja uma assimetria à direita na distribuição.

A assimetria fica clara no histograma, mostrando que poucos dos testes realizados resultaram em estatísticas com valores altos, ou seja, provavelmente houve poucos testes que rejeitaram a hipótese de normalidade. Também é possível perceber, por meio desse gráfico, que as estatísticas estão concentradas em valores logo abaixo de 0,2.

Observando a tabela com as medidas descritivas, descobre-se uma mediana de 0,197 e uma média de 0,209 que foi influenciada pelos valores extremos na direita. Com um terceiro quartil de 0,245, conclui-se que 75% das estatísticas estão abaixo desse valor. Disso infere-se que no mínimo 75% dos testes resultaram em aceitação da normalidade, já que o valor crítico tabelado para o teste de Kolmogorov (para amostras de tamanho 15) é de 0,338 a um nível descritivo de 0,05. Os dois coeficientes de assimetria positivos são evidências concretas da assimetria à direita. O coeficiente de achatamento com seu valor de 0,26 indica que a distribuição é praticamente mesocúrtica.

O boxplot, novamente, reflete a assimetria da distribuição com a diferença visível entre as distâncias do limite inferior à mediana e da mediana ao limite superior. Além disso, mostra uma grande quantidade de valores extremos, que são representados pelos asteriscos.

Comparando os quantis calculados com aqueles tabelados para amostras de tamanho 15, percebe-se que os resultados são similares. O quantil 80% teve valor de 0,263 para as mil amostras, já seu valor tabelado é de 0,266. O quantil de 90% para as dadas amostras resultou em 0,3, enquanto seu valor na tabela é de 0,304. O valor teórico do quantil 95%, como já foi dito, é de 0,338, mas seu valor dado pela

tabela acima é de 0,333. Os quantis que mais diferem dos valores tabelados são os de 98% (0,377 no Conover, 0,363 baseado nas amostras) e 99% (0,404 no Conover, 0,390 baseado nas amostras), mas ainda são suficientemente próximos.

2.2 Estatística do Teste de Lilliefors para Normalidade

Tabela 4: Distribuição de Frequências das Estatísticas de Lilliefors em Mil Amostras Aleatórias- Software R, 2020.

Classes de Valores das Estatísticas de Lilliefors	Frequência Absoluta
(0.069, 0.09]	14
(0.09, 0.111]	103
(0.111, 0.132]	239
(0.132, 0.153]	204
(0.153, 0.174]	175
(0.174, 0.195]	150
(0.195, 0.216]	64
(0.216, 0.237]	23
(0.237, 0.258]	14
(0.258, 0.279]	9
(0.279, 0.3]	5

Fonte: Função `rnorm,` Software R.

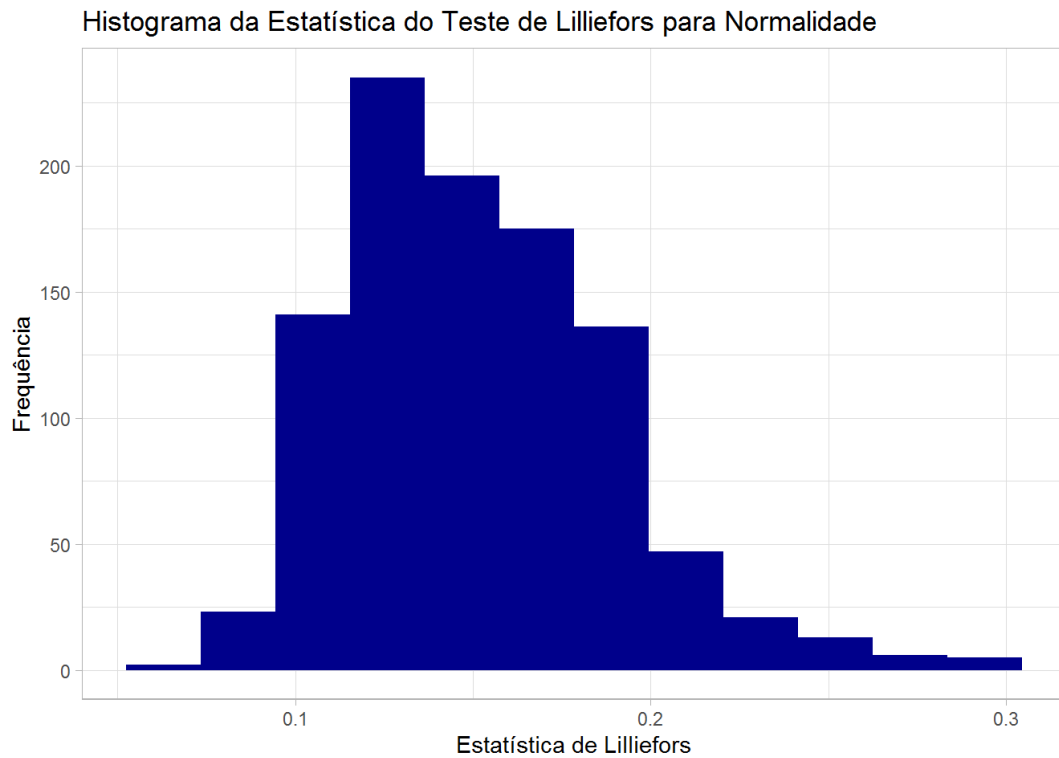


Tabela 5: Valores de Medidas Descritivas para as Estatísticas de Lilliefors em Mil Amostras Aleatórias, Software R, 2020.

Medidas Descritivas	Valores
Média	0.152
Primeiro Quartil	0.124
Mediana/ Segundo Quartil	0.147
Terceiro Quartil	0.176
Quarto Quartil	0.296
Amplitude	0.227
Intervalo Interquartil	0.051
Variância	0.001
Desvio Padrão	0.037
Coeficiente de Variação	0.246
Coeficiente de Assimetria de Pearson	0.369
Coeficiente Quartil de Assimetria	0.109
Coeficiente de Achatamento	0.289

Fonte: Função summarize, pacote dplyr, Software R.

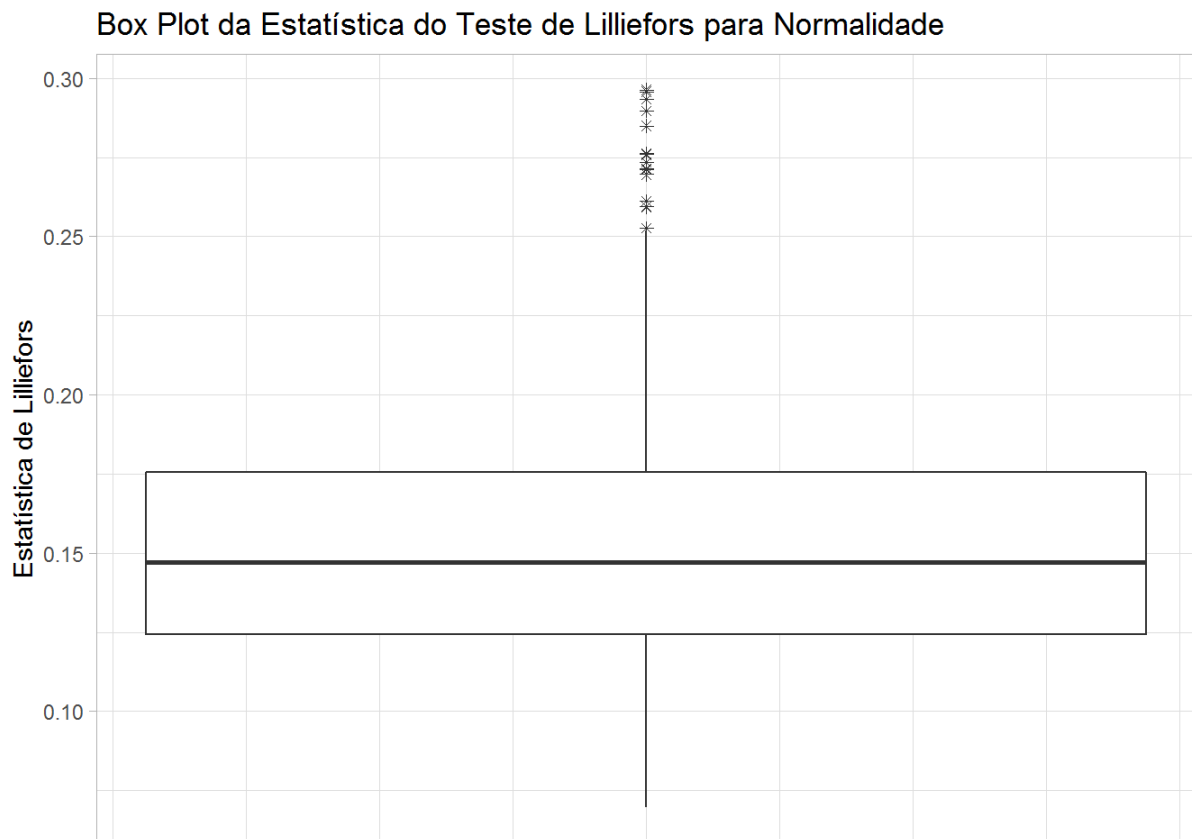


Tabela 6: Quantis da Distribuição das Estatísticas de Lilliefors em Mil Amostras Aleatórias, Software R, 2020.

Quantis	Valores
80%	0.182
85%	0.190
90%	0.198
95%	0.216
97.5%	0.240
99%	0.270

Fonte: Função summarize, pacote dplyr, Software R.

O teste de Lilliefors para normalidade foi feito para checar se as mil amostras seguiam distribuições normais sem especificação de parâmetros (média e desvio padrão são estimados para cada amostra). As mil estatísticas de teste resultantes formam sua própria distribuição, a partir da qual foram feitos os gráficos e as tabelas acima.

Há, novamente, uma maior concentração de frequências nas classes iniciais, indicando uma possível assimetria à direita. A tabela de distribuição de frequências também mostra que a classe mais frequente é a que vai de 0,111 a 0,132.

O histograma mostra que a distribuição está centrada em torno de 0,15. Também é possível enxergar a assimetria que havia sido indicada na tabela de frequências, mas ela parece ser menos acentuada do que a da distribuição da estatística tradicional de Kolmogorov.

A tabela de medidas descritivas revela uma mediana de 0,147 e uma média de 0,152, confirmando a impressão que o histograma deu sobre o centro da distribuição. Com um valor máximo de 0,296, percebe-se que as estatísticas de Lilliefors não chegaram nem a 0,3, que é o quantil 90% das estatísticas de Kolmogorov-Smirnov. O coeficiente de variação de 0,246 quando comparado ao da distribuição anterior (0,31) mostra que as estatísticas de Lilliefors resultaram em uma distribuição menos variável. Os coeficientes de assimetria continuam positivos, ratificando a assimetria à direita, porém são menores, em confirmação da suspeita de que essa assimetria é menos acentuada na nova distribuição. O coeficiente de achatamento de 0,289 classifica a distribuição como platicúrtica.

O boxplot ilustra as medidas de posição já vistas na tabela (mediana de 0,147, primeiro quartil de 0,124, terceiro quartil de 0,176) e também evidencia a existência de vários valores extremos no topo.

Os quantis tabelados para o teste de Lilliefors para normalidade são menores que os do teste de Kolmogorov, pois foram ajustados para considerar a estimação de parâmetros. De novo foram achados valores próximos aos da tabela do Conover. O quantil 80% encontrado foi de 0,182, enquanto seu valor teórico seria de 0,181. O quantil de 85% resultou em 0,192, já seu valor tabelado é de 0,190. Para o quantil de 90%, as amostras indicaram o valor de 0,198, ainda próximo do valor teórico de

0,202. O valor da tabela do Conover para o quantil de 95% é de 0,219, enquanto o valor encontrado a partir das mil amostras é de 0,216. Novamente, o quantil mais diferente do esperado (0,256) foi o de 99% (0,270), mas ainda é próximo o suficiente.

3 Testes para Normalidade

3.1 Teste de Shapiro-Wilk

Tabela 7: Resultado do Teste de Shapiro-Wilk para a Amostra 1, Software R, 2020.

Medidas	Valores
Média Amostral	120.423
Desvio Padrão Amostral	9.167
Estatística do Teste	0.932
P-valor	0.288

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 8: Resultado do Teste de Shapiro-Wilk para a Amostra 2, Software R, 2020.

Medidas	Valores
Média Amostral	117.169
Desvio Padrão Amostral	10.084
Estatística do Teste	0.941
P-valor	0.397

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 9: Resultado do Teste de Shapiro-Wilk para a Amostra 3, Software R, 2020.

Medidas	Valores
Média Amostral	119.650
Desvio Padrão Amostral	13.773
Estatística do Teste	0.956
P-valor	0.630

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 10: Resultado do Teste de Shapiro-Wilk para a Amostra 4, Software R, 2020.

Medidas	Valores
Média Amostral	119.833
Desvio Padrão Amostral	8.682
Estatística do Teste	0.875
P-valor	0.040

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 11: Resultado do Teste de Shapiro-Wilk para a Amostra 5, Software R, 2020.

Medidas	Valores
Média Amostral	114.470
Desvio Padrão Amostral	14.298
Estatística do Teste	0.949
P-valor	0.505

Fonte: Função gofTest, pacote EnvStats, Software R.

O teste de Shapiro-Wilk é específico para o estudo da normalidade. Se sua estatística tiver um valor próximo de 1, significa que a distribuição se aproxima da normal. Ou seja, quanto menor o valor da estatística do teste, maior é a evidência contra a normalidade.

Como se pode ver pelas tabelas, os primeiros três testes, assim como o último, resultaram em estatísticas próximas de 1 (0,932, 0,941, 0,956, 0,949) e p-valores grandes (0,288, 0,397, 0,630, 0,505). Portanto, em nenhum desses testes houve a rejeição da hipótese de normalidade. Esse resultado já era esperado, uma vez que as amostras de fato foram geradas para uma variável aleatória normalmente distribuída.

A amostra 4 foi a única com um resultado diferente. Sua estatística do teste foi de 0,875, um valor bem mais baixo, e seu p-valor foi de 0,04, permitindo a rejeição da hipótese nula a um nível descritivo de 0,05. Isso pode ter sido apenas uma peculiaridade da amostra. Apesar da contradição na quarta amostra, num geral os testes apontam para a normalidade, o que está de acordo com as estatísticas do tipo Kolmogorov analisadas anteriormente. Como foi visto nas distribuições dessas estatísticas, seus valores se concentraram em classes mais baixas e a maioria não permitiu a rejeição da hipótese de normalidade.

3.2 Teste de Anderson-Darling

Tabela 12: Resultado do Teste de Anderson-Darling para a Amostra 1, Software R, 2020.

Medidas	Valores
Média Amostral	120.423
Desvio Padrão Amostral	9.167
Estatística do Teste	0.473
P-valor	0.207

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 13: Resultado do Teste de Anderson-Darling para a Amostra 2, Software R, 2020.

Medidas	Valores
Média Amostral	117.169
Desvio Padrão Amostral	10.084
Estatística do Teste	0.419
P-valor	0.284

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 14: Resultado do Teste de Anderson-Darling para a Amostra 3, Software R, 2020.

Medidas	Valores
Média Amostral	119.650
Desvio Padrão Amostral	13.773
Estatística do Teste	0.331
P-valor	0.470

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 15: Resultado do Teste de Anderson-Darling para a Amostra 4, Software R, 2020.

Medidas	Valores
Média Amostral	119.833
Desvio Padrão Amostral	8.682
Estatística do Teste	0.982
P-valor	0.010

Fonte: Função gofTest, pacote EnvStats, Software R.

Tabela 16: Resultado do Teste de Anderson-Darling para a Amostra 5, Software R, 2020.

Medidas	Valores
Média Amostral	114.470
Desvio Padrão Amostral	14.298
Estatística do Teste	0.363
P-valor	0.393

Fonte: Função gofTest, pacote EnvStats, Software R.

O teste de Anderson-Darling é uma modificação do teste de Kolmogorov e, assim como nele, uma estatística de teste grande representa evidência contra a hipótese de normalidade.

Os resultados foram parecidos com os do teste de Shapiro-Wilk: as três primeiras amostras, assim como a última, apresentaram estatísticas do teste pequenas (0,473, 0,419, 0,331, 0,363) e p-valores grandes (0,207, 0,284, 0,470, 0,393). Portanto, para essas quatro amostras não houve a rejeição da normalidade.

A quarta amostra, entretanto, resultou em uma estatística de 0,982, um valor bem maior do que os outros, e um p-valor de 0,01, permitindo a rejeição da hipótese nula a um nível de significância de 0,05. De forma geral, esses resultados estão de acordo com os relativos ao teste de Shapiro-Wilk e com as tendências observadas na análise descritiva da distribuição das estatísticas do tipo Kolmogorov.

4 Considerações Finais

Foi simulado, em menores dimensões, o processo de análise e tabelamento das estatísticas do tipo Kolmogorov. O resultado foi próximo do esperado, considerando os quantis similares àqueles teóricos. As distribuições dessas estatísticas possuem assimetria à direita, indicando que a maioria assume valores baixos e, portanto, não permite rejeição da hipótese de normalidade. Isso ocorre pois a variável em questão foi construída em cima de uma distribuição normal.

Utilizando cinco amostras dessa mesma variável, foram feitos os testes de Shapiro-Wilk e Anderson-Darling para testar normalidade. De um total de dez testes realizados, apenas dois (referentes a uma mesma amostra) resultaram na rejeição da hipótese nula.

Em suma, se são utilizadas as mesmas amostras de uma mesma variável, os testes de aderência tendem a concordar em questão de resultado, mesmo cada um tendo um método de análise e decisão diferente. Conclui-se também que, apesar de uma variável seguir uma determinada distribuição, pode acontecer de um teste de ajustamento negar esse fato. Isso ocorre porque seus resultados dependem de diversos fatores, como, por exemplo, a amostragem.