

Normalidade

Normalidade (variáveis contínuas/métricas)

- Distribuição normal (características)
 - Simétrica
 - Mesocúrtica
 - Permitindo o cálculo da média e do desvio-padrão como os valores mais representativos da localização e da dispersão
- Distribuição teórica necessária em muitos procedimentos estatísticos
- Detecção
 - Testes de normalidade
 - Rej. ou não da normalidade
 - Gráficos de Normalidade
 - Identificar as observações que se desviam da distribuição

- Teste não paramétrico de aderência à normal
K-S (Kolmogorov-Smirnov)
 - H_0 : A distribuição da variável é normal
 - H_1 : A distribuição da variável não é normal

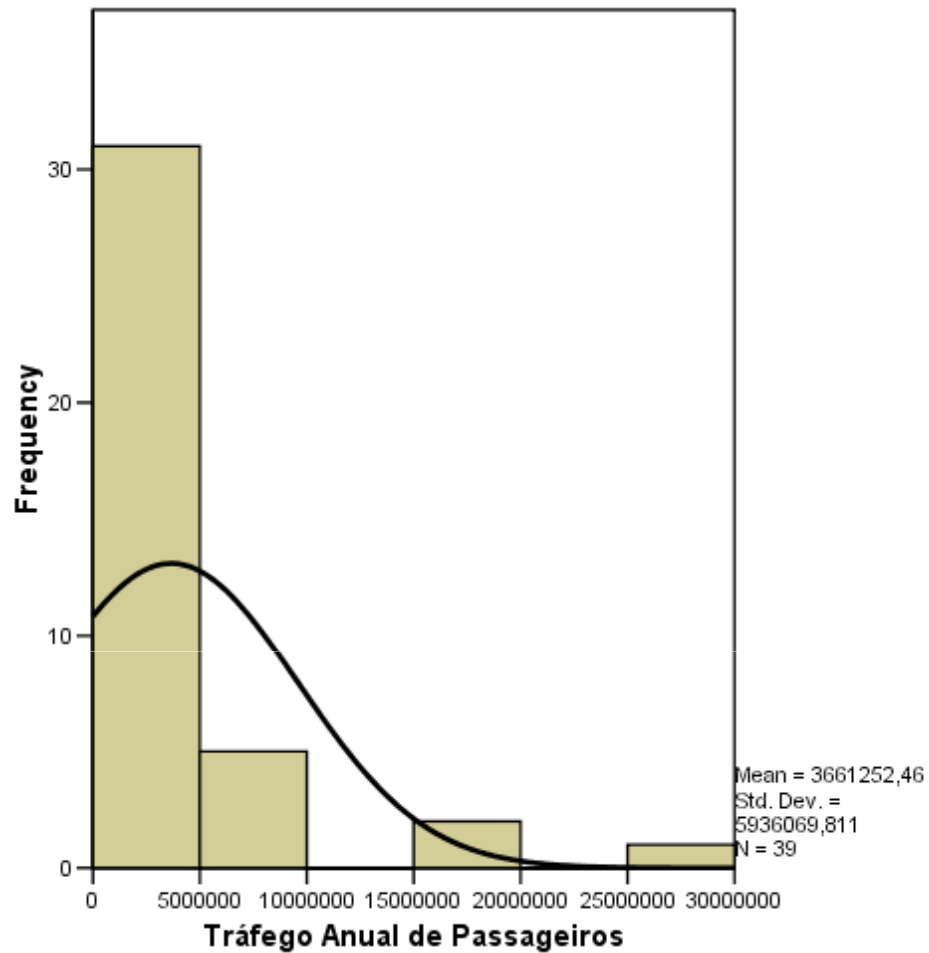


Gráficos de Normalidade

- Q-Q Plot
 - As observações devem distribuir-se junto à recta oblíqua para a distribuição ser considerada normal
- Detrended normal Q-Q Plot
 - As observações devem distribuir-se de forma aleatória à volta de uma linha recta horizontal 0 para a distribuição ser normal

Exemplo 1

Tráfego Anual de Passageiros



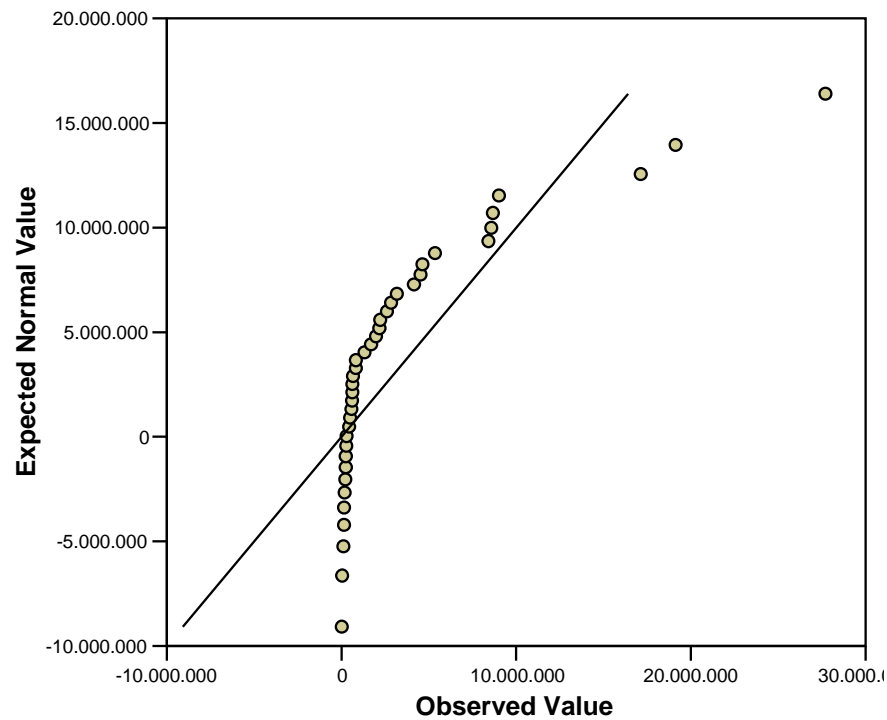
É rejeitada a hipótese da distribuição ser Normal

Tests of Normality

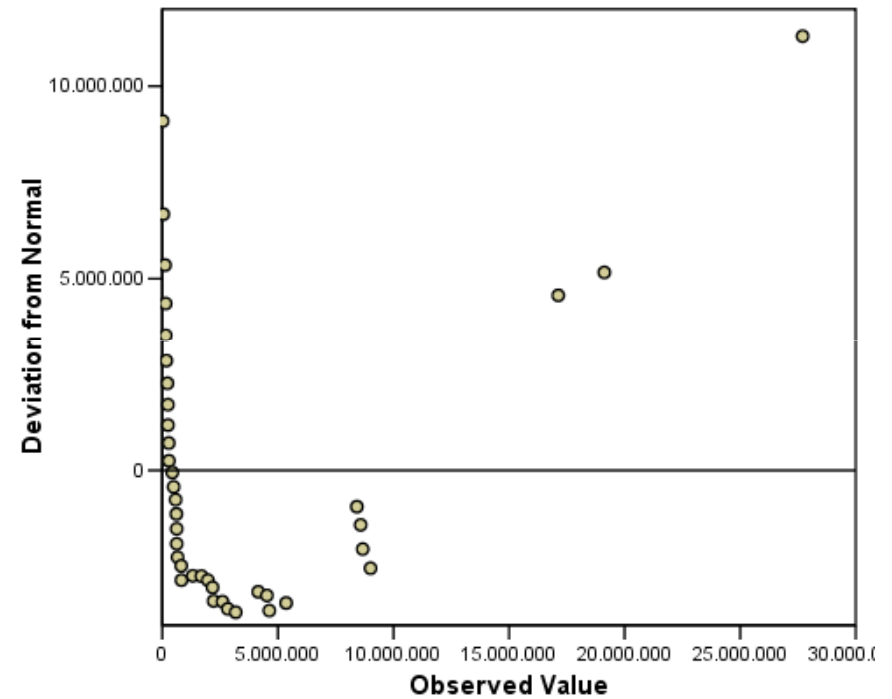
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Tráfego Anual de Passageiros	,270	39	,000	,636	39	,000

a. Lilliefors Significance Correction

Normal Q-Q Plot of Tráfego Anual de Passageiros



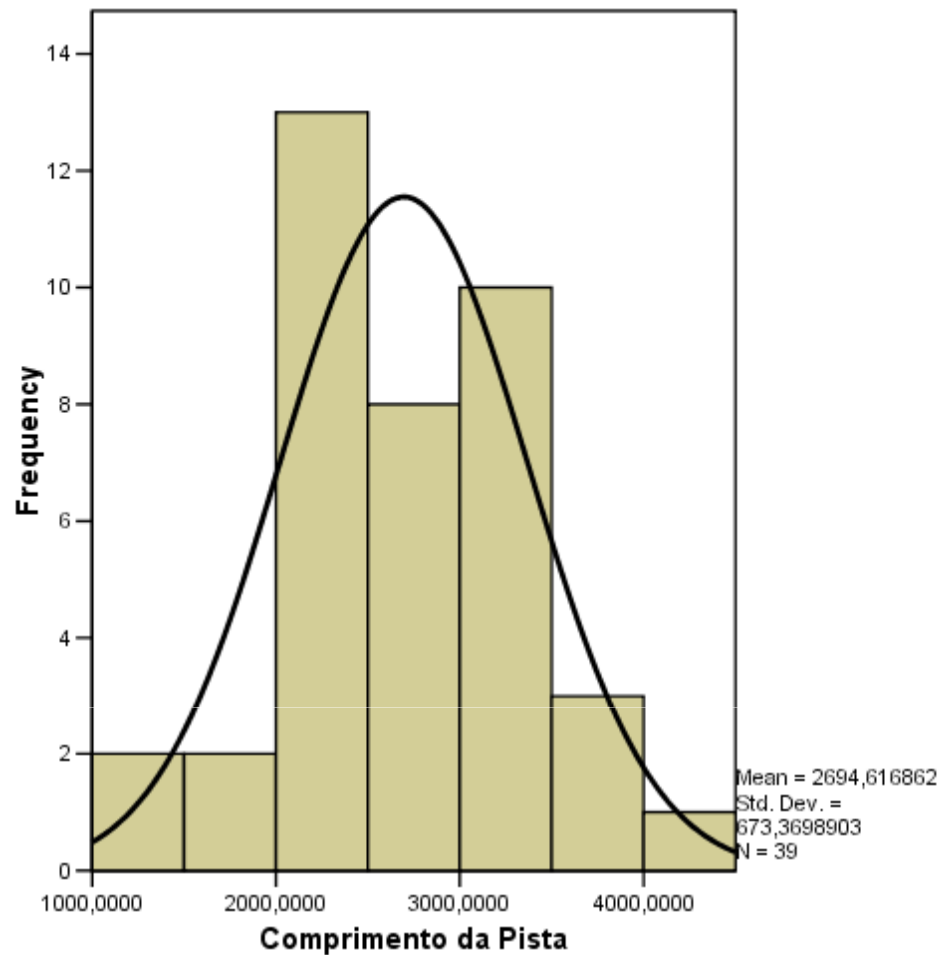
Detrended Normal Q-Q Plot of Tráfego Anual de Passageiros



Identificação das observações que se desviam da distribuição

Exemplo 2

Comprimento da Pista



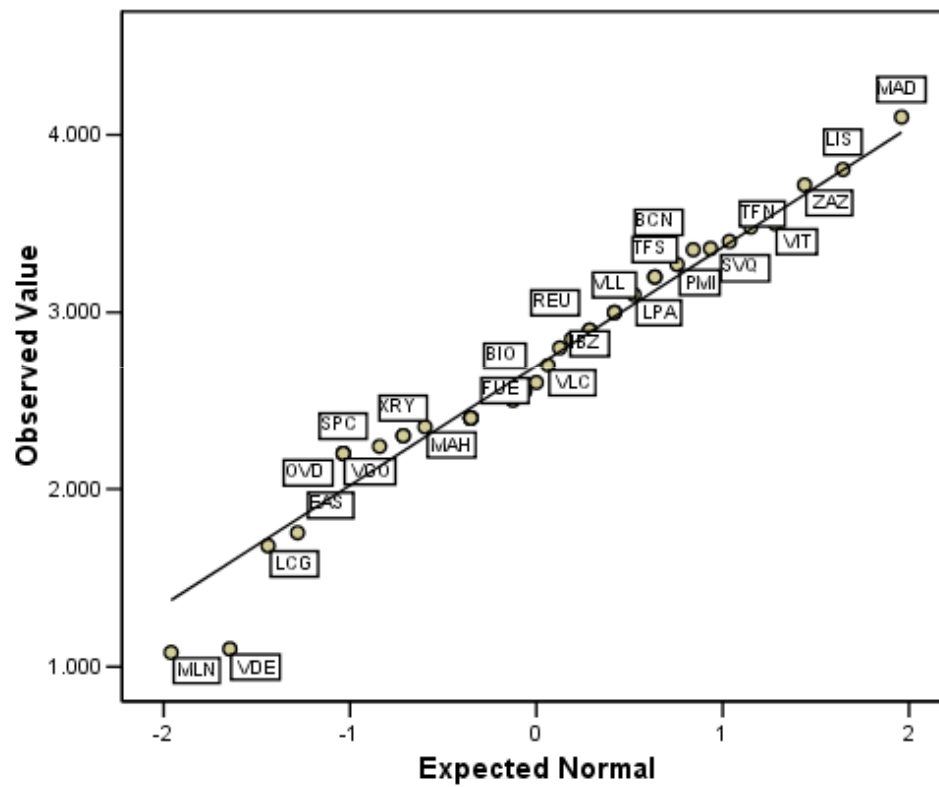
Como o nível de significância do teste é 0,102, superior ao valor 0,05, não se rejeita a hipótese da distribuição ser normal

Tests of Normality

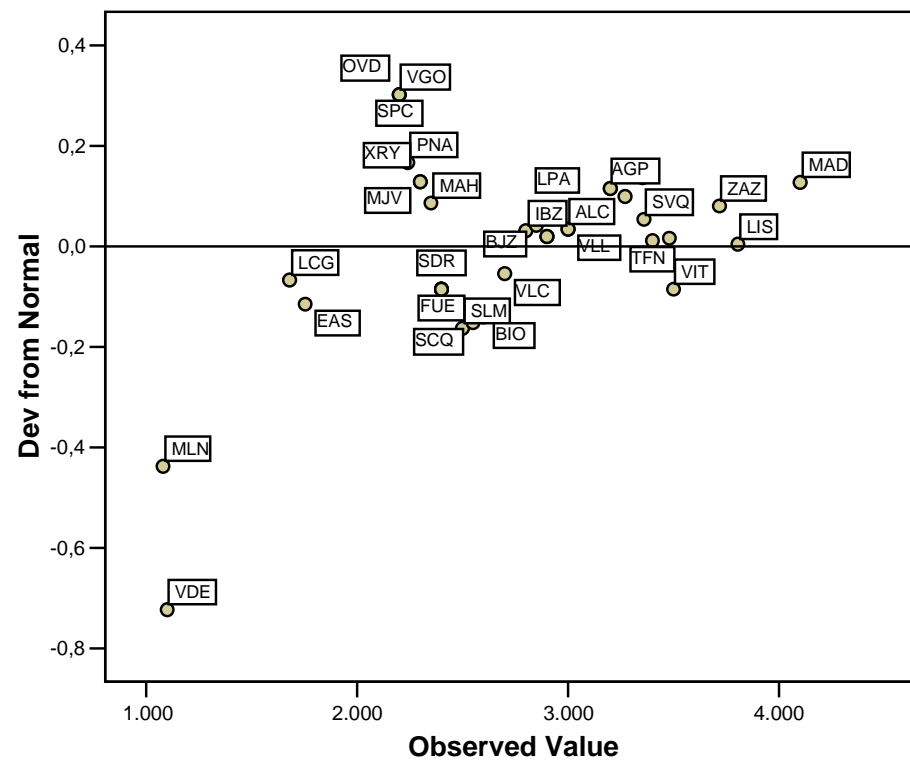
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Comprimento da Pista	,129	39	,102	,969	39	,356

a. Lilliefors Significance Correction

Normal Q-Q Plot of Comprimento da Pista



Detrended Normal Q-Q Plot of Comprimento da Pista



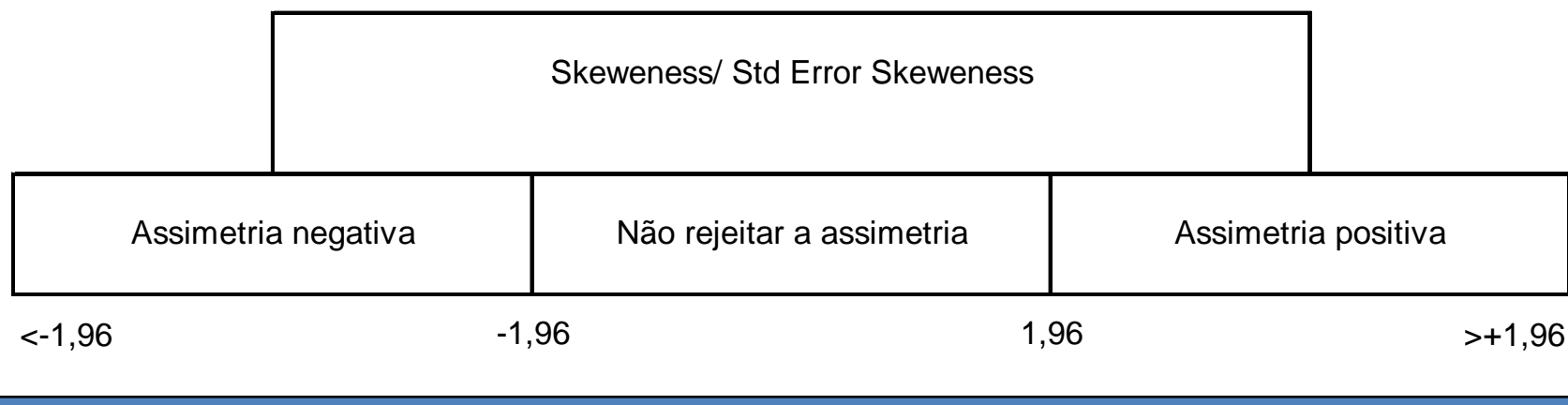
Análise dos desvios à normalidade

Medida de assimetria

- Distribuição simétrica/ moderadamente assimétrica
 - A mediana está perto do meio da caixa de bigodes
 - A média, mediana e média aparada a 5% apresentam valores semelhantes
 - A mediana pertence ao intervalo de confiança a 95% para a média
 - O quociente de assimetria aproxima-se de 2

- Distribuição não simétrica
 - As medidas de tendência central aproximam-se mais do valor mínimo (assimetria +) ou do valor máximo (assimetria -)
 - A mediana não pertence ao intervalo de confiança a 95% para a média nem está perto dos extremos do intervalo
 - O quociente de assimetria (em módulo) é superior a 2
 - Nas distribuições assimétricas os dados podem ser transformados por forma a obter a simetria ou usarem-se estatísticas robustas.

Medida de assimetria [Skewness]

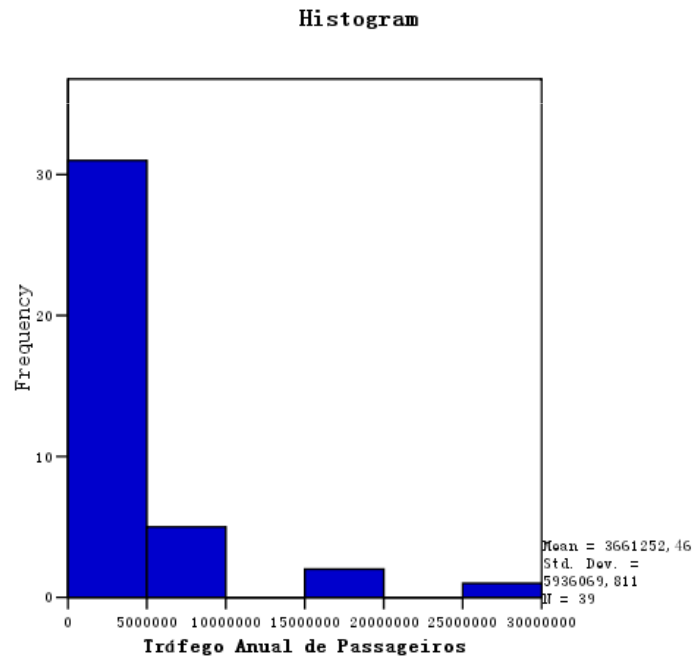


$$skewness = \frac{N \sum_{i=1}^N F_i (x_i - \bar{x})^3}{(N-1)(N-2)s^3}$$

$$stdError_{Skewness} = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}}$$

Tráfego Anual de Passageiros

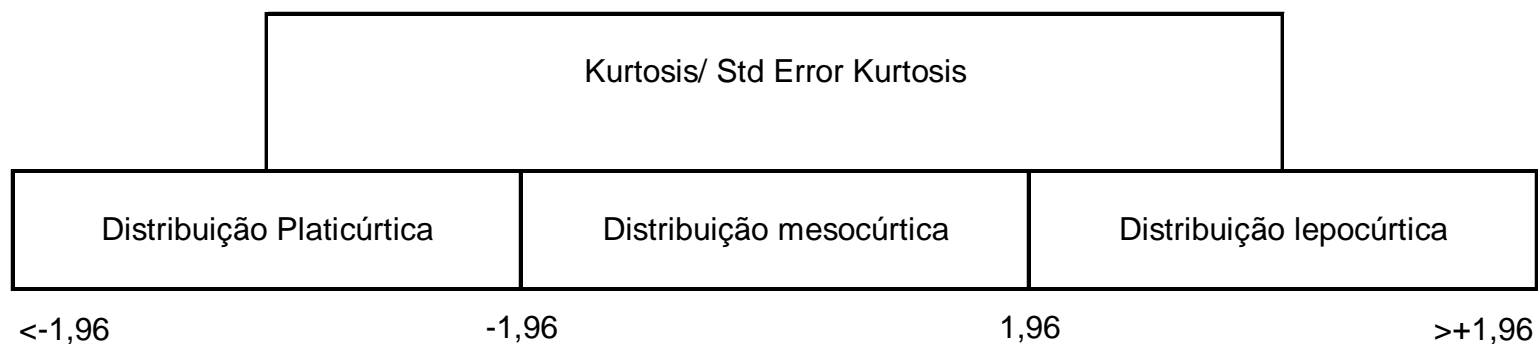
Skewness	Std.Error	Skewness /Std. Error
2,62	0,38	6,94



A distribuição do Tráfego Anual de Passageiros é assimétrica positiva ou enviesada à esquerda

Medida de achatamento

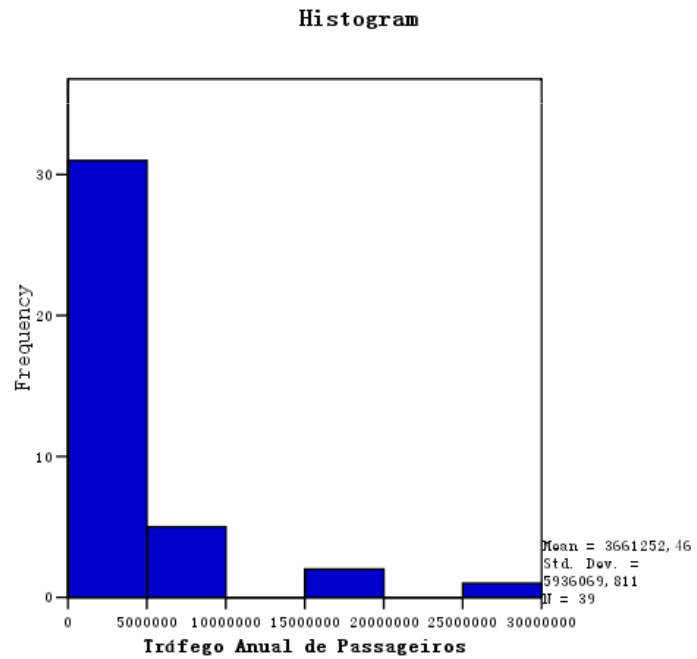
Medida de achatamento [Kurtosis]



$$Kurtosis = \frac{N(N+1) \sum_{i=1}^N F_i (x_i - \bar{x})^4 - 3s^2(N-1)}{(N-1)(N-2)(N-3)s^4}$$

Tráfego Anual de Passageiros

Kurtosis	Std.Error	Kurtosis /Std. Error
7,40	0,74	9,98



A distribuição do Tráfego é leptocúrtica