



PRO2514 - Pesquisa Quantitativa em Gestão de Operações

Normalidade dos dados

Prof. Dr. Renato de Oliveira Moraes



Avaliar a Normalidade das variáveis do Hatco

Variáveis

- X1: Delivery Speed
- X2: Price Level
- X3: Price Flexibility
- X4: Manufacturer Image
- X5: Service
- X5: Salesforce Image
- X7: Product Quality

Teste de Normalidade

$$\begin{cases} H_0: A \text{ variável tem distribuição Normal} \\ H_1: A \text{ variável } \mathbf{N\tilde{A}O} \text{ tem distribuição Normal} \end{cases}$$

Vários testes disponíveis

- Shapiro-Wilk (nativo no R)
- Kolmogorov-Smirnov
- Anderson-Darling
- Shapiro-Francia



X1: Delivery Speed

```
shapiro.test(Hatco$x1)
```

```
Shapiro-Wilk normality test
```

```
data: Hatco$x1
```

```
W = 0.98542, p-value = 0.3406
```

NÃO Rejeitar H_0
(A distribuição é Normal)



X2: Price Level

```
shapiro.test(Hatco$x2)
```

```
Shapiro-Wilk normality test
```

```
data: Hatco$x2
```

```
W = 0.96851, p-value = 0.01702
```

**Rejeitar H_0
(Normalidade)**



Testes de Normalidade

- Instalar o Pacote

```
install.packages('nortest')
```

```
library(nortest)
```

- Testes

Kolmogorov-Smirnov: `lillie.test(Hatco$x2)`

Shapiro-Francia: `sf.test(Hatco$x2)`

Anderson-Darling: `ad.test(Hatco$x2)`



Lilliefors (Kolmogorov-Smirnov) normality test

data: Hatco\$x2

$D = 0.094549$, p-value = 0.02789

Shapiro-Francia normality test

data: Hatco\$x2

$W = 0.97228$, p-value = 0.0337

Anderson-Darling normality test

data: Hatco\$x2

$A = 1.0048$, p-value = 0.01147



Transformação de Box-Cox

Suponha que a variável X não tenha distribuição Normal (sua distribuição está distante da normalidade).

$$X_{Nova} = \frac{X^\lambda - 1}{\lambda}$$

Mas qual deve ser o valor de λ para X_{Nova} tenha uma distribuição mais próxima da Normalidade?



- Instalação dos pacotes para outros testes de Normalidade

```
install.packages('car')
```

```
library(car)
```

```
install.packages('fpp')
```

```
library(fpp)
```

- Achar a potência para a transformação

```
lambda <- BoxCox.lambda(Hatco$x2, method = c('loglik'), lower = -5,  
upper = 5)
```

```
print(lambda)
```




- Fazendo a transformação

```
nova_var <- BoxCox (Hatco$x2, lambda)
```

- Verificando a Normalidade da nova variável

```
shapiro.test(nova_var)
```

```
Shapiro-Wilk normality test
```

```
data: nova_var
```

```
W = 0.98972, p-value = 0.6419
```

NÃO Rejeitar H_0
(A distribuição é Normal)

**A transformação levou a normalidade !
(nem sempre funciona)**



Verificando a Normalidade das demais variáveis

X1: Delivery Speed (Normal)

X2: Price Level (Normal depois da transformação de Box-Cox)

- X3: Price Flexibility (?)
- X4: Manufacturer Image (?)
- X5: Service (?)
- X5: Salesforce Image (?)
- X7: Product Quality (?)