



*Big Data Analytics com R e Microsoft Azure Machine Learning 3.0*

# Big Data Analytics com R e Microsoft Azure Machine Learning Versão 3.0

## Estatística na Prática 2 Definindo o Teste F



O Teste F é um tipo de teste de hipótese, baseado na distribuição f de Snedecor, sob a hipótese nula. O teste é realizado quando não se sabe se as duas populações têm a mesma variância.

O Teste F também pode ser usado para verificar se os dados estão de acordo com um modelo de regressão, que é adquirido por meio da análise de mínimos quadrados. Uma das suposições do Teste F é que os dados seguem uma distribuição normal.

O Teste F recebe seu nome da sua estatística de teste, a estatística F. A estatística F é simplesmente uma razão de duas variâncias. As variâncias são uma medida de dispersão, ou até que ponto os dados estão dispersos em relação à sua média. Valores maiores representam maior dispersão.

A variância é o quadrado do desvio padrão. Para nós humanos, os desvios padrão são mais fáceis de entender do que as variâncias pois estão nas mesmas unidades que os dados, e não nas unidades quadradas.

E por que precisamos do Teste F para este estudo de caso?

Porque uma das suposições do Teste t é que as médias das amostras têm a mesma variância (ou que não há diferença significativa na variância das médias). Vamos aplicar um Teste F para comparar as variâncias das amostras (extraídas de uma população normalmente distribuída).

### Hipóteses do Teste F:

- H0 (Hipótese Nula) = As médias de dados extraídos de uma população normalmente distribuída têm a mesma variância.
- Ha (Hipótese Alternativa) = As médias de dados extraídos de uma população normalmente distribuída não têm a mesma variância.

Aplicamos o Teste F, interpretamos o resultado e então decidimos se rejeitamos ou não a H0. Somente após validar as suposições poderemos aplicar o Teste t.