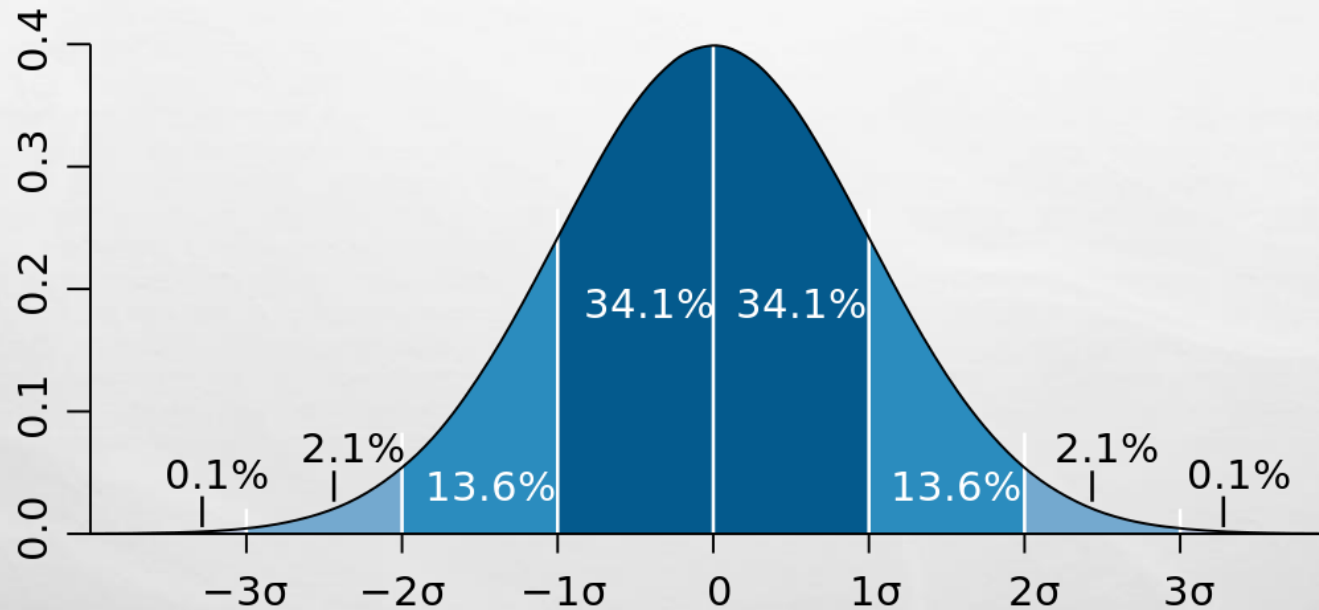


# FORMAÇÃO CIENTISTA DE DADOS

ESTATÍSTICA I: DISTRIBUIÇÃO NORMAL

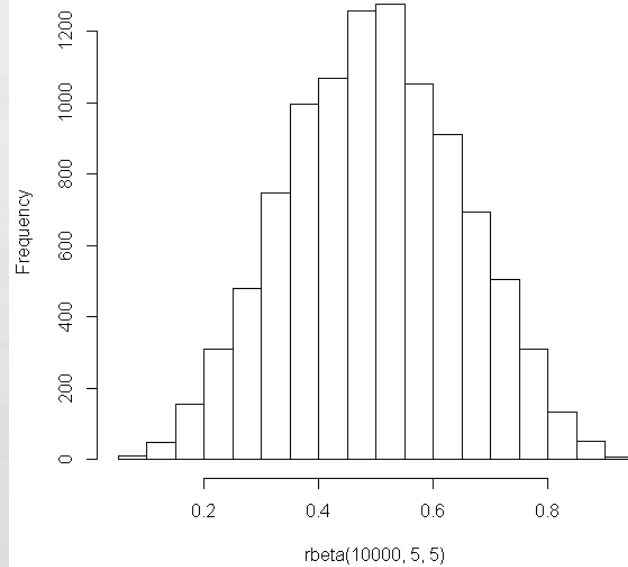


# Como saber se é Normal?

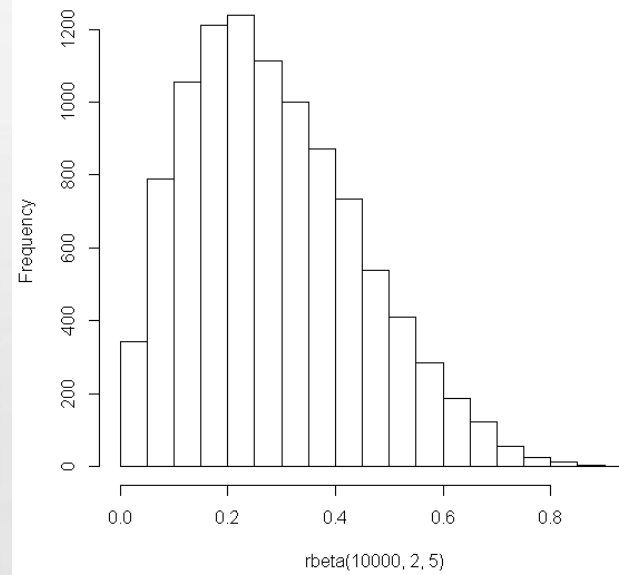


# Histograma

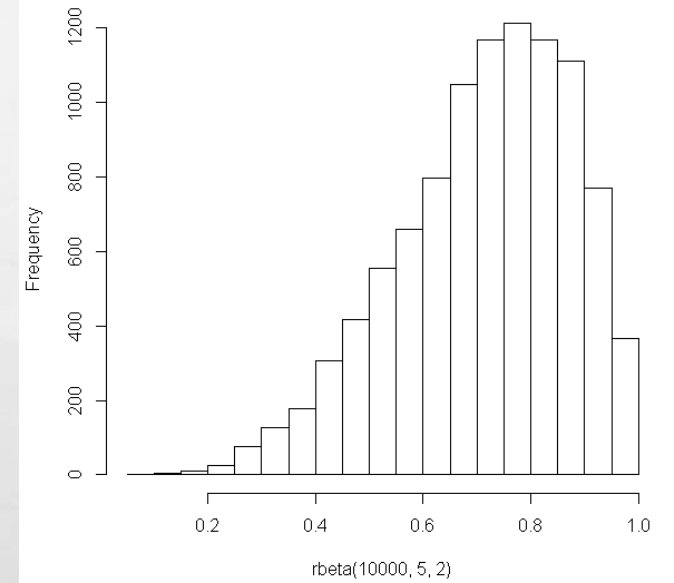
Histogram of  $\text{rbeta}(10000, 5, 5)$



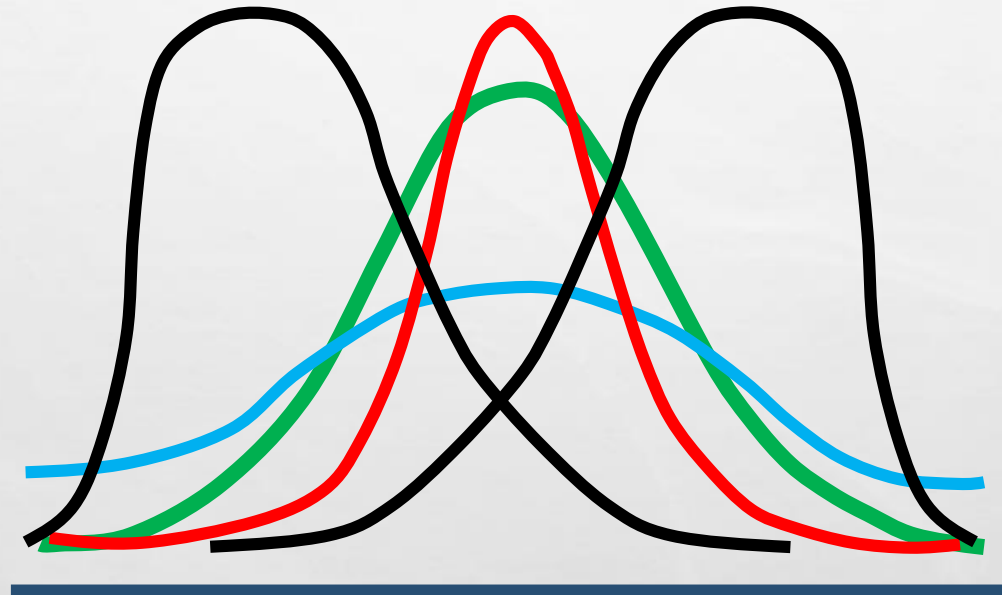
Histogram of  $\text{rbeta}(10000, 2, 5)$



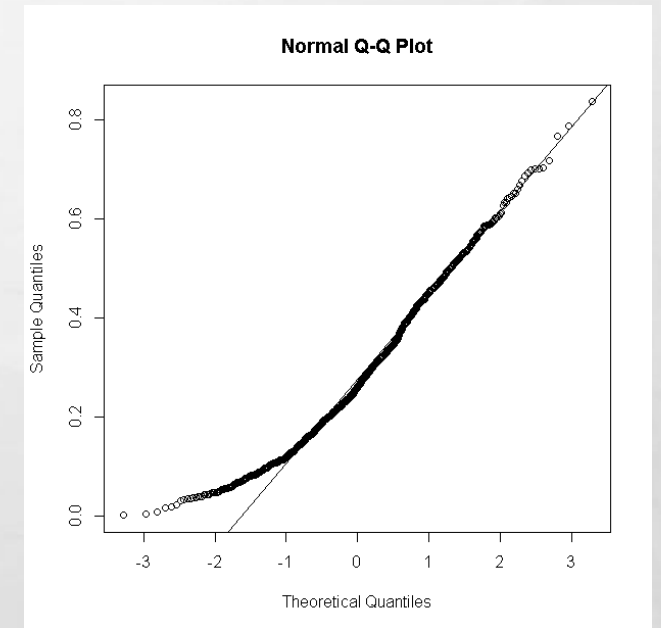
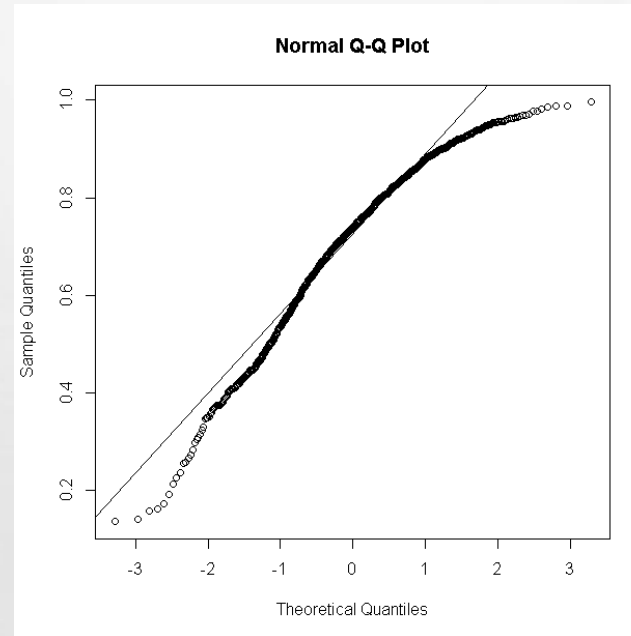
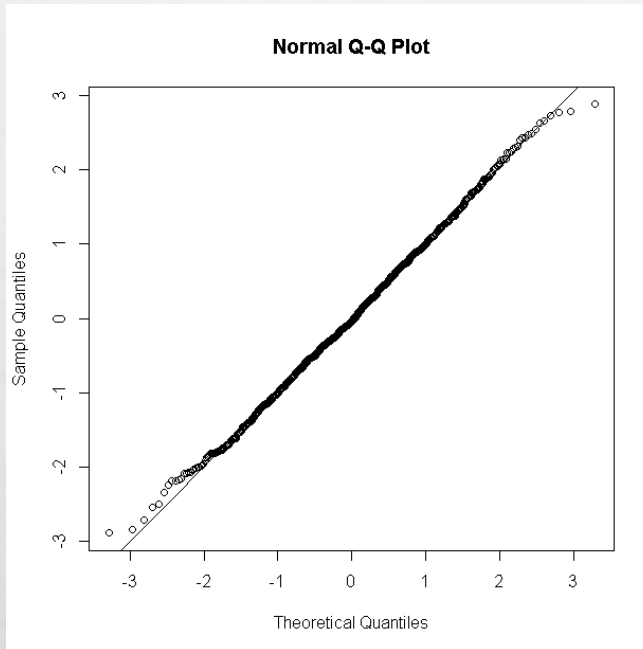
Histogram of  $\text{rbeta}(10000, 5, 2)$



# Outras Distribuições



# Diagrama de Probabilidade Normal



# Teste de Shapiro–Wilk

- **Teste de Hipótese**
- **$H_0$  = Dados estão normalmente distribuídos**
- **$\alpha = 0,05$**
- **valor de  $p \leq 0,05$ : rejeita hipótese nula**
- **valor de  $p > 0,05$ : não é possível rejeitar a hipótese nula**
  
- **p-value = 0.001722 ( $p \leq 0,05$ ) - “não normal”**
- **p-value = 0.05696 ( $p > 0,05$ ) - “normal”**

Valor de p	
$\leq 0,05$	Não
$> 0,05$	Sim

# Qual Rigor?

- **Nem sempre os dados precisam ser rigorosamente normais. Dados aproximadamente normais são suficientes para a maioria dos casos.**



```
> x = rnorm(100)
```

```
> qqnorm(x)
```

```
> qqline(x)
```

```
> shapiro.test(x)
```