

Paulo Jorge Silveira Ferreira

**Filho, não é um bicho:
chama-se Estatística!**

Estatística aplicada – uma abordagem prática

FICHA TÉCNICA

EDIÇÃO: Paulo Ferreira

TÍTULO: Filho, não é um bicho: chama-se Estatística! Estatística aplicada – uma abordagem prática

AUTOR: Paulo Ferreira

CAPA: Sítio do Livro, Lda.

PAGINAÇÃO: Paulo Ferreira

1.ª EDIÇÃO
LISBOA, 2010

IMPRESSÃO E ACABAMENTO: Agapex

ISBN: 978-989-96688-1-2

DEPÓSITO LEGAL: 316073/10

© Paulo Ferreira

PUBLICAÇÃO E COMERCIALIZAÇÃO

Sítio do Livro, Lda.

Lg. Machado de Assis, lote 2 — 1700-116 Lisboa

Índice

0. Estatística: aquele bicho feio!	11
0.1. O que é a estatística?	11
0.2. População vs amostra	13
0.3. Tipo de dados	16
0.4. Softwares estatísticos	17
0.5. Organização do livro	18
1. A diferença de tratamento entre diferentes tipos de dados	19
1.1. Utilização de tabelas para representação de dados	19
1.1.1. Tabelas de frequências para dados univariados	19
1.1.2. Tabelas de contingência para dados bivariados	26
1.2. A representação gráfica dos diferentes tipos de dados	27
1.2.1. Gráficos utilizados em dados univariados	28
1.2.2. Gráficos utilizados em dados bivariados	35
2. Estatística descritiva com dados univariados	39
2.1. Medidas de localização	39
2.1.1. Média	40
2.1.2. Moda	44
2.1.3. Mediana	45
2.1.4. Quantis	48
2.2. Medidas de dispersão	51
2.2.1. Amplitude total e amplitude inter-quartil	51
2.2.2. Variância, desvio-padrão e desvio médio absoluto	52
2.2.3. Coeficientes de dispersão e de variação	56
2.3. Medidas de assimetria	57
2.4. Medidas de achatamento	60
2.5. A caixa de bigodes	61
3. Estatística descritiva com dados bivariados	65
3.1. Covariância e coeficiente de correlação linear de Pearson	65
3.2. Coeficiente de correlação de Spearman	69
3.3. Concentração entre variáveis	72
3.3.1. Curva de Lorenz	72
3.3.2. Índice de Gini	75
4. Noções básicas de probabilidades	79
4.1. Conceitos elementares de probabilidade	79
4.2. Propriedades e álgebra das probabilidades	81
4.3. Teorema da probabilidade total e teorema de Bayes	85

5. Variáveis aleatórias	89
5.1. Variáveis aleatórias unidimensionais	89
5.1.1. Definição de uma variável aleatória unidimensional	89
5.1.2. Função de probabilidade e função de distribuição	90
5.1.3. Parâmetros de uma variável aleatória unidimensional	93
5.2. Variáveis aleatórias bidimensionais	95
5.2.1. Definição de variável aleatória bidimensional	96
5.2.2. Função de probabilidade conjunta e função de distribuição conjunta	96
5.2.3. Função de probabilidade marginal e função de probabilidade condicionada	98
5.2.4. Parâmetros de uma variável aleatória bidimensional	100
5.2.5. Independência de variáveis aleatórias	102
5.3. Um exemplo para exercitar!	103
6. Distribuições de probabilidade	105
6.1. Distribuições de probabilidade discretas	105
6.1.1. Distribuição uniforme	106
6.1.2. Distribuição de Bernoulli e distribuição binomial	108
6.1.3. Distribuição binomial negativa	111
6.1.4. Distribuição multinomial	112
6.1.5. Distribuição hipergeométrica	113
6.1.6. Distribuição de Poisson	115
6.2. Distribuições de probabilidade contínuas	118
6.2.1. Distribuição uniforme	120
6.2.2. Distribuição normal	123
6.2.3. Distribuição t-student	132
6.2.4. Distribuições qui-quadrado e F-Snedecor	133
6.2.5. Distribuição exponencial	136
6.3. Aproximações de distribuições de probabilidade	138
6.3.1. Aproximação da distribuição binomial à normal	140
6.3.2. Aproximação da distribuição de Poisson à normal	141
7. Distribuições amostrais de probabilidade	143
7.1. Distribuição amostral da média	146
7.2. Distribuição amostral da proporção	150
7.3. Distribuição amostral da variância	152
7.4. Distribuição amostral da diferença de médias	153
7.5. Distribuição amostral da diferença de proporções	157
7.6. Distribuição amostral da razão de variâncias	158
7.7. Resumo das distribuições amostrais	160

8. Estimação intervalar	163
8.1. Uma metodologia de abordagem aos intervalos de confiança	163
8.2. Intervalos de confiança para a média	164
8.3. Intervalos de confiança para a proporção	168
8.4. Intervalos de confiança para a variância	169
8.5. Intervalos de confiança para a diferença de médias	170
8.6. Intervalos de confiança para a diferença de proporções	173
8.7. Intervalos de confiança para a razão de variâncias	175
8.8. Intervalos de confiança para o coeficiente de correlação	176
8.9. Quadro resumo	178
9. Testes de hipóteses paramétricos	181
9.1. Uma metodologia de abordagem aos testes de hipóteses	184
9.2. Erros nos testes de hipóteses, potência de teste e p-value	185
9.3. Teste de hipóteses para a média	187
9.4. Teste de hipóteses para a proporção	193
9.5. Teste de hipóteses para a variância	194
9.6. Teste de hipóteses para a diferença de duas médias	196
9.7. Teste de hipóteses para a diferença de proporções	200
9.8. Teste de hipóteses para a razão de variâncias	201
9.9. Teste de hipóteses para o coeficiente de correlação	203
9.10. Teste de hipóteses para igualdade de múltiplas médias:	
Análise de variância – ANOVA	205
9.10.1. Análise de variância – 1 factor	208
9.10.2. Análise de variância – 2 factores	212
9.10.3. Testes de comparação múltipla	219
9.11. Quadro resumo	219
10. Teste de hipóteses não paramétricos	223
10.1. Teste de ajustamento do qui-quadrado	223
10.2. Testes de associação	230
10.2.1. Teste de independência do qui-quadrado	230
10.2.2. Teste de correlação ordinal de Spearman	234
10.3. Outros testes não paramétricos – testes de localização	236

11. Regressão linear simples e regressão linear múltipla	239
11.1. O método dos mínimos quadrados na regressão linear simples	239
11.2. O método dos mínimos quadrados na regressão linear múltipla	244
11.3. A tabela Anova e o coeficiente de determinação	246
11.4. Premissas e propriedades dos estimadores	248
11.5. Inferência	251
11.5.1. Testes de hipóteses sobre os parâmetros	251
11.5.2. Intervalos de confiança	257
11.5.3. Previsão pontual e cálculo do resíduo de estimação	258
11.6. Um exemplo de um modelo de regressão linear simples	259
11.7. Um exemplo de um modelo de regressão linear múltipla	261
11.8. Utilização do modelo para inferência	263
11.9. Regressão não linear	265
Anexo – Tabelas estatísticas	269
A.1. Distribuição Normal	270
A.2. Distribuição t-student	271
A.3. Distribuição F-Snedecor	272
A.4. Distribuição qui-quadrado	276
A.5. Transformação de r em Z_r	278
Bibliografia	279