

António Carvalho Pedrosa
Sílvio Marques A. Gama

INTRODUÇÃO COMPUTACIONAL À
**PROBABILIDADE
E ESTATÍSTICA**
COM EXCEL[®]

Índice

Prefácio	XI
Abreviaturas	XIV
Glossário de símbolos	XV
Capítulo 1 Estatística Descritiva	
1.1 Introdução e Terminologia	2
1.2 Organização dos Dados	7
1.2.1 Dados Qualitativos	7
1.2.1.1 Distribuição de Frequência	7
1.2.1.2 Representação Gráfica de Distribuições Qualitativas	9
1.2.2 Dados Quantitativos Discretos	17
1.2.2.1 Distribuição de Frequência	17
1.2.2.2 Representação Gráfica	18
1.2.2.3 Distribuição de Frequência Acumulada	21
1.2.3 Dados Quantitativos Contínuos	25
1.2.3.1 Distribuição de Frequência	25
1.2.3.2 Representação Gráfica	27
1.2.3.3 Distribuição de Frequência Acumulada	33
1.2.4 Usar o Excel para Organizar e Mostrar os Dados	37
1.3 Estatísticas Amostrais – Medidas Descritivas	43
1.3.1 Medidas de Posição ou Localização Central	43
1.3.1.1 Média Aritmética ou Média	43
1.3.1.2 Média Geométrica, Harmônica e Quadrática	46
1.3.1.3 Mediana	49
1.3.1.4 Moda	52
1.3.1.5 Comparação entre Média, Mediana e Moda	54
1.3.2 Medidas de Posição Não Central – Quantis	55
1.3.2.1 Quantil	56
1.3.2.2 Percentil ou Centil	57
1.3.2.3 Decil	57
1.3.2.4 Quartil	58
1.3.2.5 Relação entre Quartis, Decis e Percentis	58
1.3.3 Medidas de Variabilidade	59
1.3.3.1 Amplitude Total e Amplitude Interquartil	59
1.3.3.2 Desvio Absoluto Médio	60
1.3.3.3 Variância e Desvio Padrão	61
1.3.3.4 Coeficiente de Variação	66
1.3.4 Medidas de Assimetria e de Curtose	68
1.3.4.1 Momentos	68
1.3.4.2 Medidas de Assimetria	69
1.3.4.3 Medidas de Curtose	71
1.3.5 Usar o Excel para Obter Medidas Descritivas	73
EXERCÍCIOS	74

Capítulo 2 Probabilidade

2.1	Introdução	80
2.2	Métodos de Contagem	88
2.3	Probabilidade de um Evento	104
2.3.1	Introdução Histórica	104
2.3.2	Definição de Probabilidade	104
2.4	Reunião de Eventos e Regras Aditivas	109
2.5	Probabilidade Condicional	116
2.5.1	Introdução e Definições	116
2.5.2	Independência de Eventos	118
2.6	Intersecção de Eventos e Regras Multiplicativas	121
2.7	Teorema de Bayes	122
2.8	Introdução à Simulação	127
2.8.1	Motivação	127
2.8.2	Natureza da Simulação e Geração de Números Aleatórios	127
2.8.3	Exemplos de Aplicação	129
	EXERCÍCIOS	135

Capítulo 3 Variáveis Aleatórias

3.1	Definição de Variável Aleatória	140
3.2	Variáveis Aleatórias Discretas e Contínuas	142
3.3	Distribuições de Probabilidade Discretas	143
3.4	Distribuições de Probabilidade Contínuas	149
3.5	Distribuições de Probabilidade Conjuntas	154
3.5.1	Variáveis Aleatórias Multidimensionais	154
3.5.2	Distribuições Bidimensionais	155
3.5.3	Função de Distribuição Acumulada Conjunta	157
3.5.4	Distribuições Marginais	158
3.5.5	Distribuições Condicionais	160
3.5.6	Independência de Variáveis Aleatórias	162
3.6	Funções de Variáveis Aleatórias	164
3.7	Média ou Valor Esperado de uma Variável Aleatória	168
3.8	Variância, Coeficiente de Variação, Covariância e Coeficiente de Correlação	174
3.8.1	Variância, Desvio Padrão e Coeficiente de Variação	174
3.8.2	Covariância e Coeficiente de Correlação	176
3.9	Momentos e Função Geradora de Momentos	180
3.9.1	Momentos	180
3.9.2	Função Geradora de Momentos	183
3.10	Variáveis Aleatórias Standardizadas	186
3.11	Desigualdade de Chebyshev	186
3.12	Simulação no Cálculo de Probabilidades	188
	EXERCÍCIOS	196

Capítulo 4 Distribuições Discretas

4.1	Parametrização das Distribuições	204
4.2	Distribuição Discreta Uniforme – $DU(i, j)$	205
4.2.1	Aplicações	205
4.2.2	Definição	205
4.2.3	Propriedades	206
4.2.4	Exemplos de Aplicação	207
4.3	Distribuição de Bernoulli – Bernoulli(p)	209
4.3.1	Processo de Bernoulli	209
4.3.2	Aplicações	209
4.3.3	Definição	209
4.3.4	Propriedades	210
4.3.5	Exemplos de Aplicação	211
4.4	Distribuição Binomial – Bin(n, p)	212
4.4.1	Aplicações	212
4.4.2	Definição	212
4.4.3	Propriedades	214
4.4.4	Exemplos de Aplicação	215
4.5	Distribuição Geométrica – Geom(p)	217
4.5.1	Aplicações	217
4.5.2	Definição	217
4.5.3	Propriedades	218
4.5.4	Exemplos de Aplicação	219
4.6	Distribuição Binomial Negativa – BinNeg(s, p)	223
4.6.1	Aplicações	223
4.6.2	Definição	223
4.6.3	Propriedades	225
4.6.4	Exemplos de Aplicação	226
4.7	Distribuição Hipergeométrica – HipGeom(N, K, n)	229
4.7.1	Aplicações	229
4.7.2	Definição	229
4.7.3	Propriedades	231
4.7.4	Exemplos de Aplicação	232
4.8	Distribuição de Poisson – Poisson(λ)	235
4.8.1	Aplicações	235
4.8.2	Definição	235
4.8.3	Propriedades	238
4.8.4	Exemplos de Aplicação	239
4.9	Relação entre as Distribuições Discretas	242
4.10	Cálculo de Probabilidades Usando o Excel	243
	EXERCÍCIOS	245

Capítulo 5 Distribuições Contínuas

5.1	Distribuição Uniforme – $U(a, b)$	250
5.1.1	Aplicações	250
5.1.2	Definição	250
5.1.3	Propriedades	251
5.1.4	Exemplos de Aplicação	252
5.2	Distribuição Normal – $N(\mu, \sigma^2)$	254
5.2.1	Aplicações	254
5.2.2	Definição	254
5.2.3	Propriedades	257
5.2.4	Tabulação da Distribuição Normal e Aplicações	260
5.2.5	Distribuição Normal Bidimensional – $N(\mu_x, \sigma_x^2; \mu_y, \sigma_y^2; \rho)$	272
5.2.6	Distribuição Normal e sua Aproximação à Binomial e à de Poisson	273
5.3	Distribuição Exponencial – $\text{Exp}(\beta)$	278
5.3.1	Aplicações	278
5.3.2	Definição	278
5.3.3	Propriedades	279
5.3.4	Exemplos de Aplicação	280
5.4	Distribuição Gama – $\text{Gama}(\alpha, \beta)$	283
5.4.1	Aplicações	283
5.4.2	Definição	283
5.4.3	Propriedades	285
5.4.4	Exemplo de Aplicação	286
5.5	Distribuição Qui-Quadrado – $\chi^2(v)$	287
5.5.1	Aplicações	287
5.5.2	Definição	287
5.5.3	Propriedades	288
5.5.4	Exemplos de Aplicação	290
5.6	Distribuição t -Student – $T(v)$	292
5.6.1	Aplicações	292
5.6.2	Definição	292
5.6.3	Propriedades	294
5.6.4	Exemplos de Aplicação	297
5.7	Distribuição F – $F(v_1, v_2)$	298
5.7.1	Aplicações	298
5.7.2	Definição	298
5.7.3	Propriedades	300
5.7.4	Exemplos de Aplicação	303
5.8	Distribuição de Cauchy – $\text{Cauchy}(\mu, \sigma)$	304
5.9	Relação entre as Distribuições Contínuas	305
5.10	Cálculo de Probabilidades Usando o Excel	306
	EXERCÍCIOS	309

Capítulo 6 Amostragem e Distribuições Amostrais

6.1	Introdução	314
6.2	Métodos de Amostragem	316
6.2.1	Introdução à Amostragem Aleatória	318
6.2.2	Amostragem Aleatória Simples	319
6.2.3	Outros Métodos de Amostragem Aleatória	325
6.2.4	Amostragem Determinística ou Empírica	328
6.3	Realização de Amostragem Aleatória	330
6.3.1	Geração de Números Aleatórios	330
6.3.2	Geração de Realizações de Variáveis Aleatórias	334
6.3.3	Geração de Observações Normais, Binomiais e de Poisson	338
6.4	Distribuições Amostrais	340
6.4.1	Introdução	340
6.4.2	Estatísticas Amostrais	340
6.4.3	Distribuição da Média Amostral	345
6.4.4	Distribuição da Diferença entre Médias Amostrais de Duas Populações Distintas	350
6.4.5	Distribuição da Variância Amostral	353
6.4.6	Distribuição de $\frac{\bar{X} - \mu}{S/\sqrt{n}}$	355
6.4.7	Distribuição do Quociente de Variâncias Amostrais	357
6.4.8	Distribuição da Proporção Amostral	358
6.4.9	Distribuição da Diferença entre Proporções Amostrais	360
6.5	Geração de Realizações de Variáveis Aleatórias Usando o Excel	362
	EXERCÍCIOS	368

Capítulo 7 Estimação

7.1	Introdução	374
7.2	Estimador e Estimativa	375
7.3	Métodos para Determinar Estimadores	376
7.3.1	Método dos Momentos	376
7.3.2	Método da Máxima Verosimilhança	378
7.4	Propriedades dos Estimadores	385
7.4.1	Erro Padrão	385
7.4.2	Enviesamento	385
7.4.3	Eficiência e Erro Quadrático Médio	390
7.4.4	Consistência	393
7.4.5	Informação de Fisher	397
7.4.6	Estatísticas Suficientes	398
7.4.7	Propriedades de \bar{X} , S^2 e \hat{P}	402
7.5	Intervalos de Confiança	403
7.5.1	Definição e Interpretação	403
7.5.2	Estimação por Intervalo para a Média	405
7.5.3	Estimação por Intervalo para a Diferença entre Médias	411

7.5.4	Estimação por Intervalo para Proporções	418
7.5.5	Estimação por Intervalo para Variâncias	421
7.6	Variáveis Fulcrais	425
7.7	Usar o Excel na Estimação de Parâmetros	428
	EXERCÍCIOS	433

Capítulo 8 Teste de Hipóteses

8.1	Introdução	440
8.2	Testes Paramétricos	441
8.2.1	Notação e Metodologia	441
8.2.1.1	Hipóteses Nula e Alternativa	441
8.2.1.2	Estatística de Teste	443
8.2.1.3	Regra de Decisão Estatística	444
8.2.1.4	Erros de Inferência	445
8.2.1.5	Testes e Pressupostos	451
8.2.1.6	Valor de Prova	452
8.2.1.7	Metodologia dos Testes	453
8.2.1.8	Relação entre IC e Testes de Hipóteses	454
8.2.1.9	Significado Prático dos Testes Estatísticos	454
8.2.2	Teste ao Valor Esperado de uma População	455
8.2.2.1	Variância Conhecida	455
8.2.2.2	Variância Desconhecida	460
8.2.3	Teste à Diferença entre Médias	463
8.2.3.1	Variâncias Iguais	463
8.2.3.2	Variâncias Diferentes	466
8.2.4	Teste de Hipóteses a Variâncias	468
8.2.4.1	Teste à Variância de uma População	468
8.2.4.2	Teste às Variâncias de duas Populações	471
8.2.5	Testes de Hipóteses a Proporções	474
8.2.5.1	Teste a uma Proporção	474
8.2.5.2	Teste à Diferença entre duas Proporções	476
8.2.6	Teste Mais Potente – Hipóteses Simples	478
8.2.7	Teste Uniformemente Mais Potente – Hipóteses Compostas	482
8.3	Testes Não-Paramétricos	486
8.3.1	Testes de Aderência ou de Qualidade de Ajuste	486
8.3.2	Testes de Independência Estatística e de Homogeneidade	492
8.3.2.1	Metodologia e Tabela de Contingência	492
8.3.2.2	Teste de Independência	494
8.3.2.3	Teste de Homogeneidade	496
8.4	Usar o Excel em Testes de Hipóteses	499
	EXERCÍCIOS	504

Capítulo 9 Regressão Linear

9.1	Motivação	510
9.2	Natureza da Relação entre Variáveis	511
9.3	Modelo de Regressão Linear Simples	512
9.4	Método dos Mínimos Quadrados para Estimar β_0 e β_1	515
9.5	Estimadores Pontuais $\hat{\beta}_0$, $\hat{\beta}_1$ e \hat{Y}_i e suas Propriedades	520
9.6	Análise de Variância: ANOVA	523
9.6.1	Partição da Variabilidade Total	523
9.6.2	Graus de Liberdade e Médias de Quadrados de Desvios	525
9.6.3	Tabela ANOVA	526
9.6.4	Regressão e Filtragem	528
9.7	Coefficientes de Determinação e de Correlação Amostral	529
9.7.1	Coefficiente de Determinação	529
9.7.2	Coefficiente de Correlação Amostral	530
9.8	Previsão da Resposta	532
9.8.1	Distribuição Amostral de \hat{Y}_0	532
9.8.2	Intervalo de Confiança para a Resposta Média	533
9.8.3	Previsão de uma Nova Resposta Y_0	534
9.9	Inferências sobre β_1	538
9.9.1	Distribuição Amostral de $\hat{\beta}_1$	538
9.9.2	Intervalo de Confiança para β_1	538
9.9.3	Teste de Hipóteses sobre β_1	539
9.10	Inferências sobre β_0	543
9.10.1	Distribuição Amostral de $\hat{\beta}_0$	543
9.10.2	Intervalo de Confiança para β_0	543
9.10.3	Teste de Hipóteses sobre β_0	544
9.11	Reflexões sobre a Regressão	547
9.12	Usar o Excel na Regressão Linear	551
	EXERCÍCIOS	555

Anexo A

A.1	Teorema de Cramer-Rao	561
A.2	Desigualdade de Cauchy-Schwarz	564
A.3	Fórmula Alternativa para $I(\theta)$	565
A.4	Factorização de Neyman	566
A.5	Amplitude Mínima de um IC	569
A.6	Teorema de Neyman-Pearson	571
A.7	Justificação de Resultados de Regressão Linear	573

Anexo B

	Funções do Excel Utilizadas	576
--	-----------------------------	-----

Bibliografia

583

Índice remissivo

585

Prefácio da terceira edição

Esta terceira edição sofreu algumas alterações importantes. Apesar de se manter a filosofia e organização, não só se atualizou a obra a nível gráfico mas também a versão do Excel utilizada – e, conseqüente, todos os conteúdos relacionados –, tendo ainda sido efetuadas pequenas correções e acertos.

Gostaríamos de agradecer às muitas pessoas que têm usado e aconselhado este manual e, em particular, a todos os que nos ajudaram a melhorá-lo. Também agradecemos à Porto Editora por nos ter permitido atualizar a obra, o que levou à concretização desta nova edição.

Janeiro de 2016

Prefácio

Objectivo

O objectivo principal de *Introdução Computacional à Probabilidade e Estatística* é proporcionar um texto, em língua portuguesa, de métodos probabilísticos e estatísticos com uma abordagem rigorosa e inovadora.

Destina-se a ser usado quer como livro de texto de disciplinas de Probabilidade e Estatística de cursos superiores de Ciências, Engenharia, Economia, Gestão, Marketing, Contabilidade, Sociologia, Biologia e áreas afins, quer como texto para consulta de utilizadores e profissionais de estatística.

Motivação

O projecto arrancou a partir de apontamentos preparados para responder às necessidades dos alunos de disciplinas de Probabilidade e Estatística que leccionámos, procurando incentivar os alunos a usar ferramentas computacionais, como as folhas de cálculo, para resolver vários problemas.

As folhas de cálculo têm, entre outras, a vantagem de facilitar a formulação computacional dos modelos matemáticos, num ambiente agradável, e de fornecer as soluções de uma forma simples e eficiente. Além disso, a quase totalidade dos alunos tem, no presente, acesso fácil a computadores com folhas de cálculo já instaladas.

Abordagem

A abordagem dos diversos tópicos, baseada na apresentação de *problemas tipo*, é intuitiva mas sem sacrificar a correcção estatística à conveniência da exposição. Além disso e ao contrário do habitual, incentiva-se a familiarização dos leitores com o uso de *software* apropriado (folhas de cálculo) para resolver muitos dos problemas propostos.

A abordagem intuitiva deste texto combinada com uma apresentação leve e atraente torna-o acessível a uma vasta gama de leitores. Este livro foi concebido para possibilitar leituras diferentes consoante a curiosidade matemática do leitor. Com esta finalidade, optámos por demonstrar quase tudo (isto é, sempre que os conhecimentos requeridos para a demonstração estejam dentro do âmbito deste texto) tendo rejeitado métodos *ad hoc*. Deste modo, as demonstrações deverão ser entendidas como um modo de cimentar, enriquecer e aprofundar os conhecimentos contidos nos teoremas. No entanto, um leitor com um maior sentido prático poderá omitir as demonstrações sem prejuízo de poder compreender os conceitos fundamentais e os aplicar. Algumas das demonstrações mais complexas encontram-se nos Anexos. As secções e outras partes do texto onde são tratados tópicos tecnicamente mais difíceis são assinaladas com asterisco (*).

A abordagem computacional da matéria é uma característica importante deste livro e é particularmente adequada em probabilidade e estatística. O computador, além de tornar a aprendizagem mais aliciante, pode funcionar como um laboratório para simular a experimentação, possibilitando aos estudantes adquirir maior intuição e compreender mais facilmente a matéria. Por outro lado, o computador permite resolver problemas cuja solução não pode ser obtida de forma analítica.

Folhas de cálculo

As folhas de cálculo actuais, incluindo o Microsoft Excel usado neste livro, são ferramentas preciosas no tratamento estatístico. Utilizamos o Excel da versão portuguesa do Microsoft Office 2010. O uso de outras versões do Excel poderá ocasionar alterações menores ao que se menciona no decorrer deste livro. A escolha do Excel deveu-se ao facto de ser facilmente acessível a estudantes e docentes de escolas universitárias e a outros profissionais. Contudo o Excel não é condição essencial para uma boa leitura e acompanhamento deste livro. Na realidade, existem no domínio público ferramentas informáticas (*freeware* ou *shareware*) compatíveis com o Excel e que correm em plataformas Linux, Unix ou Windows nas diversas configurações. Como exemplo destas ferramentas, citamos o OpenOffice (ver em www.openoffice.org) e o LibreOffice (ver em www.libreoffice.org). Gostaríamos, ainda, de alertar o leitor para a existência da linguagem de programação *R* e de alguns pacotes *Python* dedicados à probabilidade e estatística.

Organização

O livro está organizado em nove capítulos. Em todo o texto, exemplos resolvidos e explicados vão surgindo à medida que os diversos assuntos vão sendo apresentados, terminando cada capítulo com inúmeros exercícios de dificuldade variável.

O leitor que unicamente pretenda aprender os conceitos básicos e métodos para classificar, organizar e sintetizar dados necessita de estudar o primeiro capítulo (Estatística Descritiva).

Os quatro capítulos seguintes (Probabilidade, Variáveis Aleatórias, Distribuições Discretas, Distribuições Contínuas), como um todo, formam um curso introdutório de probabilidade e distribuições de variáveis aleatórias.

Os restantes quatro capítulos (Amostragem e Distribuições Amostrais, Estimação, Teste de Hipóteses, Regressão Linear) correspondem ao que normalmente é leccionado em um segundo curso semestral de estatística.

Agradecimentos

Gostaríamos de agradecer a todas as pessoas e instituições que tornaram possível esta obra. Em particular, queremos mencionar a Engenheira Maria José Mata pelas suas sugestões e pela ajuda prestada na resolução dos exercícios propostos. A organização e conteúdo deste livro beneficiou ainda dos comentários e conversas mantidas com outras pessoas, nomeadamente, o Professor Doutor Bruce Schmeiser da Universidade de Purdue, EUA.

Finalmente, queremos agradecer às nossas famílias que sempre nos apoiaram e encorajaram.

António Carvalho Pedrosa
Sílvio Marques A. Gama

1

ESTATÍSTICA DESCRITIVA

1.1 Introdução e Terminologia

1.2 Organização dos Dados

1.2.1 Dados Qualitativos

1.2.2 Dados Quantitativos Discretos

1.2.3 Dados Quantitativos Contínuos

1.2.4 Usar o Excel para Organizar e Mostrar os Dados

1.3 Estatísticas Amostrais – Medidas Descritivas

1.3.1 Medidas de Posição ou Localização Central

1.3.2 Medidas de Posição Não Central – Quantis

1.3.3 Medidas de Variabilidade

1.3.4 Medidas de Assimetria e de Curtose

1.3.5 Usar o Excel para Obter Medidas Descritivas

EXERCÍCIOS

1.1 Introdução e Terminologia

A palavra *estatística* tem dois significados. O primeiro, e mais comum, refere-se a factos numéricos e está ligado à origem latina da palavra *status* (estado). Os números – tais como o número de acidentes de trabalho por dia em Portugal, o número de habitantes do Porto num dado momento, o salário médio mensal de um trabalhador, a percentagem de alunos do 3.º ciclo que reprovaram num dado ano, o número de cursos de ensino superior em Portugal num dado ano, a percentagem de grandes penalidades convertidas nos jogos da 1.ª Liga da época passada – que usamos para caracterizar certas situações ou comportamentos são estatísticas. Um segundo significado refere-se à área científica *Estatística*.

Definição 1.1 – Estatística

É uma área científica cujo objectivo principal é observar um fenómeno, recolher, analisar e interpretar os dados e auxiliar a formulação de decisões.

Muitos dos problemas que queremos resolver, na nossa vida do dia-a-dia, têm uma natureza incerta, em que não conseguimos definir completamente uma relação de causa-efeito entre as variáveis que julgamos condicionantes e os resultados, e, por isso, atribuímos importância decisiva ao factor sorte. Os métodos estatísticos ajudam-nos a resolver estes problemas, proporcionando-nos os modelos matemáticos que nos permitem prever os resultados, estimando o factor sorte.

Por exemplo, os partidos políticos e a comunicação social estão constantemente interessados em conhecer a opinião dos eleitores acerca dos mais variados aspectos, particularmente, para poderem prever os resultados eleitorais. Por isso, realizam-se sondagens usando diversos meios (telefone, inquéritos de rua ou outros) e, depois de uma análise e tratamento dos dados, recorre-se a modelos matemáticos para extrapolar os resultados para a população. Contudo, nem sempre se consegue prever com a exactidão desejada os resultados pretendidos, porque a opinião da população evolui constantemente e a amostra pode não ser representativa da população. Numa amostra, todos os estratos sociais e escalões etários devem estar devidamente considerados nas mesmas proporções em que existem na população. Ora, na prática, esse problema não é fácil de resolver e está na base da formulação de estimativas erradas.

Há igualmente situações em que se usa a Estatística para analisar o impacto económico de muitas decisões. Por exemplo, é frequente as empresas efectuarem sondagens para estimar o efeito de uma campanha publicitária ou para conhecerem as preferências dos consumidores antes de desenvolverem e lançarem um novo produto para o mercado.

Na medicina também se recorre frequentemente à Estatística para, depois de se analisar o efeito de um tratamento novo sobre um conjunto de doentes,

se poder tirar conclusões mais seguras acerca dos benefícios do referido tratamento. Mas os exemplos do uso da Estatística multiplicam-se e estendem-se a todas as áreas da actividade económica, científica, política e social.

A Estatística é normalmente dividida em duas áreas: Estatística Descritiva e Inferência Estatística.

Definição 1.2 – Estatística descritiva

É um conjunto de métodos cujo objectivo é sintetizar e representar de forma compreensível a informação contida nos dados.

Usam-se frequentemente medidas, tabelas e gráficos para sintetizar e representar a informação.

Quanto à Inferência Estatística, antes de procedermos à sua definição, é conveniente introduzir os conceitos de população, observação e amostra.

Definição 1.3 – População

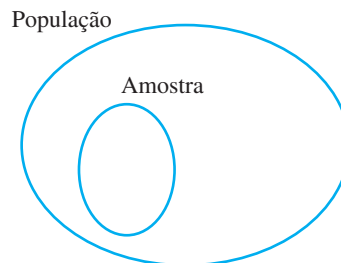
É o conjunto de todos os objectos cuja(s) característica(s) pretendemos estudar.

Definição 1.4 – Amostra

É um subconjunto finito da população.

O número de elementos que constituem a amostra chama-se **tamanho da amostra**.

Figura 1.1
População e amostra.

**Definição 1.5 – Inferência Estatística**

É um conjunto de métodos que permitem fazer estimativas e tirar conclusões sobre uma população a partir da informação contida numa amostra.

As populações podem, no que respeita ao seu tamanho, ser classificadas em populações finitas ou infinitas. Uma **população é finita** se o número de elementos que a constitui é finito. Em caso contrário, a **população é infinita**.

Como exemplos de populações finitas, indicamos as peças que constituem um jogo de dominó ou os alunos de uma dada escola. Em ambos os casos, o número total de objectos (peças do dominó, no primeiro caso; alunos, no segundo caso) é finito. Uma população finita pode ser estudada através de um censo ou de uma sondagem.

Definição 1.6 – Censo

É um estudo que inclui todos os elementos da população (por isso, só é possível realizar censos de populações finitas).

Definição 1.7 – Sondagem

É um estudo (da população) efectuado a partir da análise de uma amostra.

Como exemplo de populações infinitas, citamos as sucessões de faces e coroas obtidas no processo de lançamento consecutivo de uma moeda ao ar ou os resultados obtidos no processo de medição diária da altura do nível da água de uma albufeira. Outro exemplo é o processo de fabrico de memórias de computador a operar indefinidamente sob as mesmas condições que gera a população infinita constituída pelas memórias que seriam produzidas.

De um modo geral, todo o processo que opera indefinidamente sob as mesmas condições gera objectos que conduzem a populações infinitas.

A maioria das vezes as decisões são tomadas com base em amostras, em virtude de se conseguir economizar dinheiro e tempo. Além disso, no caso de populações infinitas, é impossível estudar toda a população.

A amostragem deve ser feita cuidadosamente. Para o estudo estatístico ter validade é conveniente que a amostra seja representativa da população. Voltaremos a este assunto no Capítulo 6.

Exemplo 1.1

Considere as seguintes situações. Identifique, em cada caso, população e amostra.

- (a) Pretende-se estimar a percentagem de votantes num candidato (ou partido) em determinadas eleições.
- (b) Pretende-se determinar se uma encomenda de dezenas de milhar de processadores de computador deve ou não ser aceite.
- (c) Pretende-se estimar a percentagem de lâmpadas defeituosas produzidas por uma fábrica.
- (d) Pretende-se estimar o rendimento médio das famílias portuguesas.

Resolução

- (a) A população consiste em todos os eleitores registados e a amostra (na ordem das centenas ou milhares de elementos) nos eleitores seleccionados.
- (b) A população é o conjunto de todos os processadores que constituem a encomenda e a amostra é o subconjunto dos processadores inspeccionados (escolhidos entre os que foram fornecidos).