

Estadística Aplicada **2**

Para quem se inicia na estatística.

Para quem, tendo já alguns conhecimentos, necessita esclarecer, aprofundar ou rever tópicos específicos.

Com casos e exemplos do dia a dia que garantem uma aprendizagem fácil e eficaz.



**Elizabeth Reis
Paulo Melo
Rosa Andrade
Teresa Calapez**

5ª Edição
Revista e Aumentada



EDIÇÕES SÍLABO

ESTADÍSTICA APLICADA

Volume 2

ESTATÍSTICA APLICADA

Volume 1

**Probabilidades, Variáveis aleatórias,
Distribuições Teóricas**

Volume 2

**Amostragem, Estimação pontual e por intervalos,
Ensaio de hipóteses paramétricos e não paramétricos**

EXERCÍCIOS DE ESTATÍSTICA APLICADA

Volume 1

**Probabilidades, Variáveis aleatórias,
Distribuições Teóricas**

Volume 2

**Amostragem, Estimação pontual e por intervalos,
Ensaio de hipóteses paramétricos e não paramétricos**

ESTATÍSTICA APLICADA

Volume 2

**Amostragem, Estimação pontual
e por intervalos, Ensaio de hipóteses
paramétricos e não paramétricos**

*Elizabeth Reis
Paulo Melo
Rosa Andrade
Teresa Calapez*

5ª Edição
Revista e Aumentada



É expressamente proibido reproduzir, no todo ou em parte, sob qualquer forma ou meio, **NOMEADAMENTE FOTOCÓPIA**, esta obra. As transgressões serão passíveis das penalizações previstas na legislação em vigor.

Visite a Sílabo na rede
www.silabo.pt

Editor: Manuel Robalo

FICHA TÉCNICA:

Título: Estatística Aplicada – Vol. 2

Autores: Elizabeth Reis, Paulo Melo, Rosa Andrade, Teresa Calapez

© Edições Sílabo, Lda.

1ª Edição – Lisboa, fevereiro de 1996.

5ª Edição – Lisboa, março de 2016.

Impressão e acabamentos: Cafilesa – Soluções Gráficas, Lda.

Depósito Legal: 262878/07

ISBN: 978-972-618-841-4

EDIÇÕES SÍLABO, LDA.

R. Cidade de Manchester, 2

1170-100 Lisboa

Tel.: 218130345

Fax: 218166719

e-mail: silabo@silabo.pt

www.silabo.pt

Índice

NOTA INTRODUTÓRIA À SEGUNDA EDIÇÃO	13
PREFÁCIO	15

Capítulo V – O processo de amostragem

1. INTRODUÇÃO	21
2. ALGUNS CONCEITOS IMPORTANTES NA TEORIA DA AMOSTRAGEM	23
3. QUESTÕES PRÉVIAS AO PROCESSO DE AMOSTRAGEM	26
4. AS FASES DO PROCESSO DE AMOSTRAGEM	27
4.1. A identificação da população alvo / população inquirida	28
4.2. Os métodos de seleção da amostra	30
4.2.1. Métodos de amostragem aleatória	31
4.2.1.1. Amostragem aleatória simples	32
4.2.1.2. Amostragem casual sistemática	35
4.2.1.3. Amostragem estratificada	36
4.2.1.4. Amostragem por clusters	39
4.2.1.5. Amostragem multi-etapas	40
4.2.1.6. Amostragem multi-fásica	41
4.2.2. Métodos de amostragem dirigida	43
4.2.2.1. Amostragem por conveniência	43
4.2.2.2. Amostragem intencional	44
4.2.2.3. Amostragem <i>snowball</i>	45
4.2.2.4. Amostragem sequencial	45
4.2.2.5. Amostragem por quotas	46
EXERCÍCIOS PROPOSTOS	49

Capítulo VI – Distribuições amostrais

1. INTRODUÇÃO	53
1.1. Amostra aleatória	54
1.2. Parâmetros e estatísticas	57

1.3. Lei dos grandes números	59
1.4. Teorema do limite central	61
2. DISTRIBUIÇÕES AMOSTRAIS TEÓRICAS	63
2.1. Distribuição normal	63
2.2. Outras distribuições	66
2.2.1. Distribuição do Qui-quadrado	66
2.2.1.1. Principais características da distribuição do χ^2	67
2.2.1.2. Alguns teoremas	67
2.2.2. Distribuição t de Student	68
2.2.2.1. Principais características da distribuição t de Student	69
2.2.2.2. Alguns teoremas	69
2.2.3. Distribuição do F de Snedecor	70
2.2.3.1. Principais características da distribuição F	71
2.2.3.2. Alguns teoremas	71
3. DISTRIBUIÇÕES AMOSTRAIS DAS ESTATÍSTICAS MAIS IMPORTANTES	73
3.1. Populações Bernoulli	73
3.1.1. Distribuição de uma proporção amostral	75
3.1.2. Distribuição da diferença entre duas proporções amostrais	77
3.2. Populações normais	78
3.2.1. Distribuição da média amostral (\bar{X}) quando a variância σ^2 é conhecida	78
3.2.2. Distribuição da variância amostral (S^2)	79
3.2.3. Distribuição da média amostral (\bar{X}) quando a variância σ^2 não é conhecida	80
3.2.4. Distribuição do quociente de variâncias amostrais ($S_1'^2 / S_2'^2$)	81
3.2.5. Distribuição da diferença entre médias amostrais ($\bar{X}_1 - \bar{X}_2$)	82
3.3. Distribuições amostrais dos extremos	84
3.3.1. Distribuição do máximo da amostra	84
3.3.2. Distribuição do mínimo da amostra	86
EXERCÍCIOS PROPOSTOS	88

Capítulo VII – Estimação de parâmetros

1. INTRODUÇÃO	95
2. ESTIMAÇÃO PONTUAL	96
2.1. Estimadores e estimativas	96
2.2. Propriedades dos estimadores	97
2.3. Métodos de estimação pontual	109
2.3.1. O método da máxima verosimilhança	109
3. ESTIMAÇÃO POR INTERVALOS	117
EXERCÍCIOS PROPOSTOS	130

Capítulo VIII – Ensaio de hipóteses

1. A NECESSIDADE DOS ENSAIOS DE HIPÓTESES	139
2. HIPÓTESES E ERROS	141
3. COMO FAZER UM ENSAIO DE HIPÓTESES	143
4. ERROS NOS ENSAIOS DE HIPÓTESES	151
4.1. Análise de erros	153
4.1.1. O erro tipo I	154
4.1.2. O erro tipo II	157
4.1.3. Minimização dos erros	161
4.2. Função potência do ensaio	167
5. ESCOLHA DA ESTATÍSTICA ADEQUADA AO ENSAIO	173
5.1. Introdução	173
5.2. Ensaio de hipóteses com uma amostra	174
5.2.1. Ensaio para a média μ do universo	174
5.2.1.1. A população é normal e a variância do universo é conhecida	174
5.2.1.2. A população é normal e a variância do universo é desconhecida	174
5.2.1.3. A população é desconhecida	178
5.2.2. Ensaio para a proporção	179
5.2.3. Ensaio para a variância	180

5.3. Ensaio de hipóteses com duas amostras	182
5.3.1. Ensaio para a diferença de médias	182
5.3.1.1. Populações normais e variâncias conhecidas	183
5.3.1.2. Qualquer população, variâncias desconhecidas, mas amostras grandes	183
5.3.1.3. Amostras pequenas, populações normais e variâncias desconhecidas mas iguais	186
5.3.1.4. Amostras pequenas, populações normais e variâncias desconhecidas mas diferentes	189
5.3.1.5. Amostras emparelhadas	189
5.3.2. Ensaio para a diferença de proporções	193
5.3.3. Ensaio para comparação de duas variâncias	196
5.4. Ensaio de hipóteses para mais de duas amostras	201
5.4.1. Ensaio para a diferença de k médias – análise de variância simples	202
5.4.2. Testes de comparação múltipla	208
5.4.3. Ensaio para a diferença de k variâncias	215
EXERCÍCIOS PROPOSTOS	219

Capítulo IX – Testes não-paramétricos

1. INTRODUÇÃO	227
2. TESTES DE AJUSTAMENTO	231
2.1. Teste de ajustamento do qui-quadrado	233
2.2. Teste de Kolmogorov-Smirnov	242
3. TABELAS DE CONTINGÊNCIA	248
3.1. Teste do Qui-quadrado de Independência	248
3.2. Medidas de Associação	255
4. TESTES À IGUALDADE DE DUAS OU MAIS DISTRIBUIÇÕES	258
4.1. Testes à igualdade de distribuições em duas amostras independentes	260
4.1.1. Teste de Mann-Whitney	260
4.1.2. Teste de Kolmogorov-Smirnov para duas amostras	269
4.2. Teste à igualdade de distribuições em mais de duas amostras independentes — o teste de Kruskal-Wallis	273

5. COMPARAÇÕES ENTRE DUAS AMOSTRAS EMPARELHADAS	281
5.1. Teste de McNemar ou de mudança de opinião	282
5.2. Teste do sinal	287
5.3. Teste de Wilcoxon	290
EXERCÍCIOS PROPOSTOS	296

Apêndice – Tabelas de distribuição

Distribuição binomial	301
Distribuição de Poisson	306
Distribuição normal padrão	313
Distribuição do qui-quadrado	314
Distribuição <i>t</i> de Student	315
Distribuição <i>F</i> de Snedecor	316
Valores críticos da distribuição do studentized range para comparações múltiplas (nível de significância 0,05)	318
Valores críticos da distribuição do studentized range para comparações múltiplas (nível de significância 0,01)	320
Quantis da estatística de Kolmogorov-Smirnov para uma amostra	322
Quantis da estatística de Mann-Whitney	323
Quantis da estatística de Kolmogorov-Smirnov para duas amostras de igual dimensão	326
Quantis da estatística de Kolmogorov-Smirnov para amostras de dimensões diferentes	327
Quantis da estatística de Kruskal-Wallis para pequenas amostras	329
BIBLIOGRAFIA	331

Índice Volume 1

NOTA À SEXTA EDIÇÃO

PREFÁCIO

Capítulo I – Introdução

1. DUAS RAZÕES PARA SE ESTUDAR ESTATÍSTICA
2. A NECESSIDADE DA ESTATÍSTICA NAS ÁREAS PROFISSIONAIS E CIENTÍFICAS
3. MÉTODO ESTATÍSTICO DE RESOLUÇÃO DE UM PROBLEMA
4. ESTATÍSTICA DESCRITIVA E INFERÊNCIA ESTATÍSTICA
5. ESCALAS DE MEDIDA DOS DADOS ESTATÍSTICOS
6. ALGUMAS CONSIDERAÇÕES FINAIS
7. UTILIZAÇÃO DO COMPUTADOR

Capítulo II – Teoria das probabilidades

1. RESUMO HISTÓRICO
 2. CONCEITOS DA TEORIA DAS PROBABILIDADES
 3. ÁLGEBRA DOS ACONTECIMENTOS
 4. CONCEITOS DE PROBABILIDADE
 5. AXIOMAS DA TEORIA DAS PROBABILIDADES
 6. PROBABILIDADES CONDICIONADAS
 7. PROBABILIDADE DE INTERSEÇÃO DE ACONTECIMENTOS. ACONTECIMENTOS INDEPENDENTES
 8. TEOREMA DA PROBABILIDADE TOTAL E FÓRMULA DE BAYES
- EXERCÍCIOS PROPOSTOS

Capítulo III – Variáveis aleatórias

1. DEFINIÇÃO
 2. FUNÇÕES DE PROBABILIDADE E DE DISTRIBUIÇÃO DE VARIÁVEIS ALEATÓRIAS UNIDIMENSIONAIS
 3. FUNÇÕES DE PROBABILIDADE E DE DISTRIBUIÇÃO DE VARIÁVEIS ALEATÓRIAS
 5. MOMENTOS
 6. DESIGUALDADES DE MARKOV E CHEBYSHEV
- EXERCÍCIOS PROPOSTOS

Capítulo IV – Distribuições teóricas mais importantes

1. DISTRIBUIÇÕES DISCRETAS
 2. DISTRIBUIÇÕES CONTÍNUAS
- EXERCÍCIOS PROPOSTOS

APÊNDICE – TABELAS DE DISTRIBUIÇÃO

BIBLIOGRAFIA

Nota à quinta edição

Esta edição da obra *Estatística Aplicada*, constitui uma nova versão revista e aumentada. Está estruturada em dois volumes para responder às solicitações de muitos leitores, docentes e alunos, coincidindo cada um dos volumes com programas específicos de *Estatística* lecionados nas várias Universidades.

Nesta edição, o número de exercícios propostos no final de cada capítulo, com as respetivas soluções, foi significativamente aumentado. Muitos destes exercícios encontram-se resolvidos na obra dos mesmos autores: *Exercícios, Estatística Aplicada, Volume 1*, editada em 2012 pela Sílabo.

O primeiro volume, para além do capítulo introdutório, inclui um segundo capítulo sobre *Teoria das Probabilidades*, um terceiro sobre *Variáveis Aleatórias*, sendo o quarto e último sobre *Distribuições Teóricas mais Importantes*.

O segundo volume, sendo maioritariamente dedicado aos métodos de *Inferência Estatística* (capítulos VII, VIII e IX, *Estimação de Parâmetros*, *Ensaio de Hipóteses e Testes não-Paramétricos*), inclui também uma breve introdução aos *Processos de Amostragem* (capítulo V) e a apresentação das *Distribuições Amostrais* (capítulo VI).

Finalmente, agradecemos as sugestões e indicações de muitos leitores que contribuíram para a produção desta nova edição revista e aumentada que, acreditamos constituir uma mais valia para os leitores, docentes e alunos.

Os autores
Lisboa, Março de 2016

Prefácio

Este livro de Estatística Aplicada destina-se a profissionais licenciados ou não e a estudantes universitários que, na vida prática ou no processo de aprendizagem, têm necessidade de saber Estatística e de a aplicar aos problemas mais variados do dia-a-dia. Como objetivos finais, este livro pretende tornar compreensíveis a linguagem e notação estatísticas, bem como exemplificar as suas potenciais utilizações, sem descuidar os pressupostos subjacentes e o rigor teórico necessário.

Deverá referir-se que a escolha do título não foi pacífica. De entre os vários alternativos — Probabilidades e Estatística, Inferência Estatística, etc. — a preferência por Estatística Aplicada justifica-se pela abordagem diferenciada de outras obras já publicadas sobre Inferência Estatística, e que resumidamente pode ser assim descrita: mais do que «ensinar», pretende-se com este livro, a) despertar e estimular o interesse dos leitores pelo método estatístico de resolução dos problemas; b) utilizando uma linguagem simples e acessível, apresentar os conceitos e métodos de análise estatística de modo mais intuitivo e informal; c) acompanhar a apetência teórica com exemplos apropriados a cada situação.

O livro encontra-se dividido em nove capítulos. No capítulo I (Introdução) são explicitadas várias razões para que um profissional, técnico, estudante ou mero cidadão adquira um nível mínimo de conhecimentos em Estatística.

A Teoria das Probabilidades é objeto de estudo do capítulo II. Nele são apresentados os diferentes conceitos de probabilidade e a sua axiomática, dando especial relevo aos teoremas da probabilidade total e de Bayes.

Os terceiro e quarto capítulos, tal como o segundo, são essenciais para a compreensão dos seguintes, relativos à Inferência Estatística. O capítulo III respeita às Variáveis Aleatórias, sua definição, características e propriedades. No quarto capítulo estudam-se em pormenor as distribuições de algumas variáveis aleatórias de importância maior nas áreas de aplicação das ciências sócio-económicas como sejam as distribuições de Bernoulli, binomial, Poisson, binomial negativa, hipergeométrica, multinomial, uniforme e normal.

O capítulo V é dedicado ao estudo dos processos de amostragem, incluindo os diferentes métodos de recolha de uma amostra, enquanto que no capítulo VI se apresentam as distribuições amostrais mais importantes.

Os três últimos capítulos são dedicados à Inferência Estatística propriamente dita. No capítulo VII apresentam-se métodos de estimação de parâmetros, com ênfase especial para o método de máxima verosimilhança. Inclui-se ainda a estimação por intervalos. Os capítulos VIII e IX destinam-se à apresentação, respetivamente, dos ensaios de hipóteses paramétricos e não-paramétricos.

Com exceção do primeiro, todos os restantes capítulos são finalizados com um conjunto de exercícios não resolvidos, acompanhados geralmente das respetivas soluções.

No Apêndice estão incluídas as Tabelas (das distribuições) necessárias à compreensão do texto e à resolução dos exemplos e dos exercícios propostos.

Este livro é o resultado de alguns anos de experiência docente dos seus autores na equipa de Estatística do ISCTE e da tentativa de responder às necessidades sentidas por muitos — alunos e docentes de variadas licenciaturas, docentes do ensino secundário, profissionais e técnicos de diferentes áreas científicas (gestão, economia, sociologia, psicologia, medicina, enfermagem, engenharia, informática, etc.) — que, no decorrer destes anos, e na falta de uma obra que os ajudasse a encontrar as soluções estatísticas apropriadas aos seus problemas, procuraram ajuda junto dos autores.

Sem dúvida que a responsabilidade desta obra é assumida pelos seus autores, mas a sua concretização só se tornou possível com a ajuda, apoio e disponibilidade de muitos. Por isso, não deixando de agradecer a todos os que, directa ou indirectamente, contribuíram para a sua realização, gostaríamos de, nominalmente, dar uma palavra especial de agradecimento aos seguintes docentes de Estatística do ISCTE: Ana Cristina Ferreira, Ana Paula Marques, António Robalo, Fátima Ferrão, Fátima Salgueiro, Graça Trindade, Helena Carvalho, Helena Pestana, João Figueira, J.C. Castro Pinto, J.J. Dias Curto, Margarida Perestrelo e Paula Vicente.

Finalmente, uma palavra de apreço a todos os alunos, quer das licenciaturas do ISCTE, quer dos mestrados do INDEG/ISCTE, cujas sugestões, dúvidas e problemas certamente contribuíram para enriquecer este livro.

Os autores

Capítulo V

O processo de amostragem

Introdução

1

A amostragem e em particular os processos de amostragem aplicam-se em variadíssimas áreas do conhecimento e constituem, muitas vezes, a única forma de obter informações sobre uma determinada realidade que importa conhecer.

A teoria da amostragem é assim um dos instrumentos que possibilita esse conhecimento científico da realidade (sempre complexa), onde outros processos ou métodos alternativos, por razões diversas, não se mostram adequados ou até mesmo possíveis.

Ainda que as pessoas não vejam esta temática, em particular os princípios da teoria da amostragem, como algo banalizado, a verdade é que eles suportam (ou deviam suportar) muitas das mensagens que no seu quotidiano lhes são transmitidas nas mais variadas situações. Senão vejamos:

«Neste último mês foi-me pedido para colaborar em dois inquéritos de rua e até num pelo telefone».

«A telenovela e os programas desportivos continuam a ter as maiores audiências em todo o país».

«Os valores Amizade e Liberdade alteraram-se substancialmente na última década».

«O líder do partido A tem visto nos últimos meses aumentar o seu prestígio em detrimento dos líderes dos partidos B e C».

«A opinião dos consumidores sobre o nosso produto é bastante desfavorável, dadas as razões da sua preferência quanto às diferentes características dos que existem no mercado».

«Nunca tinha pensado que as razões principais do divórcio fossem as que esse artigo refere».

«O lote entregue pelo nosso fornecedor não satisfaz a qualidade a que se comprometeu, pelo que não deverá ser aceite».

«Os nossos concorrentes têm como pontos fortes o cumprimento dos prazos de entrega e as condições de pagamento».

«O índice de preços no consumidor tem baixado substancialmente nos últimos anos».

«De acordo com o interesse manifestado pelos utentes, a Carris vai proceder à reestruturação de algumas carreiras em várias zonas da cidade».

«O baixo clima social existente na empresa poderá ser bastante diminuído por uma comunicação mais cuidada, em particular no que respeita aos quadros superiores e intermédios».

«Quando a estenose aórtica se manifesta por angina de peito, a média de sobrevida não ultrapassa os 5 anos».

Uma boa parte das *mensagens* atrás descritas aparecem como conclusões sobre determinada realidade em que se aplicou a *Inferência Indutiva* — isto é — a partir dos resultados de experiências ou inquéritos que fornecem dados estatísticos sobre determinada investigação, formulam-se conclusões que ultrapassam o âmbito das experiências ou inquéritos efetuados. Ou seja, faz-se a extensão do particular para o geral.

Mas, então, põe-se a questão: serão válidas as conclusões a que se chega?

A Estatística Indutiva fornece as técnicas que permitem realizar as inferências indutivas e controlar e até medir o grau de incerteza que aquelas conclusões possam conter.

Alguns conceitos importantes na teoria da amostragem

O problema da Inferência Indutiva é, do ponto de vista da Estatística, encarado da seguinte forma: a finalidade da investigação é descobrir algo sobre determinada *população* ou *universo*.

Importa assim que se definam alguns conceitos fundamentais na teoria da amostragem:

- *População* ou *universo*

Conjunto de unidades com características comuns.

O conjunto dos utentes da Carris, das famílias moradoras em certos bairros, dos alunos do ISCTE, das peças produzidas por uma máquina em determinado período, dos resultados obtidos no lançamento de um dado, são exemplos de populações ou universos.

Refira-se que os exemplos atrás mencionados referem-se a *populações reais*, com exceção para o conjunto de resultados obtidos com o lançamento de um dado em que tal *universo* ou *população* se diz *hipotética*.

A unidade básica de uma população denomina-se *elemento da população*.

- *Amostra*

Sub-conjunto do universo ou população.

A obtenção de informação sobre parte de uma população denomina-se *amostragem*.

Em geral, o investigador está interessado em certa(s) característica(s) específica(s) da população em estudo. Define-se então uma certa variável X que representará a característica que se pretende avaliar.

A variável X poderá designar o número de filhos, o rendimento disponível ou o atributo de ser trabalhador por conta de outrém ($X = 1$) ou trabalhador por conta própria ($X = 0$) das famílias moradoras em certo bairro (população).

A característica X poderá ser uma variável discreta ou contínua, mas, desde que o elemento tenha sido escolhido ao acaso da população, é uma variável aleatória com uma certa distribuição de probabilidade.

Embora a variável aleatória X designe uma característica de uma população, é frequente utilizar no âmbito da teoria da amostragem a designação X para a própria população.

No estudo das variáveis aleatórias e distribuições, parte-se sempre de determinado modelo probabilístico e a partir dele calculam-se probabilidades de certos resultados e observações.

Na Inferência Estatística, o processo é, como alguns autores afirmam, o inverso — isto é, parte-se de certos resultados ou observações fornecidas para uma amostra e procura-se chegar a um modelo probabilístico.

Exemplo 1

Suponha-se que a população em estudo é constituída por 10 mil famílias residentes em determinada região.

Aquelas famílias utilizam diferentes marcas de óleo alimentar que se encontram à disposição no mercado.

A característica em estudo é o atributo *utilizar o óleo A* ($X = 1$) ou *não utilizar o óleo A* ($X = 0$).

Seja p a proporção das famílias que utilizam o óleo A.

Escolhem-se ao acaso 100 famílias e pretende-se determinar a probabilidade de, no conjunto das 100 famílias, encontrar 30 que utilizem o óleo A (e as restantes 70 utilizarem um outro óleo).

Convém aqui distinguir duas situações:

- *Situação 1*

A proporção das famílias que utilizam o óleo A é conhecida, isto é, o p é conhecido, supondo-se igual a 0,4.

Então, para determinar aquela probabilidade, bastaria aplicar o modelo probabilístico adequado.

Trata-se de uma distribuição hipergeométrica (ou binomial sem reposição), desde que as 100 famílias tenham sido selecionadas sem reposição — o que aliás é a situação que realisticamente tem mais sentido — já que se pressupõe que uma mesma família não pode ser selecionada mais que uma vez.

Se por exemplo $p = 0,4$, isto é, se das 10000 famílias, 4000 utilizam o óleo A, então a probabilidade pedida será dada por¹:

$$\frac{\binom{4000}{30} \binom{6000}{70}}{\binom{10000}{100}}.$$

● *Situação 2*

A proporção das famílias que utilizam o óleo A é desconhecida, isto é, p é desconhecido.

Esta é a situação que, na prática, sucede na maioria das vezes e o objectivo é diferente do da situação anterior.

Ao serem seleccionadas as 100 famílias, o objetivo consiste em tirar conclusões sobre a verdadeira percentagem das famílias que utilizam o óleo A, no total das 10000 famílias. Ou seja, a partir dos resultados de uma amostra, pretende-se concluir para o universo ou população que neste caso é constituído pelas 10000 famílias residentes em determinada região.

É óbvio que as conclusões a que se chega conterão, em maior ou menor grau, uma certa dose de incerteza — que, no entanto, respeitadas certas condições, pode ser medida e controlada.

Não se pode dizer que tais conclusões são verdadeiras ou falsas, a não ser que fossem inquiridas as 10000 famílias e depois se verificasse qual a proporção das que utilizam o óleo A.

Na situação 2 está-se no âmbito da inferência indutiva onde se pretende — utilizando toda a informação disponível a partir da amostra (do particular) — concluir para o universo ou população em estudo (o geral).

Ora, a observação de toda a população (as 10000 famílias) teria um preço demasiado elevado para se obter uma resposta sem qualquer grau de incerteza.

Quando a população é conceptualmente infinita, a sua enumeração torna-se até impossível.

Noutros casos, o processo de amostragem é destrutivo — a numeração completa do Universo é possível, mas teria custos demasiado elevados².



¹ Aquela probabilidade poderá ser dada de forma aproximada por $\binom{10000}{100} 0,4^{30} 0,6^{70}$ dado que p se mantém quase fixo de prova para prova (de tiragem em tiragem), o que corresponderá à aplicação da distribuição binomial. Poder-se-ia ainda fazer a aproximação à distribuição normal já que n é suficientemente grande e p tem um valor intermédio.

² A generalidade dos testes de controlo de qualidade dos produtos ou materiais quanto à resistência, durabilidade, etc., são exemplos disto.

3

Questões prévias ao processo de amostragem

Uma definição clara dos objetivos do estudo a efetuar é fundamental e deve ser feita numa fase anterior ao início daquilo a que chamamos o processo de amostragem.

Definidos os objetivos, nomeadamente as características da População que se pretende estudar, há que efetuar um levantamento e sistematização da informação disponível que no caso se torna relevante.

A formulação e resposta àquelas questões prévias é por demais importante já que pode sugerir um quadro geral de alternativas cuja escolha acaba por condicionar alguma ou algumas fases de qualquer processo de amostragem.

Exemplifique-se:

- i*) Se a informação disponível sobre as variáveis (ou características) em estudo for bastante escassa, as alternativas que se põem na escolha da População, do método de amostragem e na dimensão da amostra serão em mais reduzido número.
- ii*) Se a informação estatística obtida permitir concluir da existência de uma grande variabilidade na(s) característica(s) em estudo, dever-se-á utilizar uma amostra de maior dimensão.



PAULO MELO licenciou-se em Economia pelo ISCTE em 1977. Desde então tem lecionado várias disciplinas na área de Métodos Quantitativos, no ISCTE e na Faculdade de Economia da Universidade Nova de Lisboa, bem como em cursos de pós-graduação promovidos pelo INDEG/ISCTE. Em 1985 concluiu o mestrado em Economia, na área de Economia das Telecomunicações, pela Faculdade de Economia da UNL. Colaborou com a Direção de Planeamento do Banco Fonseca & Burnay e com a Direção de Marketing da TMN. Nos CTT – Telecomunicações foi responsável pelas áreas de Política Tarifária e de Planeamento de Marketing.

ROSA ANDRADE licenciou-se em Economia pelo ISCTE, em 1977. Em 1982 completou a parte escolar do mestrado em Métodos Matemáticos Aplicados à Economia e Gestão de Empresas do ISEG. Lecionou em diversas unidades curriculares na área dos Métodos Quantitativos no ISCTE, tendo iniciado a partir de 1986, uma carreira profissional na área do Marketing e do Planeamento e Controlo. Atualmente é professora auxiliar convidada no ISCTE-IUL onde leciona Estatística I e Estatística II na licenciatura de Gestão e é consultora na Autoridade Nacional de Comunicações (ICP-ANACOM) tendo desempenhado, desde 1993, várias funções técnicas e de chefia nesta organização.

TERESA CALAPEZ licenciou-se em Matemáticas Aplicadas, em 1986, pela Faculdade de Ciências de Lisboa. Concluiu o mestrado em Estatística e Investigação Operacional em 1991 pela mesma Faculdade e o doutoramento em Métodos Quantitativos, na especialidade de Análise Multivariada de Dados, pelo ISCTE-IUL, em 2004. Atualmente é professora auxiliar no ISCTE-IUL onde leciona e coordena várias unidades curriculares na área da Estatística e da Análise de Dados, nos diversos ciclos de ensino. É membro da Sociedade Portuguesa de Estatística, investigadora da BRU/UNIDE-IUL, Business Research Unit do ISCTE-IUL e colabora regularmente com o Dinâmia/CET-IUL.

ELIZABETH REIS licenciou-se em Economia pela Faculdade de Economia da Universidade do Porto em 1979. Fez o primeiro curso de pós-graduação em Economia Europeia (1981) na Universidade Católica Portuguesa. Em 1984 concluiu o M.Sc. em Social Statistics na Universidade de Southampton e, em 1987, o Ph.D. na mesma área científica. É doutora em Métodos Quantitativos para Gestão pela Universidade Técnica de Lisboa, tendo obtido a Agregação em Métodos Quantitativos pelo ISCTE-IUL em 1999, instituição onde atualmente é Professora Catedrática e leciona e coordena unidades curriculares de Estatística, Análise de Dados, Metodologias de Recolha de Informação, Métodos de Investigação e Pesquisa de Mercados nos níveis de Licenciatura, Mestrado e Doutoramento, em diversos domínios da Gestão Empresarial e das Ciências Sociais.

Este livro de Estatística sintetiza e beneficia da experiência acumulada pelos autores em mais de três décadas de ensino em licenciaturas, mestrados e doutoramentos.

Depois de várias edições e reimpressões, que atestam a sua popularidade e excelente acolhimento pelo seu público-alvo, surge agora em sexta edição, revista e ampliada, com mais casos práticos, exercícios, exemplos e a introdução de melhorias na exposição das matérias teóricas apresentadas.

Numa linguagem simples e acessível, profusamente ilustrado, com exemplos e situações do dia a dia este livro desperta e estimula o interesse dos leitores que, quer por motivos académicos quer por motivos profissionais, tenham de aprender ou aplicar a estatística.

As matérias abordadas são as que fazem parte da generalidade das disciplinas de estatística dos diversos graus de ensino.

Profissionais e estudantes de gestão, economia, sociologia, psicologia, antropologia, medicina, engenharia, informática, comunicação social e muitos outros dispõem agora de um novo manual, revisto e aumentado, a que poderão sempre recorrer.

Estatística Aplicada 2

ISBN 978-972-618-841-4



9 789726 188414