

Manual de
Auto-aprendizagem

Estatística Descritiva

Carla Santos

3^a Edição
Revista e Aumentada



EDIÇÕES SÍLABO

ESTATÍSTICA DESCRITIVA

MANUAL DE AUTO-APRENDIZAGEM

Carla Maria Lopes da Silva Afonso dos Santos

3ª Edição
Revista e Aumentada

EDIÇÕES SÍLABO

É expressamente proibido reproduzir, no todo ou em parte, sob qualquer forma ou meio gráfico, electrónico ou mecânico, inclusive fotocópia, este livro. As transgressões serão passíveis das penalizações previstas na legislação em vigor. Não participe ou encoraje a pirataria electrónica de materiais protegidos. O seu apoio aos direitos dos autores será apreciado.

Visite a Sílabo na rede:

www.silabo.pt

FICHA TÉCNICA:

Título: Estatística Descritiva – Manual de Auto-aprendizagem

Autora: Carla Maria Lopes da Silva Afonso dos Santos

© Edições Sílabo, Lda.

Capa: Pedro Mota

1ª Edição – Lisboa, Janeiro de 2008.

3ª Edição – Lisboa, Setembro de 2018.

Impressão e acabamentos: Europress, Lda.

Depósito Legal: 445693/18

ISBN: 978-972-618-968-8



EDIÇÕES SÍLABO, Lda.

Publicamos conhecimento

Editor: Manuel Robalo

R. Cidade de Manchester, 2
1170-100 LISBOA

Telf.: 218130345

e-mail: silabo@silabo.pt

www.silabo.pt

ÍNDICE

Apresentação	11
---------------------	----

CAPÍTULO 1

Introdução

1. Âmbito e objectivo da Estatística	15
2. Etapas do Método Estatístico. Estatística Descritiva	16
3. Amostragem	17
3.1. Amostragem aleatória simples	18
3.2. Amostragem sistemática	22
3.3. Amostragem estratificada	23
3.4. Amostragem por conglomerados	25
4. Metodologia da recolha de dados	26
5. Dados qualitativos e dados quantitativos	28

CAPÍTULO 2

Técnicas de tratamento de dados

1. Distribuição de frequências. Representação gráfica	35
1.1. Tabela de frequências	35
1.2. Gráficos	48
1.2.1. Gráfico linear	50
1.2.2. Gráfico de barras	50
1.2.3. Pictograma	52

1.2.4. Sectograma ou gráfico circular	53
1.2.5. Histograma	58
2. Medidas de localização	61
2.1. Medidas de tendência central	61
2.1.1. Média	61
2.1.2. Mediana	66
2.1.3. Moda	73
2.2. Medidas de tendência não central	82
3. Medidas de dispersão	90
3.1. Medidas de dispersão absoluta	90
3.1.1. Intervalo de variação	90
3.1.2. Intervalo inter-quartis	91
3.1.3. Desvio médio absoluto	91
3.1.4. Desvio médio quadrático (variância) e desvio padrão	93
3.2. Medidas de dispersão relativa	99
4. Medidas de forma	101
4.1. Assimetria	101
4.2. Achatamento ou curtose	104
5. Diagrama de extremos e quartis	106
6. Medidas de concentração. Curva de Lorenz e Índice de Gini	110

CAPÍTULO 3

Distribuições bidimensionais

1. Tabela de Contingência. Distribuição conjunta. Distribuições marginais	125
2. Diagrama de dispersão	138
3. Análise de regressão	139
3.1. Regressão linear simples. Método dos mínimos quadrados	141
3.2. Coeficiente de correlação. Coeficiente de determinação	145

CAPÍTULO 4

Séries Cronológicas

1. Representação gráfica	157
2. Modelos clássicos de decomposição de séries cronológicas	160
2.1. Componentes de uma série cronológica	160
2.2. Modelos de decomposição: aditivo e multiplicativo	161
2.2.1. Modelo aditivo	163
2.2.2. Modelo multiplicativo	174

CAPÍTULO 5

Números Índices

1. Índices simples	185
2. Índices compostos	192
2.1. Índices de Laspeyres e Paasche	193
2.2. Índice de Fisher	195
3. Deflação de séries cronológicas	202

CAPÍTULO 6

Exercícios Resolvidos

1. Técnicas de tratamento de dados unidimensionais	207
2. Distribuições bidimensionais	231
3. Séries cronológicas	240
4. Números índices	249

Soluções dos exercícios propostos	255
--	-----

Exercícios resolvidos com SPSS	261
---------------------------------------	-----

Anexo: propriedades dos somatórios	293
---	-----

Bibliografia	297
---------------------	-----

APRESENTAÇÃO

Pretendendo não ser mais um livro de Estatística Descritiva, igual a tantos outros, este texto foi concebido de forma a poder servir um público-alvo mais vasto, permitindo que qualquer pessoa, independentemente da sua formação académica, possa compreender e usar a Estatística, mesmo estudando sozinha.

Sendo um livro de fácil leitura, em que os conceitos aparecem, numa sequência natural, explicados pormenorizadamente com objectividade e clareza, revela-se de grande utilidade, como documento de apoio, para profissionais das mais diferentes áreas, Gestão, Medicina, Engenharia, Ciências Sociais ou outras, que tenham necessidade de usar a Estatística na sua actividade profissional.

Como livro de apoio, a disciplinas cujo programa aborde a Estatística Descritiva, a sua adopção permite colocar em pé de igualdade alunos com diferentes graus de formação matemática, uma vez que apenas exige conhecimentos matemáticos básicos para a sua compreensão.

Com base na vasta experiência no ensino de Matemática/Estatística, ciente das dificuldades sentidas por quem se encontra no início da sua formação em Estatística e/ou tem reduzida formação em Matemática, o texto foi estruturado de forma que a componente teórica, subjacente a cada conceito, esteja sempre acompanhada por uma ou mais aplicações práticas que permitam consolidar a sua assimilação. Para cada item foram incluídos exercícios resolvidos, explicados passo-a-passo, seguidos de exercícios propostos que permitem, ao leitor, aplicar os conceitos estudados.

O primeiro capítulo, é um capítulo introdutório onde se apresenta um pequeno resumo de quais os procedimentos preliminares necessários à realização de uma pesquisa estatística, nomeadamente o tipo de amostragem usada e a metodologia da recolha de dados.

É no segundo capítulo que se entra na Estatística Descritiva propriamente dita, com a descrição pormenorizada das técnicas de tratamento de dados unidimensionais, desde a sua apresentação em tabelas e gráficos até ao estudo das medidas estatísticas que permitem caracterizar rigorosamente uma distribuição.

O terceiro capítulo é dedicado ao estudo de variáveis bidimensionais, com destaque para a análise de regressão.

No capítulo quatro, são apresentados os modelos clássicos de decomposição de séries cronológicas.

O capítulo cinco, de especial interesse para o estudo de fenómenos de natureza económica, aborda os números índices.

Como complemento, mas particularmente importante para quem pretenda estudar grandes volumes de dados, alguns dos exercícios são também resolvidos recorrendo ao *software* de análise estatística, IBM SPSS Statistics 24, em português. Estas resoluções incluem, para além dos relatórios gerados pelo programa, uma descrição sequencial de quais as funções e menus usados em cada caso.

A autora

■ Ícones identificativos dos exercícios



Exercícios resolvidos



Exercícios propostos



Exercícios resolvidos com SPSS

1

CAPÍTULO

INTRODUÇÃO

1

ÂMBITO E OBJECTIVO DA ESTATÍSTICA

Desde sempre as sociedades sentiram necessidade de conhecimento numérico dos recursos existentes. Desta forma, as primeiras operações estatísticas, tinham como objectivo conhecer as características das populações através de contagens.

Das muitas referências a levantamentos estatísticos presentes na história, poder-se-ão destacar, a mais antiga, devida a Heródoto, que refere um estudo realizado no Egipto em 3050 a.C. com o objectivo de averiguar quais os recursos, humanos e económicos, disponíveis para a construção das pirâmides e as existentes na Bíblia que testemunham a realização de dois recenseamentos, um ordenado por Deus a Moisés, em 1490 a.C., e outro que alguns autores defendem ter sido o motivo da viagem de Maria e José a Belém, aquando do nascimento de Jesus.

Mantendo-se durante muitos séculos, quase exclusivamente, como uma técnica de contagem, usada para traduzir numericamente factos e fenómenos, a Estatística teve o seu ponto de viragem, no sec. XVII, quando passou a incluir a análise dos fenómenos, através da construção de modelos para o comportamento e evolução desses fenómenos, possibilitando a realização de previsões.

Hoje em dia, num mundo cada vez mais dependente da informação, a Estatística, tornou-se uma ferramenta imprescindível na tomada de decisões, em áreas tão diversas como a Agricultura, a Medicina, a Engenharia ou o Marketing, entre muitas outras.

Do conjunto de procedimentos e técnicas que compõem a Estatística distinguem-se os que servem para recolher, organizar, sintetizar e descrever os dados, que formam a Estatística Descritiva, e os que, com base na Teoria das Probabilidades, permitem a análise e a interpretação dos dados, assim como efectuar inferências sobre uma população com base no estudo de uma amostra: Estatística Indutiva ou Inferência Estatística.

2**ETAPAS DO MÉTODO ESTATÍSTICO.
ESTATÍSTICA DESCRITIVA**

Para que um estudo estatístico produza resultados fiáveis é necessário cumprir diversas etapas: a identificação do problema, o planeamento do estudo, a recolha de dados, a apresentação dos dados e a sua análise e interpretação.

A fase inicial de qualquer estudo estatístico consiste em escolher, rigorosamente, o tema, formulando correctamente o problema a ser estudado.

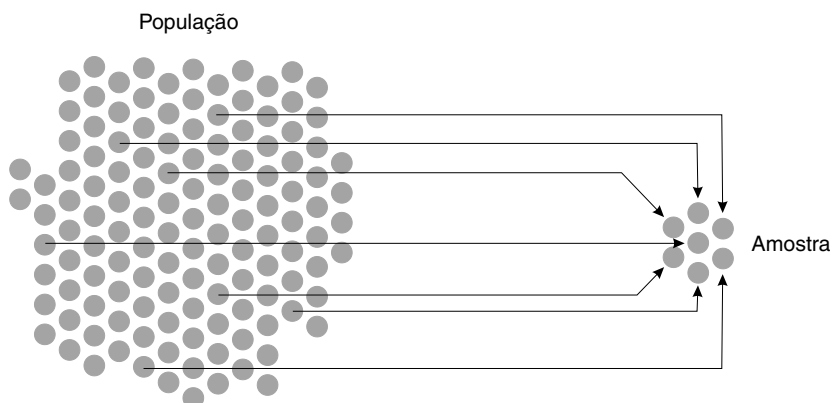
Após a definição do problema há que determinar quais os procedimentos para obter a informação necessária para a resolução do problema. Um dos aspectos a ter em consideração é a definição exacta de quais os indivíduos (pessoas, objectos, ...) que interessa estudar.

Ao conjunto de indivíduos que apresentam uma ou mais características em comum e sobre os quais recai o estudo dá-se o nome de **POPULAÇÃO** (ou Universo). Cada um dos elementos da população, sobre os quais recai o estudo, designa-se por **UNIDADE ESTATÍSTICA**.

Ao realizar um estudo estatístico, muitas vezes não é possível efectuar um **RECEASEAMENTO**, isto é, estudar toda a população, quer por ela ser infinita, por falta de meios, por questões de custo ou muito simplesmente por a sua observação ser destrutiva. Nestes casos, o estudo será feito por **AMOSTRAGEM** e basear-se-á na observação de apenas uma parte da população, a **AMOSTRA**.

Para garantir a qualidade do estudo, o subconjunto da população, que constitui a amostra, deverá ser representativo dessa mesma população para que os resultados obtidos através da observação da amostra sejam, tanto quanto possível, semelhantes com os que se obteriam se tivesse sido estudada toda a população. Em traços gerais, a amostra será representativa se incluir elementos que tenham todas as características que figuram na população e nas devidas proporções. Desta forma é possível estimar certas características desconhecidas da população e testar hipóteses ou afirmações acerca dessa população.

FIGURA 1.1. POPULAÇÃO E AMOSTRA



3

AMOSTRAGEM

Ao optar por um estudo por amostragem há que definir quais as unidades estatísticas que vão pertencer à amostra. Para tal é necessário saber quantas unidades se devem observar, como vão ser seleccionadas essas unidades e qual o grau de precisão pretendido ao extrapolar os resultados da amostra para a população.

Intuitivamente, assume-se que quanto maior for a dimensão da amostra melhores serão os resultados obtidos, mas este não é o único factor que intervém na precisão dos resultados. A forma como são escolhidas as unidades estatísticas, que irão compor a amostra, é também muito importante.

Atendendo a que, para realizar um estudo por amostragem é necessário despender quantias avultadas, a decisão de qual a dimensão da amostra a observar deve basear-se num «equilíbrio» entre precisão de resultados e custo. O uso de técnicas de amostragem adequadas permite aumentar a precisão dos resultados sem aumentar os custos, conseguindo-se

determinar a dimensão da amostra ideal para determinada precisão pretendida (e vice-versa).

MÉTODOS DE AMOSTRAGEM

Não probabilísticos (empíricos)	Amostragem por conveniência
	Amostragem por julgamento
	Amostragem por quotas
Probabilísticos (aleatórios)	Amostragem aleatória simples
	Amostragem sistemática
	Amostragem estratificada
	Amostragem por conglomerados

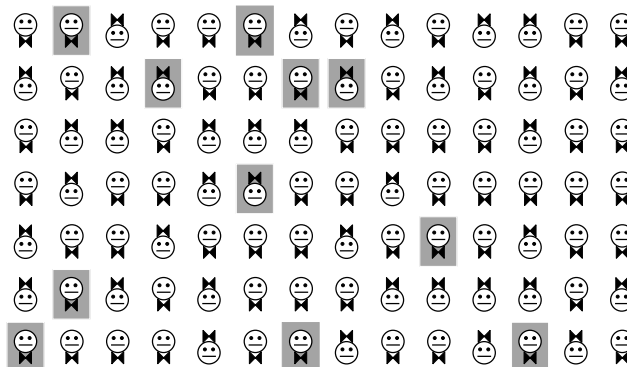
Os métodos de amostragem probabilísticos (aleatórios) são mais complexos, morosos e dispendiosos que os métodos empíricos (não probabilísticos), uma vez que exigem um conhecimento prévio da população em estudo, no entanto, como a selecção é feita de forma aleatória e cada unidade estatística da população tem uma probabilidade conhecida e não nula de pertencer à amostra, têm a vantagem de permitirem controlar a precisão dos resultados.

3.1. AMOSTRAGEM ALEATÓRIA SIMPLES

Quando a população do estudo é homogénea e todos os elementos dessa população têm a mesma probabilidade de vir a pertencer à amostra, usando uma lista numerada desses elementos, a selecção da amostra pode ser feita por sorteio ou usando uma tabela ou gerador de números aleatórios. A esse tipo de selecção chama-se AMOSTRAGEM ALEATÓRIA SIMPLES.

Apesar de ser, à partida, o mais simples dos métodos probabilísticos, o método de amostragem aleatória simples exige uma completa listagem e enumeração da população, tornando-se, muitas vezes, impraticável pelo seu carácter moroso e trabalhoso.

FIGURA 1.2. AMOSTRA ALEATÓRIA SIMPLES

**EXERCÍCIO 1.1**

A direcção da fábrica de esferográficas Grafiline deseja conhecer a opinião dos seus operários quanto ao trabalho por turnos. Para tal pretende realizar um inquérito, usando uma amostra com 10% dos seus 130 operários.

Indique quais os operários que deverão ser inquiridos, com base nos números dos cartões de identificação, numerados de 1 a 130.

Resolução

Para obter uma amostra representativa da população e considerando que todos os elementos da população têm o mesmo «grau de importância», ou seja, a mesma probabilidade de serem seleccionados para a amostra, o método escolhido será a amostragem aleatória simples.

A utilização dos cartões de identificação dos operários (numerados de 1 a 130) dispensa a etapa de numeração dos elementos da população.

Antes do uso de ferramentas informáticas, para geração de números aleatórios, se ter generalizado a selecção dos elementos a incluir na amostra requeria a utilização de uma tabela de números aleatórios, como a que é apresentada a seguir. Usando essa tabela, escolha dos elementos da amostra consiste em escolher um valor como ponto de partida na tabela e percorrer a linha a que pertence esse valor e as seguintes até ter obtido a dimensão desejada da amostra.

Os números obtidos serão o número de ordem (número do cartão de identificação) dos operários que vão compor a amostra.

TABELA DE NÚMEROS ALEATÓRIOS

013	581	704	400	988	100	938	997	298	856
623	023	137	118	929	567	939	964	963	752
490	083	021	121	378	551	866	913	807	504
339	358	318	108	069	677	437	740	568	911
788	770	497	267	700	869	369	114	836	241
492	291	887	676	412	898	843	850	656	196
893	761	037	810	468	719	324	854	469	783
537	160	210	070	665	264	100	820	073	287
605	648	400	391	511	860	203	953	036	272
153	115	795	410	046	868	179	512	423	321
164	239	068	327	070	488	181	099	333	237
489	988	790	798	093	081	523	410	319	759
790	565	366	895	084	982	020	822	827	618
226	750	758	647	791	774	529	789	008	138
549	919	473	901	594	338	884	673	235	631
094	570	597	509	211	043	490	543	018	747
439	199	498	092	644	079	740	644	408	765
012	029	055	194	288	490	873	945	993	761
119	362	265	419	603	912	506	347	898	686
504	497	702	447	912	581	371	138	357	863

A amostra pretendida tem dimensão $n = 13$ ($130 \times 10\%$), pelo que é necessário recolher 13 números da tabela. Como a população disponível tem 130 elementos, todos os números superiores à dimensão da população ou repetidos serão ignorados.¹

Partindo, por exemplo, do primeiro valor da 3ª linha e percorrendo os valores da esquerda para a direita:

(1) Se for usada uma tabela com números de 5 algarismos (por exemplo), para uma população de dimensão 130, a escolha dos elementos da amostra é feita da mesma forma que foi descrita considerando-se apenas os três últimos algarismos de cada número.

CARLA MARIA LOPES DA SILVA AFONSO DOS SANTOS é doutora em Matemática Aplicada. Docente do ensino superior desde 1999, é, actualmente, professora adjunta do Departamento de Matemática e Ciências Físicas do Instituto Politécnico de Beja onde lecciona unidades curriculares de Estatística/Matemática a cursos de Licenciatura e Mestrado, nas áreas das Engenharias, Gestão, Segurança, Ciências Sociais, Educação e Saúde. É autora de diversas comunicações e publicações científicas na área da Estatística e fervorosa divulgadora da Matemática.

Os conceitos e a prática da Estatística Descritiva estruturados e apresentados de forma clara e simples, permitindo que qualquer pessoa, mesmo em auto-aprendizagem e independentemente da sua formação académica, possa compreender e utilizar a Estatística. Sendo os conceitos apresentados num encadeamento intuitivo que facilita a sua absorção e ilustrados com aplicações e exemplos retiradas do dia-a-dia, esta obra, para além de constituir uma excelente ferramenta de apoio para profissionais das mais diferentes áreas (gestão, medicina, engenharia, ciências sociais ou quaisquer outras que tenham necessidade de utilizar a estatística na sua actividade profissional) poderá servir como manual de suporte em disciplinas cujo programa aborde a Estatística Descritiva. Alguns dos exercícios apresentados, para além da sua resolução *clássica*, são resolvidos recorrendo ao programa estatístico SPSS.

- **Exercícios Resolvidos e Propostos**
- **Aplicações em SPSS**

A nossa civilização pode ser definida como a «civilização da informação». Ora a arte de perante um conjunto de dados extrair deles a informação contida é o objectivo central da Estatística. Assim deve-se dar a todos uma competência básica em Estatística. O presente livro representa um esforço bem estruturado nesse sentido. (...) Estamos pois convictos do interesse deste livro não só para alunos como para todos que queiram ter uma cidadania participativa.

João Tiago Mexia

Professor Catedrático
Faculdade de Ciências e Tecnologia
da Universidade Nova de Lisboa

