ESTATÍSTICA APLICADA À GESTÃO

João Veríssimo Lisboa
Mário Gomes Augusto
Pedro Lopes Ferreira

VidaEconómica

Nota prévia

O ensino da estatística em qualquer dos ciclos do ensino da Gestão tem estado sempre presente, refletindo, assim, a importância desta matéria na formação dos gestores. É abundante a literatura existente sobre esta disciplina, sendo também volumosa a quantidade de temas a tratar. Faltava no entanto, em português, um livro que reunisse não só aquelas matérias tradicionalmente incluídas nos *curricula* dos cursos de gestão ao nível do primeiro ciclo, mas também conteúdos de natureza mais avançada, que nestes últimos anos têm vindo a ser objeto de uma grande utilização, quer ao nível da investigação, quer em trabalhos aplicados.

Partimos do princípio que os utilizadores deste livro já têm alguns conhecimentos básicos de estatística, sem descurar no entanto aqueles leitores que, por terem terminado os seus estudos superiores de gestão há mais tempo ou por lacunas na aprendizagem desta disciplina, não têm presentes alguns conceitos considerados fundamentais para a compreensão de matérias mais avançadas. Incluímos para este efeito os três capítulos iniciais que procuram precisamente ir ao encontro destes leitores e ao mesmo tempo refrescar a memória daqueles que já são mais conhecedores dos temas ali tratados.

Este é um livro de características essencialmente pedagógicas, pelo que o seu público alvo serão os alunos que frequentam as disciplinas de estatística, incluídas não só nos planos curriculares das licenciaturas em Gestão, mas também dos cursos de pós-graduação, nomeadamente MBA e programas de doutoramento nesta área. No entanto, dada a natureza aplicada do seu conteúdo, também poderá ser um instrumento útil para os responsáveis pela gestão das empresas, nomeadamente na análise de situações que incluem cenários de incerteza.

Considerando a enorme abrangência da estatística e o público alvo, decidimos incluir neste texto, para além dos quatro capítulos clássicos geralmente destinados aos alunos do primeiro ciclo, o capítulo V, dedicado à análise da variância (Métodos de Taguchi), o capítulo VI, que aborda a estatísticas não paramétricas, o capítulo VII, que trata o tema das componentes principais e análise fatorial, e o capítulo VIII destinado ao estudo dos modelos de equações estruturais, capítulos estes especialmente dirigidos aos alunos de MBA e doutoramento.

O livro está assim organizado em oito capítulos, tendo cada um deles, no final, um conjunto de exercícios de aplicação que permitirão aos alunos testar os conhecimentos adquiridos.

Antes de terminar esta nota introdutória, gostaríamos de deixar uma palavra de agradecimento a um conjunto de instituições, fundamentais na formação académica dos autores, que tornaram possível este livro, nomeadamente a Universidade de Clemson (EUA), a Universidade de Wisconsin (EUA), o Instituto de Sistemas e Robótica (Pólo de Coimbra), a Universidade de Coimbra, a Universidade da Beira Interior, a Fundação Luso-Americana e o Centro de Estudos e Investigação em Saúde da Universidade de Coimbra.

Uma palavra de agradecimento à Editora "Vida Económica" pela disponibilidade e a coragem que teve para a edição deste livro.

Os autores
Coimbra, outubro de 2011

JVL
MGA
PLF

CAPÍTULO I
Funções de distribuição

I – Funções de distribuição

1.1 – Introdução

No dia a dia de uma empresa ocorrem numerosos acontecimentos aos quais podemos associar valores numéricos, por exemplo as vendas mensais, o volume de produção diário, o número de peças produzidas com defeito, etc. Os valores que estes acontecimentos assumem são na maior parte dos casos incertos, isto é, só se conhecem depois de se verificarem. Do mesmo modo, quando simulamos um sistema ou parte dele, os resultados a que chegamos são inicialmente desconhecidos. O modelo que construímos dá origem a resultados que esperamos que se aproximem daqueles que seriam encontrados se pudéssemos ter observado a realidade. Para que esta situação se verifique, terá que haver o cuidado de elaborar um modelo que não só possua uma estrutura que traduza a realidade física do sistema, mas também que inclua um conjunto de instruções de modo a que os componentes do modelo se comportem de acordo com as leis da incerteza que regulam o sistema real. Se assim acontecer, o modelo produzirá resultados que se aproximam do sistema ou subsistema real que procura simular e, por conseguinte, também desconhecidos antes da sua realização. Só sabemos os resultados do exercício económico de uma empresa depois do encerramento do ano, do mesmo modo que só depois de termos simulado durante um determinado período de tempo o processo produtivo de uma empresa, poderemos estimar o número de peças defeituosas que se produziram. A única diferença quando utilizamos simulação é no maior controlo não só sobre as componentes do modelo, mas também do tempo necessário para a obtenção dos resultados que pretendemos. Em qualquer dos casos, seja ela uma situação real ou simulada, há acontecimentos que dão origem a resultados que são desconhecidos antes da sua realização. Aos acontecimentos que se encontram nestas condições chamam-se "acontecimentos aleatórios". À variável numérica que assume valores relacionados com os resultados possíveis dum acontecimento ou experiência aleatória, chama-se "variável aleatória".

As variáveis aleatórias podem classificar-se em discretas e contínuas. Uma variável aleatória diz-se discreta se os valores numéricos associados a cada acontecimento forem numeráveis; diz-se contínua se puder assumir quaisquer valores dentro de um determinado intervalo. Exemplos de variáveis aleatórias discretas são a produção diária de uma máquina, o número de encomendas atrasadas durante uma semana, a quantidade de encomendas recebidas num determinado dia, etc. São exemplos de variáveis aleatórias contínuas a quantidade de líquido contido numa garrafa de refrigerante, o tempo de duração de uma lâmpada, o intervalo de tempo entre a chegada consecutiva de dois clientes, etc.

Perante estes fenómenos, torna-se assim importante conhecer o comportamento de algumas variáveis aleatórias que lhe estão associadas, nomeadamente aquelas que se encontram ligadas a acontecimentos muito frequentes na gestão corrente das empresas. Vejamos em primeiro lugar algumas características gerais das variáveis aleatórias.

1.2 Conceitos básicos sobre probabilidades

1.2.1 O conceito de probabilidade

A definição clássica diz-nos que a probabilidade de um acontecimento é um número compreendido entre 0 e 1 (inclusive), que reflete a possibilidade de ocorrência desse acontecimento, quando o fenómeno a observar pode ser repetido nas mesmas condições e num número indefinido de vezes. De acordo com a definição de Laplace[1], a medida

1 - Pierre-Simon, Marquês de Laplace (1749-1827), famoso matemático e astrónomo francês. Em 1812 publica o livro "Théorie analytique des probabilités".

dessa possibilidade de ocorrência, que se denomina de probabilidade do acontecimento, obtém-se através do quociente entre o número de casos favoráveis e o número total de casos igualmente possíveis. Designando por *A* um acontecimento qualquer, a probabilidade da sua ocorrência *P(A)* será então dada por

$$P(A)=\frac{n}{N}$$

representando *n* o número de casos favoráveis e *N* o de casos possíveis. Por acontecimento, que passaremos a representar por letras maiúsculas *A*, *B*, *C*, etc., designamos um subconjunto do universo das observações, espaço de resultados ou população (Ω). O próprio universo é um acontecimento certo, bem como o conjunto vazio (Ø) correspondente a um acontecimento impossível.

Em termos matemáticos podemos também dizer que a probabilidade é uma função *P* que satisfaz as seguintes condições:

$P(\varnothing) = 0$ e $P(\Omega) = 1$; $P(A) \in [0\ \ 1]$.

Se A_1 e A_2 forem dois acontecimentos disjuntos, isto é, $A_1 \cap A_2 = \varnothing$, então

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

Destas condições podemos obter outras, como por exemplo:

- A soma da probabilidade de um acontecimento com a probabilidade do acontecimento contrário é igual à unidade, isto é, $P(\overline{A}) = 1 - P(A)$;
- Se $A \subset B$ então $P(A) \leq P(B)$;
- Quaisquer que sejam *A* e *B*, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \Leftrightarrow A_i \cap A_j = \varnothing, \ \forall_{i,j \leq n}$ com $i \neq j$.

Exemplo 1.1

De um universo de 500 empréstimos bancários destinados à compra de automóveis foi retirada a seguinte informação:

	Montante emprestado (euros)			
Prazo do empréstimo (meses)	<1000 (B_1)	[1000 - 4000[(B_2)	[4000 - 6000[(B_3)	≥ 6000 (B_4)
12 (A1)	24	5	0	0
24 (A2)	3	14	1	0
36 (A3)	0	28	90	3
42 (A4)	0	52	95	49
48 (A5)	0	0	112	24

Suponhamos que pretendemos determinar a probabilidade de observar o seguinte acontecimento – "Prazo de empréstimo de 36 meses ou montante emprestado compreendido entre 4000 e 6000 euros". Teremos, então,

$$P(A_3 \cup B_3) = P(A_3) + P(B_3) - P(A_3 \cap B_3) = \frac{121}{500} + \frac{298}{500} - \frac{90}{500} = 0{,}658.$$

Ou se pretendermos determinar a probabilidade de selecionar um "empréstimo com um prazo de 24 meses e um montante compreendido entre 1000 e 4000 euros". Teremos, neste caso,

$$P(A_2 \cap B_2) = \frac{14}{500} = 0{,}028.$$

Ou selecionar um empréstimo "não superior a 6000 euros", ou seja, o complementar de B_4:

$$P(\overline{B}_4) = 1 - P(B_4) = 1 - \frac{76}{500} = 0{,}848.$$

Ou ainda se pretendermos calcular a probabilidade de encontrar um empréstimo com um "prazo inferior ou igual a 42 meses". Neste caso, teremos

$$P\left(\overline{A_4 \cup A_5}\right) = 1 - P(A_4 \cup A_5) = 1 - [P(A_4) + P(A_5)] = 1 - 0{,}392 - 0{,}272 = 0{,}336.$$

1.2.2. Leis de De Morgan

Dados dois acontecimentos A e B, a não realização de um ou outro leva-nos a poder concluir da não realização de nenhum deles. Esta afirmação pode ser traduzida em termos matemáticos da seguinte maneira,

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \qquad [1.1]$$

Podemos também afirmar que a não realização conjunta de nenhum deles leva-nos a poder concluir ser equivalente a que não se realize pelo menos um deles. Em termos matemáticos, teremos

$$\overline{A \cap B} = \overline{A} \cup \overline{B} \qquad [1.2]$$

Estas duas expressões são conhecidas por leis de De Morgan[2] e têm muita utilidade na demonstração de relações de lógica e no cálculo de probabilidades.

Exemplo 1.2

Suponhamos que A e B são dois acontecimentos independentes e que $P(A)=0{,}4$ e $P(B) = 0{,}5$. Verifiquemos as leis de De Morgan:

2 - Augustus De Morgan (1806 - 1871), matemático inglês.

A primeira

$$P(\overline{A \cup B}) = 1 - [P(A \cup B)] = 1 - [P(A) + P(B) - P(A) \times P(B)] = 1 - (0{,}4 + 0{,}5 - 0{,}2) = 0{,}30$$

$$P(\overline{A} \cap \overline{B}) = P(\overline{A}) \times P(\overline{B}) = [1 - P(A)] \times [1 - P(B)] = 0{,}6 \times 0{,}5 = 0{,}30$$

A segunda

$$P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - P(A) \times P(B) = 1 - 0{,}4 \times 0{,}5 = 0{,}8$$

$$P(\overline{A} \cup \overline{B}) = P(\overline{A}) + P(\overline{B}) - P(\overline{A}) \times P(\overline{B}) = 0{,}6 + 0{,}5 - 0{,}6 \times 0{,}5 = 0{,}8$$

1.2.3 Probabilidade condicionada

Sejam A e B dois acontecimentos quaisquer, com $P(B)>0$, então a probabilidade condicionada de A dado B, que se representa por $P(A/B)$, é definida pela expressão

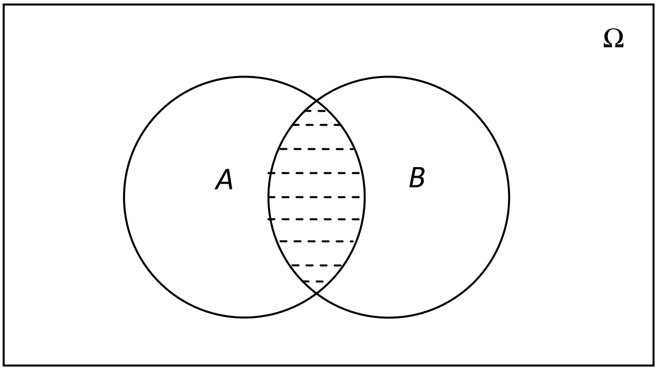
$$P(A/B) = \frac{P(A \cap B)}{P(B)} \qquad [1.3]$$

Desta definição resulta imediatamente que

$$P(A \cap B) = P(A/B) \times P(B) \qquad [1.4]$$

A probabilidade condicionada encontra-se ilustrada no diagrama de Venn (Figura 1.1).

Figura 1.1: Probabilidade condicionada



Uma vez ocorrido o acontecimento *B*, qualquer outro pertencente a A, mas não incluído em $A \cap B$, não poderá ter sido observado. Logo, a probabilidade condicionada de *A*, dado *B*, é obtida dividindo $P(A \cap B)$ por $P(B)$. A expressão que traduz a probabilidade condicionada possui as seguintes propriedades.

1. $P(\varnothing/B) = 0$ e $P(B/B) = 1$.

2. $P(A/B) \geq 0, \quad \forall A$.

3. Se $A_i \cap A_j = 0 \;\; P[(A_i \cup A_j)/B] = P(A_i/B) + P(A_j/B)$.

Exemplo 1.3

Utilizemos de novo a informação do exemplo 1.1, mas admitamos agora que estamos interessados em calcular a probabilidade de observar um empréstimo com um prazo de 24 meses, dado que se selecionaram todos aqueles com um montante compreendido entre 1000 e 4000 euros. Teremos, assim,

$$P(A_2/B_2) = \frac{P(A_2 \cap B_2)}{P(B_2)} = \frac{\frac{14}{500}}{\frac{9}{500}} = \frac{14}{9} = 0{,}141.$$

Dois acontecimentos dizem-se independentes se

$P(A/B) = P(A)$,

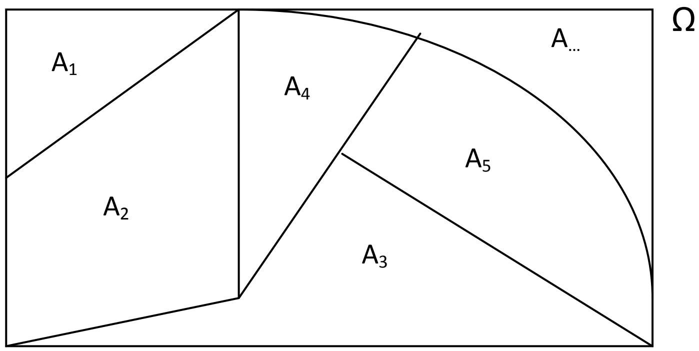
o que significa que o conhecimento do acontecimento B não tem qualquer influência na observação de A. Se dois acontecimentos forem independentes da expressão 1.1, retira-se imediatamente que

$P(A \cap B) = P(A) \times P(B)$.

1.2.4 Teorema de Bayes

Suponhamos que os acontecimentos $A_1, A_2, A_3, \ldots A_n$ constituem uma partição do espaço de resultados Ω (Figura 1.2).

Figura 1.2: Partição do espaço de resultados Ω



Como aqueles subconjuntos constituem uma partição de Ω, então

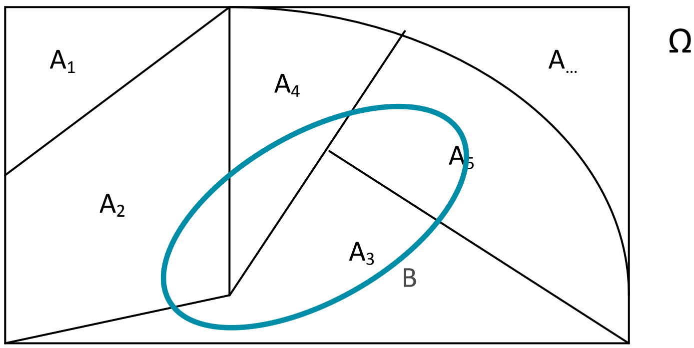
$$\underset{i,j\leq n}{\forall} A_i \cap A_j = \varnothing \quad \text{com} \quad (i \neq j),$$

verificando-se também que

$$\bigcup_{i=1}^{n} \mathrm{A_i} = \Omega.$$

Consideremos agora em Ω um acontecimento B que se sobrepõe total ou parcialmente aos restantes A_i, conforme ilustra a Figura 1.3.

Figura 1.3: Sobreposição do acontecimento B



Das interceções de B com A_i $(B \cap A_i)$ podemos determinar as probabilidades condicionadas $P(B/A_i)$. De facto, as probabilidades condicionadas de $P(A_i/B)$, por definição, são dadas por

$$P(A_i/\mathrm{B}) = \frac{P(A_i \cap B)}{P(B)} \tag{1.5}$$

Mas vejamos como podemos exprimir esta expressão apenas em função das probabilidades condicionadas $P(B/A_i)$ e $P(A_i)$.

Ora, como sabemos, $B \subset \Omega$, logo $\Omega \cap B = B$. Mas $\bigcup_{i=1}^{n} A_i = \Omega$, então

$$
\begin{aligned}
B &= (A_i \cup A_2 \cup A_3 \cup \ldots A_n) \cap B \\
&= (A_i \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B) \cup \ldots (A_n \cap B)
\end{aligned}
$$

Mas como os acontecimentos A_i são uma partição de Ω, $(A_i \cap B)$, para $i = 1, 2, 3, \ldots, n)$, são acontecimentos disjuntos, logo

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + \ldots + P(A_n \cap B) .$$

Mas $P(A_i \cap B) = P(A_i) P(B/A_i)$, para $i = 1, 2, 3, \ldots, n$, logo a expressão 1.5 pode ser escrita da seguinte maneira

$$P(A_i/B) = \frac{P(A_i)\, P(B/A_i)}{P(A_1)\, P(B/A_1) + P(A_2)\, P(B/A_2) + P(A_3)\, P(B/A_3) + \ldots + P(A_n)\, P(B/A_n)}$$

ou seja

$$P(A_i/B) = \frac{P(A_i)\, P(B/A_i)}{\sum_{i=1}^{n} P(A_i)\, P(B/A_i)} \qquad [1.6]$$

Expressão que nos permite calcular as probabilidades $P(A_i/B)$ se forem conhecidas $P(B/A_i)$ e que se denomina por fórmula de Bayes[3].

3 - Thomas Bayes (1702 - 1761), matemático inglês.

Índice Sistemático

CAPÍTULO I: Funções de distribuição

CAPÍTULO II: Intervalos de confiança e ensaios de hipóteses

CAPÍTULO III: Regressão linear simples

CAPÍTULO IV: O modelo de regressão linear múltipla

CAPÍTULO V: Análise da Variância

CAPÍTULO VI: Estatísticas não paramétricas

Apendíces

Tabelas

ESTATÍSTICA APLICADA À GESTÃO

O ensino da estatística em qualquer dos ciclos do ensino da Gestão tem estado sempre presente, refletindo, assim, a importância desta matéria na formação dos gestores. Apesar de ser um livro de características essencialmente pedagógicas, o seu público alvo será os alunos que frequentam as disciplinas de estatística em qualquer dos níveis do ensino superior. No entanto, dada a natureza aplicada do seu conteúdo, também poderá ser um instrumento útil para os responsáveis pela gestão das empresas, nomeadamente na análise de situações que incluem cenários de incerteza.

Considerando a enorme abrangência da estatística e o público a que nos dirigimos, neste texto, para além de quatro capítulos clássicos geralmente destinados aos alunos de licenciatura, decidimos também incluir um que trata o tema das componentes principais e análise fatorial e, um outro, destinado ao estudo dos modelos de equações estruturais, capítulos estes, especialmente dirigidos aos alunos de MBA, doutoramento e investigadores em geral.

Em cada um dos capítulos são apresentados exemplos práticos para melhor compreensão dos assuntos abordados.

www.vidaeconomica.pt

ISBN: 978-972-788-448-3

9 789727 884483 >