Informação Probabilística, Informação Assimétrica

Duarte Trigueiros

ISCTE

Mestrado em Sistemas Integrados

Indice

Capítulo 1

Informação Probabilística, Informação Assimétrica

A difinição de informação dada no início do capítulo anterior apenas contempla o caso em que o dado é capaz de remover *toda* a incerteza porventura existente. Porém, especialmente ao nível da decisão estratégica, os dados são frequentemente de tipo probabilístico e não removem senão parte da incerteza. Por exemplo, o dado que consiste em saber que existem 80% de probabilidade de as vendas crescerem no próximo ano, deixa em aberto a possibilidade (embora pouco verisímil) de as vendas não crescerem. A incerteza, portanto, foi apenas parcialmente removida: diminuiu mas não desapareceu.

Esta informação parcial ou probabilística tem uma expressão matemática simples e, em muitas circunstâncias, útil para os gestores. Tal expressão é, além disso, importante no contexto das Finanças Empresariais.

1.1 Quantidade de Informação

Imagine-se um jogo de dados. A probabilidade de que venha a sair qualquer das caras é a mesma, 1/6, e a informação sobre qual delas irá sair é nula. Se fosse possível viciar os dados de modo a que uma das caras tivesse mais probabilidades de sair do que as outras, a informação sobre qual delas sairia deixava de ser nula: passava a haver *alguma* informação sobre o futuro, mas era uma informação *incompleta*. Se os dados fossem de tal modo enviesados que, em todas as jogadas, apenas pudesse sair uma determinada cara, as probabilidades de que as outras caras saissem seriam zero e a informação sobre o futuro seria então *completa*. Este exemplo serve para introduzir o conceito de *quantidade de informação*, importante para se compreender a natureza da informação estratégica a que os gestores têm acesso. A informação, ao nível esratégico, raramente é completa.

1 Viu-se que informação era qualquer dado capaz de remover incerteza. Quando a incerteza é totalmente removida por um novo dado, está-se perante informação completa. Pode acontecer porém que o dado não remova toda a incerteza: ela diminui mas não desaparece. Nesse caso, a informação que esse dado trouxe consigo é incompleta e faz sentido perguntar se a quantidade de informação que esse dado trouxe foi muita ou pouca (qual o *ganho em informação*), e ainda quanta incerteza falta remover para se ter uma informação completa (qual a *entropia* ainda existente).

As probabilidades são apenas uma entre várias possíveis maneiras de expressar expectativas ou tendências. Existem expectativas quando há informação disponível sobre um desenlace futuro mas esta é incompleta. Existem certezas quando a informação é completa. Estas duas situações devem ser tratadas como distintas. Na primeira, é costume falar-se de uma situação de *risco*.

2. Variedade. Pode medir-se a quantidade de informação de que alguem está carecido, notando que ela é igual ao número de dígitos necessários para distinguir um acontecimento de entre todos os possíveis. Por exemplo, se existem 9 acontecimentos possíveis, é preciso um dígito decimal para comunicar a informação sobre qual deles acabou por acontecer. Se fossem 99 os acontecimentos possíveis, seriam precisos dois dígitos decimais para identificar qual deles tinha saido; 999 acontecimentos iriam requerer três dígitos decimais e por aí fora.

Sendo assim, a quantidade de informação que é precisa para distinguir um de entre N acontecimentos possíveis é igual ao número de dígitos necessários para escrever e comunicar o número N, como referido. Por sua vez, o número de dígitos é igual ao logaritmo desse número N. De facto, lembremos que o logaritmo de 10 é 1, o logaritmo de 100 é 2, e por aí fora. Diz-se pois que $\log N$ é a *variedade* de uma colecção de N acontecimentos possíveis.

$$\text{Variedade } = \log N \tag{1}$$

A variedade é a quantidade de informação que é precisa para comunicar qual dos N acontecimentos possíveis acabou por dar-se.

3. Ganho. Quando existe alguma regularidade na colecção dos N acontecimentos possíveis, o conhecimento desse dado traz consigo informação adicional sobre qual deles acabará por dar-se. Nesse caso, a quantidade de informação que é precisa para distingui-los deixa de ser $\log N$. Deu-se um *ganho* em informação ao saber-se que, por exemplo, k_i dos N acontecimentos possuem um atributo comum. Assim, se existem 99 atletas em competição, são precisos dois dígitos para transmitir a informação de qual deles venceu; mas se esses 99 atletas representam apenas 9 paises, o facto de se saber, à partida, o pais de cada atleta, traz consigo um ganho em informação. A probabilidade de 9/10 associada à extracção de uma bola branca de uma urna dá ideia de uma forte expectativa: existe informação quase completa sobre o desenlace. Isto deve-se ao ganho em informação obtido com o conhecimento de que a urna contem 90 bolas brancas e dez pretas.

Sendo assim, na generalidade,

$$\text{Ganho } = \sum_i \frac{k_i}{N} \log k_i \tag{2}$$

onde a soma de todos os $k_1, \cdots, k_N$ é N. O ganho é pois a quantidade de informação média que uma classificação, previamente conhecida, traz consigo.

4. A Entropia. A quantidade de informação extra obtida por classificação, o ganho, deve ser subtraida à variedade, $\log N$, sempre que se pretenda saber a quantidade de informação que ainda falta para se prever com certeza um desenlace futuro.

$$\text{Informação que ainda falta } = \text{Variedade} - \text{Ganho} \tag{3}$$

Quando, como vimos, existam atributos comuns aos N possíveis desenlaces, a quantidade de informação que ainda falta conhecer para identificar qual deles acabará por dar-se, será pois a diferença, H, entre a informação que faltava antes da classificação (a variedade) e a quantidade de informação média que tal classificação trouxe consigo (o ganho). Isto é,

$$H = \log N - \sum_i \frac{k_i}{N} \log k_i \tag{4}$$

Esta diferença, H, é conhecida pelo nome de *entropia*. A entropia mostra a um gestor a quantidade de informação que ainda lhe falta obter para remover toda a incerteza sobre um acontecimento futuro. É portanto uma medida do grau em que a informação que obteve é incompleta.

Quando $H = 0$, não há falta de informação. O dado obtido pela empresa foi capaz de dissipar toda a incerteza anteriormente existente. Isto deu-se porque o ganho em informação foi igual à variedade $\log N$. Era informação completa. Por exemplo, depois de um estudo de mercado, o gestor fica a saber que a procura de um produto será alta com 100% de probabilidades. No polo oposto, quando H se mantem igual a $\log N$, o dado que o gestor obteve não trouxe consigo nenhum ganho em informação. Neste caso, a incerteza é máxima porque a irregularidade da colecção de acontecimentos possíveis é também máxima. Seria o caso de um gestor que, depois de um estudo de mercado, fosse informado de que a probabilidade de se verificar uma subida na procura era igual à probabilidade de se verificar uma descida. Como tanto uma coisa como outra são igualmente possíveis, esse estudo não acrescentou nada ao conhecimento que o gestor já tinha do futuro.

A situação de $H = 0$ dá-se quando cada um dos acontecimentos pode ser completamente descrito pelos seus atributos, de tal modo que quem conhecer os atributos fica também a saber qual é o desenlace. Continuando com o exemplo anterior, aquilo que permitiria a um estudo de mercado ser tão taxativo em relação ao futuro, seria a observação de determinados indicadores que, infalivelmente, estariam associados ao crescimento. Claro que este tipo de certezas absolutas não se dá na vida real. A situação oposta, $H = \log N$, dá-se quando os atributos porventura existentes, os tais indicadores, não trazem consigo nenhuma informação sobre o desenlace futuro. Entre estes dois extremos, qualquer situação é possível.

5 Viu-se que a entropia era a diferença entre a variedade e o ganho em informação obtido com o conhecimento prévio de certas regularidades existentes na colecção de acontecimentos. Este conhecimento prévio é também designado por *informação a-priori* e está contido em colecções de probabilidades. Por exemplo, no caso de dados viciados, a informação a-priori vem dada pela colecção de probabilidades associadas à saída de cada uma das caras. Porém, a forma como colecções de probabilidades medem expectativas não é sugestiva. Para que um gestor fique de posse de todos os dados necessários à tomada de uma decisão, precisará de ter em consideração, além das probabilidades associadas a cada um dos acontecimentos incertos, o número desses acontecimentos e a sua estrutura. Por exemplo, uma probabilidade de $1/2$ associada a uma procura baixa mede algo muito diferente consoante se tenham considerado dois, ou três possíveis desenlaces. No primeiro caso, ela expressa expectativas nulas ou ausência de qualquer informação a-priori. No segundo, esta mesma probabilidade expressa uma tendência ou expectativa a favor da procura ser baixa, isto é, revela a existência de uma certa quantidade de informação.

6. Entropia e Probabilidades a-priori. Seria fácil de ver que H em (4) pode ser escrita como a média ou valor esperado da informação que falta para conhecer completamente um desenlace:

$$H = -\sum_{i=1}^{N} p_i \log p_i \tag{5}$$

onde p_i é a probabilidade de ocorrência de cada um dos N possíveis desenlaces.

Considerar a entropia ou o ganho em vez de uma colecção de probabilidades e sua estrutura, simplifica e faz mais realista a tomada de decisões. Para entender porquê, considerar-se-á uma decisão à qual se seguem, para cada possível movimento do gestor, um acontecimento

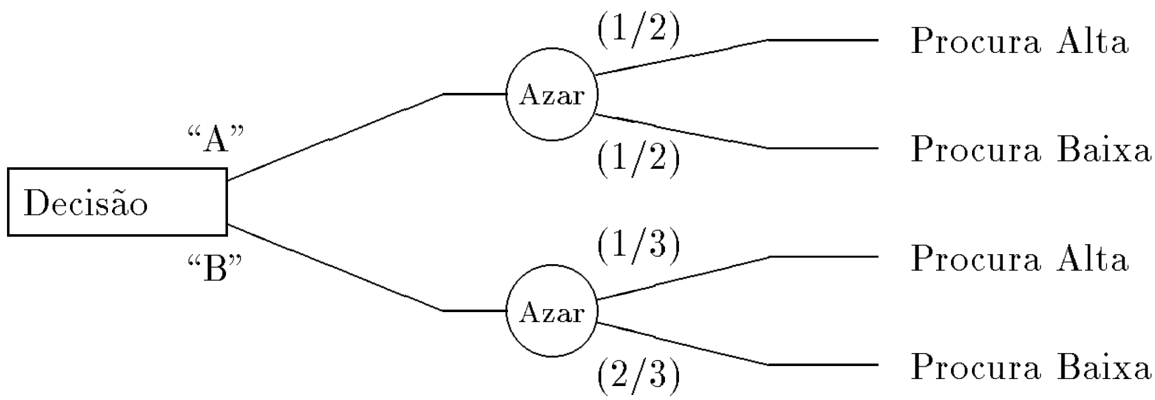

Figura 1: O elemento básico de qualquer decisão estratégica com incerteza. À decisão de um gestor ("A" ou "B") segue-se uma procura incerta. As probabilidades associadas a cada desenlace estão entre parêntesis.

Prob.	Entropia
1/2 e 1/2	0.30
1/3 e 2/3	0.28
1/4 e 3/4	0.24

Prob.	Entropia
1/5 e 4/5	0.22
1/10 e 9/10	0.14
1/20 e 19/20	0.09

Tabela 1: Relação entre entropia e probabilidades num jogo com dois desenlaces.

incerto. Esta simples estrutura, esquematicamente representada na figura 1, pode considerar-se como o elemento básico de qualquer decisão estratégica. Neste caso, o uso da entropia dá ao gestor a possibilidade de comparar, com mais realismo do que se usasse colecções de probabilidades, cada uma das possíveis opções que enfrenta.

A figura 1 mostra uma decisão à qual se segue um entre dois acontecimentos incertos envolvendo dois possíveis desenlaces com probabilidades de ocorrência de p_1 e $p_2 = 1 - p_1$. Ao contrário do que seria intuitivo, a diferença entre uma incerteza dada por

$$p_1 = \frac{1}{2},\ p_2 = \frac{1}{2} \quad \text{e a incerteza gerada por} \quad p_1 = \frac{1}{3},\ p_2 = \frac{2}{3}$$

é negligível e não merece ser tida em consideração, excepto quando a decisão tem que repetir-se muitas vezes. Tal facto é claramente visível quando se usa a entropia de cada um deles como uma estimação da incerteza. A tabela 1 mostra o valor de H para pares de probabilidades associados a um jogo com dois possíveis desenlaces. As diferenças só começam a ser importantes a partir de 1/3. Na figura 1, apesar das probabilidades parecerem indicar uma quebra na incerteza quando a decisão é "B", esta é na realidade mínima.

7 A relação entre probabilidades e entropia — a informação que falta conhecer sobre um desenlace — não é linear. Em redor de valores de $p = 1/N$, grandes variações de p conduzem a pequenas variações da incerteza. E nos extremos, quando p se aproxima de

zero, qualquer pequena variação de p produz uma grande variação em incerteza. Ao gestor interessa muito mais comparar entropias ou ganhos pois, ao contrário das probabilidades, estas medidas são capazes de lhe dizer directamente qual o poder de que ele dispõe sobre cada um dos desenlaces.

1.2 Numeração Binária e Informação

Até aqui não se falou da base dos logaritmos a usar para calcular o ganho ou a entropia quando a informação é incompleta. Em teoria, não importa muito qual a base escolhida (decimal ou neperiana) desde que seja sempre a mesma. Na prática, porém, tem um grande interesse usar a base 2 para este tipo de estimativas já que assim a quantidade de informação resultante virá expressa em *número de bits*.

Como não é fácil encontrar tabelas de logaritmos na base binária, lembramos que

$$\log_2 x = \frac{\log_n x}{\log_n 2} \tag{6}$$

Por exemplo, qual seria a quantidade de informação necessária para transmitir o resultado do lançamento de uma moeda ao ar? Aplicando (5), com $p_{\text{caras}} = 0.5$ e $p_{\text{coroas}} = 0.5$ vem:

$$H = -\sum_{i=1}^{2} p_i \log p_i = -0.5 \log 0.5 - 0.5 \log 0.5$$

Caso fosse usada a base decimal, H valeria 0.301; uma base natural daria $H = 0.693$; e uma base binária daria $H = 1$. Um bit é pois a quantidade de informação que falta para remover totalmente a incerteza que rodeia um jogo de moeda ao ar. Este valor é bastante mais intuitivo do que o mesmo noutras bases.

A variedade, $\log N$, passa a medir o número de bits necessário para identificar uma de entre N possibilidades, desde que se usem logaritmos de base 2. Para identificar um de entre 2 acontecimentos diferentes será preciso um bit, quatro acontecimentos requerem dois bits, oito requerem três e por aí fora.

1.3 Informação Assimétrica e Ganho Esperado

Viu-se no capítulo anterior como a informação tinha valor e podia portanto gerar ou acrescentar os ganhos (retornos) de um investimento. O ganho G em informação que se obtém ao conhecer, com exclusividade, uma tendência ou enviezamento é, como se viu,

$$G = \max H - H = \log N - (-\sum_{i=1}^{N} p_i \log p_i)$$

Probabilidade q	Estratégia ω	Retorno esperado g
50%	0%	0%
60%	20%	2.9%
70%	40%	11.9%
80%	60%	27.8%
90%	80%	53.1%
100%	100%	100%

Tabela 2: Relação entre informação parcial (exclusiva) e retorno esperado num jogo com dois desenlaces. ω é a estratégia que conduz a melhores retornos

Chama-se *ganho relativo* ao ganho percentual relativo à incerteza original ou variedade. Será portanto o quociente

$$g = \frac{G}{\log N} \tag{7}$$

Por sua vez, pode provar-se que este ganho relativo g coincide com o maximo retorno esperado por um investidor que possua informação parcial G não acessível a outros investidores. É esta portanto a relação entre ganho em informação e ganho esperado.

A posse exclusiva de informação é um fenómeno frequente e encontra-se bem estudado pelos economistas. Os gestores de um negócio, por exemplo, têm informação exclusiva não acessível aos próprios donos desse negócio. Esta posse exclusiva dá orígem a problemas de *assimetria informativa* como o referido problema da agência.

8. Estratégia de Maior Ganho Para obter o ganho g, um investidor com informação exclusiva terá que usar uma estratégia específica que a seguir se descreve para o caso simples de $N = 2$ (portanto $\log_2 N = 1$).

Suponha-se que a probabilidade de um dado acontecimento vir a dar-se é q e que esta probabilidade é claramente maior do que 50%. Um investidor que possua o conhecimento exclusivo desta assimetria, deverá, para obter o maximo retorno de tal conhecimento, investir a proporção

$$\omega = 2q - 1 \tag{8}$$

do seu capital jogando contra todos os outros investidores (para quém $q = 50\%$). Caso o faça em sucessivas ocasiões, o seu retorno esperado será, como descrito em (7),

$$g = G = 1 + q\log_2 q + (1-q)\log_2(1-q)$$

Repare-se, (tabela 2), como um ω óptimo se relaciona com os retornos esperados, sugerindo a estratégia simples de investir tanto mais quanto menor é a incerteza.

Quanto maior a incerteza, quanto menor a vantagem do investidor em termos de informação exclusiva, menor deverá ser a proporção do capital total a investir–e vice versa. Esta estratégia constitui uma regra geral da teoria financeira e nunca deve ser esquecida. Iremos encontrar exemplos de tal regra em muitas outras ocasiões. Quém segue esta regra, prospera; quém esquece esta regra acaba na bancarrota com probabilidade 1.

Resumo

Este capítulo estabeleceu a relação entre informação incompleta e probabilidades a-priori. Os aspectos quantitativos da informação só merecem ser tidos em consideração no manejo da informação estratégica e especialmente na tomada de decisões. A informação táctica e operacional é geralmente completa. Porém, é importante lembrar que o ganho em informação representa, no caso de ela ser incompleta, o poder que o gestor tem nas suas mãos para causar um dado desenlace.

O capítulo também mostrou o valor da posse exclusiva de informação e lembrou a importância que podem assumir os problemas de assimetria informativa.