

O valor das previsões — sondagem conjuntural

Eden Gonçalves de Oliveira *

1. Introdução; 2. Exame das previsões em termos informacionais; 3. Previsão — realização; 4. Resultados.

1. Introdução ¹

As previsões e observações dos informantes da sondagem conjuntural junto à indústria brasileira de transformação (realizada pelo Centro de Estudos Industriais do Instituto Brasileiro de Economia — Fundação Getulio Vargas), relacionam-se com características ao nível das firmas individuais ou ao nível da própria indústria de transformação.

As informações fornecidas pelos empresários traduzem níveis de cada variável (forte, normal, fraco) ou tendências de evolução da correspondente variável (em alta, estável, em baixa).

Examina-se, neste estudo, em que medida projeções do tipo *em alta*, *estável* ou *em baixa* têm sido bem sucedidas; far-se-á uma análise das previsões em termos informacionais. ²

* O autor é economista do Centro de Estudos Industriais do IBRE/ FGV.

¹ O autor agradece a colaboração de Jorge Ferreira Braga na realização dos cálculos complementares procedidos.

² Este trabalho se baseia em desenvolvimentos apresentados por Theil, H. *Economics and information theory e Applied economic forecasting*.

2. Exame das previsões em termos informacionais

Seja E um evento, e defina-se o *conteúdo de informação* da mensagem que declara a ocorrência de E .

Estabeleça-se que:

- (i) o conteúdo de informação h , de uma determinada mensagem, depende somente de uma variável (probabilidade x do evento E);
- (ii) h é função contínua de x , $0 < x \leq 1$;
- (iii) $h(0) = +\infty$, $h(1) = 0$.

Traduzem, respectivamente, surpresa infinita pela ocorrência de um evento com probabilidade 0, e nenhuma surpresa pela ocorrência de um evento com probabilidade 1.

- (iv) h é monótona decrescente;
- (v) $h(x_1 x_2) = h(x_1) + h(x_2)$, $0 \leq x_1, x_2 \leq 1$;

Ou seja, existe aditividade no caso de eventos independentes. O uso de logaritmos com base 2 é comum em teoria da informação; tem a propriedade de ter a mensagem que declara a ocorrência de um evento E (com a probabilidade $1/2$ de ocorrer) conteúdo de informação $h(1/2) = 1$.

Diz-se que o conteúdo de informação é expresso em *bits*, se se usa base logarítmica 2, e em *nits*, se a base é neperiana. Essas unidades assim se relacionam: $1 \text{ nit} = 1,443 \text{ bit}$

Seja X_1, X_2, \dots, X_n uma seqüência de projeções de probabilidades da ocorrência do evento E ; dir-se-á, simplesmente, uma seqüência de projeções de probabilidades de E .

Define-se *ganho de informação* de uma projeção como:

$$h(X_i) - H(X_{i+1}) = \log \frac{X_{i+1}}{X_i}, \quad (i = 1, 2, \dots, n-1)$$

Pode-se estabelecer a seguinte generalização para a definição do *conteúdo de informação*:

Conteúdo de informação
de uma mensagem

$$= \frac{\log \text{probabilidade } ex \text{ post}}{\text{probabilidade } ex \text{ ante}}$$

Probabilidade *ex post*

$$= \text{probabilidade do evento após a mensagem ter sido recebida}$$

Probabilidade *ex ante*

= probabilidade do evento antes de mensagem ter sido recebida

Considere-se um conjunto de n eventos E_1, E_2, \dots, E_n , que ocorrem, respectivamente, com probabilidades X_1, X_2, \dots, X_n , sendo certo que exatamente um desses eventos ocorra.

Uma mensagem que declare que E_i ocorreu, tem conteúdo de informação $h(X_i) = \log \frac{1}{X_i} = -\log X_i$.

Antes de a mensagem ter sido recebida desconhece-se a magnitude do conteúdo de informação associado, podendo ser $h(X_1), h(X_2), \dots, h(X_n)$.

Denote-se por $h(X)$ ao conteúdo médio de informação (antes de a mensagem ser recebida):

$$H(X) = \sum_{i=1}^n X_i \cdot h(X_i) = - \sum_{i=1}^n X_i \cdot \log X_i \text{ e ponha-se } X_i \cdot \log X_i = 0 \text{ se } X_i = 0$$

O valor mínimo de H é zero (assumido quando $X_i = 1$ para algum $i, X_j = 0$ para todo $j \neq i$).

Ou seja, nenhuma informação é esperada da emissão de uma mensagem que declara que um dos n eventos ocorreu, quando se sabe que algum evento tem probabilidade 1.

O conteúdo médio de informação é máximo quando os eventos são equi-prováveis. Ou seja, há um máximo de incerteza quando:

$$Pr \{E_i\} = \frac{1}{n}, \quad (i = 1, 2, \dots, n).$$

Pode-se mostrar que $0 \leq H(X) \leq \log n$, em que $\log n$ é o máximo de H .

A informação esperada de uma distribuição é chamada *entropia da distribuição*.

3. Previsão-realização

Objetiva-se agora examinar as projeções dos informantes acerca do estado de seus negócios, tendo em vista as informações sobre realizações por eles fornecidas.

Represente-se por f_{ij} a frequência relativa dos casos em que os informantes tenham previsto tipo i de comportamento para uma variável num período, e tenham observado que o comportamento efetivo da variável foi do tipo j .

Previsão-realização

	Realização		Marginal	
	Em alta	Estável	Em baixa	
Previsão: Em alta	f_{11}	f_{12}	f_{13}	$f_{1.}$
Estável	f_{21}	f_{22}	f_{23}	$f_{2.}$
Em baixa	f_{31}	f_{32}	f_{33}	$f_{3.}$
Marginal	$f_{.1}$	$f_{.2}$	$f_{.3}$	1

Quadro 1

Previsão-realização

Mão-de-obra — período: Jan./mar. — 1973

	Realização		Marginal	
	Em alta	Estável	Em baixa	
Previsão: Em alta	0,262	0,105	0,008	0,375
Estável	0,139	0,437	0,016	0,592
Em baixa	0,003	0,018	0,012	0,033
Marginal	0,404	0,560	0,036	1,000

O quadro 1 mostra que 37,5% dos informantes previram nível de absorção de mão-de-obra *em alta* no período janeiro/março, e que, dessas projeções, 26,2% foram corretas.

Esse quadro mostra ainda que antes de a projeção ter sido procedida, o tipo de comportamento *em alta* tinha probabilidade $f_{.1} = 0,404$. Considerando a projeção, a realização, *em alta* tem probabilidade $\frac{f_{11}}{f_{1.}} = 0,698$.

O ganho de informação na projeção *nível de absorção de mão-de-obra em alta* é:

$$h(f_{.1}) - h\left(\frac{f_{11}}{f_{1.}}\right) = \log \frac{f_{11}}{f_{1.} \cdot f_{.1}} = 0,548$$

Assim, o ganho de informação de uma projeção de aumento, estabilidade ou declínio é, respectivamente, dado por:

$$h(f_{.i}) - h\left(\frac{f_{ii}}{f_{i.}}\right) = \log \frac{f_{ii}}{f_{i.} \cdot f_{.i}}, \quad (i = 1, 2, 3)$$

Define-se *ganho relativo de informação* como a seguinte fração do valor máximo do ganho de informação $h(f_{.i})$:

$$\frac{h(f_{.i}) - h\left(\frac{f_{ii}}{f_{i.}}\right)}{h(f_{.i})} = \frac{\log\left(\frac{f_{ii}}{f_{i.} \cdot f_{.i}}\right)}{\log\left(\frac{1}{f_{i.}}\right)}, \quad (i = 1, 2, 3)$$

O valor máximo do ganho de informação $H(f_{.i})$ ocorre para $f_{ii}=f_{i.}$ (o que requer que todas as projeções tipo i tenham sido, necessariamente, corretas).

O conteúdo médio de informação de uma projeção tipo i ($i = 1, 2, 3$) é dado por:

$$\frac{1}{f_{j.}} \sum_{j=1}^3 f_{ij} \cdot \log\left(\frac{f_{ij}}{f_{i.} \cdot f_{.j}}\right)$$

O conteúdo médio de informação do conjunto dos três tipos de projeções é definido por:

$$\sum_{i=1}^3 \sum_{j=1}^3 f_{ij} \cdot \log\left(\frac{f_{ij}}{f_{i.} \cdot f_{.j}}\right)$$

4. Resultados

Os quadros 2 e 3 apresentam elementos destinados à avaliação do mérito das projeções (em termos informacionais) .

Os ganhos absolutos e relativos de informação são apresentados para cada uma das cinco variáveis selecionadas e para cada um dos três tipos de previsão (quadro 2).

Quadro 2
Ganhos de informações
1973

Variável	Ganho absoluto (N/I.)			Ganho relativo		
	Em alta	Estável	Em baixa	Em alta	Estável	Em baixa
Jan-Maí						
Mão-de-obra	0,548	0,276	2,316	0,604	0,476	0,696
Produção	0,288	0,495	1,606	0,498	0,446	0,724
Demanda interna	0,168	0,500	1,800	0,453	0,367	0,616
Demanda externa	0,442	0,350	1,449	0,519	0,438	0,691
Preços produtos industriais	0,240	0,362	3,597	0,475	0,370	0,675
Abr-Jun						
Mão-de-obra	0,528	0,258	1,968	0,556	0,443	0,674
Produção	0,316	0,502	1,431	0,519	0,443	0,730
Demanda interna	0,226	0,403	1,523	0,450	0,339	0,629
Demanda externa	0,313	0,306	1,389	0,371	0,377	0,670
Preços produtos industriais	0,332	0,270	2,959	0,483	0,371	0,668
Jul-Set						
Mão-de-obra	0,645	0,200	1,737	0,536	0,426	0,663
Produção	0,484	0,359	0,894	0,526	0,351	0,628
Demanda interna	0,415	0,323	0,929	0,495	0,310	0,601
Demanda externa	0,353	0,295	1,166	0,381	0,357	0,649
Preços produtos industriais	0,306	0,325	2,567	0,445	0,446	0,601

Os ganhos de informação das projeções de *baixa* são ponderavelmente maiores que aqueles das projeções de *estabilidade* e de *alta*. Como as fre-

quências f_3 (de *baixa*) são relativamente pequenas, sobram muitos *nits* para serem ganhos.

As três últimas colunas do quadro 2 mostram os ganhos de informação como uma fração de seu máximo. Os ganhos de informação das projeções de *baixa* variam entre 60% e 73% do máximo atingível, mostrando-se portanto acima dos ganhos relativos das projeções de *alta* e de *estabilidade*.

O conteúdo de informação de uma previsão (suponha-se de aumento) é a informação esperada dessa previsão quando ela é considerada como uma mensagem indireta. O quadro 3 mostra que o conteúdo de informação das previsões de *baixa* é bem maior que aquele das previsões de *estabilidade* e *alta*.

Quadro 3
Conteúdo médio de informação (NIT)
1973

Variável	Por tipo de projeção			Global
	Em alta	Estável	Em baixa	
Jan-Mar				
Mão-de-obra	0,179	0,069	0,697	0,131
Produção	0,147	0,097	0,629	0,162
Demanda interna	0,042	0,069	0,348	0,066
Demanda externa	0,113	0,070	0,466	0,128
Preços produtos industriais	0,060	0,060	0,455	0,064
Abr-Jun				
Mão-de-obra	0,145	0,058	0,509	0,115
Produção	0,087	0,089	0,567	0,131
Demanda interna	0,053	0,071	0,382	0,085
Demanda externa	0,052	0,050	0,371	0,082
Preços produtos industriais	0,082	0,049	0,538	0,069
Jul-Set				
Mão-de-obra	0,160	0,044	0,455	0,118
Produção	0,129	0,051	0,292	0,124
Demanda interna	0,104	0,040	0,258	0,108
Demanda externa	0,059	0,046	0,358	0,082
Preços produtos industriais	0,068	0,074	0,455	0,074

Note-se a diferença conceitual entre ganhos de informação e conteúdo de informação de uma projeção. O ganho de informação $\log \left(\frac{f_{ii}}{f_{i.} f_{.i}} \right)$ confina-se à frequência de projeções corretas f_{ii} , enquanto que o conteúdo de informação considera também as projeções incorretas.

Quando as previsões e realizações são estocasticamente independentes, o conteúdo de informação da previsão é zero — este é o mínimo, pois o conteúdo de informação não pode ser negativo.

O conteúdo médio de informação (global) é o conteúdo de informação esperada dos três tipos de previsões — é também uma medida da incerteza.

Os dados mostram que o conteúdo médio de informação (global) tem sido maior nas previsões de evolução dos contingentes de mão-de-obra e dos volumes de produção das empresas, do que naquelas da evolução da demanda interna, da demanda externa e dos preços de seus produtos.

A Escola de Pós-Graduação em Economia
(EPGE)

da Fundação Getulio Vargas
realizará

entre os dias 21 de abril e 8 de maio
um curso intensivo e prático
sobre

**A CORREÇÃO MONETÁRIA NO SISTEMA
ECONÔMICO-FINANCEIRO NACIONAL**

Inscrições na **EPGE** da Fundação Getulio Vargas
Praia de Botafogo, 190,
10.º andar, sala 1011
Tel.: 246-1787 (direto)