NOTAS

A DECOMPOSIÇÃO DO ÍNDICE DE HIRSCHMAN-HERFINDAHL

Rodolfo Hoffmann

RESUMO

Inicialmente é deduzida a relação entre o coeficiente de variação e o índice de Hirschman-Herfindahl relativo à concentração da distribuição das empresas de uma indústria. A seguir, admitindo que as empresas de uma indústria estão divididas em k grupos, mostra-se como o índice de Hirschman-Herfindahl para a concentração nessa indústria depende da concentração entre grupos e dos níveis de concentração dentro de cada grupo. Os índices de Hirschman-Herfindahl relativos à concentração entre grupos e dentro de grupos são definidos de maneira que a decomposição do índice de concentração da indústria seja coerente com a correspondente decomposição da variância.

SUMMARY: DECOMPOSITION OF THE HIRSCHMAN-HERFINDAHL INDEX

Firstly, a relation between the coefficient of variation and the Hirschman-Herfindahl index of concentration of an industry is deduced. Then, assuming that the firms in the industry are distributed in k groups, it is shown that the Hirschman-Herfindahl index is a function of the concentration among groups and the concentration within groups. The Hirschman-Herfindahl indexes for concentration among groups and within groups are defined so that the decomposition of the measure of concentration of the whole industry is coherent with the corresponding decomposition of the population variance.

1. INTRODUÇÃO

O objetivo principal deste trabalho é analisar a decomposição do índice de concentração de Hirschman-Herfindahl.

Uma vez que esse índice é freqüentemente utilizado no estudo da concentração industrial, vamos admitir que a variável em análise (x) seja o valor da produção (ou valor adicionado) das empresas de uma indústria.

Vamos supor que o universo das empresas da indústria considerada esteja dividido em k grupos (conforme sua localização regional, por exemplo). Seja n_h (com $h = 1, \dots, k$) o número de empresas no h -ésimo grupo e seja x_{hi} (com $h = 1, \dots, k$ e $i = 1, \dots, n_h$) o valor da produção da i -ésima empresa do h -ésimo grupo. O número total de empresas da indústria é

$$N = \sum_{h=1}^k n_h$$

Se o valor médio da produção por empresa é μ , a fração do valor total da produção da indústria que corresponde à i -ésima empresa do h -ésimo grupo é

$$y_{hi} = \frac{x_{hi}}{N\mu} \quad (1)$$

A proporção de empresas que pertence ao h -ésimo grupo é

$$\pi_h = \frac{n_h}{N} \quad (2)$$

(*) Professor da ESALQ/USP. O autor agradece ao Prof. Geraldo Sant'Ana de Camargo Barros pelos comentários a uma versão preliminar deste trabalho.

e a correspondente proporção do valor total da produção da indústria é

$$y_h = \frac{n_h}{\sum_{i=1}^{n_h} y_{hi}} \quad (3)$$

O valor médio da produção por empresa no h -ésimo grupo é

$$\mu_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi} = \frac{y_h}{\pi_h} \mu \quad (4)$$

Por definição, o índice de Hirschman-Herfindahl para a indústria é

$$H = \sum_{h=1}^k \sum_{i=1}^{n_h} y_{hi}^2 \quad (5)$$

O valor mínimo de H é $1/N$ e ocorre quando todos os valores de x são iguais a μ . O valor máximo de H , com $x_{hi} > 0$, é 1 e ocorre quando um dos valores de x é igual a $N\mu$ e todos os outros são iguais a $\underline{ze}$ ro (ver Barbosa, 1981, pp. 42-43).

2. CONCENTRAÇÃO E DESIGUALDADE

Se σ^2 e μ são, respectivamente, a variância e a média de uma distribuição, o coeficiente de variação é $C = \sigma/\mu$. Antes de analisar a decomposição do índice de Hirschman-Herfindahl, vamos deduzir a relação entre este índice e o coeficiente de variação. Dessa maneira poderemos relacionar a decomposição do índice de Hirschman-Herfindahl com a decomposição da variância.

Tendo em vista a notação introduzida na seção anterior, a variância de x na indústria considerada é

$$\sigma^2 = \frac{1}{N} \sum_{h=1}^k \sum_{i=1}^{n_h} (x_{hi} - \mu)^2$$

ou

$$\sigma^2 = \frac{1}{N} \sum \sum x_{hi}^2 - \mu^2$$

Então, lembrando (1), obtemos

$$C^2 = \frac{\sigma^2}{\mu^2} = N \sum \sum y_{hi}^2 - 1 \quad (6)$$

Comparando (5) e (6) conclui-se que

$$C^2 = \frac{\sigma^2}{\mu^2} = NH - 1,$$

$$H = \frac{1}{N} (C^2 + 1) \quad (7)$$

ou

$$\sigma^2 = \mu^2 (NH - 1) \quad (8)$$

A expressão (7) relaciona uma medida de *concentração* (H) e uma medida de *desigualdade* (C). Um alto grau de desigualdade significa que uma pequena *porcentagem* das empresas é responsável por grande parte da produção da indústria. Por outro lado, um alto grau de concentração significa que grande parte da produção da indústria se origina de um pequeno *número* de empresas. A distinção entre esses conceitos é analisada em recente trabalho de Braga e Mascolo (1981), os quais concluem que, "dado o grau de desigualdade, a concentração decresce com o aumento do número de firmas"⁽¹⁾. Para o caso particular da relação (7), verifica-se que, dado o valor do coeficiente de variação, o índice de Hirschman-Herfindahl (H) é *inversamente* proporcional ao número (N) de empresas.⁽²⁾

(1) A distinção entre desigualdade e concentração é, sem dúvida, essencial no estudo da concentração industrial. Entretanto, em estudos sobre distribuição de renda, onde o conceito específico de concentração não se aplica, usa-se o termo concentração como sinônimo de desigualdade.

(2) Há uma relação semelhante entre o índice de Gini (G, uma medida de desigualdade) e o índice de Rosenbluth (R, uma medida de concentração), que é descrita no apêndice.

3. DECOMPOSIÇÃO DO ÍNDICE DE HIRSCHMAN-HERFINDAHL

De acordo com (1) e (5) temos

$$H = \sum_h \sum_i \left(\frac{x_{hi}}{N\mu} \right)^2$$

Multiplicando e dividindo por $n_h^2 \mu_h^2$ obtemos

$$H = \sum_h \left(\frac{n_h \mu_h}{N\mu} \right)^2 \sum_i \left(\frac{x_{hi}}{n_h \mu_h} \right)^2$$

ou

$$H = \sum_h Y_h^2 \sum_i \left(\frac{x_{hi}}{n_h \mu_h} \right)^2 \quad (9)$$

Uma vez que $x_{hi}/(n_h \mu_h)$ representa a proporção da produção total do h -ésimo grupo que se origina da i -ésima empresa desse grupo,

$$H_h = \sum_i \left(\frac{x_{hi}}{n_h \mu_h} \right)^2 \quad (10)$$

é o índice de Hirschman-Herfindahl para a concentração dentro desse grupo. Substituindo (10) em (9) obtemos

$$H = \sum_h Y_h^2 H_h \quad (11)$$

Se não houver nenhuma desigualdade ou concentração dentro dos grupos, temos $x_{hi} = \mu_h$ e, conseqüentemente, $H_h = 1/n_h$ para todo h . Neste caso o índice de Hirschman-Herfindahl se reduz a

$$H_e = \sum_h Y_h^2 \frac{1}{n_h} = \frac{1}{N^2 \mu^2} \sum_h n_h \mu_h^2 \quad (12)$$

que corresponde à concentração *entre* grupos. Se todos os grupos t_i verem a mesma produção média por empresa, isto é, se $\mu_h = \mu$ para todo h , então $H_e = 1/N$, que é o valor mínimo de H .

De acordo com (11) e (12) temos

$$H = H_e + \sum_h Y_h^2 (H_h - \frac{1}{n_h}), \quad (13)$$

mostrando como, em termos do índice de Hirschman-Herfindahl, a concentração na indústria depende da concentração entre grupos (H_e) e dos níveis de concentração dentro de cada grupo (H_h).

Barbosa (1981) define o índice de Hirschman-Herfindahl relativo à concentração entre grupos como

$$H_o = \sum_h Y_h^2,$$

obtendo

$$H = H_o \sum_h w_h H_h, \quad (14)$$

onde $w_h = Y_h^2 / H_o$. O autor assinala que a decomposição dada por (14) é do tipo multiplicativo. "O índice total, H , é igual ao produto do índice inter-grupos, H_o , por uma média ponderada dos índices intra-grupos, $\sum_h w_h H_h$, com a soma dos pesos, w_h , igual à unidade". Essa decomposição, aparentemente mais simples do que aquela dada por (13), apresenta sérios inconvenientes:

1º) se não houver desigualdade dentro dos grupos temos $H_h = 1/n_h$ e o valor de H não se reduz a H_o ;

2º) se $\mu_h = \mu$ para todo h , temos $H_o = (1/N^2) \sum_h n_h^2$ e H não será uma média ponderada dos valores de H_h ;

3º) não há correspondência com a decomposição da variância.

Mostraremos, a seguir, que a decomposição do índice de Hirschman-Herfindahl dada por (13) é coerente com a decomposição da variância.

É bastante conhecido o fato de que, quando uma população é dividida em k grupos, a variância total (σ^2) pode ser decomposta em uma variância entre grupos (σ_e^2) e uma média ponderada das variâncias dentro dos grupos (σ_h^2 , com $h = 1, \dots, k$), ou seja,

$$\sigma^2 = \sigma_e^2 + \sum_h \pi_h \sigma_h^2, \quad (15)$$

onde

$$\sigma_e^2 = \frac{1}{N} \sum_h n_h (\mu_h - \mu)^2 \quad (16)$$

e

$$\sigma_h^2 = \frac{1}{n_h} \sum_i (x_{hi} - \mu_h)^2 \quad (17)$$

Pode-se verificar que existem relações equivalentes a (8) para a variância entre grupos e para as variâncias dentro dos grupos. A partir de (12) e (16) deduz-se que

$$\sigma_e^2 = \mu^2 (N H_e - 1) \quad (18)$$

Analogamente, a partir de (10) e (17) verifica-se que

$$\sigma_h^2 = \mu_h^2 (n_h H_h - 1) \quad (19)$$

Substituindo (8), (18) e (19) em (15) obtemos

$$\mu^2 (NH - 1) = \mu^2 (N H_e - 1) + \sum_h \pi_h \mu_h^2 (n_h H_h - 1)$$

Após algumas operações algébricas essa relação se transforma na expressão (13), mostrando como a decomposição do índice de Hirschman Herfindahl aí indicada corresponde à decomposição da variância dada em (15).

APÊNDICE

RELAÇÃO ENTRE O ÍNDICE DE GINI E O ÍNDICE DE ROSENBLUTH

Consideremos uma indústria com n empresas e seja x_j o valor da produção da j -ésima empresa. Vamos admitir que $x_1 \geq x_2 \geq \dots \geq x_n$. Por definição, o índice de Rosenbluth, redescoberto por Hall e Tideman (1967), é

$$R = (2 \sum_j y_j - 1)^{-1},$$

onde $y_j = x_j / (n \mu)$. Se admitirmos que $x_1 \leq x_2 \leq \dots \leq x_n$, o índice de x passa a ser $i = n - j + 1$ e a expressão de definição do índice de Rosenbluth fica

$$R = (2n + 1 - 2 \sum_i y_i)^{-1}$$

De acordo com a expressão (16.10), p.277, em Hoffmann (1980), temos

$$nG = 2 \sum_i y_i - 1 - n,$$

onde G é o índice de Gini da distribuição.

A partir dessas duas últimas expressões obtemos

$$R = \left[n (1 - G) \right]^{-1}$$

Essa expressão mostra que, dado G , o índice de Rosenbluth é inversamente proporcional a n . A existência de uma relação entre G e R é assinalada por Hart (1971).

BIBLIOGRAFIA

- BARBOSA, Fernando de Holanda (1981). Medidas de Concentração. *Revista de Econometria* 1(1): 31-53.
- BRAGA, Helson C. e João L. MASCOLO (1981). Mensuração da Concentração Industrial no Brasil. Trabalho apresentado no III Encontro Brasileiro de Econometria. Olinda, PE, 8 a 10 de dezembro de 1981.
- HALL, Marshall e Nicolaus TIDEMAN (1967). Measures of Concentration. *Journal of the American Statistical Association* 162-168.
- HART, P.E. (1971). Entropy and Other Measures of Concentration. *Journal of the Royal Statistical Society, Series A (General)* 134(1): 73-85.
- HOFFMANN, Rodolfo (1980). *Estatística* São Paulo.