Arq. Fac. Hig. (São Paulo)
17 *(2): 217-221, dez. 1963*

SÔBRE A DISTRIBUIÇÃO DO QUOCIENTE DE DOIS COEFICIENTES ♀

RUBENS MURILLO MARQUES*

1. INTRODUÇÃO

Um dos problemas freqüentes no campo da Saúde Pública consiste em indagar se a queda ou o aumento percentual sofrido por um coeficiente numa localidade em duas épocas distintas se comporta de maneira análoga à que se verifica numa região maior da qual a localidade em estudo é parte integrante.

Assim, por exemplo, poderíamos estar interessados em saber se, no Município de São Paulo, a mortalidade infantil caiu, no período de tempo de 1955 a 1961 de m%, onde m caracteriza a queda sofrida pela mortalidade infantil em todo o Estado de São Paulo, no período de tempo considerado.

O presente trabalho visa apresentar o tratamento estatístico do aludido problema.

2. O PROBLEMA TRATADO DO PONTO DE VISTA ESTATÍSTICO

Sejam C_1 e C_2, respectivamente, os verdadeiros coeficientes de mortalidade infantil nos instantes t_1 e t_2 para uma região R.

Seja $m = \frac{C_1 - C_2}{C_1} . 100$, a queda percentual sofrida em R pela mortalidade infantil, no período considerado.

Se chamarmos, para uma parte de R, de c_1 e c_2 os coeficientes de mortalidade infantil nos instantes t_1 e t_2, respectivamente, então esta é a pergunta que nos propomos a responder: é esta queda percentual, $\frac{c_1 - c_2}{c_1} . 100$, observada nesta parte de R, significantemente diferente de m%? Ou, equivalentemente, é a relação $\frac{c_2}{c_1} . 100$ significantemente diferente de $(1 - m)\%$?

À primeira vista, poder-se-ia pensar que a variável $\frac{c_2}{c_1} . 100$ tem dis-

Recebido para publicação em 6-5-1963.
♀ Trabalho da Cadeira de Bioestatística (Profa. Elza Berquó) da Faculdade de Higiene e Saúde Pública da Universidade de São Paulo.
* Instrutor da Cadeira.

tribuição normal e resolver o problema dessa forma. Entretanto, isto não é verdade. Conquanto o tratamento estatístico do problema, quando certas condições que apresentaremos em 3, estiverem satisfeitas, seja feito utilizando a curva normal, Geary[1] demonstrou que isto é conseguido não pelo emprêgo da variável $\dfrac{c_2}{c_1} \cdot 100$ mas sim pela variável :

$$y = \frac{\dfrac{c_2}{c_1} - \dfrac{C_2}{C_1}}{\dfrac{1}{C_1} \sqrt{\sigma^2_{c_1} \cdot \dfrac{c_2}{c_1} - 2 \cdot \rho \cdot \sigma_{c_1} \cdot \sigma_{c_2} \cdot \dfrac{c_2}{c_1} + \sigma^2_{c_2}}} \qquad (I)$$

que tem distribuição aproximadamente normal reduzida. Em (I), $\sigma^2_{c_1}$ e $\sigma^2_{c_2}$ são, respectivamente, as variâncias a que estão sujeitas as variáveis aleatórias c_1 e c_2 e ρ é o coeficiente de correlação existente entre elas.

Nestas condições, a pergunta feita seria respondida substituindo-se em (I) $\dfrac{C_2}{C_1}$ por $(1 - m)$ e $\dfrac{c_2}{c_1}$ pelo valor observado do quociente dos coeficientes de mortalidade infantil nos instantes 1955 e 1961 para o Município de São Paulo, supondo que $\sigma^2_{c_1}$, $\sigma^2_{c_2}$ e ρ são conhecidos. O valor obtido para y seria comparado com o valor crítico, para o nível de significância fixado, dado pela tabela da curva normal.

No caso mais comum na prática, os parâmetros $\sigma^2_{c_1}$ $\sigma^2_{c_2}$ e ρ são desconhecidos, porém y ainda poderá ser tratada como uma variável normal reduzida quando se substituir em (I) os referidos parâmetros por suas respectivas estimativas.

3. CONDIÇÕES DE APLICABILIDADE DO MÉTODO PROPOSTO

Além das pressuposições enumeradas a seguir, o método descrito, devido a Geary, supõe que nos instantes t_1 e t_2 as populações na parte considerada de R nada mais são do que amostras casuais de R tomadas nas respectivas datas. São as seguintes as demais pressuposições :

a) — as variáveis c_1 e c_2 são normalmente distribuidas com médias μ_{c_1}, μ_{c_2} e variâncias $\sigma^2_{c_1}$ e $\sigma^2_{c_2}$, respectivamente.

b) — o coeficiente de variação $\dfrac{\sigma_{c_1}}{\mu_{c_1}}$ da variável c_1 deve ser menor que 1/3.

No problema proposto as variáveis c_1 e c_2 (coeficientes de mortalidade infantil) não têm, a rigor, distribuição normal mas sim binomial. Sabemos todavia, ser lícita a aproximação da distribuição binomial simétrica pela normal. Para os nossos propósitos a simetria será admitida se o número de nascidos vivos nos instantes t_1 e t_2 na parte de R, quando sob a mortalidade infantil de R, fôr maior que $\frac{5}{C_1}$ e $\frac{5}{C_2}$, respectivamente. É fácil ver que esta condição neste tipo de problema é amplamente satisfeita.

A condição b, implica em :

$$\frac{\mu_{c_1}}{\sigma_{c_1}} < \frac{1}{3}$$

ou seja

$$\frac{\sqrt{\dfrac{C_1\,(1 - C_1)}{n_1}}}{C_1} < \frac{1}{3} \qquad \text{(II)}$$

onde n_1 é o tamanho da amostra da parte de R no instante t_1 ou, no caso presente, o número de nascidos vivos no Município de São Paulo em 1955. De (II) resulta que n_1 deve satisfazer à condição :

$$n_1 > \frac{9.(1 - C_1)}{C_1} \qquad \text{(III)}$$

também fàcilmente satisfeita em problemas dessa natureza.

4. APLICAÇÃO

No Estado de São Paulo (R) a mortalidade infantil em 1955[2] e 1961[3] foi, respectivamente, $C_1 = 104{,}09‰$ e $C_2 = 76{,}83‰$ mostrando, portanto, uma queda percentual nestes sete anos de $\frac{C_1 - C_2}{C_1} . 100 = m = 26{,}2\%$.

Com o intuito de saber se o Município de São Paulo se comportou quanto à queda, no mesmo período, de forma análoga a todo o Estado de São Paulo, formulamos as seguintes hipóteses :

H_0 : No Município de São Paulo a queda percentual da mortalidade infantil no período de 1955 a 1961 foi *igual a* 26,2%.

H_1 : No Município de São Paulo a queda percentual da mortalidade infantil no período de 1955 a 1961 foi *diferente de* 26,2%.

O teste de H_0 contra H_1 será feito através da variável y dada pela (I) após a verificação das condições de aplicabilidade do método proposto.

Condição a :

Conforme o estabelecido em 3., as distribuições das variáveis aleatórias c_1 (coeficiente de mortalidade infantil observável no Município de São Paulo em 1955) e c_2 (coeficiente de mortalidade infantil observável no Município de São Paulo em 1961) poderão ser consideradas simétricas. De fato, para tanto, o número de nascidos vivos em 1955 no Município de São Paulo (n_1) deveria ser maior ou igual a $\frac{5}{C_1}$, ou seja, aproximadamente, 50 e em 1961 (n_2) maior ou igual a $\frac{5}{C_2}$ ou seja, aproximadamente 65, condições òbviamente satisfeitas, desde que os valores observados para n_1 e n_2 foram, respectivamente, iguais a 93.432 e 123.933 nascidos vivos.

Condição b :

Devemos ter :

$$n_1 > \frac{9.(1 - C_1)}{C_1} = \frac{9.(0{,}89591)}{0{,}10409} = 77{,}46$$

que também está satisfeita.

Verificadas as condições de aplicabilidade do método, resta ainda o problema, já anteriormente apontado em 2., de estimar ρ pois no nosso caso $\sigma^2_{c_1}$ e $\sigma^2_{c_2}$ são conhecidos e, respectivamente, dados por :

$$\sigma^2_{c_1} = \frac{C_1(1 - C_1)}{n_1} = 0{,}0998/100.000$$

$$\sigma^2_{c_2} = \frac{C_2(1 - C_2)}{n_2} = 0{,}0572/100.000$$

Estimativa da correlação entre a mortalidade infantil no Município de São Paulo em 1955 e em 1961.

Esta estimativa foi feita através da correlação observada entre a mortalidade infantil em 1955 e 1961 nos 62 municípios do Estado de São Paulo que tinham em 1955 população igual ou superior a 30.000 habitantes. O valor encontrado para tal correlação foi de 0,59.

Nestas condições, desde que $\frac{c_2}{c_1} = \frac{60{,}21\%}{86{,}84\%} = 0{,}693$, temos para y:

MARQUES, R. M. . Sôbre a distribuição do quociente de dois coeficientes. Arq. Fac. Hig. (São Paulo) **17**: 217-221, 1963.

$$y = \frac{0{,}693 - 0{,}738}{\frac{1}{0{,}10409} \cdot \sqrt{\frac{0{,}998}{1.000.000} \cdot 0{,}693 - 2.\ 0{,}59 \sqrt{\frac{0{,}998}{1.000}} \sqrt{\frac{0{,}572}{1.000}}\ 0{,}693 + \frac{0{,}572}{1.000.000}}} = -\ 5{,}82$$

o qual comparado com o valor —1,95 nos leva a aceitar H_1, ou seja que ao nível de significância de 5%, existe uma diferença significante entre o comportamento do Município de São Paulo e todo o Estado de São Paulo quanto à queda percentual da mortalidade infantil no período de 1955 a 1961.

RESUMO

O presente trabalho refere-se à distribuição do quociente de duas variáveis aleatórias bem como à análise das condições de aplicabilidade ao caso em que estas variáveis são coeficientes.

O método é ilustrado estudando se a queda percentual sofrida pela mortalidade infantil no Município de São Paulo, no período de 1955 a 1961, comportou-se como a observada no mesmo período em todo o Estado.

SUMMARY

The present paper considers the distribution of the quocient of two random variables as well as the possibility of applying the method to the particular case where the random variables are rates. This method is illustrated by studying the percentual decrease in the infantile mortality rate in the county of São Paulo from 1955 to 1961 as compared to the one observed in the whole State during the same period of time.

REFERÊNCIAS

1. GEARY, R. C. The frequency distribution of the quocient of two normal variates. *J. royal statist. Soc.* **93**: 442-446, 1930.
2. SÃO PAULO. DEPARTAMENTO DE ESTATÍSTICA DO ESTADO. Resumo anual. São Paulo, 1955. p. 18-25.
3. ———— Anuário 1961. São Paulo, 1963. p. 31-85.