Algumas propriedades da distribuição binomial negativa

por **António Dorival Campos** (*) e **Euclydes Custódio de Lima Filho**
Departamento de Matemática Aplicada à Biologia
Faculdade de Medicina de Ribeirão Preto Universidade de São Paulo — Brasil

1. Consideremos uma população infinita π com uma proporção P desconhecida de elementos com uma característica específica A.

Amostremos π através de uma sequência de repetições independentes, com probabilidade constante, de sucessos e denotemos por X o número total de fracassos nesta sequência antes do $r^{\text{ésimo}}$ sucesso, isto é, $X+r$ é o número de ensaios necessários para produzir na amostra $r>1$ elementos com a característica A.

Temos:

$$(1.1)\qquad g(x;r,P)=\binom{x+r-1}{x}P^r(1-P)^x \quad x=0,1,2,\cdots.$$

Usando-se $\hat{P}=\dfrac{r-1}{X+r-1}$ como estimador de P, o qual foi proposto por Haldane [2], demonstraremos que:

i) X é uma estatística suficiente para P;

ii) $\{g(x;r,P),\ 0<P<1\}$ é completa;

iii) $E[\hat{P}]=P$.

Demonstração:

i) Imediata, pois, $g(x;r,P)$ satisfaz o critério de fatoração.

ii) Seja $u(x)$ uma função contínua, independente de P, tal que:

$$(1.2)\qquad E[u(X)]=0$$

para todo P, $0<P<1$.

De (1.2), segue que:

$$(1.3)\qquad E[u(X)]=\sum_{x=0}^{\infty}u(x)\cdot\binom{x+r-1}{x}P^r(1-P)^x=0$$

para todo P, $0<P<1$.

Desenvolvendo-se (1.3) e tomando-se $1-P=Q$, temos:

$$E[u(X)]=(1-Q)^r\left[u(0)+r\,u(1)\,Q+\frac{r(r+1)}{2}u(2)\,Q^2+\cdots\right]=0.$$

Sendo $1-Q\neq 0$ para todo $0<Q<1$, segue que a série em Q entre colchetes converge a 0 para todo $0<Q<1$, o que implica $u(x)=0$ para todo $x=0,1,\cdots$.

iii)

$$E[\hat{P}]=\sum_{x=0}^{\infty}\frac{r-1}{x+r-1}\binom{x+r-1}{x}P^r(1-P)^x=P\sum_{x=0}^{\infty}\binom{x+r-2}{x}P^{r-1}(1-P)^x.$$

Donde:

$$E[\hat{P}]=P.$$

(*) Bolsista do I. A. C. no Centro de Matemáticas Aplicadas, Lisboa.

Satisfeitas as três condições acima, concluímos que o estimador $\hat{P}$ assim definido, tem variância mínima entre todos os estimadores não viciados de P.

2. Determinação da variância de $\hat{P}$.

$$(2.1) \qquad E[\hat{P}^2] =$$

$$= \sum_{x=0}^{\infty} \left(\frac{r-1}{x+r-1}\right)^2 \binom{x+r-1}{x} P^r (1-P)^x =$$

$$= (r-1)^2 P^r \sum_{x=0}^{\infty} \frac{1}{(x+r-1)^2} \binom{x+r-1}{x} (1-P)^x$$

Usando-se o facto que se a série

$$\psi(x) = \sum_{k=0}^{\infty} b_k x^k$$

é diferenciada m vezes dentro do seu intervalo de convergência, então,

$$\frac{\psi^{(m)}(x)}{m!} = \sum_{k=0}^{\infty} b_{k+m} \binom{k+m}{k} x^k,$$

de (2.1), temos:

$$(2.2) \qquad E[\hat{P}^2] = (r-1)^2 P^r \frac{1}{(r-1)!} \cdot$$

$$\cdot \frac{d^{r-1}}{d\,Q^{r-1}} \left[\sum_{x=1}^{\infty} \frac{1}{x^2} Q^x\right].$$

Por outro lado, temos:

$$\frac{d}{d\,Q}\left[\sum_{x=1}^{\infty} \frac{1}{x^2} Q^x\right] = \sum_{x=1}^{\infty} \frac{1}{x} Q^{x-1} =$$

$$= \frac{1}{Q}[-\log(1-Q)].$$

Donde:

$$(2.3) \qquad \frac{d^{r-1}}{d\,Q^{r-1}}\left[\sum_{x=1}^{\infty} \frac{1}{x^2} Q^x\right] =$$

$$= -\frac{d^{r-2}}{d\,Q^{r-2}}\left[\frac{\log(1-Q)}{Q}\right].$$

Substituindo-se (2.3) em (2.2), segue que:

$$(2.4) \qquad E[\hat{P}^2] = -\frac{(r-1)P^r}{(r-2)!} \cdot$$

$$\cdot \frac{d^{r-2}}{d\,Q^{r-2}}\left[\frac{\log(1-Q)}{Q}\right].$$

Sabemos que:

$$\frac{d^{r-2}}{d\,Q^{r-2}}\left[\frac{\log(1-Q)}{Q}\right] =$$

$$\sum_{i=0}^{r-2} \binom{r-2}{i} \left[\frac{d^i}{d\,Q^i} \log(1-Q)\right]\left[\frac{d^{r-2-i}}{d\,Q^{r-2-i}}\left(\frac{1}{Q}\right)\right]$$

Donde:

$$(2.5) \qquad \frac{d^{r-2}}{d\,Q^{r-2}}\left[\frac{\log(1-Q)}{Q}\right] =$$

$$= \frac{(-1)^{r-2}(r-2)!\log(1-Q)}{Q^{r-1}} +$$

$$+ (r-2)! \sum_{i=1}^{r-2} \frac{(-1)^{r-1-i}}{i(1-Q)^i Q^{r-1-i}}.$$

Substituindo-se (2.5) em (2.4), obtemos:

$$E[\hat{P}^2] = \frac{(-1)^{r-1}(r-1)P^r \log(1-Q)}{Q^{r-1}} +$$

$$+ (r-1)P^r \sum_{i=1}^{r-2} \frac{(-1)^{r-i}}{i(1-Q)^i Q^{r-1-i}} =$$

$$= \frac{(r-1)P^r}{Q^{r-1}}\left[(-1)^{r-1}\log(1-Q) + \right.$$

$$\left. + \sum_{i=1}^{r-2} \frac{(-1)^{r-i}}{i} \cdot \left(\frac{Q}{1-Q}\right)^i\right].$$

Portanto,

$$\operatorname{Var}[\hat{P}] = \frac{P^r}{(1-P)^{r-1}}(r-1)\left[(-1)^{r-1}\log P + \right.$$

$$\left. + \sum_{j=3}^{r} \frac{(-1)^{j-1}}{r-j+1} \cdot \left(\frac{1-P}{P}\right)^{r-j+1}\right] - P^2 .$$

Outras expressões obtidas através de séries infinitas da variância de $\hat{P}$ são mencionadas no trabalho já citado de HALDANE [2] e também em KENDALL and STUART [3].

3. Determinação de um estimador não viciado e de variância mínima para a razão $R = \dfrac{1-P}{P}$.

A partir de (1. 1) o estimador de máxima verossimilhança para P é expresso por:

$$\hat{P} = \frac{r}{r+X} . \tag{3. 1}$$

De acordo com a propriedade de invariância dos estimadores de máxima verossimilhança, substituindo-se (3. 1) na expressão de R, temos:

$$\hat{R} = \frac{X}{r} . \tag{3. 2}$$

Dado que X distribui-se segundo uma binomial negativa, temos que as condições *i*), *ii*) de 1) estão satisfeitas.

Por outro lado, temos:

$$E[\hat{R}] = \frac{1}{r} E[X] = \frac{1}{r} \cdot \frac{r(1-P)}{P} . \tag{3. 3}$$

Donde:

$$E[\hat{R}] = \frac{1-P}{P} = R . \tag{3. 4}$$

Das expressões (3. 2) e (3. 4), concluimos que $\hat{R}$ também satisfaz as condições *i*), *ii*) e *iii*), respectivamente.

Portanto, entre todos os estimadores não viciados de R, o definido por (3. 2) é de variância mínima, cuja expressão é dada por:

$$\operatorname{Var}[\hat{R}] = \frac{1}{r^2} \operatorname{Var}[X] = \frac{1}{r^2} \cdot \frac{r(1-P)}{P^2} .$$

Donde:

$$\operatorname{Var}[\hat{R}] = \frac{1}{r} \cdot \frac{(1-P)}{P^2} = \frac{(1+R)R}{r} . \tag{3. 5}$$

Devemos acrescentar ainda que estes resultados podem ser obtidos trabalhando-se com o estimador de máxima verossimilhança de $\dfrac{1}{P}$ e decompondo-se $R = \dfrac{1}{P} - 1$.

Podemos observar também que esses resultados não são válidos quando consideramos amostras de tamanho fixo.

BIBLIOGRAFIA

[1] W. FELLER (1957): *An Introduction to Probability Theory and Its Applications*. Vol. 1-2nd Edition John Wiley and Sons, New York.

[2] J. B. S. HALDANE (1945): *On a method of estimating frequencies*. Biometrika, Vol. 33, 222-225.

[3] M. G. KENDALL and A. STUART (1961): *The Advanced Theory of Statistics*. Vol. 2-3rd Edition, Charles Griffin and Company Limited, London.