ESTIMATION OF THE ANNUAL SURVIVAL RATE OF A STATIONARY POPULATION

BU-106-M D. S. Robson April 29, 1959

Introduction

Some natural populations in which reproduction occurs annually reach a state of approximate equilibrium where the number of births each year is just sufficient to balance out the number of deaths at all ages during the previous year. Total population size measured on annual aniversaries therefore remains constant, and the population is said to be stationary. This phenomenon has been observed, for example, in fish populations of inland waters. In many instances it has also been observed that the number of fish in each age group diminishes geometrically with age, implying that the annual survival rate for each age group is the same. The model commonly used in this case to describe the age-frequency distribution in the population is then

$$(1) \qquad f(x)=as^x$$

where s is the annual survival rate and a=1-s is the annual mortality rate. The age X of an individual randomly selected from this population is therefore a chance variable having the geometric probability distribution (1). We consider here the problem of estimating the annual survival rate s on the basis of n independent observations on the chance variable X; that is, on the basis of the ages observed in a random sample of n individuals from this population.

Estimation of the constant annual survival rate

Since the chance variables $X_1,\cdots,X_n$ are independent then their joint probability distribution is

$$P(X_1=x_1,\cdots,X_n=x_n)=P(X_1=x_1)\cdots P(X_n=x_n)=a^n s^{\sum_1^n x_i} ,$$

from which it is readily seen that the sum $T=\Sigma X_i$ is a sufficient statistic. The distribution of T is given by the n-fold convolution of the geometric distribution, which is well known to be the Pascal distribution

$$g(t)=\binom{n+t-1}{t}a^n s^t \quad . \tag{2}$$

As seen from (2), T is also a complete statistic; i.e., if h(T) is a statistic with expectation identically 0 then the function h(t) is identically 0, for if

$$Eh(T)=(1-s)^n \sum_{t=0}^{\infty} h(t)\binom{n+t-1}{t}s^t \underset{0<s<1}{\equiv} 0$$

then the coefficient of s^t in the convergent power series

$$\sum_{t=0}^{\infty}[h(t)\binom{n+t-1}{t}]s^t=0$$

must be 0 for all t.

Existence of a complete sufficient statistic implies that if s is estimable then a uniformly minimum variance unbiased estimator of s exists, and this estimator is the unique function of T which is an unbiased estimator of s. Such a function is uniquely defined by the identity

$$Eh(T)=(1-s)^n \sum_{t=0}^{\infty} h(t)\binom{n+t-1}{t}s^t \underset{0<s<1}{\equiv} s$$

or

$$\sum_{t=0}^{\infty} h(t)\binom{n+t-1}{t}s^t \equiv s(1-s)^{-n} \equiv \sum_{t=0}^{\infty} \binom{n+t-1}{t}s^{t+1}$$

Since the coefficients of s^t on both sides of this identity must be equal then

$$h(t)=\frac{\binom{n+t-2}{t-1}}{\binom{n+t-1}{t}}=\frac{t}{n+t-1}$$

The statistic

$$\hat{s}=\frac{T}{n+T-1}$$

is therefore the unique minimum variance unbiased estimator of s.

Sampling variance of the estimator

The exact variance of $\hat{s}$ is not expressible in closed form; however, an unbiased estimator of $\mathrm{var}(\hat{s})$ is easily constructed using the fact that

$$E\,\frac{T(T-1)\cdots(T-r+1)}{(n+T-1)\cdots(n+T-r)}=(1-s)^n\sum_{t=r}^{\infty}\binom{n+t-r-1}{t-r}s^t=s^r\quad.$$

The variance of $\hat{s}$ may thus be expressed

$$\mathrm{var}(\hat{s})=E\left(\frac{T}{n+T-1}\right)^2-E\,\frac{T(T-1)}{(n+T-1)(n+T-2)}$$

showing that

$$\widehat{\mathrm{var}}(\hat{s})=\hat{s}\left(\hat{s}-\frac{T-1}{n+T-2}\right)$$

is an unbiased estimator of $\mathrm{var}(\hat{s})$, Moreover, by completeness of T, this is the minimum variance unbiased estimator of $\mathrm{var}(\hat{s})$.

An approximation revealing the magnitude of $\mathrm{var}(\hat{s})$ is given by the variance of the asymptotic distribution of the chance variable $\sqrt{n}(s^*-s)$, where

$$s^* = \frac{T}{n+T}$$

is the maximum likelihood estimator of s. Thus, for large n, the variance of $\hat{s}$ is approximately

$$-\frac{1}{nE\left(\dfrac{d^2\log f}{ds^2}\right)}=\frac{s(1-s)^2}{n}\quad.$$