ESTIMATING THE MEAN OF A RANDOM BINOMIAL PARAMETER

G. MORRIS SOUTHWARD
INTERNATIONAL PACIFIC HALIBUT COMMISSION
and
J. VAN RYZIN
UNIVERSITY OF WISCONSIN

1. Introduction

In studying biological phenomenon, one often observes random variables which are the result of other randomly occurring unobservable events. This is usually the case in the observation of genetic traits. The measurable trait in question has a probability distribution for the population of animals under study. Each individual member of the population of animals carries a value of the measurable trait, but it may or may not (and often is not) directly observable. It is not difficult to envision the probability distribution of the trait in the population as being continuous, while the distribution of the visible expression of the trait is a discrete count depending on the value of the measurable trait.

Such a problem came to the authors' attention during discussion with a poultry scientist who was interested in the probability distribution governing the frequency with which blood spotted eggs occur. Poultrymen wish to determine from examination of a small number of eggs laid early in the life of each hen what the average probability of laying blood spotted eggs is for the flock.

The problem can be conceptualized as follows. The distribution of blood spots in eggs for a given chicken is taken as binomial. That is, if p represents the probability of a given chicken to lay a blood spotted egg and m eggs are laid, then X = number of blood spotted eggs is binomially distributed with parameters m and p assuming the eggs are laid independently. However, the probability p (or propensity) for laying blood spotted eggs (the trait in question), differs from chicken to chicken and can be thought of as having a continuous distribution on the unit interval. The probability distribution of the blood spotting trait p in the population is not directly observable. That is, one might postulate that the binomial parameter p (or trait) has a distribution on the unit interval and that the values of this probability carried by each bird in the flock are independently allocated according to this distribution, denoted $G(p)$. Rarely, if ever, are values of p directly observable.

The investigation of J. Van Ryzin was supported in part by USPHS Research Grant No. GM-10525-08, National Institute of Health, Public Health Service, and in part by NSF Contract No. GP-9324.