Statistical Methodology 33 (2016) 192–202

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

Estimating the integer mean of a normal model related to binomial distribution

Rasul A. Khan

Cleveland State University, United States

ARTICLE INFO

Article history:
Received 29 October 2015
Received in revised form
4 March 2016
Accepted 29 September 2016
Available online 7 October 2016

Keywords:
Asymptotic
Estimation
MLE
Sequential probability ratio test
Stopping time

ABSTRACT

A problem for estimating the number of trials n in the binomial distribution $B(n, p)$, is revisited by considering the large sample model $N(\mu, c\mu)$ and the associated maximum likelihood estimator (MLE) and some sequential procedures. Asymptotic properties of the MLE of n via the normal model $N(\mu, c\mu)$ are briefly described. Beyond the asymptotic properties, our main focus is on the sequential estimation of n. Let $X_1, X_2, \ldots, X_m, \ldots$ be iid $N(\mu, c\mu)(c > 0)$ random variables with an unknown mean $\mu = 1, 2, \ldots$ and variance $c\,\mu$, where c is known. The sequential estimation of μ is explored by a method initiated by Robbins (1970) and further pursued by Khan (1973). Various properties of the procedure including the error probability and the expected sample size are determined. An asymptotic optimality of the procedure is given. Sequential interval estimation and point estimation are also briefly discussed.

1. Introduction

The estimation of n in the binomial distribution $B(n, p)$ was considered by Feldman and Fox [7]. However, motivated by large sample approximations they considered the normal model $N(\mu, \mu)$ and concentrated on the estimation of μ for the following reason. Consider a sequence of binomial random variables $X_j, j = 1, 2, \ldots$ with known probability p and unknown parameter n. Since $EX_j = n\,p$, $var(X_j) = np\,q$. $(q = 1-p)$, and $(X_j - np)/\sqrt{npq}$ can be approximated in distribution by standard normal distribution for large n, it follows that $\frac{X_j}{q}$ is approximately $N(\frac{np}{q}, \frac{np}{q}) = N(\mu, \mu)$ for large n. This is the normal model considered by Feldman and Fox [7], however, they treated μ as continuous

E-mail address: r.khan@csuohio.edu.

http://dx.doi.org/10.1016/j.stamet.2016.09.004
1572-3127/

parameter, and using a fixed sample size they studied the estimation of μ. If we make a change as $\frac{X_j}{p}$, then it can be approximated by $N(n, n\frac{q}{p}) = N(\mu, c\mu)$ where $\mu = 1, 2, 3, \ldots$, and c is a known positive factor. A similar problem was considered by McCabe [12] for estimating the number of terms in a sum by an adhoc sequential method. To outline this second problem, consider an iid sequence $\{X_j, j \geq 1\}$ with known mean μ and variance σ^2, then $Y_n = X_1 + \cdots + X_n$ has mean $n\mu$ and variance $n\sigma^2$. Consequently, $Y_n^* = Y_n/\mu$ has mean n and variance $n(\sigma/\mu)^2 = nc(c > 0)$. Thus Y_n^* is approximately $N(n, c\,n) = N(\mu, c\,\mu)$ for large n, where $\mu = 1, 2, \ldots$ and c is a known positive factor. Thus both problems are included in this later model. This is the motivation for considering the model $N(\mu, c\mu)$ where μ is an unknown positive integer and c is a known factor (see Section 4.3 in Banerjee and Mukhopadhyay [1] for another approach to sequential estimation of μ when $c = 1$). Also, there is a duality here, namely, sequential testing about μ leads to the estimation of μ.

In what follows we consider the model $N(\mu, c\,\mu)$ $(c > 0)$, $\mu = 1, 2, \ldots$. The underlying data $X_1, X_2, \ldots$ observed sequentially follow $N(\mu, c\mu)$ distribution. The problem is to find an estimate of μ in such a way that the error probability of incorrectly estimating μ is uniformly bounded by a small preassigned level. That is, $\widehat{\mu}_N$ as an estimate based on a random sample of size N is such that $\sup_j P_j(\widehat{\mu}_N \neq j) \leq \epsilon$ $(0 < \epsilon < 1)$. The estimation procedure given here is equivalent to a sequence of SPRTs for a countable number of simple hypotheses. The procedure is developed by defining a suitable stopping time N, and its standard properties such as termination, error probability, expected stopping time, and asymptotic optimality are discussed. The idea was initiated and inspired by Robbins [14] for estimating the integer mean of a normal distribution with known variance. In our case the mean is restricted to positive integer, and the variance is a known factor times the unknown mean. The procedure is simple and interesting, and inherently enjoys all the desirable properties of Wald's SPRT.

Here is a synopsis of the paper. In Section 2 we describe the asymptotic properties of the MLE based on a fixed sample size. Section 3 contains the sequential procedure for the model $N(\mu, c\mu)$ and its properties are given in Section 4, where a numerical comparison with the estimate of Feldman and Fox [7] is also given. Section 5 briefly discusses the sequential point and interval estimation of the binomial parameter N when the success probability p is known.

2. The asymptotic properties of the MLE

Let $X_1, \ldots, X_m$ be iid random variables having binomial distribution $B(n, p)$ where p is known but n is an unknown parameter. The parameter n is estimated by the MLE given below, whose asymptotic properties will be discussed. But we first provide some required central moments of X_1. Clearly, $EX_1 = np$, and let $q = 1 - p$, and $\mu_r = E(X_1 - np)^r$. Then it is known (cf. Cramer [6]) that $\mu_2 = npq$, $\mu_3 = npq(q - p)$, and $\mu_4 = 3n^2p^2q^2 + npq(1 - 6npq)$. Now let $\xi_1 = X_1^2$, and note that $E\xi_1 = (np)^2 + npq$. Let $\gamma = np$ and further note that $EX_1^4 = \mu_4 + 4\mu_3\gamma + 6\mu_2\gamma^2 + \gamma^4$. Using this and the central moments one can show that

$$\sigma^2 = var(\xi_1) = var(X_1^2) = 4n^3p^3q + 2n^2p^2q(3q - 2p) + npq(1 - 6pq). \tag{1}$$

Now let $\overline{S}_m(2) = \frac{1}{m}\sum_{i=1}^m X_i^2 = \sum_{i=1}^m \xi_i/m$, where $\xi_i = X_i^2$. Using the likelihood function L_m from the approximate distribution of each X_j as normal with mean np and variance npq it is easily seen (cf. Feldman and Fox [7]) that the MLE of n is given by

$$\hat{n} = \frac{q}{p}\left(\frac{1}{4} + \frac{\overline{S}_m(2)}{q^2}\right)^{\frac{1}{2}} - \frac{q}{2p}. \tag{2}$$

Throughout the paper we keep n fixed and all the asymptotics are as the sample size $m \to \infty$. The asymptotic behavior of the MLE $\hat{n}$ is given by the following simple but quite useful properties.

Theorem 1. *For n fixed and large m we have*

(i) $E\hat{n} = n - \frac{\sigma^2}{pm}(q + 2np)^{-3} + O\left(\frac{1}{m^2}\right)$.
and
(ii) $var(\hat{n}) = \frac{\sigma^2}{mp^2(q+2np)^2} + O\left(\frac{1}{m^2}\right)$.