HOEFFDING TYPE INEQUALITIES FOR LIKELIHOOD RATIO TEST STATISTIC

M. Radavičius
Institute of Mathematics and Informatics, Vilnius University
Vilnius, LITHUANIA
e-mail: marijus.radavicius@vu.mii.lt

Abstract

For Bernoulli trials, simple upper and lower bounds for tail probabilities of logarithmic likelihood ratio statistic are given. The bounds are exact up to a factor of 2. A problem of generalization of the results to the multinomial distribution is briefly discussed.

1 Introduction

Let $\mathbf{y} = (y_1, \ldots, y_n)$ be a random vector having the multinomial distribution

$$\mathbf{y} \sim \text{Multinomial}_n(N, \mathbf{p}), \quad \mathbf{p} = (p_1, \ldots, p_n).$$

For $n = 2$, $y_1 \sim \text{Binomial}(N, p_1)$. The maximum likelihood estimator of the unknown probabilities $\mathbf{p}$ is given by

$$\hat{\mathbf{p}} = \hat{\mathbf{p}}_N := N^{-1}\mathbf{y}.$$

Define scaled (logarithmic) likelihood ratio statistic

$$\ell_n(\hat{\mathbf{p}}, \mathbf{p}) := \sum_{i=1}^{n} \hat{p}_i \log\left(\frac{\hat{p}_i}{p_i}\right).$$

Note that for $n = 2$,

$$\ell(\hat{p}_1, p_1) := \hat{p}_1 \log\left(\frac{\hat{p}_1}{p_1}\right) + (1 - \hat{p}_1)\log\left(\frac{1 - \hat{p}_1}{1 - p_1}\right) = \ell_2(\hat{\mathbf{p}}, \mathbf{p}). \tag{1}$$

Hoeffding (1965) proved the following inequality (see also Kallenberg, 1985): for $n = 2$,

$$\mathbf{P}\{\ell_n(\hat{\mathbf{p}}_N, \mathbf{p}) \geq x\} \leq 2\mathrm{e}^{-Nx}, \quad x > 0. \tag{2}$$

It is important to stress that (2) is *universal*: it holds for all $x > 0$, all $p_1 \in [0, 1]$ and all $N = 1, 2, \ldots$. It is also *tight*, i.e. it cannot be improved without imposing some additional conditions.

The *problem* is to generalize this inequality to the case $n > 2$.

Generalizations of (2) to the case $n > 2$ have been obtained by Hoeffding himself (1965) and W.C.M. Kallenberg (1985). The inequality established by W.C.M. Kallenberg is tight up to a constant. However it holds only for $x \leq 0.15$ and impose some boundedness from below restriction on probabilities $\mathbf{p}$. Known universal bounds (i.e. bounds that are independent of $\mathbf{p}$) are loose and impractical for large n (W.C.M. Kallenberg, 1985, inequality (2.6)). The upper bound typically exceeds the corresponding lower bound by a factor of order $\sqrt{xN}$ (see W.C.M. Kallenberg, 1985, Theorem 2.1 on p. 1557). This applies to 2) as well.

2 Notation

Define the signed logarithmic likelihood ratio statistic for the binomial distribution

$$s_q(u) := \operatorname{sign}(u-q)\,\ell(u,q), \quad (u,q) \in (0,1)\times(0,1).$$

The logarithmic likelihood ratio statistic for the binomial distribution is defined in (1). The function $s_q(u)$ is strictly increasing and continuous with respect to $u \in (0,1)$. Let $\bar{s}_q$ denote the inverse function of s_q: $\bar{s}_q(s_q(u)) \equiv u$. In what follows, we reserve notation χ^2_m for a random variable which has χ^2 distribution with m degrees of freedom.

Let

$$b(t) = b(t; N, q) := \frac{\Gamma(N+1)}{\Gamma(t+1)\Gamma(N-t+1)}\, q^t (1-q)^{N-t}, \quad t \in [0, N].$$

Note that

$$b(k; N, q) = C_N^k q^k (1-q)^{N-k}, \quad k = 0, 1, \ldots, N,$$

is the binomial probability density (mass) function.

3 Results

The proposition below gives upper and lower bounds (exact up to a factor of 2) for the tail probabilities of the logarithmic likelihood ratio statistic. In contrast to (2), they depend on the success probability of the binomial distribution.

Proposition 1. *Let $\hat{p}_N := N^{-1}y$, $y \sim \mathrm{Binomial}(N,p)$. Then*

$$\begin{aligned} \mathbf{P}\{\ell(\hat{p}_N, p) \geq x\} &\leq \mathbf{P}\{\chi_1^2 \geq 2\,xN\} \\ &+ b(N\bar{s}_p(-x); N, p) + b(N\bar{s}_p(x); N, p) \\ &\leq 2\,\mathbf{P}\{\ell(\hat{p}_N, p) \geq x\}. \end{aligned} \tag{3}$$

The inequalities for the upper tail probability presented below are more apprehensible than (3). Let $x_k := s_p(k/N)$ with $k/N > p$. Then

$$\begin{aligned} \max(2^{-1}\mathbf{P}\{\chi_1^2 \geq 2\,x_k N\}, b(k; N, p)) &\leq \mathbf{P}\{s_p(\hat{p}_N) \geq x_k\} \\ &\leq 2^{-1}\mathbf{P}\{\chi_1^2 \geq 2\,x_k N\} + b(k; N, p). \end{aligned} \tag{4}$$

Note that the first term in the right-hand side of (4) is just the tail probability of the asymptotic distribution of the logarithmic likelihood ratio statistic.

The inequalities (4) as well as the Proposition 1 are simple corollaries of results by Zubkov and Serov [4].

Remark. We expect that, for arbitrary $n > 2$, exact (up to a constant factor) upper bounds for tail probabilities of logarithmic likelihood ratio statistic can be obtained by making use of the Proposition 1 and induction with respect to n.

References

[1] Hoeffding W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*. Vol. **58**, pp. 13-30.

[2] Hoeffding W. (1965). Asymptotically Optimal Tests for Multinomial Distributions. *Ann. Math. Statist.* Vol. **36**, pp. 369-401.

[3] Kallenberg W.C.M. (1985). On moderate and large deviations in multinomial distributions. *Annals of Statistics*. Vol. **13**, pp. 1554-1580.

[4] Zubkov A.M., Serov A.A. (2012). A complete proof of universal inequalities for distribution function of binomial law. *Teoriya Veroyatnostei i ee Primeneniya*. Vol. **57**, pp. 597-602.