Stochastic Inequalities
IMS Lecture Notes – Monograph Series
Volume 22 (1993)

ON THE BIAS OF THE JACKKNIFE ESTIMATE OF VARIANCE[1]

By RICHARD A. VITALE

University of Connecticut

Using machinery developed earlier for the covariances of symmetric statistics, we consider various aspects of the bias of the jackknife estimate of variance.

1. Introduction

The jackknife estimate of variance (Quenouille (1949, 1956), Tukey (1958)) can be described as follows. Given a symmetric function h of iid arguments $X_1, X_2, \ldots, X_m$, it is desired to estimate $\sigma^2 = \text{Var } h$. With an augmented supply $X_1, X_2, \ldots, X_n$ where $n = m+1$, or more generally $n \geq m+1$, one forms $Q = \binom{n-1}{m}^{-1} \sum_{|I|=m} [h(X_I) - \bar{h}]^2$ where $\bar{h} = \binom{n}{m}^{-1} \sum_{|I|=m} h(X_I)$ and $X_I \equiv (X_{i_1}, X_{i_2}, \ldots, X_{i_m})$ with $I = \{i_1, i_2, \ldots, i_m\}$. Several papers (Efron and Stein (1981), Karlin and Rinott (1982), Bhargava (1983), Vitale (1984), Steele (1986)) have considered the bias relation

$$\sigma^2 \leq EQ, \tag{1.1}$$

which has come to be known as the Efron–Stein inequality. Our purpose here is to investigate aspects of (1.1) including (i) an alternate proof with variant forms of the condition for equality, (ii) a sharpening, (iii) a complementary upper bound, and (iv) a consideration of Q as an estimator which is "contaminated" by estimators of other parameters.

2. Preliminaries

If $X_1, X_2, \ldots, X_n$ are iid random variables and h is a symmetric function of m of them with $Eh^2 < \infty$, then we set

$$r_k = \text{Cov}[h(X_I), h(X_J)]$$

[1]Work supported in part under grants Office of Naval Research N00014–90–J–1641 and National Science Foundation DMS–9002665.

AMS 1991 *subject classifications.* Primary 60B15; Secondary 62M10, 62M15.

Key words and phrases. Efron–Stein inequality, jackknife estimate of variance, symmetric statistic, U–statistic.