Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 112–114

AN EXPANSION FOR SYMMETRIC STATISTICS AND THE EFRON-STEIN INEQUALITY

BY RICHARD A. VITALE
Claremont Graduate School

The Efron-Stein inequality and a generalization by Bhargava are derived using a tensor-product basis and bounds for covariances of related symmetric statistics.

1. Introduction. Let $S(X_1, \ldots, X_n)$ be a symmetric function of its iid arguments. Its variance can be estimated by the jackknife technique as follows: assuming an augmented iid collection $X_1, \ldots, X_n, X_{n+1}$, form $S_i = S(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{n+1})$, $i = 1, \ldots, n+1$ and $\bar{S} = (n+1)^{-1}\Sigma_{i=1}^{n+1} S_i$. Then $\operatorname{Var} S(X_1, \ldots, X_n)(=\operatorname{Var} S_i)$ is estimated by $Q = \Sigma_{i=1}^{n+1} (S_i - \bar{S})^2$. As part of an extensive study, Efron and Stein (1981) showed that Q is necessarily positively biased, an observation that has come to be known as the *Efron-Stein inequality*.

THEOREM 1.

$$\operatorname{Var} S(X_1, \ldots, X_n) \leq EQ \tag{1.1}$$

with equality iff. S is linear in functions of its individual arguments.

Other proofs and extensions have been given by Bhargava (1980) and Karlin and Rinott (1982), and the inequality has already had interesting applications (Hochbaum and Steele (1982), Steele (1981), Steele (1982)). Our purpose here is to derive the inequality by using an idea exploited for other purposes in Rubin and Vitale (1980): expansion of symmetric statistics in a tensor-product basis. The approach yields attractive, concrete representations and is particularly well-adapted to proving the E-S inequality by first establishing a universal bound on the covariance of related symmetric statistics. It is an alternative to the ANOVA-type expansions used elsewhere.

2. The Efron-Stein Inequality via Covariance Bounds. If $e_0(X_1) \equiv 1$, $e_1(X_1)$, $e_2(X_1)$, ... form an orthonormal basis for the square integrable functions of X_1, then products of the type $\Pi_{i=1}^{n} e_{\nu_i}(X_i)$ form an orthonormal basis for the square integrable functions of $\mathbf{X} = (X_1, \ldots, X_n)$. For ease of notation we denote the above product by $e_\nu(\mathbf{X})$, $\nu = (\nu_1, \ldots, \nu_n)$.

THEOREM 2. *For* $i \neq j$,

$$0 \leq \operatorname{Cov}(S_i, S_j) \leq ((n-1)/n) \operatorname{Var} S_1 \tag{2.1}$$

with equality above iff. S_1 *is linear in functions of its individual arguments.*

Proof. Without loss of generality, assume that the S_i (which are identically distributed) have zero mean. Accordingly, we consider ES_1S_{n+1} as a surrogate for $\operatorname{Cov}(S_i, S_j)$, $i \neq j$. Using the basis given above and symmetry considerations yields

---

AMS 1980 subject classifications. Primary 60E15, secondary 62H20, 62J10.

Key words and phrases: ANOVA-type expansion, covariance bounds, Efron-Stein inequality, symmetric statistic, tensor-product basis, U-statistic.

$$S_1 = S(X_2, \ldots, X_n, X_{n+1}) = \Sigma\, c_\nu e_\nu(\mathbf{X}), \text{ where } \mathbf{X} = (X_2, \ldots, X_n, X_{n+1}),$$

and

$$S_{n+1} = S(X_1, \ldots, X_n) = S(X_2, \ldots, X_n, X_1) = \Sigma c_\nu e_\nu(\mathbf{X}'), \text{ where } \mathbf{X}' = (X_2, \ldots, X_n, X_1).$$

Then

$$ES_1S_{n+1} = E\Sigma c_\nu e_\nu(\mathbf{X})\, \Sigma\, c_\mu e_\mu(\mathbf{X}') = \Sigma\, c_\nu c_\mu E e_\nu(\mathbf{X}) \mathbf{e}_\mu(\mathbf{X}').$$

The expectation of $e_\nu(\mathbf{X})e_\mu(\mathbf{X}')$ is zero unless $\nu = \mu$ with $\nu_n = \mu_n = 0$, in which case it is unity. Thus $ES_1S_{n+1} = \Sigma_{\nu_n=0}\, c_\nu^2$, which displays the asserted positive correlation.

For the upper bound, we symmetrize: note that generally for summands $\{\sigma_\nu\}$ which are symmetric in ν

$$\Sigma_{\nu_n=0}\sigma_\nu = n^{-1}\, \Sigma_\nu z_\nu \sigma_\nu$$

where z_ν is the number of zero components in ν. The $\{c_\nu\}$ may be assumed symmetric in ν and hence

$$ES_1S_{n+1} = n^{-1}\, \Sigma_\nu z_\nu c_\nu^2.$$

Now $z_\nu c_\nu^2 \leq (n-1)c_\nu^2$ for every ν because of the centering of the S_i, which leads to

$$ES_1S_{n+1} \leq ((n-1)/n)\, \Sigma_\nu c_\nu^2 = ((n-1)/n) \operatorname{Var} S_1.$$

Equality occurs iff $z_\nu = n-1$ for all non-vanishing c_ν. This means that

$S_{n+1} = f(X_1) + \ldots + f(X_n)$ for some f. □

Returning to the Efron-Stein inequality, we note that expanding EQ in (1.1) yields

$$\operatorname{Var} S_1 \leq n \operatorname{Var} S_1 - n \operatorname{Cov}(S_1, S_{n+1})$$

which, upon rearrangement, is the upper inequality in (2.1).

3. A Higher-Order Construction. A natural question to ask is whether a more ample supply of randomness can lead to other estimates and inequalities. Specifically, suppose that S is a symmetric function of n iid. arguments which can now be chosen form $X_1, X_2, \ldots, X_N$ where $n < N$ ($N = n+1$ in the previous section). Proceeding by analogy, for $A = \{\nu_1, \nu_2, \ldots, \nu_n\}$ with distinct $\nu_i \in \{1,2, \ldots, N\}$, define $S_A = S(X_{\nu_1}, X_{\nu_2}, \ldots, X_{\nu_n})$ and $\bar{S} = \binom{N}{n}^{-1}\Sigma_{|A|=n} S_A$. Then an estimate for Var S_A is $Q = \binom{N-1}{n}^{-1}\Sigma_{|A|=n}(S_A - \bar{S})^2$. This is the set-up studied by Bhargava (1980), who showed that positive bias obtains here as well.

THEOREM 3. *Var $S_A \leq EQ$ with equality iff. S_A is linear in functions of its individual arguments.*

In treating this problem, we establish bounds on covariances as before. These generalize theorem 2 and show that the upper bound is linear in the number of shared arguments (cf. Bhargava (1980, p. 6)).

THEOREM 4. *For $|A \cap A'| = k$, $0 \leq \operatorname{Cov}(S_A, S_{A'}) \leq (k/n) \operatorname{Var} S_A$ with equality above iff S_A is linear in functions of its individual arguments.*

Proof. The argument parallels that of theorem 2; assuming zero mean, we compute $ES'S''$ where

$$S' = S(X_1, \ldots, X_k, Y_{k+1}, \ldots, Y_n), \quad S'' = S(X_1, \ldots, X_k, Z_{k+1}, \ldots, Z_n)$$

(the X,Y,Z variables taken together are iid.). This gives $ES'S'' = \Sigma' c_\nu^2$ where Σ' denotes summation over subscripts ν with vanishing final $n-k$ components. This can be symmetrized to the form

$$ES'S'' = \binom{n}{k}^{-1}\Sigma_\nu \binom{z_\nu}{n-k} c_\nu^2$$

where z_ν is the number of zero components of ν.

This is clearly non-negative and noting that $z_\nu c_\nu^2 \leqslant (n-1)c_\nu^2$ yields the upper bound with the condition for equality. □

Theorem 3 follows directly from the upper bound just given. We merely sketch some important points. In computing EQ, sums of the form $\Sigma_A ES_A S_{A'}$, intervene and calculate out to

$$\Sigma_{k=0}^{n} \binom{n}{k}\binom{N-n}{n-k}\left[\binom{n}{k}^{-1}\Sigma_\nu \binom{z_\nu}{n-k} c_\nu^2\right],$$

the bracketed quantity being the exact value of the covariance in theorem 4. This leads to

$$EQ = \Sigma_\nu c_\nu^2 \Sigma_{k=0}^{n} (N/(N-n)) \binom{N-n}{n-k}\binom{N}{n}^{-1}\left[\binom{n}{k} - \binom{z_\nu}{n-k}\right],$$

and a collapse to the lower bound $\Sigma_\nu c_\nu^2 = \operatorname{Var} S_A$.

Acknowledgments. I thank Mike Steele for suggesting the tensor-product basis approach to the Efron-Stein inequality.

REFERENCES

BHARGAVA, R. P. (1980). A property of the jackknife estimation of the variance when more than one observation is omitted. Tech. Report No. 140, Dept. Statist., Stanford Univ.

EFRON, B. and STEIN, C. (1981). The jackknife estimate of variance. *Ann. Statist.* *9* 586–596.

HOCHBAUM, D. and STEELE, J. M. (1982). Steinhaus's geometric location problem for random samples in the plane. *Adv. Appl. Prob.* *14* 56–67.

KARLIN, S. and RINOTT, Y. (1982). Applications of ANOVA type decompositions for comparisons of conditional variance statistics including jackknife estimates. *Ann. Statist.* *10* 485–501.

RUBIN, H. and VITALE, R. A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* *8* 165–170.

STEELE, J. M. (1981). Complete convergence of short paths and Karp's algorithm for the tsp. *Math. Oper. Res.* *6* 374–378.

STEELE, J. M. (1982). Optimal triangulation of random samples in the plane. *Ann. Probab.* *10* 548–553.