Ann. Inst. Statist. Math.
36 (1984), Part A, 129-134

TWO INEQUALITIES WITH AN APPLICATION

C. G. BHATTACHARYA

(Received Sept. 14, 1982; revised Mar. 22, 1983)

Summary

An inequality used in Brown and Cohen (1974, *Ann. Statist.*, 2(5), 963-976) and Bhattacharya (1978, *Ann. Inst. Statist. Math.*, A, **30**, 407-414) is generalized and another useful inequality derived from it. An application of the latter which provides a more elegant approach to and an improvement over a result in Shinozaki (1978, *Commun. Statist. Theor. Meth.*, A, 7, 1421-1432), is also presented.

1. Introduction

Brown and Cohen [4] and Bhattacharya [2] used and integral inequality which can be stated as follows: Let u, v, t be functions of a random variable (say x) such that v is positive with a finite expectation and t, u/v are monotonic (not necessarily strictly) in opposite directions with respect to x; Then $\mathrm{E}(tu)/\mathrm{E}(tv) \leqq \mathrm{E}\,u/\mathrm{E}\,v$, provided $\mathrm{E}(tv) > 0$. This inequality is generalized here and another useful inequality derived from it. This is done in Section 2. Section 3 is then devoted to an application of Theorem 2.2. Here, Theorem 3.1 uses Theorem 2.2 to obtain a result which is an improvement of a similar result in Shinozaki [8] and Bhattacharya [3]. It may be observed that besides the improvement mentioned the approach here is more elegant than that in the two papers just cited, where the proof was dependent upon the favourable outcome of certain computations.

2. The inequalities

Suppose f and g are functions of n random variables $x_1, x_2, \cdots, x_n$. We shall use the symbol $\mathrm{E}_r f$, where $r \leqq n$ to denote the conditional expectation of f given $x_1, x_2, \cdots, x_r$. We shall also use the abbrevia-

AMS 1970 subject classifications: Primary 26A86, 62F10; Secondary 26A51, 62K10.
Key words and phrases: Integral inequality, concave functional, common mean, unbiased estimator.

tions: $f\uparrow(x_i)$ for the statement f is increasing in x_i; $f\downarrow(x_i)$ for the statement that f is decreasing in x_i. The abbreviations: f SD $g|x_i$ would mean that f and g are monotonic in the same direction with respect to x_i in the sense that either $f\uparrow(x_i)$ and $g\uparrow(x_i)$ or, $f\downarrow(x_i)$ and $g\downarrow(x_i)$. Similarly f OD $g|x_i$ would mean that f and g are monotonic in opposite directions with respect to x_i in the sense that either $f\uparrow(x_i)$ and $g\downarrow(x_i)$ or, $f\downarrow(x_i)$ and $g\uparrow(x_i)$. We now prove,

THEOREM 2.1. *Let u, v, t be functions of random variables $x_1, x_2, \cdots, x_n$ such that v is positive with a finite expectation and* $\mathrm{E}(tv)>0$. *Let $f_r=\mathrm{E}_r(tv)/\mathrm{E}_r(v)$ and $g_r=\mathrm{E}_r(u)/\mathrm{E}_r(v)$. Then*

$$\text{(2.1)}\qquad \text{(i)}\quad f_r\,\mathrm{SD}\,g_r|x_r\ \ \forall\, r\leqq n \Longrightarrow \text{(ii)}\quad \mathrm{E}(tu)/\mathrm{E}(tv)\geqq \mathrm{E}(u)/\mathrm{E}(v)$$

$$\text{(2.2)}\qquad \text{(iii)}\quad f_r\,\mathrm{OD}\,g_r|x_r\ \ \forall\, r\leqq n \Longrightarrow \text{(iv)}\quad \mathrm{E}(tu)/\mathrm{E}(tv)\leqq \mathrm{E}(u)/\mathrm{E}(v)\,.$$

PROOF. The proof is based on Cohen [5] concerning the particular case stated in the introduction. We have,

$$\text{(2.3)}\qquad \text{(ii)}\ \Longleftrightarrow \mathrm{E}(tu)/\mathrm{E}(v)\geqq[\mathrm{E}(u)/\mathrm{E}(v)][\mathrm{E}(tv)/\mathrm{E}\,v] \Longleftrightarrow \mathrm{E}^*(tw)\geqq \mathrm{E}^*(w)\,\mathrm{E}^*(t)$$

where, $w=u/v$; E^* stands for expectation with respect to the probability measure P^* given by $P^*(A)=\mathrm{E}(vI_A)/\mathrm{E}\,v$, where I_A is the indicator function of the set A and P is the probability measure corresponding to the distribution of $(x_1, x_2, \cdots, x_n)$. Note that $\mathrm{E}_r^*(t)=f_r$; $\mathrm{E}_r^*(w)=g_r$. Hence

$$\text{(2.4)}\qquad \text{(i)}\ \Longleftrightarrow \mathrm{E}_r^*(t)\,\mathrm{SD}\,\mathrm{E}_r^*(w)\,|\,x_r\ \forall\, r\leqq n\,.$$

It is well known that (see e.g. Hardy, Littlewood and Polya [6], p. 43), if f and g are functions of a single random variable x, then

$$\text{(2.5)}\qquad f\,\mathrm{SD}\,g\,|\,x \Longrightarrow \mathrm{E}(fg)\geqq \mathrm{E}(f)\,\mathrm{E}(g)$$

$$\text{(2.6)}\qquad f\,\mathrm{OD}\,g\,|\,x \Longrightarrow \mathrm{E}(fg)\leqq \mathrm{E}(f)\,\mathrm{E}(g)\,.$$

In view of (2.4) and (2.5),

$$\text{(2.7)}\qquad \text{(i)}\ \Longrightarrow \mathrm{E}_{r-1}^*[\mathrm{E}_r^*(t)\,\mathrm{E}_r^*(w)]\geqq \mathrm{E}_{r-1}^*(t)\,\mathrm{E}_{r-1}^*(w)\ \forall\, r$$

since $\mathrm{E}_r^*(t)$ and $\mathrm{E}_r^*(w)$ are functions of $(x_1, x_2, \cdots, x_r)$ and the operation E_{r-1}^* on such functions is equivalent to taking (conditional) expectation of them considered as a function of x_r only while the others are temporarily held fixed. (For a rigorous justification of the last statement see Bahadur and Bickel [1]).

Obviously $\mathrm{E}_r^*\,\mathrm{E}_s^*=\mathrm{E}_r^*\ \forall\, r\leqq s$. Hence, it follows from (2.7) that (i) $\Longrightarrow \mathrm{E}^*(tw)\geqq \mathrm{E}^*(t)\,\mathrm{E}^*(w)$. This proves (2.1) because of (2.3).

The proof of (2.2) is similar. In this case, we use (2.6) instead of (2.5).

Remark 2.1. (2.5) and (2.6) follow from the fact that $\operatorname{cov}[f(x), g(x)]$ can be written in the form $(1/2)\,\mathrm{E}\,[\{f(x)-f(y)\}\{g(x)-g(y)\}]$, where y is a random variable distributed indentically as but independently of x. This simple argument has sometimes been overlooked giving rise to more complicated arguments with unnecessary assumptions (see e.g. Kimball [7]).

THEOREM 2.2. *Let $x_1, x_2, \cdots, x_k$ be mutually independent positive random variables such that $\mathrm{E}\,x_i^{-2}$ is finite. Let*

$$f=1/\sum p_i x_i\,; \quad 0<p_i<1\,; \quad \sum p_i=1\,.$$

Then

$$\mathrm{E}\,f/\mathrm{E}\,f^2 \geqq \mathrm{Min}\,[\mathrm{E}\,x_i^{-1}/\mathrm{E}\,x_i^{-2}:\ i=1, 2, \cdots, k]\,.$$

PROOF. The theorem is trivial for $k=1$. We shall prove it for $k=2$ from which the proof for $k>2$ will be obvious. To avoid subscripts let p, x, y stand for p_1, x_1, x_2 and let $m=\mathrm{Min}\,(\mathrm{E}\,x^{-1}/\mathrm{E}\,x^{-2}, \mathrm{E}\,y^{-1}/\mathrm{E}\,y^{-2})$. Then

$$\begin{aligned} m &\leqq p\,\mathrm{E}\,x^{-1}/\mathrm{E}\,x^{-2}+(1-p)\,\mathrm{E}\,y^{-1}/\mathrm{E}\,y^{-2} \\ &=[p\,\mathrm{E}\,x^{-1}\,\mathrm{E}\,y^{-2}+(1-p)\,\mathrm{E}\,y^{-1}\,\mathrm{E}\,x^{-2}]/(\mathrm{E}\,x^{-2}\,\mathrm{E}\,y^{-2}) \\ &=\mathrm{E}\,[px^{-1}y^{-2}+(1-p)x^{-2}y^{-1}]/\mathrm{E}\,(x^{-2}y^{-2})\,, \end{aligned} \tag{2.8}$$

since x and y are independent. Define $g=p/y+(1-p)/x$ and note that $fg=x^{-1}y^{-1}$. Then (2.8) can be written as

$$m \leqq \mathrm{E}\,(tu)/\mathrm{E}\,(tv) \tag{2.9}$$

where $t=g^2$; $u=f$, $v=f^2$. Obviously, $t\downarrow(y)$ and $u/v\uparrow(y)$. Also, $\mathrm{E}\,(tv|x)/\mathrm{E}\,(v|x)=\mathrm{E}\,(f^2g^2|x)/\mathrm{E}\,(f^2|x)=\mathrm{E}\,(x^{-2}y^{-2}|x)/\mathrm{E}\,(x^{-2}y^{-2}t^{-1}|x)=\mathrm{E}\,(y^{-2}|x)/\mathrm{E}\,(y^{-2}\cdot t^{-1}|x)$, which is decreasing in x since $t\downarrow(x)$; and $\mathrm{E}\,(u|x)/\mathrm{E}\,(v|x)=\mathrm{E}\,(f|x)/\mathrm{E}\,(f^2|x)$, which is increasing in x since its derivative with respect to x is $[2p\,\mathrm{E}\,(f|x)\,\mathrm{E}\,(f^3|x)-p\,\mathrm{E}^2\,(f^2|x)]/\mathrm{E}^2\,(f^2|x)\geqq 0$ by an well-known inequality concerning absolute moments. Hence, Theorem 2.1 applies and we have

$$m \leqq \mathrm{E}\,(tu)/\mathrm{E}\,(tv) \leqq \mathrm{E}u/\mathrm{E}v=\mathrm{E}\,f/\mathrm{E}\,f^2\,. \tag{2.10}$$

This completes the proof for $k=2$. When $k>2$, $\sum_{i=1}^{k} p_i x_i$ can be written in the form: $q_k x+(1-q_k)y$, where $x=\sum_{i=1}^{k-1} q_i x_i$, $y=x_k$, $0<q_i<1$ for every i and $\sum_{i=1}^{k-1} q_i=1$. Hence, it is easy to see that the result follows by induction.

Remark 2.2. It is plausible that Theorem 2.2 holds without the assumption of independence but a proof appears to be hard at the moment. We note, however, that the result would follow readily if one could show that the functional $\mathrm{E}\, x^{-1}/\mathrm{E}\, x^{-2}$ is concave.

3. Application

Let x_i, y_i, $i=1, 2, \cdots, k$ be mutually independent random variables such that $x_i \sim N(\mu, \theta_i)$, $y_i/\theta_i \sim \chi^2_{m(i)}$ where μ, $\theta_i(>0)$, $i=1, 2, \cdots, k$ are unknown. The problem is to estimate μ. The setting is essentially same as that for the problem of estimating the common mean of several normal populations after reduction to minimal sufficient statistics. Let

$$\text{(3.1)} \qquad \hat{\mu}=\sum w_i x_i\,; \qquad \hat{\mu}_* = \sum\nolimits_* w_{i*} x_i$$

where $w_i = c_i y_i^{-1}/\sum c_i y_i^{-1}$; $w_{i*} = c_i y_i^{-1}/\sum_* c_i y_i^{-1}$; and $\sum$, $\sum_*$ stand for summation over $i \in \{1, 2, \cdots, k\}$ and $i \in \{1, 2, \cdots, k-1\}$ respectively.

Both $\hat{\mu}$ and $\hat{\mu}_*$ are unbiased for μ and comparison of $V(\hat{\mu})$ and $V(\hat{\mu}_*)$ is interesting. This was considered by Shinozaki [8] and Bhattacharya [3]. They showed that $V(\hat{\mu}) \leqq V(\hat{\mu}_*)$ if

$$\text{(3.2)} \qquad (m_k+2)/[2(m_i-4)] \leqq c_k/c_i \leqq 2(m_k-4)/(m_i+2)\,, \qquad i=1, 2, \cdots, k-1\,.$$

(Note that notations of the two papers just cited are somewhat different e.g. Shinozaki [8] writes f_i and $c_i f_i$ for our m_i and c_i respectively.) Here, we derive the necessary and sufficient condition. It turns out that the left part of (3.2) is unnecessary but the Shinozaki-Bhattacharya proof of sufficiency fails if this is omitted. We now prove

THEOREM 3.1. *Let* $\hat{\mu}$ *and* $\hat{\mu}_*$ *be as defined in* (3.1). *Assume that* $m_k \geqq 5$. *Then* $v(\hat{\mu}) \leqq V(\hat{\mu}_*)$ *for all* $\boldsymbol{\theta} \equiv (\theta_1, \theta_2, \cdots, \theta_k)$ *if and only if*

$$\text{(3.3)} \qquad c_k/c_i \leqq 2(m_k-4)/(m_i+2)\,, \qquad i=1, 2, \cdots, k-1$$

PROOF. We have

$$V(\hat{\mu}) = \mathrm{E}\, S\,; \qquad V(\hat{\mu}_*) = \mathrm{E}\, S_*$$

where $S=\sum \theta_i w_i^2$; $S_* = \sum_* \theta_i w_{i*}^2$. Hence, $V(\hat{\mu}) \leqq V(\hat{\mu}_*)$ for all $\boldsymbol{\theta}$ iff

$$\text{(3.4)} \qquad \mathrm{E}\,(S-S_*) \leqq 0 \qquad \text{for all } \boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_k)\,.$$

It can be seen that

$$\text{(3.5)} \qquad S-S_* = w_k^2(\theta_k + S_*) - 2 w_k S_*\,.$$

Let $\theta_*=1/\sum_* \theta_i^{-1}$ and note that $\theta_* \leqq S_*$ for all $\boldsymbol{\theta}$. Note also that $w_k \leqq 1$ and hence the right-hand side of (3.5) is nonincreasing in S_*. Hence

(3.6) $$S-S_* \leqq w_k^2(\theta_k+\theta_*)-2w_k\theta_* .$$

Hence (3.4) holds if

(3.7) $$(\theta_k+\theta_*)\,\mathrm{E}\, w_k^2 \leqq 2\theta_*\,\mathrm{E}\, w_k \qquad \text{for all } \boldsymbol{\theta}$$

Let $f=(1+\theta_k/\theta_*)w_k$. Then (3.7) is equivalent to

(3.8) $$2\,\mathrm{E} f/\mathrm{E} f^2 \geqq 1 \qquad \text{for all } \boldsymbol{\theta} .$$

It is easy to see that f can be written in the form

$$f=1/[p\sum_* (q_i d_i z_i/z_k)+1-p]$$

where

$$z_i=\theta_i/y_i : \qquad p=\theta_k/(\theta_k+\theta_*) : \qquad q_i=\theta_*/\theta_i : \qquad d_i=c_i/c_k .$$

Note that $z_i^{-1} \sim \chi^2_{m(i)}$ and by the Theorem 2.2

$$\mathrm{E} f/\mathrm{E} f^2 \geqq \mathrm{Min}\,(1, c\,\mathrm{E}\, g/\mathrm{E}\, g^2)$$

where $c=\mathrm{E}\, z_k/\mathrm{E}\, z_k^2=m_k-4$; $g=1/\sum_* q_i d_i z_i$. Hence (3.8) holds if

(3.9) $$2c\,\mathrm{E}g/\mathrm{E}g^2 \geqq 1 \qquad \text{for all } \boldsymbol{q}=(q_1, q_2, \cdots, q_{k-1}) .$$

Again by Theorem 2.2

$$\mathrm{E}\, g/\mathrm{E}\, g^2 \geqq \mathrm{Min}\,\{a_i d_i :\ 1 \leqq i \leqq k-1\}$$

where $a_i=\mathrm{E}\, z_i^{-1}/\mathrm{E}\, z_i^{-2}=1/(m_i+2)$. Hence (3.9) holds if

$$2ca_i d_i \geqq 1 \qquad \text{for all } i=1, 2, \cdots, k-1$$

which is equivalent to (3.3). Thus, we have proved the sufficiency of the condition (3.3).

To prove the necessity observe that for every fixed $r<k$, as $\theta_i \to \infty$ for all $i \neq r$, we have $S_*-\theta_* \to 0$ almost sure and hence equality in (3.6) almost sure. This implies that (3.4) holds only if (3.8) holds in the limit as $\theta_i \to \infty$ for all $i \neq r$. But $\theta_i \to \infty$ for all $i \neq r \Longrightarrow p \to 1$, $q_r \to 1$, $q_i \to 0$ for $i \neq r \Longrightarrow \mathrm{E} f/\mathrm{E} f^2 \to ca_r d_r$. This proves the necessity of (3.3) and hence the proof is complete.

INDIAN STATISTISTICAL INSTISTUTE

References

[1] Bahadur, R. R. and Bickel, P. J. (1968). Substitution in conditional expectation, *Ann. Math. Statist.*, **39**, 377-378.

[2] Bhattacharya, C. G. (1978). Yates type estimators of a common mean, *Ann. Inst. Statist. Math.*, **30**, A, 407-414.

[3] Bhattacharya, C. G. (1979). A note on estimating the common mean of k normal populations, *Sankhyā*, B, **40**, 272-275.

[4] Brown, L. D. and Cohen, A. (1974). Point and confidence estimation of a common mean and recovery of inter-block information, *Ann. Statist.*, **2**, 963-976.

[5] Cohen, A. (1975). Personal communication.

[6] Hardy, G. H., Littlewood, J. E. and Polya, G. (1952). *Inequalities* (2nd ed.), Cambridge.

[7] Kimball, B. F. (1951). On dependent tests of significance in the analysis of variance, *Ann. Math. Statist.*, **22**, 600-602.

[8] Shinozaki, N. (1978). A note on estimating the common mean of K normal distributions and Stein problem, *Commun. Statist.-Theor. Meth.*, A, **7**, 1421-1432.