Sankhyā : The Indian Journal of Statistics
San Antonio Conference: selected articles
2002, Volume 64, Series A, Pt. 2, pp 282-292

AN INEQUALITY FOR THE PITMAN ESTIMATORS RELATED TO THE STAM INEQUALITY

By ABRAM KAGAN
University of Maryland, College Park, USA

SUMMARY. An inequality is proved for the variances of the Pitman estimators of a location parameter θ based on samples of a fixed size $n \geq 2$ from populations $F_1(x-\theta), F_2(x-\theta)$ and $F(x-\theta), F = F_1 * F_2$.

The inequality is a natural small sample version of the Stam inequality $1/I \geq 1/I_1 + 1/I_2$ where I_1, I_2 and I are, respectively, the Fisher information on θ contained in an observation $X_1 \sim F_1(x-\theta)$, $X_2 \sim F_2(x-\theta)$ and $X \sim F(x-\theta)$. Some related inequalities are proved.

1. Introduction

Let X_1, X_2 be independent random variables with distribution functions $F_1(x)$ and $F_2(x)$, respectively. Set $X = X_1 + X_2$ and $F = F_1 * F_2$. If F_1, F_2 are absolutely continuous with densities f_1, f_2 (then F is also absolutely continuous with density f) and the Fisher information

$$I_1 = \int (f_1'/f_1)^2 f_1 \, dx, I_2 = \int (f_2'/f_2)^2 f_2 \, dx,$$

in X_1 and X_2 (on a location parameter) is finite, then the Fisher information

$$I = \int (f'/f)^2 f dx$$

in X is also finite and the Stam inequality (Stam (1959)), see also (Blahut (1987), Ch. 7) puts the following upper bound for I:

$$1/I \geq 1/I_1 + 1/I_2 \tag{1}$$

Paper received October 2000; revised October 2001.
AMS (2000) subject classification. Primary 62F11; secondary 62B10.
Keywords and phrases. Pitman estimator, Fisher information, Stam inequality.

(which is plainly much stronger than the trivial inequality $I \le \min(I_1, I_2)$). Suppose now that $(x'_1, \dots, x'_n)$ is a sample of size n from population $F_1(x-\theta)$, θ being a parameter, $(x''_1, \dots, x''_n)$ is a sample from $F_2(x-\theta)$, and $(x_1, \dots, x_n)$ is a sample from $F(x-\theta)$. Let $\hat{t}'_n = \hat{t}'_n(x'_1, \dots, x'_n)$ be the Pitman estimator of θ with respect to the quadratic loss function, based on $(x'_1, \dots, x'_n)$, and let $\hat{t}''_n = \hat{t}''_n(x''_1, \dots, x''_n)$ and $\hat{t}_n = \hat{t}_n(x_1, \dots, x_n)$ be the Pitman estimators based on $(x''_1, \dots, x''_n)$ and $(x_1, \dots, x_n)$, respectively (see, e.g., Kagan, Linnik, Rao (1973), Ch. 7).

Under rather mild conditions on F_1, F_2, as $n \to \infty$,

$$\operatorname{var}(\hat{t}'_n) = \frac{1}{nI_1}(1+o(1)), \operatorname{var}(\hat{t}''_n) = \frac{1}{nI_2}(1+o(1)), \operatorname{var}(\hat{t}_n) = \frac{1}{nI}(1+o(1)) \quad (2)$$

(see Ibragimov and Has'minskii (1981), example to Th. 3.1)
so that for large n

$$\operatorname{var}(\hat{t}_n) \ge \operatorname{var}(\hat{t}'_n) + \operatorname{var}(\hat{t}''_n). \quad (3)$$

It turns out that under the only condition

$$\int x^2\, dF_1 < \infty, \int x^2\, dF_2 < \infty \quad (4)$$

the inequality (3) holds for *any* $n \ge 2$. This is the main result of the paper proved in Section 2. The proof of (3) for small n differs completely from Stam's proof of (1). In Section 3, a polynomial version of (1) is proved as well as (3) for the polynomial Pitman estimator of a location parameter introduced in Kagan (1966).

2. Main Result: Inequality (3) for Small n

Our only assumption is (4). Then for $F = F_1 * F_2$

$$\int x^2 dF < \infty. \quad (5)$$

Let $x_1, \dots, x_n$ be a sample of size $n \ge 2$ from $F(x-\theta)$. The Pitman estimator of θ, $\hat{t}_n = \hat{t}_n(x_1, \dots, x_n)$ with respect to the quadratic loss is, by definition, the minimum variance equivariant estimator.

Since the variance of any equivariant estimator is constant in θ,

$$\hat{t}_n = \arg\min_{t_n} \operatorname{var}(t_n) \quad (6)$$

where the variance is calculated when $\theta = 0$ and the minimum is taken over the class of equivariant estimators t_n, i. e., estimators subject to

$$t_n(x_1 + c, \dots, x_n + c) = c + t_n(x_1, \dots, x_n), \forall c \text{ real}.$$

It is easily seen that under (5), $\hat{t}_n$ can be written in the form

$$\hat{t}_n = \bar{x} - E(\bar{x}|x_1 - \bar{x}, \dots, x_n - \bar{x}) \tag{7}$$

where $\bar{x}$ is the sample mean and the expectation symbol E, here and in what follows, relates to $\theta = 0$. Actually, $\hat{t}_n$ can be represented as

$$\hat{t}_n = \theta + \bar{x} - E(\bar{x}|x_1 - \bar{x}, \dots, x_n - \bar{x}) \tag{8}$$

where $(x_1, \dots, x_n)$ is a sample from $F(x)$. Thus, in calculating $\mathrm{var}(\hat{t}_n)$ one may assume the sample $(x_1, \dots, x_n)$ drawn from $F(x)$.

Since $F = F_1 * F_2$, the sample elements $(x_1, \dots, x_n)$ have the form

$$x_i = x_i' + x_i'',\ i = 1, \dots, n \tag{9}$$

where $(x_1', \dots, x_n')$ is a sample from $F_1(x)$, $(x_1'', \dots, x_n'')$ is a sample from $F_2(x)$ and $x_1', \dots, x_n''$ are independent. Similarly to (8), the Pitman estimator of θ based on a sample from $F_1(x - \theta)$ can be represented as

$$\hat{t}_n' = \theta + \bar{x}' - E(\bar{x}'|x_1' - \bar{x}', \dots, x_n' - \bar{x}')$$

and the Pitman estimator based on a sample from $F_2(x - \theta)$ can be represented as

$$\hat{t}_n'' = \theta + \bar{x}'' - E(\bar{x}''|x_1'' - \bar{x}'', \dots, x_n'' - \bar{x}'').$$

Theorem 1 *Under the assumption (5), for any $n \geq 2$*

$$\mathrm{var}(\hat{t}_n) \geq \mathrm{var}(\hat{t}_n') + \mathrm{var}(\hat{t}_n''). \tag{10}$$

Proof. We have from (8)

$$\begin{aligned}
\mathrm{var}(\hat{t}_n) &= \mathrm{var}(\bar{x}) - \mathrm{var}\{E(\bar{x}|x_1 - \bar{x}, \dots, x_n - \bar{x})\} \\
&= \mathrm{var}(\bar{x}') + \mathrm{var}(\bar{x}'') - E\{E(\bar{x}|x_1 - \bar{x}, \dots, x_n - \bar{x})\}^2, \\
\mathrm{var}(\hat{t}_n') &= \mathrm{var}(\bar{x}') - \mathrm{var}\{E(\bar{x}'|x_1' - \bar{x}', \dots, x_n' - \bar{x}')\}, \\
\mathrm{var}(\hat{t}_n'') &= \mathrm{var}(\bar{x}'') - \mathrm{var}\{E(\bar{x}''|x_1'' - \bar{x}'', \dots, x_n'' - \bar{x}'')\}.
\end{aligned}$$

Since

$$x_1 - \bar{x} = x_1' - \bar{x}' + x_1'' - \bar{x}'', \ldots, x_n - \bar{x} = x_n' - \bar{x}' + x_n'' - \bar{x}'', \tag{11}$$

the σ-algebra generated by $(x_1 - \bar{x}, \ldots, x_n - \bar{x})$, $\sigma(x_1 - \bar{x}, \ldots, x_n - \bar{x}) = \sigma$, is a subalgebra of $\tilde{\sigma} = \sigma(x_1' - \bar{x}', x_1'' - \bar{x}'', \ldots, x_n' - \bar{x}', x_n'' - \bar{x}'')$. That is why

$$E(\bar{x}|\sigma) = E\{E(\bar{x}|\tilde{\sigma})|\sigma\}$$

and by virtue of a well known property of the conditional expectation (or of the projection),

$$E\{E(\bar{x}|\sigma)\}^2 \le E\{E(\bar{x}|\tilde{\sigma})\}^2.$$

Another known property of the conditional expectation is that if for a random variable U with $E|U| < \infty$ and arbitrary random elements V, W, the pair (U, V) is independent of W then

$$E(U|V, W) = E(U|V). \tag{12}$$

The relation (12) can be called the conditional constancy of regression of U on W given V. Actually, the random variables U and W are conditionally independent given V (and not only independent), i.e., for any $g(U)$ and $h(W)$ with finite expectations

$$E\{g(U)h(W)|V\} = E\{g(U)|V\}E\{h(W)|V\}.$$

In Meyer (1966), Th. 51 it is shown that the conditional independence of U and W given V is equivalent to that

$$E\{g(U)|V, W\} = E\{g(U)|V\}$$

for any $g(U)$ with $E|g(U)| < \infty$ of which (12) is a special case. Applying (12) first with

$$U = \bar{x}', V = (x_1' - \bar{x}', \ldots, x_n' - \bar{x}'), W = (x_1'' - \bar{x}'', \ldots, x_n'' - \bar{x}'')$$

and second with

$$U = \bar{x}'', V = (x_1'' - \bar{x}'', \ldots, x_n'' - \bar{x}''), W = (x_1' - \bar{x}', \ldots, x_n' - \bar{x}')$$

we get

$$E\{E(\bar{x}|x_1 - \bar{x}, \ldots, x_n - \bar{x})\}^2 \le E\{E(\bar{x}'|x_1 - \bar{x}', \ldots, x_n' - \bar{x}') + E(\bar{x}''|x_1'' - \bar{x}'', \ldots, x_n'' - \bar{x}'')\}^2 = E\{E(\bar{x}'|x_1 - \bar{x}', \ldots, x_n' - \bar{x}')\}^2 + E\{E(\bar{x}''|x_1'' - \bar{x}'', \ldots, x_n'' - \bar{x}'')\}^2$$

since $(x_1'-\bar{x}',\dots,x_n'-\bar{x}_n')$ and $(x_1''-\bar{x}'',\dots,x_n''-\bar{x}_n'')$ are independent. Hence,

$$\begin{aligned}\operatorname{var}(\hat{t}_n) \ge \operatorname{var}(\bar{x}') + \operatorname{var}(\bar{x}'') - \operatorname{var}\{E(\bar{x}'|x_1' - \bar{x}',\dots,x_n' - \bar{x}')\} \\ -\operatorname{var}(E(x''|x_1'' - \bar{x}'',\dots,x_n'' - \bar{x}'')\} = \operatorname{var}(\hat{t}_n') + \operatorname{var}(\hat{t}_n''). \qquad \square\end{aligned}$$

Section 4 contains some remarks on the equality sign in (10).

3. Polynomial Versions of (1) and (3)

To state the polynomial version of (1), we need first to define the polynomial score. For the case of a general (scalar- or vector-valued) parameter we refer to Kagan (1976). Here we discuss briefly the case of a location parameter.

Assume that for some integer $k \ge 1$ and $F = F_1 * F_2$,

$$\int x^{2k} dF(x) < \infty. \tag{13}$$

Denote by $L^2(F)$ the Hilbert space of all functions $h(x)$ with $\int |h|^2 dF < \infty$ with the standard inner product

$$(h_1, h_2) = \int h_1(x) h_2(x) dF(x).$$

Due to (13), any polynomial $q(x)$ of degree $\le k$ belongs to $L^2(F)$. Denote by Λ_k the subspace of $L^2(F)$ of all polynomials of degree $\le k$ and let $\hat{E}(\cdot|\Lambda_k)$ be the projection operator into Λ_k from $L^2(F)$.

Suppose for a moment that $F(x)$ is absolutely continuous with density $f(x)$ such that $\int (f'/f)^2 f dx < \infty$. In this case, the Fisher score for $F(x-\theta)$,

$$J(x,\theta) = J(x-\theta) = -\frac{f'(x-\theta)}{f(x-\theta)}$$

is well defined and in $L^2(F)$. Remind, in passing, that the classical Fisher score $J(X,\theta)$ for a family $\{f(x;\theta)\}$ of densities depending on a scalar parameter θ is characterized by the property

$$E_\theta\{J(X,\theta)h(X)\} = \frac{d}{d\theta} E_\theta\{h(X)\}$$

for every $h(X)$ with finite

$$E_\theta |h(X)|^2 = \int |h(x)|^2 f(x;\theta) dx.$$

Set

$$\hat{J}(x) = \hat{E}\{(-f'/f)|\Lambda_k\}.$$

The polynomial

$$\hat{J}(x,\theta) = \hat{J}(x-\theta)$$

is called the polynomial Fisher score for the family $\{f(x-\theta), \theta \in \mathbb{R}\}$. One can easily see that $\hat{J}(X)$ is the only polynomial of degree $\leq k$ in random variable X with distribution $F(x)$ satisfying the condition

$$E\{\hat{J}(X)X^l\} = -lE(X^{l-1}), i = 1, \ldots, k;\ E\{\hat{J}(X)\} = 0$$

or, equivalently,

$$E\{\hat{J}(X)q(X)\} = -E\{q'(X)\} \tag{14}$$

for any $q \in \Lambda_k$. If E_θ denotes the expectation with respect to $F(x-\theta)$, (14) is plainly equivalent to

$$E_\theta\{\hat{J}(X-\theta)q(X)\} = \frac{d}{d\theta}E_\theta\{q(X)\}$$

for any $q \in \Lambda_k$.

The relation (14) *does not require even the absolute continuity of* $F(x)$ and may be taken as a definition of the polynomial Fisher score for the family $\{F(x-\theta)\}$ (see Kagan (1976).) In absolutely the same way as $L^2(F)$ and Λ_k one may define the Hilbert spaces $L^2(F_1)$ and $L^2(F_2)$ and their suspaces $\Lambda_{1,k}$ and $\Lambda_{2,k}$. Then, using (14) one defines the polynomial scores $\hat{J}_1(x_1-\theta)$ and $\hat{J}_2(x_2-\theta)$ for the families $\{F_1(x_1-\theta), \theta \in \mathbb{R}\}$ and $\{F_2(x_2-\theta), \theta \in \mathbb{R}\}$, respectively. Set now

$$\hat{I} = E|\hat{J}|^2 = \int |\hat{J}(x)|^2 dF(x).$$

One can interpret $\hat{I}$ as the Fisher information on θ contained in an observation X when only the first $2k$ moments known:

$$\alpha_l(\theta) = E_\theta(X^l) = \int x^l dF(x-\theta), l = 1, \ldots, 2k,$$

but not the whole distribution $F(x-\theta)$; $\hat{I}_1, \hat{I}_2$ can be interpreted similarly, with F_1, F_2 substituted for F (see Kagan (1976).)

The following result is a straightforward analog of Stam's inequality (1).

THEOREM 2 *Under the condition (13)*

$$1/\hat{I} \geq 1/\hat{I}_1 + 1/\hat{I}_2. \tag{15}$$

Let X_1, X_2 be independent random variables with distributions $F_1(x), F_2(x)$, respectively. Denote by $L^2 = L^2(F_1, F_2)$ the Hilbert space of random variables $h(X_1, X_2)$ with

$$E|h(X_1, X_2)|^2 = \int\int |h(x_1, x_2)|^2 dF_1(x_1) dF_2(x_2) < \infty$$

and the inner product

$$(h_1, h_2) = E\{h_1(X_1, X_2) h_2(X_1, X_2)\}.$$

The spaces $L^2(F), L^2(F_1), L^2(F_2)$ introduced above are (closed) subspaces of L^2: $L^2(F)$ consists of all random variables $h(X_1 + X_2)$ with finite second moment, $L^2(F_1)$ consists of all functions of X_1 and $L^2(F_2)$ of all functions of X_2 with finite second moment. Also, $\Lambda_k, \Lambda_{1,k}, \Lambda_{2,k}$ are (closed) subspaces of L_2. The projection operators appearing in the next lemma act in L^2.

LEMMA 1

$$\hat{E}(\hat{J}_1|\Lambda_k) = \hat{J},\ \hat{E}(\hat{J}_2|\Lambda_k) = \hat{J}. \tag{16}$$

PROOF. Let us calculate

$$(\hat{J}_1, (X_1 + X_2)^l) = E\{\hat{J}_1(X_1)(X_1 + X_2)^l\} = \sum_{i=0}^{l} \binom{l}{i} E\{\hat{J}_1(X_1) X_1^i\} E(X_2^{l-i})$$

$$= -\sum_{i=1}^{l} \binom{l}{i} i E(X_1^{i-1}) E(X_2^{l-i}) = -l \sum_{i=1}^{l} \frac{(l-1)!}{(i-1)!(l-i)!} E(X_1^{i-1}) E(X_2^{l-i})$$

$$= -l \sum_{i=0}^{l-1} \frac{(l-1)!}{i!(l-1-i)!} E(X_1^i) E(X_2^{l-1-i}) = -lE\{(X_1 + X_2)^{l-1}\}$$

$$= (\hat{J}, (X_1 + X_2^l)), \tag{17}$$

where we have used independence of X_1, X_2 and (14). Since (17) holds for any $l, 0 \leq l \leq k$, it proves the first relation (16). The second is proved in the same way. □

PROOF OF THEOREM 2. By virtue of Lemma one has

$$\hat{E}(a_1 \hat{J}_1 + a_2 \hat{J}_2 | \Lambda_k) = (a_1 + a_2)\hat{J},\ \forall a_1, a_2 \text{ real}.$$

Since the norm of the projection of any element does not exceed that of the element itself,

$$E(a_1\hat{J}_1 + a_2\hat{J}_2)^2 \geq (a_1 + a_2)^2 E(\hat{J})^2$$

or

$$a_1^2\hat{I}_1 + a_2^2\hat{I}_2 \geq (a_1 + a_2)^2\hat{I}. \tag{18}$$

On choosing $a_1 = 1/\hat{I}_1, a_2 = 1/\hat{I}_2$, (18) leads to (15). $\square$

Turn now to the polynomial Pitman estimators introduced in Kagan (1966). Let now $(x_1', \ldots, x_n')$ be a sample from $F_1(x)$, $(x_1'', \ldots, x_n'')$ be a sample from $F_2(x)$ and $(x_1, \ldots, x_n)$ be a sample from $F(x), F = F_1 * F_2$. Denote by $\mathcal{L}^2$ the Hilbert space of all the functions $h(x_1', \ldots, x_n'')$ of independent random variables $x_1', \ldots, x_n''$ with finite second moment and the inner product

$$(h_1, h_2) = \int \ldots \int h_1(x_1', \ldots, x_n'') h_2(x_1', \ldots, x_n'') dF_1(x_1') \ldots dF_1(x_n')$$
$$dF_2(x_1'') \ldots dF_2(x_n'').$$

Denote by $M_{1,k}$ the space of all polynomials of degree $\leq k$ in the residuals $(x_1' - \bar{x}', \ldots, x_n' - \bar{x}')$ and by $M_{2,k}$ and M_k the similar spaces in $(x_1'' - \bar{x}'', \ldots, x_n'' - \bar{x}'')$ and $(x_1 - \bar{x}, \ldots, x_n - \bar{x})$, $x_i = x_i' + x_i''$, respectively. By $\mathcal{M}_k$ we denote the space of all polynomials of degree $\leq k$ in $2n$ residuals $(x_1' - \bar{x}', \ldots, x_n'' - \bar{x}'')$. Plainly, $M_k, M_{1,k}, M_{2,k}$ are subspaces of $\mathcal{M}_k$ which, in its turn, is a subspace of $\mathcal{L}^2$. The polynomial (of degree k) Pitman estimators of θ based on samples $(x_1, \ldots, x_n)$ from $F(x-\theta)$, $(x_1', \ldots, x_n')$ from $F_1(x-\theta)$ and $(x_1'', \ldots, x_n'')$ from $F_2(x-\theta)$ are

$$\begin{aligned} \hat{t}_{n,k} &= \bar{x} - \hat{E}(\bar{x}|M_k) \\ \hat{t}_{n,k}' &= \bar{x}' - \hat{E}(\bar{x}'|M_{1,k}) \\ \hat{t}_{n,k}'' &= \bar{x}'' - \hat{E}(\bar{x}''|M_{2,k}), \end{aligned} \tag{19}$$

respectively.
They can be represented as

$$\begin{aligned} \hat{t}_{n,k} &= \theta + \bar{x} - \hat{E}(\bar{x}|M_k), \\ \hat{t}_{n,k}' &= \theta + \bar{x}' - \hat{E}(\bar{x}'|M_{1,k}), \\ \hat{t}_{n,k}'' &= \theta + \bar{x}'' - \hat{E}(\bar{x}''|M_{2,k}), \end{aligned}$$

where the samples are from $F(x), F_1(x)$ and $F_2(x)$, respectively.

THEOREM 3 *Under the condition (13), for any $n \geq 2$*

$$var(\hat{t}_{n,k}) \geq var(\hat{t}'_{n,k}) + var(\hat{t}''_{n,k}). \tag{20}$$

PROOF. One has

$$\mathrm{var}(\hat{t}_{n,k}) = \mathrm{var}(\bar{x}) - \mathrm{var}(\hat{E}(\bar{x}|M_k)) \geq \mathrm{var}(\bar{x}) - \mathrm{var}(\hat{E}(\bar{x}|\mathcal{M}_k)).$$

By virtue of (11), any polynomial Q from $\mathcal{M}_k$ can be written as

$$Q = \sum A_{j_1 \ldots j_n}(x''_1 - \bar{x}'', \ldots, x''_n - \bar{x}'')(x'_1 - \bar{x}')^{j_1} \cdots (x'_n - \bar{x}')^{j_n}.$$

with the summation taken over $j_1, \ldots, j_n$ with $j_1 + \ldots + j_n \leq k$ whence

$$\hat{E}(Q|M_{1,k}) = \sum \bar{A}_{j_1 \ldots j_n}(x'_1 - \bar{x}')^{j_1} \cdots (x'_n - \bar{x}')^{j_n}$$

where $\bar{A}_{j_1 \ldots j_n} = E\{A_{j_1 \ldots j_n}(x''_1 - \bar{x}'', \ldots, x''_n - \bar{x}'')\}$. Since $\bar{x}'$ is independent of $(x''_1 - \bar{x}'', \ldots, x''_n - \bar{x}'')$,

$$E(\bar{x}'Q) = E\{\bar{x}'\hat{E}(Q|M_{1,k})\}$$

thus proving that

$$\hat{E}(\bar{x}'|\mathcal{M}_k) = \hat{E}(\bar{x}'|M_{1,k})$$

and, similarly,

$$\hat{E}(\bar{x}''|\mathcal{M}_k) = \hat{E}(\bar{x}''|M_{2,k}).$$

Hence,

$$\mathrm{var}\{\hat{E}(\bar{x}|\mathcal{M}_k)\} = \mathrm{var}\{\hat{E}(\bar{x}'|M_{1,k}) + \hat{E}(\bar{x}''|M_{2,k})\}$$

and

$$\mathrm{var}(\hat{t}_{n,k}) \geq \mathrm{var}(\bar{x}') + \mathrm{var}(\bar{x}'') - \mathrm{var}\{\hat{E}(\bar{x}'|M_{1,k})\}$$
$$-\mathrm{var}\{\hat{E}(\bar{x}''|M_{2,k})\} = \mathrm{var}(\hat{t}'_{n,k}) + \mathrm{var}(\hat{t}''_{n,k})$$

which is exactly (20). □

4. A Related Analytical Problem

An interesting open analytical problem is to find out when the equality sign holds in (3).

As to the Stam inequality (1), if I_1, I_2 are finite then the equality sign holds if and only if X_1 and X_2 are Gaussian (Stam (1959)). If X_1, X_2 are discrete random variables then the Fisher information on a location parameter contained in each of $X_1, X_2, X = X_1 + X_2$ is infinite (see Huber (1981), Ch. 4 for the definition of the Fisher information on a location parameter in an observation with arbitrary, not necessarily absolutely continuous, distribution.) In this case (3) becomes the trivial identity $0 = 0$.

The equality sign in (3) holds if and only if

$$E(\bar{x}' + \bar{x}''|x_1' - \bar{x}' + x_1'' - \bar{x}'', \dots, x_n' - \bar{x}' + x_n'' - \bar{x}'') = \\ E(\bar{x}'|x_1' - \bar{x}', \dots, x_n' - \bar{x}') + E(\bar{x}''|x_1'' - \bar{x}'', \dots, x_n'' - \bar{x}'') \qquad (21)$$

which is a Cauchy type equation

$$\Psi(u_1 + v_1, \dots, u_n + v_n) = \Psi_1(u_1, \dots, u_n) + \Psi_2(v_1, \dots, v_n) \qquad (22)$$

holding for $(u_1, \dots, v_n)$ subject to

$$u_1 + \dots + u_n = 0,\ v_1 + \dots + v_n = 0. \qquad (23)$$

However, in (21) $x_1', \dots, x_n''$ are random variables and the relation holds for almost all (and not necessarily all) their values. Had (22) with $n \geq 3$ held for all $(u_1, \dots, v_n)$ subject to (23) and were Ψ_1, Ψ_2 differentiable, one would get from (22) by differentiating both sides first with respect to u_i and second with respect to v_j that all the second derivatives of Ψ vanish and thus, on taking into account the symmetry of Ψ,

$$\Psi(w_1, \dots, w_n) = \psi(w_1) + \dots + \psi(w_n)$$

for some ψ. In this case, Ψ_1, Ψ_2 have the same form:

$$\Psi_1(u_1, \dots, u_n) = \psi_1(u_1) + \dots + \psi_1(u_n),\ \Psi_2(v_1, \dots, v_n) = \psi_2(v_1) + \dots + \psi_2(v_n)$$

and (22) reduces to

$$\psi(u + v) = \psi_1(u) + \psi_2(v)$$

which is the standard Cauchy equation. Thus, Ψ_1, Ψ_2 are linear. Due to the symmetry of $E(\bar{x}'|x_1' - \bar{x}', \dots, x_n' - \bar{x}')$ in $(x_1 - \bar{x}', \dots, x_n - \bar{x}')$, the linearity of $E(\bar{x}'|x_1' - \bar{x}', \dots, x_n' - \bar{x}')$ simply means that

$$E(\bar{x}'|x_1' - \bar{x}', \dots, x_n' - \bar{x}') = \text{const.} \qquad (24)$$

According to the KLR-theorem (Kagan, Linnik, Rao (1965)), (24) holds if and only if x_i' are Gaussian. Thus, it seems very plausible that under rather mild conditions on F_1, F_2 the equality sign in (3) is a characteristic property of Gaussian F_1, F_2.

Acknowledgement. My collaboration with C. R. Rao started thirty five years ago with the KLR-theorem. It gives me a special pleasure to dedicate this paper to C. R. Rao on the occasion of his 80-th birthday.

References

Blahut, R.E. (1987). *Principles and Practice of Information Theory*, Addison-Wesley, New York.

Huber, P.J. (1981). *Robust Statistics*, Wiley, New York.

Ibragimov, I.A. and Has'minskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*, Springer, New York.

Kagan, A.M., Linnik, Yu.V. and Rao, C.R. (1965). On a characterization of the normal law based on a property of the sample average. *Sankhyā Ser. A*, **27**, 405-406.

Kagan, A.M. (1966). On the estimation theory of location parameter. *Sankhyā Ser. A*, **28**, 335-352.

Kagan, A.M., Linnik, Yu.V. and Rao, C.R. (1973). *Characterization Problems in Mathematical Statistics*, Wiley, New York.

Kagan, A.M. (1976). Fisher information contained in a finite dimensional linear space and a correctly posed version of the method of moments. *Problems of Information Transmission*, **12**, 98-115.

Meyer, P.A. (1966). *Probability and Potentials*, Blaisdell.

Stam, A. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, **2**, 101-112.

Abram Kagan
Department of Mathematics (Statistics Program)
University of Maryland
College Park, MD 20742, USA
E-mail: amk@math.umd.edu