Publ. RIMS, Kyoto Univ.
Vol. 5 (1969), pp. 41-50

Some Inequalities on Means and Covariance

By

Shuzo IZUMI*

The purpose of this paper is, in brief, to show generalizations of Kantrovich's inequality:

(*) "If $0<h\leqq x_i\leqq H$, $p_i\geqq 0$ and $p_1+p_2+\cdots+p_n=1$, then $(p_1x_1+p_2x_2+\cdots+p_nx_n)(p_1/x_1+p_2/x_2+\cdots+p_n/x_n)\leqq(H+h)^2/4Hh$."

One such generalization is to take (*) for an estimate of the covariance of variables x and $1/x$:

$$C(x, 1/x)\geqq -(H-h)^2/4Hh\,.$$

We shall study the bound of the covariance $C(x, y)$ in general case. At the same time (*) can be seen as an estimate of the ratio of the arithmetical mean and the harmonic mean. In this direction G. T. Cargo and O. Shisha [1] have showed the best estimate for the ratio of means with degree r and degree s by the supremum H and infimum h of the variable. We shall show the best estimates for the difference, the ratio etc. of two "comparable" means by H, h and a mean, from which we can derive the results in [1]. As to the difference of arbitrary two means, we have rough estimates using the estimate for the covariance stated above.

The main results are stated in theorem 2, 3 and proposition 3. We use integral notation for means, but nothing essential is lost to restrict ourselves to finite cases.

I wish to thank Professor S. Hitotumatu who encouraged me to complete this study.

Received March 24, 1969.
Communicated by S. Hitotumatu.
* Mathematical Institute, Faculty of Science, Kyoto University.

1. Generalized means and absoluate inequalities.

Let $\Omega(\mathfrak{B}, \mu)$ be a completely additive measure space with measure 1. We take a monotone increasing continuous function f on an interval I. Then we can define the mean $\mathcal{M}_f(x)$ of $x(t)$ with respect to f for any function $x(t)$ on Ω satisfying the following conditions:

i) $x(t)$ has values in I,

ii) $f\circ x(t)$ is summable on Ω.

We put

$$\mathcal{M}_f(x) = f^{-1}\left[\int_\Omega f\circ x(t)d\mu\right].$$

We can easily verify the following

Proposition 1. i) *$\mathcal{M}_f(x)$ is contained in the convex hull of the essential range of $x(t)$. Consequently, if $x(t)$ is essentially equal to a constant c, $\mathcal{M}_f(x)=c$.*

ii) *If $x(t)\geqq y(t)$ (a.e.), $\mathcal{M}_f(x)\geqq\mathcal{M}_f(y)$. Furthermore if $\mu\{t\in\Omega\,|\,x(t)>y(t)\}>0$ then $\mathcal{M}_f(x)>\mathcal{M}_f(y)$.*

iii) *$\mathcal{M}_{af+b}(x)=\mathcal{M}_f(x)$ for $a\neq 0$.*

iv) *If $\{f_n(x)\}$ is uniformly convergent to $f(x)$, and $\{x_n(t)\}$ is convergent to $x(t)$ almost everywhere, and if there is a summable function $F(t)$ on Ω such that $|f_n\circ x_n(t)|\leqq F(t)$, then $\lim_{n\to\infty}\mathcal{M}_{f_n}(x_n)=\mathcal{M}_f(x)$.*

v) *Let T be a measure preserving transformation on Ω, then $\mathcal{M}_f(x\circ T)=\mathcal{M}_f(x)$.*

iv) *If g is an increasing continuous function on the convex hull of the essential range of $f\circ x$, $\mathcal{M}_{g\circ f}(x)=f^{-1}[\mathcal{M}_g(f\circ x)]$.*

As to iii), we see in next theorem that the converse is valid.

If Ω is finite, $\mathcal{M}_x(x)$ and $M_{\log x}(x)$ stand for the weighted arithmetical and geometrical mean respectively. We shall often denote them by $\mathcal{A}(x)$ and $\mathcal{G}(x)$. $\mathcal{A}(x)$ is defined just for summable functions on Ω. It is obvious that $\mathcal{M}_f(x)=f^{-1}[\mathcal{A}(f\circ x)]$.

Let $\mathfrak{F}=\mathfrak{F}(I)$ be the totality of the monotone increasing continuous functions on I, and $\mathfrak{S}=\mathfrak{S}(I)$ the totality of the measurable functions on Ω whose ranges are relatively compact in I. Then the mean $\mathcal{M}_f(x)$ is defined for any $f(x)\in\mathfrak{F}$ and $x(t)\in\mathfrak{S}$. For $f, g\in\mathfrak{F}$ we define $f\succ g$ when $\mathcal{M}_f(x)\geqq\mathcal{M}_g(x)$ for any $x(t)\in\mathfrak{S}$. This is a pseudo-order relation in $\mathfrak{F}$.

Theorem 1. *If $\Omega(\mathfrak{B}, \mu)$ is not trivial, that is if $\mathfrak{B}$ has a set ω with measure $\lambda\neq 0, 1$, then for $f, g\in\mathfrak{F}$, $f\succ g$ if and only if $f\circ g^{-1}$ is a convex function. Whence $\mathcal{M}_f(x)=\mathcal{M}_g(x)$ for any $x(t)\in\mathfrak{S}$ if and only if $g=af+b$ for $a\neq 0$.*

This is easily proved by approximating x with step-functions. We have defined the relation ($\succ$) regarding $\mathcal{M}_f$ only as a functional on $\mathfrak{S}$, but if $f\succ g$ so far as $\mathcal{M}_f(x)$ and $\mathcal{M}_g(x)$ make sense the inequality $\mathcal{M}_f(x)\geqq\mathcal{M}_g(x)$ still holds.

2. Estimate of covariance.

Let $x(t)$, $y(t)$ be essentially bounded measurable functions on Ω. We use the following notations:

$$H = \text{ess.}\sup x(t),\ h = \text{ess.}\inf x(t),\quad D(x) = H-h\,,$$
$$K = \text{ess.}\sup y(t),\ k = \text{ess.}\inf y(t),\quad D(y) = K-k\,,$$
$$C(x, y) = \mathcal{A}(xy)-\mathcal{A}(x)\mathcal{A}(y) = \mathcal{A}\{[x-\mathcal{A}(x)]\cdot[y-\mathcal{A}(y)]\}\,,$$
$$\mathcal{V}(x) = C(x, x)\,.$$

$C(x, y)$ is the covariance of x and y. $\mathcal{V}(x)$ is the variance of x.

Lemma. *If $P, Q\geqq 0$; $p, q\leqq 0$,*

i) $\min(PQ, pq)\leqq(P-p)(Q-q)/4$,

ii) $\max(Pq, pQ)\geqq-(P-p)(Q-q)/4$.

In either case equality holds if and only if $P+p=Q+q=0$ or $P=p=0$ or $Q=q=0$.

Proof. i) Assume that

$$4PQ>(P-p)(Q-q),\ 4pq>(p-P)(Q-q)\,,$$

then

$$(-4Pp)(-4Qq)>(P-p)^2(Q-q)^2\,.$$

This contradicts to the obvious inequalities:

$$(*)\qquad (P-p)^2\geqq-4Pp(\geqq 0)\,,\quad (Q-q)^2\geqq-4Qq(\geqq 0)\,.$$

If $P=p=0$ or $Q=q=0$, equality holds. Otherwise, $(P-p)(Q-q) > 0$ and taking (*) into account we have the equality condition $4PQ=4pq=(P-p)(Q-q)$. This implies

$$P+p=Q+q=0.$$

ii) Replace Q, q respectively $-q, -Q$ and apply i).

Theorem 2. $|C(x, y)| \equiv |\mathcal{A}(xy) - \mathcal{A}(x)\mathcal{A}(y)| \leqq D(x)D(y)/4$. *Equality holds in either case* i), ii) *or* iii):

i) *x or y is essentially constant; then* $C(x, y)=D(x)D(y)/4=0$.

ii) *There exist $B, B' \in \mathfrak{B}$ such that*

$$\mu(B) = \mu(B') = 1/2,$$
$$x(t) = H,\ y(t) = K \text{ on } B,$$
$$x(t) = h,\ y(t) = k \text{ on } B';$$

then $C(x, y)=D(x)D(y)/4$.

iii) *There exist $B, B' \in \mathfrak{B}$ such that*

$$\mu(B) = \mu(B') = 1/2,$$
$$x(t) = h,\ y(t) = K \text{ on } B,$$
$$x(t) = H,\ y(t) = k \text{ on } B';$$

then $C(x, y)=-D(x)D(y)/4$.

Proof. Let $x'=x-\mathcal{A}(x)$, $y'=y-\mathcal{A}(y)$. Then $\mathcal{A}(x')=\mathcal{A}(y')=0$ and

$$C(x, y) = C(x', y') = C(x'-\alpha, y'-\beta)$$
$$= \mathcal{A}[(x'-\alpha)(y'-\beta)]-\alpha\beta.$$

Put $\alpha=H'=\text{ess. sup}\, x'(t)$, $\beta=K'=\text{ess. sup}\, y'(t)$, then

$$C(x, y) \geqq -H'K'.$$

Similarly we obtain

$$C(x, y) \geqq -h'k',$$
$$C(x, y) \leqq -H'k',\ -h'K'.$$

On the other hand $\mathcal{A}(x')=\mathcal{A}(y')=0$ means $H', K' \geqq 0$; $h', k' \leqq 0$. So we can apply lemma:

$$\mathcal{C}(x, y) \geqq -\min(H'K', h'k') \geqq -D(x')D(y')/4 = -D(x)D(y)/4 .$$
$$\mathcal{C}(x, y) \leqq -\max(H'k', h'K') \leqq D(x')D(y')/4 = D(x)D(y)/4 .$$

As to the equality conditions, i) is trivial. So we assume

$$H-h = H'-h' > 0, \quad K-k = K'-k' > 0 .$$

Equality

$$\min(H'K', h'k') = D(x)D(y)/4$$

holds if and only if $H'+h'=K'+k'=0$. And then equality

$$\mathcal{C}(x, y) = -\min(H'K', h'k') = -H'K'$$

holds if and only if

$$(x'-H')(y'-K') = 0$$

almost everywhere, that is,

$$\mu[\{t \mid x'(t) = H'\} \cup \{t \mid y'(t) = K'\}] = 1 .$$

Because of $H'+h'=K'+k'=0$ and $\mathcal{A}(x')=\mathcal{A}(y')=0$, this is equivalent to the following condition:

$$\begin{cases} \mu\{t \mid x'(t) = H'\} = \mu\{t \mid y'(t) = K'\} = 1/2 , \\ x'(t) = h', \text{ almost everywhere on } \Omega - \{t \mid y'(t) = K'\} , \\ y'(t) = k', \text{ almost everywhere on } \Omega - \{t \mid x'(t) = H'\} . \end{cases}$$

Then we are only necessary to put

$$B = \{t \mid x'(t) = h', y'(t) = K'\}, \; B' = \{t \mid x'(t) = H', y'(t) = k'\} ,$$

and we have the equality condition iii). ii) is similarly proved.

Corollary. i) *If* $\mathcal{A}(xy) = \mathcal{A}(wz)$,

$$|\mathcal{A}(x)\mathcal{A}(y) - \mathcal{A}(w)\mathcal{A}(z)| \leqq [D(x)D(y) + D(w)D(z)]/4 .$$

ii) *If* $h>0$,

$$\left|\mathcal{A}\left(\frac{y}{x}\right) - \frac{\mathcal{A}(y)}{\mathcal{A}(x)}\right| \leqq \frac{D(x)D(y/x)}{4\mathcal{A}(x)} .$$

iii) $|\mathcal{V}(\sum_{i=1}^{n} x_i) - \sum_{i=1}^{n} \mathcal{V}(x_i)| \leqq \sum_{i \neq j} D(x_i)D(x_j)/4 .$

ii) is practical in calculating the approximate value of the mean of quotients.

3. Estimates of differences of means.

Now we treat the estimate of the difference of $\mathcal{M}_f(x)$ and $\mathcal{M}_g(x)$ using the result of section 2.

Proposition 2. *Let $f(x)$ be a continuous function on an interval containing the essential range of $x(t)$. Then*

$$|\mathcal{A}(f\circ x)-f[\mathcal{A}(x)]|\leqq D(x)(L-l)/4\leqq D(x)D(df/dx)/4,$$

where df/dx is any of the right or left, upper or lower differential coefficient of $f(x)$. And if x is essentially equal to a constant, $L=l=0$, otherwise

$$L=\operatorname*{ess.sup}_{t\in\Omega,\ x(t)\neq\mathcal{A}(x)}\frac{f\circ x(t)-f[\mathcal{A}(x)]}{x(t)-\mathcal{A}(x)},$$
$$l=\operatorname*{ess.inf}_{t\in\Omega,\ x(t)\neq\mathcal{A}(x)}\frac{f\circ x(t)-f[\mathcal{A}(x)]}{x(t)-\mathcal{A}(x)}.$$

N. B. Here f is not required to be monotone increasing.

Proof. $\mathcal{A}(f\circ x)-f[\mathcal{A}(x)]=\mathcal{A}\{f\circ x-f[\mathcal{A}(x)]\}$

$$=\mathcal{A}\left\{[x-\mathcal{A}(x)]\cdot\frac{f\circ x-f[\mathcal{A}(x)]}{x-\mathcal{A}(x)}\right\}.$$

For such t as $x(t)=\mathcal{A}(x)$, we define

$$\frac{f\circ x(t)-f[\mathcal{A}(x)]}{x(t)-\mathcal{A}(x)}=\frac{L+l}{2}.$$

Taking account of the equality: $\mathcal{A}[x(t)-\mathcal{A}(x)]=0$, we have only to apply theorem 2 to the variables

$$x(t)-\mathcal{A}(x)\quad\text{and}\quad\frac{f\circ x(t)-f[\mathcal{A}(x)]}{x(t)-\mathcal{A}(x)}.$$

Proposition 3. *For $f(x), g(x)\in\mathfrak{F}(I)$ and $x(t)\in\mathfrak{S}(I)$, we have*

$$|f[\mathcal{M}_f(x)]-f[\mathcal{M}_g(x)]|\leqq D(g\circ x)(L-l)/4$$
$$\leqq D(g\circ x)D(df/dg)/4,$$

where df/dg denotes any of the right or left, upper or lower differential coefficient of the function $f\circ g^{-1}$ on $g(I)$. And if x is essentially equal to a constant, $L=l=0$, otherwise

$$L = \operatorname*{ess.\,sup}_{g\circ x(t)\neq g[\mathcal{M}_g(x)]} \frac{f\circ x(t)-f[\mathcal{M}_g(x)]}{g\circ x(t)-g[\mathcal{M}_g(x)]},$$
$$l = \operatorname*{ess.\,inf}_{g\circ x(t)\neq g[\mathcal{M}_g(x)]} \frac{f\circ x(t)-f[\mathcal{M}_g(x)]}{g\circ x(t)-g[\mathcal{M}_g(x)]}.$$

Proof. We apply proposition 2 to $e=f\circ g^{-1}\in\mathfrak{F}[g(I)]$ and $y=g\circ x\in\mathfrak{S}[g(I)]$, then

$$\begin{aligned}|f[\mathcal{M}_f(x)]-f[\mathcal{M}_g(x)]| &= |\mathcal{A}(e\circ y)-e[\mathcal{A}(y)]| \\ &\leqq D(y)(L-l)/4\,.\end{aligned}$$

Corollary 1.
$$\begin{aligned}&|\mathcal{M}_f(x)-\mathcal{M}_g(x)| \\ &\leqq D(g\circ x)(L-l)/4[\inf_{\mathcal{M}_f\leqq x\leqq\mathcal{M}_g} f'(x)] \\ &\leqq D(g\circ x)D(df/dg)/4[\inf_{\mathcal{M}_f\leqq x\leqq\mathcal{M}_g} f'(x)]\,.\end{aligned}$$

Corollary 2. *If* $0<h<H$,

i) $0<\mathcal{A}(x)-\mathcal{G}(x)<\dfrac{(H-h)}{4}\log\dfrac{H}{h}$,

ii) $1<\dfrac{\mathcal{A}(x)}{\mathcal{G}(x)}<\exp\dfrac{(H-h)^2}{4Hh}$.

4. The best estimates of difference of two comparable means.

The results in section 3 are rather simple, but not the best estimates, for equality signs do not hold if $h\neq H$.

If we restrict ourselves to the case $f>g$, we have the most accurate estimate for the difference of $\mathcal{M}_f(x)$ and $\mathcal{M}_g(x)$.

Lemma 2. *Assume that* Ω *has the following property*:

(*) *For arbitrary* $\omega\in\mathfrak{B}$ *with measure* η *and arbitrary* η' *satisfying* $0\leqq\eta'\leqq\eta$, $\mathfrak{B}$ *contains a set* $\omega'\subset\omega$ *with measure* η'.

Take $e, f, g, \varphi\in\mathfrak{F}(I)$ *satisfying* $f>e>g$. *Then*

$$\varphi[\mathcal{M}_f(x)]-\varphi[\mathcal{M}_g(x)]$$

attains its maximum value $\Phi(H, h, m)$ *under the conditions*;

$$\operatorname{ess.\,sup} x = H\,,\quad \operatorname{ess.\,inf} x = h\,,\quad \mathcal{M}_e(x) = m\,,$$

for such x *as takes only* H *and* h *as its values.*

Proof. It is sufficient to prove to restrict x to vary in step-functions. If x takes $L \neq H, h$ as a value on a set ω with measure $\eta > 0$, define a new step-function x^* as follows. Divide ω into measurable sets ω_1, ω_2 with measure η_1, η_2 respectively such that

$$e(H)\eta_1 + e(h)\eta_2 = e(L)\eta .$$

And let

$$x^* = \begin{cases} H & \text{on } \omega_1 , \\ h & \text{on } \omega_2 , \\ x & \text{on } \omega^c . \end{cases}$$

Then we have

$$\mathcal{M}_e(x^*) = \mathcal{M}_e(x) , \quad \mathcal{M}_f(x^*) \geqq \mathcal{M}_f(x) , \quad \mathcal{M}_g(x^*) \leqq \mathcal{M}_g(x) .$$

So that,

$$\varphi[\mathcal{M}_f(x^*)] - \varphi[\mathcal{M}_g(x^*)] \geqq \varphi[\mathcal{M}_f(x)] - \varphi[\mathcal{M}_g(x)] .$$

Repeating such a modification, we arrive at a variable x as stated above.

Theorem 3. *For $e, f, g, \varphi \in \mathfrak{F}$ satisfying $f \succ e \succ g$*

$$\varphi[\mathcal{M}_f(x)] - \varphi[\mathcal{M}_g(x)] \leqq \Phi[H, h, \mathcal{M}_e(x)] ,$$

where

$$\Phi(H, h, m) = \varphi \circ f^{-1} \left\{ \frac{e(m)[f(H) - f(h)] - [f(H)e(h) - f(h)e(H)]}{e(H) - e(h)} \right\}$$
$$- \varphi \circ g^{-1} \left\{ \frac{e(m)[g(H) - g(h)] - [g(H)e(h) - g(h)e(H)]}{e(H) - e(h)} \right\} .$$

Equality holds if

$$\mu\{t \mid x = H\} + \mu\{t \mid x = h\} = 1 .$$

If $f \circ g^{-1}$ is everywhere properly convex, this is also the necessary condition for equality.

If we assume

$$e(h) = f(h) = g(h), \; e(H) = f(H) = g(H) ,$$

Φ becomes simpler:

$$\Phi(H, h, m) = \varphi \circ f^{-1} \circ e(m) - \varphi \circ f^{-1} \circ e(m) .$$

Proof. If Ω satisfies the condition (*) in lemma 2, take a function x stated in lemma 2 with e-mean m. And we have the desired Φ by direct calculation of $\varphi[\mathcal{M}_f(x)]-\varphi[\mathcal{M}_g(x)]$.

If Ω does not satisfy the condition (*), take the product space Ω' of Ω and the interval $[0, 1]$ with ordinary Lebesgue measure and let π be the natural projection to Ω. Ω' satisfies the condition (*) and the established inequality for $x'=x\circ\pi$ means one for x.

The equality condition shows that if Ω has the property (*) this estimate is the best one as by H, h and $\mathcal{M}_e(x)$. In order to have the best estimate only by H and h, we are only necessary to find the maximum value of $\Phi(H, h, m)$ leaving H and h fixed.

Finally we show some applications. Put

$$f(x)=\varphi(x)=x\,,\quad e(x)=g(x)=\log x$$

in theorem 3, and assume $h>0$, then

$$\Phi(H, h, m)=\frac{(H-h)\log m-H\log h+h\log H}{\log H-\log h}-m\,.$$

This as a function of m takes the maximum value

$$h(-1+\log\Gamma-\log\log\Gamma)/\log\Gamma$$

for $m=h/\log\Gamma$, where

$$\Gamma=\left(\frac{H}{h}\right)^{h/(H-h)}.$$

Similarly, putting

$$f(x)=x\,,\quad e(x)=g(x)=\varphi(x)=\log x$$

we have

$$\Phi(H, h, m)=\log\frac{(H-h)\log m-H\log h+h\log H}{\log H-\log h}-m$$

and its maximum value is

$$\Phi(H, h, eh/\Gamma)=\log(\Gamma/e\log\Gamma)\,.$$

Thus we have proved the following

Proposition 4. *Assume that* $h>0$ *then*

i) $$\mathcal{A}(x)-\mathcal{G}(x) \leqq \frac{(H-h)\log \mathcal{G}(x)-H\log h+h\log H}{\log H-\log h}-\mathcal{G}(x)$$
$$\leqq h(-1+\log \Gamma-\log\log \Gamma)/\log \Gamma .$$

The first equality sign holds if and only if

$$\mu\{t \mid x(t)=H\}+\mu\{t \mid x(t)=h\}=1 .$$

And both equality signs hold together if and only if

$$\begin{cases} \mu\{t \mid x(t)=H\} = -\log\log\Gamma/\log(H/h) , \\ \mu\{t \mid x(t)=h\} = 1+\log\log\Gamma/\log(H/h) . \end{cases}$$

ii) $$\frac{\mathcal{A}(x)}{\mathcal{G}(x)} \leqq \frac{(H-h)\log \mathcal{G}(x)-H\log h+h\log H}{\mathcal{G}(x)(\log H-\log h)} \leqq \frac{\Gamma}{e\log \Gamma} .$$

The first equality sign holds if and only if

$$\mu\{t \mid x=H\}+\mu\{t \mid x=h\}=1 .$$

And both equality signs hold together if and only if

$$\begin{cases} \mu\{t \mid x(t)=H\} = 1/\log(H/h)-h/(H-h) , \\ \mu\{t \mid x(t)=h\} = H/(H-h)-1/\log(H/h) . \end{cases}$$

The inequality

$$\mathcal{A}(x)/\mathcal{G}(x) \leqq \Gamma/e\log \Gamma$$

is one of the results by G. T. Cargo and O. Shisha [1], the rests of which can also be proved using theorem 3.

References

[1] Cargo, G. T. and O. Shisha, Bounds on ratios of means, J. Res. Nat. Bur. Standards **66B** (1962), 169–170.

[2] Diaz, J. B. and F. T. Metcalf, Stronger forms of a class of inequalities of G. Pölya-G. Szegó and L. V. Kantrovich, Bull. Amer. Math. Soc. **69** (1963), 415–418.

Note added in proof (July 2, 1969):

Lately I found that there were some works closely related to § 4 of thd present paper. For instance, I note the following

[3] Shisha, O. and G. T. Cargo, On comparable means, Pacific J. Math. **14** (1964), 1053–1058.

[4] Shisha, O. and B. Mond, Differences of means, Bull. Amer. Math. Soc. **73** (1967), 328–333.