ON QUANTILES WITH MINIMUM ASYMPTOTIC GENERALIZED VARIANCE

By

CHANDAN KUMAR MUSTAFI

Indian Institute of Management, Calcutta

(Received in September, 1969)

INTRODUCTION :

The problem of finding the quantile with smallest asymptotic variance has been considered by Sen [3]. In particular, he showed that under certain regularity conditions the sample median has the smallest asymptotic variance among the class of sample quantiles. In this paper, we shall consider the bivariate analogue of the problem using generalized variance [1, p. 166] as a criterion. The bivariate normal distribution will be studied in this perspective. Denoting by x and y as the two variables, it will be shown that the (asymptotic) generalized variance corresponding to the pth sample quantile of x and the qth sample quantile of y $(0<p, q<1)$ is minimized when $p=q=\frac{1}{2}$. While proving this result, we shall also derive the conditions under which the pair of sample medians will have the smallest generalized variance asymptotically for any general bivariate distribution.

SAMPLE QUANTILES WITH SMALLEST GENERALIZED VARIANCE :

Suppose, $F(x, y)$ is a bivariate distribution function with density $f(x, y)$ which has continuous partial derivatives of order two. Let $F_1(x)$ and $F_2(y)$ be the marginal distribution functions of x and y respectively. The corresponding density functions will be denoted by $f_1(x)$ and $f_2(y)$. The population quantile ξ_p of order p for x and the population quantile η_q of order q for y (assumed to be unique) will be defined as

$$F_1(\xi_p)=p; \quad F_2(\eta_q)=q; \quad (0<p, q<1). \qquad ...(1)$$

Let us also define the quantities c_1, c_2, d_{11}, d_{22}, d_{12}, A_{11}, A_{22}, A_{12} as follows :

$$c_1=\left[f_1^{-2}(x)\frac{\partial^2 f_1(x)}{\partial x^2}\right]_{x=\xi_{\frac{1}{2}}} ;$$

$$c_2=\left[f_2^{-2}(y)\frac{\partial^2 f_2(y)}{\partial y^2}\right]_{y=\eta_{\frac{1}{2}}} ;$$

$$d_{11}=\left[f_1^{-2}(x)\frac{\partial^2 F(x,y)}{\partial x^2}\right]_{x=\xi_{\frac{1}{2}},\ y=\eta_{\frac{1}{2}}} ;$$

$$d_{22}=\left[f_2^{-2}(y)\frac{\partial^2 F(x,y)}{\partial y^2}\right]_{x=\xi_{\frac{1}{2}},\ y=\eta_{\frac{1}{2}}} ;$$

$$d_{12}=\left[f_1^{-1}(x)f_2^{-1}(y)f(x,y)\right]_{x=\xi_{\frac{1}{2}},\ y=\eta_{\frac{1}{2}}} ;$$

$$A_{11}=-2c_1f_1^{-1}(\xi_{\frac{1}{2}})+[\tfrac{1}{2}F(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})-F^2(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})]^{-1}$$
$$[\tfrac{1}{2}d_{11}-\tfrac{1}{2}-2F(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})d_{11}] \qquad \ldots(2\cdot1)$$

$$A_{22}=-2c_2f_2^{-1}(\eta_{\frac{1}{2}})+[\tfrac{1}{2}F(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})-F^2(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})]^{-1}$$
$$[\tfrac{1}{2}d_{22}-\tfrac{1}{2}-2F(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})d_{22}] \qquad \ldots(2\cdot2)$$

$$A_{12}=[\tfrac{1}{2}F(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})-F^2(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})]^{-1}$$
$$[2F(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})-\tfrac{1}{2}-2F(\xi_{\frac{1}{2}},\eta_{\frac{1}{2}})d_{12}+\tfrac{1}{2}d_{12}] \qquad \ldots(2\cdot3)$$

In an independent sample of size n, we define the sample quantile of order p for x and the sample quantile of order q for y in the usual way. Then, the asymptotic generalized variance V_{pq} can be obtained from a result of Siddiqui [4, p. 148] as :

$$V_{pq}=\eta^{-2}f_1^{-2}(\xi_p)f_2^{-2}(\eta_q)[pq(1-p)(1-q)-(F(\xi_p,\eta_q)-pq)^2] \quad \ldots(3)$$

We first prove the following result :

Lemma

Suppose, $f_1(x)$ has an extremum at $x=\xi_{\frac{1}{2}}$ and $f_2(y)$ has an extremum at $y=\eta_{\frac{1}{2}}$. Then V_{pq} is minimized for $p=q=\frac{1}{2}$ if the following three conditions are satisfied :—

(*i*) $P\{y\leq\eta_{\frac{1}{2}} \mid x=\xi_{\frac{1}{2}}\}=P\{x\leq\xi_{\frac{1}{2}} \mid y=\eta_{\frac{1}{2}}\}=\frac{1}{2}$;

(*ii*) $A_{11}>0$; $A_{22}>0$;

(*iii*) $A_{11}A_{22}-A_{12}^2>0$.

Proof

We shall deal with log V_{pq} instead of V_{pq} without affecting the results. From (3) $\log V_{pq} = -2\log n - 2\log f_1 - 2\log f_2$

$$+\log\,[pq(1-p-q+2F)-F^2] \qquad \ldots(4)$$

where $\quad f_1 = f_1(\xi_p)\,;\, f_2 = f_2(\eta_q)\,;\; F = F(\xi_p,\, \eta_q).$

Differentiating (4) partially with respect to p and q, we obtain

$$\frac{\partial \log V_{pq}}{\partial p} = -2f_1^{-1}\frac{\partial f_1}{\partial p} + \Big[pq(1-p-q+2F)-F^2\Big]^{-1}$$

$$\Big[q(1-p-q+2F)+pq\Big(-1+\frac{2\partial F}{\partial p}\Big)-2F\frac{\partial F}{\partial p}\Big] \qquad \ldots(5)$$

$$\frac{\partial \log V_{pq}}{\partial q} = -2f_2^{-1}\frac{\partial f_2}{\partial q} + \Big[pq(1-p-q+2F)-F^2\Big]^{-1}$$

$$\Big[p(1-p-q+2F)+pq\Big(-1+\frac{2\partial F}{\partial q}\Big)-2F\frac{\partial F}{\partial q}\Big] \qquad \ldots(6)$$

Since $f_1(x)$ has an extremum at

$$x = \xi_{\frac{1}{2}} \quad \left(\frac{\partial f_1(x)}{\partial x}\right)_{x=\xi_{\frac{1}{2}}} = 0$$

$$\frac{\partial f_1}{\partial p} = \left[\frac{\partial f_1(x)}{\partial F_1(x)}\right]_{x=\xi_p} = \left[\frac{\partial f_1(x)}{\partial x}\cdot\frac{1}{f_1(x)}\right]_{x=\xi_p} = 0 \text{ for } p=\tfrac{1}{2} \qquad \ldots(7)$$

$$\frac{\partial F}{\partial p} = \left[\frac{\partial F(x,y)}{\partial F_1(x)}\right]_{x=\xi_p,\; y=\eta_q}$$

$$= \left[\frac{\partial F(x,y)}{\partial x}\cdot\frac{1}{f_1(x)}\right]_{x=\xi_p,\; y=\eta_q}$$

$$= P[y \leq \eta_q \mid x=\xi_p]$$

$$= \tfrac{1}{2} \text{ for } p=q=\tfrac{1}{2} \text{ by condition } (i) \qquad \ldots(8)$$

Introducing (7) and (8) in (5), we observe that

$$\frac{\partial \log V_{pq}}{\partial p} = 0 \text{ for } p=q=\tfrac{1}{2}$$

Similarly, $\quad \dfrac{\partial \log V_{pq}}{\partial q} = 0 \text{ for } p=q=\tfrac{1}{2}$

Simple calculations lead to :

$$\left(\frac{\partial^2 f_1}{\partial p^2}\right)_{p=\frac{1}{2}}=c_1\ ;\ \left(\frac{\partial^2 f_2}{\partial q^2}\right)_{q=\frac{1}{2}}=c_2\ ;\ \left(\frac{\partial^2 F}{\partial p^2}\right)_{p=q=\frac{1}{2}}=d_{11}$$

$$\left(\frac{\partial^2 F}{\partial p^2}\right)_{p=q=\frac{1}{2}}=d_{22} \qquad \ldots(9)$$

$$\left(\frac{\partial^2 F}{\partial p \partial q}\right)_{p=q=\frac{1}{2}}=d_{12}$$

From (5), (6) and (9) we obtain

$$\left(\frac{\partial^2 \log V_{pq}}{\partial p^2}\right)_{p=q=\frac{1}{2}}=A_{11}\ ;\ \left(\frac{\partial^2 \log V_{pq}}{\partial q^2}\right)_{p=q=\frac{1}{2}}=A_{22}\ ;$$

$$\left(\frac{\partial^2 \log V_{pq}}{\partial p \partial q}\right)_{p=q=\frac{1}{2}}=A_{12}$$

Conditions (*ii*) and (*iii*) thus ensure that V_{pq} is minimized at

$$p=q=\tfrac{1}{2}.$$

This completes the proof of the lemma.

Theorem

Suppose, (x, y) has a bivariate normal distribution with

$$E(x)=\mu_1;\ E(y)=\mu_2;\ V(x)=\sigma_1^2;\ V(x_2)=\sigma_2^2;$$

$$\text{cov}\ (x, y)=\rho\sigma_1\sigma_2(0<|\rho|<1).$$

Then V_{pq} is minimized for $p=q=\frac{1}{2}$.

Proof

We have $\xi_{\frac{1}{2}}=\mu_1$ and $\eta_{\frac{1}{2}}=\mu_2$. Using the properties of the conditional distribution of a bivariate normal distribution, it is easy to see that condition (*i*) of the lemma is satisfied. To complete the proof of the theorem we have to show that conditions (*ii*) and (*iii*) of the lemma are satisfied in this case. Simple calculations lead to :

$$c_1=-(2\pi)^{\frac{1}{2}}\ \sigma_1^{-1};\ c_2=-(2\pi)^{\frac{1}{2}}\ \sigma_2^{-1}$$

$$\frac{\partial^2 F(x, y)}{\partial x^2}=\int_{-\infty}^{y}\left[\frac{\partial f(x, v)}{\partial x}\right]dv, \qquad \ldots(10)$$

where

$$f(x, v)=(2\pi)^{-1}(1-\rho^2)^{-\frac{1}{2}}\sigma_1^{-1}\sigma_2^{-1}\exp\left[-2^{-1}(1-\rho^2)^{-1}\right.$$
$$\left.\{(x-\mu_1)^2\sigma_1^{-2}-2\rho(x-\mu_1)\sigma_1^{-1}\sigma_2^{-1}+(v-\mu_2)^2\sigma_2^{-2}\}\right]$$

Hence

$$\frac{\partial}{\partial x}f(x, v)=f(x, v)[-(x-\mu_1)\sigma_1^{-2}(1-\rho^2)^{-1}$$
$$+\rho(v-\mu_2)\sigma_1^{-1}\sigma_2^{-1}(1-\rho^2)^{-1}]$$

Thus,

$$\left[\frac{\partial^2 F(x,y)}{\partial x^2}\right]_{x=\mu_1,\ y=\mu_2} = \left[(2\pi)^{-1}\rho(1-\rho^2)^{-3/2}\sigma_1^{-2}\sigma_2^{-2}\right]$$

$$\int_{-\infty}^{\mu_2} (v-\mu_2)\exp\left[-2^{-1}(1-\rho^2)^{-1}\sigma_2^{-2}(v-\mu_2)^2\right]dv$$
$$=-(2\pi)^{-1}\rho(1-\rho^2)^{-\frac{1}{2}}\sigma_1^{-2}.$$

The expression for

$$\left[\frac{\partial^2 F(x,y)}{\partial y^2}\right]_{x=\mu_1,\ y=\mu_2}$$

may be obtained in a similar way. These lead to :

$$d_{11}=d_{22}=-\rho(1-\rho^2)^{-\frac{1}{2}}\ ;\ d_{12}=(1-\rho^2)^{-\frac{1}{2}} \quad \ldots(11)$$

It is well know [2, p. 290]

$$F(\xi_{\frac{1}{2}}, \eta_{\frac{1}{2}})=F(\mu_1, \mu_2)=\tfrac{1}{4}+(2\pi)^{-1}(\text{arc}\sin\rho). \quad \ldots(12)$$

From (2), (10) and (11), we obtain

$$A_{11}=A_{22}=4\pi+[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)]^{-1}$$
$$\left[-\frac{\rho}{2}(1-\rho^2)^{-\frac{1}{2}}-\tfrac{1}{2}+2F(\mu_1, \mu_2)\rho(1-\rho^2)^{-\frac{1}{2}}\right] \quad \ldots(13)$$

$$A_{12}=[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)]^{-1}$$
$$[2F(\mu_1, \mu_2)-\tfrac{1}{2}-2(1-\rho^2)^{-\frac{1}{2}}F(\mu_1, \mu_2)+\tfrac{1}{2}(1-\rho^2)^{-\frac{1}{2}}] \quad (\ldots 14)$$

From (12), we observe that for

$$0<|\rho|<1\ \tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)$$
$$=\frac{1}{16}-(2\pi)^{-2}(\text{arc}\sin\rho)^2>0.$$

Thus, $A_{11}>0$ if and only if

$$[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)]A_{11}>0.$$

From (12) and (13) we obtain

$$[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)]A_{11}$$

$$=4\pi[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)] - \frac{\rho}{2}(1-\rho^2)^{-\frac{1}{2}}$$

$$+2F(\mu_1, \mu_2)\rho(1-\rho^2)^{-\frac{1}{2}}-\tfrac{1}{2}$$

$$=-4\pi F^2(\mu_1, \mu_2)+2F(\mu_1, \mu_2)\left[\pi+\rho(1-\rho^2)^{-\frac{1}{2}}\right]-\frac{\rho}{2}(1-\rho^2)^{-\frac{1}{2}}-\tfrac{1}{2}$$

$$=-4\pi(\tfrac{1}{4}+(2\pi)^{-1}\text{ arc sin }\rho)^2+2(\tfrac{1}{4}+(2\pi)^{-1}\text{ arc sin }\rho)$$

$$(\pi+\rho(1-\rho^2)^{-\frac{1}{2}})-\frac{\rho}{2}(1-\rho^2)^{-\frac{1}{2}}-\tfrac{1}{2}$$

$$=-\frac{\pi}{4}+\frac{\pi}{2}-\tfrac{1}{2}+\pi^{-1}(\text{arc sin }\rho)[\rho(1-\rho^2)^{-\frac{1}{2}}-\text{arc sin }\rho]$$

$$>\pi^{-1}(\text{arc sin }\rho)B(\rho), \qquad \ldots(15)$$

where $B(\rho)=\rho(1-\rho^2)^{-\frac{1}{2}}-\text{arc sin }\rho.$

$$B'(\rho)=\rho^2(1-\rho^2)^{-\frac{3}{2}}>0.$$

Hence, $B(\rho)$ is monotone increasing. Since $B(0)=0$, it fo'lows

$$B(\rho)>0 \text{ for } 0<\rho<1$$

$$B(\rho)<0 \text{ for } -1<\rho<0 \qquad \ldots(16)$$

Introducing (16) in (15) we obtain

$$A_{11}>0. \qquad \ldots(17)$$

Thus, condition (*ii*) of the lemma is satisfied. Also, from (13) and (14)

$$[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)]^2(A_{11}A_{22}-A_{12}^2)$$

$$=[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)]^2(A_{11}^2-A_{12}^2)$$

$$=(C+D)(C-D)$$

Where

$$C=4\pi[\tfrac{1}{2}F(\mu_1, \mu_2)-F^2(\mu_1, \mu_2)]-\frac{\rho}{2}(1-\rho^2)^{-\frac{1}{2}}-\tfrac{1}{2}$$

$$+2F(\mu_1, \mu_2)\rho(1-\rho^2)^{-\frac{1}{2}}$$

$$=\frac{\pi}{4}-\frac{1}{2}+\pi^{-1}(\text{arc sin }\rho)[\rho(1-\rho^2)^{-\frac{1}{2}}-\text{arc sin }\rho]>0, \qquad \ldots(18)$$

and $D=2F(\mu_1,\mu_2)-\tfrac{1}{2}-2(1-\rho^2)^{-\frac{1}{2}}F(\mu_1, \mu_2)+\tfrac{1}{2}(1-\rho^2)^{-\frac{1}{2}}$

$$=\pi^{-1}(\text{arc sin }\rho)[1-(1-\rho^2)^{-\frac{1}{2}}]. \qquad \ldots(19)$$

Since, for

$$0 < |\rho| < 1 \qquad 1-(1-\rho^2)^{-\frac{1}{2}} < 0,$$

it follows from (19) that D is positive if and only if ρ is negative. In this case, since

$$C > 0, \; A_{11}^2 - A_{12}^2$$

is positive if and only if $C-D>0$.

Now

$$C-D = \frac{\pi}{4} - \frac{1}{2} + \pi^{-1} (\text{arc sin } \rho) G(\rho). \qquad \text{...(20)}$$

where

$$G(\rho) = (1+\rho)(1-\rho^2)^{-\frac{1}{2}} - \text{arc sin } \rho - 1.$$

$$G'(\rho) = (\rho+\rho^2)(1-\rho^2)^{-\frac{3}{2}} < 0 \text{ for } -1<\rho<0.$$

Thus $G(\rho)$ is monotone decreasing in this interval. Also, by L'Hospital's rule,

$$G(-1) = \frac{\pi}{2} - 1.$$

Hence, for $-1<\rho<0$.

$$0 = G(0) < G(\rho) < \frac{\pi}{2} - 1. \qquad \text{...(21)}$$

Further, for $-1<\rho<0$,

$$-\tfrac{1}{2} < \pi^{-1} (\text{arc sin } \rho) < 0. \qquad \text{...(22)}$$

Introducing (21) and (22) in (20), we obtain

$$C-D > \frac{\pi}{4} - \frac{1}{2} - \frac{1}{2}\left(\frac{\pi}{2} - 1\right) = 0.$$

Again, it follows from (19) that D is negative if and only if ρ is positive. In this case, $A_{11}^2 - A_{12}^2$ is positive if and only if $C+D>0$.

Now,

$$C+D = \frac{\pi}{4} - \frac{1}{2} + \pi^{-1} (\text{arc sin } \rho)\, H(\rho), \qquad \text{...(23)}$$

where

$$H(\rho) = 1 - \text{ arc sin } \rho + (\rho-1)(1-\rho^2)^{-\frac{1}{2}}$$

$$H'(\rho) = (\rho^2-\rho)(1-\rho^2)^{-\frac{3}{2}} < 0 \text{ for } 0<\rho<1,$$

Thus $H(\rho)$ is monotone decreasing in this interval. Also, by L'Hospital's rule

$$H(1)=1-\frac{\pi}{2}.$$

Hence, for $0<\rho<1$,

$$1-\frac{\pi}{2}<H(\rho)<H(0)=0. \quad \ldots(24)$$

Futher, for $0<\rho<1$

$$0<\pi^{-1}(\arcsin \rho)<\tfrac{1}{2} \quad \ldots(25)$$

Introducing (24) and (25) in (23), we obtain

$$C+D>\frac{\pi}{4}-\frac{1}{2}+\frac{1}{2}\left(1-\frac{\pi}{2}\right)=0.$$

Thus, condition (*iii*) of the lemma is satisfied. This completes the proof of the theorem.

SUMMARY

This article studies the generalized variance of the pairs of sample quantiles taken from a bivariate population. Under certain regularity conditions, it has been proved that this generalized variance is minimized when the quantiles are the two medians. The bivariate normal distribution is shown to satisfy these regularity conditions.

REFERENCES

1. Anderson, T.W. (1958) : *An Introduction to the Multivariate Statistical Analysis*, J. Wiley & Sons, New York.

2. Cramer, H. (1946) : *Mathematical Methods of Statistics*, Princeton University Press.

3. Sen, P.K. (1961) : On Some Properties of the Asymptotic Variance of the Sample Quantiles and Mid-ranges, J.R S.S. (B), *23*, 453-459.

4. Siddiqui, M.M. (1960) : Distribution of Quantiles in Samples from a Bivariate Population. J. Research N.B.S. (Math. and Math. Phys.), *64B*, 145-160.

N.B ; Research partially supported by NSF Grant No. GP—8478,