BIVARIATE EXTREME STATISTICS, I

By MASAAKI SIBUYA

(Received Jan. 20, 1960)

0. Introduction and Summary

The largest and the smallest value in a sample, and other statistics related to them are generally named extreme statistics. Their sampling distributions, especially the limit distributions, have been studied by many authors, and principal results are summarized in the recent Gumbel's book [1].

The author extends here the notion of extreme statistics into bivariate distributions and considers the joint distributions of maxima of components in sample vectors. This Part I treats asymptotic properties of the joint distributions.

In the univariate case the limit distributions of the sample maximum were limited to only three types. In the bivariate case, however, types of the limit joint distributions are various: Theorem 5 in Chapter 2 shows that infinitely many types of limit distributions may exist. For a wide class of distributions, two maxima are asymptotically independent or degenerate on a curve. Theorems 2 and 4 give the attraction domains for such limits. In bivariate normal case, two maxima are asymptotically independent unless the correlation coefficient is equal to one.

Throughout these arguments we remark only the dependence between marginal distributions, whose behaviours are well established. For this purpose a fundamental notion of "dependence function" is introduced and discussed in Section 1.

A practical application will be considered in the subsequent paper.

The author is very grateful to Messrs. M. Motoo, K. Isii and K. Takeuchi for their suggestions and discussions, by which theorems in Section 2 are considerably improved.

1. Dependence function

Let $(X_i, Y_i)(i = 1, \cdots, n)$ be a sample from the population with the distribution function $F(x, y)$. The maxima of the components, $X_{\max}$ and $Y_{\max}$, have the joint distribution function

$$\Pr(X_{\max} < x,\ Y_{\max} < y) = F^n(x, y)\ . \tag{1.1}$$

We write the marginal distributions of $F(x, y)$ as

$$G(x)=F(x,\ \infty)\ ,$$
$$H(y)=F(\infty,\ y)\ . \tag{1.2}$$

The distribution functions of $X_{\max}$, $G^n(x)$, and of $Y_{\max}$, $H^n(y)$, are marginal distributions of $F^n(x, y)$.

We introduce the function $\Omega^*(x, y)$ defined by

$$F(x, y)=\Omega^*(x, y)G(x)H(y)\ .$$

From the relation

$$F^n(x, y)=\Omega^{*n}(x, y)G^n(x)H^n(y)\ ,$$

the asymptotic dependence between $X_{\max}$ and $Y_{\max}$ is clarified from the behaviour of $\Omega^{*n}(x, y)$ when $n\rightarrow\infty$. To simplify the argument, we redefine this function as follows:

Definition 1. The dependence function $\Omega(G, H)$, $0<G\leqq 1$, $0<H\leqq 1$ of the distribution $F(x, y)$ (or of the r.v. (X, Y)) is the function that satisfies

$$F(x, y)=\Omega(G(x), H(y))G(x)H(y)\ . \tag{1.3}$$

For $G=0$ or $H=0$, $\Omega(G, H)$ may be defined if the limit

$$\Omega(0, H)=\lim_{G(x)\to 0}\frac{F(x, y)}{G(x)H(y)}\ ,$$

or

$$\Omega(G, 0)=\lim_{H(y)\to 0}\frac{F(x, y)}{G(x)H(y)}\ , \tag{1.4}$$

exists.

Remark 1. Although $G(x)=G_0$ and $H(x)=H_0$ (G_0, H_0 are constants) do not determine a point (x, y) uniquely, the value of $F(x, y)$ is the same for any $(x, y)\in\{(x, y);\ G(x)=G_0,\ H(x)=H_0\}$. Therefore, $\Omega(G, H)$ is a one-valued function defined for all possible values of G and H.

Remark 2. Clearly $\Omega(G,\ H)=1$, if and only if X and Y are independent. From the definition

$$\Omega(G,\ H)=\frac{\Pr(X<x,\ Y<y)}{\Pr(X<x)\Pr(Y<y)}\ ,$$

$\Omega(G, H)>1(<1)$ corresponds to the positive (negative) association between the events $(X<x)$ and $(Y<y)$.

Remark 3. $\Omega(G, H)$ is ordinally invariant. That is, if $\varphi(x)$ and $\psi(y)$ are monotone non-decreasing functions, the dependence function of $(\varphi(X), \psi(Y))$ is also that of (X, Y), because $F(\phi(x), \psi(y))=\Omega(G(\varphi(x)), H(\psi(y))G(\varphi(x))H(\psi(y))$. The domains of two dependence functions are not always the same, unless both $\varphi(x)$ and $\psi(y)$ are strictly increasing.

Remark 4. Considering the r.v. $(G(X), H(Y))$ we see that; a necessary and sufficient condition for $\Omega(G, H)$ to be a dependence function, is that $\Omega(G, H)GH$ is a distribution function defined on $0<G, H\leqq 1$, whose marginal distributions are uniform.

THEOREM 1.

$$L(G, H)\leqq\Omega(G, H)\leqq U(G, H), \quad 0<G, H\leqq 1 , \tag{1.5}$$

where

$$U(G, H)=\min\left(\frac{1}{G}, \frac{1}{H}\right), \tag{1.6}$$

$$L(G, H)=\max\left(0, \frac{G+H-1}{GH}\right). \tag{1.7}$$

The right hand equality holds when the distribution $F(x,y)$ degenerates on a non-decreasing curve on the x-y plane, and the left hand equality on a non-increasing curve.

PROOF. The right hand inequality:

$$F(x, y)\leqq F(x, \infty)=G(x) ,$$

$$F(x, y)\leqq F(\infty, y)=H(y) .$$

Therefore,

$$F(x, y)=\Omega(G(x), H(y))G(x)H(y)\leqq \min(G(x), H(y)) .$$

The equality holds, if and only if $F(x, y)= \min(G(x), H(y))$, which mean the above mentioned condition.

The left hand inequality:

$$G(x)+H(y)-\Omega(G(x), H(y))G(x)H(y)=1-\Pr\{X>x, Y>y\}\leqq 1 .$$

Therefore,

$$\Omega(G, H) \geqq \frac{G+H-1}{GH} .$$

$\Omega(G, H)$ is apparently not negative. The equality holds if and only if $\Pr(X>x,\ Y>y)=0$ for $G(x)+H(y)\geqq 1$.

See Fig. 1.

Example 1. Consider the probability density function defined by

$$f(x, y)=\begin{cases} 2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y , \\ 0, & \text{otherwise.} \end{cases}$$

As

$$F(x, y)=\begin{cases} (x+y-1)^2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y , \\ 0, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1>x+y , \end{cases}$$

and

$$G(x)=x^2,\ H(y)=y^2 ,$$

we have

$$\Omega(G, H)=\begin{cases} \dfrac{(\sqrt{G}+\sqrt{H}-1)^2}{GH}, & \sqrt{G}+\sqrt{H}\geqq 1, \\ 0, & \text{otherwise} . \end{cases} \tag{1.8}$$

$\Omega(G, H)\leqq 1$, because $\sqrt{G}+\sqrt{H}-1\leqq\sqrt{GH}$. X and Y have the negative association.

Example 2. Let Z_i $(i=1, \cdots, k)$ be a sample from the population with the continuous distribution function $K(x)$, and put $Z_{\min}=X$ and $Z_{\max}=Y$. We consider the distribution of (X, Y).

$$F(x, y)=K^k(y)-(K(y)-K(x))^k , \quad x\leqq y ,$$

$$G(x)=1-(1-K(x))^k , \quad K(x)=1-(1-G)^{1/k} ,$$

$$H(y)=K^k(y) , \quad K(y)=H^{1/k} ,$$

$$\Omega(G, H)=\begin{cases} \dfrac{H-(H^{1/k}+(1-G)^{1/k}-1)^k}{GH}, & G\leqq H , \\ 0, & \text{otherwise} . \end{cases} \tag{1.9}$$

As $H^{1/k}+(1-G)^{1/k}-1\leqq H^{1/k}(1-G)^{1/k}$, we have

$$\Omega(G, H)\geqq\frac{H-H(1-G)}{GH}=1 , \qquad G\leqq H .$$

$Z_{\min}$ and $Z_{\max}$ have the positive association.

$$\Omega(G,\ H)=\frac{1}{GH}\left\{H-\left(1+\frac{\log H+\log(1-G)}{k}+O\left(\frac{1}{k^2}\right)\right)^k\right\}$$

$$\longrightarrow 1\ (k\to\infty)\,,\qquad G\leqq H\,,$$

which is a well-known result.

Example 3. Consider a bivariate normal distribution.

$$\phi(x,\ y;\ \rho)=\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)\right\}\,,$$

$$\Phi(x,\ y;\ \rho)=\int_{-\infty}^{x}\int_{-\infty}^{y}\phi(u,\ v)\,dudv\,.$$

As

$$\frac{d\Omega}{d\rho}=\frac{1}{GH}\frac{d\Phi}{d\rho}=\frac{1}{GH}\phi>0 \tag{1.10}$$

(see Appendix) and $\Omega=1$ when $\rho=0$, we have $\Omega>1$, <1 corresponding to $\rho>0$, <0.

The value $\Omega(G, 0)$ and $\Omega(0, H)$ were defined as the limits. When $F(x, y)$ has the probability density function $f(x, y)$, the limit (say) $\Omega(0, H)$ is:

$$\Omega(0,\ H)=\lim_{G\to 0}\frac{F(x,\ y)}{G(x)H(y)}=\lim_{x\to\infty}\frac{\int_{-\infty}^{y}f(x,\ y)dy}{H(y)\int_{-\infty}^{\infty}f(x,\ y)dy}=\frac{1}{H}\lim_{x\to-\infty}H(y\mid x)\,, \tag{1.11}$$

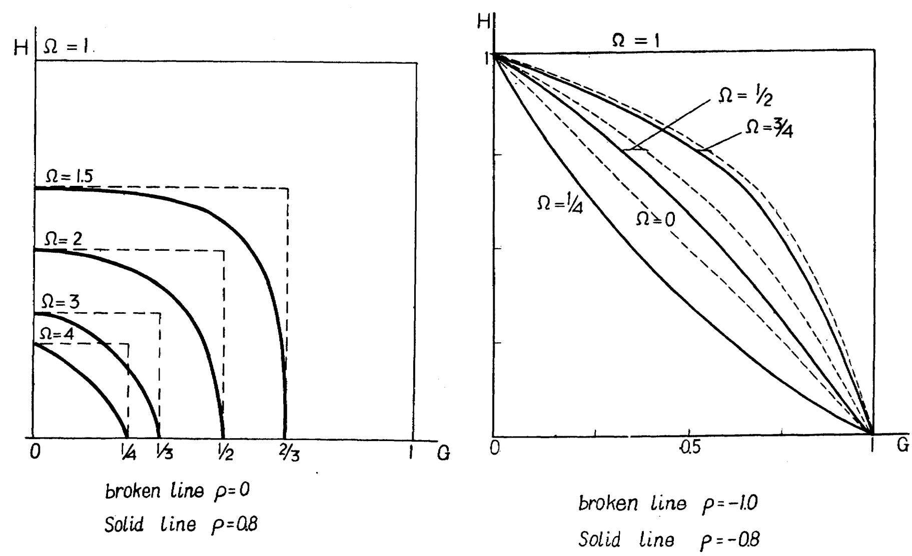

Fig. 1. $\Omega(G, H)$ for bivariate normal distributions.

where $H(y \mid x)$ is the conditional distribution function of Y given $X=x$. Like the bivariate normal distribution, if X has the positive density in a semi-infinite interval $(-\infty, b)$, and if the conditional distribution of Y given $X=x$ has the finite dispersion, monotone increase or decrease of the regression curve corresponds to $\lim H(y \mid x)=1$ or 0, that is to $\Omega(0, H)=H^{-1}$ or 0, which is the same as the value of $U(0, H)$ or $L(0, H)$. $\Omega(G, H)$ for normal distributions are shown in Fig. 1.

2. Asymptotic property

From the previously mentioned relation

$$F^n(x, y)=\Omega^n(G(x), H(y))G^n(x)H^n(y) , \qquad (2.1)$$

the dependence function of maxima $(X_{\max}, Y_{\max})$ in sample of size n is $\Omega^n(G^{1/n}, H^{1/n})$. In this section the asymptotic behaviour of

$$\Omega_n(G, H)\equiv\Omega^n(G^{1/n}, H^{1/n}) \qquad (2.2)$$

for $(n\to\infty)$ is studied. We assume the continuity of $F(x, y)$ or of $\Omega(G, H)$. If $G(x)$ and/or $H(y)$ have a positive jump at the point where $G(x)=1$ or $H(y)=1$, then the limit distribution is trivial. If the distribution has infinitely many points with positive probability, slight modifications in the following arguments are necessary. At first, we examine the examples in Section 1.

Example 1.

$$\Omega(G, H)=\begin{cases}\dfrac{(\sqrt{G}+\sqrt{H}-1)^2}{GH}, & \sqrt{G}+\sqrt{H}\geqq 1, \\ 0, & \text{otherwise} .\end{cases}$$

$$\Omega_n(G, H)=\frac{(G^{1/2n}+H^{1/2n}-1)^{2n}}{GH}$$

$$=\frac{1}{GH}\left(1+\frac{\log G+\log H}{n}+O\left(\frac{1}{n^2}\right)\right)^{2n}, \quad G^{1/2n}+H^{1/2n}\geqq 0 .$$

Therefore,

$$\lim_{n\to\infty}\Omega_n(G, H)=\frac{1}{GH}\exp\{\log G+\log H\}=1 , \quad G, H>0 .$$

The convergence is uniform for $G, H\geqq\varepsilon>0$, where ε is any constant. This is true throughout the following convergence statements.

Example 2

$$\Omega(G, H) \geqq \frac{G+H-1}{GH} .$$

$\Omega(G, H)$ is apparently not negative. The equality holds if and only if $\Pr(X>x,\ Y>y)=0$ for $G(x)+H(y) \geqq 1$.

See Fig. 1.

Example 1. Consider the probability density function defined by

$$f(x, y)=\begin{cases} 2, & 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1,\ 1 \leqq x+y , \\ 0, & \text{otherwise.} \end{cases}$$

As

$$F(x, y)=\begin{cases} (x+y-1)^2, & 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1,\ 1 \leqq x+y , \\ 0, & 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1,\ 1 > x+y , \end{cases}$$

and

$$G(x)=x^2,\ H(y)=y^2 ,$$

we have

$$\Omega(G, H)=\begin{cases} \dfrac{(\sqrt{G}+\sqrt{H}-1)^2}{GH}, & \sqrt{G}+\sqrt{H} \geqq 1, \\ 0, & \text{otherwise} . \end{cases} \tag{1.8}$$

$\Omega(G, H) \leqq 1$, because $\sqrt{G}+\sqrt{H}-1 \leqq \sqrt{GH}$. X and Y have the negative association.

Example 2. Let Z_i $(i=1, \cdots, k)$ be a sample from the population with the continuous distribution function $K(x)$, and put $Z_{\min}=X$ and $Z_{\max}=Y$. We consider the distribution of (X, Y).

$$F(x, y)=K^k(y)-(K(y)-K(x))^k , \quad x \leqq y ,$$

$$G(x)=1-(1-K(x))^k , \quad K(x)=1-(1-G)^{1/k} ,$$

$$H(y)=K^k(y) , \quad K(y)=H^{1/k} ,$$

$$\Omega(G, H)=\begin{cases} \dfrac{H-(H^{1/k}+(1-G)^{1/k}-1)^k}{GH}, & G \leqq H , \\ 0, & \text{otherwise} . \end{cases} \tag{1.9}$$

As $H^{1/k}+(1-G)^{1/k}-1 \leqq H^{1/k}(1-G)^{1/k}$, we have

$$\Omega(G, H) \geqq \frac{H-H(1-G)}{GH}=1 , \qquad G \leqq H .$$

$Z_{\min}$ and $Z_{\max}$ have the positive association.

$$\Omega(G, H) \geqq \frac{G+H-1}{GH} .$$

$\Omega(G, H)$ is apparently not negative. The equality holds if and only if $\Pr(X>x,\ Y>y)=0$ for $G(x)+H(y)\geqq 1$.

See Fig. 1.

Example 1. Consider the probability density function defined by

$$f(x,\ y)=\begin{cases}2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y ,\\ 0, & \text{otherwise.}\end{cases}$$

As

$$F(x,\ y)=\begin{cases}(x+y-1)^2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y ,\\ 0, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1>x+y ,\end{cases}$$

and

$$G(x)=x^2,\ H(y)=y^2 ,$$

we have

$$\Omega(G, H)=\begin{cases}\dfrac{(\sqrt{G}+\sqrt{H}-1)^2}{GH}, & \sqrt{G}+\sqrt{H}\geqq 1,\\ 0, & \text{otherwise} .\end{cases} \tag{1.8}$$

$\Omega(G,\ H)\leqq 1$, because $\sqrt{G}+\sqrt{H}-1\leqq\sqrt{GH}$. X and Y have the negative association.

Example 2. Let Z_i $(i=1, \cdots, k)$ be a sample from the population with the continuous distribution function $K(x)$, and put $Z_{min}=X$ and $Z_{max}=Y$. We consider the distribution of (X, Y).

$$F(x,\ y)=K^k(y)-(K(y)-K(x))^k ,\quad x\leqq y ,$$

$$G(x)=1-(1-K(x))^k ,\quad K(x)=1-(1-G)^{1/k} ,$$

$$H(y)=K^k(y) ,\quad K(y)=H^{1/k} ,$$

$$\Omega(G, H)=\begin{cases}\dfrac{H-(H^{1/k}+(1-G)^{1/k}-1)^k}{GH}, & G\leqq H ,\\ 0, & \text{otherwise} .\end{cases} \tag{1.9}$$

As $H^{1/k}+(1-G)^{1/k}-1\leqq H^{1/k}(1-G)^{1/k}$, we have

$$\Omega(G, H)\geqq\frac{H-H(1-G)}{GH}=1 ,\qquad G\leqq H .$$

Z_{min} and Z_{max} have the positive association.

$$\Omega(G, H) \geqq \frac{G+H-1}{GH} .$$

$\Omega(G, H)$ is apparently not negative. The equality holds if and only if $\Pr(X>x,\ Y>y)=0$ for $G(x)+H(y)\geqq 1$.

See Fig. 1.

Example 1. Consider the probability density function defined by

$$f(x, y)=\begin{cases} 2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y , \\ 0, & \text{otherwise.} \end{cases}$$

As

$$F(x, y)=\begin{cases} (x+y-1)^2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y , \\ 0, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1> x+y , \end{cases}$$

and

$$G(x)=x^2,\ H(y)=y^2 ,$$

we have

$$\Omega(G, H)=\begin{cases} \dfrac{(\sqrt{G}+\sqrt{H}-1)^2}{GH}, & \sqrt{G}+\sqrt{H}\geqq 1, \\ 0, & \text{otherwise} . \end{cases} \qquad (1.8)$$

$\Omega(G, H)\leqq 1$, because $\sqrt{G}+\sqrt{H}-1\leqq\sqrt{GH}$. X and Y have the negative association.

Example 2. Let Z_i $(i=1, \cdots, k)$ be a sample from the population with the continuous distribution function $K(x)$, and put $Z_{\min}=X$ and $Z_{\max}=Y$. We consider the distribution of (X, Y).

$$F(x, y)=K^k(y)-(K(y)-K(x))^k ,\quad x\leqq y ,$$

$$G(x)=1-(1-K(x))^k ,\quad K(x)=1-(1-G)^{1/k} ,$$

$$H(y)=K^k(y) ,\quad K(y)=H^{1/k} ,$$

$$\Omega(G, H)=\begin{cases} \dfrac{H-(H^{1/k}+(1-G)^{1/k}-1)^k}{GH}, & G\leqq H , \\ 0, & \text{otherwise} . \end{cases} \qquad (1.9)$$

As $H^{1/k}+(1-G)^{1/k}-1\leqq H^{1/k}(1-G)^{1/k}$, we have

$$\Omega(G, H)\geqq\frac{H-H(1-G)}{GH}=1 ,\qquad G\leqq H .$$

$Z_{\min}$ and $Z_{\max}$ have the positive association.

$$\Omega(G, H) \geqq \frac{G+H-1}{GH} .$$

$\Omega(G, H)$ is apparently not negative. The equality holds if and only if $\Pr(X>x,\ Y>y)=0$ for $G(x)+H(y)\geqq 1$.

See Fig. 1.

Example 1. Consider the probability density function defined by

$$f(x,\ y)=\begin{cases} 2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y\ , \\ 0, & \text{otherwise.} \end{cases}$$

As

$$F(x,\ y)=\begin{cases} (x+y-1)^2, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1\leqq x+y\ , \\ 0, & 0\leqq x\leqq 1,\ 0\leqq y\leqq 1,\ 1> x+y\ , \end{cases}$$

and

$$G(x)=x^2,\ H(y)=y^2\ ,$$

we have

$$\Omega(G, H)=\begin{cases} \dfrac{(\sqrt{G}+\sqrt{H}-1)^2}{GH}, & \sqrt{G}+\sqrt{H}\geqq 1, \\ 0, & \text{otherwise}\ . \end{cases} \tag{1.8}$$

$\Omega(G,\ H)\leqq 1$, because $\sqrt{G}+\sqrt{H}-1\leqq\sqrt{GH}$. X and Y have the negative association.

Example 2. Let Z_i $(i=1, \cdots, k)$ be a sample from the population with the continuous distribution function $K(x)$, and put $Z_{min}=X$ and $Z_{max}=Y$. We consider the distribution of (X, Y).

$$F(x,\ y)=K^k(y)-(K(y)-K(x))^k\ ,\quad x\leqq y\ ,$$

$$G(x)=1-(1-K(x))^k\ ,\quad K(x)=1-(1-G)^{1/k}\ ,$$

$$H(y)=K^k(y)\ ,\quad K(y)=H^{1/k}\ ,$$

$$\Omega(G, H)=\begin{cases} \dfrac{H-(H^{1/k}+(1-G)^{1/k}-1)^k}{GH}, & G\leqq H\ , \\ 0, & \text{otherwise}\ . \end{cases} \tag{1.9}$$

As $H^{1/k}+(1-G)^{1/k}-1\leqq H^{1/k}(1-G)^{1/k}$, we have

$$\Omega(G, H)\geqq\frac{H-H(1-G)}{GH}=1\ ,\qquad G\leqq H\ .$$

Z_{min} and Z_{max} have the positive association.

is necessary and sufficient for the continuous bounded function $\varphi(x)$ *to be convex in* $0 \leqq x < \infty$.

PROOF. (2.13) is necessary. Because, if $\varphi(x)$ is convex, for any $0 \leqq u_1 < u_2 < u_3$, we have

$$\frac{\varphi(u_2)-\varphi(u_1)}{u_2-u_1} \leqq \frac{\varphi(u_3)-\varphi(u_1)}{u_3-u_1} \leqq \frac{\varphi(u_3)-\varphi(u_2)}{u_3-u_2} . \tag{2.14}$$

By definition, $x_1+\Delta x_1 < x_2+\Delta x_2$. In the case $x_1+\Delta x_1 = x_2$, the relation is equivalent to (2.14), and if $x_1+\Delta x_1 \neq x_2$,

$$\frac{\varphi(x_1+\Delta x_1)-\varphi(x_1)}{\Delta x_1} \leqq \frac{\varphi(x_1+\Delta x_1)-\varphi(x_2)}{(x_1+\Delta x_1)-x_2} \leqq \frac{\varphi(x_2+\Delta x_2)-\varphi(x_2)}{\Delta x_2} .$$

Conversely, if (2.13) is true, putting $x_1+\Delta x_1 = x_2$ and $x_2+\Delta x_2 = x_3$,

$$\frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1} \leqq \frac{\varphi(x_3)-\varphi(x_2)}{x_3-x_2}$$

holds true for any $0 < x_1 < x_3$ and their geometric mean $x_2 = \sqrt{x_1 x_3}$. Applying the inequality successively to the points $(x_1, x_1^{(1)} = \sqrt{x_1 x_2}, x_2)$, $(x_2, x_2^{(1)} = \sqrt{x_2 x_3}, x_3)$; $(x_1, x_1^{(2)} = \sqrt{x_1 x_1^{(1)}}, x_1^{(1)})$; $(x_1^{(1)}, x_2^{(2)} = \sqrt{x_1^{(1)} x_2}, x_2)$, $(x_2, x_3^{(2)} = \sqrt{x_2 x_2^{(1)}}, x_2^{(1)})$, $(x_2^{(1)}, x_4^{(2)} = \sqrt{x_2^{(1)} x_3}, x_3)$; and so on, we see that at the geometric means the function $\varphi(x)$ is below or on the cord joining $(x_1, \varphi(x_1))$ and $(x_3, \varphi(x_3))$. As the series of geometric means is dense in (x_1, x_3) and $\varphi(x)$ is continuous, $\varphi(x)$, $x_1 < x < x_3$, is below or on the cord. $0 < x_1 < x_3$ is arbitrary and $\varphi(x)$ is continuous at $x=0$, so $\varphi(x)$ is convex in $0 \leqq x < \infty$.

Now we give the proof of the theorem.

PROOF. At first we prove that $\Omega(G, H)$ must have the above form. Put

$$\xi = \log G , \quad \eta = \log H, \quad 0 \geqq \xi, \quad \eta > -\infty , \tag{2.15}$$

$$\omega(\xi, \eta) = \log \Omega(e^{\xi}, e^{\eta}), \tag{2.16}$$

then the stability condition may be expressed as

$$\omega(\xi, \eta) = \omega\left(\frac{\xi}{n}, \frac{\eta}{n}\right) .$$

That is, $\omega(\xi, \eta)$ is Euler's homogeneous function of 1st order, and must have the form (see for example [2])

$$\omega(\xi, \eta) = \xi \theta\left(\frac{\eta}{\xi}\right) ,$$

which is the same as (2.11).

For

$$\Omega(G, H)=G^{\chi(\log H/\log G)}$$

to be a dependence function, ΩGH must be a distribution with uniform marginal distributions. We transform G and H into ξ and η by (2.15). In other words, we consider the distributions with marginal distributions $\exp \xi$ and $\exp \eta$. Then

$$\begin{aligned} F(\xi, \eta) &= \exp\left\{\xi\left(\chi\left(\frac{\eta}{\xi}\right)+1\right)+\eta\right\} \\ &\equiv \exp\{\mu(\xi, \eta)\} . \end{aligned} \tag{2.17}$$

The conditions $F(\xi, 0)=\exp\xi$ and $F(0, \eta)=\exp\eta$ pose the restriction

$$\chi(0)=0 , \qquad \chi(\alpha)=o(\alpha) \qquad (\alpha\to\infty) \tag{2.18}$$

$$\begin{aligned} \Delta^2_{\xi\eta}F(\xi, \eta) &\equiv F(\xi, \eta)-F(\xi-\Delta\xi, \eta)-F(\xi, \eta-\Delta\eta)+F(\xi-\Delta\xi, \eta-\Delta\eta) \\ &= F(\xi, \eta)[1-\exp(-\Delta_\xi\mu(\xi, \eta))-\exp(-\Delta_\eta\mu(\xi, \eta)) \\ &\qquad +\exp(\mu(\xi-\Delta\xi, \eta-\Delta\eta)-\mu(\xi, \eta))] \\ &= F(\xi, \eta)\exp(-\Delta_\xi\mu(\xi, \eta)-\Delta_\eta\mu(\xi, \eta)) \\ &\qquad \times[\{\exp(\Delta_\xi\mu(\xi, \eta))-1\}\{\exp(\Delta_\eta\mu(\xi, \eta))-1\} \\ &\qquad +\exp(\Delta^2_{\xi\eta}\mu(\zeta, \eta))-1] \\ &= F(\xi, \eta)\exp(-\Delta_\xi\mu(\xi, \eta)-\Delta_\eta\mu(\xi, \eta)) \\ &\qquad \times[\Delta_\xi\mu(\xi, \eta)\cdot\Delta_\eta\mu(\xi, \eta)\cdot\exp(\theta_1(\xi, \eta)) \\ &\qquad +\Delta^2_{\xi\eta}\mu(\xi, \eta)\cdot\exp(\theta_2(\xi, \eta))] , \end{aligned} \tag{2.19}$$

where $\theta_1(\xi, \eta)$ and $\theta_2(\xi, \eta)$ are functions of the value of factors before exponential functions.

$$\begin{aligned} \Delta_\xi\mu(\xi, \eta) &= \xi\left(\chi\left(\frac{\eta}{\xi}\right)+1\right)-(\xi-\Delta\xi)\left(\chi\left(\frac{\eta}{\xi-\Delta\xi}\right)+1\right) \\ &= (\xi-\Delta\xi)\left(\chi\left(\frac{\eta}{\xi}\right)-\chi\left(\frac{\eta}{\xi-\Delta\xi}\right)+1\right) . \end{aligned}$$

If $\xi\neq0$ and $\eta\neq0$, putting

$$\frac{\eta}{\xi}=\alpha,\ \frac{\eta}{\xi}-\frac{\eta}{\xi-\Delta\xi}=\frac{-\eta\Delta\xi}{\xi(\xi-\Delta\xi)}=\Delta\alpha>0 ,$$

we have

$$\Delta_\xi\mu(\xi, \eta)=-\alpha\cdot\Delta\xi\left(\frac{\chi(\alpha)-\chi(\alpha-\Delta\alpha)}{\Delta\alpha}-\frac{\chi(\alpha)+1}{\alpha}\right) , \tag{2.20}$$

is necessary and sufficient for the continuous bounded function $\varphi(x)$ *to be convex in* $0 \leqq x < \infty$.

Proof. (2.13) is necessary. Because, if $\varphi(x)$ is convex, for any $0 \leqq u_1 < u_2 < u_3$, we have

$$\frac{\varphi(u_2)-\varphi(u_1)}{u_2-u_1} \leqq \frac{\varphi(u_3)-\varphi(u_1)}{u_3-u_1} \leqq \frac{\varphi(u_3)-\varphi(u_2)}{u_3-u_2} . \tag{2.14}$$

By definition, $x_1+\Delta x_1 < x_2+\Delta x_2$. In the case $x_1+\Delta x_1 = x_2$, the relation is equivalent to (2.14), and if $x_1+\Delta x_1 \neq x_2$,

$$\frac{\varphi(x_1+\Delta x_1)-\varphi(x_1)}{\Delta x_1} \leqq \frac{\varphi(x_1+\Delta x_1)-\varphi(x_2)}{(x_1+\Delta x_1)-x_2} \leqq \frac{\varphi(x_2+\Delta x_2)-\varphi(x_2)}{\Delta x_2} .$$

Conversely, if (2.13) is true, putting $x_1+\Delta x_1 = x_2$ and $x_2+\Delta x_2 = x_3$,

$$\frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1} \leqq \frac{\varphi(x_3)-\varphi(x_2)}{x_3-x_2}$$

holds true for any $0 < x_1 < x_3$ and their geometric mean $x_2 = \sqrt{x_1 x_3}$. Applying the inequality successively to the points $(x_1,\ x_1^{(1)} = \sqrt{x_1 x_2},\ x_2)$, $(x_2,\ x_2^{(1)} = \sqrt{x_2 x_3},\ x_3)$; $(x_1,\ x_1^{(2)} = \sqrt{x_1 x_1^{(1)}},\ x_1^{(1)})$; $(x_1^{(1)},\ x_2^{(2)} = \sqrt{x_1^{(1)} x_2},\ x_2)$, $(x_2,\ x_3^{(2)} = \sqrt{x_2 x_2^{(1)}},\ x_2^{(1)})$, $(x_2^{(1)},\ x_4^{(2)} = \sqrt{x_2^{(1)} x_3},\ x_3)$; and so on, we see that at the geometric means the function $\varphi(x)$ is below or on the cord joining $(x_1, \varphi(x_1))$ and $(x_3, \varphi(x_3))$. As the series of geometric means is dense in (x_1, x_3) and $\varphi(x)$ is continuous, $\varphi(x)$, $x_1 < x < x_3$, is below or on the cord. $0 < x_1 < x_3$ is arbitrary and $\varphi(x)$ is continuous at $x = 0$, so $\varphi(x)$ is convex in $0 \leqq x < \infty$.

Now we give the proof of the theorem.

Proof. At first we prove that $\Omega(G, H)$ must have the above form. Put

$$\xi = \log G , \quad \eta = \log H, \quad 0 \geqq \xi, \quad \eta > -\infty , \tag{2.15}$$

$$\omega(\xi,\ \eta) = \log \Omega(e^{\xi},\ e^{\eta}), \tag{2.16}$$

then the stability condition may be expressed as

$$\omega(\xi, \eta) = \omega\left(\frac{\xi}{n},\ \frac{\eta}{n}\right).$$

That is, $\omega(\xi, \eta)$ is Euler's homogeneous function of 1st order, and must have the form (see for example [2])

$$\omega(\xi,\ \eta) = \xi\theta\left(\frac{\eta}{\xi}\right),$$

which is the same as (2.11).

For

$$\Omega(G, H)=G^{\chi(\log H/\log G)}$$

to be a dependence function, ΩGH must be a distribution with uniform marginal distributions. We transform G and H into ξ and η by (2.15). In other words, we consider the distributions with marginal distributions $\exp \xi$ and $\exp \eta$. Then

$$\begin{aligned} F(\xi, \eta) &= \exp\left\{\xi\left(\chi\left(\frac{\eta}{\xi}\right)+1\right)+\eta\right\} \\ &\equiv \exp\{\mu(\xi, \eta)\} . \end{aligned} \tag{2.17}$$

The conditions $F(\xi, 0)=\exp \xi$ and $F(0, \eta)=\exp \eta$ pose the restriction

$$\chi(0)=0 , \qquad \chi(\alpha)=o(\alpha) \qquad (\alpha\to\infty) \tag{2.18}$$

$$\begin{aligned} \Delta^2_{\xi\eta}F(\xi, \eta) &\equiv F(\xi, \eta)-F(\xi-\Delta\xi, \eta)-F(\xi, \eta-\Delta\eta)+F(\xi-\Delta\xi, \eta-\Delta\eta) \\ &= F(\xi, \eta)[1-\exp(-\Delta_\xi\mu(\xi, \eta))-\exp(-\Delta_\eta\mu(\xi, \eta)) \\ &\qquad +\exp(\mu(\xi-\Delta\xi, \eta-\Delta\eta)-\mu(\xi, \eta))] \\ &= F(\xi, \eta)\exp(-\Delta_\xi\mu(\xi, \eta)-\Delta_\eta\mu(\xi, \eta)) \\ &\qquad \times[\{\exp(\Delta_\xi\mu(\xi, \eta))-1\}\{\exp(\Delta_\eta\mu(\xi, \eta))-1\} \\ &\qquad +\exp(\Delta^2_{\xi\eta}\mu(\zeta, \eta))-1] \\ &= F(\xi, \eta)\exp(-\Delta_\xi\mu(\xi, \eta)-\Delta_\eta\mu(\xi, \eta)) \\ &\qquad \times[\Delta_\xi\mu(\xi, \eta)\cdot\Delta_\eta\mu(\xi, \eta)\cdot\exp(\theta_1(\xi, \eta)) \\ &\qquad +\Delta^2_{\xi\eta}\mu(\xi, \eta)\cdot\exp(\theta_2(\xi, \eta))] , \end{aligned} \tag{2.19}$$

where $\theta_1(\xi, \eta)$ and $\theta_2(\xi, \eta)$ are functions of the value of factors before exponential functions.

$$\begin{aligned} \Delta_\xi\mu(\xi, \eta) &= \xi\left(\chi\left(\frac{\eta}{\xi}\right)+1\right)-(\xi-\Delta\xi)\left(\chi\left(\frac{\eta}{\xi-\Delta\xi}\right)+1\right) \\ &= (\xi-\Delta\xi)\left(\chi\left(\frac{\eta}{\xi}\right)-\chi\left(\frac{\eta}{\xi-\Delta\xi}\right)+1\right) . \end{aligned}$$

If $\xi\neq0$ and $\eta\neq0$, putting

$$\frac{\eta}{\xi}=\alpha, \quad \frac{\eta}{\xi}-\frac{\eta}{\xi-\Delta\xi}=\frac{-\eta\Delta\xi}{\xi(\xi-\Delta\xi)}=\Delta\alpha>0 ,$$

we have

$$\Delta_\xi\mu(\xi, \eta)=-\alpha\cdot\Delta\xi\left(\frac{\chi(\alpha)-\chi(\alpha-\Delta\alpha)}{\Delta\alpha}-\frac{\chi(\alpha)+1}{\alpha}\right) , \tag{2.20}$$

is necessary and sufficient for the continuous bounded function $\varphi(x)$ to be convex in $0 \leqq x < \infty$.

Proof. (2.13) is necessary. Because, if $\varphi(x)$ is convex, for any $0 \leqq u_1 < u_2 < u_3$, we have

$$\frac{\varphi(u_2)-\varphi(u_1)}{u_2-u_1} \leqq \frac{\varphi(u_3)-\varphi(u_1)}{u_3-u_1} \leqq \frac{\varphi(u_3)-\varphi(u_2)}{u_3-u_2} . \tag{2.14}$$

By definition, $x_1+\Delta x_1 < x_2+\Delta x_2$. In the case $x_1+\Delta x_1 = x_2$, the relation is equivalent to (2.14), and if $x_1+\Delta x_1 \neq x_2$,

$$\frac{\varphi(x_1+\Delta x_1)-\varphi(x_1)}{\Delta x_1} \leqq \frac{\varphi(x_1+\Delta x_1)-\varphi(x_2)}{(x_1+\Delta x_1)-x_2} \leqq \frac{\varphi(x_2+\Delta x_2)-\varphi(x_2)}{\Delta x_2} .$$

Conversely, if (2.13) is true, putting $x_1+\Delta x_1 = x_2$ and $x_2+\Delta x_2 = x_3$,

$$\frac{\varphi(x_2)-\varphi(x_1)}{x_2-x_1} \leqq \frac{\varphi(x_3)-\varphi(x_2)}{x_3-x_2}$$

holds true for any $0 < x_1 < x_3$ and their geometric mean $x_2 = \sqrt{x_1 x_3}$. Applying the inequality successively to the points $(x_1,\ x_1^{(1)} = \sqrt{x_1 x_2},\ x_2)$, $(x_2,\ x_2^{(1)} = \sqrt{x_2 x_3},\ x_3)$; $(x_1,\ x_1^{(2)} = \sqrt{x_1 x_1^{(1)}},\ x_1^{(1)})$; $(x_1^{(1)},\ x_2^{(2)} = \sqrt{x_1^{(1)} x_2},\ x_2)$, $(x_2,\ x_3^{(2)} = \sqrt{x_2 x_2^{(1)}},\ x_2^{(1)})$, $(x_2^{(1)},\ x_4^{(2)} = \sqrt{x_2^{(1)} x_3},\ x_3)$; and so on, we see that at the geometric means the function $\varphi(x)$ is below or on the cord joining $(x_1, \varphi(x_1))$ and $(x_3, \varphi(x_3))$. As the series of geometric means is dense in (x_1, x_3) and $\varphi(x)$ is continuous, $\varphi(x)$, $x_1 < x < x_3$, is below or on the cord. $0 < x_1 < x_3$ is arbitrary and $\varphi(x)$ is continuous at $x=0$, so $\varphi(x)$ is convex in $0 \leqq x < \infty$.

Now we give the proof of the theorem.

Proof. At first we prove that $\Omega(G, H)$ must have the above form. Put

$$\xi = \log G , \quad \eta = \log H, \quad 0 \geqq \xi, \quad \eta > -\infty , \tag{2.15}$$

$$\omega(\xi, \eta) = \log \Omega(e^{\xi}, e^{\eta}), \tag{2.16}$$

then the stability condition may be expressed as

$$\omega(\xi, \eta) = \omega\left(\frac{\xi}{n}, \frac{\eta}{n}\right) .$$

That is, $\omega(\xi, \eta)$ is Euler's homogeneous function of 1st order, and must have the form (see for example [2])

$$\omega(\xi, \eta) = \xi \theta\left(\frac{\eta}{\xi}\right) ,$$

where $\phi(t)$ is the p.d.f. of the standard normal, and therefore

$$\Phi(x,\ y;\ \rho)=\int_{-\infty}^{x}\Phi\left(\frac{y-\rho u}{\sqrt{1-\rho^2}}\right)\phi(u)\,du\ .$$

Using them, we get

$$\frac{d\Phi(x,\ y;\ \rho)}{d\rho}=\int_{-\infty}^{x}\phi(u)\left\{\phi\left(\frac{y-\rho u}{\sqrt{1-\rho^2}}\right)\frac{\rho(y-\rho u)-u(1-\rho^2)}{(1-\rho^2)\sqrt{1-\rho^2}}\right\}du$$

$$=\frac{1}{\sqrt{1-\rho^2}}\int_{-\infty}^{x}\frac{-(u-\rho y)}{\sqrt{1-\rho^2}}\frac{1}{\sqrt{1-\rho^2}}\phi\left(\frac{y-\rho u}{\sqrt{1-\rho^2}}\right)\phi(u)\,du$$

$$=\frac{1}{\sqrt{1-\rho^2}}\int_{-\infty}^{x}\frac{-(u-\rho x)}{\sqrt{1-\rho^2}}\frac{1}{\sqrt{1-\rho^2}}\phi\left(\frac{u-\rho y}{\sqrt{1-\rho^2}}\right)\phi(y)\,du$$

$$=\frac{1}{\sqrt{1-\rho^2}}\phi\left(\frac{x-\rho y}{\sqrt{1-\rho^2}}\right)\phi(y)=\phi(x,\ y;\ \rho)\ .$$

For general $N(\mu_x,\mu_y;\ \sigma_x^2,\sigma_y^2;\ \rho)$, by the same procedure we have $d\Phi/d\rho=\sigma_x\sigma_y\phi$.

THE INSTUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Gumbel, E. J., *Statistics of Extremes*, Columbia University Press, N. Y., 1958.

[2] Aczél, J., "Some general methods in the theory of one variable functional equation. New applications of functional equations," *Uspekhi Math. Nauk*, 11 (1956), pp. 3–68 (in Russian).

[3] Owen, D. B., "Tables for computing bivariate normal probabilities," *Ann. Math. Stat.*, 27 (1956), pp. 1075–1090.

While the paper was in press the author became aware of Konijn's paper

[4] Konijn, H. S., "Positive and negative dependence of two random variables," *Sankhya* 21 (1959), pp. 269–280

which treats Theorem 1 of Section 1 more rigorously.