CONTRIBUTIONS TO THE PROBLEM OF MAXIMAL CORRELATION[1]

by

PÉTER CSÁKI and JÁNOS FISCHER

Introduction

The classical indices of bivariate stochastic connection are far from being perfect. The maximal correlation may be regarded as a good measure of the correlation in its broadest sense, i. e. the intensity of stochastic connection. This paper aims at giving some contributions to the approach of the problem of maximal correlation.

In this paper our notations of [1] will be used. A short survey of the main notions used therein seems desirable.

Let ξ and η be arbitrary measurable functions defined on the probability space $(\Omega, S, \mathbf{P})$ and S_ξ denote the smallest σ-algebra with respect to which ξ is measurable, further $L_\xi^2 = L^2(\Omega, S_\xi, \mathbf{P})$. The subspace of L_ξ^2 consisting of its elements with zero expected values will be denoted by $L_{\xi,0}^2$. Naturally, S_η, L_η^2 and $L_{\eta,0}^2$ can be analogously defined. The symbol $\mathbf{A}_\xi$ denotes the operator of the orthogonal projection of the elements of $L_{\eta,0}^2$ on $L_{\xi,0}^2$ (conditional expected value, i. e. regression curve) and $\mathbf{A}_\eta$ analogously the projection of the elements of $L_{\xi,0}^2$ on $L_{\eta,0}^2$.

Thus the correlation ratio of a standard random variable ζ on ξ is

$$\theta_\xi(\zeta) = \|\mathbf{A}_\xi \zeta\|, \tag{0.1}$$

the maximal correlation of ξ and η

$$\mathbf{S}(\xi, \eta) = \sup_{\substack{f \in L_{\xi,0}^2,\ \|f\|=1 \\ g \in L_{\eta,0}^2,\ \|g\|=1}} (f, g) = \|\mathbf{A}_\xi\| = \|\mathbf{A}_\eta\| \tag{0.2}$$

and the mean square contingency

$$\mathbf{C}(\xi, \eta) = |||\mathbf{A}_\xi||| = |||\mathbf{A}_\eta|||. \tag{0.3}$$

If the number λ and the pair of random variables f, g satisfy both equations

$$\begin{cases} \mathbf{A}_\xi g = \lambda f \\ \mathbf{A}_\eta f = \lambda g \end{cases} \qquad \begin{cases} f \in L_{\xi,0}^2 \\ g \in L_{\eta,0}^2, \end{cases} \tag{0.4}$$

we call λ an eigenvalue and the pair f, g a pair of eigenfunctions belonging to λ.

---

[1] A previous version of this paper has been read on 7. Sept. 1959 at the Biometric Symposium (Budapest).

In § 1 some questions concerning the maximal correlation and the linearity of correlation are dealt with. Further a method of solving (0.4) in particular cases is shown. In § 2 some examples are presented.

§ 1. Maximal correlation and conditional expectation

1.1. Both theoretical and practical considerations make it desirable to have linear correlation. Thus the problem of linearizing the regressions often arises. In the following some remarks on this question are given.

Theorem 1. *For two standard random variables $f \in L^2_{\xi,0}$ and $g \in L^2_{\eta,0}$ the following three statements are equivalent:*

1° *f and g form a pair of eigenfunctions.*

2° $\theta_\xi(g) = \theta_\eta(f) = (f, g)$.

3° *f and g are linearly correlated and $\theta_f(g) = \theta_\xi(g)$; $\theta_g(f) = \theta_\eta(f)$.*

Proof. Let us suppose at first that 1° holds. Then from (0.1) and considering the norms in (0.4) 2° follows.

For the second, if 2° holds then as from $\mathbf{A}_f g = \mathbf{A}_f \mathbf{A}_\xi g$ the inequality $\| \mathbf{A}_f g \| \leqq \| \mathbf{A}_\xi g \|$ follows whence

$$(f, g) \leqq \theta_f(g) \leqq \theta_\xi(g)\,, \tag{1.1}$$

— according to our assumption — equalities in (1.1) are obtained. On the analogy of (1.1) we have equalities in

$$(f, g) \leqq \theta_g(f) \leqq \theta_\eta(f) \tag{1.2}$$

as well.

Equalities on the left side in both (1.1) and (1.2) imply the linear correlation between f and g while those on the right side the other assumptions of 3°.

Finally, let us consider 3°. The linear correlation provides $\mathbf{A}_f g = \lambda f$. On the other hand, $\theta_f(g) = \theta_\xi(g)$ implies $\mathbf{A}_f g = \mathbf{A}_\xi g$, with respect to $\mathbf{A}_f g = \mathbf{A}_f \mathbf{A}_\xi g$, as the norms of a function and its projection may equal only if this function is a fix element of the actual projection $\mathbf{A}_f$. From the above

$$\mathbf{A}_\xi g = \lambda f\,.$$

Similarly

$$\mathbf{A}_\eta f = \lambda g$$

and so we obtain 1°.

As 2° has been deduced from 1°, 3° from 2° and 1° from 3°, our Theorem is proved.

It is noteworthy that the linear correlation between the members of a pair of eigenfunctions was already pointed out by H. O. Hirschfeld [2] for the finite discrete case.

Corollary. *If the standard random variables $f = f(\xi)$ and $g = g(\eta)$ are linearly correlated further $f(x)$ and $g(y)$ are univalent functions then f and g form a pair of eigenfunctions.*

Proof. In consequence of the univalence $L^2_f = L^2_\xi$ and $L^2_g = L^2_\eta$, 3° of Theorem 1 is satisfied, thus 1° holds, too.

1.2. The distance of the standard eigenfunctions belonging to the maximal correlation, turns out to be $2(1 - \mathbf{S}(\xi, \eta))$ as if $f \in L^2_{\xi,0}$, $g \in L^2_{\eta,0}$ are standard variables, then

$$\|f - g\|^2 = 2(1 - (f, g)),$$

hence

$$\inf_{\substack{f \in L^2_{\xi,0},\, \|f\|=1 \\ g \in L^2_{\eta,0},\, \|g\|=1}} \|f - g\|^2 = 2\Big(1 - \sup_{\substack{f \in L^2_{\xi,0},\, \|f\|=1 \\ g \in L^2_{\eta,0},\, \|g\|=1}} (f, g)\Big) = 2\big(1 - \mathbf{S}(\xi, \eta)\big). \tag{1.3}$$

Let the preceding infimum problem be modified as

$$\inf_{\substack{f \in L^2_{\xi,0},\, \|f\|=1 \\ g \in L^2_{\eta,0},\, \|g\|=1 \\ -\infty<\lambda<\infty}} \|f - \lambda g\|^2 = \inf_{\substack{f \in L^2_{\xi,0},\, \|f\|=1 \\ g \in L^2_{\eta,0},\, \|g\|=1 \\ -\infty<\lambda<\infty}} \big(1 - 2\lambda(f, g) + \lambda^2\big) = 1 - \mathbf{S}^2(\xi, \eta) \tag{1.4}$$

Obviously the infimum in (1.4) is lower than in (1.3) save the case $\mathbf{S}(\xi, \eta) = 1$, when they coincide.

It is evident from (1.3) that $\mathbf{S}(\xi, \eta) = 1$ holds if and only if the distance of the unit-spheres of $L^2_{\xi,0}$ and $L^2_{\eta,0}$ equals zero. This case may be treated in four subcases, namely:

a) The unit-spheres are disjoint (example see in [4]). Here maximal correlation is not attainable (supremum but not maximum in (0.2)).

b) Both differences of the unit-spheres are non-empty sets. Now $f(\xi) = g(\eta)$ holds but both $f(x)$ and $g(y)$ have to be non-invertible functions.

c) One of the unit-spheres contains the other. Now one of ξ and η is a non-invertible function of the other.

d) The unit-spheres are coincident ($L^2_{\xi,0} = L^2_{\eta,0}$). This means that ξ and η are univalent functions of each other.

The relation

$$\theta_\eta(\zeta) \leqq \theta_\xi(\zeta) \qquad \text{for all } \zeta \in L^2$$

is necessary and sufficient for $\eta = f(\xi)$ (cases c) and d)) as follows from a theorem, see e. g. A. C. ZAANEN ([5], p. 250).

1.3. The value of (1.4) is equal to the mean conditional variance of the pair of eigenfunctions belonging to the maximal correlation, as

$$\inf_{\substack{f \in L^2_{\xi,0} \\ \|f\|=1}} \|f - \mathbf{A}_\eta f\|^2 = \inf_{\substack{f \in L^2_{\xi,0} \\ \|f\|=1}} (1 - \|\mathbf{A}_\eta f\|^2) = \inf_{\substack{f \in L^2_{\xi,0} \\ \|f\|=1}} \big(1 - \theta^2_\eta(f)\big) = 1 - \mathbf{S}^2(\xi, \eta). \tag{1.5}$$

For practical purposes the homoscedasticity (constant conditional variance), i. e. for standard ζ the relation $\mathbf{A}_\xi \zeta^2 - (\mathbf{A}_\xi \zeta)^2 \equiv 1 - \theta^2_\xi(\zeta)$, is desired. As to this we can state

Theorem 2. *For the linearly correlated standard random variables ξ and η the following statements are equivalent:*

1° *ξ and η are homoscedastically correlated.*

2° *$\xi^2 - 1$ and $\eta^2 - 1$ form a pair of eigenfunctions belonging to $(\xi, \eta)^2$.*

Proof. Both $\mathbf{A}_\xi \eta = (\xi, \eta)\, \xi$ and $\mathbf{A}_\eta \xi = (\xi, \eta)\, \eta$ are true. Let us first suppose 1°. From this

$$\text{(1.6)}\qquad \begin{cases} \mathbf{A}_\xi \eta^2 - (\xi,\eta)^2 \xi^2 = \mathbf{A}_\xi \eta^2 - (\mathbf{A}_\xi \eta)^2 \equiv 1 - \theta_\xi^2(\eta) = 1 - (\xi,\eta)^2 \\ \mathbf{A}_\eta \xi^2 - (\xi,\eta)^2 \eta^2 = \mathbf{A}_\eta \xi^2 - (\mathbf{A}_\eta \xi)^2 \equiv 1 - \theta_\eta^2(\xi) = 1 - (\xi,\eta)^2 \end{cases}$$

which implies

$$\text{(1.7)}\qquad \begin{cases} \mathbf{A}_\xi(\eta^2 - 1) = (\xi,\eta)^2 (\xi^2 - 1) \\ \mathbf{A}_\eta(\xi^2 - 1) = (\xi,\eta)^2 (\eta^2 - 1)\,. \end{cases}$$

Conversely, from the linear correlation and (1.7) by (1.6) 1° follows.

Corollary. *If the correlation of the standard variables ξ and η is both linear and homoscedastic, further if $0 < |(\xi, \eta)| < 1$ then the third moments of ξ and η vanish.*

Proof. According to Theorem 2 $(\xi, \xi^2 - 1) = (\eta, \eta^2 - 1) = 0$ since eigenfunctions belonging to different eigenvalues are orthogonal, and thus

$$\mathbf{M}(\xi^3) = \mathbf{M}(\xi) = \mathbf{M}(\eta^3) = \mathbf{M}(\eta) = 0\,.$$

Remark. In case of the bivariate normal distribution the conditions of Theorem 2 and the Corollary are evidently fulfilled.

1.4. The calculation of the value of the maximal correlation is in general rather complicated and practically intractable. In special cases however, — as will be seen in the following — it can be managed fairly easy.

Let $\varphi_1, \varphi_2, \ldots$ resp. $\psi_1, \psi_2, \ldots$ be linearly independent systems in the spaces $L^2_{\xi,0}$, resp. $L^2_{\eta,0}$. In this case we have the following

Theorem 3. *The functions*

$$\text{(1.8)}\qquad f_n = \sum_{k=1}^{n} a_{kn} \varphi_k\,, \quad g_n = \sum_{k=1}^{n} a'_{kn} \psi_k \qquad a_{nn} a'_{nn} \neq 0\,;\ n = 1, 2, \ldots$$

are the eigenfunctions of the pair of operators $\mathbf{A}_\xi$, $\mathbf{A}_\eta$ if and only if such coefficients b_{kn} and b'_{kn} exist that

$$\text{(1.9)}\qquad \mathbf{A}_\eta \varphi_n = \sum_{k=1}^{n} b_{kn} \psi_k\,, \quad \mathbf{A}_\xi \psi_n = \sum_{k=1}^{n} b'_{kn} \psi_k \qquad n = 1, 2, \ldots$$

and in this case the appropriate eigenvalues are

$$\text{(1.10)}\qquad \lambda_n = \sqrt{b_{nn} b'_{nn}}\,, \qquad n = 1, 2, \ldots$$

moreover, for the coefficients a_{kn} and a'_{kn} the equations

$$\text{(1.11)}\qquad \sum_{k=i}^{n} b_{ik} a_{kn} = \lambda_n a'_{in}\,, \quad \sum_{k=i}^{n} b'_{ik} a'_{kn} = \lambda_n a_{in} \qquad i = 1, \ldots, n;\ n = 1, 2, \ldots$$

are fulfilled.

Proof. If the functions (1.8) are the eigenfunctions belonging to the eigenvalues λ_n, that is

$$\mathbf{A}_\eta f_n = \lambda_n g_n \quad \text{and} \quad \mathbf{A}_\xi g_n = \lambda_n f_n \qquad n = 1, 2, \ldots$$

then

$$\sum_{k=1}^{n} a_{kn} \mathbf{A}_\eta \varphi_k = \lambda_n \sum_{k=1}^{n} a'_{kn} \psi_k$$

and

$$\sum_{k=1}^{n} a'_{kn} \mathbf{A}_\xi \psi_k = \lambda_n \sum_{k=1}^{n} a_{kn} \varphi_k ;$$

this shows that both $\mathbf{A}_\eta \varphi_n$ and $\mathbf{A}_\xi \psi_n$ are linear combinations of $\psi_1, \ldots, \psi_n$, resp. $\varphi_1, \ldots, \varphi_n$. Formulae (1.11) result from the linear independence of the systems $\{\varphi_n\}$, $\{\psi_n\}$, wherefrom for $i = n$ $a_{nn} b_{nn} = \lambda_n a'_{nn}$ and $a'_{nn} b'_{nn} = \lambda_n a_{nn}$ which implies (1.10).

Conversely, if (1.9) and (1.11) hold then for the functions (1.8)

$$\mathbf{A}_\eta f_n = \sum_{k=1}^{n} a_{kn} \mathbf{A}_\eta \varphi_n = \sum_{k=1}^{n} \sum_{i=1}^{k} a_{kn} b_{ik} \psi_i = \sum_{i=1}^{n} \sum_{k=i}^{n} a_{kn} b_{ik} \psi_i = \lambda_n \sum_{i=1}^{n} a'_{in} \psi_i = \lambda_n g_n$$

and similarly $\mathbf{A}_\xi g_n = \lambda_n f_n$, where $\lambda_n = \sqrt{b_{nn} b'_{nn}}$, thus our statements are verified.

If $\{\lambda_n\}$ provides all the non-zero eigenvalues then — in case of maximum in (0.2) —

$$\mathbf{S}(\xi, \eta) = \max_n \sqrt{b_{nn} b'_{nn}}$$

holds.

If the joint distribution is symmetric in its variables the equalities (1.9) and (1.11) reduce to $a'_{kn} = a_{kn}$ and $b'_{kn} = b_{kn}$, consequently $\lambda_n = b_{nn}$.

Corollary. *This theorem may be applied if the linearly independent functions can be chosen so that*

$$\varphi_n = \xi^n - \mathbf{M}(\xi^n) \in L^2_{\xi,0}, \quad \psi_n = \eta^n - \mathbf{M}(\eta^n) \in L^2_{\eta,0} \qquad n = 1, 2, \ldots$$

In these particular cases, our Theorem implies that the eigenfunctions are polynomials if and only if for each n the n^{th} conditional moment is an at most n^{th}-degree polynomial of the conditioning variable.

In consequence of Theorem 3 the eigenvalues can be found without any further computation and the coefficients of the eigenfunctions can be determined from a linear system of equations provided that in both spaces $L^2_{\xi,0}$ and $L^2_{\eta,0}$ respective systems of linearly independent functions are known such that their conditional expected values can be written in the form (1.9).

§ 2. Examples

Some examples of calculating the maximal correlation will be presented. The following symbols will be used:

$h(x, y)$ for the joint density function of ξ and η,
$h_1(x)$ for the density function of ξ,
$h_2(y)$ for the density function of η,
$h_1(x \mid y)$ for the conditional density function of ξ on η,
$h_2(y \mid x)$ for the conditional density function of η on ξ,
p_{ik} for the joint probability of $\xi = i$ and $\eta = k$,
$p_{i\cdot}$ for the probability of $\xi = i$ and
$p_{\cdot k}$ for the probability of $\eta = k$.

1. In this example the maximal correlation coincides with the correlation coefficient. Let be

$$h(x,y)=\begin{cases}\log\dfrac{1}{(1-x)\,y} & 0\leq x\leq y\leq 1\\[2ex] \log\dfrac{1}{x(1-y)} & 0\leq y\leq x\leq 1\end{cases}$$

In this case the joint distribution is symmetric, i. e. $h(x, y) = h(y, x)$. Hence, one-sided treatment of the problem is sufficient. Now

$$h_1(x) = 1 \qquad 0 \leq x \leq 1$$

and

$$h_1(x \mid y) = h(x, y) \qquad 0 \leq x \leq 1, \qquad 0 \leq y \leq 1\,.$$

Therefore, using Theorem 3

$$\mathbf{A}_n \xi^n = \frac{1}{n+1}\left(\frac{1}{n+1} + \sum_{i=1}^{n} \frac{\eta^i}{i}\right) \qquad n = 1, 2, \ldots$$

and

$$\lambda_n = b_{nn} = \frac{1}{n(n+1)} \qquad n = 1, 2, \ldots$$

The functions $\xi^n - \mathbf{M}(\xi^n)$ form a complete system of linearly independent polynomials in $L^2_{\xi,0}$. Thus $\{\lambda_n\}$ is the set of all eigenvalues, from which we obtain

$$\mathbf{S}(\xi, \eta) = \lambda_1 = \frac{1}{2}\,.$$

2. a) Trinomial distribution.

$$p_{ik} = \frac{N!}{i!\,k!(N-i-k)!}\, p_1^i\, p_2^k (1-p_1-p_2)^{N-i-k}\,,$$

where $p_1 > 0$, $p_2 > 0$, $p_1 + p_2 < 1$; $i = 0, 1, \ldots, N$; $k = 0, 1, \ldots, N$; $i + k \leq N$. Then the marginal distributions are

$$p_{i\cdot} = \binom{N}{i} p_1^i (1-p_1)^{N-i} \qquad i = 0, 1, \ldots, N,$$

$$p_{\cdot k} = \binom{N}{k} p_2^k (1-p_2)^{N-k} \qquad k = 0, 1, \ldots, N$$

and the conditional distributions

$$\frac{p_{ik}}{p_{\cdot k}} = \binom{N-k}{i}\left(\frac{p_1}{1-p_2}\right)^i \left(\frac{1-p_2-p_1}{1-p_2}\right)^{N-k-i} \qquad i = 0, \ldots, N-k,$$

$$\frac{p_{ik}}{p_{i\cdot}} = \binom{N-i}{k}\left(\frac{p_2}{1-p_1}\right)^k \left(\frac{1-p_1-p_2}{1-p_1}\right)^{N-i-k} \qquad k = 0, \ldots, N-i.$$

Now applying Theorem 3:

$$\mathbf{A}_\eta \xi^n = \sum_{j=1}^{n} a_{jn} \left(\frac{p_1}{1-p_2}\right)^j (N-\eta)(N-\eta-1)\ldots(N-\eta-j+1) \quad n = 1, \ldots, N,$$

$$\mathbf{A}_\xi \eta^n = \sum_{j=1}^{n} a_{jn} \left(\frac{p_2}{1-p_1}\right)^j (N-\xi)(N-\xi-1)\ldots(N-\xi-j+1) \quad n = 1, \ldots, N$$

where $a_{nn} = 1$; thus the eigenvalues are

$$\lambda_n = \left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right)^{\frac{n}{2}} \qquad n = 1, \ldots, N$$

and the maximal correlation is

$$\mathbf{S}(\xi, \eta) = \sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}}\,.$$

b) Trihypergeometric distribution.

$$p_{ik} = \frac{\binom{N_1}{i}\binom{N_2}{k}\binom{N-N_1-N_2}{n-i-k}}{\binom{N}{n}},$$

where n, N_1, N_2, N are positive integers, $N_1 + N_2 < N$; $i = 0, 1, \ldots, n$; $k = 0, 1, \ldots, n$; $i + k \leqq n$. Similarly to a) we obtain the eigenvalues

$$\lambda_m = \left[\frac{\binom{N-N_1-m}{N_2-m}\binom{N-N_2-m}{N_1-m}}{\binom{N-N_1}{N_2}\binom{N-N_2}{N_1}}\right]^{\frac{1}{2}}$$

and especially

$$\mathbf{S}(\xi, \eta) = \lambda_1 = \sqrt{\frac{N_1 N_2}{(N-N_1)(N-N_2)}}\,.$$

$\left(p_1 = \frac{N_1}{N},\ p_2 = \frac{N_2}{N}\right.$ result for $\mathbf{S}(\xi, \eta)$ in the same formula as obtained in a). $\Big)$

3. A case of $\mathbf{S}(\xi, \eta) = \mathbf{C}(\xi, \eta)$.

Let us consider the symmetric density function

$$h(x, y) = p_1\, a(x)\, a(y) + p_2\, [a(x)\, b(y) + b(x)\, a(y)] + p_3\, b(x)\, b(y) \tag{2.1}$$

where $a(x)$ and $b(x)$ are linearly independent. For sake of simplicity let us suppose that $a(x)$ and $b(x)$ are density functions, further $p_1 \geqq 0$, $p_2 \geqq 0$, $p_3 \geqq 0$ and $p_1 + 2p_2 + p_3 = 1$. Then

$$h_1(x) = pa(x) + qb(x) \qquad p = p_1 + p_2;\ q = p_2 + p_3\,.$$

The construction of $h(x, y)$ allows at most one non-zero eigenvalue, whence

$$\mathbf{S}(\xi, \eta) = \mathbf{C}(\xi, \eta)\,.$$

$\mathbf{C}^2(\xi, \eta)$ being equal to the square-integral of

$$\frac{h(x,y) - h_1(x)\,h_1(y)}{\sqrt{h_1(x)\,h_1(y)}} = P\,\frac{a(x) - b(x)}{\sqrt{pa(x) + qb(x)}} \cdot \frac{a(y) - b(y)}{\sqrt{pa(y) + qb(y)}} \quad \text{where } P = p_1 p_3 - p_2^2,$$

we have

$$\mathbf{C}(\xi, \eta) = |P| \int_{-\infty}^{\infty} \frac{[a(x) - b(x)]^2}{pa(x) + qb(x)}\,dx = \frac{|P|}{pq}\left\{1 - \int_{-\infty}^{\infty} \frac{a(x)\,b(x)}{pa(x) + qb(x)}\,dx\right\}.$$

Therefore

$$\mathbf{S}(\xi, \eta) = \frac{|P|}{pq}\left\{1 - \int_{-\infty}^{\infty} H\left(\frac{a(x)}{2\,q}, \frac{b(x)}{2\,p}\right) dx\right\}, \tag{2.2}$$

where $H(x, y) = \dfrac{2}{\dfrac{1}{x} + \dfrac{1}{y}}$, the related eigenfunction being $\dfrac{a(x) + b(x)}{pa(x) + qb(x)}$.

In the particular case $p_2 = 0$, we have

$$\mathbf{S}(\xi, \eta) = 1 - \int_{-\infty}^{\infty} H\left(\frac{a(x)}{2\,p_3}, \frac{b(x)}{2\,p_1}\right) dx$$

and in case $p_2 = \dfrac{1}{2}$ $(p_1 = p_3 = 0)$

$$\mathbf{S}(\xi, \eta) = 1 - \int_{-\infty}^{\infty} H\big(a(x), b(x)\big)\,dx\,.$$

Let us consider now

a) an example due to A. Rényi ([3] p. 317), where

$$h(x,y) = \frac{1}{2\,\pi}\left\{\left(\sqrt{2}\,e^{-\frac{x^2}{2}} - e^{-x^2}\right) e^{-y^2} + \left(\sqrt{2}\,e^{-\frac{y^2}{2}} - e^{-y^2}\right) e^{-x^2}\right\}.$$

In this case the marginal distributions are normal and have the density functions

$$h_1(x) = h_2(x) = \frac{1}{\sqrt{2\,\pi}}\,e^{-\frac{x^2}{2}}\,.$$

They are uncorrelated, but not independent. In this example $p_1 = p_3 = 0$, $p_2 = \dfrac{1}{2}$ and

$$a(x) = \frac{1}{\sqrt{\pi}}\left(\sqrt{2}\,e^{-\frac{x^2}{2}} - e^{-x^2}\right), \quad b(x) = \frac{1}{\sqrt{\pi}}\,e^{-x^2},$$

hence

$$\mathbf{S}(\xi,\eta) = 1 - \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \left(\sqrt{2}\, e^{-x^2} - e^{-\frac{3x^2}{2}}\right) dx = \frac{2}{\sqrt{3}} - 1 \approx 0{,}1547 .$$

b) A further simple special case:

$$h(x, y) = 4p_1 xy + 2p_2 (x + y) + p_3 \qquad 0 \leqq x \leqq 1;\ 0 \leqq y \leqq 1,$$

where

$$a(x) = 2x, \qquad b(x) = 1 \qquad\qquad 0 \leqq x \leqq 1;$$

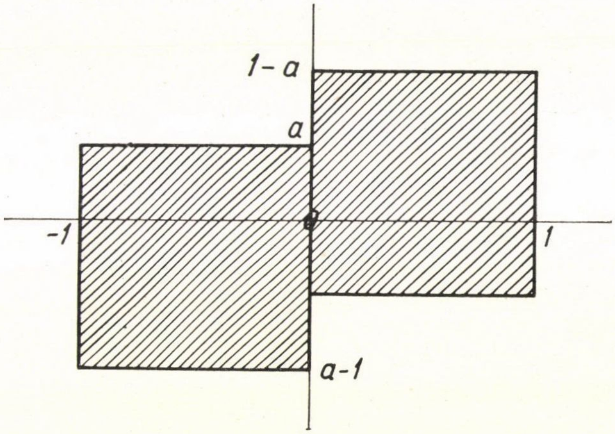

Figure 1.

the maximal correlation is

$$\mathbf{S}(\xi,\eta) = \frac{|P|}{p^3} (\operatorname{arth}\ p - p) .$$

E. g. if $h(x, y) = x + y$ or $h(x, y) = 2xy + \frac{1}{2}$, then

$$\mathbf{S}(\xi, \eta) = \log 3 - 1 \approx 0{,}0986 .$$

c) An example where the parallelity between the intensity of connection and the value of the maximal correlation is manifest.

Let us consider a domain T consisting of two squares of unit area

$$[-1 \leqq x \leqq 0;\ a - 1 \leqq y \leqq a] \text{ and } [0 \leqq x \leqq 1;\ -a \leqq y \leqq 1 - a]$$

where $0 \leqq a \leqq \frac{1}{2}$ (see Fig. 1) and let the joint distribution of ξ and η be uniform in T. This may be transformed in a symmetric distribution with density function of type (2.1)

$$\bar{h}(x, y) = \int_{-\infty}^{\infty} \frac{h(x, t)\, h(y, t)}{h_2(t)}\, dt . \tag{2.3}$$

The eigenvalues of the joint density function (2.3) are the squares of the original ones (see [1] (2.18)).

In the present case

$$p_1 = p_3 = \frac{1-a}{2}, \quad p_2 = \frac{a}{2}; \quad a(x) = 1 \quad -1 \leqq x \leqq 0, \; b(x) = 1 \quad 0 \leqq x \leqq 1.$$

Thus by (2.2)

$$\mathbf{S}(\xi, \eta) = \sqrt{\frac{|P|}{pq}} = \sqrt{1-2a}.$$

4. An example for multiple eigenvalues. (See A. C. Zaanen [5] p. 539.) Let us consider the symmetric joint density function (generating a completely continuous operator)

$$h(x, y) = k(x-y) \qquad 0 \leqq x \leqq 2\pi, \quad 0 \leqq y \leqq 2\pi$$

where the function $k(x)$ has the following properties: $k(x) \geqq 0$ with period 2π, summable over $(0,2\pi)$ and $\int_0^{2\pi} k(x)\, dx = \frac{1}{2\pi}$, further $k(-x) = k(x)$. Then

$$h_1(x) = \frac{1}{2\pi} \qquad 0 \leqq x \leqq 2\pi$$

and

$$h_1(x \mid y) = 2\pi\, k(x-y) \quad 0 \leqq x \leqq 2\pi, \quad 0 \leqq y \leqq 2\pi.$$

Now the eigenvalues are

$$\lambda_n = 2\pi \int_0^{2\pi} k(x) \cos nx\, dx \qquad n = 1, 2, \ldots$$

and the respective eigenfunctions

$$f_{n1} = \cos n\xi, \quad f_{n2} = \sin n\xi \qquad n = 1, 2, \ldots$$

i. e. the λ_n-s are double eigenvalues. Therefore

$$\mathbf{S}(\xi, \eta) = \max_n |\lambda_n|.$$

5. Let the domain T be defined on the plane $[x, y]$ by

$$T = \{(x, y) : |x|^p + |y|^q \leqq 1\} \qquad p > 0, \; q > 0$$

and let ξ, η be uniformly distributed in T.[2]

---

[2] This example is the generalization of P. Bártfai's unpublished solution for $p = q = 2$.

We shall see that — in case $p = q$ — to the narrower T-s higher maximal correlations are attached. Now,

$$h(x,y) = \frac{1}{t} \quad \text{for} \quad (x,y) \in T\,,$$

where

$$t = \iint_T dx\,dy = 4\int_0^1 \int_0^{(1-y^q)^{\frac{1}{p}}} dx\,dy = 4\int_0^1 (1-y^q)^{\frac{1}{p}}\,dy = \frac{4}{q}\,B\left(\frac{1}{q}, \frac{1}{p}+1\right).$$

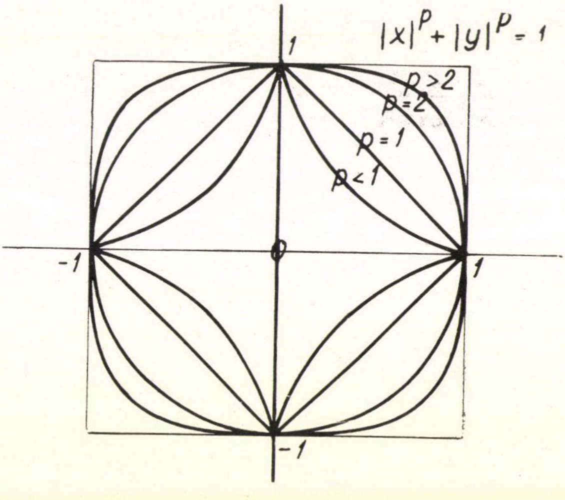

Figure 2.

The density functions of the marginal distributions are

$$h_1(x) = \frac{2}{t}\,(1-|x|^p)^{\frac{1}{q}} \qquad |x| \leq 1,$$

$$h_2(y) = \frac{2}{t}\,(1-|y|^q)^{\frac{1}{p}} \qquad |y| \leq 1$$

and the conditional density functions

$$h_1(x\,|\,y) = \frac{1}{2(1-|y|^q)^{\frac{1}{p}}} \qquad |x| \leq (1-|y|^q)^{\frac{1}{p}}, \quad |y| \leq 1\,,$$

$$h_2(y\,|\,x) = \frac{1}{2(1-|x|^p)^{\frac{1}{q}}} \qquad |y| \leq (1-|x|^p)^{\frac{1}{q}}, \quad |x| \leq 1\,.$$

Let us choose the linearly independent functions (see Theorem 3) as

$$\varphi_n = |\xi|^{pn} - \mathbf{M}(|\xi|^{pn}), \quad \psi_n = |\eta|^{qn} - \mathbf{M}(|\eta|^{qn}) \qquad n = 1, 2, \ldots$$

Then as results from

$$\int_{-1}^{1} |x|^{pn} h_1(x \mid y)\, dx = \frac{1}{(1-|y|^q)^{\frac{1}{p}}} \int_0^{(1-|y|^q)^{\frac{1}{p}}} x^{pn}\, dx = \frac{1}{pn+1}(1-|y|^q)^n,$$

we have

$$\mathbf{A}_\eta \varphi_n = \frac{1}{pn+1}(1-|\eta|^q)^n - \mathbf{M}(|\xi|^{pn}), \quad \mathbf{A}_\xi \psi_n = \frac{1}{qn+1}(1-|\xi|^p)^n - \mathbf{M}(|\eta|^{qn})$$

consequently

$$b_{nn} = \frac{(-1)^n}{pn+1}, \quad b'_{nn} = \frac{(-1)^n}{qn+1} \qquad n = 1, 2, \ldots$$

whence the eigenvalues are

$$\lambda_n = \frac{1}{\sqrt{(pn+1)(qn+1)}} \qquad n = 1, 2, \ldots.$$

The eigenfunctions belonging to such eigenvalues do not by all means form complete systems and thus $\mathbf{C}(\xi, \eta)$ is to be determined. These are, however, all the non-zero eigenvalues as

$$1 + \mathbf{C}^2(\xi, \eta) = \iint_T \frac{h^2(x, y)}{h_1(x)\, h_2(y)}\, dx\, dy = \frac{1}{4} \iint_T \frac{dx\, dy}{(1-|x|^p)^{\frac{1}{q}}(1-|y|^q)^{\frac{1}{p}}} =$$

$$= \int_0^1 \int_0^{(1-y^q)^{\frac{1}{p}}} \frac{dx\, dy}{(1-|x|^p)^{\frac{1}{q}}(1-|y|^q)^{\frac{1}{p}}} = \int_0^1 \frac{1}{(1-y^q)^{\frac{1}{p}}} \int_0^{(1-y^q)^{\frac{1}{p}}} \sum_{n=0}^{\infty} (-1)^n \binom{-\frac{1}{q}}{n} x^{pn}\, dx\, dy =$$

$$= \int_0 \sum_{n=0}^{\infty} (-1)^n \binom{-\frac{1}{q}}{n} \frac{1}{pn+1}(1-y^q)^n\, dy =$$

$$= \sum_{n=0}^{\infty} \binom{n+\frac{1}{q}-1}{n} \frac{1}{pn+1} \frac{1}{q} \int_0^1 u^{\frac{1}{q}-1}(1-u)^n\, du =$$

$$= \sum_{n=0}^{\infty} \frac{1}{(pn+1)(qn+1)} = 1 + \sum_{n=1}^{\infty} \lambda_n^2,$$

where binominal expansion and beta-function were applied. Accordingly, the maximal correlation proves to be

$$\mathbf{S}(\xi, \eta) = \max_n \frac{1}{\sqrt{(pn+1)(qn+1)}} = \frac{1}{\sqrt{(p+1)(q+1)}}$$

and the related eigenfunctions

$$f_1 = |\xi|^p - \frac{1}{p+2}, \quad g_1 = |\eta|^q - \frac{1}{q+2}.$$

It is to be noted that in this example:

a) the correlation coefficient as well as the correlation ratios vanish.

b) in calculating the maximal correlation — if it is attainable — ξ and η may be replaced by $|\xi|$ and $|\eta|$, respectively, if and only if to $\mathbf{S}(\xi, \eta)$ there belongs a pair of even eigenfunctions. Consequently, the domain T may be reduced e. g. to its upper half or even to its positive quadrant without altering the value of the maximal correlation.

(Received April 28, 1960.)

REFERENCES

[1] Csáki, P.—Fischer, J.: „On bivariate stochastic connection". *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **5** (1960) 311—323.

[2] Hirschfeld, H. O.: „A connection between correlation and contingency". *Proc. of the Camb. Phil. Soc.* **31** (1935) 520—524.

[3] Rényi A.: *Valószínűségszámítás.* Tankönyvkiadó, Budapest, 1954.

[4] Rényi, A.: „On measures of dependence". *Acta Math. Acad. Sci. Hung.* **10** (1959) 441—451.

[5] Zaanen, A. C.: *Linear analysis.* North.-Holland Publishing Co., Amsterdam, 1956.

ЗАМЕЧАНИЕ О ПРОБЛЕМЕ МАКСИМАЛЬНОЙ КОРРЕЛЯЦИИ

P. CSÁKI и J. FISCHER

Резюме

Работа занимается максимальной корреляцией, являющейся со многих точек зрения наилучшей мерой силы стохастической связи. В § 1 даются необходимые и достаточные условия того, чтобы стандартные случайные величины $f = f(\xi)$ и $g = g(\eta)$ были парой собственных функций пары операторов $\mathbf{A}_\xi$ и $\mathbf{A}_\eta$ (теорема 1). В дальнейшем характеризуется случай $\mathbf{S}(\xi, \eta) = 1$ и дается необходимое и достаточное условие того, чтобы η была функцией от ξ: $\theta_\eta(\zeta) \leqq \theta_\xi(\zeta)$ для всех ζ с конечной дисперсией.

Для случая взаимно линейной регрессии дается необходимое и достаточное условие того, чтобы и условная дисперсия регрессий была постоянной (теорема 2).

Кроме того дается метод вычисления собственных значений и функций пары операторов $\mathbf{A}_\xi$, $\mathbf{A}_\eta$ в том случае, когда известны линейно независимые системы функций $\varphi_1, \varphi_2, \ldots$ и $\psi_1, \psi_2, \ldots$ величин ξ и η с конечной дисперсией, удовлетворяющих (1.9) (теорема 3). В качестве специального случая получается, что собственные функции в том и только в том случае многочлены, если n-ые условные моменты суть многочлены не выше n-ой степени условной величины.

В § 2 приводится ряд примеров вычисления максимальной корреляции.