International Journal of Theoretical Physics, Vol. 26, No. 1, 1987

Joint Probability Distribution of Composite Quantum Systems

Silviu Guiasu[1]

Received June 16, 1986

We determine the joint probability distribution for two observables attached to two systems in weak interaction, by minimizing the entropic measure of interdependence subject to constraints given by marginal expected values and by the correlation coefficient between the two observables.

1. INTRODUCTION

Many believe that the joint probability distribution of two observables attached to the *same* quantum system does not deserve special attention. Indeed, if the corresponding operators commute, then they are independent and the joint probability distribution of their values is simply the product of the corresponding marginal probability distributions. On the other hand, if the two operators do not commute, then they cannot be determined simultaneously and therefore it is senseless to look for a joint distribution in such a case. For two different systems weakly interacting, however, the joint behavior of two weakly correlated observables can apparently be analyzed by using a joint probability distribution on the possible values.

The objective of this paper is to determine the joint probability distribution for two observables attached to two systems in weak interaction. When we know the marginal expected values, the marginal variances, and the covariance, or the correlation coefficient, between the two observables, then the corresponding joint probability distribution of the two observables is not uniquely determined. To remove this uncertainty we construct the joint probability distribution that is the closest one to independence subject to the above-mentioned constraints, the closeness between two probability distributions being measured by the Kullback–Leibler number from information theory. The exact solution can be considerably simplified just when

[1]Department of Mathematics, York University, Toronto, Canada.

0020-7748/87/0100-0011$05.00/0 © 1987 Plenum Publishing Corporation

the two observables are weakly correlated. The minimization of the divergence from independence gives the most random model of the joint behavior of the two observables compatible with the constraints expressed by first- and second-order average values of the observables and by a covariance between them. The philosophy behind such a variational principle is similar to the well-known maximum entropy techniques used in statistical mechanics, originated from the work done by von Neumann (1932), Shannon (1948), and Jaynes (1957). For details see Guiasu (1977), and Guiasu and Shenitzer (1985).

2. COMPOSITE SYSTEMS

Let S_1 and S_2 be two quantum systems and let S be the composite system obtained if S_1 and S_2 are regarded as interacting together. If $\{\xi_i\}$ is a complete orthonormal set from the Hilbert space H_1 corresponding to S_1 and $\{\eta_j\}$ is a complete orthonormal set from the Hilbert space H_2 corresponding to S_2, then the general state of the composite system S has the form

$$\psi = \sum_{i,j} a_{ij}\xi_i\eta_j$$

with

$$\sum_{i,j} a_{ij}^* a_{ij} = 1$$

Suppose that $\{\xi_i\}$ and $\{\eta_j\}$ are complete sets of eigenfunctions of the observables (Hermitian operators) U and V, respectively, and let $\{u_i\}$ and $\{v_j\}$ be the corresponding eigenvalues. In the standard probabilistic interpretation, see Messiah (1969), $|a_{ij}|^2$ represents the joint probability that U takes on the value u_i in S_1 and V takes on the value v_j in S_2, i.e.,

$$|a_{ij}|^2 = a_{ij}^* a_{ij} = P(U = u_i,\ V = v_j) = P(u_i,\ v_j)$$

Generally, we do not know the numbers $\{a_{ij}\}$ and, therefore, the above joint probability is not determined. Suppose that we know the marginal probability distributions

$$P_1(U = u_i) = P_1(u_i), \qquad P_2(V = v_j) = P_2(v_j) \tag{1}$$

the mean values $\langle U\rangle$ and $\langle V\rangle$ of the two observables, their variances σ_U^2 and σ_V^2, and the correlation coefficient ρ between U and V. We have

$$\langle U\rangle = \sum_i u_i P_1(u_i), \qquad \langle V\rangle = \sum_j v_j P_2(v_j) \tag{2}$$

$$\sigma_U^2 = \sum_i (u_i - \langle U\rangle)^2 P_1(u_i), \qquad \sigma_V^2 = \sum_j (v_j - \langle V\rangle)^2 P_2(v_j) \tag{3}$$

The correlation coefficient is

$$\rho = c/\sigma_U\sigma_V$$

where c is the covariance between U and V, namely,

$$c = \sum_i \sum_j (u_i - \langle U\rangle)(v_j - \langle V\rangle)P(u_i, v_j) \tag{4}$$

The joint probability distribution $P(u_i, v_j)$ completely determines the marginal probability distributions $P_1(u_i)$ and $P_2(v_j)$ and the marginal and mixed moments $\langle U\rangle$, $\langle V\rangle$, σ_U^2, σ_V^2, and c, but the converse statement is not true. Knowing the marginal probability distributions (1) and the moments $\langle U\rangle$, $\langle V\rangle$, σ_U^2, σ_V^2, and c is not enough for uniquely determining the joint probability distribution $P(u_i, v_j)$. From the family of joint probability distributions compatible with the above-mentioned moments we select the probability distribution that is the closest one to the direct (independent) product of the two marginal probability distributions subject to the constraints imposed by $\langle U\rangle$, $\langle V\rangle$, σ_U^2, σ_V^2, and c. Such a joint probability distribution on the values taken on by U and V is the most random one, subject to these constraints. A good measure of the closeness between two probability distributions is the Kullback–Leibler (1951) number

$$I(P\colon P_1P_2) = \sum_i \sum_j P(u_i, v_j) \ln \frac{P(u_i, v_j)}{P_1(u_i)P_2(v_j)}$$

When we know nothing about the interdependence between the two observables, then the most random model for describing their joint behavior is obtained by supposing that they are independent. But when we know the correlation between the two observables, which is a second-order moment of the two random variables, then the most random model of their joint behavior is obtained by looking for the joint probability distribution that is the closest one to the independent direct product of the marginal distributions subject to the constraint on their dependence expressed by the given correlation. The philosophy behind this approach is similar to the principle of maximum entropy, where we determine the most random probability distribution subject to some known mean values of some random variables. This is a special form of Kullback's principle of minimum discrimination applied to joint probabilities. Details may be found in Guiasu, Leblanc, and Reischer (1982).

Therefore, we have to solve the variational problem

$$\min_P I(P\colon P_1P_2)$$

subject to the constraints (2)-(4) and

$$\sum_i \sum_j P(u_i, v_j) = 1 \tag{5}$$

Applying the standard Lagrange multipliers technique, we obtain the solution

$$P(u_i, v_j) = \frac{1}{\Phi(\beta_1, \ldots, \beta_5)} P_1(u_i) P_2(v_j)\, e^{-F_{ij}(\beta_1, \ldots, \beta_5)} \tag{6}$$

where

$$\begin{aligned} F_{ij}(\beta_1, \ldots, \beta_5) &= \beta_1 u_i + \beta_2 v_j + \beta_3 (u_i - \langle U \rangle)^2 \\ &\quad + \beta_4 (v_j - \langle V \rangle)^2 + \beta_5 (u_i - \langle U \rangle)(v_j - \langle V \rangle) \end{aligned}$$

$$\Phi(\beta_1, \ldots, \beta_5) = \sum_i \sum_j P_1(u_i) P_2(v_j)\, e^{-F_{ij}(\beta_1, \ldots, \beta_5)} \tag{7}$$

The Lagrange multipliers $\beta_1, \ldots, \beta_5$ must be determined from the system of equations obtained by introducing the solution (6) into the constraints (2)-(4), namely

$$\partial \ln \Phi / \partial \beta_1 = -\langle U \rangle, \qquad \partial \ln \Phi / \partial \beta_2 = -\langle V \rangle \tag{8}$$

$$\partial \ln \Phi / \partial \beta_3 = -\sigma_U^2, \qquad \partial \ln \Phi / \partial \beta_4 = -\sigma_V^2 \tag{9}$$

$$\partial \ln \Phi / \partial \beta_5 = -c \tag{10}$$

Particularly, if $c = 0$, which means that the observables U and V are not correlated, then the Kullback-Leibler number $I(P\colon P_1 P_2)$ is minimized by

$$P(u_1, v_2) = P_1(u_i) P_2(v_j) \tag{11}$$

which is compatible with the constraints (2)-(5).

3. SOME OBJECTIONS

Two main objections could be raised against the joint probability distribution (6). First, its expression is quite cumbersome and it is not easy at all to solve the system of exponential equations (8)-(10) for determining the Lagrange multipliers $\beta_1, \ldots, \beta_5$. Second, if S_1 and S_2 represent the same system and therefore U and V are two observables related to the same quantum system, then, according to the widely accepted Copenhagen interpretation of the quantum mechanics, the formula (6) is of no use because the corresponding operators U and V either commute, in which case they are independent and formula (11) completely characterizes their joint probability distribution, or they do not commute, in which case they cannot be simultaneously determined and therefore the utility of a joint probability distribution on their values could be questioned. When S_1 and S_2 are two distinct quantum systems, then the only case when the joint probability distribution (6) may be useful is when there is a weak interaction between the systems S_1 and S_2, or, more exactly, when the dependent observables U and V in the two systems are weakly correlated. Luckily, in such a case the cumbersome solution (6) may be well approximated by a very simple and elegant formula.

4. A QUADRATIC APPROXIMATION

Using the first two terms from Taylor's formula applied to e^x and e^{xy}, we get the approximations

$$e^x \simeq 1 + x, \qquad e^{xy} \simeq 1 + xy$$

accurate when x and xy, respectively, are close to zero. With these approximations, (7) becomes

$$\Phi = 1 - \beta_1 \langle U \rangle - \beta_2 \langle V \rangle - \beta_3 \sigma_U^2 - \beta_4 \sigma_V^2$$

and, from (8)-(10), we get

$$\beta_1 \langle U \rangle + \beta_2 \langle V \rangle + \beta_3 \sigma_U^2 + \beta_4 \sigma_V^2 = 0$$

which gives

$$\Phi = 1 \tag{12}$$

With the same approximations and taking into account (12), the expression (6) becomes

$$P(u_i, v_j) = [1 - F_{ij}(\beta_1, \ldots, \beta_5)] P_1(u_i) P_2(v_j) \tag{13}$$

which, introduced into (2) and (3), gives a linear system of four equations with four unknowns β_1, β_2, β_3, β_4, which has the solution

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

and (13) becomes

$$P(u_i, v_j) = [1 - \beta_5 (u_i - \langle U \rangle)(v_j - \langle V \rangle)] P_1(u_i) P_2(v_j) \tag{14}$$

Introducing this expression into the constraint (14), we get

$$\beta_5 = -\frac{c}{\sigma_U^2 \sigma_V^2} = -\frac{\rho}{\sigma_U \cdot \sigma_V}$$

which, introduced into (14), gives

$$P(u_i, v_j) = \left[1 + \rho \frac{u_i - \langle U \rangle}{\sigma_U} \frac{v_j - \langle V \rangle}{\sigma_V}\right] P_1(u_i) P_2(V_j) \tag{15}$$

Obviously, (15) is compatible with all the constraints (2)-(4) and with the marginals too. Also, when $\rho = 0$, (15) reduces to (11). It is not, however, always nonnegative. For the quadratic form

$$A(x, y) = \rho \frac{x - \langle U \rangle}{\sigma_U} \frac{y - \langle V \rangle}{\sigma_V}$$

the frontier of the set

$$\mathbf{N} = \{(x, y) \,|\, |A(x, y)| \le 1\}$$

is a pair of conjugate rectangular hyperbolas with the center of symmetry at the point $(\langle U \rangle, \langle V \rangle)$. The circle of maximum area, centered at $(\langle U \rangle, \langle V \rangle)$, which is entirely contained in the set $\mathbf{N}$, has the radius equal to $(2\sigma_U \sigma_V / |\rho|)^{1/2}$. Thus, for ρ different from zero, this circle has a positive area if and only if

$$\sigma_U \sigma_V \ge \varepsilon > 0$$

On the other hand, when $|\rho| \to 0$, the set $\mathbf{N}$ approaches the whole space R^2.

Suppose that, with a probability larger than $1 - \delta_1$, we have

$$m_1 \le u_i - \langle U \rangle \le M_1 \qquad (m_1 \le 0 \le M_1) \tag{16}$$

and, with a probability larger than $1 - \delta_2$, we have

$$m_2 \le v_j - \langle V \rangle \le M_2 \qquad (m_2 \le 0 \le M_2) \tag{17}$$

Let us take into account only those values of i and j for which (16) and (17) hold. Denote

$$m = \min\{M_1 m_2, m_1 M_2\}, \qquad M = \max\{m_1 m_2, M_1 M_2\}$$

$$K = \max\{|m|, M\}$$

Then, with an error smaller than $\delta_1+\delta_2$, (15) is a probability distribution if

$$|\rho| \leq \sigma_U \sigma_V / K$$

5. EXAMPLE

Let us take two harmonic oscillators and as observables U and V we consider their energies E_1 and E_2, respectively. The possible values of these observables are

$$\begin{aligned} E_{1,i} &= h\nu_1(i+\tfrac{1}{2}), \qquad i=0,1,2,\ldots \\ E_{2,j} &= h\nu_2(j+\tfrac{1}{2}), \qquad j=0,1,2,\ldots \end{aligned} \tag{18}$$

respectively, where ν_1 and ν_2 are two frequencies and h is Planck's constant. Suppose that except for the set of possible values (18), we know only the mean energy $\langle E_s \rangle$ of the harmonic oscillator s $(s=1,2)$. Then, using the von Neumann–Jaynes approach, we can determine the probability distribution

$$P_s(E_{s,k})>0, \qquad \sum_{k=0}^{\infty} P_s(E_{s,k})=1$$

which maximizes the corresponding entropy

$$H(P_s) = -\sum_{k=0}^{\infty} P_s(E_{s,k}) \ln P_s(E_{s,k})$$

subject to the constraint

$$\sum_{k=0}^{\infty} E_{s,k} P_s(E_{s,k}) = \langle E_s \rangle$$

Using the Lagrange multipliers method, we obtain the solution

$$P_s(E_{s,k}) = \frac{h\nu_s(\langle E_s \rangle - \frac{1}{2}h\nu_s)^k}{(\langle E_s \rangle + \frac{1}{2}h\nu_s)^{k+1}}, \qquad k=0,1,2,\ldots$$

The variance of such a probability distribution may be calculated without any difficulty and we get

$$\sigma_s^2 = (\langle E_s \rangle - 3h\nu/2)(\langle E_s \rangle - h\nu_s/2)$$

If

$$\langle E_s \rangle = h\nu_s(N_s + \tfrac{1}{2}), \qquad s = 1, 2$$

then we obtain

$$P_s(E_{s,k}) = \frac{1}{N_s+1}\left(\frac{N_s}{N_s+1}\right)^k$$

and

$$\sigma_s^2 = (h\nu_s)^2 N_s(N_s - 1)$$

In such a case

$$E_{1,i} - \langle E_1 \rangle = h\nu_1(1 - N_1), \qquad E_{2,j} - \langle E_2 \rangle = h\nu_2(j - N_2)$$

and

$$\sum_{k=n_s}^{\infty} P_s(E_{s,k}) = \frac{1}{N_s+1} \sum_{k=n_s}^{\infty} \left(\frac{N_s}{N_s+1}\right)^k = \left(\frac{N_s}{N_s+1}\right)^{n_s}$$

Let us take

$$n_s = \max\left\{\frac{\ln \delta_s}{\ln[N_s/(N_s+1)]}, 2N_s\right\}$$

where δ_s, $s = 1, 2$, are arbitrary (small) positive numbers, and let us put

$$M_s = h\nu_s(n_s - N_s), \qquad m_s = -h\nu_s N_s$$

If the correlation coefficient between the energies of the two harmonic oscillators satisfies

$$|\rho| < \frac{[N_1(N_1-1)N_2(N_2-1)]^{1/2}}{(n_1 - N_1)(n_2 - N_2)}$$

then, with an error smaller than $\delta_1 + \delta_2$, the joint probability distribution for the energies of the two harmonic oscillators is

$$P(E_{1,i}, E_{2,j}) = \left\{1 + \frac{(i - N_1)(j - N_2)}{[N_1(N_1-1)N_2(N_2-1)]^{1/2}}\right\} \times \frac{1}{N_1+1}\left(\frac{N_1}{N_1+1}\right)^i \frac{1}{N_2+1}\left(\frac{1}{N_2+1}\right)^j$$

Numerical Example. Let us take

$$\delta_1 = \delta_2 = 0.025, \qquad N_1 = 50 \text{ ergs}, \qquad N_2 = 20 \text{ ergs}, \qquad \rho = 0.3$$

$$h = 6.62 \times 10^{-27} \text{ erg sec}$$

$$\nu_1 = 3 \times 10^{18} \text{ sec}^{-1}, \qquad \nu_2 = 2 \times 10^{18} \text{ sec}^{-1}$$

In such a case, (19) is a probability distribution for any correlation coefficient $|\rho| < 0.3081$. We have $n_1 = 187$ and $n_2 = 76$. With an error smaller than 0.05

we have for the joint probability

$$P(E_1 = 19.86(i+\tfrac{1}{2}),\ E_2 = 13.28(j+\tfrac{1}{2}))$$
$$= 0.0009[1+0.0003(i-50)(j-20)](0.9804)^i(0.9524)^j$$

for $i = 0, 1, \ldots, 187$; $j = 0, 1, \ldots, 76$.

6. REMARKS

1. The formula (15) for the joint probability distribution may be generalized for n systems in weak interaction, namely

$$P(u_{1,i_1}, \ldots, u_{n,i_n}) = \left[1 + \sum_{j\neq k} \rho_{jk} \frac{u_{j,i_j} - \langle U_j\rangle}{\sigma_{Uj}} \frac{u_{k,i_k} - \langle U_k\rangle}{\sigma_{U_k}}\right] P_1(u_{1,i_1}) \cdots P_n(u_{n,i_n})$$

2. Another possible approach is the following: instead of approximating the exact solution (6), we approximate the Kullback–Leibler indicator $I(P\colon P_1P_2)$ by

$$D(P\colon P_1P_2) = \sum_{i,j} \frac{P^2(u_i, v_j)}{P_1(u_i)P_2(v_j)} - 1$$

and minimize $D(P\colon P_1P_2)$ subject to (2)–(5) and to the additional constraint

$$P(u_i, v_j) \geq 0, \qquad \text{for all } i \text{ and } j$$

assuming that both U and V take on only a finite number of possible values. Such a nonlinear program with equality and inequality constraints can be solved using Kuhn–Tucker's method and we get again (15) as the solution of the problem.

7. CONCLUSIONS

There are many objections against even talking about a joint probability distribution for two observables attached to the *same* quantum system whose corresponding operators do not commute. Apparently, such a joint probability distribution may be considered when we are dealing with two weakly correlated observables attached to two different systems in weak interaction. The paper deals with the construction of a probabilistic model for such a case, using as available data only first- and second-order moments of some random variables (observables) accessible at the macroscopic level as a consequence of an incomplete measurement process. When the only information is supplied by the mean values and the variances of two observables and by the covariance (correlation coefficient) between them, the solution is obtained by minimizing the divergence from independence. The model thus obtained is the most random one subject to the above-mentioned constraints. Fortunately, just for weakly correlated observables from two systems in weak interaction, for which apparently the concept of joint distribution has a sense, the exact solution can be approximated by a quadratic function that can be easily used in computation.

ACKNOWLEDGMENT

The author is indebted to NSERC Canada for research grant A-5712.

REFERENCES

Guiasu, S. (1977). *Information Theory with Applications*, McGraw-Hill, New York.

Guiasu, S., and Shenitzer, A. (1985). The principle of maximum entropy, *The Mathematical Intelligencer*, **7**, 42-48.

Guiasu, S., Leblanc, R., and Reischer, C. (1982). On the principle of minimum interdependence, *Journal of Information and Optimization Sciences*, **3**, 149-172.

Jaynes, E. T. (1957). Information theory and statistical mechanics, *Physical Review*, Part I, **106**, 620-630; Part II, **108**, 171-190.

Kullback, S., and Leibler, R. A. (1951). On information and sufficiency, *Annals of Mathematical Statistics*, **22**, 79-86.

Messiah, A. (1969). *Mechanique Quantique*, Vol. 1, Dunod, Paris.

Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Technical Journal*, Part I, **27**, 379-423; Part II, **27**, 623-656.

Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag, Heidelberg.