The Annals of Mathematical Statistics
1968, Vol. 39, No. 4, 1354–1357

CONSTRUCTION OF JOINT PROBABILITY DISTRIBUTIONS

BY L. F. KEMP, JR.

Polytechnic Institute of Brooklyn

1. Introduction. In (1), (2), (4), (5), and (6), there are constructions of joint probability distributions having given marginal distributions. By a generalization of an exercise from Halmos' *Measure Theory*, we construct a class of doubly stochastic measures as the Lebesgue integral of special simple and elementary functions whose values must be positive and satisfy a dependent system of linear equations. Employing this construction, we get an additional method for generating joint distributions with given marginal distributions.

Let $F_1(x)$ and $F_2(y)$ be the distribution functions of two random variables. Frechet proved that the family of joint distributions having $F_1(x)$ and $F_2(y)$ as marginal distributions collapses to $F_1(x)F_2(y)$ if and only if either $F_1(x)$ or $F_2(y)$ is a unit step function. We rephrase his result in terms of abstract probability measures and with the aid of the above construction of doubly stochastic measures, we show his result is equivalent to the statement that Cartesian product measure is an extreme point of the set of doubly stochastic measures.

2. The construction. (X, S, μ) and (Y, T, ν) are two abstract probability triples. $(X \times Y, S \times T)$ is the Cartesian cross product measure space of the measure spaces (X, S) and (Y, T). λ is called a *doubly stochastic* measure on $(X \times Y, S \times T)$ if

$$\lambda(A \times Y) = \mu(A), \qquad \text{for all} \quad A \quad \text{in} \quad S;$$

$$\lambda(X \times B) = \nu(B), \qquad \text{for all} \quad B \quad \text{in} \quad T.$$

Cartesian product measure $\mu \times \nu$ is a doubly stochastic measure.

THEOREM 1. *Let $\{A_i\}$ and $\{B_j\}$ be finite or countably infinite measurable partitions of X and Y, respectively, then the set function*

$$\lambda(E) = \int_E \sum_{i,j} \alpha_{ij} K_{A_i \times B_j}\, d(\mu \times \nu), \qquad \text{for} \quad E \quad \text{in} \quad S \times T$$

is a doubly stochastic measure, if and only if,

(1) $\alpha_{ij} \geqq 0$, all i, j;
(2) $\sum_i \alpha_{ij}\mu(A_i) = 1$, for every j;
(3) $\sum_j \alpha_{ij}\nu(B_j) = 1$, for every i;

where $K_{A_i \times B_j}$ is the characteristic function of the rectangle $A_i \times B_j$.

PROOF. Before we begin the proof, let us assume without loss of generality that $\mu(A_i)\nu(B_j) > 0$ all i, j since $\lambda \ll \mu \times \nu$.

Sufficiency: If $\alpha(x, y) = \sum_{i,j} \alpha_{ij} K_{A_i \times B_j}(x, y)$, then $\alpha(x, y)$ is a positive measurable elementary function so that λ is a measure. If $\alpha_n(x, y) =$

Received 21 September 1967.

$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} K_{A_i \times B_j}(x, y)$, then each α_n is a simple function and α_n converges everywhere to α. Since α is the limit of a monotone increasing sequence of simple functions α_n, which converge everywhere to α, we have that

$$
\begin{aligned}
&|\lambda(A \times Y) - \mu(A)| \\
&\quad = |\textstyle\int_{A \times Y} \alpha \, d(\mu \times \nu) - \mu(A)| \\
&\quad \leqq |\textstyle\int_{A \times Y} \alpha \, d(\mu \times \nu) - \int_{A \times Y} \alpha_n \, d(\mu \times \nu)| + |\int_{A \times Y} \alpha_n \, d(\mu \times \nu) - \mu(A)| \\
&\quad = |\textstyle\int_{A \times Y} \alpha \, d(\mu \times \nu) - \int_{A \times Y} \alpha_n \, d(\mu \times \nu)| + \mu(A) \\
&\qquad - \textstyle\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} \mu(AA_i) \nu(B_j) \\
&\quad \leqq |\textstyle\int_{A \times Y} \alpha \, d(\mu \times \nu) - \int_{A \times Y} \alpha_n \, d(\mu \times \nu)| + \mu(A) \\
&\qquad - \textstyle\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} \mu(AA_i) \nu(B_j) \qquad \text{with } m \leqq n \\
&\quad \leqq |\textstyle\int_{A \times Y} \alpha \, d(\mu \times \nu) - \int_{A \times Y} \alpha_n \, d(\mu \times \nu)| + |\mu(A) \\
&\qquad - \textstyle\sum_{i=1}^{m} \mu(AA_i)| + |\sum_{i=1}^{m} (\mu(AA_i)(\sum_{j=1}^{n} \alpha_{ij} \nu(B_j) - 1))| \\
&\quad \leqq |\textstyle\int_{A \times Y} \alpha \, d(\mu \times \nu) - \int_{A \times Y} \alpha_n \, d(\mu \times \nu)| + |\mu(A) \\
&\qquad - \mu(\textstyle\bigcup_{i=1}^{m} AA_i)| + \max_{1 \leqq i \leqq m} (1 - \sum_{j=1}^{n} \alpha_{ij} \nu(B_j)).
\end{aligned}
$$

For a given arbitrary $\epsilon > 0$, choose and fix m such that the middle term is less than $\epsilon/3$; then the preceding expression is less than

$$|\textstyle\int_{A \times Y} \alpha \, d(\mu \times \nu) - \int_{A \times Y} \alpha_n \, d(\mu \times \nu)| + \epsilon/3 + \max_{1 \leqq i \leqq m} (1 - \sum_{j=1}^{n} \alpha_{ij} \nu(B_j)).$$

Now choose n such that $n \geqq m$ and the other two terms are each less than $\epsilon/3$; hence,

$$|\lambda(A \times Y) - \mu(A)| < \epsilon/3 + \epsilon/3 + \epsilon/3.$$

Similarly, $|\lambda(X \times B) - \nu(B)| < \epsilon$, so that λ is doubly stochastic.
Necessity: λ is a measure implies $\alpha_{ij} = \lambda(A_i \times B_j)/\mu(A_i)\nu(B_j) \geqq 0$ all i, j. To show (2) and (3) consider

$$
\begin{aligned}
&|\textstyle\sum_{i=1}^{n} \alpha_{ij} \mu(A_i) - 1| \\
&\quad = |\textstyle\sum_{i=1}^{n} \alpha_{ij} \mu(A_i)(\sum_{j=1}^{n} \nu(B_j) + \sum_{j=n+1}^{\infty} \nu(B_j)) - 1| \\
&\quad \leqq 1 - \lambda(\textstyle\bigcup_{i=1}^{n} \bigcup_{j=1}^{n} A_i \times B_j) + \sum_{i=1}^{n} \alpha_{ij} \mu(A_i)(\sum_{j=n+1}^{\infty} \nu(B_j)) \\
&\quad = 1 - \lambda(\textstyle\bigcup_{i=1}^{n} \bigcup_{j=1}^{n} A_i \times B_j) + (\sum_{j=n+1}^{\infty} \nu(B_j))(\lambda(\bigcup_{i=1}^{n} A_i \times B_j)/\nu(B_j)) \\
&\quad \leqq 1 - \lambda(\textstyle\bigcup_{i=1}^{n} \bigcup_{j=1}^{n} A_i \times B_j) + (\nu(B_j))^{-1} \sum_{j=n+1}^{\infty} \nu(B_j) \\
&\quad < \epsilon/2 + \epsilon/2, \qquad \text{for all } \; n > N(\epsilon, j).
\end{aligned}
$$

Similarly, $|\sum_{j=1}^{n} \alpha_{ij} \nu(B_j) - 1| < \epsilon$ for all $n > N(\epsilon, i)$, which ends the proof.

APPLICATION. If $f_1(x)$ and $f_2(y)$ are the probability density functions of the random variables X and Y, respectively, then the function $f(x, y) =$

$\alpha(x, y)f_1(x)f_2(y)$ is a joint probability density function for the random variables X and Y if $\alpha(x, y)$ is a step function satisfying

(1) $\alpha(x, y) = \alpha_{ij} \geqq 0$, (x, y) in $(x_{i-1}, x_i] \times (y_{j-1}, y_j]$;
(2) $\sum_{i=1}^{\infty} \alpha_{ij} \int_{x_{i-1}}^{x_i} f_1(x)\,dx = \int_{-\infty}^{\infty} \alpha(x, y)f_1(x)\,dx = 1$;
(3) $\sum_{j=1}^{\infty} \alpha_{ij} \int_{y_{j-1}}^{y_j} f_2(y)\,dy = \int_{-\infty}^{\infty} \alpha(x, y)f_2(y)\,dy = 1$;

where

$$-\infty = x_0 < x_1 < \cdots < x_n \to \infty,$$

$$-\infty = y_0 < y_1 \cdots < y_n \to \infty.$$

EXAMPLE. $f_1(x) = \exp(-x^2/2)/(2\pi)^{\frac{1}{2}}$, $f_2(y) = \exp(-y^2/2)/(2\pi)^{\frac{1}{2}}$,

$$f(x, y) = 2f_1(x)f_2(y) = \exp(-\tfrac{1}{2}(x^2 + y^2))/\pi \quad (x, y) \text{ in 1st or 3rd quadrants,}$$
$$= 0 \quad (x, y) \text{ in 2nd or 4th quadrants.}$$

3. Frechet's theorem. We will call a probability measure μ trivial if for every measurable set A in S either $\mu(A) = 0$ or $\mu(A) = 1$.

THEOREM 2. *Cartesian product measure $\mu \times \nu$ is the only doubly stochastic measure, if and only if, either μ or ν is trivial.*

PROOF. Sufficiency: Suppose μ is trivial, then either $\mu \times \nu(A \times B) = 0$ or $\mu \times \nu(A \times B) = \nu(B)$ for any measurable rectangle $A \times B$. When $\mu \times \nu(A \times B) = 0$, $\lambda(A \times B) = 0$, and when $\mu \times \nu(A \times B) = \nu(B)$, $\lambda(A \times B) = \nu(B)$. Therefore, every doubly stochastic measure λ agrees with $\mu \times \nu$ on the class of all measurable rectangles. Furthermore, we have agreement on the ring R of all finite disjoint unions of measurable rectangles which implies by the uniqueness of the extension of the measure $\mu \times \nu$ to the σ-algebra $S(R) = S \times T$ that the only doubly stochastic measure is $\mu \times \nu$.
Necessity: If μ and ν are both non-trivial, then there exist measurable sets $A_1, A_2 = A_1^c, B_1, B_2 = B_1^c$ such that $0 < \mu(A_i) < 1, i = 1, 2$ and $0 < \nu(B_j) < 1$, $j = 1, 2$. If λ is the set function defined by

$$\lambda = \int \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha_{ij} K_{A_i \times B_j}\, d(\mu \times \nu);$$

(i) $\alpha_{11} = (\mu(A_1) + \nu(B_1) - 1 + \alpha_{22}\mu(A_2)\nu(B_2))/\mu(A_1)\nu(B_1)$;
(ii) $\alpha_{12} = (1 - \alpha_{22}\mu(A_2))/\mu(A_1)$;
(iii) $\alpha_{21} = (1 - \alpha_{22}\nu(B_2))/\nu(B_1)$;
$\alpha_{22} = \min(1/\mu(A_2), 1/\nu(B_2))$, then λ is doubly stochastic and $\lambda \neq \mu \times \nu$.

THEOREM 3. *Cartesian product measure $\mu \times \nu$ is an extreme point of the set of doubly stochastic measures if and only if $\mu \times \nu$ is the only doubly stochastic measure.*

PROOF. Sufficiency: If $\mu \times \nu$ is the only doubly stochastic measure, $\mu \times \nu$ is trivially an extreme point of the set of doubly stochastic measures.
Necessity: We assume that the set of doubly stochastic measures contains more than $\mu \times \nu$ and show that $\mu \times \nu$ can be expressed as a convex combination of two different doubly stochastic measures. By Theorem 2, there exist sets $A_1, A_2 = A_1^c$

in S and B_1, $B_2 = B_1^c$ in T with $0 < \mu(A_2) < 1, 0 < \nu(B_2) < 1$. Let

$$\alpha'_{22} = \max(0, 1 - (1 - 1/\mu(A_2))(1 - 1/\nu(B_2)));$$
$$\alpha''_{22} = \min(1/\mu(A_2), 1/\nu(B_2));$$
$$\lambda' = \int \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha'_{ij} K_{A_i \times B_j} \, d(\mu \times \nu);$$
$$\lambda'' = \int \sum_{i=1}^{2} \sum_{j=1}^{2} \alpha''_{ij} K_{A_i \times B_j} \, d(\mu \times \nu);$$

where the α'_{ij} and α''_{ij} satisfy (i), (ii), and (iii) of Theorem 2 when the α_{ij} of Theorem 2 are primed and double primed, respectively. λ' and λ'' are different doubly stochastic measures both unequal to $\mu \times \nu$. If $\alpha = (\alpha''_{22} - 1)/(\alpha''_{22} - \alpha'_{22})$ then $0 < \alpha < 1$ and $\mu \times \nu(A \times B) = \alpha\lambda'(A \times B) + (1 - \alpha)\lambda''(A \times B)$ for all $A \times B$ in $S \times T$. It follows $\mu \times \nu = \alpha\lambda' + (1 - \alpha)\lambda''$ on the ring R of all finite disjoint unions of measurable rectangles, and hence by the unique extension of measures that

$$\mu \times \nu = \alpha\lambda' + (1 - \alpha)\lambda'' \quad \text{on} \quad S(R) = S \times T.$$

REFERENCES

[1] Frechet, M. (1951). Sur les tableaux de correlation dont les marges sont donnees. *Ann. Univ. Lyon, Sect. A* (3) **14** 66–77.

[2] Gumbel, E. J. (1958). Distributions a plusieurs variables dont les marges sont donnees. *C. R. Acad. Sci., Paris.* **246** 2717–2719.

[3] Halmos, P. R. (1950). *Measure Theory.* Van Nostrand, New York.

[4] Nataf, Andre. (1962). Determination des distributions de probabilites dont les marges sont donnees. *C. R. Acad. Sci., Paris.* **255** 42–43.

[5] Parzen, Emmanuel. (1960). *Modern Probability Theory and Its Applications.* Wiley, New York.

[6] Rao, K. M. (1966). An imbedding of a stochastic process $\{X_n\}$ in a Markov chain. *J. Math. Mech.* **15** 925–935.