MIXTURE OF DISTRIBUTIONS

By Herbert Robbins

Department of Mathematical Statistics, University of North Carolina

1. Summary. Mixtures of measures or distributions occur frequently in the theory and applications of probability and statistics. In the simplest case it may, for example, be reasonable to assume that one is dealing with the mixture in given proportions of a finite number of normal populations with different means or variances. The mixture parameter may also be denumerably infinite, as in the theory of sums of a random number of random variables, or continuous, as in the compound Poisson distribution.

The operation of Lebesgue-Stieltjes integration, $\int f(x)\, d\mu$, is linear with respect to both integrand $f(x)$ and measure μ. The first type of linearity has as its continuous analog the theorem of Fubini on interchange of order of integration; the second type of linearity has a corresponding continuous analog which is of importance whenever one deals with mixtures of measures or distributions, and which forms the subject of the present paper. Other treatments of the same subject have been given ([1], [2]; see also [3], [4]) but it is hoped that the discussion given here will be useful to the mathematical statistician.

A general measure theoretic form of the fundamental theorem is given in Section 2, and in Section 3 the theorem is formulated in terms of finite dimensional spaces and distribution functions. The operation of convolution as an example of mixture is treated briefly in Section 4, while Section 5 is devoted to random sampling from a mixed population.

We shall refer to *Theory of the Integral* by S. Saks (second edition, Warszawa, 1937) as [*S*], and the *Mathematical Methods* of *Statistics* by H. Cramér (Princeton, 1946) as [*C*].

2. Mixture of measures in general. Let $X(Y)$ be a space with points $x(y)$ and let $\mathfrak{X}(\mathfrak{Y})$ be a σ-field of subsets of $X(Y)$. Let ν be a measure on $\mathfrak{Y}$. Let μ_y be for a. e. (ν) y a measure on $\mathfrak{X}$, such that $\mu_y(S)$ is for every S in $\mathfrak{X}$ a measurable $(\mathfrak{Y})$ function of y. Define for every S in $\mathfrak{X}$,

$$\mu(S) = \int_Y \mu_y(S)\, d\nu. \tag{1}$$

Theorem 1. *μ is a measure on $\mathfrak{X}$. If $\nu(Y) = \mu_y(X) = 1$, then $\mu(X) = 1$.*

Proof. Clear.

Theorem 2. *If $f(x)$ is any non-negative or non-positive function measurable $(\mathfrak{X})$ then the function*

$$g(y) = \int_X f(x)\, d\mu_y \tag{2}$$

is measurable ($\mathfrak{Y}$), and

$$\int_X f(x)\, d\mu = \int_Y g(y)\, d\nu. \tag{3}$$

PROOF. First let $f_0(x)$ be any non-negative simple function [S, p. 7] of the form

$$f_0(x) = \{a_1, S_1; \cdots; a_k, S_k\} \tag{4}$$

where the S_i are disjoint sets in $\mathfrak{X}$ such that $X = \sum_1^k S_i$ and the a_i are non-negative constants. Then

$$g_0(y) = \int_X f_0(x)\, d\mu_y = \sum_1^k a_i\, \mu_y(S_i) \tag{5}$$

is a non-negative function measurable ($\mathfrak{Y}$), and from (1) it follows that each side of (3) is equal to $\sum_1^k a_i\mu(S_i)$. Hence the theorem holds in this case.

Next let $f(x)$ be any non-negative function measurable ($\mathfrak{X}$); then [S, p. 14] there exists a sequence $f_n(x)$ of simple functions such that for every x,

$$0 \leq f_1(x) \leq f_2(x) \leq \cdots; \qquad \lim_{n\to\infty} f_n(x) = f(x). \tag{6}$$

Setting

$$g_n(y) = \int_X f_n(x)\, d\mu_y, \qquad g(y) = \int_X f(x)\, d\mu_y, \tag{7}$$

it follows from the theorem of monotone convergence [S, p. 28] and from the preceding paragraph that

$$\int_X f(x)\, d\mu = \lim_{n\to\infty} \int_X f_n(x)\, d\mu = \lim_{n\to\infty} \int_Y g_n(y)\, d\nu, \tag{8}$$

$$g(y) = \lim_{n\to\infty} \int_X f_n(x)\, d\mu_y = \lim_{n\to\infty} g_n(y). \tag{9}$$

From (6) and (9) it follows that for a.e. $(\nu)y$,

$$0 \leq g_1(y) \leq g_2(y) \leq \cdots; \qquad \lim_{n\to\infty} g_n(y) = g(y). \tag{10}$$

Hence $g(y)$ is measurable ($\mathfrak{Y}$), and from the theorem of monotone convergence,

$$\int_Y g(y)\, d\nu = \lim_{n\to\infty} \int_Y g_n(y)\, d\nu. \tag{11}$$

Equation (3) now follows from (8) and (11).

By passing from $f(x)$ to $-f(x)$ we establish (3) when $f(x)$ is any non-positive function measurable ($\mathfrak{X}$). This completes the proof of Theorem 2.

If $f(x)$ is an arbitrary function measurable ($\mathfrak{X}$) we define

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}, \qquad f^-(x) = \begin{cases} f(x) & \text{if } f(x) \leq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{12}$$

so that

$$f(x) = f^+(x) + f^-(x) \tag{13}$$

is the sum of two functions measurable ($\mathfrak{X}$) of constant sign. By Theorem 2 the functions

$$g_1(y) = \int_X f^+(x)\, d\mu_y, \qquad g_2(y) = \int_X f^-(x)\, d\mu_y \tag{14}$$

are measurable ($\mathfrak{Y}$) and

$$0 \le \int_X f^+(x)\, d\mu = \int_Y g_1(y)\, d\nu \le \infty, \tag{15}$$

$$0 \ge \int_X f^-(x)\, d\mu = \int_Y g_2(y)\, d\nu \ge -\infty. \tag{16}$$

The integral $\int_X f(x)\, d\mu$ exists if and only if at least one of the two quantities (15) and (16) is finite [S, p. 20].

THEOREM 3. *A necessary and sufficient condition that*

$$\int_X f(x)\, d\mu = \int_Y \left\{ \int_X f(x)\, d\mu_y \right\} d\nu \tag{17}$$

is that at least one of the two quantities (15) *and* (16) *be finite.*

PROOF. By the remark preceding Theorem 3 the condition is clearly necessary. Now suppose, e.g., that (15) is finite; we must show that (17) holds. By hypothesis,

$$\int_X f^+(x)\, d\mu < \infty, \qquad \int_X f(x)\, d\mu = \int_X f^+(x)\, d\mu + \int_X f^-(x)\, d\mu. \tag{18}$$

From (18) and (15) it follows that $0 \le g_1(y) < \infty$ for a.e. $(\nu)y$; hence

$$\int_X f(x)\, d\mu_y = \int_X f^+(x)\, d\mu_y + \int_X f^-(x)\, d\mu_y = g_1(y) + g_2(y) \tag{19}$$

exists for a.e. $(\nu)y$. From the finiteness of (15) it follows that

$$\int_Y (g_1(y) + g_2(y))\, d\nu = \int_Y g_1(y)\, d\nu + \int_Y g_2(y)\, d\nu \tag{20}$$

exists. Hence from (19), the integral

$$\int_Y \left\{ \int_X f(x)\, d\mu_y \right\} d\nu = \int_Y (g_1(y) + g_2(y))\, d\nu \tag{21}$$

exists. Equation (17) now follows from (21), (20), (15), and (18). This completes the proof of Theorem 3.

COROLLARY 1. *If $\mu(X) < \infty$, and if $f(x)$ is bounded from above or from below, then both sides of* (17) *exist and the equality holds.*

PROOF. If, say, $f(x) \leq C < \infty$, then

$$0 \leq \int_X f^+(x)\, d\mu \leq C \cdot \mu(X) < \infty,$$

and the result follows from Theorem 3.

We shall now show by an example that the existence and even the finiteness of the right side of (17) does not imply the existence of the left side.

Let $X = Y = \{1, 2, \cdots, n, \cdots\}$ and let $\mathfrak{X}(\mathfrak{Y})$ consist of all subsets of $X(Y)$. Let ν be the measure which assigns mass c_n to n, where the c_n are positive constants such that $\sum_1^\infty c_n = 1$. Let μ_n assign the mass $1/2n$ to each of the points $1, 2, \cdots, 2n$. Let $f(x)$ be such that $f(1) = b_1, f(2) = -b_1, f(3) = b_2, f(4) = -b_2, \cdots$ where the b_n are positive constants. Then

$$\int_X f(x)\, d\mu_n = 0 \qquad (n = 1, 2, \cdots),$$

so that

$$\int_Y \left\{ \int_X f(x)\, d\mu_n \right\} d\nu = 0.$$

The measure μ defined by (1) assigns to each n a positive value $\mu(n)$ given by

$$\mu(1) = \mu(2) = c_1 \cdot (2)^{-1} + c_2 \cdot (2 \cdot 2)^{-1} + c_3 \cdot (2 \cdot 3)^{-1} + \cdots$$

$$\mu(3) = \mu(4) = c_2 \cdot (2 \cdot 2)^{-1} + c_3 \cdot (2 \cdot 3)^{-1} + \cdots$$

$$\cdots$$

where $\mu(X) = \sum_1^\infty \mu(n) = \sum_1^\infty c_n = 1$.

Now fix the b_n and c_n in such a way that

$$b_1 \cdot \mu(1) + b_2 \cdot \mu(3) + b_3 \cdot \mu(5) + \cdots = \infty.$$

Then

$$\int_X f^+(x)\, d\mu = -\int_X f^-(x)\, d\mu = \infty,$$

so that the left side of (17) does not exist, even though $\nu(Y) = \mu_y(X) = \mu(X) = 1$ and the right side of (17) exists and is equal to zero.

3. A restatement of the preceding results in the form most useful in probability theory. Let $x = (x_1, \cdots, x_n)$ be a point in the n-dimensional Euclidean space R_n, and let B_n denote the σ-field of Borel sets in R_n. Let S_x denote the half-open interval in R_n consisting of all points $(w_1, \cdots, w_n)$ in R_n satisfying the inequalities

$$w_1 \leq x_1, \cdots, w_n \leq x_n; \tag{22}$$

then if μ is any probability measure on B_n the function

$$F(x) = \mu(S_x) \tag{23}$$

is the distribution function corresponding to μ. Conversely, if $F(x)$ is any distribution function in R_n [C, p. 80] there is a unique probability measure μ on B_n such that (23) holds. As a matter of notation we write for any Borel measurable $f(x)$,

$$\int_{R_n} f(x)\, d\mu = \int_{-\infty}^{\infty} f(x)\, dF(x) \tag{24}$$

provided the integral on the left exists.

Now let $y = (y_1, \cdots, y_m)$ be a point in R_m, let $G(y)$ be a distribution function, and let ν denote the corresponding probability measure on B_m. Let $F(x,y)$ be for a.e. $(\nu)y$ a distribution function in x, and for every x a Borel measurable function of y, and let μ_y be the corresponding probability measure on B_n.

THEOREM 4. *The function*

$$H(x) = \int_{-\infty}^{\infty} F(x, y)\, dG(y) \tag{25}$$

is a distribution function in R_n. Let μ denote the corresponding probability measure on B_n. Then for any S in B_n, $\mu_y(S)$ is a Borel measurable function of y and

$$\mu(S) = \int_{-\infty}^{\infty} \mu_y(S)\, dG(y). \tag{26}$$

PROOF. Let C denote the class of all Borel sets S in R_n such that $\mu_y(S)$ is a Borel measurable function of y. We shall show that C is a normal class [S, p. 83].

(i) If $S_1, S_2, \cdots$ is a sequence of disjoint sets in C and if $S = \sum_1^\infty S_n$, then

$$\mu_y(S) = \mu_y\left(\sum_1^\infty S_n\right) = \sum_1^\infty \mu_y(S_n)$$

is a convergent series of Borel measurable functions and is therefore itself a Borel measurable function.

(ii) If $S_1 \supset S_2 \supset \cdots$ is a decreasing sequence of sets in C and if $S = \prod_1^\infty S_n$, then

$$\mu_y(S) = \mu_y\left(\prod_1^\infty S_n\right) = \lim_{n\to\infty} \mu_y(S_n)$$

is the limit of a sequence of Borel measurable functions and is therefore a Borel measurable function.

Hence C is a normal class. But C contains every interval S_x, for $\mu_y(S_x) = F(x, y)$ was assumed to be a Borel measurable function of y for every x. It follows [S, p. 85] that $C = B_n$.

It now follows from Theorem 1 that the set function $\mu(S)$ defined by (26) is a probability measure on B_n. The corresponding distribution function is the function $H(x)$ defined by (25). Thus Theorem 4 is proved.

Let $f(x) = f^+(x) + f^-(x)$ be any Borel measurable function. Then from Theorem 2, the integrals

$$\begin{aligned}\int_{-\infty}^{\infty} f^+(x)\, dH(x) &= \int_{-\infty}^{\infty} f^+(x)\, d_x\left\{ \int_{-\infty}^{\infty} F(x, y)\, dG(y) \right\} \\ &= \int_{-\infty}^{\infty}\left\{ \int_{-\infty}^{\infty} f^+(x)\, d_x F(x, y) \right\} dG(y),\end{aligned} \tag{27}$$

$$\begin{aligned}\int_{-\infty}^{\infty} f^-(x)\, dH(x) &= \int_{-\infty}^{\infty} f^-(x)\, d_x\left\{ \int_{-\infty}^{\infty} F(x, y)\, dG(y) \right\} \\ &= \int_{-\infty}^{\infty}\left\{ \int_{-\infty}^{\infty} f^-(x)\, d_x F(x, y) \right\} dG(y)\end{aligned} \tag{28}$$

exist. The following theorem is an immediate consequence of Theorem 3 and Corollary 1.

THEOREM 5. *A necessary and sufficient condition that*

$$\int_{-\infty}^{\infty} f(x)\, d_x\left\{ \int_{-\infty}^{\infty} F(x, y)\, dG(y) \right\} = \int_{-\infty}^{\infty}\left\{ \int_{-\infty}^{\infty} f(x)\, d_x F(x, y) \right\} dG(y) \tag{29}$$

is that the left side of (29) *exist; i.e. that at least one of the quantities* (27) *and* (28) *be finite. This will be true in particular if* $f(x)$ *is bounded from above or from below.*

4. The operation of convolution. An example of the general mixture (25) of distribution functions is the operation of convolution: if $F(x)$, $G(x)$ are two distribution functions in R_1 then $F(x, y) = F(x - y)$ satisfies the conditions of Theorem 4, so that

$$H(x) = \int_{-\infty}^{\infty} F(x - y)\, dG(y) \tag{30}$$

is also a distribution function in R_1, denoted by

$$H(x) = F(x) * G(x). \tag{31}$$

Corresponding to any distribution function $F(x)$ in R_1 is the characteristic function

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x) \tag{32}$$

which in turn uniquely determines $F(x)$ [C, p. 93].

THEOREM 6. *Let* $F(x)$, $G(x)$, $H(x)$ *be distribution functions in* R_1 *and let* $\varphi_1(t)$, $\varphi_2(t)$, $\varphi(t)$ *be the corresponding characteristic functions. Then*

$$H(x) = F(x) * G(x) \tag{33}$$

if and only if

$$\varphi(t) = \varphi_1(t) \cdot \varphi_2(t). \tag{34}$$

PROOF. Assume (33) holds. Since $|e^{itx}| \leq 1$ we have from Theorem 5,

$$
\begin{aligned}
\varphi(t) &= \int_{-\infty}^{\infty} e^{itx}\, d_x\left\{\int_{-\infty}^{\infty} F(x-y)\, dG(y)\right\} \\
&= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} e^{itx}\, d_x F(x-y)\right\} dG(y) \\
&= \int_{-\infty}^{\infty} e^{ity}\left\{\int_{-\infty}^{\infty} e^{it(x-y)}\, d_x F(x-y)\right\} dG(y) \\
&= \int_{-\infty}^{\infty} e^{ity}\left\{\int_{-\infty}^{\infty} e^{itw}\, dF(w)\right\} dG(y) = \varphi_1(t)\cdot\varphi_2(t).
\end{aligned} \tag{35}
$$

The converse implication now follows from the fact that the characteristic function of a distribution determines the latter uniquely.

The importance of the operation $*$ in probability theory arises from the fact that if X, Y are *independent* random variables with respective distribution functions $F(x)$, $G(x)$, and if $Z = X + Y$, then the distribution function $H(x)$ of Z satisfies (33), since for any value of a,

$$
\begin{aligned}
H(a) &= P[X + Y \leq a] = \iint\limits_{x+y\leqq a} dF(x)\, dG(y) \\
&= \int_{-\infty}^{\infty}\left\{\int_{x\leqq a-y} dF(x)\right\} dG(y) = \int_{-\infty}^{\infty} F(a-y)\, dG(y) = F(a) * G(a),
\end{aligned} \tag{36}
$$

the evaluation of the double integral by an iterated integral following from Fubini's theorem [S, pp. 76–88]. However, (33) may hold without X, Y being independent, and Theorem 6 shows that (34) will then hold also, and conversely.

An example where $H(x) = F(x) * G(x)$ without X, Y being independent has been given by Cramér [C, p. 317, exercise 2]. We shall give another. Let points $0, A, \cdots, F$ in the (x, y)-plane be defined as follows:

$$
O = (0, 0),\ A = (1, 1),\ B = (1/2, 1),\ C = (0, 1/2),\ D = (1, 0),
$$
$$
E = (1, 1/2),\ F = (1/2, 0).
$$

Let $f(x, y)$ have the value 2 inside the quadrilateral $OABC$ and the triangle DEF, and 0 elsewhere. Then if $f(x, y)$ is the joint frequency function of X, Y it is easily seen that X and Y have uniform distributions on the intervals $0 \leq x \leq 1$, $0 \leq y \leq 1$ respectively and that $Z = X + Y$ has the triangular distribution given by (33), although X and Y are not independent.

It would be interesting to know what distribution functions $F(x)$ are such that if $X, Y, Z = X + Y$ are random variables with the distribution functions $F(x)$, $F(x)$, $F(x) * F(x)$ respectively, then X and Y are necessarily independent. A rather trivial example of such a distribution function is the step function $F(x)$ with jumps of $\frac{1}{2}$ at the points $x = 0$ and $x = 1$. It can be shown (oral communication by W. Hoeffding), in generalization of Cramér's example, that no absolutely continuous distribution function (e.g. the normal distribution function) has this property.

5. The problem of random sampling from a mixed population. Let $G(v)$ be a distribution function in the real variable v, and let $F(u, v)$ be for a.e. (relative to the measure corresponding to G) v a distribution function in the real variable u, and for every u a Borel measurable function of v. Let

$$H(u) = \int_{-\infty}^{\infty} F(u, v)\, dG(v); \tag{37}$$

then by Theorem 4 $H(u)$ is a distribution function in R_1. Now define for $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n)$

$$\begin{aligned} \bar{H}(x) &= H(x_1) \cdots H(x_n), \\ \bar{G}(y) &= G(y_1) \cdots G(y_n). \end{aligned} \tag{38}$$

Both $\bar{H}(x)$ and $\bar{G}(y)$ are then distribution functions in R_n. In particular, $\bar{H}(x)$ is the distribution function of a random sample of n independent variates each with the distribution function (37). Set

$$\bar{F}(x, y) = F(x_1, y_1) \cdots F(x_n, y_n); \tag{39}$$

then for a. e. (relative to the measure corresponding to $\bar{G}$) y, $\bar{F}(x, y)$ is a distribution function in x, and for every x, $\bar{F}(x, y)$ is a Borel measurable function of y. By Fubini's theorem we have

$$\begin{aligned} \bar{H}(x) &= \int_{-\infty}^{\infty} F(x_1, y_1)\, dG(y_1) \cdot \cdots \cdot \int_{-\infty}^{\infty} F(x_n, y_n)\, dG(y_n) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F(x_1, y_1) \cdots F(x_n, y_n)\, dG(y_1) \cdots dG(y_n) \\ &= \int_{-\infty}^{\infty} \bar{F}(x, y)\, d\bar{G}(y). \end{aligned} \tag{40}$$

Thus $\bar{H}(x)$ is itself a mixture in the sense of Theorem 4. It follows from Theorem 5 that for any Borel measurable function $f(x)$,

$$\int_{-\infty}^{\infty} f(x)\, d\bar{H}(x) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(x)\, d_x \bar{F}(x, y) \right\} d\bar{G}(y), \tag{41}$$

if and only if the left side of (41) exists. When written out in full (41) becomes

$$\begin{aligned} &\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n)\, d_{x_1} \left\{ \int_{-\infty}^{\infty} F(x_1, y_1)\, dG(y_1) \right\} \\ &\quad \cdots d_{x_n} \left\{ \int_{-\infty}^{\infty} F(x_n, y_n)\, dG(y_n) \right\} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \right. \\ &\quad \left. \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n)\, d_{x_1} F(x_1, y_1) \cdots d_{x_n} F(x_n, y_n) \right\} dG(y_1) \cdots dG(y_n). \end{aligned} \tag{42}$$

Equation (41) is of particular interest in connection with the distribution of a statistic $t = t(x_1, \cdots, x_n) = t(x)$. For any distribution function $J(x)$ let $K(t \mid J)$ denote the distribution function of t when x has the distribution function $J(x)$. If we set

$$f(x) = \begin{cases} 1 & \text{if } t(x) \leq t, \\ 0 & \text{otherwise,} \end{cases} \tag{43}$$

then

$$K(t \mid J) = \int_{-\infty}^{\infty} f(x)\, dJ(x). \tag{44}$$

Hence from (41),

$$\begin{aligned} K(t \mid H(x_1) \cdots H(x_n)) &= K(t \mid \bar{H}) = \int_{-\infty}^{\infty} K(t \mid \bar{F}(x, y))\, d\bar{G}(y) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} K(t \mid F(x_1, y_1) \cdots F(x_n, y_n))\, dG(y_1) \cdots dG(y_n). \end{aligned} \tag{45}$$

As an example, let $t(x)$ be Student's ratio

$$t = n^{\frac{1}{2}} \cdot \bar{x}/s, \tag{46}$$

let

$$F(u, v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}(y-v)^2}\, dy, \tag{47}$$

and let

$$G(v) = \begin{cases} 0 \text{ for } & v < -a, \\ \frac{1}{2} \text{ for } & -a \leq v < a, \\ 1 \text{ for } & a \leq v. \end{cases} \tag{48}$$

Then $H(u)$ will be the distribution function of a mixture in equal proportions of two normal populations with unit variances and with means $-a$, a respectively, and $K(t \mid H(x_1) \cdots H(x_n))$ will be the distribution function of t in random samples of n from this non-normal population. On the other hand, $K(t \mid F(x_1, y_1) \cdots F(x_n, y_n))$ will be the distribution function of t in sampling from successive normal populations with unit variances and means $y_1, \cdots, y_n$ respectively. Relation (45) now becomes

$$K(t \mid H(x_1) \cdots H(x_n)) = \sum_{y_1, \cdots, y_n} K(t \mid F(x_1, y_1) \cdots F(x_n, y_n))/2^n, \tag{49}$$

where the summation is over all 2^n sets $(y_1, \cdots, y_n)$, each y_i being either $-a$ or a. Due to the complexity of $K(t \mid F(x_1, y_1) \cdots F(x_n, y_n))$ (the frequency function of which is discussed in a forthcoming paper by the author), relation

(49) is not very useful. In other cases (45) may afford a considerable simplification in the evaluation of the distribution function of a statistic obtained in random sampling from a mixed population.

REFERENCES

[1] W. FELLER, "On the integro-differential equations of purely discontinuous Markoff processes," *Am. Math. Soc. Trans.*, Vol. 48 (1940), p. 488.

[2] R. H. CAMERON AND W. T. MARTIN, "An unsymmetric Fubini theorem," *Am. Math. Soc. Bull.*, Vol. 47 (1941), p. 121.

[3] P. R. HALMOS, "The decomposition of measures," *Duke Math. Jour.*, Vol. 8 (1941), p. 386.

[4] W. FELLER, "On a general class of 'contagious' distributions," *Annals of Math. Stat.*, Vol. 14 (1943), p. 389.