**TRANSACTIONS OF SOCIETY OF ACTUARIES
1986 VOL. 38**

THE BOUNDS OF BIVARIATE DISTRIBUTIONS THAT LIMIT THE VALUE OF LAST-SURVIVOR ANNUITIES

JACQUES F. CARRIÈRE AND LAI K. CHAN*

ABSTRACT

Traditionally, actuaries have assumed that the individual lives of joint-life and last-survivor statuses are independent when calculating the probabilities of survival or failure of these statuses. However, we usually find that individuals covered by joint-life insurance or last-survivor annuity contracts are associated in some manner. This implies that the individual lifetimes are not independent.

The dependent relationship between two lives is readily analyzed within the context of probability theory. Before proceeding with the analysis, we will present some useful relationships between certain actuarial and probability functions.

This paper will then explore some measures of dependence and examine a general class of bivariate distributions that possesses some desirable properties with respect to these measures. We then present a special case of this class of distributions that simplifies the calculation of probabilities and annuities. Next, we show how large an increase or reduction is possible in the value of certain annuities when the individual lives are not independent.

PROBABILITY AND ACTUARIAL FUNCTIONS

Let T be the time until death of a life aged 0 with the distribution function

$$F(t) = P(T \leq t).$$

Most of the standard actuarial functions may be derived from $F(t)$. For example, the life function is equal to $\ell_t = \ell_0 (1 - F(t))$ where ℓ_0 is the radix of the table. Also ${}_tq_x = \dfrac{F(x+t) - F(x)}{1 - F(x)}$. Let us denote ${}_tq_x$ by $F_x(t)$.

Figure 1 illustrates $F(t)$ for a female life using three mortality bases. The mortality rates are from the Ga-1951 table, the 1971 GAM, and the 1983 Table a. These mortality tables will be used later for the calculation of annuity values.

* Dr. Chan, not a member of the Society, is Professor and Head of the Department of Statistics at the University of Manitoba in Winnipeg, Manitoba.

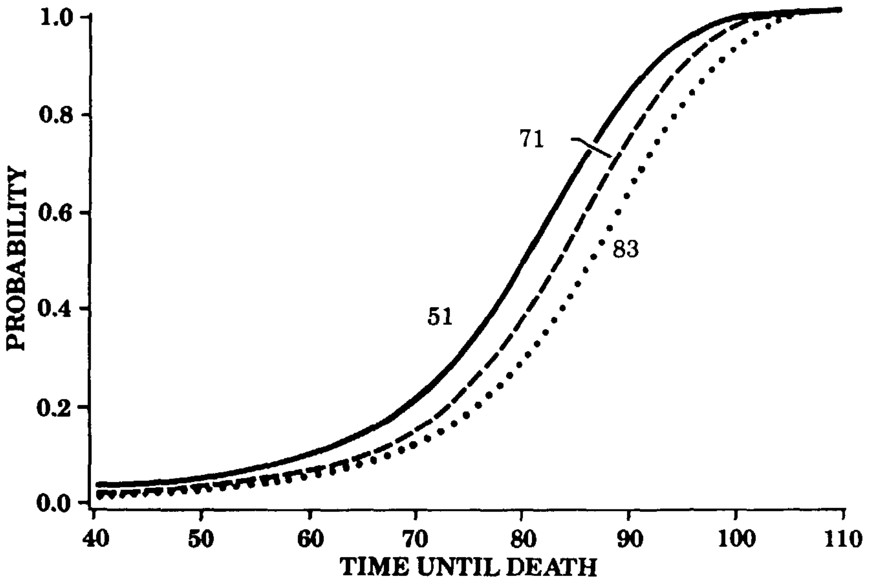

FIG. 1—The distribution function of the time until death of a female aged 0 using various mortality tables.

Let us examine the relationships between bivariate distributions and the joint- and last-survivor probabilities ${}_tq_{xy}$ and ${}_tq_{\overline{xy}}$. Let T_x be a random variable of the time until death of a life aged x, and let T_y be the random variable for the second life. The bivariate distribution function of T_x and T_y is defined as:

$$F_{xy}(t_x,t_y) = P(T_x \leq t_x, T_y \leq t_y).$$

In this case, the marginal distribution function of T_x is $F_x(t_x) = F_{xy}(t_x,\infty)$ which is ${}_tq_x$ in actuarial notation. Now ${}_tq_{\overline{xy}}$ is the distribution of the random variable $\mathrm{Max}(T_x,T_y)$ and ${}_tq_{xy}$ is the distribution of the random variable $\mathrm{Min}(T_x,T_y)$. These probabilities will be used later for the calculation of annuities. It can be shown that:

$${}_tq_{\overline{xy}} = F_{xy}(t,t)$$

$${}_tq_{xy} = F_x(t) + F_y(t) - F_{xy}(t,t).$$

Under the assumption of independence, the bivariate distribution is equal to $F_{xy}(t_x,t_y) = F_x(t_x)F_y(t_y)$. Consequently, the joint- and last-survivor probabilities are simplified into the form:

$${}_tq_{\overline{xy}} = {}_tq_x\,{}_tq_y \text{ and } {}_tq_{xy} = 1 - (1 - {}_tq_x)(1 - {}_tq_y).$$

What does a bivariate distribution look like? We could plot the distribution function, but illustrating its density is more instructive. If the density exists then it is equal to:

$$f(t_x,t_y) = \frac{\partial^2}{\partial t_x \, \partial t_y} F_{xy}(t_x,t_y).$$

Figures 2A, 2B, and 2C illustrate the densities of bivariate distributions whose marginals follow the Gompertz law of mortality when T_x and T_y are independent, negatively related and positively related. In these figures we set $x = y = 20$. The densities were derived by the translation method which will be outlined later.

A more detailed exposition of some of these ideas may be found in Bowers et al. [1].

MEASURES OF ASSOCIATION

There are a variety of ways in which the dependence between two random variables can be summarized by a single measure. These measures are usually defined to lie between -1 and $+1$ and to be equal to zero under independence.

Let us suppose that the death of a life, aged x, has no effect on the death or survival of another life, aged y. Then we will usually say that there is no association between the two lives. Stochastically, this means that the corresponding random variables T_x and T_y are independent, that

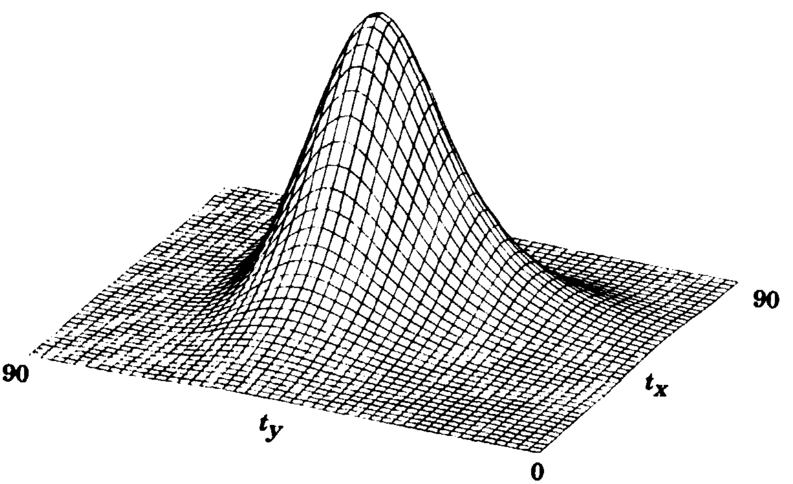

FIG. 2A—Density of an independent bivariate distribution that follows the Gompertz law of mortality.

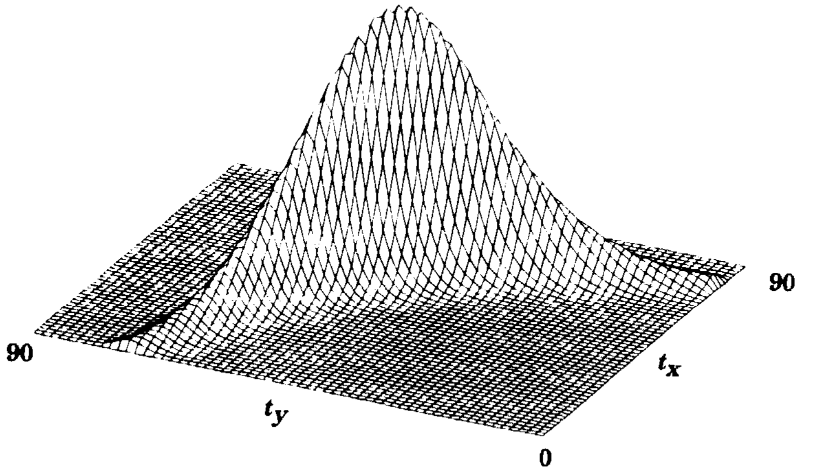

FIG. 2B—Density of a distribution with negative correlation that follows the Gompertz law of mortality.

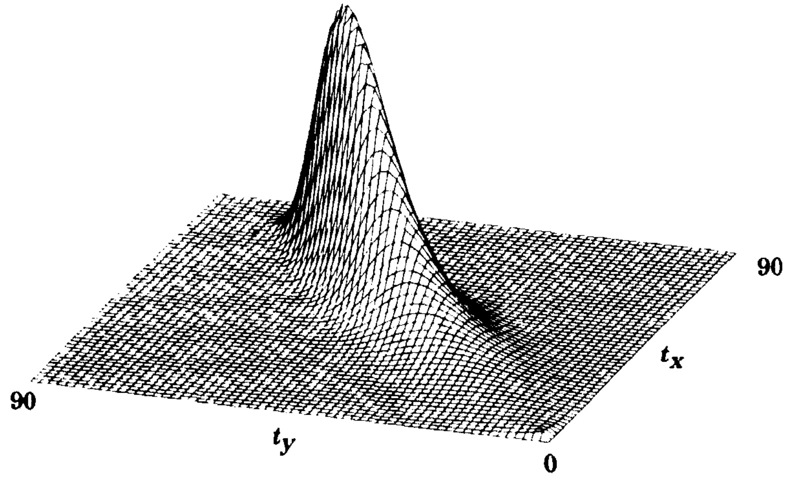

FIG. 2C—Density of a distribution with positive correlation that follows the Gompertz law of mortality.

is $F_{xy}(t_x,t_y) = F_x(t_x)F_y(t_y)$ for all values of t_x and t_y. In this case, the usual measure of association will be zero.

We will say that a pair of lives are in agreement if a long life for one is associated with a long life for the other. In this case, the measure will be positive. We will say that the two lives are in perfect agreement if there exists a strictly increasing function $I(\cdot)$ such that $T_y = I(T_x)$. For example, if $T_y = T_x$, then the lives will die at exactly the same time.

We will say that the lives are in disagreement if a short life for one is associated with a long life for the other. In this case, the measure will be negative. The lives will be in perfect disagreement if there exists a strictly decreasing function $D(\cdot)$ such that $T_y = D(T_x)$.

The usual parametric measure of association is the linear correlation coefficient. Let E be the expectation operator and let μ_x and σ_x^2 be the mean and variance of the random variable T_x. The linear correlation coefficient is defined as:

$$\rho_p = E\left[\left(\frac{T_x - \mu_x}{\sigma_x}\right)\left(\frac{T_y - \mu_y}{\sigma_y}\right)\right].$$

Another measure is the grade correlation coefficient. This measure is realized by first transforming T_x and T_y and then applying the previous formula. Let $M = F_x(T_x)$ and $W = F_y(T_y)$ and let μ_M and σ_M^2 be the mean and variance of the random variable M. The grade correlation coefficient is defined as:

$$\rho_s = E\left[\left(\frac{M - \mu_M}{\sigma_M}\right)\left(\frac{W - \mu_W}{\sigma_W}\right)\right].$$

It is well-known that $|\rho_p| = 1$ if and only if the functional relation between the variables is linear. This measure is deficient in the sense that if the variables are monotonically related and nonlinear, then $|\rho_p| < 1$. But ρ_s is better behaved in the sense that if there exists an increasing function $I(\cdot)$ such that $T_y = I(T_x)$, then $\rho_s = 1$. Similarly, $\rho_s = -1$ if and only if there exists a decreasing function $D(\cdot)$ such that $T_y = D(T_x)$. (See Theorems 1 to 3 in Appendix I for details.) We will later see that the grade correlation coefficient is an ideal measure because if $|\rho_s| = 1$, then the bivariate distribution $F_{xy}(t_x,t_y)$ attains certain bounds.

Suppose that we have a sample of n independent pairs of observations from the distribution $F_{xy}(t_x,t_y)$. Then the estimates for ρ_p and ρ_s are the Pearson product moment correlation coefficient and Spearman's rank correlation coefficient. (The formulas for these estimators are given in Appendix II.)

Additional information about these measures of association may be found in Kruskal [8].

A GENERAL CLASS OF BIVARIATE DISTRIBUTIONS

In practice, the only probabilities available for the calculation of joint-life insurance and last-survivor annuity values are ${}_tq_x$ and ${}_tq_y$. Therefore, it would be desirable to have a general bivariate model that is a function of the marginal distribution functions $F_x(t_x)$ and $F_y(t_y)$. Initially, we want a model that is not restricted to special laws of mortality, such as Gompertz's Law.

Note that the marginal distributions $F_x(t_x)$ and $F_y(t_y)$ can be generated from many different bivariate distributions. Consequently, we first shall introduce a general class of one-parameter bivariate distributions with these marginals. Each distribution in the class can be represented as:

$$F_{xy}(t_x,t_y) = H[F_x(t_x), F_y(t_y);\rho], \quad -1 \leq \rho \leq 1$$

subject to the conditions that:

a. $H[F_x(t_x), F_y(t_y); +1] = \text{Min}[F_x(t_x), F_y(t_y)] = U(t_x,t_y)$

b. $H[F_x(t_x), F_y(t_y);0] = F_x(t_x)\, F_y(t_y) = I(t_x,t_y)$

c. $H[F_x(t_x), F_y(t_y); -1] = \text{Max}[0, F_x(t_x)+F_y(t_y)-1] = L(t_x,t_y)$.

The parameter ρ gives some measure of association between T_x and T_y. Let us now discuss the derivation of these conditions and justify their use for our model. First, if $\rho = 0$, then we want the distribution to be independent, that is, $F_{xy}(t_x,t_y) = F_x(t_x)F_y(t_y)$. Second, if $\rho = +1$ then we want $T_y = I(T_x)$, which is a necessary and sufficient condition for the bivariate distribution to equal $\text{Min}[F_x(t_x),F_y(t_y)]$. Third, if $\rho = -1$ then we want $T_y = D(T_x)$, which is necessary and sufficient for the bivariate distribution to equal $\text{Max}[0,F_x(t_x)+F_y(t_y)-1]$. (For further details, see Theorems 4 and 5 in Appendix I.)

$L(t_x,t_y)$ and $U(t_x,t_y)$ are called boundary distributions because any bivariate distribution lies within these bounds. It is easy to prove that:

$$L(t_x,t_y) \leq F_{xy}(t_x,t_y) \leq U(t_x,t_y).$$

These boundary distributions were first introduced by Fréchet [2] and are directly related to the grade correlation coefficient. Specifically, if $|\rho_s| = 1$, then $F_{xy}(t_x,t_y)$ assumes Fréchet's bounds. Also, note that the density of a bivariate distribution does not exist at these bounds.

There are an uncountable number of one-parameter distributions which will satisfy the conditions of the previous model. Many of these were outlined by Johnson and Tenenbein [4], Kimeldorf and Sampson [6], [7], and Mardia [9].

MIXTURES OF DISTRIBUTIONS

One easy way of obtaining probabilities for all degrees of association is to construct a distribution that is a mixture of $L(t_x,t_y)$, $I(t_x,t_y)$, and $U(t_x,t_y)$. The mixture arises from the assumption that the population consists of three groups. We will assume that the proportion of paired lives in perfect disagreement within the population is P_1. Similarly, the proportion in perfect agreement is P_3, and the proportion that is independent is P_2. The model is defined as:

$$H[F_x(t_x),F_y(t_y);\rho] = P_1(\rho)L(t_x,t_y) + P_2(\rho)I(t_x,t_y) + P_3(\rho)U(t_x,t_y).$$

The proportions $P_1(\rho)$, $P_2(\rho)$, and $P_3(\rho)$ are weight functions that satisfy the following conditions:

a. $P_1(\rho) \geq 0$, $P_2(\rho) \geq 0$, and $P_3(\rho) \geq 0$
b. $P_1(\rho) + P_2(\rho) + P_3(\rho) = 1$
c. $P_1(-1) = P_2(0) = P_3(+1) = 1$.

The density for this model does not exist, but the strength of this distribution lies in its simplicity. Later we will show how easy it is to calculate annuities for all levels of association with mixtures of distributions. Another important feature of this model is that the grade correlation coefficient is equal to $\rho_s = P_3(\rho) - P_1(\rho)$. (See Theorem 6 in Appendix I.)
A general class of weight functions can have the form:

$$P_1(\rho) = \frac{|\rho|^k(1-\rho^{\frac{2c+1}{2d+1}})}{2}$$

$$P_2(\rho) = 1 - |\rho|^k$$

$$P_3(\rho) = \frac{|\rho|^k(1+\rho^{\frac{2c+1}{2d+1}})}{2}$$

where $k \geq 0$, $c = 0,1,2,\ldots,d = 0,1,2,\ldots$.
The grade correlation for this model is:

$$\rho_s = |\rho|^k\, \rho^{\frac{2c+1}{2d+1}}.$$

As a special case of this model, Mardia [9] used the weight functions:

$$P_1(\rho) = \rho^2\,\frac{(1-\rho)}{2},\ P_2(\rho) = 1 - \rho^2,\ P_3(\rho) = \rho^2\,\frac{(1+\rho)}{2}.$$

The grade correlation for this model is $\rho_s = \rho^3$.
Another model has the weight functions:

$$P_1(\rho) = |\rho|^{2/15}\,\frac{(1-\rho^{1/5})}{2},\ P_2(\rho) = 1 - |\rho|^{2/15},\ P_3(\rho) = |\rho|^{2/15}\,\frac{(1+\rho^{1/5})}{2}.$$

The grade correlation for this model is $\rho_s = \rho^{1/3}$.

Estimating a mixture distribution is a matter of estimating the marginal distributions $F_x(t_x)$ and $F_y(t_y)$ and the proportions P_1, P_2, and P_3. If the true form of the distribution is not a mixture, it may still be possible to approximate the true distribution with a mixture.

AN ANALYSIS OF THE EFFECT ON ANNUITIES

In this section, we examine how the value of certain annuities are affected when the probabilities assume Fréchet's bounds. We will consider some standard annuities found in Jordan [5]:

a_{xy} = immediate annuity payable yearly until the first death

$$= \sum_{t=1}^{\infty} v^t(1 - {}_tq_{xy}) = \sum_{t=1}^{\infty} v^t(1 - {}_tq_x - {}_tq_y + {}_tq_{\overline{xy}});$$

$a_{\overline{xy}}$ = immediate annuity payable yearly until the second death

$$= \sum_{t=1}^{\infty} v^t(1 - {}_tq_{\overline{xy}}); \text{ and}$$

a_{jr} = immediate annuity payable yearly and reducing to 2/3 of the principal amount on the death of the principal life y

$$= \sum_{t=1}^{\infty} v^t[1 - ({}_tq_y + 2{}_tq_{\overline{xy}})/3].$$

The female life is aged x and the male life is aged y and $v^t = (1+i)^{-t}$, where i = effective annual rate of interest. If we have a mixture of distributions, then

$$\begin{aligned} {}_tq_{\overline{xy}} &= F_{xy}(t,t) = H({}_tq_x, {}_tq_y; \rho) \\ &= P_1(\rho)L(t,t) + P_2(\rho)I(t,t) + P_3(\rho)U(t,t). \end{aligned}$$

The last-survivor annuity assumes the following form when the distribution is a mixture:

$$\begin{aligned} a_{\overline{xy}} = P_1(\rho) &\sum_{t=1}^{\infty} v^t(1 - L(t,t)) \\ + P_2(\rho) &\sum_{t=1}^{\infty} v^t(1 - I(t,t)) \\ + P_3(\rho) &\sum_{t=1}^{\infty} v^t(1 - U(t,t)). \end{aligned}$$

That is, $a_{\overline{xy}}$ is the weighted average of the last-survivor annuities corresponding to the distributions $L(t,t)$, $I(t,t)$ and $U(t,t)$ with weights P_1, P_2, and P_3. Also, a_{xy} and a_{jr} can be similarly obtained for any mixture.

Figure 3 illustrates the value of a last-survivor annuity for two lives aged 60 at 5 percent interest using the 1971 GAM table. Plot I shows the value of $a_{\overline{xy}}$ for all values of the parameter ρ, when we use Mardia's weight functions. The weight functions for plot II are those of the general class when $k = 2/15$, $c = 0$, and $d = 2$. The grade correlation for the distributions underlying plots I and II are $\rho_s = \rho^3$ and $\rho_s = \rho^{1/3}$, respectively.

Tables 1A and 1B give the values of a_{xy}, $a_{\overline{xy}}$, and a_{jr} at four interest rates (0%, 5%, 10%, 15%) and three mortality bases (Ga-1951, 1971 GAM, 1983 Table $\mathcal{a}$) for pairs of lives aged 40, 60, and 80. Annuity values at the upper and lower bounds are tabulated, as well as the independent case. On examining the tables, one notes that changing the mortality basis is about as significant as changing the correlation from zero to one or negative one. Also, note that discounting for interest lessens the effect of changing the parameter ρ and that increasing the ages of the pair of lives increases the effect of changing the correlation.

These tables can also be used to calculate annuity values for any correlation ρ_s, if we assume that ${}_tq_{\overline{xy}}$ is a mixture of distributions. For example,

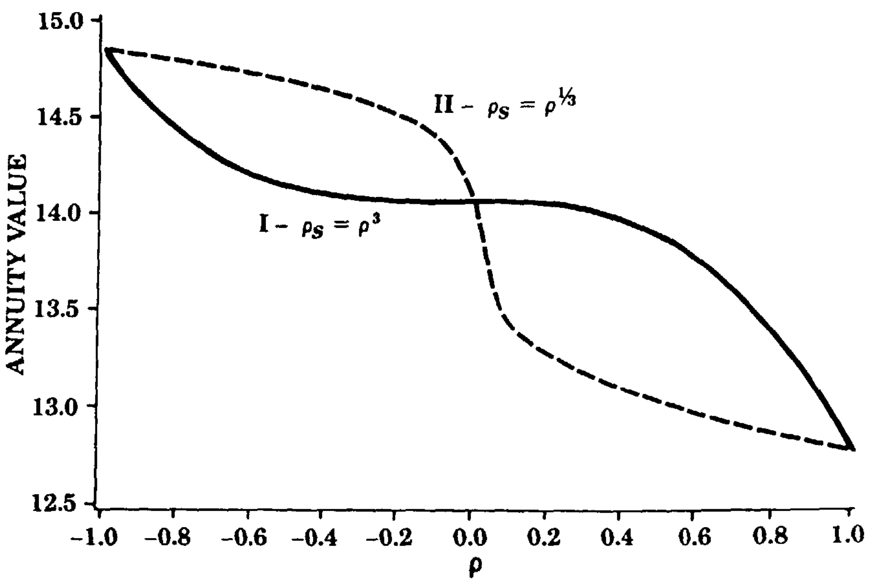

FIG. 3—Value of a last-survivor annuity for mixtures of distributions based on the 1971 GAM table.

TABLE 1A

ANNUITY VALUES AT THE BOUNDS

TABLE	AGE	ρ	i = 0%			i = 5%		
			a_{xy}	$a_{\overline{xy}}$	a_{jr}	a_{xy}	$a_{\overline{xy}}$	a_{jr}
Ga-1951	40	−1	27.124	45.723	41.827	14.083	17.779	17.000
		0	29.657	43.189	40.138	14.518	17.344	16.710
		+1	34.035	38.812	37.220	15.443	16.418	16.093
	60	−1	11.449	26.426	23.359	8.150	14.324	13.061
		0	13.614	24.261	21.916	9.016	13.458	12.484
		+1	17.225	20.651	19.509	10.535	11.939	11.471
	80	−1	2.455	10.599	9.024	2.195	7.873	6.783
		0	3.687	9.367	8.203	3.086	6.981	6.189
		+1	5.875	7.180	6.745	4.604	5.463	5.177
1971 GAM	40	−1	29.191	47.920	43.785	14.624	18.005	17.255
		0	31.654	45.458	42.143	15.004	17.625	17.001
		+1	35.515	41.597	39.569	15.754	16.875	16.501
	60	−1	12.812	28.450	25.053	8.851	14.855	13.550
		0	14.995	26.267	23.598	9.655	14.051	13.015
		+1	18.259	23.003	21.422	10.942	12.765	12.157
	80	−1	2.983	11.898	10.100	2.626	8.604	7.407
		0	4.320	10.560	9.208	3.552	7.678	6.789
		+1	6.504	8.377	7.752	5.012	6.217	5.816
1983 Table *a*	40	−1	32.658	51.610	47.657	15.350	18.339	17.707
		0	35.213	49.055	45.954	15.683	18.006	17.486
		+1	39.751	44.517	42.929	16.444	17.245	16.978
	60	−1	15.915	32.022	28.720	10.280	15.688	14.577
		0	18.206	29.731	27.192	10.998	14.971	14.099
		+1	22.115	25.822	24.586	12.355	13.613	13.194
	80	−1	4.272	14.300	12.354	3.631	9.845	8.642
		0	5.788	12.784	11.343	4.591	8.885	8.002
		+1	8.462	10.110	9.561	6.237	7.239	6.905

the value of a_{xy}, using the 1983 Table *a* when $i = .10$, $x = y = 60$, $\rho = .9$ (i.e., $\rho_s = .729$), with Mardia's weight functions is equal to:

$$.9^2 \frac{(1-.9)}{2} \times 7.273 + (1-.9^2) \times 7.526$$
$$+ .9^2 \frac{(1+.9)}{2} \times 8.110 = 7.965.$$

Now, let us consider the following problem. What is the change in the interest assumption that would yield an annuity value equal to the change in the correlation from zero to one or negative one? Table 2 shows the change required from a 5 and 15 percent interest assumption so that the value of the annuity when $\rho = 0$ is equal to the value when $\rho = \pm 1$ for various annuities and mortality bases, at ages 50 through 90. Note that at

TABLE 1B

ANNUITY VALUES AT THE BOUNDS

TABLE	AGE	ρ	$i = 10\%$			$i = 15\%$		
			a_{xy}	$a_{\overline{xy}}$	a_{jr}	a_{xy}	$a_{\overline{xy}}$	a_{jr}
Ga-1951	40	−1	8.829	9.856	9.638	6.263	6.652	6.569
		0	8.920	9.766	9.577	6.286	6.630	6.554
		+1	9.201	9.484	9.390	6.403	6.513	6.476
	60	−1	6.150	9.103	8.496	4.858	6.461	6.129
		0	6.530	8.723	8.242	5.040	6.278	6.007
		+1	7.282	7.971	7.741	5.465	5.853	5.724
	80	−1	1.979	6.130	5.338	1.798	4.951	4.352
		0	2.643	5.466	4.896	2.305	4.444	4.014
		+1	3.754	4.355	4.155	3.154	3.595	3.448
1971 GAM	40	−1	9.011	9.883	9.688	6.340	6.656	6.585
		0	9.083	9.812	9.640	6.357	6.640	6.574
		+1	9.297	9.598	9.497	6.443	6.554	6.517
	60	−1	6.543	9.259	8.665	5.096	6.511	6.198
		0	6.871	8.931	8.446	5.244	6.363	6.100
		+1	7.477	8.325	8.042	5.574	6.033	5.880
	80	−1	2.336	6.566	5.723	2.098	5.223	4.602
		0	3.000	5.902	5.280	2.588	4.733	4.276
		+1	4.037	4.865	4.589	3.361	3.960	3.760
1983 Table *a*	40	−1	9.208	9.919	9.765	6.410	6.661	6.606
		0	9.262	9.865	9.729	6.421	6.649	6.598
		+1	9.459	9.668	9.598	6.496	6.575	6.548
	60	−1	7.273	9.477	9.021	5.507	6.573	6.351
		0	7.526	9.223	8.852	5.608	6.473	6.284
		+1	8.110	8.640	8.463	5.907	6.173	6.084
	80	−1	3.137	7.250	6.454	2.747	5.621	5.064
		0	3.771	6.615	6.031	3.183	5.185	4.774
		+1	4.863	5.524	5.304	3.951	4.416	4.261

the younger ages, changing the correlation from zero to one or negative one is insignificant relative to the interest assumption. But as the lives grow older, the correlation assumption increases in importance relative to the interest assumption. And, eventually the correlation assumption dominates.

Another way to grasp the effect that the bounds have on annuity values is to calculate the relative change of the annuity at the bound compared to the independent case. That is,

$$\text{Relative Change} = 100\left(\frac{\text{Boundary Value}}{\text{Independent Value}} - 1\right).$$

Figures 4A and 4B illustrate the relative change of last-survivor annuities, $a_{\overline{xy}}$, at four interest rates, using the 1971 GAM table, for all ages between 20 and 80. Note, that as the age of the pair of lives increases, the relative

TABLE 2

THE CHANGE REQUIRED IN THE INTEREST ASSUMPTION FOR MAINTAINING THE ANNUITY VALUE AT $\rho = 0$ WHEN ρ IS CHANGED TO -1 OR 1

TABLE	AGE	ρ	$i = 5\%$			$i = 15\%$		
			a_{xy}	$a_{\overline{xy}}$	a_{jr}	a_{xy}	$a_{\overline{xy}}$	a_{jr}
Ga-1951	50	−1	0.59	−0.32	−0.24	0.22	−0.17	−0.12
		+1	−0.98	0.67	0.48	−0.68	0.57	0.39
	60	−1	1.44	−0.61	−0.46	0.80	−0.49	−0.36
		+1	−2.04	1.24	0.91	−1.67	1.25	0.88
	70	−1	4.04	−1.18	−0.92	3.03	−1.28	−0.97
		+1	−4.65	2.50	1.85	−4.38	2.83	2.04
	80	−1	11.95	−2.16	−1.73	10.81	−2.71	−2.13
		+1	−10.32	5.01	3.73	−10.65	6.04	4.41
	90	−1	34.70	−3.62	−2.91	29.53	−4.57	−3.65
		+1	−20.22	10.12	7.37	−21.54	11.93	8.62
1983 Table *a*	50	−1	0.35	−0.22	−0.16	0.10	−0.08	−0.06
		+1	−0.67	0.49	0.35	−0.43	0.38	0.26
	60	−1	0.80	−0.40	−0.30	0.35	−0.25	−0.18
		+1	−1.29	0.86	0.62	−0.97	0.79	0.55
	70	−1	2.05	−0.78	−0.60	1.29	−0.71	−0.52
		+1	−2.71	1.58	1.16	−2.34	1.66	1.18
	80	−1	6.05	−1.52	−1.20	4.95	−1.76	−1.35
		+1	−6.44	3.36	2.48	−6.26	3.84	2.78
	90	−1	17.95	−2.76	−2.20	17.10	−3.49	−2.74
		+1	−15.04	7.78	5.62	−15.83	9.23	6.60

change in $a_{\overline{xy}}$ increases and that the effect is less significant at higher interest rates.

Figures 5A and 5B illustrate the relative change of $a_{\overline{xy}}$, at four ages, using the 1971 GAM table, for all interest rates between 0 and 20 percent. Note that as interest increases, the relative change in $a_{\overline{xy}}$ decreases and that the effect is more significant at higher ages.

SOME OTHER MODELS

Consider some models developed by other authors. First, Fréchet [2] suggested the following one-parameter bivariate distribution:

$$F_{xy}(t_x,t_y) = (1-\rho)\ \mathrm{Min}\ [F_x(t_x),F_y(t_y)] + \rho\ \mathrm{Max}\ [0,F_x(t_x)+F_y(t_y)-1]$$

for $0 \leq \rho \leq 1$.

This model is simply a linear interpolation formula between his upper and lower bounds. The weakness with this model is that it does not include the case of independence.

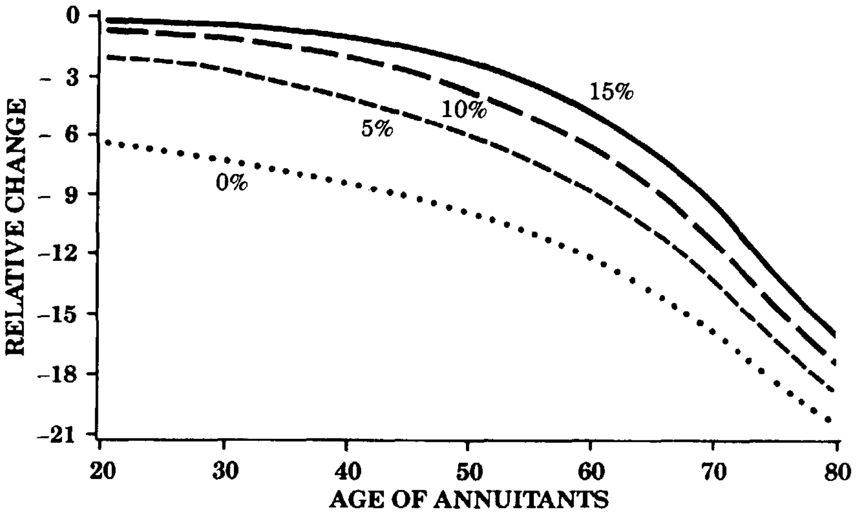

FIG. 4A—Effect of the upper bound on a last-survivor annuity at four interest rates using the 1971 GAM table.

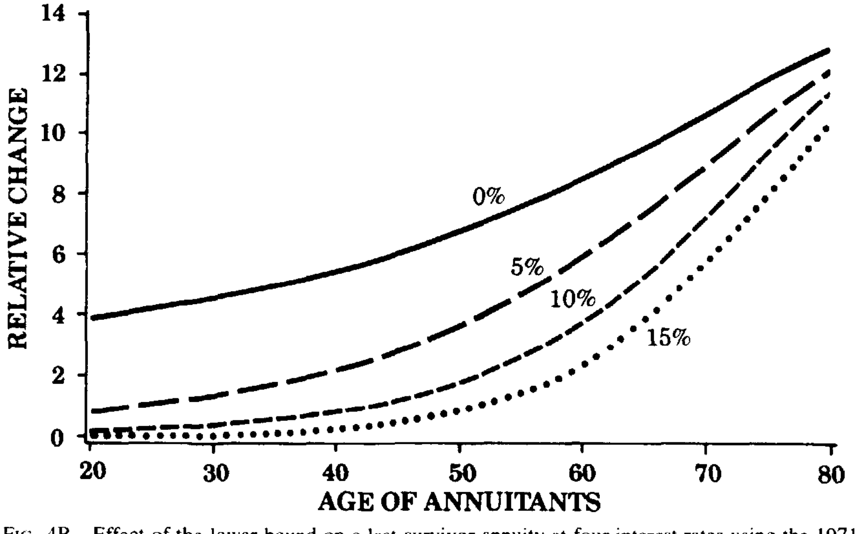

FIG. 4B—Effect of the lower bound on a last-survivor annuity at four interest rates using the 1971 GAM table.

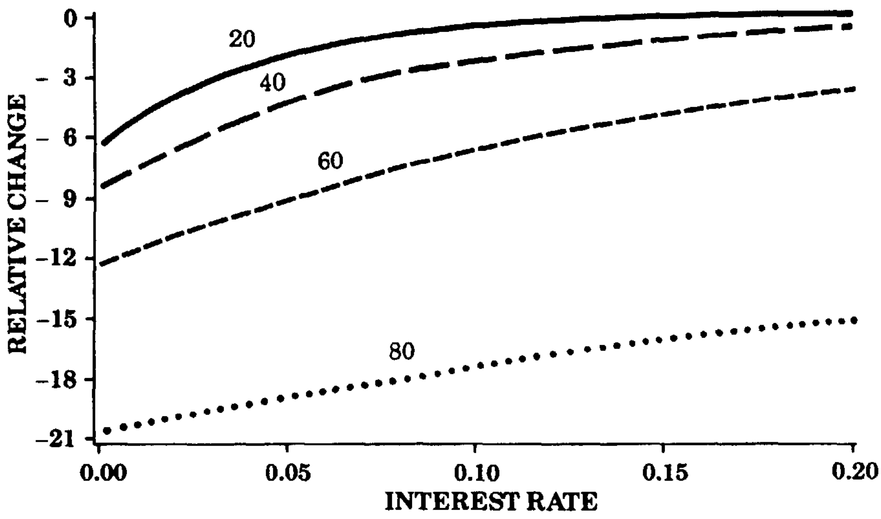

FIG. 5A—Effect of the upper bound on a last-survivor annuity at four ages using the 1971 GAM table.

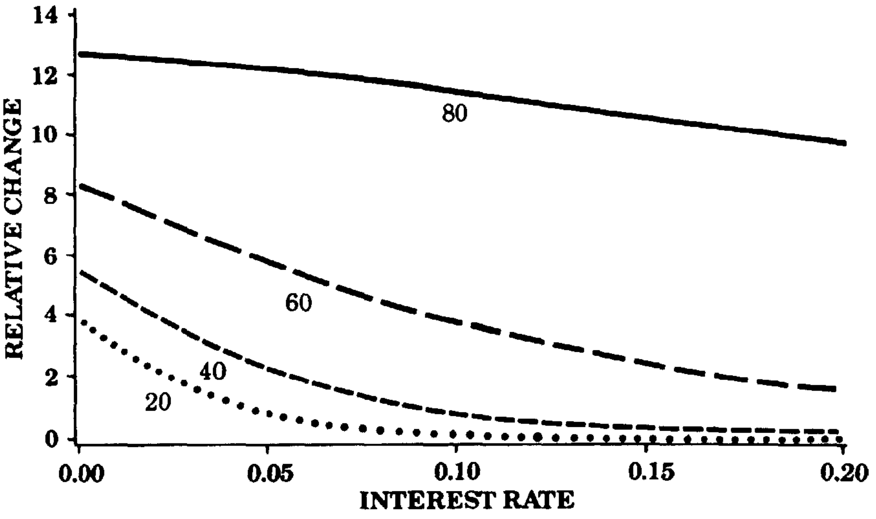

FIG. 5B—Effect of the lower bound on a last-survivor annuity at four ages using the 1971 GAM table.

Second, a one-parameter model that does include independence but not Fréchet's bounds was given by Morgenstern [10]. He suggested the following formula:

$$F_{xy}(t_x, t_y) = F_x(t_x)F_y(t_y)[1 + \rho(1 - F_x(t_x))(1 - F_y(t_y))]$$
$$\text{for } -1 \leq \rho \leq +1.$$

Third, Plackett [11] introduced a one-parameter model which includes Fréchet's bounds and also the case of independence. He pointed out that solving for $F_{xy}(t_x,t_y)$ in the following equation will yield a proper bivariate distribution:

$$\epsilon = \frac{F_{xy}(t_x,t_y)[1-F_x(t_x)-F_y(t_y)-F_{xy}(t_x,t_y)]}{[F_x(t_x)-F_{xy}(t_x,t_y)][F_y(t_y)-F_{xy}(t_x,t_y)]} \quad \text{for } \epsilon > 0.$$

Fourth, Mardia [9] suggested another one-parameter model which satisfies Fréchet's bounds and the case of independence. This model is generated by the translation method and uses the fact that the bivariate normal distribution assumes Fréchet's bounds when $|\rho| \to 1$. Figures 2A, 2B, and 2C were created with this method. To generate the distribution of $F_{xy}(t_x,t_y)$, we apply the transformation $T_x = F_x^{-1}[\phi(U)]$, $T_y = F_y^{-1}[\phi(V)]$ to the standard bivariate distribution $H_{uv}(u,v;\rho)$

$$= \int_{-\infty}^{u}\int_{-\infty}^{v} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}(r^2+s^2-2\rho rs)\right\} dsdr$$

with a marginal distribution $\phi(u) = \int_{-\infty}^{u} \frac{e^{-\frac{r^2}{2}}}{\sqrt{2\pi}}\, dr.$

This yields $F_{xy}(t_x,t_y) = P[T_x \le t_x, T_y \le t_y] =$

$$P[F_x^{-1}[\phi(U)] \le t_x, F_y^{-1}[\phi(V)] \le t_y] = H_{uv}(\phi^{-1}[F_x(t_x)], \phi^{-1}[F_y(t_y)];\rho).$$

The density is then derived by taking the partial derivatives of $F_{xy}(t_x,t_y)$ with respect to t_x and t_y.

Many other models may be found in Johnson and Kotz [3].

Finally, another measure of association that has properties analogous to the grade correlation coefficient is the difference sign correlation coefficient. Let (T_x',T_y') and (T_x,T_y) be independent pairs of identically distributed random variables, then this measure is defined as:

$$\rho_t = 2P[(T_x'-T_x)(T_y'-T_y)>0] - 1.$$

CONCLUSION

The usual calculation of joint-life or last-survivor annuities assumes that the individual lives are independent. This paper shows that dependence could have significant effects. The maximum effect occurs when the bivariate distribution of the random variables T_x and T_y attains Fréchet's bounds, or equivalently, when the grade correlation coefficient attains its bounds of

± 1. A model is presented that will facilitate the computation of annuity values for lives wtih various degrees of association.

But further work may still be done. The problem of identifying the true form of the bivariate distribution and its estimation is one project. Another project would be to study how these measures of association change as the pair of lives grow older. Still another project would be to analyze the effect the bounds may have on pension plan liabilities. Also, a natural extension would be the case of multilife statuses and their distributions.

REFERENCES

1. BOWERS N., GERBER, H., HICKMAN, J., JONES, D., AND NESBITT, C. *Actuarial Mathematics*, (40-11-84), Itasca, Ill.: Society of Actuaries, 1984.
2. FRÉCHET, M. "Sur les tableaux de corrélation dont les marges sont données," *Annales de l'Université de Lyon,* Section A, Serie 3, 14, 53–77, 1951.
3. JOHNSON, N., AND KOTZ, S. *Distributions in Statistics: Continuous Multivariate Distributions*, New York City: John Wiley & Sons, Inc., 1972.
4. JOHNSON, M.E., AND TENENBEIN, A. "Bivariate Distributions with Given Marginals and Fixed Measures of Dependence," Technical Report No. LA-7700-MS, Los Alamos Scientific Laboratory, 1979.
5. JORDAN, C.W. *Life Contingencies*, Chicago: The Society of Actuaries, 1967.
6. KIMELDORF, G., AND SAMPSON, A. "One-Parameter Families of Bivariate Distributions with Fixed Marginals," *Communications in Statistics*, 4(3) (1975), 293–301.
7. ———. "Uniform Representations of Bivariate Distributions," *Communications in Statistics*, 4(7) (1975), 617–27.
8. KRUSKAL, W.H. "Ordinal Measures of Association," *Journal of the American Statistical Association*, Volume 53 (1958), 814–59.
9. MARDIA, K.V. *Families of Bivariate Distributions*, London: Charles Griffin & Company Limited, 1970.
10. MORGENSTERN, D. "Einfache Beispiele zweidimensionaler Verteilungen," *Mitteilingsblatt fur Mathematische Statistik*, 8 (1956), 234–35.
11. PLACKETT, R.L. "A Class of Bivariate Distributions," *Journal of the American Statistical Association*, 60 (1965), 516–22.

APPENDIX I

THEOREMS AND PROOFS*

THEOREM 1. *Let X and Y be two random variables with means μ_x and μ_y and variances σ_x^2 and σ_y^2 and let*

$$\rho_p = E\left[\left(\frac{X-\mu_x}{\sigma_x}\right)\left(\frac{Y-\mu_y}{\sigma_y}\right)\right] \quad \textit{Then}$$

* The marginal probability distributions are assumed to be bijective and in many cases random variables are only equal up to sets of measure zero.

a) $\rho_p = +1$ iff $\dfrac{X-\mu_x}{\sigma_x} = \dfrac{Y-\mu_y}{\sigma_y}$,

b) $\rho_p = -1$ iff $\dfrac{X-\mu_x}{\sigma_x} = \dfrac{\mu_y - Y}{\sigma_y}$,

c) *if X and Y are independent, then* $\rho_p = 0$.

THEOREM 2. *Let X and Y be two random variables with distribution functions $F(x)$ and $G(y)$ and let*

$$\rho_s = E\left[\left(\frac{F(X)-E(F(X))}{\sqrt{V(F(X))}}\right)\left(\frac{G(Y)-E(G(Y))}{\sqrt{V(G(Y))}}\right)\right].$$ *Then*

a) $\rho_s = 12\,E[F(X)G(Y)] - 3$,

b) $\rho_s = +1$ iff $F(X) = G(Y)$,

c) $\rho_s = -1$ iff $F(X) = 1 - G(Y)$,

d) *If X and Y are independent, then* $\rho_s = 0$.

Proof.

a) Since $F(X)$ and $G(Y)$ are uniformly distributed over $(0,1)$, $E(F(X)) = E(G(Y)) = \dfrac{1}{2}$ and $V(F(X)) = V(G(Y)) = \dfrac{1}{12}$, therefore $\rho_s = 12\,E(F(X)G(Y)) - 3$;

b) and c) are direct results of Theorem 1.

THEOREM 3. *Let X and Y be two random variables with distributions $F(x)$ and $G(y)$ and let $I(\cdot)$ and $D(\cdot)$ be strictly increasing and decreasing functions. Then*

a) $Y = I(X)$ iff $F(X) = G(Y)$,

b) $Y = D(X)$ iff $F(X) = 1 - G(Y)$.

Proof.

a) If $F(X) = G(Y)$, then $Y = G^{-1}(F(X))$ which implies that $Y = I(X)$. Now, if $Y = I(X)$, then $F(x) = P(X \le x) = P(I^{-1}(Y) \le x) = P(Y \le I(x)) = G(I(x))$ which implies that $G^{-1}(F(\cdot)) = I(\cdot)$; therefore $F(X) = G(Y)$.

b) If $F(X) = 1 - G(Y)$, then $Y = G^{-1}(1 - F(X))$ which implies that $Y = D(X)$. Now, if $Y = D(X)$, then $F(x) = P(X \le x) = P(D^{-1}(Y) \le x) = P(Y > D(x)) = 1 - G(D(x))$ which implies that $G^{-1}(1 - F(\cdot)) = D(\cdot)$; therefore $F(X) = 1 - G(Y)$.

THEOREM 4. *Let A and B be two events. Then*

$\mathrm{Max}(0, P(A) + P(B) - 1) \le P(A \cap B) \le \mathrm{Min}(P(A), P(B))$.

Proof. $P(A) + P(B) - P(A \cap B) = P(A \cup B) \le 1$. This implies that $P(A) + P(B) - 1 \le P(A \cap B)$. But $P(A \cap B) \ge 0$. Therefore $P(A \cap B) \ge \mathrm{Max}(0, P(A) + P(B) - 1)$. Now, $P(A \cap B) \le P(A)$ and $P(A \cap B) \le P(B)$. Therefore $P(A \cap B) \le \mathrm{Min}(P(A), P(B))$.

THEOREM 5. *Let X and Y be two random variables with a bivariate distribution function $F_{xy}(x,y)$, and marginal distribution functions $F(x)$ and $G(y)$ and let $I(\cdot)$ and $D(\cdot)$ be strictly increasing and decreasing functions. Then*

a) $F_{xy}(x,y) = \mathrm{Min}[F(x),G(y)]$ iff $Y = I(X)$,

b) $F_{xy}(x,y) = \mathrm{Max}[0,F(x)+G(y)-1]$ iff $Y = D(X)$.

Proof.

a) If $Y = I(X)$, then $F_{xy}(x,y) = P[X \le x, Y \le y] = P[X \le x, I(X) \le y] = P[X \le x, X \le I^{-1}(y)] = \mathrm{Min}[P(X \le x), P(X \le I^{-1}(y))] = \mathrm{Min}[P(X \le x), P(I(X) \le y)] = \mathrm{Min}[F(x),G(y)]$. Now, if $F_{xy}(x,y) = \mathrm{Min}[F(x),G(y)]$, then $E[F(X)G(Y)] = E(MW)$ where $M = F(X)$, $W = G(Y)$ and where the bivariate distribution of M and W is $F_{mw}(m,w) = \mathrm{Min}[m,w]$. But it can be shown that $E(MW) = \int_0^1 \int_0^1 F_{mw}(m,w)dmdw$. So $E(MW) = \int_0^1 \int_0^w m\,dm\,dw + \int_0^1 \int_0^m w\,dw\,dm = \frac{1}{3}$. Therefore $\rho_s = 12\,E(MW) - 3 = 1$. And by Theorems 2 and 3 this implies that $Y = I(X)$.

b) If $Y = D(X)$, then $F_{xy}(x,y) = P[X \le x, D(X) \le y] = P[X \le x, X \ge D^{-1}(y)]$. Now if $x \le D^{-1}(y)$, then $F_{xy}(x,y) = 0$ and if $x > D^{-1}(y)$ then $F_{xy}(x,y) = P[X \le x] - P[X < D^{-1}(y)] = P[X \le x] - P[D(X) > y] = F(x) - (1 - G(y))$. Therefore $F_{xy}(x,y) = \mathrm{Max}[0,F(x)+G(y)-1]$. Now if $F_{xy}(x,y) = \mathrm{Max}[0,F(x)+G(y)-1]$, then $F_{mw}(m,w) = \mathrm{Max}[0,m+w-1]$ and $E(MW) = \int_0^1 \int_{1-w}^1 (w+m-1)dmdw = \frac{1}{6}$, therefore $\rho_s = -1$. And by Theorems 2 and 3 this implies that $Y = D(X)$.

THEOREM 6. *Let X and Y be two random variables with marginal distribution functions $F(x)$ and $G(y)$ and with a bivariate distribution $H_{xy}(x,y) = P_1(\rho)L(x,y) + P_2(\rho)I(x,y) + P_3(\rho)U(x,y)$ where $P_1(\rho)$, $P_2(\rho)$,* and *$P_3(\rho)$ are suitable weight functions. Then the grade correlation coefficient is equal to $\rho_s = P_3(\rho) - P_1(\rho)$.*

Proof.

$$E[F(X)G(Y)] = E(MW) = \int_0^1 \int_0^1 H_{mw}(m,w)dmdw$$

$$= \int_0^1 \int_0^1 [P_1(\rho)\mathrm{Max}(0,m+w-1) + P_2(\rho)mw + P_3(\rho)\mathrm{Min}(m,w)]dmdw$$

$$= P_1(\rho) \int_0^1 \int_{1-w}^1 (m+w-1)dmdw + P_2(\rho) \int_0^1 \int_0^1 mwdmdw + P_3(\rho) \int_0^1 \int_0^w 2mdmdw$$

$$= \frac{P_1(\rho)}{6} + \frac{P_2(\rho)}{4} + \frac{P_3(\rho)}{3}.$$ But $P_1(\rho) + P_2(\rho) + P_3(\rho) = 1$.

Therefore $E(MW) = \dfrac{P_3(\rho) - P_1(\rho) + 3}{12}$ and $\rho_s = P_3(\rho) - P_1(\rho)$.

APPENDIX II

STANDARD ESTIMATES OF CORRELATION

Let $(x_1,y_1),(x_2,y_2), \ldots, (x_n,y_n)$ be a random sample of n pairs of observations. Then the estimate of the linear correlation coefficient $\rho_p = E\left[\left(\frac{X-\mu_x}{\sigma_x}\right)\left(\frac{Y-\mu_y}{\sigma_y}\right)\right]$ is Pearson's product-moment correlation:

$$r_p = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad \text{where } \bar{x} = \sum_{i=1}^{n}\frac{x_i}{n} \text{ and } \bar{y} = \sum_{i=1}^{n}\frac{y_i}{n}.$$

Let $R_1,R_2, \ldots, R_n$ denote the ranks of the $x_1, x_2, \ldots, x_n$ and let $S_1,S_2, \ldots, S_n$ denote the ranks of the $y_1,y_2, \ldots ,y_n$. Then the estimate of the grade correlation coef-

$$\left[\left(\frac{M-\mu_M}{\sigma_M}\right)\left(\frac{W-\mu_W}{\sigma_W}\right)\right]$$

$G(y)$ are the distribution functions of X and Y, is Spearman's rank correlation:

$$r_s = \frac{\sum_{i=1}^{n}(R_i-\bar{R})(S_i-\bar{S})}{\sqrt{\sum_{i=1}^{n}(R_i-\bar{R})^2\sum_{i=1}^{n}(S_i-\bar{S})^2}} \qquad \text{where } \bar{R} = \sum_{i=1}^{n}\frac{R_i}{n} \text{ and } \bar{S} = \sum_{i=1}^{n}\frac{S_i}{n}.$$

DISCUSSION OF PRECEDING PAPER

ELIAS S.W. SHIU:

The authors are to be thanked for this interesting survey on the bounds of bivariate distributions. It is perhaps useful to point out that the theory discussed in the paper can be applied to life insurance as well as annuities. Given the random variable $T(x, y)$, we have

$$E(v^{\lceil mT(x,y)\rceil/m}) = A_{xy}^{(m)}, \tag{1}$$

$$E(\ddot{a}^{(m)}_{\overline{\lceil mT(x,y)\rceil/m}|}) = \ddot{a}_{xy}^{(m)},$$

$$E(a^{(m)}_{\overline{\lfloor mT(x,y)\rfloor/m}|}) = a_{xy}^{(m)}, \text{ and so on,}$$

where, for a real number t, $\lceil t\rceil$ denotes the least integer greater than or equal to t, and $\lfloor t\rfloor$ denotes the greatest integer less than or equal to t (see [2]). Relationships such as

$$A_{xy} = 1 - d\ddot{a}_{xy}$$

and

$$A_{\overline{xy}} = 1 - d\ddot{a}_{\overline{xy}}$$

may be used to convert the numerical annuity values given in the paper into life insurance values, although different mortality tables usually are used to compute insurance values.

A key feature in the new *Actuarial Mathematics* textbook [1] is the consistent application of the assumption of a *uniform distribution of deaths (UDD) throughout each year of age* to express values payable at the end of $1/m$ of a year in terms of those payable once a year. Under the independence assumption, such formulas for the joint-life status are quite complicated. Without the independence assumption, as advocated in this paper, these formulas will be even more difficult to derive. Let me illustrate the complexity in deriving such formulas with one example [1,(8.7.8)]. I shall show that under the UDD and independence assumptions,

$$A_{xy}^{(m)} = (i/i^{(m)})\{A_{xy} + [(m+1)/m + 2(1/i - 1/d^{(m)})] \sum_{k\geq 0} v^{k+1}\, {}_{k|}q_x\, {}_{k|}q_y\}. \tag{2}$$

By (1)

$$A_{xy}^{(m)} = \int_0^\infty v^{\lceil mt\rceil/m} d_t q_{xy}. \tag{3}$$

By the independence assumption, the differential $d_t q_{xy}$ becomes

$$_t p_y \; d_t q_x \; + \; _t p_x \; d_t q_y,$$

which, in turn, by the UDD assumption, can be written as

$$\begin{aligned}
&[{}_{\lfloor t\rfloor}p_y - (t-\lfloor t\rfloor)_{\lfloor t\rfloor|}q_y]_{\lfloor t\rfloor|}q_x \, dt + [{}_{\lfloor t\rfloor}p_x - (t-\lfloor t\rfloor)_{\lfloor t\rfloor|}q_x]_{\lfloor t\rfloor|}q_y \, dt \\
&= [{}_{\lfloor t\rfloor}p_y \; {}_{\lfloor t\rfloor|}q_x + {}_{\lfloor t\rfloor}p_x \; {}_{\lfloor t\rfloor|}q_y - 2(t-\lfloor t\rfloor)_{\lfloor t\rfloor|}q_x \; {}_{\lfloor t\rfloor|}q_y] \, dt \\
&= \{[{}_{\lfloor t\rfloor}p_y \; {}_{\lfloor t\rfloor|}q_x + {}_{\lfloor t\rfloor}p_x \; {}_{\lfloor t\rfloor|}q_y - {}_{\lfloor t\rfloor|}q_x \; {}_{\lfloor t\rfloor|}q_y] + [1-2(t-\lfloor t\rfloor)]_{\lfloor t\rfloor|}q_x \; {}_{\lfloor t\rfloor|}q_y\} \, dt \\
&= \{{}_{\lfloor t\rfloor|}q_{xy} + [1-2(t-\lfloor t\rfloor)]_{\lfloor t\rfloor|}q_x \; {}_{\lfloor t\rfloor|}q_y\} \, dt.
\end{aligned}$$

(In the preceding derivation, we ignore the points where the values of t are integers; these points form a set of measure zero.) Hence, the right side of (3) becomes

$$\begin{aligned}
&\int_0^\infty v^{\lceil mt\rceil/m}\{{}_{\lfloor t\rfloor|}q_{xy} \; + \; [1 \; - \; 2(t \; - \lfloor t\rfloor)]_{\lfloor t\rfloor|}q_x \; {}_{\lfloor t\rfloor|}q_y\} \, dt \\
&= \int_0^\infty v^{\lceil mt\rceil/m} \; {}_{\lfloor t\rfloor|}q_{xy} \, dt \; + \; \int_0^\infty v^{\lceil mt\rceil/m}[1 \; - \; 2(t \; - \; \lfloor t\rfloor)]_{\lfloor t\rfloor|}q_x \; {}_{\lfloor t\rfloor|} \, q_y \, dt.
\end{aligned} \tag{4}$$

The integrals in (4) will be evaluated using the *Average Value Theorem* [2, p. 579] (also see [1, section 3.6]).

The first integral in (4) is

$$\begin{aligned}
&\int_0^\infty (1 \; + \; i)^{\lceil t\rceil - \lceil mt\rceil/m} \, v^{\lceil t\rceil} \; {}_{\lfloor t\rfloor|}q_{xy} \, dt \\
&= \left(\int_0^1 (1 \; + \; i)^{\lceil t\rceil - \lceil mt\rceil/m} \, dt\right) \int_0^\infty v^{\lceil t\rceil} \; {}_{\lfloor t\rfloor|}q_{xy} \, dt \\
&= \left(\int_0^1 (1 \; + \; i)^{1 - \lceil mt\rceil/m} dt\right) A_{xy} \\
&= s_{\overline{1}|}^{(m)} A_{xy} \\
&= (i/i^{(m)}) \, A_{xy}.
\end{aligned}$$

The second integral in (4) can be written as

$$\int_0^\infty (1+i)^{\lceil t\rceil - \lceil mt\rceil/m}\, v^{\lceil t\rceil}\,[1-2(t-\lfloor t\rfloor)]_{\lfloor t\rfloor|}q_x\ _{\lfloor t\rfloor|}q_y\,dt. \tag{5}$$

The average value of the periodic function in the integrand in (5) is

$$\int_0^1 (1+i)^{\lceil t\rceil - \lceil mt\rceil/m}[1-2(t-\lfloor t\rfloor)]dt$$

$$= \int_0^1 (1+i)^{1-\lceil mt\rceil/m}(1-2t)dt$$

$$= (i/i^{(m)}) - 2(1+i)\int_0^1 v^{\lceil mt\rceil/m}t\,dt.$$

Now,

$$\int_0^1 v^{\lceil mt\rceil/m}t\,dt = \int_0^1 v^{\lceil mt\rceil/m}\,[\lceil mt\rceil/m - (\lceil mt\rceil/m - t)]dt$$

$$= (I^{(m)}a)^{(m)}_{\overline{1|}} - \int_0^1 v^{\lceil mt\rceil/m}\,(\lceil mt\rceil/m - t)dt$$

$$= (d/i^{(m)})[1/d^{(m)} - 1/i] - (1/2m)a^{(m)}_{\overline{1|}}$$

$$= (d/i^{(m)})[1/d^{(m)} - 1/i - 1/2m].$$

Thus, the second integral in (4) is equal to

$$(i/i^{(m)})[1-2(1/d^{(m)} - 1/i - 1/2m)]\int_0^\infty v^{\lceil t\rceil}\ _{\lfloor t\rfloor|}q_x\ _{\lfloor t\rfloor|}q_y\,dt,$$

and equation (2) follows.

REFERENCES

1. BOWERS, N.L., GERBER, H.U., HICKMAN, J.C., JONES, D.A., AND NESBITT, C.J. *Actuarial Mathematics*, Itasca, Illinois: Society of Actuaries, 1987.
2. SHIU, E.S.W. "Integer Functions and Life Contingencies," *TSA*, XXXIV (1982), 571–90; Discussion, 591–600.

(AUTHORS' REVIEW OF DISCUSSION)

JACQUES F. CARRIÈRE AND LAI K. CHAN:

A main objective of this project is to illustrate, through annuity examples, the effect of dependence on actuarial calculations. We appreciate that Dr.

Shiu, using his broad knowledge of actuarial research and profound mathematical skill, has pointed in other possible research directions.

From the viewpoint of theoretical statistics, stochastic independence of random variables simplifies algebraic derivatives of formulas considerably. But in practice it is not always realistic to assume independence. With dependent random variables, actuarial models become more difficult to deal with, perhaps even intractable analytically. However, with the advances of computing capabilities and algorithms, approximate solutions to these models usually can be obtained through numerical computations or simulation. Mathematical complexity should no longer be considered as the major hurdle for developing more realistic actuarial models.