TRANSACTIONS OF SOCIETY OF ACTUARIES
1994 VOL. 46

DEPENDENT DECREMENT THEORY

JACQUES F. CARRIERE

ABSTRACT

Currently, multiple decrement theory is based on the assumption that competing causes of decrement are stochastically independent, even though this assumption is usually not true in reality. This paper presents well-known results in the theory of dependent competing risks that are fundamental in extending multiple decrement theory towards a dependent decrement theory. First, the state of the art is examined and the results that are based on the independence assumption are identified. Next, we use the well-developed theory of copula functions to model dependence, and we present a theorem that characterizes the mathematical relationship between the crude and net probabilities when the decrements are dependent. We also discuss the issue of identifiability and the related issue of measuring the effect of removing causes of decrement. Finally, we use an identifiability result to analyze the effect of removing heart/cerebrovascular diseases from the U.S. population when these diseases are correlated to other causes of death.

I. INTRODUCTION

The book *Actuarial Mathematics* [4] develops multiple decrement theory under the convenient assumption that the competing causes of decrement are stochastically independent. Hooker and Longley-Cook [12] state that it has long been known that every decrement is "selective" or dependent to a greater or less degree. We believe that the errors in analysis due to the independence assumption are unacceptable. To rectify this situation, this paper presents some results in the theory of dependent competing risks that are fundamental in extending multiple decrement theory towards a dependent decrement theory.

We begin the discussion by redefining the standard actuarial functions in multiple decrement theory in terms of *latent* random variables. Thus we can use standard actuarial notation when restating some well-known results in the theory of competing risks. Specifically, we identify those relationships in the current theory that are based on the independence assumption and those that are invariant to it.

Next, we show how to characterize the dependence structure of any continuous multivariate probability distribution with a *copula* function. This allows us to generalize the current independent decrement theory to a *dependent decrement theory* and also provides some insight into the identifiability problem that biostatisticians have been investigating. We discover that in the nonparametric case identifiability is only possible if the copula is fixed.

Finally, we investigate the effect of removing heart/cerebrovascular diseases as a cause of death from the U.S. population, assuming that heart/cerebrovascular diseases are dependent on the other causes. We discover that if the correlation between decrements is negative, then removing a cause of death extends the median lifetime more than if the correlation is positive. This result is qualitatively similar to a result reported in Yashin, Manton and Stallard [25] in which a stochastic process model is investigated.

2. DEFINITIONS AND BASIC RESULTS

In this section, we redefine the standard actuarial functions in multiple decrement theory in terms of *latent* random variables. Using actuarial notation, we restate some well-known results in the theory of competing risks, and we identify those relationships in the current theory that are based on the independence assumption. Many of the definitions in this section can be found in the book *Actuarial Mathematics* [4].

Following the example of Elandt-Johnson [7], let $0 \le T_j < \infty$ for $j = 1, \ldots, m$ be the *latent* random time of withdrawal, due to cause j, for a life aged $a \ge 0$. Note that these random variables may be stochastically dependent and they are not observable in a competing risks model. Usually, biostatisticians and actuaries assume that some (but not all) of these random variables may be defective; that is, $\Pr(T_j < \infty) < 1$. This assumption leads to some thorny theoretical problems that complicate the analysis while contributing nothing practical to the model. In fact, this assumption is not even testable, and so we always assume that $\Pr(T_j < \infty) = 1$.

Denote the radix of a multiple decrement table as $l_a^{(\tau)} > 0$ and let $x \ge a$ denote the attained age. Also, let $t \ge 0$ and $t_j \ge 0$ for $j = 1, \ldots, m$. Define a multivariate survival function as

$$S(t_1, \ldots, t_m) = \Pr(T_1 > t_1, \ldots, T_m > t_m). \tag{2.1}$$

See Tucker [23] for a good discussion of multivariate probability distributions. Another way to define a multivariate survival function is given in Yashin, Manton and Stallard [25]. Throughout the paper, we assume that $S(t_1, \ldots, t_m)$ is absolutely continuous. That is, there exists a function $f(t_1, \ldots, t_m)$ such that

$$S(t_1, \ldots, t_m) = \int_{t_1}^{\infty} \cdots \int_{t_m}^{\infty} f(s_1, \ldots, s_m)\, ds_m, \ldots ds_1. \tag{2.2}$$

In reality, some decrements occur only at year-ends and so the absolute continuity assumption is not valid in all cases, but we believe that this is a good approximating assumption. A more precise, albeit complicated, theory can be constructed with Riemann-Stieltjes integrals, as Shiu [21] suggests.

A. The Overall Survival Function

In this section, we present certain actuarial functions associated with the random variable $\min(T_1, \ldots, T_m)$. Using this random variable, we define the *overall* survival function as $S(t, \ldots, t) = \Pr[\min(T_1, \ldots, T_m) > t]$. Whenever the functions exist, define

$$S^{(\tau)}(t) = S(t, \ldots, t) \tag{2.3a}$$

$$l_x^{(\tau)} = l_a^{(\tau)}\, S^{(\tau)}(x - a), \quad x \geq a \tag{2.3b}$$

$${}_t d_x^{(\tau)} = l_x^{(\tau)} - l_{x+t}^{(\tau)} \tag{2.3c}$$

$${}_t q_x^{(\tau)} = \frac{{}_t d_x^{(\tau)}}{l_x^{(\tau)}} \tag{2.3d}$$

$${}_t p_x^{(\tau)} = \frac{l_{x+t}^{(\tau)}}{l_x^{(\tau)}} \tag{2.3e}$$

$$\mu_{x+t}^{(\tau)} = -\frac{d}{dt} \log_e(l_{x+t}^{(\tau)}) \tag{2.3f}$$

$$f_x^{(\tau)}(t) = {}_t p_x^{(\tau)}\, \mu_{x+t}^{(\tau)}. \tag{2.3g}$$

Using the definitions, it is easy to verify that

$$\frac{d}{dt}\,{}_tq_x^{(\tau)} = f_x^{(\tau)}(t) \qquad \text{and} \qquad {}_tq_x^{(\tau)} = \int_0^t {}_sp_x^{(\tau)}\,\mu_{x+s}^{(\tau)}\,ds. \tag{2.4}$$

B. The Crude Survival Function

In this section, we present certain actuarial functions associated with the joint distribution of $\min(T_1, \ldots, T_m)$ and the index random variable

$$J = \sum_{j=1}^{m} jI[\min(T_1, \ldots, T_m) = T_j]. \tag{2.5}$$

In this case, the *crude* survival function is equal to $\Pr[\min(T_1, \ldots, T_m)>t, J=j]$. To ensure that all the mass of J is on the integers $1, \ldots, m$, we assume that $\Pr(T_j=T_i)=0$ whenever $j\neq i$. Whenever the functions exist, define

$$S^{(j)}(t) = \Pr[\min(T_1, \ldots, T_m) > t, J = j] \tag{2.6a}$$

$$l_x^{(j)} = l_a^{(\tau)}\,S^{(j)}(x-a), x \geq a \tag{2.6b}$$

$${}_td_x^{(j)} = l_x^{(j)} - l_{x+t}^{(j)} \tag{2.6c}$$

$${}_tq_x^{(j)} = \frac{{}_td_x^{(j)}}{l_x^{(\tau)}} \tag{2.6d}$$

$${}_tp_x^{(j)} = \frac{l_{x+t}^{(j)}}{l_x^{(\tau)}} \tag{2.6e}$$

$$\mu_{x+t}^{(j)} = \frac{-\dfrac{d}{dt}\,l_{x+t}^{(j)}}{l_{x+t}^{(\tau)}} \tag{2.6f}$$

$$f_x^{(j)}(t) = {}_tp_x^{(\tau)}\,\mu_{x+t}^{(j)}. \tag{2.6g}$$

A few observations about the definitions in (2.6a–g) are in order. First, ${}_tq_x^{(j)}$ is usually a defective cumulative distribution function. Second, it is easy to verify that

$$\frac{d}{dt}\,{}_tq_x^{(j)} = f_x^{(j)}(t) \qquad \text{and} \qquad {}_tq_x^{(j)} = \int_0^t {}_sp_x^{(\tau)}\,\mu_{x+s}^{(j)}\,ds. \tag{2.7}$$

Third, our definition of ${}_tp_x^{(j)}$ is equivalent to the one given in Promislow [19], where ${}_tp_x^{(j)}={}_\infty q_x^{(j)}-{}_tq_x^{(j)}$, because ${}_\infty q_x^{(j)}={}_\infty d_x^{(j)}/l_x^{(\tau)}$ and ${}_\infty d_x^{(j)}=l_x^{(j)}$. It may be instructive to note that $S^{(j)}(t)={}_tp_a^{(j)}$. Finally, let $g(\cdot)$ denote any of the functions $S^{(\cdot)}(t)$, $l_x^{(\cdot)}$, ${}_td_x^{(\cdot)}$, ${}_tq_x^{(\cdot)}$, ${}_tp_x^{(\cdot)}$, $\mu_{x+t}^{(\cdot)}$, or $f_x^{(\cdot)}(t)$. Then it is easy to verify that the *additive property*,

$$g(\tau) = g(1) + \cdots + g(m), \tag{2.8}$$

is true regardless of any dependence assumption. The additive property holds because

$$\Pr\left\{\bigcap_{j=1}^{m} [J = j]\right\} = 0,$$

which means that the events are almost surely mutually exclusive.

The following lemma is an important result in the theory of competing risks. This result gives a representation of the crude survival function that will be useful later. An alternative proof of this lemma can be found in Tsiatis [22].

Lemma 1:

If $S(t_1, \ldots, t_m)$ is differentiable with respect to $t_j>0$ for all $j=1, \ldots, m$, then

$$S^{(j)}(t) = \int_t^\infty - S_j(r, \ldots, r)\, dr, \tag{2.9a}$$

where

$$S_j(r, \ldots, r) = \frac{\partial}{\partial t_j} S(t_1, \ldots, t_m)\big|_{t_k=r, \forall k}. \tag{2.9b}$$

Proof:

$$\begin{aligned} S^{(j)}(t) &= \Pr[\min(T_1, \ldots, T_m) > t, \quad J = j] \\ &= \Pr[(T_k > t \quad \text{and} \quad T_j \le T_k \ \forall\, k)] \\ &= \Pr[T_j > t \quad \text{and} \quad (T_k > T_j \ \forall\, k \ne j)] \\ &= \int_t^\infty \left\{ \int_{t_j}^\infty \cdots \int_{t_j}^\infty f(t_1, \ldots, t_m) \prod_{k \ne j} dt_k \right\} dt_j \end{aligned}$$

$$= \int_t^\infty \left\{ \frac{-\partial}{\partial t_j} S(t_1, \ldots, t_m)\big|_{t_k = t_j, \forall k} \right\} dt_j,$$

which is exactly equal to (2.9a–b). □

Sometimes the random variable J is stochastically independent of $\min(T_1, \ldots, T_m)$. This is true if and only if

$$S^{(j)}(s) = \Pr(J = j) \times \Pr[\min(T_1, \ldots, T_m) > s] = S^{(\tau)}(s)\, {}_\infty q_a^{(j)},$$

which is true if and only if ${}_s q_a^{(j)} = {}_s q_a^{(\tau)}\, {}_\infty q_a^{(j)}$, if and only if $f_a^{(j)}(s) = f_a^{(\tau)}(s)_\infty\ q_a^{(j)}$, if and only if $\mu_{a+s}^{(j)} = \mu_{a+s}^{(\tau)} \times {}_\infty q_a^{(j)}$ for all $s>0$. Note that ${}_\infty q_a^{(j)} = S^{(j)}(0)$. If we let $s = x+t-a$, then we can summarize the result as follows.

Lemma 2:

The random variable J is independent of $\min(T_1, \ldots, T_m)$, if and only if

$$\mu_{x+t}^{(j)} = \mu_{x+t}^{(\tau)} \times {}_x q_a^{(j)}\ \forall\ x, t, j. \tag{2.10}$$

If the identity in (2.10) holds, then we have the so-called proportional hazards model of the theory of competing risks. The result in Lemma 2 is a restatement, using actuarial notation, of a theorem given in Kochar and Proschan [14].

C. The Net Survival Function

In this section, we present certain actuarial functions associated with the random variable T_j. In this case, the *net* survival function is equal to $\Pr(T_j > t)$. Whenever the functions exist, define

$$S'^{(j)}(t) = \Pr(T_j > t) \tag{2.11a}$$

$$l_x'^{(j)} = l_a'^{(j)}\, S'^{(j)}(x - a),\ x \geq a \tag{2.11b}$$

$${}_t d_x'^{(j)} = l_x'^{(j)} - l_{x+t}'^{(j)} \tag{2.11c}$$

$${}_t q_x'^{(j)} = \frac{{}_t d_x'^{(j)}}{l_x'^{(j)}} \tag{2.11d}$$

$${}_t p_x'^{(j)} = \frac{l_{x+t}'^{(j)}}{l_x'^{(j)}} \tag{2.11e}$$

$$\mu_{x+t}^{\prime(j)} = -\frac{d}{dt}\log_e(l_{x+t}^{\prime(j)}) \tag{2.11f}$$

$$f_x^{\prime(j)}(t) = {}_tp_x^{\prime(j)}\,\mu_{x+t}^{\prime(j)} \tag{2.11g}$$

A few observations about the definitions in (2.11a–g) are in order. First, note that we introduced a new symbol $\mu_{x+t}^{\prime(j)}$. The book *Actuarial Mathematics* [4] defines

$${}_tp_x^{\prime(j)} = \exp\left\{-\int_0^t \mu_{x+s}^{(j)}\,ds\right\},$$

which implies that $\mu_{x+t}^{(j)}=\mu_{x+t}^{\prime(j)}$. In Theorem 3, we show that independence is a sufficient condition for this identity to hold. Finally, it is easy to verify that

$$\frac{d}{dt}\,{}_tq_x^{\prime(j)} = f_x^{\prime(j)}(t) \qquad \text{and} \qquad {}_tq_x^{\prime(j)} = \int_0^t {}_sp_x^{\prime(j)}\,\mu_{x+s}^{\prime(j)}\,ds. \tag{2.12}$$

D. The Consequences of Independence

We end this section by giving a theorem and a corollary. The theorem proves that $\mu_{x+t}^{(j)}=\mu_{x+t}^{\prime(j)}$ when $T_1, \ldots, T_m$ are stochastically independent. The corollary characterizes ${}_tp_x^{\prime(j)}$ in terms of ${}_tp_x^{(\tau)}$ and ${}_\infty q_a^{(j)}$ when J is independent of $\min(T_1, \ldots, T_m)$. These results are a restatement, using actuarial notation, of some theorems given in Elandt-Johnson and Johnson [8].

Theorem 3:
If $T_1, \ldots, T_m$ are stochastically independent, then

(*i*) $$\mu_{x+t}^{(j)} = \mu_{x+t}^{\prime(j)} \tag{2.13}$$

(*ii*) $${}_tp_x^{(\tau)} = \prod_{j=1}^{m} {}_tp_x^{\prime(j)}. \tag{2.14}$$

Proof:
(*i*) Lemma 1 implies that

$$f_a^{(j)}(s) = -\frac{d}{ds}S^{(j)}(s) = -S_j(s, \ldots, s).$$

Using stochastic independence, we get

$$f_a^{(j)}(s) = -\frac{\partial}{\partial t_j}\prod_{i=1}^{m} S'^{(i)}(t_i)\Big|_{t_i=s,\forall i} = \mu_{a+s}'^{(j)}\prod_{i=1}^{m} S'^{(i)}(s) = \mu_{a+s}'^{(j)} \times S^{(\tau)}(s).$$

Using our definitions, we know that $S^{(\tau)}(s) = {}_sp_a^{(\tau)}$ and that $f_a^{(j)}(s) = {}_sp_a^{(\tau)}\,\mu_{a+s}^{(j)}$. Therefore $\mu_{a+s}^{(j)} = \mu_{a+s}'^{(j)}$ for all $s>0$. The result follows by letting $s=x+t-a$.

(*ii*) Independence implies that

$$S(t_1, \ldots, t_m) = \prod_{j=1}^{m} S'^{(j)}(t_j).$$

So

$${}_tp_x^{(\tau)} = \frac{l_{x+t}^{(\tau)}}{l_x^{(\tau)}} = \frac{S^{(\tau)}(x+t-a)}{S^{(\tau)}(x-a)} = \prod_{j=1}^{m}\frac{S'^{(j)}(x+t-a)}{S'^{(j)}(x-a)}$$

$$= \prod_{j=1}^{m}\frac{l_{x+t}'^{(j)}}{l_x'^{(j)}} = \prod_{j=1}^{m} {}_tp_x'^{(j)}. \square$$

Corollary 4:

If $T_1, \ldots, T_m$ are independent and J is independent of $\min(T_1, \ldots, T_m)$, then

$${}_tp_x'^{(j)} = \{{}_tp_x^{(\tau)}\}^{\infty q_a^{(j)}}. \tag{2.15}$$

Proof:

Using the results in Lemma 2 and Theorem 3, we find that

$${}_tp_x'^{(j)} = \exp\{-\textstyle\int_0^t \mu_{x+s}^{(j)}\,ds\} = \exp\{-\int_0^t \mu_{x+s}^{(\tau)} \times {}_\infty q_a^{(j)}\,ds\} = \{{}_tp_x^{(\tau)}\}^{\infty q_a^{(j)}}. \square$$

3. COPULAS AND MEASURES OF ASSOCIATION

In this section, we show how to characterize the dependence structure of any continuous multivariate probability distribution. This allows us to generalize the current independent decrement theory to a *dependent decrement theory* and also provides some insight into the identifiability problem of the theory of competing risks.

Let $\mathbf{u}=(u_1, \ldots, u_m)' \in [0, 1]^m$. Following the example of Schweizer and Sklar [20], we define a *copula*, $C(\mathbf{u})$, as a multivariate cumulative

distribution function that has uniform marginals with support on the hypercube $[0, 1]^m$. This means that for all $j=1, \ldots, m$, we have $C(u_1, \ldots, u_{j-1}, 0, u_{j+1}, \ldots, u_m)=0$ and $C(1, \ldots, 1, u_j, 1, \ldots, 1)=u_j$. An example of a copula is

$$C(\mathbf{u}) = \prod_{j=1}^{m} u_j$$

and another is $C(\mathbf{u})=\min(u_1, \ldots, u_m)$. Some other examples of two-dimensional copulas can be found in Barnett [3]. The copula function is very useful in understanding the dependence structure of multivariate probability distributions because of the following result.

Lemma 5:
Let

$$C(\mathbf{u}) = \Pr\left\{\bigcap_{j=1}^{m} [S'^{(j)}(T_j) \leq u_j]\right\}.$$

Then $C(\mathbf{u})$ is the only copula such that

$$S(t_1, \ldots, t_m) = C[S'^{(1)}(t_1), \ldots, S'^{(m)}(t_m)], \forall\, t_j \geq 0 \text{ and } j = 1, \ldots, m. \quad (3.1)$$

Proof:
It is well-known that if $S'^{(j)}(t)$ is a continuous survival function, then the transformed random variable $S'^{(j)}(T_j)$ has a uniform distribution on [0, 1]. Therefore, $C(\mathbf{u})$ is a copula because it is a cumulative distribution function with uniform marginals. Next,

$$\begin{aligned} C[S'^{(1)}(t_1), \ldots, S'^{(m)}(t_m)] &= \Pr\left\{\bigcap_{j=1}^{m} [S'^{(j)}(T_j) \leq S'^{(j)}(t_j)]\right\} \\ &= \Pr\left\{\bigcap_{j=1}^{m} [T_j > t_j]\right\} \\ &= S(t_1, \ldots, t_m), \end{aligned}$$

because the event $[S'^{(j)}(T_j) \leq S'^{(j)}(t_j)]$ is equal to $[T_j > t_j]$, except on a set of probability 0. Let us suppose that $C(\cdot)$ is not unique, then there exists $C^*(\cdot) \neq C(\cdot)$ such that

$$S(t_1, \ldots, t_m) = C[S'^{(1)}(t_1), \ldots, S'^{(m)}(t_m)]$$
$$= C^*[S'^{(1)}(t_1), \ldots, S'^{(m)}(t_m)], \forall\, t_j \geq 0 \quad \text{and} \quad j = 1, \ldots, m.$$

Let $\mathbf{u}^*=(u_1^*, \ldots, u_m^*)' \in [0, 1]^m$ be a value such that $C^*(\mathbf{u}^*) \neq C(\mathbf{u}^*)$. Using the continuity of $S'^{(j)}(t_j)$, we know there exists $t_j^* \in [0, \infty]$ such that $S'^{(j)}(t_j^*)=u_j^*$. Therefore

$$C[S'^{(1)}(t_1^*), \ldots, S'^{(m)}(t_m^*)] \neq C^*[S'^{(1)}(t_1^*), \ldots, S'^{(m)}(t_m^*)].$$

This is a contradiction. Therefore $C(\cdot)$ is unique. □

As far as we know, Lemma 5 is a new result because of its focus on survival functions. Generalizing Lemma 5 to defective random variables would be an interesting research problem. Note that if $T_1, \ldots, T_m$ are stochastically independent, then the unique copula associated with $S(t_1, \ldots, t_m)$ is equal to $\prod_{j=1}^m u_j$. Moreover, if $T_1 = \cdots = T_m$, then the unique copula associated with $S(t_1, \ldots, t_m)$ is equal to $\min(u_1, \ldots, u_m)$. This last copula is actually an upper bound because $C(\mathbf{u}) \leq \min(u_1, \ldots, u_m)$, for any copula $C(\mathbf{u})$ and for all $\mathbf{u} \in [0, 1]^m$. For more information about copulas, consult Genest and MacKay [9], [10].

We are now in a position to give a representation of the crude survival function in terms of copulas. Using the results in Lemma 1 and Lemma 5, along with the chain rule, we get the following result, which appears in Heckman and Honore [11]. We apply this result when investigating the effect of removing heart/cerebrovascular diseases from the U.S. population.

Theorem 6:

If $C(u_1, \ldots, u_m)$ is differentiable with respect to $u_j \in (0, 1)$ and $S'^{(j)}(t_j)$ is differentiable with respect to $t_j > 0$ for all $j = 1, \ldots, m$, then

$$\frac{d}{dt} S^{(j)}(t) = C_j[S'^{(1)}(t), \ldots, S'^{(m)}(t)] \times \frac{d}{dt} S'^{(j)}(t), \tag{3.2a}$$

where

$$C_j(u_1, \ldots, u_m) = \frac{\partial}{\partial u_j} C(u_1, \ldots, u_m). \tag{3.2b}$$

Note that (3.2a) gives a nonlinear system of differential equations, where the net function $S'^{(j)}(t_j)$ can be solved if the copula $C(\cdot)$ and the crude

function $S^{(j)}(t_j)$ are given. We discuss this problem in Section 4. To understand Section 5, the reader needs to know something about measures of association. We now present a concise discussion of this important topic.

A. Measures of Association

Nonparametric measures of association are very useful for understanding the nature of the dependence in a copula. They are also useful for parametrizing families of copulas, as we will see later. Two examples of these measures are Spearman's ρ and Kendall's τ. Given a bivariate copula function $C(u_1, u_2)$, we can calculate these two correlation coefficients as follows,

$$\rho = 12 \int_{[0,1]^2} u_1 u_2 dC(u_1, u_2) - 3, \tag{3.3}$$

and

$$\tau = 4 \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1. \tag{3.4}$$

Note that $|\rho| \leq 1$, $|\tau| \leq 1$ and $\rho = \tau = -1$ if and only if $C(u_1, u_2) = \max(0, u_1 + u_2 - 1)$ and $\rho = \tau = +1$ if and only if $C(u_1, u_2) = \min(u_1, u_2)$. The copulas $\max(0, u_1 + u_2 - 1)$ and $\min(u_1, u_2)$ are called the Frechet bounds, because

$$\max(0, u_1 + u_2 - 1) \leq C(u_1, u_2) \leq \min(u_1, u_2) \quad \text{for all } C(u_1, u_2).$$

For more information, see Genest and MacKay [9], [10] or Carriere and Chan [6]. One of the most simple copulas is the Morgenstern copula. We use this copula for illustrative purposes in Section 4.

B. The Morgenstern Copula

Let $|\rho| < 1/3$ and let $0 \leq u_j \leq 1$ for $j = 1, 2$; then the bivariate Morgenstern copula is equal to

$$C(u_1, u_2) = u_1 u_2 \{1 + 3\rho(1 - u_1)(1 - u_2)\}. \tag{3.5}$$

This copula is parametrized with Spearman's ρ; that is,

$$\rho = 12 \int u_1 u_2 dC(u_1, u_2) - 3.$$

Note that this family does not include the Frechet bounds $\max(0, u_1+u_2-1)$ and $\min(u_1, u_2)$. The Morgenstern copula is an example of a one-parameter family of copulas. Other examples of one-parameter families can be found in Barnett [3], Carriere and Chan [6], Genest and MacKay [9], Kimeldorf and Sampson [13], and Mardia [16]. The Morgenstern copula is used in our discussion of identifiability, given in Section 4. A copula that is useful in Section 5 is the normal copula.

C. The Normal Copula

Let us give an example of a multivariate copula. Specifically, let us give the probability density function of the copula associated with the multivariate normal distribution. This copula is used when we investigate the effect of removing heart/cerebrovascular diseases from the U.S. population. For t, $z \in \mathcal{R}$, define

$$\Phi(t) = \int_{-\infty}^{t} \phi(z)\, dz, \tag{3.6}$$

where

$$\phi(z) = (2\pi)^{-1/2} \exp\{-z^2/2\}. \tag{3.7}$$

Next, for $u \in (0, 1)$ define $\Phi^{-1}(u)$ as the inverse function of $\Phi(t)$; that is, $\Phi[\Phi^{-1}(u)]=u$. Next, let $\mathbf{R}=\{r_{kl}\}$ denote an $m \times m$ nonsingular correlation matrix. This is actually a variance-covariance matrix in which all the diagonal elements are equal to 1. Note that $\mathbf{R}$ is a symmetric and positive definite matrix. Let $\mathbf{z}=(z_1, \ldots, z_m)' \in \mathcal{R}^m$; then the probability density function of a standardized multivariate normal distribution is

$$h(\mathbf{z}) = (2\pi)^{-m/2} |\mathbf{R}|^{-1/2} \exp\left\{ -\frac{\mathbf{z}'\mathbf{R}^{-1}\mathbf{z}}{2} \right\}. \tag{3.8}$$

See Mardia, Kent and Bibby [17] for more details about the multivariate normal distribution. Let $\mathbf{u}=(u_1, \ldots, u_m)' \in (0, 1)^m$; then the density of the normal copula is

$$\frac{\partial^m C(\mathbf{u})}{\partial u_1, \ldots, \partial u_m} = \frac{h[\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_m)]}{\phi[\Phi^{-1}(u_1)] \times \cdots \times \phi[\Phi^{-1}(u_m)]} \tag{3.9}$$

We can express the coordinates in the $m \times m$ matrix $\mathbf{R}=\{r_{kl}\}$ as functions of Spearman's correlation coefficient. Using the results in Kruskal [15],

we find that $r_{kl}=2\sin(\pi\rho_{kl}/6)$, where $\rho_{kk}=1$ and if $k\neq l$, then ρ_{kl} is Spearman's correlation for the bivariate normal copula with parameter r_{kl}.

4. IDENTIFIABILITY AND THE EFFECT OF REMOVING CAUSES

In this section, we discuss the problem of *identifiability*. A competing risk model is identifiable only if the joint survival function $S(t_1, \ldots, t_m)$ can be calculated or identified by simply knowing the net survival function $S^{(j)}(t)$. We find that identifying $S(t_1, \ldots, t_m)$ is not possible unless some simplifying assumptions are made. In the nonparametric case, we find that identifiability occurs only if the copula is fixed.

To fix our ideas, consider a group from where a person may leave due to one of m different causes. Suppose that we can estimate $S^{(j)}(t)\forall j$ by simply observing this group, and suppose we wanted to measure the effect of removing a particular cause of decrement. We claim that removing cause j is mathematically equivalent to letting $S'^{(j)}(t)=1$ for all $t\geq 0$. Let $S'^{(-j)}(t)$ denote the overall survival function with cause j removed. Then

$$S'^{(-j)}(t) = C[S'^{(1)}(t), \ldots, S'^{(j-1)}(t), 1, S'^{(j+1)}(t), \ldots, S'^{(m)}(t)]. \quad (4.1)$$

This representation of $S'^{(-j)}(t)$ reveals that we can measure the effect of removing a cause only when we know the copula and the marginals. In the case of independence, measuring the effect is a straightforward exercise because knowing $S^{(j)}(t)$ $\forall j$ is equivalent to knowing $S'^{(j)}(t)$ $\forall j$, which is equivalent to knowing $S(t_1, \ldots, t_m)$. Measuring this effect is not possible, in general. But Arnold and Brockett [1] and Bagai and Rao [2] have shown that for certain dependent and parametric models, we can still identify $S(t_1, \ldots, t_m)$ by simply knowing $S^{(j)}(t)\forall t, j$. Heckman and Honore [11] also give an identifiability theorem in a semiparametric case. An identifiability result that was applied to cancer data is given in Yashin, Manton and Stallard [25]. Let us investigate the issue of identifiability.

A. A Definition of Identifiability

In this section we give a concise definition of identifiability. Let $\mathcal{F}$ denote a family of multivariate survival functions with members denoted as $\Pr(T_1>t_1, \ldots, T_m>t_m)=S(t_1, \ldots, t_m)$. Also, let $\mathcal{G}$ denote a family of crude survival functions with members denoted as $\Pr[\min(T_1, \ldots, T_m)>t,$

$\min(T_1, \ldots, T_m)=T_j]=S^{(j)}(t)$. Now, let Ψ be a mapping from $\mathcal{F}$ onto $\mathcal{G}$, defined so that

$$\Psi[S(\cdot)] = S^{(\cdot)}(\cdot). \tag{4.2}$$

We say that $\mathcal{F}$ is identifiable by $\mathcal{G}$ whenever Ψ is an *injective* function. This means that if $\Psi(S_1)=\Psi(S_2)$, then $S_1=S_2$. It is well-known that the function Ψ is not necessarily injective. This situation usually occurs in the so-called nonparametric case.

B. The Nonparametric Case

In this case, let $\mathcal{F}$ denote a family in which the survival functions are absolutely continuous and there is no restriction on the dependence structure. Also let $\mathcal{F}_C \subset \mathcal{F}$ be a subfamily in which the copula $C(\mathbf{u})$ of each multivariate survival function in $\mathcal{F}_C$ is the same. As an example, if

$$C(\mathbf{u}) = \prod_{j=1}^{m} u_j$$

is the independent copula, then all the members of $\mathcal{F}_C$ have the form

$$S(t_1, \ldots, t_m) = \prod_{j=1}^{m} S'^{(j)}(t_j).$$

Using Theorem 3, we find that in the independent case, $\mathcal{F}_C$ is identifiable by $\mathcal{G}$. This immediately implies that $\mathcal{F}$ is not identifiable by $\mathcal{G}$.

Let us investigate the identifiability of $\mathcal{F}_C$ by $\mathcal{G}$, assuming that $C(\mathbf{u})$ is fixed and known, that it has continuous second-order partial derivatives, and that $d[S^{(j)}(t)]/dt$ is continuous. These continuity assumptions allow us to use an existence result from the theory of ordinary differential equations. Now, define $\mathbf{u}=(u_1, \ldots, u_m)'$ where $u_j=S'^{(j)}$ is an absolutely continuous survival function on $(0, \infty)$ whose probability density function is denoted as $-d[S'^{(j)}(t)]/dt$. Next, suppose we know that the crude function is equal to $S^{(j)}(t)$. Using Theorem 6, we find that we have a system of nonlinear differential equations. Specifically, we can write

$$\frac{d}{dt}\mathbf{u}(t) = \mathbf{G}[t, \mathbf{u}(t)], \tag{4.3}$$

where $\mathbf{G}(t, \mathbf{u}(t))=\{G_1[t, \mathbf{u}(t)], \ldots, G_m[t, \mathbf{u}(t)]\}'$ and

$$G_j(t, \mathbf{u}(t)) = \frac{\dfrac{d}{dt} S^{(j)}(t)}{C_j[\mathbf{u}(t)]}. \tag{4.4}$$

It is well-known (see Verhulst [24]) that with the initial condition $S'^{(j)}(0)=1 \forall j$ and with some other minor conditions, there exists a unique solution $\mathbf{u}^*$ to this system. This implies that the model is identifiable, whenever $C(\mathbf{u})$ is known. Note that $\mathbf{u}^*$ may be difficult to calculate, even if we know it exists. Finally, the identifiability result given by Yashin, Manton and Stallard [25] also requires that a system of nonlinear differential equations be solved, although their approach is totally different from ours.

C. The Parametric Case

In this section, we describe in an abstract way how identifiability can be proved for parametric models, and we demonstrate the ideas with the Morgenstern copula presented earlier. The techniques presented here were used by Arnold and Brockett [1] and by Bagai and Rao [2] to prove identifiability.

Let $S(t_1, \ldots, t_m|\boldsymbol{\theta})$ denote a multivariate survival function indexed with some parameter $\boldsymbol{\theta}$ that belongs to a parameter space $\Theta \subseteq \mathcal{R}^p$ where $p=1, 2, \ldots$. As an example, consider a Morgenstern copula with

$$S'^{(j)}(t_j) = 1 - t_j, \quad 0 \le t_j \le 1, \quad j = 1, 2.$$

In this case,

$$S(t_1, t_2|\rho) = (1 - t_1)(1 - t_2)\{1 + 3\rho t_1 t_2\}$$

and $m=2$, $p=1$, $\boldsymbol{\theta}=\rho$, and $\Theta=(-1/3, +1/3)$.

As a function of $\boldsymbol{\theta}$, $S(\cdot|\boldsymbol{\theta})$ is a mapping of Θ onto $\mathcal{F}_\theta \subset \mathcal{F}$. The composition $\Psi \circ S=\Psi[S(\cdot|\boldsymbol{\theta})]$ is a mapping of Θ onto $\mathcal{G}_\theta \subseteq \mathcal{G}$. If it is injective, then Ψ is injective on $\mathcal{F}_\theta$; thus $\mathcal{F}_\theta$ is identifiable by $\mathcal{G}_\theta$. Consider the Morgenstern example. In this case, we find that $\mathcal{F}_\rho=\{S(\cdot|\rho):|\rho|<1/3\}$ and $\mathcal{G}_\rho=\{S^{(\cdot)}(\cdot|\rho):|\rho|<1/3\}$ where $S^{(j)}(t|\rho)= 2^{-1}(1-t)^2(1+3\rho t^2)=\Psi \circ S(\rho)$. For any $t \neq 0, 1$, if $\Psi \circ S(\rho_1)=\Psi \circ S(\rho_2)$, then $\rho_1=\rho_2$ and we have injectivity immediately.

Sometimes it is difficult to prove the injectivity of $\Psi \circ S$. A technique that may simplify the analysis is to introduce an auxiliary function Λ that maps $\mathcal{G}_\theta$ onto a space $\mathcal{H}_\theta$ of statistical parameters. Let $\Gamma(\cdot|\boldsymbol{\theta}) \in \mathcal{H}_\theta$ denote

a statistical parameter. An example is the moment-generating function. The key is selecting Λ, so that the injectivity of $\Gamma=\Lambda \circ \psi \circ S$ is relatively easy to check. The injectivity of Γ immediately implies the injectivity of $\Psi \circ S$. Consider the Morgenstern example. If we let

$$\Gamma(\rho) = E[\min(T_1, \ldots, T_m)|\rho] = (10 + 3\rho)/30,$$

then $\mathcal{H}_\rho=\{\Gamma(\rho):|\rho|<1/3\}$ and $\Gamma(\rho)$ is injective, proving that $\mathcal{F}_\rho$ is identifiable by $\mathcal{G}_\rho$.

D. Conclusion

Essentially, we found that the only way that we can identify a unique survival function $S(t_1, \ldots, t_m)$ with the crude survival function $S^{(j)}(t)$ is by restricting the family of functions that $S(t_1, \ldots, t_m)$ may belong to. In the nonparametric case, we do this by assuming that the copula function is the same for all members in the class $\mathcal{F}_C$. While in the parametric case, we restrict the family by using a parametric survival function, denoted as $S(\cdot|\boldsymbol{\theta})$. In both the nonparametric and parametric cases, restricting the class $\mathcal{F}$ simply replaces the problem of identifiability with the equally thorny problem of deciding what the restriction will be. In conclusion, identifying $S(t_1, \ldots, t_m)$ is not possible without some simplifying assumptions.

5. AN APPLICATION

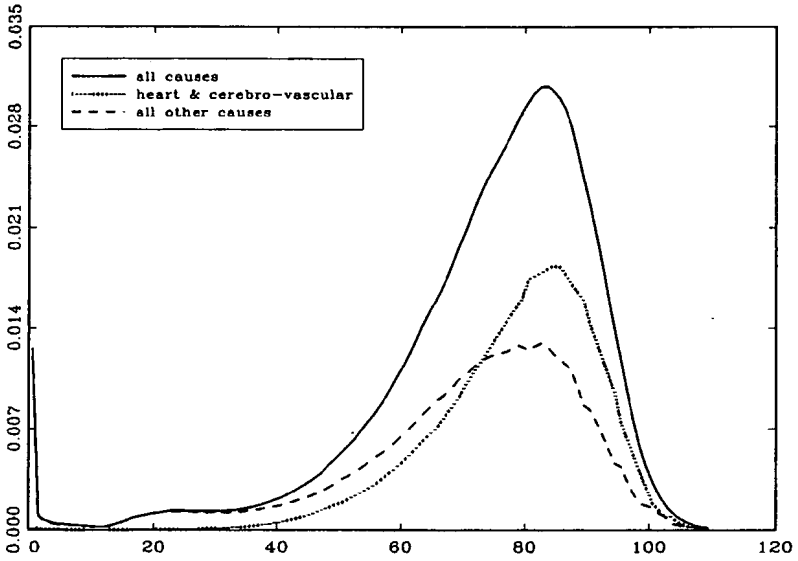
In this section, we investigate the effect of removing heart and cerebrovascular diseases as a cause of death from the U.S. population. Specifically, we calculate the net survival probabilities from the crude probabilities, assuming that these diseases are dependent on the other causes. We model the dependence with a normal copula that allows the Frechet bounds to be investigated.

The data come from the National Center for Health Statistics [18]. This publication gives the number of deaths from heart and cerebrovascular diseases and from the other causes, in five-year age groups. Using these data, we can calculate the crude survival functions $S^{(h)}(t)$ and $S^{(-h)}(t)$ for $t=0, 5, \ldots, 95, 100$. The superscript (h) denotes that heart/cerebrovascular diseases are the cause of death, while $(-h)$ denotes the other causes of death. Using the data, we find that the probability of dying from heart/cerebrovascular diseases is equal to ${}_{\infty}q_0^{(h)}=0.509$. Using a cubic polynomial, we interpolated the crude survival functions at $t=0, 1, \ldots, 104$

and we approximated the densities $f_0^{(h)}(t)$ and $f_0^{(-h)}(t)$. Figure 1 is a plot of $f_0^{(\tau)}(t)=f_0^{(h)}(t)+f_0^{(-h)}(t)$ and of $f_0^{(h)}(t)$, $f_0^{(-h)}(t)$. All the graphs and calculations were done with the statistical computing package GAUSS.

FIGURE 1

A PLOT OF THE DENSITIES $f_0^{(\tau)}$, $f_0^{(h)}$ AND $f_0^{(-h)}$ BASED ON THE 1979–81 U.S. LIFE TABLES BY CAUSE OF DEATH



The net survival functions $S'^{(h)}(t)$ and $S'^{(-h)}(t)$ can be found by solving a system of differential equations. To solve this system, we have to specify the form of the copula function. We use the normal copula given in Section 3 because it attains the Frechet bounds, when $|\rho|\to 1$, and its properties are well-documented. Currently, we cannot recommend it as an all-purpose model, but it does serve our purpose well. Consider the Equations (3.2a–b) that relate the net and crude probabilities with the copula function. With a normal copula, these equations yield the system

$$f_0^{(h)}(t) = C_1[S'^{(h)}(t), S'^{(-h)}(t)|\rho] \times f_0'^{(h)}(t),$$
$$f_0^{(-h)}(t) = C_2[S'^{(h)}(t), S'^{(-h)}(t)|\rho] \times f_0'^{(-h)}(t),$$

where

$$C_1(u, v|\rho) = \Phi\{[\Phi^{-1}(v) - r(\rho)\Phi^{-1}(u)]/\sqrt{1 - r(\rho)^2}\},$$

and $C_2(u, v|\rho)=C_1(v, u|\rho)$ and $r(\rho)=2\sin(\pi\rho/6)$. Note that we parametrized the copula with Spearman's ρ because we believe that this parametrization is informative, albeit complicated. The parameter ρ cannot be estimated because if it could, then the joint survival function

$$S(t_1, t_2|\rho) = C[S'^{(h)}(t_1), S'^{(-h)}(t_2)|\rho]$$

would be identifiable, but in Section 4 we showed that it is not.

Let us describe how we solved this system numerically. Most of the techniques that we used are given in Burden and Faires [5]. First, we transformed the differential system into a system of difference equations. We did this by letting

$$f_0^{(h)}(k + 0.5) \approx S^{(h)}(k) - S^{(h)}(k + 1),$$

$$f_0^{(-h)}(k + 0.5) \approx S^{(-h)}(k) - S^{(-h)}(k + 1),$$

$$f_0'^{(h)}(k + 0.5) \approx S'^{(h)}(k) - S'^{(h)}(k + 1),$$

$$f_0'^{(-h)}(k + 0.5) \approx S'^{(-h)}(k) - S'^{(-h)}(k + 1),$$

$$S'^{(h)}(k + 0.5) \approx 0.5 \times \{S'^{(h)}(k + 1) + S'^{(h)}(k)\},$$

$$S'^{(-h)}(k + 0.5) \approx 0.5 \times \{S'^{(-h)}(k + 1) + S'^{(-h)}(k)\},$$

for $k=0, 1, \ldots, 104$. Using the initial condition $S'^{(h)}(0)=S'^{(-h)}(0)=1$, we find that we can solve the problem recursively. Moreover, the problem reduces to finding the zeros of a sequence of nonlinear systems that were solved with Newton's method. To verify our numerical solution, we checked that

$$C[S'^{(h)}(t), S'^{(-h)}(t)|\rho] = S^{(h)}(t) + S^{(-h)}(t).$$

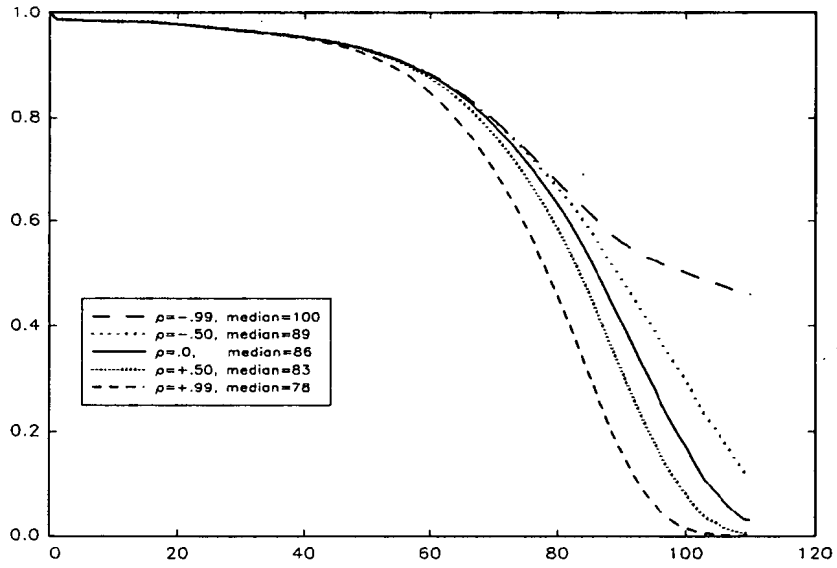
We solved the system under the assumption that Spearman's correlation is equal to $\rho=-0.99, -50, 0, +0.50, +0.99$. Note that the copula is not differentiable when ρ is equal to -1 or $+1$. If $\rho=0$, then the net probabilities are independent and we have the standard analysis. But if the correlation is $+0.99$, then this strong positive dependence means that removing heart/cerebrovascular diseases has little effect on survival. But if the correlation is -0.99, then this strong negative dependence means

that removing heart/cerebrovascular diseases will significantly increase the chances of survival.

These effects can be seen in Figure 2 in which $S'^{(-h)}(t)$ was plotted at $t=0, 1, \ldots, 110$ and $\rho=-0.99, -0.50, 0, +0.50, +0.99$. These graphs indicate that $S'^{(-h)}(t)$ increases when ρ decreases. If $\rho=-0.99$, then $S'^{(-h)}(t)$ is essentially an upper bound on the improvement in mortality that can be expected when heart/cerebrovascular diseases are removed as a cause of death. Moreover, if $\rho=+0.99$, then $S'^{(-h)}(t)$ is essentially a lower bound on the improvement in mortality that can be expected when heart/cerebrovascular diseases are removed.

FIGURE 2

A PLOT OF $S'^{(-h)}(t)$ WHEN THE HEART/CEREBROVASCULAR DISEASES ARE CORRELATED WITH THE OTHER CAUSES OF DEATH



The graph also reveals that if we remove heart/cerebrovascular diseases, then the median age at death of a newborn increases as ρ decreases. Currently, the median age at death of a newborn is 77. Under the standard analysis ($\rho=0$), removing heart/cerebrovascular diseases increases the median age at death to 86. If $\rho=+0.99$, then removing heart/cerebrovascular diseases only increases the median age to 78 but if

$\rho = -0.99$, then the median age increases to 100. Note that $\mathring{e}_0$ is not calculable because the survival function is not known after the age of 110.

6. SUMMARY

We showed that the effect of removing a cause of death depends on the copula used in the analysis. We found that if the correlation between decrements is negative, then removing a cause of death extends the median lifetime more than if the correlation is positive. We also found that a competing risks model is identifiable only when the class of potential models is greatly restricted. This means that we cannot identify $S(t_1, \ldots, t_m)$ without some simplifying assumptions. We also gave a theorem that characterizes the mathematical relationship between the crude and net probabilities when the decrements are dependent. Finally, we examined the current state of multiple decrement theory and we identified the results that depend on the independence assumption.

REFERENCES

1. ARNOLD, B.C., AND BROCKETT, P.L. "Identifiability for Dependent Multiple Decrement/Competing Risk Models," *Scandinavian Actuarial Journal* (1983): 117–27.
2. BAGAI, I., AND RAO, B.L.S.P. "Analysis of Survival Data with Two Dependent Competing Risks," *Biometrical Journal* 34 (1992): 801–14.
3. BARNETT, V. "Some Bivariate Uniform Distributions," *Communications in Statistics Part A: Theory and Methods* 9 (1980): 453–61.
4. BOWERS, N.L., GERBER, H.U., HICKMAN, J.C., JONES, D.A., AND NESBITT, C.J. *Actuarial Mathematics*. Itasca, Ill.: Society of Actuaries, 1986.
5. BURDEN, R.L., AND FAIRES, J.D. *Numerical Analysis. 4th ed.* Boston: PWS-Kent, 1989.
6. CARRIERE, J., AND CHAN, L. "The Bounds of Bivariate Distributions That Limit the Value of Last-Survivor Annuities," *TSA* XXXVIII (1986): 51–74.
7. ELANDT-JOHNSON, R.C. "Conditional Failure Time Distributions under Competing Risk Theory with Dependent Failure Times and Proportional Hazard Rates," *Scandinavian Actuarial Journal* (1976): 37–51.
8. ELANDT-JOHNSON, R.C., AND JOHNSON, N.L. *Survival Models and Data Analysis*. New York: Wiley, 1980.
9. GENEST, C., AND MACKAY, R.J. "Copules Archimediennes et Familles de Lois Bidimensionnelles Dont les Marges Sont Donnees," *Canadian Journal of Statistics* (1986): 145–59.

10. GENEST, C., AND MACKAY, J. "The Joy of Copulas: Bivariate Distributions with Uniform Marginals," *The American Statistician* 40 (1986): 280–83.
11. HECKMAN, J.J., AND HONORE, B.E. "The Identifiability of the Competing Risks Model," *Biometrika* 76 (1989): 325–30.
12. HOOKER, P.F., AND LONGLEY-COOK, L.H. *Life and Other Contingencies*. Vol. II. Cambridge, Great Britain: University Press, 1957.
13. KIMELDORF, G., AND SAMPSON, A. "Uniform Representations of Bivariate Distributions," *Communications in Statistics* 4 (1975): 617–27.
14. KOCHAR, S.C., AND PROSCHAN, F. "Independence of Time and Cause of Failure in the Multiple Dependent Competing Risks Model," *Statistica Sinica* 1 (1991): 295–99.
15. KRUSKAL, W.H. "Ordinal Measures of Association," *Journal of the American Statistical Association* 53 (1958): 814–59.
16. MARDIA, K.V. *Families of Bivariate Distributions*. London: Griffin, 1970.
17. MARDIA, K.V., KENT, J.T., AND BIBBY, J.M. *Multivariate Analysis*. New York: Academic Press, 1979.
18. NATIONAL CENTER FOR HEALTH STATISTICS, CURTIN, L.R., AND ARMSTRONG, R.J. "United States Life Tables Eliminating Certain Causes of Death, 1979–81." Vol. I. No. 2. In *U.S. Decennial Life Tables for* 1979–81. DHHS Pub. No. (PHS) 88-1150-2. Hyattsville, Md.: Public Health Service, 1988.
19. PROMISLOW, S.D. "Select and Ultimate Models in Multiple Decrement Theory," *TSA* XLIII (1991): 281–303.
20. SCHWEIZER, B., AND SKLAR, A. *Probabilistic Metric Spaces*. New York: North Holland, 1983.
21. SHIU, E. "Multiple-Decrements by Riemann-Stieltjes Integration," *ARCH* 1987.1: 1–4.
22. TSIATIS, A. "A Nonidentifiability Aspect of the Problem of Competing Risks," *Proceedings of the National Academy of Sciences U.S.A.* 72 (1975): 20–2.
23. TUCKER, H.G. *A Graduate Course in Probability*. New York: Academic Press, 1967.
24. VERHULST, F. *Nonlinear Differential Equations and Dynamical Systems*. New York: Springer-Verlag, 1990.
25. YASHIN, A., MANTON, K., AND STALLARD, E. "Dependent Competing Risks: A Stochastic Process Model," *Journal of Mathematical Biology* 24 (1986): 119–40.

DISCUSSION OF PRECEDING PAPER

STEVE CRAIGHEAD:

This paper was very enjoyable to read and will allow our multiple decrement models to reflect reality. I have three questions about the expansion and/or use of the theory.

Question 1. In most actuarial mathematical derivations, we assume that interest is independent of the decrements. With the introduction of the copula, you remove the independence requirement for the decrements from *Actuarial Mathematics* (your reference [4]). Can your theory be expanded to include the stochastic nature of interest rates?

Question 2. Can we compare your process of eliminating a decrement from the overall model to that of decomposing a vector into a linear combination of basis vectors? That is, could a multiple-decrement model be decomposed into its component decrements? Could these component decrements be manipulated by increasing or decreasing their influence to easily produce a new multiple-decrement model?

Question 3. Is the theory easily integrated into specific software packages such as S-Plus or APL?

JAMES W. DANIEL:

Dr. Carriere has made an interesting contribution to the problem of the relationship between a multiple-decrement model and single-decrement models that are associated with it in some manner. This is a difficult problem, inherent in the construction of multiple-decrement models that will be appropriate for a particular application. Although interesting, the present paper in my opinion only deals with the special case in which the various causes of decrement can each occur regardless of whether the others have occurred.

First, Dr. Carriere errs in asserting that the development of multiple-decrement theory in the textbook *Actuarial Mathematics* assumes that competing causes of decrement are stochastically independent. The assertion would be correct had *Actuarial Mathematics* presented such expressions as ${}_tp'_x(j)$ as representing probabilities associated with some event. In fact, the authors of that text are careful to point out that those symbols should not be so interpreted. A correct statement is that, *if* those expressions *were* in fact survival probabilities for independent causes, then the multiple-decrement model constructed from them would be

identical with the one from which they were derived. [This is analogous to representing a two-dimension force vector as the resultant of two perpendicular north-south and east-west forces, even if the original force had resulted from combining still other forces.] *Actuarial Mathematics* starts directly with the multiple-decrement model, rather than constructing it from competing causes. But this has little bearing on the rest of Dr. Carriere's article.

Dr. Carriere's analysis begins, not with a given multiple-decrement model, but instead with a set of so-called *latent* random variables representing survival times under various single risks. To my mind, the biggest difficulty lies right here in trying to understand what such latent variables might mean. For example, if we consider two causes of decrement from employment, namely, death and retirement, then I can understand the notion of the time until an employee dies, regardless of whether the employee retires from the company; that could well be a meaning for a latent random variable naturally associated with time until death. But I cannot intuitively understand a latent random variable naturally associated with time until retirement—surely it is not the time until retirement regardless of whether the employee dies. This makes it difficult for me to see intuitively how it is easier to construct a multiple-decrement model from latent random variables than to construct it out of whole cloth.

The preceding paragraph looked intuitively at constructing multiple-decrement models from latent random variables instead of just starting with the full multiple-decrement model. Let's look at this more mathematically.

Suppose that $\tilde{T}$ and $\tilde{J}$ are the fundamental random variables in a multiple-decrement model as described in *Actuarial Mathematics*; these would be denoted by T and J in that book. For each j, define the random variable T_j as

$$T_j = \tilde{T} + Y_j[1 - I_j(\tilde{J})]$$

where Y_j is any arbitrary positive random variable (or constant) and $I_j(\tilde{J})$ is the usual indicator function equal to 1 if $\tilde{J}=j$ and equal to 0 otherwise. Let's use these as the latent random variables T_j of Carriere. Clearly, $T_j>\tilde{T}$ for all $j\neq\tilde{J}$, while $T_j=\tilde{T}$ for $j=\tilde{J}$. That is, $\tilde{T}$ equals the minimum of the T_j, which is precisely what Dr. Carriere defines as T in his construction; moreover, the minimizing j that Dr. Carriere defines as J is in fact exactly our original $\tilde{J}$. This means that the multiple-decrement

model using T and J constructed from these latent random variables is precisely the original model using $\tilde{T}$ and $\tilde{J}$.

What's the point of the preceding construction? Just this: Since the Y_j were arbitrary positive random variables, the same multiple-decrement model can be constructed from infinitely many quite different sets of latent random variables; in fact, all sets of latent random variables producing the same multiple-decrement model must be of the form involving the Y_j. So I find it difficult to see how picking an appropriate set of latent random variables—either out of the air or from the experience in a variety of single- or multiple-decrement models—is any easier than writing down a full multiple-decrement model directly in terms of T and J.

The preceding criticisms apply only when it is unclear what might naturally be meant by the latent random variables. Of course in some cases—for example, the joint survival of two individuals, which can fail for either of two clearly understood latent causes (the death of either party)—there is a perfectly natural and meaningful set of latent random variables associated with the competing causes. In such cases Dr. Carriere's analysis in the present paper is quite informative.

ESTHER PORTNOY:

The preprint of Dr. Carriere's paper arrived as I was planning lectures for the second semester of my life-contingencies class, and I decided to try presenting the material to the graduate students (who must do some additional work to earn their higher level of credit). The experiment was quite successful, and I plan in the future to use this approach to multiple-decrement theory with undergraduates as well. My comments are restricted to the early portions of the paper, all I could present to the class. I do not mean thereby to suggest that the latter portions are uninteresting—far from it!

Dr. Carriere notes that the standard textbook *Actuarial Mathematics* develops multiple-decrement theory only under the assumption that different causes operate independently. One could make a stronger statement: the presentation obscures the fact that dependence might be an issue. I remember being confused about independence the first time I saw competing-risk models outside of Jordan, but until now I had not tried to raise the matter in class.

I presented to my graduate students the following simple example of two dependent latent variables. My primary aim was to demonstrate the

distinction between the forces of decrement $\mu^{(j)}$ and $\mu'^{(j)}$. Suppose the joint density function for two random variables (T_1, T_2) is

$$f(t_1, t_2) = c(t_1 + t_2) \text{ for } 0 < t_1 < a, 0 < t_2 < b,$$

vanishing elsewhere. We set

$$c = \left[ab \times \frac{a+b}{2} \right]^{-1}$$

so that the density integrates to 1, and for definiteness assume $a \geq b$. There is no claim that this joint distribution is at all natural or realistic; it just serves as an example for illustrative purposes.

It is easy to calculate

$$S^{(\tau)}(t) = (a-t)(b-t)\, c\left(\frac{a+b}{2} + t\right),$$

and the net survival functions

$$S'^{(1)}(t) = \Pr(T_1 > t) = c \int_t^a \int_0^b (t_1 + t_2)\, dt_2\, dt_1$$

$$= (a-t)\, bc \times \frac{1}{2}(a+b+t)$$

and

$$S'^{(2)}(t) = a(b-t)\, c \times \frac{1}{2}(a+b+t).$$

From the latter we obtain the forces

$$\mu_t'^{(1)} = \frac{b+2t}{(a-t)(a+b+t)}, \quad \mu_t'^{(2)} = \frac{a+2t}{(b-t)(a+b+t)}.$$

Calculation of the crude survival functions $S^{(1)}(t)$, $S^{(2)}(t)$ is facilitated by a geometric approach that was familiar to the class from earlier work with multilife functions. First, note that $S^{(1)}(t) = \Pr(t < T_1 \leq T_2) = 0$ if $t \geq b$. For $t < b$, we integrate the density function over the triangular region shaded in Figure 1, thus obtaining

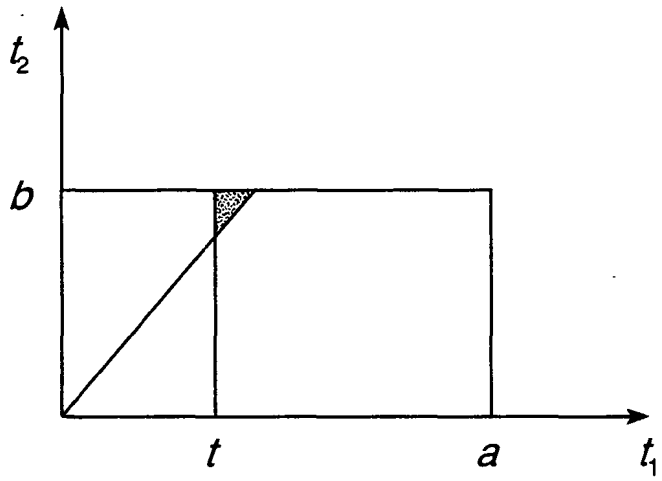
$$S^{(1)}(t) = \frac{1}{2}c(b - t)^2(b + t)$$

and

$$\mu_t^{(1)} = \frac{b + 3t}{(a - t)(a + b + 2t)}.$$

The formulas for $S^{(2)}(t)$ and $\mu_t^{(2)}$ are messier. Note that $\mu_t^{(1)}$ equals $\mu_t'^{(1)}$ only for $t=0$, and both tend to ∞ as $t \to a$.

FIGURE 1



Finally, it is not too hard to calculate the copula $C(u, v)$ associated with this joint distribution. For the sake of simplicity at this point I used specific numbers $a=20$, $b=10$, whence $c=1/3000$. I was a little surprised at the complexity of the copula associated with such a simple joint density; maybe Dr. Carriere or another reader can suggest a better example. We have

$$C(u, v) = \frac{(20 - t_1)(10 - t_2)(30 + t_1 + t_2)}{6000}$$

where

$$t_1 = \sqrt{625 - 600u} - 5,\ t_2 = \sqrt{400 - 300v} - 10.$$

Note that t_1 decreases strictly and continuously from 20 ($=a$) to 0 as u increases from 0 to 1, and t_2 decreases from 10 to 0 as v increases from 0 to 1. Clearly $C(0, v)=C(u, 0)=0$, and it takes only a bit of algebraic manipulation to confirm that $C(u, 1)\equiv u$ and $C(1, v)\equiv v$.

At this point I might note that I am puzzled by the paper's Lemma 5, or rather by its long proof. If the net survival functions $S'^{(j)}(t)$ are continuous, and none of the T_j are defective, then the equation

$$C(S'^{(1)}(t_1), \ldots, S'^{(m)}(t_m)) = S(t_1, \ldots, t_m)$$

defines C everywhere on the unit hypercube; uniqueness is obvious. One might hesitate to use this equation to define C because of the possible existence of vectors $\mathbf{t}$, $\mathbf{s}$ such that $S(\mathbf{t})\neq S(\mathbf{s})$ while $S'^{(j)}(t_j) = S'^{(j)}(s_j)$ for each j. By defining

$$C(u) = \Pr\{\cap_{j=1}^{m} [S'^{(j)}(T_j) \leq u_j]\},$$

and then showing that $C(S'^{(1)}(t_1), \ldots, S'^{(m)}(t_m))=S(t_1, \ldots, t_m)$, Dr. Carriere neatly finesses the difficulty. (In the given example, the net survival functions are one-to-one; I did not call students' attention to the fact that this might not always be the case.)

There was insufficient class time to discuss the Morgenstern and normal copulas or identifiability; these may be included in a more advanced course. Students were asked to read the application (Section 5), and they seem to have come away from the exercise with some appreciation of the issues of importance. I thank Dr. Carriere for this opportunity to show students some of the excitement of mathematical research in actuarial science.

S. DAVID PROMISLOW:

Dr. Carriere has done a great service by bringing to the attention of the actuarial profession, the work of biostatisticians in multiple-decrement theory. He illustrates that at times, our profession is in danger of being too narrowly focused. We are not always aware of developments in other fields that parallel our own interests, and we are indebted to those researchers who understand these developments and can make them known.

Although Dr. Carriere notes in several places that the paper is largely an expository piece and that many of the results are well known, he perhaps should have made this clearer. Some of his wording (the first sentence of the Abstract, for example) is misleading. It suggests that the study of multiple-decrement theory has been confined to the independent case and that the purpose of his paper is to extend this to the situation with dependent risks. The truth is, however, as the author himself shows, that this extension has already been done. The current actuarial syllabus deals only with the independent case, but the referenced work in the paper indicates that dependence has been extensively investigated by others. A casual reader of the paper may be even more confused on this point because of the distinction that Dr. Carriere seems to make between "multiple-decrement theory" and the "theory of competing risks." Actuaries typically use the former terminology, while biostatisticians use the latter, and the applications of each group are somewhat different. However, they are just different words for exactly the same mathematical model.

(AUTHOR'S REVIEW OF DISCUSSIONS)

JACQUES F. CARRIERE:

I extend my thanks to Mr. Craighead and Drs. Daniel, Portnoy and Promislow for their discussions. Let me respond to each discussant's remarks in succession.

Mr. Craighead poses three questions about the expansion and use of the theory. He asks whether the theory can easily be integrated into specific software packages such as S-Plus or APL. I found it easy to implement the theory with GAUSS, my preferred language, but the calculations can be done in almost any language. My view is that ease of integration depends on the skills of the programmer. He asks whether my process of eliminating decrements can be viewed linearly. Let me emphasize that my method requires that you solve a nonlinear system of differential equations to analyze the effect of eliminating a decrement and so the process is not linear. He asks whether the theory can be expanded to include interest rates. At first, I thought that this would be impossible, but on reflection I found that discounting for interest can be viewed as a decrement, thus allowing it to depend on other decrements. To illustrate this point, consider the net single premium

$$\int_0^\infty \exp\left\{-\int_0^t (\delta_z + \mu_{x+z}^{(\tau)})dz\right\}dt.$$

If we have m decrements, then the force of interest $\delta_z \equiv \mu_{x+z}^{(m+1)}$ can simply be viewed as another force of decrement that may depend on the others.

Dr. Daniel states that the textbook *Actuarial Mathematics* does not assume stochastic independence between the competing causes of decrement. This is true for most of the discussion in Chapter 9, but Bowers et al. implicitly make this assumption in Equation (9.5.1), where the single decrement function ${}_tp_x'^{(j)}$ is defined as

$${}_tp_x'^{(j)} \equiv \exp\left\{-\int_0^t \mu_{x+z}^{(j)}dz\right\}.$$

My article proves that this definition is actually a consequence of the assumption of independence between competing causes of decrement. Dr. Daniel also states that "the same multiple-decrement model can be constructed from infinitely many quite different sets of latent random variables." This statement is equivalent to saying that $S(t_1, \ldots, t_m)$ is not identifiable by simply knowing $S^{(j)}(t)$. This weakness of multiple-decrement models was extensively discussed in the paper. Essentially, the only way to calculate ${}_tp_x'^{(j)}$ is to make some simplifying assumption like independence, which may be unreasonable. This paper shows how to calculate this single-decrement function in the dependent case by using a copula function.

Dr. Portnoy presents a very good pedagogical discussion of the earlier parts of my paper. Dr. Portnoy suggests that the proof to Lemma 5 is rather long for a fairly obvious result. I must concur.

Finally, Dr. Promislow cautions the reader that this is not the first article that models dependence in multiple-decrement theory, because biostatisticians have been doing it for years in the guise of the theory of competing risks. Dr. Promislow would also suggest that this paper is simply an expository piece of well-known results. In my opinion, the role of the copula function for solving the nonlinear system of differential equations that relate the crude rates with the net rates was not well understood. In this paper, the copula function was used to give a representation of the system that allowed us to solve the system for the first time. Moreover, this solution was presented by an actuary, not a biostatistician.