EQUIVALENCE AND NONIDENTIFIABILITY IN COMPETING RISKS: A REVIEW AND CRITIQUE*

by

Regina C. Elandt-Johnson

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1222

April 1979

EQUIVALENCE AND NONIDENTIFIABILITY IN COMPETING RISKS: A REVIEW AND CRITIQUE*

Regina C. Elandt-Johnson

Department of Biostatistics
University of North Carolina
Chapel Hill, N.C. 27514, U.S.A.

SUMMARY

Let $X_1,\ldots,X_k$ denote hypothetical 'times due to fail' of a system which fails from a single cause. Let $X = \min(X_1,\ldots,X_k)$ and I be a random index such that I = i, if $X_i = X$. The pair (X,I) is referred to as the identified minimum. Any joint survival distribution function (SDF) - called a base SDF - determines a class of equivalent models in the sense that each member of the class has the same distribution of the pair (X, I). Each class contains an infinity of equivalent models. Among them is a unique model with independent times 'due to fail' - called the core of the class - with at least one proper marginal SDF. Contrary to the common consensus, it ia argued that in the analysis of mortality data, the marginal SDF's of the core, rather than the marginals of the base distribution have meaningful interpretations and can be estimated from mortality data. Restrictive conditions under which the specific parametric base SDF can be identified from the distribution of (X, I) are discussed and illustrated by examples.

Key Words & Phrases: Identified minimum; Equivalent models; Base and core distributions; Nonidentifiability; Waiting time distribution; Elimination of cause; Onset; Life table.

---

*This work was supported by the National Heart and Lung Institute contract NIH-NHLI-712243, and by NIH research grant number 1 R01 CA17107 from the National Cancer Institute.

1. INTRODUCTION

Consider a system subject to k modes or causes, $C_1,\ldots,C_k$, of failure. Assume that the failure of the system is due to a single cause (series system). We introduce hypothetical 'times due to fail', $X_1,\ldots,X_k$ $(X_i > 0)$, and their joint survival distribution function (briefly, SDF)

$$S_{1\ldots k}(x_1,\ldots,x_k) = \Pr\left\{\bigcap_{i=1}^{k}(X_i > x_i)\right\} .$$

Let $X = \min(X_1,\ldots,X_k)$ denote the actual failure time. It can also be referred to as age or lifetime of the system. Then $F_X(x) = \Pr\{X \le x\}$ is the cumulative failure distribution (CDF) of the system, and $S_X(x) = 1 - F_X(x)$ is its survival distribution function (SDF).

In addition to the time at failure, one can also observe the cause of failure. Let $I = i$ iff $X_i < X_j$ for $j \neq i$ be a random index identifying the actual cause, C_i, of failure (Berman (1963)).

If X alone is observed it is customary to speak about a nonidentified minimum; if the pair (X, I) is observed it is referred to as an identified minimum (Nádas (1971), Basu and Ghosh (1978)).

We introduce some further definitions. For the purpose of this paper, we will understand by a univariate family, $\mathcal{F}_i$, a set of SDF's, $S_i(x;\, \underset{\sim}{\alpha}_i)$ of the same mathematical form, but with different values of the parameter vector, $\underset{\sim}{\alpha}_i$. We write $\mathcal{F}_i = \{\underset{\sim}{\alpha}_i;\ S_i(x;\, \underset{\sim}{\alpha}_i)\}$. Sets of distributions such as negative exponential, Gompertz, lognormal, etc. represent different families of failure distributions.

In a similar way, we define a multivariate family, $\mathcal{F}_{1\ldots k} \equiv \{\underset{\sim}{\alpha};\ S_{1\ldots k}(x_1,\ldots,x_k;\, \underset{\sim}{\alpha})\}$. For example, the bivariate Farlie-Gumbel-Morgenstern (briefly, FGM) SDF, which can be represented in the general form

$$S_{12}(x_1, x_2) = S_1(x_1)S_2(x_2)[1 + \theta F_1(x_1)F_2(x_2)] ,$$

where $F_i(x) = 1 - S_i(x)$ $(i = 1, 2)$ and $|\theta| \leq 1$, is not a family according to our definition. We may call it a type of bivariate distribution. If however, $S_1(x_1)$, $S_2(x_2)$ are specified by parametric forms (usually the same, though this is not essential), then we have a specified FGM - family; for example FGM - exponential - or FGM - Weibull family.

The following problems have recently attracted attention.

(i) Given the joint distribution of the identified minimum (X, I), is it possible to identify the family $\mathcal{F}_{1\ldots k}$, and the families of marginal distributions, $\mathcal{F}_i$, $i = 1, 2, \ldots, k$?

The answer to this question is, in the general case, "no" and is summarized in Theorem 1. Proofs and practical aspects of the results are given in Sections 2 and 3.

(ii) Under what conditions would information on the distribution of (X, I) permit us to identify the specified member of the family $\mathcal{F}_{1\ldots k}$, given that the joint SDF belongs to this family?

Discussion of this problem is mainly given in Section 4. Section 5 gives an 'optimistic' view on the 'pessimistic' problem of nonidentifiability in competing risks.

2. EQUIVALENT MODELS

2.1. Preliminaries

We briefly introduce some definitions and terminology from the theory of competing risks. More details can be found in: Chiang (1961), Gail (1975), Elandt-Johnson (1976a), David and Moeschberger (1978), Birnbaum (1979) among many others.

Consider first a joint survival function, $S_{1\ldots k}(x_1,\ldots,x_k)$, where $X_1,\ldots,X_k$ are not necessarily independent. We introduce the element of gradient hazard rate (Johnson and Kotz (1975))

$$h_i(x) = -\left.\frac{\partial \log S_{1\ldots k}(x_1,\ldots,x_k)}{\partial x_i}\right|_{\{x_j=x\}} . \tag{2.1}$$

This is also called the 'crude' hazard rate (Chiang (1961)). In the actuarial literature this is called the 'crude' force of mortality and denoted by $(a\mu)_x^i$ (e.g., Neill (1977)) or briefly, by $a\mu_{ix}$ (Elandt-Johnson (1976a)).

The marginal SDF of X_i is

$$S_i(x_i) = S_{1\ldots k}(0,\ldots,x_i,\ldots,0),$$

and its hazard rate is

$$\lambda_i(x) = -\frac{d\log S_i(x)}{dx} . \tag{2.2}$$

This is sometimes called the 'net' hazard rate. In the actuarial literature it is denoted by μ_x^i (or by μ_{ix}). Note that in the general case $\lambda_i(x) \neq h_i(x)$. If, however, $X_1,\ldots,X_k$ are mutually independent,

$$\lambda_i(x) = h_i(x) . \tag{2.3}$$

The reverse is not necessarily true (Gail (1975)). A counterexample is given by Hakulinen and Rahiala (1977).

For convenience, we refer to x as age. The lifetime (overall) survival function is

$$\Pr\{\min(X_1,\ldots,X_k) > x\} = S_{1\ldots k}(x,\ldots,x) = S_X(x), \tag{2.4}$$

and the corresponding hazard rate is $h_X(x) = -d\log S_X(x)/dx$, Note that

$$h_X(x) = h_1(x) + h_2(x) + \cdots + h_k(x) \quad . \tag{2,5}$$

The probability of failure before age x from cause C_i - the so called 'crude' probability function - is

$$Q_i(x) = \Pr\{(X \le x) \cap (I = i)\} = \int_0^x h_i(t) S_X(t) dt \quad , \tag{2,6}$$

so that the proportion of failures from cause C_i, among all failures, is $\Pr\{I = i\} = Q_i(\infty) = \pi_i$.

The 'crude' probability of eventually failing after age x ('crude survival' probability) is

$$P_i(x) = \Pr\{(X > x) \cap (I = i)\} = \int_x^\infty h_i(t) S_X(t) dt, \tag{2.7}$$

with $P_i(0) = Q_i(\infty) = \pi_i$. Of course,

$$Q_i(x) + P_i(x) = Q_i(\infty) = P_i(0) \le 1, \tag{2.8}$$

so that (2.7) may be an improper SDF. The corresponding proper SDF is

$$\Pr\{(X > x) \mid (I = i)\} = \Pr(x)/P_i(0).$$

We also notice that

$$\begin{aligned} Q_1(x) + Q_2(x) + \cdots + Q_k(x) &= F_X(x) \quad , \\ P_1(x) + P_1(x) + \cdots + P_k(x) &= S_X(x) \quad , \end{aligned} \tag{2.9}$$

with $F_X(x) + S_X(x) = 1$, so that the lifetime SDF, $S_X(x)$, is always a proper SDF.

Further,

$$S_X(x) = \exp[-\int_0^x h_X(t) dt] = \prod_{i=1}^k \exp[-\int_0^x h_i(t) dt] = \prod_{i=1}^k G_i(x) \quad , \tag{2,10}$$

where

$$G_i(x) = \exp[-\int_0^x h_i(t) dt] \quad , \tag{2,11}$$

2.2. A Class of Equivalent Distributions

For convenience, we confine ourselves to absolutely continuous distributions. Let $X = \min(X_1,\ldots,X_k)$ and $X' = \min(X_1',\ldots,X_k')$ be minima of two sets of random variables and I be the random index defined in the Introduction.

DEFINITION: Two survival models are equivalent if they have the same joint distributions of the identified miminum, that is

$$\Pr\{(X > x) \cap (I = i)\} = \Pr\{(X' > x) \cap (I = i)\} , \qquad (2.12)$$

for all $i = 1,2,\ldots,k$ and all x.

Statement (2.12) is equivalent to the conditions

$$S_{X'}'(x) = S_X(x), \quad \text{and} \quad P_i'(x) = P_i(x) \qquad (2.13)$$

for all $i = 1,2,\ldots,k$ and all x. (Elandt-Johnson (1976b), and Langberg et.al. (1978) - without assumption of absolute continuity.)

We now wish to answer the question (i) posed in the Introduction. We first prove a few lemmas.

LEMMA 1. Consider a joint SDF $S_{1\ldots k}(x_1,\ldots,x_k)$ - call it a "base" SDF. There is an infinity of possible models equivalent to any base model.

PROOF: The 'crude' probability function, $Q_i(x)$ (or $P_i(x)$), cannot be affected by any alteration in the distributions of any of those 'times due to fail', which exceed the minimum. Thus, if any transformation is applied to the non-minimum set of X_j's $\{j: X_j \neq \min(X_1,\ldots,X_k)\}$, such that all of the transformed variables $X_1',\ldots,X_k'$, are still greater than $\min(X_1,\ldots,X_k)$, then the identified minimum (X, I) remains the same for every individual, and so the distributions of (X, I) and (X', I) are identical.

EXAMPLE 1. Consider a two component system with failure times, X_1, X_2, and a joint SDF, $S_{12}(x_1, x_2)$. Use the transformation

$$X_1' = X_1$$

$$X_2' = \begin{cases} X_2 & \text{if } X_2 < X_1 \\ X_1 + c(X_2 - X_1) & \text{if } X_2 \geqslant X_1 \text{, with } c > 0. \end{cases}$$

Clearly, $S_{X'}'(x) = S_X(x)$, and $P_i'(x) = P_i(x)$ $(i = 1,2)$, but $S_{12}'(x_1, x_2)$ is different from $S_{12}(x_1, x_2)$.

(Lemma 1 and Example 1 are due to N.L. Johnson (1979).)

LEMMA 2. To each joint SDF with dependent 'times due to fail' there is always an equivalent model with independent 'times due to fail' with marginal SDF's defined by (2.11). (Tsiatis (1975).)

PROOF. The equivalent model can be constructed as follows.

From any given base distribution $S_{1\ldots k}(x_1,\ldots,x_k)$, we can determine gradient hazard rate, $h_i(x)$. We then introduce independent variables, $X_1^*,\ldots,X_k^*$, with SDF's

$$S_i^*(x_i) = \exp[-\int_0^{x_i} h_i(t)dt] = G_i(x_i) . \tag{2.14}$$

The joint SDF is

$$S_{1\ldots k}^*(x_1,\ldots,x_k) = \prod_{i=1}^{k} G_i(x_i) . \tag{2.15}$$

Let $X^* = \min(X_1^*,\ldots,X_k^*)$. Then

$$S_{X^*}^*(x) = S_{1\ldots k}^*(x,\ldots,x) = \prod_{i=1}^{k} G_i(x) = S_X(x) . \tag{2.16}$$

Since $X_1^*,\ldots,X_k^*$ are independent, the gradient hazard rate, $h_i^*(x)$, is the same as the 'net' hazard rate and is equal to $h_i(x)$. Hence

$$P_i^*(x) = \int_x^\infty h_i^*(t)S_{X^*}^*(t)dt = \int_x^\infty h_i(t)S_X(t)dt = P_i(x) . \tag{2.17}$$

LEMMA 3. The joint SDF of independent failure times, associated with the given distribution of (X, I) is unique.

PROOF. Suppose that there is another joint SDF, $S^{\#}_{1,\ldots k}(x_1,\ldots,x_k)$, such that

$$S^{\#}_{1\ldots k}(x_1,\ldots,x_k) = \prod_{i=1}^{k} S^{\#}_i(x_i) \ ,$$

with the 'net' (and 'crude') hazard rates $h^{\#}_i(x)$. We must then have

$$S^{\#}_{X^{\#}}(x) = \prod_{i=1}^{k} S^{\#}_i(x) = S_X(x) = \prod_{i=1}^{k} G_i(x) \tag{2,17}$$

and

$$P^{\#}_i(x) = \int_x^{\infty} h^{\#}_i(t) S^{\#}_{X^{\#}}(t)dt = \int_x^{\infty} h_i(t) S_X(t)dt = P_i(x) \ . \tag{2.18}$$

Differentiating (2,18) and taking into account (2.17), we obtain

$$h^{\#}_i(s) = h_i(x), \tag{2,19}$$

so that

$$S^{\#}_i(x) = \exp[-\int_0^x h_i(t)dt] = G_i(x) \tag{2.20}$$

for $i = 1,2,\ldots,k$.

We will call the unique joint SDF of independent 'times due to fail', defined in (2.15), the core of the class of equivalent models. We note that

(a) The base SDF determines the core SDF, but not vice-versa.

(b) In the general case, the base and the core SDF's belong to different families. We will show, however, that the are special cases, when this is not true. (Example 3.)

LEMMA 4. For at least one X_i, we have

$$\lim_{x\to\infty} \int_0^x h_i(t)dt = \infty, \quad i = 1,2,\ldots,k. \tag{2,21}$$

PROOF: The lifetime SDF, $S_X(x)$, is a proper distribution, that is

$$S_X(0) = 1 \quad \text{and} \quad \lim_{x\to\infty} S_X(x) = 0.$$

Hence, from (2.10)

$$\lim_{x\to\infty} \prod_{i=1}^{k} G_i(x) = \lim_{x\to\infty} \prod_{i=1}^{k} \exp[-\int_0^x h_i(t)dt] ,$$

implies for at least one i

$$\lim_{x\to\infty} \int_0^x h_i(t)dt = \infty ,$$

that is, at least one $G_i(x)$ is a proper survival distribution.

We summarize the results in the following theorem.

THEOREM 1. Any joint survival function - called a base SDF - defines an infinite class of equivalent models. Each such class contains a unique model - called the core SDF for this class - in which the 'times due to fail' are independent, and at least one marginal SDF is proper.

A version of Theorem 1, replacing the assumption of absolute continuity by the assumption that the conditional distributions, $\Pr\{X > x | I = i\}$, have no common discontinuities, was proved by Miller (1977). Langberg et.al. (1978) generalized Miller's results, showing that necessary and sufficient conditions for equivalence properties are that the set of discontinuities of the distribution of (X, I) are pairwise disjoint on the range of life span.

For illustration of these results, we use the following example.

EXAMPLE 2. The bivariate Gumbel Type A (briefly, GTA) survival function is of the form

$$S_{12}(x_1\ x_2) = S_1(x_1)S_2(x_2)\exp\left\{-\alpha\left[\frac{1}{\log S_1(x_1)} + \frac{1}{\log S_2(x_2)}\right]^{-1}\right\}.$$

Suppose that $S_i(x_i) = \exp[-(R_i/a_i)(1-e^{a_ix_i})]$ $(i = 1,2)$ are Gompertz distributions, so that we consider a GTA - Gompertz family, and use it as a base SDF. The marginal SDF's are then Gompertz distributions. The 'crude' hazard rate for X_1 is

$$h_1(x_1) = R_1e^{a_1x_1}\left\{1 - \alpha\left[\frac{(R_2/a_2)(1-e^{a_2x_2})}{(R_1/a_1)(1-e^{a_1x_1}) + (R_2/a_2)(1-e^{a_2x_2})}\right]^2\right\}.$$

There is a similar expression for $h_2(x)$. Clearly, the $G_i(x)$'s are not Gompertz distributions, so the core SDF belongs to a different family.

EXAMPLE 3. Suppose that in Example 2, $a_1 = a_2 = a$, so that the hazard rates of marginal distributions are in the constant ratio $\lambda_2(t)/\lambda_1(t) = R_2/R_1 = \theta$. Then the crude hazard rates are also proportional. In fact, $h_2(x)/h_1(x) = \theta[1-\alpha(1+\theta)^{-2}]/[1-\alpha\theta^2(1+\theta)^{-2}]$ (Elandt-Johnson (1978)).

In this case, the $G_i(x)$'s are also Gompertz distributions with common parameter a, and $R_1' = R_1[1-\alpha\theta^2(1+\theta)^{-2}]$, $R_2' = \theta R_1[1-\alpha(1+\theta)^{-2}]$ - the marginal SDF's of the core belong to the same family as the marginal SDF's of the base distribution.

2.3. Interpretation of $P_i(x)$ and $G_i(x)$

Imagine a cohort of newborn individuals subject to deaths from k causes. The 'crude' probability function, $P_i(x)$, defined in (2.7) represents the expected proportion of deaths from cause C_i after age x among those alive at age zero, so that

$P_i(0) = Pr\{I = i\} = \pi_i$ is the expected proportion of all deaths from cause C_i in this cohort. Then $P_i(x)/P_i(0)$ is the expected proportion of deaths after age x from cause C_i among all deaths from C_i in the cohort - it is a proper distribution function. The $P_i(x)$'s can easily be estimated from mortality data by constructing a multiple decrement life table.

On the other hand, $G_i(x)$ as defined in (2.11) - which is also a marginal SDF of the core SDF - represents the probability that a newborn individual, exposed to risk of dying from cause C_i alone, and not from any other causes, will still be alive at age x. The value $1 - G_i(\infty) = \phi$ gives the proportion of newborn individuals who are susceptible and exposed to risk of death from cause C_i in this cohort, and $[1 - G_i(x)]/[1 - G_i(\infty)]$ represents a proper 'waiting time distribution' for those individuals. The probability function $G_i(x)$ is also estimable from population mortality data: it corresponds to a single decrement life table associated with a multiple decrement life table, using the basic assumption that the 'crude' and 'net' hazard rates are the same. (For construction see, for example, Neill (1977), Chapter 9.)

We return again to the interpretation of $G_i(x)$, in the next section.

3. ELIMINATION OF A CAUSE: WHY MARGINAL DISTRIBUTIONS?

3.1. What Does "Elimination" of a Cause Mean?

A question which seems to raise a lot of controversy is: what would the survival distribution be if one or more cause(s) were eliminated?

"Elimination" of a cause may have different interpretations, with consequent differences among the hypothesized distributions expected to result after a cause has been eliminated. (Elandt-Johnson (1976a)).

Deaths (unless accidental) are usually associated with an infectious or chronic disease. Development of a chronic disease is usually a complicated process, involving disturbances in biological and biochemical equilibria; a model - far too simplified - for it can be formulated in terms of failures of several components in a specified order. Suppose that failure of one component eliminates cause C_i. This may also alter - reduce or increase - the number of deaths from cause C_j. Even if the joint SDF were known, one cannot easily compute the appropriate survival probabilities after "elimination" of a cause.

For some reasons - in my opinion, not well justified - there is a common consensus that the marginal SDF, $S_{1\ldots k}(0,\ldots,x_i,\ldots,0)$, represents the hypothetical SDF for cause C_i acting alone. This could be true, if the 'times due to die' were real, i.e., each individual were actually endowed at birth with a set of X_i's, and they were independent. If they are not independent, $S_{1\ldots k}(x_1,\ldots,x_k)$, is not identifiable, and so neither are the marginal SDF's. Some theoretical work has been done on constructing bounds for the marginal distributions from the observable distribution of (X, I) (Peterson (1975)), or specifying additional conditions under which certain sets of marginal distributions can be identified from the $G_i(x)$'s.

(Langberg et.al. (1977)).

3.2. Age of Onset Distribution

The irrelevance of seeking for the marginal distribution, as that representing the distribution of X_i alone, is even more apparent, when one event is not lethal as for example, onset of a chronic disease, while the other is death from any cause. One may introduce hypothetical times 'due for onset', X_1, and 'due for death', X_2, and their joint SDF, $S_{12}(x_1, x_2)$. Clearly, mortality cannot be ignored; one will not consider $S_{12}(x_1, 0)$ as a realistic 'onset waiting time distribution'. A distribution of some interest would be the one for the survivors exposed to risk of onset, that is the $G_i(x)$.

Generally, not the marginal of the base SDF, $S_{1...k}(0,...,x_i,...,0)$, but rather the marginal of the core SDF, $G_i(x)$, has more meaning and can be estimated from mortality data.

3.3. Nonparametric Estimation of $P_i(x)$ and $G_i(x)$

(i) From population mortality rates, one can construct a multiple decrement life table to estimate $P_i(x)$, and single decrement life tables associated with it to estimate $G_i(x)$, using the same mortality rates, in a standard manner (e.g. Neill (1977), Chapter 9).

(ii) Follow-up data. For simplicity, suppose that the data are complete: we follow a cohort of n individuals until the last member dies. Denote here by T (not by X - to distinguish it from age) the survival time since $t_0 = 0$, and $T_1,...,T_k$ the hypothetical 'times due to die' from causes $C_1,...,C_k$, respectively. We have

$T = \min(T_1, \ldots, T_k)$.

Let

$$t_1' < t_2' < \ldots < t_r' < \ldots < t_n'$$

be the actual times at death, and

$$\delta_{ij} = \begin{cases} 1 \text{ if individual (j) dies from cause } C_i \\ 0 \text{ otherwise.} \end{cases}$$

Then

$$S_T^o(t) = \frac{n-r}{n} \quad \text{for} \quad t_r' \leq t < t_{r+1} \tag{3.1}$$

is the empirical overall SDF - an estimator of $S_T(t)$.

The empirical SDF

$$P_i^o(t) = \frac{1}{n} \sum_{j=r}^{n} \delta_{ij} \quad \text{for} \quad t_r \leq t < t_{r+1} \tag{3.2}$$

is an estimator of $P_i^o(t)$, with $P_i^o(0) = n_i/n$.

Finally, an estimator of $G_i(t)$ is

$$G_i^o(t) = \prod_{j=1}^{r} \left(\frac{n-j}{n-j+1}\right)^{\delta_{ij}} \quad \text{for} \quad t_r \leq t < t_{r+1} \text{ ,} \tag{3.3}$$

which is the product-limit (Kaplan-Meier) estimator, with deaths from other causes treated as withdrawals. Note that if the T_i's were independent, the empirical SDF in (3.3) would coincide with the empirical marginal (Hoel (1972), Peterson (1977) - without assumption of absolute continuity).

Appropriate adjustment for incomplete (censored) data is straightforward.

4. IDENTIFIABILITY OF THE PARAMETERS OF JOINT DISTRIBUTION FROM THE IDENTIFIED MINIMUM

Another aspect of nonidentifiability of distributions, referred to in question (ii) in the Introduction is of a different nature, namely:

Given the parametric form of a base joint SDF, can one identify this SDF uniquely (i.e., identify its parameters) from the distribution of (X, I)?

It has been shown, for example, that the bivariate normal distribution can be uniquely identified given the distribution of (X, I). (Nádas (1971), Anderson and Ghurye (1977), Basu and Ghosh (1978).) For $k > 2$, the problem becomes more difficult. Anderson and Ghurye (1977) showed that if $X_1, \ldots, X_k$ are independent, a sufficient condition for unique identifiability is that for any possible pair of densities which belong to the same family, one tends to zero faster than the other, as $x \to \infty$ (Anderson and Ghurye (1977), p. 338). This condition is not, however, necessary as has been shown by the same authors by means of a counterexample.

We also notice that the Gumbel Type A distribution with Gompertz marginal SDF's and proportional hazard rates, discussed in Example 3 cannot be uniquely defined from the distribution of (X, I). In fact, there are two equations relating the observable parameters R_1' and R_2' with the unobservable parameters R_i, α and θ. It has been shown [Elandt-Johnson (1978)] that for $\alpha = 0$ (i.e., when X_1 and X_2 are independent), the (Gompertz) $G_i(x)$'s are the upper bounds for the (Gompertz) marginal $S_i(x)$'s.

Basu and Ghosh (1978) has shown nonidentifiability of Gumbel Type A when the marginals are exponential. (Clearly, the hazard rates are proportional.)

It appears from Elandt-Johnson (1978) (though no formal proof is given) that Gumbel Type A distributions with proportional hazard rates are not identifiable from the distribution of (X, I), whatever the family of the marginals; moreover the upper bounds of marginals are obtained, when X_1 and X_2 are independent.

More examples showing identifiability of the parameters or its lack can be found in Basu and Ghosh (1978).

Using the sufficient criterion given by Anderson and Ghurye (1977), one can construct still new examples. For a general answer, when $X_1, \ldots, X_n$ are not independent, the necessary conditions have yet to be investigated.

5. DISCUSSION

The problems of equivalence and nonidentifiability in competing risks have two broad aspects: (i) mathematical and (ii) practical.

(i) It has been shown that if only the distribution of the identified miminum (X, I) is available, one can identify neither joint nor marginal distributions. If additionally, however, the form of the joint distribution is given, it is possible - for certain families and under certain conditions-to identify also the parameter values. Independence is one such condition. In case of nonidentifiability of the parameter values, one can establish bounds for the marginal distributions.

(ii) The practical aspects and relevance of these results seem sometimes to have been misinterpreted.

The base SDF is usually a joint distribution of well recognized form, or a new one constructed from some physical background in biology, engineering, etc. There is a strong though, in my opinion, unjustified belief, that the marginal SDF's of the base SDF are the 'key' distributions. Since these cannot be identified, we have the unsolved problem of assessing the SDF when one or more causes are "eliminated". Without further biological background (unlikely to be available) this problem cannot be satisfactorily solved, anyway.

But something about the 'waiting time distribution' can be done, without this background and without assumption of independence of hypothetical 'times due to die'. The thesis of the present paper is that the 'key' distributions for cause C_i are: the 'crude' probability function

$$P_i(x) = \int_x^\infty h_i(t)S_X(t)dt \ , \tag{5.1}$$

giving the expected proportion of deaths from cause C_i after age x, in a cohort of newborn individuals, and the 'waiting time distribution'

$$G_i(x) = \exp[-\int_0^x h_i(t)dt] \ , \tag{5.2}$$

giving the probability of eventually dying after age x for those who are exposed to risk of dying from C_i, and not from other causes, in a given population.

Since $S_X(x)$ and $h_i(x)$ are estimable from mortality data, both $P_i(x)$ and $G_i(x)$ are estimable, too. It so happens that the $G_i(x)$'s coincide with the marginal SDF's of the core distribution.

If one wishes to fit a known (and identifiable) parametric form of a specified base joint SDF, one can construct a likelihood, using either distribution (5.1) or (5.2) - they should give the same estimates of the parameters, since they are based on the same mortality data. In consequence, one can then estimate also the marginal SDF's of the base distribution. It should be realized, however, that interpretation of marginal SDF's as the survival functions when other causes are "eliminated" might still be misleading and incorrect. The multiple and single decrement life table distributions, which are analogues of $S_X(x)$, $P_i(x)$, and $G_i(x)$, respectively, are always estimable. We emphasize again that $G_i(x)$ is not the hypothetical SDF for cause C_i when all other causes were eliminated; it is the estimable SDF for the members in a cohort of newborn individuals actually exposed to risk of death just from cause C_i, without taking into account other causes; the only available and meaningful hazard rate is the estimable hazard rate, $h_i(x)$.

REFERENCES

Anderson, T.W. and Ghurye, S.G. (1977). Identification of parameters by the distribution of a maximum random variable. J. Roy. Statist. Soc. Ser. B, 39, 337-342.

Basu, A.P. and Ghosh, J.K. (1978). Identifiability of the multinormal and other distributions under competing risk model. J. Mult. Anal., 8, 413-429.

Berman, S.M. (1963). Note on extreme values, competing risks and semi-Markov processes. Ann. Math. Statist., 34, 1104-1106.

Birnbaum, Z.W. (1979). On the mathematics of competing risks. DHEW Publication No. (PHS) 79-1351. U.S. Department of Health, Education, and Welfare, pp. 1-58.

Chiang, C.L. (1961). On the probability of death from specific causes in the presence of competing risks. Fourth Berkeley Symp. Vol. 4, 169-180.

David, H.A. and Moeschberger, M.L. (1978). The Theory of Competing Risks. Griffin's Statistical Monographs and Courses No. 39, MacMillan Publishing Co., Inc., New York.

Elandt-Johnson, R.C. (1976a). Conditional failure time distributions under competing risk theory with dependent failure times and proportional hazard rates. Scand. Actu. J., 37-51.

Elandt-Johnson, R.C. (1976b). Equivalent models in the theory of competing risks. (unpublished manuscript).

Elandt-Johnson, R.C. (1978). Some properties of bivariate Gumbel Type A distributions with proportional hazard rates. J. Mult. Anal., 8, 244-254.

Gail, M. (1975). A review and critique of some models used in competing risk analysis. Biometrics, 31, 209-222.

Hakulinen, T. and Rahiala, M. (1977). An example of the risk dependence and additivity of intensities in the theory of competing risks. Biometrics, 33, 557-559.

Hoel, D.G. (1972). A representation of mortality data by competing risks. Biometrics, 28, 475-488.

Johnson, N.L. (1979). Personal communication.

Johnson, N.L. and Kotz, S. (1975). A vector-values multivariate hazard rate. J. Mult. Anal., 5, 53-66.

Langberg, N., Proschan, F., and Quinzi, A.J. (1977). Estimating dependent life lengths, with applications to the theory of competing risks. Statistics Report M438, Florida State University, Department of Statistics, Tallahassee, pp. 1-30.

Langberg, N., Proschan, F., and Quinzi, A.J. (1978). Converting dependent models into independent ones, preserving essential features. Ann. Prob., 6, 174-181.

Miller, D.R. (1977). A note on independence of multivariate lifetimes in competing risks. Ann. Statist., 5, 576-579.

Nádas, A. (1971). The distribution of the identified minimum of a normal pair determines the distribution of the pair. Technometrics, 13, 201-202.

Neill, A. (1977). Life Contingencies. Heineman, London. Chapter 9.

Peterson, A.V. (1975). Nonparametric estimation in the competing risk. Technical Report, No. 13, Stanford University.

Peterson, A.V. (1977). Expressing the Kaplan-Meier estimator a function of empirical sub-survival functions. J. Amer. Statist. Assoc., 72, 854-858.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. Proc. Nat. Acad. Ser., Wash., 72, 20-22.