Testing dependence between the failure time and failure modes: An application of enlarged filtration

Dario Gasbarra[a,*], **S. B. Kulathinal**[b], **Isha Dewan**[c], **Aulikki Nissinen**[b]

[a] *Rolf Nevanlinna Institute, University of Helsinki, P.L. 4, 00014 Helsinki, Finland,*

[*] *E-mail: dag@rni.helsinki.fi, Tel.: +358 9 191 22765, Fax: +358 9 191 22779.*

[b] *Department of Epidemiology and Health Promotion, National Public Health Institute, Mannerheimintie 166, 00300 Helsinki, Finland*

[c] *Indian Statistical Institute, New Delhi, India*

Abstract

The model in which competing risks are assumed to be independent does not provide any information for the assessment of competing failure modes, if the failure mechanisms underlying these modes are coupled. Certain models for dependent competing risks have been proposed in the literature. These can be distinguished on the basis of the monotonicity of the conditional probability of a particular failure mode given that the failure time exceeds a fixed time. There is an interesting link between the monotonicity of such conditional probability, the dependence between the failure time and the failure mode, and the crude hazard rates. In this paper, we propose tests for testing the dependence between the failure time and the failure mode using the crude hazards and using the conditional probabilities mentioned above. We establish the equivalence between the two approaches and provide an optimal weight function. The tests are applied to simulated data and to mortality follow-up data.

Keywords : Competing risks; Conditional probability; Crude hazards; Enlarged filtration; Kolmogorv-Smirnov type tests; Martingale

1 Introduction

In the follow-up study of mortality, it is observed that the contribution of the causes of death due to common cause which includes cardiovascular diseases, cancer and accident and suicide decreases with age while the contribution of other causes increases. Hence, in such situations it is of interest to compare the probabilities of dying due to a common cause and due to other cause given that a person has survived upto a certain age. It is also of interest to test whether such conditional probabilities increase or decrease with age. Dewan *et al.* (2004) give several examples where the conditional probabilities are of interest.

In this paper, we study the relationship between the crude hazards and the conditional probabilities in the case of two competing risks. We develop test procedures using the crude hazards and the Kolmogorv-Smirnov type test for testing independence of the failure mode and the failure time. For a specific choice of local alternative, the two tests are shown to be equally efficient and an optimal kernel is given. A test based on crude hazards can then be easily extended to include more than two risks and also censoring. The methods are illustrated by simulated data and also by a real data on mortality follow-up conducted in Finland.

The competing risks data consist of the failure time, T and an indicator of failure mode, δ which can have one of the values $\{0, 1\}$.

Define the joint distribution of (T, δ) through the subsurvival functions

$$S_i(t) = P[T \geq t, \delta = i], i = 0, 1,$$

leading to the overall survival function of the failure time

$$S(t) = P[T \geq t] = S_0(t) + S_1(t).$$

Let $F_i(t)$ and $F(t)$ denote the corresponding subdistribution and distribution functions. Throughout the paper, we assume that $F_i(.)$ and $F(.)$ are continuous and $f_i(.)$ and $f(.)$

denote the corresponding subdensity and density functions. Also, define the conditional probability of failure due to the first risk given that there is no failure upto time t as

$$\Phi_1(t) = P[\delta = 1 \mid T \geq t]. \tag{1}$$

Equivalently, we can define $\Phi_0(t) = P[\delta = 0 \mid T \geq t] = 1 - P[\delta = 1 \mid T \geq t]$. It is interesting to note that $\Phi_1(t) = P[\delta = 1] = \phi, \forall\, t > 0$ is equivalent to independence of T and δ. In general dependence set-up, the analysis of competing risks data is carried out using the subsurvival functions $S_i(t), i = 0, 1$, and hence if T and δ are independent then $S_i(t) = S(t)P[\delta = i]$. Thus, the hypothesis of equality of incidence functions is equivalent to testing whether $P[\delta = 1] = P[\delta = 0] = 1/2$.

Let $\Lambda_i(t)$ and $\tilde{A}_i(t)$ be the cumulative cause-specific and cumulative crude hazards for failure mode i, and are given by

$$\Lambda_i(t) = \int_0^t \frac{dF_i(u)}{S(u)}, \quad \tilde{A}_i(t) = \int_0^t \frac{dF_i(u)}{S_i(u)}.$$

Here, we consider the testing problems $H_0 : \Phi_1(t) = \phi$ against $H_1 : \Phi_1(t)$ is not constant, and $H_2 : \Phi_1(t)$ is increasing in t which is the same as in Dewan *et al.* (2004). Their test U_3 is shown to be asymptotically equivalent to a test proposed here, for a special choice of the weight function. It is interesting to note that the null hypothesis in terms of cause-specific hazards is

$$\frac{\Phi_1(t)}{\Phi_0(t)} = \frac{P(\delta = 1|T \geq t)}{P(\delta = 0|T \geq t)} = \frac{d\Lambda_1}{d\Lambda_0}(t) = \theta = \frac{\phi}{1 - \phi}$$

and is equivalent to testing $a_1(t) = a_0(t) = a(t)$, where $a_i(t) = d\tilde{A}_i(t)/dt, i = 0, 1$ are the crude hazards. The alternative hypothesis that $\Phi_1(t)$ is increasing in t is equivalent to $a_1(t) \leq a_0(t)$.

In section 2, we propose a test based on crude hazards and a weighted Kolmogorov-Smirnov type of test for testing the above hypotheses. We also prove the equivalence between the optimal tests obtained in these two classes of tests. In section 3, simulated data and mortality follow-up data are used to illustrate the proposed tests. We also compare the optimal weight function and the weight function suggested by Harrington and Fleming in case of bivariate exponential distribution, which can be used in the general situation.

2 Test of significance

Let $(T_j, \delta_j),\ j = 1, 2, ..., n$ be the competing risks data obtained from n independent and identical copies of the system. Define the counting processes

$$\begin{aligned} N_i(t) &= \sum_{j=1}^{n} I[T_j \leq t, \delta_j = i], i = 0, 1, \\ N_.(t) &= N_0(t) + N_1(t), Y_.(t) = \sum_{j=1}^{n} I[T_j \geq t]. \end{aligned}$$

Note that $N_i(t)$ counts the number of failures due to competing risk i by time t and $Y_.(t)$ is the number of units at risk just prior to time t. The natural estimates of the subsurvival functions are given by their empirical counterparts

$$\begin{aligned} \hat{F}_{in}(t) &= \frac{N_i(t-)}{n},\ \ \hat{S}_{in}(t) = \frac{n_i}{n} - \hat{F}_{in}(t), \\ \hat{S}_n(t) &= \frac{Y_.(t)}{n},\ \ \hat{\phi}_n = \frac{n_1}{n}, \end{aligned}$$

where $n_1 = \sum_{j=1}^{n} \delta_j$ and $n_0 = n - n_1$.

2.1 Test based on crude hazards

Let $\mathcal{F}_t^{N,Y}$ be the filtration generated by (N_0, N_1, Y). Consider the enlarged filtration $\mathcal{G}_t = \mathcal{F}_t^{N,Y} \vee \sigma(N_1(\infty))$. It can be shown that, for $i = 0, 1$,

$$\tilde{M}_i(t) = N_i(t) - \int_0^t \frac{N_i(\infty) - N_i(s-)}{P(\delta = i | T \geq s)} d\Lambda_i(s)$$

is a $\{\mathcal{G}_t\}$-martingale with predictable variation process

$$\langle \tilde{M}_i \rangle_t = \int_0^t \frac{N_i(\infty) - N_i(s-)}{P(\delta = i | T \geq s)} d\Lambda_i(s)$$

$$= \int_0^t (N_i(\infty) - N_i(s-)) d\tilde{A}_i(s). \quad (2)$$

It is interesting to note that the conditional probability of interest, $P(\delta = i|T \geq t)$, appears in the compensator.

We can split the group of n individuals into a group of $N_1(\infty)$ individuals, those which will fail from cause 1 and a group of $n - N_1(\infty)$, those who will fail from cause 0. So, we have two independent counting processes

$$N_i(t) = \int_0^t \frac{N_i(\infty) - N_i(s-)}{P(\delta = i|T \geq s)} d\Lambda_i(s) + \tilde{M}_i(t) = \int_0^t Y_i(s) d\tilde{A}_i(s) + \tilde{M}_i(t), \quad (3)$$

where $i = 0, 1$ and $Y_i(t) = N_i(\infty) - N_i(t-)$.

Testing the proportional hazards hypothesis is equivalent to testing whether the intensities of the two counting processes N_i are identical. This testing problem is the same as that discussed by Andersen *et al.* (1993) on pages 345-348.

For a weight function $K_n(t)$ such that it is non-zero whenever the risk sets corresponding to the two groups are non-empty, consider

$$V_n = \int_0^\tau K_n(s) \left(\frac{dN_1(s)}{Y_1(s)} - \frac{dN_0(s)}{Y_0(s)} \right)$$

$$= \int_0^\tau K_n(s) \left(\frac{d\tilde{M}_1(s)}{Y_1(s)} - \frac{d\tilde{M}_0(s)}{Y_0(s)} \right)$$

$$+ \int_0^\tau K_n(s) \left(d\tilde{A}_1(s) - d\tilde{A}_0(s) \right) \quad (4)$$

where $K_n(t)$ is a $\{\mathcal{G}_t\}$-predictable weight process which must be chosen in some efficient way and second equality follows due to (3). The assumption on the weight function is natural since the estimates of the crude hazards need to be compared in the environment when there are subjects acting in both environments. Note that the second term is zero under the null hypothesis. A test based on V_n can be used to test H_0 against H_1 and H_2 since large values of V_n (either negative or positive) support H_1 while large negative values support H_2.

Theorem 2.1 *Assume that $n^{-1}Y_i(t)$ converges uniformly in probability to a deterministic function $y_i(t)$ for $i = 0, 1$. Further assume that $n^{-1}K_n(t)$ converges to $k(t)$ uniformly in probability with $k(.)$ bounded on $[0, \tau]$. Under the null hypothesis, $n^{-1/2}V_n$ converges in distribution to $N(0, \sigma^2)$ where*

$$\sigma^2 = \int_0^\tau k^2(t)(\frac{1}{y_1(t)} + \frac{1}{y_0(t)})d\tilde{A}(t)$$

Proof: The result follows from (2), (3) and (4), and

$$< n^{-1/2}V_n, n^{-1/2}V_n >= \int_0^\tau n^{-1}K_n^2(t)(\frac{d\tilde{A}(t)}{Y_1(t)} + \frac{d\tilde{A}(t)}{Y_0(t)}).$$

The consistent estimator of the variance is

$$\begin{aligned} \hat{\sigma}^2 &= \int_0^\tau K_n^2(t)\left(\frac{1}{Y_1(t)Y.(t)} + \frac{1}{Y_0(t)Y.(t)}\right)dN.(t) \\ &= \int_0^\tau K_n^2(t)\{Y_1(t)Y_0(t)\}^{-1}dN.(t). \end{aligned} \quad (5)$$

Consider the sequence of local alternatives $\{P^{(n)}(\theta)\}$ for the crude hazards of the form

$$a_i^{(n)}(t) = a(t)(1 + \varepsilon_n\gamma(t)\theta_i),\ i = 0, 1 \quad (6)$$

where $\theta = (\theta_0, \theta_1) \in R^2$ is a local parameter and $\epsilon_n = O(n^{-1/2})$. Under these local alternatives, the asymptotic mean and the variance are

$$\begin{aligned} \mu &= (\theta_1 - \theta_0)\int_0^\tau k(t)a(t)\gamma(t)dt, \quad (7) \\ \sigma^2 &= \{\phi(1-\phi)\}^{-1}\int_0^\tau k^2(t)a(t)S^{-1}(t)dt. \end{aligned}$$

The noncentrality parameter is then given by

$$(\theta_1-\theta_0)^2\phi(1-\phi)(\int_0^{\tau} k(t)a(t)\gamma(t)dt)^2(\int_0^{\tau} k^2(t)a(t)S^{-1}(t)dt)^{-1}.$$

It can be easily seen that the kernel which maximizes this noncentrality parameter is $k(t)$ proportional to $\gamma(t)S(t)$ and hence this choice of kernel gives the most efficient test (see Andersen *et al.*, 1993 for details). The maximum value of the noncentrality parameter is

$$(\theta_1-\theta_0)^2\phi(1-\phi)\int_0^{\tau}\gamma^2(t)S(t)a(t)dt.$$

The above derivation is applicable to a more general sequence of local alternatives

$$a_i^{(n)}(t)=a(t)(1+\varepsilon_n\gamma_i(t)),\ i=0,1. \tag{8}$$

where $\gamma(t)=\gamma_1(t)-\gamma_0(t)$.

An alternative weight function which can be used here is the weight function introduced by Harrington and Fleming (1982) and is given by $[1-\hat{F}_n(t)]^{\rho}$, where ρ is a fixed constant between 0 and 1 and $\hat{F}_n(t)$ is an estimate of the overall incidence function.

2.2 Kolmogorv-Smirnov type test

The hypothesis $\Phi_1(t)\uparrow t$ is equivalent to $\Phi(t_1)\le\Phi(t_2)$, whenever $t_1\le t_2$. That is

$$\begin{aligned} S_1(t_1)/S(t_1) &\le S_1(t_2)/S(t_2) \\ S_1(t_1)S(t_2) &\le S_1(t_2)S(t_1). \end{aligned}$$

This gives $\Psi(t_1,t_2)=S_1(t_2)S(t_1)-S_1(t_1)S(t_2)=S_1(t_2)S_0(t_1)-S_1(t_1)S_0(t_2)\ge 0, t_1\le t_2$ with strict inequality for some (t_1,t_2). Let $\hat{\Psi}_n(t_1,t_2)$ be obtained by replacing the functions by their empirical counter parts

$$\begin{aligned} \hat{\Psi}_n(t_1,t_2) &= \hat{S}_{1n}(t_2)\hat{S}_n(t_1)-\hat{S}_{1n}(t_1)\hat{S}_n(t_2) \\ &= \hat{F}_{1n}(\infty)(\hat{F}_n(t_2)-\hat{F}_n(t_1))+\hat{F}_{1n}(t_1)(1-\hat{F}_n(t_2))-\hat{F}_{1n}(t_2)(1-\hat{F}_n(t_1)). \end{aligned}$$

A Kolmogorov-Smirnov type of test to test H_0 against H_1 can be defined as $\sqrt{n}\hat{D}_n = \sqrt{n}\sup_{u \le v} \mid \hat{\Psi}_n(u,v) \mid$ and large values of the test statistic support H_1. One sided test can be used to test H_0 against H_2, $\sqrt{n}\hat{D}_{1n} = \sqrt{n}\sup_{u \le v}(\hat{\Psi}_n(u,v))$ and large positive values support H_2. Similarly, a hypothesis that $\Phi_1(t)$ is decreasing can be tested and large negative values support the hypothesis.

The following theorem is proved in the Appendix I by the functional delta method.

Theorem 2.2 *As n tends to ∞, $\sqrt{n}(\hat{\Psi}_n(u,v) - \Psi(u,v))$ converges to a zero-mean Gaussian random field $Z(u,v)$ with covariance structure*

$$
\begin{aligned}
cov(Z(u_1,v_1), Z(u_2,v_2)) &= 0 \; if \, u_1 \le v_1 \le u_2 \le v_2 \; or \, u_2 \le v_2 \le u_1 \le v_1 \\
&= \phi(1-\phi)(1-F(max(v_1,v_2)) \\
&\quad (1-F(min(u_1,u_2))(F(min(v_1,v_2)-F(max(u_1,u_2)) \\
&\qquad otherwise
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
var(Z(u,v)) &= \phi(1-\phi)(1-F(v))(1-F(u))(F(v)-F(u)) \\
&= \phi(1-\phi)(1-t)(1-s)(t-s)
\end{aligned} \tag{10}
$$

where $F(u) = s$ and $F(v) = t$, and hence $0 \le s \le t \le 1$.

The above theorem can be used to test the hypothesis of interest but this point is not elaborated here since our interest in the distribution of $Z(u,v)$ is for defining a class of tests in the next section.

2.3 A class of weighted Kolmogorov-Smirnov type of tests

For some weight function $K(u,v)$, we consider

$$
\begin{aligned}
\Delta &= \iint_{0<u\le v<\infty} K(u,v)(S_1(v)S_0(u) - S_1(u)S_0(v))dudv \\
&= \iint_{u \le v} K(u,v)\Psi(u,v)dudv.
\end{aligned} \tag{11}
$$

A weighted Kolmogorov-Smirnov type test statistic for testing H_0 against H_1 and H_2 is defined as

$$\begin{aligned}\hat{\Delta}_n &= \iint\limits_{0<u\le v<\infty} K_n(u,v)(\hat{S}_{1n}(v)\hat{S}_{0n}(u) - \hat{S}_{1n}(u)\hat{S}_{0n}(v))dudv \\ &= \iint\limits_{u\le v} K_n(u,v)\hat{\Psi}_n(u,v)dudv, \qquad (12)\end{aligned}$$

where we assume that

$$\sqrt{n}\iint\limits_{u\le v}(K_n(u,v) - K(u,v))\Psi(u,v)dudv \xrightarrow{\mathbf{P}} 0 \text{ as } n\to\infty.$$

The asymptotic distribution can be obtained by using the covariance structure (9) and (10) of the Kolmogorov-Smirnov type test.

Theorem 2.3 *As n tends to ∞, $\sqrt{n}(\hat{\Delta}_n - \Delta)$ converges in distribution to a Gaussian random variable with mean zero and variance*

$$\sigma^2 = \iint\limits_{u\le v}\iint\limits_{u'\le v'} K(u,v)K(u',v')cov(Z(u,v),Z(u',v'))dudvdu'dv'.$$

The U-statistic, U_3 in Dewan *et al.* (2004) can be obtained from (12) by selecting $K_n(u,v)$ such that the limit $K(u,v)dudv = dF_1(u)dF_1(v)$.

To check the efficiency of this test and also to compare it with a simple test based on the crude hazards, V_n, we consider the same local alternatives $\{P^{(n)}(\theta)\}$ as in (6).

We define

$$\begin{aligned}A(t) &= \int_0^t a(s)ds,\ \Gamma(t) = A^{-1}(t)\int_0^t a(s)\gamma(s)ds, \\ S(t) &= \exp(-A(t)),\ V(t) = \exp(-A(t)\Gamma(t)).\end{aligned}$$

Then the corresponding cumulative hazards are given by

$$A_i^{(n)}(t) = A(t)(1+\varepsilon_n\Gamma(t)\theta_i),\ i=0,1 \qquad (13)$$

and the sequence of subsurvival functions are

$$S_1^{(n)}(t) = \phi S(t)V(t)^{\varepsilon_n\theta_1},\ S_0^{(n)}(t) = (1-\phi)S(t)V(t)^{\varepsilon_n\theta_0},\ S^{(n)}(t) = S_1^{(n)}(t) + S_0^{(n)}(t).$$

To obtain the asymptotic mean of $\sqrt{n}\hat{\Delta}_n$, μ under the local alternatives, consider

$$\sqrt{n}\iint_{u\leq v} K(u,v)(S_1^{(n)}(v)S_0^{(n)}(u) - S_1^{(n)}(u)S_0^{(n)}(v))dudv =$$
$$\sqrt{n}\phi(1-\phi)\iint_{u\leq v} K(u,v)S(v)S(u)\Big(V(v)^{\theta_1\varepsilon_n}V(u)^{\theta_0\varepsilon_n} - V(v)^{\theta_0\varepsilon_n}V(u)^{\theta_1\varepsilon_n}\Big)dudv.$$

As n tends to ∞, the above expression goes to the limit

$$\begin{aligned}\mu &= \phi(1-\phi)\iint_{u\leq v} K(u,v)S(v)S(u)(\log V(v) - \log V(u))(\theta_1-\theta_0)dudv \\ &= \phi(1-\phi)(\theta_0-\theta_1)\iint_{u\leq v} K(u,v)S(v)S(u)(A(v)\Gamma(v) - A(u)\Gamma(u))dudv. \qquad (14)\end{aligned}$$

To obtain the noncentrality parameter of the limiting test, we must square μ and divide by

$$\| K \|_{\mathcal{H}}^2 = \mathrm{Var}\left(\iint_{u\leq v} K(u,v)Z(u,v)dudv\right) =$$
$$\iint_{u\leq v}\iint_{u'\leq v'} K(u,v)K(u',v')\ \mathrm{Cov}\ (Z(u,v),Z(u',v'))du'dv'dudv$$

where $\|\cdot\|_{\mathcal{H}}$ is the norm in the corresponding reproducing kernel Hilbert space.

A kernel $K(u,v)$ is efficient under the sequence of local alternatives (6) or equivalently (13) when it maximizes

$$\| K \|_{\mathcal{H}}^{-1}\iint_{u\leq v} K(u,v)S(v)S(u)(A(v)\Gamma(v) - A(u)\Gamma(u))dudv. \qquad (15)$$

If we denote $L(u,v) = S(u)S(v)(A(v)\Gamma(v) - A(u)\Gamma(u))$ and for a generic function $G(u,v)$ we define the convolution operator

$$(\mathcal{R}G)(u,v) = \iint_{u'\leq v'} G(u',v')\ \mathrm{Cov}\ (Z(u,v),Z(u',v'))du'dv',$$

then we can rewrite expression (15) as the square root of

$$\frac{(K, \mathcal{R}^{-1}L)^2_{\mathcal{H}}}{(K, K)_{\mathcal{H}}}.$$

It is clear that this is maximal when $K(u, v) = (\mathcal{R}^{-1}L)(u, v)$.

In other words, the asymptotically efficient kernel $K(u, v)$ satisfies

$$(16) \qquad (\mathcal{R}K)(u, v) = \iint_{u' \leq v'} K(u', v') \text{ Cov } (Z(u, v), Z(u', v'))du'dv' = L(u, v).$$

It is shown in Appendix II that an optimal weight function which maximizes the noncentrality parameter is of the form

$$(17) \qquad K(u, v) = k(u)\dot{\delta}_u(v) = \gamma(u)S^{-1}(u)\dot{\delta}_u(v)$$

where $\dot{\delta}_u$ is the derivative of the delta function in the sense of distributions for a smooth function f,

$$\int_{-\infty}^{\infty} \dot{\delta}_u(v)f(v)dv = -f'(u)$$

where $f'(u)$ is the derivative of $f(u)$ with respect to u. Note that (17) satisfies (16). This kernel can be approximated by a sequence of smooth kernels, and for such sequences the weighted Kolmogorov-Smirnov test (12) approximates the asymptotically efficient test (2) based on crude hazards.

3 Illustrations

3.1 Simulation study

Consider a bivariate exponential distribution with the density function

$$f(x, y) = \lambda_1\lambda_2 exp(-\lambda_1 x - \lambda_2 y)[1 + \alpha(2exp(-\lambda_1 x) - 1)(2exp(-\lambda_2 y) - 1)]$$

and the survival function

$$S(x,y) = exp(-\lambda_1 x - \lambda_2 y)[1+\alpha(1-exp(-\lambda_1 x))(1-exp(-\lambda_2 y))]$$

The survival function of $T = \min(X,Y)$ is

$$S(t) = S(t,t) = exp(-\lambda_1 t - \lambda_2 t)[1+\alpha(1-exp(-\lambda_1 t))(1-exp(-\lambda_2 t))].$$

It is clear that, for $\alpha = 0$, $S(t) = S(t,t) = exp(-\lambda_1 t - \lambda_2 t)$ and that corresponds to the independence of X and Y.

We fix λ_1 and λ_2 such that $\lambda_1 \neq \lambda_2$ and vary α. Consider the crude hazards

$$a_1(t) = a_1(t,\alpha) = a(t,\lambda_1,\lambda_2,\alpha) = \frac{d\tilde{A}_1(t)}{dt}$$
$$= \frac{\Big(1+\alpha(1-exp(-\lambda_1 t))(1-exp(-\lambda_2 t))\Big)\Big(1-\frac{\alpha e^{-\lambda_1 t}(1-e^{-\lambda_2 t})}{1+\alpha(1-e^{-\lambda_1 t})(1-e^{-\lambda_2 t})}\Big)}{\Big\{\frac{1+\alpha(1+e^{-(\lambda_1+\lambda_2)t})}{\lambda_1+\lambda_2} - \frac{2\alpha e^{-\lambda_1 t}}{2\lambda_1+\lambda_2} - \frac{\alpha e^{-\lambda_2 t}}{\lambda_1+2\lambda_2}\Big\}}$$

and $a_0(t) = a_0(t,\alpha) = a(t,\lambda_2,\lambda_1,\alpha)$ defined analogously by interchanging the role of λ_1 and λ_2. When $\alpha = 0$, $a(t,\lambda_1,\lambda_2,0) = a(t,\lambda_2,\lambda_1,0) = \lambda_1+\lambda_2$ and a is continous in its arguments.

The sequence of local alternatives obtained by expanding the crude hazards $a(t,\lambda_1,\lambda_2,\alpha_n)$ and $a(t,\lambda_2,\lambda_1,\alpha_n)$ around the point $\alpha = 0$ is

$$a_1^{(n)}(t) = (\lambda_1+\lambda_2)(1+\alpha_n\gamma_1(t)), \quad a_0^{(n)}(t) = (\lambda_1+\lambda_2)(1+\alpha_n\gamma_2(t))$$

where $\alpha_n = cn^{-1/2}$ and c is a constant such that $-1 \le \alpha_n \le 1$ and

$$\begin{aligned}
\gamma_1(t) &= \frac{\partial}{\partial\alpha} a(t,\lambda_1,\lambda_2,\alpha)_{\alpha=0} \\
&= e^{-(\lambda_1+\lambda_2)t} - \frac{2\lambda_1}{2\lambda_1+\lambda_2}e^{-\lambda_1 t} - \frac{\lambda_2}{\lambda_1+2\lambda_2}e^{-\lambda_2 t} \\
\gamma_2(t) &= \frac{\partial}{\partial\alpha} a(t,\lambda_2,\lambda_1,\alpha)_{\alpha=0} \\
&= e^{-(\lambda_1+\lambda_2)t} - \frac{2\lambda_2}{2\lambda_2+\lambda_1}e^{-\lambda_2 t} - \frac{\lambda_1}{\lambda_2+2\lambda_1}e^{-\lambda_1 t}
\end{aligned}$$

Let

$$\gamma(t) = \gamma_1(t) - \gamma_2(t) = \frac{\lambda_2}{\lambda_1 + 2\lambda_2} e^{-\lambda_2 t} - \frac{\lambda_1}{2\lambda_1 + \lambda_2} e^{-\lambda_1 t}.$$

The optimal kernel for testing $a_1(t) = a_0(t)$ is proportional to $S(t)\gamma(t)$.

We consider the optimal weight function $S(t)\gamma(t)$ and also the weight function $[1-F_n(t)]^\rho$ with $\rho = 1$. The level of significance used throughout is 0.05. The parameters used for the simulation are $\lambda_1 = 1$, $\lambda_2 = 3$ and $\alpha = 0$ for the null hypothesis. A sample of size 500 was generated with 1000 repetitions. Figure 1 gives the empirical distribution of the test statistic under the null hypothesis and also when α takes values -0.22, -0.44, -0.67 and -0.89 along with the true standard normal distribution. The empirical distribution corresponding to $\alpha = 0$ is quite close to the true distribution and as α goes away from 0, the distributions look like shifted normal and the curves move away from the true distribution as expected.

To compare the two weight functions and U-statistic U_3 proposed in Dewan *et al.* (2004), empirical distributions of the three statistics are computed using $\alpha = -0.894$. Figure 2 shows these three empirical distributions and also the standard normal distribution. It is clear that the test based on the Harrington and Fleming type weight function has power similar to the test based on optimal weight function. The non-centrality parameter of the U-statistic is smaller than that of the test based on crude hazards. In practice, when one does not want to make assumptions about the structure of the alternative hypothesis, the Harrington and Fleming weight function is a good choice.

3.2 Mortality follow-up study

We analyse the mortality follow-up data from the Finnish cohorts which was a part of the Seven Countries Study in which men in the age-group of 40-59 were examined during 1958-1964 (see Keys *et al.*, 1966 and Karvonen *et al.*, 1970 for the details of the study). There were two Finnish cohorts: one from Ilomantsi in the eastern Finland and one from Pöytyä and Mellilä in the south-western Finland, consisting mainly of rural agricultural populations. The original cohort consist of 823 men from the eastern Finland and 888 men from the south-western Finland. Here, we analyse 40-years of mortality follow-up data of 1560 men who died during the follow-up. The mortality follow-up data give the date of death and underlying cause of death. A death due to common causes, that is coronary heart disease, stroke, cancer, accidents and suicide, is defined as cause 1 and a death due to any other causes is defined as cause 0. The number of deaths due to cause 1 is 621 and that due to cause 2 is 939. Figure 3 shows the empirical conditional probability functions $\Phi_1(t)$ and $\Phi_0(t)$ and Figure 4 shows the corresponding estimates of the crude hazards. It can be seen from the Figure 3 that the probability of dying due to common causes given that a person has survived upto certain age is a decreasing function of age and hence the probability of dying due to other causes is increasing with age. After the age of 85, there is no clear trend. In fact, there are several ages when the rate of change in the Φ function changes. It can be seen that $\Phi_1(t) \leq \Phi_1(0)$. Here the hypothesis of interest is whether $\Phi_1(t)$ is decreasing that is $a_0(t) \leq a_1(t)$ for all t. The value of the test statistic using Harrington and Fleming type of weight function is 5.2411. We accept the hypothesis that $\Phi_1(t)$ is decreasing at 5% level of significance and hence it may be concluded that probability of dying due to common causes given the survival upto a certain age decreases with age and hence the chances of dying due to other causes increases.

4 Discussion

It is shown that the most efficient test based on crude hazards, (2) is equivalent to the most efficient test in the class of the weighted Kolmogorov-Smirnov type tests, (12) for a specific choice of local alternatives. A simple well-known test for comparing hazards of two counting processes can be efficiently applied in the present situation. This allows a straight forward extension of testing hazards for two sample to k sample, in case of k failure modes. A $k-$ sample test for comparing hazards given on pages 345-348 in Andersen *et al.* (1993) can be used in case of $k-$ failure modes or competing risks.

It is demonstrated using the simulated data that Harrington and Fleming type of weight function performs satisfactorily when compared to the optimal weight function. In general when the form of the optimal weight function is not known, Harrington and Fleming type of weight function can be used.

It is easy to check that the equality of crude hazards in the absence of censoring gives the equality of crude hazard in the presence of independent censoring. Hence, in case of right-censored comepting risks data with independent censoring, the above methods can be applied without any changes. We refere to Example V.2.1, Chapter V, Andersen *et al.* (1993) for the discussion regarding censored survival data.

Acknowledgements

We would like to thank Professor J. V. Deshpande, Department of Statistics, University of Pune, India for useful discussion and comments.

References

Andersen, P. K., Borgan, O , Gill, R. D. and Keiding, N. (1993). *Statistical models based on counting processes.* Springer-Verlag: New York.

Breslow,N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Annals of statistics* **2**: 437-453.

Dewan, I., Deshpande, J. V. and Kulathinal, S. B. (2004). On testing dependence between time to failure and cause of failure via conditional probabilities. *Scan. J. Statist.* (in press).

Karvonen, M. J., Orma, E., Punsar, S. et al. (1970). Coronary heart disease in seven countries. Five-year experience in Finland. *Circulation*, **41** (Suppl. I): I-52-I-62.

Keys A., Aravanis, C., Blackburn, H. W. et al. (1966). Epidemiological studies related to coronary heart disease: characteristics of men aged 40-59 in seven countries. *Acta Med Scand* (Suppl. 460): 1-392.

Appendix I: Proof of Theorem 2.2

It is shown in Breslow *et al.* (1974) that as n tends to ∞, $\sqrt{n}(\hat{F}_{1n} - F_1, \hat{F}_n - F)$ converges jointly in $D([0,\infty)) \times D([0,\infty))$ with the Skorokhod topology to zero mean Gaussian processes (X_1, X) with covariance structure, for $s \leq t$,

$$\begin{aligned} cov(X_1(s), X_1(t)) &= F_1(s)(1 - F_1(t)), \\ cov(X(s), X(t)) &= F(s)(1 - F(t)), \\ cov(X_1(s), X(t)) &= F_1(s)(1 - F(t)), \\ cov(X_1(t), X(s)) &= F_1(s) - F_1(t)F(s). \end{aligned}$$

Note that X is a time changed Brownian bridge that is $X(\infty) = 0$, but X_1 is not a Gaussian bridge, that is the limit $X_1(\infty)$ is random. The martingale decomposition for (X_1, X) can be written as

$$dX = dM - \frac{X}{(1-F)}dF, \quad dX_1 = dM_1 - \frac{X}{(1-F)}dF_1,$$

where M and M_1 are Gaussian martingales with

$$d\langle M, M\rangle = (1-F)dF \text{ and } d\langle M, M_1\rangle = d\langle M_1, M_1\rangle = (1-F)dF_1.$$

If we denote $X_0 = X - X_1$, $M_0 = M - M_1$, $F_0 = F - F_1$, we get the linear system of stochastic differential equations

$$dX_0 = dM_0 - \frac{(X_0+X_1)}{(1-F)}dF_0$$
$$dX_1 = dM_1 - \frac{(X_0+X_1)}{(1-F)}dF_1$$

where M_0 and M_1 are orthogonal Gaussian martingales with $d\langle M_i\rangle = (1-F)dF_i$. The solution can be given explicitely in terms of (M_0, M_1) and matrix exponentials. Note that $X_0(\infty) + X_1(\infty) = 0$.

By functional delta method, it can be shown that

$$\begin{aligned}
&\sqrt{n}(\hat{\Psi}_n(u,v) - \Psi(u,v)) = \\
&\sqrt{n}(\hat{F}_{1n}(\infty) - F_1(\infty))(F(v) - F(u)) + \\
&\hat{F}_{1n}(\infty)\sqrt{n}[(\hat{F}_n(v) - F(v)) - (\hat{F}_n(u) - F(u))] + \\
&\sqrt{n}(\hat{F}_{1n}(u) - F_1(u))(1 - F(v)) - \hat{F}_{1n}(u)\sqrt{n}(\hat{F}_n(v) - F(v)) \\
&-\sqrt{n}(\hat{F}_{1n}(v) - F_1(v))(1 - F(u)) + \hat{F}_{1n}(v)\sqrt{n}(\hat{F}_n(u) - F(u))
\end{aligned}$$

converges to

$$\begin{aligned}
&Z(u,v) = X_1(\infty)(F(v) - F(u)) + F_1(\infty)(X(v) - X(u)) + \\
&X_1(u)(1 - F(v)) - F_1(u)X(v) - X_1(v)(1 - F(u)) + F_1(v)X(u).
\end{aligned}$$

We can express

$$Z(u,v) = \int_0^\infty f(u,v,t)dX_1(t) + \int_0^\infty g(u,v,t)dX(t),$$

so that

$$\begin{aligned}
&\mathrm{Cov}(Z(u_1,v_1), Z(u_2,v_2)) = \\
&\mathrm{Cov}\Big(f(u_1,v_1,\tau)I(\eta=1) + g(u_1,v_1,\tau), f(u_2,v_2,\tau)I(\eta=1) + g(u_2,v_2,\tau)\Big),
\end{aligned}$$

where

$$f(u,v,t) = (F(v)-F(u)) + I_{[0,u]}(t)(1-F(v)) - I_{[0,v]}(t)(1-F(u)), \tag{18}$$

$$g(u,v,t) = I_{[0,u]}(t)F_1(v) - I_{[0,v]}(t)F_1(u) + F_1(\infty)I_{[u,v]}(t). \tag{19}$$

Under H_0, $F_1(t) = F_1(\infty)F(t) = \phi F(t)$ and hence,

$$\begin{aligned} f(u,v,t) &= (F(v)-F(u)) + I_{[0,u]}(t)(1-F(v)) - I_{[0,v]}(t)(1-F(u)), \\ g(u,v,t) &= \phi(I_{[0,u]}(t)F(v) - I_{[0,v]}(t)F(u) + I_{[0,v]}(t) - I_{[0,u]}(t)) \\ &= \phi(-I_{[0,u]}(t)(1-F(v)) + I_{[0,v]}(t)(1-F(u))) \\ &= \phi(F(v)-F(u)-f(u,v,t)). \end{aligned}$$

Note that

$$\int_0^\infty f(u,v,t)dF_1(t) = \phi \int_0^\infty f(u,v,t)dF(t) = 0$$

and

$$\int_0^\infty g(u,v,t)dF(t) = \phi \int_0^\infty (F(v)-F(u)-f(u,v,t))dF(t) = \phi(F(v)-F(u)).$$

It can be verified using simple calculations that

$$\begin{aligned} cov(Z(u_1,v_1),Z(u_2,v_2)) &= 0 \text{ if } u_1 \le v_1 \le u_2 \le v_2 \text{ or } u_2 \le v_2 \le u_1 \le v_1 \\ &= \phi(1-\phi)(1-F(max(v_1,v_2)) \\ &\quad (1-F(min(u_1,u_2))(F(min(v_1,v_2)-F(max(u_1,u_2)), \\ &\quad \text{otherwise}, \\ var(Z(u,v)) &= \phi(1-\phi)(1-F(v))(1-F(u))(F(v)-F(u)) \\ &= \phi(1-\phi)(1-t)(1-s)(t-s), \end{aligned}$$

where $F(u) = s$ and $F(v) = t$, and hence $0 \le s \le t \le 1$.

Hence the Theorem 2.2.

Appendix II: Optimal weight function

To find the asymptotic noncentrality parameter under the sequence of local alternatives $\{P^{(n)}(\theta)\}$ given in (6), we need to compute the asymptotic mean and variance of the weighted Kolmogorov-Smirnov test (12) for a sequence of possibly random kernels $K_n(u,v)$ approximating $K(u,v) = k(u)\dot{\delta}_u(v)$, so that

$$\sqrt{n} \iint_{u \le v} (K_n(u,v) - k(u)\dot{\delta}_u(v))\Psi^{(n)}(u,v)dudv =$$

$$\sqrt{n}\left(\iint_{u \le v} K^n(u,v)\Psi^{(n)}(u,v)dudv - \int_0^\infty k(u)S_1^{(n)}(u)S_0^{(n)}(u)(dA_1^{(n)}(u) - dA_0^{(n)}(u))\right)$$

$$\overset{\mathbf{P}^{(n)}(\theta)}{\rightarrow} 0.$$

Using (14) and (15), it is easy to verify that the asymptotic mean and variance of $\sqrt{n}\hat{\Delta}_n$ is

$$\mu = (\theta_1 - \theta_0)\phi(1-\phi)\int_0^\infty k(u)S^2(u)a(u)\gamma(u)du,$$

and

$$\text{Var}\left(\iint_{u \le v} K^*(u,v)Z(u,v)dudv\right) =$$

$$\iint\iint I(u \le v)I(u' \le v')k(u)\dot{\delta}_u(v)k(u')\dot{\delta}_{u'}(v') \text{ Cov } (Z(u,v), Z(u',v'))du'dv'dudv.$$

Note that

$$-\int \dot{\delta}_{u'}(v')I(u' \le v')R(u,v,u',v')dv' \quad = \quad \frac{\partial}{\partial v'}I(u' \le v')R(u,v,u',v')|_{v'=u'},$$

where

$$\begin{aligned} R(u,v,u',v') &= cov(Z(u,v), Z(u',v')) \\ &= \phi(1-\phi)S(v')S(min(u,u'))(S(max(u,u')) - S(v))I(v \le v')I(u' \le v) \\ &\quad +\phi(1-\phi)S(v)S(min(u,u'))(S(max(u,u')) - S(v'))I(v' \le v)I(u \le v'). \end{aligned}$$

Finally,

$$\sigma^2 = \mathrm{Var}\left(\iint_{u\leq v} K^*(u,v)Z(u,v)dudv\right) \quad = \quad \phi(1-\phi)\int k^2(u)S^3(u)a(u)du. \tag{20}$$

The value of the noncentrality parameter is

$$(\theta_1-\theta_0)^2\phi(1-\phi)(\int_0^\infty k(u)S^2(u)a(u)\gamma(u)du)^2(\int_0^\infty k^2(u)S^3(u)a(u)du)^{-1}$$

and is maximised when $k(u) = \gamma(u)S^{-1}(u)$ and the maximum value of the noncentrality parameter is

$$(\theta_1-\theta_0)^2\phi(1-\phi)\int_0^\infty \gamma^2(u)S(u)a(u)du$$

which is exactly the same as that for the test (2) based on crude hazards.

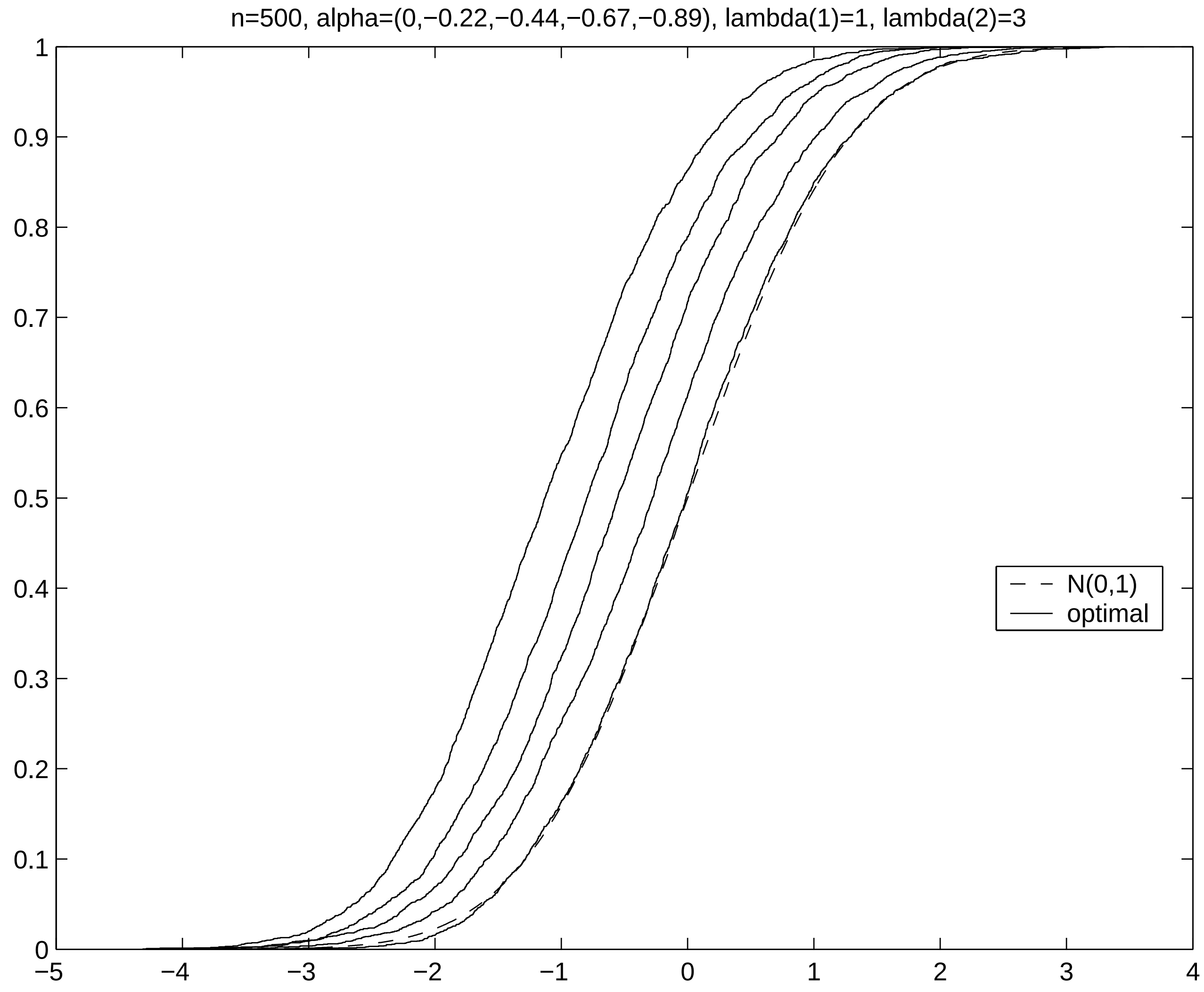

Figure 1: Empirical distributions of the test statistic based on crude hazards for various values of α $n = 500$, $\alpha = (0, -0.22, -0.44, -0.67, -0.89)$, $\lambda_1 = 1, \lambda_2 = 3$

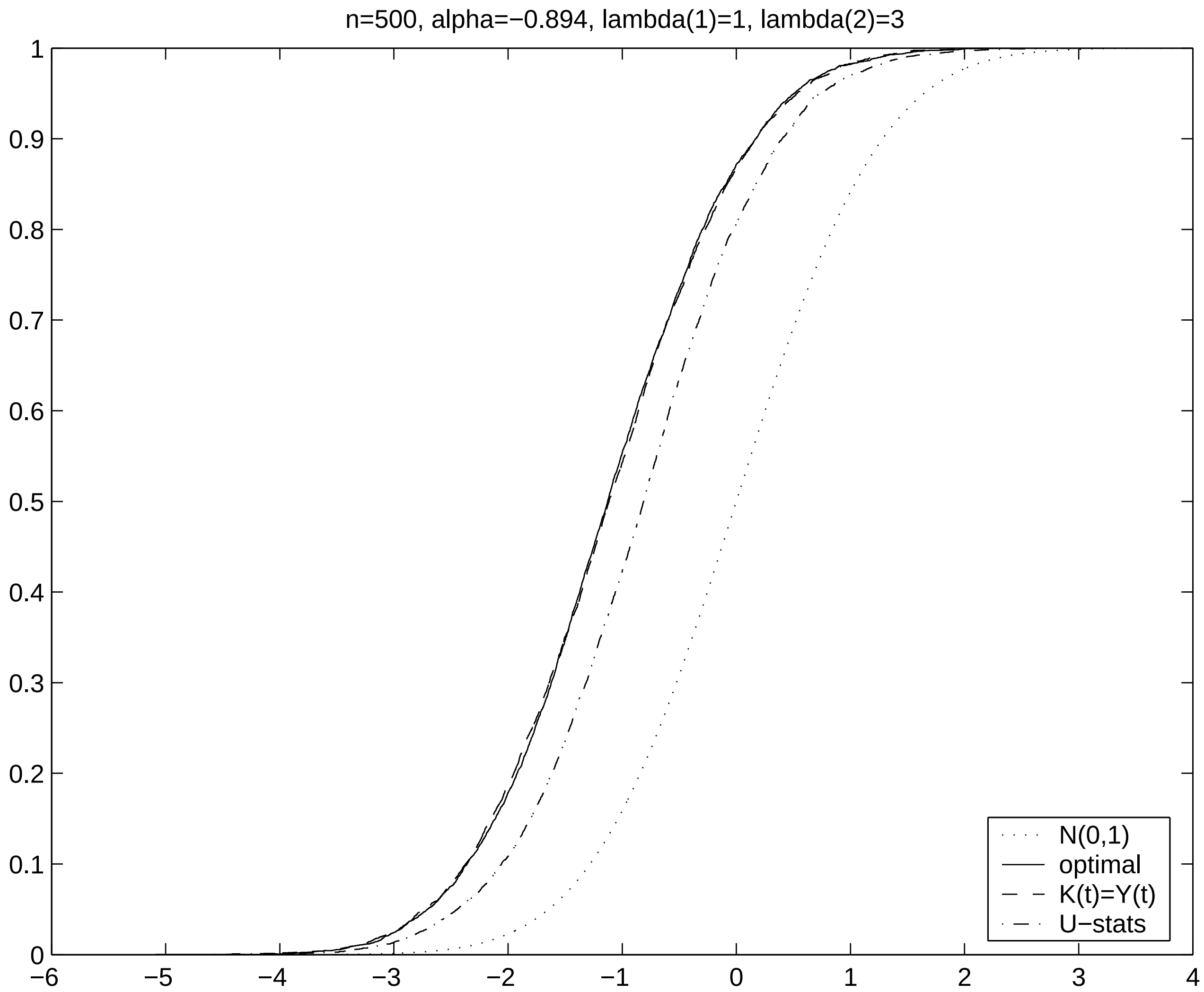

Figure 2: Empirical distributions of test statistic based on crude hazards using optimal kernel and Harrington-Flemming type kernel, and U_3 test $n = 500$, $\alpha = -0.89$, $\lambda_1 = 1, \lambda_2 = 3$

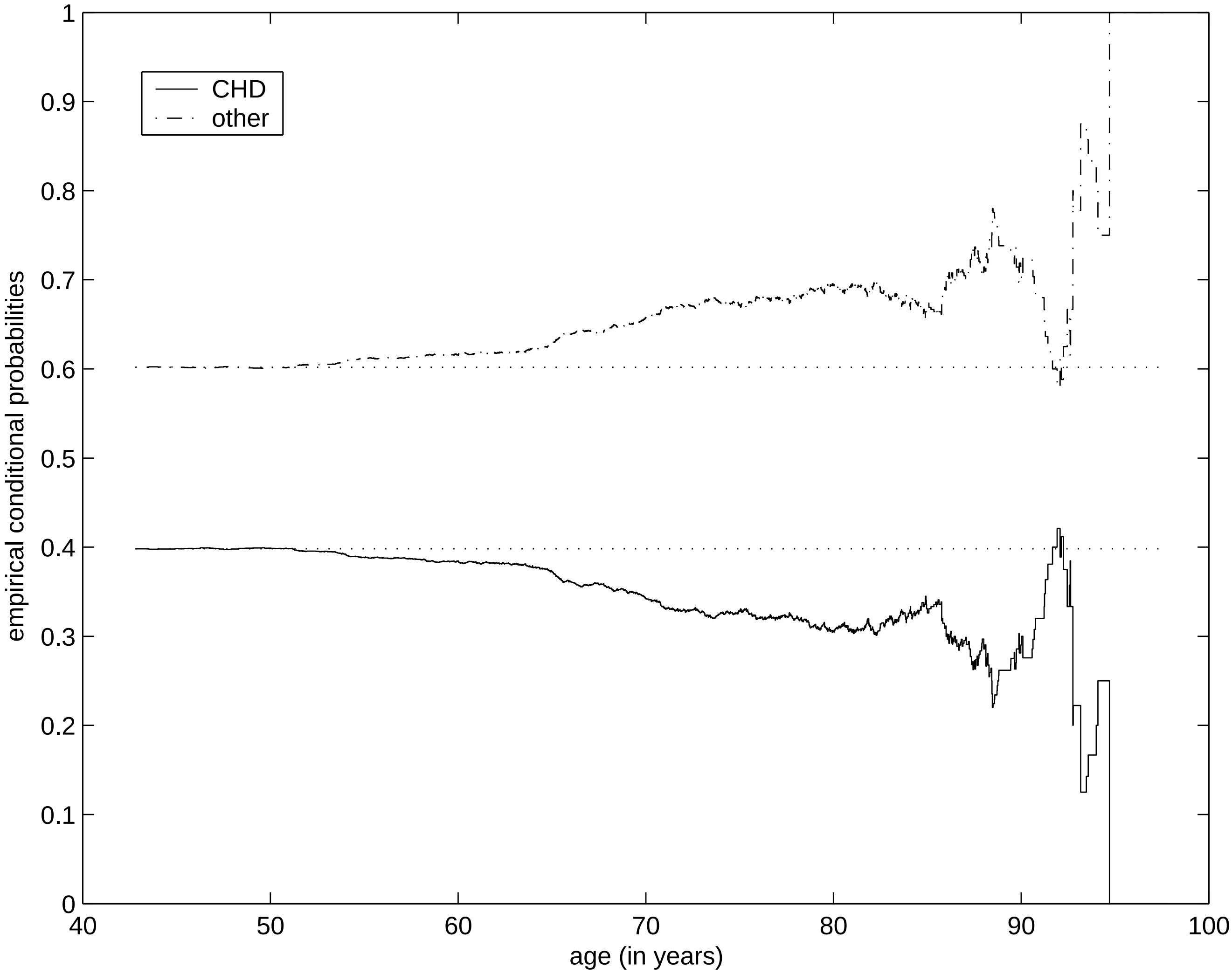

Figure 3: Empirical conditional probabilities for two competing causes of death

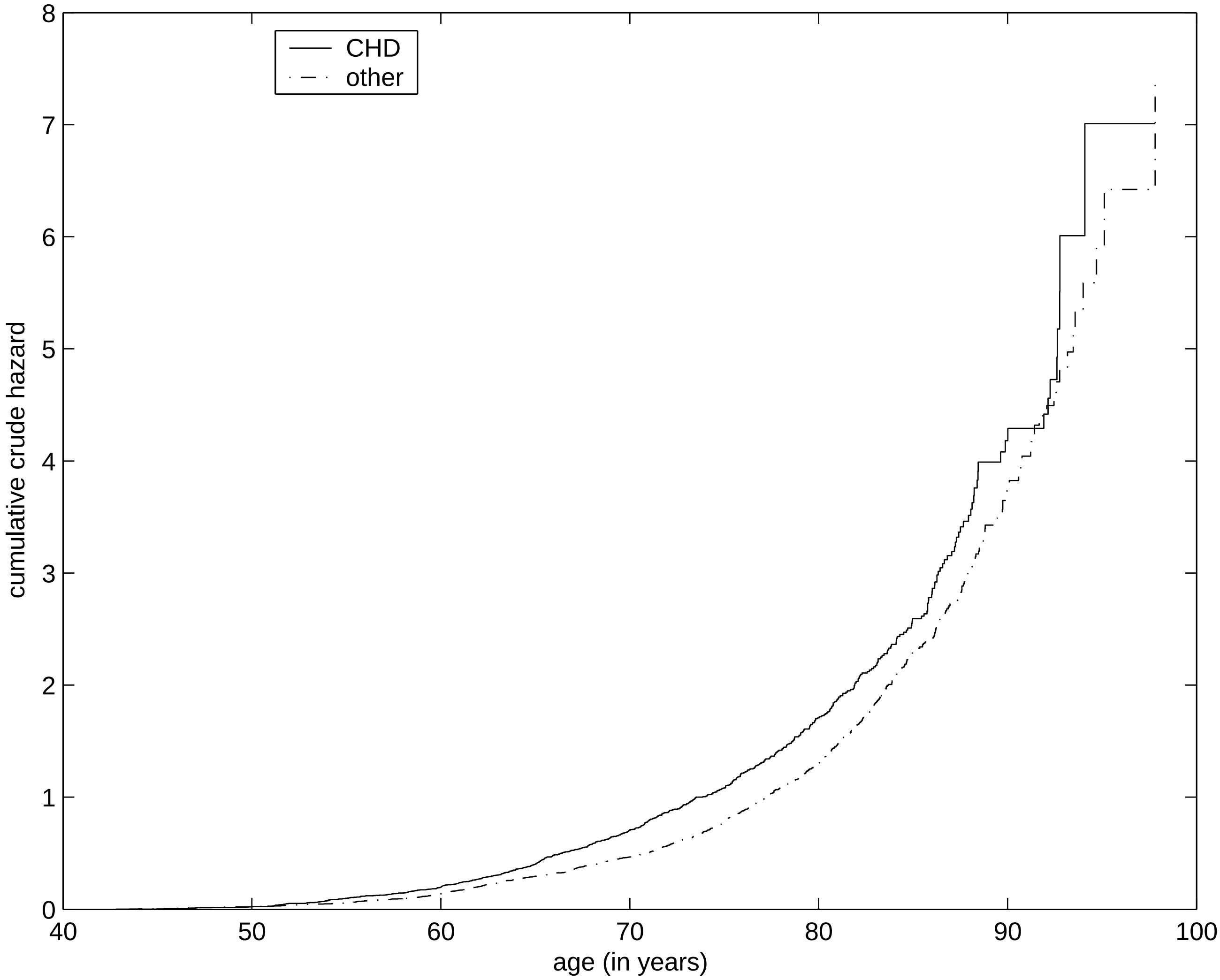

Figure 4: Nelson-Aalen estimates of cumulative crude hazards for two competing causes of death