TECHNICAL REPORT NO. 433

January 1976

MODIFICATION OF OPTIMAL $C(\alpha)$ TEST WITH APPLICATION TO CENSORED SURVIVAL STUDIES.

by

Anastasios A. Tsiatis* and Harold S. Javitz†

*Dept. of Statistics and Preventive Medicine
Univ. of Wis., Madison, Wisconsin 53706

†Dept. of Statistics, Univ. of California
Berkeley, California 94720

SUMMARY

This paper modifies the optimal $C(\alpha)$ test derived by Neyman for use in comparative censored survival studies. In particular, the problem of comparing the survival experiences of two populations (treatment and control) which are subject to random censoring is considered. The survival times of individuals are assumed to follow a probability density which is a function of certain parameters, whereas no functional form for the censoring time probability density is assumed. The problem of identifying whether treatment is beneficial is reduced to the problem of testing the value of a parameter ξ. It is found that the regression coefficients of the optimal $C(\alpha)$ test are functions of the unknown censoring time probability densities. However, these regression coefficients may be estimated in such a fashion that the resulting test statistic retains the same optimality properties as the optimal $C(\alpha)$ test and therefore the problem has an explicit solution.

1. INTRODUCTION

In "comparative survival studies" two populations of individuals, a "control" group and a "treated" group, are observed over a length of time. In the simplest of such studies (in the absence of censoring) the recorded observations consist of times until death of the participant from a specified disease or cause. In most human clinical studies the "control" group is itself receiving the standard therapy while the "treated" group is receiving an additional or different therapy. The statistician and medical personnel are interested in comparing these therapies and recommending one as superior. The length of life of any individual in the study is assumed to be independent of the lengh of life of any other individual and is assumed to follow a density which is a function of certain parameters. These parameters reflect the group to which the individual belongs. For example, the lifetime of the control group participant might be assumed to follow an exponential density $\theta \exp(-\theta t)$, $t > 0$, and the lifetime of the treated group participant assumed to follow an exponential density $(\theta+\xi) \exp[-(\theta+\xi)t]$, where neither θ nor ξ is known. If the parameter ξ is smaller than zero the therapy applied to the treated group is to be preferred. Unfortunately in practice and particularly in human clinical follow up studies, some participants are often removed from observation. Such individual are said to be censored from the study.

Some of the causes of censoring might be reasonably assumed to be independent of the "force of mortality". It is these causes of censoring in which we are interested.

Loosely speaking, if in addition to the assumption of independence of the force of mortality and the force of censoring, the statistician were capable of specifying the exact form of the force of censoring then the theory of $C(\alpha)$ tests as formulated by Neyman (1959)might be applied to determine whether the control therapy or treatment therapy were superior. However, the statistician may be unwilling or unable to further specify a parametric form for the force of censoring. This paper proves that an optimal $C(\alpha)$ test may still be derived for deciding which therapy is superior.

2. FORMULAE AND NOTATION

In the experimental situation which we will discuss in detail individuals are assigned to the control or treated group at random with known probability π of being assigned to the control group. For each of n individuals a three-tuple $(Y_i, \alpha_i, \delta_i)$ is observed. Let

Y_i = time at which ith individual dies or is censored from the study.

α_i = zero if ith individual was assigned to the control group, and one otherwise.

δ_i = zero if ith individual died from the specified disease, and one if he was censored from the study.

For each participant, the time at which he is censored or dies is assumed to be the minimum of two independent random variables. One of these random variables represents the hypothetical time of death of the individual in the absence of censoring and the other random variable represents the hypothetical time of censoring in the absence of death. More exactly, introduce the unobservable random variables

T_{cd} = time of death of participant if he were assigned to the control group.

T_{cs} = time of censoring of the participant if he were assigned to the control group.

T_{td} = time of death of the participant if he were assigned to the treated group.

T_{ts} = time of censoring of the participant if he were assigned to the treated group.

Note that the letter s is used as a mnemonic for censoring.

As stated T_{cd} and T_{cs} are independent, as are T_{td} and T_{ts}. If α equals zero, $Y = \min(T_{cd}, T_{cs})$ and if α equals one, $Y = \min(T_{td}, T_{ts})$.

The random variables T_{cs} and T_{ts} are assumed to have arbitrary densities $h_0(t)$ and $h_1(t)$ respectively. Let the corresponding cumulative distribution function be denoted by $H_0(t)$ and $H_1(t)$.

The random variables T_{cd} and T_{td} are assumed to have exponential densities with parameters θ and $(\theta+\xi)$ respectively. Neither θ nor ξ are assumed known. The analysis which is undertaken here does not depend on the assumption that T_{cd} and T_{td} have exponetial densities. The reader should find it relatively easy to extend these results to more complex distributions such as the Weibull.

For convenience let $\underset{\sim}{X} = (Y, \alpha, \delta)$. The density of $\underset{\sim}{X}$ will be denoted $p(x|\pi,\theta,\xi)$ and is given by

$$p(x|\pi,\theta,\xi) = [\pi\cdot\{\theta\exp(-\theta y)(1-H_0(y))\}^{1-\delta}\ \{h_1(y)\exp(-\theta y)\}^{\delta}]^{1-\alpha}$$

$$\cdot[(1-\pi)\{(\theta+\xi)\exp(-(\theta+\xi)y)(1-H_1(y))\}^{1-\delta}\ \{h_1(y)\exp(-(\theta+\xi)y)\}^{\delta}]^{\alpha} \qquad (2.1)$$

3. NEYMAN'S C(α) TEST

The statistician desires to test H_0: $\xi \geq 0$ against H_1: $\xi < 0$, at some predetermined level α (e.g. .01). If H_0 is rejected the standard treatment is abandoned in favor of the new treatment.

The C(α) test introduced by Neyman (1959) is a "locally asymptotically optimal" test procedure for testing H_0 against H_1.

Assume that $p(x|\pi,\theta,\xi)$ satisfies the Cramér-type conditions in Neyman (1959). In the customary notation, the optimal C(α) test rejects H: $\xi \geq 0$ whenever

$$\frac{n^{-\frac{1}{2}}}{\sigma(\hat{\theta}_n)} \sum_{i=1}^{n} [\phi_\xi(X_i,\hat{\theta}_n) - a^o(\hat{\theta}_n)\phi_\theta(X_i,\hat{\theta}_n)] \leq z_{-\alpha} \qquad (3.1)$$

where

(a) $z_{-\alpha}$ is the lower cut off point of a unit normal distribution.

(b) $\phi_\xi(x,\theta)$, $\phi_\theta(x,\theta)$, are defined as the partial derivatives of the log density function ln $p(x|\pi,\theta,\xi)$ with respect to the test parameter ξ and the nuisance parameter θ. These partial derivatives are evaluted at $\xi = 0$.

(c) The coefficient $a^{o}(\theta)$ is so selected as to minimize the variance of

$$\phi_{\xi}(X,\theta) - a(\theta)\phi_{\theta}(X,\theta) \tag{3.2}$$

on the assumption that $\xi = 0$.

(d) The expression $\sigma(\theta)$ stands for the S.D. of (3.2) on the assumption that $\xi = 0$.

(e) The $\hat{\theta}_n$ is a "root -n" consistent estimator of θ on the assumption that $\xi = 0$.

Direct calculation yields

$$\phi_{\xi}(x,\theta) = \left.\frac{\partial \ln p(x|\pi,\theta,\xi)}{\partial \xi}\right|_{\xi=0} = \frac{\alpha(1-\delta)}{\theta} - \alpha y \tag{3.3}$$

$$\phi_{\theta}(x,\theta) = \left.\frac{\partial \ln p(x|\pi,\theta,\xi)}{\partial \theta}\right|_{\xi=0} = \frac{(1-\alpha)(1-\delta)}{\theta} + \frac{\alpha(1-\delta)}{\theta} - y \tag{3.4}$$

The $C(\alpha)$ test only requires a "root -n" consistent estimator $\hat{\theta}_n$ of θ. That is, $n^{\frac{1}{2}}(\hat{\theta}_n-\theta)$ stays bounded in probability as $n \to \infty$ on the assumption $\xi = 0$. (Neyman defines a root -n consistent estimator in a different but equivalent fashion Javitz (1975)).

Let

N_{cd} = total number of deaths in control group

N_{cs} = total number of censoring in control group

N_{td} = total number of deaths in treated group

N_{ts} = total number of censorings in treated group

Y_c = sum of Y_i for control group participants

Y_t = sum of Y_i for treated group participants.

Under the assumed conditions, the maximum likelihood estimator $\hat{\theta}$

$$\hat{\theta}_n = \frac{N_{cd} + N_{td}}{Y_c + Y_t} \tag{3.5}$$

is root -n consistent for θ and will be used in the remainder of the paper.

Using (3.1), (3.3), (3.4), the optimal $C(\alpha)$ test rejects H_0: $\xi \geq 0$ whenever

$$\frac{n^{-\frac{1}{2}}}{\sigma(\hat{\theta}_n)} \{[\frac{N_{td}}{\hat{\theta}_n} - Y_t] - a^o(\hat{\theta}_n)[\frac{N_{cd}+N_{td}}{\hat{\theta}_n} - (Y_t+Y_c)]\} < z_{-\alpha} \tag{3.6}$$

where

$$a^o(\theta) = \frac{Cov[\phi_\xi(X,\theta),\phi_\theta(X,\theta)|\pi,\theta,\xi=0]}{Var[\phi_\theta(X,\theta)|\pi,\theta,\xi=0]}$$

$$\sigma^2(\theta) = Var[\phi_\xi(X,\theta|\pi,\theta,\xi=0] - a^o(\theta)Cov[\phi_\xi(X,\theta),\phi_\theta(X,\theta)|\pi,\theta,\xi=0].$$

Under the assumed conditions

$$\mathrm{Cov}[\phi_\xi(X,\theta),\phi_\theta(X,\theta)|\pi,\theta,\xi=0] = -E\left[\frac{\partial^2 \ln p(X|\pi,\xi,\theta)}{\partial\xi\partial\theta}\bigg|_{\xi=0}\bigg|\pi,\theta,\xi=0\right]$$

$$= E\left[\frac{\alpha(1-\delta)}{\theta^2}\right] = \frac{1-\pi}{\theta^2}\int_0^\infty \theta\exp(-\theta x)[1-H_1(x)]dx \qquad (3.7)$$

$$\mathrm{Var}[\phi_\theta(X,\theta)|\pi,\theta,\xi=0] = -E\left[\frac{\partial^2 \ln p(X|\pi,\theta,\xi)}{\partial\theta^2}\bigg|_{\xi=0}\bigg|\theta,\xi=0\right]$$

$$= E\left[\frac{\alpha(1-\delta)}{\theta^2} + \frac{(1-\alpha)(1-\delta)}{\theta^2}\right] \qquad (3.8)$$

$$= \frac{1-\pi}{\theta^2}\int_0^\infty \theta\exp(-\theta x)[1-H_1(x)]dx + \frac{\pi}{\theta^2}\int_0^\infty \theta\exp(-\theta x)[1-H_o(x)]dx$$

$$\mathrm{Var}[\phi_\xi(X,\theta)|\pi,\theta,\xi=0] = -E\left[\frac{\partial^2 \ln p(X|\pi,\theta,\xi)}{\partial\xi^2}\bigg|_{\xi=0}\bigg|\pi,\theta,\xi=0\right]$$

$$= E\left[\frac{\alpha(1-\delta)}{\theta^2}\right] = \frac{1-\pi}{\theta^2}\int_0^\infty \theta\exp(-\theta x)[1-H_1(x)]dx \qquad (3.9)$$

Since $a^o(\theta)$ and $\sigma^2(\theta)$ are functions of arbitrary $h_o(x)$ and $h_1(x)$ they cannot be calculated.

4. MODIFICATION OF NEYMAN'S OPTIMAL C(α) TEST

The following theorem states that if $\hat{a}_n^o$ and $\hat{\sigma}_n$ are consistent estimator's of $a^o(\theta)$ and $\sigma(\theta)$ on the assumption that $\xi = 0$, these estimators may be substituted into the optimal C(α) test (3.1) without changing its local asymptotic optimality properties or its distribution.

Theorem 1. If $\hat{a}_n^o$ and $\hat{\sigma}_n$ are consistent estimators of $a^o(\theta)$ and $\sigma(\theta)$ on the assumption $\xi = 0$. Then

$$\frac{n^{-\frac{1}{2}}}{\sigma(\hat{\theta}_n)} \sum_{i=1}^{n} [\phi_\xi(X_i,\hat{\theta}_n) - a^o(\hat{\theta}_n)\phi_\theta(X_i,\hat{\theta}_n)]$$

$$- \frac{n^{-\frac{1}{2}}}{\hat{\sigma}_n} \sum_{i=1}^{n} [\phi_\xi(X_i,\hat{\theta}_n) - \hat{a}_n^o\phi_\theta(X_i,\hat{\theta}_n)] \tag{4.1}$$

converges in probability to zero on the assumption $\xi = 0$ and on the sequence of asymptotic alternatives considered by Neyman (1959).

(The proof is given in the appendix)

By the strong law of large numbers and Slutzky's theorem the covariance (3.7) can be consistently estimated by $N_{td}/(n\hat{\theta}_n^2)$, and the variances (3.8) and (3.9) can be consistently estimated by $(N_{td}+N_{cd})/(n\hat{\theta}_n^2)$ and $N_{td}/(n\hat{\theta}_n^2)$. Consequently consistent estimators for $a^o(\theta)$ and $\sigma^2(\theta)$ are, respectively,

$$\hat{a}_n^o = \frac{N_{td}}{N_{td}+N_{cd}} \tag{4.2}$$

$$\hat{\sigma}_n^2 = \frac{N_{td}N_{cd}}{n\hat{\theta}_n(N_{td}+N_{cd})} \tag{4.3}$$

Substituting (3.5), (4.2) and (4.3) into (3.6) we reject H_o: $\xi \geq 0$ whenever

$$\frac{N_{td}Y_c - N_{cd}Y_t}{Y_c+Y_t} \left(\frac{N_{td}+N_{cd}}{N_{td}N_{cd}}\right)^{\frac{1}{2}} < z_{-\alpha} \tag{4.4}$$

5. EXAMPLE

As an example of the application of the modified $C(\alpha)$ test (4.4) data from the 44 month prospective study sponsored by the American Cancer Society and reported by Hammond and Horn (1958) was examined. Some 200,000 men in the age range 50-70 years were interviewed and a statement obtained from each to his smoking habits. Periodically, inquiry was made and the time and cause of death as stated on the death certificate were recorded. Table 1 summarizes the essential data for the group of men 60-65 years of age at the time of the original inquiry. Deaths are dichotomized into two classes: deaths from cancer and death from other causes.

TABLE 1

MALES 60-65 YEARS OF AGE

Non Smokers		Smokers	
Man yrs. exposure to risk =	27,817	man yrs. exposure to risk =	75,557
No. of cancer deaths	67	No. of cancer deaths	428
other causes	371	other causes	1,588
Total	438	Total	2,016

The terms in this paper and their associated meaning in this experimental study are identified as follows.

control group:	smokers
treatment group:	non-smokers
death:	death from cancer
censoring:	death from any other cause, survival until end of study, loss to follow up, etc.

The expression "man years exposure to risk" is the total number of years lived in the 44 month period of study.

From Table 1 we obtain that

$$N_{td} = 67 \quad , \quad N_{cd} = 428$$
$$Y_t = 27{,}817, \quad Y_c = 75{,}557 \tag{5.1}$$

In order to use the specific test criterion (4.4), we are of course assuming that the risk of death from cancer is constant amongst men 60-65 years of age during the 44 month period of the study.

Substituting the values of (5.1) into the formula for the optimal $C(\alpha)$ test (4.4), the value of the test statistic obtained is -8.70. We conclude therefore, that there is a highly significant difference in the rate of cancer amongst smoker and non-smokers.

ACKNOWLEDGEMENTS

I would like to thank Dr. Avinash Singh from the University of California at Berkeley for his help on some of the contiginty arguments used in this paper. This investigation was partially supported by NIH Research Grant No. GM-10525, National Institute of Health, Public Health Service, U.S. Office of Naval Research, contract number ONR N00014 69A 0200 1051, and partially supported by the auspices of the U.S. Energy Research and Development Agency.

REFERENCES

Hajek, J., Sidak, Z. (1967). Theory of Rank Tests. Academic Press, New York.

Hammond, E.C., Horn, D. (1958). Smoking and death rates-report on forty-four months of follow up of 187,783 men, I. Total mortality. Journal of the American Medical Association 166, 1159-72.

Hammond, E.C., Horn, D. (1958). Smoking and death rates-report on forty-four months of follow up on 187,783 men, II. Death rates by cause. Journal of the American Medical Association 166, 1294-1308.

Javitz, H.S. (1975). Generalized smooth tests of goodness of fit, independence, and the quality of distribution. Unpublished Ph.D. dissertation, Department of Statistics, University of California at Berkeley.

Neyman, J. (1959). Optimal asymptotic tests of composite statistical hypotheses. In: Probability and Statistics. (The Harald Cramér Volume) Almquist and Wiksells, Uppsalla, Sweden.

Tsiatis, A.A. (1974). Evaluation of competing risks. Unpublished Ph.D. dissertation, Department of Statistics, University of California at Berkeley.

APPENDIX

Let $p_n(\theta,\xi)$ denote the joint density of n iid random variables each with density $p(\theta,\xi)$. If we show that the sequence of densities $p_n(\theta,\xi=0)$ is contiguous to the sequence of densities $p_n(\theta,\xi=n^{-\frac{1}{2}}\tau)$, then it suffices to prove Theorem 1 under the assumption $\xi = 0$.

Theorem A.1 (Le Cam's Second Lemma)

If $L_n = \ln \left[\frac{p_n(\theta,\xi=n^{-\frac{1}{2}}\tau)}{p_n(\theta,\xi=0)}\right]$ is *asymptotically normal* $(-\frac{1}{2}\sigma^2,\sigma^2)$ *then the densities* $p_n(\theta,\xi=0)$ *and* $p_n(\theta,\xi=n^{-\frac{1}{2}}\tau)$ *are contiguous.*

Proof Hajek and Sidak (1967, pp. 203-208)

Theorem A.2 *Under the assumed Cramér type conditions in Neyman* (1959) *the sequence of densities* $p_n(\theta,\xi=0)$ *and* $p_n(\theta,\xi=n^{-\frac{1}{2}}\tau)$ are contiguous.

Proof By Theorem A.1 it suffices to show that

$$\sum_{i=1}^{n} \ln p(X_i,\theta,n^{-\frac{1}{2}}\tau) - \ln p(X_i,\theta,0) \xrightarrow{\mathcal{L}} N(-\tfrac{1}{2}\sigma^2,\sigma^2)$$

under the assumption that $\xi = 0$.

Expanding in a Taylor series about $\xi = 0$ yields

$$\tau n^{-\frac{1}{2}} \sum_{i=1}^{n} \phi_\xi(X_i,\theta,0) + \frac{\tau^2}{2} n^{-1} \sum_{i=1}^{n} \phi_{\xi\xi}(X_i,\theta,\xi^*) \qquad (A.1)$$

where

$$\phi_\xi(X,\theta,\xi) = \frac{\partial \ln p(X,\theta,\xi)}{\partial \xi}$$

$$\phi_{\xi\xi}(X,\theta,\xi) = \frac{\partial^2 \ln p(X,\theta,\xi)}{\partial \xi^2}$$

and $0 \leq \xi_n^* \leq n^{-\frac{1}{2}}\tau$

The first term in (6.1) converges in law to $N[0,\tau^2 \text{Var}[\phi_\xi(X,\theta)|\theta,\xi=0]]$ by the Central limit theorem.

One of the Cramér conditions states that there exists $\psi(x,\theta) \geq 0$ satisfying

$$|\phi_{\xi\xi}(X,\theta,\xi^*)| \leq \psi(X,\theta) \text{ for all } \xi^* \text{ in a vicinity, } \gamma(0), \text{ of } 0,$$

and

$$E[\psi(X,\theta)|\theta,\xi=0] < \infty.$$

Mourier's theorem yields

$$\sup_{\xi^*\epsilon\gamma(0)} \left| \frac{1}{n} \sum_{i=1}^{n} \phi_{\xi\xi}(X_i,\theta,\xi^*) - E[\phi_{\xi\xi}(X,\theta,\xi^*)|\theta,\xi=0] \right| \xrightarrow{a.s.} 0.$$

By this theorem and continuity of $E[\phi_{\xi\xi}(X,\theta,\xi)|\theta,\xi=0]$ in $\xi\varepsilon\gamma(0)$, we obtain that

$$\frac{1}{n}\sum_{i=1}^{n} \phi_{\xi\xi}(X_i,\theta,\xi_n^*) - E[\phi_{\xi\xi}(X,\theta,0)|\theta,\xi=0] \overset{p}{\to} 0.$$

It is easily shown that

$$E[\phi_{\xi\xi}(X,\theta,0)|\theta,\xi=0] = -\mathrm{Var}[\phi_{\xi}(X,\theta,0)|\theta,\xi=0]$$

therefore the second term in (6.1) converges in probability to $-\frac{\tau^2}{2}\mathrm{Var}[\phi_{\xi}(X,\theta,0)|\theta,\xi=0]$. Using Slutsky's theorem L_n converges in law to

$$N[-\frac{\tau^2}{2}\mathrm{Var}[\phi_{\xi}(X,\theta,0)|\theta,\xi=0],\ \tau^2\mathrm{Var}[\phi_{\xi}(X,\theta,0)|\theta,\xi=0]]. \quad \text{QED.}$$

<u>Theorem</u> A.3

$$\frac{n^{-\frac{1}{2}}}{\sigma(\hat{\theta}_n)}\sum_{i=1}^{n}[\phi_{\xi}(X_i,\hat{\theta}_n) - a^o(\hat{\theta}_n)\phi_{\theta}(X_i,\hat{\theta}_n)]$$

$$-\frac{n^{-\frac{1}{2}}}{\hat{\sigma}_n}\sum_{i=1}^{n}[\phi_{\xi}(X_i,\hat{\theta}_n) - \hat{a}_n^o\phi_{\theta}(X_i,\hat{\theta}_n)]$$

<u>converges in probability to zero on the assumption</u> $\xi = 0$.

Proof Rewriting the above we obtain

$$\overset{(A)}{\left[1-\frac{\sigma(\hat{\theta}_n)}{\hat{\sigma}_n}\right]}\overset{(B)}{\left[\frac{n^{-\frac{1}{2}}}{\sigma(\hat{\theta}_n)}\sum_{i=1}^{n}[\phi_\xi(X_i,\hat{\theta}_n)-a^o(\hat{\theta}_n)\phi_\theta(X_i,\hat{\theta}_n)]\right]} \tag{A.2}$$

$$+\overset{(C)}{(\hat{a}_n^o-a^o(\hat{\theta}_n))}\overset{(D)}{\left[\frac{n^{-\frac{1}{2}}}{\hat{\sigma}_n}\sum_{i=1}^{n}\phi_\theta(X_i,\hat{\theta}_n)\right]}.$$

Clearly terms (A) and (C) converge in probability to zero. Term B which is the optimal $C(\alpha)$ test derived by Neyman (1959) converges in law to a unit normal. To complete the proof we need to show that term (D) is bounded in probability. Expanding by Taylor series

$$n^{-\frac{1}{2}}\sum_{i=1}^{n}\phi_\theta(X_i,\theta)+(n^{\frac{1}{2}}(\hat{\theta}_n-\theta))[n^{-1}\sum_{i=1}^{n}\phi_{\theta\theta}(X_i,\theta^*)] \tag{A.3}$$

where θ^* falls beteen θ and $\hat{\theta}_n$.

The first term of (A.3) converges in law to a normal distribution with 0 mean and finite variance. The estimate $\hat{\theta}_n$ is root n consistent therefore $n^{\frac{1}{2}}(\hat{\theta}_n-\theta)$ is bounded in probability. Finally by Cramér's conditions and Mourier's Theorem $n^{-1}\sum_{i=1}^{n}\phi_{\theta\theta}(X_i,\theta^*)$ converges in probability to $-\text{Var}[\phi_\theta(X,\theta)|\theta,\xi=0]$. Term (D) is therefore bounded in probability and the proof is complete.