COMMUN. STATIST.—SIMULA., 23(1), 1-16 (1994)

A GENERALIZED STEEL PROCEDURE FOR COMPARING SEVERAL TREATMENTS WITH A CONTROL UNDER RANDOM RIGHT-CENSORSHIP

Yuh-Ing Chen

Graduate Institute of Statistics, National Central University
Chung-Li, Taiwan 32054

Key Words and Phrases: Monte Carlo study; one-way layout; random right-censorship; treatments verus control procedure.

ABSTRACT

Following Gehan (1965) and Breslow (1970), a generalization of Steel's (1959) test for comparing several treatments with a control in a one-way layout when observations are subject to the same pattern of random right-censorship is proposed. The proposed test is constructed mainly for testing against simple-tree alternatives. However, based on the test, a multiple testing procedure for deciding which treatments (if any) are better than the control is suggested. The relative level and power performances of the proposed testing procedure and the ones suggested respectively by Magel (1988) and Chakraborti (1990) are examined in a Monte Carlo study.

1. INTRODUCTION

A variety of nonparametric procedures have been developed for comparing several treatments with a control in a one-way layout with complete observations. In particular, Steel (1959) suggested a multiple comparison rank sum test based on pairwise rankings for comparing several treatments with a control. Slivka (1970) extended the two-sample control median test proposed by Mathisen (1943) to the case of several treatments with a control. Fligner and Wolfe (1982) further proposed an extension of the two-sample Mann-Whitney (1949) test, by considering the control group as one sample and all treatment groups as the other sample, to the treatments versus control setting.

In a clinical trial or life-testing experiment for survival analysis, however, subjects who randomly enter the experiment at different times may be lost to follow-up randomly or, owing to time limitation, the experiment may be terminated at a preassigned time. In these cases only randomly right-censored data are available. Since, there are some practical situations of clinical trials where the assumption of equal censorship is tenable, Magel (1989) generalized the Fligner-Wolfe (1982) test based on Gehan's (1965) scores for the setting where observations are subject to the same pattern of random right-censorship. For the same setting, to terminate the study as early as possible when the cost of the experiment is high, Chakraborti (1990) suggested a generalization of Slivka's test. Chakraborti and Desu (1991) further considered a class of linear rank tests for comparing several treatments with a control when data are subject to different censoring patterns.

In section 2 we describe the treatments versus control setting with randomly right-censored data under consideration in this paper and discuss previously proposed testing procedures. In section 3 we propose a generalization of Steel's ptest when observations are subject to an arbitrary right censorship. A multiple testing procedure for deciding which treatments (if any) are better than the control is also suggested. In section 4 an illustrative example of studying the effect of various doses of Red Dye No. 40 on the development of reticuloendothelial tumours is provided. In section 5 we present the results of a Monte Carlo simulation investigation of the relative level and power performances of these competing testing procedures for a variety of treatment effects configurations.

2. THE SETTING, NOTATION, AND PREVIOUS WORK

For the ith sample (i=0, 1, ..., k), let $T_{i1}, \ldots, T_{in_i}$ be independent and identically distributed (i.i.d.) random variables each with a continuous distribution function F_i, and $C_{i1}, \ldots, C_{in_i}$ be i.i.d. random variables each with a continuous distribution function G_i, where C_{is} is the censoring time associated with the life time T_{is}. Suppose that the zero population (i=0) is the control and the other k populations are treatments. Furthermore, assume that the k+1 samples are independent of each other and the C_{is}'s are distributed independently of the T_{is}'s. In such a setting, we often only observe the bivariate vectors (X_{is}, δ_{is}), where $X_{is}=\min(T_{is}, C_{is})$, $\delta_{is}=1$, if $X_{is} = T_{is}$, and 0 otherwise. In this paper, specifically, we are concerned with testing the null hypothesis H_0: $[F_i = F_o, i=1, ..., k]$ against the simple-tree alternative hypothesis H_A: $[F_i < F_o$ for at least one i$]$ when $G_o = G_1 = ... = G_k$. The problem of estimating the treatment i for which $F_i < F_o$ is also considered.

For the two-sample problem with censored data, Gehan (1965) defined the statistics

$$a_{st}^{ij} = \begin{cases} +1 & \text{if } X_{is} < X_{jt};\ \delta_{is} = 1 \\ -1 & \text{if } X_{is} > X_{jt};\ \delta_{jt} = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

and
$$U_{ij} = \sum_{s=1}^{n_i} \sum_{t=1}^{n_j} a_{st}^{ij}. \tag{2.2}$$

For testing H_0 against the simple-tree alternative H_A, Magel (1988) considered the statistic

$$W = \sum_{i=1}^{k} U_{oi} \tag{2.3}$$

along with its permutation variance estimate

$$\mathrm{Var}_o(W) = \frac{n_o(N-n_o)}{N(N-1)} \sum_{j=1}^{t} \{m_j M_{j-1} (M_j+1) + \iota_j M_j (M_j+1) + m_j (N- M_j- L_{j-1}) (N- 3M_{j-1}- m_j- L_{j-1}- 1)\},$$

where $N = \sum_{i=o}^{k} n_i$, $M_j = \sum_{i=1}^{j} m_i$, $M_o= 0$, $L_j = \sum_{i=1}^{j} \iota_i$, $L_o= 0$, m_i is the number of uncensored observations at rank i in the rank ordering of uncensored observations with distinct values, and ι_i is the number of right censored observations with values greater than observations at rank i but less than observations at rank i+1. Suppose that $N \to\infty$ in such a way that $n_o/N\to\lambda_o$, with $0< \lambda_o< 1$. Under the assumption of $G_i= G_o$ and $F_i= F_o$ for i=1, ..., k, the results of Gehan (1965) imply directly that

$W/\{Var_0(W)\}^{1/2}$ has an asymptotic ($N \to \infty$) standard normal distribution. Magel then obtained an approximate level α test for H_0

$$\text{reject } H_0 \text{ if } W/\{Var_0(W)\}^{1/2} \geq z(\alpha), \qquad (2.4)$$

where $z(\alpha)$ is the upper α percentile of a standard normal distribution.

When the cost of experimentation for survival analysis is high, the experimenter may want to terminate the experiment as soon as enough data become available to reach a decision. To this end, Chakraborti (1990) proposed to estimate first the median of the control population θ_0 through the linearized version of the Kaplan-Meier estimator of its distribution function $\hat{F}_0$ (see, for example, Brookmeyer and Crowley (1982)), that is, $\hat{\theta}_0 = \hat{F}_0^{-1}(1/2)$. For the case of $n_0 = n$ and $n_i = nc$, i=1, ..., k, let $V_i = nc[1/2 - \hat{F}_i(\hat{\theta}_0)]$, where the $\hat{F}_i$ is the linearized version of the Kaplan-Meier estimator of F_i, for i=1, ..., k. Suppose that $N \to \infty$ in such a way that $n/N \to \lambda$, with $0 < \lambda < 1$. Chakraborti pointed out that, for $G_i = G_0$ and $F_i = F_0$, i=1, ..., k, the asymptotic distribution of the random vector $N^{-1/2}(V_1, ..., V_k)$ is a k-dimensional normal distribution with mean zero vector and covariance matrix $\Sigma_0 = (\sigma_{0ij})$, which can be consistently estimated, based on observations in the control sample, by $\hat{\Sigma}_0 = (\hat{\sigma}_{0ij})$, where

$$\hat{\sigma}_{0ii} = c(c+1)(1+kc)^{-1}\hat{\beta},$$
$$\hat{\sigma}_{0ij} = c^2(1+kc)^{-1}\hat{\beta},$$
$$\hat{\beta} = (n/4)\,\Sigma\, d_{0j}\,[R_{0j}(R_{0j} - d_{0j})]^{-1},$$

d_{0j} is the number of uncensored observations at the jth distinct values denoted by X_{0j}, R_{0j} is the number of observations at risk at X_{0j}, and the summation is over all X_{0j} being less than or equal to $\hat{\theta}_0$. Therefore, Chakraborti suggested to reject H_0 if

$$\{c(c+1)(1+kc)^{-1}\hat{\beta}\,N\}^{-1/2}\max(V_1, ..., V_k) \geq g(\alpha; k, \rho), \qquad (2.5)$$

where $\rho = c(c+1)^{-1}$ and $g(\alpha; k, \rho)$ is the upper α percentile of the maximum of k equally correlated standard normal variates with common correlation ρ. Gupta (1963) has tabled $g(\alpha; k, \rho)$ for various values of ρ. Note that, in certain types of life-testing experiments, where observations become available in a naturally sequential (time ordered) manner, the experiment can be terminated and Chakraborti's test can be applied as soon as the median of the control sample is observed.

For testing against the simple-tree alternative when the data are subject to unequal patterns of censorship, Chakraborti and Desu (1991) further considered a class of linear rank tests of the form

$$T(\omega) = \sum_{i=1}^{k} \omega_i U_{oi},$$

where the U_{oi}'s are given in equation (2.2) and the $\omega = (\omega_1, ..., \omega_k)^t$ is a vector of nonnegative weights. Note that when the ω_i are all 1, the statistic is, in fact, the one proposed by Magel (1988) as stated in equation (2.3). Let, for i=0, 1, ..., k,

$$Q_i(t) = \Pr(X_{i1} \le t) = 1 - [1 - F_i(t)][1 - G_i(t)]$$

and

$$Q_{iu}(t) = \Pr(X_{i1} \le t, \delta_{i1} = 1) = \int_{-\infty}^{t} [1 - G_i(s)])\, dF_i(s).$$

Suppose that $N \to \infty$ in such a way that $n_i/N \to \lambda_i$, with $0 < \lambda_i < 1$, i=0, 1, ..., k. Under the assumption of $F_o = F_1 = ... = F_k$, Chakraborti and Desu proved that the asymptotic ($N \to \infty$) distribution of the random vector $N^{-3/2}(U_{o1}, ..., U_{ok})$ is a k-dimensional normal distribution with mean zero vector and covariance matrix $\Sigma_1 = (\sigma_{1ij})$, where

$$\sigma_{1ii} = \lambda_o \lambda_i \int_o^\infty [1 - Q_o(t)]\, [1 - Q_i(t)]\, d[\lambda_o Q_{ou}(t) + \lambda_i Q_{iu}(t)]$$

and

$$\sigma_{1ij} = (1/2)\, \lambda_0 \lambda_i \lambda_j \{ \int_o^\infty [1 - Q_o(t)]\, [1 - Q_i(t)]\, dQ_{ju}(t) + \int_o^\infty [1 - Q_o(t)]\, [1 - Q_j(t)]\, dQ_{iu}(t) \}.$$

By replacing the λ_i, $Q_i(t)$ and $Q_{iu}(t)$ with their empirical versions, namely,

$$\hat{\lambda}_i = \frac{n_i}{N},$$

$$\hat{Q}_i(t) = \frac{1}{n_i} \sum_{s=1}^{n_i} I\, (X_{is} \le t)$$

and

$$\hat{Q}_{iu}(t) = \frac{1}{n_i} \sum_{s=1}^{n_i} I\, (X_{is} \le t, \delta_{is} = 1),$$

they obtained a consistent estimator of Σ_1, denoted by $\hat{\Sigma}_1$. Therefore, they proposed to reject H_o if

$$\{N^3 \omega^t \hat{\Sigma}_1 \omega\}^{-1/2}\, T(\omega) \ge z(\alpha), \tag{2.6}$$

where, again, $z(\alpha)$ is the upper α percentile of a standard normal distribution. Moreover, to estimate the treatment in which $F_i < F_o$, they proposed a conservative procedure based on the Slepian inequality (see, for example, Koziol and Reid (1977)). They then suggested to

$$\text{claim } F_i < F_o \text{ if } Z_i = \{N^3 \hat{\sigma}_{1ii}\}^{-1/2}\, U_{oi} \ge z(b) \text{ for } i=1, ..., k,$$

where $\alpha = 1 - (1-b)^k$. Note that the Chakraborti-Desu procedure can be used for the more general case of unequal patterns of censorship and the pairwise follow-up test is convenient to use since the required critical values come from the standard normal tables.

3. THE GENERALIZED STEEL PROCEDURE

To generalize Steel's (1959) test for censored data, we consider, in this section, the random vector $(U_{01}, ..., U_{ok})$, where the U_{oi}'s are given in equation (2.2). We obtain, directly from the results in Chakraborti and Desu (1991), that, under the assumption of $G_o = G_1 = ... = G_k$, the asymptotic $(N \to \infty)$ null (H_0) distribution of the random vector $N^{-3/2}(U_{01}, ..., U_{ok})$ is a k-dimensional normal distribution with mean zero vector and covariance matrix $\Sigma = (\sigma_{ij})$, where

$$\sigma_{ii} = \lambda_o \lambda_i (\lambda_o + \lambda_i) \tau$$

$$\sigma_{ij} = \lambda_o \lambda_i \lambda_j \tau$$

and
$$\tau = \int_o^\infty [1 - Q(t)]^2 \, dQ_u(t),$$

with $Q(t) = \sum_{i=o}^{k} \lambda_i Q_i(t)$ and $Q_u(t) = \sum_{i=o}^{k} \lambda_i Q_{iu}(t)$.

In the following we base on all observations to find consistent estimators of the σ_{ij}'s. By replacing λ_i, $Q(t)$ and $Q_u(t)$, respectively, with their empirical versions $\hat{\lambda}_i = n_i / N$, $\hat{Q}(t) = \sum_{i=o}^{k} \sum_{s=1}^{n_i} I\,(X_{is} \le t)/ N$ and $\hat{Q}_u(t) = \sum_{i=o}^{k} \sum_{s=1}^{n_i} I\,(X_{is} \le t, \delta_{is} = 1)/ N$, a consistent estimator of τ is given by

$$\hat{\tau} = N^{-3} \Sigma\Sigma\, d_{is} R_{is} (R_{is} - d_{is}), \tag{3.1}$$

where d_{is} is the number of uncensored observations at the jth distinct values denoted by X_{is}, R_{is} is the number of observations at risk at X_{is}, and the summation is over all X_{is} in the k+1 samples combined. Consistent estimators of the σ_{ij}'s are then obtained as

$$\hat{\sigma}_{ii} = \hat{\lambda}_o \hat{\lambda}_i (\hat{\lambda}_o + \hat{\lambda}_i)\, \hat{\tau} \text{ and } \hat{\sigma}_{ij} = \hat{\lambda}_o \hat{\lambda}_i \hat{\lambda}_j\, \hat{\tau}.$$

Let, for i=1, ..., k,

$$U^*_{oi} = [n_o n_i (n_o + n_i)\, \hat{\tau}]^{-1/2}\, U_{oi}.$$

We observe, by applying Slutsky's theorem, that, under the assumption of $G_o = G_1 = ... = G_k$, the asymptotic $(N \to \infty)$ null (H_0) distribution of the random vector $(U^*_{o1}, ..., U^*_{ok})$ is a k-dimensional normal distribution with mean zero vector and covariance matrix $\Sigma^* = (\sigma^*_{ij})$, where

$$\sigma^*_{ii} = 1 \text{ and } \sigma^*_{ij} = \{\lambda_i\lambda_j / [(\lambda_o+\lambda_i)(\lambda_o+\lambda_j)]\}^{1/2}.$$

It can be seen that, for the special case of $n_o = n$ and $n_i = nc$, $i=1, ..., k$, σ^*_{ij} is $\rho = c(c+1)^{-1}$. Therefore, we propose to reject H_o in favor of the simple-tree alternative H_A if

$$S_{max} = \max(U^*_{o1}, ..., U^*_{ok}) \geq g(\alpha; k, \rho), \quad (3.2)$$

where $g(\alpha; k, \rho)$ is given in (2.5).

Note that $\sigma^*_{ij} = b_i b_j$, where $b_i = [\lambda_i/(\lambda_o+\lambda_i)]^{1/2}$ and $b_j = [\lambda_j/(\lambda_o+\lambda_j)]^{1/2}$. We can use the computer program developed by Dunnett (1989) to evaluate the joint probability of a multivariate normally distributed random vector with mean zero vector and such a special form of covariance matrix. Thus the approximate p-value of the test based on S_{max} can then be obtained even when sample sizes are different. Therefore, the test is in fact applicable in the general case of unequal sample sizes.

If the test based on S_{max} rejects H_o, one would wish to determine which treatments are more effective than the control. For the case of $n_o = n$ and $n_i = nc$, $i=1, ..., k$, we suggest to

$$\text{claim } F_i < F_o \text{ if } U^*_{oi} \geq g(\alpha; k, \rho) \text{ for } i=1, ..., k.$$

It is obvious that the experimentwise error rate for this procedure is approximately controlled since, under the assumption of $G_i = G_o$ for $i=1, ..., k$,

$$\alpha \cong \Pr\{\max(U^*_{o1}, ..., U^*_{ok}) \geq g(\alpha; k, \rho) | H_o\}$$

$$\geq \Pr\{ U^*_{oi} \geq g(\alpha; k, \rho) \text{ for at least one } i=1, ..., k | H_o\}.$$

4. AN EXAMPLE

To determine whether FD&C Red No. 40, Red 40, a colour additive widely used in foods in the U.S., has any effect on the development of reticuloendothelial, RE, tumours, which can be detected only at death, a lifetime feeding experiment involving mice was undertaken (Lagakos and Mosteller (1981)). It was generally believed that RE tumours kill their mouse hosts shortly after onset and hence the time to RE death approximates time to RE tumour onset. The observations (fictional data is used) given in the following are the time to death of 40 mice receiving various doses of Red 40, and the time to death of those mice without RE tumours is treated as the censoring variable.

TABLE I

Time to death of mice receiving Red Dye No. 40

Zero-Dose Control	Low Dosage	Medium Dosage	High Dosage
70	59	30	34
77	70	37	$\underline{36}$
$\underline{83}$	73	56	45
87	77	$\underline{65}$	$\underline{48}$
92	$\underline{80}$	76	$\underline{65}$
92	84	83	90
93	87	$\underline{87}$	$\underline{91}$
$\underline{96}$	$\underline{90}$	$\underline{90}$	$\underline{92}$
100	91	$\underline{95}$	$\underline{95}$
$\underline{102}$	$\underline{95}$	97	98

Note: The time to death of the mice with RE tumours is underlined.

Since the higher the dose of Red 40 applied, the shorter the time to RE death will be, we calculate the U_{io} in the following:

$$U_{1o}= 8+ 6+ 3- 5= 12,$$

$$U_{2o}= 10+ 6+ 6+3 - 4- 1= 20$$

and

$$U_{3o}= 10+ 10+ 10+ 6+ 4- 5- 1= 34.$$

The Magel's statistic is $W=\sum_{i=1}^{3} U_{io}= 66$. After computing the m_j, M_j, ι_j and L_j, the permutation variance estimate of W is obtained as $Var_o(W)= 1{,}339.42$ and then $W/\{Var_o(W)\}^{1/2}= 1.80$. Therefore, we observe, from a standard normal table, that the p-value of Magel's test is about 0.0359. To calculate Chakraborti's statistic, we need to estimate the median of the control population θ_0 through the linearized version of the Kaplan-Meier estimator of its distribution function $\hat{F}_0$. Since $\hat{F}_0(96)= 0.417$ and $\hat{F}_0(102)= 1.000$, the estimated median of the control population is $\hat{\theta}_0= 96.86$. The statistics $V_i= n[\hat{F}_i(\hat{\theta}_0)-1/2]$ are then obtained as $V_1= 5.00$, $V_2= 2.86$ and $V_3 =3.41$. Based on the control sample only we have $\hat{\beta}= \frac{10}{4}(\frac{1}{8*7} + \frac{1}{3*2})= 0.461$. Hence, Chakraborti's test statistic is obained as 1.65 and the corresponding p-value is greater than 0.10. Finally, to compute the generalized Steel test proposed in this paper, we need to estimate the parameter τ. After calculating the d_{is} and R_{is} based on all observations, we have $\hat{\tau}= 0.107$ and the statistics $U^*_{io}= [2n^3 \hat{\tau}]^{-1/2} U_{io}$ are then obtained as $U^*_{1o}= 0.82$, $U^*_{2o}= 1.37$ and $U^*_{3o}= 2.32$. Therefore, the value of S_{max} is 2.32 and its p-value is about 0.0252. It is clearly that both the test based on S_{max} and Magel's test indicate that, comparing to the zero-dose control, Red 40 has a significant effect on the development of RE tumours, while Chakraborti's test does not. Futhermore, since $U^*_{3o}= 2.32 > g(0.05; 3, 0.5) = 2.064$, but both the U^*_{1o} and U^*_{2o} are less than 2.064, we conclude, at the 5% significance level, that only the high dosage of Red 40 has more effect on the development of RE tumours than does the zero-dose control.

5. MONTE CARLO STUDY

5.1 DISCUSSION OF STUDY

To examine the relative level and power performances of Magel's test in (2.4), Chakraborti's test in (2.5) and the generalized Steel's test in (3.2) for comparing several treatments with a control when observations are subject to the same pattern of random right-censorship, we conducted a Monte Carlo study. We considered k=3 treatments with sample sizes $n_o= n_1= n_2= n_3= n= 10$, 20 and 30 in the level study and n= 20 and 30 in the power study.

Exponential and Weibull distributions were considered as life time distributions for their wide application in survival analysis. The uniform distribution over (0, R)

TABLE II

Estimated levels for $\alpha = 0.05$, $n_0 = n_1 = n_2 = n_3 = n$ and uniform censoring distribution U(0, R)

(a) Exponential

n	R	S_{max}	Chakraborti	Magel
10	9.9995	0.056	0.041	0.048
	4.9651	0.052	0.039	0.051
	3.1971	0.050	0.035	0.050
20	9.9995	0.052	0.048	0.048
	4.9651	0.052	0.043	0.045
	3.1971	0.053	0.041	0.049
30	9.9995	0.046	0.045	0.046
	4.9651	0.047	0.042	0.046
	3.1971	0.048	0.039	0.046

$f_i(x) = \exp(-x)$

(b) Weibull

n	R	S_{max}	Chakraborti	Magel
10	12.5331	0.054	0.038	0.050
	6.2666	0.054	0.029	0.051
	4.1777	0.055	0.025	0.049
20	12.5331	0.053	0.045	0.045
	6.2666	0.054	0.040	0.046
	4.1777	0.054	0.037	0.047
30	12.5331	0.051	0.043	0.046
	6.2666	0.044	0.039	0.044
	4.1777	0.045	0.037	0.046

$f_i(x) = (2x) \exp(-x^2)$

was used as the censoring distribution. In the level study, the standard exponential distribution and the Weibull distribution with shape parameter 2 and scale parameter 1 were considered. In the power study, we used exponential distributions with various values of location or scale parameters and Weibull distributions with shape parameter 2 but scale parameters varied. To investigate the effect of different

TABLE III
Estimated powers for $\alpha = 0.05$, exponential life-time distribution and uniform censoring distribution U(0, R)

(a) $n_o = n_1 = n_2 = n_3 = 20$

ε_o	ε_1	ε_2	ε_3	R	S_{max}	Chakraborti	Magel
0	0	0	0.25	9.9995	0.215	0.189	0.137
				4.9651	0.222	0.174	0.142
				3.1971	0.233	0.166	0.146
0	0	0.25	0.25	9.9995	0.315	0.278	0.291
				4.9651	0.327	0.258	0.296
				3.1971	0.351	0.250	0.314
0	0	0.25	0.50	9.9995	0.588	0.520	0.422
				4.9651	0.607	0.492	0.437
				3.1971	0.635	0.467	0.451
0	0.25	0.25	0.25	9.9995	0.381	0.322	0.474
				4.9651	0.401	0.300	0.502
				3.1971	0.440	0.290	0.535
0	0.25	0.25	0.50	9.9995	0.611	0.602	0.624
				4.9651	0.639	0.583	0.649
				3.1971	0.679	0.565	0.681

(b) $n_o = n_1 = n_2 = n_3 = 30$

ε_o	ε_1	ε_2	ε_3	R	S_{max}	Chakraborti	Magel
0	0	0	0.25	9.9995	0.295	0.231	0.156
				4.9651	0.313	0.218	0.161
				3.1971	0.329	0.212	0.168
0	0	0.25	0.25	9.9995	0.407	0.320	0.351
				4.9651	0.436	0.307	0.372
				3.1971	0.465	0.304	0.397
0	0	0.25	0.50	9.9995	0.769	0.679	0.532
				4.9651	0.794	0.655	0.559
				3.1971	0.815	0.635	0.576
0	0.25	0.25	0.25	9.9995	0.486	0.380	0.597
				4.9651	0.522	0.364	0.633
				3.1971	0.560	0.356	0.667
0	0.25	0.25	0.50	9.9995	0.764	0.696	0.765
				4.9651	0.793	0.675	0.786
				3.1971	0.823	0.653	0.818

$f_i(x; \varepsilon_i) = \exp\{-(x - \varepsilon_i)\}$

TABLE IV

Estimated powers for $\alpha = 0.05$, exponential life-time distribution and uniform censoring distribution U(0, R)

(a) $n_o = n_1 = n_2 = n_3 = 20$

γ_o	γ_1	γ_2	γ_3	R	S_{max}	Chakraborti	Magel
1	1	1	2	9.9995	0.360	0.201	0.165
				4.9651	0.319	0.177	0.158
				3.1971	0.274	0.163	0.151
1	1	2	2	9.9995	0.544	0.325	0.400
				4.9651	0.493	0.287	0.366
				3.1971	0.437	0.267	0.332
1	1	2	3	9.9995	0.766	0.485	0.539
				4.9651	0.689	0.444	0.482
				3.1971	0.614	0.401	0.440
1	2	2	2	9.9995	0.653	0.404	0.697
				4.9651	0.598	0.359	0.643
				3.1971	0.554	0.334	0.594
1	2	2	3	9.9995	0.810	0.537	0.815
				4.9651	0.748	0.490	0.765
				3.1971	0.689	0.449	0.706

(b) $n_o = n_1 = n_2 = n_3 = 30$

γ_o	γ_1	γ_2	γ_3	R	S_{max}	Chakraborti	Magel
1	1	1	2	9.9995	0.504	0.302	0.196
				4.9651	0.445	0.276	0.186
				3.1971	0.385	0.262	0.167
1	1	2	2	9.9995	0.697	0.447	0.521
				4.9651	0.630	0.417	0.476
				3.1971	0.566	0.399	0.425
1	1	2	3	9.9995	0.907	0.685	0.698
				4.9651	0.854	0.652	0.629
				3.1971	0.784	0.616	0.568
1	2	2	2	9.9995	0.794	0.545	0.848
				4.9651	0.743	0.514	0.775
				3.1971	0.687	0.488	0.742
1	2	2	3	9.9995	0.936	0.730	0.933
				4.9651	0.892	0.696	0.896
				3.1971	0.838	0.661	0.851

$f_i(x; \gamma_i) = (1/\gamma_i) \exp(-x/\gamma_i)$

TABLE V

Estimated powers for $\alpha = 0.05$, Weibull life-time distribution and uniform censoring distribution U(0, R)

(a) $n_o = n_1 = n_2 = n_3 = 20$

γ_o	γ_1	γ_2	γ_3	R	S_{max}	Chakraborti	Magel
1	1	1	1.5	12.5331	0.469	0.237	0.192
				6.2666	0.411	0.200	0.181
				4.1777	0.346	0.185	0.160
1	1	1.5	1.5	12.5331	0.665	0.384	0.490
				6.2666	0.604	0.335	0.443
				4.1777	0.534	0.308	0.394
1	1	1.5	2	12.5331	0.911	0.618	0.683
				6.2666	0.859	0.562	0.616
				4.1777	0.779	0.509	0.538
1	1.5	1.5	1.5	12.5331	0.764	0.475	0.814
				6.2666	0.718	0.422	0.758
				4.1777	0.657	0.392	0.691
1	1.5	1.5	2	12.5331	0.931	0.665	0.925
				6.2666	0.893	0.609	0.884
				4.1777	0.830	0.563	0.820

(b) $n_o = n_1 = n_2 = n_3 = 30$

γ_o	γ_1	γ_2	γ_3	R	S_{max}	Chakraborti	Magel
1	1	1	1.5	12.5331	0.694	0.378	0.240
				6.2666	0.573	0.344	0.219
				4.1777	0.496	0.322	0.190
1	1	1.5	1.5	12.5331	0.826	0.553	0.642
				6.2666	0.764	0.512	0.578
				4.1777	0.694	0.482	0.504
1	1	1.5	2	12.5331	0.985	0.844	0.842
				6.2666	0.967	0.812	0.778
				4.1777	0.925	0.787	0.690
1	1.5	1.5	1.5	12.5331	0.904	0.659	0.933
				6.2666	0.860	0.620	0.894
				4.1777	0.804	0.582	0.843
1	1.5	1.5	2	12.5331	0.991	0.872	0.988
				6.2666	0.979	0.842	0.973
				4.1777	0.951	0.816	0.942

$f_i(x; \gamma_i) = (2x/\gamma_i^2) \exp\{-(x/\gamma_i)^2\}$

degrees of censorship on the performance of a test, we considered several different values of R which correspond to the probability of censorship p as 0.10, 0.20 and 0.30 in the level study. For example, when life time distribution is the standard exponential distribution and p=0.3, R is 3.1971. For Weibull distribution with shape parameter 2 and scale parameter 1, R is 4.1777 corresponding to p=0.3. Note that these uniform distributions were then employed as censoring distributions in the power study.

For each of these settings, appropriate uniform, exponential and Weibull variates were generated by using the IMSL routines RNUN, RNEXP and RNWIB. In each case we used 5,000 replications to obtain the estimated error rate or power under the nominal level $\alpha = 0.05$. Therefore, the maximum standard error for the estimator is about 0.007. (In fact, we are guaranteed a standard error no greater than 0.003 for estimating the error rate.) The estimated error rates are presented in Table II and power estimates are reported in Tables III, IV and V. The designated treatment effects configurations correspond to values of $\gamma_0, \gamma_1, \gamma_2$ and γ_3 ($\varepsilon_0, \varepsilon_1, \varepsilon_2$ and ε_3), where the γ_i (ε_i) are the scale (location) parameters of the life time distributions.

5.2 DISCUSSION OF RESULTS

It is evident, upon examination of Table II, that the proposed test and Magel's test hold their levels quite well across all situations considered in this paper, while the level performance of Chakraborti's test depends heavily on the probability of censorship. In fact, Chakraborti's test holds its level only for the cases of light censoring as p=0.1 and large sample sizes about 20.

The power study presented in Tables III, IV and V shows that the proposed test is in general superior to Chakraborti's test for comparing several treatments with a control. Magel's test, which is a sum of score test, provides a better test than does the proposed test when the treatments are equally yet more effective than the control. However, the proposed test based on S_{max} is seen to be more powerful when there is at least one treatment equally effective as the control and the rest are more effective than the control. For situations in between, the two tests appear to perform rather similarly. Note that these results also point out the fact that the problem of a choice of a particular weighting function that combines the k two-sample score statistics into an overall test statistic is an important and interesting one. This issue has been partly addressed in Chakraborti and Desu (1991).

In comparing several treatments with a control, experimenters are usually more interested in deciding which treatments (if any) are better than the control. In such cases, the generalized Steel procedure considered in this paper is useful when the treatment groups are subject to an identical pattern of random right-censorship. If data are subject to unequal patterns of censorship, however, experimenters may use the pairwise follow-up test proposed by Chakraborti and Desu (1991).

ACKNOWLEDGEMENTS

This work was supported in part by National Science Council of Taiwan. The author wishes to thank the associate editor and referee for valuable comments and suggestions which substantially improve the presentation of this paper.

BIBLIOGRAPHY

Breslow, N. (1970). A generalized Kruskal-Wallis test for comparing k samples subject to unequal patterns of censorship. Biometrika, 57, 579-594.

Brookmeyer, R. and Crowley, J. (1982). A k-sample median test for censored data. Journal of the American Statistical Association. 77, 433-440.

Chakraborti, S. (1990). A one-sided test of homogeneity against simple tree alternative for right-censored data. Communications in Statistics: Simulation and Computation, 19, 879-889.

Chakraborti, S. and Desu, M. M. (1991). Linear rank tests for comparing treatments with a control when data are subject to unequal patterns of censorship. Statistica Neerlandica, 45, 227-254.

Dunnett, C. (1989). Multivariate normal probability integrals with product correlation structure, Algorithm AS251, Applied Statistics, 38, 564-579.

Fligner, M. A. and Wolfe, D. A. (1982). Distribution-free tests for comparing several treatments with a control. Statistica Neerlandica, 36, 119-127.

Gehan, E. A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika, 52, 203-223.

Koziol, J. A. and Reid, N. (1977). On multiple comparisons among K sample subject to unequal patterns of censorship. Communications in Statistics: Theory and Methods, 6, 1149-1164.

Lagakos, S. W. and Mosteller, F. (1981). A case study of statistics in the regulatory process: The FD&C Red No. 40 experiments. Journal of National Cancer Institution, 66, 197-212.

Magel, R. (1988). A test for the equality of k medians against the simple tree alternative under right censorship. Communications in Statistics: Simulation and Computation, 17, 917-925.

Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. Annals of Mathematical Statistics, 18, 50-60.

Mathisen, H. C. (1943). A method for testing the hypothesis that two samples are from the same population. Annals of Mathematical Statistics, 14, 188-194.

Slivka, J. (1970). A one-sided nonparametric multiple comparison control percentile test: treatments versus control. Biometrika, 57, 431-438.

Received August 1992; Revised June 1993.