

AD-A077 439

SOUTH CAROLINA UNIV COLUMBIA DEPT OF MATHEMATICS COM--ETC F/G 12/1
CONDITIONALLY DISTRIBUTION-FREE TEST FOR CENSORED BIVARIATE OBS--ETC(U)
JUL 79 W J PADGETT , L J WEI F49620-79-C-0140

UNCLASSIFIED

TR-41

AFOSR-TR-79-1037

NL

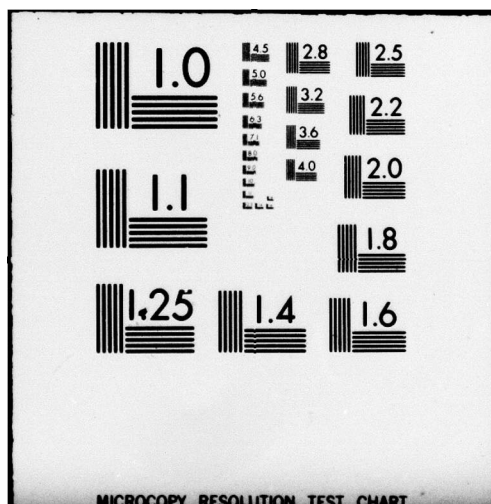
| OF |
ADA
077439



END
DATE
FILMED

4 -80

DDC



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. AEROSPACE REPORT NUMBER 18 AFOSR/TR-79-1037	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) CONDITIONALLY DISTRIBUTION-FREE TEST FOR CENSORED BIVARIATE OBSERVATIONS	5. TYPE OF REPORT & PERIOD COVERED Interim rept.		
6. AUTHOR(s) W. J. Padgett and L. J. Wei	7. PERFORMING ORG. REPORT NUMBER		
8. CONTRACT OR GRANT NUMBER(s) TR-41	9. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0140		
10. PERFORMING ORGANIZATION NAME AND ADDRESS University of South Carolina Department of Mathematics, Computer Science, & Statistics Columbia, S.C. 29208	11. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 16 61102F/2304/A5		
12. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, D.C. 20332	13. REPORT DATE July 1979		
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. NUMBER OF PAGES 10		
	16. SECURITY CLASS. (of this report) UNCLASSIFIED		
	17. DECLASSIFICATION/DOWNGRADING SCHEDULE		
18. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited			
19. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) DDC			
20. SUPPLEMENTARY NOTES RECEIVED NOV-30 1979 A			
21. KEY WORDS (Continue on reverse side if necessary and identify by block number) Marshall and Olkin bivariate exponential model; Right censorship; Permutation test.			
22. ABSTRACT (Continue on reverse side if necessary and identify by block number) A conditionally distribution-free test is proposed for testing the symmetry of a bivariate distribution function with observations which are subject to arbitrary right censorship. In a numerical study, this new test is shown to be more powerful than the sign test under Marshall and Olkin's bivariate exponential model.			

DD FORM 1473 1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD A 077439

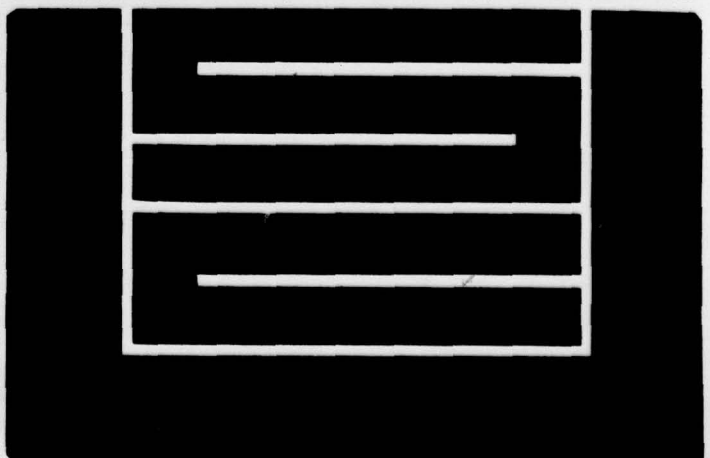
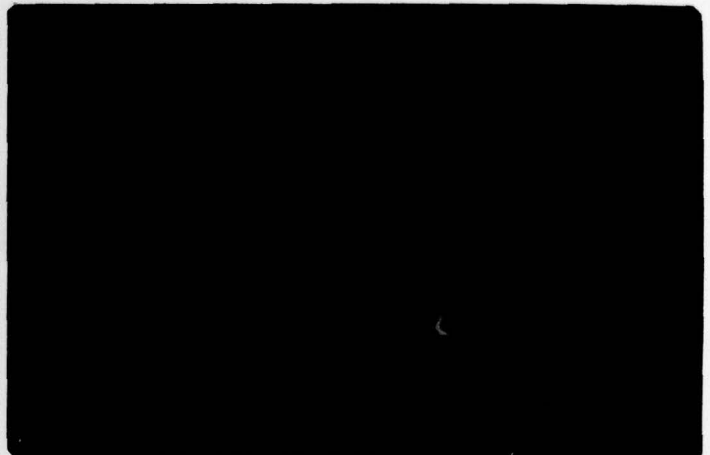
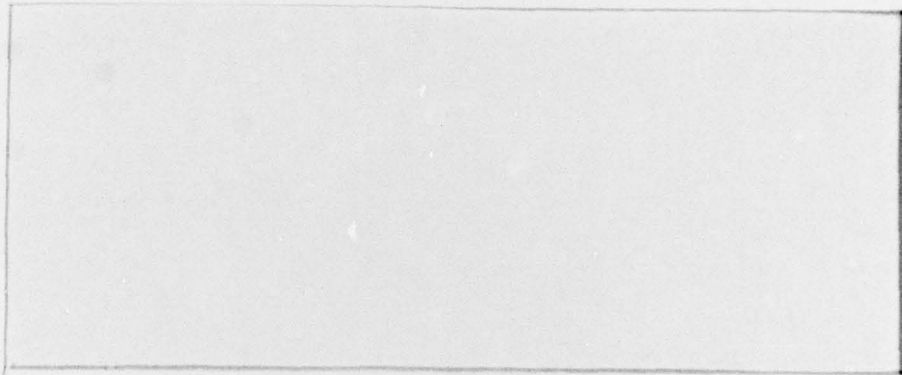
DDC FILE COPY

410 422

J03

AFOSR-TR- 79 - 1037

Department of Mathematics,
Computer Science and Statistics
The University of South Carolina
Columbia, South Carolina 29208



**CONDITIONALLY DISTRIBUTION-FREE TEST FOR
CENSORED BIVARIATE OBSERVATIONS***

by

W. J. Padgett and L. J. Wei
University of South Carolina
Statistics Technical Report No. 41
~~62C10~~

July, 1979

Department of Mathematics, Computer Science, and Statistics
University of South Carolina
Columbia, South Carolina 29208

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AF3C)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

* Research supported by the United States Air Force Office of Scientific
Research under Contract No. F49620-79-C-0140.

CONDITIONALLY DISTRIBUTION-FREE TEST FOR
CENSORED BIVARIATE OBSERVATIONS

W. J. Padgett and L. J. Wei

Department of Mathematics, Computer Science,
and Statistics

University of South Carolina

A conditionally distribution-free test is proposed for testing the symmetry of a bivariate distribution function with observations which are subject to arbitrary right censorship. In a numerical study, this new test is shown to be more powerful than the sign test under Marshall and Olkin's (1967) bivariate exponential model.

Accession for	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification	
By	
Distribution	
Availability Codes	
Dist	Avail and/or special
A	

Key words and phrases: Right censorship; Marshall and Olkin bivariate exponential model; permutation test.

1. INTRODUCTION

In many situations, the comparisons between two treatments based on paired observations which are censored in one or both variates may arise. For example, Hammond (1964) has used a matched pair analysis to study smoking in relation to mortality in the United States. Batchelor and Hackett (1970) gave a comparison of the survival times between HL-A closely matched and poorly matched skin allografts on the same badly burned patients. Also, in life testing it may be desirable to compare the life times of two components in a system.

Suppose that

$$(X_1^0, Y_1^0)', (X_2^0, Y_2^0)', \dots, (X_n^0, Y_n^0)' \quad (1.1)$$

are independent, identically distributed random vectors, having $H^0(s, t)$ as their distribution function (d.f.) and having $F^0(s)$ and $G^0(t)$ as their marginal d.f.'s, respectively, where ' denotes vector transpose. The null hypothesis, which is to be tested is

$$H_0: H^0(s, t) = H^0(t, s), \text{ for } (s, t)' \in R^2.$$

Since X_1^0 and Y_j^0 may be censored from the right by variables U_1 and V_j , respectively, (1.1) cannot always be observed. The observations available to the experimenter actually consist of the minima

$$\begin{aligned} X_1 &= \min(X_1^0, U_1), \dots, X_n = \min(X_n^0, U_n), \\ Y_1 &= \min(Y_1^0, V_1), \dots, Y_n = \min(Y_n^0, V_n), \end{aligned} \quad (1.2)$$

and two random sequences $\{\delta_1, \dots, \delta_n\}$ and $\{\epsilon_1, \dots, \epsilon_n\}$, where

$$\delta_i = \begin{cases} 1, & \text{if } x_i = x_i^0, \\ 0, & \text{if } x_i < x_i^0, \end{cases}$$

$$\epsilon_j = \begin{cases} 1, & \text{if } y_j = y_j^0, \\ 0, & \text{if } y_j < y_j^0. \end{cases}$$

For $i \neq j$, the censoring variables U_i and V_j are assumed to be independent random variables with a common d.f. J . To have the same censoring mechanism for both variates is quite common in paired studies. It is also assumed that $(U_i, V_i)'$ and $(X_i^0, Y_i^0)'$ are independent, $i = 1, 2, \dots, n$.

In the parametric case, the procedure for testing H_0 is rather complicated and no useful results have been obtained. However, Holt and Prentice (1974) used the proportional hazards model (Cox (1972)) to analyze the data by Batchelor and Hackett (1970), and Wei (1979) proposed an asymptotically distribution-free test for H_0 based on paired observations which are subject to arbitrary right censorship. Since the sample size in paired studies is frequently small, a distribution-free test is highly desirable. Although the sign test is a conditionally distribution-free test for testing H_0 , it is rather inefficient when there are too many censored pairs in the data. As an extreme case, for the data $(3^+, 4)'$, $(6, 5^+)'$, $(2^+, 4)'$, $(9, 7^+)'$, $(8^+, 6^+)'$, where "+" denotes censoring, the sign test leads to no conclusion about the null hypothesis H_0 .

In this article, a conditionally distribution-free test for H_0 is presented in Section 2. In a numerical study, it is shown that the new test

is more powerful than the sign test under Marshall and Olkin's bivariate exponential model (Marshall and Olkin (1967)).

We note that all the results of this article can be easily extended to the case of arbitrarily restricted observations (Mantel (1967)).

2. THE TEST STATISTIC

Let $Z_i = X_i$ and $Z_{n+i} = Y_i$, $i = 1, \dots, n$. We will say that Z_i is definitely greater than Z_j if $Z_i > Z_j$ and Z_j is observed, and Z_i is definitely less than Z_j if $Z_i < Z_j$ and Z_i is observed. Now, let $\xi_i(\eta_i)$, $i = 1, 2, \dots, n$, be the number of the remaining $(2n-1)$ Z 's than which $X_i(Y_i)$ is definitely greater minus the number than which it is definitely less. The original observations (1.2) are then replaced by $(\xi_1, \eta_1)', \dots, (\xi_n, \eta_n)'$. Under H_0 and the assumption of an equal censoring mechanism for both variates, all the arrangements of the form

$$(R_1, R_2)', \dots, (R_{2n-1}, R_{2n})'$$

are equally likely, where $(R_{2i-1}, R_{2i}) = (\xi_i, \eta_i)$ or (η_i, ξ_i) , $i = 1, \dots, n$.

The statistic proposed here for testing H_0 is

$$W = \sum_{i=1}^n R_{2i-1}.$$

Small or large values of $\sum_{i=1}^n \xi_i$ lead to the rejection of H_0 . Note that $\sum_{i=1}^n \xi_i$ is Gehan's (1965) two-sample statistic. Mantel (1967) gave a simple routine to calculate ξ_i and η_i . It should also be noted that scores other than Gehan's (ξ_i, η_i) can be utilized.

The drawback to any permutation test such as W is the usually long and tedious calculations required when the sample size n is large. Fortunately, an asymptotically distribution-free test is available for large sample cases (Wei (1979)). In the rest of this article, we concentrate on the small-sample performance of the W test.

3. THE POWER STUDY

In this section, we study a special alternative hypothesis H_1 : $F^0(s) \leq G^0(s)$ for all s and $F^0(s') < G^0(s')$ for some s' , and compare the W test with the sign test under Marshall and Olkin's bivariate exponential model. The survival function of this model with parameters λ_1 , λ_2 , and λ_{12} can be written as:

$$P\{X^0 \geq s, Y^0 \geq t\} = \exp[-\lambda_1 s - \lambda_2 t - \lambda_{12} \max(s, t)] . \quad (3.1)$$

The two marginal means are

$$EX^0 = 1/(\lambda_1 + \lambda_{12}) \quad \text{and} \quad EY^0 = 1/(\lambda_2 + \lambda_{12}) .$$

Under this model, the hypotheses to be tested become $H_0 : \lambda_1 = \lambda_2$ against $H_1 : \lambda_1 < \lambda_2$.

Three censoring schemes are considered in this comparison:

- (A) $J(s)$ is a uniform distribution over $(0, EX^0)$;
- (B) $J(s)$ is a uniform distribution over $(0, 2EX^0)$; and
- (C) $J(s)$ is a uniform distribution over $(0, 4EX^0)$.

In this numerical study, the censoring variables U_1 and V_1 are assumed to be independent.

For the sign test and the W test, the proportions of times in the 1,000 Monte Carlo samples generated that H_0 was rejected at the $\alpha = .05$ level were calculated for samples of sizes $n = 10$ and 15 from (3.1) with various values of λ_1 , λ_2 , and λ_{12} . Tables 1 and 2 give the results. As we expected, the W test is uniformly more powerful than the sign test. In addition to the drawback which was illustrated by an example in Section 1, another disadvantage of the sign test is the actual probability of Type I error is far below the specified α value. For example, when $n = 10$, under the severe censorship (A), the empirical levels of the sign test are only .01 as compared with the nominal value $\alpha = .05$.

Table 1. Proportion of times in 1,000 samples of size $n = 10$ that H_0 was rejected at $\alpha = .05$ by sign test and W test.

		Censoring Scheme					
		(A)		(B)		(C)	
λ_{12}	$\lambda_1 \lambda_2$	W	Sign	W	Sign	W	Sign
.1	.1 1.0	.85	.60	.91	.73	.92	.76
	.2	.58	.34	.69	.45	.74	.49
	.3	.38	.17	.46	.24	.51	.29
	.4	.25	.09	.32	.14	.36	.18
	.5	.21	.06	.25	.11	.26	.14
	.6	.15	.05	.18	.07	.19	.09
	.7	.09	.02	.11	.05	.13	.06
	.8	.06	.02	.07	.03	.08	.04
	.9	.05	.02	.06	.02	.05	.03
	1.0	.04	.01	.04	.02	.04	.03
.5	.1 1.0	.44	.16	.57	.34	.64	.47
	.2	.33	.09	.42	.20	.48	.30
	.3	.23	.06	.28	.13	.34	.20
	.4	.18	.04	.21	.09	.25	.13
	.5	.15	.02	.18	.06	.19	.09
	.6	.12	.01	.13	.04	.14	.05
	.7	.07	.01	.09	.03	.10	.04
	.8	.06	.01	.07	.02	.07	.03
	.9	.06	.01	.05	.02	.06	.02
	1.0	.04	.01	.05	.01	.04	.02

Table 2. Proportion of times in 1,000 samples of size $n = 15$ that H_0 was rejected at $\alpha = .05$ by sign test and W test.

		Censoring Scheme					
		(A)		(B)		(C)	
λ_{12}	$\lambda_1 \lambda_2$	W	Sign	W	Sign	W	Sign
.1	.2 1.0	.79	.55	.86	.66	.88	.74
	.4	.38	.18	.47	.25	.51	.32
	.6	.17	.06	.21	.10	.23	.13
	.8	.09	.03	.10	.04	.11	.05
	1.0	.04	.01	.04	.02	.04	.02
.5	.2 1.0	.48	.25	.60	.42	.67	.50
	.4	.27	.10	.32	.17	.37	.22
	.6	.13	.04	.16	.06	.19	.09
	.8	.07	.01	.09	.03	.09	.04
	1.0	.05	.01	.05	.02	.05	.03

REFERENCES

- Batchelor, J. R. and Hackett, M. (1970). HL-A matching in treatment of burned patients with skin allografts. Lancet 2, 581-3.
- Cox, D. R. (1972). Regression models and life tables (with Discussion). J. R. Statist. Soc. 34B, 187-220.
- Gehan, E. (1965). A generalized Wilcoxon test for comparing arbitrarily single censored samples. Biometrika 52, 203-23.
- Hammond, E. C. (1964). Smoking in relation to mortality and morbidity. Findings in first thirty-four months of follow-up in a prospective study started in 1959. J. Nat. Cancer Inst. 32, 1161-88.
- Holt, J. D. and Prentice, R. L. (1974). Survival analyses in twin studies and matched pair experiments. Biometrika 61, 17-30.
- Mantel, N. (1967). Ranking procedures for arbitrarily restricted observations. Biometrics 23, 65-78.
- Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. J. Am. Statist. Assoc. 62, 30-44.
- Wei, L. J. (1979). A generalized Gehan and Gilbert test for paired observations which are subject to arbitrary right censorship. T.R. Department of Math., Computer Science, and Statistics, University of South Carolina.