Journal of the Korean
Statistical Society
Vol. 23, No. 2, 1994

# Testing for Failure Rate Ordering between Survival Distributions

**ChulGyu Park**[1]

**ABSTRACT**

We develop in this paper the likelihood ratio test(LRT) for testing $H_1 : F_1 \preceq F_2$ against $H_2 - H_1$ where $H_2$ imposes no restriction on $F_1$ and $F_2$ and '$\preceq$' means failure rate ordering. Both one and two-sample problems will be considered. In the one-sample case, one of the two distributions is known, while we assume in the other case both are unknown. We derive the asymptotic null distribution of the LRT statistic which will be of chi-bar-square type. The main issue here is to determine the least favorable distribution which is stochastically largest within the class of null distributions.

**KEYWORDS:** Failure rate ordering, Likelihood ratio test, Chi-bar-square distribution, Level probability, Least favorable distribution.

[1]Department of Applied Statistics, The University of Suwon, Suwon, 445–743, Korea.

# 1. INTRODUCTION

Failure rate ordering between two survival distributions has been widely studied in statistical inference, reliability theory and actuarial science. When we compare the survival times of two populations(or groups), a typical measure for comparison is mean survival time. Another type of comparison is based upon the ordinary stochastic ordering which is defined by the inequality of distribution functions($F_X(z) \leq F_Y(z), z \in R^1$). This is a stronger concept in the sense that stochastic ordering implies mean value ordering. More stringent comparison will be based upon failure rate ordering which is stronger than ordinary stochastic ordering.

Let $Z$ be a continuous random variable with its probability density function $f$. The failure rate of $Z$ is defined as $\gamma(t) = f(t)/\bar{F}(t)$ for $t \in (-\infty, F^{-1}(1))$ where $F^{-1}(1) = \sup\{t : F(t) < 1\}$ and $\bar{F}(t) = 1 - F(t)$. The discrete version of this with finite support $S = \{t_1, t_2, \ldots, t_k\}$ will be $\gamma(t_j) = p_j / \sum_{i=j}^{k} p_i, t_j \in S$ where $-\infty < t_1 < t_2 < \cdots < t_k < \infty$. Suppose that $\gamma_1$ and $\gamma_2$ are the failure rate functions of random variables $Z_1$ and $Z_2$ with their CDF's $F_1$ and $F_2$ respectively. Let the binary relation '$\preceq$' denote failure rate ordering between $Z_1$ and $Z_2$. That is, the relation $Z_1 \preceq Z_2(\text{or} F_1 \preceq F_2)$ is equivalent to $\gamma_1(t) \geq \gamma_2(t)$ for any $t$.

Testing problems related to failure rate ordering are discussed in many papers. Kochar(1979, 1981) proposes a generalized $U$-statistic and a linear rank statistic for testing the equality of failure rates against ordered failure rates in two populations. He shows that these tests achieve high Pitman asymptotic relative efficiency over a broad spectrum of alternatives. For the same problem, Cheng(1985) develops another rank test which is asymptotically unbiased and consistent. He demonstrates by Monte Carlo study that his test is more powerful than Kochar's rank test(1981) and Savage's $U$-test(1956) when sample sizes are small. Joe and Proschan(1984) and Aly(1988) note that the failure rate ordering ($\gamma_F \leq \gamma_G$) can be expressed as the inequality between 100 percentile residual life functions($q_{\alpha,F} \leq q_{\alpha,G}$). Depending upon the value of $\alpha$, they suggest a class of rank tests which contains Kochar's(1981) as a special case. Bagai and Kochar(1986) introduce a measure of deviation between two distributions $F$ and $G$ which is given by $\Delta_k(F,G) = \int\int_{x>y} \delta(x,y)dF(x)dG(y)$ where $\delta(x,y) = [\bar{F}(x)/\bar{F}(y)]^k - [\bar{G}(x)/\bar{G}(y)]^k$. Based upon $\Delta_k(F,G)$, they develop a class of distribution-free tests and show that the test with $k = 0.6$ has good overall performance.

Likelihood ratio principle can be also applied to the similar problems. Dykstra, Kochar and Robertson(1990) consider the competing risks model and test the proportionality of two cause specific hazard rates against the monotonicity in their ratio. For grouped data, they derive a chi-bar-square type of asymptotic distribution for the LRT statistic while a test based upon $U$-statistic is proposed in the continuous case. In another joint work by Dykstra, Kochar and Robertson(1991), nonparametric maximum likelihood estimates of $N$ survival functions are obtained under uniform stochastic ordering constraints and used to test equality of distributions against uniform stochastic ordering alternative.

As reviewed earlier, most of previous studies are concentrated on testing equality of distributions under the presumption that they are ordered in terms of failure rate($\gamma_1 \geq \gamma_2$). However, testing procedure for this presumption has never been suggested. In this context, we consider failure rate ordering between two distributions as the null hypothesis. First, we discuss the consistency of the maximum likelihood estimates of the distributions under failure rate ordering. And then, we will determine the least favorable distribution in the class of the asymptotic null distributions of the LRT statistic. Both of one and two-sample problems will be solved.

## 2. CONSISTENCY OF MLE

Let $S = \{1, 2, \ldots, k\}$ be the common support of two discrete random variables which follow distributions $F_1$ and $F_2$ respectively. Based upon a random sample, $X_{i1}, X_{i2}, \ldots, X_{in}$, from the distribution $F_i$, we define $n_{ij}$ and $m_{ij}$ as follows : $n_{ij} = \#\{l : X_{il} = j\}$ and $m_{ij} = \sum_{l=j}^{k} n_{il}$, $j \in S, i = 1, 2$. Then, the random vector $(n_{i1}, n_{i2}, \ldots, n_{ik})$ follows a multinomial distribution with parameters $n_i, p_{i1}, p_{i2}, \ldots, p_{ik}$ where $p_{ij} = Pr[X_i = j]$. In the case of one-sample problem, we assume that $F_2$ is known and the sample is taken from $F_1$ only.

### 2.1. One-sample case.

Given a random sample of size $n_1$ from a unknown distribution $F_1$, we will find the maximum likelihood estimate of $F_1$ under $F_1 \preceq F_2$(or equivalently $\gamma_1 \geq \gamma_2$), where $F_2$ is known.

The likelihood function under the previous sampling scheme is given by

$$L(F_1) = \prod_{j=1}^{k} \left[\bar{F}_1(j-1) - \bar{F}_1(j)\right]^{n_{1j}}. \tag{2.1}$$

From the definition, the failure rate ordering constraints are

$$\frac{p_{1j}}{\sum\limits_{l=j}^{k} p_{1l}} \geq \frac{p_{2j}}{\sum\limits_{l=j}^{k} p_{2l}}, \quad j = 1, 2, \ldots, k-1. \tag{2.2}$$

Since $p_{ij} / \sum_{l=j}^{k} p_{il} = 1 - \sum_{l=j+1}^{k} p_{il} / \sum_{l=j}^{k} p_{il}$, the constraints in (2.2) are equivalent to

$$\theta_{1j} \leq \theta_{2j}, \quad j = 1, 2, \ldots, k-1, \tag{2.3}$$

where $\theta_{ij} = \bar{F}_i(j)/\bar{F}_i(j-1)$. In this one-sample problem, we assume that $\theta_{2j}$'s are known values.

Noting that $\bar{F}_i(0) = 1$, we can express $\bar{F}_i(j)$ as $\bar{F}_i(j) = \prod_{r=1}^{j} \theta_{ir}$. Thus, the likelihood function (2.1) becomes

$$\mathcal{L}(\underline{\theta}_1) = \sum_{j=1}^{k-1} \theta_{1j}^{m_{1j} - n_{1j}} (1 - \theta_{1j})^{n_{1j}}, \tag{2.4}$$

where $m_{1j} = \sum_{r=j}^{k} n_{1r}$.

By maximizing the log-likelihood function, we get the unrestricted MLE of $\theta_{1j}$, i.e.,

$$\hat{\theta}_{1j} = \frac{m_{1j} - n_{1j}}{m_{1j}}, \quad j = 1, 2, \ldots, k-1. \tag{2.5}$$

Since $k-1$ constraints in (2.3) are working independently, the restricted MLE of $\theta_{1j}$ is obtained by maximizing $\theta_{1j}^{m_{1j}-n_{1j}}(1-\theta_{1j})^{n_{1j}}$ individually under a single constraint $\theta_{1j} \leq \theta_{2j}$. Therefore, the MLE of $\theta_{1j}$ under (2.3) is

$$\theta_{1j}^{*} = \min(\hat{\theta}_{1j}, \theta_{2j}). \tag{2.6}$$

The expression in (2.6) can be denoted by $P(\hat{\theta}_{1j}|\jmath_j)$ which represents the closest point of the set $\jmath_j = \{x \in R^1 : x \leq \theta_{2j}\}$ to $\hat{\theta}_{1j}$. By the invariance property of MLE, the MLE of $\bar{F}_1$ under $F_1 \preceq F_2$ is expressed as

$$\bar{F}_1^{*}(x) = \prod_{\{j: j \leq x\}} \theta_{1j}^{*}. \tag{2.7}$$

It can be shown from the strong law of large numbers that both restricted and unrestricted MLE's of $\theta_{1j}$ are strongly consistent when $F_1 \preceq F_2$. Based upon this almost sure convergence of the estimates, we can establish the following theorem.

**Theorem 2.1.** Suppose $F_1$ and $F_2$ are discrete CDF's with common finite support $S$. Let $F_1^{*}$ be the MLE of $F_1$ under failure rate ordering $F_1 \preceq F_2$ with $F_2$ known. Then, if $F_1 \preceq F_2, F_1^{*}$ converges uniformly to $F_1$ as the sample size $n_1$ goes to infinity.

### 2.2. Two-sample case.

Assuming that distributions $F_1$ and $F_2$ are both unknown, we will find their MLE's under $F_1 \preceq F_2$ and also discuss the consistency of those MLE's when the ordering is true. Consider random samples of sizes $n_1$ and $n_2$ respectively from $F_1$ and $F_2$ with common support $S$. Define $n_{ij}$ as the number of the observations in the $i$th sample which are equal to $j$. Then, the likelihood function is

$$L(\underset{\sim}{F}) = \prod_{i=1}^{2} \prod_{j=1}^{k} \left[\bar{F}_i(j-1) - \bar{F}_i(j)\right]^{n_{ij}}. \tag{2.8}$$

As we did in the previous section, we reparameterize by setting $\theta_{ij} = \frac{\bar{F}_i(j)}{\bar{F}_i(j-1)}$ for any $i$ and $j$. With this reparameterization, the likelihood function is expressed as

$$\begin{aligned}\mathcal{L}(\theta) &= \prod_{i=1}^{2}\Big[(1-\theta_{i1})^{n_{i1}}\prod_{j=2}^{k}\Big(\prod_{r=1}^{j-1}\theta_{ir}-\prod_{r=1}^{j}\theta_{ir}\Big)^{n_{ij}}\Big]\\ &= \prod_{i=1}^{2}\Big[(1-\theta_{i1})^{n_{i1}}\prod_{j=2}^{k}\Big\{\Big(\prod_{r=1}^{j-1}\theta_{ir}\Big)^{n_{ij}}(1-\theta_{ij})^{n_{ij}}\Big\}\Big].\end{aligned}$$

Now,

$$\prod_{j=2}^{k}\prod_{r=1}^{j-1}\theta_{ir}^{n_{ij}} = \prod_{r=1}^{k-1}\prod_{j=r+1}^{k}\theta_{ir}^{n_{ij}} = \prod_{r=1}^{k-1}\theta_{ir}^{m_{ir}-n_{ir}},$$

where $m_{ir} = \sum_{l=r}^{k} n_{il}$. Thus, we can rewrite the likelihood function as

$$\mathcal{L}(\theta) = \prod_{j=1}^{k-1}\Big[\prod_{i=1}^{2}\theta_{ij}^{m_{ij}-n_{ij}}(1-\theta_{ij})^{n_{ij}}\Big]. \tag{2.9}$$

As was shown earlier, the failure rate ordering constraints are equivalent to

$$\theta_{1j} \le \theta_{2j}, \quad j = 1, 2, \ldots, k-1, \tag{2.10}$$

where $\theta_{1j}$ and $\theta_{2j}$ are both unknown.

The unrestricted MLE's of $\theta_{ij}$'s are easily obtained by maximizing the log-likelihood function and these are

$$\hat{\theta}_{ij} = \frac{m_{ij}-n_{ij}}{m_{ij}}, \quad 1, 2, \ldots, k-1 \text{ and } i = 1, 2. \tag{2.11}$$

As discussed in Dykstra et al(1991), the restricted MLE's of $\theta_{ij}$'s are obtained simply by maximizing the $j$th term $\prod_{i=1}^{2}\theta_{ij}^{m_{ij}-n_{ij}}(1-\theta_{ij})^{n_{ij}}$ under the $j$-th constraint $\theta_{1j} \le \theta_{2j}$ because the constraints in (2.10) work independently. For each $j$, this is a bioassay problem discussed in Example 1.5.1 of Robertson, Wright and Dykstra(1988). Thus, the solution $(\theta^*_{1j}, \theta^*_{2j})$ is the isotonic regression of the unrestricted MLE $(\hat{\theta}_{1j}, \hat{\theta}_{2j})$ with weights $(m_{1j}, m_{2j})$. That is,

$$\theta^*_{ij} = \begin{cases} \hat{\theta}_{ij}, & \text{if } \hat{\theta}_{1j} \le \hat{\theta}_{2j} \\ \frac{m_{1j}\hat{\theta}_{1j}+m_{2j}\hat{\theta}2j}{m_{1j}+m_{2j}}, & \text{if } \hat{\theta}_{1j} > \hat{\theta}_{2j} \end{cases}, \quad j = 1, 2, \ldots, k-1. \tag{2.12}$$

Let $\underset{\sim}{\theta}_j = (\theta_{1j}, \theta_{2j})'$ and $\underset{\sim}{m}_j = (m_{1j}, m_{2j})'$. Then, $\underset{\sim}{\theta}_j^*$ can be also denoted by $P_{\underset{\sim}{m}_j}(\hat{\underset{\sim}{\theta}}_j | \jmath_j)$ which is the closest point of $\jmath_j = \{\underset{\sim}{x} \in R^2 : x_1 \leq x_2\}$ to $\hat{\underset{\sim}{\theta}}_j$ in the sense of the weighted Euclidian distance. These restricted and unrestricted estimates of $\theta_{ij}$'s are strongly consistent if the failure rate ordering is true. Let $\underset{\sim}{n} \to \infty$ mean that sample sizes $n_1$ and $n_2$ go to infinity such that $n_2/n_1 \to \gamma(> 0)$. Then, using the strong consistency of those MLE's under $F_1 \preceq F_2$, we can prove the following theorem.

**Theorem 2.2.** Let $F_i^*$ be the MLE of CDF $F_i$ under the failure rate ordering $F_1 \preceq F_2$ where $F_1$ and $F_2$ are both unknown. Then, if $F_1 \preceq F_2$, $F_i^*$ converges uniformly to $F_i$ as $\underset{\sim}{n} \to \infty$.

## 3. HYPOTHESIS TESTING

### 3.1. One-sample problem.

Assuming that $F_2$ is known, we consider in this section a problem of testing $H_1 : F_1 \preceq F_2$ against $H_2 - H_1$ where $H_2$ puts no restriction on $F_1$. If we set $\theta_{ij} = \bar{F}_i(j)/\bar{F}_i(j-1)$, we can rewrite the hypotheses as

$$\begin{cases} H_1 : & \theta_{1j} \leq \theta_{2j}, \quad j = 1, 2, \ldots, k-1 \\ H_2 : & \text{No restriction on } \underset{\sim}{\theta}_i = (\theta_{11}, \theta_{12}, \ldots, \theta_{1,k-1})' \end{cases}. \tag{3.1}$$

Under this reparameterization, the likelihood function becomes

$$\mathcal{L}(\underset{\sim}{\theta}_1) = \prod_{j=1}^{k-1} \theta_{1j}^{m_{1j}-n_{1j}} (1 - \theta_{1j})^{n_{1j}}. \tag{3.2}$$

Using the MLE's of $\theta_{1j}$ in (2.5) and (2.6), we can construct a test based upon the likelihood ratio, $\Lambda_{12} = \mathcal{L}(\underset{\sim}{\theta}_1^*)/\mathcal{L}(\hat{\underset{\sim}{\theta}}_1)$, which rejects the null hypothesis $H_1$ for the large values of $T_{12} = -2 \ln \Lambda_{12}$. This LRT statistic is

$$T_{12} = -2 \sum_{j=1}^{k-1} \Big[ (m_{1j} - n_{1j})(\ln \theta_{1j}^* - \in \hat{\theta}_{1j}) + n_{1j} \Big\{ \ln(1 - \theta_{1j}^*) - \ln(1 - \hat{\theta}_{1j}) \Big\} \Big]. \tag{3.3}$$

Expanding $\ln \theta^*_{ij}$ about $\hat{\theta}_{1j}$, we have

$$\ln \theta^*_{1j} = \ln \hat{\theta}_{1j} + \frac{1}{\hat{\theta}_{1j}}(\theta^*_{1j} - \hat{\theta}_{1j}) - \frac{1}{2\alpha^2_{1j}}(\theta^*_{1j} - \hat{\theta}_{1j})^2, \tag{3.4}$$

where $\alpha_{1j}$ is a value between $\hat{\theta}_{1j}$ and $\theta^*_{1j}$.
Similarly, we can express $\ln(1-\theta^*_{1j})$ as

$$\ln(1-\theta^*_{1j}) = \ln(1-\hat{\theta}_{1j}) + \frac{1}{(1-\hat{\theta}_{1j})}(\hat{\theta}_{1j} - \theta^*_{1j}) - \frac{1}{2\beta^2_{1j}}(\hat{\theta}_{1j} - \theta^*_{1j})^2, \tag{3.5}$$

where $\beta_{1j}$ is a value between $1-\hat{\theta}_{1j}$ and $1-\theta^*_{1j}$.

Putting (3.4) and (3.5) into (3.3), we have

$$\begin{aligned} T_{12} &= -2\sum_{j=1}^{k-1}\Big[(m_{1j}-n_{1j})\Big\{\frac{1}{\hat{\theta}_{1j}}(\theta^*_{1j}-\hat{\theta}_{1j}) - \frac{1}{2\alpha^2_{1j}}(\theta^*_{1j}-\hat{\theta}_{1j})^2\Big\} \\ &\quad + n_{1j}\Big\{\frac{1}{(1-\hat{\theta}_{1j})}(\hat{\theta}_{1j}-\theta^*_{1j}) - \frac{1}{2\beta^2_{1j}}(\hat{\theta}_{1j}-\theta^*_{1j})^2\Big\}\Big]. \end{aligned}$$

Since $\hat{\theta}_{1j} = (m_{1j}-n_{1j})/m_{1j}$, the linear terms are canceled. That is,

$$\begin{aligned} &(m_{1j}-n_{1j})(\theta^*_{1j}-\hat{\theta}_{1j})/\hat{\theta}_{1j} + n_{1j}(\hat{\theta}_{1j}-\theta^*_{1j})/(1-\hat{\theta}_{1j}) \\ &= m_{1j}(\theta^*_{1j}-\hat{\theta}_{1j}) + m_{1j}(\hat{\theta}_{1j}-\theta^*_{1j}) = 0. \end{aligned}$$

Therefore, the LRT statistic becomes

$$T_{12} = \sum_{j=1}^{k-1}\Big[(\theta^*_{1j}-\hat{\theta}_{1j})^2\Big(\frac{m_{1j}-n_{1j}}{\alpha^2_{1j}} + \frac{n_{1j}}{\beta^2_{1j}}\Big)\Big]. \tag{3.6}$$

Let $T^{(j)}_{12} = (\theta^*_{1j}-\hat{\theta}_{1j})^2\Big(\frac{m_{1j}-n_{1j}}{\alpha^2_{1j}} + \frac{n_{1j}}{\beta^2_{1j}}\Big)$, $j = 1, 2, \ldots, k-1$. Using $\hat{\theta}_{1j} = (m_{1j} - n_{1j})/m_{1j}$, we can rewrite $T^{(j)}_{12}$ as

$$T^{(j)}_{12} = n_1(\theta^*_{1j}-\hat{\theta}_{1j})^2\Big(\frac{m_{1j}}{n_1}\Big)\Big(\frac{\hat{\theta}_{1j}}{\alpha^2_{1j}} + \frac{1-\hat{\theta}_{1j}}{\beta^2_{1j}}\Big). \tag{3.7}$$

Now, we introduce a new random variable, $\mathring{T}_{12}^{(j)}$, given by

$$\mathring{T}_{12}^{(j)} = n_1 \left[P(\hat{\theta}_{1j}|\jmath_j) - \hat{\theta}_{1j}\right]^2 \frac{\bar{F}_1(j-1)}{\theta_{1j}(1-\theta_{1j})}, \tag{3.8}$$

where $P(\cdot|\jmath_j)$ is the least square projection operator onto $\jmath_j = \{x \in R^1 : x \leq \theta_{2j}\}$. Since $m_{1j}/n_1 \to \bar{F}_1(j-1)$, $\alpha_{1j} \to \theta_{1j}$ and $\beta_{1j} \to 1 - \theta_{1j}$ almost surely under $H_1$ , it follows from Slutsky's theorem that $T_{12}^{(j)}$ and $\mathring{T}_{12}^{(j)}$ have the same asymptotic distribution under $H_1$.

Let $\underset{\sim}{p_1} = (p_{11}, p_{12}, \ldots, p_{1k})'$ where $p_{ij} = \Pr[X_1 = j]$. The unrestricted MLE of $\underset{\sim}{p_1}$ is $\hat{\underset{\sim}{p_1}} = (\hat{p}_{11}, \hat{p}_{12}, \ldots, \hat{p}_{1k})'$ where $\hat{p}_{1j} = n_{1j}/n_1$, $j = 1, 2, \ldots, k$. By multivariate central limit theorem, $\sqrt{n_1}(\hat{\underset{\sim}{p_1}} - \underset{\sim}{p_1})$ converges weakly to a multivariate normal random vector which has mean vector $\underset{\sim}{0}$ and covariance matrix $D_{\underset{\sim}{p_1}} - \underset{\sim}{p_1}\underset{\sim}{p_1}'$ with $D_{\underset{\sim}{p_1}} = diag\{p_{11}, p_{12}, \ldots, p_{1k}\}$. Recall that $\theta_{1j} = \bar{F}_1(j)/\bar{F}_1(j-1), j = 1, 2, \ldots, k-1$. Applying the multivariate $\delta$-method, we can show that $\sqrt{n_1}(\hat{\underset{\sim}{\theta_i}} - \underset{\sim}{\theta_i})$ converges in distribution to a random vector $\underset{\sim}{V}$ which has a multivariate normal distribution $N_{k-1}(\underset{\sim}{0}, \Sigma)$ where the $(i,j)$th element of $\Sigma$ is

$$\sigma_{ij} = \begin{cases} p_{11}(1-p_{11}), & i = j = 1 \\ p_{1j}(1 - \sum_{l=1}^{j} p_{1l})/(1 - \sum_{l=1}^{j-1} p_{1l})^3, & 2 \leq i = j \leq k-1 \\ 0, & i \neq j \end{cases} .$$

Therefore, $\sqrt{n_1}(\hat{\theta}_{1j} - \theta_{1j})$, $j = 1, 2, \ldots, k-1$, are asymptotically independent and normally distributed with mean 0. By converting $p_{1j}$'s into $\theta_{1j}$'s, the variance of the $j$th variable is expressed as $Var(V_j) = \theta_{1j}(1-\theta_{1j})/\prod_{r=1}^{j-1}\theta_{1r} = \theta_{1j}(1-\theta_{1j})/\bar{F}_1(j-1)$, $j = 1, 2, \ldots, k-1$. Since $\mathring{T}_{12}^{(j)}$ is a function of $\hat{\theta}_{1j}$ only, it follows that $T_{12}^{(j)}$, $j = 1, 2, \ldots, k-1$, are also asymptotically independent. The following lemma will be used to find the asymptotic least favorable distribution of $T_{12}$. We should note that the test based upon the critical value from the stochastically largest distribution of $T_{12}$ in $H_1$ is the least favorable in the sense that it has the smallest power.

**Lemma 3.1.** Let $\Omega = \left\{\underset{\sim}{\theta}_1 \in R^{k-1} : \theta_{1j} \le \theta_{2j},\ j = 1, 2, \ldots, k-1\right\}$ where $\theta_{2j}$'s are known. Then, we have for $x > 0$

$$\sup_{\underset{\sim}{\theta}_1 \in \Omega} \lim_{n_1 \to \infty} P_{\underset{\sim}{\theta}_1}\left[T_{12}^{(j)} > x\right] = \frac{1}{2} \Pr\left[\chi_1^2 > x\right]. \tag{3.9}$$

This supremum is achieved when $\theta_{1j} = \theta_{2j}$.

**Proof.** First note that $T_{12}^{(j)}$ and $\mathring{T}_{12}^{(j)}$ have the same asymptotic distribution. Suppose $\theta_{1j} < \theta_{2j}$. Since $\hat{\theta}_{1j}$ converges to $\theta_{1j}$ almost surely, $\hat{\theta}_{1j} < \theta_{2j}$ eventually as $n_1 \to \infty$. Thus, for sufficiently large $n_1$ , $P(\hat{\theta}_{1j}|\jmath_j) = \hat{\theta}_{1j}$ and $\mathring{T}_{12}^{(j)} = 0$. And so, $\lim_{n_1 \to \infty} P_{\underset{\sim}{\theta}_1}\left[\mathring{T}_{12}^{(j)} \ge x\right] = 0$ for $x > 0$. Now, suppose $\theta_{1j} = \theta_{2j}$. Defining $\jmath_o = \{x \in R^1 : x \le 0\}$, we can express $\mathring{T}_{12}^{(j)}$ as

$$\begin{aligned}
\mathring{T}_{12}^{(j)} &= n_1\left[\left\{P(\hat{\theta}_{1j}|\jmath_j) - \theta_{1j}\right\} - (\hat{\theta}_{1j} - \theta_{1j})\right]^2 \frac{\bar{F}_1(j-1)}{\theta_{1j}(1-\theta_{1j})} \\
&= n_1\left[P(\hat{\theta}_{1j} - \theta_{1j}|\jmath_o) - (\hat{\theta}_{1j} - \theta_{1j})\right]^2 \frac{\bar{F}_1(j-1)}{\theta_{1j}(1-\theta_{1j})} \\
&= \left[P\left(\frac{\sqrt{n_1}(\hat{\theta}_{1j} - \theta_{1j})}{\sqrt{\theta_{1j}(1-\theta_{1j})/\bar{F}_1(j-1)}}\Big|\jmath_o\right) - \frac{\sqrt{n_1}(\hat{\theta}_{1j} - \theta_{1j})}{\sqrt{\theta_{1j}(1-\theta_{1j})/\bar{F}_1(j-1)}}\right]^2 \\
&\xrightarrow{L} \left[P(Z|\jmath_0) - Z\right]^2 \quad \text{where } Z \sim N(0,1).
\end{aligned}$$

Therefore, for $x > 0$

$$\begin{aligned}
\lim_{n_1 \to \infty} \Pr\left[\mathring{T}_{12}^{(j)} > x\right] &= \Pr\left[\left\{P(Z|\jmath_o) - Z\right\}^2 > x\}\right] \\
&= \Pr\left[Z^2 > x, Z > 0\right] \\
&= \frac{1}{2}\Pr\left[\chi_1^2 > x\right].
\end{aligned}$$

This completes the proof.

The following theorem provides the asymptotic least favorable distribution for the LRT statistic which is a chi-bar-square distribution with binomial level probabilities.

**Theorem 3.1.** Let $\Omega = \left\{\underset{\sim}{\theta}_1 \in R^{k-1} : \theta_{1j} \leq \theta_{2j},\ j = 1, 2, \ldots, k-1\right\}$ where $\theta_{2j}$'s are known. Then, for $x > 0$, we have

$$\sup_{\underset{\sim}{\theta}_1 \in \Omega} \lim_{n_1 \to \infty} P_{\underset{\sim}{\theta}_1}\left[T_{12} > x\right] = \sum_{l=1}^{k} \binom{k-1}{l-1} \left(\frac{1}{2}\right)^{k-1} \Pr\left[\chi^2_{k-l} \geq x\right] \tag{3.10}$$

and the supremum occurs when $\theta_{1j} = \theta_{2j}$, $j = 1, 2, \ldots, k-1$.

**Proof.** From Lemma 3.1, each $T_{12}^{(j)}$ becomes stochastically the largest in asymptotic sense when $\theta_{1j} = \theta_{1j}$. Since $T_{12} = \sum_{j=1}^{k-1} T_{12}^{(j)}$ , $T_{12}$ is stochastically maximized when $\underset{\sim}{\theta}_i \in \Omega_0$. Furthermore, Lemma 3.1 implies that $\lim_{n_1 \to \infty} P_{\underset{\sim}{\theta}_1 \in \Omega_o}\left[T_{12}^{(j)} = 0\right] = \frac{1}{2}$ because $T_{12}^{(j)} \geq 0$. Hence, the asymptotic moment generating function of $T_{12}^{(j)}$ under $\Omega_o$ is

$$\phi_j(t) = \frac{1}{2} + \frac{1}{2}(1-2t)^{-\frac{1}{2}}, \quad j = 1, 2, \ldots, k-1.$$

Since $T_{12}^{(j)}, j = 1, 2, \ldots, k-1$ are asymptotically independent, the asymptotic moment generating function of $T_{12}$ is expressed as

$$\begin{aligned}
\phi(t) &= [\frac{1}{2} + \frac{1}{2}(1-2t)^{-\frac{1}{2}}]^{k-1} \\
&= \sum_{l=0}^{k-1} \binom{k-1}{l} \left(\frac{1}{2}\right)^{k-1} (1-2t)^{-\frac{1}{2}(k-1-l)} \\
&= \sum_{l=1}^{k} \binom{k-1}{l-1} \left(\frac{1}{2}\right)^{k-1} (1-2t)^{-\frac{1}{2}(k-l)}.
\end{aligned}$$

The proof is completed.

### 3.2 Two-sample problem.

We assume in this case that $F_1$ and $F_2$ are both unknown, and consider the problem of testing $H_1 : F_1 \preceq F_2$ against $H_2 - H_1$ where $H_2$ imposes no restriction on $F_i$'s. Under the same reparameterization as in the previous section, the likelihood function is given by

$$\mathcal{L}(\theta) = \prod_{j=1}^{k-1}\Big[\prod_{i=1}^{2}\theta_{ij}^{m_{ij}-n_{ij}}(1-\theta_{ij})^{n_{ij}}\Big]. \tag{3.11}$$

We can express the hypotheses in terms of $\theta_{ij}$'s as follows:

$$\begin{cases} H_1: & \theta_{1j} \le \theta_{2j}, \quad j = 1, 2, \dots, k-1 \\ H_2: & \text{No restriction on } \theta \end{cases} . \tag{3.12}$$

The unrestricted and restricted maximum likelihood estimates of $\theta_{ij}$ are given in (2.11) and (2.12) respectively. Based upon these estimates, the LRT statistic is expressed as

$$T_{12} = -2\sum_{j=1}^{k-1}\sum_{i=1}^{2}\Big[(m_{ij}-n_{ij})(\ln\theta_{ij}^{*}-\ln\hat{\theta}_{ij})+n_{ij}\big\{\ln(1-\theta_{ij}^{*})-\ln(1-\hat{\theta}_{ij})\big\}\Big]. \tag{3.13}$$

Expanding $\ln\theta_{ij}^{*}$ and $\ln(1-\theta_{ij}^{*})$ about $\hat{\theta}_{ij}$ and $1-\hat{\theta}_{ij}$ yields

$$T_{12} = \sum_{j=1}^{k-1}\sum_{i=1}^{2}\Big[(\theta_{ij}^{*}-\hat{\theta}_{ij})^2\Big(\frac{m_{ij}-n_{ij}}{\alpha_{ij}^2}+\frac{n_{ij}}{\beta_{ij}^2}\Big)\Big],$$

where $\alpha_{ij}$ and $\beta_{ij}$ are the values between $\theta_{ij}^{*}$ and $\hat{\theta}_{ij}$ and $1-\theta_{ij}^{*}$ and $1-\hat{\theta}_{ij}$ respectively.

Let $T_{12}^{(j)}$ denote the $j$-th summation term in $T_{12}$. Then, we can rewrite $T_{12}^{(j)}$ as

$$T_{12}^{(j)} = \sum_{i=1}^{2} n_i\Big[(\theta_{ij}^{*}-\hat{\theta}_{ij})^2\Big(\frac{m_{ij}}{n_i}\Big)\Big(\frac{\hat{\theta}_{ij}}{\alpha_{ij}^2}+\frac{1-\hat{\theta}_{ij}}{\beta_{ij}^2}\Big)\Big]. \tag{3.14}$$

Define $\gamma_i = \lim_{n_1\to\infty}\frac{n_i}{n_1}(>0), i = 1, 2$ and let $\jmath = \big\{\underset{\sim}{x} \in R^2 : x_1 \le x_2\big\}$. When $\underset{\sim}{\theta}_j \in \jmath$, we can show by Slutsky's theorem that $T_{12}^{(j)}$ has the same asymptotic distribution as

$$\mathring{T}_{12}^{(j)} = \sum_{i=1}^{2} n_1\big[P_{\underset{\sim}{w}_j}(\hat{\underset{\sim}{\theta}}_j|\jmath)_i - \hat{\theta}_{ij}\big]^2\frac{\gamma_i\bar{F}_i(j-1)}{\theta_{ij}(1-\theta_{ij})}, \tag{3.15}$$

where $P_{\underset{\sim}{w}_j}(\hat{\underset{\sim}{\theta}}_j|\jmath)_i$ is the $i$th element of the closest point of $\jmath$ to $\hat{\underset{\sim}{\theta}}_j$ with weight vector $\underset{\sim}{w}_j$. Notice that the $i$th element of $\underset{\sim}{w}_j$ is $w_{ij} = \gamma_i \prod_{r=1}^{j-1} \theta_{ir} = \gamma_i \bar{F}_i(j-1)$. The next lemma provides the stochastically largest one among the asymptotic distributions of $T_{12}^{(j)}$ over the parameter set from the null hypothesis.

**Lemma 3.2.** Let $\Omega = \left\{\theta : \theta_{1j} \leq \theta_{2j}, j = 1, 2, \ldots, k-1\right\}$. Then, for $x > 0$, we have

$$\sup_{\theta \in \Omega} \lim_{\underset{\sim}{n} \to \infty} P_\theta\left[T_{12}^{(j)} > x\right] = \frac{1}{2} \Pr\left[\chi_1^2 > x\right]. \tag{3.16}$$

The supremum is achieved when $\theta_{1j} = \theta_{2j}$.

**Proof.** Suppose $\theta_{1j} < \theta_{2j}$. Since $\hat{\theta}_{1j}$ and $\hat{\theta}_{2j}$ converge almost surely to $\theta_{1j}$ and $\theta_{2j}$ respectively, $\hat{\theta}_{1j}$ becomes strictly smaller than $\hat{\theta}_{2j}$ eventually as $\underset{\sim}{n} \to \infty$. This implies that $P_{\underset{\sim}{w}_j}(\hat{\underset{\sim}{\theta}}_j|\jmath) = \hat{\underset{\sim}{\theta}}_j$ eventually. Hence, $T_{12}^{(j)} = 0$ almost surely as $\underset{\sim}{n} \to \infty$. Suppose that $\theta_{1j} = \theta_{2j}$. Let $V_{ij} = \frac{\sqrt{n_1}(\hat{\theta}_{ij} - \theta_{ij})}{\sqrt{\theta_{ij}(1-\theta_{ij})}}$. As we discussed in the paragraph before Lemma 3.1, it follows that $\underset{\sim}{V}_j = (V_{1j}, V_{2j})' \xrightarrow{L} \underset{\sim}{U}_j = (U_{1j}, U_{2j})'$ where $U_{ij}$ are independent and normally distributed with $E[U_{ij}] = 0$ and $Var[U_{ij}] = \frac{1}{\gamma_i \bar{F}_i(j-1)}$. Thus, $T_{12}^{(j)} \xrightarrow{L} \sum_{l=1}^{2}\left[P_{\underset{\sim}{w}_j}(\underset{\sim}{U}_j|\jmath)_i - U_{ij}\right]^2 w_{ij}$ as $\underset{\sim}{n} \to \infty$. Note that $P_{\underset{\sim}{w}_j}(\underset{\sim}{U}_j|\jmath)$ is equivalent to the isotonic regression of $\underset{\sim}{U}_j$ with weight vector $\underset{\sim}{w}_j$ as we discussed in Section 2. Applying the corollary of Theorem 2.3.1 in Robertson et al(1988) yields

$$\lim_{\underset{\sim}{n} \to \infty} P_\theta\left[T_{12}^{(j)} > x\right] = \sum_{l=1}^{2} p(l, 2 : \underset{\sim}{w}_j) \Pr\left[\chi_{2-l}^2 > x\right] \quad \text{for } x > 0. \tag{3.17}$$

where $p(j, 2 : \underset{\sim}{w}_j)$ is the probability that the elements of the isotonic regression have exactly $j$ distinct values. The level probabilities, $p(j, N : \underset{\sim}{w})$, are generally unknown. However, it was shown in Section 2.4 of Robertson et al(1988) that $p(l, 2 : \underset{\sim}{w}_j) = \frac{1}{2}$, $l = 1, 2$ regardless of weight vector $\underset{\sim}{w}_j$. Thus, the limiting tail probability in (3.17) becomes

$$\lim_{\underset{\sim}{n} \to \infty} P_\theta\left[T_{12}^{(j)} > x\right] = \frac{1}{2} \Pr\left[\chi_1^2 > x\right] \quad \text{for } x > 0. \tag{3.18}$$

The right hand side of (3.18) is positive and does not depend upon parameter $\theta$. Therefore, the limiting distribution of $T_{12}^{(j)}$ becomes stochastically largest when $\theta_{1j} = \theta_{2j}$.

This lemma has a key role in finding out the asymptotic least favorable distribution of $T_{12}$ in the null hypothesis $H_1$. Let $\Omega_o = \{\theta : \theta_{1j} = \theta_{2j},\ j = 1, 2, \ldots, k-1\}$. The following theorem claims that any parameter in $\Omega_o$ is the least favorable for the asymptotic of $T_{12}$. Surprisingly, the asymptotic least favorable distributions of the LRT statistics are the same in both one and two-sample problems.

**Theorem 3.2.** Let $\Omega = \{\theta : \theta_{1j} \le \theta_{2j},\ j = 1, 2, \ldots, k-1\}$. Then, for any $x > 0$, we have

$$\sup_{\theta \in \Omega} \lim_{\underset{\sim}{n} \to \infty} P_\theta\left[T_{12} > x\right] = \sum_{l=1}^{k} \binom{k-1}{l-1} \left(\frac{1}{2}\right)^{k-1} \Pr\left[\chi^2_{k-l} > x\right] \tag{3.19}$$

and this supremum occurs when $\theta \in \Omega_o$.

**Proof.** Suppose $\theta \in \Omega_o$. Then, $\theta_{1j} = \theta_{2j}$, $j = 1, 2, \ldots, k-1$. From Lemma 3.2, each $T_{12}^{(j)}$ has the asymptotic distribution given in (3.18) which is stochastically largest under the null hypothesis. By convoluting those distributions as we did in the proof of Theorem 3.1, we get the least favorable distribution provided in the theorem.

## 4. EXAMPLE

We apply in this section our testing method to the 'oropharynx' data which are grouped into seven intervals in Dykstra et al(1991). The patients are classified into four populations according to the amount of lymph node deterioration at the beginning of the study. However, we will consider the first two populations as shown in Table 4.1. Population 0 has no evidence of lymph node metastases and Population 1 is confronted with tumors. Since our inference does not allow censoring, censored observations are dropped from the original data set. Thus, in this example, we are interested in testing whether or not there is failure rate ordering ($F_0 \succeq F_1$) between these uncensored populations.

Suppose that complete observations from both of the samples occur on a subset of times $S_1 < S_2 < \cdots < S_k (S_o = 0, S_{k+1} = \infty)$. We will use the following notations:

$n_{ij} =$ number of complete observations from the $i$-th population at $S_j$

$m_{ij} = \sum_{r=j}^{k} n_{ir} =$ number of observations from the $i$-th population surviving just prior to $S_j$.

**Table 4.1.** Number of Survivals and Deaths for Grouped Data

| | | Pop 0 | | Pop 1 | |
|---|---|---|---|---|---|
| Group | Interval | $m$ | $n$ | $m$ | $n$ |
| I | 0–160 | 29 | 3 | 19 | 2 |
| II | 161–260 | 26 | 5 | 17 | 2 |
| III | 261–360 | 21 | 5 | 15 | 6 |
| IV | 361–540 | 16 | 9 | 9 | 2 |
| V | 541–700 | 7 | 4 | 7 | 4 |
| VI | 701–900 | 3 | 2 | 3 | 1 |
| VII | above 900 | 1 | 1 | 2 | 2 |

**Table 4.2.** Maximum Likelihood Estimates of $\theta_{ij}$'s

| $i \backslash j$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 0 | 0.8966 | 0.8077 | 0.7619 | 0.4375 | 0.4286 | 0.3333 |
| | (0.8958) | (0.8077) | (0.6944) | (0.4375) | (0.4286) | (0.3333) |
| 1 | 0.8947 | 0.8824 | 0.6000 | 0.7778 | 0.4286 | 0.6667 |
| | (0.8958) | (0.8871) | (0.6944) | (0.7778) | (0.4286) | (0.6667) |

Note : 1) Restricted MLE's are given in ( ).
2) $T_{12} = 1.073$ ($p$-value $= 0.645$).

For this grouped data, the restricted and unrestricted MLE's are easily obtained by the reparameterization discussed in Section 2. These estimates are given in Table 4.2. The $p$-value of the LRT statistic for failure rate ordering is 0.645 as given below Table 4.2. This large $p$-value seems somewhat surprising because two violations in ordering are found between the elements of the unrestricted MLE of $\underset{\sim}{\theta}_j$. However, those violations do not result in large value of test statistic because the reversals are coming up with small discrepancy. This fact will be the main reason for not showing clear evidence for rejecting failure rate ordering.

## ACKNOWLEDGEMENT

The author wishes to thank the referees for helpful comments.

## REFERENCES

( **1**) Aly, E.E. (1988). Comparing and testing under order relations between percentile residual life functions. *The Canadian Journal of Statistics*, Vol. **16**, **337–369**.

( **2**) Bagai, I. and Kochar, S.C. (1986). On tailordering and comparison of failure rates. *Communications in Statistics Theory and Method*, Vol. **15**, **1377–1388**.

( **3**) Cheng, K.F. (1985). Tests for the equality of failure rates. *Biometrika*, Vol. **72**, **211–215**.

( **4**) Dykstra, R.L., Kochar, S., and Robertson, Tim. (1990). Testing for the proportionality of cause specific hazard rates against ordered alternatives in the competing risks model. Technical Report 184, Department of Statistics and Actuarial Science, University of Iowa.

( 5) Dykstra, R. L., Kochar, S., and Robertson, Tim. (1991). Statistical inference for uniform stochastic ordering in several populations. *The Annals of Statistics*, Vol. **19**, **870–888**.

( 6) Joe, H. and Proschan, F. (1984). Comparison of life distributions on the basis of their percentile residual life functions. *The Canadian Journal of Statistics*, Vol. **12**, **91–97**.

( 7) Kochar, S. (1979). Distribution-free comparison of two probability distributions with reference to their hazard rates. *Biometrika*, Vol. **66**, **437–441**.

( 8) Kochar, S. (1981). A new distribution-free test for the equality of two failure rates. *Biometrika*, Vol. **68**, **423–426**.

( 9) Robertson, Tim, Wright, F.T., and Dykstra, R.L. (1988). *Order Restricted Statistical Inference*, New York, John Wiley.

(10) Ross, S.M. (1983). *Stochastic Process*, New York, John Wiley.

(11) Savage, I.R. (1956). Contribution to rank order statistics: The two-sample case. *Annals of Mathematical Statistics*, Vol. **27**, **590–615**.