Testing for Ordered Failure Rates under General Progressive Censoring

Bhaskar Bhattacharya
Department of Mathematics, Southern Illinois University
Carbondale, IL 62901-4408, USA

Abstract

For exponentially distributed failure times under general progressive censoring schemes, testing procedures for ordered failure rates are proposed using the likelihood ratio principle. Constrained maximum likelihood estimators of the failure rates are found. The asymptotic distributions of the test statistics are shown to be mixtures of chi-square distributions. When testing the equality of the failure rates, a simulation study shows that the proposed test with restricted alternative has improved power over the usual chi-square statistic with an unrestricted alternative. The proposed methods are illustrated using data of survival times of patients with squamous carcinoma of the oropharynx.

Key words and phrases: Chi-bar square distribution, clinical trials, nondecreasing order, two-parameter exponential distribution, life-testing.
1 Introduction

Progressive censoring schemes are very useful in clinical trials and life-testing experiments. For example, in a clinical trial study, suppose the survival times of patients with squamous carcinoma of the oropharynx are being compared. The patients are placed in different groups depending upon the degree of lymph node deterioration. It is inherently plausible that the disease is further advanced in those patients with more lymph node deterioration, and hence their survival times would generally be shorter. Some unobserved failures might exist in any group before the study officially begins. As the study progresses, with each failure some patients are possibly censored for various reasons, e.g., patients may leave because they are doing well physically, move out of the region for personal reasons, etc. Study would stop at any predetermined time or when the experimenter believes that enough information has already been collected, at which point all remaining patients are censored. Progressive censoring is also useful in a life-testing experiment because the ability to remove live units from the experiment saves time and money.

Sen (1985) describes the progressive censoring schemes in a time sequential view and points out that statistical monitoring plays a major role in the termination of the study. Chatterjee and Sen (1972) and Majumder and Sen (1978) suggest a general class of nonparametric testing procedures under progressive censoring schemes. Sen (1985) also addresses the nonparametric testing procedures against restricted alternatives. Cohen and Whitten (1988) and Cohen (1991) have summarized the likelihood inference under progressive censoring for a wide range of distributions. Mann (1971) and Thomas and Wilson (1972) discuss the best linear invariant estimates. Balakrishnan and Sandhu (1996) has derived the best linear unbiased and maximum likelihood estimates (MLEs) under general progressive type II censored samples from exponential distributions. Viveros and Balakrishnan (1994) has developed the exact conditional inference based on progressive type II censored samples. For a general
account under the progressive censoring scheme, see Sen (1981, Ch 11). The monograph by Balakrishnan and Aggarwala (2000) provides a wealth of information on inferences under progressive censoring sampling.

We consider a general (type-II) progressive censoring scheme as follows (we use a life-testing format, but our description fits equally well in a clinical trial set-up): for the $i$th population ($1 \leq i \leq k$), suppose $N_i$ randomly selected units are placed on a life-test; the failure times of the first $r_i$ units to fail are not observed; at the time of the $(r_i + j)$th failure, $R_{i,r_i+j}$ number of surviving units are withdrawn from the test randomly, for $j = 1, \ldots, m_i - r_i - 1$; and finally at the $m_i$th failure, the remaining $R_{i,m_i}$ units are withdrawn from the test where $R_{i,m_i} = N_i - m_i - R_{i,r_i+1} - R_{i,r_i+2} - \ldots - R_{i,m_i-1}$, $1 \leq i \leq k$. Note that

$$\sum_{j=r_i+1}^{m_i} (1 + R_{i,j}) = N_i - r_i, \ 1 \leq i \leq k.$$ (1.1)

This identity will be used several times in the sequel.

Suppose the lifetimes of the completely observed units to fail from the $i$th population are $X_{r_i+1, N_i} \leq X_{r_i+2, N_i} \leq \ldots \leq X_{m_i, N_i}, \ 1 \leq i \leq k$. If the failure times from the $i$th (continuous) population have a cumulative distribution function $F_i(x)$ and a probability density function $f_i(x)$, then the joint probability density function of $X_{r_i+1, N_i}, X_{r_i+2, N_i}, \ldots, X_{m_i, N_i}, \ 1 \leq i \leq k$ is given by

$$c \prod_{i=1}^{k} \left\{ \left[ F_i(x_{r_i+1}) \right]^{r_i} \prod_{j=r_i+1}^{m_i} f_i(x_j) \left( 1 - F_i(x_j) \right)^{R_{i,j}} \right\}$$ (1.2)

where

$$c = \prod_{i=1}^{k} \left( \frac{N_i}{r_i} \right) (N_i - r_i) \prod_{j=r_i+2}^{m_i} \left( N_i - \sum_{s=r_i+1}^{j-1} R_{i,s} - j + 1 \right)$$ (1.3)

Balakrishnan and Aggarwala (2000). In this paper, we assume the failure times follow the two-parameter exponential distributions. These distributions are well-known to be very useful when modeling survival, life-testing and/or reliability data.
We consider \( k \) general progressively censored random samples from independent two-parameter exponential distributions when the scale parameters \((\theta_i, 1 \leq i \leq k)\) satisfy a nondecreasing restriction. Such a restriction amounts to a nonincreasing nature of the failure rates \((\theta_i^{-1})\) among the populations involved. In Section 2, we consider the MLEs of all parameters under such restrictions. It is well-known that the variance of the unrestricted MLEs does not depend on the censoring schemes (Balakrishnan and Sandhu, 1996). However, in Lemma 2.1, we show, rather surprisingly, that the estimates themselves are free of the censoring scheme. In Section 3, we consider the likelihood ratio tests for testing homogeneity of the scale parameters against the nondecreasing order. In Section 4, simulation studies are performed to show how the restricted estimates depend on \( r_i \), which is the number of initially unobserved failed units from the \( i \)th population and the sample size \( n_i \). It turns out that smaller values of \( r_i \)'s yield a more reliable inference than larger ones in both estimation and testing. We also analyze a data set of survival times of patients with squamous carcinoma of the oropharynx from three groups using the procedures developed in this paper.

## 2 Maximum Likelihood Estimation

We consider the two parameter exponential distribution with probability density function given by

\[
 f(x; \mu, \theta) = \frac{1}{\theta} e^{-\frac{x-\mu}{\theta}}, \quad x \geq \mu > 0, \ \theta > 0. \tag{2.1}
\]

We assume that independent samples are available from \( k \) different exponential distributions with location parameters \( \mu_i \) and scale parameters \( \theta_i, 1 \leq i \leq k \). Let \( \mu = (\mu_1, \ldots, \mu_k), \ \theta = (\theta_1, \ldots, \theta_k) \). The joint likelihood of the \( k \) samples from (1.2) and (2.1) is given by

\[
 L(\mu, \theta) = c^* \prod_{i=1}^{k} \left\{ \left[ 1 - e^{-\frac{x_{i,N_i-\mu_i}}{\theta_i}} \right]^{r_i} \prod_{j=r_i+1}^{m_i} \left( \frac{1}{\theta_i} e^{-\frac{x_{j,N_i-\mu_i}}{\theta_i}} \right) e^{-\frac{x_{j,N_i-\mu_i}}{\theta_i}} \right\} R_{i,j} \tag{2.2}
\]

where the constant \( c^* \) may be calculated from (1.3) and (2.1).
The unrestricted maximum likelihood estimates (Balakrishnan and Sandhu, 1996) of the parameters \( \mu_t, \theta_t, 1 \leq t \leq k \) are given by

\[
\hat{\mu}_t = x_{r_t+1,N_t} + \hat{\theta}_t \ln \left( 1 - \frac{r_t}{N_t} \right), \quad 1 \leq t \leq k,
\]

\[
\hat{\theta}_t = \frac{\sum_{s=r_t+2}^{m_t}(1+R_{t,s})(x_s,N_t - x_{r_t+1,N_t})}{m_t - r_t}, \quad 1 \leq t \leq k.
\]

(2.3)

Let \( \hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_k) \) and \( \hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k) \). Balakrishnan and Sandhu (1996) has shown that the variances of \( \hat{\theta}_i \)'s do not depend on the censoring scheme. Here we show that the estimates in (2.3) themselves do not depend on the censoring scheme.

To the best of our knowledge, this fact has not been explicitly derived in the literature earlier.

**Lemma 2.1.** The unrestricted estimates \( \hat{\mu}_t, \hat{\theta}_t, 1 \leq t \leq k \) in (2.3) do not depend on the censoring scheme \( \{R_{t,s}, r_t + 1 \leq s \leq m_t, 1 \leq t \leq k\} \).

**Proof.** Using Theorem 3.4 of Balakrishnan and Aggarwala (2000), when \( N \) items are put on test with \( r \) initial failures not observed, if \( Y_{r+1}, \ldots, Y_m \) denote a general progressively censored sample from the exponential distribution with location parameter \( \mu \) and scale parameter \( \theta \) with censoring scheme \( R_{r+1}, \ldots, R_m \), the generalized spacings defined by

\[
Z_{r+1} = (N - r)(Y_{r+1} - \mu),
\]

\[
Z_{r+2} = (N - r - R_{r+1} - 1)(Y_{r+2} - Y_{r+1}),
\]

\[
Z_{r+3} = (N - r - R_{r+1} - R_{r+2} - 2)(Y_{r+3} - Y_{r+2}),
\]

\[
\ldots = \ldots
\]

\[
Z_m = (N - r - R_{r+1} - \ldots - R_{m-1} - m + r + 1)(Y_m - Y_{m-1})
\]

are independent random variables, with \( Z_{r+2}, \ldots, Z_m \) being one-parameter exponential random variables with mean \( \theta \), and \( Z_{r+1}/(N-r) \) being distributed as the \((r+1)\)th usual order statistic from a sample of size \( N \) from the same distribution. Since \( \sum_{j=r+1}^{m}(1 + R_j) = N - r \), it follows from (2.4) using algebra that

\[
\sum_{s=r+1}^{m} Z_s = \sum_{j=1}^{m-r} (1 + R_{r+j})Y_{r+j} - (N - r)\mu.
\]
Now, with \( Z_{t,s}, X_{r,t}\), playing the role of \( Z_s, Y_{r,t} \) (respectively) above and noting that the numerator of \( \hat{\theta}_t \) in (2.3) can be expressed as

\[
\sum_{s=r_t+1}^{m_t} (1 + R_{t,s})X_{s,N_t} - \left( \sum_{s=r_t+1}^{m_t} (1 + R_{t,s}) \right)X_{r_t+1,N_t}
\]

\[
= \sum_{s=r_t+1}^{m_t} (1 + R_{t,s})X_{s,N_t} - (N_t - r_t)\mu_t - [(N_t - r_t)X_{r_t+1,N_t} - (N_t - r_t)\mu_t]
\]

\[
= \sum_{s=r_t+1}^{m_t} Z_{t,s} - Z_{t,r_t+1}
\]

\[
= \sum_{s=r_t+2}^{m_t} Z_{t,s}
\]

we can write

\[
\hat{\theta}_t = \frac{\sum_{s=r_t+2}^{m_t} Z_{t,s}}{m_t - r_t}, \quad 1 \leq t \leq k.
\]

Since \( Z_{t,s} \) are random variables whose distributions are free of \( R_{t,s} \), it follows that \( \hat{\theta}_t \) is free of \( R_{t,s} \). Also from (2.3), it follows that \( \hat{\mu}_t \) is free of \( R_{t,s} \). \( \square \)

It follows from (2.3) and the proof of Lemma 2.1 that \( \hat{\mu}_t \) and \( \hat{\theta}_t \) are independent when \( r_t = 0 \). However, this is not the case when \( r_t > 0 \). Also, when \( \mu_t \) is known and \( r_t > 0 \), the maximum likelihood estimate of \( \theta_t \) is obtained by solving an equation numerically (Balakrishnan and Sandhu, 1996). In this latter case, the MLE of \( \theta_t \) is not free of the censoring scheme.

To find the MLEs of the parameters \( \mu_i, \theta_i, 1 \leq i \leq k \) subject to the constraints

\[ H_0 : \theta_1 = \theta_2 = \ldots = \theta_k \tag{2.5} \]

the log-likelihood from (2.2) can be expressed as (except for the constant term)

\[
\ln L(\bm{\mu}, \theta) = \sum_{i=1}^{k} \left\{ r_i \ln \left( 1 - e^{-\frac{x_{r_i+1,N_i} - \mu_i}{\theta}} \right) - (m_i - r_i) \ln \theta - \frac{1}{\theta} \sum_{j=r_i+1}^{m_i} (1 + R_{i,j})(x_{j,N_i} - \mu_i) \right\}
\]

where \( \theta \) is the common value of the \( \theta_i \)'s under \( H_0 \).

Differentiating \( \ln L(\bm{\mu}, \theta) \) with respect to \( \mu_i \) and \( \theta \), we obtain

\[
\frac{\partial \ln L}{\partial \mu_i} = -\frac{r_i}{\theta} e^{-\frac{x_{r_i+1,N_i} - \mu_i}{\theta}} + \frac{1}{\theta} \sum_{j=r_i+1}^{m_i} (1 + R_{i,j}) \tag{2.6}
\]
and

\[
\frac{\partial \ln L}{\partial \theta} = -\frac{r_i}{\theta} e^{\frac{x_{r_i+1,N_t} - \mu_i}{\theta}} x_{r_i+1,N_t} - \mu_i - m_i - r_i + \frac{1}{\theta^2} \sum_{j=r_i+1}^{m_i} (1 + R_{i,j})(x_{j,N_t} - \mu_i).
\]

(2.7)

Setting the derivatives (2.6) and (2.7) equal to zero and simplifying them using (1.1), we get the MLEs under $H_0$ as follows

\[
\mu_i^0 = x_{r_t+1,N_t} + \theta^0 \ln \left(1 - \frac{r_i}{N_t}\right), \quad 1 \leq t \leq k,
\]

(2.8)

and

\[
\theta^0 = \frac{\sum_{i=1}^k \sum_{j=r_i+1}^{m_i} (1 + R_{i,j})(x_{j,N_t} - x_{r_i+1,N_t})}{\sum_{i=1}^k (m_i - r_i)}.
\]

(2.9)

Let $\mu^0 = (\mu_1^0, \ldots, \mu_k^0)$ and $\theta^0 = (\theta_1^0, \ldots, \theta_k^0)$.

Next, we like to find the MLEs of the parameters $\mu_i, \theta_i, 1 \leq i \leq k$ subject to the constraints that

\[
H_1 : \theta_1 \leq \theta_2 \leq \ldots \leq \theta_k.
\]

(2.10)

From (2.2), the log-likelihood can be simplified as (except for the constant term)

\[
\ln L(\mu, \theta) = \sum_{i=1}^k \left\{ r_i \ln \left(1 - e^{-\frac{x_{r_i+1,N_t} - \mu_i}{\theta_i}}\right) - (m_i - r_i) \ln \theta_i - \frac{1}{\theta_i} \sum_{j=r_i+1}^{m_i} (1 + R_{i,j})(x_{j,N_t} - \mu_i) \right\}.
\]

Differentiating $\ln L(\mu, \theta)$ with respect to $\mu_i$ and $\theta_i$, we obtain

\[
\frac{\partial \ln L}{\partial \mu_i} = -\frac{r_i}{\theta_i} e^{\frac{x_{r_i+1,N_t} - \mu_i}{\theta_i}} x_{r_i+1,N_t} - \mu_i - m_i - r_i + \frac{1}{\theta_i^2} \sum_{j=r_i+1}^{m_i} (1 + R_{i,j})(x_{j,N_t} - \mu_i),
\]

and

\[
\frac{\partial \ln L}{\partial \theta_i} = -\frac{r_i}{\theta_i} e^{\frac{x_{r_i+1,N_t} - \mu_i}{\theta_i}} x_{r_i+1,N_t} - \mu_i - m_i - r_i + \frac{1}{\theta_i^2} \sum_{j=r_i+1}^{m_i} (1 + R_{i,j})(x_{j,N_t} - \mu_i).
\]

(2.11)

(2.12)

To maximize $\ln L$ under $H_1$, we can equivalently minimize $B(\mu, \theta) = -\ln L(\mu, \theta)$ subject to $\theta_i \leq \theta_{i+1}, 1 \leq i \leq k - 1$. For solution of this optimization problem, we appeal to the Kuhn-Tucker necessary conditions. Setting

\[
c_t(\mu_t, \theta_t) = \frac{\partial B(\mu, \theta)}{\partial \mu_t}, \quad d_t(\mu_t, \theta_t) = \frac{\partial B(\mu, \theta)}{\partial \theta_t}, \quad 1 \leq t \leq k,
\]

7
with some algebra, the Kuhn-Tucker conditions for this minimization problem are equivalent to

\[
\sum_{t=1}^{i} d_t(\mu_t, \theta_t) + v_i = 0, \quad 1 \leq i \leq k - 1, \quad \sum_{t=1}^{k} d_t(\mu_t, \theta_t) = 0, \tag{2.13}
\]

\[
c_t(\mu_t, \theta_t) = 0, \quad 1 \leq t \leq k, \tag{2.14}
\]

\[
v_i(\theta_i - \theta_{i+1}) = 0, \quad v_i \geq 0, \quad \theta_i - \theta_{i+1} \leq 0, \quad 1 \leq i \leq k - 1, \tag{2.15}
\]

where \(v_i\) are the Lagrange multipliers corresponding to the inequality constraints. Let the solutions to (2.13 - 2.15) be denoted by \(\mu^* = (\mu_1^*, \ldots, \mu_k^*), \theta^* = (\theta_1^*, \ldots, \theta_k^*), v^* = (v_1^*, \ldots, v_k^*)\), which are the desired estimates under \(H_1\).

Let \(Av(i, j), i \leq j\) be the solution \(\theta_0\) of the equations

\[
c_t(\mu_t, \theta_0) = 0, \quad t = i, \ldots, j, \tag{2.16}
\]

\[
\sum_{t=i}^{j} d_t(\mu_t, \theta_0) = 0. \tag{2.17}
\]

Using the identity (1.1), it follows from (2.11) and (2.16) that

\[
\mu_t = x_{r_t+1,N_t} + \theta_0 \ln \left(1 - \frac{r_t}{N_t}\right), \quad 1 \leq t \leq k. \tag{2.18}
\]

Using this value of \(\mu_t\), it follows from (2.12) and (2.17) that

\[
\sum_{t=i}^{j} \left(\frac{1}{\theta_0} \sum_{s=r_t+1}^{m_t} (1 + R_{i,s}) \frac{x_{r_t+1,N_t} - \mu_t}{\theta_0} + m_t - r_t \frac{1}{\theta_0^2} \sum_{s=r_t+1}^{m_t} (1 + R_{t,s})(x_{s,N_t} - \mu_t)\right) = 0
\]

which solving for \(\theta_0\) (\(= Av(i,j)\)) yields

\[
\theta_0 = \frac{\sum_{t=i}^{j} \sum_{s=r_t+1}^{m_t} (1 + R_{i,s})(x_{s,N_t} - x_{r_t+1,N_t})}{\sum_{t=i}^{j} (m_t - r_t)}. \tag{2.19}
\]

The estimates under \(H_1\) are given by the following theorem.

**Theorem 2.1.** The constrained estimates of \(\mu_t, \theta_t\) under \(H_1\) are given by

\[
\mu_t^* = x_{r_t+1,N_t} + \theta_t^* \ln \left(1 - \frac{r_t}{N_t}\right), \quad t = 1, \ldots, k, \tag{2.20}
\]

\[
\theta_t^* = \max_{i \leq t} \min_{j \geq t} \frac{\sum_{h=i}^{j} \sum_{s=r_h+1}^{m_h} (1 + R_{h,s})(x_{s,N_h} - x_{r_h+1,N_h})}{\sum_{h=i}^{j} (m_h - r_h)}, \quad t = 1, \ldots, k. \tag{2.21}
\]
Proof. It is easy to verify that (2.20) satisfies (2.14). Let \( v_t^* = -\sum_{a=1}^t d_a(\mu_a^*, \theta_a^*), \) \( 1 \leq t \leq k - 1. \) The first part of (2.13) is satisfied, and \( \theta_i^* \) are nondecreasing, so the last part of (2.15) is satisfied.

For the level set \( \{i, i+1, \ldots, j\} \) so that \( \theta_{i-1}^* < \theta_i^* = \theta_{i+1}^* = \ldots = \theta_j^* < \theta_{j+1}^* \), we have \( \theta_i^* = Av(i, j), \) \( t = i, i+1, \ldots, j, \) and from the definition of \( \theta_i^* \) (or \( Av(i, j) \)) and (2.17), we have
\[
\sum_{t=1}^{i-1} \sum_{r=1}^j d_t(\mu_r^*, \theta_r^*) = 0, \quad \sum_{t=1}^j d_t(\mu_r^*, \theta_r^*) = 0, \quad (2.22)
\]
which implies that \( v_{i-1}^* = v_j^* = 0. \) Thus \( v_t^*(\theta_{i+1}^* - \theta_i^*) = 0, \) \( 1 \leq i \leq k - 1. \) Also using \( j = k, \) the second part of (2.13) holds.

It suffices to prove that \( v_t^* \geq 0, \) \( t = i, \ldots, j - 1. \) From (2.22) we have \( v_t^* = -\sum_{a=1}^t d_a(\mu_a^*, \theta_a^*), \) \( t = i, i+1, \ldots, j - 1. \) Using (2.11), (2.12) and \( \theta_0 = Av(i, j) \) we have
\[
v_t^* = -\sum_{a=1}^t \left[ \left( \frac{1}{\theta_0^2} \sum_{h=r_a+1}^{m_a} (1 + R_{a,h}) \right) (x_{r_a+1,N_a} - \mu_a^*) - \frac{m_a - r_a}{\theta_0} \right.
\]
\[
+ \frac{1}{\theta_0^2} \sum_{h=r_a+1}^{m_a} (1 + R_{a,h})(x_{h,N_a} - \mu_a^*) \right] \right]
\]
\[
= \sum_{a=i}^t \left( \frac{N_a - r_a}{\theta_0} \right) \ln \left( \frac{1 - r_a}{N_a} \right) - \sum_{a=i}^t \frac{m_a - r_a}{\theta_0}
\]
\[
+ \frac{1}{\theta_0^2} \sum_{a=i}^t \sum_{h=r_a+1}^{m_a} (1 + R_{a,h}) \left( x_{h,N_a} - x_{r_a+1,N_a} - \theta_0 \ln \left( \frac{1 - r_a}{N_a} \right) \right)
\]
\[
= \sum_{a=i}^t \frac{m_a - r_a}{\theta_0} - \frac{1}{\theta_0^2} \sum_{a=i}^t \sum_{h=r_a+1}^{m_a} (1 + R_{a,h}) (x_{h,N_a} - x_{r_a+1,N_a})
\]
\[
= \sum_{a=i}^t \frac{(m_a - r_a)}{\theta_0^2} \left( \frac{\sum_{a=i}^t \sum_{h=r_a+1}^{m_a} (1 + R_{a,h}) (x_{h,N_a} - x_{r_a+1,N_a})}{\sum_{a=i}^t (m_a - r_a)} - \theta_0 \right)
\]
which is nonnegative because \( Av(i, t) \geq Av(i, j), \) \( \forall t = i, \ldots, j - 1. \) This completes the proof of the theorem. □
The estimates in (2.21) can be obtained by isotonic regression of \( \hat{\theta} \) onto the cone of nondecreasing vectors with weights \( m = (m_1, \ldots, m_k) \). These estimates are computed easily by using the pooled adjacent violators algorithm (Robertson, et. al., 1988). We will use these constrained estimates in the next section to construct the test statistics.

3 Likelihood Ratio Tests

We assume \( r_i > 0, \forall i \). To test \( H_0 \) versus \( H_1 - H_0 \), the likelihood ratio test statistic may be expressed as

\[
T_{01} = 2[\ln L(\mu^*, \theta^*) - \ln L(\mu^0, \theta^0)].
\] (3.1)

However, the exact distribution of \( T_{01} \) seems intractable; hence, we appeal to the asymptotic theory. From Lemma C of Serfling (1980, page 154), we have

\[
2[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu, \theta)] = Q(\mu, \theta) + o_p(1)
\] (3.2)

where

\[
Q(\mu, \theta) = (\hat{\mu} - \mu', \hat{\theta} - \theta')I(\mu, \theta)(\hat{\mu} - \mu', \hat{\theta} - \theta')'
\]

where \( I(\mu, \theta) \) is the information matrix.

Since \( L(\mu^0, \theta^0) = \sup_{H_0} L(\mu, \theta) \) and \( L(\mu^*, \theta^*) = \sup_{H_1} L(\mu, \theta) \), we can write

\[
\inf_{H_0}[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu, \theta)] = \ln L(\hat{\mu}, \hat{\theta}) - \sup_{H_0} \ln L(\mu, \theta)
\]

\[
= \ln L(\mu, \theta) - \ln L(\mu^0, \theta^0)
\]

and

\[
\inf_{H_1}[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu, \theta)] = \ln L(\hat{\mu}, \hat{\theta}) - \sup_{H_1} \ln L(\mu, \theta)
\]

\[
= \ln L(\mu, \theta) - \ln L(\mu^*, \theta^*).
\]

Minimizing both sides of (3.2) under \( H_0 \) and under \( H_1 \), we get,

\[
2[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu^0, \theta^0)] = \inf_{H_0} Q(\mu, \theta) + o_p(1),
\] (3.3)

\[
2[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu^*, \theta^*)] = \inf_{H_1} Q(\mu, \theta) + o_p(1),
\]
respectively. Thus we can rewrite $T_{01}$ as

\[
T_{01} = 2[\ln L(\mu^*, \theta^*) - \ln L(\mu^0, \theta^0)] \\
= 2[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu^0, \theta^0)] - 2[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu^*, \theta^*)] \\
= \inf_{H_0} Q(\mu, \theta) - \inf_{H_1} Q(\mu, \theta) + o_p(1).
\]

Now in the expression of $Q(\mu, \theta)$, replacing $I(\mu, \theta)$ by $I(\hat{\mu}, \hat{\theta})$ we define

\[
Q^*(\mu, \theta) = (\hat{\mu}' - \mu', \hat{\theta}' - \theta')I(\hat{\mu}, \hat{\theta})(\hat{\mu}' - \mu', \hat{\theta}' - \theta')'.
\]

Since $(\hat{\mu}', \hat{\theta}') \to (\mu', \theta')$ a.s., so $I(\hat{\mu}, \hat{\theta}) \to I(\mu, \theta)$ a.s., and $Q^*(\mu, \theta) - Q(\mu, \theta) \to 0$ a.s.

Let $(\mu_0, \theta_0)$ and $(\mu_1, \theta_1)$ denote the estimates of $(\mu, \theta)$ obtained by minimizing $Q^*(\mu, \theta)$ under $H_0$ and $H_1$, respectively. Then from (3.4), we can say that

\[
T_{01} - (Q^*(\mu_0, \theta_0) - Q^*(\mu_1, \theta_1)) \to 0, \text{ a.s.}
\]

Thus the asymptotic distribution of $T_{01}$ is same as that of $Q^*(\mu_0, \theta_0) - Q^*(\mu_1, \theta_1)$.

If we express the constraints in (2.5) and (2.10) as $H_0: C(\mu', \theta')' = 0$ and $H_1: C(\mu', \theta')' \geq 0$ respectively, where the $(k-1 \times 2k)$ matrix $C$ is given by

\[
C = \begin{bmatrix}
0 & \cdots & 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -1
\end{bmatrix}
\]

and $0$ is a column vector $k-1$ zeroes, then the asymptotic distribution of $T_{01}$ is given by Theorem 3.1 below.

When testing the order restrictions $H_1$ as a null hypothesis against the alternative $H_2 - H_1$ where $H_2: \text{no restriction among } \theta_i$'s, using (3.3) the test statistic is given by

\[
T_{12} = 2[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu^*, \theta^*)] \\
= \inf_{H_1} Q(\mu, \theta) + o_p(1).
\]
From earlier discussions, it follows that the asymptotic distribution of $T_{12}$ is the same as that of $Q^*(\mu_1, \theta_1)$, whose least favorable distribution under $H_0$ is given by Theorem 3.1 below.

For $1 \leq i \leq k-1$, let $P(i, k-1, CI^{-1}C')$, the level probabilities, be the probability that $C\theta$ has $i$ distinct positive components under $H_0$. The proof of Theorem 3.1 follows from the work of Shapiro (1985, 1988) and Kudô (1963).

**Theorem 3.1.** For a constant $u_1$, the asymptotic distribution of $T_{01}$ under $H_0$ is given by

$$\lim_{n_i \to \infty, \forall i} P(T_{01} \geq u_1) = \sum_{i=0}^{k-1} P(i, k-1, CI^{-1}C')P(\chi_i^2 \geq u_1)$$

where $\chi_i^2$ is a chi-square random variable with $i$ degrees of freedom with $\chi_0^2 \equiv 0$.

For $T_{12}$, $H_0$ is least favorable within $H_1$, and for a constant $u_2$, its asymptotic distribution under $H_0$ is given by

$$\lim_{n_i \to \infty, \forall i} P(T_{12} \geq u_2) = \sum_{i=0}^{k-1} P(k-i, k-1, CI^{-1}C')P(\chi_i^2 \geq u_2).$$

\[\square\]

Now we consider approximations for the level probabilities. Partition

$$I = \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix}$$

where each $I_{ij}$ is a diagonal matrix and is given by

$$I_{11} = -\left(\frac{\partial^2 \ln L}{\partial \mu_i^2}\right) = \text{Diag}(a_1, \ldots, a_k),$$

$$I_{22} = -\left(\frac{\partial^2 \ln L}{\partial \theta_i^2}\right) = \text{Diag}(b_1, \ldots, b_k),$$

$$I_{12} = -\left(\frac{\partial^2 \ln L}{\partial \mu_i \partial \theta_i}\right) = \text{Diag}(c_1, \ldots, c_k),$$

where

$$a_i = -\frac{\partial}{\partial i} w_i (1 + w_i),$$

$$b_i = \frac{2}{\partial i} g_i w_i - \frac{\partial}{\partial i} g_i^2 w_i (1 + w_i) + \frac{m - r_i}{\partial i} - \frac{2}{\partial i} \sum_{j=r_{i+1}}^{m_i} (1 + R_{i,j}) (x_{j,N_i} - \mu_i),$$

$$c_i = \frac{r_i}{\partial i} w_i - \frac{\partial}{\partial i} g_i w_i (1 + w_i) - \frac{1}{\partial i} \sum_{j=r_{i+1}}^{m_i} (1 + R_{i,j}).$$
where \( w_i = e^{-(x_{r_i+1,N_i} - \mu_i)}/\theta_i / (1 - e^{-(x_{r_i+1,N_i} - \mu_i)}/\theta_i) \), \( g_i = x_{r_i+1,N_i} - \mu_i \). To approximate the level probabilities under \( H_0 \), we replace \( \mu_i, \theta_i \) by \( \mu_i^0, \theta_i^0 \), respectively, in above expressions.

Let \( W = CI^{-1}C' \). When \( k = 2 \), we have \( P(0, 1, W) = P(1, 1, W) = .5 \). When \( k = 3 \), we have \( P(0, 2, W) = .5 - (\cos^{-1}(\rho_{12})/2\pi), \ P(1, 2, W) = .5, \ P(2, 2, W) = (\cos^{-1}(\rho_{12})/2\pi) \), where \( \rho_{12} \) is the (1,2)th element of the matrix \([\text{diag}(W)]^{-1/2}W\) [\text{diag}(W)]^{-1/2}, \ and \ can \ be \ expressed \ as \ \rho_{12} = -d_2/\sqrt{(d_1 + d_2)(d_2 + d_3)} \) where \( d_i = a_i/(a_i b_i - c_i^2) \).

For \( k = 4 \), we have
\[
P(0, 3, W) = 1/2 - (\cos^{-1}(\rho_{12} + \cos^{-1}(\rho_{13} + \cos^{-1}(\rho_{23}))/4\pi),
\]
\[
P(1, 3, W) = 3/4 - (\cos^{-1}(\rho_{123} + \cos^{-1}(\rho_{132} + \cos^{-1}(\rho_{231}))/4\pi),
\]
\[
P(2, 3, W) = 1/2 - P(0, 3, W), \ \text{and} \ P(3, 3, W) = 1/2 - P(1, 3, W)
\]

where \( \rho_{ij-k} = (\rho_{ij} - \rho_{ik}\rho_{jk})/\sqrt{(1 - \rho_{ik}^2)(1 - \rho_{jk}^2)} \) with
\[
\rho_{12} = -d_2/\sqrt{(d_1 + d_2)(d_2 + d_3)}, \ \rho_{13} = 0, \ \rho_{23} = -d_3/\sqrt{(d_2 + d_3)(d_3 + d_4)}.
\]

For \( k \geq 5 \), expressions for the level probabilities are available in terms of orthant probabilities for a multivariate normal distribution. However, numerical techniques are needed to compute these arbitrary orthant probabilities. For this purpose, the programs of Bohrer and Chow (1978) and Sun (1988) are useful.

**Remark 3.1.** The usual likelihood ratio test for \( H_0 \) against unrestricted alternative \( H_2 - H_0 \) is given by
\[
T_{02} = 2[\ln L(\hat{\mu}, \hat{\theta}) - \ln L(\mu^0, \theta^0)]
\]
which has an asymptotically chi-square distribution with \( k - 1 \) degrees of freedom.

We compare in Section 4 the performance of the two tests \( T_{01} \) and \( T_{02} \).

**Remark 3.2.** When \( r_i = 0 \forall i \) (progressive type II right censoring), then the log-likelihood reduces to (except for the constant term)
\[
\ln L(\mu, \theta) = \sum_{i=1}^{k} \left\{ -m_i \ln \theta_i - \frac{1}{\theta_i} \left[ \sum_{j=1}^{m_i} (1 + R_{i,j})(x_j, N_i - \mu_i) \right] \right\}.
\]
The unrestricted MLEs are \( \hat{\mu}_i = x_{1,N_i} \) and \( \hat{\theta}_i = \sum_{j=2}^{m_i} (R_{i,j} + 1)(x_{j,N_i} - x_{1,N_i})/m_i, 1 \leq i \leq k \) (Cohen, 1991). From the proof of Lemma 2.1 (with \( Z_{j,N_i} \) playing the role of \( Z_j \) in (2.4)), it follows that \( \hat{\mu}_i = Z_{1,N_i}/N_i + \mu_i, \hat{\theta}_i = \sum_{j=2}^{m_i} Z_{j,N_i}/m_i, 1 \leq i \leq k \) and \( m_i \hat{\theta}_i \) has a gamma distribution with shape parameter \( m_i - 1 \) and scale parameter \( \hat{\theta}_i \), \( i = 1, \ldots, k \).

Since \( Z_{j,N_i} \)'s are independent, it follows that \( \hat{\mu}_i \) and \( \hat{\theta}_i \) are independent, and hence for testing hypothesis concerning \( \theta_i \)'s, it is enough to work with the distribution of \( \hat{\theta}_i \)'s.

It is easily seen that under \( H_0 \), the MLE of \( \theta_i \) is \( \hat{\theta}_0 = \sum_{i=1}^{k} \sum_{j=2}^{m_i} (R_{i,j} + 1)(x_{j,N_i} - x_{1,N_i})/\sum_{i=1}^{k} m_i \). Under \( H_1 \), the MLE of \( \theta = (\theta_1, \ldots, \theta_k) \) is \( \theta^* = (\theta^*_1, \ldots, \theta^*_k) \), which is the isotonic regression of \( \hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k) \) onto the cone of nondecreasing vectors with weights \( m = (m_1, m_2, \ldots, m_k) \). It follows from Lemma 2.1 (with \( r = 0 \)) that \( \theta^* \) is free of the censoring scheme as well. The likelihood ratio tests of \( H_0 \) versus \( H_1 - H_0 \) and \( H_1 \) versus \( H_2 - H_1 \) reduce to those described in pages 174-175 of Robertson et al. (1988). The related test statistics can be simplified to

\[
T_{01} = 2 \sum_{i=1}^{k} (m_i - 1) \ln \left( \frac{\hat{\theta}_0}{\hat{\theta}^*_i} \right) \quad \text{and} \quad T_{12} = 2 \sum_{i=1}^{k} (m_i - 1) \ln \left( \frac{\hat{\theta}^*_i}{\hat{\theta}_i} \right),
\]

respectively. The next theorem gives the asymptotic distributions of these statistics (from Theorem 4.1.1 of Robertson et al., 1988).

**Theorem 3.2.** For a constant \( u_1 \), the asymptotic distribution of \( T_{01} \) under \( H_0 \) is given by

\[
\lim_{M \to \infty} P(T_{01} \geq u_1) = \sum_{i=1}^{k} P(i, k, \mathbf{w}) P(\chi^2_{i-1} \geq u_1)
\]

where \( M = \sum_{i=1}^{k} m_i \), \( \mathbf{w} = (w_1, \ldots, w_k) \) and \( w_i = \lim_{m_i \to \infty} m_i/M \).

For \( T_{12} \), \( H_0 \) is least favorable within \( H_1 \), and for a constant \( u_2 \), its asymptotic distribution under \( H_0 \) is given by

\[
\lim_{m_i \to \infty, \forall i} P(T_{12} \geq u_2) = \sum_{i=1}^{k} P(k - i, k, \mathbf{w}) P(\chi^2_{i-1} \geq u_2).
\]

For \( k \leq 4 \), the level probabilities \( P(i, k, \mathbf{w}) \) can be found in Tables A1 - A3 of Robertson et al. (1988). For \( k > 4 \), they can be simulated by estimating \( w_i \).
with \( m_i/M \). For the case of equal \( m_i \)'s, better approximations to the asymptotic distribution are available using the results of Bain and Engelhardt (1975). We direct the reader to equations (4.1.19) and (4.1.20) of Robertson et. al. (1988) for those results.

**Remark 3.3.** When \( r_i = 0 \forall i \) and \( \mu_i = 0 \forall i \) (or, equivalently, \( \mu_i \)'s are known), (progressive type II right censoring with given guarantee periods), then the log-likelihood reduces to (except for the constant term)

\[
\ln L(\theta) = \sum_{i=1}^{k} \left\{ -m_i \ln \theta_i - \frac{1}{\theta_i} \sum_{j=1}^{m_i} (1 + R_{i,j}) x_{j,i} \right\}.
\]

The unrestricted MLEs are \( \hat{\theta}_i = \sum_{j=1}^{m_i} (R_{i,j} + 1)x_{j,i}/m_i, 1 \leq i \leq k \) (Cohen, 1991). Here \( m_i \hat{\theta}_i \) has a gamma distribution with shape parameter \( m_i \) and scale parameter \( \theta_i, i = 1, \ldots, k \) (this follows from the proof of Lemma 2.1 by setting \( r = 0 \)).

Under \( H_0 \), the MLE of \( \theta_i \) is \( \hat{\theta}_0 = \sum_{j=1}^{m_i} (R_{i,j} + 1)x_{j,i}/\sum_{i=1}^{k} m_i \). Under \( H_1 \), the MLE of \( \theta = (\theta_1, \ldots, \theta_k) \) is \( \theta^* = (\theta^*_1, \ldots, \theta^*_k) \), which is the isotonic regression of \( \hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k) \) onto the cone of nondecreasing vectors with weights \( m = (m_1, m_2, \ldots, m_k) \). It follows from Lemma 2.1 (with \( r = 0 \)) that all these estimates are free of the censoring scheme as well. The likelihood ratio tests and their asymptotic distributions can be obtained as in Remark 3.2 by replacing \( m_i - 1 \) with \( m_i \).

### 4 Simulation and Example

Since the MLEs are only asymptotically efficient and our testing results are based on the large sampling theory, it is necessary to observe the small sample behavior of the estimates and the tests under the null and the alternative hypotheses. As shown earlier, the MLEs for the parameters of the general progressively censored (with \( r_i > 0 \)) exponential distributions are free of the censoring schemes; hence, for the purpose of the simulation, we have set \( R_{t,s} = 0, \forall t, s \). To study the dependence
of MLEs on $r_i$, we consider four exponential populations and take random samples according to the simulation scheme given on page 37 of Balakrishnan and Aggarwala (2000) with 10,000 replications. Since we are interested in the scale parameters $\theta_i$’s only, the location (nuisance) parameters are kept fixed at $\mu_i = 1.0, \forall i$ (considered unknown). For simplicity, we have used equal sample sizes ($n_i$) and equal numbers of initially unobserved failures ($r_i$) for each group. Table 1 shows the estimated average bias and mean square errors for the restricted estimators $\hat{\theta}_i$’s for different values of $\theta$’s and different sample sizes.

The restricted MLEs are known to be biased (Robertson, et. al., 1988, p42). In light of this, two important discoveries are made in examining Table 1. As $r_i$ gets larger for a given $n_i$, the biases and the MSEs get much larger. For smaller $r_i$’s, both bias and MSE get smaller when sample size increases, but for larger $r_i$’s bias and MSE are almost unaffected by the sample sizes ($n_i = 10, 20, 30$). Both biases and MSEs get larger when $\theta_i$’s are further apart. Other combinations of sample sizes and $r_i$ values revealed the same information as reported above.

The results of Table 1 suggest that a progressive censoring study with a smaller proportion of initial unobserved failures is more reliable than one with a very large proportion. It is known in group-testing literature that the cost of obtaining individuals is small compared to the cost of testing (Tebbs and Swallow, 2003). We expect a similar situation with progressive censoring as well, and recommend that the study begin early enough so that $r_i$ values are still relatively smaller compared to $n_i$ values as permitted by the study.

In Table 2 the simulated sizes and powers of the restricted and unrestricted test statistics ($T_{01}$ and $T_{02}$) with 10,000 replications are listed using $\alpha = .05$. We observe that the simulated sizes are quite close to the nominal size of $\alpha = .05$ when $r_i$’s are relatively smaller within any of the $n_i$’s. But as $r_i$ gets larger within any $n_i$, we observe larger empirical sizes. These deviations are almost unaffected by the sample
sizes (except when sample sizes are very small, e.g. 10). The situation improves for larger sample sizes. The simulated sizes of the $T_{02}$ test are found to be further away from .05 than those of $T_{01}$ test.

Also in Table 2, one observes under $H_1$, as $r_i$ gets larger relative to $n_i$, the power gets smaller. However, the actual value of the power depends on the configuration of the actual value of $\theta$. We observe higher powers as the number of inequality signs among $\theta_i$’s increase as well as when the $\theta_i$’s are further apart. Also, higher power is obtained for larger sample sizes. The powers of the $T_{01}$ test are higher than those of the $T_{02}$ test, except when $r_i$’s are very large. However, these latter values are clearly not very reliable as demonstrated by the size calculations in the previous paragraph. But for an alternative $\theta$ value such as (2, 4, 6, 8) with $n_i = 50$, $\forall i$, we found the $T_{01}$ test to perform uniformly better than the $T_{02}$ test (not reported in Table 2 for brevity).

Now we apply our procedure on a data set of survival times (some censored) of patients with squamous carcinoma of the oropharynx (Kalbfleisch and Prentice, 1980). The patients were placed in three groups depending upon the degree of lymph node deterioration (or $N$-stage tumor classification). If the three populations correspond to the three lymph node categories, then the survival rates ($1/\theta_i$) are nonincreasing if and only if the $\theta_i$’s are nondecreasing.

We assume that the samples are from exponential populations. To apply our methods on progressive censoring on this data, we have assumed that if an item is censored, then that censoring has taken place at the previous failure time. Although all the initial failures are observed (i.e. $r_i = 0$) in this data, to illustrate the general applicability of our procedure we assume that $r_1 = r_2 = r_3 = 3$. We note that different outcomes would arise if different values of $r_i$’s are chosen. For this data we have $n_1 = 29$, $n_2 = n_3 = 11$, $m_1 = 25$, $m_2 = m_3 = 8$, and, the progressive censoring scheme is given by $R_{1,14} = 1$, $R_{1,22} = 3$, $R_{2,5} = 3$, $R_{3,4} = 1$, $R_{3,5} = 2$, all
other $R_{i,j} = 0$.

The unrestricted estimates are $\hat{\theta}_1 = 391.91$, $\hat{\theta}_2 = 566.60$, $\hat{\theta}_3 = 813.40$, $\hat{\mu}_1 = 84.20$, $\hat{\mu}_2 = 214.56$, $\hat{\mu}_3 = 87.97$. Since the $\hat{\theta}_i$’s satisfy the constraints in $H_1$, they are also the restricted estimates under $H_1$. The estimates under $H_0$ are $\theta^0 = 485.06$, $\mu^0_1 = 74.03$, $\mu^0_2 = 240.53$, $\mu^0_3 = 192.53$. Using these values we find $T_{02} = 2.66$. Since 2.66 is smaller than 4.61, which is the critical value from a chi-square distribution with 2 d.f. with $\alpha = .1$, we cannot reject $H_0$ against the unrestricted alternative $H_2 - H_0$ at 10% significance level. However, when testing $H_0$ against $H_1 - H_0$, we find $T_{01} = 2.66$, which is larger than 2.43 (where $\rho_{12} = .6536$, see discussion of level probabilities prior to Remark 3.1), the critical value at $\alpha = .1$ and hence we reject $H_0$ in this case at the same level. Also, when testing $H_1$ against $H_2 - H_1$, we get $T_{12} = 0$. Thus data support the fact that the survival rates are nonincreasing.

Acknowledgments

The author thanks the referees, the associate editor and the executive editor for their helpful comments.
References


Thomas, D.R. and Wilson, W.M. (1972), Linear order statistic estimation for the two-parameter Weibull and extreme-value distributions from type - II progressively censored samples, *Technometrics*, 14, 679-691.

<table>
<thead>
<tr>
<th>$n_i$</th>
<th>$r_i$</th>
<th>$\hat{\theta}_1$</th>
<th>$\hat{\theta}_2$</th>
<th>$\hat{\theta}_3$</th>
<th>$\hat{\theta}_4$</th>
<th>$\hat{\theta}_1$</th>
<th>$\hat{\theta}_2$</th>
<th>$\hat{\theta}_3$</th>
<th>$\hat{\theta}_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bias</td>
<td>MSE</td>
<td>Bias</td>
<td>MSE</td>
<td>Bias</td>
<td>MSE</td>
<td>Bias</td>
<td>MSE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\theta = (2, 2, 2)$</td>
<td>$\theta = (2, 2, 4)$</td>
<td>$\theta = (2, 3, 4)$</td>
<td>$\theta = (2, 4, 6, 8)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.5316</td>
<td>0.5242</td>
<td>0.5128</td>
<td>0.5306</td>
<td>0.3897</td>
<td>0.3860</td>
<td>0.3783</td>
<td>0.4101</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>0.7627</td>
<td>0.7504</td>
<td>0.7055</td>
<td>0.7077</td>
<td>0.5769</td>
<td>0.5626</td>
<td>0.5021</td>
<td>0.5533</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>1.3812</td>
<td>1.3693</td>
<td>1.3101</td>
<td>1.3168</td>
<td>1.1631</td>
<td>1.1410</td>
<td>1.0652</td>
<td>1.0903</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0.2631</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.1903</td>
<td>0.1433</td>
<td>0.1315</td>
<td>0.2114</td>
</tr>
<tr>
<td>30</td>
<td>5</td>
<td>0.3919</td>
<td>0.1982</td>
<td>0.1903</td>
<td>0.2880</td>
<td>0.2843</td>
<td>0.2433</td>
<td>0.2509</td>
<td>0.2599</td>
</tr>
<tr>
<td>30</td>
<td>8</td>
<td>1.3897</td>
<td>1.3693</td>
<td>1.3101</td>
<td>1.3168</td>
<td>1.1635</td>
<td>1.1410</td>
<td>1.0652</td>
<td>1.0903</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.1433</td>
<td>0.1433</td>
<td>0.1315</td>
<td>0.2114</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0.3919</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.2843</td>
<td>0.2433</td>
<td>0.2509</td>
<td>0.2599</td>
</tr>
<tr>
<td>50</td>
<td>8</td>
<td>1.3897</td>
<td>1.3693</td>
<td>1.3101</td>
<td>1.3168</td>
<td>1.1635</td>
<td>1.1410</td>
<td>1.0652</td>
<td>1.0903</td>
</tr>
<tr>
<td>30</td>
<td>15</td>
<td>0.3919</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.2843</td>
<td>0.2433</td>
<td>0.2509</td>
<td>0.2599</td>
</tr>
<tr>
<td>30</td>
<td>28</td>
<td>1.3897</td>
<td>1.3693</td>
<td>1.3101</td>
<td>1.3168</td>
<td>1.1635</td>
<td>1.1410</td>
<td>1.0652</td>
<td>1.0903</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0.2631</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.1903</td>
<td>0.1433</td>
<td>0.1315</td>
<td>0.2114</td>
</tr>
<tr>
<td>30</td>
<td>5</td>
<td>0.3919</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.2843</td>
<td>0.2433</td>
<td>0.2509</td>
<td>0.2599</td>
</tr>
<tr>
<td>30</td>
<td>8</td>
<td>1.3897</td>
<td>1.3693</td>
<td>1.3101</td>
<td>1.3168</td>
<td>1.1635</td>
<td>1.1410</td>
<td>1.0652</td>
<td>1.0903</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.1433</td>
<td>0.1433</td>
<td>0.1315</td>
<td>0.2114</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0.3919</td>
<td>0.1982</td>
<td>0.1982</td>
<td>0.2880</td>
<td>0.2843</td>
<td>0.2433</td>
<td>0.2509</td>
<td>0.2599</td>
</tr>
<tr>
<td>50</td>
<td>8</td>
<td>1.3897</td>
<td>1.3693</td>
<td>1.3101</td>
<td>1.3168</td>
<td>1.1635</td>
<td>1.1410</td>
<td>1.0652</td>
<td>1.0903</td>
</tr>
</tbody>
</table>

![Table 1: The bias and MSE's of the restricted estimates](image)
Table 2: The size and power of the restricted ($T_{01}$) and unrestricted ($T_{02}$) likelihood ratio test statistics

Sizes of the $T_{01}$ and $T_{02}$ tests

<table>
<thead>
<tr>
<th>$n_i$ = 10</th>
<th>$n_i$ = 20</th>
<th>$n_i$ = 30</th>
<th>$n_i$ = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_i$</td>
<td>$T_{01}$</td>
<td>$T_{02}$</td>
<td>$r_i$</td>
</tr>
<tr>
<td>1</td>
<td>.0708</td>
<td>.0795</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>.0753</td>
<td>.0972</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>.1214</td>
<td>.1892</td>
<td>15</td>
</tr>
</tbody>
</table>

Powers of the $T_{01}$ and $T_{02}$ tests

<table>
<thead>
<tr>
<th>$n_i$ = 10</th>
<th>$n_i$ = 20</th>
<th>$n_i$ = 30</th>
<th>$n_i$ = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_i$</td>
<td>$T_{01}$</td>
<td>$T_{02}$</td>
<td>$r_i$</td>
</tr>
<tr>
<td>1</td>
<td>.2468</td>
<td>.1653</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>.2612</td>
<td>.3714</td>
<td>16</td>
</tr>
</tbody>
</table>

$\theta = (2, 2, 2, 2)$

$\theta = (2, 2, 2, 3)$

$\theta = (2, 3, 3, 4)$

$\theta = (2, 3, 4, 5)$