# The Modified Sudden Death Test: Planning Life Tests with a Limited Number of Test Positions

**Francis G. Pascual**
Department of Statistics
St. Cloud State University
139 Engineering and Computing Center
720 4th Ave S
St. Cloud, MN 56301-4498

**William Q. Meeker**
Department of Statistics and
Center for Nondestructive Evaluation
Iowa State University
Ames, IA 50011

**ABSTRACT**: We present modified sudden death test (MSDT) plans to address the problem of limited testing positions in life tests. A single MSDT involves testing $k$ specimens simultaneously until the $r$th failure. The traditional sudden death test (SDT) is a special case when $r = 1$. The complete MSDT plan consists of $g$ single MSDTs run in sequence. When $r > 1$, there can be up to $r - 1$ idle test positions at any time. We propose testing "standby" specimens in the idle positions and use simulation to gauge the improvement over the basic MSDT plan. We evaluate test plans with respect to the asymptotic variance of maximum likelihood estimators of quantities of interest, total experiment duration and sample size. In contrast to traditional experimental plans, shorter total testing time and smaller sample sizes are possible under MSDT plans.

**KEYWORDS**: Cornish-Fisher expansions, limited test positions, maximum likelihood methods, modified sudden death test, sudden death test, Type I and Type II censoring, Weibull distribution.

# 1 Introduction

## 1.1 Motivation

In fatigue life tests of materials, experimenters are constrained not only by time and the number of test specimens but also by the additional constraint of limited testing positions. The cost of purchasing and maintaining test stands and other equipment limits the number of units that can be tested simultaneously. In fatigue testing shops, for instance, typically only a small number of test machines are available. Laboratory testing of components in an automobile engine requires the use of expensive "test stands." Dynamic testing of rf power devices requires expensive and complicated electronic circuitry to drive each device. When interest centers on the lower part of life distribution, it is unnecessary and even detrimental to accuracy to run all units until failure. In this paper, we study a test procedure addressing these issues, particularly, the issue of a limited number of test positions.

## 1.2 Related Work

Gertsbakh [1] obtains optimal test plans with a limited number of test positions and Type I censoring under an exponential regression model. A complete test plan consists of $g$ stages of fixed lengths. The sum of these lengths serves as another test constraint. In each stage, devices are immediately restored upon failure and runouts (right-censored observations) occur when the stage duration expires. To evaluate plans, he uses the criterion of minimizing the sum of the asymptotic variances of estimators of the regression model coefficients.

A solution to the problem of limited test positions is the traditional sudden death experiment where $k$ units are put on test until the first failure. Johnson [2] discusses how sudden death experiments can significantly reduce testing time and still yield estimates of Weibull quantiles that are just as precise as when all observations are failures. Kececioglu [3] illustrates how sudden death testing can be used to estimate life distribution quantiles for Weibull distributions. We shall see in the discussions below, however, that the tests designed to stop at the second, third, or some subsequent failure in the group can provide a test that is better than the sudden death test.

Suzuki, Ohtsuka and Ashitate [4] study test plans that consist of $g$ simultaneous sudden death experiments with $k$ units. Assuming that fatigue life is distributed Weibull, they investigate plans under different values of $g$ and $k$ through maximum likelihood methods. They also incorporate the idea of Type II censoring in the plans as a generalization. That is, they terminate the experiment at the $p$th sudden death failure. They use a transformed expression for the total test length $L$ to compare different values of $g, k$ and $p$.

We generalize the concept of a sudden death test (SDT) by considering the modified sudden death test (MSDT) that tests $k$ units until the $r$th failure. The MSDT includes the SDT as a special case when $r = 1$. The test plans we consider below consist of running $g$ MSDTs in sequence. We illustrate these test plans by relating them to actual life data sets. We use efficiency, sample size, and total testing time as a set of simple criteria for choosing reasonable test plans.

The modified sudden death test plans discussed below are related to Type II or "failure" censoring where specimens are removed from testing when a certain number has failed. Halperin [5] and Battacharyya [6] show that, under certain regularity conditions, maximum likelihood estimators based on Type II censored data are consistent, asymptotically normally distributed and efficient. Halperin [5] mentions that Type II censoring in destructive tests helps maintain the total monetary loss within budget restrictions. Escobar and Meeker [7] study experimental test plans for accelerated life tests with Type II censored data. They mention that Type II censoring provides more control of the amount of information obtained from the experiment. Escobar and Meeker [8] give an algorithm to compute the variance factors for the Fisher information matrix for the extreme value, normal and logistic distributions with censoring. We use this algorithm to compute large-sample approximate variances for estimators from modified sudden death tests.

## 1.3 Approach

In practice, it is common to test specimens in sequence so that failures are replaced as soon as they occur and nonfailing units are removed after a predetermined length of time (e.g., 100 thousand cycles). We shall refer to this as the traditional experiment and use this as a reference point in studying MSDT plans. For fixed values of $k$ and $r$, we determine the number $g$ of modified sudden death tests required to achieve precision similar to that of a traditional experiment. We will measure precision in terms of asymptotic variances of maximum likelihood (ML) estimators of quantiles of the life distribution. We vary the values of $g$ and $r$ in the test plans to study the tradeoffs between sample size, estimation accuracy, and total duration of testing. We investigate situations where MSDT plans provide improvements over the traditional plan.

When $r > 1$, there is a maximum of $r - 1$ idle test positions at any time during testing. As an improvement on the modified sudden death test, we propose the use of "standby" specimens to be tested in idle test positions. We gauge the improvement over the basic modified sudden death test through simulation studies.

The distribution of test length $L$ under MSDT plans does not have a simple closed form. We approximate quantiles of $L$ by Cornish-Fisher expansion approximations which use cumulants of the life distribution. Cornish and Fisher [9] and Fisher and Cornish [10] provide formulas for approximating quantiles of random variables whose cumulants are known. Johnson, Kotz and Balakrishnan [11] provide formulas for the probability density function (pdf), moments and cumulants of Weibull order statistics.

### 1.4 Overview

Section 2 describes the modified sudden death test and discusses notation and distributional assumptions. Section 4 discusses the distribution of total testing time under an MSDT plan. In Section 5, we present the results of simulation studies to evaluate small-sample properties of MSDT plans. The similarities between small and large-sample properties suggest that large-sample approximations provide a computationally efficient method of comparing MSDT plans. In Section 6, we include standby specimens in MSDT plans to utilize idle test positions and, thus, improve the efficiency of the plans. In Section 7, we apply MSDT plans to practical situations and discuss advantages of using these plans. We evaluate plans in terms of asymptotic and small-sample efficiency of the maximum likelihood estimators, total testing time, and sample size. Section 8 outlines possible areas for further research.

## 2 The Modified Sudden Death Test, Notation and Distributional Assumptions

In traditional tests, specimens are tested in sequence and failures are replaced as soon as they occur. Unfailed units (runouts) are removed after a certain length of time (time or Type I censoring). There is, however, some difficulty in deciding when to take specimens off testing. First, predetermined censoring times do not guarantee enough failures to carry out analysis at a desirable level of accuracy. Second, choosing a Type I censoring time requires knowledge of the life distribution and test results are highly sensitive to the choice. Unless prior knowledge of the life distribution is available or the censoring times are determined dynamically in the progress of the experiment, the above strategy can lead to results that fall short of expectations. Below we study the modified sudden death test (extended Type II censoring) as an alternative to the extended Type I strategy in traditional tests. We also specify the scope of the problem and introduce notation that we will be using henceforth.

A single MSDT involves testing $k$ specimens simultaneously until the $r$th failure occurs at which time all remaining specimens are removed. This results in $r$ failures and $k - r$ runouts. The traditional SDT is a special case when $r = 1$. The complete MSDT, proposed here, consists of running $g$ single MSDTs in sequence. Thus, the test has a total of $gr$ failures and $gk - gr$ runouts out of $gk$ specimens. We will use the notation MSDT($g,\ k,\ r$) to denote this experiment. A single MSDT corresponds to MSDT($1,\ k,\ r$).

The special case MSDT($1,\ k,\ r$) is known as Type II or "failure" censoring in the literature. In

Type II censoring, testing is terminated when a certain number of specimens on test has failed. In general, an MSDT($g,\ k,\ r$) can be viewed as a sequence of Type II censored life tests. Thus, the properties of estimators from MSDTs follow from those of Type II censored life tests.

We use the following approach to study the performance of MSDT($g,\ k,\ r$) plans under different values of $g, k$ and $r$. Suppose we are interested in estimating the $q$ quantile of life distribution under a fixed set of experimental conditions. Let $Y_{ij}$ be the $j$th observation in the $i$th single MSDT for $i = 1, \ldots, g$ and $j = 1, \ldots, k$. Assume that $Y_{ij}, i = 1, \ldots, g, j = 1, \ldots, k$ are identically and independently distributed Weibull with scale and shape parameters $\alpha$ and $\beta$, respectively, where $\alpha, \beta > 0$. The probability density and cumulative distribution functions of $Y_{ij}$ are given by

$$f_Y(y) = \beta\left(\frac{y}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{y}{\alpha}\right)^{\beta}\right],\ y \geq 0,$$

and

$$F_Y(y) = 1 - \exp\left[-\left(\frac{y}{\alpha}\right)^{\beta}\right],\ y \geq 0,$$

respectively. For the discussions that follow, we will assume, without loss of generality, that $\alpha = 1$ since the desired scale is achieved by multiplying the appropriate constant to $Y_{ij}$.

# 3 The Asymptotic Variance of the Maximum Likelihood Estimator of Population Quantiles

Suppose that the objective of the life test is to estimate the $q$ quantile of $Y_{ij}$ by ML methods. Let $y_q$ be the $q$ quantile of $Y_{ij}$ and $\widehat{y}_q$ be its ML estimator. The asymptotic variance of $\log(\widehat{y}_q)$ is given by the equation

$$gk\beta^2\,\mathrm{AVar}(\log\widehat{y}_q) = \frac{1}{f_{11}f_{22} - f_{12}^2}\{f_{22} + f_{11}[\log(-\log(1-q))]^2 - 2f_{12}[\log(-\log(1-q))]\} \tag{1}$$

where

$$gk\beta^2\begin{bmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{bmatrix}$$

is the Fisher information matrix under MSDT($g,\ k,\ r$) and log denotes natural logarithm. Escobar and Meeker [8] give numerical algorithms to compute the $f_{ij}$'s. The right hand side of (1) depends on the proportion failing $p$ through the $f_{ij}$'. For MSDT plans, $p = r/k$. Figure 1 plots $gk\beta^2\mathrm{AVar}(\log\widehat{y}_q)$ versus $q$ for different values of $p$. Our results will show that a general rule of thumb in selecting a "good" MSDT plan is to choose the smallest $r$ so that $r/k$ is at least $q$ and, if possible, as large as $2q$.

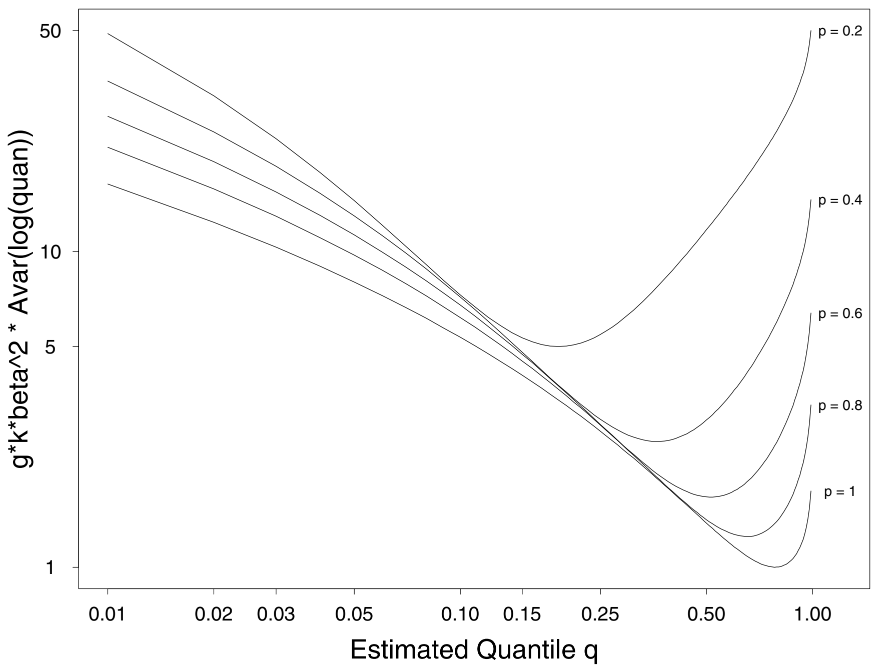


Figure 1: *Plot of Variance Factor versus q for the ML Estimators* $\log(\widehat{y}_q)$

# 4 Total Testing Time $L$ under MSDT$(g,\ k,\ r)$

This section describes the distribution of total testing time $L$ under the MSDT$(g,\ k,\ r)$ plan. We give expressions for $L$ in terms of the sample data and formulas for its mean and variance. We obtain approximations of the quantiles of $L$ by Cornish-Fisher expansions.

Let $Y_{i(r)}$ denote the $r$th order statistic of $Y_{i1}, \ldots, Y_{ik}$ for $i = 1, 2, \ldots, g$. Johnson et al. [11] provide formulas for the pdf and moments of Weibull order statistics. The $s$th moment of $Y_{i(r)}$ is given by

$$m_{s(r)} \equiv \mathrm{E}[Y_{i(r)}^{s}] = \left(1 + \frac{s}{\beta}\right) \sum_{j=0}^{r-1} \frac{(-1)^j \binom{r-1}{j}}{(k-r+j+1)^{1+\frac{s}{\beta}}}. \tag{2}$$

The mean and variance of $Y_{i(r)}$ are given by

$$\mu_{(r)} = m_{1(r)} \tag{3}$$

and

$$\sigma_{(r)}^2 = m_{2(r)} - [m_{1(r)}]^2, \tag{4}$$

respectively.

The total length $L$ of MSDT$(g,\ k,\ r)$ can be written as

$$L = \sum_{i=1}^{g} Y_{i(r)}.$$

The mean and variance of $L$ are $\mu_L = g\mu_{(r)}$ and $\sigma_L^2 = g\sigma_{(r)}^2$, respectively.

Under the MSDT$(g,\ k,\ r)$ plan, the distribution of $L$ does not have a simple form. We approximate the quantiles of $L$ by Cornish-Fisher expansions which use the cumulants of $L$. Let $L'$ be the standardized version of $L$, that is, $L' = (L - \mu_L)/\sigma_L$. Let $\{\xi_i\}_{i=1}^{\infty}$ be the cumulants of $L'$. It can be shown that

$$\begin{aligned}
\xi_1 &= 0 \\
\xi_2 &= 1 \\
\xi_3 &= \frac{g}{\sigma_{(r)}^3}[m_{3(r)} - 3m_{1(r)}m_{2(r)} + 2m_{1(r)}^3] \\
\xi_4 &= \frac{g}{\sigma_{(r)}^4}[m_{4(r)} - 4m_{1(r)}m_{3(r)} - 3m_{2(r)}^2 + 12m_{1(r)}^2 m_{2(r)} - 6m_{1(r)}^4]
\end{aligned}$$

Let $L'_q$ and $z_q$ denote, respectively, the $q$ quantiles of $L'$ and a standard normal random variable. A Cornish-Fisher expansion approximation of $L'_q$ is given by

$$L'_q \doteq z_q + \frac{1}{6}\xi_3(z_q^2 - 1) + \frac{1}{24}\xi_4(z_q^3 - 3z_q) - \frac{1}{36}\xi_3^2(2z_q^3 - 5z_q). \qquad (5)$$

Cornish and Fisher [9] and Fisher and Cornish [10] provide the derivation of this approximation. An approximation for $L_q$, the $q$ quantile of $L$, is $L_q \doteq \mu_L + \sigma_L L'_q$. This provides a more computationally efficient method of obtaining quantiles than simulation.

# 5 Simulation Studies to Evaluate MSDT$(g,\ k,\ r)$ Plans under the Weibull Distribution

This section uses simulation to present a broader study of small-sample behavior of the ML estimators of quantiles under MSDT plans. The results here also justify the use of asymptotic variance as a computationally efficient tool for comparing MSDT designs.

For the simulation study below, we use the Weibull scale $\alpha = 19.59$ and shape $\beta = 2.35$ (from the laminate panel example below) as planning values. We are interested in the MSDT$(10,\ 5,\ r)$ plans for estimating the $q$ quantile of the life distribution. In one simulation replication, we draw 10 random samples of size 5 from the Weibull distribution and obtain the corresponding observations (failures/runouts) under MSDT$(10,\ 5,\ r)$ for $r = 1, \ldots, 5$. For each $r$, we compute the ML estimate of $\log(y_q)$ where $y_q$ is the q quantile. We repeat this procedure 4000 times. We use the variance of the 4000 estimates to compare plans and gauge the improvement that larger values of $r$ have over smaller ones. For purposes of consistency and comparisons, our approach here parallels that used to construct Figure 1.

Under the plan MSDT(10, 5, $r$), there are 50 specimens tested yielding $10r$ failures and $10(k-r)$ runouts. Table 1 gives the mean $\mu_L$, standard deviation $\sigma_L$ and quantiles of length $L$ of testing using formulas in Section 4. Fatigue life is given in millions of cycles.

Table 1: *Quantiles, Mean and Standard Deviation of Test Length under the Plans MSDT*$(10, 5, r)$ *for* $r = 1, \ldots, 5$

| Plan | $L_{05}$ | $L_{50}$ | $L_{95}$ | $\mu_L$ | $\sigma_L$ |
|---|---|---|---|---|---|
| MSDT(10, 5, 1) | 67 | 87 | 109 | 88 | 13 |
| MSDT(10, 5, 2) | 110 | 131 | 153 | 131 | 13 |
| MSDT(10, 5, 3) | 147 | 169 | 192 | 169 | 14 |
| MSDT(10, 5, 4) | 186 | 210 | 236 | 211 | 15 |
| MSDT(10, 5, 5) | 239 | 269 | 301 | 269 | 19 |

Figure 2 gives a plot of the simulated values of $gk\beta^2\mathrm{Var}(\log \widehat{y}_q)$ versus $q$ for MSDT$(10,\ 5, r)$ plans with $r = 1, \ldots, 5$. The similarity between Figures 1 and 2 suggests that the asymptotic variance of ML estimators provides an adequate guideline for comparing MSDT plans.

Figure 2 shows that the sudden death ($r = 1$) plan does not perform as well as the alternative plans, although it competes well for $q$ quantiles for $q$ in the vicinity of 0.10 to 0.20. As expected, larger values of $r$ are necessary to estimate larger quantiles with improved precision. The intuitive rule of choosing the smallest $r$ so that the proportion failing $r/5$ exceeds the value $q$ of interest is illustrated in Figure 2.

Based on precision and length of testing, MSDT(10, 5, 2) and MSDT(10, 5, 3) plans are reasonable. They are competitive with the other plans, particularly if interest is in lower quantiles, as is often the case in actual applications.

## 6 Improving the Efficiency of the MSDT Plans

When failures occur under the MSDT($g$, $k$, $r$) plan, the corresponding test positions are idle until the $r$th failure. In general, when $r > 1$, there are at most $r - 1$ idle test positions at any given time during testing. This causes some inefficiency.

To improve the efficiency of MSDT plans we consider testing "standby" specimens in test positions when they become vacant. At the start of the experiment, we divide specimens into two

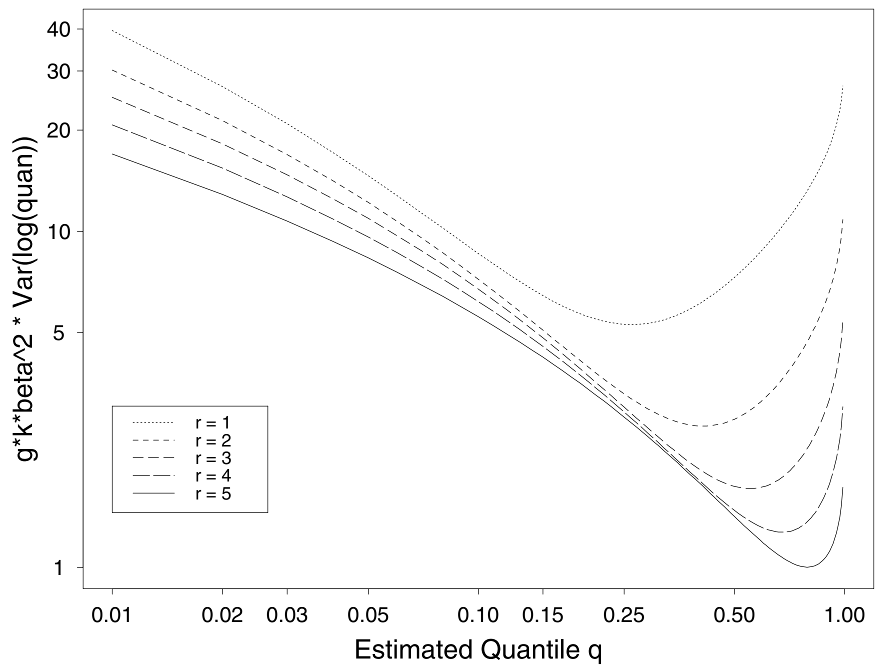


Figure 2: *Plot of Variance Expression Versus q for Simulated ML Estimates of* $\log(y_q)$ *under the Plans MSDT*$(10, 5, r)$ *for* $r = 1, \ldots, 5$

groups. Group 1 consists of the $gk$ units to be tested under MSDT$(g,\ k,\ r)$ and Group 2, a number of units called "standbys" to be tested in idle positions. We propose the following procedure to test the standby units.

- When a failure (not the $r$th) occurs, take a standby specimen from Group 2 and test it until it fails or until the $r$th failure from the original set of specimens occurs.
- If a standby fails before the $r$th failure, replace it with another standby specimen.
- When the $r$th failure occurs, remove all units including standbys and test a fresh batch of $k$ specimens from Group 1.
- Nonfailing standby specimens will continue to be tested in the same test stands in which they were first tested, as soon as their stands become idle again. Each standby specimen will be tested until a specified amount of running time (or number of cycles) $t_{q_c}$.
- The experiment ends when the $r$th failure occurs in the $g$th batch.

The sample size and the number of failures are random under this procedure. On the other hand, the improved plan yields $g(k - r) + r - 1$ runouts. The distribution of test length $L$ remains the same as before because the standbys are tested without adding testing time to the original plan.

Consider censoring standbys at the $q_c$ quantile $t_{q_c}$ of the life distribution for different values of $q_c$. We use IMSDT$(g,\ k,\ r,\ q_c)$ to denote the improved experimental plan that combines MSDT$(g,\ k,\ r)$ and standbys censored at $t_{q_c}$. Note that IMSDT$(g,\ k,\ r,\ 0)$ is equivalent to MSDT$(g,\ k,\ r)$ and IMSDT$(g,\ k,\ r,\ 1)$ is an experimental plan in which standbys are not censored at all except at the $g$th (last) batch in the test. Figure 3 illustrates a possible experimental scenario under the plan IMSDT$(g,\ k = 5,\ r = 3,\ q_c = 1)$.

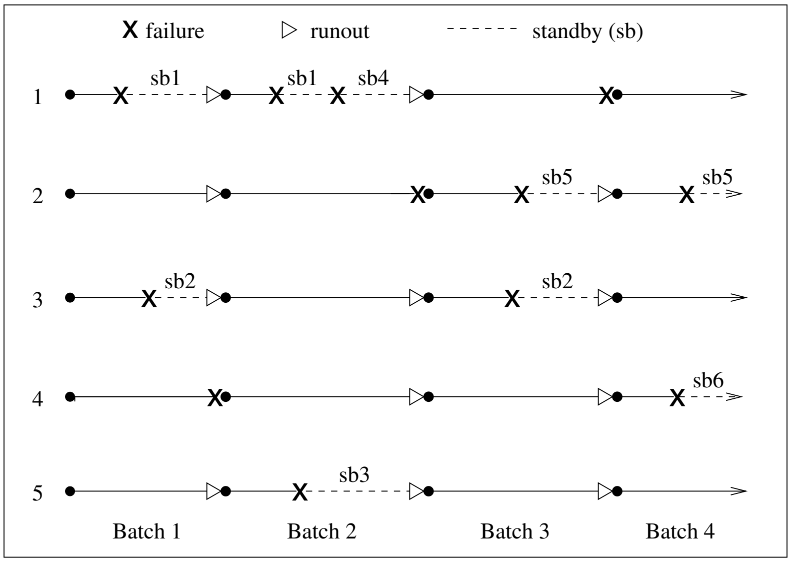


Figure 3: *An Experimental Scenario under IMSDT(*$g,\ k = 5,\ r = 3,\ q_c = 1$*)*

Recall the simulated tests in the previous section for the MSDT$(10,\ 5,\ r)$ plans. We use a similar simulation to evaluate the improvements provided by IMSDT$(10,\ 5,\ r,\ q_c)$ over MSDT$(10,\ 5,\ r)$ for $q_c = 0.40, 0.60, 0.80, 1$. For this purpose, we use the Weibull planning values $\alpha = 1$ and $\beta = 2.35$, and assume that specimen replacement is instantaneous. To measure the improvement, we compute the percent decrease in the variance of the ML estimate of $\log(y_q)$ for each $r$ relative to the MSDT plan.

Table 2 gives information on mean sample sizes $\bar{n}$ and mean proportions failing $\bar{p}_f$ under IMSDT$(10,\ 5,\ r,\ q_c)$ for $r = 2, 3, 4, 5$ based on simulation. The first row of the table corresponds to MSDT$(10,\ 5,\ r)$.

Figure 4 gives a plot of the percent decrease in the variance of the ML estimates under IMSDT(10, 5, 2, $q_c$). Figure 5 is an analogous plot for IMSDT$(10,\ 5,\ 5,\ q_c)$. The figures indicate greater improvements for larger values of $r$, as expected. For these values of $r$, there are more idle positions in which to test standby specimens. For instance, Table 2 shows that there are 26 to 37 more specimens tested under IMSDT$(10,\ 5,\ 5,\ q_c)$ than under IMSDT$(10,\ 5,\ 2,\ q_c)$. If testing time and availability and cost of specimens are not restrictive, the plan IMSDT$(10,\ 5,\ r,\ q_c)$ for large $r$

Table 2: *Average Sample Sizes* $\bar{n}$ *and Proportions Failing* $\bar{p}_f$ *under the Plans IMSDT*$(10, 5, r, q_c)$ *for* $r = 1, \ldots, 5$ *and* $q_c = 0.40, 0.60, 0.80, 1$

| | $r=2$ | | $r=3$ | | $r=4$ | | $r=5$ | |
|---|---|---|---|---|---|---|---|---|
| $q_c$ | $\bar{n}$ | $\bar{p}_f$ | $\bar{n}$ | $\bar{p}_f$ | $\bar{n}$ | $\bar{p}_f$ | $\bar{n}$ | $\bar{p}_f$ |
| 0.00 (MSDT) | 50.00 | 0.40 | 50.00 | 0.60 | 50.00 | 0.80 | 50.00 | 1.00 |
| 0.40 | 53.95 | 0.39 | 60.42 | 0.56 | 70.63 | 0.67 | 89.57 | 0.72 |
| 0.60 | 53.47 | 0.40 | 59.12 | 0.58 | 68.06 | 0.73 | 84.40 | 0.81 |
| 0.80 | 53.27 | 0.41 | 58.47 | 0.60 | 66.58 | 0.77 | 81.52 | 0.89 |
| 1.00 | 53.11 | 0.42 | 58.09 | 0.62 | 65.83 | 0.80 | 79.94 | 0.95 |

provides the best results.

For the example above, when test specimens are inexpensive and testing standby units is convenient, useful gains in efficiency are possible. Figures 4 and 5 show that $q_c = 1$ is generally a good choice for estimating low and high quantiles. For intermediate quantiles, other choices for $q_c$ are better. For larger values of $r$, $q_c = 0.40$ is a competitive alternative to $q_c = 1$. When $r = 5$, censoring standbys at the 0.40 quantile is best for estimating quantiles below the 0.75 quantile.

The results here do not take anything away from the practicality of MSDT plans with small values of $r$. If time is constrained and if experimental units are expensive, small values of $r$ provide appropriate plans for estimating small quantiles. When $r$ is small, investing in standbys may not yield worthwhile dividends in terms of improved estimation precision because test stand idle times are short and standbys will not increase sample size significantly. MSDT plans in this case are adequate.

We saw above that IMSDT improvements over MSDT vary with the choice of quantile at which to censor standbys. To optimize the use of standbys in testing, one must have good distribution planning values because quantiles depend heavily on these values. If there is a high degree of uncertainty in one's planning values, choosing $q_c = 1$ is, in general, a conservative strategy to follow.

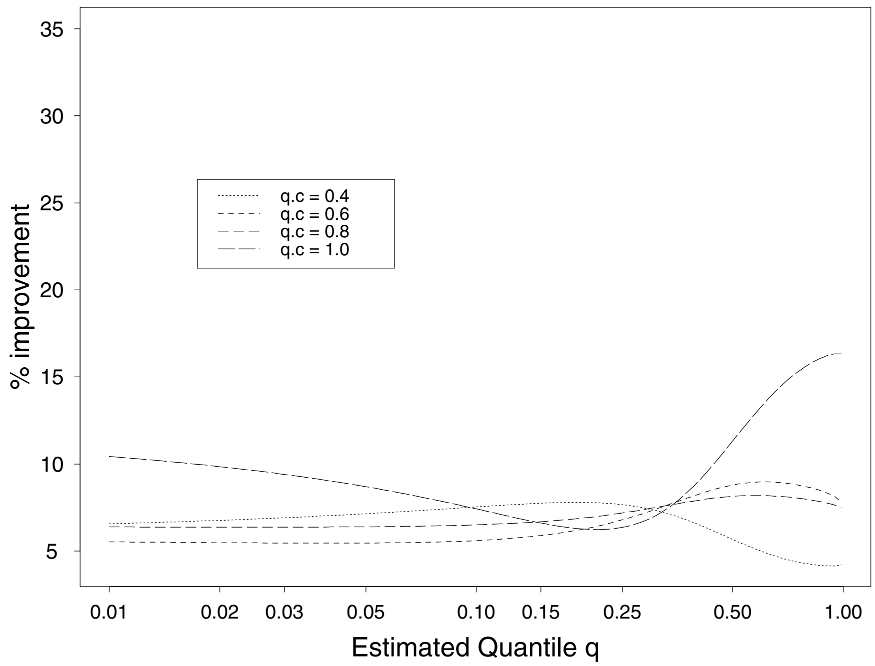


Figure 4: *Plot of Percent Decrease in the Variance of ML Estimates of* $\log(y_q)$ *under the Plan IMSDT(*$10, 5, 2, q_c$*)*

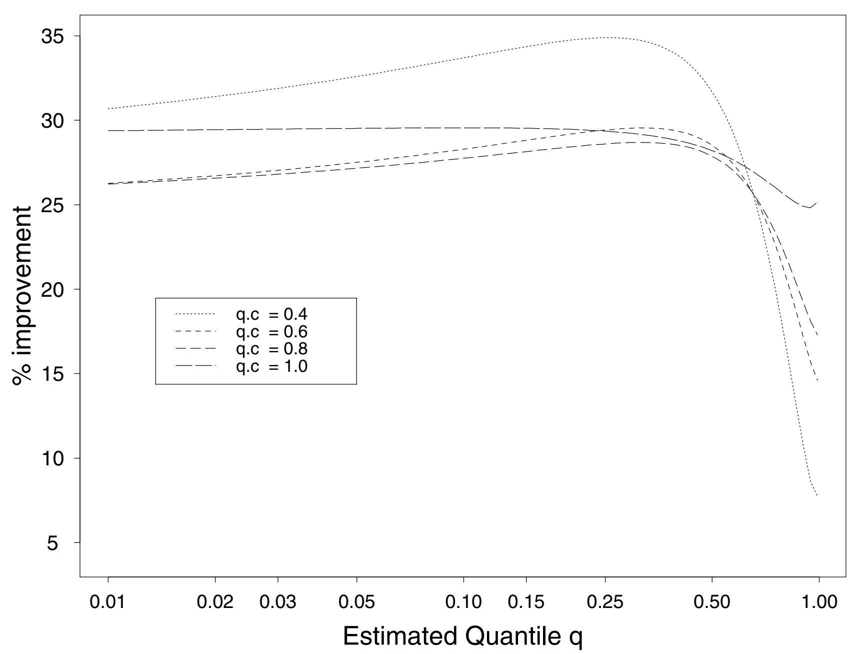


Figure 5: *Plot of Variance Expression versus q for ML Estimates of* $\log(y_q)$ *under the Plan IMSDT(*$10, 5, 5, q_c$*) for* $q_c = 0.40, 0.60, 0.80, 1$

# 7 MSDT and IMSDT Plans to Estimate the $q$ Quantile of Life Distribution

Below we study the MSDT($g,\ k,\ r$) and IMSDT($g,\ k,\ r,\ q_c$) plans in several situations chosen to correspond with actual applications. We relate these test plans to actual life test data sets assuming a particular number of test positions. We determine the values of $g$ so that the MSDT plans achieve about the same precision as the actual life test as measured by asymptotic variance of ML estimators. We also study corresponding IMSDT plans to investigate improvements over the MSDT plans. The examples below show instances where MSDT and IMSDT have advantages over traditional test plans. They also illustrate tradeoffs between precision, sample size, and test duration in determining feasible plans.

## 7.1 Numerical Examples

For traditional tests, we assume that test specimens are tested in sequentially in $k$ test positions so that failures are replaced as soon as they occur and surviving units are removed after a predetermined length of time $t_c$. If planning values for model parameters are available at the planning stage, the following procedure can be used to select appropriate MSDT and IMSDT plans.

1. For $r = 1, \ldots, k$, we determine the value of $g$ so that MSDT($g,\ k,\ r$) achieves approximately the same precision [i.e., AVar($\log(\widehat{y}_q)$)] as the traditional experiment. Let $f(p)$ be the right hand side of (1) for proportion failing $p$. Suppose that in the traditional experiment, $n$ is the sample size and $p_f$ is the expected proportion failing. We compute $g$ using

$$g = \frac{nf(\frac{r}{k})}{kf(p_f)}.$$

   Smaller sample sizes or, equivalently, smaller values of $g$, are desirable because of constraints on both time and number of test specimens. If sample size is not restrictive, we can consider higher values of $g$.

2. We reduce test lengths under these plans by using smaller values of $g$ or $r$. Comparisons provide insight about the tradeoffs between test length and relative efficiency.

3. We improve the efficiency of the MSDT plans by considering the corresponding IMSDT plans.

### 7.1.1 Sensitivity of Traditional and MSDT Plans to Model Misspecification

The censoring time $t_c$ in traditional plans often corresponds to a proportion failing $p_f$. Because $p_f$ depends on the model parameters, an appropriate choice for $t_c$ relies heavily on the planning values. If under the planning values, $p_f$ is smaller than its true value (there is more censoring than expected at $t_c$), the value of $gk\beta^2\text{AVar}(\log\widehat{y}_q)$ will be higher than expected. This variance expression is constant for MSDT plans because these plans are based on a fixed proportion failing and not on a fixed censoring time. Figure 6 plots $gk\beta^2\text{AVar}(\log\widehat{y}_q)$ versus the proportion failing $p_f$ for traditional and MSDT plans with $k = 5$ test positions. It is clear from the plot that traditional plans are not robust to model misspecification in which $p_f$ is overestimated. There is more control of the amount of information derived from MSDT plans in that $gk\beta^2\text{AVar}(\log\widehat{y}_q)$ is already known at the planning stage.

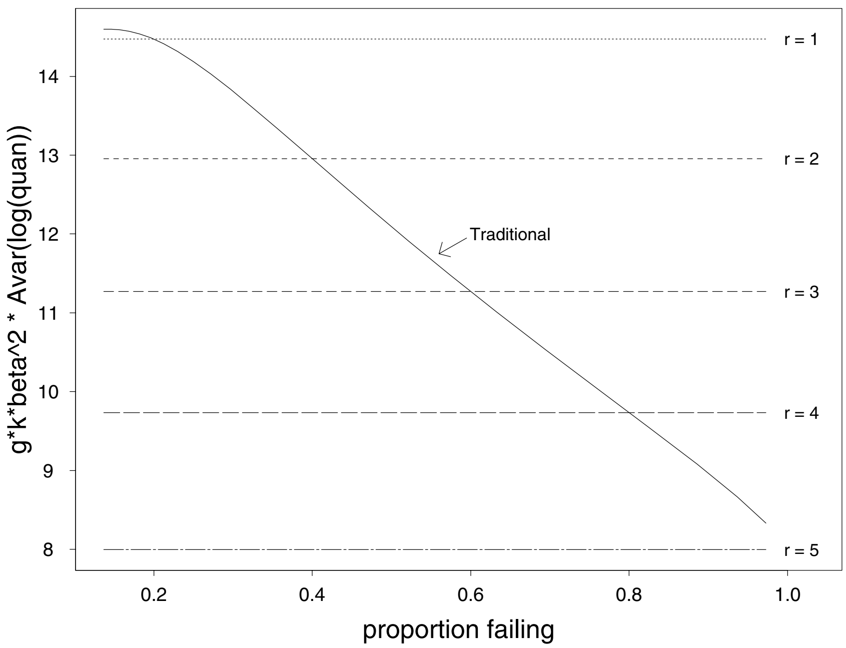


Figure 6: *Plot of Variance Factor versus* $p_f$ *for the ML Estimators* $\log(\widehat{y}_{0.05})$ *under Traditional and MSDT Plans (*$k = 5$*)*

#### 7.1.2 Example 1: Laminate Panel Fatigue Data

Consider the laminate panel data given by Shimokawa and Hamaguchi [12]. This data set was the result of four-point out-of-plane bending tests of carbon eight-harness-satin/epoxy laminate specimens. For our purposes, we will use the 25 observations taken at stress 270 MPa. Seventeen specimens failed, while 8 were censored at about 20 million cycles. Fitting a Weibull distribution gives ML estimates of the scale and shape parameters $\widehat{\alpha} = 19.59$ (million cycles) and $\widehat{\beta} = 2.35$, respectively. These estimates will be used as planning values.

For the traditional test plan, 25 specimens are tested until failure or until 20 million cycles.

Under this plan with the planning values given above, the proportion failing is $p_f = 0.65$. Suppose that there are $k = 5$ test positions available. Table 3 gives the values of $g$ needed and the resulting sample sizes $n$ for MSDT($g$, 5, $r$) to achieve the same precision as the traditional test in estimating the 0.05 quantile of the life distribution.

Table 3: *Sample Sizes, Quantiles, Mean and Standard Deviation of Test Length, and Asymptotic Variance of the ML Estimator of* $\log(\widehat{y}_{0.05})$ *under the Plans MSDT*$(g, k = 5, r)$ *for* $r = 1, \ldots, 5$

| Plan | $n$ | $L_{05}$ | $L_{50}$ | $L_{95}$ | $\mu_L$ | $\sigma_L$ | $CV(\%)$ | AVar($\log(\widehat{y}_{0.05})$) |
|---|---|---|---|---|---|---|---|---|
| Traditional | 25 | 74 | 84 | 93 | 84 | 5.8 | 6.9 | 0.0788 |
| MSDT(7, 5, 1) | 35 | 45 | 61 | 79 | 61 | 10.5 | 17.1 | 0.0749 |
| MSDT(6, 5, 2) | 30 | 62 | 78 | 96 | 79 | 10.1 | 12.8 | 0.0782 |
| MSDT(6, 5, 3) | 30 | 84 | 101 | 119 | 102 | 10.6 | 10.5 | 0.0681 |
| MSDT(5, 5, 4) | 25 | 88 | 105 | 123 | 105 | 10.7 | 10.2 | 0.0705 |
| MSDT(4, 5, 5) | 20 | 89 | 107 | 128 | 108 | 12.0 | 11.1 | 0.0725 |

The table includes the 0.05 quantile $L_{05}$, median $L_{50}$, 0.95 quantile $L_{95}$, mean $\mu_L$ and standard deviation $\sigma_L$ of total testing time $L$ under MSDT plans and the traditional experiment. For the MSDT plans, the test length mean and variance are computed using formulas given in Section 4 and the quantiles are approximated by Cornish-Fisher expansions. Because there is no systematic unit-replacement scheme in the traditional experiment, we simulate it 1000 times and compute the mean, standard deviation and quantiles of the total test length. Fatigue life is in millions of cycles. The table gives the coefficient of variation $CV$, the standard deviation as a percentage of the mean. The $CV$ is a unitless quantity that is useful in comparing relative variabilities of testing lengths under different test plans. The asymptotic variance of the ML estimator of the 0.05 quantile is also given for each test plan.

Table 3 shows that any MSDT plan has competitive sample size and test length. For 10 specimens more, the SDT plan MSDT(7, 5, 1) provides the same precision as the traditional test in less time on the average. The plan MSDT(6, 5, 2) still has a smaller $\mu_L$ and requires only 5 specimens more than the traditional plan. MSDT(4, 5, 5) reduces sample size from 25 to 20, but requires more time.

We investigate MSDT plans with $g = 4$ or 5 and improve upon them by considering the corresponding IMSDT plans. Recall that MSDT and IMSDT plans have the same test length. Below, we choose $q_c = 1$ for the IMSDT plans. Figures 4 and 5 suggest that other values of $q_c$ may yield more improvement depending on $r$ and the quantile being estimated. However, $q_c = 1$ is a conservative strategy to follow.

Table 4 provides information on test length distributions under MSDT(4, 5, $r$) (or IMSDT(4, 5, $r$, $q_c = 1$)) and MSDT(5, 5, $r$) (or IMSDT(4, 5, $r$, $q_c = 1$)) for $r = 1, \ldots, 5$ based on Cornish-Fisher expansion approximations. MSDT(4, 5, 3) and MSDT(5, 5, 2) yield shorter test lengths than the traditional plan on the average. MSDT(5, 5, 3) has mean test length equal to that of the traditional plan. The reductions in test length under MSDT plans, however, are at the price of losing efficiency in estimating the 0.05 quantile. We improve the efficiency by testing standby specimens in idle positions.

Table 4: *Quantiles, Mean and Standard Deviation of Test Length and Variance of* 1000 *ML Estimates of* $\log(\widehat{y}_{0.05})$ *under Traditional, MSDT and IMSDT* ($q_c = 1$) *Plans with* $k = 5$

| Plan | | $L_{05}$ | $L_{50}$ | $L_{95}$ | $\mu_L$ | $\sigma_L$ | $CV(\%)$ | Var($\log(\widehat{y}_{0.05})$) MSDT | Var($\log(\widehat{y}_{0.05})$) IMSDT |
|---|---|---|---|---|---|---|---|---|---|
| Traditional | | 74 | 84 | 93 | 84 | 5.8 | 6.9 | 0.0850 | 0.0850 |
| $g = 4$ | $r = 1$ | 22 | 35 | 49 | 35 | 7.9 | 22.6 | 0.1276 | – |
| | $r = 2$ | 39 | 52 | 66 | 52 | 8.2 | 15.7 | 0.1089 | 0.1003 |
| | $r = 3$ | 54 | 68 | 82 | 68 | 8.7 | 12.8 | 0.0987 | 0.0847 |
| | $r = 4$ | 69 | 84 | 100 | 84 | 9.6 | 11.4 | 0.0890 | 0.0698 |
| | $r = 5$ | 89 | 107 | 128 | 108 | 12.0 | 11.1 | 0.0761 | 0.0517 |
| $g = 5$ | $r = 1$ | 30 | 43 | 59 | 44 | 8.9 | 20.2 | 0.1024 | – |
| | $r = 2$ | 51 | 65 | 81 | 66 | 9.2 | 14.0 | 0.0874 | 0.0812 |
| | $r = 3$ | 70 | 84 | 101 | 85 | 9.7 | 11.5 | 0.0802 | 0.0658 |
| | $r = 4$ | 88 | 105 | 123 | 105 | 10.7 | 10.2 | 0.0716 | 0.0579 |
| | $r = 5$ | 113 | 134 | 157 | 135 | 13.4 | 10.0 | 0.0612 | 0.0421 |

We simulate the traditional, MSDT and IMSDT experiments 1000 times each and obtain the ML estimates of the 0.05 quantile. Table 4 gives the variances of the estimates for the plans. Table 5 gives information on mean sample sizes $\bar{n}$ and mean proportions failing $\bar{p}_f$ under the MSDT and IMSDT plans based on the simulations.

IMSDT(4, 5, 3, $q_c = 1$) and IMSDT(5, 5, 2, $q_c = 1$) have better efficiencies and smaller test length means than the traditional plan. IMSDT(4, 5, 3, $q_c = 1$) has an average sample size smaller than the sample size of 25 in the traditional plan. The average sample size of IMSDT(5, 5, 2, $q_c = 1$) is slightly above 25. Generally, IMSDT(4, 5, 3, $q_c = 1$) and IMSDT(5, 5, 2, $q_c = 1$) are good alternatives to the traditional plan.

For estimating higher quantiles, IMSDT plans may not simultaneously yield shorter test lengths, smaller sample sizes and better efficiency relative to the traditional plan. For example, for estimating

Table 5: *Average Sample Sizes* $\bar{n}$ *and Proportions Failing* $\bar{p}_f$ *under MSDT and IMSDT* ($q_c = 1$) *Plans with* $k = 5$

| | Plan | $r=2$ | | $r=3$ | | $r=4$ | | $r=5$ | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\bar{n}$ | $\bar{p}_f$ | $\bar{n}$ | $\bar{p}_f$ | $\bar{n}$ | $\bar{p}_f$ | $\bar{n}$ | $\bar{p}_f$ |
| $g=4$ | MSDT | 20.00 | 0.40 | 20.00 | 0.60 | 20.00 | 0.80 | 20.00 | 1.00 |
| | IMSDT | 21.61 | 0.40 | 23.94 | 0.58 | 27.36 | 0.74 | 29.42 | 0.88 |
| $g=5$ | MSDT | 25.00 | 0.40 | 25.00 | 0.60 | 25.00 | 0.80 | 25.00 | 1.00 |
| | IMSDT | 26.86 | 0.40 | 29.61 | 0.59 | 33.75 | 0.76 | 41.18 | 0.90 |

the population median (0.50 quantile), IMSDT$(4, 5, 4, q_c)$ and IMSDT$(5, 5, 4, q_c)$ have test length distributions similar to that of the traditional plan, better efficiencies, but slightly larger sample sizes. IMSDT$(4, 5, 3, q_c)$, on the other hand, has shorter test lengths and smaller sample sizes in exchange for a 15% loss in efficiency.

#### 7.1.3 Example 2: Annealed Aluminum Wire Fatigue Data

Shen [13] analyzes a fatigue data set on annealed aluminum wire. There were 20 observations, all failures, at stress level 294.3 MPa. The planning values are given by the ML estimates $\widehat{\alpha} = 9.2$ thousand cycles and $\widehat{\beta} = 6.22$ of the Weibull scale and shape parameters, respectively. A large shape parameter value such as 6.22 is not typical of fatigue data. Large values of $\beta$ (small CV values) imply that failures occur close together.

Assume that $k = 5$ test positions are available and that for the traditional experiment, all units are tested until failure. Again, we consider the MSDT$(g,\ 5,\ r)$ plans and find the values of $g$ that give the same precision as the traditional test. Table 6 gives the values of $g$ needed for estimating the 0.05 quantile. It also gives information on the test length distribution and the asymptotic variance of the ML estimator of the 0.05 quantile for each plan. The test length information for the traditional plan is based on 1000 simulations of the experiment and that for MSDT plans is based on formulas in Section 4. It is clear from Table 6 that MSDT$(4,\ 5,\ 5)$ provides the best MSDT plan for estimating the 0.05 quantile because it is the shortest, on the average, and has the smallest sample size.

We study MSDT and IMSDT plans with $g = 4$ or 5. Table 7 gives information on test length under MSDT$(4,\ 5,\ r)$ (or IMSDT$(4,\ 5,\ r,\ q_c = 1)$) and MSDT$(5,\ 5,\ r)$ (or IMSDT$(5,\ 5,\ r,\ q_c = 1)$) for $r = 1,\ \ldots,\ 5$. The table also gives the variances of the ML estimators of 0.05 quantiles based on 1000 simulations of each plan. IMSDT sample sizes and proportions failing based on these simulations are given in Table 8.

IMSDT$(4,\ 5,\ 5,\ q_c = 1)$, IMSDT$(5,\ 5,\ 3,\ q_c = 1)$, IMSDT$(5,\ 5,\ 4,\ q_c = 1)$ and IMSDT$(5,\ 5,\ 5,\ q_c =$

Table 6: *Sample Sizes, Quantiles, Mean and Standard Deviation of Test Length, and Asymptotic Variance of the ML Estimator of* $\log(\widehat{y}_{0.05})$ *under the Plans MSDT*$(g, k = 5, r)$ *for* $r = 1, \ldots, 5$

| Plan | $n$ | $L_{05}$ | $L_{50}$ | $L_{95}$ | $\mu_L$ | $\sigma_L$ | $CV(\%)$ | AVar($\log(\widehat{y}_{0.05})$) |
|---|---|---|---|---|---|---|---|---|
| Traditional | 20 | 34 | 38 | 41 | 38 | 1.9 | 5.0 | 0.0103 |
| MSDT(8, 5, 1) | 40 | 47 | 53 | 58 | 53 | 3.5 | 6.6 | 0.0094 |
| MSDT(7, 5, 2) | 35 | 50 | 55 | 59 | 55 | 2.6 | 4.7 | 0.0096 |
| MSDT(6, 5, 3) | 30 | 48 | 52 | 55 | 52 | 2.1 | 4.0 | 0.0097 |
| MSDT(5, 5, 4) | 25 | 44 | 48 | 50 | 47 | 1.8 | 4.0 | 0.0101 |
| MSDT(4, 5, 5) | 20 | 38 | 41 | 44 | 41 | 1.7 | 4.2 | 0.0103 |

1) are at least as efficient as the traditional plan. But, the improved efficiency is at the cost of larger sample sizes and longer test lengths. We have similar comments about IMSDT plans for estimating the population median.

The CV column in Table 7 shows that, in comparison to the laminate panel example, there is less relative variability in test length. This is because the Weibull shape parameter is large and, thus, failures tend to occur closer together and it would be more sensible to wait for all test units to fail. All the observations failed in the actual test.

## 7.2 Discussion

From a practical perspective, there are important advantages of using MSDT plans instead of traditional experimental plans. MSDT plans provide a systematic procedure of replacing test units. Unlike traditional test plans, when censoring is used to limit testing time, MSDT plans give the experimenter control over the number of failures and, equivalently, over the accuracy of estimation (measured, for example, by asymptotic variance or confidence interval width). The control of information in MSDT plans is more robust to model parameter misspecification than in traditional plans.

MSDT plans for estimating a small quantile, say the 0.05 quantile, provide smaller sample sizes or shorter testing times than traditional plans. In the laminate panel data, MSDT plans with smaller values of $r$ resulted in shorter test lengths but increased sample sizes. IMSDT tests, however, resulted in not only smaller sample sizes but also better efficiency than the traditional plan. But, in the annealed aluminum example, there were tradeoffs between sample size, efficiency and test length.

The benefits of using MSDT and IMSDT plans should be assessed in the light of the possibility of censoring in the tests. Censoring is especially common when the failure time distribution has a large

Table 7: *Quantiles, Mean and Standard Deviation of Test Length and Variance of* 1000 *ML Estimates of* $\log(\hat{y}_{0.05})$ *under Traditional, MSDT and IMSDT* ($q_c = 1$) *Plans with* $k = 5$

| Plan | | $L_{05}$ | $L_{50}$ | $L_{95}$ | $\mu_L$ | $\sigma_L$ | $CV(\%)$ | Var($\log(\hat{y}_{0.05})$) MSDT | Var($\log(\hat{y}_{0.05})$) IMSDT |
|---|---|---|---|---|---|---|---|---|---|
| Traditional | | 34 | 38 | 41 | 37 | 1.9 | 5.1 | 0.0103 | 0.0103 |
| $g = 4$ | $r = 1$ | 23 | 27 | 30 | 26 | 2.5 | 9.4 | 0.0171 | – |
| | $r = 2$ | 28 | 31 | 34 | 31 | 1.9 | 6.2 | 0.0141 | 0.0139 |
| | $r = 3$ | 32 | 35 | 37 | 35 | 1.7 | 5.0 | 0.0125 | 0.0119 |
| | $r = 4$ | 35 | 38 | 40 | 38 | 1.6 | 4.4 | 0.0118 | 0.0115 |
| | $r = 5$ | 38 | 41 | 44 | 41 | 1.7 | 4.2 | 0.0103 | 0.0090 |
| $g = 5$ | $r = 1$ | 35 | 40 | 45 | 40 | 3.0 | 7.7 | 0.0148 | – |
| | $r = 2$ | 43 | 47 | 51 | 47 | 2.4 | 5.1 | 0.0119 | 0.0114 |
| | $r = 3$ | 48 | 52 | 55 | 52 | 2.1 | 4.1 | 0.0108 | 0.0105 |
| | $r = 4$ | 53 | 56 | 60 | 56 | 2.0 | 3.6 | 0.0099 | 0.0090 |
| | $r = 5$ | 58 | 62 | 65 | 62 | 2.1 | 3.5 | 0.0087 | 0.0070 |

coefficient of variation CV (small Weibull shape parameter $\beta$). This was the case in the laminate panel example. On the other hand, when the CV is small ($\beta$ is large), it is generally unnecessary to censor a life test, given that the test will be run until at least some failures are observed. Failures in this situation tend to occur closer together than when $\beta$ is smaller. In the aluminum wire example, the shape parameter is large and all observations are failures. Here, to achieve the same efficiency as traditional plans, MSDT plans need sample sizes at least as big as the traditional plan's. Some information on the coefficient of variation is thus useful in selecting an appropriate MSDT or IMSDT plan.

Table 8: *Average Sample Sizes* $\bar{n}$ *and Proportions Failing* $\bar{p}_f$ *under MSDT and IMSDT* ($q_c = 1$) *Plans with* $k = 5$

| Plan | | $r = 2$ $\bar{n}$ | $r = 2$ $\bar{p}_f$ | $r = 3$ $\bar{n}$ | $r = 3$ $\bar{p}_f$ | $r = 4$ $\bar{n}$ | $r = 4$ $\bar{p}_f$ | $r = 5$ $\bar{n}$ | $r = 5$ $\bar{p}_f$ |
|---|---|---|---|---|---|---|---|---|---|
| $g = 4$ | MSDT | 20.00 | 0.40 | 20.00 | 0.60 | 20.00 | 0.80 | 20.00 | 1.00 |
| | IMSDT | 21.09 | 0.38 | 22.42 | 0.55 | 24.02 | 0.71 | 26.20 | 0.85 |
| $g = 5$ | MSDT | 25.00 | 0.40 | 25.00 | 0.60 | 25.00 | 0.80 | 25.00 | 1.00 |
| | IMSDT | 26.18 | 0.39 | 27.75 | 0.57 | 29.56 | 0.73 | 32.21 | 0.88 |

## 8 Conclusions and Suggestions for Further Research

When there is a limit on the number of testing positions, MSDT and IMSDT plans provide useful alternatives to traditional test plans. Test lengths under these plans are shorter than under traditional plans on the average. When the cost and availability of specimens are not restrictive (but test positions are still limited), we can avoid MSDT machine idle times by using IMSDT plans that test standby specimens in idle test positions. The simulation studies and examples above show that IMSDT plans significantly improve efficiency. The MSDT and IMSDT plans perform best when censoring is expected (large CV) or needed in traditional plans. Relative to traditional plans, IMSDT plans have shorter test lengths, smaller sample sizes and better efficiency particularly in estimating low quantiles.

The discussions above are confined only to the Weibull distribution. Other life distributions such as the lognormal and loglogistic distributions could be investigated. We would expect to see similar results. Unlike the Weibull distribution, there are no closed forms for the moments of lognormal or loglogistic order statistics. The moments, however, can be computed numerically and simulations can be conducted without difficulty.

The MSDT plans considered here involve testing in sequence $g$ groups of $k$ specimens until the $r$th failure. This procedure could be improved by taking advantage of the sequential nature of the testing. The test plan for each batch is determined by information from previous test batches. This makes the choice of test plans dynamic and reduces the experimenter's dependence on starting values of parameters. Asymptotic theory for sequential plans is, however, much more complicated. For example, Ford, Titterington and Kitsos [14] remark that the distribution of the ML estimators is complex and its variance-covariance matrix is no longer proportional to the inverse of the Fisher information matrix. The sample information matrix can be used as a measure of the precision of the estimation of model parameters rather than as an estimated covariance matrix. Simulation studies on the large-sample distribution of estimators offer an alternative. From a practical point of view, such sequential tests would also be more difficult to administer.

## References

[1] Gertsbakh, I. (1996). "Optimal Allocation of Test Positions and Testing Time in a Factorial Life Testing Experiment." *Statistics and Probability Letters*, 26, 249-258.

[2]Johnson, L. (1964), *The Statistical Treatment of Fatigue Experiments*, New York: Elsevier Publishing Company.

[3]Kececioglu, D. (1993), *Reliability and Life Testing Handbook*, New Jersey: Prentice Hall.

[4]Suzuki, K., Ohtsuka, K., Ashitate, M. (1992), "On a Comparison Between Sudden Death Life Testing and Type II Number Fixed Life Testing - The Precisions and the Testing Times Using the Maximum Likelihood Estimators." *Journal of the Japanese Society for Quality Control* 22, 5-12.

[5]Halperin, M. (1952), "Maximum Likelihood Estimation in Truncated Samples," *Annals of Mathematical Statistics* 23, 226-238.

[6]Battacharyya, G. K. (1985), "The Asymptotics of Maximum Likelihood and Related Estimators Based on Type II Censored Data," *Journal of the American Statistical Association* 80, 398-404.

[7]Escobar, L. A. and Meeker, W. Q. (1986), "Planning Accelerated Life Tests with Type II Censored Data," *Journal of Statistical Computation and Simulation* 23, 273-297.

[8]Escobar, L. A. and Meeker, W. Q. (1994), "Fisher Information Matrix for the Extreme Value, Normal and Logistic Distributions and Censored Data." *Applied Statistics* 3, 533-554.

[9]Cornish, E. A. and Fisher R. A. (1937), "Moments and Cumulants in the Specification of Distributions." *Review of the International Statistical Institute* 5, 307-320.

[10]Fisher, R. A. and Cornish, E. A. (1960), "The Percentile Points of Distributions Having Known Cumulants." *Technometrics* 2, 209-225.

[11]Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994), *Continuous Univariate Distributions*, New York: John Wiley & Sons.

[12]Shimokawa, T. and Hamaguchi, Y. (1987), "Statistical Evaluation of Fatigue Life and Fatigue Strength in Circular-Holed Notched Specimens of a Carbon Eight-Harness-Satin/Epoxy Laminate." *Statistical Research on Fatigue and Fracture (Current Japanese Materials Research - Vol. 2)*, Tanaka T., Nishijima, S. and Ichikawa, M. editors. The Society of Materials Science, Japan, London: Elsevier Applied Science, 159-176.

[13]Shen, C. L. (1994), "Statistical Analysis of Fatigue Data," Ph.D. dissertation, University of Arizona, Tucson.

[14]Ford, I., Titterington, D. M. and Kitsos, C. P. (1989). "Recent Advances in Nonlinear Experimental Design." *Technometrics* 31, 49-60.