Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cjas20

Analyzing type II censored data obtained from repetitious experiments

Lee-Ing Tong [a] & Chien-Hui Yang [b]

[a] Department of Industrial Engineering and Management , National Chiao Tung University , Hsinchu, Taiwan

[b] Department of Business Administration , Yuanpei Institute of Science and Technology , Hsinchu, Taiwan

Published online: 22 Jan 2007.

To cite this article: Lee-Ing Tong & Chien-Hui Yang (2006) Analyzing type II censored data obtained from repetitious experiments, Journal of Applied Statistics, 33:1, 49-63, DOI: 10.1080/02664760500389673

To link to this article: http://dx.doi.org/10.1080/02664760500389673

Journal of Applied Statistics
Vol. 33, No. 1, 49–63, January 2006

Analyzing Type II Censored Data Obtained from Repetitious Experiments

LEE-ING TONG* & CHIEN-HUI YANG**

**Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu, Taiwan and **Department of Business Administration, Yuanpei Institute of Science and Technology, Hsinchu, Taiwan*

ABSTRACT *Experimental design and Taguchi's parameter design are widely employed by industry to optimize the process/product. However, censored data are often observed in product lifetime testing during the experiments. After implementing a repetitious experiment with type II censored data, the censored data are usually estimated by establishing a complex statistical model. However, using the incomplete data to fit a model may not accurately estimates the censored data. Moreover, the model fitting process is complicated for a practitioner who has only limited statistical training. This study proposes a less complex approach to analyze censored data, using the least square estimation method and Torres's analysis of unreplicated factorials with possible abnormalities. This study also presents an effective method to analyze the censored data from Taguchi's parameter design using least square estimation method. Finally, examples are given to illustrate the effectiveness of the proposed methods.*

KEY WORDS: Type II censored data, least square estimation, Torres's method, experimental design, Taguchi's parameter design

Introduction

In industrial experiments, data are frequently censored because of the time or cost consumption associated with collecting data, as well as other restrictions, such as material resources or measuring instrument limitations. If the exact value of an observation is unknown but is not less than a predetermined value, L, then this observation is believed to be right censored at L. Likewise, if the exact value of an observation is unknown but is less than or equal to L, then this observation is believed to be left censored at L (Lawless, 1982). If the left and right censoring situations arise together, then the observation is known as doubly censored. Censoring arises for different reasons and various censoring processes are developed accordingly. Three censoring models are frequently employed in reliability analysis to save experimental time: type I, type II, and interval censoring. In type I censoring, a lifetime testing experiment is operated for a fixed period and the lifetime of an individual is known only if it is less than a predetermined value. Namely, in a lifetime testing experiment, n items are placed on test and the

---

Correspondence Address: Lee-Ing Tong, Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu, Taiwan 30050; Email: litong@cc.nctu.edu.tw

0266-4763 Print/1360-0532 Online/06/010049–15
DOI: 10.1080/02664760500389673

experiment is terminated after time L has elapsed. Under this condition, lifetimes from the testing experiment are known exactly only for those items that fail by time L. Notably, with type I censoring, the number of exact lifetimes observed is random (Lawless, 1982). For instance, there are initial sampling size 100 batteries to be tested. The experimental terminating time is on seventh month. The individual lifetimes of batteries are recorded at the end of the seventh month. If the battery lifetime is longer than seven months, the datum is regarded as a censored datum; otherwise, the datum is regarded as an uncensored datum (Chiou and Tong, 2001). In type II censoring, a total of n items are tested, but, instead of continuing experiments until all n items have failed, the experiment is terminated when the rth item fails. Namely, a type II censored sample arises when only r observations in a random sample of n items are made (Lawless, 1982). For example, during 100 batteries test, the experiment is terminated in advance until the 80th lifetime datum is obtained. The remaining 20 items are regarded as censored. Interval censoring occurs when the lifetime from a testing experiment is not known exactly but is known to lie between two inspection times. Namely, interval-censored data are obtained when lifetimes are observed at inspection times. Interval-censored data does not provide the actual lifetimes (Chowdhury and Fard, 2001). For instance, during the experiment on 100 batteries, the lifetimes of failed batteries are recorded at the end of every month and the experimental terminating time is at the seventh month. The lifetime of failed batteries are obtained only when the batteries are failed between the inspection times. If the battery lifetime exceeds seven months, the datum is regarded as a censored data, and can be expressed as $(7, \infty)$ indicating that the battery lifetime exceeds seven months. In the case of censored experimental data, traditional analysis of variance (ANOVA) method is not appropriate to analyze the factor effects. Various methods estimate the censored data in experiments and are presented in the following section. However, some of these methods, such as maximum likelihood estimation (MLE) and Taguchi's minute accumulating analysis (MAA) methods are computationally complicated or have certain limitations, which also depicted in the following section.

This study proposes a novel method to analyze type II censored data from repetitious experiments using the least square estimation and performing Torres's analysis of unreplicated factorials with possible abnormalities (Torres, 1993). In addition, this study also proposes an effective method to analyze the censored data from Taguchi's parameter design using least square estimation method.

Related Work

Nelson and Hahn (1972, 1973) wrote a two-part article on linear estimation methods for the parameters in the regression models using the ordered observations of censored data. Part I (Nelson and Hahn, 1972) presented two methods for making simple (but not minimum variance) linear estimates of the parameters of a linear regression model from censored data on the dependent variable in the special case that involves a single independent variable when the observations have unequal variances and are correlated. Herein, a method for obtaining the best linear unbiased estimates may not be used because this method leads to the problem in weighted least squares estimation, which involves complicated computations. Part II (Nelson and Hahn, 1973) presents a method for making the best linear unbiased estimates of the parameters of a linear regression model using censored data on the dependent variable in the special case in which only one uncensored observation is made and the sample size under each test condition is the same. However, in both part I and part II, methods of linear estimation were applicable only to type II censoring. Furthermore, the linear estimators must employ the available tables of the variance of the rth order statistic in a sample of size n that are tabulated by Sarhan and Greenberg (1962), limiting the sample size under any test condition to the sample sizes in the tables. Hahn and Nelson (1974) reviewed graphical, maximum likelihood and linear estimation methods for analyzing censored lifetime data in a study of the relationships between stress and product lifetime. They compared the advantages and disadvantages of these methods and chose a preferred one. Graphical methods provide subjective procedures, causing different results according to different individuals. The maximum likelihood method can be applied to any model – that is, to almost any life distribution and relationships of the distribution parameters. Moreover, it is applicable to many types of censored data, including type I and type II and single and multiple censoring. However, the MLE may sometimes not exist. Linear estimation methods are computationally simpler than maximum likelihood methods. Nevertheless, linear estimation methods are applicable only to type II censoring. Krall *et al.* (1975) suggested the utilization of a forward selection procedure for selecting the most important variables associated with survival time data, using MLE. The method proposed herein involves fewer computations than the original MLE method. However, the proposed procedure still involves many computations. Schmee and Hahn (1979) suggested an iterative model selection method based on an iterative least squares procedure for estimating censored data. An initial least squares fit is obtained, in which the censored values are treated as failures. Next, the initial fit is used to estimate the expected failure time for each censored observation. These estimates are then used to obtain a revised least squares fit and new expected failure times are estimated for the censored values. This procedure is iterated until convergence is achieved. However, treating the censored data as if they were uncensored may bias the estimates of the regression line. Hahn *et al.* (1981) developed the iterative least squares method to analyze the results of a fractional factorial experiment involving censoring to the left. However, the initial model selection significantly influences the final model. Taguchi (1987) presented a minute accumulating analysis (MAA) method to deal with interval-censored data. The data are represented by 0s and 1s. A 1 is assigned to all groups that precede the group in which the unit is failed and the remaining groups are assigned 0s. Hence, a time factor whose levels correspond to these groups is created. Then, MAA performs an analysis of variance on this generated binary data, treating them as if they were obtained from a split-plot experiment. However, in MAA, censoring times are treated as actual failure times, possibly causing a problem if the unobserved failure times and censoring times differ greatly.

Hamada and Wu (1991) used the MLE method to specify a structural model for identifying the important factors that affect a quality characteristic of a product or process and choosing levels of these factors that lead to improvement when data are censored. First, data are transformed to achieve near normality. Standard methods can then be used to choose a tentative model based on the data that combine complete and imputed censored observations. Next, the current model is fitted, and the censored data are imputed again. This cycle that includes fitting, imputation, and model selection continues until the selected model stops changing. Several models may be identified and diagnostic checking can be performed to assess the effectiveness of these models. A limitation of these methods is that they rely on the existence of the MLEs because the model cannot be fitted if the MLEs of the initial model do not exist. Torres (1993) proposed an analysis based on the rank transformation of the response to deal with the analysis of unreplicated factorial or fractional factorial experiments with possible abnormalities. The ranks of the observations are analyzed as if the ranks were the original observations. Then, the normal probability plot of the effects of the ranked observations is used to determine the important factors. Lu and Unal (1994) suggested an expectation-modeling-maximization (EMM) algorithm that uses pseudo-complete data, and used a forward regression to compare all the main effects and two-factor interactions for process characterization. The best combination of controllable variables is then determined to optimize the predictions of the final model. A study of the sensitivity of the selected models indicates the importance of using appropriate models and estimation methods in EMM. Hamada and Wu (1995) adapted a Bayesian approach to estimate the censored data obtained by conducting a factorial experiment to solve the problem of an infinite or nonexistent MLE. However, that method may not be simple enough to be useable by practitioners. Tong and Su (1997) proposed an effective procedure based on the transformation of the rank of the responses and regression analysis for analyzing an experiment with singly censored data. The non-parametric method and regression analysis are used to analyze multi-factor and multi-level experimental results obtained from censored data. Their procedure not only concurrently analyzes censored data obtained in replicated and unreplicated experiments but also considers the variability of control factors. Chowdhury and Fard (2001) provided an experimental plan for analyzing interval-censored response data from an unreplicated factorial experiment to improve the reliability of the product. In analyzing an unreplicated factorial experiment that involves interval-censored data, the mean and the rank of the response data are transformed and then a normal probability plot analysis is used to identify the significant factors.

Least Square Estimation Method for Censored Experimental Data

For the censored experimental data, the least square method can be used to estimate the population mean and variance under the normality assumption. Suppose $X_{(r+1)} < X_{(r+2)} < \cdots < X_{(n-s)}$ is a doubly type II censored sample (when $r = 0$, it is known as a right type II censored, and when $s = 0$, it is known as a left type II censored), which is normally distributed with mean μ and variance σ^2. Then the cumulative distribution function of X_i, $i = r+1, \ldots, n-s$, is expressed as follows.

$$F(X_{(i)}) = \Phi\left(\frac{X_{(i)} - \mu}{\sigma}\right), \quad r+1 \le i \le n-s \tag{1}$$

where $\phi(\cdot)$ represents the standard normal cumulative distribution function.

From equation (1), $\hat{F}(X_{(i)})$ can be obtained as,

$$\hat{F}(X_{(i)}) = \Phi\left(\frac{X_{(i)} - \hat{\mu}}{\hat{\sigma}}\right), \quad r+1 \le i \le n-s \tag{2}$$

where $\hat{F}(X_{(i)})$ equals $i/(n+1)$ (D'Agostino and Stephens, 1986).

Equation (2) can be rewritten as,

$$\Phi^{-1}(\hat{F}(X_{(i)})) = \frac{X_{(i)} - \hat{\mu}}{\hat{\sigma}} = \frac{-\hat{\mu}}{\hat{\sigma}} + \frac{1}{\hat{\sigma}} X_{(i)}, \quad r+1 \le i \le n-s \tag{3}$$

Equation (3) can be regarded as a simple regression model, so the least square estimators of μ and σ can be obtained by minimizing the sum of the squared deviations:

$$\min \sum_{i=r+1}^{n-s} \varepsilon_i^2 = \min \sum_{i=r+1}^{n-s} \left(Y_{(i)} - \beta_0 - \beta_1 X_{(i)}\right)^2 \tag{4}$$

where $Y_{(i)} = \Phi^{-1}(\hat{F}(X_{(i)}))$, $\beta_0 = -\hat{\mu}/\hat{\sigma}$, $\beta_1 = 1/\hat{\sigma}$.

Partially differentiating equation (4) with respect to β_0 and β_1 yields $\hat{\beta}_0$ and $\hat{\beta}_1$ as,

$$\hat{\beta}_1 = \frac{\sum_{i=r+1}^{n-s} (X_{(i)} - \bar{X}) \cdot Y_{(i)}}{\sum_{i=r+1}^{n-s} (X_{(i)} - \bar{X})^2} = \frac{1}{\hat{\sigma}} \tag{5}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}.$$

From equation (5), $\hat{\mu}$ and $\hat{\sigma}$ can be obtained as,

$$\hat{\sigma} = \frac{\sum_{i=r+1}^{n-s} (X_{(i)} - \bar{X})^2}{\sum_{i=r+1}^{n-s} (X_{(i)} - \bar{X}) \cdot Y_{(i)}}$$

$$\hat{\mu} = \bar{X} - \hat{\sigma}\bar{Y}. \tag{6}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the least square estimators of μ and σ, respectively.

Proposed Procedure for Analyzing Censored Data from Factorial Design

The main objective of experimental design is to determine the important factors that affect the quality characteristics of a process/product and to select the optimal setting of factor-levels. Accordingly, this study presents a procedure for optimizing type II censoring experiments. The first step is to estimate the mean of every treatment using the least square method, and then Torres's analysis of unreplicated factorials (Torres, 1993) is conducted to determine the optimal factor-level setting.

Proposed Procedure for Analyzing Censored Data from Taguchi's Parameter Design

In most experiments, data are used to analyze a mean response. Taguchi, however, emphasized the importance of studying the variation of the response using the signal-to-noise (S/N) ratio. Many S/N formulas are available, but three of them are considered to be standard and are generally applicable when the response can be classified as "smaller-the-better", "larger-the-better", or "nominal-the-best" (Byrne and Taguchi, 1987).

When the response is nominal-the-best, the two-phase optimization procedure can be used to improve the process so that the mean of the response will be as close as possible to the target value. The steps of the two-phase optimization procedure are as follows (Peace, 1993).

1. The first step in enhancing the process/product performance is to reduce variability. Find the factors that have a significant effect on variability and then determine the settings of factor-levels that will be least affected by noise. The variation of the products can be reduced using the determined settings of factor-levels.

2. After the variation of the data is reduced, the second step is to shift the mean to the target value. Focusing on the factors that significantly impact the average result but have little or no effect on the variation enables the distribution of the data to be moved without disturbing the spread of the results. Restated, determining the setting of factor–levels that affect the mean but not variability can shift the mean of the process toward the desired target value.

The S/N formulas must be transformed into other forms containing the estimators of the mean and the variance of each treatment combination in cases in which responses are "larger-the-better", or "smaller-the-better". Doing so enables us to use the least square estimation method to analyze the Taguchi's parameter design with censored data. When the response is smaller-the-better, the formula for the S/N ratio can be expressed as,

$$\begin{aligned} SN_{STB} &= -10\log\left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right) \\ &= -10\log\left\{\frac{1}{n}\sum_{i=1}^{n}[(y_i-\bar{y})+\bar{y}]^2\right\} \\ &= -10\log\left(s^2+\bar{y}^2\right) \end{aligned} \tag{7}$$

where $s^2 = \sum_{i=1}^{n}(y_i-\bar{y})^2/n$.

If the response is larger-the-better, then the response is inverted to transform it into a smaller-the-better response and then equation (7) is utilized to calculate the S/N ratio. In the case in which the response is nominal-the-best, the S/N formula can be written as,

$$SN_{NTB} = 10\log\left(\frac{\bar{y}^2}{s^2}\right). \tag{8}$$

Then, the procedure for analyzing the Taguchi's parameter design with censored data is presented as follows:

1. If the response is larger-the-better, then it is inverted to transform it into a smaller-the-better response; if the response is smaller-the-better or nominal-the-best, then nothing is done in this step.
2. The mean and the standard deviation of every treatment are estimated using the least square method described in the preceding section.
3. The mean and the standard deviation estimated in step 2 are used to compute the S/N ratios for all treatments. When the response is smaller-the-better, equation (7) is used; when the response is nominal-the-best, equation (8) is used.
4. The level of a controllable factor is selected based on the value of the S/N ratio. For a smaller-the-better response, the level that corresponds to the highest average S/N ratio is selected as the optimal setting of factor-level. For a response that is nominal-the-best, a two-phase optimization procedure can be used.

A Numerical Example for Experimental Design

Problem Description

A real-world example is adopted from Montgomery (2001). A 16-run experiment was performed in a semiconductor manufacturing plant to study the effects of six factors on the curvature or camber of the substrate devices produced. Table 1 lists the six variables and their levels. Each run was replicated four times, and a camber measurement was made on the substrate. Table 2 shows all the relevant data. Suppose the largest four data cannot be observed, and right type II censored data are generated with $n = 4$, $r = 0$, and $s = 1$. Also, the response variable is assumed to be normally distributed.

Data Analysis

In the analysis of the right type II censored data in this example, first, the least square estimation is utilized to estimate the means of each treatment combination. Table 3 presents the results. The last column of Table 3 lists the ranks of the LSE. This design of running only a fraction of the complete factorial experiment is a fractional factorial design, so one effect is aliased with another. The important factors are A, B, C, D, E, F, AB, AC, AD, AE, AF, BD, BF, ABD, and ACD. Accordingly, Torres's analysis is employed as follows.

Step 1: Let **L** be a vector whose elements are the least square estimates of 16 treatments. From Table 3, the rank transformation of the vector **L** can be obtained as,

$$\mathbf{R} = [13, 4, 2, 6, 3, 16, 8, 15, 5, 7, 1, 14, 10, 9, 12, 11]'.$$

Step 2: Let **X** be a 16×16 design matrix with all entries are equal to plus or minus one. The first column is a vector of ones and the rest of the columns refer to the fifteen main effects and interaction effects.

Step 3: Let **R** be a response vector, whose regression coefficients (estimated constant term and effects) can be computed as follows.

$$\boldsymbol{\alpha}_{\mathbf{R}} = [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{R}$$

Column 1 of Table 4 shows $\boldsymbol{\alpha}_{\mathbf{R}}$.

Step 4: The estimated constant term and the effects are used to construct a normal probability plot, as shown in Figure 1.

Table 1. Factors and their levels in the experiment

Label	Factor	Low level	High level
A	Lamination temperature(°C)	55	75
B	Lamination time (s)	10	25
C	Lamination pressure (Ton)	5	10
D	Firing temperature (°C)	1580	1620
E	Firing cycle time (h)	17.5	29
F	Firing dew point (°C)	20	26

Table 2. Design and data for the experiment

Run	A	B	C	D	E	F	Camber for replicate			
							X_1	X_2	X_3	X_4
1	−1[a]	−1	−1	−1	−1	−1	167	128	149	185
2	1	−1	−1	−1	1	−1	62	66	44	20
3	−1	1	−1	−1	1	1	41	42	43	50
4	1	1	−1	−1	−1	1	73	81	39	30
5	−1	−1	1	−1	1	1	47	47	40	89
6	1	−1	1	−1	−1	1	219	258	147	296
7	−1	1	1	−1	−1	−1	121	90	92	86
8	1	1	1	−1	1	−1	191	186	162	106
9	−1	−1	−1	1	−1	1	32	23	77	69
10	1	−1	−1	1	1	1	78	158	60	45
11	−1	1	−1	1	1	−1	43	27	28	28
12	1	1	−1	1	−1	−1	186	137	159	158
13	−1	−1	1	1	1	−1	110	86	101	158
14	1	−1	1	1	−1	−1	65	109	126	71
15	−1	1	1	1	−1	1	155	158	145	145
16	1	1	1	1	1	1	93	124	110	133

[a] 1 and −1 are used to represent the high and low levels of the factors.

Figure 1 shows that the effects E, and ACD are significant. Since the response (rank) is smaller-the-better, factor E is set at its high level since its regression coefficient is negative. Additionally, since the interaction effect ACD is significant, all of the regression coefficients for A, C, D, AC, AD, and ACD must be considered because the interaction

Table 3. Censored data, LSE, and the rank

No.	$X_{(1)}$	$X_{(2)}$	$X_{(3)}$	LSE	Rank
1	128	149	167	157.993	13
2	20	44	62	52.796	4
3	41	42	43	42.512	2
4	30	39	73	59.771	6
5	40	47	47	47.000	3
6	147	219	258	237.062	16
7	86	90	92	90.914	8
8	106	162	186	172.711	15
9	23	32	69	54.778	5
10	45	60	78	69.497	7
11	27	28	28	28.000	1
12	137	158	159	158.364	14
13	86	101	110	105.238	10
14	65	71	109	95.171	9
15	145	145	155	151.836	12
16	93	110	124	116.948	11

Table 4. Estimated constant term, effects and data used in constructing a normal probability plot

Factor	α_R	Normal score	Rank	Cumulative probability
A	1.750	1.2450	14	0.8934
B	0.125	−0.2491	6.5	0.4016
C	2.000	1.7394	15	0.9590
D	0.125	−0.2491	6.5	0.4016
E	−1.875	−1.2450	2	0.1066
F	−0.750	−0.9458	3	0.1721
AB	1.125	0.9458	13	0.8279
AC	0.500	0.2491	9.5	0.5984
AD	−0.125	−0.6113	4.5	0.2705
AE	0.875	0.7137	12	0.7623
AF	0.500	0.2491	9.5	0.5984
BD	0.750	0.5150	11	0.6967
BF	−0.125	−0.6113	4.5	0.2705
ABD	0.250	0.0000	8	0.5000
ACD	−2.625	−1.7394	1	0.0410

effect ACD is significant. Table 5 lists the ranks of the response estimated from various combinations of factors A, C, D, ACD, AC, and AD. From Table 5, condition 7 (in which A and C are at their low levels and D is at its high level) will produce the smallest rank, so the optimal setting of factor-levels is A and C at their low levels and D and E at their high levels. The levels of other insignificant factors can also be determined by their regression coefficients. In this case, factor B is set at its low level and factor F is set at its high level. Hence, the optimal setting of factor-levels can be determined as $A_{-}B_{-}C_{-}D_{+}E_{+}F_{+}$ (minus and plus signs are used to represent the low and high levels of a factor, respectively).

Comparison of the Result of Analyzing Censored Data with Complete Data

When the data are complete, factors A, B, C, D, and E are significant, and the optimal setting of factor-levels of A, B, and C are all set at their low levels, while factors D and

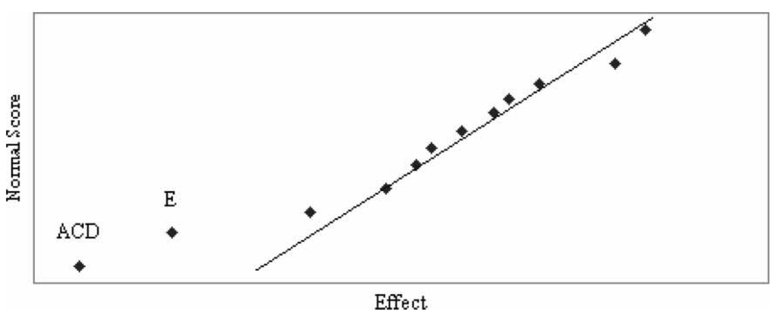

Figure 1. Normal probability plot of the estimated effects

Table 5. Rank estimated from various combinations of effects A, C, D, AC, AD, and ACD

Effect	A	C	D	ACD	AC	AD	Response (rank)
Regression coefficient	1.75	2.00	0.125	−2.625	0.5	−0.125	—
Condition 1	1	1	1	1	1	1	1.625
Condition 2	1	1	−1	−1	1	−1	6.875
Condition 3	1	−1	1	−1	−1	1	1.875
Condition 4	1	−1	−1	1	−1	−1	−3.375
Condition 5	−1	1	1	−1	−1	−1	2.625
Condition 6	−1	1	−1	1	−1	1	−3.125
Condition 7	−1	−1	1	1	1	−1	−5.625
Condition 8	−1	−1	−1	−1	1	1	−0.875

E are set at their high levels. Comparing the result where the data are complete with the result where the data are censored shown in the preceding section reveals that factor B is not included as a significant factor by the proposed method. However, according to Table 4, factor B will be set at its low level, which conclusion was also reached using complete data.

A Numerical Example for Taguchi's Parameter Design

Problem Description

The following experiment is adopted from Byrne and Taguchi (1987). An experiment was conducted to find a method to assemble economically an elastomeric connector to a nylon tube that would provide the requisite pull-off performance for use in automotive engine components. The objective of the experiment is to maximize the pull-off force.

Four controllable factors and three noise factors that could affect the assembly's pull-off force have been identified. Table 6 lists the four controllable factors and their levels. The three noise factors are conditioning time, conditioning temperature, and conditioning relative humidity. The three noise factors are not to be included in the experiments performed herein. Each run is replicated eight times, and the pull-off forces in pounds are recorded. Table 7 presents all the complete data (Byrne and Taguchi, 1987). Consider a case in

Table 6. Factors and levels of each factor in the experiment

Label	Factor	Levels		
A	Interference	Low	Medium	High
B	Connector wall thickness	Thin	Medium	Thick
C	Insertion depth	Shallow	Medium	Deep
D	Percent adhesive in connector pre-dip	Low	Medium	High

Table 7. Design and data of the experiment

Run	A	B	C	D	Pull-off force for replicate								SN
					X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	ratio
1	1[b]	1	1	1	15.6	9.5	16.9	19.9	19.6	19.6	20.0	19.1	24.045
2	1	2	2	2	15.0	16.2	19.4	19.6	19.7	19.8	24.2	21.9	25.522
3	1	3	3	3	16.3	16.7	19.1	15.6	22.6	18.2	23.3	20.4	25.335
4	2	1	2	3	18.3	17.4	18.9	18.6	21.0	18.9	23.2	24.7	25.904
5	2	2	3	1	19.7	18.6	19.4	25.1	25.6	21.4	27.5	25.3	26.908
6	2	3	1	2	16.2	16.3	20.0	19.8	14.7	19.6	22.5	24.7	25.326
7	3	1	3	2	16.4	19.1	18.4	23.6	16.8	18.6	24.3	21.6	25.711
8	3	2	1	3	14.2	15.6	15.1	16.8	17.8	19.6	23.2	24.4	24.832
9	3	3	2	1	16.1	19.9	19.3	17.3	23.1	22.7	22.6	28.6	26.152

[b]1 represents the low level of a factor, 2 represents the medium level of a factor, and 3 represents the high level of a factor.

which the largest eight data cannot be observed, and right type II censored data are formed with $n = 8$, $r = 0$, and $s = 1$.

Data Analysis

In this problem, the response is larger-the-better; therefore, according to the procedure of analyzing censored data from Taguchi's parameter design described in the preceding section, the response must be inverted so as to be smaller-the-better, and the type II right censored data in this problem changes to type II left censored data with $n = 8$, $r = 1$, and $s = 0$. Table 8 presents the transformed type II left censored data. In analyzing the type II left censored data, first, the least square estimation is utilized to estimate the means and the variances of each treatment combination. Table 9 presents the results.

This analysis can be performed in several ways. One common way is to use the ANOVA method to analyze factor effects. A conceptual method, which factor effects to be graphed and the factors that appear to be significant to be visually identified can also be used.

Figure 2 plots the average S/N ratios for each level of the four controllable factors. Figure 2 indicates that the factors A and C are more significant than factors B and D. The average S/N ratios are larger-the-better, so factor A is set at its inter-medium level, and factor C is set at its high level. The optimal setting of factor-levels can be determined as $A_2\, B_2\, C_3\, D_1$. The ANOVA method is also utilized to test the significance of four factors. Table 10 presents the results. Table 10 clearly shows that factors A and C are significant, therefore, the conclusion of this analysis is that factors A and C are significant and the optimal setting of factor-levels can also be determined as $A_2\, B_2\, C_3\, D_1$, of which factors B and D are not significant.

Comparison of the Result of Analyzing Censored Data with Complete Data

When all of the complete data are analyzed, factors A and C are determined to be significant, and the optimal setting of factor-levels can be determined as $A_2\, B_2\, C_3\, D_1$. This result is the same as that associated with the censored data in the preceding section.

Table 8. Type II left censored data of the reciprocal of the responses

	Pull-off force for replicate							
Run	$X_{(1)}$	$X_{(2)}$	$X_{(3)}$	$X_{(4)}$	$X_{(5)}$	$X_{(6)}$	$X_{(7)}$	$X_{(8)}$
1	—	0.05025	0.05102	0.05102	0.05236	0.05917	0.06410	0.10526
2	—	0.04566	0.05051	0.05076	0.05102	0.05155	0.06173	0.06667
3	—	0.04425	0.04902	0.05236	0.05495	0.05988	0.06135	0.06410
4	—	0.04310	0.04762	0.05291	0.05291	0.05376	0.05464	0.05747
5	—	0.03906	0.03953	0.03984	0.04673	0.05076	0.05155	0.05376
6	—	0.04444	0.05000	0.05051	0.05102	0.06135	0.06173	0.06803
7	—	0.04237	0.04630	0.05236	0.05376	0.05435	0.05952	0.06098
8	—	0.04310	0.05102	0.05618	0.05952	0.06410	0.06623	0.07042
9	—	0.04329	0.04405	0.04425	0.05025	0.05181	0.05780	0.06211

Table 9. Censored data, LSE and the SN ratio

No.	$X_{(2)}$	$X_{(3)}$	$X_{(4)}$	$X_{(5)}$	$X_{(6)}$	$X_{(7)}$	$X_{(8)}$	LSE of mean	LSE of variance	SN ratio
1	0.05025	0.05102	0.05102	0.05236	0.05917	0.06410	0.10526	0.05577	0.00123	23.627
2	0.04566	0.05051	0.05076	0.05102	0.05155	0.06173	0.06667	0.05196	0.00013	25.475
3	0.04425	0.04902	0.05236	0.05495	0.05988	0.06135	0.06410	0.05330	0.00011	25.300
4	0.04310	0.04762	0.05291	0.05291	0.05376	0.05464	0.05747	0.05045	0.00006	25.845
5	0.03906	0.03953	0.03984	0.04673	0.05076	0.05155	0.05376	0.04420	0.00009	26.888
6	0.04444	0.05000	0.05051	0.05102	0.06135	0.06173	0.06803	0.05309	0.00016	25.260
7	0.04237	0.04630	0.05236	0.05376	0.05435	0.05952	0.06098	0.05107	0.00010	25.675
8	0.04310	0.05102	0.05618	0.05952	0.06410	0.06623	0.07042	0.05622	0.00019	24.742
9	0.04329	0.04405	0.04425	0.05025	0.05181	0.05780	0.06211	0.04860	0.00012	26.052

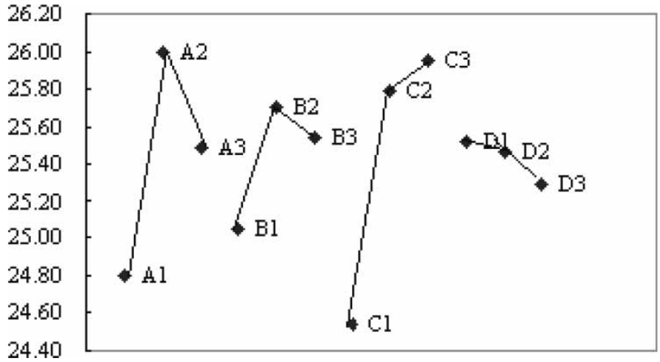

Figure 2. Average S/N ratios for each level of the four controllable factors

Conclusion

Experimental design and Taguchi methods are two useful tools for optimizing and improving products. However, data are frequently censored because of the time or cost associated with collecting data, as well as other restrictions such as material resources or measuring instrument limitations. In this situation, the traditional ANOVA method is inappropriate for analyzing the effects of factors, so a novel method is proposed in this study to overcome this problem. The proposed method can be used to perform right type II censoring, left type II censoring, or doubly type II censoring. The proposed procedure first uses the least square method to estimate the mean of each treatment, and then applies Torres's analysis to unreplicated factorials to determine the important settings of factor–levels. The proposed method does not analyze the data iteratively and does not involve complicated model, so it is simpler than conventional methods. Moreover, this study also presents an effective method to analyze the censored data from Taguchi's parameter design using least square estimation method. Notably, regardless of whether the proposed methods are used to analyze the factorial experiments or the Taguchi experiments, the censored data of each treatment combination must include at least two observations that can be observed. Finally, the proposed procedures, while lacking rigorous theoretical justification, can be easily implemented in an industrial setting.

Table 10. ANOVA table for S/N ratio

Source	Degree of freedom	SS	MS	F value	Pure sum of squares	Contribution (%)
A	2	2.1656	1.0828	5.5817	1.7776	27.28
B	2	0.6914*	—	—	—	—
C	2	3.5753	1.7877	9.2153	3.1873	48.91
D	2	0.0845*	—	—	—	—
Error	0	—	—	—	—	—
Pooling error	(4)	(0.7760)	(0.1940)		1.5519	23.81
Total	8	6.5169			6.5169	100.00

References

Byrne, D. M. & Taguchi, S. (1987) The Taguchi approach to parameter design, *Quality Progress*, 20, pp. 19–26.

Chiou, K. C. & Tong, L. I. (2001) Average type II censoring times for two parameter Pareto and Rayleigh distributions, *International Journal of Quality and Reliability Management*, 18(6), pp. 643–65.

Chowdhury, A. H. & Fard, N. S. (2001) Analysis of interval censored data from highly fractionated experiments, *International Journal of Industrial Engineering*, 8(3), pp. 241–250.

D'Agostino, R. B. & Stephens, M. A. (1986) *Goodness-of-fit Techniques* (New York: Marcel Dekker).

Hahn, G. J. & Nelson, W. (1974) A comparison of methods for analyzing censored life data to estimate relationships between stress and product life, *IEEE Transactions on Reliability*, 23, pp. 2–11.

Hahn, G. J. *et al.* (1981) The analysis of a fractional factorial experiment with censored data using iterative least squares, *Technometrics*, 23, pp. 33–36.

Hamada, M. & Wu, C. F. J. (1991) Analysis of censored data from highly fractionated experiments, *Technometrics*, 33, pp. 25–38.

Hamada, M. & Wu, C. F. J. (1995) Analysis of censored data from fractionated experiments: a Bayesian approach, *Journal of the American Statistical Association*, 90, pp. 467–477.

Krall, J. M. *et al.* (1975) A step-up procedure for selection variables associated with survival, *Biometrics*, 31, pp. 49–57.

Lawless, J. F. (1982) *Statistical Models and Methods for Lifetime Data* (New York: Wiley).

Lu, J. C. & Unal, C. (1994) Process characterization and optimization based on censored data from highly fractionated experiments, *IEEE Transactions on Reliability*, 43, pp. 145–155.

Montgomery, D. C. (2001) *Design and Analysis of Experiments Fourth Edition* (New York: Wiley).

Nelson, W. & Hahn, G. J. (1972) Linear estimation of a regression relationship from censored data, part I – simple methods and their application, *Technometrics*, 14, pp. 247–269.

Nelson, W. & Hahn, G. J. (1973) Linear estimation of a regression relationship from censored data, part II – best linear unbiased estimation and theory, *Technometrics*, 15, pp. 133–150.

Peace, G. S. (1993) *Taguchi Methods: A Hands-On Approach* (New York: Addison-Wesley Publishing Company).

Sarhan, A. E. & Greenberg, B. G. (1962) *Contributions to Order Statistics* (New York: Wiley).

Schmee, J. & Hahn, G. J. (1979) A simple method for regression analysis with censored data, *Technometrics*, 21, pp. 417–434.

Taguchi, G. (1987) *System of Experiment Design* (New York: White Plains, Unipub).

Tong, L. I. & Su, C. T. (1997) A non-parametric method for experimental analysis with censored data, *International Journal of Quality and Reliability Management*, 14(5), pp. 456–463.

Torres, V. A. (1993) Simple analysis of unreplicated factorials with possible abnormalities, *Journal of Quality Technology*, 25(3), pp. 183–187.