ISSN 1645-6726

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

REVSTAT

Statistical Journal

Volume 7, Nº. 2
June 2009

Catalogação Recomendada

REVSTAT. Lisboa, 2003-
Revstat : statistical journal / ed. Instituto Nacional de Estatística. - Vol. 1, 2003- . - Lisboa I.N.E., 2003- . - 30 cm
Semestral. - Continuação de : Revista de Estatística =
ISSN 0873-4275. - edição exclusivamente em inglês
ISSN 1645-6726

CREDITS

- **PUBLISHER**
  - *Instituto Nacional de Estatística, I.P. (INE, I.P.)*
  Av. António José de Almeida, 2
  1000-043 LISBOA
  PORTUGAL
  Tel.: +351 218 426 100
  Fax: +351 218 426 364
  Web site: http://www.ine.pt
  Customer Support Service
  (National network): 808 201 808
  (Other networks): +351 226 050 748
- **COVER DESIGN**
  - *Mário Bouçadas, designed on the stain glass window at INE, I.P., by the painter Abel Manta*
- **LAYOUT AND GRAPHIC DESIGN**
  - *Carlos Perpétuo*
- **PRINTING**
  - *Instituto Nacional de Estatística, I.P.*
- **EDITION**
  - *350 copies*
- **LEGAL DEPOSIT REGISTRATION**
  - *N.º 191915/03*

PRICE

[VAT 5% included]

- Single issue ..... **€ 10**
- Annual subscription (No. 1 Special Issue, No. 2 and No.3)... **€ 24**
- Annual subscription (No. 2, No. 3) ..... **€ 16**

REVSTAT

Statistical Journal

INDEX

REVSTAT – Statistical Journal
Volume 7, Number 2, June 2009, 135–149

FORECASTING TIME SERIES WITH BOOT.EXPOS PROCEDURE *

Authors: Clara Cordeiro
– Dept. Math, FCT, University of Algarve, Portugal
ccordei@ualg.pt

M. Manuela Neves
– CEAUL and Dept. Math, ISA, TULisbon, Portugal
manela@isa.utl.pt

Abstract:

- To forecast future values of a time series is one of the main goals in times series analysis. Many forecasting methods have been developed and its performance evaluated. In order to make a selection among an avalanche of such emerging methods they have to be compared in a kind of forecasting competition. One of these competitions is the M3 competition with its 3003 time series. The competition results in Makridakis and Hibon (2000) paper are frequently used as a benchmark in comparative studies.
The Boot.EXPOS approach developed by the authors, combines the use of exponential smoothing methods with the bootstrap methodology to forecast time series. The idea is to join these two approaches (bootstrap and exponential smoothing) and to construct a computational algorithm to obtain forecasts. It works in an automatic way and can be summarized as follows: (i) choose an exponential smoothing model, among several proposed using the mean squared error, and obtain the model components; (ii) fit an AR to the residuals of the adjusted model; the order of the AR is selected by AIC criterion; (iii) center the new residuals obtained in previous step and resample; (iv) obtain a bootstrapped replica of the time series according to the AR model and exponential smoothing components found in first step; (v) forecast future values according to model in (i); (vi) compute the point forecast as the mean or as the median of the predicted values. The performance of the procedure here proposed is evaluated by comparing it with other procedures presented in the M3 competition. Some accuracy measures are used for that comparison. All computational work is done using the R2.8.1 software (R Development Core Team, 2008).

Key-Words:

- *bootstrap; exponential smoothing; forecasting accuracy; M3 competition.*

AMS Subject Classification:

- 62F40, 60G25, 37M10.

*Paper accepted for presentation at COMPSTAT 2008.

1. INTRODUCTION

In our days it is well known the importance of time series studies. These studies provide indicators about a country economy, the unemployment rate, the export and import product rates, etc. A time series is a set of observations usually ordered in equally spaced intervals. The first step in the analysis of any time series is the description of the historic series. It includes the graphical representation of the data. When a time series is plotted, common patterns are frequently found. These patterns might be explained by many possible cause-and-effect relationships. Common components are the trend, seasonal effect, cyclic changes and randomness. The more interesting and ambitious task is to forecast future values of a series on the basis of its recorded past, and more specifically to calculate forecast intervals. Classical procedures to obtain forecast intervals assume that the distribution of the error process is known. Bootstrap approaches have been proposed to compute distribution free forecast intervals. The authors here propose a procedure (Boot.EXPOS) to forecast time series that is inspired on sieve bootstrap approach (Bühlmann, 1997). The Boot.EXPOS procedure starts by selecting the best exponential smoothing method according to the characteristics that a times series reveals, among a set of methods. After adjusting the best model, our attention is drawn to the residual part. The bootstrap is then used after an autoregressive adjustment, selected by AIC criterion. The time series is then reconstructed, adding the initial components (if they exist) to the bootstrapped residuals, see Cordeiro and Neves (2007a, 2007b) for more details. Forecasts are finally obtained using the model initially selected. The procedure runs automatically, so its possible to test it on large data sets. Surely there are advantages and disadvantages in this automatic process, but for now we are not discussing this. The issue here is to discuss measures of its performance, results and progress achieved. The computational work was done using the R software. Some of the R packages and functions were used, but new functions needed to be constructed.

2. EXPONENTIAL SMOOTHING METHODS

Exponential smoothing (EXPOS) refers to a set of forecasting methods, several of which are widely used. The EXPOS is a procedure that continually updates a forecast emphasizing the most recent experience, that is, recent observations are given more weight than the older observations. Single exponential smoothing, Holt's linear trend, Holt–Winters seasonal smoothing with either additive or multiplicative seasonality are some examples of EXPOS methods, see DeLurgio (1998) for more details. The forecasting performance of exponential

smoothing methods has been addressed by several authors. A very good reviewing of the past 25 years of time series forecasting is given by De Gooijer and Hyndman (2006). These methods are relatively simple but reveal robust approaches to forecasting and are accurate in model identification. The classical Box–Jenkins ARIMA models require the user to identify an appropriate model and to use at least 50 observations to have a good chance of success (Chatfield, 1978).

Table 1 shows four EXPOS methods addressed here, where data with or without trend and/or with or without seasonal components are considered. The goal is to choose the EXPOS model using the mean squared error criterion and then separate the pattern (trend or/and seasonality) components from the error term.

Table 1: The EXPOS methods considered.

Classification	**Method**
1	Single exponential smoothing
2	Holt's linear trend
3	Holt–Winters seasonal smoothing with additive seasonality
4	Holt–Winters seasonal smoothing with multiplicative seasonality

2.1. Holt–Winters method

The Holt–Winters forecasting method is applied whenever the data behavior is trendy and is seasonally. Relatively to the seasonal factor it can be additive or multiplicative, depending on the oscillatory movement along the time period. The *additive* Holt–Winters (classification 3 in Table 1) has the following recursive equations to estimate the trend and the seasonal factor at time t

$$\begin{aligned} T_t &= \alpha\,(X_t - S_{t-s}) + (1-\alpha)\,(T_{t-1} + b_{t-1}) \\ b_t &= \beta\,(T_t - T_{t-1}) + (1-\beta)\,b_{t-1} \\ S_t &= \gamma\,(X_t - T_t) + (1-\gamma)\,S_{t-s} \end{aligned}$$

with $\alpha, \beta, \gamma \in [0,1]$ and

- T_t smoothed value at end of period t after adjusting for seasonality
- X_t value of actual demand at end of period t
- S_{t-s} smoothed seasonal index, s periods ago
- b_t smoothed value of trend through period t

α smoothing constant used for T_t

β smoothing constant used to calculate the trend (b_t)

γ smoothing constant used for calculate the seasonal index in period t.

The prediction equation is

$$\hat{X}_t(h) = T_t + h \times b_t + S_{t+h-rs} \tag{2.1}$$

where $h = 1, 2, 3, ...$ is the forecast horizon and $r = 1$ if $1 \leq h \leq s$, $r = 2$ if $s < h \leq 2s$, etc.

The *multiplicative* Holt–Winters (classification 4 in Table 1) has as recursive equations

$$\begin{aligned} T_t &= \alpha (X_t / S_{t-s}) + (1 - \alpha)(T_{t-1} + b_{t-1}) \\ b_t &= \beta (T_t - T_{t-1}) + (1 - \beta) b_{t-1} \\ S_t &= \gamma (X_t / T_t) + (1 - \gamma) S_{t-s} \end{aligned}$$

and prediction equation

$$\hat{X}_t(h) = (T_t + h \times b_t) \times S_{t+h-rs} \ , \tag{2.2}$$

where the parameters are defined above.

The exponential smoothing parameters (α, β, γ) are estimated by minimizing the sum of squared errors and are restricted to values in (0,1). Simple exponential smoothing and Holt's method (classification 1 and 2 in Table 1) are derived from the above equations considering the corresponding exponential smoothing parameters (β, γ) to be set to zero.

3. ABOUT BOOTSTRAP

The bootstrap resampling technique (Efron, 1979) is a very popular methodology in independent data because of its simplicity and nice properties. It is a computer-intensive method that presents solutions in situations where the traditional methods fail. Efron's bootstrap classical approach has revealed inefficient in the context of dependent data, such as in the case of time series, where the dependence data arrangement should be kept during the resampling scheme. But for dependent data the generating process is often not fully specified. Then there exists no unique way for resampling.

A great development in the resampling methods area for dependent data has been observed, see Lahiri (2003). The majority of those methods suggests the use of blocks, in order to keep the dependence structure. Different versions

of blocking differ in the way as blocks are constructed. The most well known versions are: Nonoverlapping Block Bootstrap (Carlstein, 1992); Moving Block Bootstrap (Künsch, 1989); Circular Block Bootstrap (Politis and Romano, 1992) and Stationary Block Bootstrap (Politis and Romano, 1994). For a large class of stationary processes, Bühlmann (1997) presents the sieve bootstrap method based on a sieve of autoregressive processes of increasing order. Recently Alonso *et al.* (2002) extended the sieve approach in order to obtain prediction intervals. In Cordeiro and Neves (2006) several bootstrap methodologies for dependent data were compared in constructing forecast intervals and the sieve bootstrap has revealed as a good compromise for obtaining forecast intervals.

3.1. Particular approach: sieve bootstrap

In 1997, Bühlmann proposed a bootstrap scheme called sieve bootstrap. This method is based on the idea of fitting parametric models first and afterwards resampling from the residuals. However the model is chosen adaptively rather than considering a pre-fixed model. This approach is different from other bootstrap methods, the sample bootstrap is (conditionally) stationary and does not present structure of dependence. Another different feature is that the sieve bootstrap sample is not a subsample from the original data, as in other methods. Given a sample $X_1, ..., X_n$, from a stationary process, select the order $p = p(n)$ of an autoregressive approximation by AIC criterion. The autoregressive coefficients are Yule–Walker estimates. The AR(p) model is used to filter the residuals series. The residuals are then centered and the empirical cumulative distribution function of these residuals is obtained. From this distribution we get an i.i.d. resample of the centered residuals. Use the AR for obtaining a new series X_t^* by recursion. Given $X_1^*, ..., X_T^*$ compute the estimation of the autoregressive coefficients and then obtain future bootstrap observations by recursion from the new series.

3.2. Boot.EXPOS procedure

Here a different approach is proposed: first fit an EXPOS model to the data and then to proceed like the above procedure over the residuals — Boot.EXPOS (Cordeiro and Neves, 2008). The Boot.EXPOS procedure starts by selecting the best forecasting method, section 2. The seasonality and trend components (if they exist) are removed from the initial series and only added at the end to reconstruct the series. In between, the procedures go on like the sieve bootstrap approach. This general steps are described in Table 2, where it is also established a comparison between the previous and the new approach.

Table 2: The previous and the new approach.

Sieve bootstrap	**Boot.EXPOS**
	Step 0: Select the best EXPOS method; components are removed and the residuals obtained;
Step 1: Adjust an autoregressive model with increasing order p using AIC criterion;	**Step 1:** Adjust an autoregressive model with increasing order p using AIC criterion;
Step 2: Obtain the residuals;	**Step 2:** Obtain the residuals;
For B replicates:	*For B replicates:*
Step 3: Resample the centered residuals;	**Step 3:** Resample the centered residuals;
Step 4: Use AR for obtaining a new series by recursion;	**Step 4:** Use AR for obtaining a new series by recursion;
Step 5: Fit AR(p) to the new series;	**Step 5:** Add the components in Step 0 to the new series; fit EXPOS method (same type as in Step 0);
Step 6: Obtain the predicted values from the new series using the previous AR(p) fit.	**Step 6:** Obtain the predicted values from the new series using the previous EXPOS fit.

3.2.1. Dealing with some statistical issues

The initial step in the Boot.EXPOS procedure is to fit an EXPOS method selected by MSE, section 2. Then the random part is separated from the other patterns, such as trend or/and seasonality (these patterns are added at the end). Next, a test on the stationarity of the random part is performed, before the AR adjustment. If the stationarity is not accepted, transformation to the data is required.

Statistical tests used to study stationarity and data transformations:

- Test for a unit root in a time series when the model under consideration in the null hypothesis does not present autocorrelation in the error term. In such case, simple version of the Dickey–Fuller test is the most appropriate (Halkos and Kevork, 2005).
- Box–Cox transformations (Box and Cox, 1964) can impose stability on data variance stability and can make seasonal effect additive. Transformation is used whenever data exhibit multiplicative seasonality, *i.e.*,

when EXPOS model choice is 4 according to Table 1. Although the use of these transformations are not worthwhile in some cases, for example in economic data (Nelson and Granger, 1979), there can be some advantages in using it. Computational work with $\lambda = 0$, $-1 \leq \lambda \leq 1$, $0 \leq \lambda \leq 1$, $0 \leq \lambda \leq 0.5$, $0 \leq \lambda \leq 0.9$ and $0 \leq \lambda \leq 0.99$, was performed and it has revealed a good option when $0 \leq \lambda \leq 1$.

- Differencing can transform a non-stationary series to a stationary series. The KPSS (Kwiatkowski *et al.*, 1992) procedure tests for the null hypothesis that a time series has a stationary root against a unit-root.

Boot.EXPOS application makes all these procedures whenever the non-stationarity is detected. But if the data goes over to the above processes and the presence of the non-stationarity is still detected in the data, the best EXPOS method is used to obtain forecast without bootstrapping.

4. FORECAST ACCURACY MEASURES

To evaluate the performance of Boot.EXPOS procedure some accuracy measures are used. In M3 competition some of those measures are calculated. Concerning this subject it is also interesting to read the article proposed by Hyndman and Koehler (2006).

Let X_t denote the observation at time t and $\hat{X}_t$ the forecast of X_t. The forecast error is defined by $e_t = X_t - \hat{X}_t$. The forecasts are computed for a hold-out period. Thus the out-of-sample forecasts $\hat{X}_n(1), ..., \hat{X}_n(h)$ are computed based on the data from time $t = 1, ..., n$. Accuracy measures are then computed in order to compare our results with those presented in Makridakis and Hibon (2000) in `http://www.forecastingprinciples.com/m3-competition.html`. The following accuracy measures are here considered:

Table 3: Two accuracy measures.

Acronyms	Name	Definition
sMAPE	Symmetric Mean Absolute Percentage Error	$\text{mean}\left(200\,\dfrac{\lvert e_t\rvert}{X_t + \hat{X}_t}\right)$
RMSE	Root Mean Squared Error	$\sqrt{\text{mean}(e_t^2)}$

5. A COMPETITIVE EXAMPLE

In this section, the performance and evaluation of the Boot.EXPOS procedure is analyzed. A 3003 time series data — M3 Competition — is used in this analysis. This huge data set has been used by many researchers as a powerful tool to test new forecasting methods and this is what we shall present next.

5.1. The competing data

The M3 Competition data set involves 3003 time series selected on a quota basis of 6 different types of series: micro, industry, finance, demographic and other; and 4 different time intervals between successive observations: yearly, quarterly, monthly and other. The historical values of each series are at least 14 observations for yearly data, 16 for quarterly data, 48 for monthly data and 60 observations for other data. The time horizons of forecasting are 6 periods for yearly data, 8 periods for quarterly and other data, and 18 periods for monthly data. Table 4 shows the number of series in each category and Figure 1 shows the plot of examples of some types.

Table 4: The 3003 time series distribution.

Period	Type of times series data						Total
	Demographic	Finance	Industry	Macro	Micro	OTHER	
Monthly	111	145	334	312	474	52	1428
OTHER	0	29	0	0	4	141	174
Quarterly	57	76	83	336	204	0	756
Yearly	245	58	102	83	146	11	645
Total	413	308	519	731	828	204	3003

5.2. Some tools use in competition

An automatic procedure to analyze each time series is necessary as there is a large set of time series. All the intensive computational work is performed using R 2.8.1 software (R Development Core Team, 2008). Packages such as **car**, **FitAR**, **forecast**, **Mcomp**, **tseries**, among others, are widely used in the implementation of the new procedures in R language: **best.EXPOS()** to select the best EXPOS method and **boot.EXPOS()** a procedure using bootstrap and EXPOS.

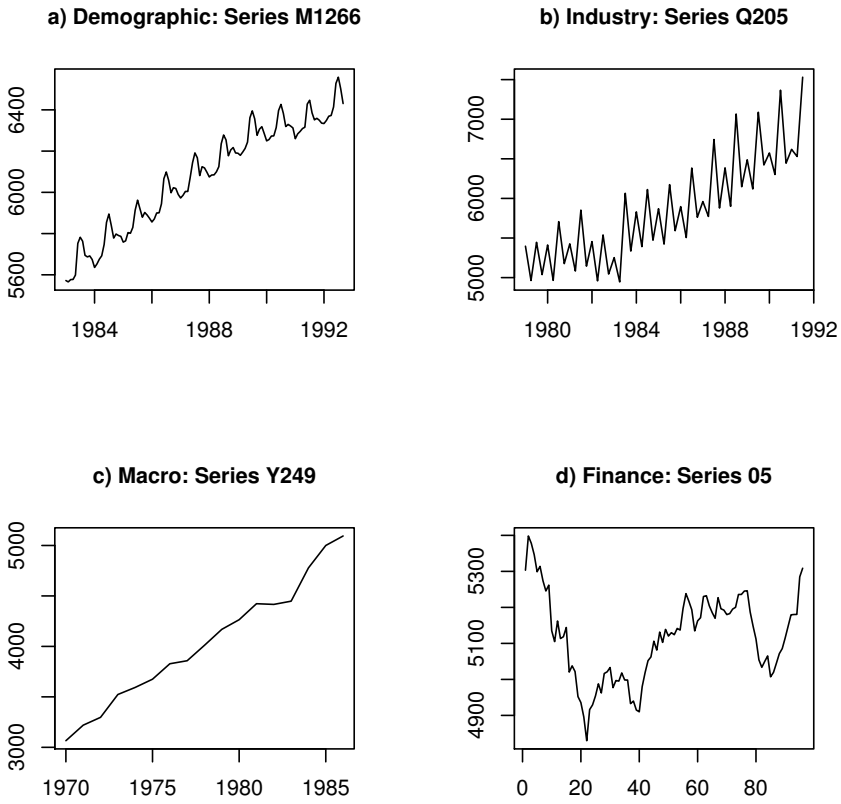

Figure 1: M3 competition category: monthly (a), quarterly (b), yearly (c) and "other" (d).

5.3. EXPOS *vs* Boot.EXPOS

All time series considered are strictly positive. For each category, different forecasting periods are considered: 6 for yearly, 8 for quarterly and "other", 18 for monthly. For the M3 competition a comparison between the forecasts produced by any of the four EXPOS methods and the forecast produced by Boot.EXPOS procedure is showed in Figure 2. Symmetric mean absolute percentage error is one of the accuracy measures used in this study. It can be seen that for the time period interval yearly and Other the resampling technique is not a favorable procedure. In the monthly and quarterly cases the scenario is better. It seems that for time series with components trend and seasonality the Boot.EXPOS procedure can be a good compromise in forecasting.

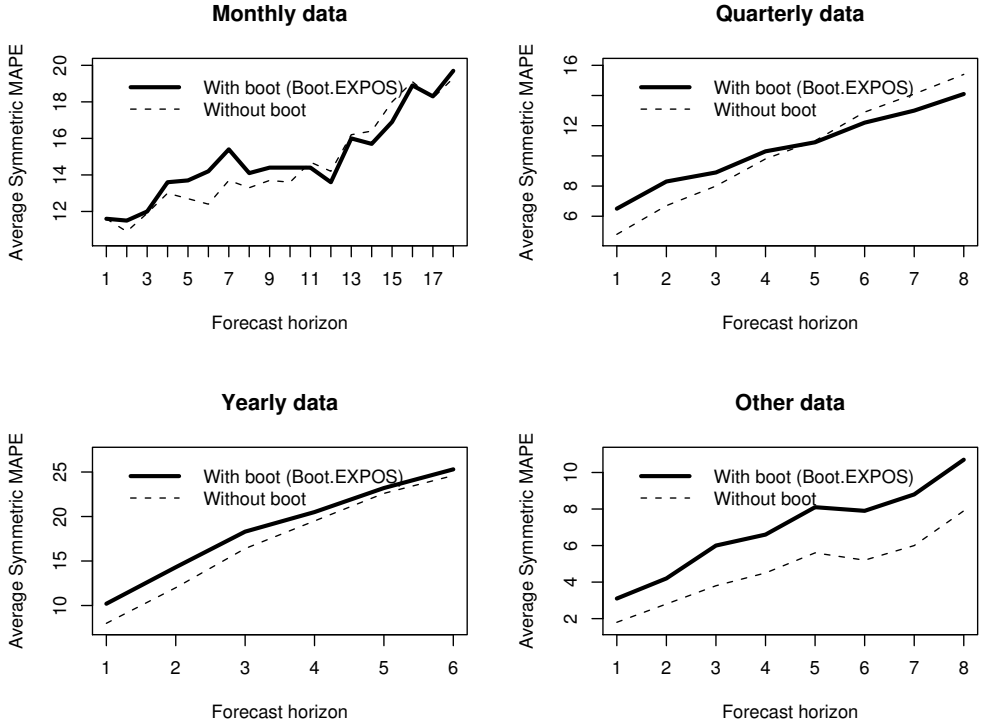

Figure 2: EXPOS and Boot.EXPOS comparison using M3 competition data.

5.4. Boot.EXPOS *vs* six competition methods

Each M3 competition series was classified according to Table 1, using function **best.EXPOS**(). This function selects the model that presents the minimum MSE. For monthly time series the distribution is showed in Table 5. Here, to illustrate Boot.EXPOS only **monthly time series** will be considered. For each series the procedure scheme in Table 2, right, goes from Step 0 to Step 6. One thousand replications ($B = 1000$) are carried out and forecasts are obtained for each replication. Two forecast estimates are obtained: the average and median of the B forecasted values. Only the average is used here and is based in the 95% central forecast simulations.

Table 5: Number of monthly time in each method.

Category	Classification			
	1	2	3	4
Monthly	515	300	345	268

In these competitions the participating experts are asked to obtain a given number of forecast values for each data period. Their forecasts are latter compared with the actual values and the accuracy of such forecasts is calculated. Makridakis and Hibon (2000) lists 24 methods of forecasting used in the M3 competition. Naive2, Box–Jenkins automatic, ForecastPro, THETA, RBH and ForecastX are the six methods that present the best performance in M competition (Makridakis *et al.*, 1982) and M3 competition (Makridakis and Hibon, 2000). These well behaved 6 methods (Hyndman *et al.*, 2002) are used in comparison to our approach. For the monthly case the sMAPE and RMSE are present in Tables 6 and 7, respectively.

Table 6: Average symmetric MAPE: 1428 monthly series.

Method	Forecasting horizon																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Naive2	15.0	13.5	15.7	17.0	14.9	14.4	16.0	15.6	15.9	16.0	16.7	16.0	18.1	18.4	19.3	21.3	19.6	20.7
B-J automatic	12.3	11.7	12.8	14.3	12.7	12.3	13.6	13.0	13.4	13.4	14.5	14.1	16.2	16.9	17.8	19.7	18.1	19.3
ForecastPRO	11.5	10.7	11.8	13.0	11.9	12.0	13.0	12.7	13.0	13.0	13.9	13.3	15.3	15.4	16.4	18.2	16.8	18.3
THETA	11.2	10.7	11.8	12.4	12.2	12.2	13.6	12.7	13.2	13.4	13.5	13.2	15.4	15.3	16.5	17.9	17.0	18.5
RBF	13.7	12.3	13.7	14.3	12.3	12.5	14.6	13.5	14.2	14.5	14.1	14.1	16.1	15.8	17.3	18.3	16.8	17.8
ForecastX	11.6	11.2	12.6	14.0	12.4	12.0	13.0	12.8	13.1	13.4	13.9	14.0	15.8	16.6	17.8	19.4	18.1	18.8
Boot.EXPOS	11.6	11.5	12.0	13.6	13.7	14.2	15.4	14.1	14.4	14.4	14.4	13.6	16.0	15.7	16.9	18.9	18.3	19.7

Table 7: Root mean squared error: 1428 monthly series.

Method	Forecasting horizon										Average 1–18
	1	2	3	4	5	6	8	12	15	18	
Naive2	1144	1367	1466	1643	1363	1201	1453	1329	1766	1673	1448
B-J automatic	864	942	934	1061	1006	1100	1107	1208	1454	1563	1185
ForecastPRO	812	905	913	1068	1032	990	1157	1135	1411	1463	1146
THETA	810	936	1067	1181	1130	979	1170	1138	1445	1487	1168
RBF	984	1636	1468	1850	1503	1000	1355	1197	1764	1651	1459
ForecastX	794	977	920	1087	1008	966	1175	1169	1457	1510	1163
Boot.EXPOS	840	988	1214	1627	1370	1540	1702	1149	1580	1828	1405

In a first look the results have a reasonable classification among the six methods in study. But if the results are separated into classification 1, 2, 3 and 4 (Table 1) the scenario gets better. Results for the above six methods and for the Boot.EXPOS procedure are given in Tables 8, 9, 10 and 11, corresponding to each classification.

The performance of the various methods depends upon the length of the forecasting horizon. Our method presents sMAPE among the values of the selected methods (Table 6), particularly it has a very good performance when the time series model classification is 4 (Table 11). When the time series model is the Holt–Winters additive (classification 3) Table 10 shows a reasonable behavior. Regarding the other classification (1 and 2) Tables 8 and 9 show poor results.

It is evident that the procedure is best succeeded when the time series is seasonal and trendy. For time series in classifications 1 and 2, the authors thought that the classical EXPOS methods could produce accurate forecasts but in fact a simulation computational study has revealed the opposite. For some of these time series whenever stationarity was achieved the Boot.EXPOS forecasts were better values.

Table 8: Average symmetric MAPE: 1.

Method	Forecasting horizon																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Naive2	16.2	16.2	17.9	20.5	18.5	18.4	20.8	20.4	19.2	19.3	20.4	19.4	21.5	22.2	23.5	26.3	22.7	26.0
B-J automatic	12.3	12.9	15.3	16.8	16.3	16.3	17.4	18.0	18.3	17.0	18.4	17.1	20.4	20.5	22.6	23.4	21.5	24.7
ForecastPRO	12.1	12.4	14.5	16.1	15.3	15.9	16.6	17.8	17.2	16.5	17.8	16.4	20.0	19.8	21.7	23.0	20.7	23.8
THETA	12.0	12.8	15.0	15.9	15.8	15.8	17.7	17.9	16.8	16.5	17.3	16.1	19.9	19.6	21.3	22.6	19.9	23.5
RBF	14.9	14.8	16.8	18.3	16.2	16.8	18.7	18.7	17.2	17.3	17.6	17.0	20.4	20.0	22.3	23.5	20.4	22.8
ForecastX	11.1	12.9	14.3	16.2	15.3	15.9	16.6	17.7	16.6	16.4	17.2	16.5	19.4	19.6	21.9	23.2	20.7	23.3
Boot.EXPOS	14.8	14.8	15.5	17.6	16.5	15.9	17.2	17.2	16.8	17.0	19.2	17.4	20.1	21.2	21.7	24.4	21.8	24.4

Table 9: Average symmetric MAPE: 2.

Method	Forecasting horizon																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Naive2	1.4	2.0	2.7	3.8	3.9	4.3	4.9	5.2	5.4	5.9	6.5	7.2	7.7	7.6	8.6	9.1	9.5	10.1
B-J automatic	1.3	1.7	2.3	3.4	3.5	3.7	4.3	4.6	4.7	5.2	5.6	6.2	6.6	6.6	7.5	8.2	8.5	9.0
ForecastPRO	1.3	1.7	2.1	3.2	3.2	3.6	4.0	4.4	4.5	5.0	5.6	5.9	6.5	6.3	7.5	8.1	8.6	9.3
THETA	1.3	1.7	2.3	3.3	3.3	3.6	4.2	4.4	4.6	5.1	5.8	6.3	6.7	6.7	7.8	8.2	8.6	9.2
RBF	3.0	3.3	3.7	4.3	3.9	4.1	4.7	4.9	4.8	5.4	6.0	6.2	6.6	6.4	7.3	7.6	7.7	8.3
ForecastX	1.3	1.7	2.3	3.4	3.5	3.8	4.5	4.8	5.0	5.5	6.0	6.6	7.1	7.2	8.2	8.8	9.3	9.9
Boot.EXPOS	4.6	5.0	5.8	6.9	6.2	7.7	7.6	7.8	8.3	7.6	8.3	9.2	10.6	10.5	12.0	11.7	11.6	12.3

Table 10: Average symmetric MAPE: 3.

Method	Forecasting horizon																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Naive2	12.8	13.3	13.9	16.0	13.7	12.7	14.8	13.8	15.4	15.7	16.7	14.0	16.8	15.7	17.6	17.3	17.4	18.3
B-J automatic	11.0	11.4	10.5	12.6	11.4	10.7	12.3	11.6	13.2	12.7	13.8	12.7	14.3	14.1	15.7	15.8	15.2	16.2
ForecastPRO	10.5	10.6	10.4	12.4	11.0	10.7	11.9	11.3	13.3	12.9	13.7	12.0	13.6	13.4	14.8	15.6	14.8	15.9
THETA	10.5	10.6	10.7	12.2	11.6	11.0	12.5	11.4	14.1	12.7	14.1	12.2	13.6	13.3	15.0	15.2	15.3	16.7
RBF	12.8	12.5	12.3	13.5	11.1	11.2	14.3	12.3	14.7	14.8	15.3	13.6	16.1	15.1	16.6	17.0	16.1	17.1
ForecastX	10.7	10.9	10.4	12.7	11.4	10.4	12.3	11.6	13.2	12.9	14.3	12.9	14.5	14.0	15.9	15.6	16.3	16.7
Boot.EXPOS	9.5	10.3	10.9	12.8	14.6	15.4	17.9	15.7	16.4	13.6	13.4	12.0	13.9	12.2	15.5	16.8	17.7	19.2

Table 11: Average symmetric MAPE: 4.

Method	Forecasting horizon																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Naive2	23.7	18.1	23.0	22.8	19.3	18.5	19.6	19.3	19.7	19.4	19.8	20.1	22.5	24.0	23.7	27.7	25.0	24.8
B-J automatic	19.8	16.9	18.9	20.0	16.2	15.6	17.3	15.3	14.8	15.9	17.3	17.4	20.3	22.6	21.6	26.8	23.6	23.5
ForecastPRO	17.9	14.6	16.5	16.7	14.9	14.8	16.2	14.7	14.4	15.0	15.8	16.3	18.1	19.0	18.8	22.5	20.2	21.2
THETA	16.8	14.4	15.8	15.1	15.0	15.2	16.7	14.4	14.6	16.5	14.5	15.8	18.5	18.6	19.0	21.9	21.1	21.6
RBF	19.6	15.4	18.3	17.6	15.2	15.2	17.4	15.4	16.7	17.2	15.2	16.6	18.3	18.6	19.7	21.5	19.7	19.8
ForecastX	18.6	15.4	19.1	19.4	16.1	15.0	15.8	14.6	14.8	16.2	15.6	17.2	19.0	22.0	21.9	25.6	22.8	22.2
Boot.EXPOS	16.2	14.0	13.6	14.2	15.7	16.5	17.2	13.4	14.2	18.1	13.5	13.7	17.0	15.5	15.2	18.9	20.0	19.7

6. CLOSING COMMENTS

An automatic procedure based on EXPOS and bootstrap methodology is presented. Our methodology was applied to the M3 competition data. Accuracy measures such as sMAPE, RMSE were calculated and compared with the measures obtained for the best six competition methods. The method revealed a very good performance for series with seasonal and trendy components. Short length time series showed difficulties in the forecasting procedure. In fact Billah *et al.* (2006) decide to separate the short series in order to obtain... *plausible results with satisfactory level of statistical reliability.*

Some research is in progress for improving the results obtained, as well as for considering new methods for the initial fitting.

ACKNOWLEDGMENTS

The authors are grateful to Professor Michèle Hibon for kindly providing the forecasts submitted to the M3 competition.

REFERENCES

[1] Alonso, A.M.; Peña, D. and Romo, J. (2002). Forecasting time series with sieve bootstrap, *Journal of Statistical Planning and Inference*, **100**, 1–11.

[2] Billah, B.; King, M.; Snyder, R. and Koehler, A. (2006). Exponential smoothing model selection for forecasting, *International Journal of Forecasting*, **22**, 239–247.

[3] Box, E.P. and Cox, D.R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society*, **B**.

[4] Bühlmann, P. (1997). Sieve bootstrap for time series, *Bernoulli*, **3**, 123–148.

[5] Carlstein, E. (1992). Resampling techniques for stationary time-series: some recent developments, *New Directions in Time Series Analysis*, **27**, 75–85.

[6] Chatfield, C. (1978). The Holt–Winters forecasting procedure, *Applied Statistics*, **27**(3), 264–279.

[7] Cordeiro, C. and Neves, M.M. (2006). *The Bootstrap methodology in time series forecasting.* In "Proceedings of CompStat2006" (J. Black and A. White, Eds.), Springer Verlag, 1067–1073.

[8] Cordeiro, C. and Neves, M.M. (2007a). *The Bootstrap prediction intervals: a case-study*. In "Proceedings of the 22nd International Workshop on Statistical Modelling (IWSM2007)" (J. Castillo, A. Espinal and P. Puig, Eds.), Springer Verlag, 191–194.

[9] Cordeiro, C. and Neves, M.M. (2007b). *Resampling techniques in time series prediction: a look at accuracy measures*. In "Proceedings of the 56th Session of the International Statistical Institute (ISI 2007)" (M.I. Gomes, D. Pestana and P. Silva, Eds.), Springer Verlag, 353 in CD-ROM.

[10] Cordeiro, C. and Neves, M.M. (2008). *Bootstrap and exponential smoothing working together in forecasting time series*. In "Proceedings in Computational Statistics (COMPSTAT 2008)" (Paula Brito, Editor), Physica-Verlag, 891–899 in CD-ROM.

[11] De Gooijer, J.G. and Hyndman, R.J. (2006). 25 years of time series forecasting, *International Journal of Forecasting*, **22**, 443–473.

[12] DeLurgio, S.A. (1998). *Forecasting Principles and Applications*, McGraw-Hill International Editions.

[13] Efron, B. (1979). Bootstrap methods: another look at the Jackknife, *The Annals of Statistics*, **7**, 1–26.

[14] Halkos, G. and Kevork, I. (2005). A comparison of alternative unit root tests, *Journal of Applied Statistics*, **32**(1), 45–60.

[15] Hyndman, R.; Koehler, A.; Snyder, R. and Grose, S. (2002). A state framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, **18**, 439–454.

[16] Hyndman, R. and Koehler, A. (2006). Another look at measures of forecast accuracy, *International Journal of Forecasting*, **22**, 679–688.

[17] Künsch, H. (1989). The Jackknife and the Bootstrap for general stationary observations, *The Annals of Statistics*, **17**(3), 1217–1241.

[18] Kwiatkowski, D.; Phillips, P.; Schmidt, P. and Schin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, **54**, 159–178.

[19] Lahiri, S.N. (2003). *Resampling Methods for Dependent Data*, Springer Verlag.

[20] Makridakis, S.; Anderson, A.; Carbone, R.; Fildes, R.; Hibon, M.; Lewandowski, R.; Newton, J.; Parzen, E. and Winkler, R. (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition, *Journal of Forecasting*, **1**, 111–153.

[21] Makridakis, S. and Hibon, M. (2000). The M3-Competition: results, conclusions and implications, *International Journal of Forecasting*, **16**, 451–476.

[22] Nelson, H.L. and Granger, C. (1979). Experience with using the Box–Cox transformation when forecasting economic time series, *Journal of Econometrics*, **10**, 57–69.

[23] Politis, D.R. and Romano, J.P. (1992). *A Circular Block-Resampling Procedure for Stationary Data*, Wiley, New York, Springer, 225–270.

[24] Politis, D.R. and Romano, J.P. (1994). The stationary bootstrap, *Journal of the American Statistical Association*, **89**(428), 1303–1313.

[25] R development core team (2008). *R: A language and environment for statistical computing*, R Foundation for statistical computing, Vienna, Austria, `http://CRAN.R-project.org`.

REVSTAT – Statistical Journal
Volume 7, Number 2, June 2009, 151–170

MONITORING INDUSTRIAL PROCESSES WITH ROBUST CONTROL CHARTS

Authors: Fernanda Figueiredo
– C.E.A.U.L. and Fac. de Economia da Universidade do Porto, Portugal
otilia@fep.up.pt

M. Ivette Gomes
– F.C.U.L. (D.E.I.O.) and C.E.A.U.L., Universidade de Lisboa, Portugal
ivette.gomes@fc.ul.pt

Received: July 2008 Revised: October 2008 Accepted: December 2008

Abstract:

- The Shewhart control charts, used for monitoring industrial processes, are the most popular tools in *Statistical Process Control* (SPC). They are usually developed under the assumption of independent and normally distributed data, an assumption rarely true in practice, and implemented with estimated control limits. But in general, we essentially want to control the process mean value and the process standard deviation, independently of the data distribution. In order to monitor these parameters, it thus seems sensible to advance with control charts based on *robust statistics*, because these statistics are expected to be more resistant to moderate changes in the underlying process distribution. In this paper, we investigate the advantage of using control charts based on *robust statistics*. Apart from the traditional control charts, the sample *mean* and the sample *range* charts, we consider robust control charts based on the *total median* and on the *total range* statistics, for monitoring the process mean value and the process standard deviation, respectively. Through the use of Monte Carlo simulations, we compare these charts in terms of robustness and performance.

Key-Words:

- *statistical process control; control charts; robust estimation; Monte Carlo methods.*

AMS Subject Classification:

- 62G05, 62G35, 62P30, 65C05.

1. INTRODUCTION

The most commonly used charts for monitoring industrial processes, or more precisely, a quality characteristic X at the targets μ_0 and σ_0, the desired mean value and standard deviation of X, respectively, are the Shewhart control charts with 3-sigma control limits. These charts are usually developed under the assumptions of independent and normally distributed data, and have control limits (CL's) of the form

$$LCL_W = \mathbb{E}(W) - 3\sqrt{\mathbb{V}(W)}\ , \qquad UCL_W = \mathbb{E}(W) + 3\sqrt{\mathbb{V}(W)}$$

where W, LCL, UCL, $\mathbb{E}$ and $\mathbb{V}$ denote the control statistic, the lower control limit, the upper control limit, the expected value operator and the variance operator, respectively. More precisely, to monitor the process mean value μ at $\mu = \mu_0$, it is common to implement a two-sided sample mean chart, $\overline{X}$, also denoted M-chart, with lower and upper control limits given by

$$LCL_M = \mu_0 - 3\,\sigma_0/\sqrt{n}\ , \qquad UCL_M = \mu_0 + 3\,\sigma_0/\sqrt{n}\ . \tag{1.1}$$

To monitor the process standard deviation σ at $\sigma = \sigma_0$, it is common to implement a sample range chart, R, with lower and upper control limits given by

$$LCL_R = d_2\,\sigma_0 - 3\,d_3\,\sigma_0\ , \qquad UCL_R = d_2\,\sigma_0 + 3\,d_3\,\sigma_0\ , \tag{1.2}$$

where d_2 and d_3 are constants tabulated for standard normal data, and presented in Table 2 (Section 2.1) for the most common rational subgroups size, n. General details about control charts can be found in Ryan (2000) and Montgomery (2005), among others.

For normal data and when it is not necessary to estimate the control limits, the Shewhart control charts exhibit a reasonable high performance to detect moderate to large changes in the process parameters. However, despite of the importance of the normal distribution in *Statistical Process Control* (SPC), the experience tells us that even in potential normal situations there is some possibility of having an underlying non-normal distribution, with moderate to strong asymmetry and with tails heavier than the normal tail, as well as a significant correlation between the observations.

Additionally, the target values μ_0 and σ_0 are not usually fixed given values, and we have to estimate them, in order to determine the control limits of the chart. Several studies refer that, even for normal data, we are able to obtain control charts with estimated control limits with the same properties as the corresponding charts with true limits, only if we use a large number of initial rational subgroups in the estimation. Moreover, we should determine the control limits in a robust way, in order to minimize the effect of possible outliers in the initial subgroups.

The effect of the estimation of the control limits and of the non-normality in the performance of the usual control charts can be found in Rocke (1989, 1992), Quesenberry (1993), Amin and Lee (1999), Chakraborti (2000, 2006), Nedumaran and Pignatiello (2001), Champ and Jones (2004), Figueiredo and Gomes (2004, 2006) and Jensen *et al.* (2006), among others. Schilling and Nelson (1976), Bai and Choi (1995) and Castagliola (2000), also among others, provide different corrections to the control limits of the usual control charts in order to maintain the expected false alarm rate, whenever monitoring non-normal data.

To sum up, the traditional control charts must be used carefully. If the model underlying the process is far from the normal, we can decide for the implementation of a control chart associated with the specific distribution underlying the process, whenever this seems necessary and feasible. Alternatively, we can decide for the implementation of a robust control chart, less sensitive to the normality assumption. In this paper, we shall investigate the benefits of using control charts based on robust control statistics, so that we do not have either a very high or a very low false alarm rate whenever the parameters to be controlled are close to the targets, although the data is no longer normal, together with the use of robust estimates of the upper and lower control limits. Some considerations about "robust" estimation can be found in Hampel (1971), Hoaglin *et al.* (1983), Lax (1985), Hampel *et al.* (1986), Figueiredo (2003a, 2003b) and Figueiredo and Gomes (2004), among others.

In Section 2, we provide some information about the total median and the total range statistics, analyzing the robustness and efficiency of these location and scale estimators, as well as their sampling distribution. These are the robust statistics considered in this study, used in the estimation and monitoring of the process mean value and the process standard deviation, respectively, alternatively to the usual sample mean and sample range statistics. In Section 3, we present some simulation results about the robustness and the comparative performance of control charts based on classical and robust estimation of mean values and standard deviations.

2. THE TOTAL MEDIAN AND THE TOTAL RANGE STATISTICS

Let us denote $(X_1, X_2, ..., X_n)$ a random sample of size n taken from a process X with distribution function (d.f.) F, and $(X_{1:n}, X_{2:n}, ..., X_{n:n})$ the random sample of the associated ascending order statistics (o.s.). Given an observed sample $(x_1, x_2, ..., x_n)$, the associated bootstrap random sample, $(X_1^*, X_2^*, ..., X_n^*)$, is a random sample of independent, identically distributed replicates from a random variable X^*, with d.f. equal to the empirical d.f. of our observed sample, i.e.,

given by

$$F_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} I_{\{x_i \le x\}} , \qquad \text{with} \quad I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

the indicator function of the set A. We next define the *bootstrap median* and the *bootstrap range* as the median and the range, respectively, of the bootstrap random sample. The *bootstrap median* is thus given by

$$BMd = \begin{cases} X^*_{m:n} & \text{if } \; n = 2\,m-1\,, \\ \big(X^*_{m:n} + X^*_{m+1:n}\big)/2 & \text{if } \; n = 2\,m, \quad m = 1, 2, ... \end{cases}$$

and the *bootstrap range* is given by $BR = X^*_{n:n} - X^*_{1:n}$.

Remark 2.1. Note that given an observed sample $(x_1, x_2, ..., x_n)$, the support of the bootstrap median is the set $\big\{(x_{i:n} + x_{j:n})/2,\, 1 \le i \le j \le n\big\}$, and the support of the bootstrap range is the set $\big\{x_{j:n} - x_{i:n},\, 1 \le i \le j \le n\big\}$.

Let us denote α_{ij} and β_{ij} the following probabilities:

$$\alpha_{ij} := \mathbb{P}\Big(BMd = \frac{x_{i:n} + x_{j:n}}{2}\Big) , \qquad 1 \le i \le j \le n , \tag{2.1}$$

$$\beta_{ij} := \mathbb{P}\Big(BR = x_{j:n} - x_{i:n}\Big) , \qquad 1 \le i < j \le n , \tag{2.2}$$

with $\mathbb{P}(A)$ denoting the probability of the event A.

Definition 2.1. The total median statistic, denoted TMd, is given by

$$TMd := \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{ij} \Big(\frac{X_{i:n} + X_{j:n}}{2}\Big) =: \sum_{i=1}^{n} a_i\, X_{i:n} , \tag{2.3}$$

and the total range statistic, denoted TR, is given by

$$TR := \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} \Big(X_{j:n} - X_{i:n}\Big) =: \sum_{i=1}^{n} b_i\, X_{i:n} , \tag{2.4}$$

where the coefficients a_i and b_i are thus given by

$$a_i = \frac{1}{2}\left(\sum_{j=i}^{n} \alpha_{ij} + \sum_{j=1}^{i} \alpha_{ji}\right) \quad \text{and} \quad b_i = \sum_{j=1}^{i-1} \beta_{ji} - \sum_{j=i+1}^{n} \beta_{ij} , \qquad 1 \le i \le n . \tag{2.5}$$

Cox and Iguzquiza (2001) and Figueiredo and Gomes (2004, 2006) present explicit expressions for α_{ij} and β_{ij} in (2.1) and (2.2), respectively, which enable the computation of the weights a_i and b_i, $1 \le i \le n$, in (2.3) and (2.4), respectively, through the use of (2.5).

Remark 2.2. Note that the coefficients a_i and b_i are independent of the underlying model F, and only depend on the sample size n. A linear combination of the sample o.s., with weights given by these coefficients, such as the *TMd* and the *TR* statistics, in (2.3) and (2.4), respectively, define a kind of "robust" trimmed-mean, where the percentage of trimming is determined independently of the underlying distribution of the data, and a "robust" range. The extreme observations have a smaller influence in these statistics than in the sample mean and in the sample range. They can thus be used to estimate the location and the scale parameters whenever there is a possibility of disturbances in the data, such as outliers or contaminated data. In Table 1 we present, for each entry i, the values of the coefficients a_i and b_i with three decimal figures, for the most usual rational subgroups size, n, in SPC.

Table 1: Coefficients a_i and b_i, $a_i = a_{n-i+1}$ and $b_i = -b_{n-i+1}$, $1 \le i \le n$.

i	n	3	4	5	6	7	8	9	10
1	a_i	0.259	0.156	0.058	0.035	0.010	0.007	0.001	0.001
	b_i	−0.750	−0.690	−0.672	−0.666	−0.661	−0.657	−0.653	−0.652
2	a_i	0.482	0.344	0.259	0.174	0.098	0.064	0.029	0.019
	b_i	0.000	−0.198	−0.240	−0.246	−0.245	−0.244	−0.242	−0.241
3	a_i			0.366	0.291	0.239	0.172	0.115	0.078
	b_i			0.000	−0.058	−0.073	−0.077	−0.078	−0.079
4	a_i					0.306	0.257	0.221	0.168
	b_i					0.000	−0.016	−0.020	−0.022
5	a_i							0.268	0.234
	b_i							0.000	−0.004

2.1. Location and scale estimators: robustness and efficiency

The skewness of a model is often measured through two different coefficients, the Fisher and the Bowley skewness coefficients. The Fisher skewness coefficient of a d.f. F, denoted γ, is given by

$$\gamma := \mu_3/\mu_2^{3/2}, \tag{2.6}$$

where μ_r denotes the r-th central moment of F. The Bowley skewness coefficient (also called quartile skewness coefficient), denoted $\gamma_{\rm B}$, is given by

$$\gamma_{\rm B} := \frac{\left(F^{-1}(0.75) - F^{-1}(0.5)\right) - \left(F^{-1}(0.5) - F^{-1}(0.25)\right)}{F^{-1}(0.75) - F^{-1}(0.25)}, \tag{2.7}$$

where F^{-1} denotes the inverse functions of F.

The tail-weight coefficient of a distribution F here considered, and denoted τ, is given by

$$\tau := \frac{1}{2} \frac{\frac{F^{-1}(0.99)-F^{-1}(0.5)}{F^{-1}(0.75)-F^{-1}(0.5)} + \frac{F^{-1}(0.5)-F^{-1}(0.01)}{F^{-1}(0.5)-F^{-1}(0.25)}}{\frac{\Phi^{-1}(0.99)-\Phi^{-1}(0.5)}{\Phi^{-1}(0.75)-\Phi^{-1}(0.5)}} , \tag{2.8}$$

where F^{-1} and Φ^{-1} denote the inverse functions of F and of the standard normal d.f. Φ, respectively. For symmetric distributions we have $\tau = \big(F^{-1}(0.99)/F^{-1}(0.75)\big)\big/\big(\Phi^{-1}(0.99)/\Phi^{-1}(0.75)\big)$, the tail-weight coefficient defined in Hoaglin *et al.* (1983).

Several Monte Carlo simulation studies have been carried out to evaluate the efficiency and the robustness of different location and scale estimators, including the total median and the total range statistics. Some of these studies have been presented in Figueiredo (2003a, 2003b) and in Figueiredo and Gomes (2004, 2006), for a reasonably large set of symmetric and asymmetric distributions, with different skewness and tail-weight. It was then possible to conclude that the *TMd* statistic can be used to estimate the median value of a distribution F, as well as the mean value of a symmetric or approximately symmetric distribution. The *TR* statistic can be used to estimate the process standard deviation, in the case of rational subgroups of small to moderate size. However, both R and *TR* are biased estimators of the standard deviation. In order to get unbiased estimates, whenever the underlying model F is normal, it is necessary to consider, as usual, standardized versions of these statistics, obtained by the division of $R = X_{n:n} - X_{1:n}$ and $TR = X^*_{n:n} - X^*_{1:n}$ by appropriate scale constants. These constants are equal to the expected values of the statistics for the standard normal distribution (here denoted by $d_2 \equiv d_{2,R}$ and $d_{2,TR}$, respectively). For the most common values of n, they are given in Table 2, together with the statistics standard deviations (here denoted by $d_3 \equiv d_{3,R}$, $d_{3,TR}$ and $d_{3,TMd}$).

Table 2: Expected value, $d_{2,\bullet}$, and standard deviation, $d_{3,\bullet}$, of R, *TR* and *TMd* for a standard normal distribution ($d_{2,TMd} = 0$).

Constants	3	4	5	6	7	8	9	10
d_2	1.693	2.059	2.326	2.534	2.704	2.847	2.970	3.078
$d_{2,TR}$	1.269	1.538	1.801	2.027	2.210	2.364	2.491	2.610
d_3	0.888	0.880	0.864	0.848	0.833	0.820	0.808	0.797
$d_{3,TR}$	0.666	0.653	0.657	0.659	0.656	0.650	0.641	0.636
$d_{3,TMd}$	0.583	0.507	0.464	0.425	0.401	0.375	0.359	0.340

2.1.1. The class of models under consideration

To analyze the robustness of the above mentioned statistics to slight deviations of the normal model, and following the methodology presented in Figueiredo (2003b) and Figueiredo and Gomes (2004), we have considered several symmetric distributions, related with the standard normal distribution, and with different tail-weights τ, the indicator defined in (2.8). More precisely, we have considered standardized data from the following set $\mathcal{D}$ of symmetric distributions:

1. the standard normal, $\mathrm{N}(0,1)$;

2. the standard Laplace, $\mathrm{Laplace}(0,1)$;

3. the contaminated normal distributions, $\mathrm{CN}(\alpha \times 100\%)$, in which each observation has a $(1-\alpha)\times 100\%$ probability of being drawn from the $\mathrm{N}(0,1)$ and $\alpha \times 100\%$ probability of being drawn from the $\mathrm{N}(0,k)$, with a standard deviation $k=3$ and percentages of contamination $\alpha = 0.01$, 0.025, 0.05, 0.075, 0.10, 0.125 and 0.15.

The d.f. of the standard Laplace model is given by

$$F(x) = \begin{cases} e^x/2\,, & x \le 0 \\ 1 - e^{-x}/2\,, & x > 0 \end{cases}$$

and the d.f. of the contaminated normal model $\mathrm{CN}(\alpha \times 100\%)$, is given by

$$F(x) = \alpha\, \Phi(x/k) + (1-\alpha)\, \Phi(x)\ ,$$

where Φ denotes the d.f. of the standard normal distribution, given by $\Phi(x) = \int_{-\infty}^{x} \exp(-t^2/2)\, dt/\sqrt{2\pi},\ x \in \mathbb{R}$.

Remark 2.3. Note that even in potential normal situations there is some possibility of having disturbances in the data, and one of the previous distributions in $\mathcal{D}$, for instance, can describe the process data in a more reliable way.

2.1.2. The methodology

- To compare the efficiency of the different location estimators, we have used their mean square error. Since this measure is affected by the scaling of the estimator, we have used the variance of the logarithm of the estimator in the comparison of the scale estimators. Details about performance measures of scale estimators can be found in Lax (1985).
- To select the most robust estimator among the estimators under study, in the set $\mathcal{D}$ of models under consideration, we have applied a *Max/Min* criterion, following the steps below:

S_1 – for every distribution in $\mathcal{D}$, obtain the most efficient estimator, among the ones considered;

S_2 – then, compute the efficiency of the other estimators relatively to the best one, previously selected in step S_1;

S_3 – next, for each estimator, save the obtained minimum relative efficiency along all the considered distributions in $\mathcal{D}$, the so-called "degree of robustness" of the estimator;

S_4 – finally, the most robust estimator is the one with the highest "degree of robustness".

2.1.3. Results

Apart from the sample mean $M \equiv \overline{X}$ and the total median TMd, both location estimators, we have considered another location estimator, the sample median Md. Apart from the range R and the total range TR, both scale estimators, we have also considered the sample standard deviation estimator, S. In Figure 1 we present the most efficient estimator for the mean value (at the left) and for the standard deviation (at the right) of a distribution F in $\mathcal{D}$, for rational subgroups of size $n = 3$ up to 10.

		sample size							
$!_F$	*F*	3	4	5	6	7	8	9	10
1,717	CN(15%)	Md	Md	TMd	TMd	TMd	TMd	TMd	M
1,642	CN(12,5%)	TMd	Md	TMd	TMd	TMd	TMd	TMd	M
1,636	Laplace	TMd	Md	TMd	TMd	TMd	TMd	TMd	TMd
1,532	CN(10%)	TMd	Md	TMd	TMd	TMd	TMd	TMd	M
1,376	CN(7,5%)	TMd	Md	TMd	TMd	TMd	TMd	TMd	M
1,205	CN(5%)	TMd	TMd	TMd	TMd	TMd	TMd	TMd	M
1,08	CN(2,5%)	TMd	TMd	TMd	TMd	TMd	TMd	M	M
1,028	CN(1%)	M	TMd	M	M	M	M	M	M
1	N(0,1)	M	M	M	M	M	M	M	M

		sample size							
$!_F$	*F*	3	4	5	6	7	8	9	10
1,717	CN(15%)	TR	TR	TR	TR	TR	TR	S	S
1,642	CN(12,5%)	TR	TR	TR	TR	TR	TR	S	S
1,636	Laplace	TR	TR	TR	TR	TR	S	S	S
1,532	CN(10%)	TR	TR	TR	TR	TR	TR	S	S
1,376	CN(7,5%)	TR	TR	TR	TR	TR	TR	S	S
1,205	CN(5%)	TR	TR	TR	TR	TR	TR	S	S
1,08	CN(2,5%)	S	TR	TR	TR	TR	TR	S	S
1,028	CN(1%)	S	S	S	S	S	S	S	S
1	N(0,1)	S	S	S	S	S	S	S	S

Figure 1: Most efficient estimator for the mean value (*left*) and for the standard deviation (*right*).

The TMd and the TR estimators are the most efficient to estimate the mean value and the standard deviation, respectively, of a moderate-to-heavy-tailed distribution, whenever we consider rational subgroups of moderate size. We advise the use of the M and the S estimators only for distributions with small tail-weight and moderate-to-large sample sizes. In the extreme case of small samples and too heavy-tailed distributions, the sample median Md and the total range TR turn out to be the most efficient location and scale estimators, respectively. For $n = 3$, the Md estimator is worse than the TMd-estimator for high degrees of contamination of a normal model, due to the fact that there is only one central observation, instead of the two central observations when $n = 4$.

This is the reason for the discontinuity point in the graph of Figure 1 (*left*). The R-estimator is not at all competitive, despite of the fact that, in SPC, the range control chart based on the R-statistic is much more popular to monitor the standard deviation than the standard deviation control chart, based on the S-statistic.

In Figure 2 we picture the degree of robustness of the above-mentioned estimators.

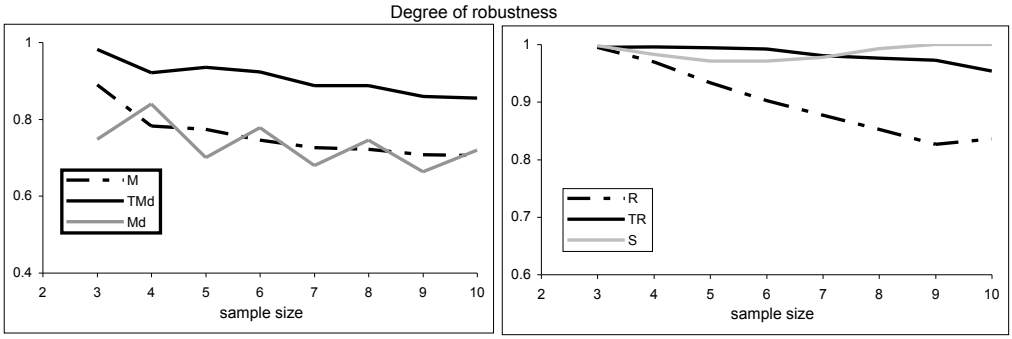

Figure 2: Degree of robustness of the location (*left*) and scale (*right*) estimators under study.

From Figure 2 (*left*), we can observe that the TMd-estimator is much more robust to changes in the underlying distribution F than the sample mean and the sample median estimators, M and Md, respectively. Indeed, the degree of robustness of TMd is always higher than the ones of either M or Md. The TR-estimator and the S-estimator present a similar degree of robustness, whenever we consider any d.f. F in $\mathcal{D}$, and are more robust than the R-estimator.

2.2. The sampling distribution

In order to get information about the sampling distribution of the previous location and scale statistics $M \equiv \overline{X}$, TMd, R and TR, here generically denoted by W, we have generated 50,000 values of each of the statistics W, for rational subgroups of size $n = 5$ and $n = 10$ from d.f.'s in $\mathcal{D}$. We have simulated their sampling distribution, and we have estimated the tail-weight, τ, defined in (2.8), as well as the asymmetry, through the use of the Fisher and of the Bowley skewness coefficients, γ and $\gamma_{\rm B}$, defined in (2.6) and (2.7), respectively.

The obtained estimates of τ, γ and $\gamma_{\rm B}$, and of the quantiles $\chi_p \equiv F^{-1}(p)$, $p = 0.1\%, 1\%, 25\%, 50\%, 75\%, 99\%$ and 99.9%, of the sampling distribution of the different statistics under study, are presented in Tables 3–6.

Table 3: Estimates of the mean values of τ, γ, γ_{B}, $\chi_{0.1\%}$, $\chi_{1\%}$, $\chi_{25\%}$, $\chi_{50\%}$, $\chi_{75\%}$, $\chi_{99\%}$, $\chi_{99.9\%}$, for the statistic M and subgroups of size $n = 5, 10$.

F	n	τ	γ	γ_{B}	$\chi_{0.1\%}$	$\chi_{1\%}$	$\chi_{25\%}$	$\chi_{50\%}$	$\chi_{75\%}$	$\chi_{99\%}$	$\chi_{99.9\%}$
N(0,1)	5	1.00	0.01	0.00	−1.398	−1.039	−0.302	−0.001	0.297	1.046	1.410
	10	0.99	−0.01	0.00	−0.977	−0.733	−0.217	−0.001	0.212	0.734	0.962
CN(1%)	5	1.03	0.02	−0.01	−1.582	−1.090	−0.312	0.000	0.305	1.099	1.576
	10	1.00	−0.02	0.00	−0.968	−0.746	−0.212	0.002	0.216	0.731	0.975
CN(2.5%)	5	1.08	0.02	−0.00	−1.772	−1.175	−0.322	−0.003	0.312	1.192	1.801
	10	0.99	−0.01	0.01	−0.995	−0.736	−0.213	−0.001	0.214	0.730	0.956
CN(5%)	5	1.14	−0.02	0.01	−2.025	−1.316	−0.331	0.001	0.336	1.309	1.942
	10	0.99	−0.01	0.01	−0.999	−0.737	−0.213	0.000	0.215	0.731	0.975
CN(7.5%)	5	1.19	−0.00	−0.01	−2.114	−1.447	−0.352	0.000	0.347	1.427	2.169
	10	0.99	−0.01	0.00	−0.982	−0.739	−0.216	−0.002	0.214	0.737	0.964
CN(10%)	5	1.21	0.00	−0.01	−2.370	−1.523	−0.373	−0.007	0.355	1.526	2.235
	10	0.99	0.00	0.00	−0.957	−0.736	−0.214	0.000	0.214	0.733	0.964
Laplace(0,1)	5	1.13	−0.01	0.01	−2.241	−1.547	−0.391	0.001	0.403	1.538	2.205
	10	1.06	0.01	0.00	−1.455	−1.062	−0.293	0.001	0.293	1.074	1.448
CN(12.5%)	5	1.21	0.03	0.01	−2.280	−1.598	−0.377	0.002	0.392	1.617	2.449
	10	1.00	0.01	0.00	−0.969	−0.741	−0.216	−0.003	0.211	0.734	0.992
CN(15%)	5	1.21	−0.00	0.01	−2.426	−1.703	−0.407	−0.004	0.403	1.688	2.484
	10	1.00	0.00	0.00	−0.971	−0.739	−0.215	−0.002	0.213	0.729	0.977

From the values in Table 3 we observe that the sampling distribution of the M-statistic is approximately symmetric for the models under study. When we consider underlying models F with small tail-weight, such as the normal and the CN(1%) models, the sampling distribution of M presents the same tail-weight as the normal distribution; for distributions F with moderate-to-heavy tails, such as the CN(10%), the CN(12.5%), the CN(15%) and the Laplace$(0, 1)$, the sampling distribution of M has tails heavier than the normal tail, but not so heavy as the tails of the underlying distribution. Moreover, this tail-weight decreases as the sample size n increases. Note that although the Laplace and the CN(10%) distributions have similar tail-weight, the tail-weight of the sampling distribution of M is similar when we consider the Laplace and the CN(5%) distribution instead of the CN(10%). For non-normal models, the obtained lower quantiles of the sampling distribution of M, $\chi_{0.1\%}$ and $\chi_{1\%}$, are smaller than the corresponding quantiles obtained in the normal case, and the upper quantiles, $\chi_{99\%}$ and $\chi_{99.9\%}$, are larger than the corresponding normal quantiles. This reveals the weak robustness of the M statistic to changes in the underlying model, mainly for small values of n. Finally, the interval of variation of the sampling distribution of the M statistic for non-normal data is larger than in the normal case, but the inter-quartile range is almost the same for all the distributions, except in the case

of distributions with very heavy tails. The main differences between the several sampling distributions are in the tails, and this is a very important feature when we are interested in the estimation of high quantiles, as usually happens in SPC.

From Table 4, we notice that the sampling distribution of the *TMd*-statistic is approximately symmetric for all the models under study, even when we consider heavy-tailed underlying models F, such as the CN(15%), for instance. The chance of having an extreme value from the *TMd* sampling distribution is smaller than the chance of obtaining it from the M sampling distribution. For large rational subgroups size of contaminated normal data, say $n = 10$, the lower and the upper quantiles of the *TMd* distribution are similar to the corresponding normal quantiles, but for $n = 5$ there are significant differences. Consequently, we do not advise the use of the *TMd* statistic in SPC for very small values of n, when there is some possibility of having contaminated normal data. For the contaminated normal models the interval of variation of the sampling distribution of the *TMd* statistic is larger than in the normal case, as well as the interquartile range for small values of n, but even so, the differences to the normal case are smaller when we consider the *TMd* instead of the M statistic. Note also that the sampling distribution of *TMd* presents the highest tail-weight for the Laplace distribution.

Table 4: Estimates of the mean values of τ, γ, γ_{B}, $\chi_{0.1\%}$, $\chi_{1\%}$, $\chi_{25\%}$, $\chi_{50\%}$, $\chi_{75\%}$, $\chi_{99\%}$, $\chi_{99.9\%}$, for the statistic *TMd* and subgroups of size $n = 5, 10$.

F	n	τ	γ	γ_{B}	$\chi_{0.1\%}$	$\chi_{1\%}$	$\chi_{25\%}$	$\chi_{50\%}$	$\chi_{75\%}$	$\chi_{99\%}$	$\chi_{99.9\%}$
N(0,1)	5	1.01	0.01	0.00	−1.445	−1.080	−0.312	0.000	0.310	1.088	1.457
	10	1.00	0.00	0.00	−1.044	−0.797	−0.231	0.000	0.229	0.791	1.055
CN(1%)	5	1.01	0.01	−0.01	−1.489	−1.100	−0.321	0.000	0.313	1.105	1.498
	10	1.01	−0.02	−0.01	−1.052	−0.805	−0.228	0.004	0.232	0.792	1.052
CN(2.5%)	5	1.02	0.02	0.00	−1.545	−1.126	−0.328	−0.005	0.316	1.140	1.538
	10	1.00	−0.01	−0.01	−1.043	−0.791	−0.230	0.001	0.230	0.790	1.055
CN(5%)	5	1.03	−0.01	0.01	−1.686	−1.183	−0.328	0.000	0.335	1.179	1.633
	10	0.99	0.00	0.01	−1.050	−0.790	−0.229	0.000	0.231	0.784	1.052
CN(7.5%)	5	1.06	−0.01	0.00	−1.765	−1.246	−0.344	−0.003	0.339	1.246	1.723
	10	0.99	−0.01	0.00	−1.052	−0.791	−0.232	−0.002	0.229	0.787	1.034
CN(10%)	5	1.07	−0.02	0.00	−1.879	−1.297	−0.357	−0.001	0.341	1.276	1.887
	10	1.00	−0.01	0.01	−1.052	−0.793	−0.229	0.000	0.232	0.795	1.030
Laplace(0,1)	5	1.19	−0.02	0.00	−2.080	−1.429	−0.342	0.004	0.352	1.418	2.054
	10	1.14	0.02	0.00	−1.275	−0.923	−0.236	0.001	0.236	0.932	1.344
CN(12.5%)	5	1.08	0.03	0.01	−1.966	−1.330	−0.354	0.004	0.368	1.363	2.041
	10	1.00	0.01	0.00	−1.066	−0.802	−0.232	−0.004	0.227	0.787	1.079
CN(15%)	5	1.09	0.00	0.00	−2.132	−1.413	−0.378	−0.003	0.373	1.421	2.138
	10	0.99	0.00	0.01	−1.039	−0.786	−0.230	−0.003	0.229	0.776	1.078

From Tables 5–6, we notice that the sampling distributions of the R and of the TR statistics are highly positively skewed, even in the normal case. For contaminated normal models, even with a moderate percentage of contamination, the sampling distributions of R and TR are heavy-tailed, with high positive skewness, and present some asymmetry even in the central part of the distribution, as it is indicated by the obtained value of the quartile skewness coefficient, γ_{B}. However, the distribution of the TR statistic is less asymmetric than the distribution of the R statistic, with a not so long right tail. In all the cases the skewness as well as the tail-weight decrease with the increase of the sample size n, and we thus advise the use of the TR statistic for large rational subgroups size. The tail-weight of the sampling distribution of the statistics R and TR is approximately equal to the tail-weight of the normal distribution when we consider the Laplace model, and its asymmetry is much smaller than the asymmetry of the sampling distribution of R and TR for the contaminated normal models here considered.

The histograms obtained, not pictured, confirm the symmetry of the sampling distributions of M and TMd, and the visible asymmetry of the distributions of R and TR, mainly for small samples. The increase of n leads us, in some cases, to a quasi-symmetric distribution.

Table 5: Estimates of the mean values of τ, γ, γ_{B}, $\chi_{0.1\%}$, $\chi_{1\%}$, $\chi_{25\%}$, $\chi_{50\%}$, $\chi_{75\%}$, $\chi_{99\%}$, $\chi_{99.9\%}$, for the statistic R and subgroups of size $n = 5, 10$.

F	n	τ	γ	γ_{B}	$\chi_{0.1\%}$	$\chi_{1\%}$	$\chi_{25\%}$	$\chi_{50\%}$	$\chi_{75\%}$	$\chi_{99\%}$	$\chi_{99.9\%}$
N(0,1)	5	0.97	0.48	0.06	0.351	0.663	1.699	2.252	2.875	4.628	5.542
	10	1.00	0.39	0.04	1.057	1.459	2.513	3.028	3.582	5.151	5.985
CN(1%)	5	1.08	1.19	0.06	0.351	0.664	1.719	2.289	2.927	5.204	8.039
	10	0.98	0.38	0.04	1.080	1.473	2.514	3.028	3.591	5.133	5.922
CN(2.5%)	5	1.25	1.59	0.08	0.355	0.670	1.749	2.332	3.015	6.287	9.222
	10	0.99	0.40	0.05	1.076	1.479	2.516	3.026	3.584	5.160	5.930
CN(5%)	5	1.38	1.82	0.10	0.387	0.707	1.793	2.413	3.164	7.477	10.383
	10	0.99	0.39	0.05	1.059	1.470	2.515	3.026	3.589	5.156	5.974
CN(7.5%)	5	1.35	1.78	0.13	0.399	0.708	1.840	2.499	3.352	8.115	11.121
	10	0.99	0.39	0.04	1.072	1.464	2.514	3.026	3.585	5.139	6.020
CN(10%)	5	1.30	1.68	0.16	0.400	0.723	1.901	2.585	3.536	8.516	11.470
	10	0.99	0.38	0.05	1.075	1.464	2.515	3.027	3.589	5.144	5.973
Laplace(0,1)	5	1.03	1.11	0.12	0.379	0.677	2.011	2.902	4.034	8.151	10.823
	10	1.06	0.95	0.10	1.068	1.602	3.253	4.227	5.409	9.692	12.152
CN(12.5%)	5	1.23	1.61	0.18	0.419	0.744	1.953	2.697	3.770	9.013	11.929
	10	0.99	0.39	0.04	1.082	1.466	2.518	3.033	3.589	5.156	5.957
CN(15%)	5	1.17	1.50	0.20	0.423	0.755	2.016	2.809	4.007	9.385	12.268
	10	1.00	0.39	0.04	1.106	1.456	2.514	3.030	3.586	5.159	5.917

Table 6: Estimates of the mean values of τ, γ, γ_{B}, $\chi_{0.1\%}$, $\chi_{1\%}$, $\chi_{25\%}$, $\chi_{50\%}$, $\chi_{75\%}$, $\chi_{99\%}$, $\chi_{99.9\%}$, for the statistic TR and subgroups of size $n = 5, 10$.

F	n	τ	γ	γ_{B}	$\chi_{0.1\%}$	$\chi_{1\%}$	$\chi_{25\%}$	$\chi_{50\%}$	$\chi_{75\%}$	$\chi_{99\%}$	$\chi_{99.9\%}$
N(0,1)	5	0.95	0.43	0.05	0.273	0.519	1.321	1.749	2.223	3.507	4.143
	10	0.99	0.28	0.02	0.924	1.277	2.164	2.583	3.023	4.219	4.830
CN(1%)	5	1.05	1.00	0.06	0.282	0.524	1.339	1.775	2.264	3.910	5.808
	10	0.98	0.27	0.02	0.946	1.283	1.163	2.585	3.025	4.199	4.753
CN(2.5%)	5	1.20	1.35	0.07	0.280	0.526	1.361	1.810	2.329	4.616	6.615
	10	0.99	0.29	0.03	0.943	1.286	2.165	2.581	3.024	4.221	4.767
CN(5%)	5	1.31	1.59	0.10	0.307	0.551	1.401	1.871	2.440	5.420	7.326
	10	0.98	0.27	0.03	0.939	1.276	2.167	2.584	3.029	4.209	4.801
CN(7.5%)	5	1.28	1.61	0.12	0.299	0.555	1.433	1.937	2.577	5.834	8.040
	10	0.98	0.28	0.04	0.944	1.280	2.166	2.580	3.026	4.203	4.830
CN(10%)	5	1.25	1.52	0.15	0.318	0.562	1.477	2.000	2.711	6.157	8.100
	10	0.98	0.27	0.03	0.937	1.280	2.163	2.583	3.029	4.198	4.859
Laplace(0,1)	5	1.01	1.01	0.11	0.290	0.528	1.555	2.221	3.060	5.949	7.812
	10	1.06	0.95	0.10	0.929	1.386	2.723	3.493	4.383	7.394	8.995
CN(12.5%)	5	1.20	1.48	0.16	0.333	0.582	1.517	2.086	2.875	6.554	8.518
	10	0.98	0.28	0.03	0.941	1.282	2.164	2.585	3.030	4.215	4.819
CN(15%)	5	1.14	1.39	0.19	0.330	0.592	1.569	2.165	3.046	6.786	8.821
	10	0.98	0.28	0.03	0.965	1.280	2.162	2.581	3.028	4.218	4.796

3. CONTROL CHARTS SIMULATED BEHAVIOUR

Whenever implementing a control chart, a practical advice is that 3-sigma control limits should be avoided whenever the distribution of the control statistic is very asymmetric. In such a case, it is preferable to fix the control limits of the chart at adequate probability quantiles of the control statistic distribution. However, the analytical determination of these quantiles is in general impossible to obtain, as well as its estimation, because we do not have sufficient observations for doing it accurately.

The results presented in Subsection 2.2 justify the use, in this study, of two-sided control charts with 3-sigma control limits to monitor the process mean value at a target μ_0. Thus, to detect increases or decreases in the process mean value μ, we have implemented the classical M-chart with control limits given in (1.1), and the TMd chart with lower and upper control limits given by

$$LCL_{TMd} = \mathbb{E}(TMd) - 3\sqrt{\mathbb{V}(TMd)}\ , \qquad UCL_{TMd} = \mathbb{E}(TMd) + 3\sqrt{\mathbb{V}(TMd)}\ .$$

For standard normal data the limits of the *TMd*-chart are given by

$$LCL_{TMd} = -3\, d_{3,TMd} \ , \qquad UCL_{TMd} = 3\, d_{3,TMd} \ , \tag{3.1}$$

where $d_{3,TMd}$ has been tabulated in Table 2. Here in order to obtain the same false alarm rate for the M and the *TMd* charts, whenever the underlying model F is normal, we have replaced in (3.1), $d_{3,TMd}$ by $d^*_{3,TMd} = 0.4643$ for $n = 5$ and $d^*_{3,TMd} = 0.3407$ for $n = 10$.

To monitor the process standard deviation at a target σ_0, and noting that the main interest is to detect increases in σ and not decreases in σ, we have implemented one-sided control charts, with lower control limits placed at 0. The R-chart has an upper control limit given in (1.2), and the TR chart has the upper control limit given by

$$UCL_{TR} = \mathbb{E}(TR) + 3\sqrt{\mathbb{V}(TR)} \ .$$

For standard normal data the upper control limits of the *TR*-chart is thus given by

$$UCL_{TR} = d_{2,TR} + 3\, d_{3,TR} \ , \tag{3.2}$$

where $d_{2,TR}$ and $d_{3,TR}$ have also been tabulated in Table 2. To obtain the same false alarm rate for the R and the TR charts, whenever the underlying model F is normal, we have considered a slightly different value for $d_{3,TR}$. More precisely, we have replaced in (3.2), $d_{3,TR}$ by $d^*_{3,TR} = 0.6465$ for $n = 5$ and $d^*_{3,TR} = 0.611$ for $n = 10$.

3.1. Robustness versus performance

The ability of a generic W control chart to detect process changes is usually measured by the expected number of samples taken before the chart signals, i.e., by its *ARL* (*Average Run Length*), or alternatively, in some cases, by its power function, together with the standard deviation of the Run Length distribution, *SDRL*.

When the successive values of the control statistic W are independent, and when we do not have to estimate the control limits of the chart, the RL variable (i.e., the number of samples taken before the chart signals) has a geometric distribution, and the *ARL* is given by

$$ARL_W(\theta) = \frac{1}{1 - P\big(W \in C \,|\, \theta\big)} =: \frac{1}{\pi_W(\theta)} \ , \tag{3.3}$$

where θ denotes the parameter to be controlled at $\theta = \theta_0$, with $\pi_W(\theta)$ the *power function* of the W-chart. The *SDRL* is given by

$$SDRL_W(\theta) = \frac{\sqrt{1 - \pi_W(\theta)}}{\pi_W(\theta)} \ . \tag{3.4}$$

Remark 3.1. Assuming that the process changes from the in-control state, $\theta = \theta_0$, to an out-of-control state, θ, a value in the space parameter, the power function of the chart is thus the probability of detection of that change in any arbitrary sample.

When the process is in-control, the power function gives us the *false alarm rate* of the chart, also called the α-*risk*, given by

$$\alpha \;=\; P\big(W \notin C \,|\, IN\big) \;=\; P\big(W \notin C \,|\, \theta = \theta_0\big) \;=\; \pi_{_W}(\theta_0) \;=\; 1/ARL(\theta_0)\;. \tag{3.5}$$

Remark 3.2. The control limits of a W-chart are usually determined in order to have a chart with a small fixed false alarm rate (a large in-control ARL) and we hope to obtain high power function values (small out-of-control ARL) for the shifts the chart must detect.

Remark 3.3. When we have to estimate process parameters in order to obtain the control limits of the chart or when the successive values of the control statistic W are not independent, the distribution of the random variable RL is not geometric, but a more right-skewed distribution. Some authors, see for instance Chakraborti (2006, 2007), refer that in this case the ARL and the $SDRL$ parameters in (3.3) and (3.4), respectively, are not the best measures of performance of a control chart. They also suggest the use of the Median Run Length, MRL, as a measure of performance, and the 5-th and the 95-th percentiles of the RL distribution to represent the spread of the RL. Additionally, for a more complete understanding of the chart performance, Chakraborti (2000, 2006, 2007) and Jensen *et al.* (2006) suggest the analysis of the RL distribution conditional on the observed estimates (i.e., the conditional RL distribution), together with the analysis of the marginal RL distribution. Such a marginal distribution is computed by integrating the conditional RL distribution over the range of the parameter estimators, and thus, it takes into account the random variability introduced into the charting procedure through parameter estimation, without requiring the knowledge of the observed estimates.

In the following study, to analyze the robustness to the normality assumption of any of the previous control charts, implemented with exact control limits, we have implemented the following algorithm:

- S_1^* – consider standardized data of the symmetric distributions in set $\mathcal{D}$ (see Subsection 2.2), as adequate to describe the data process;
- S_2^* – next, implement the charts with the control limits given in (1.1), (1.2), and the mentioned modifications of (3.1) and (3.2), for rational subgroups of sizes $n = 5$ and $n = 10$;
- S_3^* – compute the false alarm rates, α, defined in (3.5), through the use of Monte Carlo simulation techniques, using a sample of 500,000 values

of the control statistic for each of the 30 replicates of the simulation experiment (such a procedure allows us to present the α values with a precision of 4 decimal figures);

S_4^* – finally, compare them with the expected value α, obtained for normal data, and register the smallest α-risk.

The obtained simulated false alarm rates are presented in Tables 7–8 for rational subgroups of sizes $n = 5$ and 10. In each line we underline the α-value associated with the most robust chart, i.e., the one with smallest α-risk. From the obtained values we conclude that neither the *TMd*-chart nor the *TR*-chart can be considered robust to the normality assumption, but even so, they are more robust than the $\overline{X}$ and the R charts, respectively. We also conclude that we should preferably consider rational subgroups of size $n = 10$, instead of $n = 5$. Consequently, when there is a chance of having contaminated normal data, it is better to implement the *TMd* and the *TR* charts for rational subgroups of size $n = 10$. The α-values of the charts $\overline{X}$ and *TMd* (R and *TR*) for the Laplace distribution are similar to the α-values of these charts for the CN(1%) (CN(2.5%)) distribution, although the Laplace model has much heavier tails than these contaminated normal models.

Table 7: False Alarm rates of the $\overline{X}$ and *TMd* charts.

Model F	τ	$\overline{X}_{n=5}$	$TMd_{n=5}$	$\overline{X}_{n=10}$	$TMd_{n=10}$
N(0,1)	1.000	.00270	.00270	.00270	.00270
CN(1%)	1.028	.00540	<u>.00334</u>	.00478	<u>.00300</u>
Laplace(0,1)	1.636	.00618	<u>.00283</u>	.00474	<u>.00088</u>
CN(2.5%)	1.080	.00849	<u>.00443</u>	.00725	<u>.00346</u>
CN(5%)	1.205	.01198	<u>.00588</u>	.01018	<u>.00426</u>
CN(7.5%)	1.376	.01409	<u>.00723</u>	.01213	<u>.00506</u>
CN(10%)	1.532	.01543	<u>.00841</u>	.01347	<u>.00586</u>
CN(12.5%)	1.642	.01621	<u>.00938</u>	.01442	<u>.00662</u>
CN(15%)	1.717	.01668	<u>.01019</u>	.01505	<u>.00731</u>

Table 8: False Alarm rates of the R and *TR* charts.

Model F	τ	$R_{n=5}$	$TR_{n=5}$	$R_{n=10}$	$TR_{n=10}$
N(0,1)	1.000	.00453	.00453	.00423	.00423
CN(1%)	1.028	.01397	<u>.01323</u>	.02175	<u>.01994</u>
Laplace(0,1)	1.636	.02679	<u>.02236</u>	.04608	<u>.03589</u>
CN(2.5%)	1.080	.02538	<u>.02377</u>	.04268	<u>.03890</u>
CN(5%)	1.205	.03905	<u>.03646</u>	.06734	<u>.06185</u>
CN(7.5%)	1.376	.04788	<u>.04477</u>	.08294	<u>.07699</u>
CN(10%)	1.532	.05354	<u>.05022</u>	.09269	<u>.08702</u>
CN(12.5%)	1.642	.05684	<u>.05356</u>	.09820	<u>.09334</u>
CN(15%)	1.717	.05848	<u>.05536</u>	.10057	<u>.09690</u>

The analysis of the performance of the previous charts to detect changes in the process parameters is evaluated in terms of the simulated power function $\pi_W(\theta) = 1/ARL_W(\theta)$ for a few different magnitude changes θ, with $ARL_W(\theta)$ defined in (3.3).

The obtained simulated power function values are presented in Tables 9–10 for rational subgroups of size $n = 10$. For every magnitude of change and each model, we underline the highest obtained value, which corresponds to the most powerful chart. These results lead us to the following conclusions: the *TMd* (the *TR*) chart is more robust than the $\overline{X}$ (the *R*) chart, with smaller false alarm rates. Consequently, we were not at all expecting higher values for the power. However, the *TR*-chart overpass the *R*-chart in terms of performance to detect small to large changes for all the contaminated normal models, and even for normal data. Only for the Laplace distribution the *R*-chart overpass the *TR*-chart to detect very small changes. In this case the *TMd*-chart is able to overpass the $\overline{X}$-chart to detect large changes. Despite of the fact that the $\overline{X}$-chart presents higher power function values than the *TMd*-chart, we cannot conclude that the $\overline{X}$-chart is the most powerful chart, because the α-values of the *TMd*-chart are smaller than the α-values of the $\overline{X}$-chart. The conclusions are similar for $n = 5$, although we obtain larger power function values when the charts are implemented for rational subgroups of size $n = 10$. However the differences between the power function values of the $\overline{X}$ and of the *TMd* charts are smaller when we consider $n = 5$ instead of $n = 10$.

Table 9: Power function values of the charts ($n = 10$, $\mu \to \mu_1$, $\sigma = 1$).

μ_1	$\overline{X}$	*TMd*	$\overline{X}$	*TMd*	$\overline{X}$	*TMd*	$\overline{X}$	*TMd*
	N(μ_1,1)		Laplace(0,1)		CN(1%)		CN(2.5%)	
0.5	<u>.0780</u>	.0626	<u>.0758</u>	.0273	<u>.0841</u>	.0640	<u>.0917</u>	.0662
1	<u>.5646</u>	.4741	<u>.5670</u>	.4641	<u>.5630</u>	.4745	<u>.5612</u>	.4745
2	<u>.9996</u>	.9980	.9990	<u>.9994</u>	<u>.9990</u>	.9977	<u>.9983</u>	.9974
2.5	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>
3.0	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>
	CN(5%)		CN(7.5%)		CN(10%)		CN(15%)	
0.5	<u>.1010</u>	.0699	<u>.1076</u>	.0735	<u>.1122</u>	.0764	<u>.1180</u>	.0816
1	<u>.5591</u>	.4747	<u>.5573</u>	.4753	<u>.5561</u>	.4759	<u>.5547</u>	.4761
2	<u>.9975</u>	.9969	<u>.9970</u>	.9964	<u>.9966</u>	.9959	<u>.9962</u>	.9950
2.5	.9999	<u>1.000</u>	.9999	<u>1.000</u>	.9999	<u>1.000</u>	<u>.9999</u>	<u>.9999</u>
3.0	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>

Table 10: Power function values of the charts ($n = 10$, $\mu = 0$, $\sigma \to \sigma_1$).

σ_1	R	TR	R	TR	R	TR	R	TR
	N(0,σ_1)		Laplace(0,1)		CN(1%)		CN(2.5%)	
1.25	.0616	.0758	.1481	.1379	.0904	.1044	.1269	.1409
1.5	.2277	.2800	.2949	.2962	.2589	.3100	.2990	.3489
2	.6458	.7178	.5899	.6126	.6636	.7323	.6871	.7516
2.5	.8730	.9114	.7864	.8117	.8799	.9162	.8891	.9227
3	.9564	.9721	.8928	.9116	.9588	.9738	.9621	.9759
	CN(5%)		CN(7.5%)		CN(10%)		CN(15%)	
1.25	.1734	.1882	.2074	.2235	.2320	.2500	.2626	.2847
1.5	.3519	.4007	.3924	.4409	.4235	.4721	.4659	.5159
2	.7185	.7776	.7433	.7983	.7629	.8147	.7912	.8385
2.5	.9016	.9317	.9114	.9388	.9194	.9445	.9309	.9529
3	.9666	.9788	.9702	.9812	.9731	.9830	.9773	.9858

ACKNOWLEDGMENTS

Research partially supported by FCT/POCTI and POCI/FEDER.

REFERENCES

[1] Amin, R.W. and Lee, S.J. (1999). The effects of autocorrelation and outliers on two-sided tolerance limits, *J. Quality Technology*, **31**(3), 286–300.

[2] Bai, D.S. and Choi, I.S. (1995). $\overline{X}$ and R control charts for skewed populations, *J. of Quality Technology*, **27**(2), 120–131.

[3] Castagliola, P. (2000). $\overline{X}$ control chart for skewed populations using a scaled weighted variance method, *International J. of Reliability, Quality and Safety Engineering*, **7**(3), 237–252.

[4] Chakraborti, S. (2000). Run length, average run length and false alarm rate of Shewhart $\overline{X}$ chart: exact derivations by conditioning, *Communications in Statistics – Simulation and Computation*, **29**1, 61–81.

[5] Chakraborti, S. (2006). Parameter estimation and design considerations in prospective applications of the $\overline{X}$ Chart., *Journal of Applied Statistics*, **33**4, 439–459.

[6] Chakraborti, S. (2007). Run length distribution and percentiles: the Shewhart chart with unknown parameters, *Quality Engineering*, **19**, 119–127.

[7] Champ, W.C. and Jones, A.L. (2004). Design phase I $\overline{X}$ charts with small sample sizes, *Quality and Reliability Engineering International*, **20**, 497–510.

[8] Cox, M.G. and Iguzquiza, E.P. (2001). The total median and its uncertainity. In *Advanced Mathematical and Computational Tools in Metrology*, **V**, (Ciarlini *et al.*, Eds.), 106–117.

[9] Figueiredo, F. (2003a). *Controlo Estatístico da Qualidade e Métodos Robustos*, Ph.D. Thesis, Faculty of Science, Lisbon University.

[10] Figueiredo, F. (2003b). Robust estimators for the standard deviation based on a bootstrap sample, *Proc. 13th European Young Statisticians Meeting*, 53–62.

[11] Figueiredo, F. and Gomes, M.I. (2004). The total median is Statistical Quality Control, *Applied Stochastic Models in Business and Industry*, **20**, 339–353.

[12] Figueiredo, F. and Gomes, M.I. (2006). Box-Cox Transformations and Robust Control Charts in *SPC*, In *Advanced Mathematical and Computational Tools in Metrology*, **VII** (Pavese *et al.*, Eds.), 35–46.

[13] Hampel, F.R. (1971). A general qualitative definition of robustness, *Annals of Mathematical Statistics*, **42**, 1887–1896.

[14] Hampel, F.R.; Ronchetti, E.M.; Rousseew, P.J. and Stahel, W. (1986). *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York.

[15] Hoaglin, D.M.; Mosteller, F. and Tukey, J.W. (1983). *Understanding Robust and Exploratory Data Analysis*, Wiley, New York.

[16] Jensen, W.A.; Jones-Farmer, L.A.; Champ, C.H. and Woodall, W.H. (2006). Effects of parameter estimation on control chart properties: a literature review, *J. Quality Technology*, **38**, 349–364.

[17] Lax, D.A. (1985). Robust estimators of scale: finite sample performance in long-tailed symmetric distributions, *J. Amer. Statist. Assoc.*, **80**, 736–741.

[18] Montgomery, D.C. (2005). *Introduction to Statistical Quality Control*, Wiley, New York.

[19] Nedumaran, G. and Pignatiello, J.J. (2001). On estimating $\overline{X}$ control limits, *J. Quality Technology*, **33**(2), 206–212.

[20] Quesenberry, D.C. (1993). The effect of sample size on estimated limits for $\overline{X}$ and X control charts, *J. Quality Technology*, **25**(4), 237–247.

[21] Rocke, D.M. (1989). Robust control charts, *Technometrics*, **31**(2), 173–184.

[22] Rocke, D.M. (1992). $\overline{X}_Q$ and R_Q charts: robust control charts, *The Statistician*, **41**, 97–104.

[23] Ryan, T.P. (2000). *Statistical Methods for Quality Improvement*, Wiley, New York.

[24] Schilling, E.G. and Nelson, P.R. (1976). The effect of non-normality on the control limits of the $\overline{X}$ charts, *J. Quality Technology*, **8**(4), 183–188.

REVSTAT – Statistical Journal
Volume 7, Number 2, June 2009, 171–187

A BAYESIAN SHRINKAGE APPROACH IN WEIBULL TYPE-II CENSORED DATA USING PRIOR POINT INFORMATION

Authors: Gyan Prakash
– Department of Statistics, Harish Chandra P. G. College, Varanasi, U.P., India
ggyanji@yahoo.com

D.C. Singh
– Department of Statistics, Harish Chandra P. G. College, Varanasi, U.P., India
dc_singh52@rediffmail.com

Received: June 2008 Revised: March 2009 Accepted: May 2009

Abstract:

- In the present paper we study the performance of the Bayes Shrinkage estimators for the scale parameter of the Weibull distribution under the squared error loss and the LINEX loss functions in the presence of a prior point information of the scale parameter when Type-II censored data are available. The properties of the minimax estimators are also discussed.

Key-Words:

- *Bayes shrinkage estimator; Uniformly Minimum Variance Unbiased (UMVU) estimator; Minimax estimator and LINEX loss function.*

AMS Subject Classification:

- 62A15, 62C10, 62C20.

1. INTRODUCTION

The 'time of failure' and 'average life' of a component, measured from some specified time until it fails, is represented by a continuous random variable. Extensively in recent years, one distribution that has been used as a model to deal with such problems for product life is the Weibull distribution. Its applications in life-testing problems and survival analysis have been widely advocated (Weibull, 1951; Berrettoni, 1964). It has been used as model with diverse types of items such as ball bearing (Lieblein & Zelen, 1956), vacuum tube (Kao, 1959) and electrical isolation (Nelson, 1972). Mittnik & Reachev (1993) found that the Weibull distribution might be an adequate statistical model for stock returns. Mann (1968) gave a variety of situations in which the distribution is used for other types of failure data. Whittemore & Altschuler (1976) used it as a model in biomedical applications.

The probability density function of the Weibull distribution is given by

$$f(x; v, \theta) = \frac{v}{\theta}\, x^{v-1} e^{-\frac{x^v}{\theta}}\ , \qquad x > 0, \quad v > 0, \quad \theta > 0\ , \tag{1.1}$$

where the parameters v and θ are referred to as the shape and scale parameters of the distribution, respectively.

For the special case $v = 1$, the Weibull distribution is the Exponential distribution. For $v = 2$, is the Rayleigh distribution. For shape parameter values in the range $3 \leq v \leq 4$, the shape of the Weibull distribution is close to that of the Normal distribution and for a large values of v, say $v \geq 10$, the shape of the Weibull distribution is close to that of a smallest extreme value distribution.

Pandey (1983), Pandey *et al.* (1989), Pandey & Singh (1993) considered the estimation of the Weibull shape parameter in censored data. The prediction problems in the Weibull distribution have been discussed by Engelhardt & Bain (1973), Nigm (1989), Dellaportas & Wright (1991) and others. Montanari *et al.* (1997), Singh & Shukla (2000), Hisada & Arizino (2002), Singh *et al.* (2002), Tsionas (2002) and others considered the Weibull distribution in different contexts.

In many situations, the experimenter has some prior information about the parameter in the form of a point guess value. To utilize this guess value, the shrinkage estimators have been discussed by a number of authors, for details see the article Casella & Lehmann (1998), Prakash & Singh (2006, 2008), Singh *et al.* (2007). The shrinkage estimator performs better than the usual estimator when a guess value is approximately the true value of the parameter and sample size is small. A shrinkage estimator (Thompson, 1968) for the parameter θ when a

prior point guess value θ_0 of θ is available, is defined as

$$S = k\,\hat{\theta} + (1-k)\,\theta_0\ , \qquad 0 \le k \le 1\ . \tag{1.2}$$

Here $\hat{\theta}$ is any usual estimator of the parameter θ. The shrinkage procedure has been applied in numerous problems, including mean survival time in epidemiological studies (Harries & Shakarki, 1979), forecasting of the money supply (Tso, 1990), estimating mortality rates (Marshall, 1991) and improved estimation in sample surveys (Wooff, 1985).

When positive and negative errors have different consequences, the use of squared error loss function (SELF) in Bayesian estimation may not be appropriate. To overcome this difficulty, Varian (1975) and Zellner (1986) proposed an asymmetric loss function known as the LINEX loss function (LLF). The invariant version of LLF for any parameter θ is given by

$$L(\Delta) = e^{a\Delta} - a\Delta - 1\ , \qquad a \neq 0 \quad \text{and} \quad \Delta = \frac{\hat{\theta}-\theta}{\theta}\ . \tag{1.3}$$

The sign and magnitude of 'a' represents the direction and degree of asymmetry respectively. The positive (negative) value of 'a' is used when overestimation is more (less) serious than underestimation. The loss function (1.3) is approximately square error and almost symmetric if $|a|$ is near to zero. A number of authors have discussed the estimation procedures under LLF criterion. A Few recent works under the Bayesian and/or the LLF criterions are Nigm *et al.* (2003), Bellhouse (2004), Xu & Shi (2004), Ahmadi *et al.* (2005), Prakash & Singh (2006), Son & Oh (2006), Singh *et al.* (2007), Ahmad *et al.* (2007), Prakash & Singh (2008), among others.

Let $x_1, x_2, ..., x_n$ be the life times of n items put to test under model (1.1). The maximum likelihood estimate of θ (when v is known) is given by

$$\hat{\theta} = \sum_{i=1}^{n} \frac{x_i^v}{n}\ . \tag{1.4}$$

Consider Type-II censored sampling, where the test terminates as soon as the r^{th} item fails ($r \le n$). Let $x_1, x_2, ..., x_r$ be the observed failure times for the first r components. Then the likelihood function for the r failure items is

$$L\big(x_1, x_2, ..., x_r \,|\, \theta\big) \;=\; \frac{v^r}{\theta^r} \prod_{i=1}^{r} x_i^{v-1}\, \exp\!\left\{-\frac{r\,T_r}{\theta}\right\}\ , \tag{1.5}$$

where $T_r = \frac{1}{r}\left\{\sum_{i=1}^{r} x_i^v + (n-r)\,x_{(r)}^v\right\}$ is a UMVU estimator (Sinha, 1986) of the parameter θ and $\frac{2\,r\,T_r}{\theta} \backsim \chi^2_{2r}$.

The risks under the SELF and the LLF for T_r are given as

$$R_{(S)}(T_r) = \frac{\theta^2}{r} \qquad \text{and} \qquad R_{(L)}(T_r) = e^{-a}\left(\frac{r}{r-a}\right)^r - 1\ ,$$

where suffix S and L respectively, denote the risk taken under the SELF and the LLF criterions.

If parameter v is known, the natural family of conjugate prior of θ is taken as the inverted Gamma distribution with probability density function

$$g_1(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, \theta^{-\alpha-1} e^{-\frac{\beta}{\theta}} , \qquad \alpha > 0, \quad \beta > 0 . \tag{1.6}$$

In the situation where the researchers have no prior information about the parameter θ, one may use the uniform, quasi or improper prior. A family of priors is given by

$$g_2(\theta) = \theta^{-d}\, e^{-\frac{c\,d}{\theta}} , \qquad d > 0, \quad c > 0 . \tag{1.7}$$

If $d = 0$, we get a diffuse prior and if $d = 1$, $c = 0$ a non-informative prior is obtained. For a set of values of d and c, that satisfies the equality $\Gamma(d-1) = (c\,d)^{d-1}$, makes $g_2(\theta)$ as a proper prior.

If both of the parameters θ and v are unknown in model (1.1), the joint prior distribution (Sinha, 1986) is considered as

$$g(\theta, v) = g_1(\theta).\, h(v), \quad h(v) = \frac{1}{\vartheta} , \qquad 0 < v < \vartheta, \quad \vartheta > 0 . \tag{1.8}$$

In the present paper, we suggest some Bayes shrinkage estimators for the scale parameter of the two-parameter Weibull distribution in presence of a prior point information when Type-II censored data is available under the SELF and the LLF. The properties have been studied in terms of the relative efficiencies when compared with the UMVU estimator. The properties of the minimax estimator are also discussed in the last section.

2. THE BAYES SHRINKAGE ESTIMATORS (KNOWN SHAPE PARAMETER)

The posterior density of the parameter θ under prior density $g_1(\theta)$, is

$$Z_1(\theta) = \frac{(r\,T_r + \beta)^{\alpha+r}}{\Gamma(\alpha + r)}\, e^{-\frac{(r\,T_r+\beta)}{\theta}}\, \theta^{-\alpha-r-1} , \tag{2.1}$$

which is again an inverted Gamma distribution with the parameters $(\alpha + r)$ and $(r\,T_r + \beta)$. The Bayes estimator of the parameter θ under the SELF is obtained as

$$\hat{\theta}_1 = E_p(\theta) = \varphi_1(r\,T_r + \beta) , \qquad \varphi_1 = (\alpha + r - 1)^{-1} . \tag{2.2}$$

Here, the suffix p indicates that the expectation is taken under posterior density. We choose the parameters of the prior distribution $g_1(\theta)$ such that $E(\hat{\theta}_1) = \theta_0$,

where θ_0 is the point guess value of θ. This gives $\beta = (\alpha - 1)\theta_0$. Substituting β in (2.2), we obtain the Bayes estimator for θ as

$$\bar{\theta}_1 = \lambda_1 T_r + (1 - \lambda_1)\theta_0 \ , \qquad \lambda_1 = r\,\varphi_1, \quad (\alpha + r) > 1 \ . \tag{2.3}$$

This is similar to the shrinkage estimator defined in (1.2). We termed $\bar{\theta}_1$ as the Bayes shrinkage estimator.

The Bayes estimate of θ under the LLF (1.3) is obtained by simplifying the equality

$$\begin{aligned} E_p\left(\frac{1}{\theta}\, e^{\frac{a\,\hat{\theta}_2}{\theta}}\right) = e^a\, E_p\left(\frac{1}{\theta}\right) \quad \Longrightarrow \\ \Longrightarrow \quad \hat{\theta}_2 = \varphi_2(r\,T_r + \beta), \quad \varphi_2 = \frac{1}{a}\left(1 - \exp\left(-\frac{a}{\alpha + r + 1}\right)\right). \end{aligned} \tag{2.4}$$

Again,

$$E(\hat{\theta}_2) = \theta_0 \quad \Longrightarrow \quad \beta = \theta_0\,(1 - r\,\varphi_2)\,\varphi_2^{-1} \ .$$

Hence, the Bayes shrinkage estimator for θ under the LLF with this choice of constant is given by

$$\bar{\theta}_2 = \lambda_2\, T_r + (1 - \lambda_2)\,\theta_0 \ , \qquad \lambda_2 = r\,\varphi_2 \ . \tag{2.5}$$

The expressions of the risks of these estimators under the SELF and the LLF are obtained as

$$R_{(S)}(\hat{\theta}_i) \ = \ r\,\theta^2\,\varphi_i^2 + \left(\theta\,(r\,\varphi_i - 1) + \beta\,\varphi_i\right)^2 \ , \tag{2.6}$$

$$R_{(L)}(\hat{\theta}_i) \ = \ \exp\left(a\left(\frac{\varphi_i\,\beta}{\theta} - 1\right)\right)\left(1 - a\,\varphi_i\right)^{-r} - 1 - a\left(r + \frac{\beta}{\theta} - 1\right), \tag{2.7}$$

$$R_{(S)}(\bar{\theta}_i) \ = \ \theta^2\left\{\lambda_i^2\left(\frac{r+1}{r} + \delta\,(\delta - 2)\right) + (1 - \delta)^2\,(1 - 2\,\lambda_i)\right\} \tag{2.8}$$

and

$$R_{(L)}(\bar{\theta}_i) \ = \ e^{a((1-\lambda_i)\,\delta - 1)}\,(1 - \frac{a\,\lambda_i}{r})^{-r} - 1 \ + \ a\,(1 - \delta)\,(1 - \lambda_i) \ , \tag{2.9}$$

where $\delta = \frac{\theta}{\theta_0},\ i = 1, 2.$

The posterior density of θ corresponding to $g_2(\theta)$ is given as

$$Z_2(\theta) \ = \ \frac{(r\,T_r + c\,d)^{r+d-1}}{\Gamma(r + d - 1)}\ e^{-\frac{(r\,T_r + c\,d)}{\theta}}\ \theta^{-r-d} \ . \tag{2.10}$$

This posterior distribution has the same form as the posterior (2.1). The only change is that in the place of α and β there are $d - 1$ and $c\,d$, respectively. All the results discussed in Section 3 hold if we substitute $d = (\alpha + 1)$ and $c = \frac{\beta}{(\alpha+1)}$.

3. NUMERICAL ANALYSIS

The relative efficiencies of the Bayes shrinkage estimator $\bar{\theta}_i$ $(i=1,2)$ relative to the UMVU estimator T_r under the SELF and the LLF criterions are defined as

$$RE_{(S)}(\bar{\theta}_i, T_r) = \frac{R_{(S)}(T_r)}{R_{(S)}(\bar{\theta}_i)}$$

and

$$RE_{(L)}(\bar{\theta}_i, T_r) = \frac{R_{(L)}(T_r)}{R_{(L)}(\bar{\theta}_i)}\ , \qquad i=1,2\ .$$

The expressions of relative efficiencies are the functions of r, a, δ and α whereas $RE_{(S)}(\bar{\theta}_i, T_r)$ is independent with 'a'. For the selected set of values $r = 04\,(02)10$; $a = 0.50(0.50)2.00$; $\delta = 0.25\,(0.25)1.75$ and $\alpha = 1.25, 1.50, 2.50, 5.00, 10, 20$, the relative efficiencies have been calculated and presented in Tables 1–4, respectively. The numerical findings are presented here only for $r = 04$ when risk criterion is the LLF.

Table 1: $RE_{(S)}(\bar{\theta}_1, T_r)$.

r	δ	α					
		1.25	1.50	2.50	5.00	10.00	20.00
04	0.25	1.1191	1.2226	1.4362	1.2308	0.8525	0.6387
	0.50	1.1245	1.2462	1.6575	2.0000	1.7423	1.4032
	0.75	1.1278	1.2607	1.8264	3.2000	4.6621	4.9788
	1.00	1.1289	1.2656	1.8906	4.0000	10.562	33.062
	1.25	1.1278	1.2607	1.8264	3.2000	4.6621	4.9788
	1.50	1.1245	1.2462	1.6575	2.0000	1.7423	1.4032
	1.75	1.1191	1.2226	1.4362	1.2308	0.8525	0.6387
06	0.25	1.0787	1.1467	1.2903	1.1111	0.7273	0.4983
	0.50	1.0823	1.1615	1.4286	1.6667	1.4286	1.0823
	0.75	1.0844	1.1706	1.5267	2.3810	3.3898	3.6470
	1.00	1.0851	1.1736	1.5625	2.7778	6.2500	17.361
	1.25	1.0844	1.1706	1.5267	2.3810	3.3898	3.6470
	1.50	1.0823	1.1615	1.4286	1.6667	1.4286	1.0823
	1.75	1.0787	1.1467	1.2903	1.1111	0.7273	0.4983
08	0.25	1.0588	1.1094	1.2175	1.0588	0.6744	0.4317
	0.50	1.0614	1.1202	1.3175	1.5000	1.2788	0.9275
	0.75	1.0630	1.1267	1.3858	2.0000	2.7656	2.9816
	1.00	1.0635	1.1289	1.4102	2.2500	4.5156	11.390
	1.25	1.0630	1.1267	1.3858	2.0000	2.7656	2.9816
	1.50	1.0614	1.1202	1.3175	1.5000	1.2788	0.9275
	1.75	1.0588	1.1094	1.2175	1.0588	0.6744	0.4317
10	0.25	1.0469	1.0872	1.1739	1.0316	0.6497	0.3947
	0.50	1.0490	1.0957	1.2521	1.4000	1.1934	0.8389
	0.75	1.0502	1.1008	1.3042	1.7818	2.3967	2.5827
	1.00	1.0506	1.1025	1.3225	1.9600	3.6100	8.4100
	1.25	1.0502	1.1008	1.3042	1.7818	2.3967	2.5827
	1.50	1.0490	1.0957	1.2521	1.4000	1.1934	0.8389
	1.75	1.0469	1.0872	1.1739	1.0316	0.6497	0.3947

Table 1 shows that the Bayes shrinkage estimator $\bar{\theta}_1$ performs uniformly well for small $\alpha \leq 5.00$ with respect to the UMVU estimator T_r under the SELF. The effective interval (the interval in which the relative efficiency is more than one) decreases with the sample size r as well as α increases under the SELF. The efficiency attains maximum at the point $\delta = 1.00$ and the gain in efficiency decreases as r increases for all considered values of δ when other parametric values are fixed. Further, the gains in efficiencies increase as α increases in the interval $0.75 \leq \delta \leq 1.25$ with other fixed parametric values.

On the other hand, when the risk criterion is the LLF (Table 2) the estimator $\bar{\theta}_1$ performs uniformly well with respect to T_r when sample size is small $r\,(\leq 06)$ for all considered values of parametric space but for a large sample size, this property holds in the interval $0.50 \leq \delta \leq 1.50$. The gain in efficiency increases when 'a' increases for all considered values of δ with small sample size $r\,(\leq 06)$ and in the interval $\delta \leq 1.25$ otherwise, under other fixed parametric values. Other properties are similar to the SELF-criterion.

Table 2: $RE_{(L)}(\bar{\theta}_1, T_r)$.

$r = 04$		α					
a	δ	1.25	1.50	2.50	5.00	10.00	20.00
0.50	0.25	1.1374	1.2835	1.6460	1.4888	1.0422	1.0173
	0.50	1.1448	1.2954	1.8404	2.3580	2.0777	1.6762
	0.75	1.1501	1.2968	1.9546	3.6144	5.4006	5.7685
	1.00	1.1534	1.3007	1.9569	4.2400	11.414	36.177
	1.25	1.1281	1.2613	1.8306	3.2413	4.8409	5.2728
	1.50	1.1169	1.2311	1.6228	1.9921	1.7805	1.4433
	1.75	1.1039	1.1738	1.3638	1.2135	1.1517	1.1352
1.00	0.25	1.1540	1.3182	1.9506	1.8884	1.3370	1.0178
	0.50	1.1702	1.3510	2.0770	2.9135	2.6054	2.1057
	0.75	1.1847	1.3757	2.1064	4.2662	6.5833	7.0413
	1.00	1.1975	1.3914	2.1565	4.6862	12.955	41.695
	1.25	1.1363	1.2782	1.8866	3.4259	5.2801	5.8795
	1.50	1.1172	1.2321	1.6349	2.0697	1.9090	1.5603
	1.75	1.1068	1.1814	1.3725	1.2466	1.1890	1.1618
1.50	0.25	1.1845	1.3822	2.3035	2.5672	1.8402	1.4108
	0.50	1.2102	1.4372	2.4297	3.8543	3.5120	2.8440
	0.75	1.2348	1.4861	2.4822	5.3708	8.6319	9.2565
	1.00	1.2582	1.5275	2.5244	5.5108	15.782	51.739
	1.25	1.1580	1.3225	2.0301	3.8550	6.1813	7.0553
	1.50	1.1307	1.2594	1.7207	2.2881	2.1938	1.8118
	1.75	1.1128	1.1942	1.4223	1.3606	1.1996	1.1738
2.00	0.25	1.2425	1.5052	2.7492	3.8948	2.8309	2.1846
	0.50	1.2790	1.5879	3.0769	5.6791	5.3001	4.3013
	0.75	1.3153	1.6677	3.2667	7.1635	12.675	13.647
	1.00	1.3512	1.7432	3.2835	7.4945	21.480	71.977
	1.25	1.2060	1.4211	2.3516	4.7975	8.0839	9.4870
	1.50	1.1695	1.3370	1.9500	2.7951	2.8120	2.3530
	1.75	1.1332	1.2540	1.5866	1.6381	1.2402	1.1918

The Bayes shrinkage estimator $\bar{\theta}_2$ performs well for all considered values of the parametric space when $\alpha \leq 10.00$ with respect to T_r under the SELF. The gain in efficiency increases when 'a' increases in the interval $0.75 \leq \delta \leq 1.25$ for all considered parametric values when $\alpha \leq 10.00$. Other properties of the estimator $\bar{\theta}_2$ are similar to the estimator $\bar{\theta}_1$ under the SELF.

Table 3: $RE_{(S)}(\bar{\theta}_2, T_r)$.

$r = 04$		α					
a	δ	1.25	1.50	2.50	5.00	10.00	20.00
0.50	0.25	1.4051	1.3797	1.2604	1.0111	1.0132	0.6178
	0.50	1.8994	1.9341	1.9980	1.9079	1.6312	1.3636
	0.75	2.4076	2.5485	3.0790	4.0782	4.8817	4.9465
	1.00	2.6433	2.8504	3.7566	6.5691	14.537	39.850
	1.25	2.4076	2.5485	3.0790	4.0782	4.8817	4.9465
	1.50	1.8994	1.9341	1.9980	1.9079	1.6312	1.3636
	1.75	1.4051	1.3797	1.2604	1.0111	1.0132	0.6178
1.00	0.25	1.3786	1.3502	1.2295	1.0921	1.0766	0.6157
	0.50	1.9353	1.9611	2.0000	1.8922	1.6203	1.3596
	0.75	2.5542	2.6919	3.2053	4.1529	4.8971	4.9425
	1.00	2.8589	3.0737	4.0111	6.9016	15.026	40.651
	1.25	2.5542	2.6919	3.2053	4.1529	4.8971	4.9425
	1.50	1.9353	1.9611	2.0000	1.8922	1.6203	1.3596
	1.75	1.3786	1.3502	1.2295	1.0921	1.0766	0.6157
1.50	0.25	1.3482	1.3183	1.1988	1.1738	1.0991	0.6137
	0.50	1.9626	1.9806	1.9978	1.8760	1.6096	1.3557
	0.75	2.7013	2.8351	3.3291	4.2240	4.9113	4.9383
	1.00	3.0888	3.3113	4.2797	7.2479	15.528	41.465
	1.25	2.7013	2.8351	3.3291	4.2240	4.9113	4.9383
	1.50	1.9626	1.9806	1.9978	1.8760	1.6096	1.3557
	1.75	1.3482	1.3183	1.1988	1.1738	1.0991	0.6137
2.00	0.25	1.3153	1.2850	1.1687	1.2563	1.1524	0.6117
	0.50	1.9820	1.9930	1.9919	1.8595	1.5991	1.3518
	0.75	2.8480	2.9772	3.4498	4.2914	4.9241	4.9341
	1.00	3.3336	3.5638	4.5629	7.6084	16.044	42.293
	1.25	2.8480	2.9772	3.4498	4.2914	4.9241	4.9341
	1.50	1.9820	1.9930	1.9919	1.8595	1.5991	1.3518
	1.75	1.3153	1.2850	1.1687	1.2563	1.1524	0.6117

Under the LLF criterion (Table 4), the estimator $\bar{\theta}_2$ also performs well for $\alpha \leq 10.00$ with respect to T_r and the gain in efficiency increases as 'a' increases for all considered values of parametric space. Other properties of $\bar{\theta}_2$ are similar to the Bayes shrinkage estimator $\bar{\theta}_1$ under the LLF criterion.

The gain in efficiency is larger for the Bayes shrinkage estimator $\bar{\theta}_2$ under the LLF-criterion with respect to the SELF-criterion.

Remark 3.1. One may obtain the results for the complete sample case by replacing only the censored sample size r with the complete sample size n.

Table 4: $RE_{(L)}(\bar{\theta}_2, T_r)$.

$r = 04$		α					
a	δ	1.25	1.50	2.50	5.00	10.00	20.00
0.50	0.25	1.6696	1.6475	1.5223	1.2316	1.0475	0.7625
	0.50	2.1863	2.2395	2.3499	2.2710	1.9461	1.6295
	0.75	2.6513	2.8215	3.4666	4.6905	5.6626	5.7285
	1.00	2.7676	2.9917	3.9753	7.0418	15.778	43.665
	1.25	2.4211	2.5656	3.1147	4.1820	5.1063	5.2472
	1.50	1.8671	1.9049	1.9857	1.9310	1.6731	1.4025
	1.75	1.3451	1.3242	1.2403	1.0074	1.0002	0.6137
1.00	0.25	2.0557	2.0308	1.8866	1.5459	1.2075	0.9837
	0.50	2.7072	2.7745	2.9142	2.8197	2.4254	2.0421
	0.75	3.2546	3.4681	4.2746	5.7837	6.9391	6.9833
	1.00	3.2684	3.5344	4.7001	8.3263	18.629	51.436
	1.25	2.6878	2.8427	3.4322	4.5902	5.6420	5.8586
	1.50	1.9502	1.9870	2.0702	2.0351	1.7908	1.5118
	1.75	1.3584	1.3390	1.2455	1.0308	1.0102	0.6363
1.50	0.25	2.7217	2.6918	2.5135	2.0847	1.6518	1.3611
	0.50	3.6007	3.6926	3.8833	3.7633	3.2503	2.7518
	0.75	4.2805	4.5695	5.6579	7.6678	9.1522	9.1678
	1.00	4.3134	4.6534	5.9421	10.564	23.661	65.268
	1.25	3.1686	3.3496	4.0386	5.4076	6.7074	7.0433
	1.50	2.1639	2.2055	2.3070	2.3004	2.0581	1.7507
	1.75	1.4374	1.4204	1.3358	1.1282	1.0425	0.7064
2.00	0.25	4.0363	3.9961	3.7493	3.1458	2.5266	2.1040
	0.50	5.3596	5.5001	5.7924	5.6237	4.8782	4.1525
	0.75	6.2841	6.7235	8.3740	11.384	13.536	13.598
	1.00	6.7483	6.8356	8.3882	14.978	33.669	92.906
	1.25	4.1355	4.3749	5.2875	7.1208	8.9341	9.4941
	1.50	2.6588	2.7148	2.8614	2.9034	2.6444	2.2679
	1.75	1.6843	1.6704	1.5919	1.3743	1.0963	0.8737

4. THE BAYES SHRINKAGE ESTIMATORS (UNKNOWN SHAPE PARAMETER)

When both parameters of the model (1.1) are unknown, the joint posterior density with respect to $g(\theta, v)$, in (1.8), is obtained as

$$Z_3(\theta, v) = \frac{v'\,\theta^{-(\alpha+r+1)}\,e^{-\frac{(r\,T_r+\beta)}{\theta}}}{\Gamma(\alpha+r)\int_0^{\vartheta} v'(r\,T_r+\beta)^{-\alpha-r}\,dv}\,, \qquad v' = v^r\prod_{i=1}^{r} x_i^{v-1}\,. \tag{4.1}$$

The marginal density of θ is obtained as

$$Z_4(\theta) = \frac{\theta^{-(\alpha+r+1)} \int_0^{\vartheta} v' e^{-\frac{(r\,T_r+\beta)}{\theta}}\, dv}{\Gamma(\alpha+r) \int_0^{\vartheta} v' (r\,T_r+\beta)^{-\alpha-r}\, dv} \,. \tag{4.2}$$

Hence, the Bayes estimate of the parameter θ under the SELF is obtained as

$$\hat{\theta}_3 = \frac{I\big(v',\,(\alpha+r-1)\big)}{(\alpha+r-1)\; I\big(v',\,(\alpha+r)\big)} \,, \tag{4.3}$$

where $I(\omega_1,\omega_2) = \int_0^{\vartheta}(\omega_1)\,(r\,T_r+\beta)^{-\omega_2}\, dv$.

Similarly, the Bayes estimate of the parameter θ under the LLF is

$$I\big(v'',\,(\alpha+r-1)\big) = e^a\, I\big(v'',\,(\alpha+r)\big) \,, \quad v'' = v'\left(1 - \frac{a\,\hat{\theta}_4}{r\,T_r+\beta}\right)^{-\alpha-r+1}. \tag{4.4}$$

The Bayes estimates of the parameter θ under the SELF and the LLF criterions do not exist in the close form. Therefore, the risks under both risks criterion also do not exist. Hence, the Bayes shrinkage estimator is not obtained in this case. However, the numerical findings of the Bayes estimates and their risks under both risks criterion are presented here by using the following example.

Example 4.1. Mann and Fertig (1973) give failure times of airplane components subjected to a life test. The Weibull distribution has often been found a suitable model in such situations. The data are Type-II censored: 13 components were placed on test and test was terminated at time of 10^{th} failure. Failure times (in hours) of the 10 components that failed were

$$0.22 \quad 0.50 \quad 0.88 \quad 1.00 \quad 1.32 \quad 1.33 \quad 1.54 \quad 1.76 \quad 2.50 \quad 3.00\,.$$

The expressions of the Bayes estimates of θ and their risks under the SELF and LLF risk criterion involve a, α, β, θ, n, ϑ and r. For the similar set of selected values of 'a' and α as considered earlier with $\beta = 0.50, 2.00, 5.00, 10.00, 20.00$; $\theta = 02, 04, 06$ and $\vartheta = 02, 04, 06, 10$, the numerical findings have been obtained and we present them in Table 5–6.

Table 5 presents the numerical values of the Bayes estimate $\hat{\theta}_3$ (SELF) only for $\theta = 2.00$ and their risks under the SELF and the LLF for $(\theta, \vartheta, a) = 2.00$ only. It is observed form the table that the magnitude of the risks under both risk criterion increase (decrease) as $\beta(\alpha)$ increases when other considered values of the parametric space are fixed. It also has been seen that the risks under the LLF-criterion increases when 'a' increase with other fixed parametric values. It is noted that there is a smaller magnitude of the risks under the LLF-criterion than under the SELF-criterion.

Table 5: The values of the Bayes estimate $\hat{\theta}_3$.

ϑ	$\beta\downarrow$ $\alpha\rightarrow$	1.25	1.50	2.50	5.00	10.00	20.00
2.00	0.50	3.0896	2.9711	2.5581	1.8471	1.1469	0.6381
	2.00	3.3315	3.2086	2.7774	2.0228	1.2637	0.7056
	5.00	3.7889	3.6581	3.1954	2.3646	1.4946	0.8398
	10.00	4.4901	4.3479	3.8411	2.9064	1.8722	1.0612
	20.00	5.7463	5.5829	4.9988	3.8992	2.6010	1.4982
4.00	0.50	3.2633	3.1129	2.6198	1.8538	1.1469	0.6380
	2.00	3.5547	3.3931	2.8622	2.0336	1.2637	0.7059
	5.00	4.1258	3.9425	3.3383	2.3879	1.4950	0.8400
	10.00	5.0506	4.8327	4.1115	2.9660	1.8744	1.0611
	20.00	6.8305	6.5474	5.6050	4.0890	2.6171	1.4983
6.00	0.50	3.2631	3.1126	2.6195	1.8548	1.1446	0.6430
	2.00	3.5546	3.3929	2.8618	2.0341	1.2629	0.7096
	5.00	4.1260	3.9426	3.3382	2.3873	1.4966	0.8390
	10.00	5.0507	4.8328	4.1118	2.9657	1.8749	1.0581
	20.00	6.8306	6.5474	5.6047	4.0894	2.6163	1.5010
10.00	0.50	3.2469	3.0962	2.6104	1.8782	1.1311	0.6470
	2.00	3.5399	3.3760	2.8443	2.0500	1.2632	0.7035
	5.00	4.1247	3.9370	3.3183	2.3802	1.5201	0.8204
	10.00	5.0715	4.8522	4.1168	2.9371	1.8978	1.0412
	20.00	6.8295	6.5524	5.6308	4.1048	2.5871	1.5249

The values of risks of the Bayes estimator							
$(a,\vartheta)=2.00$	$\beta\downarrow$ $\alpha\rightarrow$	1.25	1.50	2.50	5.00	10.00	20.00
$R_{(S)}(\hat{\theta}_3)$	0.50	64.744	63.185	57.786	48.327	38.039	29.313
$R_{(L)}(\hat{\theta}_3)$		2.9978	2.9517	2.7875	2.4814	2.1149	1.7671
$R_{(S)}(\hat{\theta}_3)$	2.00	67.093	65.451	59.765	49.816	39.013	29.874
$R_{(L)}(\hat{\theta}_3)$		3.0694	3.0215	2.8510	2.5331	2.1525	1.7912
$R_{(S)}(\hat{\theta}_3)$	5.00	71.919	70.104	63.826	52.862	41.000	31.011
$R_{(L)}(\hat{\theta}_3)$		3.2130	3.1615	2.9784	2.6369	2.2278	1.8396
$R_{(S)}(\hat{\theta}_3)$	10.00	80.342	78.222	70.896	58.142	44.421	32.956
$R_{(L)}(\hat{\theta}_3)$		3.4530	3.3957	3.1915	2.8106	2.3540	1.9205
$R_{(S)}(\hat{\theta}_3)$	20.00	98.617	95.818	86.170	69.469	51.678	37.022
$R_{(L)}(\hat{\theta}_3)$		3.9354	3.8663	3.6201	3.1603	2.6083	2.0837

Table 6 presents the numerical values of the Bayes estimate $\hat{\theta}_4$ (LLF) only for $\theta = 2.00$, $\vartheta = 2.00, 10.00$ and $a = 0.50, 2.00$ and their risks under the SELF and the LLF for $(\theta, \vartheta, a) = 2.00$ only. It is observed form the table that the magnitudes of the risks under both risk criterions increase when 'a' increases with other fixed parametric values. Other properties are similar to the Bayes estimator $\hat{\theta}_3$.

Table 6: The values of the Bayes estimate $\hat{\theta}_4$.

ϑ	a	$\beta\downarrow$ $\alpha\rightarrow$	1.25	1.50	2.50	5.00	10.00	20.00
2.00	0.50	0.50	2.1562	2.2002	2.3762	2.8162	3.6963	5.4564
		2.00	2.0055	2.0465	2.2102	2.6195	3.4381	5.0752
		5.00	1.7637	1.7997	1.9436	2.3036	3.0234	4.4632
		10.00	1.4743	1.5044	1.6248	1.9256	2.5274	3.7309
		20.00	1.1174	1.1402	1.2314	1.4595	1.9156	2.8278
	2.00	0.50	0.5390	0.5500	0.5940	0.7040	0.9241	1.3641
		2.00	0.5014	0.5116	0.5525	0.6549	0.8595	1.2688
		5.00	0.4409	0.4499	0.4859	0.5759	0.7559	1.1158
		10.00	0.3686	0.3761	0.4062	0.4814	0.6318	0.9327
		20.00	0.2794	0.2851	0.3079	0.3649	0.4789	0.7069
10.0	0.50	0.50	2.6221	2.6756	2.8896	3.4247	4.4950	6.6354
		2.00	2.4649	2.5152	2.7164	3.2195	4.2255	6.2377
		5.00	2.2106	2.2557	2.4362	2.8874	3.7897	5.5942
		10.00	1.9024	1.9412	2.0965	2.4848	3.2613	4.8143
		20.00	1.5133	1.5442	1.6677	1.9766	2.5942	3.8296
	2.00	0.50	0.6555	0.6689	0.7224	0.8562	1.1237	1.6589
		2.00	0.6162	0.6288	0.6791	0.8049	1.0564	1.5594
		5.00	0.5527	0.5639	0.6091	0.7218	0.9474	1.3986
		10.00	0.4756	0.4853	0.5241	0.6212	0.8153	1.2036
		20.00	0.3783	0.3860	0.4169	0.4941	0.6486	0.9574

The values of risks of the Bayes estimator							
$(a,\vartheta)=2.00$	$\beta\downarrow$ $\alpha\rightarrow$	1.25	1.50	2.50	5.00	10.00	20.00
$R_{(S)}(\hat{\theta}_4)$	0.50	13.445	13.395	13.198	12.713	11.774	10.022
$R_{(L)}(\hat{\theta}_4)$		0.9928	0.9899	0.9785	0.9501	0.8941	0.7853
$R_{(S)}(\hat{\theta}_4)$	2.00	13.550	13.502	13.313	12.846	11.939	10.241
$R_{(L)}(\hat{\theta}_4)$		0.9989	0.9961	0.9852	0.9580	0.9042	0.7995
$R_{(S)}(\hat{\theta}_4)$	5.00	13.734	13.690	13.514	13.079	12.232	10.634
$R_{(L)}(\hat{\theta}_4)$		1.0095	1.0070	0.9969	0.9717	0.9219	0.8246
$R_{(S)}(\hat{\theta}_4)$	10.00	13.983	13.944	13.786	13.395	12.632	11.179
$R_{(L)}(\hat{\theta}_4)$		1.0237	1.0215	1.0125	0.9901	0.9457	0.8586
$R_{(S)}(\hat{\theta}_4)$	20.00	14.341	14.308	14.177	13.853	13.215	11.987
$R_{(L)}(\hat{\theta}_4)$		1.0440	1.0421	1.0348	1.0164	0.9798	0.9078

5. THE MINIMAX ESTIMATORS AND THEIR PROPERTIES

The basic principle of this approach is to minimize the loss. The derivation depends primarily on a theorem, which is due to Hodge & Lehmann (1950) and can be stated as follows.

Let $\tau = \{F_\theta : \theta \in \Theta\}$ be a family of distribution functions and D be a class of estimators of the parameter θ. Suppose that $d^* \in D$ is a Bayes estimator against a prior distribution $\pi(\theta)$ on the parameter space Θ. Then the Bayes estimator d^* is said to be the Minimax estimator if the risk function of d^* is independent on Θ.

When the shape parameter v is considered to be known, the Bayes estimator for the parameter θ corresponding to the SELF and LLF are given respectively in equations (2.2) and (2.4). Further, the expressions of the risk for these Bayes estimators corresponding to the considered loss criterion are given in equations (2.6) and (2.7) respectively.

Both expressions of the risk involve the parameter θ. Hence, the Bayes estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ are not Minimax estimators. Thus, under the natural family of the conjugate prior the Minimax estimators do not exist.

Now, the Bayes estimators corresponding to the posterior $Z_2(\theta)$, given in (2.10), are obtained respectively under both loss criteria as

$$\hat{\theta}_5 = \varphi_5 \, r \, T_r \ , \qquad \varphi_5 = (d + r - 2)^{-1} \tag{5.1}$$

and

$$\hat{\theta}_6 = \varphi_6 \, r \, T_r \ , \qquad \varphi_6 = \frac{1}{a}\left(1 - \exp\Big(-\frac{a}{d+r}\Big)\right). \tag{5.2}$$

The risks of these Bayes estimators corresponding to the SELF and the LLF are given respectively as

$$R_{(S)}(\hat{\theta}_i) \ = \ \theta^2 \left(r\,(r+1)\,\varphi_i^2 \ + 1 - 2\,r\,\varphi_i\right) \tag{5.3}$$

and

$$R_{(L)}(\hat{\theta}_i) \ = \ e^{-a}\,(1 - a\,\varphi_i)^{-r} - 1 - a\,(r\,\varphi_i - 1) \ , \qquad i = 5, 6 \ . \tag{5.4}$$

It is observed that the Bayes estimators $\hat{\theta}_5$ and $\hat{\theta}_6$ are not the minimax estimators corresponding to the loss criterion SELF. However, the risk of Bayes estimators $\hat{\theta}_5$ and $\hat{\theta}_6$ are independent of the parameter θ under the LLF criterion. Hence, both estimators $\hat{\theta}_5$ and $\hat{\theta}_6$ are Minimax estimators under the LLF loss criterion.

The following statistical problem (Minimax Estimation) is equivalent to some two person zero sum game between the Statistician (Player-II) and Nature (Player-I). Here the pure strategies of Nature are the different values of θ in the interval $(0, \infty)$ and the mixed strategies of Nature are the prior densities of θ in the interval $(0, \infty)$. The pure strategies of Statistician are all possible decision functions in the interval $(0, \infty)$.

The expected value of the loss function is the risk function and it is the gain of the Player-I. Further, the Bayes risk is defined as

$$R^*(\eta, \hat{\theta}_B) \ = \ E_\theta \, R(\hat{\theta}_B) \ .$$

Here, the expectation has been taken under the prior density of parameter θ. If the loss function is continuous in both the estimator $\hat{\theta}_B$ and the parameter θ, and convex in $\hat{\theta}_B$ for each value of θ then there exist measures η^* and $\hat{\theta}_B^*$ for all θ and $\hat{\theta}_B$ so that, the following relation holds:

$$R^*(\eta, \hat{\theta}_B^*) \leq R^*(\eta^*, \hat{\theta}_B^*) \leq R^*(\eta^*, \hat{\theta}_B) \ .$$

The number $R^*(\eta^*, \hat{\theta}_B^*)$ is known as the value of the game, and η^* and $\hat{\theta}_B^*$ are the corresponding optimum strategies of the Player I and II. In statistical terms η^* is the least favorable prior density of θ and the estimator $\hat{\theta}_B^*$ is the minimax estimator. In fact, the value of the game is the loss of the Player-II. Hence, the optimum strategy of Player-II and the value of game are given as

Optimum Strategy	Corresponding Loss	Value of Game
$\hat{\theta}_5 = \varphi_5\, r\, T_r$	LLF	$e^{-a}(1-a\,\varphi_5)^{-r} - 1 - a\,(r\,\varphi_5 - 1)$
$\hat{\theta}_6 = \varphi_6\, r\, T_r$	LLF	$e^{-a}(1-a\,\varphi_6)^{-r} - 1 - a\,(r\,\varphi_6 - 1)$

ACKNOWLEDGMENTS

The authors acknowledge helpful suggestions by the referee and the editor of the paper that have considerable improved the earlier version.

REFERENCES

[1] Ahmad, A.A.; Mohammad, Z.R. and Mohamed, T.M. (2007). Bayesian prediction intervals for the future order statistics from the generalized Exponential distribution, *JIRSS*, **6**(1), 17–30.

[2] Ahmadi, J.; Doostparast, M. and Parsian, A. (2005). Estimation and prediction in a two-parameter Exponential distribution based on k-record values under LINEX loss function, *Communications in Statistics – Theory and Methods*, **34**, 795–805.

[3] Bellhouse, D.R. (2004). The reverend THOMAS Bayes FRS: a biography to celebrate the tercentenary of his birth (with discussion), *Statistical Science*, **3**, 3–43.

[4] Berrettoni, J.A. (1964). Practical applications of the Weibull distribution, *Industrial Quality Control*, **21**, 71–79.

[5] Casella, G. and Lehmann, E. (1998). *Theory of Point Estimation*, Springer Verlag.

[6] Dellaportas, P and Wright, D.E. (1991). Numerical prediction for the two-parameter Weibull distribution, *The Statistician*, **40**, 365–372.

[7] Engelhardt, M. and Bain, L.J. (1973). Some complete and censored sampling results for the Weibull or extreme value distribution, *Technometrics*, **15**, 541–549.

[8] Harris, E. and Shakarki, G. (1979). Use of the population distribution to improve estimation of individual mean in epidemiological studies, *Journal of Chronical Disease*, **32**, 233–243.

[9] Hisada, K. and Arizino, I. (2002). Reliability tests for Weibull distribution with varying shape-parameter based on complete data, *IEEE Transactions on Reliability*, **51**(3), 331–336.

[10] Hodge, J.I. and Lehmann, E.L. (1950). Some problems in Minimax estimation, *Annals of Mathematical Statistics*, **21**, 182–197.

[11] Kao, J.H.K. (1959). A graphical estimation of mixed Weibull parameters in life-testing electron tubes, *Technometrics*, **4**, 309–407.

[12] Lieblein, J. and Zelen, M. (1956). Statistical investigation of the fatigue life of deep groove ball bearings, *Journal of Res. Nat. Bur. Std.*, **57**, 273–315.

[13] Mann, N.R. (1968). *Results on statistical estimation and hypothesis testing with application to the Weibull and extreme value distribution*, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio.

[14] Mann, N.R. and Fertig, K.W. (1973). Tables for obtaining confidence bounds and tolerance bounds based on best linear invariant estimates of parameters of the extreme value distribution, *Technometrics*, **15**, 87–101.

[15] Marshall, R.J. (1991). Mapping disease and mortality rates using empirical Bayes estimators, *Journal of Applied Statistics*, **40**, 283–294.

[16] Mittnik, S. and Rachev, S.T. (1993). Modeling asset returns with alternative stable distribution, *Economic Reviews*, **12**, 261–330.

[17] Montanari, G.C.; Mazzanti, G.; Cacciari, J. and Fothergill, J.C. (1997). In search of convenient techniques for reducing bias in the estimation of Weibull parameters for uncensored tests, *IEEE Transactions on Dielectrics and Electrical Insulation*, **4**(3), 306–313.

[18] Nelson, W.B. (1972). Graphical analysis of accelerated life test data with the inverse power law model, *IEEE Transaction on Reliability*, **R-21**, 2–11.

[19] Nigm, A.M. (1989). An informative Bayesian prediction for the Weibull life-time distribution, *Communication Statistics – Theory and Methods*, **18**, 897–911.

[20] Nigm, A.M.; AL-Hussaini, E.K. and Jaheen, Z.F. (2003). Bayesian one-sample prediction of future observations under Pareto distribution, *Statistics*, **37**(6), 527–536.

[21] Pandey, B.N.; Malik, H.J. and Srivastava, R. (1989). Shrinkage testimators for the shape parameter of Weibull distribution under Type-II censoring, *Communications in Statistics – Theory and Methods*, **18**, 1175–1199.

[22] Pandey, M. (1983). Shrunken estimators of Weibull shape parameter in censored samples, *IEEE Transactions on Reliability*, **R-32**, 200–202.

[23] Pandey, M. and Singh, U.S. (1993). Shrunken estimators of Weibull shape parameter from Type-II censored samples, *IEEE Transactions on Reliability*, **42**(1), 81–86.

[24] Prakash, G. and Singh, D.C. (2006). Shrinkage testimators for the inverse dispersion of the inverse Gaussian distribution under the LINEX loss function, *Austrian Journal of Statistics*, **35**(4), 463–470.

[25] Prakash, G. and Singh, D.C. (2008). Shrinkage Estimation in Exponential Type-II Censored data Under the LINEX Loss function, *Journal of the Korean Statistical Society*, **37**(1), 53–61.

[26] Singh, D.C.; Prakash, G. and Singh, P. (2007). Shrinkage testimators for the shape parameter of Pareto distribution using the LINEX loss function, *Communication in Statistics – Theory and Methods*, **36**(4), 741–753.

[27] Singh, H.P. and Shukla, S.K. (2000). Estimation in the two-parameters Weibull distribution with prior information, *IAPQR Transactions*, **25**(2), 107–118.

[28] Singh, U.; Gupta, P.K. and Upadhyay, S.K. (2002). Estimation of exponentiated Weibull shape parameters under LINEX loss function, *Communication in Statistics – Simulation*, **31**(4), 523–537.

[29] Sinha, S.K. (1986). *Reliability and Life-Testing*, Wiley Eastern, New Delhi.

[30] Son, Y.S. and Oh, M. (2006). Bayesian estimation of the two-parameter Gamma distribution, *Communications in Statistics – Simulation and Computation*, **35**, 285–293.

[31] Thompson, J.R. (1968). Some shrinkage techniques for estimating the mean, *Journal of American Statistical Association*, **63**, 113–122.

[32] Tsionas, E.G. (2002). Bayesian analysis of finite mixtures of Weibull distributions, *Communications in Statistics – Theory and Methods*, **31**(1), 37–48.

[33] Tso, G. (1990). Forecasting money supply in Hong Kong with a multiple shrinkage estimator, *Proceeding of the ASA Section on Business and Commerce.*

[34] Varian, H.R. (1975). A Bayesian approach to real estate assessment. In "Studies in Bayesian Econometrics and Statistics in Honor of L.J. Savage" (S.E. Feinberg and A. Zellner, Eds.), Amsterdam North Holland, 195–208.

[35] Weibull, W. (1951). A statistical distribution functions of wide applicability, *Journal of Applied Math. and Mech.*, **18**, 293–297.

[36] Wooff, D.A. (1985). Bounds on reciprocal moments with applications and developments in Stein estimation and post stratification, *Journal of Royal Statistical Society*, **B-47**, 362–371.

[37] Whittemore, A. and Altschuler, B. (1976). Lung cancer incidence in cigarette smokes: further analysis of the Doll and Hill's data for British physicians, *Biometrics*, **32**, 805–816.

[38] Xu, Y. and Shi, Y. (2004). Empirical Bayes test for truncation parameters using LINEX loss, *Bulletin of the Institute of Mathematics Academia SINICA*, **32**(3), 207–220.

[39] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss function, *Journal of American Statistical Association*, **81**, 446–451.

REVSTAT – Statistical Journal
Volume 7, Number 2, June 2009, 189–201

A SADDLEPOINT APPROXIMATION TO A DISTRIBUTION-FREE TEST FOR STOCHASTIC ORDERING IN THE COMPETING RISKS MODEL

Author: Hidetoshi Murakami
– Department of Industrial and Systems Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, 112-8551 Japan
murakami@indsys.chuo-u.ac.jp

Received: March 2009 Revised: May 2009 Accepted: May 2009

Abstract:

- The approximation for the distribution function of a test statistic is extremely important in statistics. A distribution-free test for stochastic ordering in the competing risks model has been proposed by Bagai *et al.* (1989). Herein, we performed a standard saddlepoint approximation in the tails for the Bagai statistic under finite sample sizes. We then compared the saddlepoint approximation with the Bagai approximation to obtain the exact critical value. The table of the critical values was extended by using the saddlepoint approximation. Additionally, the orders of errors of a saddlepoint approximation were derived.

Key-Words:

- *saddlepoint approximation; moment generating function; Bagai statistic; distribution-free test.*

AMS Subject Classification:

- 62F03, 62G10.

1. INTRODUCTION

Testing hypotheses is one of the most important challenges in performing nonparametric statistics. Various nonparametric statistics have been proposed and discussed over the course of many years. However, the single most challenging testing problem lies in calculating the exact critical test statistic value for small data sets. It is also difficult to obtain the exact critical value when the sample sizes are moderate to large in size. Under these circumstances, we must estimate the exact critical value with an approximation method. Hence, considering approximations for evaluating the density or distribution function of the test statistic remains one of the most important topics in statistics. For the approximation presented in this study, we used a saddlepoint formula proposed by Daniels (1954, 1987) and of the type developed by Lugannani and Rice (1980). The saddlepoint approximation can be obtained for any statistic that admits a cumulant generating function. Additionally, for small sample sizes, the saddlepoint can generate accurate probabilities in the distribution tails. Saddlepoint approximations have been used with great success by many authors, and excellent discussions of their applications to a range of distributional problems are provided by Reid (1988), Jensen (1995), Goutis and Casella (1999), Huzurbazar (1999), Kolassa (2006) and Butler (2007). Additionally, Easton and Ronchetti (1986) have discussed saddlepoint approximations by using expansions of the cumulant generating function. For conducting a distribution-free test, Giles (2001) and Chen and Giles (2008) compared saddlepoint approximations with the limiting distribution of the Anderson–Darling (1952, 1954) and the Sinclair and Spurr (1988) tests and found that the saddlepoint approximations were better than both. In nonparametric statistics, researchers are very interested in considering approximations under finite sample sizes. Froda and van Eeden (2000) proposed a uniform saddlepoint expansion to the null distribution of the Wilcoxon–Mann–Whitney test (Gibbons and Chakraborti, 2003). Additionally, Bean *et al.* (2004) compared a saddlepoint approximation of the Wilcoxon–Mann–Whitney test with that of Edgeworth, and determined normal and uniform approximations under finite sample sizes.

In addition to assessing distributions, nonparametric statistics are used to test the competing risks model. Various authors have proposed the test statistics for the case in which there are competing risks and without censoring. For example, Bagai *et al.* (1989) developed distribution-free rank tests for stochastic ordering in the two independent competing risks model. Yip and Lam (1992) suggested a class of weighted log-rank-type statistics, and Neuhaus (1991) constructed the asymptotically optimal rank tests for q competing risks against stochastic ordering without censoring. Hu and Tsai (1999) considered the linear rank tests for a competing risks model. In this paper, we considered a saddlepoint approximation to the small size sample distribution of the statistic proposed by Bagai *et al.* (1989), and we estimated the exact critical value of the Bagai statistic for large

sample sizes, also using a saddlepoint approximation. We expect to apply the saddlepoint approximation to the other statistics for testing stochastic ordering in the competing risks model. In Section 2, we introduce Bagai's (1989) statistic, namely V, and the moment-generating function of the V statistic. In Section 3, we report on a saddlepoint approximation to the distribution of the V statistic and compare a saddlepoint approximation with that of Bagai. In addition, we extend the table of the critical values of the Bagai statistic using the saddlepoint approximation. In Section 4, we report on the derivation of the orders of the errors of a saddlepoint approximation.

2. THE BAGAI STATISTIC

In this section, we introduce a distribution-free test for stochastic ordering in a two independent competing risks model. We assumed that X and Y are independent and absolutely continuous random variables. Let $X_1, ..., X_n$ and $Y_1, ..., Y_n$ be two random samples of independent observations of size n, each of which has a continuous distribution described as F and G, respectively, denoting the hypothetical times to failure of the n individuals in the sample under the two risks. We observed only $(T_1, \delta_1), ..., (T_n, \delta_n)$, where $T_i = \min(X_i, Y_i)$ denotes the time to failure and $\delta_i = I(X_i > Y_i)$ indicates the cause of failure of the i-th unit. On the basis of these data types, we were interested in testing the hypothesis:

$$H_0 \colon F(x) = G(x) \qquad \text{against} \qquad H_1 \colon F(x) \leqslant G(x) \ .$$

Subsequently, Bagai *et al.* (1989) proposed a test statistic, namely V, as

$$V \ = \ 2 \sum_{i=1}^{n} (2\,n - 1 - R_i)\,\delta_i \ - \ \frac{3\,n\,(n-1)}{2} \ .$$

Here, R_i denotes the rank of T_i among $T_1, ..., T_n$. In addition, the moment-generating function of the V statistic is given by Bagai *et al.* (1989) as follows:

$$M^*(s) \ = \ 2^{-n} \exp\!\left(\frac{-3\,n\,(n-1)\,s}{2}\right) \prod_{j=2}^{n+1} \Big\{1 - \exp\big(2\,s\,(2n-j)\big)\Big\} \ .$$

However, there is a typo in the $M^*(s)$ formula given by Bagai *et al.* (1989), and we should use the moment-generating function $M(s)$ as follows:

$$M(s) \ = \ 2^{-n} \exp\!\left(\frac{-3\,n\,(n-1)\,s}{2}\right) \prod_{j=2}^{n+1} \Big\{1 + \exp\big(2\,s\,(2n-j)\big)\Big\} \ .$$

By using the moment-generating function $M(s)$, the mean and variance of the V statistic are respectively given by

$$\mathrm{E}(V) = 0 \qquad \text{and} \qquad \mathrm{var}(V) = \frac{n\,(n-1)\,(14\,n-13)}{6}$$

under the null hypothesis. We consider a saddlepoint approximation by using the moment-generating function $M(s)$ in the next section.

3. SADDLEPOINT APPROXIMATION

3.1. Saddlepoint approximation to Bagai statistic

In this section, we considered a saddlepoint approximation (Daniels, 1954, 1987) to the distribution of the V statistic. In the previous section, the moment-generating function $M(s)$ was given by

$$M(s) \;=\; 2^{-n} \exp\!\left(\frac{-3\,n\,(n-1)\,s}{2}\right) \prod_{j=2}^{n+1} \Big\{1 + \exp\big(2\,s\,(2n-j)\big)\Big\}\,.$$

The cumulant generating function of the V statistic, namely $\kappa(s)$, is

$$\kappa(s) \;=\; \log\big[M(s)\big] \;=\; -n\log 2 - \frac{3\,n\,(n-1)\,s}{2} + \sum_{j=2}^{n+1} \log\Big\{1 + \exp\big(2\,s\,(2n-j)\big)\Big\}\,.$$

To obtain the saddlepoint approximation, we evaluated the first two derivatives of the cumulant generating function as

$$\kappa^{(1)}(s) \;=\; -\frac{3\,n\,(n-1)}{2} + \sum_{j=2}^{n+1} \frac{2\,(2n-j)\exp\big(2\,s\,(2n-j)\big)}{1 + \exp\big(2\,s\,(2n-j)\big)}$$

and

$$\kappa^{(2)}(s) \;=\; \sum_{j=2}^{n+1} \frac{4\,(j-2n)^2 \exp\big(2\,s\,(j+2n)\big)}{\big(\exp(2\,j\,s) + \exp(4n\,s)\big)^2}\;,$$

where $\kappa^{(i)}(\cdot)$ denotes the i-th derivative. A highly lucid account of the generalized Lugannani and Rice formula for nonnormal distributions was suggested by Wood *et al.* (1993). Then, to determine the saddlepoint approximation to $\Pr(V \geq v)$, we solved the saddlepoint equation, $\kappa^{(1)}(s) = v$, and used the unique solution $(s = \hat{s})$ to calculate

$$\hat{w} = \sqrt{2\left(\hat{s}\,v - \kappa(\hat{s})\right)}\,\operatorname{sgn}(\hat{s}) = \sqrt{2\left(\hat{s}\,\kappa^{(1)}(\hat{s}) - \kappa(\hat{s})\right)}\,\operatorname{sgn}(\hat{s}) \quad \text{and} \quad \hat{u} = \hat{s}\sqrt{\kappa^{(2)}(\hat{s})}\;,$$

given by Wood *et al.*, where $\operatorname{sgn}(\hat{s}) = \pm 1, 0$ if $\hat{s}$ is positive, negative, or zero. The saddlepoint approximation to the cumulative distribution function of the V statistic is

$$\Pr(V \geq v) \;\approx\; 1 - \Phi(\hat{w}) + \phi(\hat{w})\left\{\frac{1}{\hat{u}} - \frac{1}{\hat{w}}\right\},$$

where $\phi(\cdot)$ is the standard normal density function and $\Phi(\cdot)$ is the corresponding cumulative distribution function.

3.2. Numerical results

In this section, we report on the evaluation of the tail probability using the saddlepoint approximation. For this test, we listed the exact probability of the V statistic derived by Bagai *et al.*, namely E_P, the Bagai's approximation, namely A_B, and a saddlepoint approximation, namely A_S, given in Tables 1 and 2.

Table 1-1: Numerical results for 1% significance level.

n	v	E_P	A_B	A_S
7	51	0.0156	0.0183	0.016256
8	68	0.0117	0.0126	0.010587
9	84	0.0117	0.0113	0.009430
10	99	0.0107	0.0117	0.010029
11	115	0.0102	0.0118	0.010393
12	134	0.0105	0.0186	0.009538
13	152	0.0102	0.0109	0.009701
14	169	0.0106	0.0116	0.010537
15	191	0.0103	0.0107	0.009674
16	210	0.0105	0.0111	0.010161
17	232	0.0100	0.0181	0.009895
18	255	0.0100	0.0105	0.009595
19	275	0.0104	0.0110	0.010205
20	298	0.0104	0.0109	0.010201

Table 1-2: Difference between E_P and approximations.

n	$\lvert E_P - A_B \rvert$	$\lvert E_P - A_S \rvert$
7	0.0027	**0.000656**
8	**0.0009**	0.001113
9	**0.0004**	0.002270
10	0.0010	**0.000671**
11	0.0016	**0.000193**
12	0.0081	**0.000962**
13	0.0007	**0.000499**
14	0.0010	**0.000063**
15	**0.0004**	0.000626
16	0.0006	**0.000339**
17	0.0081	**0.000105**
18	0.0005	**0.000405**
19	0.0006	**0.000195**
20	0.0005	**0.000199**

The numeric results for 1% and 5% significance levels are listed in Tables 1-1 and 2-1, respectively. The difference between the exact probability of the V statistic and the approximations is given in Tables 1-2 and 2-2, respectively, where the best result is in bold. Note that v and n denote the exact critical value of the Bagai statistic and the sample size, respectively. We treated the cases $7 \leq n \leq 20$ at a 1% significance level and $5 \leq n \leq 20$ at 5% significance, which represent the same cases as presented in Bagai *et al.* (1989).

Table 2-1: Numerical results for 5% significance level.

n	v	E_P	A_B	A_S
5	22	0.0625	0.0522	0.057870
6	31	0.0625	0.0499	0.051279
7	41	0.0547	0.0463	0.047075
8	50	0.0508	0.0499	0.050831
9	62	0.0508	0.0461	0.046546
10	73	0.0508	0.0472	0.047642
11	83	0.0527	0.0513	0.051902
12	98	0.0500	0.0467	0.046961
13	108	0.0528	0.0516	0.052143
14	123	0.0511	0.0494	0.049753
15	137	0.0516	0.0495	0.049831
16	150	0.0523	0.0512	0.051650
17	166	0.0519	0.0500	0.050448
18	181	0.0516	0.0506	0.050878
19	197	0.0516	0.0505	0.050748
20	214	0.0511	0.0499	0.050178

Table 2-2: Difference between E_P and approximations.

n	$\lvert E_P - A_B \rvert$	$\lvert E_P - A_S \rvert$
5	0.0103	**0.004630**
6	0.0126	**0.011221**
7	0.0084	**0.007625**
8	0.0009	**0.000031**
9	0.0047	**0.004254**
10	0.0036	**0.003158**
11	0.0014	**0.000798**
12	0.0033	**0.003039**
13	0.0012	**0.000657**
14	0.0017	**0.001347**
15	0.0021	**0.001769**
16	0.0011	**0.000650**
17	0.0019	**0.001452**
18	0.0010	**0.000722**
19	0.0011	**0.000852**
20	0.0012	**0.000922**

The results of Table 1 revealed that the saddlepoint approximation to the distribution of the V statistic is more suitable than the Bagai's approximation at a 1% significance level. For the cases of $n = 8$, 9 and 15, Bagai's approximation is better than the saddlepoint approximation. However, the saddlepoint approximation is conservative for the exact probability of the V statistic for $n = 8$ and 15. In addition, Table 2 indicates that the saddlepoint approximation to the distribution of the V statistic is better than the Bagai's approximation at the 5% significance level. We then estimated the exact critical values of the Bagai statistic using the approximation A_S for large sample sizes because it is difficult to derive an exact critical value otherwise.

Table 3: Critical values of the V statistic by saddlepoint approximation.

n	v	Probability	v	Probability
21	323	0.00991	231	0.05002
22	347	0.00997	248	0.05017
23	372	0.00994	266	0.04992
24	397	0.01001	284	0.04994
25	423	0.01000	302	0.05018
30	561	0.00997	400	0.04999
35	711	0.00998	506	0.05012
40	872	0.01001	621	0.04993
45	1044	0.01001	743	0.04993
50	1226	0.01002	872	0.04994

4. ORDERS OF ERRORS OF SADDLEPOINT APPROXIMATION

In this section, we consider the error orders of a saddlepoint approximation. From Section 3, we developed a standardized cumulant generating function of the V statistic as follows:

$$\kappa_*(s) \,=\, -n\log 2 - \frac{3\,n\,(n-1)\,s}{2\,\sigma} + \sum_{j=2}^{n+1} \log\Big\{1 + \exp\big(2\,s\,(2n-j)/\sigma\big)\Big\}\ ,$$

where

$$\sigma^2 \,=\, \frac{n\,(n-1)\,(14\,n-13)}{6}\ .$$

Then, we obtained the first four derivatives of the standardized cumulant generating as follows:

$$\kappa_*^{(1)}(s) \,=\, -\frac{3\,n\,(n-1)}{2\,\sigma} + \sum_{j=2}^{n+1} \frac{C_j\,\exp(s\,C_j)}{1+\exp(s\,C_j)}\ ,$$

$$\kappa_*^{(2)}(s) \;=\; \sum_{j=2}^{n+1} \frac{C_j^2 \exp(s\,C_j)}{\big\{1+\exp(s\,C_j)\big\}^2}\;,$$

$$\kappa_*^{(3)}(s) \;=\; \sum_{j=2}^{n+1} \frac{C_j^3 \exp(s\,C_j)\big\{1-\exp(s\,C_j)\big\}}{\big\{1+\exp(s\,C_j)\big\}^3}$$

and

$$\kappa_*^{(4)}(s) \;=\; \sum_{j=2}^{n+1} \frac{C_j^4 \exp(s\,C_j)\big\{1-4\exp(s\,C_j)+\exp(2s\,C_j)\big\}}{\big\{1+\exp(s\,C_j)\big\}^4}\;,$$

where

$$C_j \;=\; \frac{2\,(2\,n-j)}{\sigma}\;.$$

The standardized skewness and standardized kurtosis was then given by

$$\text{standardized skewness:}\quad \frac{\kappa_*^{(3)}(0)}{\kappa_*^{(2)}(0)^{3/2}} = 0$$

$$\text{standardized kurtosis:}\quad \frac{\kappa_*^{(4)}(0)}{\kappa_*^{(2)}(0)^{2}} = \frac{-12\big(186\,n^3-489\,n^2+421\,n-119\big)}{5\,n\,(n-1)\,(14n-13)^2}\;. \tag{4.1}$$

Bagai *et al.* noted that the normal approximation was appropriate for $n>20$ but that the difference of the standardized kurtosis from zero is (4.1).

We next derived the orders of the errors of the V statistic. By using an expansion for the standardized cumulant generating function, we approximated the $\kappa_*(s)$ as follows:

$$\begin{aligned}
\kappa_*(s) \;\approx\; & -n\log 2 - \frac{3\,n\,(n-1)\,s}{2\,\sigma} + \sum_{j=2}^{n+1}\left\{\log 2 + \frac{2\,s\,(2n-j)}{2\,\sigma} + \frac{4s^2\,(2n-j)^2}{8\,\sigma^2}\right.\\
& \left. - \frac{16s^4(2n-j)^4}{192\,\sigma^4} + \frac{64s^6(2n-j)^6}{2880\,\sigma^6} + \cdots\right\}\\
\approx\; & \frac{ns^2(n-1)\,(14n-13)}{12\,\sigma^2} - \frac{ns^4(n-1)\,(186\,n^3-489\,n^2+421\,n-119)}{360\,\sigma^4}\\
& + \frac{ns^6(n-1)\,(76\,n^5-3207\,n^4+5256\,n^3-494n^2+1637\,n-253)}{1890\,\sigma^6}\\
\approx\; & \frac{s^2}{2} - \frac{1}{n}\left\{\frac{s^4\big(186\,n^3-489\,n^2+421\,n-119\big)}{10\,(n-1)\,(14n-13)^2}\right\}\\
& + \frac{1}{n^2}\left\{\frac{4s^6\big(762\,n^5-3207\,n^4+5256\,n^3-4194n^2+1637\,n-253\big)}{35\,(n-1)^2\,(14n-13)^3}\right\}\\
& + O(n^{-3})\;.
\end{aligned}$$

We then approximated the first four derivatives of the standardized cumulant generating function by

$$\begin{aligned}
\kappa_*^{(1)}(s) \;\approx\; & s - \frac{1}{n}\left\{\frac{2s^3\left(186\,n^3-489\,n^2+421\,n-119\right)}{5\,(n-1)\,(14n-13)^2}\right\} \\
& + \frac{1}{n^2}\left\{\frac{24s^5\left(762\,n^5-3207\,n^4+5256\,n^3-4194\,n^2+1637\,n-253\right)}{35\,(n-1)^2\,(14n-13)^3}\right\} \\
& + O(n^{-3})\ , \\
\kappa_*^{(2)}(s) \;\approx\; & 1 - \frac{1}{n}\left\{\frac{6s^2\left(186\,n^3-489\,n^2+421\,n-119\right)}{5\,(n-1)\,(14n-13)^2}\right\} \\
& + \frac{1}{n^2}\left\{\frac{24s^4\left(762\,n^5-3207\,n^4+5256\,n^3-4194\,n^2+1637\,n-253\right)}{7\,(n-1)^2\,(14n-13)^3}\right\} \\
& + O(n^{-3})\ , \\
\kappa_*^{(3)}(s) \;\approx\; & -\frac{1}{n}\left\{\frac{12\,s\left(186\,n^3-489\,n^2+421\,n-119\right)}{5\,(n-1)\,(14n-13)^2}\right\} \\
& + \frac{1}{n^2}\left\{\frac{96s^3\left(762\,n^5-3207\,n^4+5256\,n^3-4194\,n^2+1637\,n-253\right)}{7\,(n-1)^2\,(14n-13)^3}\right\} \\
& + O(n^{-3})
\end{aligned}$$

and

$$\begin{aligned}
\kappa_*^{(4)}(s) \;\approx\; & -\frac{1}{n}\left\{\frac{12\left(186\,n^3-489\,n^2+421\,n-119\right)}{5\,(n-1)\,(14n-13)^2}\right\} \\
& + \frac{1}{n^2}\left\{\frac{288\,s^2\left(762\,n^5-3207\,n^4+5256\,n^3-4194\,n^2+1637\,n-253\right)}{7\,(n-1)^2\,(14n-13)^3}\right\} \\
& + O(n^{-3})\ .
\end{aligned}$$

By expanding for $\kappa_*(s)$ in a Taylor series, we then determined

$$0 = \kappa_*(0) = \kappa_*(s) - s\kappa_*^{(1)}(s) + \frac{s^2\kappa_*^{(2)}(s)}{2} - \frac{s^3\kappa_*^{(3)}(s)}{6} + \frac{s^4\kappa_*^{(4)}(s)}{24} + \cdots\ .$$

Then, substituting to the w, we determined

$$\begin{aligned}
w^2 \;&=\; 2\left\{s\,\kappa_*^{(1)}(s) - \kappa_*(s)\right\} \\
&=\; 2\left\{s\,\kappa_*^{(1)}(s) - s\,\kappa_*^{(1)}(s) + \frac{s^2\kappa_*^{(2)}(s)}{2} - \frac{s^3\kappa_*^{(3)}(s)}{6} + \frac{s^4\kappa_*^{(4)}(s)}{24} + \cdots\right\} \\
&=\; s^2\kappa_*^{(2)}(s)\left\{1 - \frac{s\,\kappa_*^{(3)}(s)}{3\,\kappa_*^{(2)}(s)} + \frac{s^2\kappa_*^{(4)}(s)}{12\,\kappa_*^{(2)}(s)} + \cdots\right\}.
\end{aligned}$$

Therefore, we obtained

$$\begin{aligned}\frac{1}{w} &= \frac{1}{s\sqrt{\kappa_*^{(2)}(s)}}\left\{1-\frac{s\,\kappa_*^{(3)}(s)}{3\,\kappa_*^{(2)}(s)}+\frac{s^2\kappa_*^{(4)}(s)}{12\,\kappa_*^{(2)}(s)}+\cdots\right\}^{-\frac{1}{2}}\\ &= \frac{1}{s\sqrt{\kappa_*^{(2)}(s)}}\left\{1+\frac{s\,\kappa_*^{(3)}(s)}{6\,\kappa_*^{(2)}(s)}-\frac{s^2\kappa_*^{(4)}(s)}{24\,\kappa_*^{(2)}(s)}+\frac{s^2\big(\kappa_*^{(3)}(s)\big)^2}{24\big(\kappa_*^{(2)}(s)\big)^2}+\cdots\right\}\\ &= \frac{1}{s\sqrt{\kappa_*^{(2)}(s)}}+O(n^{-1})\end{aligned}$$

by applying the binomial theorem and substituting the i-th standardized cumulant. To determine the saddlepoint approximation to $\Pr\big((V-\mathrm{E}(V))/\sigma \geq v_*\big)$, we solved the saddlepoint equation, $\kappa_*^{(1)}(s) = v_*$, and used the unique solution $(s = \hat{s})$ to calculate

$$\hat{w} = \sqrt{2\left(\hat{s}\,v_* - \kappa_*(\hat{s})\right)}\,\mathrm{sgn}(\hat{s}) \qquad \text{and} \qquad \hat{u} = \hat{s}\,\sqrt{\kappa_*^{(2)}(\hat{s})}\,,$$

where $\mathrm{sgn}(\hat{s}) = \pm 1, 0$ if $\hat{s}$ is positive, negative, or zero. Therefore, we determined

$$\begin{aligned}\Pr\left(\frac{V-\mathrm{E}(V)}{\sigma} \geq v_*\right) &\approx 1-\Phi(\hat{w})+\phi(\hat{w})\left\{\frac{1}{\hat{u}}-\frac{1}{\hat{w}}+O(n^{-1})\right\}\\ &= 1-\Phi(\hat{w})+\phi(\hat{w})\left(\frac{1}{\hat{u}}-\frac{1}{\hat{w}}\right)+O(n^{-1})\,.\end{aligned}$$

Note that $\phi(\hat{w}) \approx \text{Constant} + O(n^{-1})$. Typically, an approximation of the above form has the relative error $O(n^{-3/2})$. However, the exact distribution of the standardized Bagai statistic is discrete, so the discrete distribution of the standardized Bagai statistic may be approximated at its support point by a smooth function that behaves similarly to a distribution function. Therefore, $\Phi(r^*(w))$ approximates the distribution function of the Bagai statistic with a relative error of $O(n^{-1})$ in a normal deviation region in which $r^*(w) = w + w^{-1}\log(u/w)$; Barndorff-Nielsen and Cox (1994). The use of saddlepoint approximation as a technique for smoothing discrete distributions is discussed by Davison and Wang (2002).

5. CONCLUDING REMARKS

In this paper, we considered the saddlepoint approximation to the distribution of the Bagai statistic V (1989). The standard saddlepoint formula provided an accurate approximation to the distribution of the V statistic. From the numerical results, we determined that the approximation precision of the saddlepoint approximation is superior to the Bagai's approximation using finite sample

sizes. The orders of the errors of a saddlepoint approximation were also derived. In future work, we intend to 1) compare the orders of the errors of the higher-order saddlepoint approximation, Bagai's approximation, and other approximations, 2) be able to apply the saddlepoint approximation to other statistics for testing the independent competing risks model, and 3) consider the saddlepoint approximation to the distribution of the V statistic for cases of dependent competing risks models, *e.g.* Aly, Kochar and McKeague (1994), Dykstra, Kochar and Robertson (1995).

ACKNOWLEDGMENTS

The author would like to give his thank to the referee and the editor for their valuable comments and suggestions. It is a pleasure to acknowledge conversations with Professor T. Yanagimoto.

REFERENCES

[1] Aly, E.A.A.; Kochar, S.C. and McKeague, L.W. (1994). Some tests for comparing cumulative incidence functions and cause-specific hazard rates, *Journal of the American Statistical Association*, **89**, 994–999.

[2] Anderson, T.W. and Darling, D.A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes, *Annals of Mathematical Statistics*, **23**, 193–212.

[3] Anderson, T.W. and Darling, D.A. (1954). A test of goodness of fit, *Journal of the American Statistical Association*, **49**, 765–769.

[4] Bagai, I.; Deshpandé, J.V. and Kochar, S.C. (1989). Distribution free tests for stochastic ordering in the competing risks model, *Biometrika*, **76**(4), 775–781.

[5] Barndorff-Nielsen, O.E. and Cox, D.R. (1994). *Inference and Asymptotic*, Chapman and Hall.

[6] Bean, R.; Froda, S. and van Eeden, C. (2004). The normal, Edgeworth, saddlepoint and uniform approximations to the Wilcoxon–Mann–Whitney null-distribution: A numerical comparison, *Journal of Nonparametric Statistics*, **16**(1–2), 279–288.

[7] Butler, R.W. (2007). *Saddlepoint Approximations with Applications*, Cambridge University Press.

[8] Chen, Q. and Giles, D.E. (2008). General saddlepoint approximations: Application to the Anderson–Darling test statistic, *Communications in Statistics – Simulation and Computation*, **37**, 789–804.

[9] Daniels, H.E. (1954). Saddlepoint approximations in statistics, *Annals of Mathematical Statistics*, **25**, 631–650.

[10] Daniels, H.E. (1987). Tail probability approximations, *International Statistical Review*, **55**, 37–48.

[11] Davison, A.C. and Wang, S. (2002). Saddlepoint approximations as smoothers, *Biometrika*, **89**(4), 933–938.

[12] Dykstra, R.; Kochar, S. and Robertson, T. (1995). Likelihood based inference for cause specific hazard rates under restrictions, *Journal of Multivariate Analysis*, **54**, 163–174.

[13] Easton, G.S. and Ronchetti, E. (1986). General saddlepoint approximations with applications to L statistics, *Journal of the American Statistical Association*, **81**, 420–430.

[14] Froda, S. and van Eeden, C. (2000). A uniform saddlepoint expansion for the null-distribution of the Wilcoxon–Mann–Whitney statistic, *Canadian Journal of Statistics*, **28**, 137–149.

[15] Gibbons, J.D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference*, Fourth Edition, Dekker, New York.

[16] Giles, D.A.E. (2001). A saddlepoint approximation to the distribution function of the Anderson–Darling test statistic, *Communications in Statistics – Simulation and Computation*, **30**, 899–905.

[17] Goutis, C. and Casella, G. (1999). Explaining the saddlepoint approximation, *American Statistician*, **53**, 216–224.

[18] Hu, X.S. and Tsai, W.Y. (1999). Linear rank tests for competing risks model, *Statistica Sinica*, **9**, 971–983.

[19] Huzurbazar, S. (1999). Practical saddlepoint approximations, *American Statistician*, **53**, 225–232.

[20] Jensen, J.L. (1995). *Saddlepoint Approximations*, Oxford University Press.

[21] Kolassa, J.E. (2006). *Series Approximation Methods in Statistics*, Springer-Verlag.

[22] Lugannani, R. and Rice, S.O. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables, *Advances in Applied Probability*, **12**, 475–490.

[23] Neuhaus, G. (1991). Some linear and nonlinear rank tests for competing risks models, *Communications in Statistics – Theory and Methods*, **20**, 667–701.

[24] Reid, N. (1988). Saddlepoint methods and statistical inference (with discussion), *Statistical Science*, **3**, 213–238.

[25] Sinclair, C.D. and Spurr, B.D. (1988). Approximations to the distribution function of the Anderson–Darling test statistic, *Journal of the American Statistical Association*, **83**, 1190–1191.

[26] Wood, A.T.A.; Booth, J.G. and Butler, R.W. (1993). Saddlepoint approximations to the CDF of some statistics with nonnormal limit distributions, *Journal of the American Statistical Association*, **88**, 680–686.

[27] Yip, P. and Lam, K.F. (1992). A class of non-parametric tests for the equality of failure rates in a competing risks model, *Communications in Statistics – Theory and Methods*, **21**, 2541–2556.

REVSTAT – Statistical Journal

Background

Statistical Institute of Portugal (INE, I.P.), well aware of how vital a statistical culture is in understanding most phenomena in the present-day world, and of its responsibility in disseminating statistical knowledge, started the publication of the scientific statistical journal *Revista de Estatística*, in Portuguese, publishing three times a year papers containing original research results, and application studies, namely in the economic, social and demographic fields.

In 1998 it was decided to publish papers also in English. This step has been taken to achieve a larger diffusion, and to encourage foreign contributors to submit their work.

At the time, the Editorial Board was mainly composed by Portuguese university professors, being now composed by national and international university professors, and this has been the first step aimed at changing the character of *Revista de Estatística* from a national to an international scientific journal.

In 2001, the *Revista de Estatística* published three volumes special issue containing extended abstracts of the invited contributed papers presented at the 23rd European Meeting of Statisticians.

The name of the Journal has been changed to REVSTAT – STATISTICAL JOURNAL, published in English, with a prestigious international editorial board, hoping to become one more place where scientists may feel proud of publishing their research results.

— The editorial policy will focus on publishing research articles at the highest level in the domains of Probability and Statistics with emphasis on the originality and importance of the research.

— All research articles will be refereed by at least two persons, one from the Editorial Board and another, external.

— The only working language allowed will be English.

— Three volumes are scheduled for publication, one in April, one in June and the other in November.

— On average, four articles will be published per issue.

Aims and Scope

The aim of REVSTAT is to publish articles of high scientific content, in English, developing innovative statistical scientific methods and introducing original research, grounded in substantive problems.

REVSTAT covers all branches of Probability and Statistics. Surveys of important areas of research in the field are also welcome.

Abstract / indexed in

REVSTAT is expected to be abstracted / indexed at least in *Current Index to Statistics*, *Statistical Theory and Method Abstracts* and *Zentralblatt für Mathematik*.

Instructions to Authors, special-issue editors and publishers

Papers may be submitted in two different ways:

— By sending a paper copy to the Executive Editor and one copy to one of the two Editors or Associate Editors whose opinion the author(s) would like to be taken into account, together with a postscript or a PDF file of the paper to the e-mail: revstat@fc.ul.pt.

— By sending a paper copy to the Executive Editor, together with a postscript or a PDF file of the paper to the e-mail: revstat@fc.ul.pt.

Submission of a paper means that it contains original work that has not been nor is about to be published elsewhere in any form.

Submitted manuscripts (text, tables and figures) should be typed only in black, on one side, in double spacing, with a left margin of at least 3 cm and not have more than 30 pages.

The first page should include the name, affiliation and address of the author(s) and a short abstract with the maximum of 100 words, followed by the key words up to the limit of 6, and the AMS 2000 subject classification.

Authors are obliged to write the final version of accepted papers using LaTeX, in the REVSTAT style.

This style (REVSTAT.sty), and examples file (REVSTAT.tex), which may be download to *PC Windows System* (Zip format), *Mackintosh*, *Linux* and *Solaris Systems* (StuffIt format), and *Mackintosh System* (BinHex Format), are available in the REVSTAT link of the National Statistical Institute's Website: http://www.ine.pt/revstat/inicio.html

Additional information for the authors may be obtained in the above link.

Accepted papers

Authors of accepted papers are requested to provide the LaTeX files and also a postscript (PS) or an acrobat (PDF) file of the paper to the Secretary of REVSTAT: liliana.martins@ine.pt.

Such e-mail message should include the author(s)'s name, mentioning that it has been accepted by REVSTAT.

The authors should also mention if encapsulated postscript figure files were included, and submit electronics figures separately in .tiff, .gif, .eps or .ps format. Figures must be a minimum of 300 dpi.

Also send always the final paper version to:

Maria José Carrilho
Executive Editor, REVSTAT – STATISTICAL JOURNAL
Instituto Nacional de Estatística, I.P.
Av. António José de Almeida
1000-043 LISBOA
PORTUGAL

Copyright