The total median statistic to monitor contaminated normal data*

Fernanda Figueiredo† M. Ivette Gomes‡

Abstract

Despite of the advantages of the use of the normal distribution in Statistical Quality Control the normality assumption is too restrictive for modeling real data sets, which usually exhibit asymmetry or tails heavier than the normal tails. But even in potential normal situations there is often a small to moderate percentage of contamination in the data. In this paper we analyze the efficiency and robustness of the total median statistic comparatively to the sample mean and the sample median to estimate the mean value of symmetric contaminated normal distributions, close to the normal, but with heavier-than-normal tails. We also compare the performance of the total median and the sample mean charts to monitor the mean value of such processes. The simulation results lead us to suggest the use of the total median statistic due to its efficiency and degree of robustness.

*Research partially supported by national funds through the Fundação Nacional para a Ciência e Tecnologia, Portugal – FCT, under the PEst-OE/MAT/UI0006/2011 project.

† Corresponding author. Faculty of Economics, University of Porto and CEAUL; Rua Dr. Roberto Frias, 4200-464 Porto, Portugal. E-mail: otilia@fep.up.pt

‡ FCUL, DEIO and CEAUL, University of Lisbon; Bloco C4, Campo Grande, 1749-016 Lisboa, Portugal. E-mail: ivette.gomes@fc.ul.pt

AMS 2000 subject classification: 62G05, 62G35, 62P30, 65C05.

Keywords: Statistical Process Control; Control Charts; Robust Estimators; Order statistics.

1 Introduction

Most of the data sets from diverse industrial processes, for instance, in the areas of telecommunications, reliability, insurance, finance and economics among others, exhibit asymmetry and/or tails heavier than the normal tail, and thus, the normality assumption is too restrictive for modeling this kind of data. But even in potential normal situations, when we are working with real applications there is often a possibility of having disturbances in the data, for instance, a small to moderate percentage of contamination.

There are several approaches in the literature to accommodate the possibility of having non-normal data. Figueiredo and Gomes (2013) suggest the use of the parametric family of skew-normal distributions to model real data sets. They enhance the flexibility of this family to accommodate uncontrollable disturbances in the data, such as some level of asymmetry or non-normal tail behavior, and provide some control charts to monitor industrial processes based on this class of distributions. For a detailed study on the properties of the skew-normal distribution see, for instance, Azzalini (1985, 1986, 2005). Bai and Choi (1995), Castagliola (2000) and Chan and Heng (2003), among others, propose control charts for skewed populations. For purposes of robust control charting, we refer Langenberg and Iglewicz (1986), Chan *et al.* (1988), Rocke (1989, 1992), Castagliola (2001), Figueiredo and Gomes (2004, 2009), Jensen *et al.* (2006), Chakraborti *et al.* (2009) and Human and Chakraborti (2010), among others. For a recent overview on the latest developments on nonparametric control charts see, for instance, Chakraborti *et al.* (2011) and references therein. This paper presents a different approach to monitor contaminated-normal data, based on the use of the total median statistic.

Contaminated distributions have been used in many areas of application, including medicine and clinical chemistry in hematology studies, engineering and physical sciences to model polymer chains length and particles size generated by multiple mechanisms, genetics in gene mutation and microarray studies, biology in chain branching processes and fishery to model lengths of fish in overlapped populations. For details on applications of contaminated distributions see, for instance, Brownie *et al.* (1983), Gleason (1993), McLaren (1996), Liang and Rathouz (1999), Chen *et al.* (2008) and Ghosh and Chinnaiyan (2009).

Although the sample mean is the most efficient estimator for the mean value of a normal process its efficiency decreases substantially when we have some deviations to the normality and we can not consider it as a robust estimator. Similarly, the mean chart, usually designed under the assumption of independent and identically normal distributed observations, is not the most appropriate chart to monitor the mean value of processes that exhibit asymmetry and/or heavy-tails. This chart is not robust to deviations to the normality assumption, i.e., it easily exhibits average run-length (ARL) values very different (higher or smaller) from the expected value when we have non-normal data or even small disturbances in the data. Moreover, although the mean chart has a reasonable performance to detect moderate to large changes in the mean value of a normal process, its efficiency can be drastically affected if the data come from a distribution with moderate to heavy tails. For details see, among others, Schilling and Nelson (1976), Balakrishnan and Kocherlakota (1986), Chan *et al.* (1988), Chan and Heng (2003) and Figueiredo and Gomes (2004, 2009).

Following Figueiredo and Gomes (2004, 2009), but for other symmetric distributions close to the normal with heavier-than-normal tails, we have carried out a simulation study to analyze the efficiency and robustness of the total median statistic, here denoted TMd, and of the usual estimators for the mean value, the sample mean and the sample median, here denoted $M \equiv \overline{X}$ and Md, respectively. The plan of the paper is as follows. In Section 2 we present a brief description of the total median statistic. In Section 3 we analyze the efficiency and robustness of the total median estimator under a wide variety of distributional situations and we compare it with the usual estimators for the process mean value. The results of a simulation experiment about the performance of the mean and the total median charts are presented in Section 4. Finally we include some concluding remarks in Section 5.

2 Description of the total median statistic

Let $(X_1, X_2, ..., X_n)$ be a random sample of size n from a distribution F and let us denote $X_{i:n}$, $1 \leq i \leq n$, the random sample of the associated ascending order statistics ($o.s.$). The bootstrap sample, $(X_1^*, X_2^*, ..., X_n^*)$, is obtained by randomly sampling n times, with replacement, from the observed sample $(x_1, x_2, ..., x_n)$. Denoting BMd the median of the bootstrap sample, the probabilities $\alpha_{ij} = P\left(BMd = \frac{x_{i:n}+x_{j:n}}{2}\right)$, $1 \leq i \leq j \leq n$, are given by

$$\alpha_{ij} = \begin{cases} \frac{1}{n^n} \sum_{k=0}^{(n-1)/2} \frac{n!(i-1)^k}{k!(n-k)!} \sum_{r=[n/2]-k+1}^{n-k} \frac{(n-k)!(n-i)^{n-k-r}}{r!(n-k-r)!}, \ 1 \leq i = j \leq n \\ \frac{n!\left\{i^{n/2}-(i-1)^{n/2}\right\}\left\{(n-j+1)^{n/2}-(n-j)^{n/2}\right\}}{n^n((n/2)!)^2}, \ n \text{ even and } 1 \leq i < j \leq n \\ 0, \ n \text{ odd and } 1 \leq i < j \leq n, \end{cases}$$

where $[x]$ denotes the integer part of x, with $\alpha_{ij} = \alpha_{n-j+1,n-i+1}$, $\forall k$. Note that these probabilities are independent of the underlying model F and they depend only on the sample size n. Details to obtain the $\alpha_{i,j}$ can be found in Cox and Iguzquiza (2001) and Figueiredo and Gomes (2004).

The total median statistic, TMd, is given by

$$TMd := \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{ij} \, \frac{X_{i:n} + X_{j:n}}{2} = \sum_{i=1}^{n} a_i X_{i:n}, \tag{2.1}$$

with $a_i = \frac{1}{2}\left(\sum_{j=i}^{n} \alpha_{ij} + \sum_{j=1}^{i} \alpha_{ji}\right)$, $1 \leq i \leq n$. These coefficients a_i are independent of F, and for the most common sample sizes in SPC are given in Table 1.

Table 1: Coefficients a_i for sample sizes n from 3 up to 10.

n	1	2	3	4	5
3	0.259	0.482			
4	0.156	0.344			
5	0.058	0.259	0.366		
6	0.035	0.174	0.291		
7	0.010	0.098	0.239	0.306	
8	0.007	0.064	0.172	0.257	
9	0.001	0.029	0.115	0.221	0.268
10	0.001	0.019	0.078	0.168	0.234

We easily observe that the TMd statistic is a linear combination of the sample order statistics, where the most extreme observations of the sample have smaller weights than the other observations, which enables us to obtain a robust estimator to small disturbances in the sample. We also note that the TMd statistic converges for the median value of the underlying distribution, and thus, for symmetric distributions it is an unbiased estimator for the process mean value. In general it is not possible to obtain the exact distribution of the TMd statistic, but we easily obtain accurate quantiles by simulation. But when the samples come from an exponential distribution, the TMd statistic can be written as a mixture of independent exponential variables or as a linear combination of independent chi-square variables, and we are able to obtain the exact distribution of the TMd. Note that the transformation $Y = -\ln(1 - F(X))$ enables us to transform the observations from a process X with distribution function ($d.f.$) F into exponential standard data Y.

3 Efficiency and robustness of the total median statistic

To analyze the efficiency and the robustness of the total median (TMd), the sample mean (M) and the sample median (Md) estimators for the mean value, apart from the standard normal distribution, $N(0, 1)$, we consider symmetric contaminated normal distributions, F, close to the normal, but with heavier-than-normal tails. More precisely, the following models F usually used for creating outliers and modeling data sets that exhibit heavy tails:

- the scaled-contaminated normal distributions, here denoted $CN(\alpha, \lambda)$, for $\alpha = 5\%, 10\%, 15\%, 20\%$ and $\lambda = 3$, with $d.f.$

$$F(x) = (1 - \alpha)\Phi(x) + \alpha\Phi(x/\lambda);$$

- the student-contaminated normal distribution, here denoted $CN(\alpha, t_k)$, for $\alpha = 5\%, 10\%, 15\%, 20\%$ and $k = 3$, with $d.f.$

$$F(x) = (1 - \alpha)\Phi(x) + \alpha F_{t_k}(x),$$

where F_{t_k} denotes the $d.f.$ of the Student-t distribution, t_k.

Note that a distribution that is stretched relative to the Gaussian is said to have tails heavier than the normal tails, i.e., its quantiles change more rapidly than the Gaussian quantiles. To measure the weight of the tails of a symmetric distribution F, Hoaglin *et al.* (1983) proposed the tail-weight coefficient, τ, given by

$$\tau = \left(\frac{F^{-1}(0.99)-F^{-1}(0.5)}{F^{-1}(0.75)-F^{-1}(0.5)}\right) / \left(\frac{\Phi^{-1}(0.99)-\Phi^{-1}(0.5)}{\Phi^{-1}(0.75)-\Phi^{-1}(0.5)}\right),$$

where F^{-1} and Φ^{-1} denote the inverse of the $d.f.$ F and of the standard normal $d.f.$ Φ.

Let us generally denote the estimators under study by T_n. To compare the efficiency of the estimators we evaluate their variance, $Var(T_n)$, by a Monte Carlo simulation experiment of size 1000000. The most efficient estimator is the one that presents the minimum value of $Var(T_n)$.

To select a robust estimator among the estimators under study, we have proceeded as follows:

- first, for every considered distribution we have obtained the most efficient estimator, among the ones under study;
- next, we have computed the relative efficiency of the other estimators relatively to the best one selected previously (given by the quotient of their variances), and we have retained the smallest value.

The degree of robustness of an estimator is given by this minimum efficiency, and the robust estimator is the one with the highest minimum efficiency.

Table 2 presents the most efficient estimator for the mean value of the distributions under study and for samples of size 3 up to 10. Table 3 indicates their efficiency relatively to the best one, for samples of size 5. From these tables, and with some exceptions, the total median estimator turns out to be the most efficient estimator for the mean value as the tail-weight of the underlying distribution increases. The sample mean is the most efficient only when we consider samples of $N(0,1)$ data or moderate-to-large samples of $CN(5\%, t_3)$ data. The sample median is the most efficient in the extreme cases of very small samples from $CN(15\%, 3)$ and $CN(20\%, 3)$ distributions, the ones with the highest tail-weight. For samples of size 5, we obtain several relative efficiencies of the sample mean and of the sample median significantly smaller than 100%, with minimum values of 75.02% and 69.75%, respectively, while the minimum value obtained for the relative efficiency of the total median estimator is 92.94%.

The degree of robustness of the M, TMd and Md estimators is presented in Table 3. These values are pictured in Figure 1 and allow us to easily identify the most robust estimator for the mean value and see how distant it is from the other estimators. Without any doubt, the total median statistic is the most robust estimator for the mean value among the ones here considered, and the sample mean is not at all robust.

Table 2: Most efficient estimator for sample sizes 3 up to 10

Distribution	τ	3	4	5	6	7	8	9	10
$N(0,1)$	1.000	M	M	M	M	M	M	M	M
$CN(5\%, t_3)$	1.029	TMd	TMd	M	M	M	M	M	M
$CN(10\%, t_3)$	1.062	TMd	TMd	TMd	TMd	TMd	TMd	TMd	TMd
$CN(15\%, t_3)$	1.098	TMd	TMd	TMd	TMd	TMd	TMd	TMd	TMd
$CN(20\%, t_3)$	1.139	TMd	TMd	TMd	TMd	TMd	TMd	TMd	TMd
$CN(5\%, 3)$	1.205	TMd	TMd	TMd	TMd	TMd	TMd	TMd	TMd
$CN(10\%, 3)$	1.532	TMd	Md	TMd	TMd	TMd	TMd	TMd	TMd
$CN(15\%, 3)$	1.717	Md	Md	TMd	TMd	TMd	TMd	TMd	TMd
$CN(20\%, 3)$	1.802	Md	Md	TMd	TMd	TMd	TMd	TMd	TMd

Table 3: Relative efficiency of the estimators for samples of size 5.

Distribution	τ	M	TMd	Md
$N(0,1)$	1.000	1.0000	0.9294	0.6975
$CN(5\%, t_3)$	1.029	1.0000	0.9950	0.7597
$CN(10\%, t_3)$	1.062	0.9568	1.0000	0.7730
$CN(15\%, t_3)$	1.098	0.9074	1.0000	0.7861
$CN(20\%, t_3)$	1.139	0.8813	1.0000	0.7946
$CN(5\%, 3)$	1.205	0.8935	1.0000	0.7992
$CN(10\%, 3)$	1.532	0.8086	1.0000	0.8445
$CN(15\%, 3)$	1.717	0.7679	1.0000	0.8849
$CN(20\%, 3)$	1.802	0.7502	1.0000	0.9163

Table 4: Degree of robustness of the M, TMd and Md estimators.

Sample size n	M	TMd	Md	Robust estimator
3	0.8699	0.9647	0.7435	TMd
4	0.7647	0.9073	0.8389	TMd
5	0.7502	0.9294	0.6975	TMd
6	0.7258	0.9230	0.7764	TMd
7	0.6964	0.8899	0.6780	TMd
8	0.6856	0.8885	0.7433	TMd
9	0.6747	0.8650	0.6701	TMd
10	0.6691	0.8622	0.7217	TMd

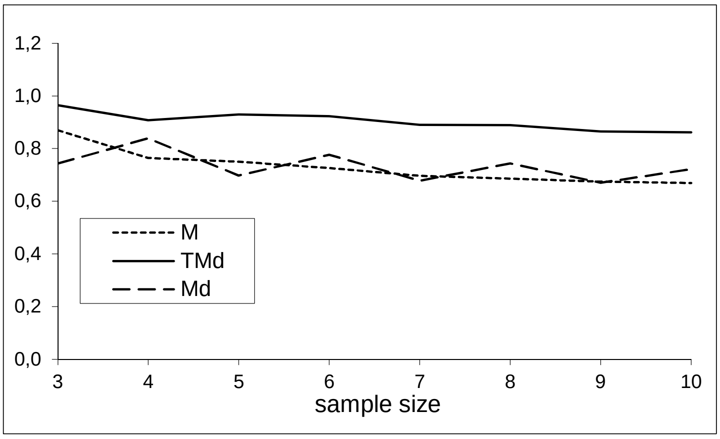

Figure 1: Degree of robustness of the estimators under study

4 The mean and the total median charts

We have implemented the mean and the total median charts, here denoted, M-chart and the TMd-chart, for rational subgroups of size 5, with lower and upper control limits, LCL and UCL, placed at the quantiles $\chi_{0.001}$ and $\chi_{0.999}$ of the simulated distribution of the control statistic. When the process is in-control state we assume its mean value is μ_0 and its standard deviation is σ_0. When the process is out-of-control, we assume $\mu = \mu_0 + \delta\sigma_0$ with $\delta \neq 0$, and that the standard deviation maintains equal to σ_0.

To analyze the ability of the M and TMd charts to detect changes in the process mean value we have computed, through a Monte Carlo simulation experiment of size 1000000, the ARL (*Average Run Length*) value, ie., the expected number of samples taken before the chart signals, for a few different magnitude changes $\Delta = \delta\sigma_0$. For the symmetric distributions here considered, already referred in Section 3, the distribution of the TMd statistic is nearly symmetric, and thus, we have only computed the ARL values for $\delta > 0$. The in-control ARL is approximately 500. The obtained ARL estimates are presented in Tables 5-6, together with the upper and the lower control limits of the charts. The comparison of the ARL-behavior of the M and TMd charts is easily understood through the ARL values pictured in Figures 2-3.

For normal processes, despite of the fact that the M-chart presents smaller ARL values than the ones obtained for the TMd-chart, as expected, the differences are not too large. For the other models, the TMd chart exhibits a considerable better performance than the M-chart, and detects small to moderate changes in the process mean value very fast. The results obtained for the scaled-contaminated normal distributions can be explained by the fact that changes in α and λ affects the scale of the distribution, and thus, the tail-weight effect is mystified by these changes in scale.

Table 5: ARL estimates of the M and TMd charts implemented for samples of size 5 and $\Delta = \delta\sigma_0$ magnitude changes in the process mean value. Normal and scaled-contaminated normal in-control distributions.

	$N(0,1)$		$CN(5\%,3)$		$CN(10\%,3)$		$CN(15\%,3)$	
δ	M	TMd	M	TMd	M	TMd	M	TMd
0.0	501.3	501.5	499.8	499.8	499.5	499.8	499.5	500.0
0.1	405.5	395.9	452.5	426.1	452.1	430.5	446.0	451.9
0.2	234.6	242.8	357.3	276.2	340.7	313.9	339.2	336.6
0.3	127.6	137.1	268.1	159.8	238.8	201.3	232.1	228.1
0.4	71.2	77.0	186.4	91.7	162.8	119.7	152.1	143.3
0.5	41.5	45.2	126.5	53.5	110.6	70.5	100.1	86.9
0.6	25.1	27.9	84.7	31.8	74.7	41.6	66.8	52.6
0.7	15.8	17.6	56.3	19.5	50.8	25.0	44.9	31.9
0.8	10.4	11.6	37.4	12.4	34.4	15.5	30.4	19.5
0.9	7.1	8.0	24.6	8.2	23.4	9.9	20.8	12.1
1.0	5.1	5.7	16.3	5.6	15.9	6.5	14.2	7.8
1.1	3.8	4.2	10.9	4.1	10.9	4.5	9.9	5.3
1.2	2.9	3.3	7.5	3.1	7.5	3.3	6.9	3.7
1.3	2.3	2.6	5.3	2.4	5.3	2.5	5.0	2.7
1.4	1.9	2.1	3.9	1.9	3.9	2.0	3.6	2.1
1.5	1.7	1.8	2.9	1.6	2.9	1.7	2.8	1.7
2.0	1.1	1.1	1.3	1.1	1.3	1.1	1.2	1.1
LCL	-1.3801	-1.4318	-1.9683	-1.6452	-2.2397	-1.8841	-2.4543	-2.1285
UCL	1.3840	1.4348	1.9754	1.6397	2.2308	1.8775	2.4303	2.1182

Table 6: ARL estimates of the M and TMd charts implemented for samples of size 5 and $\Delta = \delta\sigma_0$ magnitude changes in the process mean value. Student-contaminated normal in-control distributions.

	$CN(5\%, t_3)$		$CN(10\%, t_3)$		$CN(15\%, t_3)$		$CN(20\%, t_3)$	
δ	M	TMd	M	TMd	M	TMd	M	TMd
0.0	499.3	500.0	499.8	500.0	499.8	500.0	500.0	499.8
0.1	455.8	401.6	474.2	411.4	488.8	422.1	491.6	424.8
0.2	351.1	252.1	408.0	260.8	449.2	279.3	460.2	286.6
0.3	239.3	142.9	333.6	148.8	391.5	161.9	408.7	167.1
0.4	149.3	79.7	247.3	84.1	325.6	91.1	346.1	95.8
0.5	88.6	46.5	170.4	48.3	252.8	52.1	278.4	55.0
0.6	52.8	28.3	109.4	28.9	186.4	30.9	210.7	32.3
0.7	32.1	17.8	68.4	18.0	130.8	19.0	156.1	19.6
0.8	20.1	11.7	42.2	11.6	86.2	12.1	107.3	12.3
0.9	12.9	7.9	26.3	7.8	55.1	8.1	71.5	8.1
1.0	8.7	5.6	16.6	5.5	34.7	5.6	46.5	5.6
1.1	6.0	4.1	10.9	4.0	22.0	4.0	29.7	4.0
1.2	4.4	3.1	7.4	3.0	14.2	3.0	18.9	3.0
1.3	3.3	2.5	5.2	2.4	9.4	2.4	12.3	2.3
1.4	2.6	2.2	3.8	2.0	6.4	1.9	8.2	1.9
1.5	2.1	1.7	2.9	1.7	4.6	1.6	5.7	1.6
2.0	1.2	1.1	1.3	1.1	1.5	1.1	1.6	1.1
LCL	-1.5958	-1.4794	-1.8388	-1.5234	-2.0810	-1.5818	-2.1931	-1.6109
UCL	1.5972	1.4805	1.8286	1.5252	2.0792	1.5808	2.2262	1.6280

5 Concluding remarks

The main conclusions of this study can be summarized as follows. The sample mean is not the most efficient/robust estimator for the process mean value of distributions with heavier-than-normal tails or even when we have small disturbances in potential normal data. Consequently, the mean chart is not appropriate to monitor the process mean value in many practical situations, and we suggest the use of the TMd chart as an alternative, for instance to detect small-to-moderate changes in the process mean value of contaminated normal data.

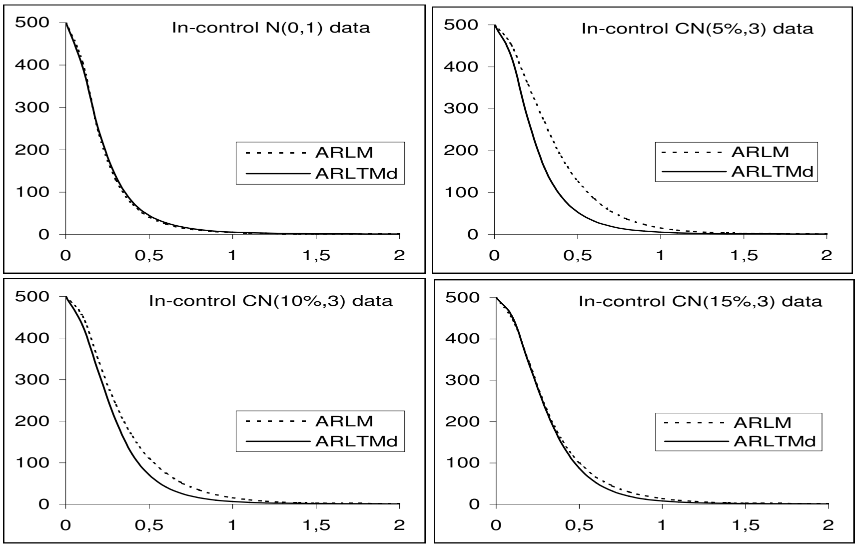

Figure 2: ARL estimates for normal and scaled-contaminated normal data.

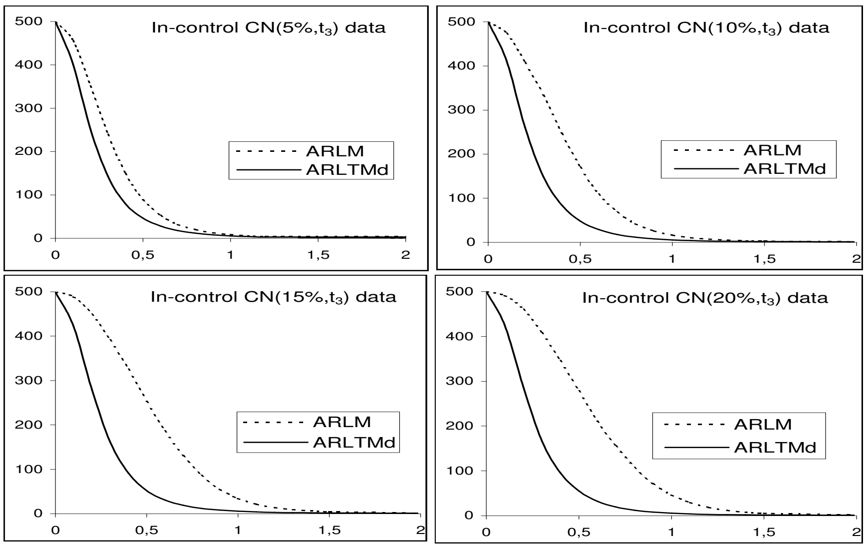

Figure 3: ARL estimates for student-contaminated normal data.

References

[1] Azzalini A. A Class of distributions which includes the normal Ones. *Scandinavian J. of Statistics* 1985, **12**: 171–178.

[2] Azzalini A. Further results on a class of distributions which includes the normal ones. *Statistica* 1986, **XLVI**: 199–208.

[3] Azzalini A. The skew-normal distribution and related multivariate families. *Scandinavian J. of Statistics* 2005, **32**: 159–188.

[4] Bai DS, Choi IS. $\overline{X}$ and R control charts for skewed populations, *J. of Quality Technology* 1995, **27**(2): 120-131.

[5] Balakrishnan N, Kocherlakota S. Effects of nonnormality on $\overline{X}$ charts: single and assignable cause model. *Sankhyã: The Indian Journal of Statistics* 1986, **48**(Series B): 439-444.

[6] Brownie C, Habicht J-P, Robson DS. An estimation procedure for the contaminated normal distributions arising in clinical chemistry. *JASA* 1983, **78**382: 228-237.

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

[8] Castagliola P. *Recent Advances in Reliability and Quality Engineering*, World Scientific Publishers, chapter Control Charts for Data Having a Symmetrical Distribution with a Positive Kurtosis, 2001, 1-16.

[9] Chan LK, Hapuarachchi KP, Macpherson BD. Robustness of $\overline{X}$ and R charts. *IEEE Transactions on Reliability* 1988, **37**(1): 117-123.

[10] Chan LK, Heng JK. Skewness Correction $\overline{X}$ and R charts for skewed distributions. *Naval Research Logistics* 2003, **50**: 555-573.

[11] Chakraborti S, Human SW, Graham MA. Phase I Statistical Process Control Charts: An Overview and Some Results. *Quality Engineering* 2009, **21**(1): 52-62.

[12] Chakraborti S, Human SW, Graham MA. Nonparametric (distribution-free) quality control charts. In N. Balakrishnan, Ed., *Methods and Applications of Statistics: Engineering, Quality Control, and Physical Sciences*, 2011, 298-329.

[13] Chen J, Tan X, Zhang R. Inference for normal mixtures in mean and variance. *Statist. Sinica* 2008, **18**: 443-465.

[14] Cox MG, Iguzquiza EP. The total median and its uncertainity. In Ciarlini et al. (eds.), *Advanced Mathematical and Computational Tools in Metrology* 2001, **5**: 106-117.

[15] Figueiredo F, Gomes MI. The total median is Statistical Quality Control. *Applied Stochastic Models in Business and Industry* 2004, **20**: 339-353.

[16] Figueiredo F, Gomes MI. Monitoring industrial processes with robust control charts. *Revstat – Statistical Journal* 2009, **7**(2): 151-170.

[17] Figueiredo F, Gomes MI. The skew-normal distribution in SPC. *Revstat – Statistical Journal* 2013 (in press).

[18] Gleason JR. Understanding Elongation: The Scale Contaminated Normal Family. *JASA* 1993, **88**: 327-337.

[19] Ghosh D, Chinnaiyan AM. Genomic outlier profile analysis: Mixture models, null hypotheses and nonparametric estimation. *Biostatistics* 2009, **10**: 60-69.

[20] Hoaglin DM, Mosteller F, Tukey, JW. *Understanding Robust and Exploratory Data Analysis*, Wiley, New York, 1983.

[21] Human SW, Chakraborti S. A Unified Approach for Shewhart-Type Phase I Control Charts for the Mean. *International Journal of Reliability, Quality and Safety Engineering* 2010, **17**(3): 199-208.

[22] Jensen WA, Jones-Farmer LA, Champ CH, Woodall WH. *Effects of Parameter Estimation on Control Chart Properties: A Literature Review. J. Quality Technology* 2006, **38**: 349-364.

[23] Langenberg P, Iglewicz B. Trimmed Mean $\overline{X}$ and R Charts. *J. Quality Technology* 1986, **18**3: 152-161.

[24] Liang Ky, Rathouz P. Hypothesis testing under mixture models: application to genetic linkage analysis. *Biometrics* 1999, **55**: 65-74.

[25] McLaren CE. Mixture models in hematology: A series of case studies. *Stat. Methods Med. Res.* 1996, **5**: 129-153.

[26] Rocke DM. Robust control charts. *Technometrics* 1989, **31**(2): 173-184.

[27] Rocke DM. $\overline{X}_Q$ and R_Q charts: robust control charts. *The Statistician* 1992, **41**: 97-104.

[28] Schilling EG, Nelson PR. The Effect of Non-Normality on the Control Limits of the $\overline{X}$ Charts. *J. Quality Technology* 1976, **8**(4): 183-188.