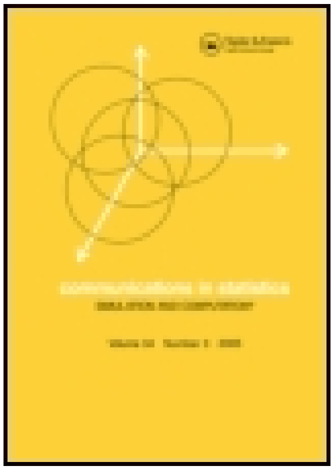

Communications in Statistics - Simulation and Computation
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lssp20

Median Control Charts Based on Bootstrap Method

Hyo-Il Park [a]

[a] Department of Statistics , Chong-ju University , Chong-ju, Choong-book, South Korea

Published online: 06 Feb 2009.

Communications in Statistics—Simulation and Computation®, 38: 558–570, 2009

ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1080/03610910802585824

Median Control Charts Based on Bootstrap Method

HYO-IL PARK

Department of Statistics, Chong-ju University, Chong-ju,
Choong-book, South Korea

In this study, we propose a median control chart. In order to determine the control limits, we consider using an estimate of the variance of sample median. Also, we consider applying the bootstrap methods. Then we illustrate the proposed median control chart with an example and compare the bootstrap methods by simulation study. Finally, we discuss some peculiar features for the median control chart as concluding remarks.

Keywords Bootstrap method; Confidence interval; Median.

Mathematics Subject Classification 62P30; 62F40.

1. Introduction

In order to enhance and maintain the quality of products on the production floor, the control charts have played substantial roles in the quality control. The $\overline{X}$-chart which is widely used to monitor the process mean quality level can be characterized as follows (cf. Khoo, 2005):

(i) it is constructed and maintained under the assumption that the distribution of the production process is normal;
(ii) it is sensitive to the outliers.

The sensitivity of the $\overline{X}$-chart due to the outliers may disappear with the increase of the number of observations (sample size). Then, this in turn may incur some economic burden such as the increase of sampling cost. Also, many statisticians and engineers agree with the idea that it is not rare to encounter the data with heavy-tailed or skewed distribution on the floor (cf. Gunter, 1989). Also, they consider that the occurrence of the outliers might be mainly due to the digress from the normality (cf. Alwan, 2000). In this way or another, when the assumption for the normality in the process may be questionable, some alternative statistical procedure, which should be robust in some sense may be considered. Among them,

Received November 30, 2007; Accepted October 28, 2008
Address correspondence to Hyo-Il Park, Department of Statistics, Chong-ju University, Chong-ju, Choong-book 360-764, South Korea; E-mail: hipark@cju.ac.kr

the nonparametric procedure has been most successfully developed and widely applied. For the nonparametric method, the median, as a location parameter, is used instead of the mean in order to provide appropriate inferential tools for the data analysis under the various underlying distributions such as the Cauchy distribution. In this study, we consider to construct a control chart based on the median with the assumption that the process distribution is not known but continuous. In fact, the control chart based on the median is not a brand-new subject. Since the existing median control charts contain the intrinsic drawbacks which we will review later, they have hardly been used in the production floor for the real applications. Now we review some selected papers which are related to the median control charts in the following.

1.1. *Control Charts Based on Sample Median*

Nelson (1982), Janacek and Meikle (1997), and Khoo (2005) proposed the median control charts using sample median for the center line. Nelson (1982) considered the case that the underlying distribution is normal and so could obtain the $3-\sigma$ control limits easily. Janacek and Meikle (1997) derived the beta-binomial distribution as a distribution of sample median using the so-called double expectation theorem (cf. Bickel and Doksum, 1977). However, as the authors pointed out, the form of the beta-binomial distribution is awkward to deal with. Moreover, because the beta-binomial distribution is discrete, it is not easy to obtain the exact $3-\sigma$ control limits. Also, Khoo (2005) considered obtaining the control limits using the distribution for sample median when the underlying distribution is normal with the variable transformation technique. The main embedded obstacle for the application of the median control charts is not easy to decide the control limits without some specific distributional assumption since the distribution of sample median contains the underlying distribution function (df) and the corresponding probability density function (pdf). Even though one can take the asymptotic approach with the large sample approximation theory, the limiting variance still contains its pdf. All the reviewed results have tried to evade this obstacle but not been much successful.

1.2. *Control Chart Using the Hodges–Lehmann Estimate and Related Interval Estimate*

Alloway and Raghavachari (1991) used the Hodges–Lehmann estimate of the median for the center line and related nonparametric interval estimate for the control limits based on the Wilcoxon signed-rank statistic. Since the distribution of the Wilcoxon signed-rank statistic is also discrete, it is not easy to obtain the exact $3-\sigma$ control limits. Moreover, this median control chart requires the minimal sample size to achieve its $3-\sigma$ control limits.

1.3. *Control Charts Using the Quantile Function*

Grimshaw and Alt (1997) proposed a control chart using the quantile function. The proposed control chart is different from the traditional control charts in the sense that their method checks on the changes of the process through the approach of the hypotheses test. Therefore, one can hardly visualize any change of the process directly and so may use the control chart for the goodness-of-fit test rather than the detection of changes of the process. For this reason, it is hard to apply the control chart to the production floor. Kanji and Arif (2000) proposed the median rankit control chart based on the logistic quantile distribution, which is a generalized form of the lambda distribution (cf. Gilchrist, 1997). The control limits including the center line can be obtained directly from the quantile function. However, this approach involves some estimation procedure for the parameters which characterize the logistic quantile distribution. Also, Kanji and Arif (2001) considered the application of the median rankit control chart to the Weibull data.

Also, there have been several nonparametric control charts based on nonparametric statistics (cf. Bakir, 2006) and order statistics (cf. Albers and Kallenberg, 2006) for the center line or control limits. Chakraborti et al. (2004) considered the performance of the median control charts by deriving the average run length (ARL) and false alarm rate (FAR) with emphasizing the fact that the ARL and FAR are the same for all continuous distribution under the nonparametric approach. Further, Chakraborti and Eryilmaz (2007) generalized Bakir's nonparametric control chart with incorporating some run rules. Gupta and Sengupta (2008) considered control charts based on repeated median filters to remain unaffected due to the non normality and outliers. Therefore, it would be worthwhile to propose a new median control chart. In this study, since we assume that df is unknown, a possible and simple way to wade through this situation is that we may apply the normal approximation with variance estimation as an asymptotic approach. This subject will be taken in the next section. However, for the estimation of the variance of sample median, we consider the estimate proposed by Maritz and Jarrett (1978) rather than that of the limiting form. Since the estimate does not involve estimation of the unknown density but is based on the estimates of moments, one may expect that the estimate by Maritz and Jarrett (1978) should be more accurate than that of limiting form.

Since Efron (1979) introduced the bootstrap method, it has been developed and evolved in various ways to provide powerful tools for the statistical inferences including the interval estimation and hypothesis test. Also, the applications have been expanded into almost all the statistical fields. However, surprisingly enough, the application to the control charts or quality control has rarely been carried out. Seppala et al. (1995) considered the application of the bootstrap method to the $\overline{X}$-chart to obtain the control limits for the skewed distribution using the bootstrap percentile (BP) method. Liu and Tang (1996) considered the application to the dependent process. Jones and Woodall (1998) compared the performance of the several bootstrap control charts. We note that all the results for the control charts with the bootstrap method up to now have been confined to the $\overline{X}$-chart. Also very recently, Teyarachkul et al. (2007) considered a modified the bootstrap method, which they called as a direct method for obtaining control limits of the $\overline{X}$-chart but the form of data should be integer. Lio and Park (2008) proposed the percentile control charts by applying the bootstrap method under the Birnbaum–Saunders distribution.

In this study, first of all, we consider the construction of a median control chart based on the estimate of the variance of sample median. Then we review several bootstrap methods for obtaining confidence intervals to decide the control limits for the median control chart. Then we illustrate the proposed median control charts using the data of Ford Motor Company (cf. Alloway and Raghavachari, 1991) and compare the bootstrap methods through simulation study with three kinds of distributions. Finally, we discuss some peculiar features for the median control charts as concluding remarks.

2. Median Control Chart

Suppose that we have a sample $X_1, \ldots, X_n$ from a production process with an unknown continuous df F with median θ. Then a sample median $\widetilde{X} = med\{X_1, \ldots, X_n\}$ can be defined as:

$$\widetilde{X} = \begin{cases} X_{(k+1)} & \text{if } n = 2k+1 \\ (X_{(k)} + X_{(k+1)})/2 & \text{if } n = 2k \end{cases},$$

where $X_{(k)}$ is the kth order statistic from $X_1, \ldots, X_n$. When $n = 2k+1$, the pdf of $\widetilde{X}$, $g(x)$, is of the form

$$g(x) = \frac{n!}{(k!)^2}[F(x)]^k[1 - F(x)]^k f(x), \tag{1}$$

where f is the corresponding pdf of F. Then one could obtain the control limits using g if F and f were fully known. Since we assumed that F is unknown, we cannot derive the control limits using g directly. Instead, one may take an asymptotic approach as follows. We note that the limiting distribution of $\sqrt{n}(\widetilde{X} - \theta)$ is normal with mean 0 and variance $1/(4f^2(\theta))$. If $\hat{f}(\widetilde{X})$ is a consistent estimate of $f(\theta)$, one may obtain the lower and upper control limits (LCL and UCL) for the $3 - \sigma$ control limit as:

$$LCL = \widetilde{X} - 3/(2\sqrt{n}\hat{f}(\widetilde{X})) \quad \text{and} \quad UCL = \widetilde{X} + 3/(2\sqrt{n}\hat{f}(\widetilde{X})).$$

We note that the estimate $1/(4n\hat{f}^2(\widetilde{X}))$ is that of the limiting variance of $\widetilde{X}$. If we can estimate the exact form of the variance of $\widetilde{X}$ for each n, the resulting estimate would be expected to be more accurate than $1/(4n\hat{f}^2(\widetilde{X}))$. In this direction, Maritz and Jarrett (1978) have proposed an estimate for the variance of $\widetilde{X}$ as follows. Since

$$V(\widetilde{X}) = E(\widetilde{X}^2) - [E(\widetilde{X})]^2,$$

if one can estimate $E(\widetilde{X}^d)$ consistently, $d = 1, 2$, for each n, one may obtain a consistent estimate $\widehat{V}(\widetilde{X})$ in some sense. For this, we define the quantile function F^{-1} of F as follows: for $0 < u < 1$, let

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}.$$

Using this notation, we have with (1) that for $n = 2k+1$

$$\begin{aligned} E(\widetilde{X}^d) &= \int_{-\infty}^{\infty} x^d g(x)dx \\ &= \sum_{i=1}^{n} \frac{n!}{(k!)^2} \int_{(i-1)/n}^{i/n} (F^{-1}(u))^d u^k (1-u)^k du. \end{aligned} \tag{2}$$

Equation (2) has been obtained through the variable transformation technique. Let F_n be the empirical distribution function based on $X_1, \ldots, X_n$. Also, let F_n^{-1} be the corresponding sample quantile function of F_n. Then we note that for any u such as $(i-1)/n < u \le i/n$,

$$F_n^{-1}(u) = \inf\{x : F_n(x) \ge u\} = X_{(i)}, \tag{3}$$

where $X_{(i)}$ is the ith order statistic from $X_1, \ldots, X_n$, $i = 1, \ldots, n$. Then by substituting F_n^{-1} for F^{-1} in (2), we have the following estimates of $E(\widetilde{X}^d)$, $d = 1, 2$:

$$\begin{aligned}\widehat{E}(\widetilde{X}^d) &= \sum_{i=1}^{n} \widetilde{X}_{(i)}^d \frac{n!}{(k!)^2} \int_{(i-1)/n}^{i/n} u^k (1-u)^k du \\ &= \sum_{i=1}^{n} \widetilde{X}_{(i)}^d W_i, \quad \text{say.}\end{aligned}$$

Then some algebraic manipulation for the estimate of variance, $\widehat{V}(\widetilde{X}) = \widehat{E}(\widetilde{X}^2) - [\widehat{E}(\widetilde{X})]^2$, leads the following quadratic form:

$$\widehat{V}(\widetilde{X}) = (X_{(1)}, \ldots, X_{(n)}) Q_n (X_{(1)}, \ldots, X_{(n)})^T, \tag{4}$$

where $(\cdot)^T$ is the transpose of a vector and $Q_n = (q_{ij})_{i,j=1,\ldots,n}$ is the $n \times n$ matrix with

$$q_{ii} = W_i - W_i^2$$

and

$$q_{ij} = q_{ji} = -W_i W_j.$$

For $n = 2k$, since $\widetilde{X} = (X_{(k)} + X_{(k+1)})/2$, we note that

$$V(\widetilde{X}) = \frac{1}{4} E(X_{(k)}^2 + X_{(k+1)}^2) + \frac{1}{2} E(X_{(k)} X_{(k+1)}) - \frac{1}{4}\{E(X_{(k)} + X_{(k+1)})\}^2. \tag{5}$$

Then using the two marginal pdf's and a joint pdf for $X_{(k)}$ and $X_{(k+1)}$, one may obtain estimates of the three components in (5). Also, the same algebraic manipulation yields again (4). In this case, $n = 2k$, the components of $Q_n = (q_{ij})_{i,j=1,\ldots,n}$ become as

$$\begin{aligned}q_{ii} = U_i - U_i^2 - \frac{n!}{4(k!)^2}\Bigg[&\left(\frac{i}{n}\right)^k \left(\frac{n-i}{n}\right)^k + \left(\frac{i-1}{n}\right)^k \left(\frac{n+1-i}{n}\right)^k \\ &-2\left(\frac{i-1}{n}\right)^k \left(\frac{n-i}{n}\right)^k\Bigg]\end{aligned}$$

and

$$q_{ij} = q_{ji} = -U_i U_j + \frac{n!}{4(k!)^2}\left[\left(\frac{i}{n}\right)^k - \left(\frac{i-1}{n}\right)^k\right]\left[\left(\frac{n+1-j}{n}\right)^k - \left(\frac{n-j}{n}\right)^k\right] \quad (i < j),$$

where $U_i = \frac{n!}{k!(k-1)!} \int_{(i-1)/n}^{i/n} u^{k-1}(1-u)^{k-1} du$, $i = 1, \ldots, n$.

With the assumption that $E(\widetilde{X}^2) < \infty$, it is easy to show that

$$(\widetilde{X} - \theta)/\sqrt{\widehat{V}(\widetilde{X})} \to N(0, 1)$$

in distribution. The derivation will appear briefly in the Appendix. Therefore, one may obtain the LCL and UCL under $3 - \sigma$ control limits as:

$$LCL = \widetilde{X} - 3\sqrt{\widehat{V}(\widetilde{X})} \quad \text{and} \quad UCL = \widetilde{X} + 3\sqrt{\widehat{V}(\widetilde{X})}, \tag{6}$$

respectively. Maritz and Jarrett (1978) provided the values of U_i's and W_i's with varying n from 3–20. Also, we note that the estimate $\widehat{V}(\widetilde{X})$ coincides with the bootstrap estimate (cf. Shao and Tu, 1995). Therefore, one may consider the LCL and UCL in (6) as bootstrap control limits. This method is known as the bootstrap standard (BS) method. In the next section, we consider to construct the median control charts by obtaining the control limits using the various bootstrap methods.

3. Bootstrap Median Control Charts

Already, we had a bootstrap median control chart using the BS method since the estimate, $\widehat{V}(\widetilde{X})$ coincides with the bootstrap estimate. Before we further proceed the discussion of construction of the other bootstrap median control charts, we review some bootstrap re-sampling procedure briefly, which is common for all the bootstrap methods. Any bootstrap method considers the sample $X_1, \ldots, X_n$ as a population and re-samples from $X_1, \ldots, X_n$ with replacement. Then for the exact bootstrap distribution, we have to consider all the n^n number of configurations of the bootstrap re-samples. Since it may be impossible to consider all the n^n re-sampling configurations even for the moderate sample sizes, one should take the Monte-Carlo approach. Hence, all the bootstrap control limits, which we will consider later, will be obtained by the approximation of the bootstrap distribution based on the Monte-Carlo approach. Now we state the Monte-Carlo bootstrap procedure for obtaining an approximate bootstrap distribution for sample median as follows:

(i) from the original sample $X_1, \ldots, X_n$, we obtain a bootstrap sample $X_1^*, \ldots, X_n^*$ with replacement;
(ii) from the bootstrap sample, $X_1^*, \ldots, X_n^*$, we obtain bootstrap sample median $\widetilde{X}^*$;
(iii) we repeat the above two steps (i) and (ii) with B times.

The number B will be usually a big one such as at least 1,000. Then we have B number of bootstrap sample medians,

$$\widetilde{X}_1^*, \widetilde{X}_2^*, \ldots, \widetilde{X}_B^*.$$

We note that a bootstrap estimate $\widehat{V}^*(\widetilde{X}^*)$ of $V(\widetilde{X})$ can be obtained from $\widetilde{X}_1^*, \widetilde{X}_2^*, \ldots, \widetilde{X}_B^*$ with $\overline{\widetilde{X}}^* = \frac{1}{B}\sum_{b=1}^{B} \widetilde{X}_b^*$ as

$$\widehat{V}^*(\widetilde{X}^*) = \frac{1}{B}\sum_{b=1}^{B}\left(\widetilde{X}_b^* - \overline{\widetilde{X}}^*\right)^2. \tag{7}$$

Also, we note that $\lim_{B\to\infty} \widehat{V}^*(\widetilde{X}^*) = \widehat{V}(\widetilde{X})$. Therefore, $\widehat{V}^*(\widetilde{X}^*)$ is an approximate of $\widehat{V}(\widetilde{X})$ based on the Monte-Carlo approach (cf. Shao and Tu, 1995).

Now we review several bootstrap methods to obtain confidence intervals for the median based on $\widetilde{X}_1^*, \widetilde{X}_2^*, \ldots, \widetilde{X}_B^*$ to determine the control limits. For this subject, Efron and Tibshirani (1993) and Shao and Tu (1995) provided excellent basis with concrete materials and explained systematically all the bootstrap methods, which we will review. In the following, the empirical bootstrap distributions based on $\widetilde{X}_1^*, \widetilde{X}_2^*, \ldots, \widetilde{X}_B^*$ will be used as the estimates of the corresponding sample distributions. Then one may wish to distinguish the exact bootstrap distribution from approximate one. However the distinction will be neglected in this study since the two concepts can be used interchangeably in the applied field without any confusion and moreover the Monte-Carlo approach is a basic tool for applying the bootstrap method.

3.1. *Bootstrap Percentile (BP) Method*

The BP method, which is easy to handle and so widely used, can be applied in the following manner. We note that to obtain a confidence interval for θ based on $\widetilde{X}$, we need the distribution

$$G_n(x) = \Pr\{\widetilde{X} \le x\}.$$

Then the bootstrap estimate G_{boot} of G_n can be approximated by:

$$G_{boot}(x) = \frac{1}{B}\sum_{b=1}^{B} I(\widetilde{X}_b^* \le x),$$

where $I(\cdot)$ is the indicate function. Thus, one may obtain the $(1-2\alpha)\times 100\%$ BP confidence interval for θ as follows:

$$(G_{boot}^{-1}(\alpha), G_{boot}^{-1}(1-\alpha)), \tag{8}$$

where G_{boot}^{-1} is the quantile function of G_{boot}. From the relation between quantiles and order statistics as in (3), the BP confidence interval in (8) can be re-expressed directly by using ordered bootstrap medians as follows:

$$(\widetilde{X}^*_{([\alpha\times B])}, \widetilde{X}^*_{([(1-\alpha)\times B])}),$$

where $\widetilde{X}^*_{(i)}$ denotes the ith ordered bootstrap median among $\widetilde{X}_1^*, \widetilde{X}_2^*, \ldots, \widetilde{X}_B^*$ and $[c]$ stands for the largest integer part, which does not exceed the real number c.

3.2. *Bootstrap Bias-Corrected Percentile (BBCP) Method*

Even though the BP method is easy to use and widely applied, some severe critics have been drawn because of the possibility of the intrinsically embedded bias (cf. Hall, 1988). In order to take the bias into consideration, a modified percentile bootstrap method has been proposed and is called as BBCP method. Efron (1982) considered the following assumption to accommodate a bias term b:

$$\Pr\left\{\frac{\widetilde{X}-\theta+b}{\sqrt{V(\widetilde{X})}} \le x\right\} = \Phi(x),$$

where Φ is the cumulative standard normal df. Let $b_0 = b/\sqrt{V(\widetilde{X})}$. Then one could obtain the lower bound $\underline{\theta}$ as

$$\underline{\theta} = \widetilde{X} + (b_0 + z_\alpha)\sqrt{V(\widetilde{X})}$$

if b_0 were known, where $\Phi(z_\alpha) = \alpha$. Then in order to obtain b_0, we note that $\Phi(b_0)$ can be estimated by $G_{boot}(\widetilde{X})$ since

$$\Phi(b_0) = \Pr\left\{\frac{\widetilde{X}-\theta}{\sqrt{V(\widetilde{X})}} + b_0 \le b_0\right\} = \Pr\{\widetilde{X} \le \theta\} = G_n(\theta).$$

Then an estimate $\hat{b}_0$ of b_0 can be obtained by

$$\hat{b}_0 = \Phi^{-1}(G_{boot}(\widetilde{X})), \tag{9}$$

where G_{boot} was defined in the previous subsection. Also, we note that since

$$1-\alpha = \Phi(-z_\alpha) = G_{boot}\big(\widetilde{X} - (b_0 + z_\alpha)\sqrt{V(\widetilde{X})}\big),$$

we have

$$G_{boot}^{-1}(1-\alpha) = \widetilde{X} - (b_0 + z_\alpha)\sqrt{V(\widetilde{X})}. \tag{10}$$

Since α was arbitrarily chosen in (10), for any $0 < x < 1$, we have

$$G_{boot}^{-1}(x) = \widetilde{X} - (b_0 - \Phi^{-1}(x))\sqrt{V(\widetilde{X})}.$$

By taking $\Phi^{-1}(x) = 2b_0 + z_\alpha$, we see that:

$$G_{boot}^{-1}(\Phi(2b_0 + z_\alpha)) = \widetilde{X} + (b_0 + z_\alpha)\sqrt{V(\widetilde{X})} = \underline{\theta}.$$

Thus with (9), we have that:

$$\underline{\theta} \approx G_{boot}^{-1}(\Phi(2\Phi^{-1}(G_{boot}(\widetilde{X})) + z_\alpha)).$$

Also for the upper bound, using the same steps with the lower case, we see that

$$\overline{\theta} \approx G_{boot}^{-1}(\Phi(2\Phi^{-1}(G_{boot}(\widetilde{X})) + z_{1-\alpha})).$$

Then we have the following $(1-2\alpha)\times 100\%$ BBCP confidence interval:

$$(G_{boot}^{-1}(\Phi(2\Phi^{-1}(G_{boot}(\widetilde{X}))+z_\alpha)), G_{boot}^{-1}(\Phi(2\Phi^{-1}(G_{boot}(\widetilde{X}))+z_{1-\alpha}))).$$

3.3. *Bootstrap-t (BT) Method*

As another bootstrap method for deriving a confidence interval for θ, the BT method is also widely used. However, this method requires a consistent estimate for the variance of $\widetilde{X}$. Then using the consistent estimate $\widehat{V}(\widetilde{X})$, let

$$H_n(x) = \Pr\left\{(\widetilde{X}-\theta)/\sqrt{\widehat{V}(\widetilde{X})} \le x\right\},$$

be the distribution of the studentized form for $\widetilde{X}$. Then one may obtain a bootstrap estimate H_{boot} for H_n by

$$H_{boot}(x) = \Pr\left\{(\widetilde{X}^*-\widetilde{X})/\sqrt{\widehat{V}^*(\widetilde{X}^*)} \le x\right\},$$

where $\widehat{V}^*(\widetilde{X}^*)$ is a bootstrap estimate of $V(\widetilde{X})$ as in (7). Let H_{boot}^{-1} be the quantile function of H_{boot}. Then $(1-2\alpha)\times 100\%$ BT confidence interval for θ becomes

$$\left(\widetilde{X} - H_{boot}^{-1}(1-\alpha)\sqrt{\widehat{V}(\widetilde{X})}, \widetilde{X} - H_{boot}^{-1}(\alpha)\sqrt{\widehat{V}(\widetilde{X})}\right).$$

3.4. *Bootstrap Hybrid (BH) Method*

As another bootstrap method, the BH method is easy to use and known as relatively accurate. In particular, BH method does not require any estimate of the variance but is based on a similar approximate pivotal quantity like the BT method. Let

$$K_n(x) = \Pr\left\{\sqrt{n}(\widetilde{X}-\theta) \le x\right\}.$$

Then it is obvious that

$$K_{boot}(x) = \Pr\left\{\sqrt{n}(\widetilde{X}^*-\widetilde{X}) \le x\right\}$$

is the bootstrap approximation of $K_n(x)$. Also, let K_{boot}^{-1} be the quantile function of K_{boot}. Since for any $0<\alpha<1$

$$\Pr\left\{\sqrt{n}(\widetilde{X}-\theta) \le K_{boot}^{-1}(\alpha)\right\} \approx \alpha,$$

the $(1-2\alpha)\times 100\%$ BH confidence interval for θ is

$$\left(\widetilde{X} - K_{boot}^{-1}(1-\alpha)/\sqrt{n}, \widetilde{X} - H_{boot}^{-1}(\alpha)/\sqrt{n}\right).$$

In passing, we note that the term *hybrid* comes from the points of view that the approximate pivotal quantity $\sqrt{n}(\widetilde{X}-\theta)$ resembles the studentized form of the BT method and the use of the percentiles of K_{boot} does the BP method. Hence it may be called as the bootstrap hybrid method.

Using one of the bootstrap methods which have been reviewed up to now, one may determine the control limits in the following way. For this we assume that we have r number of sub-samples with sample size n each.

(i) Determine the center line (CL) by $CL = med\{\widetilde{X}_i, \ldots, \widetilde{X}_r\}$, where $\widetilde{X}_i$ is sample median of the ith sub-sample.
(ii) Obtain r number of the 99.74% bootstrap confidence intervals (L_i^*, U_i^*), $i = 1, \ldots, r$ for the $3-\sigma$ control limits.
(iii) Then LCL and UCL can be chosen as:

$$LCL = med\{L_1^*, \ldots, L_r^*\} \quad \text{and} \quad UCL = med\{U_1^*, \ldots, U_r^*\}.$$

4. An Example and Simulation Study

In order to give an illustration for the proposed control charts, we consider the data of the primary thickness problem from the Ford Motor Company (cf. Alloway and Raghavachari, 1991). All the results are summarized in Table 1 for the $3-\sigma$ control limits. In the tables, EXACT means that the BS method uses the estimate $\widehat{V}(\widetilde{X})$ in (4), while APPRO implies that the BS method uses the bootstrap estimate $\widehat{V}^*(\widetilde{X}^*)$ in (7). Therefore, the results of the EXACT and APPRO should be close but appear somewhat different. However, the difference seems negligible compared with other bootstrap methods. In this example, the APPRO method yields the shortest control limits but BT method does the longest control limits. The control chart proposed by Alloway and Raghavachari (1991) gave 1.119 and (0.950, 1.310) as its CL and (LCL, UCL), respectively. Also, we note that the $\overline{X}$-chart under the normality assumption yields 1.121 and (1.014, 1.230) as CL and (LCL, UCL), respectively. Therefore, the proposed median control chart can be considered as a compromise between the $\overline{X}$-chart and median control chart of Alloway and Raghavachari (1991).

Also for the comparison of the performance among the bootstrap methods, we carried out the following simulation study by obtaining the $3-\sigma$ control limits. For this purpose, we considered three different distributions, normal, exponential and Cauchy with sub-sample size 10 for each case. All the simulated results are summarized in Tables 2–4. In Table 4, l and s stand for the parameters of the location and scale of the Cauchy distribution, respectively. In this study, a peculiar feature is that the BT method also shows relatively longer intervals of control limits for all three cases. In the normal case, the BH method achieves the shortest control

Table 1
$3-\sigma$ control limits for various bootstrap methods

Bootstrap method		CL	LCL	UCL
BS	EXACT	1.1225	0.995	1.246
	APPRO		1.004	1.241
BP			0.998	1.255
BBCP			1.010	1.283
BT			0.945	1.343
BH			0.963	1.250

Table 2
Standard normal distribution

Bootstrap method		*LCL*	*UCL*
BS	EXACT	−1.145	1.113
	APPRO	−1.133	1.120
BP		−1.088	1.087
BBCP		−0.979	1.197
BT		−1.805	1.786
BH		−1.073	1.077

Table 3
Exponential distribution with $\lambda = 1$

Bootstrap method		*LCL*	*UCL*
BS	EXACT	−0.204	1.658
	APPRO	−0.168	1.641
BP		0.164	2.023
BBCP		0.177	2.235
BT		−0.561	2.575
BH		−0.533	1.210

Table 4
Cauchy distribution with $l = 0$ and $s = 1$

Bootstrap method		*LCL*	*UCL*
BS	EXACT	−2.297	2.293
	APPRO	−2.447	2.430
BP		−2.665	2.449
BBCP		−1.767	3.150
BT		−2.937	2.938
BH		−2.638	2.790

limits. However for the overall performance, the BS method behaves relatively well. Finally, we note that the BP and BBCP methods yield realistic results for the exponential distribution since *LCL*'s are non negative values. This phenomenon comes from the fact that control limits for BP and BBCP cases are quantiles of the bootstrap empirical distribution comprised with bootstrap medians.

5. Some Concluding Remarks

If we take the traditional nonparametric approach, we note that the distributions of the corresponding statistics are all discrete type. Thus, we may have difficulty when we try to obtain exact quantile point for any given probability. Whereas as we have already seen, if we apply any one of the bootstrap methods, we may obtain relatively close quantile points by increasing the number of re-sampling. Therefore, the requirement of the minimal sample size disappear for any bootstrap method. Also, we note that except the BS method, we obtained control limits from the bootstrap empirical distribution. Therefore, especially the bootstrap method may be appropriate to apply when the underlying distribution is skewed.

When we considered a median control chart, we were concerned about the estimate of variance of $\widetilde{X}$. Since $1/(4n\hat{f}^2(\widetilde{X}))$ is the estimate of the limiting form of the variance of $\widetilde{X}$, the accuracy may be somehow questionable when the sample size n is small. On the other hand, since $\widehat{V}(\widetilde{X})$ is the estimate of $V(\widetilde{X})$ for each n, one may expect that $\widehat{V}(\widetilde{X})$ is more accurate than $1/(4n\hat{f}^2(\widetilde{X}))$, especially when n is small. Also, as we have seen in the previous example, the median control chart using $\widehat{V}(\widetilde{X})$had shorter control limits than the one proposed by Alloway and Raghavachari (1991). In passing we note that $\widehat{V}(\widetilde{X})$ is a moment estimate and can be considered a weighted estimate assigning more weight to the observation close to $\widetilde{X}$.

We have reviewed several bootstrap methods for the construction of the confidence interval for the median. However, we did not review the bootstrap methods in an exhaustive manner. Therefore, there remain more bootstrap methods such as the bootstrap accelerated bias-corrected percentile (BABCP) method. This bootstrap method includes estimation of the highly technical extra parameter, the acceleration constant, which takes account of the skewness. However, the improvement of BABCP over the simple BP method relies on the smoothness of the estimate of the interested parameter. Especially, Hall and Martin (1989) showed that sample median cannot be improved through the BABCP method since sample median is a non smooth estimate.

From the simulation results, we note that even though the normal and Cauchy distributions are symmetric, the BBCP method yielded very asymmetric control limits. Therefore, one should be careful when one chooses a bootstrap method for the real use. Also, it would be interesting to consider the extension of this median control chart to the vector case. Then this subject will be one of our future research topics.

Appendix

In this Appendix, we derive the asymptotic normality of $(\widetilde{X}-\theta)/\sqrt{\widehat{V}(\widetilde{X})}$ in a descriptive manner. With the assumption that $E(\widetilde{X}^2) < \infty$, first of all, we note that

$$(\widetilde{X}-\theta)/\sqrt{V(\widetilde{X})} \rightarrow N(0, 1)$$

in distribution (cf. Bickel and Doksum, 1977). Also, we note that if $E(\widetilde{X}^2) < \infty$, $\widehat{V}(\widetilde{X}) - V(\widetilde{X}) \rightarrow 0$ in probability (cf. Maritz and Jarrett, 1978). Thus, the result follows with the application of Slutsky's Theorem.

Acknowledgment

The author would like to express his sincere appreciation to the referee who suggests valuable constructive advices to improve the quality of this article.

References

Alloway, J. A., Jr., Raghavachari, M. (1991). Control chart based on the Hodges–Lehmann estimator. *Journal of Quality Technology* 23:336–347.

Albers, W., Kallenberg, C. M. W. (2006). Self-adapting control charts. *Statistica Neerlandica* 60:292–308.

Alwan, L. (2000). *Statistical Process Analysis*. Boston: Mc-Graw-Hill.

Bakir, S. T. (2006). Distribution-free quality control charts based on signed-rank-like statistics. *Communications in Statistics—Theory and Methods* 35:743–757.

Bickel, P. J., Doksum, K. A. (1977). *Mathematical Statistics-Basic Ideas and Selected Topics*. San Francisco: Holden-Day Inc.

Chakraborti, S., Eryilmaz, S. (2007). A nonparametric Shewhart-type signed rank control chart based on runs. *Communications in Statistics—Simulation and Computation* 36:335–356.

Chakraborti, S., van der Laan, P., van de Wiel, M. A. (2004). A class of distribution-free control charts. *Journal of the Royal Statistical Society Ser. C—Applied Statistics* 53:443–462.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics* 7:1–26.

Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. CMBS-NSF Regional Conference Series in Applied Mathematics. Vol. 38. Philadelphia: SIAM.

Efron, B., Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Gilchrist, W. (1997). Modelling with quantile distribution functions. *Journal of Applied Statistics* 24:113–122.

Grimshaw, S. D., Alt, F. B. (1997). Control charts for quantile function values. *Journal of Quality Technology* 29:1–7.

Gunter, B. H. (1989). The use and abuse of C_{pk}, part 2. *Quality Progress* 22:108–109.

Gupta, A., Sengupta, S. (2008). Online control charts for process averages based on repeated median filters. *Communications in Statistics—Simulation and Computation* 37:178–202.

Hall, P. (1988). Theoretical comparisons of bootstrap confidence intervals (with discussions). *Annals of Statistics* 16:927–953.

Hall, P., Martin, M. A. (1989). A note on the accuracy of bootstrap quantile variance estimation. *Probability Theory and Related Fields* 80:197–200.

Janacek, G. J., Meikle, S. E. (1997). Control charts based on medians. *Journal of the Royal Statistical Society Series. D—The Statistician* 46:19–31.

Jones, L. A., Woodall, W. H. (1998). The performance of bootstrap control charts. *Journal of Quality Technology* 30:362–375.

Kanji, G. K., Arif, O. H. (2000). Median Rankit control chart by the quantile approach. *Journal of Applied Statistics* 27:757–770.

Kanji, G. K., Arif, O. H. (2001). Median Rankit control chart for Weibull distribution. *Total Quality Management* 12:629–642.

Khoo, M. B. C. (2005). A control chart based on sample median for the detection of a permanent shift in the process mean. *Quality Engineering* 17:243–257.

Lio, Y. L., Park, C. (2008). A bootstrap control chart for Birnbaum–Saunders percentiles. *Quality and Reliability Engineering International* 24:585–600.

Liu, R. Y., Tang, J. (1996). Control charts for dependent and independent measurements based on bootstrap methods. *Journal of the American Statistical Association* 91:1694–1700.

Maritz, J. S., Jarrett, R. G. (1978). A note on estimating the variance of the sample median. *Journal of the American Statistical Association* 73:194–196.

Nelson, L. S. (1982). Control chart for medians. *Journal of Quality Technology* 14:226–227.

Seppala, T., Moskowitz, H., Plante, R., Tang, J. (1995). Statistical process control via the subgroup bootstrap. *Journal of Quality Technology* 27:139–153.

Shao, J., Tu, D. (1995). *The Jackknife and Bootstrap*. New York: Springer.

Teyarachkul, S., Chand, S., Tang, J. (2007). Estimating the limits for statistical process control charts: a direct method improving upon the bootstrap. *European Journal of Operation Research* 178:472–481.