DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 7 | Issue 1 Article 23

5-1-2008

# Two-Stage Short-Run (X, MR) Control Charts

Matthew E. Elam
*Texas A&M University-Commerce*, Matthew_Elam@tamu-commerce.edu

Kenneth E. Case
*Oklahoma State University*, kcase@okstate.edu

Journal of Modern Applied Statistical Methods
May, 2008, Vol. 7, No. 1, 275-285

1538 – 9472/08/$95.00

# Two-Stage Short-Run (X, MR) Control Charts

Matthew E. Elam
Texas A&M University-Commerce

Kenneth E. Case
Oklahoma State University

This article is the first in a series of two articles that applies two-stage short-run control charting to (X, MR) charts. Theory is developed and then used to derive the control chart factor equations. In the sequel, the control chart factor calculations are computerized and an example is presented.

Key words: control chart, short-run statistical process control, two-stage control charting, probability integral of the range, probability integral of the studentized range, distribution of the mean moving range

## Introduction

The statistical analysis of sample data often requires the sample to be random. In a random sample, each value comes from the same population distribution. Many situations exist in which it is difficult to obtain a random sample. One of these is when the population is not well-defined, as is the case when studying on-going processes, which are often encountered in manufacturing situations.

A statistical technique for establishing data as random in this situation is control charting. The upper and lower control limits and center line for control charts are constructed from data collected as some number m of subgroups, each having size n. Subgroup statistics are then plotted on the control charts. If these statistics plot between the control limits in a random pattern, then the data is likely random. If not, a procedure is invoked to remove the offending subgroups to establish the data as random. The focus of this article is control charting in data limited (short-run) situations when using n=1.

Short-run control charting, as described by Hillier (1969), is necessary in the initiation of a new process, during the startup of a process just brought into statistical control again, and for a process whose total output is not large enough to use conventional control chart constants. Each of these is an example of a short-run situation. A short-run situation is one in which little or no historical information is available about a process in order to estimate process parameters to begin control charting. Consequently, the initial data obtained from the early run of the process must be used for this purpose.

When control charting in a short-run situation, Hillier (1969) gave a two-stage procedure that must be followed to set control limits that result in both the desired probability of a false alarm and a high probability of detecting a special cause signal. In the first stage, m initial subgroups of size n are drawn from the process and are used to determine the control limits. The initial subgroups are plotted against the control limits to retrospectively test if the process was in control while the initial subgroups were being drawn. Once control is established, the procedure moves to the second stage, where the subgroups that were not deleted in the first stage are used to determine the control limits for testing if the process remains in control while future subgroups are drawn. Each stage uses a different set of control chart

Matthew E. Elam is an Associate Professor of Industrial Engineering and Technology at Texas A&M University-Commerce and is an ASQ Certified Quality Engineer. Email him at Matthew_Elam@tamu-commerce.edu. Kenneth E. Case is Regents Professor Emeritus of Industrial Engineering and Management at Oklahoma State University. He is a fellow and past president of the ASQ and IIE, and is an Academician in the International Academy for Quality.

factors called first-stage short-run control chart factors and second-stage short-run control chart factors.

Hillier (1969) presented a two-stage short-run theory initially for $(\overline{X}, R)$ control charts (R is the range of a subgroup) and gave extensive results for first- and second-stage short-run control chart factors for $(\overline{X}, R)$ charts, but for n=5 only. Pyzdek (1993) and Yang (1995, 1999, 2000) attempted to expand Hillier's (1969) results for two-stage short-run $(\overline{X}, R)$ control charts, but their results contained incorrect values. Elam and Case (2001), as well as Elam (2001), described the development and execution of a computer program that overcomes the problems associated with Hillier's (1969), Pyzdek's (1993), and Yang's (1995, 1999, 2000) efforts to present two-stage short-run control chart factors for $(\overline{X}, R)$ charts.

The second application of Hillier's (1969) two-stage short-run theory was to $(\overline{X}, v)$ and $(\overline{X}, \sqrt{v})$ control charts (v is the variance of a subgroup). Yang and Hillier (1970) followed Hillier's (1969) theory to derive equations for calculating the factors required to determine two-stage short-run control limits for $(\overline{X}, v)$ and $(\overline{X}, \sqrt{v})$ charts. The tables of factors Yang and Hillier (1970) presented (see their Tables 1-6) were for several values for number of subgroups, $\alpha$ for the $\overline{X}$ chart, and $\alpha$ for the v and $\sqrt{v}$ charts both above the upper control limit and below the lower control limit ($\alpha$ is the probability of a false alarm). However, as in Hillier (1969), the results were for n=5 only. Elam and Case (2003a, 2003b) addressed issues concerning Yang and Hillier's (1970) results.

The third application of Hillier's (1969) two-stage short-run theory was to $(\overline{X}, s)$ control charts (s is the standard deviation of a subgroup). The difference between $(\overline{X}, \sqrt{v})$ and $(\overline{X}, s)$ control charts is that the former are constructed using the statistic $\overline{\sqrt{v}}$ and the latter are constructed using the statistic $\overline{s}$. Elam and Case (2005a) developed the theory that was needed to apply Hillier's (1969) two-stage short-run theory to $(\overline{X}, s)$ control charts. They then used this theory to derive the equations for calculating the factors required to determine two-stage short-run control limits for $(\overline{X}, s)$ charts. In a second article, Elam and Case (2005b) used the equations presented in Elam and Case (2005a) to develop a computer program that accurately calculates first- and second-stage short-run control chart factors for $(\overline{X}, s)$ charts regardless of the subgroup size, number of subgroups, $\alpha$ for the $\overline{X}$ chart, and $\alpha$ for the s chart both above the upper control limit and below the lower control limit.

Problem

It seems that no attempt appears in the literature to derive equations for calculating the factors required to determine two-stage short-run control limits for (X, MR) charts (MR is the moving range for two individual values). Del Castillo and Montgomery (1994) and Quesenberry (1995) both pointed out this deficiency. The application of (X, MR) control charts is desirable because, in a short-run situation, it may be difficult to form subgroups (Del Castillo & Montgomery, 1994).

Pyzdek (1993) attempted to present two-stage short-run control chart factors for (X, MR) charts for several values for number of subgroups and one value each for $\alpha$ for the X chart and $\alpha$ for the MR chart above the upper control limit. However, all of Pyzdek's (1993) Table 1 results for subgroup size one are incorrect because he used invalid theory (this is explained in detail in the Conclusion section).

Solution

First, the theory is developed that is needed to apply Hillier's (1969) two-stage short-run theory to (X, MR) control charts. It is then used to derive the equations for calculating the factors required to determine two-stage short-run control limits for (X, MR) charts. In the second article, Elam and Case (2006) used the equations presented in this article to develop a computer program that accurately calculates first- and second-stage short-run control chart factors for (X, MR) charts regardless of the number of subgroups, $\alpha$ for the X chart, and $\alpha$ for the MR chart both above the upper control limit and below the lower control limit.

Outline

The probability integrals of the range and the studentized range are presented, both for subgroup size two. These are essential in the application of Hillier's (1969) theory to (X, MR) control charts. Next, Patnaik's (1950) theory is used to develop an approximation to the distribution of the mean moving range. From this result, equations for calculating the factors required to determine two-stage short-run control limits for (X, MR) charts are derived by following the work in the appendix of Hillier (1969). Also, equations to calculate conventional control chart constants for (X, MR) charts are derived. This article concludes with a discussion of its corrections to the literature.

## Methodology

The Probability Integral of the Range for Subgroup Size Two

The probability integral (or cumulative distribution function (cdf)) of the range for subgroups of size two sampled from a standard Normal population was given by Pachares (1959) as equation (1) (with some modifications in notation):

$$P(W) = 2 \times \int_{-\infty}^{\infty} f(x) \times (F(x+W) - F(x))dx \quad (1)$$

W represents the (standardized) range w/σ, where w is the range of a subgroup and σ is the population standard deviation. Throughout this article, F(x) is the cdf of the standard Normal probability density function (pdf) f(x).

The mean of the distribution of the range $W = (w/\sigma)$ for subgroups of size two sampled from a Normal population with mean μ and variance equal to one given by Harter (1960) is equation (2) (with some modifications in notation):

$$d2 = 2/\pi^{0.5} \quad (2)$$

The value d2 is the control chart constant denoted by $d_2$ (see Table M in the appendix of Duncan, 1974). The equation for d2 for subgroup size two for any value of σ was given by Johnson, Kotz, and Balakrishnan (1994). Equations (1) and (2) are the forms used in the computer program in Elam and Case (2006).

The Probability Integral of the Studentized Range for Subgroup Size Two

The probability integral of the studentized range for subgroups of size two sampled from a Normal population was given by Harter, Clemm, and Guthrie (1959) as equation (3a):

$$P3(z) = (5/z) \times \exp(c\nu) \times (P1(z) + P2(z)) \quad (3a)$$

where

$$c\nu = \ln(2) + (\nu/2) \times \ln(\nu/2) - (\nu/2) - \text{gammln}(\nu/2) \quad (3b)$$

$$P1(z) = \int_0^{11} \left[5 \times (W/z) \times \exp\left((z^2 - 25 \times W^2)/(2 \times z^2)\right)\right]^{\nu-1} \times \exp\left((z^2 - 25 \times W^2)/(2 \times z^2)\right) \times P(W)dW \quad (3c)$$

$$P2(z) = (z/5) \times \int_{55/z}^{\infty} \left(x \times \exp\left((1-x^2)/2\right)\right)^{\nu-1} \times \exp\left((1-x^2)/2\right)dx \quad (3d)$$

The variable z is equal to $5 \times Q$. Q represents the studentized range w/s, where w is the range of a subgroup and s is an independent estimate (based on ν degrees of freedom) of the population standard deviation. The equation for determining ν is derived in the next subsection. The equation for cν (equation (3b)) is the natural logarithm of the equation for C(ν) given by Harter, Clemm, and Guthrie (1959). It is derived in Appendix I: Derivations of Elam and Case (2001). The function gammln represents the natural logarithm of the gamma (Γ) function. In equation (3c), P(W) is the probability integral of the range $W = (w/\sigma)$ for subgroup size two (see equation (1)). Equations (3a)-(3d) are the forms used in the computer program in Elam and Case (2006) because they allow for large values of ν (hence large values for m (the number of subgroups)) in the program.

As $\nu \to \infty$ (i.e., as $m \to \infty$), the distribution of the studentized range $Q = (w/s)$ for subgroup size two converges to the distribution of the range $W = (w/\sigma)$ for subgroup size two (see Pearson and Hartley, 1943). This fact is used to derive equations to calculate $\alpha$-based conventional control chart constants for the MR chart.

The Distribution of the Mean Moving Range

Consider the situation in which the mean of a statistic is calculated by averaging m values of the statistic, each of which is calculated from a subgroup of size n. Patnaik (1950) investigated this situation when the statistic was the range and developed an approximation to the distribution of the mean range $\overline{R}/\sigma$. The resulting distribution was the $(\chi \times d_2^*)/\sqrt{\nu}$ distribution, which is a function of the $\chi$ distribution with $\nu$ degrees of freedom (the $\chi$ distribution with $\nu$ degrees of freedom and its moments about zero may be found in Johnson and Welch, 1939).

Equations for $\nu$ and $d_2^*$ were derived from results obtained by equating the squared means as well as the variances of the distribution of the mean range $\overline{R}/\sigma$ and the $(\chi \times d_2^*)/\sqrt{\nu}$ distribution with $\nu$ degrees of freedom. Hillier (1964, 1967) used Patnaik's (1950) theory to derive equations to calculate short-run control chart factors for $\overline{X}$ and R charts, respectively. Hillier (1969) then incorporated the two-stage procedure into his short-run control chart factor calculations for $(\overline{X}, R)$ charts.

Consider the situation in which the statistic is the moving range of size two and the distribution of interest is the distribution of the mean moving range $\overline{MR}/\sigma$. Evidence exists in the literature that $\overline{MR}/\sigma$ may be approximated by a distribution that is a function of either the $\chi^2$ or the $\chi$ distribution. Sathe and Kamat (1957) used results given by Cadwell (1953, 1954) to approximate the distribution of the mean successive difference (i.e., the distribution of the mean moving range $\overline{MR}/\sigma$) by a distribution that is a function of a power of the $\chi^2$ distribution. Roes, Does, and Schurink (1993) used theory similar to Patnaik's (1950) theory to approximate the distribution of the mean moving range $\overline{MR}/\sigma$ (with $\sigma$=1.0) by a distribution that is a function of the $\chi$ distribution.

In order to be able to use Hillier's (1969) theory to derive equations for calculating the factors required to determine two-stage short-run control limits for (X, MR) charts, Patnaik's (1950) theory was applied to approximate the distribution of the mean moving range $\overline{MR}/\sigma$ by the $(\chi \times d_2^*(MR))/\sqrt{\nu}$ distribution with $\nu$ degrees of freedom (this $\nu$ is the same as the one that appears in equation (3a)). The equation for $d_2^*(MR)$ is derived in the Appendix and is given as equation (4) (note: $\text{d2starMR} \equiv d_2^*(MR)$):

$$\text{d2starMR} = \left(d2^2 + d2^2 \times r\right)^{0.5} \qquad (4)$$

The equation for the control chart constant d2 for subgroup size two is given earlier as equation (2). The value r represents the variance of $\overline{MR}/d2$. Its equation is given later as equation (7a). Equation (4) is the form used in the computer program in Elam and Case (2006).

Using results from Prescott (1971), the equation for $\nu$ is determined by equating the ratio of the variance to the squared mean, both of the $\chi$ distribution with $\nu$ degrees of freedom, to the ratio of the variance to the squared mean, both of the distribution of the mean moving range $\overline{MR}/\sigma$. The resulting equation for $\nu$ is equation (5):

$$d(x) = h(x) - r \qquad (5)$$

The exact value for $\nu$ is the value of x such that d(x) is equal to zero. The function h(x) is the ratio of the variance to the squared mean, both of the $\chi$ distribution with x degrees of freedom (x replaces $\nu$). The mean and variance of the $\chi$ distribution with $\nu$ degrees of freedom are given in the Appendix. The equation for h(x), which is derived in Appendix I: Derivations of Elam and Case (2001), is given as equation (6):

$$h(x) = (x \times \exp(2 \times (\text{gammln}(0.5 \times x) - \text{gammln}(0.5 \times x + 0.5))) - 2)/2 \quad (6)$$

The value r is the ratio of the variance to the squared mean, both of the distribution of the mean moving range $\overline{MR}/\sigma$. The mean and the variance of the distribution of the mean moving range $\overline{MR}/\sigma$ are derived in the Appendix. The equation for r was given by Palm and Wheeler (1990) as equation (7a):

$$r = (b \times (m-1) - c)/(m-1)^2 \quad (7a)$$

where

$$b = 2 \times \pi/3 - 3 + 3^{0.5} \quad (7b)$$

$$c = \pi/6 - 2 + 3^{0.5} \quad (7c)$$

Cryer and Ryan (1990) gave an equivalent form for equation (7a). Hoel (1946) gave an equation for the variance of $\overline{MR}$ which, when multiplied by $1/d2^2$, gives the same results as those obtained by using equation (7a). It should be noted that an equivalent form (also based on Patnaik's (1950) theory) of equation (5) may be found in Palm and Wheeler (1990), who used their result to calculate equivalent degrees of freedom for population standard deviation estimates based on consecutive overlapping moving ranges of size two. Equations (5), (6), and (7a)-(7c) are the forms used in the computer program in Elam and Case (2006).

Approximating the distribution of the mean moving range $\overline{MR}/\sigma$ by the $(\chi \times d_2^*(MR))/\sqrt{\nu}$ distribution with ν degrees of freedom works well. In fact, based on how $d_2^*(MR)$ is derived in the Appendix, the means and variances of these two distributions are equal.

## Results

Because the $(\chi \times d_2^*(MR))/\sqrt{\nu}$ distribution with ν degrees of freedom approximates the distribution of the mean moving range $\overline{MR}/\sigma$, the derivation of equations to calculate first- and second-stage short-run control chart factors for (X, MR) charts follows the work in the appendix of Hillier (1969). E22, the second-stage short-run control chart factor for the X chart, is derived in almost the same manner as Hillier's (1969) $A_2^*$. Differences are that n=1 and X, $\overline{X}$, E22, $\overline{MR}$, and $d_2^*(MR)$ in this article replace $\overline{X}$, $\overline{\overline{X}}$, $A_2^*$, $\overline{R}$, and c, respectively, in Hillier (1969). The resulting equation for E22 is given as equation (8) (note: $\text{d2starMR} \equiv d_2^*(MR)$):

$$E22 = (\text{crit\_t}/\text{d2starMR}) \times ((m+1)/m)^{0.5} \quad (8)$$

The value crit_t is the critical value for a cumulative area of $1-(\text{alphaInd}/2)$ under the Student's t curve with ν degrees of freedom (alphaInd is the probability of a false alarm on the X control chart). Equation (8) is the form used in the computer program in Elam and Case (2006).

E21, the first-stage short-run control chart factor for the X chart, is derived in almost the same manner as Hillier's (1969) $A_2^{**}$. Differences are that E21, $X_i$, $\overline{X}$, $\overline{MR}$, and $d_2^*(MR)$ in this article replace $A_2^{**}$, $\overline{X}_i$, $\overline{\overline{X}}$, $\overline{R}$, and c, respectively, in Hillier (1969). The resulting equation for E21 is given as equation (9):

$$E21 = (\text{crit\_t}/\text{d2starMR}) \times ((m-1)/m)^{0.5} \quad (9)$$

The value crit_t has the same meaning here as in equation (8). Equation (9) is the form used in the computer program in Elam and Case (2006).

D42, the second-stage short-run upper control chart factor for the MR chart, is derived in the Appendix. Other than differences in notation, this derivation follows that for Hillier's

(1969) $D_4^*$. The resulting equation for D42 is given as equation (10):

$$D42 = qD4 / d2starMR \qquad (10)$$

The value qD4 is the 1-alphaMRUCL percentage point of the distribution of the studentized range $Q = (w/s)$ for subgroup size two with ν degrees of freedom (alphaMRUCL is the probability of a false alarm on the MR chart above the upper control limit). Equation (10) is the form used in the computer program in Elam and Case (2006).

D32, the second-stage short-run lower control chart factor for the MR chart, is derived in a manner similar to D42. Differences are that D32, qD3, and alphaMRLCL replace D42, qD4, and 1-alphaMRUCL, respectively (alphaMRLCL is the probability of a false alarm on the MR chart below the lower control limit). The resulting equation for D32 is given as equation (11):

$$D32 = qD3 / d2starMR \qquad (11)$$

The value qD3 is the alphaMRLCL percentage point of the distribution of the studentized range $Q = (w/s)$ for subgroup size two with ν degrees of freedom. Equation (11) is the form used in the computer program in Elam and Case (2006).

D41, the first-stage short-run upper control chart factor for the MR chart, is derived in almost the same manner as Hillier's (1969) $D_4^{**}$. Differences are that $D41, MR_i$, D42, and $\overline{MR}$ in this paper replace $D_4^{**}$, $R_i$, $D_4^*$, and $\overline{R}$, respectively, in Hillier (1969). D41 is given as equation (12):

$$D41 = m \times qD4prevm / (d2starMRprevm \times (m-1) + qD4prevm) \qquad (12)$$

The value qD4prevm is the 1-alphaMRUCL percentage point of the distribution of the studentized range $Q = (w/s)$ for subgroup size two with νprevm degrees of freedom (the value νprevm has the same meaning as ν, except it is for m-1 subgroups). The value d2starMRprevm has the same equation as d2starMR (given earlier as equation (4)), except m is replaced with m-1. Equation (12) is the form used in the computer program in Elam and Case (2006).

The equation for D31, the first-stage short-run lower control chart factor for the MR chart, is derived in almost the same manner as Hillier's (1969) $D_3^{**}$. Differences are that $D31, MR_i$, D32, and $\overline{MR}$ in this article replace $D_3^{**}$, $R_i$, $D_3^*$, and $\overline{R}$, respectively, in Hillier (1969). The resulting equation for D31 is given as equation (13):

$$D31 = m \times qD3prevm / (d2starMRprevm \times (m-1) + qD3prevm) \qquad (13)$$

The value qD3prevm is the alphaMRLCL percentage point of the distribution of the studentized range $Q = (w/s)$ for subgroup size two with νprevm degrees of freedom. Equation (13) is the form used in the computer program in Elam and Case (2006).

The equation for E2, the conventional control chart constant for the X chart, may be obtained by taking the limit of either E21 or E22 as m→∞ (i.e., as ν→∞). The resulting equation for E2 is given as equation (14):

$$E2 = crit\_z / d2 \qquad (14)$$

The value crit_z is the critical value for a cumulative area of $1-(alphaInd/2)$ under the standard Normal curve. The equation for the control chart constant d2 for subgroup size two is given earlier as equation (2). Equation (14) is the form used in the computer program in Elam and Case (2006).

The equation for D4, the α-based conventional upper control chart constant for the MR chart, may be obtained by taking the limit of either D41 as m→∞ (i.e., as νprevm→∞) or D42 as m→∞ (i.e., as ν→∞). The resulting equation for D4 is given as equation (15):

$$D4 = wD4 / d2 \qquad (15)$$

The value wD4 is the 1-alphaMRUCL percentage point of the distribution of the range $W = (w/\sigma)$ for subgroup size two. Equation (15) is the form used in the computer program in Elam and Case (2006).

The equation for D3, the $\alpha$-based conventional lower control chart constant for the MR chart, may be obtained by taking the limit of either D31 as m→∞ (i.e., as νprevm →∞) or D32 as m→∞ (i.e., as ν→∞). The resulting equation for D3 is given as equation (16):

$$D3 = wD3 / d2 \qquad (16)$$

The value wD3 is the alphaMRLCL percentage point of the distribution of the range $W = (w/\sigma)$ for subgroup size two. Equation (16) is the form used in the computer program in Elam and Case (2006).

## Conclusion

As mentioned in the Problem subsection of the Introduction, all of Pyzdek's (1993) Table 1 results for subgroup size one are incorrect because he used invalid theory. This is true for two reasons. The first is that he used degrees of freedom based on Patnaik's (1950) approximation applied to the distribution of the mean range $\overline{R}/\sigma$, where $\overline{R}$ is the average of m values of R, each based on a subgroup of size two, not the distribution of the mean moving range $\overline{MR}/\sigma$. In the latter case, the degrees of freedom reflect the fact that serial correlation exists among consecutive overlapping moving ranges of size two, which means that the average of these overlapping MRs reflects that serial correlation. The result is that degrees of freedom based on Patnaik's (1950) approximation applied to the distribution of the mean moving range $\overline{MR}/\sigma$ is less than that from applying Patnaik's (1950) approximation to the distribution of the mean range $\overline{R}/\sigma$, where R is the range of a subgroup of size two.

The second is that Pyzdek (1993) used the equation for $d_2^*$ (i.e., d2star) instead of that for d2starMR (given earlier as equation (4)). The equation for $d_2^*$ is given as equation (17):

$$d_2^* = \left(d_2^2 + d_3^2 / m\right)^{0.5} \qquad (17)$$

where $d_2$ and $d_3$ are the mean and standard deviation, respectively, of the distribution of the range $W = (w/\sigma)$. Equations to calculate $d_2$ and $d_3$ for any subgroup size as well as the equation for $d_2^*$ may be found in Elam and Case (2001). The difference between equations (4) and (17) is that equation (4) has $d2^2 \times r$, which is the variance of the distribution of the mean moving range $\overline{MR}/\sigma$, instead of $d_3^2/m$, which is the variance of the distribution of the mean range $\overline{R}/\sigma$. The equation for r in $d2^2 \times r$ reflects the fact that serial correlation exists among consecutive overlapping moving ranges of size two, which means that the average of these overlapping MRs reflects that serial correlation. The result is that values for d2starMR are less than those for d2star for subgroup size two; but, as m→∞, both converge to d2. It should be noted that d2starMR for m=2 is equal to d2star for n=2 and m=1 (see Table A1 in Appendix III: Tables of Elam and Case, 2001).

One last issue regarding Pyzdek's (1993) Table 1 results is that he gave second-stage short-run control chart factors for number of subgroups equal to one. This is impossible because one must have two subgroups in order to calculate one moving range. For first-stage short-run control chart factors for the individuals and moving range charts, m must be at least two and three, respectively. The reason is E21 (see equation (9)) includes d2starMR (see equation (4)), which includes r, which, according to equation (7a), must have m at least two. Also, in equations (12) and (13), D41 and D31, respectively, include d2starMRprevm, which includes rprevm (r for m-1 subgroups), which must have m at least three. For second-stage short-run control chart factors for the individuals and moving range charts, m must be at least two.

## References


Cadwell, J. H. (1953). Approximating to the distributions of measures of dispersion by a power of $\chi^2$. *Biometrika*, *40*, 336-346.

Cadwell, J. H. (1954). The statistical treatment of mean deviation. *Biometrika*, *41*, 12-18.

Cryer, J. D. & Ryan, T. P. (1990). The estimation of sigma for an X chart: $\overline{\mathrm{MR}}/d_2$ or $S/c_4$?. *Journal of Quality Technology*, *22(3)*, 187-192.

Del Castillo, E. & Montgomery, D. C. (1994). Short-run statistical process control: Q-chart enhancements and alternative methods. *Quality and Reliability Engineering International*, *10*, 87-97.

Duncan, A. J. (1974). *Quality control and industrial statistics*. 4th ed. Homewood, IL: Richard D. Irwin, Inc.

Elam, M. E. (2001). *Investigation, extension, and generalization of a methodology for two stage short run variables control charting*. Ph.D. dissertation. Stillwater, Oklahoma: Oklahoma State University,

Elam, M. E. & Case, K. E. (2001). A computer program to calculate two-stage short-run control chart factors for $(\overline{X}, R)$ charts. *Quality Engineering*, *14(1)*, 77-102.

Elam, M. E. & Case, K. E. (2003a). Two-stage short-run $(\overline{X}, v)$ and $(\overline{X}, \sqrt{v})$ control charts. *Quality Engineering*, *15(3)*, 441-448.

Elam, M. E. & Case, K. E. (2003b). A computer program to calculate two-stage short-run control chart factors for $(\overline{X}, v)$ and $(\overline{X}, \sqrt{v})$ charts. *Quality Engineering*, *15(4)*, 609-638.

Elam, M. E. & Case, K. E. (2005a). Two-stage short-run $(\overline{X}, s)$ control charts. *Quality Engineering*, *17(1)*, 95-107.

Elam, M. E. & Case, K. E. (2005b). A computer program to calculate two-stage short-run control chart factors for $(\overline{X}, s)$ charts. *Quality Engineering*, *17(2)*, 259-277.

Elam, M. E. & Case, K. E. (2006). A computer program to calculate two-stage short-run control chart factors for (X, MR) charts. *Journal of Statistical Software, 15(11),* http://www.jstatsoft.org/.

Harter, H. L. (1960). Tables of range and studentized range. *Annals of Mathematical Statistics*, *31*, 1122-1147.

Harter, H. L., Clemm, D. S., & Guthrie, E. H. (1959). *The probability integrals of the range and the studentized range – probability integral and percentage points of the studentized range; critical values for Duncan's new multiple range test*. Wright Air Development Center Technical Report 58-484, Vol. II. Defense Technical Information Center (DTIC) Document AD Number 231733.

Hillier, F. S. (1964). $\overline{X}$ chart control limits based on a small number of subgroups. *Industrial Quality Control*, *20(8)*, 24-29.

Hillier, F. S. (1967). Small sample probability limits for the range chart. *Journal of the American Statistical Association*, *62(320)*, 1488-1493.

Hillier, F. S. (1969). $\overline{X}$- and R-chart control limits based on a small number of subgroups. *Journal of Quality Technology*, *1(1)*, 17-26.

Hoel, P. G. (1946). The efficiency of the mean moving range. *Annals of Mathematical Statistics*, *17*, 475-482.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions - volume 1*. 2nd ed. New York, NY: John Wiley and Sons, Inc.

Johnson, N. L. & Welch, B. L. (1939). On the calculation of the cumulants of the χ-distribution. *Biometrika*, *31*, 216-218.

Pachares, J. (1959). Table of the upper 10% points of the studentized range. *Biometrika*, *46*, 461-466.

Palm, A. C. & Wheeler, D. J. (1990). Equivalent degrees of freedom for estimates of the process standard deviation based on Shewhart control charts. Unpublished Manuscript.

Patnaik, P. B. (1950). The use of mean range as an estimator of variance in statistical tests. *Biometrika*, *37*, 78-87.

Pearson, E. S. & Hartley, H. O. (1943). Tables of the probability integral of the studentized range. *Biometrika*, *33*, 89-99.

Prescott, P. (1971). Use of a simple range-type estimator of σ in tests of hypotheses. *Biometrika*, *58(2)*, 333-340.

Pyzdek, T. (1993). Process control for short and small runs. *Quality Progress*, *26(4)*, 51-60.

Quesenberry, C. P. (1995). Response. *Journal of Quality Technology*, *27(4)*, 333-343.

Roes, K. C. B., Does, R. J. M. M., & Schurink, Y. (1993). Shewhart-type control charts for individual observations. *Journal of Quality Technology*, *25(3)*, 188-198.

Sathe, Y. S. & Kamat, A. R. (1957). Approximations to the distributions of some measures of dispersion based on successive differences. *Biometrika*, *44*, 349-359.

Yang, S.-F. (1995). *Control policy for small and short production runs*. Taiwan, ROC: National Science Council Research Report.

Yang, S.-F. (1999). An approach to controlling process variability for short production runs. *Total Quality Management*, *10(8)*, 1123-1129.

Yang, S.-F. (2000). Statistical process control for short run manufacturing systems. *Process Control and Quality*, *11(5)*, 433-439.

Yang, C.-H. & Hillier, F. S. (1970). Mean and variance control chart limits based on a small number of subgroups. *Journal of Quality Technology, 2*(1), 9-16.

## Appendix

Derive:

$$\mathrm{d2starMR} = \left(\mathrm{d2}^2 + \mathrm{d2}^2 \times \mathrm{r}\right)^{0.5}$$

First, the mean and variance of the distribution of the mean moving range $\overline{\mathrm{MR}}/\sigma$ need to be determined. Note: By definition,

$$\mathrm{E}(\mathrm{MR}/\sigma) = \mathrm{d2} \Rightarrow (1/\sigma)\times \mathrm{E}(\mathrm{MR}) = \mathrm{d2} \Rightarrow \mathrm{E}(\mathrm{MR}) = \mathrm{d2}\times\sigma$$

$$E\left(\overline{MR}/\sigma\right)$$
$$= \left(1/\sigma\right)\times E\left(\overline{MR}\right)$$
$$= \left(1/\sigma\right)\times E\left(\sum_{i=1}^{m-1} MR_i/(m-1)\right)$$
$$= \left(1/\sigma\right)\times\left(1/(m-1)\right)\times E\left(\sum_{i=1}^{m-1} MR_i\right)$$
$$\Rightarrow E\left(\overline{MR}/\sigma\right)$$
$$= \left(1/\sigma\right)\times\left(1/(m-1)\right)\times\sum_{i=1}^{m-1} E\left(MR_i\right)$$
$$= \left(1/\sigma\right)\times\left(1/(m-1)\right)\times\sum_{i=1}^{m-1} (d2\times\sigma)$$

because

$$\mathrm{E}(\mathrm{MR}) = \mathrm{d2}\times\sigma.$$

$$\Rightarrow \mathrm{E}\left(\overline{\mathrm{MR}}/\sigma\right) = (1/\sigma)\times(1/(\mathrm{m}-1))\times((\mathrm{m}-1)\times\mathrm{d2}\times\sigma) = \mathrm{d2}$$

$$\mathrm{Var}\left(\overline{\mathrm{MR}}/\sigma\right) = \left(1/\sigma^2\right)\times\mathrm{Var}\left(\overline{\mathrm{MR}}\right)$$

From Palm and Wheeler (1990),

$$\mathrm{Var}\left(\overline{\mathrm{MR}}/\mathrm{d2}\right) = \sigma^2\times\mathrm{r}$$

where

$$\mathrm{r} = (\mathrm{b}\times(\mathrm{m}-1)-\mathrm{c})/(\mathrm{m}-1)^2,$$

with

$$\mathrm{b} = 2\times\pi/3 - 3 + \sqrt{3}$$

and

$$\mathrm{c} = \pi/6 - 2 + \sqrt{3}$$

$$\Rightarrow r = \left(1/\sigma^2\right)\times Var\left(\overline{MR}/d2\right)$$
$$= \left(1/\sigma^2\right)\times\left(1/d2^2\right)\times Var\left(\overline{MR}\right)$$

$$\Rightarrow d2^2 \times r = \left(1/\sigma^2\right) \times Var\left(\overline{MR}\right)$$

$$\Rightarrow Var\left(\overline{MR}/\sigma\right) = d2^2 \times r$$

According to Johnson and Welch (1939), the mean of the $\chi$ distribution with $\nu$ degrees of freedom is calculated using the following equation (with some modifications in notation):

$$E(\chi) = \sqrt{2} \times \Gamma(0.5\times\nu + 0.5)/\Gamma(0.5\times\nu)$$

$$\Rightarrow E\left(\chi \times d2starMR/\sqrt{\nu}\right) = \left(d2starMR/\sqrt{\nu}\right) \times E(\chi) = \sqrt{2} \times \left(d2starMR/\sqrt{\nu}\right) \times \left(\Gamma(0.5\times\nu+0.5)/\Gamma(0.5\times\nu)\right)$$

Equating the squared means of the distribution of the mean moving range $\overline{\mathrm{MR}}/\sigma$ and the $(\chi \times \mathrm{d2starMR})/\sqrt{\nu}$ distribution with $\nu$ degrees of freedom results in the following:

$$d2^2 = 2\times\left(d2starMR^2/\nu\right) \times \left(\Gamma(0.5\times\nu+0.5)/\Gamma(0.5\times\nu)\right)^2$$

$$\Rightarrow d2starMR^2 = d2^2 \times (\nu/2) \times \left(\Gamma(0.5\times\nu)/\Gamma(0.5\times\nu+0.5)\right)^2 \quad \text{(A.1)}$$

Appendix 7 of Elam and Case (2005a) gave the variance of the $\chi$ distribution with $\nu$ degrees of freedom as follows:

$$Var(\chi) = \nu - 2\times\left(\Gamma(0.5\times\nu+0.5)/\Gamma(0.5\times\nu)\right)^2$$

$$\Rightarrow Var\left(\chi \times d2starMR/\sqrt{\nu}\right) = \left(d2starMR^2/\nu\right) \times Var(\chi) = \left(d2starMR^2/\nu\right) \times \left(\nu - 2\times\left(\Gamma(0.5\times\nu+0.5)/\Gamma(0.5\times\nu)\right)^2\right)$$

Equating the variances of the distribution of the mean moving range $\overline{\mathrm{MR}}/\sigma$ and the $(\chi \times \mathrm{d2starMR})/\sqrt{\nu}$ distribution with $\nu$ degrees of freedom results in the following:

$$d2^2 \times r = \left(d2starMR^2/\nu\right) \times \left(\nu - 2\times\left(\Gamma(0.5\times\nu+0.5)/\Gamma(0.5\times\nu)\right)^2\right)$$

$$\Rightarrow \left(\Gamma(0.5\times\nu+0.5)/\Gamma(0.5\times\nu)\right)^2 = (d2^2 \times r \times \nu / d2starMR^2 - \nu)/(-2)$$

$$\Rightarrow \left(\Gamma(0.5\times\nu)/\Gamma(0.5\times\nu+0.5)\right)^2 = 2/\left(\nu\times\left(1 - d2^2\times r/d2starMR^2\right)\right) \quad \text{(A.2)}$$

Substituting equation (A.2) into equation (A.1) gives the following equation:

$$d2starMR^2 = d2^2 \times (\nu/2) \times \left(2/\left(\nu\times\left(1 - d2^2\times r/d2starMR^2\right)\right)\right)$$

$$\Rightarrow d2starMR^2 = d2^2/(1 - d2^2\times r/d2starMR^2) = d2starMR^2 \times d2^2/(d2starMR^2 - d2^2\times r)$$

$$\Rightarrow 1 = d2^2/(d2starMR^2 - d2^2\times r)$$

$$\Rightarrow d2starMR^2 = d2^2 + d2^2\times r$$

$$\Rightarrow \mathrm{d2starMR} = \left(\mathrm{d2}^2 + \mathrm{d2}^2 \times \mathrm{r}\right)^{0.5}$$

Derive:

$D42 = (qD4/d2starMR)$, where qD4 is the 1-alphaMRUCL percentage point of the distribution of the studentized range $Q = (w/s)$ for subgroup size two with ν degrees of freedom (alphaMRUCL is the probability of a false alarm on the MR chart above the upper control limit).

Notes: The ensuing derivation is based on the derivation of $D_4^*$ in the appendix of Hillier (1969). The value MR denotes the moving range of a subgroup of size two drawn while in the second stage of the two-stage procedure.

The value D42 needs to be determined such that the following holds:

$$P(MR \le D42 \times \overline{MR}) = 1 - alphaMRUCL$$
$$\Rightarrow P(MR / \overline{MR} \le D42) = 1 - alphaMRUCL$$

To do this, the probability distribution of $MR/\overline{MR}$ needs to determined. Notice that $MR/\sigma$ is the statistic for the distribution of the range $W = (w/\sigma)$ for subgroup size two. An independent estimate of σ based on $\overline{MR}$ is now needed. Replacing σ with this independent estimate results in the statistic for the distribution of the studentized range $Q = (w/s)$ for subgroup size two, which has ν degrees of freedom. The equation to calculate ν is based on the fact that the Patnaik (1950) approximation has been applied to the distribution of the mean moving range. As a result, $\overline{MR}/d2starMR$ needs to be used.

$$\Rightarrow MR / \sigma$$
$$= MR / (\overline{MR} / d2starMR)$$
$$= MR \times d2starMR / \overline{MR}$$

where

$$(MR \times d2starMR) / \overline{MR}$$

is the statistic for the distribution of the studentized range $Q = (w/s)$ for subgroup size two with ν degrees of freedom.

$$\Rightarrow 1 - alphaMRUCL$$
$$= P\left(MR \times d2starMR / \overline{MR} \le qD4\right)$$
$$= P\left(MR / \overline{MR} \le qD4 / d2starMR\right)$$

Setting

$$D42 = qD4 / d2starMR \Rightarrow 1 - alphaMRUCL$$
$$= P\left(MR / \overline{MR} \le D42\right) = P\left(MR \le D42 \times \overline{MR}\right).$$