Journal of Modern Applied Statistical Methods

Volume 5 | Issue 1 Article 17

5-1-2006

Two New Unbiased Point Estimates Of A Population Variance

Matthew E. Elam
The University of Alabama, melam@bama.ua.edu

Recommended Citation
Elam, Matthew E. (2006) "Two New Unbiased Point Estimates Of A Population Variance," *Journal of Modern Applied Statistical Methods*: Vol. 5 : Iss. 1 , Article 17.
DOI: 10.22237/jmasm/1146456960
Available at: http://digitalcommons.wayne.edu/jmasm/vol5/iss1/17

Journal of Modern Applied Statistical Methods
May, 2006, Vol. 5, No. 1, 194-216

1538 – 9472/06/$95.00

Two New Unbiased Point Estimates Of A Population Variance

Matthew E. Elam
Department of Industrial Engineering
The University of Alabama

Two new unbiased point estimates of an unknown population variance are introduced. They are compared to three known estimates using the mean-square error (MSE). A computer program, which is available for download at http://program.20m.com, is developed for performing calculations for the estimates.

Key words: Unbiased, point estimate, variance, range, standard deviation, moving range.

Introduction

The statistical analysis of sample data often involves determining point estimates of unknown population parameters. A desirable property for these point estimates is that they be unbiased. An unbiased point estimate has an expected value (or mean) equal to the unknown population parameter it is being used to estimate. For example, consider the mean $\bar{x}$ and variance v calculated from a random sample of size n (x_1, x_2, …, x_n) obtained from a population with unknown mean μ and variance σ^2. The equations for these two statistics are equations (1) and (2):

$$\bar{x} = \sum_{i=1}^{n} x_i / n \qquad (1)$$

$$v = \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 / (n-1) \qquad (2)$$

It is well known that $\bar{x}$ and v are unbiased point estimates of μ and σ^2, respectively (e.g., see Theorems 8.2.1 and 8.2.2, respectively, in Bain & Engelhardt, 1992). This means the expected value of the sampling distribution of $\bar{x}$ is equal to μ (i.e., $E(\bar{x})=\mu$) and the expected value of the sampling distribution of v is equal to σ^2 (i.e., $E(v)=\sigma^2$).

It is important to have a sample that is random when calculating unbiased point estimates of unknown population parameters. In a random sample, each value comes from the same population distribution. If the values come from different population distributions (i.e., populations with different distributions, means, and/or variances), then the point estimates they are used to calculate will be inaccurate. For example, if the values come from population distributions with different means, then v calculated from this sample using equation (2) will be inflated.

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. On-going processes are often encountered in manufacturing situations. An approach to obtain unbiased point estimates of unknown population parameters from these types of processes is to collect data as some number m of subgroups, each having size n. This is the procedure that is used when constructing control charts to monitor the centering and/or spread of a process. The idea is for the data within a subgroup to come from the same process distribution. If any changes are to occur in the process distribution, it is desirable for them to show up between subgroups. An additional procedure in control chart construction, which may be called a delete-and-revise (D&R) procedure, is performed as an additional safeguard to ensure data within subgroups has the same distribution.

Matthew E. Elam is Assistant Professor of Industrial Engineering at The University of Alabama. He is a member of the ASQ and IIE, and is an ASQ Certified Quality Engineer. Email him at melam@bama.ua.edu.

Two new unbiased point estimates of an unknown population variance are introduced. They are derived assuming the sample data is drawn from an on-going process as m subgroups, each of size n. The Methodology section has an example showing how the control charting procedure works. Also, it presents the three known unbiased point estimates used in the situation considered in this article, it derives the two new unbiased point estimates, and it explains a Mathcad (1999) computer program that performs calculations for the unbiased point estimates. The Results section has mean-square error (MSE) results for the unbiased point estimates. These are useful for the purpose of comparing the unbiased point estimates. The Conclusion section summarizes the interpretations of the analyses in the Results section.

Methodology

Control Charting Procedure. Consider the data in Table 1 obtained from a normally distributed process with μ=100.0 and σ=7.0 (the data was generated in Minitab (2003) and a few changes were made to simulate a process with a nonconstant mean). The true unknown variability for the process is estimated using within subgroup variability. A control chart for spread may be used to determine if data within a subgroup comes from the same process distribution. The control chart for spread used here is the range (R) chart. It is constructed using equations (3a)-(3c):

$$\mathrm{UCL} = D_4 \times \overline{R} \tag{3a}$$

$$\mathrm{CL} = \overline{R} \tag{3b}$$

$$\mathrm{LCL} = D_3 \times \overline{R} \tag{3c}$$

UCL, CL, and LCL are the upper control limit, center line, and lower control limit, respectively, for the R chart. Values for the control chart factors D_4 and D_3 for various n are widely available in control chart factor tables (e.g., see Table M in the appendix of Duncan, 1974). The value $\overline{R}$ (Rbar) is the mean of the m subgroup ranges. The subgroup ranges are calculated for each subgroup as the maximum value in the subgroup minus the minimum value in the subgroup (these calculations are in the "R" column of Table 1). Equations (4a)-(4c) are the R chart control limit calculations for the data in Table 1:

$$\mathrm{UCL} = D_4 \times \overline{R} = 2.282 \times 13.584 = 30.999 \tag{4a}$$

$$\mathrm{CL} = \overline{R} = 13.584 \tag{4b}$$

$$\mathrm{LCL} = D_3 \times \overline{R} = 0.0 \times 13.584 = 0.0 \tag{4c}$$

Figure 1 is the R control chart generated in Minitab (2003).

The delete-and-revise (D&R) procedure involves identifying any subgroup ranges that are greater than the UCL or less than the LCL. The identified subgroups are then removed from the analysis as long as, in this case, each identified subgroup was an indication of a shift in the process mean. The R chart control limits are recalculated using the remaining subgroups. For the Table 1 data, the range (R) for subgroup seven is above the UCL (see the point marked with a "1" in Figure 1). The new value for $\overline{R}$ calculated using the remaining m=19 subgroups after subgroup seven is removed is shown as the Revised $\overline{R}$ in Table 1. The revised control limits are calculated in equations (5a)-(5c):

$$\mathrm{UCL} = D_4 \times \overline{R} = 2.282 \times 12.604 = 28.762 \tag{5a}$$

$$\mathrm{CL} = \overline{R} = 12.604 \tag{5b}$$

$$\mathrm{LCL} = D_3 \times \overline{R} = 0.0 \times 12.604 = 0.0 \tag{5c}$$

Because all of the remaining subgroup ranges are between the revised control limits, the conclusion is that the data within each subgroup comes from the same process distribution.

Table 1. Data Collected as m=20 Subgroups, Each of Size n=4

Subgroup	X_1	X_2	X_3	X_4	R
1	89.558	99.593	99.069	91.211	10.035
2	98.263	98.745	96.959	102.132	5.173
3	93.246	108.054	98.811	102.767	14.808
4	95.493	94.852	109.277	98.418	14.425
5	109.667	108.467	88.994	105.678	20.673
6	94.636	105.764	93.755	88.376	17.388
7	88.000	108.000	113.203	81.000	32.203
8	112.215	104.877	97.752	104.484	14.463
9	87.578	90.221	108.198	99.202	20.620
10	100.029	92.639	96.211	94.332	7.390
11	97.998	101.717	98.704	92.989	8.728
12	107.147	102.370	103.020	95.581	11.566
13	94.597	105.221	103.527	94.565	10.656
14	110.381	93.632	103.740	102.841	16.749
15	96.551	104.145	102.043	102.206	7.594
16	108.505	100.040	99.048	110.904	11.856
17	107.918	104.065	94.514	93.943	13.975
18	114.000	116.000	121.000	123.000	9.000
19	109.304	99.160	97.338	114.353	17.015
20	96.920	104.280	100.290	101.984	7.360
				$\overline{R}$	13.584
				Revised $\overline{R}$	12.604

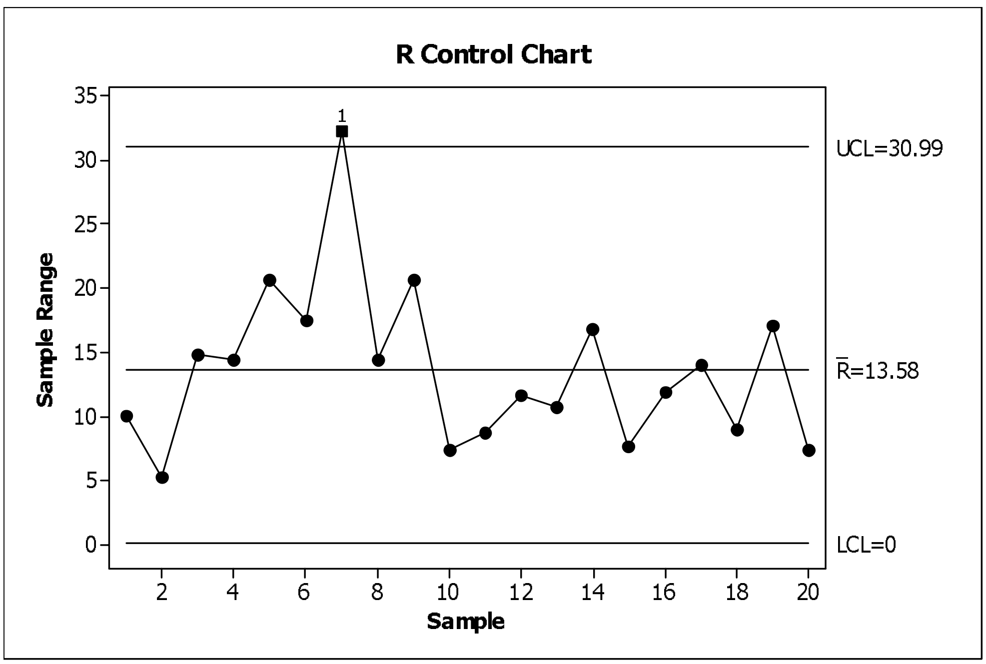

Figure 1. R Control Chart for the Data in Table 1

The next two subsections, Known Unbiased Point Estimates of σ^2 and Two New Unbiased Point Estimates of σ^2, explain how data collected and cleaned in this manner is used to obtain an unbiased point estimate of an unknown process variance using the following statistics:

- $\bar{v}$, the mean of the subgroup variances, where each subgroup variance is calculated using equation (2).
- v_c, the variance of the m×n data values grouped together as one sample. It is calculated using equation (2) with n replaced by m×n and with $\bar{x}$ calculated using equation (1), also with n replaced by m×n. It should be noted that v_c cannot be used when cleaning subgrouped data using a delete-and-revise (D&R) procedure as explained in this subsection. The reason is it would include between subgroup variability, which would inflate its value if the process from which the data is collected is operating under multiple distributions.
- $\bar{R}$, as previously demonstrated.
- $\bar{s}$, the mean of the subgroup standard deviations, where each subgroup standard deviation is calculated using the square root of equation (2).
- $\overline{MR}$, the mean of the moving ranges. When data is collected as m individual values, m-1 moving ranges may be calculated as the absolute value of the difference between consecutive individual values. In this case, the subgroup size n is taken to be two. For example, if the first three individual values are 5.1, 5.3, and 4.8, the first two moving ranges are |5.1-5.3|=0.2 and |5.3-4.8|=0.5.

Known Unbiased Point Estimates of σ^2

The three known unbiased point estimates of σ^2 calculated from data collected as m subgroups, each of size n, considered in this article are $\bar{v}$, v_c, and $\left(\bar{R}/d_2^*\right)^2$. The unbiasedness of $\bar{v}$ is shown in the Appendix of Elam and Case (2003). Wheeler (1995), in his Tables 3.6, 3.7, and 4.2, indicated the unbiasedness for $\bar{v}$ (listed as the pooled variance) as well as for $\left(\bar{R}/d_2^*\right)^2$. The value d_2^* may be called an unbiasing factor, as $\left(\bar{R}\right)^2$ by itself is a biased point estimate of σ^2. The value d_2^* is tabled for various m and n (e.g., see Table D3 in the appendix of Duncan, 1974).

David (1951) gave the equation for d_2^* (i.e., d2star) as equation (6):

$$\text{d2star} = \sqrt{\text{d2}^2 + \text{d3}^2/m} \tag{6}$$

The value d2 (i.e., d_2) is the mean of the distribution of the range W. Its values for various n are widely available in control chart factor tables. Assuming a normal population with mean μ and variance equal to one, Harter (1960) gave the equation for d2 as equation (7) (with some modifications in notation):

$$\text{d2} = n\times(n-1)\times\int_{-\infty}^{\infty}\left[\int_0^{\infty} W\times(F(x+W)-F(x))^{n-2}\times f(x+W)dW\right]\times f(x)dx \tag{7}$$

The function F(x) is the cumulative distribution function (cdf) of the standard normal probability density function (pdf) f(x). The value d3 (i.e., d_3) is the standard deviation of the distribution of the range W. Its values for various n are widely available in control chart factor tables. It is calculated using equation (8):

$$\text{d3} = \sqrt{\text{EW2} - \text{d2}^2} \tag{8}$$

Harter (1960) gave the equation for EW2, the expected value of the second moment of the distribution of the range W for subgroups of size n sampled from a normal population with mean μ and variance equal to one, as equation (9) (with some modifications in notation):

$$\text{EW2} = n\times(n-1)\times\int_{-\infty}^{\infty}\left[\int_0^{\infty} W^2\times(F(x+W)-F(x))^{n-2}\times f(x+W)dW\right]\times f(x)dx \tag{9}$$

Equations (6)-(9) are the forms used in the Mathcad (1999) computer program explained in the Mathcad (1999) Computer Program subsection.

Two New Unbiased Point Estimates of σ^2

Elam and Case (2005a), in their Appendix 7, derived the equation for the factor that allows for an unbiased point estimate of σ^2 to be calculated using $\bar{s}$. Elam and Case (2005a) denoted this factor as c_4^* (i.e., c4star) and gave its equation as equation (10):

$$\text{c4star} = \sqrt{\text{c4}^2 + \text{c5}^2 / \text{m}} \quad (10)$$

The fact that $\left(\bar{s}/c_4^*\right)^2$ is an unbiased point estimate of σ^2 is shown in the Appendix. In equation (10), the value c4 (i.e., c_4) is the mean of the distribution of the standard deviation. Its values for various n are widely available in control chart factor tables. Mead (1966) gave the equation for c4 as equation (11) when σ=1.0 (with some modifications in notation):

$$\text{c4} = \sqrt{2/(\text{n}-1)} \times \exp(\text{gammln}(\text{n}/2) - \text{gammln}((\text{n}-1)/2)) \quad (11)$$

The equivalency of this form to that given by Mead (1966) is shown in Appendix 3 of Elam and Case (2005a). The function gammln represents the natural logarithm of the gamma (Γ) function. The value c5 (i.e., c_5) is the standard deviation of the distribution of the standard deviation. Mead (1966) also gave the equation for c5 as equation (12) when σ=1.0 (with some modifications in notation):

$$\text{c5} = [(2/(\text{n}-1)) \times [\exp(\text{gammln}((\text{n}+1)/2) - \text{gammln}((\text{n}-1)/2)) - \exp(2 \times (\text{gammln}(\text{n}/2) - - \text{gammln}((\text{n}-1)/2)))]]^{0.5} \quad (12)$$

The equivalency of this form to that given by Mead (1966) is shown in Appendix 4 of Elam and Case (2005a). The value c5 is also equal to $\sqrt{1-c_4^2}$. Equations (10)-(12) are the forms used in the Mathcad (1999) computer program explained in the Mathcad (1999) Computer Program subsection.

Elam and Case (2006a), in Appendix 1, derived the equation for the factor that allows for an unbiased point estimate of σ^2 to be calculated using $\overline{\text{MR}}$. Elam and Case (2006a) denoted this factor as $d_2^*(\text{MR})$ (i.e., d2starMR) and gave its equation as equation (13):

$$\text{d2starMR} = \sqrt{\text{d2n2}^2 + \text{d2n2}^2 \times \text{r}} \quad (13)$$

The fact that $\left(\overline{\text{MR}}/d_2^*(\text{MR})\right)^2$ is an unbiased point estimate of σ^2 is shown in the Appendix. In equation (13), the value d2n2 is d2 when n is equal to two. Harter (1960) gave the equation for d2n2 as equation (14) (with some modifications in notation):

$$\text{d2n2} = 2/\sqrt{\pi} \quad (14)$$

The value r is the ratio of the variance to the squared mean, both of the distribution of the mean moving range $\overline{\text{MR}}/\sigma$, an approximation to which is derived in Elam and Case (2006a). Palm and Wheeler (1990) gave the equation for r as equation (15):

$$\text{r} = ((4 \times \pi - 18 + 2 \times 3^{1.5}) \times (\text{m}-1) - \pi + 12 - 2 \times 3^{1.5})/(6 \times (\text{m}-1)^2) \quad (15)$$

Equations (13)-(15) are the forms used in the Mathcad (1999) computer program explained in the Mathcad (1999) Computer Program subsection.

Mathcad (1999) Computer Program

A computer program was coded in Mathcad (1999) with the Numerical Recipes Extension Pack (1997) in order to calculate the unbiasing factors d_2^*, c_4^*, and $d_2^*(\text{MR})$ in equations (6), (10), and (13), respectively, regardless of the number of subgroups m and the subgroup size n. The program is in the Appendix and is named UEFactors.mcd. It is on one page which is divided into seven sections. Download instructions for the program are available at http://program.20m.com.

The first section of the program is the data entry section. The program requires the user to enter m (number of subgroups) and n (subgroup size). Before a value can be entered, the cursor must be moved to the right side of the appropriate equal sign. This may be done using the arrow keys on the keyboard or by moving the mouse arrow to the right side of the equal sign and clicking once with the left mouse button. The program is activated by paging down once the last entry is made. The user is allowed to immediately page down to the output section of the program (explained later) after the last entry is made.

In section 1.1 of the program, the value TOL is the tolerance. The calculations that use this value will be accurate to ten places to the right of the decimal. The functions dnorm(x, 0, 1) and pnorm(x, 0, 1) in Mathcad (1999) are the pdf and cdf, respectively, of the standard normal distribution.

Section 1.2 of the program has the equations for d2, d3, and EW2 given earlier as equations (7), (8), and (9), respectively. Section 1.3 of the program has the equations for c4 and c5 given earlier as equations (11) and (12), respectively. The function gammln is a numerical recipe in the Numerical Recipes Extension Pack (1997). Using it in equations (11) and (12) allows for c4 and c5, respectively, to be calculated for large values of n. Section 1.4 of the program has the equations for d2n2 and r, given earlier as equations (14) and (15), respectively. Section 1.5 of the program has the equations for d2star, c4star, and d2starMR, given earlier as equations (6), (10), and (13), respectively.

The last section of the program has the output. The two values entered at the beginning of the program are given. Accurate values for the unbiasing factors d_2^*, c_4^*, and $d_2^*(\mathrm{MR})$ are also given. The value for $d_2^*(\mathrm{MR})$ is always calculated for n=2, regardless of the value for n entered at the beginning of the program. To copy results into another software package (like Excel), follow the directions from Mathcad's (1999) help menu or highlight a value and copy and paste it into the other software package. When highlighting a value with the mouse arrow, place the arrow in the middle of the value, depress the left mouse button, and drag the arrow to the right. This will ensure just the numerical value of the result is copied and pasted.

Results

The two new unbiased point estimates of σ^2 are compared to the three known unbiased point estimates of σ^2 using the mean-square error (MSE) calculation in equation (16), which is based on Luko's (1996) equation (A3):

$$\mathrm{MSE}(\hat{\sigma}^2) = \mathrm{Var}(\hat{\sigma}^2) + [\mathrm{E}(\hat{\sigma}^2) - \sigma^2]^2 \qquad (16)$$

The value $\hat{\sigma}^2$ represents $\bar{v}$, v_c, $\left(\bar{R}/d_2^*\right)^2$, $\left(\bar{s}/c_4^*\right)^2$, or $\left(\overline{\mathrm{MR}}/d_2^*(\mathrm{MR})\right)^2$, and Var represents the variance as calculated in equation (2). Because these five point estimates of σ^2 are all unbiased, $\mathrm{E}(\hat{\sigma}^2) - \sigma^2 = 0$. Therefore, calculating their MSEs is identical to calculating their variances. Better point estimates are those with smaller MSEs.

MSEs for $\bar{v}$, v_c, $\left(\bar{R}/d_2^*\right)^2$, and $\left(\bar{s}/c_4^*\right)^2$ are calculated using the FORTRAN (1994) computer program named "simulate" in the Appendix. The program simulates the random sampling of m subgroups (m: 1-20, 25, 30, 50, 75, 100, 150, 200, 250, 300), each of size n (n: 2-8, 10, 25, 50), from a standard normal distribution (uniform (0, 1) random variates are generated using the Marse-Roberts code (1983)). This process is repeated 5000 times for each combination of m and n in order to generate 5000 values each of $\bar{v}$, v_c, $\left(\bar{R}/d_2^*\right)^2$, and $\left(\bar{s}/c_4^*\right)^2$ so that their variances can be determined. The necessary values for d_2^* and c_4^* are taken from Table A1 in Appendix III: Tables of Elam and Case (2001) and Table A.1 in Appendix II of Elam and Case (2005b), respectively.

MSEs for $\left(\overline{\mathrm{MR}}/d_2^*(\mathrm{MR})\right)^2$ are calculated using the FORTRAN (1994) computer program named "simulate_MR" in the Appendix. The program simulates the random sampling of m subgroups (m: 2-20, 25, 30, 50, 75, 100, 150, 200, 250, 300) from a standard normal distribution (uniform (0, 1) random variates are generated using the Marse-Roberts code (1983)). This process is repeated 5000 times for each m in order to generate 5000 $\left(\overline{MR}/d_2^*(MR)\right)^2$ values so that the variance can be determined. The necessary values for $d_2^*(MR)$ are taken from Table A.1 in Appendix 2 of Elam and Case (2006b).

The Appendix has the MSE results for $\bar{v}$, v_c, $\left(\bar{R}/d_2^*\right)^2$, $\left(\bar{s}/c_4^*\right)^2$, and $\left(\overline{MR}/d_2^*(MR)\right)^2$ in its Tables A.1-A.5, respectively. As m increases for any n, or as n increases for any m, the MSEs in Tables A.1-A.4 decrease. As m increases, the MSEs decrease in Table A.5. This is not surprising because as more information about the process is at hand, the unbiased estimates should perform better. Only the MSEs for $\bar{v}$, v_c, $\left(\bar{R}/d_2^*\right)^2$, and $\left(\bar{s}/c_4^*\right)^2$ when n=2 and m=1 can be compared to the MSE for $\left(\overline{MR}/d_2^*(MR)\right)^2$ when m=2. In this case, the moving range is interpreted to be the same as the range. These results are the same.

Tables A.6-A.8 in the Appendix have the percent change in MSE ($\bar{v}$) over MSE (v_c), $MSE\left[\left(\bar{s}/c_4^*\right)^2\right]$ over MSE ($\bar{v}$), and $MSE\left[\left(\bar{R}/d_2^*\right)^2\right]$ over $MSE\left[\left(\bar{s}/c_4^*\right)^2\right]$, respectively. The calculations in Tables A.6-A.8 were performed using Excel's full accuracy. Because most of the percentages in these tables are zero or positive, it can be stated that, in general, MSE (v_c) $\leq$ MSE ($\bar{v}$) $\leq MSE\left[\left(\bar{s}/c_4^*\right)^2\right] \leq MSE\left[\left(\bar{R}/d_2^*\right)^2\right]$. The following additional conclusions can be drawn from Tables A.6-A.8:

- In Tables A.6 and A.7, the percent changes decrease as n increases for any m. This means the MSEs for $\bar{v}$, v_c, and $\left(\bar{s}/c_4^*\right)^2$ converge to each other as n increases for any m.
- The MSEs for $\bar{v}$, v_c, and $\left(\bar{s}/c_4^*\right)^2$ are the same when m=1.
- The MSE for $\left(\bar{R}/d_2^*\right)^2$ when n=2 and m=1 is almost identical to that for $\bar{v}$, v_c, and $\left(\bar{s}/c_4^*\right)^2$; however, as n gets larger for m=1 (or any m), the MSEs for $\left(\bar{R}/d_2^*\right)^2$ grow larger than those for $\bar{v}$, v_c, and $\left(\bar{s}/c_4^*\right)^2$. This is because of the well known fact that the range calculation loses efficiency as the size of the sample from which it is calculated increases.
- The MSEs for $\left(\bar{R}/d_2^*\right)^2$ and $\left(\bar{s}/c_4^*\right)^2$ when n=2 are almost identical. This is because the range and standard deviation calculations differ by only a constant when n=2.

Conclusion

From the analyses in the Results section, it may be concluded that $\left(\bar{s}/c_4^*\right)^2$ is at least as good of an unbiased point estimate of σ^2 as $\left(\bar{R}/d_2^*\right)^2$. In fact, as n increases for any m, $\left(\bar{s}/c_4^*\right)^2$ becomes a much better unbiased point estimate of σ^2 than $\left(\bar{R}/d_2^*\right)^2$. Also, the performance of $\left(\bar{s}/c_4^*\right)^2$ approaches that of $\bar{v}$ and v_c as n increases for any m. Additionally, $\left(\overline{MR}/d_2^*(MR)\right)^2$ appears to be an adequate unbiased point estimate of σ^2, as indicated by its reasonably small MSE values. This means that, for the first time, there is an alternative to equation (2) for obtaining an unbiased point estimate of σ^2 from individual values.

Program: UEFactors.mcd

ENTER the following 2 values:

(1) $m := 5$ **(number of subgroups)**

(2) $n := 5$ **(subgroup size)**

Please PAGE DOWN to begin the program.

(1.1) $TOL := 10^{-10}$ $f(x) := \text{dnorm}(x,0,1)$ $F(x) := \text{pnorm}(x,0,1)$

(1.2)
$$d2 := n \times (n-1) \times \int_{-\infty}^{\infty} \left[\int_{0}^{\infty} W \times (F(x+W) - F(x))^{n-2} \times f(x+W)\, dW \right] \times f(x)\, dx$$

$$EW2 := n \times (n-1) \times \int_{-\infty}^{\infty} \left[\int_{0}^{\infty} W^2 \times (F(x+W) - F(x))^{n-2} \times f(x+W)\, dW \right] \times f(x)\, dx$$

$$d3 := \sqrt{EW2 - d2^2}$$

(1.3)
$$c4 := \sqrt{\frac{2}{n-1}} \times \left(e^{\text{gammln}\left(\frac{n}{2}\right) - \text{gammln}\left(\frac{n-1}{2}\right)} \right)$$

$$c5 := \left[\left(\frac{2}{n-1} \right) \times \left[e^{\text{gammln}\left(\frac{n+1}{2}\right) - \text{gammln}\left(\frac{n-1}{2}\right)} - e^{2 \times \left(\text{gammln}\left(\frac{n}{2}\right) - \text{gammln}\left(\frac{n-1}{2}\right) \right)} \right] \right]^{0.5}$$

(1.4) $d2n2 := \frac{2}{\sqrt{\pi}}$

$$r := \frac{\left(4 \times \pi - 18 + 2 \times 3^{1.5}\right) \times (m-1) - \pi + 12 - 2 \times 3^{1.5}}{6 \times (m-1)^2}$$

(1.5) $d2star := \sqrt{d2^2 + \frac{d3^2}{m}}$ $c4star := \sqrt{c4^2 + \frac{c5^2}{m}}$ $d2starMR := \sqrt{d2n2^2 + d2n2^2 \times r}$

FINAL RESULTS:

(1) $m = 5$ **$(\text{Rbar} / \text{d2star})^2$** $d2star = 2.35781$ **$(\text{MRbar} / \text{d2starMR})^2$ (valid for n=2 only)** $d2starMR = 1.23124$

(2) $n = 5$ **$(\text{sbar} / \text{c4star})^2$** $c4star = 0.95229$

```
program simulate
implicit none
INTEGER, parameter :: DOUBLE=SELECTED_REAL_KIND(p=15)
real(kind=double) :: mean, sd, pi, d2star, c4star, r1, r2, X, large, small, v, s, R, vc
real(kind=double) :: sumvc, sumvc2, sumvbar, sumvbar2, sumsbar2, sumsbar22, sumRbar2, sumRbar22
real(kind=double) :: sumX, sumX2, sumv, sums, sumR, sumXsv, sumX2sv
real(kind=double) :: vbar, sbar2, Rbar2, varvc, varvbar, varsbar2, varRbar2
INTEGER :: c, b, a, rep, i, j, seed = 1973272912
integer, dimension(1:29) :: m
integer, dimension(1:10) :: n
open(unit=1, file="simulate.txt")
open(unit=2, file="d2star.txt")
open(unit=3, file="c4star.txt")

mean = 0.0
sd = 1.0
pi = ACOS(-1.0)
m = (/ (c, c = 1, 20), 25, 30, 50, 75, 100, 150, 200, 250, 300 /)
n = (/ 2, 3, 4, 5, 6, 7, 8, 10, 25, 50 /)

write(1, 5) "n", "m", "c4star", "d2star", "varvc", "varvbar", "varsbar2", "varRbar2"
5 format(2X, A, 3X, A, 2X, A, 2X, A, 5X, A, 8X, A, 5X, A, 5X, A)

do b = 1, 10
! n loop

  do a = 1, 29
!   m loop

    sumvc = 0.0
    sumvc2 = 0.0
    sumvbar = 0.0
    sumvbar2 = 0.0
    sumsbar2 = 0.0
    sumsbar22 = 0.0
    sumRbar2 = 0.0
    sumRbar22 = 0.0

    read(2, *) d2star
    read(3, *) c4star

    do rep = 1, 5000
!     replication loop

      sumX = 0.0
      sumX2 = 0.0
      sumv = 0.0
      sums = 0.0
      sumR = 0.0

      do i = 1, m(a)

        sumXsv = 0.0
        sumX2sv = 0.0

!       new subgroup

        do j = 1, n(b)

          call random(r1, seed)
          call random(r2, seed)

          X = mean + sd * ((SQRT(-2. * LOG(r1))) * (COS(2. * pi * r2)))

          sumX = sumX + X
          sumX2 = sumX2 + X**(2.0)
          sumXsv = sumXsv + X
          sumX2sv = sumX2sv + X**(2.0)

          if (j == 1) then
            large = X
            small = X
          else
            if (X > large) large = X
            if (X < small) small = X
          end if

        end do
```

```
        v = (sumX2sv - ((sumXsv)**(2.0)) / n(b)) / (n(b)-1)
        s = v**(0.5)
        R = large - small

        sumv = sumv + v
        sums = sums + s
        sumR = sumR + R

      end do

      vc = (sumX2 - ((sumX)**(2.0)) / (m(a)*n(b))) / (m(a)*n(b)-1.0)
      vbar = sumv / m(a)
      sbar2 = ((sums / m(a))/c4star)**2
      Rbar2 = ((sumR / m(a))/d2star)**2

      sumvc = sumvc + vc
      sumvc2 = sumvc2 + vc**(2.0)
      sumvbar = sumvbar + vbar
      sumvbar2 = sumvbar2 + vbar**(2.0)
      sumsbar2 = sumsbar2 + sbar2
      sumsbar22 = sumsbar22 + sbar2**(2.0)
      sumRbar2 = sumRbar2 + Rbar2
      sumRbar22 = sumRbar22 + Rbar2**(2.0)

!     replication loop
    end do

    varvc = (sumvc2 - ((sumvc)**(2.0)) / (rep - 1.0)) / (rep - 2.0)
    varvbar = (sumvbar2 - ((sumvbar)**(2.0)) / (rep - 1.0)) / (rep - 2.0)
    varsbar2 = (sumsbar22 - ((sumsbar2)**(2.0)) / (rep - 1.0)) / (rep - 2.0)
    varRbar2 = (sumRbar22 - ((sumRbar2)**(2.0)) / (rep - 1.0)) / (rep - 2.0)

    write(1, 10) n(b), m(a), c4star, d2star, varvc, varvbar, varsbar2, varRbar2
10 format(1X, I2, 1X, I3, 1X, F7.5, 1X, F7.5, 1X, F12.10, 1X, F12.10, 1X, F12.10, 1X, F12.10)

!   m loop
  end do

! n loop
end do

stop

contains

subroutine random(uniran, seed)
!
! *******************************************************
! ***** This subroutine generates Uniform (0, 1)     *****
! ***** random variates using the Marse-Roberts code *****
! *******************************************************
!
    implicit none
    INTEGER, parameter :: DOUBLE=SELECTED_REAL_KIND(p=15)
    REAL(KIND=DOUBLE), INTENT(OUT) :: uniran
    INTEGER, INTENT(IN OUT) :: seed
    INTEGER :: hi15, hi31, low15, lowprd, ovflow
    INTEGER, PARAMETER :: mult1 = 24112, mult2 = 26143, &
                          b2e15 = 32768, b2e16 = 65536, &
                          modlus = 2147483647
!
    hi15 = seed / b2e16
    lowprd = (seed - hi15 * b2e16) * mult1
    low15 = lowprd / b2e16
    hi31 = hi15 * mult1 + low15
    ovflow = hi31 / b2e15
    seed = (((lowprd - low15 * b2e16) - modlus) + &
            (hi31 - ovflow * b2e15) * b2e16) + ovflow
!
    if (seed < 0) seed = seed + modlus
```

```
!
    hi15 = seed / b2e16
    lowprd = (seed - hi15 * b2e16) * mult2
    low15 = lowprd / b2e16
    hi31 = hi15 * mult2 + low15
    ovflow = hi31 / b2e15
    seed = (((lowprd - low15 * b2e16) - modlus) + &
            (hi31 - ovflow * b2e15) * b2e16) + ovflow
!
    if (seed < 0) seed = seed + modlus
!
    uniran = (2 * (seed / 256) + 1) / 16777216.0
!
    return
  end subroutine random
!
end program simulate



program simulate_MR
implicit none
INTEGER, parameter :: DOUBLE=SELECTED_REAL_KIND(p=15)
real(kind=double) :: mean, sd, pi, d2starMR, r1, r2, X, first, second, MR
real(kind=double) :: sumMRbar2, sumMRbar22, sumMR, MRbar2, varMRbar2
INTEGER :: c, a, rep, i, seed = 1973272912
integer, dimension(1:28) :: m
open(unit=1, file="simulate_MR.txt")
open(unit=2, file="d2starMR.txt")

mean = 0.0
sd = 1.0
pi = ACOS(-1.0)
m = (/ (c, c = 2, 20), 25, 30, 50, 75, 100, 150, 200, 250, 300 /)

write(1, 5) "m", "d2starMR", "varMRbar2"
5 format(3X, A, 2X, A, 3X, A)

do a = 1, 28
! m loop

  sumMRbar2 = 0.0
  sumMRbar22 = 0.0

  read(2, *) d2starMR

  do rep = 1, 5000
!   replication loop

    sumMR = 0.0

    do i = 1, m(a)

      call random(r1, seed)
      call random(r2, seed)
      X = mean + sd * ((SQRT(-2. * LOG(r1))) * (COS(2. * pi * r2)))

      if (i == 1) then
        first = X
      else
        second = X
        MR = abs(first - second)
        sumMR = sumMR + MR
        first = second
      end if

    end do
```

```
    MRbar2 = ((sumMR / (m(a) - 1))/d2starMR)**2
    sumMRbar2 = sumMRbar2 + MRbar2
    sumMRbar22 = sumMRbar22 + MRbar2**(2.0)

!   replication loop
  end do

  varMRbar2 = (sumMRbar22 - ((sumMRbar2)**(2.0)) / (rep - 1.0)) / (rep - 2.0)

  write(1, 10) m(a), d2starMR, varMRbar2
10 format(1X, I3, 2X, F7.5, 2X, F12.10)

! m loop
end do

stop

contains

subroutine random(uniran, seed)
!
! ***********************************************************
! ***** This subroutine generates Uniform (0, 1)      *****
! ***** random variates using the Marse-Roberts code *****
! ***********************************************************
!
    implicit none
    INTEGER, parameter :: DOUBLE=SELECTED_REAL_KIND(p=15)
    REAL(KIND=DOUBLE), INTENT(OUT) :: uniran
    INTEGER, INTENT(IN OUT) :: seed
    INTEGER :: hi15, hi31, low15, lowprd, ovflow
    INTEGER, PARAMETER :: mult1 = 24112, mult2 = 26143, &
                          b2e15 = 32768, b2e16 = 65536, &
                          modlus = 2147483647
!
    hi15 = seed / b2e16
    lowprd = (seed - hi15 * b2e16) * mult1
    low15 = lowprd / b2e16
    hi31 = hi15 * mult1 + low15
    ovflow = hi31 / b2e15
    seed = (((lowprd - low15 * b2e16) - modlus) + &
            (hi31 - ovflow * b2e15) * b2e16) + ovflow
!
    if (seed < 0) seed = seed + modlus
!
    hi15 = seed / b2e16
    lowprd = (seed - hi15 * b2e16) * mult2
    low15 = lowprd / b2e16
    hi31 = hi15 * mult2 + low15
    ovflow = hi31 / b2e15
    seed = (((lowprd - low15 * b2e16) - modlus) + &
            (hi31 - ovflow * b2e15) * b2e16) + ovflow
!
    if (seed < 0) seed = seed + modlus
!
    uniran = (2 * (seed / 256) + 1) / 16777216.0
!
    return
  end subroutine random
!
end program simulate_MR
```

Table A.1. MSE of $\bar{v}$

m	n									
	2	3	4	5	6	7	8	10	25	50
1	1.752	0.988	0.670	0.476	0.410	0.331	0.288	0.222	0.084	0.040
2	1.039	0.504	0.313	0.244	0.184	0.169	0.145	0.111	0.042	0.021
3	0.667	0.334	0.223	0.165	0.135	0.107	0.099	0.071	0.027	0.013
4	0.527	0.245	0.167	0.127	0.096	0.086	0.074	0.056	0.021	0.011
5	0.395	0.202	0.132	0.102	0.080	0.067	0.057	0.045	0.016	0.008
6	0.338	0.163	0.112	0.084	0.067	0.056	0.048	0.037	0.014	0.007
7	0.294	0.145	0.094	0.071	0.055	0.047	0.040	0.031	0.012	0.006
8	0.245	0.127	0.085	0.063	0.050	0.042	0.036	0.027	0.011	0.005
9	0.224	0.109	0.074	0.054	0.044	0.039	0.031	0.025	0.009	0.005
10	0.200	0.098	0.067	0.050	0.039	0.034	0.028	0.022	0.008	0.004
11	0.181	0.094	0.062	0.046	0.037	0.031	0.025	0.020	0.008	0.004
12	0.163	0.086	0.056	0.043	0.035	0.027	0.023	0.018	0.007	0.003
13	0.151	0.077	0.050	0.038	0.031	0.025	0.022	0.017	0.006	0.003
14	0.142	0.072	0.047	0.036	0.028	0.023	0.021	0.015	0.006	0.003
15	0.134	0.068	0.045	0.033	0.026	0.022	0.020	0.015	0.006	0.003
16	0.127	0.062	0.042	0.032	0.025	0.020	0.017	0.014	0.005	0.003
17	0.118	0.059	0.039	0.030	0.023	0.020	0.017	0.013	0.005	0.002
18	0.110	0.057	0.038	0.027	0.022	0.018	0.016	0.012	0.004	0.002
19	0.101	0.053	0.035	0.026	0.021	0.018	0.015	0.012	0.004	0.002
20	0.100	0.051	0.033	0.025	0.019	0.017	0.014	0.011	0.004	0.002
25	0.079	0.041	0.027	0.020	0.016	0.013	0.012	0.009	0.003	0.002
30	0.066	0.034	0.022	0.017	0.014	0.012	0.010	0.007	0.003	0.001
50	0.041	0.020	0.014	0.010	0.008	0.006	0.006	0.004	0.002	0.001
75	0.028	0.013	0.009	0.007	0.005	0.004	0.004	0.003	0.001	0.001
100	0.021	0.010	0.007	0.005	0.004	0.003	0.003	0.002	0.001	0.000
150	0.013	0.007	0.004	0.003	0.003	0.002	0.002	0.001	0.001	0.000
200	0.010	0.005	0.003	0.002	0.002	0.002	0.001	0.001	0.000	0.000
250	0.008	0.004	0.003	0.002	0.002	0.001	0.001	0.001	0.000	0.000
300	0.007	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.000	0.000

Table A.2. MSE of v_c

m	n									
	2	3	4	5	6	7	8	10	25	50
1	1.752	0.988	0.670	0.476	0.410	0.331	0.288	0.222	0.084	0.040
2	0.683	0.400	0.268	0.216	0.168	0.156	0.136	0.104	0.041	0.021
3	0.396	0.247	0.184	0.143	0.120	0.097	0.092	0.066	0.026	0.013
4	0.300	0.181	0.134	0.107	0.086	0.077	0.068	0.051	0.021	0.011
5	0.214	0.142	0.104	0.085	0.070	0.059	0.052	0.041	0.016	0.008
6	0.178	0.116	0.087	0.071	0.056	0.049	0.043	0.034	0.014	0.007
7	0.159	0.100	0.073	0.056	0.047	0.042	0.036	0.028	0.011	0.006
8	0.132	0.085	0.065	0.051	0.043	0.037	0.032	0.025	0.010	0.005
9	0.118	0.076	0.057	0.045	0.037	0.034	0.028	0.023	0.009	0.005
10	0.106	0.067	0.052	0.041	0.033	0.029	0.025	0.020	0.008	0.004
11	0.095	0.064	0.047	0.038	0.031	0.027	0.023	0.018	0.007	0.004
12	0.087	0.060	0.042	0.035	0.029	0.024	0.021	0.017	0.007	0.003
13	0.078	0.054	0.040	0.030	0.026	0.022	0.019	0.015	0.006	0.003
14	0.077	0.050	0.035	0.029	0.024	0.021	0.018	0.014	0.006	0.003
15	0.070	0.045	0.033	0.027	0.022	0.019	0.017	0.013	0.005	0.003
16	0.067	0.042	0.032	0.025	0.022	0.018	0.015	0.013	0.005	0.002
17	0.061	0.039	0.030	0.024	0.019	0.017	0.015	0.011	0.005	0.002
18	0.058	0.038	0.028	0.022	0.018	0.016	0.014	0.011	0.004	0.002
19	0.052	0.036	0.026	0.021	0.018	0.016	0.013	0.011	0.004	0.002
20	0.050	0.034	0.025	0.020	0.016	0.015	0.012	0.010	0.004	0.002
25	0.041	0.028	0.020	0.016	0.013	0.011	0.010	0.008	0.003	0.002
30	0.033	0.023	0.017	0.014	0.012	0.010	0.008	0.007	0.003	0.001
50	0.020	0.014	0.010	0.008	0.007	0.005	0.005	0.004	0.002	0.001
75	0.014	0.009	0.007	0.005	0.004	0.004	0.003	0.003	0.001	0.001
100	0.010	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.001	0.000
150	0.007	0.004	0.003	0.003	0.002	0.002	0.002	0.001	0.001	0.000
200	0.005	0.003	0.003	0.002	0.002	0.001	0.001	0.001	0.000	0.000
250	0.004	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.000	0.000
300	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.000	0.000

Table A.3. MSE of $\left(\overline{R}/d_2^*\right)^2$

m	n									
	2	3	4	5	6	7	8	10	25	50
1	1.752	1.000	0.693	0.517	0.452	0.378	0.331	0.273	0.131	0.085
2	1.089	0.531	0.331	0.262	0.204	0.190	0.165	0.135	0.064	0.044
3	0.709	0.356	0.240	0.181	0.150	0.125	0.113	0.088	0.042	0.028
4	0.586	0.264	0.181	0.136	0.108	0.098	0.086	0.068	0.032	0.022
5	0.441	0.222	0.142	0.113	0.089	0.076	0.067	0.055	0.025	0.017
6	0.366	0.180	0.122	0.093	0.075	0.064	0.055	0.046	0.022	0.014
7	0.333	0.158	0.103	0.079	0.061	0.052	0.047	0.038	0.019	0.012
8	0.289	0.137	0.092	0.068	0.055	0.048	0.042	0.033	0.016	0.011
9	0.250	0.121	0.081	0.060	0.048	0.044	0.037	0.030	0.014	0.009
10	0.222	0.107	0.073	0.055	0.044	0.039	0.032	0.026	0.013	0.008
11	0.205	0.104	0.067	0.051	0.041	0.035	0.030	0.025	0.012	0.008
12	0.182	0.093	0.060	0.048	0.038	0.031	0.027	0.022	0.010	0.007
13	0.178	0.084	0.055	0.041	0.035	0.029	0.026	0.021	0.010	0.006
14	0.163	0.078	0.051	0.040	0.032	0.026	0.024	0.019	0.009	0.006
15	0.154	0.074	0.049	0.036	0.029	0.025	0.023	0.018	0.009	0.006
16	0.144	0.067	0.046	0.035	0.029	0.023	0.020	0.017	0.008	0.005
17	0.132	0.064	0.043	0.032	0.025	0.022	0.020	0.015	0.008	0.005
18	0.124	0.062	0.042	0.031	0.025	0.021	0.018	0.015	0.007	0.005
19	0.113	0.058	0.039	0.028	0.023	0.021	0.017	0.014	0.007	0.005
20	0.111	0.056	0.036	0.028	0.022	0.019	0.016	0.013	0.007	0.004
25	0.090	0.045	0.030	0.022	0.018	0.015	0.013	0.011	0.005	0.003
30	0.076	0.037	0.024	0.019	0.016	0.013	0.011	0.009	0.004	0.003
50	0.046	0.022	0.015	0.011	0.009	0.007	0.006	0.005	0.003	0.002
75	0.032	0.014	0.010	0.008	0.006	0.005	0.004	0.004	0.002	0.001
100	0.024	0.011	0.007	0.006	0.005	0.004	0.003	0.003	0.001	0.001
150	0.015	0.007	0.005	0.004	0.003	0.002	0.002	0.002	0.001	0.001
200	0.011	0.005	0.004	0.003	0.002	0.002	0.002	0.001	0.001	0.000
250	0.009	0.004	0.003	0.002	0.002	0.002	0.001	0.001	0.001	0.000
300	0.008	0.004	0.002	0.002	0.002	0.001	0.001	0.001	0.000	0.000

Table A.4. MSE of $\left(\bar{s}/c_4^*\right)^2$

m	n									
	2	3	4	5	6	7	8	10	25	50
1	1.752	0.988	0.670	0.476	0.410	0.331	0.288	0.222	0.084	0.040
2	1.089	0.517	0.322	0.250	0.187	0.171	0.147	0.113	0.042	0.021
3	0.709	0.354	0.234	0.171	0.139	0.111	0.102	0.072	0.027	0.013
4	0.586	0.259	0.174	0.131	0.100	0.089	0.076	0.058	0.021	0.011
5	0.441	0.220	0.138	0.107	0.083	0.069	0.058	0.046	0.016	0.008
6	0.366	0.177	0.118	0.087	0.069	0.058	0.049	0.038	0.014	0.007
7	0.333	0.156	0.100	0.075	0.056	0.048	0.042	0.032	0.012	0.006
8	0.289	0.135	0.090	0.066	0.052	0.043	0.037	0.028	0.011	0.005
9	0.250	0.120	0.079	0.057	0.045	0.040	0.033	0.026	0.010	0.005
10	0.222	0.106	0.071	0.052	0.041	0.035	0.028	0.022	0.008	0.004
11	0.205	0.103	0.065	0.049	0.039	0.032	0.026	0.021	0.008	0.004
12	0.182	0.093	0.059	0.046	0.036	0.029	0.024	0.019	0.007	0.003
13	0.178	0.083	0.053	0.040	0.033	0.026	0.023	0.017	0.006	0.003
14	0.163	0.078	0.050	0.038	0.030	0.024	0.021	0.016	0.006	0.003
15	0.154	0.073	0.047	0.035	0.027	0.023	0.020	0.015	0.006	0.003
16	0.144	0.066	0.044	0.033	0.027	0.021	0.018	0.015	0.005	0.003
17	0.132	0.064	0.042	0.031	0.023	0.020	0.017	0.013	0.005	0.002
18	0.124	0.062	0.041	0.029	0.023	0.019	0.016	0.012	0.005	0.002
19	0.113	0.057	0.038	0.027	0.021	0.019	0.016	0.012	0.004	0.002
20	0.111	0.056	0.035	0.027	0.020	0.018	0.014	0.011	0.004	0.002
25	0.090	0.045	0.029	0.021	0.016	0.014	0.012	0.009	0.003	0.002
30	0.076	0.036	0.024	0.018	0.015	0.012	0.010	0.008	0.003	0.001
50	0.046	0.022	0.014	0.011	0.008	0.007	0.006	0.004	0.002	0.001
75	0.032	0.014	0.009	0.007	0.006	0.005	0.004	0.003	0.001	0.001
100	0.024	0.011	0.007	0.005	0.004	0.003	0.003	0.002	0.001	0.000
150	0.015	0.007	0.005	0.004	0.003	0.002	0.002	0.002	0.001	0.000
200	0.011	0.005	0.004	0.003	0.002	0.002	0.002	0.001	0.000	0.000
250	0.009	0.004	0.003	0.002	0.002	0.001	0.001	0.001	0.000	0.000
300	0.008	0.004	0.002	0.002	0.001	0.001	0.001	0.001	0.000	0.000

Table A.5. MSE of $\left(\overline{MR}/d_2^*(MR)\right)^2$

m	MSE
2	1.752
3	1.498
4	1.015
5	0.790
6	0.677
7	0.519
8	0.440
9	0.397
10	0.366
11	0.340
12	0.297
13	0.270
14	0.248
15	0.242
16	0.213
17	0.199
18	0.199
19	0.183
20	0.178
25	0.135
30	0.118
50	0.068
75	0.043
100	0.032
150	0.022
200	0.017
250	0.014
300	0.011

Table A.6. Percent change in MSE($\bar{v}$) (Table A.1) over MSE(v_c) (Table A.2)

m	n									
	2	3	4	5	6	7	8	10	25	50
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	52.073	25.903	16.884	13.289	9.206	8.055	6.425	6.544	2.664	0.815
3	68.193	35.431	20.908	15.835	13.093	10.186	7.706	7.619	2.964	1.380
4	75.557	35.226	24.624	18.797	12.587	12.943	9.253	9.351	1.935	0.722
5	84.482	41.662	27.244	20.039	15.126	13.957	10.916	9.449	2.340	1.101
6	89.940	40.801	28.364	17.146	19.364	14.511	10.580	9.076	3.501	0.988
7	84.869	45.270	28.358	25.440	17.214	10.842	11.563	10.332	4.594	1.879
8	85.890	48.805	30.047	22.209	16.430	14.215	12.691	10.746	3.352	2.133
9	89.357	43.492	31.116	20.060	18.264	14.980	11.704	9.735	3.795	1.924
10	88.726	44.652	30.283	23.127	16.977	16.780	11.599	9.772	2.714	1.527
11	90.284	47.833	31.328	20.932	17.115	15.487	12.731	10.912	3.942	1.449
12	87.522	43.637	33.993	23.695	18.684	14.152	13.556	10.073	2.775	1.944
13	92.710	43.511	26.581	26.125	19.541	13.470	14.819	9.337	2.875	2.287
14	85.025	44.385	32.747	22.168	17.461	14.061	12.865	9.206	3.346	1.852
15	92.841	51.279	34.034	21.948	19.869	16.353	13.868	9.921	3.877	0.501
16	88.865	48.512	31.479	24.520	17.975	15.630	12.666	9.818	3.403	2.457
17	93.852	49.103	30.251	22.397	18.714	15.667	13.909	10.301	4.064	1.888
18	90.367	49.454	35.143	23.567	19.157	12.699	13.300	9.184	2.637	2.173
19	93.210	46.948	33.387	23.265	16.551	17.442	14.703	10.343	3.472	2.220
20	98.648	51.753	31.433	24.018	19.463	16.540	12.286	11.093	3.394	1.117
25	92.231	50.113	32.890	24.573	19.684	15.712	15.426	10.608	3.755	1.647
30	101.498	47.193	30.602	20.961	19.572	14.935	12.462	10.124	3.892	2.687
50	99.336	49.047	32.104	25.351	19.752	17.257	14.353	11.368	3.573	1.625
75	104.021	42.990	31.257	27.672	19.255	14.066	13.853	12.540	4.104	2.447
100	103.253	48.012	32.163	27.019	20.347	15.836	14.674	9.487	4.611	2.196
150	99.622	49.317	32.110	25.578	19.767	14.875	13.895	12.096	2.737	2.173
200	99.086	48.312	33.452	27.408	20.831	18.136	14.511	11.227	4.837	1.806
250	98.234	50.075	28.355	24.681	18.559	14.259	14.870	11.612	3.553	1.313
300	95.180	48.556	33.210	26.037	20.520	16.140	14.207	10.015	4.702	2.797

Table A.7. Percent change in $\text{MSE}\left[\left(\bar{s}/c_4^*\right)^2\right]$ (Table A.4) over $\text{MSE}(\bar{v})$ (Table A.1)

m	n									
	2	3	4	5	6	7	8	10	25	50
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	4.881	2.705	3.097	2.504	1.781	1.503	1.720	2.002	0.189	0.251
3	6.383	6.081	5.074	3.414	2.595	3.238	2.916	1.927	1.082	0.419
4	11.288	5.917	4.175	3.449	3.254	2.920	2.719	3.323	0.639	0.536
5	11.460	8.756	4.641	5.167	3.533	2.564	1.907	2.180	0.679	0.462
6	8.302	8.807	5.859	3.842	3.447	2.976	2.243	2.439	0.578	0.187
7	13.171	7.632	5.779	5.650	2.772	2.706	3.247	2.562	1.147	0.338
8	18.372	6.216	6.211	5.034	3.692	2.920	2.442	2.343	0.496	0.307
9	11.715	9.304	5.771	4.618	2.924	3.732	4.079	2.461	0.581	0.643
10	11.246	8.445	5.271	3.988	4.625	2.571	3.133	1.124	1.383	0.094
11	13.623	8.562	6.059	6.324	4.598	3.544	2.660	2.845	0.916	0.507
12	11.798	8.239	4.983	5.454	3.129	4.335	3.421	2.974	0.628	0.576
13	17.774	8.531	6.361	4.593	3.890	2.999	3.081	1.622	0.779	0.217
14	14.660	8.359	6.573	4.902	3.807	3.132	2.695	3.011	1.096	0.684
15	14.841	8.254	6.022	5.661	4.008	3.416	2.771	2.536	0.759	0.491
16	14.061	6.759	5.745	5.270	4.246	3.923	2.975	2.291	1.029	0.149
17	12.719	9.391	7.298	4.528	3.963	3.591	3.355	1.873	1.131	0.211
18	12.691	8.229	5.910	5.397	4.580	4.292	2.302	3.021	1.005	0.694
19	12.543	8.205	7.148	3.732	3.448	4.753	2.423	2.115	0.903	0.100
20	11.823	8.412	4.982	5.763	4.035	3.235	3.461	2.196	1.116	0.705
25	14.414	8.806	8.196	5.137	4.691	3.476	3.517	2.900	1.079	0.311
30	14.684	8.077	5.816	5.742	4.475	3.819	3.185	3.060	1.102	0.633
50	13.124	8.101	5.945	5.744	4.827	3.670	2.727	3.491	0.909	0.284
75	14.217	8.798	6.347	4.371	4.961	3.834	3.423	3.404	1.392	0.477
100	11.679	8.384	6.871	5.844	3.480	2.562	3.648	2.768	1.351	0.920
150	11.609	10.482	7.403	5.020	5.131	3.924	3.077	3.138	0.797	1.006
200	13.577	9.599	5.333	5.454	4.021	3.456	4.161	2.549	0.938	0.571
250	14.803	9.349	8.450	5.100	4.652	3.561	3.576	2.826	1.459	0.940
300	12.505	8.421	6.809	5.323	3.496	3.308	3.730	1.700	0.698	0.437

Table A.8. Percent change in $\mathrm{MSE}\left[\left(\overline{\mathrm{R}}/d_2^*\right)^2\right]$ (Table A.3) over $\mathrm{MSE}\left[\left(\bar{s}/c_4^*\right)^2\right]$ (Table A.4)

m	n									
	2	3	4	5	6	7	8	10	25	50
1	0.001	1.215	3.480	8.627	10.131	13.936	15.067	22.798	55.428	113.664
2	-0.001	2.594	2.563	4.613	9.264	11.179	11.860	18.848	52.969	107.533
3	0.003	0.573	2.603	6.292	8.040	12.618	11.690	22.479	54.757	108.073
4	0.004	1.784	4.013	3.941	8.350	10.524	13.311	17.925	51.426	105.104
5	0.000	1.115	3.106	5.287	7.480	10.261	14.058	19.329	54.490	107.180
6	-0.001	1.379	3.138	6.893	7.902	9.331	12.450	20.489	52.802	109.893
7	0.002	0.942	3.337	5.896	8.455	9.545	12.367	18.134	55.828	102.097
8	0.001	1.123	2.172	3.487	7.175	11.319	13.533	19.499	53.487	105.157
9	0.002	1.121	2.672	6.139	7.143	10.362	14.281	17.587	47.735	103.686
10	-0.001	0.825	3.198	4.816	7.233	10.044	12.150	17.570	53.828	105.455
11	-0.004	1.292	3.164	5.275	6.896	9.743	13.373	18.760	57.409	100.255
12	0.001	0.775	1.920	4.957	7.255	9.834	11.513	19.786	49.903	106.588
13	-0.003	0.558	2.048	3.359	6.969	10.947	13.622	20.044	58.465	97.759
14	0.002	0.246	3.045	5.150	8.640	8.495	12.859	17.647	55.959	99.568
15	-0.003	0.448	2.745	5.215	8.112	8.094	12.096	17.428	57.369	107.824
16	-0.001	1.045	2.793	4.187	7.972	10.813	14.090	15.777	53.930	109.306
17	-0.001	0.356	2.518	4.946	7.901	8.322	12.944	18.603	53.725	107.640
18	0.002	0.776	3.205	6.324	7.750	11.144	13.361	19.721	57.326	105.909
19	0.001	0.745	2.892	3.913	7.910	8.911	10.971	17.007	56.335	107.476
20	-0.001	1.182	2.800	4.045	7.385	9.055	11.573	19.389	56.966	100.242
25	-0.002	0.191	2.892	4.748	8.011	9.757	13.276	17.238	57.042	98.729
30	0.001	0.959	2.271	5.115	6.917	9.134	13.342	17.543	51.071	99.970
50	0.003	0.887	2.441	5.394	5.941	10.419	12.549	19.035	53.174	111.775
75	0.004	1.585	2.236	4.956	6.513	10.382	13.174	18.330	57.105	96.958
100	0.003	0.592	2.768	5.044	7.278	10.218	10.040	16.489	52.726	100.217
150	0.001	0.667	1.895	4.526	6.441	8.006	12.889	16.511	51.540	106.852
200	-0.002	1.128	2.557	4.633	7.094	8.680	12.273	16.959	56.797	109.206
250	-0.004	1.189	2.730	4.952	5.964	8.974	12.622	19.308	54.066	106.406
300	-0.001	1.226	3.078	4.402	6.802	12.732	15.062	17.502	51.441	103.615

References

Bain, L. J. & Engelhardt, M. (1992). *Introduction to probability and mathematical statistics* (2nd ed.). Belmont, CA: Duxbury Press.

David, H. A. (1951). Further applications of range to the analysis of variance. *Biometrika*, *38*, 393-409.

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

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. (2003). 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. (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. (2006a). Two-stage short-run (X, MR) control charts. *Journal of Modern Applied Statistical Methods*, *5*(2), in press.

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

FORTRAN PowerStation v4.0 in Microsoft Developer Studio (1994). Redmond, WA: Microsoft Corporation.

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

Luko, S. N. (1996). Concerning the estimators $\overline{R}/d_2$ and $\overline{R}/d_2^*$ in estimating variability in a normal universe. *Quality Engineering*, *8(3)*, 481-487.

Marse, K. & Roberts, S. D. (1983). Implementing a portable FORTRAN uniform (0, 1) generator. *Simulation*, *41(4)*, 135-139.

Mathcad 2000 Professional (1999). Cambridge, MA: MathSoft, Inc.

Mead, R. (1966). A quick method of estimating the standard deviation. *Biometrika*, *53*, 559-564.

Minitab Release 14.1 Statistical Software (2003). State College, PA: Minitab Inc.

Numerical Recipes Extension Pack (1997). Cambridge, MA: MathSoft, Inc.

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

Wheeler, D. J. (1995). *Advanced Topics in Statistical Process Control*. Knoxville, TN: SPC Press, Inc.

Appendix

Show: $\left(\bar{s}/c_4^*\right)^2$ is an unbiased point estimate of σ^2; i.e., show $E\left[\left(\bar{s}/c_4^*\right)^2\right] = \sigma^2$

$$E\left[\left(\frac{\bar{s}}{c_4^*}\right)^2\right] = \left(\frac{1}{\left(c_4^*\right)^2}\right)\cdot E\left[\left(\ \bar{s}\ \right)^2\right]$$

$$= \left(\frac{1}{\left(c_4^*\right)^2}\right)\cdot E\left[\left(\frac{\sum_{i=1}^{m} s_i}{m}\right)^2\right]$$

$$= \left(\frac{1}{\left(c_4^*\right)^2}\right)\cdot\left(\frac{1}{m^2}\right)\cdot E\left[\left(\sum_{i=1}^{m} s_i\right)^2\right],$$

$$= \left(\frac{1}{\left(c_4^*\right)^2}\right)\cdot\left(\frac{1}{m^2}\right)\cdot\left[Var\left(\sum_{i=1}^{m} s_i\right) + \left[E\left(\sum_{i=1}^{m} s_i\right)\right]^2\right]$$

$$= \left(\frac{1}{\left(c_4^*\right)^2}\right)\cdot\left(\frac{1}{m^2}\right)\cdot\left[\sum_{i=1}^{m} Var\left(s_i\right) + \left(\sum_{i=1}^{m} E\left(s_i\right)\right)^2\right]$$

because the s_i's are independent.

$$\Rightarrow E\left[\left(\frac{\bar{s}}{c_4^*}\right)^2\right] = \left(\frac{1}{\left(c_4^*\right)^2}\right)\cdot\left(\frac{1}{m^2}\right)\cdot\left[\begin{array}{l}\sum_{i=1}^{m}\left(c_5^2\cdot\sigma^2\right)\\ +\left[\sum_{i=1}^{m}\left(c_4\cdot\sigma\right)\right]^2\end{array}\right],$$

because $Var(s) = c_5^2 \cdot \sigma^2$ and $E(s) = c_4 \cdot \sigma$ (by definition,

$$Var\left(\frac{s}{\sigma}\right) = c5^2$$

$$\Rightarrow \left(\frac{1}{\sigma^2}\right)\cdot Var(s) = c5^2 \Rightarrow Var(s) = c5^2\cdot\sigma^2;$$

by definition,

$$E\left(\frac{s}{\sigma}\right) = c4 \Rightarrow \left(\frac{1}{\sigma}\right)\cdot E(s) = c4 \Rightarrow E(s) = c4\cdot\sigma\text{).}$$

$$\Rightarrow E\left[\left(\frac{\bar{s}}{c_4^*}\right)^2\right] = \left(\frac{1}{(c_4^*)^2}\right)\cdot\left(\frac{1}{m^2}\right)\cdot\left[m\cdot c_5^2\cdot\sigma^2 + (m\cdot c_4\cdot\sigma)^2\right]$$

$$= \left(\frac{1}{(c_4^*)^2}\right)\cdot\left(\frac{1}{m^2}\right)\cdot\left(m\cdot c_5^2\cdot\sigma^2 + m^2\cdot c_4^2\cdot\sigma^2\right)$$

$$= \left(\frac{c_5^2\cdot\sigma^2}{m\cdot(c_4^*)^2}\right) + \left(\frac{c_4^2\cdot\sigma^2}{(c_4^*)^2}\right) = \sigma^2\cdot\left(\frac{c_4^2 + \frac{c_5^2}{m}}{(c_4^*)^2}\right)$$

$$= \sigma^2\cdot\left(\frac{(c_4^*)^2}{(c_4^*)^2}\right), \text{ since } c_4^* = \left(c_4^2 + \frac{c_5^2}{m}\right)^{0.5}$$

$$\Rightarrow E\left[\left(\frac{\bar{s}}{c_4^*}\right)^2\right] = \sigma^2\cdot(1) = \sigma^2$$

Show: $\left(\overline{MR}/d_2^*(MR)\right)^2$ is an unbiased estimate of σ^2; i.e., show $E\left[\left(\overline{MR}/d_2^*(MR)\right)^2\right] = \sigma^2$. One first needs to determine the variance of the distribution of the mean moving range $\overline{MR}/\sigma$.

$$Var\left(\frac{\overline{MR}}{\sigma}\right) = \left(\frac{1}{\sigma^2}\right)\cdot Var\left(\overline{MR}\right)$$

From Palm and Wheeler (1990),

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

where

$$r = \frac{b\cdot(m-1) - c}{(m-1)^2}$$

with

$$b = \frac{2\cdot\pi}{3} - 3 + \sqrt{3}$$

and

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

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

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

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

$$E\left[\left(\frac{\overline{MR}}{d_2^*(MR)}\right)^2\right] = \left(\frac{1}{\left(d_2^*(MR)\right)^2}\right)\cdot E\left[\left(\overline{MR}\right)^2\right]$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2}\right)\cdot\left[Var\left(\overline{MR}\right) + \left(E\left(\overline{MR}\right)\right)^2\right]$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2}\right)\cdot\left[d_2^2\cdot r\cdot\sigma^2 + \left[E\left(\frac{\sum_{i=1}^{m-1} MR_i}{m-1}\right)\right]^2\right]$$

because

$$Var\left(\overline{MR}/\sigma\right) = d_2^2\cdot r \Rightarrow \left(1/\sigma^2\right)\cdot Var\left(\overline{MR}\right)$$
$$= d_2^2\cdot r \Rightarrow Var\left(\overline{MR}\right) \quad .$$
$$= d_2^2\cdot r\cdot\sigma^2$$

$$\Rightarrow E\left[\left(\frac{\overline{MR}}{d_2^*(MR)}\right)^2\right]$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2}\right)\cdot\left[d_2^2\cdot r\cdot\sigma^2 + \left(\frac{1}{(m-1)^2}\right)\cdot\left[E\left(\sum_{i=1}^{m-1} MR_i\right)\right]^2\right],$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2}\right)\cdot\left[d_2^2\cdot r\cdot\sigma^2 + \left(\frac{1}{(m-1)^2}\right)\cdot\left(\sum_{i=1}^{m-1} E\left(MR_i\right)\right)^2\right]$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2} \right) \cdot \left[d_2^2 \cdot r \cdot \sigma^2 + \left(\frac{1}{(m-1)^2} \right) \cdot \left[\sum_{i=1}^{m-1} (d_2 \cdot \sigma) \right]^2 \right]$$

because

$$E(MR) = d_2 \cdot \sigma \text{ (by definition,}$$
$$E\left(\frac{MR}{\sigma}\right) = d2$$
$$\Rightarrow \left(\frac{1}{\sigma}\right) \cdot E(MR) = d2 \Rightarrow E(MR) = d2 \cdot \sigma).$$

$$\Rightarrow E\left[\left(\frac{\overline{MR}}{d_2^*(MR)} \right)^2 \right]$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2} \right) \cdot \left[d_2^2 \cdot r \cdot \sigma^2 + \left(\frac{1}{(m-1)^2} \right) \cdot \left((m-1) \cdot d_2 \cdot \sigma\right)^2 \right]$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2} \right) \cdot \left(d_2^2 \cdot r \cdot \sigma^2 + d_2^2 \cdot \sigma^2 \right)$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2} \right) \cdot \sigma^2 \cdot \left(d_2^2 + d_2^2 \cdot r \right)$$

$$= \left(\frac{1}{\left(d_2^*(MR)\right)^2} \right) \cdot \sigma^2 \cdot \left(d_2^*(MR) \right)^2$$

because

$$d_2^*(MR) = \left(d_2^2 + d_2^2 \cdot r\right)^{0.5}.$$

$$\Rightarrow E\left[\left(\frac{\overline{MR}}{d_2^*(MR)} \right)^2 \right] = \sigma^2 \cdot (1) = \sigma^2$$

QED