Reference : INTRODUCTION TO STATISTICAL QUALITY CONTROL (7e)

By D. C. Montgomery, pp. 433-444

Exponentially Weighted Moving Average (EWMA) Control Charts for Monitoring Gradual Shifts in a Process Mean

As stated by D. C. Montgomery on his p. 433, "the EWMA control charts was first introduced by W. S. Roberts (1959), "*Technometrics*, Vol. 42(1), pp. 97-102". The EWMA statistic for individual measurements (n = 1) at time t is defined as

$$G_t = wx_t + (1-w)G_{t-1}, \qquad (12a)$$

where $0 < w \le 1$ is the weight placed on the present-time observation x_t, and $(1-w)$ is the weight placed on all other observations in the past. At w = 1, the EWMA chart reduces to the ordinary Shewhart chart. Further, x_t (for all t) is assumed distributed as Laplace-Gaussian, i.e., normally distributed random variable with either known (or specified, or desired) variance σ^2, or unknown variance. The 1st problem that occurs is what value should be assigned to G_0 for which t = 1. There are two possibilities:

(1) G_0 is set at a desired known target μ_0, i.e., CNTL = μ_0, and σ^2 is either known, or specified.

(2) No initial value of G_0 is available and σ^2 is unknown. Then, the trial CNTL must be estimated from the first m = 20 to m = 50 observations, i.e., CNTL = G_0 = $\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i$, and for n = 1, σ^2 should be estimated from $\overline{MR}/1.128379$. Some authors use S as an estimate of σ, but I agree with Minitab that uses $\hat{\sigma}_x$ = $\overline{MR}/1.1284$, because S also explains some of the variation between samples when n = 1.

We first discuss the simplest case described under condition (1) above.

Eq. (12a), clearly shows that $G_1 = wx_1 + (1-w)\mu_0$, $G_2 = wx_2 + (1-w)G_1 = wx_2 + (1-w) \times[wx_1 + (1-w)\mu_0] = wx_2 + w(1-w)x_1 + (1-w)^2\mu_0$, $G_3 = wx_3 + (1-w)G_2 = wx_3 + (1-w)\times[wx_2 + w(1-w)x_1 + (1-w)^2\mu_0] = wx_3 + w(1-w)x_2 + w(1-w)^2x_1 + (1-w)^3\mu_0$, …., $G_t = wx_t + w(1-w)x_{t-1} +$

$w(1-w)^2x_{t-2} + w(1-w)^3x_{t-3} + \ldots + w(1-w)^{t-1}x_1 + (1-w)^t\mu_0$. The reader should now observe that I have used G_t to represent the EWMA statistic because of its Geometric-Series nature. Thus, Eq. (12a) generalizes as follows:

$$G_t = w\sum_{i=1}^{t}(1-w)^{t-i}x_i + (1-w)^t x_0, \text{ where } x_0 = \mu_0 \quad (12b)$$

In order to obtain the control limits at time (or stage t), we need to apply the variance operator to Eq. (12b).

$$V(G_t) = V\left[w\sum_{i=1}^{t}(1-w)^{t-i}x_i + (1-w)^t x_0\right] \quad (13a)$$

Because the observations x_t ($t = 1, 2, \ldots, m$) are randomly selected, thus independent, then Eq. (13a) reduces to

$$V(G_t) = w^2 \times \left\{\sum_{i=1}^{t}[(1-w)^{2(t-i)}V(x_i)] + (1-w)^{2t}V(x_0)\right\} \quad (13b)$$

Because μ_0 is a known constant target, then the $V(x_0 = \mu_0) = 0$, and Eq. (13b) further reduces to

$$V(G_t) = w^2 \times \sum_{i=1}^{t}[(1-w)^{2(t-i)}V(x_i)] = w^2 \times \sum_{i=1}^{t}[(1-w)^{2(t-i)}\sigma_X^2]$$

$$= w^2 \times \sigma_X^2 \sum_{i=1}^{t}[(1-w)^{2(t-i)}] = w^2 \times \sigma^2 \sum_{i=0}^{t-1}(1-w)^{2i} \quad (13c)$$

Because the sum $\sum_{i=0}^{t-1}(1-w)^{2i}$ is a finite geometric series with parameters $a = 1$ and the multiplier $(1-w)^2$, it follows that $\sum_{i=0}^{t-1}(1-w)^{2i} = \frac{1-[(1-w)^2]^t}{1-(1-w)^2} = \frac{1-(1-w)^{2t}}{2w-w^2} = \frac{1-(1-w)^{2t}}{w(2-w)}$,

[Recall that $a + ar + ar^2 + ar^3 + \ldots + ar^{g-1} = a(1-r^g)/(1-r)$]. Substituting this last result into (13c), we obtain

$$V(G_t) = w^2 \times \sigma^2 \sum_{i=0}^{t-1}(1-w)^{2i} = w^2 \times \sigma^2 \times \frac{1-(1-w)^{2t}}{w(2-w)} = \frac{\sigma^2 w[1-(1-w)^{2t}]}{2-w} \quad (13d)$$

The above derived Eq. (13d) is consistent with what D. C. Montgomery (7e) provides in Eq. (9.24) on his p. 434. Thus, the standard error of G_t, for a constant G_0, is given by

$$SE(G_t) = \sigma\sqrt{\frac{w}{2-w}}\times[1-(1-w)^{2t}]^{1/2} \tag{14a}$$

If $G_t = w\bar{x}_t + (1-w)G_{t-1}$, and we randomly sample n > 1 items from a process, at equal-interval of times, then the SE in Eq. (14a) modifies to

$$SE(G_t) = \sigma\sqrt{\frac{w}{n(2-w)}}\times[1-(1-w)^{2t}]^{1/2} \tag{14b}$$

From Eqs. (14), the targeted lower and upper control limits for G_t are given by

$$LCL(G_t) = \mu_0 - L\times SE(G_t), \quad \text{and} \quad UCL(G_t) = \mu_0 + L\times SE(G_t), \tag{15}$$

where the value of L lies within the interval $2.6 \le L \le 3.1$, depending on the size of w and ARL_0. D. C. Montgomery provides Table 9.11 on his p. 437 of his Seventh Edition that give different values of L at $ARL_0 = 500$ that show the decreasing values of L as w decreases in order to maintain the same in-control $ARL_0 = 500$. The first 6 rows of his Table 9.11, p. 437, are depicted below:

Table 9.11 of D. C. Montgomery on his p. 437 [Adapted from Lucas and Saccucci (1990)]

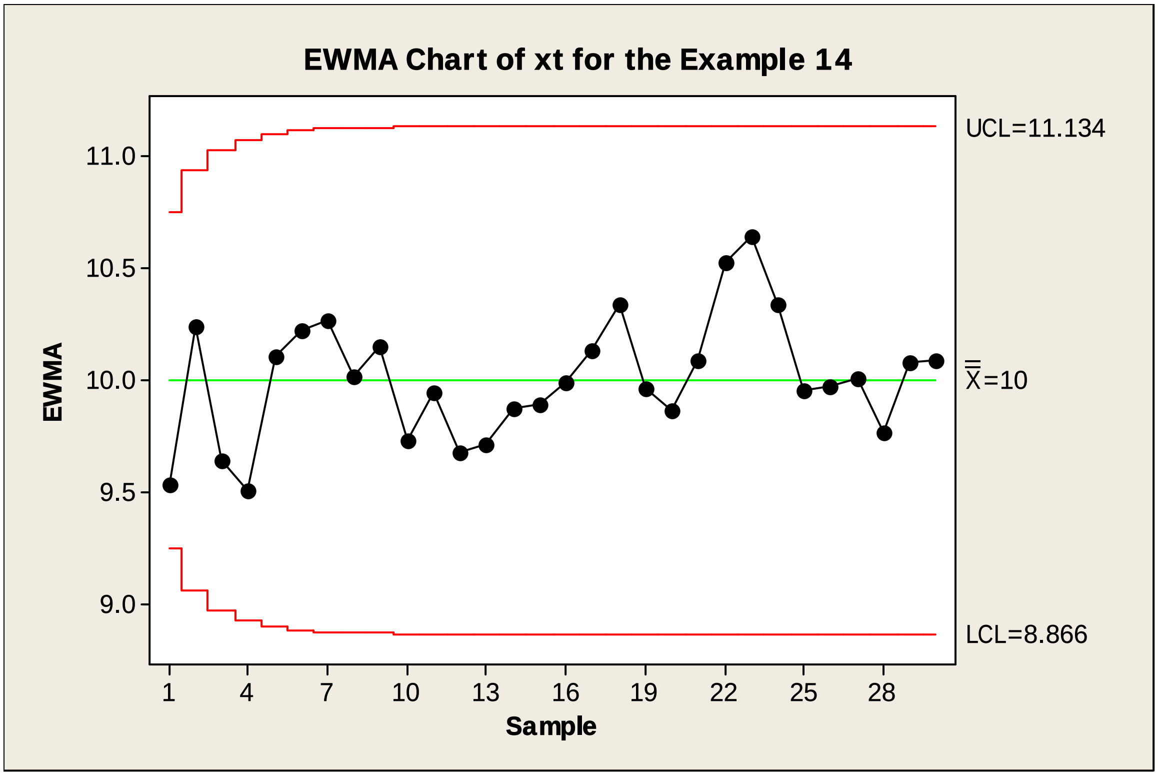
Sigma-shift in μ	L = 3.054, w = 0.40	L = 2.998, w = 0.25	L = 2.962, w = 0.20	L = 2.814, w = 0.10	L = 2.615, w = 0.05
0	$ARL_0 = 500$	500	500	500	$ARL_0 = 500$
0.25	$ARL_1 = 224$	$L_1 = 170$	$L_1 = 150$	$L_1 = 106$	$ARL_1 = 84.1$
0.50	$L_1 = 71.2$	$L_1 = 48.2$	$L_1 = 41.8$	$L_1 = 31.3$	$ARL_1 = 28.8$
0.75	$L_1 = 28.4$	$L_1 = 20.1$	$L_1 = 18.2$	$L_1 = 15.9$	$ARL_1 = 16.4$
1.00	$L_1 = 14.3$	$L_1 = 11.1$	$L_1 = 10.5$	$L_1 = 10.3$	$ARL_1 = 11.4$
1.50	$L_1 = 5.9$	$L_1 = 5.5$	$L_1 = 5.5$	$L_1 = 6.1$	$ARL_1 = 7.1$

When w is small, then the EWMA statistic at time t carries with it a great deal of inertia from the past, and hence less sensitive in catching present gradual mean shifts in the process, and vice a versa. Thus, if the EW control procedure requires a fast response to shifts in the process mean (μ), a larger value of w in the interval [0.20, 0.50] is best; however, larger values of sample size n, taken at each stage, do not need as larger values of w. For $1 \le n \le 10$, as a rule of thumb, I would recommend a maximum of w at 0.40 to w 0.20 when n = 1; it seems that this is somewhat of a gray area! As D. C. Montgomery mentions on his p. 436 of his 7th edition, the optimal EWMA chart is derived by the selection of w and L (he uses the symbol λ as the weight of x_t at time t) in such a manner that gives the desired in-control ARL_0 and out-of-control ARL_1. He further states that when $0 < w \le 0.10$, there is an advantage in reducing the width of the limits by using an L less than 3, such as $2.6 \le L \le 2.8$. Clearly, D. C. Montgomery's recommendation is made in order to increase the power (or sensitivity) of the EWMA chart in detecting gradual shifts in μ.

Just like Example 9.2 of D. C. Montgomery (pp. 435-436 of his 7e), I used Minitab to generate a random sample of size m = 30 from a N(10, 1), which is emailed to you. Then clearly the desired CNTL is $\mu_0 = 10$ with known $\sigma^2 = 1$. Further, because n =1, the Shewhart 3-sigma $LCL_x = 7$ and $UCL_x = 13$. My spreadsheet shows that every individual measurement is in control starting at time t = 1 to t = 30. On the same spreadsheet, I have also evaluated the control nature of the process based on the EWMA control chart with w = 0.25 and $L = 3.0$. The Excel file clearly shows that every point is in control from the standpoint of sudden shift (Shewhart chart) and gradual shift (from the EWMA chart). The EWMA chart for my Example 14 data is provided atop the next page from Minitab. Note that Minitab sets the default value of w at 0.20 and $L = 3$.

2. EWMA Charts When the Desired target, G_0, and σ^2 are Unknown and Have to be Estimated

Consider a Gaussian process N(unknown μ, unknown σ^2). In order to set up trial EWMA limits, we take an initial random sample of size $20 \le m \le 40$, from which we compute the statistics $\bar{\mathbf{x}}$ and $\overline{\mathbf{MR}}$, where $\overline{\mathbf{MR}}$ is the average of (m−1) moving ranges. Then the EWMA statistic at time t is defined as



$$G_t = wx_t + (1-w)G_{t-1}, \text{ where } G_0 = \bar{\bar{\mathbf{x}}} \qquad (16)$$

Note that, although D. C. Montgomery does not discuss this particular case, I find the definition in Eq. (16) somewhat troubling, because at t = 1, no value of $\bar{\bar{\mathbf{x}}}$ based on $m \geq 20$ is yet available. However, my definition in (16) seems consistent with that of Minitab's and that of R. E. Devore *et. al* (i.e., Devore, Chang, Sutherland), *Statistical Quality Design and Control*, SECOND EDITION (2007), pp. 301-313, Prentice Hall, ISBN: 0-13-041344-5.

As before, we need to compute the variance of G_t, but now the CNTL is set at $\bar{\bar{\mathbf{x}}}$. Following my development on pp. 86-88 of these notes, Eq. (13b) modifies to

$$V(G_t) \cong w^2 \times \left\{ \sum_{\mathbf{i=0}}^{t-1} \mathbf{[(1-w)^{2i} V(x_{t-i})]} + (1-w)^{2t} V(\bar{\bar{\mathbf{x}}}) \right\} \qquad (17a)$$

Eq. (17a) is an approximation because there is a small term (which can be ignored for $t > 6$) that is left out. Because, $V(\bar{\mathbf{x}}) = \sigma^2/m$, then Eq. (17a) results in

$$V(G_t) \cong \frac{\sigma^2 w[1-(1-w)^{2t}]}{2-w} + \frac{(1-w)^{2t}\sigma^2}{m} \tag{17b}$$

I checked my Eq. (17b) against that of Minitab's; it seems that Minitab ignores the second term on the far RHS of (17b), because the 2nd term on the RHS of (17b) is much smaller than the 1st term on the RHS, and roughly by a factor of 7 at t =1, and diminishes much more rapidly than the 1st term as t increases, especially when $w \geq 0.30$. Thus, the standard error of G_t is given by

$$se(G_t) \cong \hat{\sigma}_x \sqrt{\frac{w[1-(1-w)^{2t}]}{2-w} + \frac{(1-w)^{2t}}{m}} \tag{18a}$$

or

$$se(G_t) \doteq \hat{\sigma}_x \sqrt{\frac{w[1-(1-w)^{2t}]}{2-w}} \tag{18b}$$

where (18b) is a further approximation to (18a), but consistent with Minitab's *se*.

Note that R. E. Devore *et. al* (2007) use a different procedure to obtain the control limits for G_t, based on a paper by A. L. Sweet, *IIE Transactions*, Vol. 18, No. 1, 1986, pp. 26-33, which is not consistent with those of Minitab's. Their control limits on p. 305 are given by $\bar{\mathbf{x}} \pm \mathbf{A}^* S_x$, where $\mathbf{A}^*$ = 0.480, 0.688, 1.000, 1.132, 1.225, 1.342, 1.500, 1.732, and 2.121 at w = 0.05, 0.10, 0.20, 0.25, 0.286, 0.333, 0.40, 0.50, and 0.667, respectively, and S_x is the standard deviation of the sample, i.e., $S_x = \sqrt{\sum_{j=1}^{m} (x_j - \bar{x})^2 / (m-1)}$.

<u>(25 Point-Bonus Problem)</u> (a, 5-Point bonus). Explain why (18b) is still an approximation to the exact $V(G_t)$. (b, 10-Point bonus). Derive the exact $V(G_t)$. (c, 10- Point Bonus). Use the data of Example 15 on my website to obtain G_t, and the two sets of $LCL(G_t)$, and $UCL(G_t)$ at w = 0.333, and L = 3 using Eqs. (18c & d). Compare your results against those of Minitab's and also submit your Minitab output.

The Exact Variance of G_t when the CNTL = $\bar{x}$

In order to obtain the exact variance of $G_t = w\sum_{i=1}^{t}(1-w)^{t-i}x_i + (1-w)^t G_0 =$ $w\sum_{i=1}^{t}(1-w)^{t-i}x_i + (1-w)^t\bar{x}$, we must include the covariance of $w\sum_{i=1}^{t}(1-w)^{t-i}x_i$ with $(1-w)^t\bar{x}$. Thus, the exact variance is given by

$$V(G_t) = V[w\sum_{i=1}^{t}(1-w)^{t-i}x_i + (1-w)^t\bar{x}] = V[w\sum_{i=1}^{t}(1-w)^{t-i}x_i] + V[(1-w)^t\bar{x}] + 2COV[w\sum_{i=1}^{t}(1-w)^{t-i}x_i, (1-w)^t\bar{x}] \tag{19a}$$

From Eqs. (18b) and (19a), we obtain

$$V(G_t) = \frac{\sigma^2 w[1-(1-w)^{2t}]}{2-w} + \frac{(1-w)^{2t}\sigma^2}{m} + 2w(1-w)^t \times COV[\sum_{i=1}^{t}(1-w)^{t-i}x_i, \bar{x}] \tag{19b}$$

We now compute the Covariance on the RHS of (19b).

$$COV[\sum_{i=1}^{t}(1-w)^{t-i}x_i, \bar{x}] = COV[\sum_{i=1}^{t}(1-w)^{t-i}x_i, \sum_{j=1}^{m}x_j/m]$$

$$= \frac{1}{m}[\sum_{i=1}^{t}\sum_{j=1}^{m}(1-w)^{t-i}COV(x_i, x_j)] \tag{20a}$$

Because x_i is independent of x_j for all $i \neq j$, i.e., the $COV(x_i, x_j) = 0$ for all $i \neq j$, and for $i = j$, the $COV(x_i, x_i) = V(x_i) = \sigma^2$, then Eq. (20a) reduces to

$$COV[\sum_{i=1}^{t}(1-w)^{t-i}x_i, \bar{x}] = \frac{1}{m}\times[\sum_{i=1}^{t}(1-w)^{t-i}COV(x_i, x_i)] = \frac{1}{m}\times[\sum_{i=1}^{t}(1-w)^{t-i}\sigma^2]$$

$$=\frac{\sigma^2}{m}\sum_{i=t}^{i=1}(1-w)^{t-i} = \frac{\sigma^2}{m}\sum_{i=0}^{t-1}(1-w)^{i} = \frac{\sigma^2}{m}\times\frac{1-(1-w)^t}{1-(1-w)}$$

$$= \frac{\sigma^2}{m}\times\frac{1-(1-w)^t}{w} \quad (20b)$$

Substituting (20b) into Eq. (19b) results in

$$V(G_t) = \frac{\sigma^2 w[1-(1-w)^{2t}]}{2-w} + \frac{(1-w)^{2t}\sigma^2}{m} + 2w(1-w)^t \times \frac{\sigma^2}{m}\times\frac{1-(1-w)^t}{w}$$

$$= \frac{\sigma^2 w[1-(1-w)^{2t}]}{2-w} + \frac{(1-w)^{2t}\sigma^2}{m} + 2(1-w)^t \times \frac{\sigma^2}{m}\times[1-(1-w)^t]$$

$$= \sigma^2\{\frac{w[1-(1-w)^{2t}]}{2-w} + \frac{(1-w)^{2t}}{m} + \frac{2(1-w)^t}{m}\times[1-(1-w)^t]\}$$

$$= \sigma^2\{\frac{w[1-(1-w)^{2t}]}{2-w} - \frac{(1-w)^{2t}}{m} + \frac{2(1-w)^t}{m}]\} \quad (21a)$$

Letting ω = 1−w, Eq. (21a) simplifies to

$$V(G_t) = \sigma^2\{\frac{w[1-\omega^{2t}]}{2-w} - \frac{\omega^{2t}}{m} + \frac{2\omega^t}{m}]\} = \sigma^2[\frac{w(1-\omega^{2t})}{2-w} + \frac{\omega^t}{m}(2-\omega^t)] \quad (21b)$$

Consequently, the *se* of G_t, when n = 1 and the initial subgroup is of size m, is given by

$$se(G_t) = \hat{\sigma}_x[\frac{w(1-\omega^{2t})}{2-w} + \frac{\omega^t}{m}(2-\omega^t]^{1/2} = \frac{\overline{MR}}{1.1284}[\frac{w(1-\omega^{2t})}{2-w} + \frac{\omega^t}{m}(2-\omega^t)]^{1/2}$$

(21c)

Note that if n > 1 at each stage t, then Eq. (21a) has to be re-derived and $\hat{\sigma}_x$ has to be estimated either by $\bar{R}/d_2$, $\bar{S}$, or S_p.