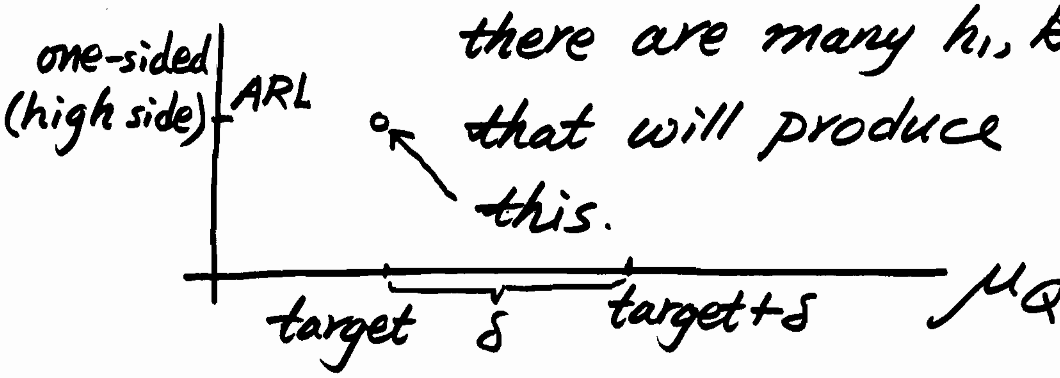
Armed with the ability to compute ARL's for CUSUMs we can consider rational choice of parameters. — This is well documented for iid $Q$.

<u>Bottom Line</u>: For fixed "target" and $\sigma_Q$, there are many $h_1, k_1$ pairs that will produce this.




I can choose among these $h_1, k_1$ pairs to get an optimal ARL at target$+\delta$.

How? Let

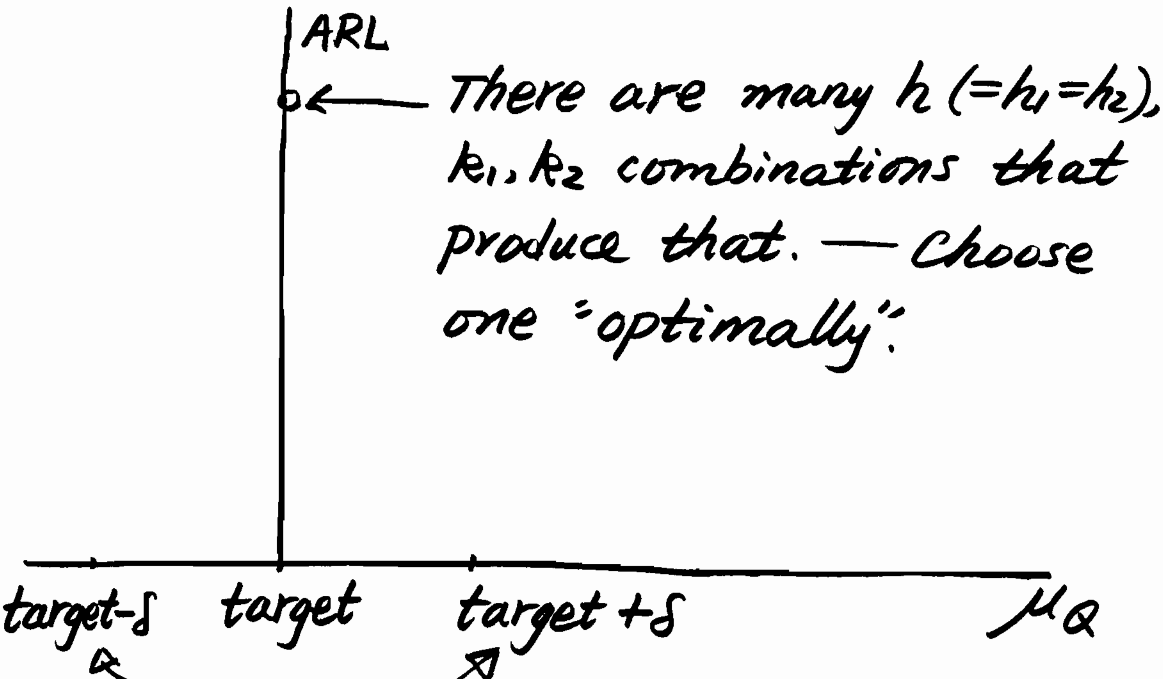
$$k_1^{opt} = \text{target} + \frac{\delta}{2}$$

$$\left(k_2^{opt} = \text{target} - \frac{\delta}{2}\right) \leftarrow \text{low side CUSUM}$$

Then choose $h_1$ ($h_2$) to get the desired on-target ARL — see pages 147–148 of V&J.

---

There is a similar story to tell for 2-sided schemes.




try to minimize ARL at these means for $Q$

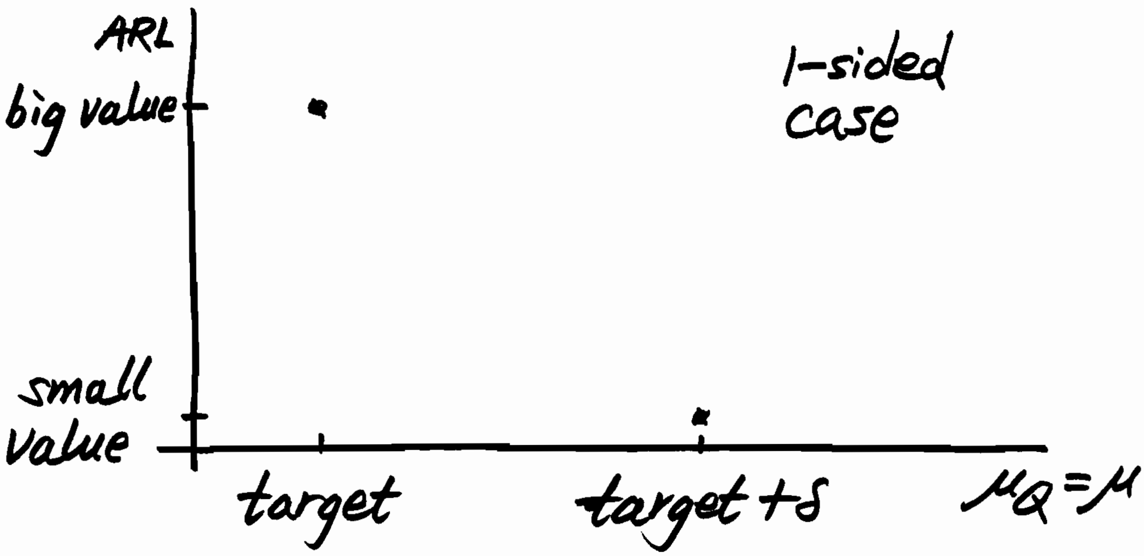
How? Choose $k_1^{opt}$, $k_2^{opt}$ as before and then find $h$ $(=h_1=h_2)$ to get the desired on-target ARL — see pages 147–148 of V&J.

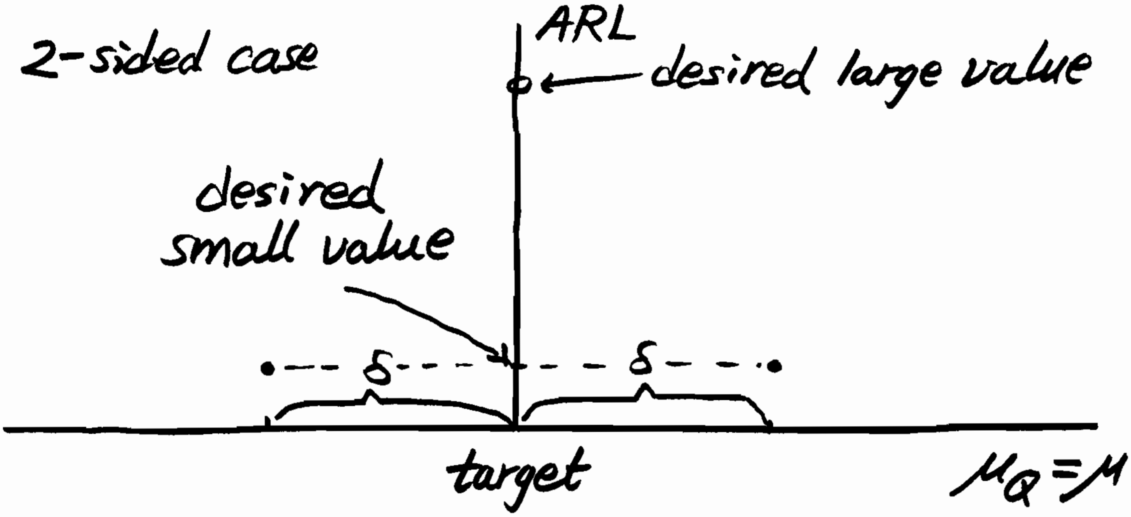
---

A few other CUSUM issues (see §4.2.2 of V&J)

1) Sample size choice/chart design for the case $Q=\bar{x}$, $\sigma_Q=\frac{\sigma}{\sqrt{n}}$.

Here if I'm choosing chart parameters <u>and</u> $n$, at least approximately I can have:




See pages 151, 152 of V&J (makes use of the $k_1^{opt}$ and/or $k_2^{opt}$ from before)

2) "FIR" CUSUMs

(fast initial response) — after an alarm and intervention there can be a fair chance of failure to fix a problem and in such cases one really wants a "quick alarm." —

With this motivation, Lucas & Crosier (1982, Technometrics) suggested

one-sided schemes: $u = \frac{h_1}{2}$ or $v = -\frac{h_2}{2}$

two-sided scheme: $u = \frac{h}{2}$ and $v = -\frac{h}{2}$

(i.e., start "half way to an alarm")

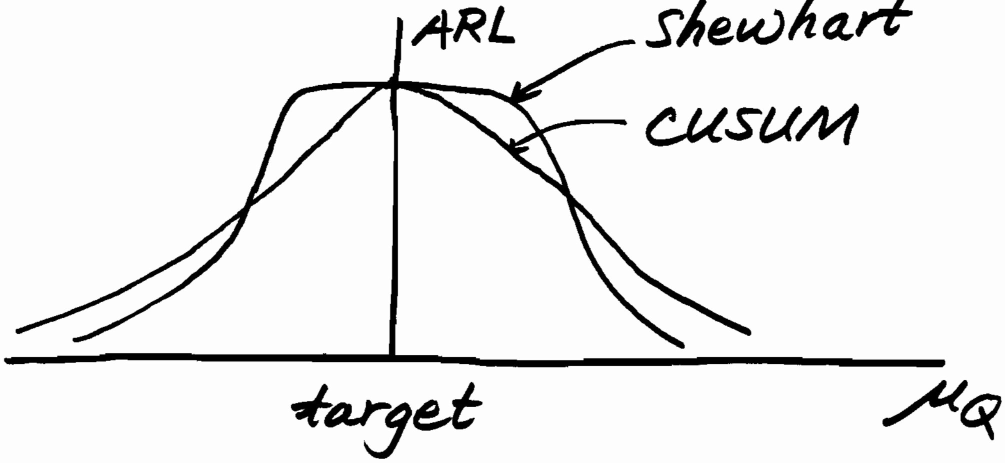
– In comparison to 0-headstart CUSUMs this requires only a very small increase in decision interval in order to maintain a given on-target ARL.

— See Tables 4.10 and 4.11 (page 155) in V&J and compare to Tables 4.5 and 4.6. (page 147)

— They provide some fairly important gains in off-target ARLs.

3) Shewhart-CUSUM combinations

CUSUMs do well at detecting small changes in $\mu_Q$ but are "poor" at detecting big changes.




Lucas suggested using a 3.5 sigma or 4 sigma Shewhart limit along with a CUSUM — (Lucas, 1982, Journal of Quality Technology)

This doesn't change ARLs much at or near target and improves them substantially off-target —

ARL behavior for Shewhart-CUSUM can be handled with MC's and with integral equations — see, e.g., problems 2.28 and 2.29 of the notes.

# EWMA Control Charts (§4.1 of V&J)

- alternative to CUSUM schemes
- in <u>some ways</u> they have similar ARL properties (at least in one way they are much inferior)
  ← EWMAs

For a $0<\lambda\le 1$, let

$$EWMA_0 = \text{target}$$
$$EWMA_i = \lambda Q_i + (1-\lambda)EWMA_{i-1}$$

Action/Control Limits: $\left.\begin{array}{l}\text{target}+h\\ \text{target}-h\end{array}\right\}$ for $h>0$

ARL properties?

Design of such a scheme? (Choice of $\lambda, h$)

Suppose $Q_1, Q_2, \cdots$ are iid with marginal probability density $f$. Let

$L(u)$ = ARL beginning from $u$

(primarily I'm interested in $L(\text{target})$)

Begin with an integral equation approach.

Notice first that there is no signal if

$$\text{target} - h < \lambda Q + (1-\lambda)u < \text{target} + h, \text{ i.e.,}$$

$$\underline{\text{target} - h - (1-\lambda)u} < \lambda Q < \underline{\text{target} + h - (1-\lambda)u}$$

or $$\underbrace{\frac{1}{\lambda}(\swarrow)}_{*} < Q < \underbrace{\frac{1}{\lambda}(\swarrow)}_{**}$$

and otherwise I get a signal.

Thus

$$L(u) = 1 \cdot P[Q_1 \notin (*, **)]$$

$$+ \int_{*}^{**} \left(1 + L(\lambda t + (1-\lambda)u)\right) f(t)\,dt$$

$$= 1 + \int_{*}^{**} L(\lambda t + (1-\lambda)u) f(t)\,dt$$

$$= 1 + \frac{1}{\lambda} \int_{\text{target}-h}^{\text{target}+h} L(y) f\left(\frac{y-(1-\lambda)u}{\lambda}\right) dy$$

(let $y = \lambda t + (1-\lambda)u$)

If I can solve for $L(u)$ I have ARLs.

To do so (to find approximate solutions) I need a quadrature rule for integrals on (target $-h$, target $+h$), i.e.,

points $\quad \text{target}-h \le a_1 \le a_2 \le \cdots \le a_m \le \text{target}+h$

and weights $\omega_j \ge 0 \quad \sum \omega_j = 2h$

So I can approximate

$$\int_{\text{target}-h}^{\text{target}+h} g(t)\,dt \approx \sum_{j=1}^{m} \omega_j\, g(a_j)$$

Then for $i = 1, 2, \cdots, m$ one can write

$$L(a_i) \approx 1 + \frac{1}{\lambda} \sum_{j=1}^{m} \omega_j\, L(a_j)\, f\left(\frac{a_j - (1-\lambda)a_i}{\lambda}\right).$$

This is a big set of linear equations in $m$ unknowns $L(a_1), L(a_2), \cdots, L(a_m)$ of the form

$$L \approx 1 + RL$$

for

$$\underset{m\times 1}{L} = \begin{pmatrix} L(a_1) \\ \vdots \\ L(a_m) \end{pmatrix} \quad \text{and} \quad \underset{m\times m}{R} = \left( \frac{\omega_j f\left(\frac{a_j-(1-\lambda)a_i}{\lambda}\right)}{\lambda} \right)$$

So if I can invert $m \times m$ matrices I write

$$L \approx (I-R)^{-1}\mathbf{1}$$

and get approximate values for $L(a_1), \ldots, L(a_m)$ and then

$$L(u) \approx 1 + \frac{1}{\lambda}\sum_{j=1}^{m} w_j L(a_j) f\left(\frac{a_j-(1-\lambda)u}{\lambda}\right)$$

---

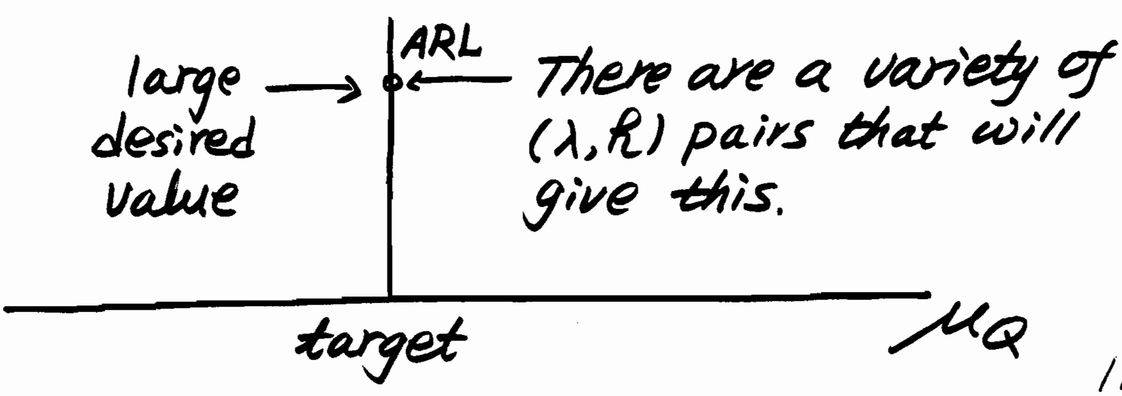
Pages 33&34 of the notes give a MC treatment of this EWMA ARL stuff and page 34 of the notes gives the approximate equivalence of the 2 methods.

Crowder computed ARL's for normal $Q$ — see his program on Professor Vardeman's Stat 531 Web page. — Table A.3 of V&J was created using Crowder's program.

---

EWMA Chart Design? (Choice of $\lambda$ and $h$)

For normal (iid) $Q$ with fixed $\sigma_Q$:



We can choose a $(\lambda, h)$ pair to optimize some off-target ARL or choose one with a "nice" $\lambda$ (for computational simplicity).

To optimize/minimize ARL at target $+\delta$ I use $\lambda$ specified in Table 4.2 (page 138) of V&J (these were obtained by Crowder using a numerical search).

$$\text{small } \delta \Rightarrow \lambda^{opt} \text{ small}$$
$$\text{large } \delta \Rightarrow \lambda^{opt} \text{ large (but } < 1)$$

Note that I expect EWMA to beat Shewhart for small changes in $\mu_Q$ from a target.

Having chosen $\lambda$ ("optimally" or not) pages 138 & 139 of V&J show how to choose $h$ to get a desired on-target ARL (assuming $EWMA_0$ = target).

CUSUM vs. EWMA:
Both provide ARL improvement over Shewhart. Some believe them to be equivalent... But they are <u>not</u>.

## Yashchin's Objection

Normal iid $Q$ with fixed $\sigma_Q$.

ARL functions for 2-sided schemes

$L_{CUSUM}(u,v|\mu_Q)$. $L_{EWMA}(u|\mu_Q)$

Design for

$L(\text{start}|\text{target}) = \text{ARL}_{\text{target}}$

$L(\text{start}|\text{target} \pm \delta) = \text{optimal}$

Then

$L_{CUSUM}(0,0|\text{target} \pm \delta) \approx L_{EWMA}(\text{target}|\text{target} \pm \delta)$

Optimally designed CUSUMs and EWMAs have about the same off-target ARLs from a place where CUSUMs have zeroed out and EWMA is half way between the control limits. However, note what happens if we get a process change when schemes are not at initial conditions:

$L_{CUSUM}(u,v|\text{target} \pm \delta) \leq L_{CUSUM}(0,0|\text{target} \pm \delta)$

i.e., regardless of where the CUSUM is at the moment of a shift, one is never worse off than if both high- and low-side schemes were zeroed out.

However, for the EWMA schemes things are different. If we get a shift of $\delta > 0$ when we are at EWMA $= u <$ target,
$L_{EWMA}(u \mid \text{target} + \delta) > L_{EWMA}(\text{target} \mid \text{target} + \delta)$,
and in fact the difference between these can be huge.

— i.e., EWMA has worse "worst case" behavior than the CUSUM.

By the way, there are some theoretical results that say CUSUMs are "optimal" as change detection procedures.

**Bottom Line:** Best available monitoring scheme is CUSUM + (loose) Shewhart

<u>§4.4 of V&J</u> one last piece of process monitoring / <u>ARL</u> X/MR

$x_1, x_2, \ldots$ individuals

At time $t \geq 1$ chart $x_t$ and for $t \geq 2$ also chart $MR_t = |x_t - x_{t-1}|$.

Alarming/signalling if

$$|x_t - \text{target}| > A \quad \text{or} \quad MR_t > B.$$

Suppose $x_i$ are iid with marginal density $f$ and let

$L(y)$ = mean # of additional observations required to cause an alarm if the current observation is $y$ (and no alarm to date).

Then

$$L(y) = 1 \cdot P\left[\, |x - \text{target}| > A \text{ or } |x - y| > B \right] + \int_{\max\{\text{target}-A,\, y-B\}}^{\min\{\text{target}+A,\, y+B\}} (1 + L(x)) f(x)\, dx$$

Picture: