An ARL-unbiased np-chart

Manuel Cabral Morais

maj@math.ist.utl.pt

Department of Mathematics & CEMAT — IST, ULisboa, Portugal

IST — Lisbon, September 2016

Agenda

1. Charts for nonconforming items
   - The $np-$chart with 3-sigma limits
   - Some variants
2. Eliminating the bias of the ARL function
   - A first attempt
   - Relating ARL-unbiased charts and UMPU tests
3. Illustrations
   - A few ARL-unbiased $np-$charts
   - A useful table and a curiosity
4. Final thoughts

In industrial processes, we can classify each **inspected item** as either conforming or **nonconforming** to a set of specifications.

The ***np*−chart with 3-sigma limits** has been historically used to detect changes in the fraction nonconforming (p):

- control statistic: number of nonconforming items in the $t-$th sample of size n, X_t
- distribution: $X_t \overset{indep.}{\sim} X \sim Binomial(n, p)$, $t \in \mathbb{N}$
- target mean: $n\, p_0$
- process mean: $n\, p = n\,(p_0 + \delta)$ $\quad$ ($\delta =$ magnitude of the shift in p)
- **3-sigma control limits**:

$$LCL = \left\lceil \max\left\{0,\ np_0 - 3\sqrt{np_0(1-p_0)}\right\} \right\rceil$$
$$UCL = \left\lfloor np_0 + 3\sqrt{np_0(1-p_0)} \right\rfloor$$

- triggers a signal and **deem the process out-of-control** at sample t if $X_t \notin [LCL, UCL]$.

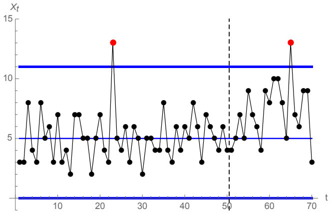
Example 1

- $n = 100$ (sample size), $p_0 = 0.05$ (target fraction nonconforming).
- Simulated data: first 50 samples — process is known to be in-control; last 20 samples — process out-of-control (increase in p, $p = p_0 + 0.006$).
- $3 - \sigma$ control limits

$$LCL = \left\lceil \max\left\{0,\ np_0 - 3\sqrt{np_0(1-p_0)}\right\} \right\rceil = 0$$

$$UCL = \left\lfloor np_0 + 3\sqrt{np_0(1-p_0)} \right\rfloor = 11$$

- $np-$chart

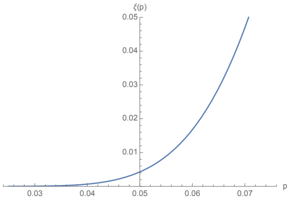


One false alarm, sample 23; one valid signal, sample 65.

Example 1 (cont'd)

Parallels with a repeated hypothesis test...

- $H_0 : p = p_0$ (process is in-control)
  $H_1 : p \neq p_0$ (process is out-of-control)
- control statistic: $T = \frac{X - n\,p_0}{\sqrt{n\,p_0(1-p_0)}} \overset{a}{\sim}_{H_0} \text{Normal}(0,1)$
- rejection region: $W = (-\infty, -3) \cup (3, +\infty)$
- exact power function: $\xi(p) = P(T \in W \mid p), \quad p \in (0,1)$



problems

- minimum of $\xi(p)$ not achieved at $p_0 \quad \Rightarrow \quad \xi(p) < \xi(p_0),\ p < p_0$
- significance level: $\xi(p_0) = 0.004274 \neq 0.0027 \simeq 1 - [\Phi(3) - \Phi(-3)]$.

Performance

- $\xi(p) = P(\text{emission of a signal} \mid p) = 1 - \sum_{x=LCL}^{UCL} \binom{n}{x} p^x (1-p)^{n-x}$.
- **Run length** (RL) — number of samples taken until a signal is triggered $RL(p) \sim \text{Geometric}(\xi(p))$.
- The performance is frequently assessed in terms of $ARL(\delta) = 1/\xi(\delta)$.
  It is desirable that false alarms (resp. valid signals) are rarely triggered (resp. emitted as quickly as possible), corresponding to a large in-control (resp. small out-of-control) ARL.
- In most practical applications $p_0 \leq 9/(9+n)$, thus $LCL = 0$ and $ARL(p) > ARL(p_0)$, $p \in (0, p_0)$, i.e., the chart triggers false alarms more frequently than valid signals in the presence of any decrease in p.
- Selecting the smallest sample size n_{min} verifying $n > 9(1-p_0)/p_0$ to deal with $LCL > 0$, can lead to impractical sample sizes (e.g., $p_0 = 0.001$, $n_{min} = 8992$).
- The 3-sigma control limits presume the adequacy of the normal approximation to the binomial distribution, often a poor approximation.

Variants to mitigate the poor performance of the $np-$chart with 3-sigma limits basically rely on:

- **transformations**[1] traced back to
  - Freeman and Tukey (1950), $y = 0.5\left[\arcsin\sqrt{x/(n+1)} + \arcsin\sqrt{(x+1)/(n+1)}\right]$
  - Hald (1952, p. 685), $y = \arcsin\sqrt{x/n}$
  - Johnson and Kotz (1969, p. 65), $y = \arcsin\sqrt{(x+3/8)/(n+3/4)}$;
- **modified control limits**[2] obtained by regression against np_0 and $\sqrt{np_0}$, for $p_0 \in (0, 0.03]$ (Ryan and Schwertman, 1997)
  - $LCL = 2.9529 + 1.01956\, np_0 - 3.2729\sqrt{np_0}$
  - $UCL = 0.6195 + 1.00523\, np_0 + 2.983\sqrt{np_0}$.

All resulting charts are ARL-biased, i.e., the ARL function does not attain a maximum at $p = p_0$.

[1] Transform the binomial data (x) so that the transformed data (y) are approximately normal, and use 3-sigma limits for the transformed data (Ryan, 1989, p. 182).

[2] Search for values of n that would lead to control limits associated with in-control tail areas very close to the nominal value of 0.0027×0.5.

Example 2

- $n = 1267,\ p_0 = 0.01$
  $\alpha^{-1} = 1/0.0027 \simeq 370.4$ (desired in-control ARL).

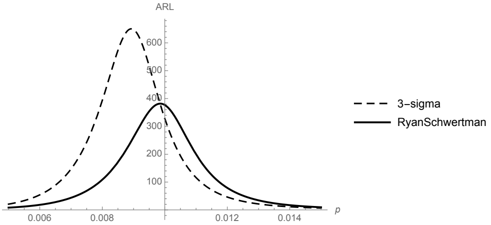


Chart	$[LCL, UCL]$	Max. of ARL	Relat. bias of ARL	In-control ARL
3-sigma	[3, 23]	650.419	−10.723%	327.976
RS	[4, 24]	381.718	−1.449%	376.811

- It takes longer, in average, to detect some shifts in p than to trigger a false alarm!

The first attempt to correct the bias of the ARL function of the $np-$chart is attributed to Acosta-Mejía (1999).

- By differentiating the probability of triggering a signal with respect to p and conditioning this derivative to be equal to zero when $p = p_0$:

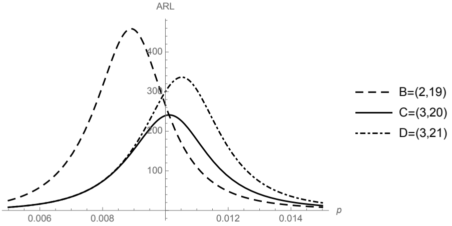
$$\frac{p_0^{LCL-1}\,(1-p_0)^{n-LCL}}{\Gamma(n-LCL+1)\,\Gamma(LCL)} = \frac{p_0^{UCL}\,(1-p_0)^{n-UCL-1}}{\Gamma(n-UCL)\,\Gamma(UCL+1)}.$$

- This equation defines the unbiased performance line (UPL) and leads in general to non-integer control limits.
- Acosta-Mejía (1999) suggested the adoption of the pair of integers closest to the intersection point of the UPL and the iso-ARL curve that defines all pairs (LCL, UCL) having the same desired in-control ARL.
- The resulting chart is ARL-biased, yet Acosta-Mejía (1999) termed it *nearly* ARL-unbiased $np-$chart.

Example 4

- $n = 1000$, $p_0 = 0.01$
- ARL curves associated with the (LCL, UCL) closest to the intersection of the UPL and the iso-ARL curve for a desired in-control ARL equal to 300:

Out[39]=



(LCL, UCL)	Maximum of ARL	Relative bias of the ARL	In-control ARL
$B = (2, 19)$	458.698	-10.901%	265.421
$C = (3, 20)$	241.056	$+1.237\%$	239.469
$D = (3, 21)$	336.472	$+5.219\%$	300.187

- The smallest relative bias corresponds to $C = (3, 20)$, however the associated $np-$chart has the in-control ARL furthest from 300.

Basic facts

- A size α test for $H_0 : p = p_0$ against $H_1 : p \neq p_0$, with power function $\xi(p)$, is said to be **unbiased** if $\xi(p_0) \leq \alpha$ and $\xi(p) \geq \alpha$, for $p \neq p_0$.
  The test is at least as likely to reject under any alternative as under H_0;

  $$ARL(p_0) \geq \alpha^{-1} \quad \text{and} \quad ARL(p) \leq \alpha^{-1},\ p \neq p_0.$$

- If we consider $\mathcal{C}$ a class of tests for $H_0 : p = p_0$ against $H_1 : p \neq p_0$, then a test in $\mathcal{C}$, with power function $\xi(p)$, is a **uniformly most powerful (UMP)** class $\mathcal{C}$ **test** if $\xi(p) \geq \xi'(p)$, for every $p \neq p_0$ and every $\xi'(p)$ that is a power function of a test in class $\mathcal{C}$.

- In this situation there is no UMP test, but there is a test which is UMP among the class of all unbiased tests — the **uniformly most powerful unbiased (UMPU) test**.

- The concept of an ARL-unbiased Shewhart-type chart is related to the notion of UMP test.

Basic facts (cont'd)

- The UMPU test derived by Lehmann (1959, pp. 128–129, Example 1) for the parameter p of the binomial distribution uses the critical function

$$\phi(x) = P(\text{Reject } H_0 | X = x) = \begin{cases} 1 & \text{if } x < LCL \text{ or } x > UCL \\ \gamma_{LCL} & \text{if } x = LCL \\ \gamma_{UCL} & \text{if } x = UCL \\ 0 & \text{if } LCL < x < UCL, \end{cases}$$

where LCL, UCL, γ_{LCL}, and γ_{UCL} are such that

$$\begin{aligned} E_{n,p_0}[\phi(X)] &= \alpha & &\text{(prob. of false alarm } = \alpha) \\ E_{n,p_0}[X\,\phi(X)] &= \alpha\, E_{n,p_0}(X) & &\text{(unbiased ARL).} \end{aligned}$$

Equivalently,

$$\begin{aligned} \gamma_{LCL} \times P_{n,p_0}(LCL) &+ \gamma_{UCL} \times P_{n,p_0}(UCL) \\ &= \alpha - \left[1 - \sum\nolimits_{x=LCL}^{UCL} P_{n,p_0}(x)\right] \\ \gamma_{LCL} \times LCL \times P_{n,p_0}(LCL) &+ \gamma_{UCL} \times UCL \times P_{n,p_0}(UCL) \\ &= \alpha \times np_0 - \left[np_0 - \sum\nolimits_{x=LCL}^{UCL} x \times P_{n,p_0}(x)\right]. \end{aligned}$$

Basic facts (cont'd)

- However, the two previous equations are not sufficient to define two control limits and two randomization probabilities.

Characterizing the ARL-unbiased $np-$chart

Inspired by this UMPU test, we defined a $np-$**chart that triggers a signal** with:

- **probability one** if the sample **number of nonconforming items, x, is below LCL or above UCL**;
- **probability** γ_{LCL} (resp. γ_{UCL}) if $x = LCL$ (resp. $x = UCL$).

Furthermore,

- **randomization probabilities**

  solution of a system of linear equations:

$$\gamma_{LCL} = \frac{d\,e - b\,f}{a\,d - b\,c} \qquad \text{and} \qquad \gamma_{LCL} = \frac{a\,f - c\,e}{a\,d - b\,c},$$

where $a = P_{n,p_0}(LCL)$, $b = P_{n,p_0}(UCL)$, $c = LCL \times P_{n,p_0}(LCL)$, $d = UCL \times P_{n,p_0}(UCL)$, $e = \alpha - 1 + \sum_{x=LCL}^{UCL} P_{n,p_0}(x)$, $f = \alpha \times n\,p_0 - n\,p_0 + \sum_{x=LCL}^{UCL} x \times P_{n,p_0}(x)$, and $a\,d - b\,c \neq 0$.

Characterizing the ARL-unbiased *np*−chart (cont'd)

- **Control limits** (and randomization probabilities)

  Bear in mind that giving protection to decreases (resp. increases) in p means a LCL (resp. UCL) as large (resp. small) as possible.

  Thus, in order to rule out control limits leading to $(\gamma_{LCL}, \gamma_{UCL}) \notin (0,1)^2$, (LCL, UCL) should be restricted to the following set of non-neg. integer:

  - $\{(LCL(\alpha),\, UCL_{LCL(\alpha)}),\, (LCL(\alpha),\, UCL_{LCL(\alpha)}+1),$
    $(LCL(\alpha)-1, UCL_{LCL(\alpha)-1}),\, (LCL(\alpha)-1, UCL_{LCL(\alpha)-1}+1), \ldots,$
    $(0, UCL_0),\, (0, UCL_0+1)\},$

  where

  - $LCL(\eta)$ is the largest non-neg. integer $LCL : P(X < LCL \mid p = p_0) \leq \eta$,
  - $\alpha_{LCL(\eta)} = P(X < LCL \mid p = p_0)$ is the lower tail in-control area associated with $LCL(\eta)$,
  - $UCL_{LCL(\eta)} = F^{-1}_{n,p_0}[1-(\alpha-\alpha_{LCL(\eta)})]$ is the corresponding UCL.

  The search for values for $(\gamma_{LCL}, \gamma_{UCL})$ starts with $(LCL(\alpha),\, UCL_{LCL(\alpha)})$ and stops as soon as an admissible solution is found (Mathematica program).

Characterizing the ARL-unbiased $np-$chart (cont'd)

- **ARL function**

  A signal is triggered by the ARL-unbiased $np-$chart with probability

$$\xi_{unbiased}(p) = \left[1 - \sum_{x=LCL}^{UCL} P_{n,p}(x)\right] + \gamma_{LCL} \times P_{n,p}(LCL) + \gamma_{UCL} \times P_{n,p}(UCL)$$

  and the corresponding ARL function is given by $1/\xi_{unbiased}(p)$.

- **Randomization of the emission of the signal**

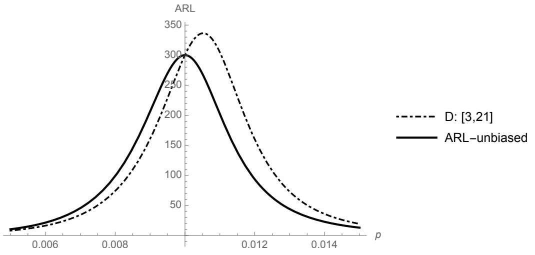
  Can be done in practice by incorporating the generation of a pseudo-random number from a Bernoulli distribution with parameter γ_{LCL} (resp. γ_{LCL}) in the software used to monitor the data fed from the production line, whenever the observed number of nonconforming items is equal to LCL (resp. UCL).

- **ARL-unbiased $p-$chart**

  The conversion to the corresponding ARL-unbiased $p-$chart is evidently made by dividing the control limits by n.

Example 5

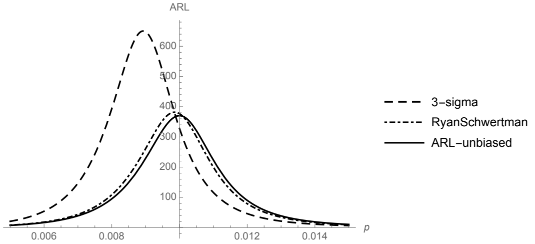
- $n = 1000,\ p_0 = 0.01,\ \alpha = 1/300$
- Acosta-Mejía's $np-$chart
  $[LCL, UCL] = [3, 21]$ (in-control ARL very close to 300)
- ARL-unbiased $np-$chart
  $[LCL, UCL] = [2, 21]$, $(\gamma_{LCL}, \gamma_{UCL}) = (0.673094, 0.853994)$



- Acosta-Mejía's $np-$chart outperforms (resp. is outperformed by) the ARL-unbiased $np-$chart in the detection of decreases (resp. increases) in p.

Example 6

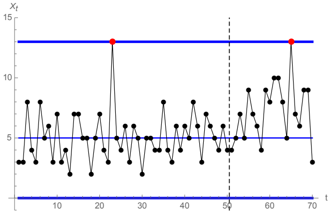
- $n = 1267$, $p_0 = 0.01$, $\alpha = 0.0027$
- $np-$chart with 3-sigma limits: $[LCL, UCL] = [3, 23]$
- Ryan & Schwertmann's $np-$chart: $[LCL, UCL] = [4, 24]$
- ARL-unbiased $np-$chart
  $[LCL, UCL] = [4, 25]$, $(\gamma_{LCL}, \gamma_{UCL}) = (0.076400, 0.713818)$



- The elimination of the bias of the ARL function is due to the adoption of the quantile based control limits and the randomization probabilities.

Example 7

- $n = 100$, $p_0 = 0.05$, $\alpha = 0.0027$
- Simulated data: first 50 samples — process is known to be in-control; last 20 samples — process out-of-control (increase in p, $p = p_0 + 0.006$).
- ARL-unbiased $np-$chart
  $[LCL, UCL] = [0, 13]$, $(\gamma_{LCL}, \gamma_{UCL}) = (0.289066, 0.524741)$



- A red • corresponds now to an obs. responsible for a signal because it is: beyond $[LCL, UCL]$; or equal to LCL (resp. UCL) and the corresp. gen. pseudo-random no. from the Bernoulli dist. with parameter γ_L (resp. γ_U) equals 1.
  One false alarm, sample 23, valid signal, sample 65, both due to randomization.

p_0	n	$[LCL, UCL]$	$(\gamma_{LCL}, \gamma_{UCL})$	$\lfloor n/10 \rfloor$	$[LCL, UCL]$	$(\gamma_{LCL}, \gamma_{UCL})$
0.005	1324	[1, 16]	(0.039089, 0.642052)	132	[0,5]	(0.004567, 0.554449)
0.01	664	[1, 16]	(0.045716, 0.646175)	66	[0,5]	(0.004573, 0.599389)
	1267	[4, 25]	(0.076399, 0.713818)	126	[0,6]	(0.007775, 0.141892)
0.02	533	[3, 22]	(0.017480, 0.683500)	53	[0,6]	(0.006534, 0.655018)
	708	[5, 27]	(0.017478, 0.712931)	70	[0,6]	(0.008927, 0.008561)
0.03	357	[3, 22]	(0.045553, 0.691577)	35	[0,6]	(0.006507, 0.795807)
	474	[5, 27]	(0.059043, 0.716629)	47	[0,6]	(0.009037, 0.027439)
	874	[13, 43]	(0.051330, 0.737303)	87	[0,9]	(0.027701, 0.485802)
	923	[14, 45]	(0.089674, 0.865971)	92	[0,9]	(0.031914, 0.228522)
0.04	218	[2, 19]	(0.038876, 0.702542)	21	[0,5]	(0.005362, 0.269840)
	268	[3, 22]	(0.062363, 0.772698)	26	[0,6]	(0.006481, 0.966599)
	393	[6, 29]	(0.029994, 0.744246)	39	[0,7]	(0.010418, 0.655596)
	620	[12, 41]	(0.084557, 0.672194)	62	[0,9]	(0.024866, 0.906530)
	755	[15, 48]	(0.990580, 0.771784)	75	[0,10]	(0.040897, 0.969212)
	893	[20, 55]	(0.071036, 0.844692)	89	[0,11]	(0.070106, 0.978386)
0.05	175	[2, 19]	(0.056816, 0.741418)	17	[0,5]	(0.005422, 0.299790)
	315	[6, 29]	(0.064331, 0.804571)	31	[0,7]	(0.010397, 0.784981)
	345	[7, 31]	(0.034659, 0.759024)	34	[0,7]	(0.011903, 0.379684)
	466	[11, 39]	(0.082700, 0.734550)	46	[0,8]	(0.021011, 0.218476)
	606	[16, 48]	(0.092161, 0.756650)	60	[0,9]	(0.041372, 0.026530)
0.1	104	[3, 21]	(0.041042, 0.732947)	10	[0,5]	(0.006285, 0.243150)
	139	[5, 26]	(0.071442, 0.813540)	13	[0,6]	(0.008391, 0.569880)
	154	[6, 28]	(0.030245, 0.745744)	15	[0,6]	(0.010116, 0.158061)
	229	[11, 38]	(0.050072, 0.800832)	22	[0,8]	(0.020094, 0.784673)
	299	[16, 47]	(0.086104, 0.864793)	29	[0,9]	(0.039811, 0.404286)
	339	[19, 52]	(0.076982, 0.853738)	33	[0,10]	(0.059491, 0.760501)

These values coincide with the ones recently obtained with the R package *ump*.

We came a long way since Shewhart proposed the $p-$chart in the 1920s...

- **An ARL-unbiased $np-$chart**
    - It has a **pre-specified in-control ARL**, as opposed to the $np-$chart with 3-sigma control limits or existing alternatives.
    - The associated **ARL curve attain a maximum when p is on target**, i.e., any shift in p leads to a valid signal triggered in less time, in average, than a false alarm.
    - It **tackles the curse of the null LCL** and detects decreases in p in a timely fashion, by relying on the **randomization probabilities**.

- **Future work**
    - Derive an ARL-unbiased version of the **CUSUM chart/scheme for binomial data**, in order to improve the detection of small-to-moderate shifts in p.
      Since the control statistics of the CUSUM chart/scheme are dependent r.v., we have to resort to different search methods to determine the control limits and randomization probabilities.

Related statistical inference papers et al. found while preparing this seminar

- Geyer, C.J. and Meeden, G.D. (2004). ump: An R package for UMP and UMPU tests. Available at `www.stat.umn.edu/geyer/fuzz/` (only binomial distribution)
- Geyer, C.J. and Meeden, G.D. (2005). Fuzzy and randomized confidence intervals and p-values. *Statistical Science* **20**, 358–366.

Related SPC papers by submission date

- Paulino, S., Morais, M.C. and Knoth, S. (2016a). An ARL-unbiased c-chart. Accepted for publication in *Quality and Reliability Engineering International*, http://onlinelibrary.wiley.com/doi/10.1002/qre.1969/epdf (Different search algorithm, DSA; R program)
- Morais, M.C. (2016a). An ARL-unbiased np-chart. *Economic Quality Control* **31**, 11–21.
- Paulino, S., Morais, M.C. and Knoth, S. (2016b). On ARL-unbiased c-charts for INAR(1) Poisson counts. Submitted for publication in *Statistical Papers*.
- Morais, M.C. (2016b). ARL-unbiased geometric and CCC_G control charts. Submitted for publication in *International Journal of Production Research*. (DSA; R program)
- Morais, M.C. and Knoth, S. (2016). On ARL-unbiased charts to monitor the traffic intensity of a single server queue. Proceedings of the XIIth. International Workshop on Intelligent Statistical Quality Control, 217-242. http://www.hsu-hh.de/compstat/index_8sVJz3C3s0oQzk3M.html