Joint distribution of longest run and number of crossings in run charts

Tore Wentzel-Larsen*[1,2], Jacob Anhøj[3]
[1]Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, [2]Norwegian Centre for Violence and Traumatic Stress Studies,
[3]Centre of Diagnostic Investigation, Rigshospitalet, University of Copenhagen, Denmark, *Presenting author, tore.wentzellarsen@gmail.com

Context

The Anhøj rules, implemented in the R package qicharts2 (Anhøj 2019), identify special cause variation in run charts based on an unusually long run or unusually few crossings of a predetermined median. The performance of these rules depends on the joint distribution of the number of crossings and the longest run.

Objectives

Compute this joint distribution of the number of crossings C and the longest run L, for any above median probability p, where p = 0.5 corresponds to the symmetric case with no real special cause variation. Use the joint distribution to propose a modification of the Anhøj rules, and compute probabilities interpretable as degree of evidence of special cause variation in a specific run chart.

Methods

Computation of the joint distribution is implemented in the R package crossrun (Wentzel-Larsen & Anhøj 2018), as detailed in the crossrun vignette. The Anhøj rules declare no special cause variation in a *box*, C ≥ ca, L ≤ la, corresponding to sufficiently many crossings and a sufficiently short longest run. The limits ca and la are set to give a high specificity, the probability of declaring no special cause variation when there is none (p = 0.5), and also a high sensitivity, the probability of declaring special cause variation when p ≠ 0.5. A modification of the Anhøj rules is proposed based on a lower threshold specificity (≥ 0.925), while maximizing sensitivity for a specified higher p = pnorm(0.8) = 0.79, corresponding to a standard normal deviate with shift 0.8 (target shift). Evidence for special cause variation is quantified as a low probability for a result at least as deviant (no more crossings or longest run not shorter) as observed, under the symmetric assumption. This is illustrated in Figure 1 below for chart size 11, the box was C ≥ 2, L ≤ 6 by the Anhøj (red) and C ≥ 3, L ≤ 7 by the modified (green) rules. Probabilities are shown in crossrun's times representation, i.e. multiplied by 2^(11-1)=1024 to avoid small numbers displayed.

Figure 1: Joint distribution of number of crossings and longest run

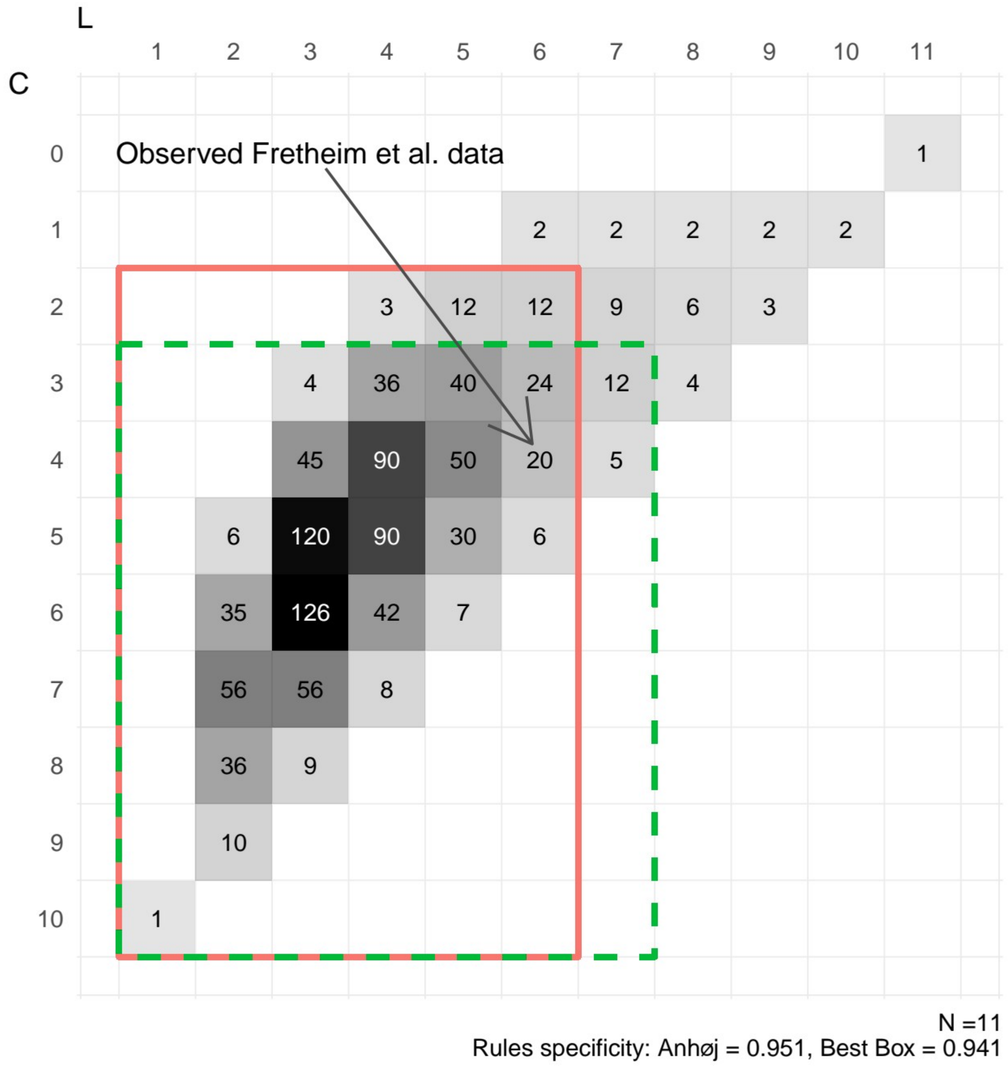
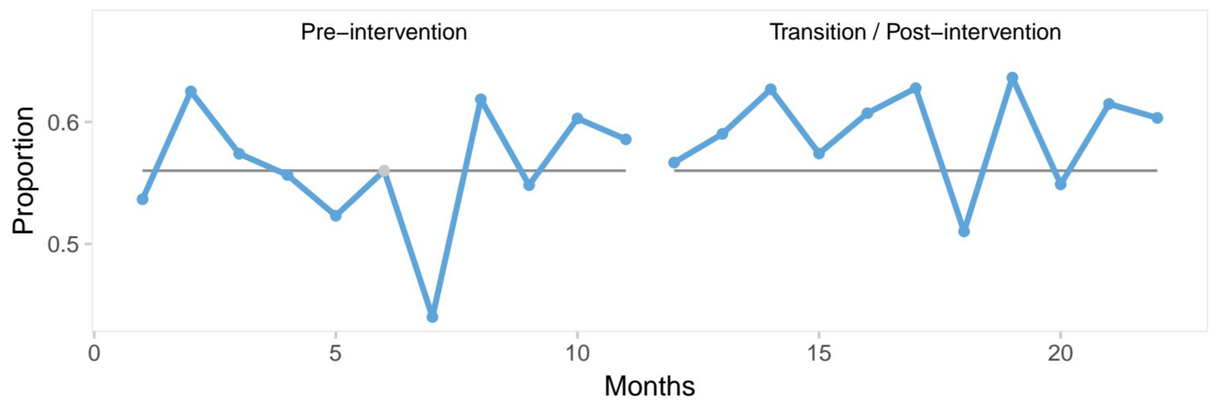


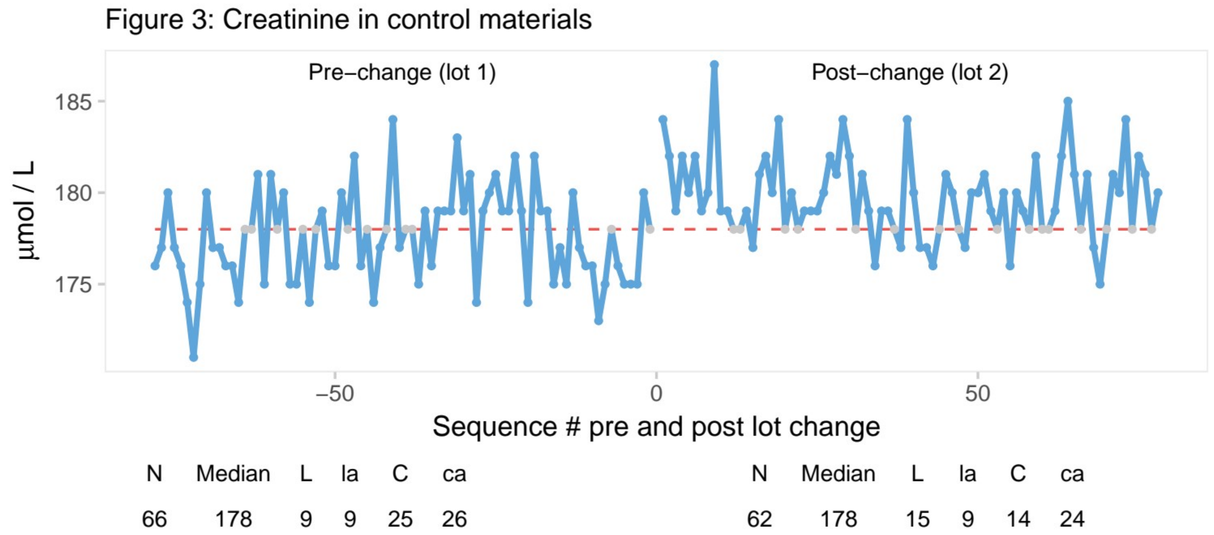
Figure 2: Patients receiving treatment goals



N	Median	L	la	C	ca
10	0.56	3	6	5	2

N	Median	L	la	C	ca
11	0.56	6	6	4	2

Figure 3: Creatinine in control materials



N	Median	L	la	C	ca
66	178	9	9	25	26

N	Median	L	la	C	ca
62	178	15	9	14	24

Results

For patients achieving treatment goals illustrated in Fretheim et al (2007, fig. 3), the pre-intervention median was 0.56 (Figure 2, data kindly made available by the authors). There were 4 crossings and longest run was 6 in the 11 transition/post-intervention period observations, within the box in Figure 1 for both rule specifications. There was no evidence for special cause variation since P(C ≤ 4 or L ≥ 6) = 0.38 assuming symmetry.

For creatinine concentrations in one control material at the Department of Medical Biochemistry and Pharmacology, Haukeland University Hospital, Norway, 78 observations after each change of reagent lot were compared with the pre-change median (data kindly made available by Marit Sverresdotter Sylte). After the first lot change (Figure 3), 62 of 78 post-change observations were different from the pre-change median. Among these, C = 14 and L = 15, which indicate special cause variation since the Anhøj rules for series length 62 declare no special cause variation when C ≥ 24 and L ≤ 9 (modified rules C ≥ 25 and L ≤ 11). The probability P(C ≤ 14 or L ≥ 15) was 0.0015 under the symmetric assumption signifying special cause variation. Even for shift 0.4 this probability is as low as 0.031. Note that special cause variation was also present in data from the first lot assuming symmetry around its own median.

Conclusions

The joint distributions are useful for assessing observed values of the number of crossings and the longest run in run charts. Further work is ongoing for refining the modified rules, and also for extending the crossrun R package with functions for the case where the median is determined from the same data series.

References

- Wentzel-Larsen T, Anhøj J (2018). crossrun: Joint Distribution of Number of Crossings and Longest Run. R package version 0.1.0.
- Anhøj J (2019). qicharts2: Quality Improvement Charts. R package version 0.6.0.
- Fretheim A, Håvelsrud K, MacLennan G, Kristoffersen DT, Oxman AD (2007). The effects of mandatory prescribing of thiazides for newly treated, uncomplicated hypertension: interrupted time-series analysis. PLOS Medicine 4(7): e232. https://doi.org/10.1371/journal.pmed.0040232.

Regional Centre for Child and Adolescent Mental Health
Eastern and Southern Norway
NORWEGIAN CENTRE FOR VIOLENCE AND TRAUMATIC STRESS STUDIES
REGION
Rigshospitalet