Change-point Problems
IMS Lecture Notes – Monograph Series (Volume 23, 1994)

A RANK-CUSUM PROCEDURE FOR DETECTING SMALL CHANGES IN A SYMMETRIC DISTRIBUTION

BY P. K. BHATTACHARYA AND HONG ZHOU
University of California, Davis

A rank-sign analogue of the Page-CUSUM procedure is introduced here for detecting location change in a sequentially observed series of data following a symmetric distribution. This Rank-CUSUM procedure is asymptotically equivalent to the Page-CUSUM procedure if the score functions used in both procedures are appropriate for the underlying density; but even otherwise, it maintains its prescribed false alarm rate and has good detection property. Results of simulation studies comparing the Rank-CUSUM and the Page-CUSUM procedures are reported.

1. Introduction. For the problem of sequential detection of change in distribution, the CUSUM chart proposed by Page (1954) is widely accepted for its simplicity. The stopping rule for the procedure also has a minimax property as shown by Lorden (1971) and Moustakides (1986). However, one has to recognize the fact that the method is based on strict distributional assumptions and may perform poorly when these assumptions fail.

In this paper, we consider a rank-sign analogue of the Page-CUSUM procedure for the situation where one wants to detect a small change in location of a symmetric density without knowing the actual form of the density. If the score function used to calculate the rank-CUSUM's is appropriately chosen for the underlying density, then this procedure is asymptotically equivalent to the Page-CUSUM procedure (based on the true density); but even otherwise, the rank-CUSUM procedure maintains its prescribed false alarm rate and also has good detection property.

In Section 2, the rank-CUSUM 's are introduced and the weak convergence properties of both types of CUSUM are discussed. Some simulation results comparing the performance of the rank-CUSUM procedure with that of the Page-CUSUM procedure are given in Section 3. Technical details of

---

AMS 1991 Subject Classification: Primary 60G40, 62G99; Secondary 62N10, 62L15.

Key words and phrases: Rank-CUSUM, Page-CUSUM, sequential detection, contiguous change, weak convergence.

the theoretical properties outlined in Section 2 will be provided in a separate publication.

2. Page-CUSUM, Rank-CUSUM and Their Asymptotic Properties. Suppose that we observe independent random variables $\{X_i\}$ sequentially, of which the Lebesgue density undergoes a location change from $f(\cdot)$ to $f(\cdot - \Delta)$ by a specified $\Delta > 0$, starting with $X_{\tau+1}$, where f is an unknown density which is symmetric about 0 and τ is an unknown change-point.

The Page-CUSUM procedure attempts to detect change from a known density f_0 to a known density f_1 by applying a likelihood ratio test for the null hypothesis $H_k : \tau \geq k$ against alternative $H'_k : 0 \leq \tau \leq k-1$ at each stage of sampling and stops at the first time k at which the log likelihood ratio crosses a constant boundary. In our framework of location change in a symmetric density, this requires the symmetric density f to be known. However, in practice, one does not know f but assumes it to be of a certain form, say g, which is not necessarily the same as the true f, and based on this symmetric density g, computes the log likelihood ratio at the k-th stage as

$$\begin{aligned} T_k(g) &= \max_{0\leq j<k} T_{jk}(g), \\ T_{jk}(g) &= \sum_{i=j+1}^{k} \log\left[\frac{g(X_i-\Delta)}{g(X_i)}\right] \end{aligned} \tag{1}$$

The Page-CUSUM stopping rule resulting from this is given by

$$N(c,g) = \min\{k : T_k(g) \geq c\} \tag{2}$$

where the constant c is so chosen as to control the false alarm rate by maintaining a specified value of $E[N(c)]$ or $P[N(c) > n]$ for given n when there is no change.

In the rank-CUSUM procedure, we use the signs and the ranks of absolute values of $(X_1, \ldots, X_k)$, viz.,

$$\begin{aligned} s(X_i) &= \text{sign}(X_i), \\ R^+_{k:i} &= \text{ rank of } |X_i| \text{ among } |X_1|, \ldots, |X_k|, \qquad 1 \leq i \leq k \end{aligned}$$

to construct a rank-sign analogue of the log likelihood ratio statistic for H_k against H'_k. For $n = 1, 2, \ldots$, let

$$a_n(i,\varphi) = E[\varphi(U_{n:i})], \qquad 1 \leq i \leq n \tag{3}$$

where φ is a square-integrable function on $[0,1]$ and $U_{n:1} < \cdots < U_{n:n}$ are the order statistics in a random sample from uniform $(0,1)$, and define

$$\begin{aligned} S_k(\varphi) &= \max_{0 \le j < k} S_{jk}(\varphi), \\ S_{jk}(\varphi) &= \sum_{i=j+1}^{k} s(X_i) a_k(R^+_{k:i}, \varphi) - \frac{1}{2}(k-j)\Delta\|\varphi\|^2. \end{aligned} \tag{4}$$

Then the Rank-CUSUM stopping rule is given by

$$N^*(c, \varphi) = \min\{k : S_k(\varphi) \ge c\}, \tag{5}$$

where the constant c is chosen as in (2) to control the false alarm rate. In practice, the score function φ will be taken to be

$$\varphi(u) = \varphi^+(u, g) = -\frac{g' \circ G^{-1}(\frac{1}{2} + \frac{1}{2}u)}{g \circ G^{-1}(\frac{1}{2} + \frac{1}{2}u)} \tag{6}$$

where $g = G'$ is a symmetric pdf which is the assumed model for the true pdf f. For simplicity of notation, we shall write

$$\begin{aligned} S_{jk}(\varphi^+(\cdot, g)) &= S_{jk}(g), \\ S_k(g) &= \max_{0 \le j < k} S_{jk}(g), \\ N^*(c, g) &= \min\{k : S_k(g) \ge c\}, \end{aligned} \tag{7}$$

using $\varphi(\cdot) = \varphi^+(\cdot, g)$ in (4) and (5).

To see the connection between the Rank-CUSUM and the Page-CUSUM procedures, let

$$R^+_{jk:i} = \text{ rank of } |X_i| \text{ among } |X_{j+1}|, \ldots, |X_k|, \qquad 0 \le j < i \le k,$$

and note that the standard approximation to the log likelihood ratio of H'_k to H_k for small $\Delta > 0$, based on ranks and signs, if g is assumed to be the true form of the density (see Hájek and Šidák (1967)), is obtained if we replace $a_k(R^+_{k:i}, \varphi)$ in formula (4) for $S_{jk}(\varphi)$ by $a_{k-j}(R^+_{jk:i}, \varphi^+(\cdot, g))$.

We now compare the Page-CUSUM stopping rule $N(c, g)$ and the rank-CUSUM stopping rule $N^*(c, g)$, both based on a symmetric pdf g, possibly different from the true f.

Let P_n and Q_n denote the distributions of $(X_1, \ldots, X_n)$ with joint densities

$$p_n(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i),$$

$$q_n(x_1, \ldots, x_n) = \prod_{i=1}^{[n\theta]} f(x_i) \prod_{i=[n\theta]+1}^{n} f(x_i - \delta/\sqrt{n})$$

corresponding to $\tau \geq n$ and $\tau = [n\theta]$ respectively, with $\Delta = \delta/\sqrt{n}$, where $\delta > 0$ and $0 < \theta < 1$. The basis for our comparison of the stopping rules will be the weak convergence properties under P_n and under the contiguous distributions Q_n, of the normalized versions

$$Y_n(s,t;g) = \frac{S_{[ns],[nt]}(g)}{\sqrt{n}\|\varphi^+(\cdot,g)\|}$$

$$Z_n(s,t;g) = \frac{T_{[ns],[nt]}(g)}{\delta\sqrt{I(g)}}$$

on $0 \leq s \leq t \leq 1$ of $\{S_{jk}(g), 0 \leq j \leq k \leq n\}$ and $\{T_{jk}(g), 0 \leq j \leq k \leq n\}$, where $I(f)$ is the Fisher-information of f and $S_{kk}(g) = T_{kk}(g) = 0$.

If in both $S_{jk}(g)$ given by (7) and $T_{jk}(g)$ given by (1), one uses $g = f$, i.e., if the score function for each procedure is based on the true f, then the weak limits of $\{Y_n(s,t;f)\}$ and $\{Z_n(s,t;f)\}$ are identical under P_n and are also identical under Q_n. The common weak limit of the two processes is

$$\left\{B(t) - B(s) - \frac{1}{2}\delta\sqrt{I(f)}(t-s), \qquad 0 \leq s \leq t \leq 1\right\} \tag{8}$$

under P_n , where $\{B(t), t \geq 0\}$ is a standard Brownian motion, and both processes have another deterministic component

$$\delta\sqrt{I(f)}[(t-\theta)I_{(\theta,1]}(t) - (s-\theta)I_{(\theta,1]}(s)]$$

added to their common limit under Q_n. Thus in the contiguous change model, the Rank-CUSUM and the Page-CUSUM procedures have identical limiting behavior both with respect to false alarm and correct detection, if they both use the score function corresponding to the true f.

Now consider the case where the symmetric density g used in $S_{jk}(g)$ and $T_{jk}(g)$ is different from f. In such a case, the weak limit of $\{Y_n(s,t;g)\}$ under P_n is obtained by replacing $\sqrt{I(f)}$ by $\sqrt{I(g)}$ in (8) and under Q_n is obtained by multiplying the deterministic component in (9) by $\rho(f,g)$ = correlation coefficient between $\varphi^+(\cdot,f)$ and $\varphi^+(\cdot,g)$. Note that if there is no change, then $\{Y_n(s,t;g)\}$ is distribution-free in the limit, i.e., its limiting distribution does not depend on f, as should be the case. On the other hand, if there is a contiguous change, then the deterministic component in (9) which drives $S_k(g) = \max_{0 \leq j < k} S_{jk}(g)$ towards the decision boundary after change gets weakened by the factor $\rho(f, g)$. This factor ρ, although less than 1, may still attain a reasonable level, and is always positive if f and g are both symmetric unimodal. This is an improvement upon the stopping rule based on sequential ranks considered by Bhattacharya and Frierson(1981), which is driven towards the decision boundary by a logarithmic drift after change.

The limiting properties of the Rank-CUSUM described above only requires f and g to have finite and positive Fisher-information and absolutely continuous derivatives. However, the limiting properties of the Page-CUSUM using $g \neq f$ hold only under more stringent conditions and the false alarm rate itself varies drastically with the choice of g, as will be demonstrated by the simulation results in the next section. The actual limiting distributions of Page-CUSUM using $g \neq f$ are somewhat complicated and will not be discussed here.

The above weak convergence properties of the Z_n-process follow in a straightforward manner from an expansion of $\log\{f(X_i - \delta n^{-1/2})/f(X_i)\}$ to the second order terms. Corresponding results for the Y_n-process are first established under $\{P_n\}$ by proving the convergence of finite-dimensional distributions (fdd) via Hájek-projection and using a martingale argument to prove tightness. By contiguity of $\{Q_n\}$ to $\{P_n\}$, convergence of fdd's of the Y_n-process under $\{Q_n\}$ is then obtained by LeCam's Third Lemma while tightness under $\{Q_n\}$ becomes automatic. Details will be given in a separate publication.

3. Simulation Results. In these simulation experiments we consider Rank-CUSUM and Page-CUSUM procedures with scores based on certain symmetric densities g and calibrate them according to certain amounts Δ of change in location, so that their false alarm rates before a truncation time $n = 200$ are controlled at certain specified levels α when the observations are generated by the same g on which the scores are based. The calibration consists of determining the cut-off points c in (2) and (7) empirically according to the specifications described above. We denote the cut-off points obtained in this manner by $C_R(\alpha, \Delta; g)$ for the Rank-CUSUM procedure and by $C_P(\alpha, \Delta; g)$ for the Page-CUSUM procedure. Thus

$$\hat{P}_g[\max_{1 \leq k \leq 200} S_k(g) \geq C_R(\alpha, \Delta; g)] = \hat{P}_g[\max_{1 \leq k \leq 200} T_k(g) \geq C_P(\alpha, \Delta; g)] = \alpha,$$

where $\hat{P}_g$ denotes relative frequency in 1000 runs.

Having determined the cut-off points, the Rank-CUSUM and the Page-CUSUM procedures, truncated at $n = 200$, are now applied on data generated without change in location (null case) and with change in location at $\tau+1 = 51$ of the amount Δ used in calibration (non-null case), from the pdf's that the procedures were calibrated for, and from other pdf's. We denote by f the pdf's used in generating data. Let N_R and N_P denote the stopping times of Rank-CUSUM and Page-CUSUM procedures respectively in the situations described above.

The false alarm rates, i.e., the relative frequencies of $N_R \leq 200$ and $N_P \leq 200$ in the null case when there is no change are given in Table 1 for scores based on g = Normal, Double Exponential and Cauchy, data generated from f = Normal, Double Exponential and Cauchy, and calibration parameters $(\Delta, \alpha) = (1, .05), (0.5, .10)$, and $(0.2, .20)$. All three distributions are scaled to make the Fisher-information 1 in each case. For the non-null case when a change occurs, the relative frequencies of stopping before change (false alarm), i.e., $N_R \leq 50$ and $N_P \leq 50$ and of stopping after change but before truncation (true detection), i.e., $51 \leq N_R \leq 200$ and $51 \leq N_P \leq 200$ are given in Table 2, together with the conditional mean and standard deviation of N_R and N_P given that true detection took place. All values given in Tables 1 and 2 are based on 1000 runs.

Table 1: False Alarm Rates of Rank-CUSUM and Page-CUSUM: Null Case
Truncation: $n = 200$, Change-Point: $\tau = 200$ (No Change)

			Data Distribution (f)					
Score (g)	Calibration		Normal		Dbl. Exp.		Cauchy	
	Δ	α	Rank	Page	Rank	Page	Rank	Page
	1.0	.05	.046	.041	.055	.655	.054	1.000
Normal	0.5	.10	.100	.101	.079	.548	.083	.998
	0.2	.20	.179	.187	.148	.575	.158	.997
	1.0	.05	.038	.046	.047	.048	.042	.063
Dbl. Exp.	0.5	.10	.117	.137	.101	.085	.104	.101
	0.2	.20	.208	.239	.161	.170	.188	.194
	1.0	.05	.038	.129	.052	.109	.036	.044
Cauchy	0.5	.10	.094	.264	.087	.150	.092	.088
	0.2	.20	.166	.356	.166	.231	.195	.189

Table 2: Stopping Times of Rank-CUSUM and Page-CUSUM: Non-Null Case. Truncation: $n = 200$, Change-Point: $\tau = 50, N =$ Stopping Time, $\bar{N}$ and sd (N) are conditional upon $51 \leq N \leq 200$. Note: Relative Frequencies of $N \geq 201$ are positive in some cases.

(a) Amount of Change: $\Delta = 1.0$, Calibration: $\Delta = 1.0$, $\alpha = .05$

Score(g)	Rel. Freq., Mean and s.d. of N	Data Distribution (f)					
		Normal		Dbl. Exp.		Cauchy	
		Rank	Page	Rank	Page	Rank	Page
	$N \leq 50$	.005	.009	.002	.214	.005	.887
Normal	$51 \leq N \leq 200$	.995	.991	.997	.786	.951	.113
	$\bar{N}$	64.5	62.9	73.0	61.3	91.9	59.0
	sd(N)	8.1	6.7	15.7	7.4	32.0	7.6
	$N \leq 50$	.004	.010	.006	.010	.008	.018
Dbl. Exp.	$51 \leq N \leq 200$	.996	.990	.993	.990	.991	.982
	$\bar{N}$	75.4	66.3	78.4	65.6	81.9	65.7
	sd(N)	15.1	8.3	18.7	7.9	22.2	7.9
	$N \leq 50$	.007	.034	.005	.021	.007	.008
Cauchy	$51 \leq N \leq 200$	.865	.966	.934	.979	.979	.992
	$\bar{N}$	120.9	64.5	108.6	64.4	95.5	64.8
	sd(N)	34.2	8.1	33.0	7.7	28.3	7.4

(b) Amount of Change: $\Delta = 0.5$, Calibration: $\Delta = 0.5$, $\alpha = .10$

Score(g)	Rel. Freq., Mean and s.d. of N	Data Distribution (f)					
		Normal		Dbl. Exp.		Cauchy	
		Rank	Page	Rank	Page	Rank	Page
	$N \leq 50$	.004	.012	.004	.126	.009	.817
Normal	$51 \leq N \leq 200$	.993	.985	.974	.873	.890	.183
	$\bar{N}$	83.9	82.1	98.8	78.5	111.2	67.1
	sd(N)	19.1	18.8	29.9	21.8	36.2	15.8
	$N \leq 50$	.015	.012	.010	.010	.023	.012
Dbl. Exp.	$51 \leq N \leq 200$	.966	.986	.979	.987	.964	.986
	$\bar{N}$	94.1	86.4	94.2	85.8	94.7	87.2
	sd(N)	27.5	22.6	27.8	21.8	27.3	22.2
	$N \leq 50$	.016	.042	.014	.022	.023	.014
Cauchy	$51 \leq N \leq 200$	.759	.957	.921	.976	.963	.986
	$\bar{N}$	119.8	81.0	103.3	82.5	96.9	84.3
	sd(N)	38.5	21.1	32.5	21.1	28.1	20.0

(c) Amount of Change: $\Delta = 0.2$, Calibration: $\Delta = 0.2$, $\alpha = .20$

Score(g)	Rel. Freq., Mean and s.d. of N	Data Distribution (f)					
		Normal		Dbl. Exp.		Cauchy	
		Rank	Page	Rank	Page	Rank	Page
	$N \leq 50$	.006	.010	.010	.117	.008	.763
Normal	$51 \leq N \leq 200$	.853	.843	.704	.807	.608	.235
	$\bar{N}$	117.9	115.0	125.2	104.8	126.3	76.7
	sd(N)	36.1	36.9	37.5	37.9	39.0	24.3
	$N \leq 50$	.014	.020	.018	.016	.012	.013
Dbl. Exp.	$51 \leq N \leq 200$	.741	.804	.804	.826	.776	.808
	$\bar{N}$	118.6	117.1	120.2	118.5	119.2	117.8
	sd(N)	37.8	37.6	37.7	36.3	37.8	37.2
	$N \leq 50$	.017	.046	.015	.019	.017	.014
Cauchy	$51 \leq N \leq 200$	.528	.834	.736	.847	.776	.854
	$\bar{N}$	124.3	110.0	121.7	113.6	118.7	116.6
	sd(N)	38.7	37.7	37.5	37.2	37.1	36.9

The simulation results given in Tables 1 and 2 demonstrate the following asymptotic properties of the two procedures:

1. The false alarm rates of the Rank-CUSUM procedure are robust against misspecification of the underlying distribution when there is no change, but the false alarm rates of the Page-CUSUM procedure are not. The Page-CUSUM false alarm rates become as high as twice the calibration parameter α if a Cauchy score is used on Normal data and are practically 1 if a Normal score is used on Cauchy data. Since the Normal score is a popular choice for Page-CUSUM , this behavior should be of serious concern.

2. If the score matches with the data distribution, then the stopping times of the two procedures in the non-null case resemble one another more and more closely as the amount of change Δ to be detected gets smaller and smaller, and are almost identical for $\Delta = 0.2$ with $\alpha = .20$.

The simulation results also seem to indicate the following tendencies in the two procedures in Table 2:

3. When the score function does not match with the data distribution, the Page-CUSUM procedure acts aggressively by stopping too early. This may explain, at least partly, the higher rates of true detection and smaller average run lengths that the Page-CUSUM procedure has in some cases.

4. For the Page-CUSUM procedure, the Double Exponential score seems to be a good choice. Indeed, this score can be thought of as a robust version of the Normal score for Page-CUSUM .

REFERENCES

BHATTACHARYA, P. K. and FRIERSON, D., JR. (1981). A nonparametric control chart for detecting small disorders. *Ann. Statist.* **9**, 544–554.

HÁJEK, J. and SĬDÁK, Z. (1967). *Theory of Rank Tests*. Academic Press.

LORDEN, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Statist.* **42**, 1897–1908.

MOUSTAKIDES, G. V. (1986). Optimal stopping times for detecting changes in distributions. *Ann. Statist.* **14**, 1379–1387.

PAGE, E. S. (1954). Continuous inspection schemes. *Biometrika* **41**, 100–115.

DIVISION OF STATISTICS
UNIVERSITY OF CALIFORNIA
DAVIS, CA 95616