STATISTICS IN TRANSITION-new series, December 2010
Vol. 11, No. 3, pp. 585—596

DISTRIBUTION APPROXIMATIONS FOR CUSUM AND CUSUMSQ STATISTICS

Reza Habibi[1]

ABSTRACT

The cumulative sum (cusum) is an important statistics in testing for a change point. This paper is concerned with the distribution approximations to the cusum statistic under the null and alternative hypotheses. We also consider distribution approximations for the cumulative sum of squares (cusumsq) test statistics. Finally, a discussion section is given.

Key words: Beta approximated; Change point; Cumulative sum; Cumulative sum of squares; Multivariate normal; Response surface regression.

1. Introduction

It is very important for economic policy to identify change points in economic and financial series. For example, Hsu (1979) tested the existence of changes in stock market data. Kim *et al*. (2000) considered the problem of multiple change point in GARCH models. Hillebrand and Schnabl (2003) studied change point detection in volatility of Japanese foreign exchange intervention under GARCH modeling. Halunga *et al*. (2009) detected changes in the order of integration of US and UK inflation. In finance, the portfolio's volatility may increase as well as risk premium rises (see e.g. Mikosch and Starica, 2003).

During the last four decades, different methods are employed for detecting change points. Page (1954) studied change point analysis in the context of quality control. Chernoff and Zacks (1964), using a quasi-Baysian approach, modeled the change points. Hinkely (1970) derived the maximum likelihood estimation of change point. Worsley (1988) constructed confidence intervals for change point in the exponential family distributions. Habibi *et al*. (2005) considered the change point detection in a general class of distributions. An excellent reference in change point problems is Csorgo and Horvath (1997). The distribution theory used for these tests is typically asymptotic. So deriving the exact distribution of

[1] Department of Statistics, Central Bank of Iran, Ferdowsi Ave., Postal Code: 1135931496, Tehran, Iran, e-mail: habibi1356@gmail.com.

test statistics is very important. One of the most useful approach to detect shift in means of observations is cumulative sum (cusum) which is described as follows. In this paper, we study the finite distributions of cusum test statistic. We also consider the shift in variance case. The cusum statistic given by

$$M_n = n^{-1/2} \max_{1\le k\le m} |s_k|,$$

is an important statistic in testing for a change point, at which

$$s_k = \sum_{i=1}^{k} e_i,$$

$e_i = x_i - \bar{x}, \bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i, i, k = 1,...,m, m = n-1$. Here, $x_1,...,x_n$ is a sequence of independent normal random variables whose means are θ_i, $i = 1,...,n$, where

$$\theta_i = \begin{cases} \theta & i = 1,2,...,k_0 \\ \theta+\delta & i = k_0+1,...,n, \end{cases}$$

with a common known variance σ^2. It is interesting to test if θ_i are changed at unknown time point k_0, that is: $H_0: \delta = 0$ against $H_1: \delta \neq 0$. The large values of M_n rejects H_0. The exact null distribution of M_n is complicated. Note that the null distribution of M_n does not depend on θ. The limiting null distribution of M_n is given by $\sigma Sup\, B(t)$ where $B(t)$ is the standard Brownian bridge on $(0,1)$ and the supremum is taken over $(0,1)$. Conniffe and Spencer (2000) proposed a central chi-squared approximation for the null distribution of M_+^2, where

$$M_+ = n^{-1/2} \max_{1\le k\le m} s_k.$$

Their approximation method works well. However, in this note, the exact null and alternative distribution of M_n is studied. We also consider the exact distribution of change point estimator. The change point in variance is considered. These problems are not considered by Conniffe and Spencer (2000). The problem is to find the quantile c_α such that $P_{H_0}(M_n > c_\alpha) = \alpha$ for finite sample sizes *n*. We note that

$$P_{H_0}(M_n > c_\alpha) = 1 - P_{H_0}(|s_k| \le \sqrt{n}c_\alpha, for\ all\ k = 1,...,m) = \alpha,$$

and

$$P_H(M_n < x) = 1 - P_H(|s_k| \le \sqrt{n}c_\alpha, \textit{for all } k = 1,...,m), H = H_0 \textit{ and } H_1.$$

As follows, we show that $s = (s_1,...,s_m)^T$ has multivariate normal distribution under the null and alternative hypotheses. That is, M_n is the maximum of absolute of a multivariate normal distribution and that $\sqrt{n}c_\alpha$ is a two sided equi-coordinate $1-\alpha$ percent quantile of multivariate normal distribution. The percentage points of M_+ is denoted by c_α^+. Then $\sqrt{n}c_\alpha^+$ is the one sided equi-coordinate quantile of multivariate normal. Genz (1992) proposed some numerical approaches to calculate the cumulative probabilities and equi-coordinate quantiles of a multivariate normal distribution with any mean vector and covariance matrix. The function *pmvnorm* in *mvtnorm* package of *R* software performs this calculations.

To prove the normality, let $x = (x_1,...,x_n)^T$ be observation vector. Under the null hypothesis of no change point, x has a $n-$variates normal with mean vector θj_n and covariance matrix $\sigma^2 I_n$ where I_n is the $n \times n$ identity matrix and j_n is $n \times 1$ vector of 1's. The vector of deviation from the mean $e = (e_1,...,e_m)^T$ equals to Ax where A is the following partitioned matrix

$$A = \left[I_m - \frac{1}{n} J_m \,\vdots\, -\frac{1}{n} j_m \right],$$

where J_m is $m \times m$ matrix of 1's. Then has a $m-$variate normal distribution with mean vector 0 and the covariance matrix

$$\sigma^2 AA^T = \sigma^2 (I_m - \frac{1}{n} J_m).$$

Let $L = (L_1,...,L_m)^T$ is the m × m matrix of vectors L_i such that $L_i = (1,...,1,0,...,0)^T$ at which the number of 1's at L_i is $i = 1,...,m$. One can see that $s = Le$, and then s is m−variate normal with mean vector 0 and the covariance matrix $\sigma^2 D$, where

$$D_{ij} = \min(i,j) - \frac{ij}{n}, i, j = 1,...,0.$$

Under the alternative hypothesis, s is again $m-$variate normal with the covariance matrix $\sigma^2 D$ but the mean vector is $\alpha = (\alpha_1,...,\alpha_m)^T$, such that

$$\alpha_k = \frac{-\delta}{n}\min(k,k_0)\{n-\max(k,k_0)\},\, k=1,\ldots,m.$$

In the next section, we compute one-sided and two-sided cv's $\sqrt{n}c_\alpha^+$ and $\sqrt{n}c_\alpha$ We compare our critical points (cv's) with the cv's obtained by Monte Carlo (MC) simulation. We study the power of test. We also consider distribution approximations for the cumulative sum of squares (cusumsq) test statistic for change point detection in variance in section 3.

2. Cv's and power of test

Tables 1 and 2 give one-sided and two-sided cv's. Without loss of generality, we assume that $\sigma^2 = 1$ and $\theta = 1$. It is seen our approximated cv's are close to true cv's estimated by Monte Carlo study. This fact shows our approximation is accurate. The absolute errors of our approximation is measured by

$$e_\alpha = \left| P_{H_0}(\max_{1\le k\le m}|s_k| \le q_\alpha) - \alpha \right|,$$

where q_α is the true quantile of $\max_{1\le k\le m}|s_k|$ obtained by Monte Carlo experiment and $P_{H_0}(\max_{1\le k\le m}|s_k| \le x)$ is computed by our normal approximation. Table 3 gives the maximum and median of 100 errors e_α, $\alpha = 0.9(0.001)0.999$ for each sample sizes. Table 3, we conclude that our approximation works well. Table 4 gives the normal approximated power of test for some selected sample sizes *n*. We let $k_0 = \frac{n}{2}, \delta = 1, 2, \sigma^2 = 1$ and $\theta = 1$.

Table 1. Comparison of normal and MC two-sided cv's

α	0.1	0.05	0.025	0.01
Normal cv, n = 5	2.166470	2.466221	2.741670	3.074471
MC cv, n = 5	2.161882	2.467639	2.739131	3.074245
Normal cv, n = 6	2.422699	2.753462	3.057174	3.418402
MC cv, n = 6	2.421451	2.753106	3.057837	3.415915
Normal cv, n = 10	3.291676	3.717723	4.107556	4.569724
MC cv, n = 10	3.293115	3.710154	4.110897	4.57586
Normal cv, n = 15	4.159806	4.680369	5.159716	5.729828
MC cv, n = 15	4.160063	4.686651	5.148775	5.726891
Normal cv, n = 20	4.891475	5.494822	6.042609	6.70567
MC cv, n = 20	4.915969	5.505689	6.045103	6.689938

Table 2. Comparison of normal and MC one-sided cv's

α	0.1	0.05	0.025	0.01
Normal cv, n = 5	1.837014	2.173590	2.456540	2.826679
MC cv, n = 5	1.837081	2.172501	2.456957	2.825337
Normal cv, n = 6	2.062609	2.423182	2.753357	3.144838
MC cv, n = 6	2.060721	2.423146	2.757961	3.150792
Normal cv, n = 10	2.812123	3.291219	3.718938	4.225157
MC cv, n = 10	2.813812	3.292955	3.729271	4.237521
Normal cv, n = 15	3.563842	4.154548	4.670302	5.298982
MC cv, n = 15	3.565857	4.154409	4.659146	5.298102
Normal cv, n = 20	4.228066	4.892632	5.493382	6.229083
MC cv, n = 20	4.233962	4.903461	5.496865	6.233975

Table 3. Errors of normal approximated probabilities

n	Maximum error	Median error
5	0.0017	0.0005
6	0.0009	0.0007
10	0.0012	0.0008
15	0.0024	0.0004
20	0.0011	0.0002

Table 4. Normal approximated power, α =0.05

n/θ	1	2
5	0.16263	0.51562
6	0.19832	0.62041
10	0.29361	0.83136
15	0.40851	0.94663
20	0.52157	0.98595

Remark 1

Ploberger and Kramer (1992) used the cusum statistic replacing $e_i = x_i - \overline{x}$, by regression residuals. Our results can be extended to regression residuals. Consider the regression model

$$Y = X\beta + \varepsilon \ ,$$

at which $Y_{n\times 1}$, $X_{n\times p}$, $\beta_{p\times 1}$ and $\varepsilon_{n\times 1}$ are observation vector, design matrix, unknown coefficients vector and residual vector, respectively. Suppose that X is of full rank and define $Q = I_n - P$ where $P = X(X^T X)^{-1} X^T$

Then the vector of estimated residuals $\in = (\in_1, ..., \in_m)^T$ is given by $\in = Q\varepsilon$ and $\sum_{i=1}^{n} \varepsilon_i = 0$.

Let $s_j^{\in} = \sum_{i=1}^{n} \varepsilon_i$ for $i = 1,...,m$ .Then $s_j^{\in} = L_i^T \in$. Then

$$\mathrm{cov}(s_i^{\in}, s_j^{\in}) = \sigma_{ij}^{\in} = L_i^T Q L_j = L_i^T (I_n - P) L_j = \min(i, j) - L_i^T Q L_j .$$

Following section 2, it can be shown that the null distribution of $s_j^{\in}$, $j = 1,...,m$ are multivariate normal with zero mean and covariance matrix $\sum \in = (\sigma_{ij}^{\in})$. The similar results can be extended to alternative distributions. The above results relates to known variance case. In most practical situations, variance will have to be estimated. This case is considered in section 4.

3. Change in Variance

In the previous section, we studied the change point in the means of observation. To detect change point in variance using the cusum statistic, let $y_i = x_i^2$. Here, under H_1, $x_1, ..., x_n$ is a sequence of independent normal observations such that $E(x_i) = 0$ and

$$\mathrm{var}(x_i) = \begin{cases} \sigma^2 & i = 1, 2, ..., k_0 \\ \xi^2 & i = k_0 + 1, ..., n, \end{cases}$$

where $\sigma^2 \neq \xi^2$. The means of y_i are changed under H_1. The cusum test statistic is

$$M_3 = \frac{n^{-1/2} \max_{1 \le k \le m} \left| \sum_{i=1}^{k} (y_i - \bar{y}) \right|}{\sigma^2}.$$

The limiting null distribution of M_3 is given by $\sqrt{2} \sup |B(t)|$. The Table 5 gives 5% quantile of M_3. A response surface regression for 5% quantile of M3 is estimated as follows. The adjusted R^2 of regression is 99.1.

$$q_n = 2.76 - 2.59 n^{-1}.$$

Table 5. 5% quantile of M_3

n	q_n	n	q_n	n	q_n	n	q_n
5	2.228	10	2.514	35	2.693	55	2.706
6	2.364	15	2.596	40	2.697	60	2.715
7	2.399	20	2.606	45	2.705	65	2.723
9	2.446	25	2.671	50	2.705	70	2.737

Following Conniffe and Spencer (2000), we propose a chi-squared approximation in the form of $a_n\chi^2_{df_n}$ $(a_n > 0\, and\, df_n > 0)$ for M^2_{3+}. The moment estimates of a_n and df_n are

$$a_n = \frac{\lambda_n^2}{2\mu_n} and\, df_n = \frac{2\mu_n^2}{\lambda_n^2}$$

where μ_n and λ_n^2 are the mean and variance of M^2_{3+} under H_0. Table 6 gives μ_n, λ_n^2, a_n and df_n for $\sigma^2 = 1$. Following Conniffe and Spencer (2000), we let n = 10, 20, 30, 40, 60, 80, 100. Table 7 gives the median and maximum of absolute errors.

Table 6. Values of μ_n, λ_n^2, a_n and df_n

n	10	20	30	40	60	80	100
μ_n	6.01634	13.7955	21.8967	30.5796	48.3933	65.6576	84.3329
λ_n^2	137.622	458.696	960.258	1711.04	3663.53	6216.98	9710.17
a_n	11.4374	16.6248	21.9269	27.9768	37.8516	47.3439	57.5704
df_n	0.52602	0.82981	0.99862	1.09303	1.27851	1.38682	1.46486

Table 7. Comparisons chi-squared with MonteCarlo probabilities

n	Maximum of error	Median of errors
10	0.0015	0.0005
20	0.0015	0.0003
30	0.0014	0.0004
40	0.0012	0.0006
60	0.0016	0.0004
80	0.0014	0.0005
100	0.0015	0.0005

Remark 2

When σ^2 is unknown it is replaced by its estimate under the null hypothesis, i.e., $\bar{y}$ and the test is proposed by

$$T_n = n^{\frac{1}{2}} \max_{1 \le k \le m} |D_k|, D_k = \frac{\sum_{i=1}^{k} X_i^2}{\sum_{i=1}^{n} X_i^2} - \frac{k}{n}, k = 1, \dots, n.$$

This statistic is the cumulative sum of square (cumsumsq) proposed by Inclan and Tiao (1994). The limiting null distribution of T_n is given by $\sqrt{2} \sup |B(t)|$. For moderate sample sizes, Sanso *et al*. (2004) estimated a response surface regression for 5% quantile of T_n as follows

$$q_{0.05} = 1.359167 - 0.737020 n^{-1/2} - 0.69155 n^{-1}.$$

Remark 3

When $E(x_i) = \mu$ (known), we let $y_i = (x_i - \mu)^2$. When μ is unknown, we can use M_3 again, since means of x_i^2 are changed. We can also use M_3 with letting $y_i = (x_i - \bar{x})^2$. Call this statistic by $\hat{M}_3$. The critical values (cv) are given in Table 8 for $\alpha = 0.05$. Parameter μ is chosen by computer. As we expect, the critical values of $\hat{M}_3$ does not depend on μ. In Table 9, we compare the power of tests for M_3 and $\hat{M}_3$ cases, $\sigma^2 = 1$, $\xi^2 = 3$ and $k_0 = \frac{n}{2}$. It seems test procedure, based on $\hat{M}_3$ works much better. We are working on null and alternative distributions of $\hat{M}_3$. For other possibilities, see discussion section.

Table 8. Critical values

n	10	20	30	40	60	80	100
μ	0.791	1.375	0.163	1.101	0.785	1.111	0.723
cv of M_3	2.770	3.935	1.918	3.364	2.844	3.457	2.664
cv of $\hat{M}_3$	1.739	1.797	1.808	1.830	1.874	1.877	1.866

Table 9. Power of tests M_3 and $\hat{M}_3$

n	Power of M_3	Power of $\hat{M}_3$
10	0.344	0.417
20	0.404	0.646
30	0.795	0.788
40	0.622	0.866
60	0.810	0.950
80	0.807	0.980
100	0.956	0.997

4. Discussion

Following referee comments, this section is added to paper to present a list of future research topics. We are working on them and they will be completed in future.

4.1. Change in mean: unknown variance

Conniffe and Spencer also developed tests for the unknown variance case. In most practical situations, variance will have to be estimated and unless sample size is very large the estimate cannot safely be treated as a known value. This is why, we considered this part. I guess, I can approximate the distribution of test statistic by multivariate t distribution.

4.2. Change in mean and variance

Sometimes when variance changes so does the mean. For example, finance theory says that if a portfolio's volatility increases the risk premium will rise and change the return. The cusum test statistics and their distributions are interesting.

4.3. Test procedures based on regression residuals

Similar to Remark 1, it is interesting to detect change point in variance using regression residuals. We are working on this topic.

4.4. Distribution of $\hat{M}_3$

The null and alternative distribution of $\hat{M}_3$ is interesting.

Acknowledgment

I thank the referee for his valuable comments on my manuscript. These comments helped me to improve the paper.

REFERENCES

CHERNOFF, H. and ZACKS, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Statist*. 35, 999–1018.

CONNIFFE, D. and SPENCER, J. E. (2000). Approximating the distribution of the maximum partial sum of normal deviates. *Journal of Statistical Planning and Inference*, 88, 19–27.

CSORGO, M., and HORVATH, L., (1997). *Limit Theorems in Change Point Analysis*, Wiley. UK.

GENZ, A. (1992). Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, 1, 141–150

HABIBI, R., SADOOGHI-ALVANDI, S. M. and NEMATOLLAHI, A. R. (2005). Change point detection in a general class of distributions. *Communications in Statistics, Theory and Methods* 34, 1935–1938.

HALUNGA, A. G., OSBORN, D. R. and M. SENSIER (2009). Changes in the order of integration of US and UK infation. *Economics Letters* 102, 30–32.

HILLEBRAND, E. and SCHNABL, G. (2003). The effects of Japanese foreign exchange intervention: GARCH estimation and change point detection. *Discussion Paper* No.6. Japan Bank for International Cooperation (JBIC).

HINKLEY, D., (1970). Inference about the change-point in a sequence of random variables. *Biometrika* 57, 1–17.

HSU, D.A., (1979). Detecting shifts of parameter in gamma sequences, with applications to stock price and air traffic flow analysis. *J. Amer. Statist. Assoc*. 74, 31–40.

INCLAN, C. and TIAO, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variances. *J. Amer. Statist. Assoc*. 89, 913–923.

KIM S., S. CHO and S. LEE (2000), On the cusum test for parameter changes in GARCH(1,1) models, *Communications in Statistics, Theory and Methods* 29, 445–462.

MIKOSCH, T. and STARICA, C. (2003). Change of structure in financial time series, long range dependence and the GARCH modeld. *Review of Economics and Statistics*, forthcoming.

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

PLOBERGER, W. and KRAMER, W. (1992). The cusum test with ols residuals. *Econometrica* 60, 271–285.

SANSO, A., ARAGO, V. and CARRION, J. L. (2004). Testing for changes in the unconditional variance of financial time series, *University of Barcelona Working Paper*.

WORSLEY, K. J. (1986). Confidence regions and test for a change point in a sequence of exponential family random variables. *Biometrika* 73, 91–104.