Working Paper

A NOTE ON THE CRITICAL LIMITS OF THE CUSUM TEST

Peter Hackl

June 1986
WP-86-26

International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

A NOTE ON THE CRITICAL LIMITS OF THE CUSUM TEST

Peter Hackl

June 1986
WP-86-26

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

This paper is one of a series embodying the outcome of 2 years of IIASA's collaborative research on techniques for the identification of structural changes. This type of research is required by economists dealing with econometric modeling in view of the well-known changes in the world economy in the previous decades. It is furthermore a topic of the IIASA/Bonn Research Project on World Economic Modeling led by Prof. W. Krelle.

In a few months IIASA will hold final meetings on this topic, and P. Hackl's paper contributes to a better understanding of the problem, which is outlined as follows:

In a number of econometric problems data are generated by different processes for neighboring time intervals. The paper concerns the simplest case, where on every interval a response function is a linear function, but the point of switching is unknown. Therefore it is necessary to estimate regression parameters and to identify the location of this point. Two test statistics are compared: the cumulative sum statistics proposed by Brown, Durbin and Evans (1975) and its modification developed by the author. Using the Monte-Carlo approach, the author has shown that the second statistics is more powerful when the number of observations is modest (say, 20 per one unknown parameter) and for the larger number of observations the first statistics gains its asymptotical optimality.

Anatoli Smyshlyaev

Peter Hackl

A NOTE ON THE CRITICAL LIMITS OF THE CUSUM TEST

1. Consider the regression model $y_t = x_t'\beta_t+u_t$, $t=-k+1,\dots,T$, where y_t is the observation of the dependent variable at time t, x_t is a k-vector of observations of the independent variables (which are assumed non-stochastic), β_t is a k-vector of unknown regression coefficients, and u_t are independently and normally distributed disturbances with mean zero and variance σ^2. Brown, Durbin and Evans (1975), in the following BDE, suggest to test the null-hypothesis H_0: $\beta_t=\beta$, $t=-k+1,\dots,T$, against a non-specific alternative by means of the cusum test: Reject H_0 if

$$S_T = \sup_t(|C_t|((1+2t/T)\sqrt{T})^{-1}) \geq a_\alpha \ .$$

Here, $C_t = (\Sigma_{j\leq t}W_j)/s$, $t=1,\dots,T$, are the cumulative sum statistics, and s is an estimate of the standard deviation σ; the recursive residuals are defined to be $W_t = r_t(y_t-x_t'b_{t-1})$, $t=1,\dots,T$; here, $r_t = (1+x_t(X_{t-1}'X_{t-1})^{-1}x_t')^{-1/2}$, X_{t-1} is a $(k+t-1)\times k$-matrix (of full rank for $t>1$) with rows x_j', $j=-k+1,\dots,t-1$, and b_{t-1} is the OLS-estimator of β, based on observations prior to t. $\{S_T\geq a_\alpha\}$ is equivalent to the event that the path of the cumulative sums C_t leaves the space between the straight lines between the points $(1,\pm a_\alpha\sqrt{T})$ and $(T,\pm 3a_\alpha\sqrt{T})$, i.e., crosses one of the straight lines $\pm a_\alpha\sqrt{T}(1+2t/T)$. These critical lines are determined by a_α which has to fulfill $P(S_T \geq a_\alpha) \leq \alpha$ under H_0. Brown, Durbin and Evans derive a_α from $P(S \geq a_\alpha) \leq \alpha$ where S is analogous to S_T with C_t substituted by a standard Wiener process. A result from the theory of the Wiener processes (Durbin, 1971) allows to calculate the probability that a sample path crosses the straight line $a_\alpha \sqrt{T}(1+2t/T)$; in the opposite way, values for a_α can be obtained for a given α if the probability of crossing both the upper and the lower critical line is neglected.

2. The way how the critical limits of the cusum test are derived is justified by the fact that S_T converges in distribution to S. In this sense, the cusum test is an asymptotic test. The effect of this approximation to the situation of finite sample sizes can hardly be assessed. Simulation studies generally indicate a reduction of the error (I) probability as compared to the nominal significance level, this reduction being decreased with increasing sample size T (Garbade, 1977; Hackl, 1980). In addition, it must be expected that the power of the test depends on the onset of the violation of the null-hypothesis.

3. An exact test based on a finite number of cumulative sum statistics can be performed by simultaneously testing each of the C_t's for significant deviations from its expectation under H_0. The C_t's are approximately normally distributed with $E(C_t)=0$ and $Cov(C_t, C_{t'}) = \min(t,t')$. As BDE mention, if it is wished to have critical limits such that under H_0 the probability that the sample path crosses the curves at any point between t=1 and t=T is constant, the curves must have the form $\pm c_\alpha \sqrt{t}$ where c_α is a

suitably chosen real number. The derivation of c_α must be based on the joint distribution of the C_t's. The application of Bonferroni's inequality implies that the dependence of the C_t's is neglected; this would - due to large correlations of the C_t's - result in a strongly conservative test. By use of Hunter's (1976) inequality, much less conservative critical limits can be derived on the basis of bivariate distributions of the respective statistics.

4. For any set of events $A_1,\ldots,A_T$, Hunter's inequality states

$$P(\cup_i A_i) \leq \Sigma_i P(A_i) - \Sigma_{(i,j)\in M} P(A_i A_j) \ .$$

The set M of pairs of indices (i,j) has a particular form: Interpreting the events A_i as knots in a simple graph with the pairs of indices (i,j) representing the edges, then M is that connected subgraph without cycles, or tree, consisting of not more than k-1 edges, which has maximal length, the lengths of the edges being measured by the joint probabilities $P(A_i A_j)$. The set M can be found by initially searching for the largest of the joint probabilities and then subsequently adding from the remaining pairs that one which has maximal joint probability and does not create a cycle with pairs already included in the tree. If the tree of maximal length is not unique, any of these trees can be used. A detailed discussion of the application of Hunter's inequality in simultaneous testing situations is given by Bauer & Hackl (1985).

5. In the following, modified cusum statistics $C_t^* = C_t/\sqrt{t}$, $t=1,\ldots,T$, will be basis of the test. They are standardized statistics with $\mathrm{Cov}(C_t^*, C_{t'}^*) = \sqrt{(\min(t,t')/\max(t,t'))}$. Hunter's inequality can be used to derive from $P(\max_{1\le t\le T}|C_t^*|\ge c_\alpha^*|H_0) = \alpha$ the critical limits c_α^*: If the right-hand-side of $P(\cup_t|C_t^*|\ge c_\alpha^*) \le \Sigma_t P(|C_t^*|\ge c_\alpha^*) - \Sigma_{(t,t')\in M} P(|C_t^*|\ge c_\alpha^*, |C_{t'}^*|\ge c_\alpha^*)$ is set equal to α, conservative Hunter-type critical limits $c_{\alpha H}^*$ are obtained for c_α^*. As $P(|C_t^*|\ge c_\alpha^*, |C_{t'}^*|\ge c_\alpha^*)$ is a monotonic function of $|\rho_{tt'}|$, $\rho_{tt'}$ being the correlation coefficient between C_t^* and $C_{t'}^*$, the set M contains the index-pairs (T-1,T),...,(T-[T/2],T-[T/2]+1) and (T-2i-2,T-2i),(T-[T/2]-i-1,T-[T/2]-i),(T-2i-3,T-2i-1), i=0,1,..., so that the number of pairs is T-1; here, [.] is the ceiling function. The gain in power as compared to the use of Bonferroni's inequality must be expected to be considerable as the cusum statistics are highly correlated: For T=20, 40, and 80, the mean of the correlation coefficients in M is 0.952, 0.976, and 0.988, respectively. In Tab.1 Hunter-type critical values $c_{\alpha H}^*$ and the corresponding probabilities $P(|C_t^*|\ge c_{\alpha H}^*) = p_{TH}$ for the overall significance levels $\alpha=0.05$ and 0.01 and for values of T in the range between 5 and 100 are given. The integration of the bivariate normal distribution has been performed by means of the Routine D01DAF of the NAG Programme Library. The accuracy level has been chosen so that the results given in Tab.1 are correct to the digits given.

Tab.1: Hunter-type Critical Value $c^*_{\alpha H}$ and Probability p_{TH} (in Percent), for Various Values of α and T.

α / T	0.05 $c^*_{\alpha H}$	p_{TH}	0.01 $c^*_{\alpha H}$	p_{TH}
5	2.430	1.510	2.994	0.275
10	2.586	0.972	3.131	0.174
15	2.666	0.767	3.202	0.136
20	2.721	0.651	3.250	0.115
25	2.762	0.574	3.286	0.101
30	2.795	0.519	3.315	0.092
35	2.823	0.476	3.339	0.084
40	2.846	0.443	3.359	0.078
45	2.868	0.413	3.377	0.073
50	2.888	0.387	3.393	0.069
60	2.926	0.344	3.420	0.063
80	2.988	0.280	3.475	0.051
100	3.044	0.233	3.521	0.043

6. Tab.2 shows the results of a Monte Carlo simulation experiment performed in order to compare the BDE cusum test and its modification. The power of the tests against a sudden shift of the expectation between $t=t^*$ and $t=t^*+1$ was investigated: The tests were applied to data obtained from $Y_t = \mu+u_t$ if $t\leq t^*$ and $Y_t = \mu+\delta\sigma+u_t$ if $t>t^*$; the disturbance terms u_t are pseudo random numbers with means zero and variances σ^2 for all t, generated by means of the function RUNNOR of the Statistical Analysis System (SAS). The parameters used were $\mu=0$, $\sigma^2=1$, T=20, 40, and 80, and $\delta=0$, 1, and 2.

In the case of non-constancy ($\delta\neq0$) the estimate $s^2 = \Sigma_t W_t{}^2/(T-1)$ overestimates σ^2, so that the power of the test is reduced by the fact that the estimated standard deviation is denominator in each of the test statistics (Hackl, 1980). To reduce this effect the variance σ^2 was estimated as $s_o{}^2 = \Sigma_t(W_t-\bar{w})^2/(T-1)$ where $\bar{w}$ is the sample mean of the recursive residuals.

Each rejection probability reported was estimated from a sample of size 2000 except for the null-hypothesis case (δ=0) where the sample size was 10000. The standard deviations of the estimates for probabilities 0.05, 0.2, and 0.5 are 0.0049, 0.0089, and 0.011, respectively, if the sample size is 2000 and less than half of these values for sample size 10000.

Tab.2: Estimated Power (in Percent) in Rejecting the Null-Hypothesis of Constancy When the True Model Contains a Shift δ of the Intercept, for Various Values of α, δ, T, and t*/T.

T	t*/T	δ	0		1		2	
		α	0.05	0.01	0.05	0.01	0.05	0.01
20	0.25	BDE-cusum			13.8	5.5	44.2	22.9
		mod.cusum			16.2	6.5	48.3	26.8
	0.50	BDE-cusum	3.26	0.78	11.7	5.0	35.0	17.7
		mod.cusum	4.21	0.97	13.9	6.3	39.0	21.3
	0.75	BDE-cusum			3.9	1.2	5.2	1.8
		mod.cusum			4.6	1.4	6.4	2.4
40	0.25	BDE-cusum			40.3	19.5	94.1	81.9
		mod.cusum			37.2	18.6	93.8	81.9
	0.50	BDE-cusum	3.93	0.75	27.7	11.7	82.4	59.1
		mod.cusum	4.10	0.88	27.2	12.6	82.4	62.5
	0.75	BDE-cusum			7.5	2.0	24.8	10.0
		mod.cusum			7.4	2.2	24.7	11.7
80	0.25	BDE-cusum			75.2	51.9	100.0	99.8
		mod.cusum			66.7	46.8	100.0	99.6
	0.50	BDE-cusum	3.85	0.68	56.5	32.6	99.6	96.2
		mod.cusum	3.47	0.76	49.7	30.3	98.9	95.6
	0.75	BDE-cusum			14.7	4.9	59.1	31.0
		mod.cusum			11.3	4.5	52.0	28.8

7. The conclusions drawn from the Table can be summarized as follows: Both tests are conservative, the difference between α and the estimated actual significance level decreasing and increasing for the BDE cusum test and the modified version, respectively. For both tests the estimated power increases with in-

creasing shift size δ and number of cusum statistics T, but decreases with increasing ratio t*/T. For T=20, the modified cusum test is more powerful than the BDE test, this superiority being increased with decreasing α. For T=40, the tests are of about equal power and for large values of T, the estimated power of the BDE cusum test exceeds that of the modified test in nearly all cases considered. For T=80, however, the estimated actual significance level of the modified cusum test lies considerably below that of the BDE cusum test.

From these results follows that the modified cusum test should be preferred to the BDE test in cases of small number T.

References

Bauer,P. and Hackl,P. (1985). The application of Hunter's inequality in simultaneous testing. Biometr.J., 27, 25-38.

Brown,R.L., Durbin,J., and Evans,J.M. (1975). Techniques for testing the constancy of regression relationships over time. J.Roy.Statist.Soc., Ser.B, 37, 149-192.

Durbin,J. (1971). Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov-Smirnov test. J.Appl.Prob., 8, 431-453.

Garbade,K. (1977). Two methods for examining the stability of regression coefficients. J.Amer.Statist.Assoc., 72, 54-63.

Hackl,P. (1980). Testing the Constancy of Regression Relationships Over Time. Göttingen: Vandenhoeck & Ruprecht.

Hunter,D. (1976). An upper bound for the probability of a union. J.Appl.Prob., 13, 597-603.