Kuwait J. Sci. **41** (2) pp. 107-125, 2014

Multiple change points tests against the umbrella alternative

FAHIMAH AL-AWADHI AND EMAD-ELDIN A. A. ALY

Department of Statistics and Operations Research, Faculty of Science, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait.
E-mail: mapfahaa@hotmail.com and eealy50@gmail.com

ABSTRACT

We consider the problem of testing the null hypothesis of no change against the alternative of multiple umbrella-type change points in a series of independent observations. We extend the tests of Mack & Wolfe (1981) to the change point set-up. We consider the two cases of known and unknown umbrella peak point. We obtain the asymptotic null distributions of the proposed tests and give approximations for their limiting critical values. We also give tables for their finite sample Monte Carlo critical values. We report the results of several Monte Carlo power studies conducted to compare the proposed tests with a number of multiple change points tests. As an illustration we applied the proposed tests using a real data set.

Keywords: Brownian bridge; limit theorems; Monte Carlo simulations.

PRELIMINARIES

Let $X_1, X_2, \cdots, X_n$ be independent random variables with continuous distribution functions (DF's), $F_1, F_2, \cdots, F_n$, respectively. Let $[y]$ be the integer part of y and $\overset{P}{\prec}$ be a specified partial ordering of the family of DF's under consideration, like, for example, the hazard rate, the stochastic and the dispersive orderings. We consider in this article the problem of testing the null hypothesis of no change

$$H_\circ : F_1 = F_2 = \cdots = F_n = F, \quad (\text{F is unknown}) \tag{1}$$

against the umbrella k-multiple change points alternative with a specified umbrella peak point $l = 2, 3, \cdots, k$,

$$\begin{aligned} H_{1l} \quad &: \quad \exists\ 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < 1 \text{ such that } F_1 = \cdots = \\ F_{[n\lambda_1]} \overset{P}{\prec} F_{[n\lambda_1]+1} \quad &= \quad \cdots = F_{[n\lambda_2]} \overset{P}{\prec} \cdots \overset{P}{\prec} F_{[n\lambda_{l-1}]+1} = \cdots = F_{[n\lambda_l]} \overset{P}{\succ} \\ F_{[n\lambda_l]+1} \quad &= \quad \cdots = F_{[n\lambda_{l+1}]} \overset{P}{\succ} \cdots \overset{P}{\succ} F_{[n\lambda_k]+1} = \cdots = F_n, \end{aligned} \tag{2}$$

or against the umbrella k-multiple change points alternative with unknown umbrella peak point,

$$H_1 : \exists\, l \text{ with } 2 \leq l \leq k \text{ such that } H_{1l} \text{ holds.} \tag{3}$$

Lombard (1987) and Aly & BuHamra (1996) considered the problem of testing $H_\circ$ of (1) against the unrestricted multiple change points alternative which corresponds to (2) when all the orderings $\overset{P}{\prec}$ are replaced by $\overset{P}{\neq}$. Aly *et al.* (2003) considered the problem of testing $H_\circ$ of (1) against the ordered multiple change points alternative which corresponds to (2) when all the orderings $\overset{P}{\prec}$ are in the same "upward" direction. The test of Mack & Wolfe (1981) is designed to test for the umbrella alternative in the k-sample setup. The tests proposed in this article, which are motivated by this test, extend it to the change point set-up. Applications and tests of this kind of hypotheses can be found in Barlow & Brunk (1972); Robertson *et al.* (1988) and Xiong *et al.* (2003).

For additional results and references on change point analysis we refer to Zacks (1983); Bhattacharyya (1984); Csörgő & Horváth (1988 a,b); Sen (1988); Lombard (1989); Hušková & Sen (1989); Csörgő & Horváth (1997); Hušková & Slabý (2001); Aly (2004); Lebarbier (2005); Lavielle & Teyssiére (2006); Lin & Kao (2008) and Döring (2010).

Let $\phi(y_1, y_2, \cdots, y_m; z_1, z_2, \cdots, z_m)$ be a kernel of degree (m, m). Assume that ϕ is symmetric in its first (resp. second) m arguments and is skew-symmetric in the sense that

$$\phi(y_1, y_2, \cdots, y_m; z_1, z_2, \cdots, z_m) = -\phi(z_1, z_2, \cdots, z_m; y_1, y_2, \cdots, y_m).$$

The generalized two-sample U-statistic based on the kernel ϕ and the samples $Y_1, \cdots, Y_{n_1}$ and $Z_1, \cdots, Z_{n_2}$ is defined as

$$U_{n_1,n_2}(Y_1, \cdots, Y_{n_1}; Z_1, \cdots, Z_{n_2}) = \frac{\sum \phi(Y_{i_1}, \cdots, Y_{i_m}; Z_{j_1}, \cdots, Z_{j_m})}{\binom{n1}{m}\binom{n2}{m}}, \tag{4}$$

where $1 \leq m \leq \min(n_1, n_2)$ and the above summation extends over $1 \leq i_1 < i_2 < \ldots < i_m \leq n_1$ and $1 \leq j_1 < j_2 < \cdots < j_m \leq n_2$.

In Section 2, we present the proposed test statistics and investigate their limiting theory. In Section 3, we give some special cases of the asymptotic results of Section 2. The results of Monte Carlo critical values studies are given in Section 4. In Section 5, we give the results of Monte Carlo power studies conducted to compare the proposed tests with some competitors. An illustrative example is discussed in Section 6. The proofs of the results of Section 2 are given in Section 7. Through out the rest of this paper, we will adopt the convention $\sum_a^b = 0$ for $a > b$.

THE PROPOSED TESTS

Let $s_0 = 0, s_{k+1} = 1$ and $\underline{s} = (0 < s_1 < \cdots < s_k < 1)$ be such that $[ns_i] - [ns_{i-1}] \geq 2$. Define $d_{i,n} = [ns_i] - [ns_{i-1}], i = 1, 2, \cdots, k+1$. Let

$$U_{d_{i,n},d_{j+1,n}} = U_{d_{i,n},d_{j+1,n}}(X_{[ns_{i-1}]+1}, \cdots, X_{[ns_i]}; X_{[ns_j]+1}, \cdots, X_{[ns_{j+1}]})$$

be the U-statistic of (4) based on the two sub-samples $X_{[ns_{i-1}]+1}, \cdots, X_{[ns_i]}$ and $X_{[ns_j]+1}, \cdots, X_{[ns_{j+1}]}$. Assume that large values of $U_{d_{i,n},d_{j+1,n}}$ are significant. Define

$$U_{i,j,n} = \frac{d_{i,n} d_{j+1,n}}{mn^2} U_{d_{i,n},d_{j+1,n}},$$

$$W_n(\underline{s}, l) = \sum_{i=1}^{l-1} \sum_{j=i}^{l-1} U_{i,j,n} - \sum_{i=l}^{k} \sum_{j=i}^{k} U_{i,j,n} \quad , l = 2, \cdots, k \tag{5}$$

and

$$\underline{W}_n(\underline{s}) = (W_n(\underline{s}, 2), \cdots, W_n(\underline{s}, k)). \tag{6}$$

Assume that $H_\circ$ of (1) holds. Since the kernel ϕ is skew-symmetric, it follows that $E(U_{i,j,n}) = 0$. Define

$$\phi_1(y) = E\phi(y, X_2, \cdots, X_m; X_{m+1}, \cdots, X_{2m}) \tag{7}$$

and assume

$$\sigma^2 = E\phi_1^2(X_1) < \infty. \tag{8}$$

Let $\{W_n(\underline{s}, l)\}$ be as in (5) and σ be as in (8). When the umbrella peak point l is known we propose the test statistics

$$T_{1,n}(k, l) := \frac{\sqrt{n}}{\sigma} \max_{\underline{s}} W_n(\underline{s}, l), \tag{9}$$

and

$$T_{2,n}(k,l) := \frac{\sqrt{n}}{\sigma} \underbrace{\int \cdots \int}_{\underline{s}} W_n(\underline{s}, l) d\underline{s}. \tag{10}$$

When the umbrella peak point l is unknown we propose the test statistics

$$T_{3,n}(k) := \frac{\sqrt{n}}{\sigma} \max_{2 \le l \le k} \max_{\underline{s}} W_n(\underline{s}, l) = \max_{2 \le l \le k} T_{1,n}(k,l), \tag{11}$$

$$T_{4,n}(k) := \max_{2 \le l \le k} T_{2,n}(k,l), \tag{12}$$

$$T_{5,n}(k) := \sum_{l=2}^{k} T_{2,n}(k,l) \tag{13}$$

and

$$T_{6,n}(k) := \left(T_{2,n}(k,2), \cdots, T_{2,n}(k,k)\right) \left(\sum\nolimits^{(k)}\right)^{-1} \left(T_{2,n}(k,2), \cdots, T_{2,n}(k,k)\right)', \tag{14}$$

where $\sum^{(k)}$ is the limiting variance-covariance matrix of $\left(T_{2,n}(k,2), \cdots, T_{2,n}(k,k)\right)$ which is given in (23).

The asymptotic distributions of $T_{1,n}(k,l), T_{2,n}(k,l), T_{i,n}(k), i = 3, \cdots, 6$ will follow from Theorem 1, Lemma 1 and Corollary 1 of the sequel.

Theorem 1. *Let $B(\cdot)$ be a Brownian bridge. Assume that $H_\circ$ of (1) and condition (8) hold. Then, as $n \to \infty$,*

$$\frac{\sqrt{n}}{\sigma} \underline{W}_n(\underline{s}) \xrightarrow{D} \underline{\Psi}_k(\underline{s}) = (\Psi_k(\underline{s}, 2), \cdots, \Psi_k(\underline{s}, k)), \tag{15}$$

where, for $l = 2, \cdots, k$,

$$\begin{aligned} \Psi_k(\underline{s}, l) &= B(s_{l-1}) - B(s_k) + \textstyle\sum_{i=2}^{k-l+1} \{s_{k-i+1} B(s_{k-i+2}) - s_{k-i+2} B(s_{k-i+1})\} \\ &\quad - \textstyle\sum_{i=k-l+3}^{k} \{s_{k-i+1} B(s_{k-i+2}) - s_{k-i+2} B(s_{k-i+1})\}. \end{aligned}$$

For $i = 2, \cdots, k$, define

$$T_2(k,i) = \underbrace{\int \cdots \int}_{\underline{s}} \Psi_k(\underline{s}, i) d\underline{s}, \tag{16}$$

$$a_{i,k} = \frac{1}{(k-i+2)(k-i)!(i-2)!}, b_{i,k} = \frac{1}{(k-i)!(i)!}, c_{i,k} = \frac{1}{(i-1)!(k-i)!}, \tag{17}$$

$$\zeta_i(k,x) = c_{i,k}(1-x)^{k-i}x^{i-1} - c_{k,k}x^{k-1} \tag{18}$$

and

$$\varphi_i(k,x) = (1-x)^{i-2}x^{k-i}\{a_{i,k}x^2 - b_{i,k}(1-x)((i-1)x+1)\}. \tag{19}$$

Lemma 1. *Assume that $H_\circ$ of (1) and condition (8) hold. Let $B(\cdot)$ be a Brownian bridge. Then,*

$$\left(T_{2,n}(k,2),\cdots,T_{2,n}(k,k)\right) \xrightarrow{D} \left(T_2^*(k,2),\cdots,T_2^*(k,k)\right), \tag{20}$$

where, for $l = 2,\cdots,k$,

$$T_2^*(k,l) = \int_0^1 \varphi(x;k,l)B(x)dx \tag{21}$$

and

$$\varphi(x;k,l) = \sum_{i=2}^{k-l+1}\varphi_i(k,x) - \sum_{k-l+3}^{k}\varphi_i(k,x) + \zeta_{l-1}(k,x). \tag{22}$$

Note that $\left(T_2^*(k,2),\cdots,T_2^*(k,k)\right)$ is a mean zero multivariate Normal vector with variance-covariance matrix $\sum^{(k)} = \left[\sum_{uv}^{(k)}\right]$, where

$$\sum\nolimits_{uv}^{(k)} = \int_0^1\int_0^1 \varphi(x;k,u)\varphi(y;k,v)(\min(x,y)-xy)dxdy. \tag{23}$$

Corollary 1. *Assume that $H_\circ$ of* (1) *and condition (8) hold. By the continuous mapping theorem,*

$$T_{1,n}(k,l) \xrightarrow{D} \sup_{\underline{s}} \Psi_k(\underline{s},l) := T_1(k,l), l = 2,\cdots,k, \tag{24}$$

$$\left(T_{2,n}(k,2),\cdots,T_{2,n}(k,k)\right) \xrightarrow{D} \left(T_2^*(k,2),\cdots,T_2^*(k,k)\right), \tag{25}$$

$$T_{3,n}(k) \xrightarrow{D} \max_{2 \le l \le k} \sup_{\underline{s}} \Psi_k(\underline{s}, l) := T_3(k), \tag{26}$$

$$T_{4,n}(k) \xrightarrow{D} \max_{2 \le l \le k} T_2^*(k, l) := T_4(k), \tag{27}$$

$$T_{5,n}(k) \xrightarrow{D} N(0, \sigma_k^2) \tag{28}$$

and

$$T_{6,n}(k) \xrightarrow{D} \chi_{k-1}^2,$$

where $T_4(k)$ is the largest coordinate of a mean zero multivariate normal vector with variance-covariance matrix $\sum^{(k)}, \sigma_k^2 = \underline{1}' \sum^{(k)} \underline{1}$ and $\underline{1}$ is a column vector of 1's.

Let $\sigma_{k,l}^2 = \sum_{l,l}^{(k)}$. Note that

$$T_{2,n}(k, l) \xrightarrow{D} N(0, \sigma_{k,l}^2). \tag{29}$$

SPECIAL CASES

1. The case of $k = 2$

$$T_{1,n}(2,2) \xrightarrow{D} \sup_{0<s_1<s_2<1} (B(s_1) - B(s_2))$$

and

$$T_{2,n}(2,2) \xrightarrow{D} \int_0^1 (1-2x)B(x)dx \stackrel{D}{=} N(0, \frac{1}{180}).$$

2. The case of $k = 3$

$$T_{2,n}(3,2) \xrightarrow{D} \int_0^1 (\frac{5}{6}x^3 - \frac{3}{2}x + \frac{1}{2})B(x)dx \stackrel{D}{=} N(0, \frac{17}{11340}),$$

$$T_{2,n}(3,3) \xrightarrow{D} \int_0^1 (x - \frac{5}{2}x^2 + \frac{5}{6}x^3 + \frac{1}{6})B(x)dx \stackrel{D}{=} N(0, \frac{17}{11340}),$$

$$\sum^{(3)} = \begin{bmatrix} \frac{17}{11340} & \frac{79}{362880} \\ \frac{79}{362880} & \frac{17}{11340} \end{bmatrix} \text{and} \quad \sigma_3^2 = \frac{89}{25920}. \tag{30}$$

3. The case of $k = 4$

$$\varphi(x;4,2) = \frac{1}{4}x^2 - \frac{2}{3}x + \frac{1}{2}x^3 - \frac{7}{24}x^4 + \frac{1}{6},$$

$$\varphi(x;4,3) = \frac{1}{2}x - \frac{7}{4}x^2 + \frac{7}{6}x^3 + \frac{1}{24},$$

$$\varphi(x;4,4) = \frac{1}{6}x - \frac{2}{3}x^3 + \frac{7}{24}x^4 + \frac{1}{24},$$

$$\sum^{(4)} = \begin{bmatrix} \frac{3127}{19958400} & \frac{247}{3628800} & -\frac{79}{1596672} \\ \frac{247}{3628800} & \frac{23}{259200} & \frac{247}{3628800} \\ -\frac{79}{1596672} & \frac{247}{3628800} & \frac{3127}{19958400} \end{bmatrix} \text{and} \; \sigma_4^2 = \frac{29}{50400}. \tag{31}$$

MONTE CARLO CRITICAL VALUES

We simulated the critical values (CV's) of $T_1(k,l)$ of (24) for $l = 2, \cdots, k$ and $k = 3,4$ and $T_3(k)$ of (26) for $k = 3,4$. These CV's are given in Table 1. The details of this simulation study for the case $k = 4$ are given in Appendix A.

The simulated CV's of $T_4(k)$ of (27) for $k = 3,4$ are given in Table 1. To obtain these CV's we followed the steps

1. Use $\sum^{(3)}$ of (30) and $\sum^{(4)}$ of (31).
2. Generate $N = 2000$ realizations of $MVN(0, \sum^{(k)})$
3. For each realization, take the maximum of each vector
4. Order the resulting N values and obtain the corresponding upper percentiles. This gives the CV's for $T_4(k)$.

Table 1. Limiting critical values for $T_{1,n}, T_{3,n}$ and $T_4(k)$ for $k = 3, 4$

α	**0.1**	**0.05**
$K = 3$		
$T_{1,n}(3,2)$	1.743	1.879
$T_{1,n}(3,3)$	1.732	1.868
$T_{3,n}(3)$	1.752	1.896
$T_4(3)$	0.063	0.076
$K = 4$		
$T_{1,n}(4,2)$	1.915	2.088
$T_{1,n}(4,3)$	1.968	2.104
$T_{1,n}(4,4)$	1.917	2.070
$T_{3,n}(4)$	1.979	2.112
$T_4(4)$	0.021	0.025

The simulated finite sample CV's of $T_{1,n}(k,l)$ of (9), $T_{2,n}(k,l)$ of (10) and $T_{i,n}(k)$ for $3 \leq i \leq 6$ defined respectively in (11)-(14) are given in Tables 2-4. In this simulation study we used the kernel.

$$\varphi(x; y) = 2I(x \leq y) - 1.$$

Note that for this kernel $m = 1$ and $\sigma^2 = \frac{1}{3}$, where σ^2 is as in (8). We explain the computation in the case $k = 4$ in Appendix B.

Table 2. Finite sample critical values for $T_{1,n}, T_{3,n}$ for $k = 3, 4$ at $\alpha = 0.05, 0.10$.

	$n =$	**40**		**50**		**100**	
	$\alpha =$	**0.10**	**0.05**	**0.10**	**0.05**	**0.10**	**0.05**
$K = 3$	$T_{1,n}(3,2)$	1.383	1.520	1.431	1.582	1.495	1.655
	$T_{1,n}(3,3)$	1.417	1.554	1.468	1.602	1.483	1.653
	$T_{3,n}(3)$	1.472	1.609	1.519	1.651	1.578	1.736
$K = 4$	$T_{1,n}(4,2)$	1.383	1.492	1.347	1.430	1.434	1.574
	$T_{1,n}(4,3)$	1.362	1.478	1.401	1.503	1.576	1.666
	$T_{1,n}(4,4)$	1.426	1.490	1.424	1.521	1.603	1.697
	$T_{3,n}(4)$	1.430	1.506	1.440	1.533	1.626	1.717

Table 3. Finite sample critical values for $T_{2,n}, T_{4,n}, T_{5,n}$ and $T_{6,n}$ for $k = 3$

$K = 3$	$n = 40$		50		100	
α	0.1	0.05	0.1	0.05	0.1	0.05
$T_{2,n}(3,2)$	1.059	1.340	1.178	1.444	1.157	1.597
$T_{2,n}(3,3)$	1.100	1.402	1.226	1.535	1.311	1.746
$T_{4,n}$	1.349	1.641	1.477	1.787	1.619	1.972
$T_{5,n}$	1.156	1.470	1.207	1.514	1.247	1.592
$T_{6,n}$	3.395	4.559	3.473	4.577	4.375	5.466

Table 4 Finite sample critical values for $T_{2,n}, T_{4,n}, T_{5,n}$ and $T_{6,n}$ for $k = 4$

$K = 4$	$n = 40$		50		100	
α	0.1	0.05	0.1	0.05	0.1	0.05
$T_{2,n}(4,2)$	0.854	1.100	0.979	1.220	1.090	1.457
$T_{2,n}(4,3)$	0.991	1.234	1.036	1.368	1.166	1.491
$T_{2,n}(4,4)$	0.975	1.219	1.109	1.421	1.249	1.687
$T_{4,n}(4)$	1.257	1.515	1.441	1.628	1.681	1.987
$T_{5,n}(4)$	0.334	0.408	1.061	1.374	1.177	1.504
$T_{6,n}(4)$	0.767	1.000	4.739	6.374	5.948	6.928

MONTE CARLO POWER COMPARISONS

We conducted a Monte Carlo power study to compare the powers of the proposed tests with those of several competing tests. In this study we used the normal (light tail) and the double-exponential (heavy tail) distributions. We considered the case of $n = 50, \alpha = 0.05$ and $k = 3$ with known $l = 2, 3$ and with unknown l. We applied several combinations of change points and jump sizes. For the change points we used the combinations $a = (5, 15, 25), b = (5, 15, 35), c = (5, 15, 45), d = (15, 25, 35)$ and $e = (15, 25, 45)$ which reflect changes close to the beginning, the middle and the end of the sample. We took the location parameter of X_1 as zero. The sizes of the location shifts at the change points $1 < r < s < t < n$ are respectively determined by solving $P_1 = P\{X_{r+1} > X_1\}, P_2 = P\{X_{s+1} > X_{r+1}\}$ and $P_3 = P\{X_{t+1} > X_{s+1}\}$. For (P_1, P_2, P_3) we used the combinations $I = (0.7, 0.2, 0.3), II = (0.7, 0.4, 0.3)$ and $III = (0.7, 0.4, 0.2)$, when $l = 2$, and the combinations $IV = (0.7, 0.8, 0.3), V = (0.7, 0.6, 0.3)$ and $VI = (0.7, 0.6, 0.2)$, when $l = 3$. Note that $P_i < 0.5$ corresponds to a downward change and $P_i > 0.5$ corresponds to an upward change. We estimated the powers of the following test statistics:

1. $T_{1,n}(3, l), T_{2,n}(3, l)$ for $l = 2, 3$ and $T_{i,n}(3), i = 3, \cdots, 6$.

2. $A(3)$ and $A^*(3)$ (resp. $T_{n1}(3)$ and $T^*_{n1}(3)$ of (2.4) and (2.5) of Aly *et al.* (2003)) which are consistent against the ordered multiple change points alternative.

3. t_{1n} and t_{2n} (adjusted for $k = 3$) of page 364 of Aly & BuHamra (1996) which are consistent against the general multiple change points alternative.

The powers were estimated using the NAG Library based on 5,000 realizations under the alternative hypotheses. The power results are summarized in Tables 5-6.

The powers of $T_{1,n}$, $T_{3,n}$ and $T_{4,n}$ are high compared to the other tests for the selected positions of the change points. The powers of the proposed tests are high, when the jump size is large. The powers of $T_{2,n}(3,2)$ and $T_{6,n}$ fluctuate with the positions of the change points and the size of the jumps. Also, the test $T_{2,n}(3,3)$ is better than $T_{2,n}(3,2)$ in detecting the change point for all cases. In general $T_{2,n}(3,3)$ is a good test. The two tests t_{1n} and t_{2n} designed for testing against the general multiple change point alternative performed good and managed to detect the change point better than the tests $A(3)$ and $A^*(3)$ which are designed for testing against the ordered multiple change point alternative.

Table 5. Power values for the normal (double exponential) distribution when $\alpha = 0.05, n = 50, k = 3$ and $l = 2$. The change points (CP) are: $a = (5,15,25), b = (5,15,35), c = (5,15,45), d = (15,25,35), e = (15,25,45)$. The probabilities of jumps (P) are: $I = (P_1, P_2, P_3) = (0.7,0.2,0.3), II = (0.7,0.4,0.3), III = (0.7,0.4,0.2)$.

CP	P	$T_1(3,2)$	$T_1(3,3)$	T_3	$T_2(3,2)$	$T_2(3,3)$	T_4	T_5	T_6	A	A^*	t_1	t_2
a	*I*	99	86	98	83	73	78	85	1	11	1	100	97
		(57)	(68)	(61)	(87)	(87)	(87)	(83)	(8)	(39)	(61)	(68)	(69)
	II	12	7	11	54	26	48	51	3	12	1	7	5
		(44)	(33)	(38)	(35)	(24)	(26)	(43)	(2)	(1)	(3)	(28)	(30)
	III	97	88	96	83	12	71	61	13	12	1	96	90
		(80)	(67)	(75)	(83)	(9)	(21)	(38)	(13)	(0)	(0)	(68)	(64)
b	*I*	96	69	95	91	99	99	93	9	17	2	95	88
		(81)	(90)	(87)	(97)	(98)	(98)	(94)	(29)	(45)	(75)	(87)	(82)
	II	13	9	10	65	64	59	62	4	16	2	6	7
		(56)	(53)	(53)	(67)	(62)	(62)	(66)	(7)	(2)	(5)	(39)	(37)
	III	94	78	92	75	68	69	77	1	16	2	90	80
		(88)	(82)	(85)	(79)	(71)	(72)	(80)	(1)	(0)	(1)	(74)	(71)

Cont. Table 5. Power values for the normal (double exponential) distribution when $\alpha = 0.05, n = 50, k = 3$ and $l = 2$. The change points (CP) are: $a = (5, 15, 25), b = (5, 15, 35), c = (5, 15, 45), d = (15, 25, 35), e = (15, 25, 45)$. The probabilities of jumps (P) are: $I = (P_1, P_2, P_3) = (0.7, 0.2, 0.3), II = (0.7, 0.4, 0.3), III = (0.7, 0.4, 0.2)$.

CP	P	$T_1(3,2)$	$T_1(3,3)$	T_3	$T_2(3,2)$	$T_2(3,3)$	T_4	T_5	T_6	A	A^*	t_1	t_2
	I	84	37	79	81	98	96	53	59	33	5	77	63
		(89)	(97)	(95)	(80)	(97)	(97)	(82)	(59)	(84)	(93)	(93)	(86)
c	*II*	12	8	10	35	54	44	29	19	33	7	6	6
		(38)	(48)	(46)	(46)	(56)	(56)	(44)	(21)	(19)	(24)	(33)	(26)
	III	47	22	42	47	54	43	28	20	33	8	29	16
		(55)	(60)	(59)	(46)	(55)	(55)	(43)	(19)	(7)	(16)	(39)	(30)
	I	96	84	95	87	92	86	24	85	64	20	96	94
		(85)	(96)	(94)	(86)	(88)	(88)	(46)	(94)	(96)	(98)	(96)	(92)
d	*II*	19	25	22	62	74	63	33	39	63	20	17	21
		(50)	(57)	(52)	(62)	(72)	(72)	(52)	(37)	(23)	(35)	(46)	(46)
	III	91	77	86	83	79	71	52	20	62	21	85	83
		(74)	(71)	(71)	(79)	(76)	(76)	(65)	(17)	(2)	(9)	(58)	(67)
	I	84	57	78	80	87	77	1	99	80	37	78	72
		(97)	(100)	(99)	(80)	(83)	(83)	(22)	(99)	(100)	(100)	(100)	(98)
e	*II*	18	24	22	50	66	54	10	74	80	36	19	24
		(66)	(79)	(76)	(63)	(65)	65	(28)	(70)	(64)	(70)	(68)	(54)
	III	81	53	70	77	70	60	11	73	80	36	35	31
		(76)	(83)	(80)	(68)	(63)	(63)	(31)	(69)	(42)	(53)	(66)	(52)

Table 6. Power values for the normal (double exponential) distribution when $\alpha = 0.05, n = 50, k = 3$ and $l = 3$. The change points (CP)

CP	P	$T_1(3,2)$	$T_1(3,3)$	T_3	$T_2(3,2)$	$T_2(3,3)$	T_4	T_5	T_6	A	A^*	t_1	t_2
	IV	68	68	66	60	73	74	85	1	11	1	58	74
		(52)	(68)	(62)	(18)	(87)	(81)	(71)	(8)	(11)	(1)	(68)	(70)
a	*V*	42	33	36	54	26	29	(45	3	12	1	39	37
		(44)	(33)	(38)	(50)	(24)	(44)	(49)	(3)	(12)	(1)	(28)	(30)
	VI	82	70	77	83	11	22	42	13	12	1	73	70
		(80)	(67)	(75)	(83)	(10)	(72)	(57)	(13)	(12)	(1)	(68)	(64)

continue next page

Cont. Table 6. Power values for the normal (double exponential) distribution when $\alpha = 0.05, n = 50, k = 3$ and $l = 3$. The change points (CP)

CP	P	$T_1(3,2)$	$T_1(3,3)$	T_3	$T_2(3,2)$	$T_2(3,3)$	T_4	T_5	T_6	A	A^*	t_1	t_2
b	*IV*	96	97	96	21	99	99	99	9	17	2	94	93
		(81)	(90)	(87)	(7)	(98)	(97)	(82)	(21)	(17)	(2)	(87)	(82)
	V	56	55	53	25	64	64	69	4	16	2	35	25
		(56)	(53)	(53)	(27)	(62)	(57)	(60)	(4)	(16)	(2)	(39)	(37)
	VI	89	91	92	45	68	69	78	1	17	2	42	29
		(88)	(82)	(85)	(49)	(72)	(70)	(78)	(1)	(17)	(2)	(74)	(71)
c	*IV*	95	98	97	21	98	98	86	68	33	5	94	87
		(89)	(97)	(95)	(10)	(97)	(94)	(50)	(59)	(33)	(5)	(93)	(86)
	V	42	52	48	25	54	54	41	19	33	7	48	50
		(39)	(48)	(46)	(26)	(56)	(44)	(27)	(10)	(33)	(7)	(33)	(23)
	VI	92	94	94	37	54	54	41	21	33	5	58	65
		(55)	(60)	(59)	(36)	(55)	(43)	(30)	(19)	(33)	(5)	(39)	(30) e
d	*IV*	86	95	93	42	92	92	62	85	64	20	91	91
		(85)	(96)	(94)	(39)	(88)	(81)	(10)	(94)	(64)	(20)	(95)	(92)
	V	54	61	58	42	74	74	53	39	63	20	71	56
		(50)	(57)	(52)	(41)	(72)	(62)	(32)	(37)	(63)	(20)	(46)	(46)
	VI	84	90	88	38	79	79	67	20	62	21	67	56
		(74)	(71)	(71)	(37)	(76)	(67)	(51)	(17)	(62)	(21)	(58)	(67)
e	*IV*	99	100	100	47	87	87	21	99	80	37	25	29
		(97)	(100)	(99)	(40)	(83)	(72)	(2)	(97)	(80)	(37)	(100)	(98)
	V	70	83	80	38	66	66	29	74	80	36	66	64
		(66)	(79)	(76)	(22)	(65)	(51)	(10)	(70)	(80)	(36)	(67)	(54)
	VI	95	97	94	35	70	70	32	73	80	36	66	56
		(76)	(83)	(80)	(21)	(63)	(52)	(10)	(69)	(80)	(36)	(66)	(52)

AN EXAMPLE

Moment magnitude (Mw) readings of earthquakes 2000 kilometers outside Kuwait for 206 days in 2002/2003 are given in Figure 1 (in logarithmic scale). The corresponding Cusum plot is given in Figure 2 which suggests that there might be changes in Mw at days 71, 109 and 173. Al-Hulail (2009) used Bayesian inference and assigned normal distribution with mean $\lambda_i, i = 1, 2, 3, 4,$ and precision τ as likelihood function. Appropriate priors are set to the parameters $\lambda_i, i = 1, 2, 3, \tau$ and the position of the three change points $1 < k_1 < k_2 < k_3 < n$. Then, the posterior and the marginal distributions for the parameters λ_i, τ, k_i are derived. Using BUGS 1.4.3 software (Bayesian inference using Gibbs Sampling) and 100, 000 iterations, Al-Hulail (2009) reported that the most posterior mode for the positions of change are 72, 139 and 173. The corresponding predicted rates are 5.79, 5.70, 5.91 and 5.69. For this data $T_{2,206}(3, 2) = 1.767$ with P-value 0.0386. This confirms the umberlla alternative with $k = 3$ and $l = 2$.

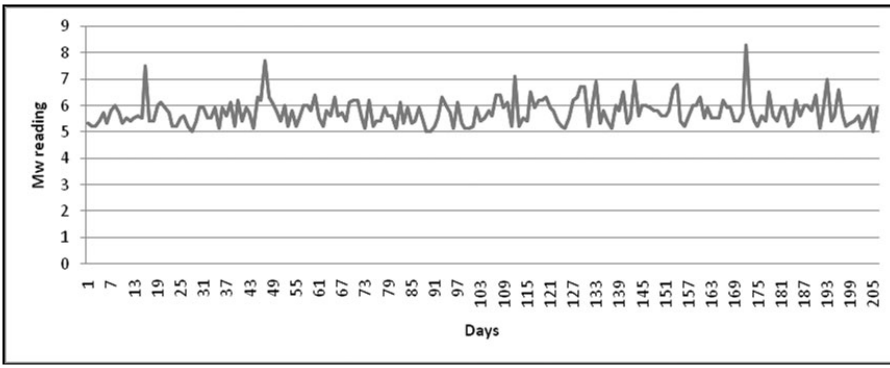

Fig.1. Mw readings of earthquakes for 206 days in 2002/2003 in Kuwait.

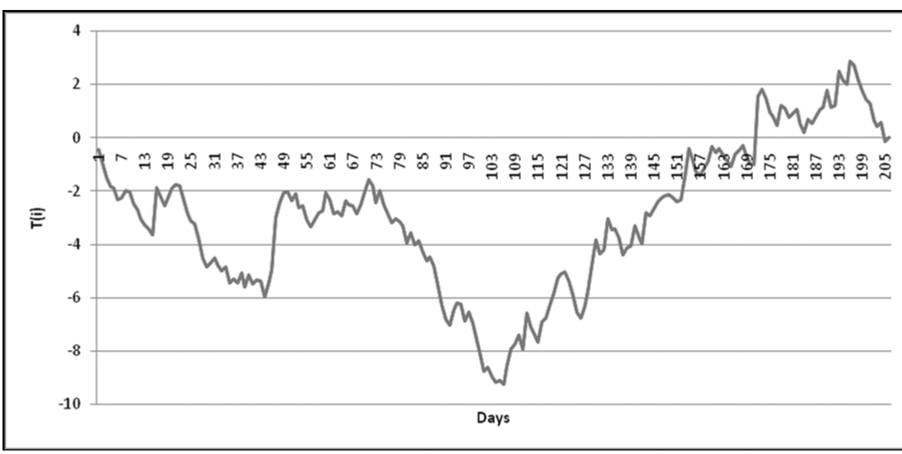

Fig. 2. The Cusum plot for the Mw readings of earthquakes for 206 days in 2002/2003 in Kuwait.

PROOFS

Let $\phi_1(\cdot)$ and σ be as in (7) and (8), respectively. Define

$$\xi_i = \sigma^{-1}\phi_1(X_i), S_i = \sum_{j=1}^{i} \xi_j, S(i) = S_i - \frac{i}{n} S_n, \tag{32}$$

$$Y_{i,j,n} = d_{j+1,n}\{S([ns_i]) - S([ns_{i-1}])\} - d_{i,n}\{S([ns_{j+1}]) - S([ns_j])\}, \tag{33}$$

$$Z_n(\underline{s}, l) = \sum_{i=1}^{l-1}\sum_{j=i}^{l-1} Y_{i,j,n} - \sum_{i=l}^{k}\sum_{j=i}^{k} Y_{i,j,n}, \tag{34}$$

$$\begin{aligned} Z_n^*(\underline{s}, l) &= \textstyle\sum_{i=2}^{k-l+1}\{[ns_{k-i+1}]S([ns_{k-i+2}]) - [ns_{k-i+2}]S([ns_{k-i+1}])\} \\ &- \textstyle\sum_{i=k-l+3}^{k}\{[ns_{k-i+1}]S([ns_{k-i+2}]) - [ns_{k-i+2}]S([ns_{k-i+1}])\} \\ &+ S([ns_{l-1}]) - S([ns_k]), \end{aligned} \tag{35}$$

$$\underline{Z}_n(\underline{s}) = (Z_n(\underline{s}, 2), \cdots, Z_n(\underline{s}, k)) \text{ and } \underline{\mathrm{Z}}_\mathrm{n}^*(\underline{\mathrm{s}}) = \left(\mathrm{Z}_\mathrm{n}^*(\underline{\mathrm{s}}, 2), \cdots, \mathrm{Z}_\mathrm{n}^*(\underline{\mathrm{s}}, \mathrm{k})\right). \tag{36}$$

It is easy to see that

$$\underline{Z}_n(\underline{s}) = \underline{Z}_n^*(\underline{s}).$$

Proof of Theorem 1: By the proof of Lemma 1 of Aly & Kochar (1997),

$$\max_{i,j:(d_{i,n} \wedge d_{j+1,n}) \geq 1} \left| \sqrt{n}\sigma^{-1} U_{i,j,n} - n^{-\frac{3}{2}} Y_{i,j,n} \right| \overset{p}{=} O(n^{-\frac{1}{2}}). \tag{37}$$

By (5) and (34)-(37),

$$\sup_{\underline{s}} \left| \underline{W}_n(\underline{s}) - n^{-\frac{3}{2}} \underline{Z}_n^*(\underline{s}) \right| \overset{p}{=} O(n^{-\frac{1}{2}}). \tag{38}$$

By (38) and the results of Section 4 of Csörgő & Horváth (1988b) (see also, Csörgő & Horváth (1997)) we obtain (15).

Proof of Lemma 1: By (15) we have

$$\left(T_{2,n}(k, 2), \cdots, T_{2,n}(k, k)\right) \xrightarrow{D} (T_2(k, 2), \cdots, T_2(k, k)),$$

where

$$T_2(k, l) = \int \cdots \int_{\underline{s}} \Psi_k(\underline{s}, l) d\underline{s}.$$

By straightforward but tedious integrations we obtain (20).

ACKNOWLEDGMENT

This research is supported by research grant SS05/01 of Kuwait University. Part of this research was done while the second author was visiting Carleton University, Canada, on a sabbatical leave from Kuwait University.

REFERENCES

Al-Hulail, D. 2009. Bayesian inference for change point problems. MSc project at Kuwait University.

Aly, E. E. 2004. Change point detection based on L-statistics. Fields Institute Communications, **44**: 293-300.

Aly, E. E., Abd-Rabou, A. S. & Al-Kandari, N. M. 2003. Tests for multiple change points under ordered alternatives. Metrika, **57**: 209-221.

Aly, E. E. & BuHamra, S. 1996. Rank tests for two change points. Computational Statistics & Data Analysis, **22**: 363-372.

Aly, E. E. & Kochar, S. C. 1997. Change point tests based on U-statistics with applications in reliability. Metrika, **45**: 259-269.

Barlow, R. E. & Brunk, H. D. 1972. The isotonic regression problem and its dual. Journal of the American Statistical Association, **67**: 140-147.

Bhattacharyya, G. K. 1984. Tests for randomness against trend or serial correlations. In: P.R. Krishnaiah & P.K. Sen, Eds., Handbook of Statistics, **V4**, Nonparametric Methods. North-Holland, Amsterdam, pp 89-111.

Csörgő, M. & Horváth, L. 1988a. Nonparametric methods for change- point problems. In: P.R. Krishnaiah and C.R. Rao, Eds., Handbook of Statistics, **V7**, Quality Control and Reliability. North-Holland, Amsterdam, pp 403-425.

Csörgő, M. & Horváth, L. 1988b. Invariance principles for changepoint problems. Journal Multivariate Analysis, **27**: 151-168.

Csörgő, M. & Horváth, L. 1997. Limit Theorems in Change-Point Analysis. John Wiley & Sons, New York.

Döring, M. 2010. Multiple change-point estimation with U-statistics. Journal of Statistical Planning and Inference, **140**: 2003-2017.

Hušková, M. & Sen, P. K. 1989. Nonparametric tests for shift and change in regression at an unknown time point. In: P. Hackl, Ed., Statistical Analysis and Forecasting of Economic Structural Change, Springer Verlag, New York, pp 71-85.

Hušková, M. & Slabý, A. 2001. Permutation tests for multiple changes. Kybernetika, **37**: 605-622.

Lavielle, M. & Teyssiére, G. 2006. Detection of multiple change-points in multivariate time series. Lithuanian Mathematical Journal, **46**: 287-306.

Lebarbier, E. 2005. Detecting multiple change-points in the mean of Gaussian process by model selection. Signal Processing, **85**: 717-736.

Lin, C. H. & Kao, T. C. 2008. Multiple structural changes in the tail behavior: Evidance from stock index futures return. Nonlinear analysis, **9**: 1702-1713.

Lombard, F. 1987. Rank tests for changepoint problems. Biometrika, **74**: 615-624.

Lombard, F. 1989. Some recent developments in the analysis of change-point data. South African Statistical Journal, **23**: 1-21.

Mack, G. A. & Wolfe, D. A. 1981. K-sample rank tests for umbrella alternative. Journal of the American Statistical Association, **76**: 175-181.

Robertson, T., Wright, T. & Dykstra, R. 1988. Order restricted Statistical Inference. John Wiley and Sons, New York.

Sen, P. K. 1988. Robust tests for change-point models. In: S. Kotz and N.L. Johnson, Eds-in-chief, Encyclopedia of Statistical Sciences, **8**, John Wiley and Sons, New York, pp 173-176.

Xiong, Ch., Yan, Y. & Ji, M. 2003. Tests of homogeneity for umbrella alternative in dose-response relationships for Poisson variables. Journal of Modern Applied Statistical Methods, **2**: 389-399.

Zacks, S. 1983. Survey of classical and Bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation. In: M.H. Rizvi, J.S. Rustagi and D. Siegmund, Eds., Recent Advances in Statistics, Academic Press, New York, pp 245-269.

Submitted : *26/11/2012*
Revised : *16/09/2013*
Accepted : *24/09/2013*

APPENDIX A: CRITICAL VALUES OF $T_1(k,l), T_3(k)$

For $M = 2000$ and $N = 2000$, we generate N realizations $(Z_1, \cdots, Z_M) \sim MVN(0, \Lambda)$, where for $i \leq j$,

$$\Lambda_{i,j} = \frac{i}{M+1}(1 - \frac{j}{M+1}).$$

We then follow the steps:

1. For each realization and for every $1 \leq i_1 < i_2 < i_3 < i_4 \leq M$, compute

$$\begin{aligned}\widetilde{\Psi}_4(i_1, i_2, i_3, i_4; l) \quad &= \frac{1}{M+1} \sum_{h=2}^{4-l+1} \left\{ i_{4-h+1} Z_{i_{4-h+2}} - i_{4-h+2} Z_{i_{4-h+1}} \right\} \\ &- \frac{1}{M+1} \sum_{h=4-l+3}^{4} \left\{ i_{4-h+1} Z_{i_{4-h+2}} - i_{4-h+2} Z_{i_{4-h+1}} \right\} \\ &+ Z_{i_{l-1}} - Z_{i_4}.\end{aligned}$$

2. Take the maximum over all $1 \leq i_1 < i_2 < i_3 < i_4 \leq M$.
3. Order the resulting N values and obtain the corresponding upper percentiles. This gives the CV's for $T_1(4, l)$.
4. Repeat steps 1 and 2 for $l = 2, 3, 4$ and take the maximum over l. Order the resulting N values and obtain the corresponding upper percentiles. This gives the CV's for $T_3(4)$.

APPENDIX B: MONTE CARLO FINITE SAMPLE CRITICAL VALUES OF $T_{1,n}(k,l)$, $T_{2,n}(k,l)$ AND $T_{j,n}(k), 3 \leq j \leq 6$

1. Generate $N = 1000$ samples of size $n(= 40, 50, 100)$ from $N(0, 1)$.
2. Let $s_0 = 0$ and $s_5 = 1$. For each realization and for every $1 \leq i_1 < i_2 < i_3 < i_4 \leq n$ take $s_1 = \frac{i_1}{n}, s_2 = \frac{i_2}{n}, s_3 = \frac{i_3}{n}, s_4 = \frac{i_4}{n}$ and compute

$$U_{r,t,n} = \frac{1}{n^2} \left\{ 2 \sum_{u=1}^{(i_r - i_{r-1})} \sum_{v=1}^{(i_{t+1} - i_t)} I\big(X_{i_{r-1}+u} \leq X_{i_t+v}\big) - (i_r - i_{r-1}) \times (i_{t+1} - i_t) \right\}$$

and

$$W_n(\underline{s}, l) = \sum_{r=1}^{l-1} \sum_{t=r}^{l-1} U_{r,t,n} - \sum_{r=l}^{4} \sum_{t=r}^{4} U_{r,t,n}.$$

3. Take the maximum of $W_n(\underline{s}, l)$ over all $1 \leq i_1 < i_2 < i_3 < i_4 \leq n$ and compute $T_{1,n}(k, l)$ of (9).
4. Order the resulting N values and obtain the corresponding upper percentiles. This gives the finite sample CV's for $T_{1,n}(k, l)$.
5. Repeat steps 1 and 2 and compute $T_{2,n}(k, l)/\sigma_{k,l}$, where $\sigma^2_{k,l}$ is as in (29) and $T_{2,n}(k, l)$ is as in (10).
6. Order the resulting N values. The corresponding upper percentiles are the finite sample CV's for $T_{2,n}(k, l)/\sigma_{k,l}$.
7. Repeat steps 1,2 and 3 and compute $T_{3,n}(k)$ of (11).
8. Order the resulting N values and obtain the corresponding upper percentiles. This gives the finite sample CV's for $T_{3,n}(k)$.
9. Repeat steps 1,2 and 5 and compute $T_{4,n}(k)/\sigma_{k,l}$, where $T_{4,n}(k)$ is as in (12).
10. Order the resulting N values. The corresponding upper percentiles are the finite sample CV's for $T_{4,n}(k)/\sigma_{k,l}$.
11. Repeat steps 1,2 and 5 and compute $T_{5,n}(k)/\sigma_{k,l}$, where $T_{5,n}(k)$ is as in (13).
12. Order the resulting N values and obtain the corresponding upper percentiles. This gives the finite sample CV's for $T_{5,n}(k)/\sigma_{k,l}$.
13. Repeat steps 1,2 and 5 and compute $T_{6,n}(k)$ of (14).
14. Order the resulting N values and obtain the corresponding upper percentiles. This gives the finite sample CV's for $T_{6,n}(k)$.

اختبارات لنقاط التغيير المتعددة للفرض البديل من نوع المظلة

فهيمة العوضي و عماد الدين علي أحمد علي
قسم الإحصاء وبحوث العمليات - جامعة الكويت - الكويت

خلاصة

في هذا البحث ندرس مشكلة اختبار فرضية العدم حيث لا يوجد تغير في مقابل خيار وجود عدة نقاط تغيير من نوعية "المظلة" وذلك في عدد من المشاهدات المستقلة. ونقوم بتطوير اختبار ماك وولف (1981) لظروف اختبارات نقطة التغيير. نأخذ بعين الاعتبار حالتي النقطتين العلويتيين المعلومة والغير معلومة.

نحصل على التوزيع التقاربي للاختبارات المقترحة تحت فرض العدم ونعطي قيم تقريبية للقيم الحرجة لهذه التوزيعات. كذلك نعطي جداول للقيم الحرجة للعينات ذات الحجم المحدود باستخدام طرق مونت كارلو. نقدم نتائج دراسات لقوة الاختبارات باستخدام طرق مونت كارلو لمقارنة الاختبارات المقترحة مع عدد من اختبارات نقاط التغيير المتعددة. نقوم بتوضيح الاختبارات المقترحة باستخدام بيانات فعلية.

المجلة العربية للعلوم الإدارية

Arab Journal of Administrative Sciences

رئيس التحرير: أ. د. آدم غازي العتيبي

- صدر العدد الأول في نوفمبر ١٩٩٣ .
- علمية محكمة تعنى بنشر البحوث الأصيلة في مجال العلوم الإدارية.
- تصدر عن مجلس النشر العلمي في جامعة الكويت ثلاثة إصدارات سنوياً (يناير - مايو - سبتمبر).
- تسهم في تطوير الفكر الإداري واختبار الممارسات الإدارية وإثرائها.
- مسجلة في قواعد البيانات العالمية.
- تخضع للتقييم الأكاديمي الخارجي.

- First issue, November 1993.
- Refereed journal publishing original research in Administrative Sciences.
- Published by Academic Publication Council,Kuwait University,3 issues a year (January, May, September).
- Contributes to developing and enriching administrative thinking and practices.
- Listed in several international databases.
- Reviewed periodically by international referees for high academic standards.

الاشــــتراكات

الكويت: 3 دنانير للأفراد - 15 ديناراً للمؤسسات - الدول العربية: 4 دنانير للأفراد - 15 ديناراً للمؤسسات
الدول الأجنبيـة: 15 دولاراً للأفراد - 60 دولاراً للمؤسسات

توجه المراسلات إلى رئيس التحرير على العنوان الآتي:

المجلة العربية للعلوم الإدارية - جامعة الكويت ص.ب: 28558 الصفاة 13146 - دولة الكويت

هاتف: Tel: (965)24827317 - أو 4734 / 4416 /(965)24984415 - فاكس: Fax: (965)24817028

E-mail: ajas@ku.edu.kw - Web Site: http://www.pubcouncil.kuniv.edu.kw/ajas