International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 3 (Special), 2004, pp 153–162

On testing for multiple change points

Emad-Eldin A. A. Aly
Department of Statistics & Operations Research,
College of Science, Kuwait University
P.O. Box 5969, Safat 13060, Kuwait

[Received May 12, 2004; Accepted August 14, 2004]

Abstract

We consider the problem of testing the null hypothesis of no change against the alternative of multiple change points in a series of independent observations. We consider the case of testing against the general multiple change point alternative and the case when the changes are in the same direction. We report the asymptotic null distribution of the considered tests. We also give approximations for their limiting critical values.

Keywords and Phrases: L-statistics, Likelihood ratio tests, Ordered alternatives, Rank tests, U-statistics.

AMS Classification: Primary 62G10 ; Secondary 60G20.

1 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. We consider here the problem of testing the null hypothesis of no change

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

against the general multiple change points alternative

$$\begin{aligned} H_1 \quad &: \quad \exists\, 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < 1 \ \text{ s.t.} \\ F_1 \quad &= \quad \cdots = F_{[n\lambda_1]} \overset{\mathrm{P}}{\neq} F_{[n\lambda_1]+1} = \\ \cdots \quad &= \quad F_{[n\lambda_2]} \overset{\mathrm{P}}{\neq} \cdots \overset{\mathrm{P}}{\neq} F_{[n\lambda_k]+1} = \cdots = F_n, \end{aligned} \tag{2}$$

or against the ordered multiple change points alternative

$$\begin{aligned} H_{11} &: \exists\, 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < 1 \text{ s.t.} \\ F_1 &= \cdots = F_{[n\lambda_1]} \overset{\text{P}}{\prec} F_{[n\lambda_1]+1} = \\ \cdots &= F_{[n\lambda_2]} \overset{\text{P}}{\prec} \cdots \overset{\text{P}}{\prec} F_{[n\lambda_k]+1} = \cdots = F_n, \end{aligned} \tag{3}$$

where $[y]$ is the integer part of y and $\overset{\text{P}}{\prec}$ is a partial ordering of the family of DF's under consideration. In H_1 and H_{11} we may consider various types of restrictions on the nature of the difference between F and G, parametric as well as nonparametric. The following partial oderings are of interest in Reliability: Hazard Rate, Stochastic, Dispersive, IFR, $IFRA$, NBU, $NBU - t_\circ$, $DMRL$, $NBUE$, $HNBUE$,$\cdots$

Consider first the problem of testing $H_\circ$ of (1) against the At Most One Change (AMOC) point alternative

$$\begin{aligned} H_1 &: \exists\, 0 < \lambda < 1 \ \text{ s.t.} \\ F_1 &= \cdots = F_{[n\lambda]} \overset{\text{P}}{\prec} F_{[n\lambda]+1} = \cdots = F_n. \end{aligned} \tag{4}$$

Assume that λ is known and $[n\lambda] = r$. In this case the above testing problem reduces to a 2-sample problem based on the samples $X_1, \cdots, X_r$ and $X_{r+1}, \cdots, X_n$. Let $S_{r,n-r}$ be an appropriate test statistic for this 2-sample problem. Define the stochastic process $S_n(.)$ on $[0,1]$ by $S_n(0) = S_n(1) = 0$ and for $l = 1, 2, \cdots, n-1$,

$$S_n(\frac{l}{n}) = a(l,n) S_{l,n-l},$$

where $a(l,n)$ is an appropriate normalizing function (non-random). Tests for the AMOC point problem are functionals of $S_n(\cdot)$. Examples are:

$$\begin{aligned} S_{1n} &= \max_{1 \le l < n} S_n(\frac{l}{n}), \\ S_{2n} &= \max_{1 \le l < n} \left| S_n(\frac{l}{n}) \right|, \\ S_{3n} &= \frac{1}{n} \sum_{1}^{n-1} S_n^2(\frac{l}{n}). \end{aligned}$$

In most situations we are able to prove that

$$S_n(.) \xrightarrow{\text{d}} B(\cdot),$$

where $B(\cdot)$ is a Brownian Bridge. Hence, by the Continuous Mapping Theorem,

$$S_{1n} \xrightarrow{\text{d}} \sup_t B(t) = S_1,$$

$$S_{2n} \xrightarrow{\mathrm{d}} \sup_t |B(t)| = S_2$$

and

$$S_{3n} \xrightarrow{\mathrm{d}} \int B^2(t)dt = S_3.$$

Critical values for S_1, S_2 and S_3 are well known.

The tests presented in the sequel are extensions of appropriate k-sample tests against multiple or ordered alternatives to the change point set-up.

2 Testing against ordered alternatives

In this section we present three families of tests for testing $H_\circ$ of (1) against H_{11} of (3).

2.1 Jonckheere-Terpstra-type tests

We describ here the work of Aly *et al* (2003). Consider the case when the hypothesis in (3) is expressed in terms of the stochastic ordering, i.e., the case of testing against the hypothesis

$$\begin{aligned} H_2 \quad &: \quad \exists\, 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < 1 \text{ s.t.} \\ F_1 \quad &= \quad \cdots = F_{[n\lambda_1]} \overset{\mathrm{ST}}{\prec} F_{[n\lambda_1]+1} = \\ \cdots \quad &= \quad F_{[n\lambda_2]} \overset{\mathrm{ST}}{\prec} \cdots \overset{\mathrm{ST}}{\prec} F_{[n\lambda_k]+1} = \cdots = F_n. \end{aligned} \tag{5}$$

Let $\underline{s} = (0 < s_1 < \cdots < s_k < 1)$. Define the processes $\{R_{i,j,n}(\underline{s})\}$, $n \geq 1$, $1 \leq i \leq j \leq n$ by

$$R_{i,j,n}(\underline{s}) = \sum_{r=[ns_{i-1}]+1}^{[ns_i]} \sum_{l=[ns_j]+1}^{[ns_{j+1}]} I\left(X_r < X_l\right), \tag{6}$$

where $s_\circ = 0, s_{k+1} = 1, I\left(A\right)$ is the indicator function of the event A and any sum with no element is defined to be zero. Define

$$V_n(\underline{s}) = n^{-\frac{3}{2}} \sum_{i=1}^{k} \sum_{j=i}^{k} \left\{ R_{i,j,n}(\underline{s}) - \frac{1}{2}\, d_{i,n} d_{j+1,n} \right\}, \tag{7}$$

where $d_{i,n} = [ns_i] - [ns_{i-1}]$. Based on $\{V_n(\underline{s})\}$ Aly *et al* (2003) proposed the Jonckheere-Terpstra-type test statistics

$$T_{n1}(k) := \max_{\underline{s}} \sqrt{12}\, V_n(\underline{s}), \tag{8}$$

and

$$T_{n1}^{*}(k) := \int \cdots \int_{\underline{s}} \sqrt{12}\, V_n(\underline{s}) d\underline{s}. \tag{9}$$

The asymptotic distributions of T_{n1} and T_{n1}^{*} follow from the following Theorem of Aly *et al* (2003).

Theorem 2.1. *Let $B(\cdot)$ be a Brownian bridge and assume that $H_\circ$ of* (1) *holds. Then, as $n \to \infty$,*

$$\sqrt{12}\, V_n(\underline{s}) \xrightarrow{\mathrm{d}} \Psi(\underline{s}), \tag{10}$$

where

$$\Psi(\underline{s}) \stackrel{\mathrm{d}}{=} \sum_{j=1}^{k} (s_{j+1} - s_{j-1})\, B(s_j). \tag{11}$$

Corollary 2.1. By the Continuous Mapping Theorem,

$$T_{n1}(k) \xrightarrow{\mathrm{d}} \sup_{\underline{s}} \Psi(\underline{s}) = T(k) \tag{12}$$

and

$$T_{n1}^{*}(k) \xrightarrow{\mathrm{d}} \int \cdots \int_{\underline{s}} \Psi(\underline{s}) d\underline{s} = T^{*}(k). \tag{13}$$

Let L_k be the DF of $T(k)$ of (12). By (12) the test

$$\text{Rejet } H_\circ \Longleftrightarrow T_{n1}(k) > c_\alpha \tag{14}$$

is an asymptotic level-α test. The DF L_k is analytically not known in the literature. Aly et al (2003) used a Monte Carlo method to approximate c_α of (14) when $k = 2$. It can be proved that for $k \geq 2$,

$$T^{*}(k) = \int_0^1 \varphi_k(t) B(t) dt,$$

where for $k = 2, 4, \cdots$

$$\varphi_k(x) = \frac{1}{k!}\left(1 - 2x^k\right) + \sum_{j=1}^{\frac{k-2}{2}} \frac{(-1)^{j+1}\left(x^j + x^{k-j}\right)}{j!(k-j)!}$$

and for $k = 3, 5, \cdots$

$$\varphi_k(x) = \frac{1}{k!} + \sum_{j=1}^{\frac{k-1}{2}} \frac{(-1)^{j+1}\left(x^j - x^{k-j}\right)}{j!(k-j)!}.$$

Hence, for $k = 2, 3, \cdots$

$$T^*(k) \stackrel{\mathrm{d}}{=} N(0, \sigma_k^2),$$

where

$$\sigma_k^2 = 2 \int_0^1 (1-y)\varphi_k(y) \int_0^y x\varphi_k(x)dxdy. \tag{15}$$

For example, $\sigma_2^2 = 4.0873 \times 10^{-2}, \sigma_3^2 = 5.9359 \times 10^{-3}$ and $\sigma_4^2 = 4.2594 \times 10^{-4}$.

Remark 2.1. The change points $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k < 1$ may be estimated by the estimates $0 < \widehat{\lambda}_1 < \widehat{\lambda}_2 < \cdots < \widehat{\lambda}_k < 1$, where

$$V_n(\widehat{\lambda}_1, \widehat{\lambda}_2, \cdots, \widehat{\lambda}_k) = \max_{\underline{s}} \; V_n(\underline{s}).$$

2.2 Tests based on 2-sample U-statistics

Let $\widetilde{U}_{i,j,n}$ be a 2-sample U-statistic based on a skew-symmetric kernel ϕ of degree (m, m) and the two samples $X_{[ns_{i-1}]+1}, \cdots, X_{[ns_i]}$ and $X_{[ns_j]+1}, \cdots X_{[ns_{j+1}]}$. The kernel ϕ is selected such that large values of $\widetilde{U}_{i,j,n}$ are significant. Define $d_{i,n} = [ns_i] - [ns_{i-1}]$,

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

and

$$W_n(\underline{s}) = \sum_{i=1}^{k} \sum_{j=i}^{k} U_{i,j,n}.$$

Assume that $H_\circ$ *of* (1) *holds.* Define

$$\phi_1(y) = E\phi(y, X_2, \cdots, X_m; X_{m+1}, \cdots, X_{2m}). \tag{16}$$

and

$$\sigma^2 = E\phi_1^2(X_1).$$

Assume that

$$E\phi^2(X_1, X_2, \cdots, X_m; X_{m+1}, \cdots, X_{2m}) < \infty. \tag{17}$$

Based on $\{W_n(\underline{s})\}$ we propose the test statistics

$$T_{n2}(k) := \frac{\sqrt{n}}{\sigma} \max_{\underline{s}} W_n(\underline{s}), \tag{18}$$

and

$$T^*_{n2}(k) := \frac{\sqrt{n}}{\sigma} \int \cdots \int_{\underline{s}} W_n(\underline{s}) d\underline{s}. \tag{19}$$

The asymptotic distributions of T_{n2} and T^*_{n2} follow from the following Theorem.

Theorem 3.1. *Let $\Psi(\underline{s})$ be as in (11). Assume that $H_\circ$ of* (1) *and condition (17) hold. Then, as $n \to \infty$,*

$$\frac{\sqrt{n}}{\sigma} W_n(\underline{s}) \xrightarrow{\text{d}} \Psi(\underline{s}). \tag{20}$$

Corollary 3.1. By the Continuous Mapping Theorem,

$$T_{n2}(k) \xrightarrow{\text{d}} \sup_{\underline{s}} \Psi(\underline{s}) = T(k) \tag{21}$$

and

$$T^*_{n2}(k) \xrightarrow{\text{d}} \int \cdots \int_{\underline{s}} \Psi(\underline{s}) d\underline{s} = T^*(k), \tag{22}$$

where $T(k)$ and $T^*(k)$ are, respectively, as in (12) and (13).

Example 1. (Alternatives expressed in terms of hazard rate ordering)
Assume that we wish to test $H_\circ$ *of* (1) against H_1 of (3) when the ordering $\overset{\text{P}}{\prec}$ is the hazard rate ordering $\overset{\text{hr}}{\prec}$, where $F \overset{\text{hr}}{\prec} G$ if and only if, $r_F(.) \geq r_G(.)$. Here we use the kernel of Kochar (1979):

$$\phi(x_1, x_2; y_1, y_2) = \begin{cases} -1 & \text{if } xxyy \text{ or } yxxy \\ 0 & \text{if } xyxy \text{ or } yxyx \\ 1 & \text{if } yyxx \text{ or } xyyx, \end{cases}$$

where, for example, $yyxx$ represents $y_1 \leq y_2 \leq x_1 \leq x_2$, $y_2 \leq y_1 \leq x_2 \leq x_1$, $y_2 \leq y_1 \leq x_1 \leq x_2$ or $y_1 \leq y_2 \leq x_2 \leq x_1$. Under $H_\circ$ and condition (17) we obtain the results of Theorem 3.1 with $\sigma^2 = \frac{8}{105}$.

2.3 Tests based on L-Statistics

Let $J(\cdot)$ be a score (weight) function such that $\int_0^1 J(u)du = 0$. Define

$$Z_i = -\int_0^1 \{I(F(X_i) \leq u) - u\} J(u) dF^{-1}(u), \tag{23}$$

$$D_{i:n} = X_{i+1:n} - X_{i:n}, \qquad i = 1, 2, \cdots, n-1,$$

and

$$\widehat{\gamma}^2 = n^{-2} \sum_{i=1}^{n-1} \sum_{j=1}^{i} j(n-i) a_{ij} D_{i:n} D_{j:n}, \tag{24}$$

where

$$a_{ij} = \begin{cases} J(\frac{i}{n+1}) J(\frac{j}{n+1}), & \text{if } i = j \\ 2J(\frac{i}{n+1}) J(\frac{j}{n+1}), & \text{if } i \neq j \end{cases} .$$

Note that under $H_\circ$, $\widehat{\gamma}^2$ is a consistent estimator of $\gamma^2 = Var(Z_i)$. Let L_i be the L-Statistic based on the observations $X_{[ns_{i-1}]+1}, \cdots, X_{[ns_i]}$ and the score function $J(\cdot)$. Define

$$L_{i,j,n} = U_{i,j,n} = \frac{d_{i,n} d_{j+1,n}}{n^2} \left(L_{j+1} - L_i \right))$$

and

$$\eta_n(\underline{s}) = \sum_{i=1}^{k} \sum_{j=i}^{k} L_{i,j,n}.$$

Based on $\{\eta_n(\underline{s})\}$ we propose the test statistics

$$T_{n3}(k) := \frac{\sqrt{n}}{\widehat{\gamma}} \max_{\underline{s}} \eta_n(\underline{s}), \tag{25}$$

and

$$T^*_{n3}(k) := \frac{\sqrt{n}}{\widehat{\gamma}} \int \cdots \int_{\underline{s}} \eta_n(\underline{s}) d\underline{s}. \tag{26}$$

The asymptotic distributions of T_{n3} and T^*_{n3} follow from the following Theorem.

Theorem 4.1. *Let $\Psi(\underline{s})$ be as in (11). Assume that $H_\circ$ of* (1) *holds true and $E\left\{|Z_1|^2\right\} < \infty$. Then, as $n \to \infty$,*

$$\frac{\sqrt{n}}{\gamma} \eta_n(\underline{s}) \xrightarrow{\text{d}} \Psi(\underline{s}). \tag{27}$$

Corollary 4.1. By the Continuous Mapping Theorem and Slutsky Theorem

$$T_{n3}(k) \xrightarrow{\text{d}} \sup_{\underline{s}} \Psi(\underline{s}) = T(k) \tag{28}$$

and

$$T^*_{n3}(k) \xrightarrow{\text{d}} \int \cdots \int_{\underline{s}} \Psi(\underline{s}) d\underline{s} = T^*(k), \tag{29}$$

where $T(k)$ and $T^*(k)$ are, respectively, as in (12) and (13)

Example 2. (Alternatives expressed in terms of dispersive ordering) Assume that we wish to test $H_\circ$ *of* (1) against H_1 of (3) when the ordering $\overset{\mathrm{P}}{\prec}$ is the dispersive ordering $\overset{\mathrm{disp}}{\prec}$, where $F \overset{\mathrm{disp}}{\prec} G$ if and only if, $G^{-1}(F(x)) - x$ is nondecreasing in x, $0 < F(x) < 1$. Here we use the tests of Aly (2004) in which $J(u) = 2u - 1$.

3 Testing against general multiple alternatives

In this section we consider the problem of testing $H_\circ$ of (1) against $H_{11}$of (2).

3.1 Rank tests

Let $h(\cdot)$ be a score function which is real-valued and differentiable on $(0, 1)$ and let h' be its derivative. Assume that $\mu = \int_0^1 h(t)dt$ and $\sigma^2 = 2\int_0^1 \int_0^y h'(x)h'(y)x(1-y)dxdy$. Let r_i be the rank of X_i among the $X's$ and define $R_j = \sum_{i=1}^{j} h(\frac{r_i}{n}), j = 1, 2, \cdots, n$. The m tests of Lombard (1987) are given by

$$m_n(k) = n^{-k} \int \cdots \int_{\underline{s}} \xi_n(\underline{s}) d\underline{s},$$

where

$$\xi_n(\underline{s}) = n^{-1}\sigma^{-2} \sum_{j=1}^{k+1} \left(R^*_{[ns_j]} - R^*_{[ns_{j-1}]} \right)^2$$

and $R^*_l = R_l - l\mu$. Lombard (1987) proved that under $H_\circ$ *of* (1)

$$\xi_n(\underline{s}) \xrightarrow{\mathrm{D}} \xi(\underline{s}) = \sum_{j=1}^{k+1} \left(B(s_j) - B(s_{j-1})\right)^2 . \tag{30}$$

By (30) we have

$$m_n(k) \xrightarrow{\mathrm{D}} m(k),$$

where

$$m(k) = \int \cdots \int_{\underline{s}} \xi(\underline{s}) d\underline{s} = \sum_{j=1}^{k+1} \int \cdots \int_{\underline{s}} \left(B(s_j) - B(s_{j-1})\right)^2 d\underline{s}.$$

Lombard (1987) gave asymptotic critical values for $m(2)$ and $m(3)$.

Aly and Buhamra (1996) proposed the test statistics

$$t_n(k) = \int \cdots \int_{\underline{s}} \zeta_n(\underline{s}) d\underline{s},$$

where

$$\zeta_n(\underline{s}) = n^{-k-1}\sigma^{-2}\left(\prod_{i=1}^{k+1} d_{i,n}\right)\left\{\sum_{j=1}^{k+1}\frac{\left(R^*_{[ns_j]} - R^*_{[ns_{j-1}]}\right)^2}{d_{j,n}} - n\mu^2\right\}.$$

Aly and Buhamra (1996) argued that under $H_\circ$ *of* (1),

$$\zeta_n(\underline{s}) \xrightarrow{\mathrm{D}} \zeta(\underline{s}),$$

where

$$\zeta(\underline{s}) = \left(\prod_{i=1}^{k+1}(s_i - s_{i-1})\right)\sum_{j=1}^{k+1}\frac{(B(s_j) - B(s_{j-1}))^2}{s_j - s_{j-1}}. \tag{31}$$

Consequently,

$$t_n(k) \xrightarrow{\mathrm{D}} t(k),$$

where

$$t(k) = \int \cdots \int_{\underline{s}} \left(\prod_{i=1}^{k+1}(s_i - s_{i-1})\right)\sum_{j=1}^{k+1}\frac{(B(s_j) - B(s_{j-1}))^2}{s_j - s_{j-1}} d\underline{s}. \tag{32}$$

Aly and Buhamra (1996) obtained asymptotic critical values for $t(2)$ by simulation.

3.2 Likelihood ratio tests

Let $f(\cdot;\theta_i)$ be the density (or probability density) function of X_i and let $g(\cdot;\theta) = \ln f(\cdot;\theta)$. The maximum likelihood estimators $\widehat{\theta}_{i,j}$ satisfy the equations

$$\sum_{l=i}^{j} g'(X_l; \widehat{\theta}_{i,j}) = 0.$$

The likelihood tests of Aly and Bouzar (1994) are given by

$$S_n(k) = \int \cdots \int_{\underline{s}} \lambda_n(\underline{s}) d\underline{s},$$

where

$$\lambda_n(\underline{s}) = 2n^{-k}\left(\prod_{i=1}^{k+1} d_{i,n}\right)\left\{\ln \prod_{j=1}^{k+1} A_{[ns_{j-1}]+1,[ns_j]} - \ln A_{1,n}\right\},$$

where

$$A_{i,j} = \prod_{r=i}^{j} f(X_r; \widehat{\theta}i, j).$$

Aly and Bouzar (1993) proved that

$$\lambda_n(\underline{s}) \xrightarrow{\text{D}} \zeta(\underline{s}),$$

where $\zeta(\underline{s})$ is as in (31). Consequently,

$$S_n(k) \xrightarrow{\text{D}} t(k),$$

where $t(k)$is as in (32).

References

[1] Aly, E.-E. (2004). Change point detection based on L-statistics. *Fields Institute Communications.*, **44**, 293-300.

[2] Aly. E.-E., Abd-Rabou, A. S. and Al-Kandari, N. M. (2003). Tests for multiple change points under ordered alternatives. *Metrika*, **57**, 209-221.

[3] Aly, E.-E. and Bouzar, N. (1993). On maximunm likelihood ratio tests for the changepoint problem. In: *Proceedings of the Theme-term Changepoint Analysis - Empirical Reliability* pp. 1-11. Carleton University, Ottawa, Canada.

[4] Aly, E.-E. and BuHamra S. (1996). Rank tests for two change points. *Comput. Statist.& Data Analy.,* **22**, 363-372.

[5] Kochar, S.C. (1979). Distribution-free comparison of two probability distributions with reference to their hazard rates. *Biometrika*, **66**, 437-441.

[6] Lombard, F. (1987). Rank tests for changepoint problems. *Biometrika*, **74**, 615-624.