Change-point tests based on Integrated Empirical Distribution Functions

Abd-Elnasser S. A. Rabou
Statistics Department, Faculty of Economics and Political Science
Cairo University. Cairo, Egypt.

Abstract

Three new nonparametric change-point test statistics, based on the integrated empirical function (IEF) are introduced. Properties and distributions of these test statistics are also discussed. We derived the limiting distributions of the tests and conducted a Monte Carlo study to compare the powers of the new tests with their non-integrated counterparts.

Keywords and Phrases: Change-point, Integrated empirical function, asymptotic theory, Strong approximations, Gaussian processes, Monte Carlo.

1. Introduction

Testing a change in the stochastic characteristics of given observations with a probabilistic model is a crucial aspect of data analysis. In practice, shift problems in areas such as , quality control, medical monitoring, econometrics, biological growth, signal processing and image processing are a few examples of the theoretical change-point problem. In this respect, much work has been done when the data are assumed to come from continuous distribution (see, e.g. Csörgő and Horváth (1993, 1997)). Many Parametric as well as nonparametric test statistics has been introduced for the detection of a suspected shift(s) in the distribution function of a sequence of observations. Empirical distribution functions, empirical quantile functions, ranks, likelihood ratios and other tools are used to construct test statistics for this change-point problem. Researches and papers covered this problem for more than five decades stretching from classical to non-classical approaches. They discussed many aspects of this problem, such as test statistics construction, change-point estimation, properties studying and the development of the statistics distribution theory. For more details we refer the reader to, Shaban (1980), Basseville and Benveniste (1986), Brodsky and Darkhovsky (1993), and Csörgő and Horváth (1993, 1997).

Let $X^{(n)}=(X_1,X_2,...,X_n)$ be a random vector with independent components. We consider the following problem of testing a hypothesis H_0 (the absence of a change-point): The independent random variables $X_1,X_2,...,X_n$ are identically continuously distributed with a distribution function $F(.)$. The alternative hypothesis H_1 (the presence of a change-point); $X_1,X_2,...,X_m$ have a common continuous distribution function $F(.)$ and $X_{m+1},X_2,...,X_n$ have a different continuous distribution function $G(.)$. The change-point $m=[nt], 0 \le t \le 1$ and the distribution functions $F(.)$ and $G(.)$ are all assumed unknown.

Empirical functions and processes are frequently used to test the null hypothesis of no change in the distribution function of a sequence of observations. Test statistics based on these functions and processes such as Kolomogrov-Simernov, Cramér von Mesis, Anderson-Darling and Erdós-Darling type statistics are some examples. Introduction of these type of tests, their stochastic theory and performance are presented mainly in Csörgő et. al. (1986), Shorack and Wellner (1986), and Csörgő and Horváth (1988, 1993).

In this paper we introduce new test statistics based on the integrated empirical distribution function (IEDF) to detect a possible change in the distribution function. Since the distribution of any random variable X with a distribution function, F(.), is uniquely determined by its integrated distribution function, (see Klar (2001)), defined by $\tilde{F}(s)=\int_{-\infty}^{s} F(x)dx,\ -\infty<s<\infty$. Then, for hypotheses concerning the distribution function F(.), we may construct test statistics based on the empirical counterpart of $\tilde{F}(s)$, namely $\tilde{F}_k(s)=\int_{-\infty}^{s} F_k(x)dx,\ -\infty<s<\infty$, k=1,2,...,n, where $F_k(.)$ is the empirical distribution function based on k sample observations. Henze and Nikitin (2000 & 2003), Gürtler and Henze (2000) and Klar (2001), proposed and studied several goodness-of-fit and two-sample test statistics based on the so called integrated empirical distribution function. In fact these types of integrated empirical test statistics have proved to be serious competitors to classical tests in case of goodness-of-fit tests, (see Klar (2001)).

Section 2, presents the problem and the new test statistics. In section 3, we derive the asymptotic distribution theory of the proposed tests. We conduct a Monte Carlo study to estimate the critical values of the proposed tests and their powers in section 4. Finally, we present some proofs for the main results in section 5.

2. Processes and Test Statistics

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent continuous random variables. Consider the change-point testing problem;

$H_0: X_1, X_2, \ldots, X_n$ have a distribution function F(.),

against,

$H_1: X_1, X_2, \ldots, X_m$ have a distribution function F(.) and

$X_{m+1}, X_{m+2}, \ldots, X_n$ have a distribution function G(.), (2.1) where the change-point $m = [nt], 0 \le t \le 1$ and the distribution functions $F(.)$ and $G(.)$ are all assumed unknown. Define the empirical distribution function based on k observations by;

$$F_k(x) = \frac{1}{k}\sum_{i=1}^{k} I\{X_i \le x\}\ ,\ \text{k} = 1,2,\ldots,\text{n} \tag{2.2}$$

and the integrated empirical distribution function based on k observations by;

$$\bar{F}_k(s) = \int_{-\infty}^{s} F_k(x)dx,\quad -\infty < s < \infty,\ \text{k}=1,2,\ldots,\text{n}. \tag{2.3}$$

Henze and Nikitin (2000), presented the following properties for IEDF in (2.3), which will be useful in introducing our processes and tests.

Theorem A (Henze and Nikitin (2000))

The integrated empirical distribution function $\bar{F}_n$ has the following properties:

1. $\bar{F}_n(x) = \frac{1}{2}(F_n^2(x) + \frac{1}{n}F_n(x))$.

2. $\bar{F}_n(x) = \frac{1}{n^2}\sum_{i,j=1}^{n} h(x_i, x_j; x)$,

where

$$h(u,v;x) = \frac{1}{2}(I\{u \le v \le x\} + I\{v \le u \le x\}),$$

that is $\bar{F}_n(x)$ *is a* V *-statistic for any fixed* x.

3. $\bar{F}_n(x) = \frac{1}{n^2}(k + \binom{k}{2}),\ x_{(k)} \le x < x_{(k+1)}, k = 0,1,\ldots,n,$

where $x_{(i)}$, *is the* i^{th} *order statistic and* $x_{(0)} = -\infty, x_{(n+1)} = \infty$.

4. $\bar{F}_n(x_{(k)}) - \bar{F}_n(x_{(k-1)} - 0) = \frac{k}{n^2},\ k = 1,2,\ldots,n,$

i.e. $\bar{F}_n(.)$ *has a jump of hieght* $\frac{k}{n^2}$ *at the* k^{th} *order statistic.*

5. $\sup_{-\infty<x<\infty} |\bar{F}_n(x) - \frac{1}{2}F^2(x)| \to 0\quad a.s.\ \ as\ n \to \infty.$

For the change-point analysis here we introduce the integrated empirical change-point process, $\{\Delta_n(s,t); -\infty < s < \infty, 0 \le t \le 1\}$, given by

$$\Delta_n(s,t) = \frac{[nt]}{\sqrt{n}}(\bar{F}_{[nt]}(s) - \bar{F}_n(s)),\qquad -\infty < s < \infty, 0 \le t \le 1, \tag{2.4}$$

where $[\theta]$ is the integer value of θ and $\bar{F}(.)$ is defined in (2.3). We also define the Kiefer process $K(s,t), 0 \le s \le 1, -\infty < t < \infty$, which is a Gaussian process with mean zero and covariance structure,

$$E\,K(s_1,t_1)K(s_2,t_2) = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2), \qquad 0 \le s_1, s_2 \le 1, -\infty < t_1, t_2 < \infty \quad , \tag{2.5}$$

and the Gaussian process,

$$\Delta(s,t) = s\,\{K(s,t) - t\,K(s,1)\}, \qquad 0 \le s,t \le 1, \tag{2.6}$$

with zero mean and covariance structure,

$$E\,\Delta(s_1,t_1)\Delta(s_2,t_2) = s_1 s_2 (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2 - t_1 t_2), \qquad 0 \le s_1, s_2 \le 1, \; -\infty < t_1, t_2 < \infty.$$

Now we state our main result, which provide an ideal (Gaussain) approximation for the change-point process in (2.4).

Theorem (2.1)

Let $\Delta_n(.,.)$ and $\Delta(.,.)$ be the processes in (2.4) and (2.6) respectively. Then we have as $n \to \infty$

$$\sup_{0 \le s \le 1} \sup_{-\infty < t < \infty} |\Delta_n(s,t) - \Delta(F(s),t)| \overset{p}{=} o(1).$$

By the continuity of the distribution function F(.) and the integral transformation $U = F(x)$, the limiting distribution in Theorem (2.1) above eventually does not depend on the unknown distribution function F(.), (see Henze and Nikitin (2003)). Thus the suggested test procedures below will produce distribution free test statistics.

For the change-point hypotheses in (2.1), we suggest the following integrated-type test statistics:

$$T_1 = \sup_{0 \le s \le 1} \sup_{-\infty < t < \infty} |\Delta_n(s,t)|,$$

$$T_2 = \int_0^1 \int_{-\infty}^{\infty} \Delta_n(s,t)\, dF_n(s)\, dt \tag{2.7}$$

and

$$T_3 = \int_0^1 \int_{-\infty}^{\infty} \Delta_n^2(s,t)\, dF_n(s)\, dt.$$

By the kernel covariance of the process $\Delta(.,.)$ in (2.6), we can easily see that;

$$\Delta(s,t) \overset{D}{=} B_1(t)\, s\, B_2(s), \quad \forall t,s \in [0,1]^2, \tag{2.8}$$

where $B_1(.)$ and $B_2(.)$ are two independent Brownian bridges.

Corollary (2.1)

Let $B_1(.)$ and $B_2(.)$ are two independent Brwonian bridges, then we have, as $n \to \infty$,

$$T_1 \overset{D}{\to} \sup_{0\le t\le 1}|B_1(t)| \sup_{0\le s\le 1}|s\, B_2(s)| = D_1,$$

$$T_2 \overset{D}{\to} \int_0^1 B_1(t)dt \int_0^1 s\, B_2(s)ds = D_2,$$

and

$$T_3 \overset{D}{\to} \int_0^1 B_1^2(t)dt \int_0^1 s^2 B_2^2(s)ds = D_3.$$

The Corollary is a direct result of Theorem (2.1), (2.8) and the continuous mapping Theorem. It is clear that each of the limiting random variables (rv's) D_1, D_2 and D_3 is a multiplication of two independent random variables. Some of these independent rv's in each of the multiplications has a known distribution or a tabulated critical values. But the distributions of D_1, D_2 and D_3 are not known in the literature, at least to the author best knowledge. Henze and Nikitin (2003), pointed out that the distributions of $\sup_{0\le s\le 1}|s\, B(s)|$ and of $\int_0^1 s^2 B^2(s)ds$ are unknown. They studied the second, $\int_0^1 s^2 B^2(s)ds$, and presented some numerical results for its distribution. The limiting rv A_3 is a multiplication of two independent centered Normal rv's with variances $\frac{1}{12}$ and $\frac{1}{45}$ respectively. Since the exact distributions of the limiting random variables of Corollary (2.1) above are unknown, we will simulate the test statistics critical values. We will also conduct a power comparisons between the proposed tests and their non-integrated competitors through a Monte Carlo study.

4. Estimated critical values and Powers

The application of the integrated-type tests in (2.7) depends on the availability of the critical points of the distributions in Corollary (2.1). But, because these asymptotic distributions are unknown in the literature their critical values are not available. For this reason we conducted a Monte Carlo study to calculate the sample critical values of the proposed tests. For each size n=10, 20, ..., 100 we generated 5,000 random samples under the null hypothesis model of no change from Normal, Exponential and Cauchy distributions. In each sample we computed the test statistics $T_1, T_2,$ *and* T_3, then ordered the 5,000 computed values and obtained the $(1-\alpha)^{th}$

percentiles for $\alpha = 0.1, 0.05, 0.01$ for each test. Tables (1), (2) and (3), below list the 90th, 95th and 99th estimated percentiles for different sample sizes and different distributions. All calculations in this section were done by a Matlab version 5.1 programs and its associated subroutines. To see how close the estimated critical values of each test statistic in different distributions, we graphed them at $\alpha = 0.05$. Looking at the critical values graph of T_1, we notice that there is a little discrepancies when the sample size is small but the values tends to agree in larger samples. In the second and third test critical values graph we see no difference approximately in all distributions in all samples. Hence we can say that these tests are really distribution free tests, especially in large samples. All test statistics critical values in all distributions started a bit higher when the sample sizes were small, then it decreased as the size increased. This show that the use of the test estimated critical values in small samples give slightly liberal tests: That is the tests in this case tend not to reject the null hypothesis in most cases. This is a logical result since the detection power of the change point tests suppose to increase with the increase of the sample size. We notice that the estimted critical values of each test converge approximately to the same values in each significance level in all distributions as the sample size increases. Finally, we can see that the variations of the first and third test critical values between samples are less than those of the second test. This is due to the fact that the former two take only positive values while the latter takes also negative values.

Table (1)

Finite sample critical values
(Normal case)

Test	T_1			T_2			T_3		
α	1%	5%	10%	1%	5%	10%	1%	5%	10%
n									
10	1.3282	1.1511	1.0499	0.3672	0.2991	0.2549	0.2414	0.1767	0.1392
20	1.3148	1.1180	1.0118	0.3119	0.2426	0.2042	0.1843	0.1238	0.0969
30	1.3121	1.0954	0.9940	0.2891	0.2257	0.1847	0.1646	0.1085	0.0841
40	1.2934	1.1028	1.0056	0.2786	0.2130	0.1773	0.1536	0.1008	0.0787
50	1.2796	1.0909	0.9809	0.2698	0.2018	0.1685	0.1451	0.0923	0.0721
60	1.2883	1.0754	0.9836	0.2666	0.1987	0.1622	0.1423	0.0909	0.0701
70	1.2671	1.0753	0.9787	0.2611	0.1906	0.1546	0.1363	0.0877	0.0674
80	1.2507	1.0694	0.9843	0.2540	0.1857	0.1540	0.1273	0.0835	0.0667
90	1.2782	1.0822	0.9803	0.2457	0.1866	0.1518	0.1272	0.0849	0.0661
100	1.2715	1.0803	0.9863	0.2520	0.1823	0.1492	0.1335	0.0821	0.0653

Table (2)

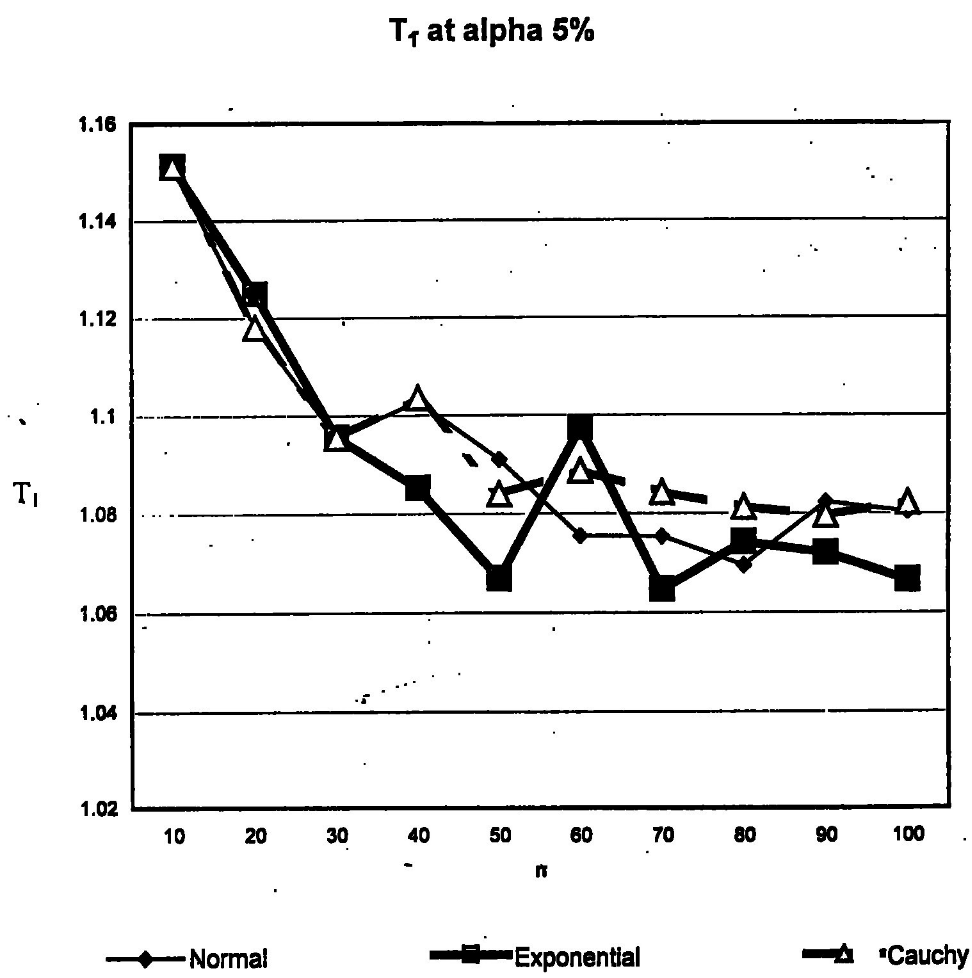
Finite sample critical values
(Exponential case)

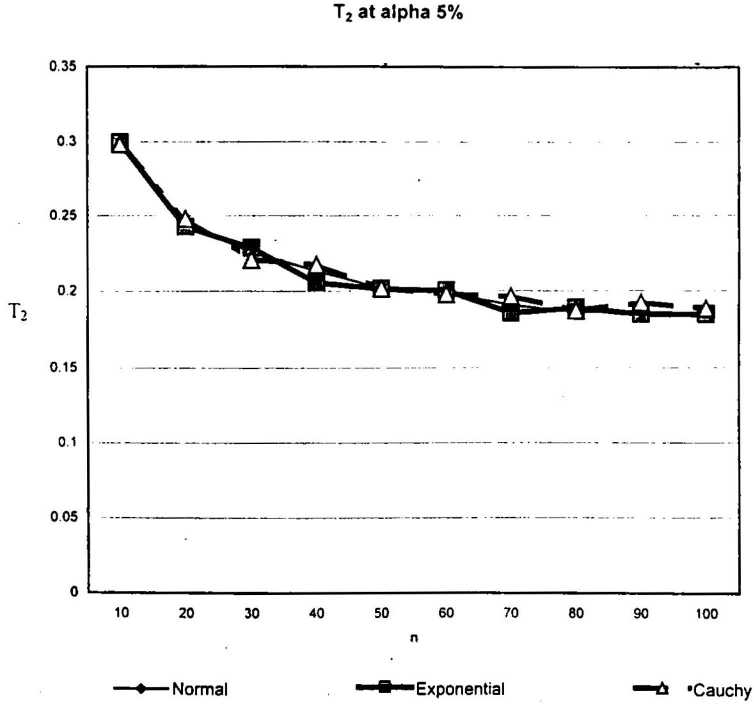
Test	T_1			T_2			T_3		
α	1%	5%	10%	1%	5%	10%	1%	5%	10%
n									
10	1.3282	1.1511	1.0499	0.3648	0.2992	0.2571	0.2375	0.1739	0.1414
20	1.3148	1.1247	1.0286	0.3152	0.2422	0.2047	0.1854	0.1220	0.0982
30	1.3065	1.0954	1.0175	0.2893	0.2282	0.1924	0.1622	0.1105	0.0855
40	1.2997	1.0851	0.9843	0.2740	0.2053	0.1701	0.1480	0.0943	0.0751
50	1.2694	1.0668	0.9777	0.2663	0.2013	0.1655	0.1407	0.0911	0.0715
60	1.2852	1.0973	0.9901	0.2615	0.2000	0.1621	0.1385	0.0906	0.0710
70	1.2710	1.0646	0.9689	0.2549	0.1853	0.1509	0.1305	0.0840	0.0650
80	1.2608	1.07411	0.9804	0.2524	0.1885	0.1552	0.1260	0.0840	0.0679
90	1.2509	1.0719	0.9644	0.2434	0.1844	0.1514	0.1239	0.0819	0.0645
100	1.2726	1.0664	0.9785	0.2477	0.1843	0.1489	0.1266	0.0832	0.0648

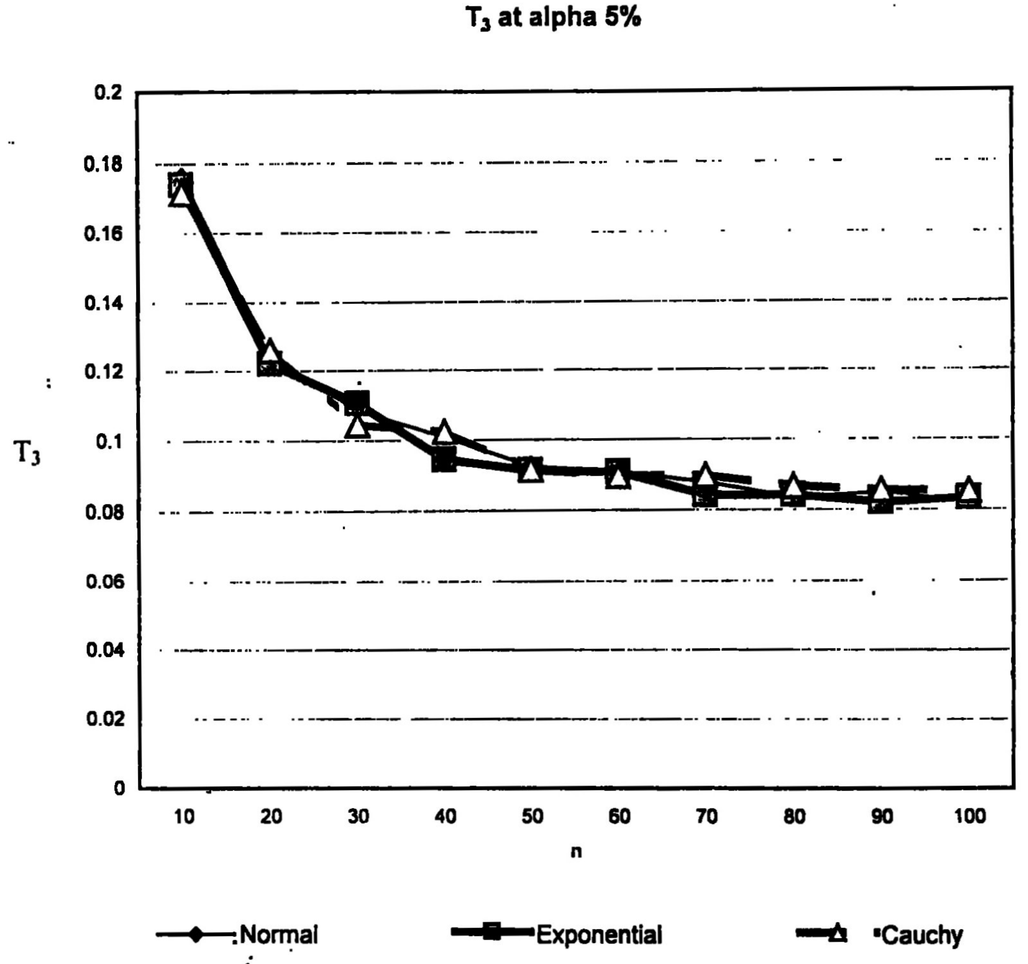
Table (3)

Finite sample critical values
(Cauchy case)

Test	T_1			T_2			T_3		
α	1%	5%	10%	1%	5%	10%	1%	5%	10%
n									
10	1.2902	1.1511	1.0499	0.3718	0.2981	0.2565	0.2418	0.1715	0.1405
20	1.3204	1.1180	1.0178	0.3237	0.2478	0.2077	0.1926	0.1259	0.0989
30	1.2938	1.0954	0.9883	0.2954	0.2206	0.1844	0.1637	0.1044	0.0817
40	1.3260	1.1036	0.9985	0.2857	0.2170	0.1778	0.1584	0.1023	0.0774
50	1.2851	1.0839	0.9860	0.2712	0.2016	0.1678	0.1438	0.0914	0.0725
60	1.2783	1.0886	0.9836	0.2551	0.1978	0.1642	0.1326	0.0895	0.0707
70	1.2840	1.0842	0.9882	0.2620	0.1962	0.1629	0.1354	0.0896	0.0698
80	1.2907	1.0812	0.9874	0.2569	0.1870	0.1532	0.1349	0.0866	0.0658
90	1.2800	1.0795	0.9859	0.2573	0.1920	0.1574	0.1365	0.0855	0.0674
100	1.2623	1.0821	0.9826	0.2595	0.1881	0.1514	0.1314	0.0852	0.0660


T1 at alpha 5%
1.16
1.14
1.12
1.1
1.08
1.06
1.04
1.02
T1
10
20
30
40
50
60
70
80
90
100
n
Normal
Exponential
Cauchy


T2 at alpha 5%
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
T2
10
20
30
40
50
60
70
80
90
100
n
Normal
Exponential
Cauchy


T_3 at alpha 5%
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
T_3
10
20
30
40
50
60
70
80
90
100
n
Normal
Exponential
Cauchy

To measure the performance of our integrated-type tests we compared their estimated powers with those of the non-integrated counterparts. The non-integrated test statistics can be easily obtained by replacing the integrated empirical distribution functions $\bar{F}(.)'s$ of (2.3) by the classical empirical distribution functions $F(.)'s$ of (2.2) in (2.4). Let $N_1, N_2,$ *and* N_3 denote the non-integrated-type test statistics. We obtained Monte Carlo powers for all the six tests. The powers were calculated for sample size n = 20, taken from Normal, Exponential and Cauchy distributions. We consider the change points m = [nr], $r \in \{0.25, 0.50, 0.75\}$. Four cases were considered for the location shift size δ at the change position m. The shift sizes were computed as the solution of the equation $P(X_{m+1} > X_m) = p$ and $p \in \{0.6, 0.7, 0.8, 0.9\}$. To calculate the powers, we simulated 5,000 realizations of samples of size n = 20 under the alternative distribution and computed the six tests in each realization. Then for each test we obtained the fraction of the number of times that the null hypothesis is rejected. The results of this power study are reported in Tables (4), (5) and (6). Other values of this design parameters, e.g. n = 40 and p = 0.85, were also considered, but these yielded the same qualitative conclusions and hence are not reported .

The main conclusions that can be drawn from the power study, reported in Tables (4), (5) and (6), below are as follows:

1. As expected the estimated powers of all tests seem to increase with the increase of p, the possibility of a change.
2. The estimated powers are always the largest whenever the change position is assumed in the middle of the sample. This is because a change in the middle of the sample gives enough observations on both sides of the change position to show the difference.
3. Compared with the non-integrated tests, T_1, T_2, and T_3 are very competitive against the distributional change regardless the unknown parent model.
4. In case of symmetric distributions, like Normal and Cauchy, the integrated-type tests have a general superiority over the non-integrated ones. But in case of skewed distributions such as Exponential, the second non-integrated-type test N_2, has a slight edge over the proposed integrated-type tests.
5. We also notice that N_2, generally has larger powers among his non-integrated-type fellow tests.

Table (4)

Percentages of 5,000 samples declared significant
n=20, Normal case

Probability Of change	Change position	T_1	T_2	T_3	N_1	N_2	N_3
P=0.6	5	12	13	12	6	12	7
	10	15	16	16	7	14	8
	15	12	13	12	6	12	7
P=0.7	5	21	25	25	12	24	13
	10	38	41	40	24	39	26
	15	23	28	26	13	26	14
P=0.8	5	41	46	46	26	45	30
	10	67	70	70	54	69	56
	15	41	48	46	27	44	30
P=0.9	5	67	69	70	54	70	57
	10	91	93	93	86	93	87
	15	72	76	76	53	70	57

Table (5)

Percentages of 5,000 samples declared significant
n=20, Exponential case

Probability Of change	Change position	T_1	T_2	T_3	N_1	N_2	N_3
P=0.6	5	9	12	11	8	13	8
	10	11	14	14	11	17	10
	15	8	11	10	8	11	7
P=0.7	5	15	23	22	15	26	17
	10	23	33	31	30	39	29
	15	15	22	20	16	24	15
P=0.8	5	28	41	40	32	46	34
	10	48	60	59	60	67	61
	15	30	41	39	32	42	30
P=0.9	5	55	70	68	61	70	65
	10	84	89	88	88	91	89
	15	72	79	76	75	79	65

Table (6)

Percentages of 5,000 samples declared significant
n=20, Cauchy case

Probability Of change	Change position	T_1	T_2	T_3	N_1	N_2	N_3
P=0.6	5	13	13	14	10	13	8
	10	19	16	17	14	16	11
	15	11	12	12	9	11	7
P=0.7	5	27	26	28	21	24	16
	10	43	38	42	38	37	31
	15	28	26	27	22	24	17
P=0.8	5	49	43	50	43	42	35
	10	74	71	72	71	65	65
	15	56	48	51	44	43	36
P=0.9	5	76	77	76	77	66	65
	10	96	96	94	96	91	93
	15	88	76	81	76	67	65

5. Proofs

For the integrated process in (2.4), we have

$$\Delta_n(s,\frac{m}{n})=\frac{m}{\sqrt{n}}(\bar{F}_m(s)-\bar{F}_n(s)),\quad 1\le m\le n,\ -\infty<t<\infty$$

$$=\frac{m}{\sqrt{n}}\{(\bar{F}_m(s)-\bar{F}(s))-(\bar{F}_n(s)-\bar{F}(s))\}$$

$$=\frac{m}{\sqrt{n}}\frac{1}{2}\{(F_m^2(s)-F^2(s))-(F_n^2(s)-F^2(s))+\frac{1}{m}F_m(s)-\frac{1}{n}F_n(s)\}$$

$$=\frac{m}{2\sqrt{n}}\{(F_m(s)+F(s))(F_m(s)-F(s))-(F_n(s)+F(s))(F_n(s)-F(s))$$

$$+\frac{1}{m}F_m(s)-\frac{1}{n}F_n(s)\}$$

$$=\sqrt{\frac{m}{n}}\frac{1}{2}(F_m(s)+F(s))\alpha_m(s)-\frac{m}{n}\frac{1}{2}(F_n(s)+F(s))\alpha_n(s)$$

$$+\frac{1}{2\sqrt{n}}F_m(s)-\frac{m}{n}\frac{1}{2\sqrt{n}}F_n(s),$$

(5.1)

where $\alpha_k(.)$ is the empirical process based on k sample observations.

By the almost sure convergence of the empirical distribution function, the last two terms of (5.1) are $o(1)$ a.s..

Let $K(.,.)$ be the Kiefer process of (2.5), then for the first term in (5.1), we get as $n \to \infty$,

$$A_1 = \max_{1\le m\le n} \sup_{-\infty<s<\infty} |\sqrt{\frac{m}{n}}\frac{1}{2}(F_m(s)+F(s))\alpha_m(s) - \frac{1}{\sqrt{n}}F(s)K(F(s),m)|$$

$$\le \max_{1\le m\le n} \sup_{-\infty<s<\infty} |\frac{1}{2}(F_m(s)+F(s))\alpha_m(s) - F(s)\frac{1}{\sqrt{m}}K(F(s),m)|$$

$$\le \max_{1\le m\le n} \sup_{-\infty<s<\infty} |\frac{1}{2}(F_m(s)+F(s))\alpha_m(s) - \frac{1}{2}(F(s)+F(s))\alpha_m(s)|$$

$$+ \max_{1\le m\le n} \sup_{-\infty<s<\infty} |F(s))\alpha_m(s) - F(s)\frac{1}{\sqrt{m}}K(F(s),m)|$$

$$\le \frac{1}{2}\max_{1\le m\le n} \sup_{-\infty<s<\infty} |\alpha_m(s)| \max_{1\le m\le n} \sup_{-\infty<s<\infty} |(F_m(s)-F(s))|$$

$$+ \max_{1\le m\le n} \sup_{-\infty<s<\infty} |F(s))\alpha_m(s) - F(s)\frac{1}{\sqrt{m}}K(F(s),m)|$$

$$\le O_{a.s.}(1)\ o_{a.s.}(1) + o_p(1),$$

(5.2)

where the last results follow form the almost sure convergence of the empirical process and empirical distribution function (see Csörgő and Horváth (1986)) and Komlós et. al (1975, 1976) respectively.

Similarly, we get for the second term in (5.1), as $n \to \infty$,

$$A_2 = \max_{1\le m\le n} \sup_{-\infty<s<\infty} |\frac{m}{n}\frac{1}{2}(F_n(s)+F(s))\ \alpha_n(s) - \frac{m}{n}F(s)\frac{1}{\sqrt{n}}K(F(s),n)|$$

$$= o_p(1).$$

(5.3)

Combining (5.1), (5.2) and (5.3), we complete the proof of Theorem (2.1).

Acknowledgements

I would like to express my gratitude to Prof. Ahmed Gad and Eng. Magdy Selim for their computer support. Many thanks goes to my mother-in-law for her morale help and prayers. Besides, the auther also thanks several anonymous refrees and editors for their constructive reviews and comments.

References

Basseville, M. and Benveniste, A. (1986): *Detection of abrupt changes in signals and dynamical systems*. New York. Basel.

Brodsky, B. and Darkhovsky, B. (1993): *Nonparametric methods in Change Point Problems*. Kluwer, Dodrecht.

Csörgő , M. and Horváth, L. (1988): Invariance principles for change point problems. *J. Mult. Anal.* 27, 151-168.

Csörgő, M. and Horváth, L. (1993): *Weighted Approximations in Probability and Statistics*. Wiley, Chichester.

Csörgő, M. and Horváth, L. (1997): *Limit Theorems in change-point analysis*. Wiley, New York.

Csörgő , M., Csörgő , S., Horváth, L. and Mason, D. M. (1986): Weighted empirical and quantile processes. *Ann. Probab.* 14, 31-85.

Gürtler, N. and Henze, N. (2000): Recent and classical goodness-of-fit tests for Poisson distribution. *J. Statist. Plan. Inf.*, 90, 207-225.

Henze, N. and Nikitin, Y., Y. (2000): A new approach to goodness-of fit testing based on the integrated empirical process. *J. Nonp. Statist.* 12, 391-416.

Henze, N. and Nikitin, Y., Y. (2003): Two-sample tests based on the integrated empirical process. *Comm. Statist. The. Meth.* 32, 1767-1788.

Klar, B. (2001): Goodness-of-fit tests for the exponential and the normal distribution based on the integrated distribution function. *Ann. Inst. Statist. Math.* 53, 338-353.

Komlós, J., Major, P. and Tusnády, G. (1975): An approximation to partial sums of independent R.V.'s and the sample DF. I. *Z. Wahrsch. Verw. Gebiete*. 32, 111-131.

Komlós, J., Major, P. and Tusnády, G. (1976): An approximation to partial sums of independent R.V.'s and the sample DF. II. *Z. Wahrsch. Verw. Gebiete*. 34, 33-58.

Shaban, S. A. (1980): Change point problem and two-phase regression, an annotated bibliography. *Inter. Statist. Rev.* 48, 83-93.

Shorack, G. and Wellner, J. (1986): *Empirical Processes with Applications to Statistics*. Wiley, New York.