The Korean Communications in Statistics
Vol. 13 No. 3, 2006, pp.607-619

Test and Estimation for Normal Mean Change[1)]

Jaehee Kim[2)] and Jong Eun Ryu[3)]

Abstract

We consider the problem of testing the existence of change in mean and estimating the change-point when the data are from the normal distribution. A change-point estimator using the likelihood ratio test statistic, Gombay and Horvath (1990) test statistic, and nonparametric change-point estimator using Carlstein (1988) empirical distribution are studied when there exists one change-point in the mean. A power study is done to compare the change test statistics. And a comparison study of change-point estimators for estimation capability is done via simulations with S-plus software.

Keywords : Brownian bridge; change-point; likelihood; mean change.

1. Introduction

Recently there has been more interest in the statistical analysis of change-point detection and estimation. It is mainly because change-point problems can be occurred in many disciplines such as economics, finance, medicine, psychology, geology, meteorology, environmental studies and etc. and even in daily lives.

In almost all classic statistical inference is based upon the assumption that there exists a fixed probabilistic mechanism of data generation. Unlike classic statistical inference, the parametric change analysis of data about the complex objects is considered. The existence of more than one data generation process is the most important characteristic of complex system.

When the hypotheses of statistical homogeneity holds true, that is, there exists only one mechanism of data generation, the law of large numbers are applied to make an inference. However if there exists change in the data generation, the

---

1) This research is supported by Korea Research Foundation R04-2004-000-10138-0.

2) Associate Professor, Department of Statistics, Duksung Women's University, Seoul 132-714, Korea.
Correspondence : jaehee@duksung.ac.kr

3) Researcher, Division of Epidemiology and Health Index Assistant Researcher, Center for Genome Science, National Institute of Health, Korea Center for Disease Control & Preven -tion(KCDC)

probabilistic law should be applied differently. In this case all data obtained should be sorted in subsamples generated by different probabilistic mechanisms. After this classification the correct inference can be made.

It is important to detect possible changes of data generation process and the appropriate statistical analysis of such data must begin with testing and decisions about possible change.

Changes happen in every field of the world. For example, the daily stock market records show that the stock price fluctuates. There are some shifts of mean price. One would want to find out the possible change and the change-point day and investigate the reason.

The quality of the products is expected to remain stable. However, for some reasons, the process might lose the control to produce the same quality. They would want to know the change-point where the quality of the products deterioration occurs.

We consider tests for the mean change and the change-point estimation in the normal distribution.

2. Univariate Normal Model

Let $X_1, X_2, \cdots, X_n$ be independent normal random variables with parameters $(\mu_1, \sigma^2), (\mu_2, \sigma^2), \cdots, (\mu_n, \sigma^2)$, respectively.

2.1 Mean Change

The mean change problem was first examined by Page (1955). Later Chernoff and Zacks (1964) studied the one change test with the Bayesian approach. Kander and Zacks (1968) extended to the problem to the one parameter exponential family of distributions. Bhattacharya and Johnson (1968) investigated a nonparametric approach to the problem of testing for a shift in the level of a process occurring at an unknown time point. Gardner (1969) considered the problem of detecting AMOC(at most one change) and the likelihood ratio for the normal random variables. Sen and Srivastava (1975) derived the test for change with the normal random variables to consider the nonparametric test.

Hinkley (1970) made an inference about the change-point problem. He examined the normal variables and derived the test and the asymptotic distributions of the likelihood ratio test statistic for testing the hypotheses about the change-point. Hawkins (1977) obtained the exact null and alternative distributions of likelihood ratio test statistic for the normal distribution.

Gombay and Horvath (1990) considered the maximum likelihood tests for change in the mean of independent random variables and proved the limit distribution as a double exponential distribution. James et al. (1987) considered testing a sequence of independent normal random variables and suggested the test statistic based on the likelihood ratio and the recursive residuals. Buckley (1991) suggested the cusum type test for the normal random variables to detect a smooth change signal.

The hypothesis of interest is defined as

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_n = \mu \text{ versus } H_1: \mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \cdots = \mu_n \tag{2.1}$$

where k is the unknown location of the single change-point. The testing procedure depends on whether the nuisance parameter σ^2 is known or unknown.

2.2 Test for Mean Change

When the variance is known, without loss of generality, assume that $\sigma^2 = 1$. The maximum likelihood ratio procedure test statistic is

$$\Lambda = \frac{L_0(\hat{\mu})}{L_1(\widehat{\mu_1}, \widehat{\mu_n})} \tag{2.2}$$

where the MLE's are

$$\hat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \widehat{\mu_1} = \overline{X}_k = \frac{1}{k}\sum_{i=1}^{k} X_i, \qquad \widehat{\mu_n} = \overline{X}_{n-k} = \frac{1}{n-k}\sum_{i=k+1}^{n} X_i$$

and

$$L_0(\mu) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n}(X_i - \mu)^2},$$

$$L_1(\mu_1, \mu_n) = \frac{1}{(\sqrt{2\pi})^n} \times exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{k}(X_i - \mu_1)^2 + \sum_{i=k+1}^{n}(X_i - \mu_n)^2\right]\right\}.$$

$-2\log\Lambda$ is considered. Therefore the likelihood procedure test statistic for testing H_0 against H_1 is

$$U^2 = \max_{1 \le k \le n} V_k. \tag{2.3}$$

LRT(likelihood ratio test) rejects H_0 if U is large. Hawkins (1977) derived the exact and asymptotic null distribution of the test $U = U_{LRT}$.

Let

$$S_k = \sum_{i=1}^{k}(X_i - \overline{X}_k)^2 + \sum_{i=k+1}^{n}(X_i - \overline{X}_{n-k})^2,,$$

$$V_k = k(\overline{X}_k - \overline{X})^2 + (n-k)(\overline{X}_{n-k} - \overline{X})^2 \tag{2.4}$$

and $S=\sum_{i=1}^{n}(X_i-\overline{X})^2$. Then $V_k=S-S_k$

Simple algebra leads an alternative expression for V_k as

$$V_k=\frac{n}{k(n-k)}\left[\sum_{i=1}^{k}(X_i-\overline{X})^2\right]^2. \tag{2.5}$$

Therefore

$$U=\max_{1\le k\le n}\sqrt{V_k}=\max_{1\le k\le n}|T_k| \tag{2.6}$$

where

$$T_k=\sqrt{\frac{n}{k(n-k)}}\left[\sum_{i=1}^{k}(X_i-\overline{X})^2\right]. \tag{2.7}$$

Note that $T_1, T_2, \cdots, T_{n-1}$ is a Markov process with $Cov(T_i, T_j)=\sqrt{\frac{i(n-j)}{j(n-i)}}$ for $i<j$ and the partial covariance between T_i and T_j when T_m is fixed equals 0. Hawkins (1977) derived the exact null distribution of U as

$$f_U(x)=2\Phi(x,0,1)\sum_{k=1}^{n-1}g_k(x,x)g_{n-k}(x,x)$$

where $\Phi(x,0,1)$ is the pdf of $N(0,1)$, $g_1(x,s)=1$ for $x,s\ge 0$, and

$$g_k(x,s)=P|T_i|<s, i=1,2,\cdots,k-1 \mid |T_k|=x, \quad \text{for } x,s\ge 0.$$

The asymptotic null distribution is based on the followings:
Let $W_k=X_1+X_2+\cdots+X_k, 1\le k\le n$. Then simple algebra leads to

$$U=\max_{1\le k\le n}\left|\frac{W_k}{\sqrt{n}}-\frac{k}{n}\frac{W_n}{\sqrt{n}}\right|/\left[\frac{k}{n}\left(1-\frac{k}{n}\right)\right]^{1/2}.$$

Let $\{B(t);0\le t<\infty\}$ is a standard Brownian motion. Then under H_0,

$$\left[\frac{W_k-k\mu}{\sqrt{n}};1\le k\le n\right]=^d\{B(k/n);1\le k\le n\}$$

where $=^d$ means "distributed as". Further

$$\begin{aligned}U&=\max_{1\le k\le n}\left|\frac{W_k}{\sqrt{n}}-\frac{k}{n}\frac{W_n}{\sqrt{n}}\right|/\left[\frac{k}{n}\left(1-\frac{k}{n}\right)\right]^{1/2}\\&=\max_{1\le k\le n}\left|\frac{W_k}{\sqrt{n}}-t\frac{W_n}{\sqrt{n}}\right|/[t(1-t)]^{1/2}\\&=^d\max_{nt=1,\cdots,n-1}|B(t)-tB(1)|/[t(1-t)]^{1/2}\end{aligned}$$

where $t=k/n, B_0=B(t)-tB(1)$ is the Brownian bridge. By the properties of Brownian motion and convergence rules from the probability theory, the asymptotic distribution of U is proved to be a Gumbel distribution by Yao and Davis (1986). The following theorem shows the limiting distribution of U based on the properties of Brownian motion.

Theorem 2.1 Under H_0, that is, with no change, for $-\infty < x < \infty$,

$$\lim_{n\to\infty} P[a_n^{-1}(U-b_n) \le x] = \exp - 2\pi e^{1/2} e^{-x},$$

where $a_n = (2\log(\log n))^{-1/2}$, $b_n = a_n^{-1} + \frac{1}{2} a_n \log(\log(\log n))$.

proof. From Chen and Gupta (2000), under H_0,

$$\begin{aligned} P[a_n^{-1}(U-b_n) \le x] &= P[U \le a_n x + b_n] \\ &= P\left(\max_{1 \le nt < \left[\frac{n}{\log n}\right]} \frac{|B(t)|}{\sqrt{t}} \le a_n x + b_n\right) \\ &\quad \bullet P\left(\max_{1 \le n(1-t) < \left[\frac{n}{\log n}\right]} \frac{|B(t)-B(1)|}{\sqrt{1-t}} \le a_n x + b_n\right) + o_p(a_n) \\ &\to \exp(-\pi e^{-1/2} e^{-x}) \bullet \exp(-\pi e^{-1/2} e^{-x}) = \exp(-2\pi e^{-1/2} e^{-x}) \end{aligned}$$

as $n \to \infty$ and by Darling and Erdos (1956) convergence properties for the Brownian motion. The limiting distribution is shown as a Gumbel distribution. □

When the variance is unknown, the likelihood based test statistic is then given by

$$V = \max_{1 \le k \le n} \frac{|T_k|}{S}$$

where $S = \sum_{i=1}^{n} (X_i - \overline{X})^2$ and T_k in (2.7). Worsley (1979) obtained the null distribution of V using Bonferroni approximation. The likelihood test rejects H_0 if $V > c$.

2.3 Change-point Estimation for Mean Change based on the Likelihood

Based on the likelihood with the known variance, the change-point can be estimated as

$$\hat{k} = \operatorname{arg}max_{1 \le k \le n} |T_k| \qquad (2.8)$$

where T_k is in (2.7).

When the variance is unknown,

$$\hat{k} = \operatorname{arg}max_{1 \le k \le n} \frac{|T_k|}{S} \qquad (2.9)$$

is equivalently $\hat{k} = \operatorname{arg}max_{1 \le k \le n} |T_k|$.

Chen and Gupta (2000) showed the distribution of the location of the change-point.

2.4 Attempted Change-point Estimation for Mean Change

Gombay and Horvath (1990) test is developed as the function of mle's. They consider the test statistic based on

$$Z_k = 2kg(\overline{X}_k) + (n-k)g(\overline{X}_{n-k}) - ng(\overline{X}_n) \tag{2.10}$$

where g is a given function. For the hypotheses (2.1), their test rejects H_0 in favor of H_1 for large values of

$$Z(i,j) = \max_{i < m < j} \frac{Z_m}{g^{(2)}(\mu)} \tag{2.11}$$

where $g^{(2)}$ is the second derivative of g and for suitably chosen i and j.

Note that the maximum occurs at the change-point when there is a change-point. Therefore we attempt change-point estimation based on Gombay and Horvath (1990) test which has a functional form of the maximum likelihood as follows:

$$\hat{k}_{GH} = \mathrm{arg}max_{1 < i < j < n} Z(i,j) \tag{2.12}$$

where $g_1(t) = t^2, g_2(t) = \exp(t)$ chosen for $g(\bullet)$ in (2.10).

Gombay and Horvath (1990) showed that the limiting distribution of their test is

$$Z(m_1, m_2)/\sigma^2 \rightarrow sup_{0 \le s \le A} |V(s)|^2$$

in distribution, where $0 < \lambda_1 \le 1 - \lambda_2 < 1$ as $n \rightarrow \infty$,

$$m_1 = n\lambda_1, m_2 = n(1-\lambda_2), A = \frac{1}{2}\{\log(1-\lambda_1)(1-\lambda_2)/\lambda_1\lambda_2\}$$

and $\{V(s), -\infty < s < \infty\}$ is an Ornstein-Uhlenbeck process, i.e. a Gaussian process with mean zero and covariance $\exp(-|t-s|)$. Therefore the distribution of the change-point is shown as

$$\mathrm{arg}max\, Z(m_1, m_2)/\sigma^2 \rightarrow \infty \{k | V_k = sup_{0 \le s \le A} |V(s)|\} \quad \text{as } n \rightarrow \infty.$$

3. Simulation

A simulation study is conducted to see the power of the tests according to the sample size, the amount of change, and the location of change. The parametric test is compared with the Gombay and Horvath (1990) test as the function of mle's.

For the change tests, the LRT based test and the Gomabay and Horvath (1990) test with $g_1(t) = t^2, g_2(t) = \exp(t)$ are compared in the simulation study.

Carlstein (1988) considered the pre-t empirical cdf $_t h(x) = \sum_{i=1}^{nt} I\{X_i \le x\}/nt$ and post-t empirical $cdf\, h_t(x) = \sum_{i=nt+1}^{n} I\{X_i \le x\}/n(1-t)$ for $t \in T_n = \{i/n : 1 \le i \le n-1\}$, with the indicator function, $I(X \le a) = 1$, if $x \le a$, if 0, $x > a$. Carlstein (1988) proposed the change-point estimators as for $j = 1,2,3$

$$T_{carlj} = \mathrm{arg}max_{1 \le t \le n}\{D_j(t)\}$$

where
$$D_1(t) = t^{0.5}(1-t)^{0.5} n^{-1} \sum_{i=1}^{n} |{}_t h(x_i) - h_i(x_i)|,$$

$$D_2(t) = t^{1/2}(1-t)^{1/2}\left[n^{-1}\sum_{i=1}^{n}({}_t h(x_i) - h_i(x_i))^2\right]^{1/2}$$

$$D_3(t) = t^{1/2}(1-t)^{1/2} sup_{1 \le i \le n} |{}_t h(x_i) - h_t(x_i)|.$$

A random sample $X_1, X_2, \cdots, X_n$ are generated from the normal distribution with mean 0 and variance $\sigma^2 = 1$. The mean level change model with one change-point is as follows:

$$X_i = \begin{cases} \mu + \epsilon_i, & i = 1, \cdots, k \\ \mu + \Delta + \epsilon_i, & i = 1, \cdots, n \end{cases} \tag{3.1}$$

where $\mu = 0$ without loss of generality. The amount of change $\Delta = -1.5, -0.5, 0, 0.5, 1, 1.5$, the sample size $n = 50$ and the location of change at $k/n = 0.3, 0.5, 0.8$ are considered. The repetition $r = 1{,}000$ were used in this simulation. The range of the points is restricted from 5th to 45th point due to boundary consideration.

For the power study, $\alpha = 0.05, 0.10$ level empirical critical values were evaluated from the empirical distribution with 10,000 repetitions. <Table 3.1> gives the simulation results of the empirical powers for testing for the existing change. Power of LRT is the most since the correct distribution was incorporated in the test statistic. In comparison, Hawkins (1977) test is

$$T_{Hwak} = \max_{1 \le k \le n} \frac{n}{k(n-k)}\left\{\sum_{i=1}^{k}(X_i - \overline{X})^2\right\}, \tag{3.2}$$

and James et al.(1987) test considered as the square root of the log likelihood ratio test statistics is

$$T_J = \max_{1 \le k \le n} \frac{\left|k\overline{X} - \sum_{i=1}^{k} X_i\right|}{\sqrt{k\left(1 - \frac{k}{n}\right)}}\ . \tag{3.3}$$

The power of tests depends on the location of the change-point. U_{LRT} does not depend on the type of change: decreasing or increasing. When the amount of change is $\Delta = 0.5, 1, 1.5$, the power of T_{GH2} is best but it depends on the type of level change.

For the comparison of change-point estimators, mean and mse(mean squared error) of each estimator were calculated. Also $prop1 = P(|\hat{k} - k| \le 1)$, $prop2 = P(|\hat{k} - k| \le 2)$ and $prop5 = P(|\hat{k} - k| \le 5)$ are calculated to know the local behavior of the change-point estimators. <Table 3.2>, <Table 3.3> and <Table 3.4> show the result.

<Table 3.1> Power comparison study of Change-point tests in Normal distribution with the sample size $n = 50$, the change-point $k = 15,25,40$ in 1,000 repetitions at $\alpha = 0.05$

change-point		k=15		k=25		k=40	
		α=0.10	α=0.05	α=0.10	α=0.05	α=0.10	α=0.05
Δ=-1.5	U_{LRT}	0.995	0.986	0.999	0.995	0.968	0.945
	T_{Hawk}	0.963	0.963	0.984	0.950	0.924	0.859
	T_J	0.809	0.687	0.902	0.813	0.782	0.668
	T_{GH1}	0.995	0.986	0.999	0.995	0.968	0.945
	T_{GH2}	0.755	0.567	0.910	0.803	0.794	0.586
Δ=-1	U_{LRT}	0.857	0.760	0.878	0.819	0.702	0.587
	T_{Hawk}	0.637	0.637	0.684	0.537	0.550	0.393
	T_J	0.546	0.424	0.637	0.533	0.490	0.361
	T_{GH1}	0.857	0.760	0.878	0.819	0.702	0.587
	T_{GH2}	0.453	0.282	0.584	0.419	0.385	0.204
Δ=-0.5	U_{LRT}	0.316	0.212	0.368	0.246	0.244	0.151
	T_{Hawk}	0.133	0.079	0.143	0.067	0.121	0.059
	T_J	0.216	0.140	0.284	0.182	0.207	0.126
	T_{GH1}	0.316	0.212	0.368	0.246	0.244	0.151
	T_{GH2}	0.150	0.079	0.174	0.088	0.111	0.048
Δ=0.5	U_{LRT}	0.313	0.205	0.374	0.257	0.246	0.163
	T_{Hawk}	0.139	0.073	0.169	0.169	0.129	0.072
	T_J	0.217	0.126	0.290	0.180	0.238	0.152
	T_{GH1}	0.313	0.205	0.374	0.257	0.246	0.163
	T_{GH2}	0.409	0.264	0.466	0.317	0.304	0.188
Δ=1	U_{LRT}	0.836	0.755	0.903	0.845	0.733	0.620
	T_{Hawk}	0.646	0.490	0.717	0.556	0.572	0.424
	T_J	0.518	0.410	0.671	0.566	0.512	0.417
	T_{GH1}	0.836	0.755	0.903	0.845	0.733	0.620
	T_{GH2}	0.946	0.904	0.958	0.922	0.820	0.731
Δ=1.5	U_{LRT}	0.997	0.988	0.999	0.997	0.972	0.957
	T_{Hawk}	0.972	0.927	0.986	0.966	0.935	0.861
	T_J	0.813	0.707	0.892	0.788	0.757	0.649
	T_{GH1}	0.997	0.988	0.999	0.997	0.972	0.957
	T_{GH2}	1.000	1.000	1.000	1.000	0.988	0.973

<Table 3.2> Comparison of change-point estimators with $n=50$ and $k=15$

change-point		k=15					
		mean	mse	prop1	prop2	prop5	95% CI
Δ=0.5	$\hat{k}_{U_{LRT}}$	20.881	171.291	0.190	0.275	0.433	(8,41)
	$\hat{k}_{T_{Hink}}$	20.844	214.040	0.145	0.215	0.362	(6,43)
	$\hat{k}_{T_J}$	10.407	120.851	0.073	0.133	0.275	(5,23)
	$\hat{k}_{T_{CS1}}$	20.881	171.291	0.190	0.275	0.433	(8,41)
	$\hat{k}_{T_{CS2}}$	22.490	200.930	0.176	0.261	0.411	(9,42)
	$\hat{k}_{T_{cut1}}$	17.737	172.601	0.165	0.227	0.374	(5,43)
	$\hat{k}_{T_{cut2}}$	18.208	176.280	0.168	0.228	0.379	(5,43)
	$\hat{k}_{T_{cut3}}$	22.229	284.639	0.088	0.135	0.264	(5,45)
Δ=1	$\hat{k}_{U_{LRT}}$	16.263	45.089	0.446	0.589	0.777	(10,23)
	$\hat{k}_{T_{Hink}}$	15.473	69.251	0.391	0.517	0.693	(7,24)
	$\hat{k}_{T_J}$	8.505	67.789	0.111	0.173	0.361	(5,14)
	$\hat{k}_{T_{CS1}}$	16.263	45.089	0.446	0.589	0.777	(10,23)
	$\hat{k}_{T_{CS2}}$	18.416	75.750	0.424	0.554	0.735	(12,30)
	$\hat{k}_{T_{cut1}}$	13.810	41.296	0.385	0.523	0.688	(6,19)
	$\hat{k}_{T_{cut2}}$	14.161	41.677	0.389	0.526	0.699	(6,20)
	$\hat{k}_{T_{cut3}}$	15.937	113.309	0.225	0.311	0.479	(5,33)
Δ=1.5	$\hat{k}_{U_{LRT}}$	15.350	10.182	0.692	0.810	0.932	(13,18)
	$\hat{k}_{T_{Hink}}$	14.528	18.038	0.643	0.755	0.891	(10,17)
	$\hat{k}_{T_J}$	8.568	55.652	0.104	0.169	0.362	(5,14)
	$\hat{k}_{T_{CS1}}$	15.350	10.182	0.692	0.810	0.932	(13,18)
	$\hat{k}_{T_{CS2}}$	17.065	31.817	0.643	0.751	0.871	(14,22)
	$\hat{k}_{T_{cut1}}$	14.270	10.632	0.650	0.761	0.894	(10,17)
	$\hat{k}_{T_{cut2}}$	14.575	11.213	0.655	0.773	0.900	(11,17)
	$\hat{k}_{T_{cut3}}$	14.912	68.592	0.378	0.465	0.614	(5,24)

<Table 3.3> Comparison of change-point estimators with $n=50$ and $k=25$

change-point		k=25					
		mean	mse	prop1	prop2	prop5	95% CI
Δ=0.5	$\hat{k}_{U_{LRT}}$	25.049	110.267	0.205	0.286	0.467	(10,41)
	$\hat{k}_{T_{Hawk}}$	25.379	163.837	0.146	0.198	0.342	(7,44)
	$\hat{k}_{T_J}$	13.263	254.283	0.062	0.085	0.168	(5,27)
	$\hat{k}_{T_{G1}}$	25.049	110.267	0.205	0.286	0.467	(10,41)
	$\hat{k}_{T_{G2}}$	26.878	116.624	0.193	0.273	0.458	(11,43)
	$\hat{k}_{T_{mt1}}$	21.149	170.115	0.156	0.214	0.342	(5,42)
	$\hat{k}_{T_{mt2}}$	21.320	165.514	0.156	0.213	0.347	(5,42)
	$\hat{k}_{T_{mt3}}$	24.151	216.815	0.100	0.141	0.239	(5,45)
Δ=1	$\hat{k}_{U_{LRT}}$	24.899	32.579	0.510	0.632	0.811	(20,30)
	$\hat{k}_{T_{Hawk}}$	25.304	68.080	0.427	0.531	0.690	(15,38)
	$\hat{k}_{T_J}$	13.718	175.404	0.071	0.106	0.232	(5,23)
	$\hat{k}_{T_{G1}}$	24.899	32.579	0.510	0.632	0.811	(20,30)
	$\hat{k}_{T_{G2}}$	27.003	42.747	0.474	0.593	0.762	(22,37)
	$\hat{k}_{T_{mt1}}$	22.332	60.486	0.447	0.558	0.722	(10,28)
	$\hat{k}_{T_{mt2}}$	22.709	56.917	0.446	0.562	0.725	(11,28)
	$\hat{k}_{T_{mt3}}$	21.585	133.467	0.242	0.331	0.498	(5,36)
Δ=1.5	$\hat{k}_{U_{LRT}}$	25.021	9.497	0.685	0.802	0.933	(23,28)
	$\hat{k}_{T_{Hawk}}$	25.040	20.742	0.630	0.739	0.878	(22,28)
	$\hat{k}_{T_J}$	13.800	164.926	0.076	0.109	0.220	(6,23)
	$\hat{k}_{T_{G1}}$	25.021	9.497	0.685	0.802	0.933	(23,28)
	$\hat{k}_{T_{G2}}$	26.504	18.198	0.619	0.733	0.872	(24,31)
	$\hat{k}_{T_{mt1}}$	24.101	17.705	0.659	0.768	0.901	(21,27)
	$\hat{k}_{T_{mt2}}$	24.307	16.035	0.659	0.769	0.902	(21,27)
	$\hat{k}_{T_{mt3}}$	22.109	97.023	0.400	0.499	0.635	(5,31)

<Table 3.4> Comparison of change-point estimators with $n=50$ and $k=40$

change-point		k=40					
		mean	mse	prop1	prop2	prop5	95% CI
Δ=0.5	$\hat{k}_{U_{LRT}}$	30.031	271.797	0.191	0.298	0.525	(8,44)
	$\hat{k}_{T_{Huck}}$	30.414	303.118	0.184	0.282	0.573	(6,45)
	$\hat{k}_{T_J}$	16.404	763.400	0.045	0.069	0.139	(5,39)
	$\hat{k}_{T_{G1}}$	30.031	271.797	0.191	0.298	0.525	(8,44)
	$\hat{k}_{T_{G2}}$	31.393	241.565	0.203	0.322	0.576	(9,44)
	$\hat{k}_{T_{cut1}}$	24.707	454.481	0.116	0.183	0.376	(5,44)
	$\hat{k}_{T_{cut2}}$	24.962	442.028	0.113	0.179	0.379	(5,44)
	$\hat{k}_{T_{cut3}}$	23.621	512.277	0.068	0.118	0.350	(5,45)
Δ=1	$\hat{k}_{U_{LRT}}$	36.537	91.065	0.483	0.614	0.810	(24,43)
	$\hat{k}_{T_{Huck}}$	37.822	92.954	0.474	0.600	0.872	(27,44)
	$\hat{k}_{T_J}$	21.239	520.055	0.086	0.120	0.208	(5,39)
	$\hat{k}_{T_{G1}}$	36.537	91.065	0.483	0.614	0.810	(24,43)
	$\hat{k}_{T_{G2}}$	38.494	55.626	0.523	0.660	0.894	(34,44)
	$\hat{k}_{T_{cut1}}$	31.351	247.787	0.381	0.488	0.650	(6,42)
	$\hat{k}_{T_{cut2}}$	31.870	230.618	0.393	0.497	0.663	(7,42)
	$\hat{k}_{T_{cut3}}$	23.933	513.727	0.193	0.256	0.414	(5,43)
Δ=1.5	$\hat{k}_{U_{LRT}}$	39.249	15.467	0.697	0.817	0.949	(37,42)
	$\hat{k}_{T_{Huck}}$	40.178	17.252	0.662	0.775	0.974	(39,43)
	$\hat{k}_{T_J}$	23.671	392.269	0.077	0.118	0.220	(7,38)
	$\hat{k}_{T_{G1}}$	39.249	15.467	0.697	0.817	0.949	(37,42)
	$\hat{k}_{T_{G2}}$	40.343	6.923	0.691	0.802	0.984	(39,43)
	$\hat{k}_{T_{cut1}}$	36.646	85.048	0.619	0.724	0.856	(28,41)
	$\hat{k}_{T_{cut2}}$	36.963	76.933	0.621	0.728	0.863	(30,41)
	$\hat{k}_{T_{cut3}}$	27.592	415.272	0.372	0.450	0.575	(5,42)

<Table 3.1> shows that the power of LRT is the best when the change-point occurs in the middle. Also there is the same trend for the power of Gombay and Horvath test. But the power of Hawkins test is best when the change-point occurs in the early part of data with decreasing change. Therefore the location of the change-point affects the power of each test.

Overall the change-point estimation with LRT is better since the parametric distributional assumption holds. In the Carlstein nonparametric estimation, $\hat{k}_{T_{CARL2}}$ is better in the sense of mse. When the change occurs in the middle of the data, the change-point estimators are better since it can have more balanced information in the estimation procedure. <Table 3.4> gives the change-point estimation results that the applied estimator of Gombay and Horvath type works better than the estimator with the likelihood when the change-point occurs in the later part of data. The estimation ability also depends on the location of the change-point. We found that the function of the MLE's can work as test statistics and change-point estimators.

4. Concluding Remark

Considered are the problems of testing change and estimating for the mean change-point when the data are from the normal distribution. Overall the change-point estimation with LRT is better since the parametric distributional assumption holds. Gombay and Horvath (1990) tests have good power as a function derived from the likelihood. Also we tried the change-point estimation based on Gombay and Horvath (1990) test statistic. Via simulation this attempted estimator has a good performance as a change-point estimator. This functional form is expected to be used for other distributions in change analysis.

References

[1] Bhattacharya, G.K. and Johnson, R.A. (1968). Nonparametric Tests for Shift at an Unknown Time Point. *Annals of Mathematical Statistics*, Vol. 39, 1731-1743.

[2] Buckley, M.J. (1991). Detecting a Smooth Signal: Optimality of Cusum based Procedures. *Biometrika*, Vol. 78, 2 253-262.

[3] Carlstein, E. (1988). Nonparametric Change-point Estimation. *Annals of Statistics*, Vol. 16, 188-197.

[4] Chen, Jie and Gupta, A.K. (2000). Parametric Statistical Change Point Analysis.

Birkhauser, Berlin.

[5] Chernoff, H. and Zacks, S. (1964). Estimating the Current Mean of a Normal Distribution Which is Subject to Changes in Time. *Annals of Mathematical Statistics*, Vol. 35, 999-1028.

[6] Gardner, L.A. (1969). On Detecting Change in the Mean of Normal Variates. *Annals of Mathematical Statistics*, Vol. 40, 116-126.

[7] Gombay, E. and Horvath, L. (1990). Asymptotic Distributions of Maximum Likelihood Tests for Change in the Mean. *Biometrika*, Vol. 77, 411-414.

[8] Gombay, E. and Horvath, L. (1996). Approximations for the Time of Change and the Power Function in Change-point Models. *Journal of Statistical Planning and Inference*, Vol. 52, 43-66.

[9] Hawkins, D.M. (1977). Testing a Sequence of Observations for a Shift in Location. *Journal of American Statistical Association*, Vol. 72, 180-186.

[10] Hinkley, D.V. (1970). Inference about the Change-point in a Sequence of Random Variables. *Biometrika*, Vol. 57. 1-16.

[11] James, B., James, K.L. and Siegmund, D. (1987). Tests for a Change-point. *Biometrika*, Vol. 74, 71-83.

[12] Kander, A. and Zacks, S. (1966). Testing Procedures for Possible Changes in Parameters of Statistical Distributions Occurring at Unknown Time Points. *Annals of Mathematical Statistics*, Vol. 37, 1196-1210.

[13] Page, E.S. (1955). A Test for a Change in a Parameter Occurring at an Unknown Point. *Biometrika*, Vol. 42, 523-527.

[14] Sen. A.K. and Srivastava, M.S. (1975). On Tests for Detecting Change in Mean. *Annals of Statistics*, Vol. 3, 98-108.

[15] Worsley, K.J. (1979). On the Likelihood Ratio Test for a Shift in Location of Normal Populations. *Journal of American Statistical Association*, Vol. 74, 365-377.

[16] Yao, Y.C. and Davis, R.A. (1986). The Asymptotic Behavior of the Likelihood Ratio Statistic for Testing Shift in Mean in a Sequence of Independent Normal Variates. *Sankhya,* 339-353.

[Received July 2006, Accepted August 2006]