Null Distributions of the Quasi-Bayesian Test Statistics

Reza Habibi
Department of Statistics, Central Bank of Iran

Summary

Consider a sequence of independent observations that is susceptible to a change in the distributions. It is interested to test if the distributions are changed after an unknown point. Two quasi-Bayesian-type statistics are derived. Both test procedures are locally most powerful. The asymptotic null distribution of test statistics are given. Numerical critical points of the test statistics are tabulated for certain selected values of the sample sizes.

Keywords: Change point; Kiefer process; Locally most powerful; Quasi-Bayesian; Response surface regressions

1 Introduction. Let $X_1, ..., X_n$ be a sequence of independent random variables with distribution functions $F_i$, $i = 1, ..., n$. In this paper, we will derive two quasi-Bayesian-type statistics to test if $F_i$'s are changed at unknown time point. The hypothesis can be written in the following way

$$H_0 : F_1 = ... = F_n = F \text{ vs. } H_1 : F_i = \begin{cases} F & i \leq k_0, \\ G & k_0 < i \leq n. \end{cases}$$

Let $k_0 = \lfloor nt_0 \rfloor$ for some unknown $t_0 \in (0, 1)$.

Since Page (1954) formally studied stopping rules in the context of quality control, there has been a surge of papers on change point problems. Chernoff and Zacks (1964) first proposed the quasi-Bayesian test statistic to detect a change point in means of independent normal observations. Kander and Zacks (1966) studied the quasi-Bayesian test in exponential family distributions. A comprehensive review of the subject can be found in Csörgö and Horváth (1997) and Chen and Gupta (2000).

Throughout this paper, let $W(t)$ and $B(t)$ denote the standard Brownian motion and Brownian bridge on $(0, 1)$ and let $K(t, F(x))$ be the Kiefer process. We define the Kiefer bridge $K_b(t, F(x))$ in the following way

$$K_b(t, F(x)) = K(t, F(x)) - tK(1, F(x)).$$

2 Quasi-Bayesian Test Statistic. First, we propose a quasi-Bayesian test statistic assuming the initial distribution $F$ is known. To derive the test statistic let $U_i = F(X_i)$, $i = 1, 2, ..., n$. Under the null hypothesis of no change point $U_i$'s are independent and identically distributed (iid) random variables distributed uniformly on $(0, 1)$ and then $E(U_i) = 0.5$. Under the alternative hypothesis, we have

$$E(U_i) = \begin{cases} 0.5 & i = 1, 2, \cdots, k_0, \\ \theta & i = k_0 + 1, \cdots, n, \end{cases}$$
where $\theta = E_G(F(X_{k_0+1}))$. It is easy to see that

$$\theta = P(X_{k_0} \leq X_{k_0+1}).$$

Under the null hypothesis of no change point $\theta = 0.5$. Under $H_1$, $\theta$ is less or greater than 0.5. For example when $F = N(\mu_1, 1)$ and $G = N(\mu_2, 1)$, then $\theta \leq 0.5$ ($\theta \geq 0.5$) if and only if $\mu_1 \leq \mu_2$ ($\mu_1 \geq \mu_2$). One can note that, under $H_1$, the mean of $U_i$’s has changed.

Following Kander and Zacks (1966) and , assuming $t_0$ is a random variable with a uniform prior distribution on $(0, 1)$, the quasi-Bayesian test statistic to detect a change point in means of observations is given by

$$n \sum_{i=1}^{n} \left( \frac{i-1}{n} \right)(U_i - \overline{U}),$$

where $\overline{U} = \sum_{i=1}^{n} U_i / n$. Generally, when $t_0$ is a random variable with the prior distribution function $\Pi(\cdot)$ on $(0, 1)$ and $\Pi(0) = 0$, $\Pi(1) = 1$, then the test statistic is

$$T_n = \sum_{i=1}^{n} \Pi \left( \frac{i-1}{n} \right)(U_i - \overline{U}),$$

Kander and Zacks (1966) showed that this test procedure is the locally most powerful. It should be noticed that null distribution of the quasi-Bayesian statistic does not depend on $F$. The limiting null distribution $T_n$ is given by

$$n^{-0.5} T_n \overset{d}{\rightarrow} \sqrt{12} \int_{0}^{1} \Pi(t) dB(t) \overset{d}{=} N(0, 12 \int_{0}^{1} (\Pi(t) - \overline{\Pi})^2 dt),$$

where $\overline{\Pi} = \int_{0}^{1} \Pi(t) dt$.

Since the exact null distribution of test statistic is difficult to obtain, a Monte Carlo experiment with 50000 repetition is performed and the sample critical points, $c_\alpha$ are computed. These values are given in Table 1. We define $\int_{c_\alpha}^{\infty} g(x) dx = \alpha$, where $g(x)$ is the density function of $n^{-0.5} T_n$. Here, we suppose that $t_0$ has uniform prior and $T_n$ is $\sum_{i=1}^{n} \left( \frac{i-1}{n} \right)(U_i - \overline{U})$. 

2
Table 1: Critical values for $n^{-0.5}T_n$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$c_{0.01}$</th>
<th>$c_{0.025}$</th>
<th>$c_{0.05}$</th>
<th>$c_{0.1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.1662844</td>
<td>0.148862</td>
<td>0.1316295</td>
<td>0.1058444</td>
</tr>
<tr>
<td>4</td>
<td>0.1730354</td>
<td>0.1541919</td>
<td>0.1343171</td>
<td>0.1082243</td>
</tr>
<tr>
<td>5</td>
<td>0.1806353</td>
<td>0.1565246</td>
<td>0.1348865</td>
<td>0.1070459</td>
</tr>
<tr>
<td>6</td>
<td>0.1829822</td>
<td>0.1580109</td>
<td>0.1354744</td>
<td>0.1074125</td>
</tr>
<tr>
<td>7</td>
<td>0.1839453</td>
<td>0.1585703</td>
<td>0.1347575</td>
<td>0.1070116</td>
</tr>
<tr>
<td>8</td>
<td>0.1863019</td>
<td>0.1612699</td>
<td>0.1373558</td>
<td>0.1074812</td>
</tr>
<tr>
<td>9</td>
<td>0.1882433</td>
<td>0.1602091</td>
<td>0.1362748</td>
<td>0.1077110</td>
</tr>
<tr>
<td>10</td>
<td>0.1888307</td>
<td>0.160797</td>
<td>0.1365997</td>
<td>0.1086998</td>
</tr>
<tr>
<td>15</td>
<td>0.1883044</td>
<td>0.1600885</td>
<td>0.1357263</td>
<td>0.1066627</td>
</tr>
<tr>
<td>20</td>
<td>0.1907633</td>
<td>0.1624800</td>
<td>0.1370582</td>
<td>0.1066665</td>
</tr>
<tr>
<td>25</td>
<td>0.1921422</td>
<td>0.1620582</td>
<td>0.1375438</td>
<td>0.1073608</td>
</tr>
<tr>
<td>30</td>
<td>0.1929953</td>
<td>0.1633384</td>
<td>0.1377394</td>
<td>0.1072576</td>
</tr>
<tr>
<td>$\infty$</td>
<td>0.1927309</td>
<td>0.1630317</td>
<td>0.1360348</td>
<td>0.1058505</td>
</tr>
</tbody>
</table>

**Remark 1.** Another approach to test $H_0$, is to apply the quasi-Bayesian test to transformed data $Y_i = \Phi^{-1}(U_i), i = 1, 2, \cdots, n$, where $\Phi^{-1}(\cdot)$ is the inverse of distribution function of standard normal distribution, that is

$$T_n^* = \sum_{i=1}^{n} \Pi(i-1/n)(Y_i - \bar{Y}).$$

One can show that, under the null hypothesis of no change point, then

$$T_n^* = \sum_{i=1}^{n}(w_i - \bar{w})Y_i \sim N(0, \sum_{i=1}^{n}(w_i - \bar{w})^2),$$

where $w_i = \Pi(i-1/n)$ and $\bar{w} = (1/n) \sum_{i=1}^{n} \Pi(i-1/n)$.

When $F$ is unknown, it is replaced with the null empirical distribution function $F_n$. In this case, $U_i = R_i/n$, at which $R_i$ the rank of $X_i$ under the null hypothesis and $\bar{U} = (n+1)/2$. The quasi-Bayesian test statistic in this case is

$$(1/n) \sum_{i=1}^{n} \Pi(i-1/n)(R_i - (n+1)/2).$$

It should be noticed that the null distribution of above rank based quasi-Bayesian statistic does not depend on $F$. Table 2 gives the exact null critical points of rank based quasi-Bayesian statistic when $\Pi(t) = t$. 
3 Another quasi-Bayesian test statistic. In connection with the quasi-Bayesian test, we define \( B_i = I(X_i \leq x) \), for fixed \( x \) and for \( i = 1, 2, \ldots, n \). The indicator \( I(y \leq x) \) is one if \( y \leq x \) and is zero if \( y > x \). Clearly, \( B_i \)'s are distributed as Bernoulli random variables with parameter of success \( p_i = F_i(x) \), \( i = 1, 2, \ldots, n \). Let \( p = F(x) \) and \( p_0 = G(x) \). The marginal log likelihood ratio function under \( H_1 \) to that under \( H_0 \) is given by

\[
\int_0^1 \{\log(p/p_0) \sum_{i=[nt]+1}^n X_i + \log(q/q_0) \sum_{i=[nt]+1}^n (1 - X_i)\} d\Pi(t),
\]

where \( q = 1 - p \) and \( q_0 = 1 - p_0 \). Following a derivation method suggested by Kander and Zacks (1966), as \( \delta = p - p_0 \to 0^+ \), the locally most powerful test statistic is given by

\[
T_n^{**}(x) = \sum_{i=1}^n \Pi\left(\frac{i-1}{n}\right)\{B_i - B\} = \sum_{i=1}^n \Pi\left(\frac{i-1}{n}\right)\{I(X_i \leq x) - F_n(x)\},
\]

where \( F_n(x) \) is the empirical distribution of \( X_i \). To remove the effect of \( x \), we use the following test statistic

\[
T_n^{**} = \sup_x |T_n^{**}(x)|.
\]

To derive the asymptotic distribution of \( T_n \), we define the sample process \( K_n(t, F_n(x)) \) by

\[
K_n(t, F_n(x)) = n^{-1/2} \sum_{i=1}^{[nt]} \{I(X_i \leq x) - F_n(x)\}.
\]

and we notice that

\[
n^{-1/2}T_n^{**}(x) = \int_0^1 \Pi(t)K_n(dt, F_n(x)).
\]

<table>
<thead>
<tr>
<th>( n )</th>
<th>( c_{0.01} )</th>
<th>( c_{0.025} )</th>
<th>( c_{0.05} )</th>
<th>( c_{0.1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.2222222</td>
<td>0.2222222</td>
<td>0.2222222</td>
<td>0.2222222</td>
</tr>
<tr>
<td>4</td>
<td>0.3125</td>
<td>0.3125</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>0.36</td>
<td>0.36</td>
<td>0.32</td>
<td>0.28</td>
</tr>
<tr>
<td>6</td>
<td>0.4305556</td>
<td>0.4027778</td>
<td>0.3472222</td>
<td>0.2916667</td>
</tr>
<tr>
<td>7</td>
<td>0.4897959</td>
<td>0.4285714</td>
<td>0.3877551</td>
<td>0.3061224</td>
</tr>
<tr>
<td>8</td>
<td>0.53125</td>
<td>0.46875</td>
<td>0.40625</td>
<td>0.328125</td>
</tr>
<tr>
<td>9</td>
<td>0.5679012</td>
<td>0.5061728</td>
<td>0.4320988</td>
<td>0.345679</td>
</tr>
<tr>
<td>10</td>
<td>0.605</td>
<td>0.535</td>
<td>0.465</td>
<td>0.365</td>
</tr>
<tr>
<td>15</td>
<td>0.7422222</td>
<td>0.6444444</td>
<td>0.5511111</td>
<td>0.4355556</td>
</tr>
<tr>
<td>20</td>
<td>0.8575</td>
<td>0.7375</td>
<td>0.63</td>
<td>0.4925</td>
</tr>
<tr>
<td>25</td>
<td>0.9776</td>
<td>0.8256</td>
<td>0.696</td>
<td>0.5504</td>
</tr>
<tr>
<td>30</td>
<td>1.061667</td>
<td>0.9016667</td>
<td>0.7638889</td>
<td>0.5972222</td>
</tr>
<tr>
<td>( \infty )</td>
<td>4.338395</td>
<td>3.665395</td>
<td>3.066574</td>
<td>2.387111</td>
</tr>
</tbody>
</table>
The empirical process $K_n(t, F_n(x))$ is decomposed to two terms as follows

$$K_n(t, F_n(x)) = U_n(t, F(x)) - tU_n(1, F(x)),$$

where

$$U_n(t, F(x)) = n^{-1/2} \sum_{i=1}^{[nt]} \{I(X_i \leq x) - F(x)\}.$$

As the sample size $n$ goes to infinity and under the null hypothesis of no change point, $U_n(\cdot, F(\cdot))$ converges to Kiefer process $K(t, F(x))$ as well as $K_n(\cdot, F_n(\cdot))$ converges to Kiefer bridge $K_b(t, F(x))$ as defined in Introduction. More precisely, as $n \to \infty$, under $H_0$, then

$$U_n(\cdot, F(\cdot)) \xrightarrow{d} K(\cdot, F(\cdot))$$

and

$$K_n(\cdot, F_n(\cdot)) \xrightarrow{d} K_b(\cdot, F(\cdot)).$$

Then the continuous mapping theorem implies that, under $H_0$,

$$n^{-1/2}T^{**}_n \xrightarrow{d} \Lambda(F(\cdot)) = \int_0^1 \Pi(t)K_b(dt, F(\cdot)).$$

Then, under $H_0$, $T^{**}_n$ converges to $\sup_x |\Lambda(F(x))|$ in distribution.

**Remark 2.** The limiting null distribution can be found in the following way. Notice that

$$\sum_{i=1}^n w_i \{I(X_i \leq x) - F_n(x)\} =$$

$$= \sum_{i=1}^n (w_i - \bar{w}) \{I(X_i \leq x) - F_n(x)\} =$$

$$= \sum_{i=1}^n (w_i - \bar{w}) \{I(X_i \leq x) - F(x)\},$$

where $w_i = \Pi(i - 1/n)$ and $\bar{w} = (1/n) \sum_{i=1}^n \Pi(i - 1/n)$. Then the test statistic is given by

$$T^{**}_n = \sup_x |\int_0^1 (\Pi(t) - \Pi)U_n(dt, x)|$$

which converges to $\sup_x |\int_0^1 (\Pi(t) - \Pi)K(dt, x)|$, where $\Pi = \int_0^1 \Pi(t)dt$.

**Remark 3.** The similar convergence results are obtained for the random weight cases. For example, for a sequence of iid random variables $\{G_{in}\}$ satisfying

$$nP(G_{1n} \in dx) \rightarrow \alpha e^{-x}x dx,$$

consider random exchangeable weights defined by

$$\Pi(i/n) - \Pi(i-1/n) = \frac{G_{i(n-1)}}{\sum_{i=1}^n G_{i(n-1)}},$$

5
\(i = 1, \cdots, n-1\), (see Zarepour and Habibi (2006)). As \(n \to \infty\), then

\[
n^{-1/2} T_n^{**}(\cdot) \xrightarrow{d} \int_0^1 \frac{S(t)}{S(1)} K_\phi(dt, F(\cdot)),
\]

where \(S(t)\) is gamma process (Ferguson and Klass (1972)). Two special case are when \(\{G_i\}\) is a sequence of iid random variables and there exists a sequence of positive constants \(a_n\) such that

\[
nP(a_n^{-1} G_i \in dx) \xrightarrow{v} \alpha x^{-\alpha-1} I(x > 0) dx.
\]

By letting \(G_{in} = G_i / a_n\), then \(n^{-1/2} T_n^{**}(\cdot)\) converges in distribution to the

\[
\int_0^1 \frac{S(t)}{S(1)} K_\phi(dt, F(\cdot)),
\]

where \(S(t)\) is stable process (Resnick (1987)). The second case, is \(\Pi(t) = U_{i:n}\), where \(U_{i:n}\) are the order statistics of a size of \(n\) sample of iid uniform random variables on \((0, 1)\) with \(U_{0:n} = 0\) and \(U_{n:n} = 1\) (see Lo (1987) in the Bayesian bootstrap setting). Notice that

\[
(U_{1:n}, U_{2:n}, \cdots, U_{n-1:n}) \overset{d}{=} \left( \frac{S_1}{S_n}, \frac{S_2}{S_n}, \cdots, \frac{S_{n-1}}{S_n} \right),
\]

where \(S_i = \sum_{j=1}^i E_j\) for \(E_1, \cdots, E_n\) a sequence of iid random variables with \(\text{EXP}(1)\) distribution (see Ferguson (1996), page 88). It is easy to show that

\[
n^{-1/2} T_n^{**} \xrightarrow{d} \sup_x \int_0^1 t K_\phi(dt, F(x)),
\]

which corresponds to the result when \(\Pi(t) = t\).

References


parameters of statistical distributions occurring at unknown time points. *Ann.

Statist.* No **4**, 2138-2148.

[8] Page, E. S. (1954). *Continuous inspection schemes*. Biometrika **41**, 100-
115.

Springer.