AMO- Advanced Modeling and Optimization, Volume 14, Number 1, 2012

A Note on Beta Approximation for Change Point Estimator

Reza Habibi

Department of Statistics, Central Bank of Iran, Tehran, Iran

Abstract. This paper is concerned with Beta approximation for distribution of change point estimator. This distribution is very important for power analysis.

Keywords: Beta approximation; Brownian bridge; Change point; Cusum procedure

1 Introduction. Let $X_{kn},\ k = 1, 2, ..., n$ be a sequence of independent observations such that

$$X_{kn} = \mu_{kn} + \varepsilon_{kn},$$

at which $\mu_{kn} = E(X_{kn}) = \theta_{0n}$, for $k = 1, ..., k_0$, and $= \theta_{1n}$, for $k = k_0 + 1, ..., n$. suppose that ε_{kn} are *iid* zero mean random variables with common variance σ_n^2. Let $\delta_n = \theta_{0n} - \theta_{1n}$ and $\sqrt{n}\delta_n/\sigma_n \to \lambda$ as $n \to \infty$. The above relations describe a shift in mean model. Since the magnitude of change (δ_n)over the standard deviation of data (σ_n) goes to zero as n goes infinity, we refer to above model as small change in mean case. Time point k_0 is unknown and it is estimated in practice. Let $k_0 = [nt_0]$ for some $t_0 \in (0, 1)$. The change point analysis has been received considerable attentions in statistical literatures. Some excellent are Csorgo and Horvath (1997), Chen and Gupta (2000) and Khodadadi and Asgharian (2004). The cusum (see Lee et al. (2004) and references therein) change point estimator of k_0 is

$$\widehat{k}_n = \mathrm{argmax}_k |S_n^X(k)|,$$

where $S_n^X(k) = \sum_{i=1}^{k}(X_{in} - \overline{X}_n)$ with $\overline{X}_n = (1/n)\sum_{i=1}^{n} X_{in}$. Let $\widehat{t}_n = \widehat{k}_n/n$. The distribution of change point estimator is well studied in the literatures. For example, when δ_n and σ_n are independent of n, Hinkley (1970) showed that the limiting distribution of $\widehat{k}_n$ is the maximizer of a two-sided random walk. Bai (1994) showed that the limiting distribution of least square change point estimator in linear process setting is minimizer of a two-sided Brownian motion with a drift. As follows, we derive the asymptotic distribution of $\widehat{t}_n$ in small change cases.

One can see that

$$S_n^X(k) = S_n^{\varepsilon}(k) + \frac{\sqrt{n}\delta_n}{\sigma_n} g_{k_0}^*(k),$$

[1]Reza Habibi is staff of Central Bank of Iran. He has a PhD in Statistics form Shiraz U.

***AMO - Advanced Modeling and Optimization. ISSN: 1841-4311**

where

$$g^*_{k_0}(k) = \begin{cases} \frac{k}{n}(1-\frac{k_0}{n}) & k \le k_0 \\ -\frac{k_0}{n}(1-\frac{k}{n}) & k \ge k_0+1. \end{cases}$$

Change k to $[nt]$ in the above formulas and let $B_n(t) = S_n^X([nt])$. It is easy to see that as $n \to \infty$, then

$$B_n(\cdot) \Longrightarrow B(\cdot) + \lambda g_{t_0}(\cdot),$$

where $B(t)$ is standard Brownian bridge on $(0, 1)$ and

$$g_{t_0}(t) = \begin{cases} t(1-t_0) & t \le t_0 \\ -t_0(1-t) & t > t_0. \end{cases}$$

Notation $\Longrightarrow$ stands for weak convergence on $D(0, 1)$. Following Kim and Pollard (1990), we conclude that

$$\widehat{t}_n \xrightarrow{d} \widehat{t} = \text{argmax}_{t\in(0,1)}|B(t) + \lambda g_{t_0}(t)|.$$

However, this is the asymptotic distribution of $\widehat{t}_n$. In the next section, we study the finite sample distributional behavior of $\widehat{t}_n$ by a Beta fitting as distribution of this estimator. This distribution is very important for power analysis.

Remark 1. Note that under the null hypothesis of no change point, then $\lambda = 0$ and $\widehat{t}_n$ converges in distribution to maximizer of $|B(t)|$ over $t \in (0, 1)$. Also, note that since $|\widehat{t}_n| \le 1$, therefore $\widehat{t}_n$ is uniformly integrable and so $E(\widehat{t}_n) \to E(\widehat{t})$. Our Monte Carlo simulation results shows that $E(\widehat{t}) = 0.5$.

2 Beta approximation. In many applications, it is necessary to approximate the distribution of complicated statistics using known and "easy to work" parametric distributions. When the target distribution is continuous and bounded, a good selection is the Beta distribution (see Habibi, 2011, and reference therein). Hereafter, we use the notation $\widehat{k}_{k_0}$ ($\widehat{t}_{t_0}$), to insist the distribution of $\widehat{k}_n(\widehat{t}_n)$ depends on k_0 (t_0). Since $\widehat{t}_{t_0}$ is between 0 and 1, suppose that it has a beta distribution $B(\alpha_{t_0}, \beta_{t_0})$. We want to find the functional forms of α_{t_0} and β_{t_0} such that $P(\widehat{k} = k)$ is well approximated by $P((k-1)/n \le \widehat{t}_{t_0} < k/n)$, that is

$$P(\widehat{k} = k) \simeq P((k-1)/n \le \widehat{t}_{t_0} < k/n),$$

for $k = 1, ..., n-1$. Under H_0 ($t_0 = 0$), variable $\widehat{t}_{t_0}$ is uniformly distributed on (0,1), then $\alpha_0 = \beta_0 = 1$. By a Monte Carlo simulation, we understand that the sampling distribution $\widehat{k}_{k_0}$ is unimodal and its mode is k_0, we also find that

$$\widehat{k}_{n-k_0} \stackrel{d}{=} n - \widehat{k}_{k_0}.$$

Therefore, we see that the sampling distribution $\widehat{t}_{t_0}$ is unimodal and its mode is t_0 and $\widehat{t}_{(1-t_0)} \stackrel{d}{=} 1-\widehat{t}_{t_0}$. We also see that

$$\widehat{t}_{(1-t_0)} \stackrel{d}{=} B(\alpha_{1-t_0}, \beta_{1-t_0}) \stackrel{d}{=} 1-\widehat{t}_{t_0} \stackrel{d}{=} B(\beta_{t_0}, \alpha_{t_0})$$

that is $\beta_{t_0} = \alpha_{1-t_0}$, for all $t_0 \in (0,1)$, and then $\widehat{t}_{t_0} \stackrel{d}{=} B(\alpha_{t_0}, \alpha_{1-t_0})$. Since $\widehat{t}_{t_0}$ is unimodal, one can conclude that $\alpha_{t_0} > 1$. The mode of $B(\alpha_{t_0}, \alpha_{1-t_0})$ is t_0, then

$$\frac{\alpha_{t_0} - 1}{\alpha_{t_0} + \alpha_{1-t_0} - 2} = t_0.$$

Some solutions are linear functions $\alpha_{t_0} = 1+ at_0$ with $a > 0$. One can see that the necessary and sufficient condition for above equation, is that $\alpha_{t_0} = 1 + tg(t_0)$,for some positive function g defined on $(0,1)$ such that

$$g(t_0) = g(1-t_0), \text{ for every } t_0 \in (0,1).$$

As follows, we want to find α_{t_0}. Let $m_{t_0} = E(\widehat{t}_{t_0})$. Then $\alpha_{t_0} = \frac{1-2t_0}{1-(1/m_{t_0})t_0}$. The Monte Carlo simulation gives m_{t_0} for some selected values of t_0. It is seen that our method works well for $P(\widehat{k} = k_0)$. In practice, a continuity correction is needed.

Examples. Here, we survey our method for some simulated examples. The size of data sequence is 100 involves independent observations, there is a change point in 40. The $P(\widehat{k} = 40)$ for both monte Carlo method and beta approximation values are given. The results are given in the following table.

Table 1: Simulation Results

dist(before)	dist(after)	Monte Carlo	Beta
$N(0,1)$	$N(-2,1)$	0.627	0.615
$N(1,1)$	$N(3,1)$	0.636	0.645
$Exp(1)$	$Exp(3)$	0.304	0.299
$Exp(1)$	$Exp(2.5)$	0.24	0.23
$N(0,1)$	$Exp(1)$	0.245	0.24

References

[1] Bai, J. (1994). Least squares estimation of a shift in linear processes. *J. Time Ser. Anal.* **15**, 453- 472.

[2] Chen, J. and Gupta, A. K. (2000). *Parametric statistical change point analysis.* Birkhäuser.

[3] Csorgo, M. and Horvath, L. (1997). *Limit theorems in change-point analysis*, N.Y., Wiley & Sons.

[4] Habibi, R. (2011). A note on approximating continuous distribution functions. *submitted*.

[5] Hinkley, D. (1970). Inference about the change-point in a sequence of random variables. *Biometrika* **57**, 1–17.

[6] Khodadadi, A. and Asgharian, M. (2004). Change point problem and regression: An annotated bibliography. *Tech. Repo.* McGill university.

[7] Kim, J. and Pollard, D. (1990). Cube root asymptotics. *Annals of Statistics* **18**, 191-219.

[8] Lee, S., Tokutsu, Y. and Maekawa, K. (2004). The CUSUM test for parameter change in regression models with ARCH errors. *J. Japan. Statist. Soc.* **3**, 173-186.