A Change-Point Test for Autoregressive Processes Using a Harmonic Mean P-Value

L. Ma [a], **G. Sofronov** [a]

[a] *Department of Mathematics and Statistics, Macquarie University, Sydney, Australia*
Email: georgy.sofronov@mq.edu.au

Abstract: In many different applications, it is important to know if time series data is generated from a single underlying mechanism or not. This problem, known as a change-point problem, can be formulated as a multiple hypotheses testing problem. In this paper, we propose a harmonic change-point test (HarmonicCPT) to identify and validate change-points in an autoregressive process. The method consists of two steps. First, we develop likelihood ratio based scan statistics on gathering the local information by comparing two adjacent sequences within each scanning window. The corresponding p-values are collected from each test. Any changes in mean, autoregressive coefficients, or variance lead to rejections of the null hypothesis that the data is generated from the same process within the scanning window. Next, we calculate a harmonic mean p-value by combining all of the tests on which the decision that whether to reject the global null hypothesis depends. The simulation study shows that the proposed scan statistic is quite sensitive to the variance change, and the harmonic mean p-value procedure is efficient in detecting the significant p-values.

Keywords: *Change-point test, autoregressive process, multiple testing, harmonic mean p-value*

1 INTRODUCTION

Change-point detection is an advanced statistical methodology that helps to understand the underlying structure of the time series data by providing change-point estimates. Various change-point detection methods have been applied to data in different fields including genomics (e.g., Li et al. (2016), Cao & Wu (2015)), astronomy (e.g., Fisch et al. (2019)), climatology (e.g., Lu et al. (2010), Li & Lund (2012)), cyber-security (e.g., Adams & Heard (2016)), neurophysiology (e.g., Messer et al. (2017)), economics (e.g., Sofronov & Ma (2017)), and finance (e.g., Zhu et al. (2013), Ma & Sofronov (2020)). Any changes in mean, variance or autoregressive (AR) coefficients could cause the significant structural change and, therefore, it is important for decision makers to know the number of change-points and their locations.

It has been a long history that the change-point problem can be formulated in form of hypothesis tests. The null hypothesis is set to be no change-point; the alternative hypothesis includes one or more than one change-point setting, categorized as single and multiple change-point detection. It is more challenging to accurately identify the change-points when the alternative hypothesis is at least one change-point. In order to estimate multiple change-points, the global alternative is usually decomposed to multiple local hypothesis tests, and then the estimations are made by aggregating local information. The previous studies from this perspective include but are not limited to Niu & Zhang (2012), Hao et al. (2013), Fryzlewicz (2014), Yau & Zhao (2016), Eichinger et al. (2018), Ma et al. (2020).

Change-point methods typically aim to identify the number of change-points and their locations. In this paper, our objective is to test if an AR process is generated from a single underlying mechanism or not. To solve this problem, we propose a harmonic change-point test (HarmonicCPT). The method consists of two steps. Firstly, we utilize likelihood ratio based scan statistics on gathering the local information by testing two symmetric sequences within each scanning window. The distribution of the local likelihood ratio test statistic was derived so that we can get the p-value from each test. Next, instead of analyzing individual p-value, we combine all the p-values using the Bonferroni correction and harmonic mean p-value method, both of them are robust to dependent p-values. The rationale of combining multiple tests is that aggregated tests are more sensitive to the extremely small or significant p-values.

In many applications, the number of change-points is unknown; hence the estimated number of change-points using different methods can vary. The methods that share the same assumptions are expected to be in agreement between each other's estimated change-points or seek external approaches to validate their estimates. The proposed HarmonicCPT method can provide solid evidence to exclude zero change-point cases, and the p-value plot after scanning procedure is applicable for validating the estimated change-points.

2 HARMONIC CHANGE-POINT TEST

2.1 Model setting and multiple hypothesis testing

Given the observations $\{X_t\}_{t=1,\ldots,T}$, it is assumed that the stationary time series

$$\{x_{L_t(h)}, \quad L_t(h) = t-h+1,\ldots,t\}, \quad \{x_{R_t(h)}, \quad R_t(h) = t+1,\ldots,t+h\} \tag{1}$$

are partial observations of stochastic process $\{X_t\}$. So that the scanning window is defined as $(L_t(h), R_t(h))$. It is of interest to test, for $t = h, h+1 \ldots, T-h$,

$$\begin{aligned} H_0(t) &: x_{L_t(h)} \text{ and } x_{R_t(h)} \text{ are generated from same stochastic process} \\ H_1(t) &: \text{Not } H_0. \end{aligned} \tag{2}$$

The model under the null hypothesis is assumed to be

$$X_t + \sum_{j=1}^{p_X} \theta_j X_{t-j} \quad = \quad \varepsilon_t, \tag{3}$$

where $\varepsilon_t \sim$ i.i.d. $N(0,\sigma^2)$, $\{X_t\}$ is assumed to be a weakly stationary AR(p_X) process.

2.2 A Likelihood Ratio Scan Statistic

Define $L_t(h)$ and $R_t(h)$ as the left and right windows at location t; in order to test if the observations $\{x_{L_t(h)}\}$ and $\{x_{R_t(h)}\}$ are from the same stationary process, we then apply the modified parametric method proposed

by Grant & Quinn (2017) to discriminate the spectral density. For $h \leq t \leq T-h$, consider the process within a scanning window

$$\begin{aligned} X_{L_t(h)}(t) + \sum_{j=1}^{p_{L_t}} \theta_{X_{L_t},j} X_{L_t(h)}(t-j) &= \varepsilon_{L_t(h)}(t), \\ X_{R_t(h)}(t) + \sum_{j=1}^{p_{R_t}} \theta_{X_{R_t},j} X_{R_t(h)}(t-j) &= \varepsilon_{R_t(h)}(t), \end{aligned} \tag{4}$$

where the left and right segments have the same length denoted by T_h (or $h-1$). It is assumed that for all t, $\left\{\varepsilon_{L_t(h)}\right\}$ and $\left\{\varepsilon_{R_t(h)}\right\}$ are i.i.d. normal.

Following the derivation from Grant (2018), the spectral densities of $\{X_{L_t(h)}(t)\}$ and $\{X_{R_t(h)}(t)\}$ can be written as,

$$f_{X_{L_t(h)}}(\omega) = \frac{\sigma^2_{L_t}}{2\pi \mid 1 + \sum_{j=1}^{p_{L_t}} \theta_{X_{L_t},j} e^{-ij\omega} \mid^2}, \quad f_{X_{R_t(h)}}(\omega) = \frac{\sigma^2_{R_t}}{2\pi \mid 1 + \sum_{j=1}^{p_{R_t}} \theta_{X_{R_t},j} e^{-ij\omega} \mid^2}.$$

There are two perspective considered for the hypothesis test in (2). By testing whether the ratio of $f_{X_{L_t(h)}}$ and $f_{X_{R_t(h)}}$ is constant, the hypothesis test in (2) is equivalent to

$$\begin{aligned} H_0(t) &: \theta_{X_{L_t},j} = \theta_{X_{R_t},j}, \quad \text{for all } j \\ H_1(t) &: \text{Not } H_0. \end{aligned} \tag{5}$$

Testing whether that $f_{X_{L_t(h)}}$ and $f_{X_{R_t(h)}}$ are the same, the hypothesis test in (2) is equivalent to

$$\begin{aligned} H_0(t) &: \theta_{X_{L_t},j} = \theta_{X_{R_t},j}, \quad \text{for all } j, \quad \sigma^2_{L_t} = \sigma^2_{R_t} \\ H_1(t) &: \text{Not } H_0. \end{aligned} \tag{6}$$

In Ma et al. (2020), we have applied modified spectral discrimination test developed by Grant & Quinn (2017) to compare time series at certain location, which can be adapted to test each time point with a sliding window. Thus, under the hypothesis test (6), the change in mean is not considered as a change-point, the scan statistics can be derived as

$$\begin{aligned} \Lambda_h(t) &= 2(\hat{ll}_{t;1} - \hat{ll}_{t;0}) \\ &= T_h \log\left(\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2_{L_t;1}}\right) + T_h \log\left(\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2_{R_t;1}}\right), \end{aligned} \tag{7}$$

where $\hat{ll}_{t,1}$ and $\hat{ll}_{t,0}$ are the maximised log-likelihood functions under H_0 and H_1, while $\widehat{\sigma}^2_{L_t;1}$ and $\widehat{\sigma}^2_{R_t;1}$ are the estimators of $\sigma^2_{L_t}$ and $\sigma^2_{R_t}$ under alternative hypothesis, $\widehat{\sigma}^2_{t;0}$ is the estimator of the residual variance of scanning window under H_0.

Theorem 1. *Under null hypothesis of* (6)*, for $t = h, h+1, \ldots, T-h$, the distribution of $\Lambda_h(t)$ converges to the χ^2_{p+1} as $T_h \to \infty$.*

The proof can be found in Grant (2018). The p-value is the probability that obtain $\Lambda_h(t)$ or more if the H_0 is true,

$$\begin{aligned} p_{(t)} &= Pr(\lambda_t \geq \Lambda_h(t) \mid \theta \in \Theta_{H_0}) \\ &= Pr(\chi^2_{p+1} \geq \Lambda_h(t) \mid \theta \in \Theta_{H_0}), \end{aligned} \tag{8}$$

where $p = \lfloor (\log T_h)^v \rfloor$, $v > 1$. In addition, it is feasible to take the mean shift into account, the model that we consider

$$\begin{aligned} (X_{L_t(h)}(t) - \mu_{L_t}) + \sum_{j=1}^{p_{L_t}} \theta_{X_{L_t},j} (X_{L_t(h)}(t-j) - \mu_{L_t}) &= \varepsilon_{L_t(h)}(t), \\ (X_{R_t(h)}(t) - \mu_{R_t}) + \sum_{j=1}^{p_{R_t}} \theta_{X_{R_t},j} (X_{R_t(h)}(t-j) - \mu_{R_t}) &= \varepsilon_{R_t(h)}(t), \end{aligned} \tag{9}$$

and the null hypothesis is

$$H_0(t) : \theta_{X_{L_t},j} = \theta_{X_{R_t},j}, \quad \text{for all } j, \quad \sigma^2_{L_t} = \sigma^2_{R_t} \quad \mu_{L_t} = \mu_{R_t}. \tag{10}$$

Unlike the hypothesis in (6), it implies that the variability in autoregressive coefficients, mean or variance is considered as a change-point. The scan statistics is derived as

$$\Lambda_h(t) \quad = \quad T_h \log\left(\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2_{L_t;1}}\right) + T_h \log\left(\frac{\widehat{\sigma}^2_0}{\widehat{\sigma}^2_{R_t;1}}\right). \tag{11}$$

Theorem 2. *Under null hypothesis of* (10)*, for* $t = h, h+1, \ldots, T-h$*, the distribution of* $\Lambda_h(t)$ *converges to the* χ^2_{p+2} *as* $T_h \to \infty$*.*

The proof can be found in Grant (2018). Similarly, the p-values of the scan tests can be obtained.

2.3 Combining Dependent P-Values

From multiple hypothesis tests, under (6) or (10), we can obtain a series of p-values, such as $p_{(h)}$, $p_{(h+1)}, \ldots, p_{(T-h)}$. Generally, there are two directions to make inference from the p-values. One option is to perform second-stage multiple tests for identifying whether there are local clustered signals, which require the change-point detection techniques for dependent data. For example, Cao & Wu (2015) proposed a boundary detection algorithm to utilize the clustered signal, allowing each p-value segment following different distributions under the alternative hypothesis, and construct test statistics based on smoothed p-values, the global null hypothesis and alternative hypothesis are assumed to be

$$\begin{aligned} H_0(t) \quad &: \quad p_{(t)} \sim \mathcal{U}(0,1) \\ H_1(t) \quad &: \quad p_{(t)} \nsim \mathcal{U}(0,1). \end{aligned} \tag{12}$$

Due to our forward-moving window spectral discrimination procedure, a series of neighbouring hypotheses tend to be rejected continuously from when the right side window runs into the change-point. In this situation, multiple rejections appear to occur in a row without any assumptions on the distance between change-points, which may be caused by one change-point or more than one change-points. So it is impossible to locate the change-point precisely. Therefore, we consider another option by making the inference from the combined p-values rather than investigating individual p-values. By combining multiple tests, the clump of significant signals could be used as an indicator to the hypothesis testing

$$\begin{aligned} H_0(t) \quad &: \quad \text{there is no change-point in } \{X_t\} \\ H_1(t) \quad &: \quad \text{there is at least one change-point in } \{X_t\}. \end{aligned} \tag{13}$$

We propose to apply the harmonic mean p-value technique in order to combine all dependent p-values. Motivated by the genome-wide association studies, Wilson (2019) developed the harmonic mean p-value (HMP) to control the familywise error rate of combined tests with improving the statistical power. Similar to the controlling procedure, Bonferroni correction, both of them can be used to compare individual tests and combine dependent tests. By using Bonferroni correction, one can compare the minimum adjusted p-values with the usual significant level. While HMP can be simply expressed as the reciprocal of the arithmetic mean of the reciprocals of the p-values. It has the desirable property to excel in detecting the grouped subtle values, it is expected to be more powerful than Bonferroni correction, which is the rationale for using the HMP approach.

Wilson (2019) provides two versions of HMP implementation: (i) one can compare the weighted HMP calculated by a set of p-values, with the harmonic significance threshold $\alpha_{\text{HMP}} \cdot w_{T^\star}$; (ii) when the asymptotically exact HMP is available, it can be directly compared with the significance level $\alpha \cdot w_{T^\star}$. The method is implemented in an `R` package `harmonicmeanp`, which can be found at `https://cran.r-project.org/web/packages/harmonicmeanp/index.html`. In this paper, we follow implementation (i). The algorithm can be described as follows.

Algorithm 1 Combining p-values procedure via HMP

1: A sequence of p-values is computed from defined scan statistics, $p_{(h)}, p_{(h+1)}, \ldots, p_{(T-h)}$.

2: Calculate p_{HMP} and HMP significant threshold $\alpha_{\text{HMP}} \cdot w_{T^\star}$,

$$p_{\text{HMP}} = \frac{\sum_{t=h}^{T-h} w_t}{\sum_{t=h}^{T-h} w_t/p_t},$$

where $w_{T^\star} = \sum_{t=h}^{T-h} w_t = 1$, w_t is the weight for each p-value, it is assumed that each test is equally weighted. $T^\star$ is the total number of p-values.

$$\alpha_{\text{HMP}} = \frac{1}{F^{-1}_{\text{Landau}}(\alpha \mid \mu, \psi)} = \frac{1}{x_\alpha}$$

we define $F_{\text{Landau}}(x_\alpha \mid \mu, \psi)$ as the cumulative distribution function of the Landau distribution,

$$F_{\text{Landau}}(x_\alpha \mid \mu = c \cdot (\log T^\star + 0.874367), \psi = c \cdot \pi/2) = \alpha$$

α is the predefined type I error rate, c is a constant with respect to p in (8). If $p + 1 > 2$, $c = \frac{T^\star}{p_{\text{HMP}}} \cdot (1 - F_{\text{Gamma}}(\log \frac{T^\star}{p_{\text{HMP}}} \mid \nu = \frac{p+1}{2}, \kappa = 1))$, where F_{Gamma} is the cumulative distribution function of a Gamma distribution.

3: Reject the global null hypothesis under (13) when $p_{\text{HMP}} \leq \alpha_{\text{HMP}} \cdot w_{T^\star}$.

3 Numerical Experiments

3.1 Single change-point in AR process

In this section, we are interested in testing the null hypothesis as follows

$$H_0(t) : \theta_{X_{L_t},j} = \theta_{X_{R_t},j}, \quad \text{for all } j, \quad \sigma^2_{L_t} = \sigma^2_{R_t} \quad \mu_{L_t} = \mu_{R_t},$$

and compare the validity between two multiple testing controlling procedure. A change-point is generated at location t_c, the time series model within the scanning window at t_c is assumed to be piecewise stationary AR(1) process with different parameters.

$$\begin{aligned}
(X_{L_{t_c}(h)}(t) - \mu_{L_{t_c}}) + \sum_{j=1}^{p_{L_{t_c}}} \theta_{X_{L_{t_c}},j}(X_{L_{t_c}(h)}(t-j) - \mu_{L_{t_c}}) &= \varepsilon_{L_{t_c}(h)}(t), \\
(X_{R_{t_c}(h)}(t) - \mu_{R_{t_c}}) + \sum_{j=1}^{p_{R_{t_c}}} \theta_{X_{R_{t_c}},j}(X_{R_{t_c}(h)}(t-j) - \mu_{R_{t_c}}) &= \varepsilon_{R_{t_c}(h)}(t).
\end{aligned} \tag{14}$$

During the scanning process, the p-values are sequentially collected from $t = h$. It is reasonable to consider the scenarios $t_c \in (h, T-h)$. In addition, we consider the cases that the shift in AR coefficients, variance, or mean respectively; the change in both AR coefficients and variance, both AR coefficients and mean, or both variance and mean; also, we consider the change in all of the parameters. In this study, 100 sequences of time series data are simulated for each case under each scenario, and $T = 1000$.

t_c/T	$\{\theta_{X_{L_t},j}, \theta_{X_{R_t},j}\}$	$\{\mu_{L_t}, \mu_{R_t}\}$	$\{\sigma^2_{L_t}, \sigma^2_{R_t}\}$	BonferroniCPT	HarmonicCPT
0.5	$\{0.5, 0.5\}$	$\{0, 1\}$	$\{1, 1\}$	0%	0%
	$\{0.4, 0.6\}$	$\{0, 0\}$	$\{1, 1\}$	4%	1%
	$\{0.5, 0.5\}$	$\{0, 0\}$	$\{1, 2\}$	99%	98%
	$\{0.4, 0.6\}$	$\{0, 1\}$	$\{1, 1\}$	2%	2%
	$\{0.4, 0.6\}$	$\{0, 0\}$	$\{1, 2\}$	98%	98%
	$\{0.5, 0.5\}$	$\{0, 1\}$	$\{1, 2\}$	98%	98%
	$\{0.4, 0.6\}$	$\{0, 1\}$	$\{1, 2\}$	99%	99%

Table 1. The table first outlines the value of parameters used in each numerical experiment, 100 sequences of piecewise stationary AR process are generated for each experiment. Also, it shows the positive detection rate (the proportion that rejects H_0 over 100 sequences) for BonferroniCPT and HarmonicCPT method.

The sequences are divided into two equal-length segments under this scenario. Table 1 shows that, in most cases, both BonferroniCPT and HarmonicCPT methods can successfully identify the change-point. It has been noticed that the procedure has a very high type II error rate for cases 1, 2, and 4. Under case 1, the coefficients and variance remain the same; the mean increased by a unit, the results show that our method lost power for this case. From case 2, we can see that our method cannot detect the change-point when there is only a weak change in the AR coefficient. Comparing cases 2 and 4, with both coefficients and mean shifting weakly together, the method still lost the power. In comparison, it is interesting to see that when the variance adds a unit regardless of the variety in coefficients and mean, our method achieves a good detection power.

3.2 Comparison with classical models

Next, we compare the Bonferroni controlling procedure with the HMP method through Models(A-I) described in Ma et al. (2020). 100 sequences from each model are generated. The scanning window setting is same as in Ma et al. (2020). The results in Table 2 show that the controlling performance of BonferroniCPT and HarmonicCPT are pretty similar except for the slightly higher type II error rate of HarmonicCPT under Model G. The reason that Model D has the lowest detection power among Models(A-I) is the location of the change-point (it is very close to the beginning of the sequence). The results from other models indicate that our method is very efficient for the global alternative.

			BonferroniCPT		**HarmonicCPT**	
	N	t_c/T	$\hat{N}$		$\hat{N}$	
Model A			0	≥ 1	0	≥ 1
$\beta = 0.4$	0	$\emptyset$	100	0	100	0
$\beta = 0.7$	0	$\emptyset$	100	0	100	0
$\beta = -0.1$	0	$\emptyset$	100	0	100	0
$\beta = -0.7$	0	$\emptyset$	99	1	99	1
			0	≥ 1	0	≥ 1
Model D	1	(50/1024)	10	90	11	89
Model I	1	(128/256)	1	99	1	99
			0	≥ 1	0	≥ 1
Model B	2	(512/1024,768/1024)	0	100	0	100
Model C	2	(400/1024,612/1024)	0	100	0	100
Model E	2	(400/1024,750/1024)	0	100	0	100
Model F	2	(400/1024,750/1024)	0	100	0	100
			0	≥ 1	0	≥ 1
Model G	3	(125/1024,532/1024,704/1024)	0	100	2	98
Model H	3	(125/1024,532/1024,704/1024)	0	100	0	100

Table 2. The simulation performance of BonferroniCPT method and HarmonicCPT method, where N represents the true number of change-points, corresponding location(s) denoted by t_c/T, $\hat{N}$ is the estimated number of change-points.

4 CONCLUSIONS

For change-point detection in autoregressive time series data, we have developed autoregression discrimination based scan statistics in order to compare the underlying structure of the data within each scanning window. The scan statistic moves through each time point from the beginning to the end of the data. Because of the nature of proposed scan statistics, any change(s) in mean, AR coefficients, or variance lead to the continuous rejections of the null hypothesis that the data is generated from the same process within the scanning window. The Bonferroni multiple testing controlling procedure and harmonic adjusted p-value procedure are used to combine all the tests. Both of them are very efficient in detecting the significant p-values, it does not seem possible to conclude which method is superior. The key messages that we can obtain from the single change-point simulated cases is that, compared with the change in autocorrelation coefficient and the mean, the proposed scan statistic is more sensitive to the variance change. Otherwise, the BonferroniCPT and HarmonicCPT produce similar results. The reason that Model D has the lowest detection power among Models(A-I) is the location of the change-point, which is very close to the beginning of the sequence. But it still has 90% positive detection rate due to the dramatic change in AR(1) coefficient. The results from other models indicate that our method is very efficient for the global alternative.

REFERENCES

Adams, N. M. & Heard, N. A. (2016), *Dynamic networks and cyber-security*, Vol. 1, World Scientific.

Cao, H. & Wu, W. (2015), 'Changepoint estimation: another look at multiple testing problems', *Biometrika* **102**(4), 974–980.

Eichinger, B., Kirch, C. et al. (2018), 'A MOSUM procedure for the estimation of multiple random change points', *Bernoulli* **24**(1), 526–564.

Fisch, A. T. M., Eckley, I. A. & Fearnhead, P. (2019), 'A linear time method for the detection of point and collective anomalies', *ArXiv/abs/1806.01947* .

Fryzlewicz, P. (2014), 'Wild binary segmentation for multiple change-point detection', *The Annals of Statistics* **42**(6), 2243–2281.

Grant, A. J. (2018), Parametric Methods for Time Series Discrimination, PhD thesis, Macquarie University, Sydney, Australia.

Grant, A. J. & Quinn, B. G. (2017), 'Parametric spectral discrimination', *Journal of Time Series Analysis* **38**(6), 838–864.

Hao, N., Niu, Y. S. & Zhang, H. (2013), 'Multiple change-point detection via a screening and ranking algorithm', *Statistica Sinica* **23**(4), 1553–1572.

Li, H., Munk, A. & Sieling, H. (2016), 'FDR-control in multiscale change-point segmentation', *Electronic Journal of Statistics* **10**(1), 918–959.

Li, S. & Lund, R. (2012), 'Multiple changepoint detection via genetic algorithms', *Journal of Climate* **25**(2), 674–686.

Lu, Q., Lund, R. & Lee, T. C. M. (2010), 'An MDL approach to the climate segmentation problem', *The Annals of Applied Statistics* **4**(1), 299–319.

Ma, L., Grant, A. & Sofronov, G. (2020), 'Multiple change point detection and validation in autoregressive time series data', *Statistical Papers* **61**(4), 1507–1528.

Ma, L. & Sofronov, G. (2020), 'Change-point detection in autoregressive processes via the Cross-Entropy method', *Algorithms* **13**(5), 1–16.

Messer, M., Costa, K. M., Roeper, J. & Schneider, G. (2017), 'Multi-scale detection of rate changes in spike trains with weak dependencies', *Journal of computational neuroscience* **42**(2), 187–201.

Niu, Y. S. & Zhang, H. (2012), 'The screening and ranking algorithm to detect DNA copy number variations', *The Annals of Applied Statistics* **6**(3), 1306–1326.

Sofronov, G. & Ma, L. (2017), Change-point detection in time series data via the Cross-Entropy method, *in* G. Syme, D. Hatton MacDonald, B. Fulton & J. Piantadosi, eds, '22nd International Congress on Modelling and Simulation', Modelling& Simulation Society Australia& New Zealand, Australia, pp. 195–201.

Wilson, D. J. (2019), 'The harmonic mean p-value for combining dependent tests', *Proceedings of the National Academy of Sciences* **116**(4), 1195–1200.

Yau, C. Y. & Zhao, Z. (2016), 'Inference for multiple change points in time series via likelihood ratio scan statistics', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **78**(4), 895–916.

Zhu, X., Li, Y., Liang, C., Chen, J. & Wu, D. (2013), 'Copula based change point detection for financial contagion in chinese banking', *Procedia Computer Science* **17**, 619 – 626.