5-15-2018

Bi-Directional Testing for Change Point Detection in Poisson Processes

Moinak Bhaduri
bhadurim@unlv.nevada.edu

BI-DIRECTIONAL TESTING FOR CHANGE POINT DETECTION IN POISSON PROCESSES

by

Moinak Bhaduri

Bachelor of Science - Statistics (Hons)
St. Xavier's College, Kolkata, India
2010

A dissertation submitted in partial fulfillment of
the requirements for the

Doctor of Philosophy - Mathematical Sciences

Department of Mathematical Sciences
College of Sciences
The Graduate College

University of Nevada, Las Vegas
May 2018

UNLV | GRADUATE COLLEGE

Dissertation Approval

The Graduate College
The University of Nevada, Las Vegas

April 5, 2018

This dissertation prepared by

Moinak Bhaduri

entitled

BI-DIRECTIONAL TESTING FOR CHANGE POINT DETECTION IN POISSON PROCESSES

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy - Mathematical Sciences
Department of Mathematical Sciences

Chih Hsiang Ho, Ph.D.
Examination Committee Chair

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Interim Dean

Amei Amei, Ph.D.
Examination Committee Member

Malwane Ananda, Ph.D.
Examination Committee Member

Kaushik Ghosh, Ph.D.
Examination Committee Member

Guogen Shan, Ph.D.
Graduate College Faculty Representative

Abstract

Bi-directional testing

by

Moinak Bhaduri

Dr. Chih-Hsiang Ho, Examination Committee Chair
Professor of Mathematical Sciences (Statistics)
University of Nevada, Las Vegas, USA

Point processes often serve as a natural language to chronicle an event's temporal evolution, and significant changes in the flow, synonymous with non-stationarity, are usually triggered by assignable and frequently preventable causes, often heralding devastating ramifications. Examples include amplified restlessness of a volcano, increased frequencies of airplane crashes, hurricanes, mining mishaps, among others. Guessing these time points of changes, therefore, merits utmost care. Switching the way time traditionally propagates, we posit a new genre of *bidirectional tests* which, despite a frugal construct, prove to be exceedingly efficient in culling out non-stationarity under a wide spectrum of environments. A journey surveying a lavish class of intensities, ranging from the tralatitious power laws to the deucedly germane rough steps, tracks the established unidirectional forward and backward test's evolution into a p-value induced dual bidirectional test, the best member of the proffered category. Niched within a hospitable Poissonian framework, this dissertation, through a prudent harnessing of the bidirectional category's classification prowess, incites a refreshing alternative to estimating changes plaguing a soporific flow, by conducting a sequence of tests. Validation tools, predominantly graphical, rid the structure of forbidding technicalities, aggrandizing the swath of applicability. Extensive simulations, conducted especially under

hostile premises of hard non-stationarity detection, document minimal estimation error and reveal the algorithm's obstinate versatility at its most unerring.

Acknowledgements

Mathematics and mathematicians thrive on proofs, and at a time and place like this, thanking one's advisor has become so formulaic a custom that gauging one's sincerity often proves elusive. To mine, Dr. Ho, however, I genuinely owe a deep debt of gratitude, especially for his ready availability, guidance (often tempered with brutal directness), and unwavering support for almost every research endeavor that took my fancy. I'm thankful for the topic he has offered me and will treasure his insights into coming to grips with life in academia.

To other members on my committee, I'm equally indebted. Dr. Amei's two courses on Stochastic Modeling have been instrumental in making several sections of this dissertation possible, while Dr. Ghosh's tips, especially on coding and formulation, had a calming influence at times of desperation. Despite unforgiving schedules, they have been generous enough to keep their doors ajar at all times. Their ways of approaching a problem and attention to details should breed several generations of imitators, but never an equal, to either. Dr. Shan's research made me appreciate the applicability of Poisson processes beyond traditional repairable systems, especially in clinical trials, and his course of Biostatistical Methods forced me to think beyond R regarding softwares. Dr. Ananda's work on reliability has, at times, served as a model to look up to. I thank them all for being indispensable parts of this project and wish them every success in their research endeavors.

A big thanks must also go to the Department of Computer Science in general, for its hospitality, and to Dr. Zhan, in particular, for offering me Research Assistantships over the last couple of semesters. I realized their way of looking at non-stationarity and showing it is fundamentally equivalent to some of the methods proposed here, is rather close to my heart.

Research focus was intensified as a result of this stint, and fruitful collaboration ensued. Also, their computers are lovely.

The Graduate College and the GPSA have been exceedingly supportive of my research travels, and appreciative of my teaching and scholarly contributions. Many thanks for letting me sit on committees, serve on student bodies, inviting me as a panelist for discussions, and be part of the Rebel Research and Mentorship Program. While touring campuses for my new Assistant Professorship position, I realized how crucial these experiences are. I wish you every success in the days to come and will be happy to be of assistance regarding anything at all.

And thus, the list threatens to go on.

I'll be failing miserably in my part however, unless I thank Dr. Asera, who has been supporting me at all my local talks with unflagging enthusiasm, my students, for aiding me sail through the dark days when research screeched to a halt, and the hosts and the audience of the conferences I've been to, for asking probing questions.

It takes a certain temerity to undertake mathematical research, and ever since I was a boy, I aspired to contribute something original. Through my school and college days, my parents, through telling stories of great scientists and discoveries, fueled the flames of that callow optimism with enviable ability. Steering a kid past the endless distractions today's world has to offer, demands vigilance. Mom and dad were always more than equal to the task. In recent times, they tolerated (or should I say "relished"?) my being away for so long, bore the brunt of my mood swings, and served as (unjustified) mute avenues for me to vent my angst and anguish of fruitless research days. Sorry for all that, I do not believe others can have better parents. Not being statisticians, you'll probably never read past this

point, but your silent support and unquestioning faith breathe life into every equation and simulation to follow. I dedicate this work to you in less than a heartbeat.

Table of Contents

List of Tables

List of Figures

Chapter 1

Rudiments

About a century ago, on 14th October 1913, a methane explosion rocked the Universal Colliery, a coalfield in Glamorgan, South Wales. 439 men and boys perished, marking the worst disaster in UK coal mining history. Had the British government paid heed to sound statistical warnings, the calamity could have been averted. At another extreme, the National Earthquake Prediction Evaluation Council (NEPEC), a branch of the United States Geological Survey (USGS), consuming a wealth of government resources and troubling scores of unsuspecting civilians, issued an unnerving statement in 1985: an earthquake clocking M6 on the Richter scale, originating in the San Andreas Fault would shake Parkfield, California around or before 1993. The dreaded earthquake never happened. This research cogitates on the mechanics of blunders such as these, interrogates the traditional wisdom on predicting these rare events, and instigates a novel way of understanding the temporal properties of rules that drive such processes.

Point processes and repairable systems, the brand of stochastic processes this dissertation will analyze, enjoy a rich, checkered, and at times, tortuous history with regards to both their theoretical underpinnings and varied applicability. Necessitated and originated by pressing problems in queuing and branching theory, these disciplines, especially over the last decade,

have carved out for themselves a hospitable niche within the formidable structure of pure and applied probability. Excellent texts, offering broad overviews, exist, that cater the needs of readers with varied levels of maturity and inclination. Rigdon and Basu (2000) [122], Bain and Engelhardt (1991)[11], Ross (1996, 2010) [131], [132], Resnick (2005, 2002) [120] [119], Knill (2009) [82], Bhat (1984) [16], Feller (2005) [47], Gertsbakh (2005) [51], Kovalenko et al. (1997) [83], Lindsey (2004) [97], among others, do a remarkable job in laying out the foundations.

Another group is geared towards specific objectives: Gamiz et al. (2011) [50] for instance, investigates the non-parametric aspects of repairable systems, Nelson (1995) [110] and Ross (1990) [130] cover simulating Point Processes in great detail. More from this specialized class will be surveyed as we go along. Regression models have been studied by Cameron and Trivedi (1998) [21], Lindsey (1995) [96], Hilbe (2014) [63]. Yet another category, research monographs along veins similar to Limnios and Nikulin (2000) [95], collects journal articles about a given unified theme. The first section of this introductory chapter thus, will be devoted to garnering the thoughts necessary to mathematize the framework to follow.

While Point Processes serve as the objects to operate on, change point identification is one of this dissertation's ultimate telos. Section 2 reviews that literature, once again, from a broad perspective, with details introduced at relevant places, especially in Chapters 3 and 4. The final section samples representative examples.

1.1 Point Processes

To admire, analyze, exploit, and at times, emulate the vagaries of nature, stock markets or the human body, a Point Process (PtP), frequently referred to as a Counting Process (CP),

often serves as an apt mathematical instrument. Crudely, it tracks the number of occurrences of events over a domain of interest. Throughout this dissertation, the evolution will be over time, although generalizations of the methods proposed can be furnished to embrace space, more complex topologies (such as the unit circle, the torus etc.) or combinations thereof. The sampling of time will dictate the index set. Unless explicitly mentioned, the process will be monitored in real time, though limitations of the measuring instrument(s), among other issues (such as the cost involved), often withhold that luxury, forcing one to check the system only at discrete (such as hourly or yearly) intervals. In the final chapter, we will outline suggestions that could be adopted under that condition. Mostly, thus, a PtP will be treated here as a continuous time, discrete space stochastic process. Some definitions are naturally, in order.

Definition 1.1. *A PtP* $\{N(t)\}_{t>0}$ *is a collection of random quantities, where* $N(t)$ *for each* $t > 0$ *counts the number of observations in the time interval* $[0, t]$. *Analogously,* $N(a, b]$ *represents the number of events in* $(a, b]$ *and with* $b > a > 0$,

$$N(a, b] = N(b) - N(a). \tag{1.1}$$

The flow can be described by specifying the joint density of $\{N(t_1), N(t_2), ..., N(t_n)\}$ for a general n and $\{t_1, t_2, ..., t_n\}$.

Definition 1.2. *The non-decreasing function* $\Lambda(.) : R^+ \to R^+$

$$\Lambda(t) = E(N(t)), \quad t > 0 \tag{1.2}$$

is the mean function of the process and outputs the average number of observations through time t.

Monotonicity of $N(t)$ induces monotonicity of $\Lambda(t)$, and this average function is always right continuous (Rigdon and Basu (2000) [122]).

Definition 1.3. *The rate of occurrence of failures (ROCOF)* $\mu(.)$ *is defined as the derivative of the mean function when the latter is differentiable:*

$$\mu(t) = \frac{d}{dt}\Lambda(t), \quad t > 0 \tag{1.3}$$

and it gives the rate of change in the average number of shocks.

A notion inescapably relevant to any discussion on PtPs is the one of an intensity function, which controls the frequency with which we observe instances, and hence, the inferences on the PtP that ensue. Several (often equivalent) definitions are prevalent, but the one most commonly used is

$$\lambda(t) = lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t] \geq 1)}{\Delta t}, \quad t > 0. \tag{1.4}$$

Thus, it gives the instantaneous probability of observing at least one failure in a small time interval. An elevated intensity, therefore, would create a highly *active* process which, depending on the context, could spell doom (such as volcanic eruptions, earthquakes or missed credit card payments) or elation (someone winning a lottery).

At times, especially when the history of the process is available, it is necessary to be pedantic, and introduce the notion of a complete intensity function

$$\lambda(t) = lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t] \geq 1 | H_t)}{\Delta t}, \quad t > 0 \tag{1.5}$$

where H_t represents the set of failure times $\{t_i : i = 1, 2, ..., N(t)\}$. An equivalent formulation of a PtP is often through a set of global times

$$0 < T_1 < T_2 < ... < T_n \tag{1.6}$$

where T_i represents the time of the ith event. This, arguably, is a *repairable system* approach, and the inter-event times are defined as

$$X_i = T_i - T_{i-1}, \quad i = 1, 2, ... \tag{1.7}$$

with $T_0 = 0$. The equivalence follows due to

Lemma 1.1. *For a given PtP* $\{N(t)\}_{t>0}$*, the following hold:*

1) $T_i > v \Leftrightarrow N(v) < i$.

2) $T_i \leq w \Leftrightarrow N(w) \geq i$.

3) $v < T_i \leq w \Leftrightarrow N(v) < i \leq N(w)$

Thus, the joint density of any of the following sets uniquely determines the joint density of the rest:

1) $N(u_1), N(u_2), ..., N(u_n)$ for arbitrary n and u_i's.

2) $T_1, T_2, ..T_n$ for arbitrary n.

3) $X_1, X_2, ..X_n$ for arbitrary n.

The proofs are standard, and we point interested readers to Rigdon and Basu (2000) [122].

A PtP gets mathematically tractable if one or both of the following hold.

Definition 1.4. *A PtP will have stationary increments if for all* $k > 0$ *and* $s > 0$

$$P\{N(t, t+s] = k\} \tag{1.8}$$

is free of t.

Put differently, the stationary increment property ensures that the probability of observing a given number of shocks in a given time interval is dependent only on the length of the interval and not on its location on the time axis.

Definition 1.5. *A PtP will enjoy independent increments if for any n and $r_1 < s_1 \leq r_2 < s_2 \leq .... \leq r_n < s_n$,*

$$P\{N(r_1, s_1] = k_1, N(r_2, s_2] = k_2, ..., N(r_n, s_n] = k_n\} = \prod_{i=1}^{n} P\{N(r_i, s_i] = k_i\}. \tag{1.9}$$

This implies that the random variables living on disjoint intervals are independent.

The discussions thus far will suffice as a working introduction to PtPs. Intricacies and classifications, however, are rife in literature. Jacobsen (2006) [74] among others, treats $\{N(t)\}_{t>0}$ as a *simple* PtP (SPtP), demarcating them from *marked* PtPs (MPtP), where in addition to recording the time of occurrence, one also registers the type of shock involved (for instance, the time of an earthquake *and* its magnitude on the Richter scale, or the time a customer enters a ticket counter *and* the gender). Snyder (1975) [136] studies the countability and uncountability of the *mark space* in considerable details. The formulation we have described implies a finite number of shocks in any finite time interval. The requirement may be relaxed by allowing equalities in (1.6), i.e., by requiring only

$$P(0 < T_1 \leq T_2 \leq ...) = 1 \tag{1.10}$$

which will make the PtP an *explosive* one (Jacobsen (2006) [74]). Resnick (2002) [119] studies Laplace functionals on PtPs. An interesting generalization, where the history of the process influences the future is studied by authors such as Snyder (1975) [136]. The orderliness restriction in (1.6) is dropped again to generate these *self-exciting* PtPs. A rich source

of inferential tools on PtPs is Karr (1986) [78] where one can examine both distribution theory and intensity theory-based inference, and how to exploit them to transform a PtP, or approximate it. Statistical models (frailty, regression etc.) based on PtPs have been discussed by Andersen et al. (1993) [4].

Unser and Tafti (2014) [144] educate readers on *sparsity* in the PtP context, touching on transformation domain and wavelet domain tools to handle it. Empirical process connections through Poisson bridges have been discussed by Shorack and Wellner (1986) [134]. Matthes et al. (1978) [104] studies PtPs that are infinitely divisible. Reiss (1993) [118] covers a host of interesting themes, including strong approximations of, and distances between PtPs, with spatial generalizations. Researchers with geometric inclinations will relish Stoyan et al. (1987) [139] who talk about random tessellations and Lowen and Teich (2005) [101] who study PtPs through self-similar objects known as fractals.

1.2 Change detection

An inferential inquiry often concerns the nature of progression of the process, with emphasis on whether events are happening more (or less) frequently as time evolves. This is distinct from say, predicting the time when the next shock will strike and has implications in learning about the random phenomenon that drives the system. Policymakers might want to know whether financial institutions such as banks are failing more frequently over the last decade, or geologists might want to know whether a given volcano is getting more restless in recent times so that people living in the vicinity may be moved to safety. Affirmative answers to these questions often imply the existence of a time point that separates two neighboring sections of the process that are (stochastically) significantly different. Such a point is usually

termed a change point, and the prospect of estimating them has attracted much attention, especially since the 1950s. Expressed in such a tone, the problem is arguably, extremely general, meriting attacks from diverse fields: the quality control community (Lai (1995) [86]) wants to efficiently identify faults in their manufacturing process, biologists are curious to locate copy number variation in genomic data (Efron and Zhang (2011)) [38]), computer scientists want to guard corruption in networks (Tartakovsky, Rozovskii, Blazek, and Kim (2006) [143]), economists fit change point defined multiple regime models to financial data (Ross (2012) [124]).

Formally, let $\{X_1, X_2, ...\}$ be a bunch of variables shocked at unknown points in time $\{\tau_1, \tau_2, ...\}$, i.e. let

$$X_i \sim \begin{cases} F_0 & \text{if } i \leq \tau_1 \\ F_1 & \text{if } \tau_1 < i \leq \tau_2 \\ F_2 & \text{if } \tau_2 < i \leq \tau_3 \\ ... & \end{cases}$$

with F_i representing the cumulative distribution function (c.d.f) in the ith interval. The X_is for us, could be taken as the inter-event times described previously. Traditionally, independence between change points is assumed. For our purpose, that would make the repairable system a *renewal* one, a special case of the general class we would consider. Gustaffson (2000) [53], however, describes how the assumption of independence is not forbidding in view of the fact that one can smother any inherent dependence, by modeling the underlying autocorrelation, and then weeding out the change points from either the residuals or one-step-ahead prediction errors. Both of them, assuming an adequate fit, should generate independently and identically distributed (iid) samples. For our simulation exercises in Chapters 3 and 4, the very nature of simulating the PtP guarantees such iid-ness.

Uncertainty about, and ignorance of F_is, be it the underlying parameters or even the functional structure, have generated a host of change detection algorithms, each tailored to a specific environment. In the face of extremely limited information on the cdf, a general Change Point Model (CPM) framework has been recently proposed, which houses a wide array of tools and statistics and implements both parametric (Student-t, Bartlett, GLR, Fisher's exact test, Exponential) and non-parametric (Mann-Whitney, Mood, Lepage, Kolmogorov-Smirnov, and Cramer-von-Mises) methods. Details may be had from Hawkins et al. (2003) [57]; Hawkins and Zamba (2005a) [58]; Zhou, Zou, Zhang, and Wang (2009) [148]; Hawkins and Deng (2010) [59]; Ross et al. (2011) [125]; Ross and Adams 2011, 2012 [126] [127]). Originated by Hawkins et al. (2003) [57] to detect changes in the average level of Normal flow, the CPM framework has been subsequently generalized to embrace a more intricate brand of changes (Zou and Tsung (2010) [149]).

Viewed broadly, change point identification problems present themselves in two distinct flavors: batch and sequential, sometimes termed Phase I and Phase II detection, which are essentially, retrospective and prospective approaches, respectively.

i) *Batch approach:* Given a fixed length sequence $\{X_1, X_2, ..., X_n\}$, the goal is using offline methods to check whether a change point exists, using the entire data set, i.e. both the sets before and *after* the probable change point. This approach weeds out anomalies with remarkable precision as long as there are not too many of them, in which case it turns out to be computationally forbidding. Heuristics are often used in the presence of a large number of change points (Inclan and Tiao (1994) [73], Hawkins (2001)[60]). Likelihood ratio testing (Hinkley and Hinkley (1970) [64]) and Bayesian inference (Stephens (1994) [137]) are popular choices with this approach. We shall work mainly under this paradigm in Chapter

3.

ii) *Sequential approach*: Here, the length of the incoming sequence varies with time. A decision about change is made at every incoming instance, and the method is entirely online. Only the past observations are used, and this proves more effective in the presence of a large number of change points. Control chart inspired Cumulative Sum (CUSUM) techniques (Page (1954) [111]), Exponential weighted moving averages (Roberts (1959) [123]), or sequential Bayesian methods (Chib (1998) [26]; Fearnhead and Liu (2007) [46]) are popular choices. We shall describe how our methods are similar to and different from this sequential approach in the latter part of Chapter 3.

The statistical software package R nests several change point algorithms for ready implementation. The `CPM` package (Ross (2015) [127]) exploits the CPM framework described previously with one-dimensional random quantities. Other packages include `bcp` (Erdman and Emerson (2007) [45]), `strucchange` (Zeileis, Leisch, Hornik, and Kleiber (2002) [147]) and `changepoint` (Killick and Eckley (2014) [79]).

1.3 Examples

There exists a class of phenomenon, both natural and artificial, for which a PtP interpretation seems most natural and there exist others that demand considerable insight and mathematical mastery to be interpreted that way. The time now, is ripe, to get introduced to the major cases that will be analyzed in considerable depth throughout the dissertation. They will serve as continuous running examples, without elaborate reference each time. A passing glance will also be cast on a few other possibilities, to be studied in future works, classified under "Other examples".

1.3.1 Mt. Etna

The largest of the three currently active Italian volcanoes, Mt. Etna, towering a staggering 3,329m, is a stratovolcano nested between the cities of Messina and Catania (Fig (1.1), taken from `http://www.konbini.com/us/inspiration/mount-etna-eruption-creates\-spectacular-fireworks-display/`). Its connection to Greek mythology, size, repute (often dubious), evidenced by its winning a place in the list of UNESCO world heritage sites, and proximity to human habitation have attracted scientists' attention and tourists' curiosity.

Figure 1.1: Mt. Etna at its full fury.

This work treats every eruption as an event from a PtP. Discrete-valued time series out of its eruption dates will be constructed in Chapter 4, when the need arises. Etna's eruption history, dating back to the mid-seventeenth century is preserved in several records. Mulargia et al. (1985) [108] and Mulargia et al. (1987) [109] store the eruption dates from 1607 to 1978, using which, Ho (1992) [65] created control chart type figures. Smethurst et al. (2009) [135] however, provides a more recent update, tracking the volcano till 2008. Chen (2010) [24] records and uses this data set for regime identification, and so shall we.

1.3.2 Kilauea and Mauna Loa

Our volcanic fascination next took us to Hawaii, where two active shield volcanoes, Kilauea and Mauna Loa are close neighbors, and take turns to torment villagers living nearby (`http://www.bbc.com/news/world-us-canada-29805102`). Geologists like Lipman (1980) [99] debate on a possible inverse relationship between the two, implying when one is active, the other is relatively dormant. Klein (1982) [81], through tests for non-randomness, suspects that the longest recorded repose time of one is associated with an intensified activity of the other. Ho and Bhaduri (2017) [70] surveys other studies involving the two protagonists.

Figure 1.2: Kilauea and Mauna Loa in close proximity.

Figure (1.2) has been taken from `http://www.photovolcanica.com/\VolcanoInfo/Kilauea/Kilauea.html`. The U.S. Geological Survey (USGS) and the Smithsonian Institution's Global Volcanism Programme at `http://www.usgs.gov/` and `http://www.volcano.si.edu/` respectively, offer excellent records of eruptions of both, including the dates, the duration, the amount of lava ejected, the area affected, etc. We have studied the period from 1750 to 1985, primarily due to the reliability of the geophysical methods generating

the information. The eruption counts for Kilauea and Mauna Loa in this 236-year duration were 63 and 40, respectively. The time series storing these counts can be found in Ho and Bhaduri (2017) [70].

With a volcanic case already being studied, another one might seem unnecessary. We must point out, however, that in the case of Mt. Etna, our goal would be to study the volcano to check whether there exists a time point separating the frequency of eruptions regardless of outside knowledge. A PtP (a *Poisson* process, to be exact) approach will be used here. On the other hand, in the latter case, we will be concerned with the *interaction* (i.e., when such outside information is available) between the two volcanoes through a graphic tool, to be proposed later. A time series approach will be used here.

Statistical tools have been used for regime identification in volcanic examples by several scholars. Wickman (1966) [145] studies a class of volcanoes termed "simple Poissonian volcanoes", for which the recurrence eruption rates are independent of time. Mulargia et al. (1987) [109] for instance, uses large values of the Kolmogorov-Smirnov (KS) statistic on all possible partition pairs to locate possible changes. In Chapter 3, we will analyze how this statistic and several others compete against our proposals.

1.3.3 Hurricane counts

The term "tropical cyclone" is generic and embraces all types of closed atmospheric circulation that forms over a tropical or subtropical ocean. If the maximum sustained wind speed exceeds 74 miles per hour, these storms are called hurricanes in the Atlantic ocean, typhoons in the Pacific and cyclones elsewhere. National Oceanic and Atmospheric Administration (NOAA) is a government organization under the United States Department of Commerce

and their Historical Hurricane Tracks webpage at `http://coast.noaa.gov/hurricanes/?redirect=301ocm#` records most of the recent storms that occur globally. Based on geographical criteria, the water mass of the earth has been partitioned into several basins such as West Atlantic, North Pacific, Gulf of Mexico, Southern Indian, Eastern Australian, etc. and data are available on the speeds, dates, and duration of storms originating in each of these basins. Additionally, based on the strength of the storms judged by the maximum sustained wind speeds (MSW), we have six major categories

Table 1.1: NOAA hurricane classification based on maximum wind speeds attained

Category	MSW
Hurricane 5	$>$ 135 kts
Hurricane 4	114-135 kts
Hurricane 3	96-113 kts
Hurricane 2	83-95 kts
Hurricane 1	64-82 kts
Trop/Subtrop	34-63 kts

and some of the records date back to 1851. But the earlier records are mostly based on eyewitnesss accounts and other less reliable methods, and hence after consultation with experts well-versed with the data collection method, we finalized on 1923 – 2013 as our observation period. We have done our preliminary analyses on storms originating in the West Atlantic basin, mainly because of its proximity to the US East Coast which has to face the wrath of these natural calamities almost every year and often with grave consequences, but also because of the fact that this basin is well studied by oceanographers and climatologists and hence would render us a chance to compare our findings to their beliefs. Similar analyses can, of course, be done on other basins as well. The start date of each tropical cyclone has been treated as a shock time from an evolving PtP.

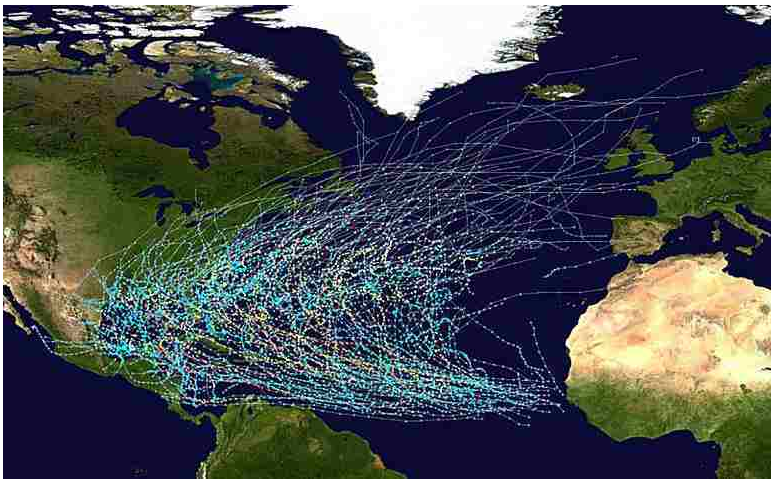


Figure 1.3: Tracks of all tropical cyclones in the West Atlantic basin during 1980 - 2005

Fig 1.3 has been taken from `https://www.thoughtco.com/global-hurricane-basins-3443941`. Emanuel (2003, 2006, 2007) [40] [41] [42] believes that it is the category 3, 4 and 5 hurricanes that cause the most damage and so, entirely for the sake of a simplified analysis, one may define these as the *strong* group of hurricanes. H2 and H1 constitute the *weak* class of hurricanes and the *tropical* category is reserved for the final class. It may be found that over the period under consideration, there has been 32 H5 storms, 84 H4 storms, 87 H3 storms, 93 H2 storms, 150 H1 storms, 271 tropical and 24 subtropical storms. Emanuel, in a candid interview with the Discovery channel `http://news.discovery.com/earth/global-warming/does-climate-change-mean-more-or-stronger-hurricanes-120907.htm` expressed that with a continually warming climate, it is increasingly difficult to start a devastating hurricane due to a hike in saturation deficit which works against its creation, but if it gets started somehow, it has the potential to become more intense. Thus, the total number of storms should decline globally, but the proportion of hurricanes which are intense should rise. For

the present purpose, we will focus only on the more interesting "Strong-Weak" interaction.

1.3.4 Dow Jones Industrial Average

Developed in 1896 by *Wall Street Journal* editor and Dow Jones and Co. co-founder Charles Dow, the Dow Jones Industrial Average (DJIA), sometimes shortened to DOW, is a stock market index that is often regarded as an adequate measure of the health of the US economy. It represents a price-weighted average of the stocks of 30 large industrial companies traded on the New York Stock Exchange (NYSE) and the NASDAQ. High values of the index represent a bull market, while low values represent a bear market.

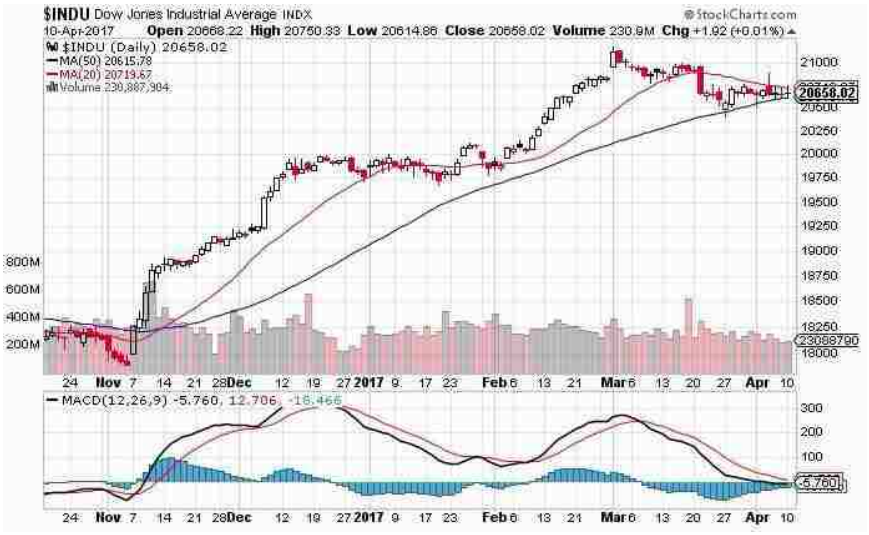


Figure 1.4: Whimsicality of the Dow Jones Industrial Average from a period in 2017

Figure (1.4) above has been taken from

`https://amigobulls.com/articles/2017-04-11-dow-jones-amazon-stock-united-continental-and-netflix-technical-analysis-trading-ideas-for-today`. The times of achieving a closing milestone may be had from several freely available sources (such as

`https://en.m.wikipedia.org/wiki/Closing_milestones_of_the_Dow_Jones_Industrial_`

`Average`) and we have used the closing times between February 1885 and June 2017 in our regime analyses to follow, accumulating a total of 283 observations.

1.3.5 Other examples

PtPs might emerge from mundane constructs, too. Rigdon and Basu (2000) [122], for example, records 18 failure times of a photocopier. Time here was measured in terms of the number of copies made. Maguire et al. (1952) [102] keeps a historical record of the dates of mining accidents in Great Britain between December 6, 1875, and May 29, 1951, claiming the lives of 10 men or more. Jelisnski and Moranda (1972) [76] gives 34 failure times of a software system. Mooley (1981) [107] records the times of cyclonic storms striking the coast of Bay of Bengal during 1877 – 1977. Bakun et al. (2005) [12] studies major (more than 6 on the Richter scale) earthquakes in Parkfield, California, since January 9, 1857. Duane (1964) [36] explores the possibility of fitting a special class of PtPs to failure patterns of several systems developed by a company called General Electric. 272 eruption times of the Old Faithful geyser in Yellowstone National Park may be found as "faithful" within the `datasets` package in R. Information on earthquakes originating in Ogata, North China Sea, and Phuket may be found within the `PtProcess` package in R.

Chapter 2

Time reversal

Reviewing the workings of a statistic often used to detect non-stationarity, this chapter will primarily meditate on the exercise that provides the fundamental spark to propel this research: switching the flow of time from the left to the right to the right to the left. The first section will recall the established test (henceforth known as the forward test Z) and touch upon its optimality, while the next section will introduce its time-reversed counterpart, christened the backward test Z_B. Section 3 will spell out our simulation strategy essential for the comparisons to follow. The last section will summarize their power performances under diverse intensity environments and will investigate favorable ramifications that ensue.

2.1 Poisson processes

The occurrences of events that can be conveniently modeled by PtPs are plentiful in nature. For instance, the examples explored in the previous introductory chapter about volcanic eruptions, hurricane counts, bank failures, strong sandstorms or earthquakes all fall in this category and under the mild regularity conditions detailed below, a *Poisson process* can offer valuable insights into the dynamics of the inherent randomness.

The counting process $\{N(t)\}_{t>0}$ will be considered Poisson if:

1. $N(0) = 0$

2. For any $a < b \leq c < d$, the variables $N(a, b]$ and $N(c, d]$ are independent.

3. There exists a function $\lambda(.)$, called the intensity of the process, such that:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t] = 1)}{\Delta t}.$$

4. Two or more failures can never happen together:

$$\lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t] \geq 2)}{\Delta t} = 0.$$

These four properties are enough to show (Rigdon and Basu (2000) [122]) that

$$N(t) \sim Pois\left(\int_0^t \lambda(x)dx\right) \text{ for all } t > 0 \tag{2.1}$$

i.e.

$$P[N(t) = n] = \exp\left(-\int_0^t \lambda(x)dx\right) \frac{\{\int_0^t \lambda(x)dx\}^n}{n!} \text{ for all } t > 0,\ n = 0, 1, 2, ... \tag{2.2}$$

Using known facts about Poisson means, it thus follows that

$$\Lambda(t) = E(N(t)) = \int_0^t \lambda(x)dx \text{ for all } t > 0. \tag{2.3}$$

Researchers often find it convenient to specialize further and impose parametric forms on the intensity of the process. For instance, adopting the following structure:

$$\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1} \tag{2.4}$$

leads to what is known as a Power law process (previously termed as the Weibull process, owing to the functional similarity to the hazard function from a non-repairable system modeled by the Weibull density) and it lends modelers the ability to nest the necessary

inferences within a mathematically tractable framework. The intensity function is flexible enough to model repairable systems which are improving (corresponding to $\beta < 1$), deteriorating ($\beta > 1$) or remaining homogeneous ($\beta = 1$) with respect to time. The route to inference on these parameters, however, differ slightly depending on the actual sampling scheme employed:

i) Failure truncated case: Under this framework, the process is continually monitored until a predetermined number, say n, of events occur. This number thus, is deterministic and the time of the last (in fact all the) occurrence T_n, is uncertain. Conditioned on this random variable, $T_1, T_2, ..., T_{n-1}$ can be shown (Rigdon and Basu (2000) [122]) to be distributed as order statistics from a uniform $(0, t_n)$ distribution, assuming the underlying process is homogeneous, and the joint density under a general intensity $\lambda(.)$ may be expressed as

$$f(t_1, t_2, ..., t_n) = \left(\prod_{i=1}^{n} \lambda(t_i)\right) \exp\left(-\int_0^{t_n} \lambda(x)dx\right), \quad 0 < t_1 < t_2 < ... < t_n. \tag{2.5}$$

Under the Power law choice of $\lambda(.)$ given in (2.4), this simplifies to

$$f(t_1, t_2, ..., t_n) = \frac{\beta^n}{\theta^{n\beta}} \left(\prod_{i=1}^{n} t_i\right)^{\beta-1} \exp\left(-\left(\frac{t_n}{\theta}\right)^{\beta}\right), \quad 0 < t_1 < t_2 < ... < t_n. \tag{2.6}$$

and maximum likelihood estimates of the parameters can then be obtained as

$$\hat{\theta} = \frac{t_n}{n^{\frac{1}{\hat{\beta}}}}, \ \hat{\beta} = \frac{n}{\sum_{i=1}^{n} \log\left(\frac{t_n}{t_i}\right)}. \tag{2.7}$$

ii) Time truncated case: Under this scheme, a random number of failures, say N, are observed and recorded till a predetermined time t. Unless $N = 0$, the m.l.e's here are given by:

$$\hat{\theta} = \frac{t}{N^{\frac{1}{\hat{\beta}}}}, \ \hat{\beta} = \frac{N}{\sum_{i=1}^{n} \log\left(\frac{t_n}{t_i}\right)} \tag{2.8}$$

These two sampling schemes are similar to Type - II and Type - I censoring in survival analysis.

2.1.1 A critique on the Forward Test

Once the point estimates are available, a natural question would be to inquire about the possibility of carrying out hypotheses tests or creating confidence intervals, especially for the β parameter because of the crucial role it plays in estimating the failure trend. Rigdon and Basu (2000) [122] show that the quantity $Z = \frac{2n\beta}{\hat{\beta}}$ is pivotal, having a chi-square distribution and can be profitably exploited to carry out these other aspects of statistical inference. In keeping with our eventual goal of working on a time-reversed version of Z, formulated by Ho (1993) [66], let us agree to term the general category of tests using this form as "forward tests" and the remainder of this section shall be devoted to a thorough and careful analysis of the merits of this class.

The rich history of the test under both sampling schemes deserves mention: Assuming the power law intensity to be valid, Crow (1974, 1982) [31], [33] respectively, develops tests for β with θ as a nuisance parameter and comes up with small sample and asymptotic confidence intervals on the mean time between failures through a novel application of the Z statistic. Finkelstein (1976) [48] chooses to concentrate on the parameter θ instead, argues that $(\frac{\hat{\theta}}{\theta})^{\beta}$ is independent of β and θ and eventually comes up with computer simulated confidence intervals for θ under the failure truncated scheme. Lee and Lee (1978) [88] demonstrates how such intervals can be constructed using numerical integration too. Bain and Engelhardt (1980) treat θ as a nuisance parameter and proves that the forward test will be uniformly most powerful unbiased (UMPU) for β by using joint complete sufficiency of N and $\hat{\beta}$ and a

theorem from Lehmann (1959, p136) [90] under the time truncated case and also constructs approximate tests and confidence intervals for θ under a similar framework.

The null and the alternate distribution

Although Rigdon and Basu (2000) [122] focus on testing whether the failure process is homogeneous (i.e. $\beta = 1$) which is what we shall do ultimately, too, for the moment, we will seek generalization and will concern ourselves with tests such as:

$$H_0 : \beta = \beta_0 \; vs \; H_a : \beta > \beta_0 \tag{2.9}$$

$$H_0 : \beta = \beta_0 \; vs \; H_a : \beta < \beta_0 \tag{2.10}$$

$$H_0 : \beta = \beta_0 \; vs \; H_a : \beta \neq \beta_0 \tag{2.11}$$

under both the failure- and time-truncated cases. Throughout the rest of this section, unless otherwise explicitly mentioned, these will be our usual greater than, less than and two tailed alternatives.

Failure truncated case

Rigdon and Basu (2000) [122] exhibit a method for constructing the probability density for the quantity $\frac{2n\beta}{\hat{\beta}}$ and owing to its immense importance in connection to the forward test, we shall first review the proof:

Theorem 2.1. *[The density for $\frac{2n\beta}{\hat{\beta}}$] For the failure truncated case, $\frac{2n\beta}{\hat{\beta}} \sim \chi^2_{2n-2}$.*

Proof. Under this framework and conditioned on the last event time $T_n = t_n$, the occurrence times $T_1 < T_2 < ... < T_{n-1}$ are distributed as $n-1$ order statistics from a distribution with

cumulative distribution function:

$$G_y = \begin{cases} 0 & y \leq 0, \\ \Lambda(y)/\Lambda(t_n) & 0 < y < t_n, \\ 1 & y \geq t_n \end{cases}$$

where $\Lambda(t) := \int_0^t \lambda(x)dx$ is the cumulative intensity function. For the power-law choice, the above boils down to

$$G_y = \begin{cases} 0 & y \leq 0, \\ (y/t_n)^\beta & 0 < y < t_n, \\ 1 & y \geq t_n. \end{cases}$$

Defining Y to be the random variable with cdf $G(.)$, we must have

$$P(Y \leq y) = G(y) = \left(\frac{y}{t_n}\right)^\beta \quad 0 < y < t_n. \tag{2.12}$$

Simultaneously,

$$P(Y \leq y) = P(Y/t_n \leq y/t_n) = P((Y/t_n)^\beta \leq (y/t_n)^\beta). \tag{2.13}$$

Comparing (2.12) and (2.13) above:

$$P((Y/t_n)^\beta \leq (y/t_n)^\beta) = \left(\frac{y}{t_n}\right)^\beta \quad 0 < y < t_n, \tag{2.14}$$

which shows that $(Y/t_n)^\beta$ is uniformly distributed on (0,1). Thus, $(T_i/t_n)^\beta \quad i = 1, 2, .., n-1$ are distributed as $n-1$ order statistics from a uniform (0,1) density. Using the fact that if U has a uniform (0,1) distribution, then $X = -\theta \log U$ has an exponential density with mean θ, we can claim that the sum

$$\sum_{i=1}^{n} -\log(t_i/t_n)^\beta = -\beta \sum_{i=1}^{n} \log(t_i/t_n)$$

is distributed as the sum of $n-1$ exponential variables, each with mean 1. Using the reproductive property of gamma, the distribution above has nothing but a gamma$(n-1, 1)$

density. Finally, since twice a gamma$(n-1,1)$ density generates a chi-square density with $2n-2$ degrees of freedom, we have

$$-2\beta\sum_{i=1}^{n}\log(t_i/t_n)=\frac{2n\beta}{\hat{\beta}}\sim\chi^2(2n-2).$$

□

Once it is shown that $\frac{2n\beta}{\hat{\beta}}\sim\chi^2(2n-2)$, the general form of the test statistic that we are going to use for any of the tests shown above is:

$$Z=\frac{2n\beta_0}{\hat{\beta}} \tag{2.15}$$

and it is imperative to find distributions of Z under both the null and the alternative hypotheses to calculate error probabilities or power. Under the null hypothesis, $Z\sim\chi^2(2n-2)$, but under the alternative, $\frac{2n\beta}{\hat{\beta}}\sim\chi^2(2n-2)$ and the distribution of Z can be obtained through a traditional change of variable technique as follows:

$Z=\frac{2n\beta_0}{\hat{\beta}}=\frac{\beta_0}{\beta}\frac{2n\beta}{\hat{\beta}}=\frac{\beta_0}{\beta}X$ (say) where $X\sim\chi^2(2n-2)$.

$$Z\ =\frac{\beta_0}{\beta}X\Rightarrow X=\frac{\beta}{\beta_0}Z=g^{-1}(Z)$$

$$\Rightarrow\frac{d}{dz}g^{-1}(z)=\frac{\beta}{\beta_0}$$

Both variables are supported on the positive half of the real line and since

$$f_X(x)=\frac{1}{\Gamma(n-1)2^{n-1}}x^{n-2}e^{-x/2},\quad x>0, \tag{2.16}$$

$$\begin{aligned}f_Z(z)&=f_X(g^{-1}(z))\left|\frac{d}{dz}g^{-1}(z)\right|\\&=\frac{1}{\Gamma(n-1)2^{n-1}}(\frac{\beta}{\beta_0}z)^{n-2}e^{-\beta z/2\beta_0}\frac{\beta}{\beta_0}\\&=\left(\frac{\beta}{\beta_0}\right)^{n-1}\frac{1}{\Gamma(n-1)2^{n-1}}z^{n-2}e^{-\beta z/2\beta_0},\quad z>0.\end{aligned} \tag{2.17}$$

This technique of one – one transformation can be used to derive the probability distribution of $\hat{\beta}$ too, by noting that setting $X = \frac{2n\beta}{\hat{\beta}}$, we have $Y = \hat{\beta} = \frac{2n\beta}{X} = g(X)$ which implies $\frac{d}{dy}g^{-1}(y) = -2n\beta\frac{1}{y^2}$. Thus, using the form of the density of X shown above, we have, upon simplifications

$$f_{\hat{\beta}}(\hat{\beta}) = \frac{1}{\Gamma(n-1)}\frac{(n\beta)^{n-1}}{\hat{\beta}^n}e^{-n\beta/\hat{\beta}}, \quad \hat{\beta} > 0. \tag{2.18}$$

The form of this density will be used in a later section in connection to optimality of the forward test, but presently we focus on extracting moments from this density

$$E(\hat{\beta}) = \int_0^\infty \frac{1}{\Gamma(n-1)}\frac{(n\beta)^{n-1}}{t^{n-1}}e^{-n\beta/t}dt.$$

Setting $\frac{\beta}{t} = x$ we have $dt = -\frac{\beta}{x^2}dx$ and thus, the above expectation reduces to

$$\begin{aligned} E(\hat{\beta}) &= \int_\infty^0 \frac{1}{\Gamma(n-1)}n^{n-1}x^{n-1}e^{-nx}\left(-\frac{\beta}{x^2}\right)dx \\ &= \frac{n^{n-1}\beta}{\Gamma(n-1)}\int_0^\infty e^{-nx}x^{n-3}dx \\ &= \frac{n^{n-1}\beta}{\Gamma(n-1)}\frac{\Gamma(n-2)}{n^{n-2}} \\ &= \frac{n\beta}{(n-2)} \end{aligned} \tag{2.19}$$

which provides a proof for the claim that $\frac{n-2}{n}\hat{\beta}$ is unbiased for β (Rigdon and Basu (2000) [122]).

Time truncated case

Analogous results can be obtained for the time truncated case by noting that the quantity $\frac{2n\beta}{\hat{\beta}}$ now follows a $\chi^2(2n)$ distribution (Rigdon and Basu (2000) [122]). Thus the null distribution of $Z = \frac{2n\beta_0}{\hat{\beta}}$ is $\chi^2(2n)$ and the alternate distribution is

$$f_Z(z) = \left(\frac{\beta}{\beta_0}\right)^n \frac{1}{\Gamma(n)2^n}z^{n-1}e^{-\beta z/2\beta_0}, \quad z > 0. \tag{2.20}$$

The null and the alternate distributions of the forward statistic are compared in Fig (2.1).

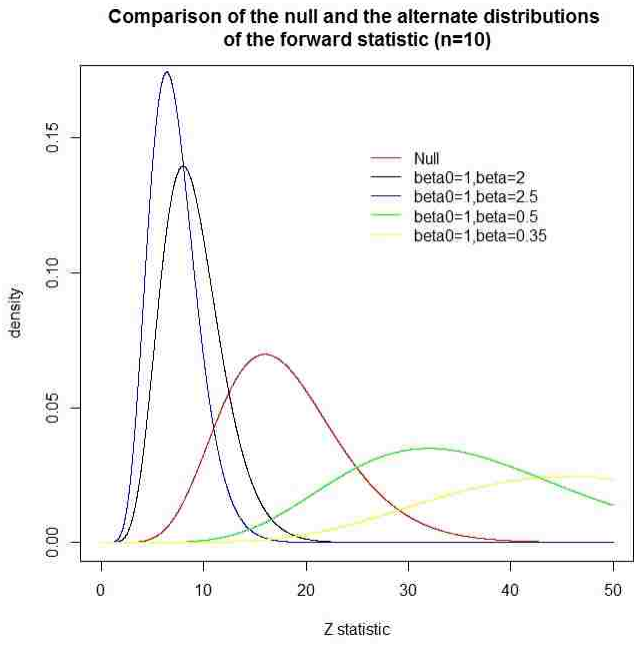


Figure 2.1: Distributions of the forward statistic corresponding to different parameter values

2.1.2 The critical regions

It is worthwhile to note that $\hat{\beta}$ tries to estimate β, and hence, under the "greater than" type alternative, large values of $\hat{\beta}$ will point to the plausibility of the alternative. Put in another way and owing to its very construction, small values of the statistic $\frac{2n\beta_0}{\hat{\beta}}$ will lend credence to the rejection of the null hypothesis. To achieve level α, the rejection region of the one tailed "greater than" type alternative under the failure truncated case will thus be

$$\left\{ z : z = \frac{2n\beta_0}{\hat{\beta}} < \chi^2_{1-\alpha}(2n-2) \right\} \tag{2.21}$$

where $\chi^2_{1-\alpha}(2n-2)$ is the lower α point of a chi-square distribution with $2n-2$ degrees of freedom. In terms of sets, the above region is, of course, equivalent to

$$\left\{ \hat{\beta} : \hat{\beta} > \frac{2n\beta_0}{\chi^2_{1-\alpha}(2n-2)} \right\}, \tag{2.22}$$

and similar ideas carry over to other types of alternatives as well. The next table summarizes the critical regions under the different cases.

Table 2.1: Critical regions under different alternatives

	$H_a : \beta > \beta_0$	$H_a : \beta < \beta_0$	$H_a : \beta \neq \beta_0$
Failure truncated	$Z < \chi^2_{1-\alpha}(2n-2)$	$Z > \chi^2_{\alpha}(2n-2)$	$Z < \chi^2_{1-\alpha/2}(2n-2)$ or $Z > \chi^2_{\alpha/2}(2n-2)$
Time truncated	$Z < \chi^2_{1-\alpha}(2n)$	$Z > \chi^2_{\alpha}(2n)$	$Z < \chi^2_{1-\alpha/2}(2n)$ or $Z > \chi^2_{\alpha/2}(2n)$

Confidence intervals for β can also be constructed by inverting the acceptance region of the two-tailed hypothesis.

2.1.3 The power functions

In connection to testing a statistical hypothesis, the power of a test is traditionally defined as the probability of making a correct decision in general and in particular, the probability of rejecting H_0 when H_0 is false. Power functions are often used to judge the quality of a test and comment on its optimal properties. Since the alternate distribution of the forward statistic has already been derived previously, we can proceed to examine the form of the power function under different alternatives.

For a "greater than" type alternative, we might recall that the critical region (w.r.t. Z) is left sided (i.e. we reject the null for extremely small values of Z). Thus adopting the notation scheme: $\pi_{X,Y}(\beta)$ for the power function under the Xth type alternative and the Yth type sampling scheme, the power function in this case will be

$$\begin{aligned} \pi_{G,F}(\beta) &= P_{H_a}(Z < \chi^2_{1-\alpha}(2n-2)) \\ &= \int_0^{\chi^2_{1-\alpha}(2n-2)} \left(\frac{\beta}{\beta_0}\right)^{n-1} \frac{1}{\Gamma(n-1)2^{n-1}} z^{n-2} e^{-\beta z/2\beta_0}\, dz. \end{aligned} \tag{2.23}$$

Renaming $\frac{\beta z}{\beta_0} = t, \Rightarrow dt = \frac{\beta}{\beta_0} dz$ and on simplifications,

$$\begin{aligned} \pi_{G,F}(\beta) &= \int_0^{\beta/\beta_0 \chi^2_{1-\alpha}(2n-2)} \frac{1}{\Gamma(n-1)2^{n-1}} t^{n-2} e^{-t/2} \, dt \\ &= \Psi_{2n-2}\left(\frac{\beta}{\beta_0}\chi^2_{1-\alpha}(2n-2)\right), \end{aligned} \tag{2.24}$$

where $\Psi_m(.)$ represents the cumulative distribution function of a chi-square distribution with m degrees of freedom, the closed form of which is difficult to explore analytically. However, software packages such as R routinely calculate these cumulative probabilities.

A more elegant way of deriving the expression should bypass the actual density and would exploit the pivotal property of the quantity $\frac{2n\beta}{\hat{\beta}}$ as follows:

$$\begin{aligned} \pi_{G,F}(\beta) &= P_{H_a}(Z < \chi^2_{1-\alpha}(2n-2)) \\ &= P_{H_a}\left(\frac{2n\beta_0}{\hat{\beta}} < \chi^2_{1-\alpha}(2n-2)\right) \\ &= P_{H_a}\left(\frac{2n\beta}{\hat{\beta}} < \frac{\beta}{\beta_0}\chi^2_{1-\alpha}(2n-2)\right). \end{aligned}$$

But under the alternative H_a, $\frac{2n\beta}{\hat{\beta}} \sim \chi^2(2n-2)$, and thus

$$\pi_{G,F}(\beta) = \Psi_{2n-2}\left(\frac{\beta}{\beta_0}\chi^2_{1-\alpha}(2n-2)\right), \tag{2.25}$$

as derived previously. Arguing along similar lines, the power functions under the less than type and the two tailed alternatives take on the following forms:

$$\pi_{L,F}(\beta) = 1 - \Psi_{2n-2}\left(\frac{\beta}{\beta_0}\chi^2_{\alpha}(2n-2)\right), \tag{2.26}$$

$$\pi_{T,F}(\beta) = \Psi_{2n-2}\left(\frac{\beta}{\beta_0}\chi^2_{1-\alpha/2}(2n-2)\right) + 1 - \Psi_{2n-2}\left(\frac{\beta}{\beta_0}\chi^2_{\alpha/2}(2n-2)\right). \tag{2.27}$$

Similar expressions can be had for the time truncated case too:

$$\pi_{G,T}(\beta) = \Psi_{2n}\left(\frac{\beta}{\beta_0}\chi^2_{1-\alpha}(2n)\right), \tag{2.28}$$

$$\pi_{L,T}(\beta) = 1 - \Psi_{2n}\left(\frac{\beta}{\beta_0}\chi^2_{\alpha}(2n)\right), \tag{2.29}$$

$$\pi_{T,T}(\beta) = \Psi_{2n}\left(\frac{\beta}{\beta_0}\chi^2_{1-\alpha/2}(2n)\right) + 1 - \Psi_{2n}\left(\frac{\beta}{\beta_0}\chi^2_{\alpha/2}(2n)\right). \tag{2.30}$$

These power curves are graphed below for different choices of the sample size n.

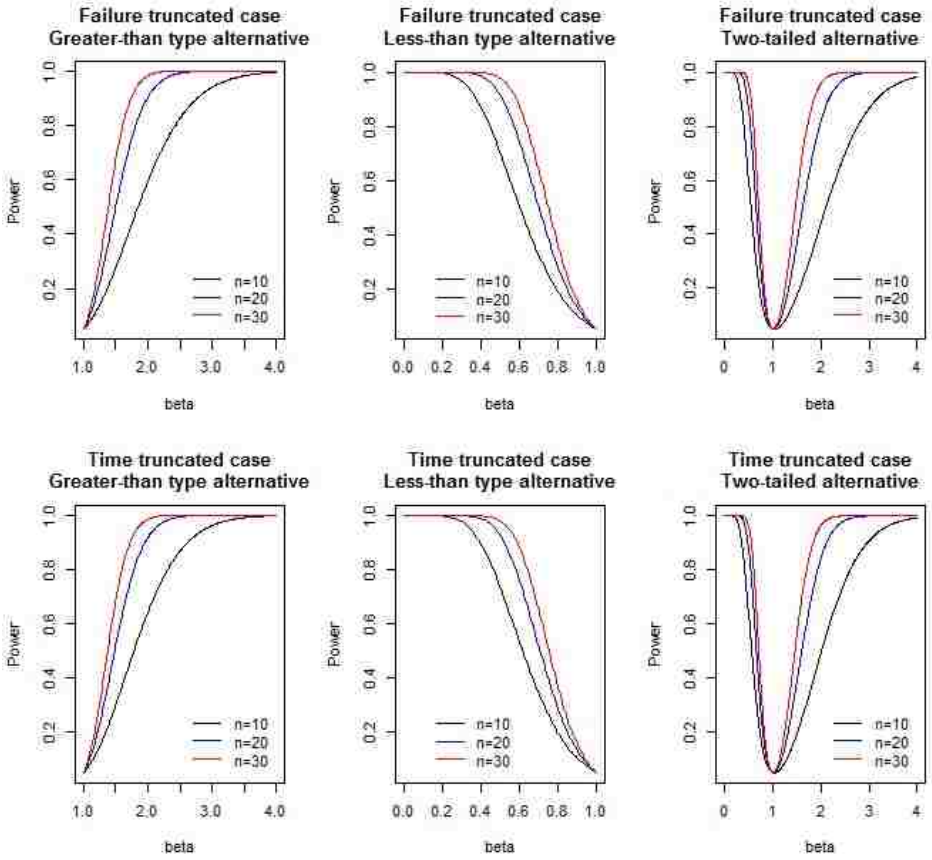


Figure 2.2: Power curves under different types of alternatives at $\alpha = 0.05$.

As expected, a large sample helps one pick the correct alternative relatively more easily.

2.1.4 A note on optimality

Upon the formulation of a statistical test, it is always instructive to discover additional properties enjoyed by it. Just as good point estimators often turn out to be sufficient, unbiased or have the minimum variance among others, properties such as most powerfulness

(MP-ness, or better still: uniformly most powerfulness, UMP-ness) or unbiasedness typically adore a statistical test. These properties date way back into the labyrinths of statistical folklore, and traditional texts such as Lehmann and Romano (2005) [90], Casella and Berger (2002) [23] among others give an excellent account of such properties. In this section, we shall endeavor to prove that the forward test is fairly optimal in more ways than one.

In particular, we will focus on the property of UMP-ness. Simply put, it says that the test considered consistently gives greater power in choosing the correct alternative compared to its competitors and at this moment, we can recollect the following notions which will be crucial to the proof:

i. Monotone-likelihood ratio (MLR): A family of densities $\{g(t|\theta) : \theta \in \Theta\}$ for a univariate random variable T with a real-valued parameter θ has a monotone likelihood ratio if $\forall \theta_2 > \theta_1$, the ratio $\frac{g(t|\theta_2)}{g(t|\theta_1)}$ is a monotone function of t on $\{t : g(t|\theta_1) > 0 \cup g(t|\theta_2) > 0\}$.

ii.

Theorem 2.2 (Karlin-Rubin)**.** *Consider testing $H_0 : \theta \leq \theta_0$ vs $H_a : \theta > \theta_0$. Suppose T is a sufficient statistic for θ and the family of densities $\{g(t|\theta) : \theta \in \Theta\}$ of T has a MLR. Then for any t_0, the test that rejects H_0 if and only if $T > t_0$ is a UMP level α test where $\alpha = P_{\theta_0}(T > t_0)$.*

We are now in a position to claim the optimality for the forward test.

Theorem 2.3 (UMP-ness of the forward test)**.** *For testing $H_0 : \beta = \beta_0$ vs $H_a : \beta > \beta_0$ the forward test using the Z statistic is conditionally UMP.*

Proof. We noted that in terms of the critical region generated under this alternative, namely $\left\{z : z = \frac{2n\beta_0}{\hat{\beta}} < \chi^2_{1-\alpha}(2n-2)\right\}$, the forward test can be equivalently referred to as one that

rejects H_0 if $\left\{\hat{\beta} : \hat{\beta} > \frac{2n\beta_0}{\chi^2_{1-\alpha}(2n-2)}\right\}$ happens. We also observe that

$$\begin{aligned} Z = \frac{2n\beta}{\hat{\beta}} &= -2\beta \sum_{i=1}^{n-1} \log\left(\frac{t_i}{t_n}\right) = -2\sum_{i=1}^{n-1} \log\left(\frac{t_i}{t_n}\right)^{\beta} \\ \Rightarrow -\frac{Z}{2} &= \sum_{i=1}^{n-1} \log\left(\frac{t_i}{t_n}\right)^{\beta} \\ \Rightarrow e^{-Z/2} &= \prod_{i=1}^{n-1} \left(\frac{t_i}{t_n}\right)^{\beta}. \end{aligned} \tag{2.31}$$

The proof now proceeds in two steps: First, we note that $\hat{\beta}$ is sufficient for β. Towards achieving that end, we use a result from Rigdon and Basu (2000) [122] which says that if the failure times of a non-homogeneous Poisson process are $T_1 < T_2 < ... < T_n$, then conditioned on $T_n = t_n$, the random variables $T_1 < T_2 < ... < T_{n-1}$ are jointly distributed as

$$f(t_1, t_2, .., t_{n-1}|t_n) = (n-1)! \prod_{i=1}^{n-1} \frac{\lambda(t_i)}{\Lambda(t_n)}, \quad 0 < t_1 < t_2 < ... < t_n. \tag{2.32}$$

For the specific power-law choice of $\lambda(.)$ this reduces to:

$$f(t_1, t_2, .., t_{n-1}|t_n) = (n-1)! \prod_{i=1}^{n-1} \frac{\beta t_i^{\beta-1}}{t_n^{\beta}}, \quad 0 < t_1 < t_2 < ... < t_n. \tag{2.33}$$

On simplification and using relation (2.31), we have

$$\begin{aligned} f(t_1, t_2, .., t_{n-1}|t_n) &= (n-1)! \prod_{i=1}^{n-1} \frac{\beta t_i^{\beta-1}}{t_n^{\beta}} \\ &= (n-1)! \prod_{i=1}^{n-1} \beta \left(\frac{t_i}{t_n}\right)^{\beta} \frac{1}{t_i} \\ &= (n-1)! \beta^{n-1} \prod_{i=1}^{n-1} \left(\frac{t_i}{t_n}\right)^{\beta} \prod_{i=1}^{n-1} \frac{1}{t_i} \\ &= (n-1)! \beta^{n-1} e^{-z/2} \prod_{i=1}^{n-1} \frac{1}{t_i} \\ &= \{\beta^{n-1} e^{-n\beta/\hat{\beta}}\} \left\{(n-1)! \prod_{i=1}^{n-1} \frac{1}{t_i}\right\}. \end{aligned}$$

This final representation is in the classical form of the factorization lemma which enables one to see that $\hat{\beta}$ is sufficient for β. In the second part, we show that $\hat{\beta}$ has an MLR which follows from the density derived in (2.8). We noted previously that this is an inverse gamma $(n-1, n\beta)$ density given by

$$f_{\hat{\beta}}(\hat{\beta}) = \frac{1}{\Gamma(n-1)} \frac{(n\beta)^{n-1}}{\hat{\beta}^n} e^{-n\beta/\hat{\beta}}, \quad \hat{\beta} > 0. \tag{2.34}$$

Choosing $\beta_2 > \beta_1$, we form the ratio:

$$\frac{f(\hat{\beta}|\beta_2)}{f(\hat{\beta}|\beta_1)} = \frac{((n\beta_2)^{n-1}/\hat{\beta}^n)e^{-n\beta_2/\hat{\beta}}}{((n\beta_1)^{n-1}/\hat{\beta}^n)e^{-n\beta_1/\hat{\beta}}} = \left(\frac{\beta_2}{\beta_1}\right)^{n-1} e^{-\frac{n}{\hat{\beta}}(\beta_2-\beta_1)} \tag{2.35}$$

and note that it is a monotone function of $\hat{\beta}$. Alternatively, this also follows from the fact that the density (2.18) belongs to the exponential family. The proof now follows by an appeal to the Karlin-Rubin theorem. □

Similar conclusions can be had for the time-truncated sampling scheme or for the "less than" type alternative. As noted previously, the UMPU-ness of the forward test has been established by Bain and Engelhardt (1980) [6] under the time truncated scheme using joint sufficiency, and a case for the failure truncated situation has been put forward in Bain and Engelhardt (1991) [11]. The proof above, however, is more detailed and appreciable.

2.2 A primer on the backward test

Since the event times sufficiently characterize the essential features of a repairable system, a study about the reliability is synonymous to a study on the failure time trends. A careful investigation of the forward statistic Z and the m.l.e. $\hat{\beta}$ for β introduced in the previous

section reveals the following connection:

$$Z = \frac{2n\beta_0}{\hat{\beta}} = -2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_i}{t_n}\right). \tag{2.36}$$

This is crucial since it gives a description of the statistic Z in terms of the trend in the event times. For instance, if the system is deteriorating ($\beta > 1$), most of the $\frac{t_i}{t_n}$ values will tend to cluster around 1, leading to a small value of the forward statistic. Some light on the reliability pattern can be shed through the reversed trend too, i.e., through the sequence of values $1 - \frac{t_i}{t_n}$. Under a similar situation of process deterioration, for instance, there will be a greater concentration of $1 - \frac{t_i}{t_n}$ values around 0 which shall contribute to an inflated value of the quantity

$$Z_B = -2\beta_0 \sum_{i=1}^{n-1} \log\left(1 - \frac{t_i}{t_n}\right). \tag{2.37}$$

This statistic Z_B, based on the idea of reversed trend, shall be termed as the *backward statistic* and tests relying on it, *backward tests*. This class was originally introduced by Ho (1993) [66] in connection to providing tests which are more powerful than the forward or the much fabled Laplace's test (introduced later) under the assumption of rough (in particular, step-like) alternate intensities. Understandably enough, unlike the forward test, history is relatively silent about Z_B. We shall explore this class of rough alternatives later in this chapter and with a greater degree of generality in Chapter 3, but for the time being, shall content ourselves in charting a course for the backward (i.e Z_B) analysis similar to the one for the forward case detailed in the last section, under the traditional smooth Power law intensities. Without making assumptions on the underlying intensity, Ho (1993) [66] analyzed an artificial data set to demonstrate how time switching could influence inferential conclusions. On a technical note, it might be worthy of mention that although not extremely

serious in (2.36), in (2.37), the sum must necessarily run till $n-1$ to keep itself from exploding.

2.2.1 An exact relationship connecting the two versions

Prior to delving deep into the labyrinths of statistical inference, we pause for a while and toy with the possibility of discovering a relationship, preferably in the form of an equation $Z_B = \phi(Z)$, that should string the two versions together. Owing to the similarity in structure between (2.36) and (2.37), one should naturally question the existence of a dependence pattern between the two. A use of the identity $\log(xy) = \log(x) + \log(y)$ provides an answer in the affirmative.

Theorem 2.4. *Irrespective of the actual sampling scheme employed, the forward and backward test statistics Z and Z_B are related through:*

$$Z_B = -2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right) + Z. \tag{2.38}$$

Proof. (Method 1) We start with the observation that $\frac{t_i}{t_n}$ admits of the following representation:

$$\frac{t_i}{t_n} = \frac{t_i}{t_n - t_i}\left(1 - \frac{t_i}{t_n}\right) = \frac{1}{\frac{t_n}{t_i} - 1}\left(1 - \frac{t_i}{t_n}\right)$$

(since pivotal to connecting Z and Z_B, would be the creation of a relation connecting their essential "kernels" as : $\frac{t_i}{t_n} = x(1 - \frac{t_i}{t_n})$. Solving for x gives $x = \frac{t_i}{t_n - t_i}$). As a consequence,

$$\log\left(\frac{t_i}{t_n}\right) = -\log\left(\frac{t_n}{t_i} - 1\right) + \log\left(1 - \frac{t_i}{t_n}\right). \tag{2.39}$$

Summing both sides and multiplying throughout by $-2\beta_0$:

$$\begin{aligned}
-2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_i}{t_n}\right) &= 2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right) - 2\beta_0 \sum_{i=1}^{n-1} \log\left(1 - \frac{t_i}{t_n}\right) \\
&\Rightarrow -2\beta_0 \sum_{i=1}^{n-1} \log\left(1 - \frac{t_i}{t_n}\right) = -2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right) - 2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_i}{t_n}\right) \\
&\Rightarrow Z_B = -2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right) + Z,
\end{aligned}$$

as claimed.

(Method 2) An alternate version of the proof might be furnished through the following observations:

$$\begin{aligned}
-\frac{Z}{2\beta_0} = \sum_{i=1}^{n-1} \log\left(\frac{t_i}{t_n}\right) &= \log \prod_{i=1}^{n-1} \left(\frac{t_i}{t_n}\right) \\
\Rightarrow \prod_{i=1}^{n-1} \left(\frac{t_i}{t_n}\right) &= e^{-\frac{Z}{2\beta_0}},
\end{aligned} \tag{2.40}$$

and similarly

$$\prod_{i=1}^{n-1} \left(1 - \frac{t_i}{t_n}\right) = e^{-\frac{Z_B}{2\beta_0}} \tag{2.41}$$

Thus, using the two above:

$$\begin{aligned}
e^{-\frac{1}{2\beta_0}(Z - Z_B)} &= \prod_{i=1}^{n-1} \frac{\frac{t_i}{t_n}}{1 - \frac{t_i}{t_n}} = \prod_{i=1}^{n-1} \frac{t_i}{t_n - t_i} = \prod_{i=1}^{n-1} \frac{1}{\frac{t_n}{t_i} - 1} \\
\Rightarrow -\frac{1}{2\beta_0}(Z - Z_B) &= \log\left(\prod_{i=1}^{n-1} \frac{1}{\frac{t_n}{t_i} - 1}\right) = \sum_{i=1}^{n-1} \log\left(\frac{1}{\frac{t_n}{t_i} - 1}\right) \\
&= -\sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right) \\
\Rightarrow Z_B = -2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right) + Z,
\end{aligned}$$

as required. □

(2.40) and (2.41) afford an alternative interpretation of these two versions: imagine a set of increasingly more likely set of independent events $\{A_i\}$'s with $P(A_i) = \frac{t_i}{t_n}$, then

$e^{-\frac{Z}{2\beta_0}}$ represents the probability that all of them should happen (i.e. $P(A_1A_2...A_{n-1})$) and $e^{-\frac{Z_B}{2\beta_0}}$ represents the probability that none of them should happen (i.e. $P(\bar{A}_1\bar{A}_2...\bar{A_{n-1}})$). A similar line of reasoning can generate a lower bound for the sum $Z + Z_B$: Consider $n-1$ independent but non-identical bernoulli trials with success probability $\frac{t_i}{t_n}$ on the i-th trial. Then the likelihood function is given by:

$$\begin{aligned} L(t|x) &= \prod_{i=1}^{n-1} \left(\frac{t_i}{t_n}\right)^{x_i} \left(1 - \frac{t_i}{t_n}\right)^{1-x_i} \\ \Rightarrow \log L(t|x) &= \sum_{i-1}^{n-1} x_i \log\left(\frac{t_i}{t_n}\right) + \sum_{i-1}^{n-1} (1-x_i) \log\left(1 - \frac{t_i}{t_n}\right) \\ \Rightarrow -2\beta_0 \log L(t|x) &= -2\beta_0 \sum_{i-1}^{n-1} x_i \log\left(\frac{t_i}{t_n}\right) + -2\beta_0 \sum_{i-1}^{n-1} (1-x_i) \log\left(1 - \frac{t_i}{t_n}\right) \\ \Rightarrow -2\beta_0 \log L(t|x) &\leq Z + Z_B \end{aligned} \tag{2.42}$$

where the last inequality follows due to the fact that the x_i's are binary variables.

It is hoped that this (possibly non-unique) exact deterministic connection (2.38) shall enable one to borrow information and relevant structure from the already existing literature on forward tests. To force the two forms (2.36) and (2.37) to be equivalent, we might require:

$$\begin{aligned} \log\left(\frac{t_n}{t_i} - 1\right) &= 0, \quad i = 1, 2, .., n-1 \\ \Rightarrow \frac{t_n}{t_i} - 1 = 1 \Rightarrow t_i &= \frac{t_n}{2}, \quad i = 1, 2, .., n-1. \end{aligned} \tag{2.43}$$

This condition (2.43), despite leading to exact equivalence, is untenable, especially under our overarching assumption of processes which are orderly or non-explosive (in the sense of (Jacobsen (2006) [74]), among others, i.e. processes where simultaneous occurrences are not possible). Nonetheless, sequences can be generated purely for academic interests, where (2.43) holds approximately, i.e. where t_i's are tightly packed around $\frac{t_n}{2}$, $i = 1, 2, ..., n-1$. For example, if a hypothetical process (under failure-truncation) exhibits the following time sequence: $\{4.95, 4.99, 5, 5.01, 5.03, 10\}$, the forward and backward statistics

take the values $Z = 6.939$ and $Z_B = 6.924$ which are extremely close and for this case, $-2\sum_{i=1}^{n-1}\log\left(\frac{t_n}{t_i}-1\right) = -0.016$. Thus, (2.43) alerts one to the existence of situations where the trend reversal through Z_B might be inconsequential. Such a situation might arise from a real example too: the Old Faithful Geyser case in Table 2.2 provides a case in point.

Applications of the forward and backward statistic (with the choice $\beta_0 = 1$, i.e. testing for homogeneity) to some of the "Other examples" considered in the introductory chapter leads us to the following table:

Table 2.2: Z and Z_B calculations on examples with diverse sample sizes ($\beta_0 = 1$)

Case studies	Sample size	Z	$-2\sum_{i=1}^{n-1} log(\frac{t_n}{t_i}-1)$	Z_B
Maguire, Pearson, Wynn (1952)	109	305.349	-148.211	157.138
Duane (1964)	14	57.954	-46.987	10.967
Jelsinki and Moranda (1972)	34	133.289	-108.057	25.232
Mooley (1981)	141	247.287	115.959	363.247
Ho (1993)	10	37.126	-28.916	8.211
Rigdon and Basu (2000)	18	33.316	2.830	36.146
Bakun et al. (2005)	7	32.161	-25.745	6.416
Wang and Liu (2014)	36	81.812	-23.993	57.819
Old faithful geyser eruptions	272	538.767	-0.761	538.007
Ogata earthquake	100	226.309	-48.013	178.296
N.China earthquake	65	129.237	4.129	133.366
Phuket earthquake	1248	2749.051	-1017.05	1732.001

Table (2.2)above covers a wide spectrum of scenarios with sample sizes ranging from as small as 7 to as large as 1248. As expected, the magnitude of the Z and Z_B statistics seems to be directly correlated with the sample size and it is gratifying to see that (2.38) is being satisfied in each case. (2.39) provides an alternative representation of $\hat{\beta}$ too:

$$\hat{\beta} = \frac{n}{-\sum_{i=1}^{n-1}\log\left(\frac{t_i}{t_n}\right)} = \frac{n}{\sum_{i=1}^{n-1}\log\left(\frac{t_n}{t_i}-1\right) - \sum_{i=1}^{n-1}\log\left(1-\frac{t_i}{t_n}\right)} \tag{2.44}$$

and in case a functional form connecting Z_B to $\hat{\beta}$ is ever sought for, one might use

$$Z_B = -2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right) + \frac{2n\beta_0}{\hat{\beta}}. \tag{2.45}$$

Returning to our case study on system deterioration, we saw that if failures are occurring more frequently in recent times, we should expect a small value for Z and a large value for Z_B. The first term $-2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{t_n}{t_i} - 1\right)$ in a way, takes care of the deficit. This term is amenable to other interpretations too: for instance, it establishes that there can be no uniform stochastic order dependence between Z and Z_B. This is in view of the fact that unlike $\log(\frac{t_i}{t_n})$ and $\log(1 - \frac{t_i}{t_n})$ which can only take on negative values due to the structural constraint $0 < \frac{t_i}{t_n} < 1, \quad i = 1, 2, .., n-1$, each term $\log(\frac{t_n}{t_i} - 1)$ can be either positive or negative, owing to the equivalent constraint $0 < \frac{t_n}{t_i} - 1 < \infty$, which in turn can make the first term as a whole, either positive or negative. Confirmation of this fact can be had from Table (2.2), and this alludes to the possibility that the algebraic sign of the first term might be taken as an indicator of system improvement or deterioration.

2.2.2 The null and alternate distributions of Z_B

Lemma 2.1. *If $X \sim Beta(m, n)$, then $E(X^{-\frac{1}{\beta}}) = \frac{B(n, m-\frac{1}{\beta})}{B(m,n)}$*

Proof. The proof might follow by an application of the change of variable technique. Defining $Y = X^{-\frac{1}{\beta}} = g(X)$ (say), we should have $X = Y^{-\beta} = g^{-1}(Y) \Rightarrow \frac{d}{dy} g^{-1}(Y) = -\beta Y^{-\beta-1}$. Thus, since:

$$f_X(x) = \frac{1}{B(m,n)} x^{m-1}(1-x)^{n-1}, \quad x \in (0,1),$$

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| \\ &= \frac{1}{B(m,n)} (y^{-\beta})^{m-1} (1 - y^{-\beta})^{n-1} \beta y^{-\beta-1} \quad , y \in (1, \infty). \end{aligned}$$

Consequently:

$$E(Y) = \frac{\beta}{B(m,n)} \int_1^\infty y^{-\beta m}(1-y^{-\beta})^{n-1} dy$$

Setting $u := 1 - y^{-\beta}$,

$$\begin{aligned} E(Y) &= \frac{1}{B(m,n)} \int_0^1 u^{n-1}(1-u)^{m-1-\frac{1}{\beta}} du \\ &= B(n, m-\frac{1}{\beta})/B(m,n) \end{aligned}$$

as claimed. □

Theorem 2.5. *Under the failure truncated case, the backward statistic Z_B is non-pivotal with expected value at least $-2\beta_0 \sum_{i=1}^{n-1} \log\left(\frac{B(n-i,i-\frac{1}{\beta})}{B(i,n-i)} - 1\right) + 2(n-1)$.*

Proof. In the previous chapter while showing the pivotal property of the Z statistic, we have argued that $(\frac{T_i}{t_n})^\beta,\ i = 1, 2, .., n-1$ are distributed as $n-1$ order statistics from a $U(0,1)$ distribution. Consequently, the marginal distribution of $(\frac{T_i}{t_n})^\beta$ would be $Beta(i, n-i)$. Thus, applying the previous lemma on $(\frac{T_i}{t_n})^\beta$, we must have

$$E\left(\frac{t_n}{T_i}\right) = \frac{B(n-i, i-\frac{1}{\beta})}{B(i,n-i)}. \tag{2.46}$$

Now since $\phi(x) = -\log(x)$ is a convex function, applying Jensen's inequality $E(\phi(X)) \geq \phi(E(X))$ with this choice of $\phi(.)$ yields:

$$E\left(-\log\left(\frac{t_n}{T_i} - 1\right)\right) \geq -\log\left(E\left(\frac{t_n}{T_i} - 1\right)\right) \tag{2.47}$$

$$\Rightarrow -E\left(\log\left(\frac{t_n}{T_i} - 1\right)\right) \geq -\log\left(E\left(\frac{t_n}{T_i}\right) - 1\right) \tag{2.48}$$

$$\Rightarrow E\left(\log\left(\frac{t_n}{T_i} - 1\right)\right) \leq \log\left(E\left(\frac{t_n}{T_i}\right) - 1\right). \tag{2.49}$$

Invoking (2.46), we have

$$E\left(\log\left(\frac{t_n}{T_i} - 1\right)\right) \leq \log\left(\frac{B(n-i, i-\frac{1}{\beta})}{B(i,n-i)} - 1\right). \tag{2.50}$$

Next, using the linearity property of expectations:

$$E\left(\sum_{i=1}^{n-1}\log\left(\frac{t_n}{T_i}-1\right)\right)\leq\sum_{i=1}^{n-1}\log\left(\frac{B(n-i,i-\frac{1}{\beta})}{B(i,n-i)}-1\right) \tag{2.51}$$

$$\Rightarrow E\left(-2\beta_0\sum_{i=1}^{n-1}\log\left(\frac{t_n}{T_i}-1\right)\right)\geq-2\beta_0\sum_{i=1}^{n-1}\log\left(\frac{B(n-i,i-\frac{1}{\beta})}{B(i,n-i)}-1\right). \tag{2.52}$$

Finally, using (2.38) and the fact that $Z\sim\chi^2(2(n-1))$, we have

$$E(Z_B)\geq-2\beta_0\sum_{i=1}^{n-1}\log\left(\frac{B(n-i,i-\frac{1}{\beta})}{B(i,n-i)}-1\right)+2(n-1), \tag{2.53}$$

as claimed. □

The first term on the right of (2.53) can be calculated numerically on softwares such as R. Specifically, we can revisit the examples shown previously and compare the lower bounds on the expected value of the backward statistic to the observed values:

Table 2.3: Z and Z_B calculations on examples described previously ($\beta_0=1$)

Case studies	Sample size	$\hat{\beta}$	Z_B	Exp lower bound at $\hat{\beta}$
Maguire, Pearson, Wynn (1952)	109	0.7139	157.138	122.775
Duane (1964)	14	0.4831	10.967	13.222
Jelsinki and Moranda (1972)	34	0.5102	25.232	18.010
Mooley (1981)	141	1.1403	363.247	332.076
Ho (1993)	10	0.5387	8.211	5.252
Rigdon and Basu (2000)	18	1.0806	36.146	38.417
Bakun et al. (2005)	7	0.4353	6.416	7.314
Wang and Liu (2014)	36	0.8801	57.819	56.829
Old faithful geyser eruptions	272	1.0097	538.007	530.077
Ogata earthquake	100	0.8837	178.296	161.500
N.China earthquake	65	1.0059	133.366	110.583
Phuket earthquake	1248	0.9079	1732.001	2263.224

Figure (2.3) below depicts the expected bound as a surface depending on n and β_0 along with its contours. It can be seen that the bound is almost always non-trivial in the sense of generating positive numbers.

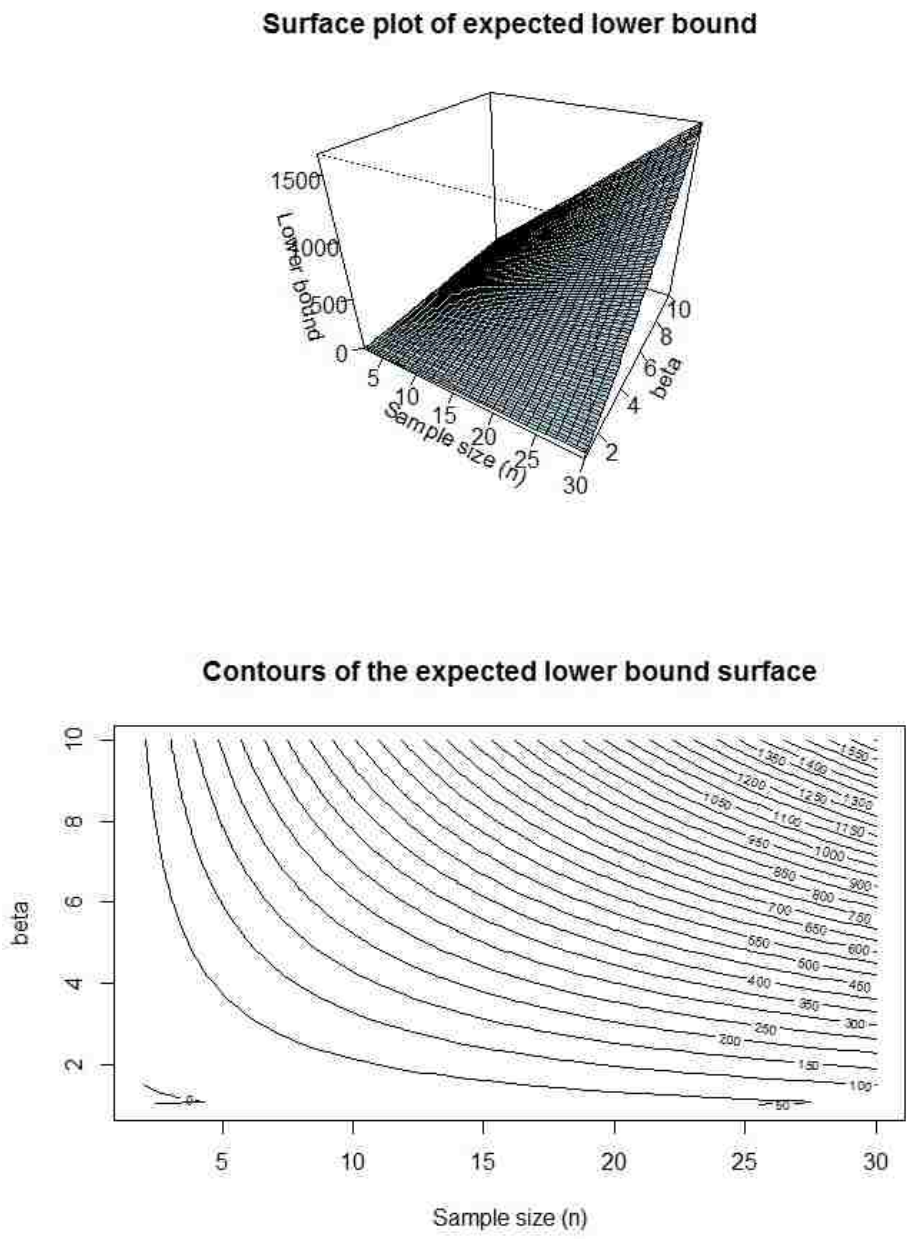


Figure 2.3: Surface diagram and contour plot for the expected lower bound of the backward statistic

To derive the null and alternate distribution of Z_B, we shall have to lay the following groundwork:

Theorem 2.6. *If $X^\beta \sim U(0,1)$, then the density of $Y = (1-X)^\beta$ is given by:*

$$f_Y(y) = y^{\frac{1}{\beta}-1}(1-y^{\frac{1}{\beta}})^{\beta-1} \quad , y \in [0,1] \tag{2.54}$$

Proof. The proof might follow from the traditional change of variable technique with $X^\beta = (1-Y^{\frac{1}{\beta}})^\beta \ = g^{-1}(Y)$ (say). Thus

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y))|\frac{d}{dy}g^{-1}(y)| \\ &= y^{\frac{1}{\beta}-1}(1-y^{\frac{1}{\beta}})^{\beta-1} \quad , y \in (0,1) \end{aligned}$$

Also, the c.d.f of Y is given by:

$$F_Y(y) = P(Y \leq y) = \int_0^y u^{\frac{1}{\beta}-1}(1-u^{\frac{1}{\beta}})^{\beta-1}du = 1-(1-y^{\frac{1}{\beta}})^{\beta-1} \quad , y \in (0,1)$$

which turns out to be the cumulative distribution function of a Kumaraswamy-Generalized distribution, studied by Pascoa et al. (2011) [112] under the baseline c.d.f of a uniform variable. □

Theorem 2.7. *If $X^\beta \sim U(0,1)$, then the density of $Z = -2ln(1-X)^\beta$ is given by:*

$$f_Z(z) = \frac{1}{2}e^{-\frac{z}{2\beta}}(1-e^{-\frac{z}{2\beta}})^{\beta-1} \quad , z \in [0,\infty) \tag{2.55}$$

Proof. We have shown that under similar conditions, the the density of $Y = (1-X)^\beta$ is given by:

$$f_Y(y) = y^{\frac{1}{\beta}-1}(1-y^{\frac{1}{\beta}})^{\beta-1} \quad , y \in [0,1] \tag{2.56}$$

Using $Z = -2lnY \quad \Rightarrow Y = e^{-Z/2} = g^{-1}(Z)$ (say), the density of Z is given by:

$$\begin{aligned} f_Z(z) &= f_Y(g^{-1}(z))|\frac{d}{dz}g^{-1}(z)| \\ &= \frac{1}{2}e^{-\frac{z}{2\beta}}(1-e^{-\frac{z}{2\beta}})^{\beta-1} \quad , z \in [0,\infty) \end{aligned}$$

□

This class of densities has been investigated extensively by Pascoa et al. (2011) [112]. Using $X = \frac{Y}{t_n}$ in the lemmas above, we can therefore conclude that $-2ln(1-\frac{Y}{t_n})^\beta$ has a density given by (2.55). Thus, arguing along lines similar to the derivation of Z's distribution, we can claim that $-2ln(1-\frac{T_i}{t_n})^\beta$ are order statistics from a distribution generated by $-2ln(1-\frac{Y}{t_n})^\beta$. Thus:

$$Z_B = -2\beta_0 \sum_{i=1}^{n-1} ln(1-\frac{T_i}{t_n}) = -2\sum_{i=1}^{n-1} ln(1-\frac{T_i}{t_n})^{\beta_0} =^d -2\sum_{i=1}^{n-1} ln(1-\frac{Y_i}{t_n})^{\beta_0} \tag{2.57}$$

where $=^d$ should be taken to mean equal in distribution. So for our purposes, it is enough to find the distribution of the sum of the i.i.d variables $-ln(1-\frac{Y_i}{t_n})^{\beta_0}$. Analytical expressions for the density of Z_B is difficult to derive as the following lines will reveal, but we can use (2.57) to prove the following claim held by Ho (1993) [66]:

Theorem 2.8. *The backward statistic Z_B is non-pivotal.*

Proof. Using (2.57):

$$Z_B =^d -2\sum_{i=1}^{n-1} ln(1-\frac{Y_i}{t_n})^{\beta_0} = \sum_{i=1}^{n-1} Z_i \ \ (say) \tag{2.58}$$

where the $\{Z_i\}$'s are i.i.d and the density of each is given by (2.55). The the m.g.f of Z_i becomes:

$$M_{Z_i}(t) = E(e^{tZ_i}) = \frac{1}{2}\int_0^\infty e^{z(t-\frac{1}{2\beta_0})}(1-e^{-\frac{z}{2\beta_0}})^{\beta_0-1}dz \tag{2.59}$$

Calling $u := 1-e^{-\frac{z}{2\beta_0}}$, this boils down to:

$$M_{Z_i}(t) = \beta_0 \int_0^1 u^{\beta_0-1}(1-u)^{-2t\beta_0}du = \beta_0 B(\beta_0, 1-2t\beta_0) \tag{2.60}$$

which is defined only on $\{t : t < \frac{1}{2\beta_0}\}$. Consequently, using the i.i.d-ness of the $\{Z_i\}$'s, the m.g.f of Z_B is:

$$M_{Z_B}(t) = [\beta_0 B(\beta_0, 1-2t\beta_0)]^{n-1} = {\beta_0}^{n-1}[\frac{\Gamma(\beta_0)\Gamma(1-2t\beta_0)}{\Gamma(1+(1-2t)\beta_0)}]^{n-1} \tag{2.61}$$

Thus, since the m.g.f. of Z_B involves β_0, the distribution of Z_B involves β_0 too and thus, in general, the backward statistic is non-pivotal. □

The m.g.f given in (2.61) is not in a form that is readily recognizable. Another direct approach to formulate the distribution of the i.i.d sum would involve convolutions of the form:

$$f_{Z_B}(z) = \int_0^z e^{-\frac{w}{\beta_0}}(1-e^{-\frac{w}{\beta_0}})^{\beta_0-1}e^{-\frac{z-w}{\beta_0}}(1-e^{-\frac{z-w}{\beta_0}})^{\beta_0-1}dw \tag{2.62}$$

However, even the above expression (with $n = 2$) is not simple to integrate analytically. Thus, we will propose two non-traditional ways to estimate the underlying distribution of Z_B: One proceeds through simulating random variables from recently proposed theoretical distributions (such as the Exponentiated Kumarswamy - G distribution or the Beta Exponential - G distribution) and adding them up, while the other relies on inverting the characteristic function of Z_B, which can, in turn, be constructed, using the mgf (2.61).

The simulation approach

To motivate this idea, let us consider a simple example: imagine we have two random variables X_1 and X_2, independently distributed of each other, with the first following a $N(1, 2)$ distribution and the second following a $N(2, 2)$ distribution. If we had to find the distribution of the sum $X_1 + X_2$ and if we were unaware of the theoretical result of normal reproductivity, we could generate a large number of observations from $N(0, 1)$, the same number of observations from $N(2, 2)$, add them up and work with the observed frequency distribution (a histogram, for instance) of this summed vector. This distribution should be centered at 3, be fairly symmetrical and in general, should replicate the essential features of, and provide an acceptable approximation to the underlying true $N(3, 4)$ model.

In view of (2.58), this idea can be profitably exploited for the task at hand: the equivalence claims that the backward statistic Z_B and the sum of the i.i.d. variables Z_i's are distributed similarly. Each of the Z_i's in turn, has a density given by (2.55):

$$f_Z(z) = \frac{1}{2} e^{-\frac{z}{2\beta}} (1 - e^{-\frac{z}{2\beta}})^{\beta - 1} \quad , z \in [0, \infty) \tag{2.63}$$

We can thus simulate a large number of observations from this distribution, add them up and look at the empirical law to get an approximation to the density of our target: Z_B. This empirical approach is unavoidable here since unlike traditional distributions such as the normal, (2.63) does not exhibit the additive property. Lemonte et al. (2013) [91] introduced the exponentiated Kumarswamy - G distribution given by:

$$f_Z(z) = abcg(z)G^{a-1}(z)[1 - G^a(z)]^{b-1}\{1 - [1 - G^a(z)]^b\}^{c-1} \quad , z \in [0, \infty) \tag{2.64}$$

where $G(.)$ is any valid c.d.f, $g(.)$ its corresponding p.d.f and a, b, c are all positive shape parameters. It can be noted that (2.63) is a special member of this generalized family corresponding to the choice: $a = 1, b = \frac{1}{2\beta}, c = \beta$ and $G(z) = 1 - e^{-z}$, the c.d.f of the exponential density.

Another recent density introduced by Alzaatreh et al. (2013) [1] is the beta exponential G distribution given by:

$$f_Z(z) = \frac{\lambda}{B(a, b)} g(z)\{1 - G(z)\}^{\lambda b - 1}\{1 - [1 - G(z)]^{\lambda}\}^{a-1} \quad , z \in [0, \infty) \tag{2.65}$$

where $G(.)$ is any valid c.d.f, $g(.)$ its corresponding p.d.f and a, b, λ are all positive shape parameters. Once again, (2.65) boils down to (2.63) when $a = \beta, b = 1, \lambda = \frac{1}{2\beta}$ and $G(z) = 1 - e^{-z}$, the c.d.f of the exponential density.

For each of the two densities introduced above, functions exist in the software R that will help one to generate random observations. We shall illustrate the findings using (2.64) only. This is owing to an economic computational time over the beta exponential G density (2.65). Similar results can be had using (2.65) too.

i) For a given β and for each i, we generate r values of the random variable Z_i using the code: `rexpkumg(r, spec=‘‘exp", a = 1, b = 1/(2*beta), c = beta)` and store them in a vector. Typically r was taken to be 100000.

ii) For a given sample size n, we add up these vectors to get $\sum_{i=1}^{n-1} Z_i :=^d Z_B$.

iii) Finally, we create a histogram of the frequency distribution of Z_B. The high value of r assures one of its closeness to the underlying probability density.

Figure (2.4) below depicts the output for different choices of β corresponding to $n = 10$:

The distribution, in general, show traces of right skewness and gets shifted to the right for higher choices of β. Thus for future exercises, if we are interested in testing $H_0 : \beta = 1$ vs $H_a : \beta = 2$, for instance, the black density can be taken as the null distribution while the green can be taken as the alternative.

The inversion approach

A one-one correspondence between the characteristic function and the density of a random variable and our ready availability of the m.g.f. of Z_B in (2.61) emboldens this alternative approach. Recall that given the characteristic function $\phi_X(.)$ of a random variable X as:

$$\phi_X(t) = E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f(x)dx, \tag{2.66}$$

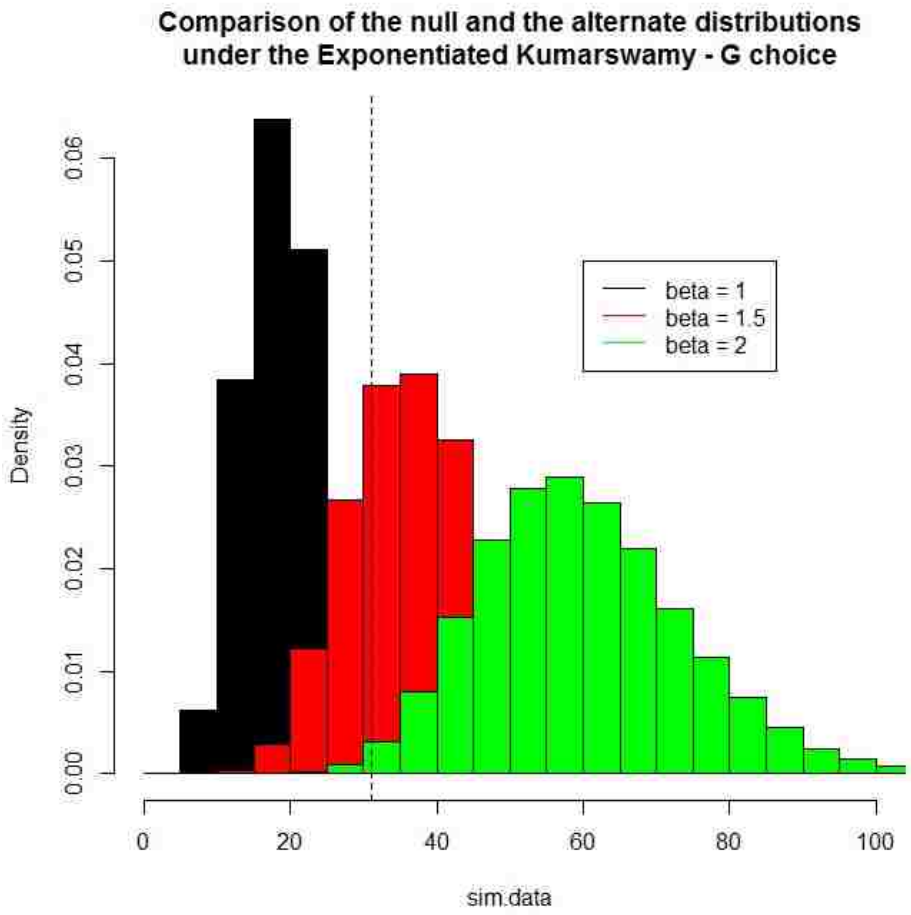


Figure 2.4: Density comparison for Z_b under different choices of β with $n = 10$ (simulations)

it is possible to recover $f(.)$ through an inverse Fourier transformation as:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t) dt. \tag{2.67}$$

Since the m.g.f and the characteristic function are connected through:

$$M_X(it) = \phi_X(t), \tag{2.68}$$

using (2.61), the characteristic function for Z_B would be:

$$\phi_{Z_B}(t) = \beta^{n-1}[B(\beta, 1 - 2it\beta)]^{n-1}, \tag{2.69}$$

and hence, the true density of Z_B would be given by:

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \beta^{n-1} [B(\beta, 1 - 2it\beta)]^{n-1} dt. \tag{2.70}$$

With our sincere efforts of finding closed form expressions for $f(.)$ ending up in smoke, we took recourse to numerical methods on the software Mathematica. In particular, the function "NIntegrate(.)" was used to simplify the expression (2.70). Once again, with $n = 10$ and with different choices of β, we have generated the following curves:

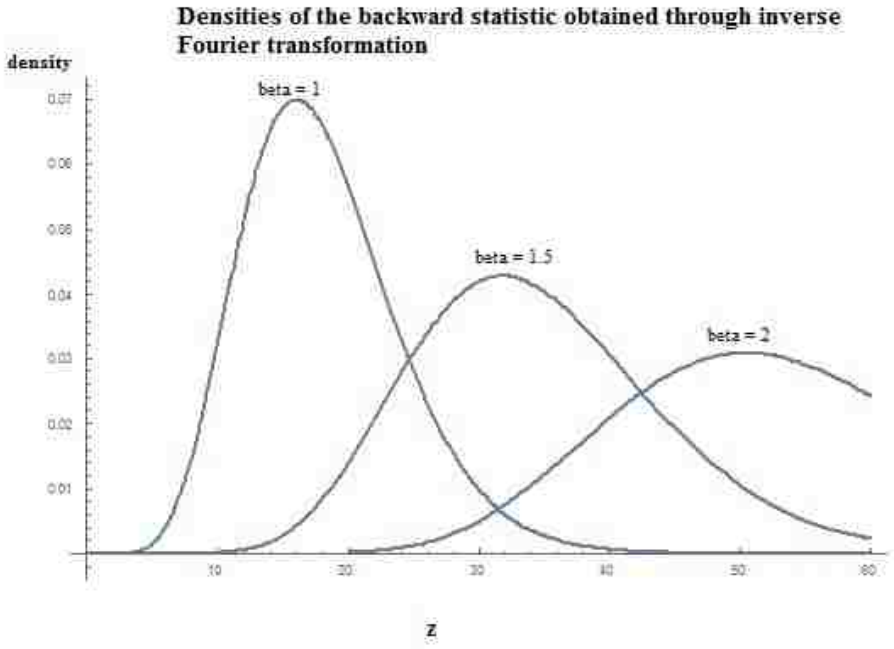


Figure 2.5: Density comparison for Z_B under different choices of β with $n = 10$ (theoteric)

Despite being generated through largely different methods, the close agreement between Figures (2.4) and (2.5) lends credence to the merit of both and the latter can be conveniently imagined to be the smoothed version of the former.

The special case of $\beta = 1$

Owing to the fact that a stable or homogeneous failure pattern results at $\beta = 1$, testing for homogeneity is often a special concern to practitioners. The m.g.f of the backward statistic with a sample of size n has been derived previously as:

$$M_{Z_B}(t) = \beta^{n-1}[B(\beta, 1-2t\beta)]^{n-1} \ , t < \frac{1}{2\beta}$$

So with $\beta = 1$:

$$M_{Z_B}(t) = [B(1, 1-2t)]^{n-1} \ , t < \frac{1}{2} \tag{2.71}$$

$$= [\frac{\Gamma(1)\Gamma(1-2t)}{\Gamma(2-2t)}]^{n-1} \ , t < \frac{1}{2} \tag{2.72}$$

$$= [\frac{\Gamma(1)\Gamma(1-2t)}{(1-2t)\Gamma(1-2t)}]^{n-1} \ , t < \frac{1}{2} \tag{2.73}$$

$$= [\frac{1}{1-2t}]^{n-1} \ , t < \frac{1}{2} \tag{2.74}$$

which is the m.g.f. of a chi-square distribution with $2(n-1)$ degrees of freedom. Thus, using the one-one correspondence between a density and its m.g.f., we can claim that the distribution of backward statistic under $\beta = 1$, which will equivalently serve as the null distribution in tests related to homogeneity, a special case of a more general structure, is a chi-square with $2(n-1)$ degrees of freedom under the failure truncated case, with a similar analog holding for the time truncated situation.

This result is particularly appealing since it will afford us the ability to get rid of numerical approximations and define critical regions *exactly* using cutoffs from the known chi-square density. Estimates of power will consequently be more reliable. While testing for other choices of β however, even to define the critical cutoffs, we shall have to fall back on the simulation from the Kumarswamy class approach or the inversion of the characteristic function

approach, both of which, with varying degrees, are inherently numerical approximations.

2.2.3 The critical regions

As a repairable system deteriorates, the majority of the t_i values will tend to cluster around t_n and consequently, $(1-\frac{t_i}{t_n})$ will be close to 0. Thus unlike the forward statistic Z, the value of the backward statistic Z_B will be inflated. This asymmetry between the two versions can also be visualized using equivalence (2.38). Consequently, we can borrow results from the previous chapter and flip the directions of the rejection regions: to be precise, previously, we rejected the null in favor of a greater than type alternative for small values of the forward statistic Z. Now we shall reject such a null for large values of the backward statistic Z_B. Table (2.4) below summarizes

Table 2.4: Critical regions from the backward test under different alternatives

	$H_a : \beta > \beta_0$	$H_a : \beta < \beta_0$	$H_a : \beta \neq \beta_0$
Failure truncated	$Z_B > f_\alpha(n)$	$Z_B < f_{1-\alpha}(n)$	$Z_B > f_{\alpha/2}(n)$ or $Z_B < f_{1-\alpha/2}(n)$
Time truncated	$Z_B > f_\alpha(n)$	$Z_B < f_{1-\alpha}(n)$	$Z_B > f_{\alpha/2}(n)$ or $Z_B < f_{1-\alpha/2}(n)$

where $f_\alpha(n)$, is the upper α point of the null distribution of Z_B. Unfortunately if $\beta \neq 1$, it is rather difficult to quantify thresholds characterizing critical regions through closed-form expressions, but one can always fall back on either of the two methods proposed in the previous section to estimate that point.

For instance, if one is interested in testing for process deterioration: $H_0 : \beta = 1$ vs $H_a : \beta > 1$ based on a sample of size 10, he can go back to the empirical red distribution of Z_B (Fig (2.4)) and measure the proportion of times the simulated values were in excess of z^* (say). Requiring this proportion be 0.05 (=α), he would estimate z^* as 30.07 (graphed as

the vertical line on Fig (2.4)). Thus, in the present instance, based on $n = 10$, the Z_B-based critical region would be $[30.07, \infty)$. Of course, under this special $\beta_0 = 1$ case, as evidenced by the previous m.g.f-based reasoning, exact chi-square calculations could have been done, which would have given the critical region as $[29, \infty)$, approximately. But this method would prove useful for checking specific non-homogeneity values such as $\beta_0 = 2, 3,$, etc.

If one insists on the inversion method, z^* can be calculated by solving:

$$\frac{1}{2\pi} \int_{z^*}^{\infty} \int_{-\infty}^{\infty} e^{-itz} 1^9 [B(1, 1 - 2it)]^9 dt \; dz = 0.05 \tag{2.75}$$

Once again, numerical methods need to be employed and we found Mathematica to be a reliable tool. We might note that in either case, the actual data set (apart from its size) has not been used to decide the cut-off.

2.2.4 The power functions

Derivation of the power curves once again, shall tread a numerical route. For a given null hypothesis and sample size, the cut-off determining the critical region should be found either through the simulation or the inversion method as described in the previous section. For a greater than type alternative, for instance, use of the simulation method would require one measure the proportion of observations (generated under the alternate choice of β) that exceed the threshold. This proportion can be taken as a reasonable approximation of the power. Using the null $H_0 : \beta = 1$ and varying the sample size, we found the following power curves using this method:

Once z^* is decided, the inversion method approach to get the power would necessitate

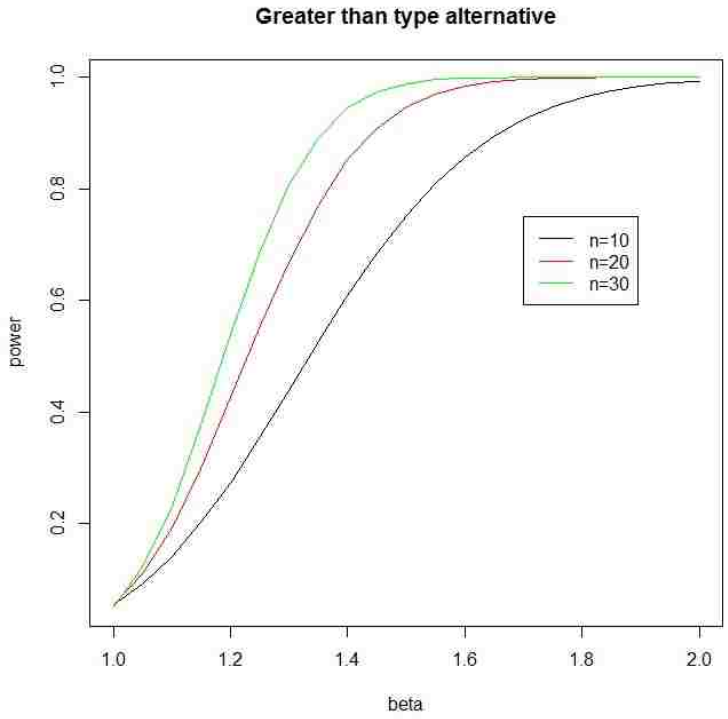


Figure 2.6: Typical power curves for the backward test for varying sample sizes

one evaluate:

$$\pi_{G,B}(\beta) = \frac{1}{2\pi} \int_{z^*}^{\infty} \int_{-\infty}^{\infty} e^{-itz} \beta^{n-1} [B(\beta, 1 - 2it\beta)]^{n-1} \, dt \, dz \tag{2.76}$$

We found the simulation results from R and the inversion results from Mathematica to be in close agreement. Similar operations can, of course, be carried out on other types of alternatives too.

2.3 On simulations

The purpose of the present section will be to put both the forward and the backward version of the Z-test in action when data are simulated from a Power Law process. It is worthwhile to recall that under such a constraint, the monotonic intensity function governing the

occurrences of the repairable system is given for $t > 0$ by

$$\lambda(t) = \left(\frac{\beta}{\theta}\right)\left(\frac{t}{\theta}\right)^{\beta-1}, \quad \beta > 0,\ \theta > 0. \tag{2.77}$$

Due to its explicit dependence on t, the Poisson process that it will give rise to will be a non-stationary or a non-homogeneous one and despite its inability to model simultaneous occurrences from explosive processes (Jacobsen (2006) [74]), owing to its smoothness, or to model a simultaneous improvement-deterioration scenario owing to its monotonicity, this form of intensity has found widespread use in reliability literature due to its mathematical tractability.

We shall agree to denote such a process by $PLP(\theta, \beta)$ and numerous methods exist in literature that enable one to simulate events from such a process, the most notable ones among them being:

i) time scale transformation of a homogeneous Poisson process (HPP).

ii) generation of the intervals between the events individually.

iii) using order statistics.

iv) by the method of thinning.

Excellent expositions on each approach can be found in such standard texts as Ross (1990) [130]. In the present instance however, we shall use the first approach which essentially exploits the connection between a non-homogeneous Poisson process (NHPP) and a HPP of rate 1:

Let $N_1(t)$ denote the HPP of rate 1. That implies that its inter-event distributions are exponentials of unit rate. Thus, if I_0 represent the inter-event variable, then:

$$P(I_0 \geq t) = \exp(-t) \Rightarrow P(I_0 \geq \Lambda(t)) = \exp(-\Lambda(t)) \tag{2.78}$$

where $\Lambda(t) = \int_0^x \lambda(x)dx$ represents the cumulative intensity function. Thus:

$$P(\Lambda^{-1}(I_0) \geq t) = \exp(-\Lambda(t)) \tag{2.79}$$

Now if $I_0^{'}$ denote the inter-events for the NHPP, then:

$$P(I_0^{'} \geq t) = \exp(-\Lambda(t)) \tag{2.80}$$

The last two equations imply that $X_1^{'}, X_2^{'}, ...$ are events from a NHPP with cumulative intensity $\Lambda(.)$ if and only if $X_1 = \Lambda(X_1^{'}), X_2 = \Lambda(X_2^{'}), ...$ are events from a unit rate HPP. In our case, with the choice of the intensity function given in (2.77):

$$\Lambda(t) = \int_0^t \left(\frac{\beta}{\theta}\right)\left(\frac{x}{\theta}\right)^{\beta-1} dx = \left(\frac{t}{\theta}\right)^{\beta} \tag{2.81}$$

Simulating events from a rate 1 HPP is rather straightforward. Once we have generated event times $t_1, t_2, ...$ from such a process, the transformation method guarantees that the required NHPP times would be given by:

$$\theta t_1^{1/\beta}, \theta t_2^{1/\beta}, ... \tag{2.82}$$

Our analyses on estimation and hypothesis testing for the NHPP would thus, naturally be based on these transformed time points.

But prior to the actual implementation of the inferential machinery, it will be instructive to look at the simulated events and check whether they are in keeping with our interpretation of the β parameter. Figure 2.7 below graphs the NHPP event points when transformations are imposed on a single parent series generated from a unit rate Poisson process. Since

$$\frac{\partial}{\partial t}\left(\frac{\beta}{\theta}\right)\left(\frac{t}{\theta}\right)^{\beta-1} = \frac{\beta}{\theta^2}(\beta-1)\left(\frac{t}{\theta}\right)^{\beta-2} > 0 \text{ for } \beta > 1, \tag{2.83}$$

the failure process gradually gets more intense as β exceeds that threshold. This is evident from the lower panels of the figure where points tend to cluster around the recent past, clearly the indication of a deteriorating system. With $\beta = 1$, i.e. corresponding to a homogeneous Poisson choice, the event points are more or less uniformly spread out with no apparent clustering tendency and with $\beta = 0.5$, failures are getting less prevalent with the advent of time: a hallmark of an improving system.

Estimates of the β and θ parameters can be obtained using the m.l. equations shown previously. But to get hypotheses testing related quantities such as the critical regions, p-values or the power curves, we can use any of the two methods described in the previous section. To motivate the simulation approach, one can fall back on the traditional example of testing for a normal mean when the variance is unknown. Under such a familiar premise, if one is interested in testing:

$$H_0 : \theta = \theta_0 \quad vs \quad H_a : \theta > \theta_0 \tag{2.84}$$

where θ is the unknown normal mean, the obvious statistic to use would be $\bar{X}$ and one would reject H_0 in favor of H_a if $\bar{X} > \theta_0 + \frac{\sigma}{\sqrt{n}}\tau_\alpha$ where τ_α is the upper α point of a $N(0,1)$ distribution, σ is the known standard deviation and n is the sample size. The power function $\pi(\theta)$, the probability of rejecting H_0 when sampling from a $N(\theta, \sigma^2)$ density is given by:

$$\pi(\theta) = P_\theta(\bar{X} > \theta_0 + \frac{\sigma}{\sqrt{n}}\tau_\alpha) = 1 - \Phi(\frac{\theta_0 - \theta}{\sigma/\sqrt{n}} + \tau_\alpha) \tag{2.85}$$

This is the theoretical expression for the power function familiar to all. The above analyses have been possible because of our knowledge of the fact that under i.i.d. $N(\theta, \sigma^2)$ sampling, $\bar{X} \sim N(\theta, \frac{\sigma^2}{n})$. We would like to investigate whether we can still get some empirical approximation to the power function if the above fact is unknown.

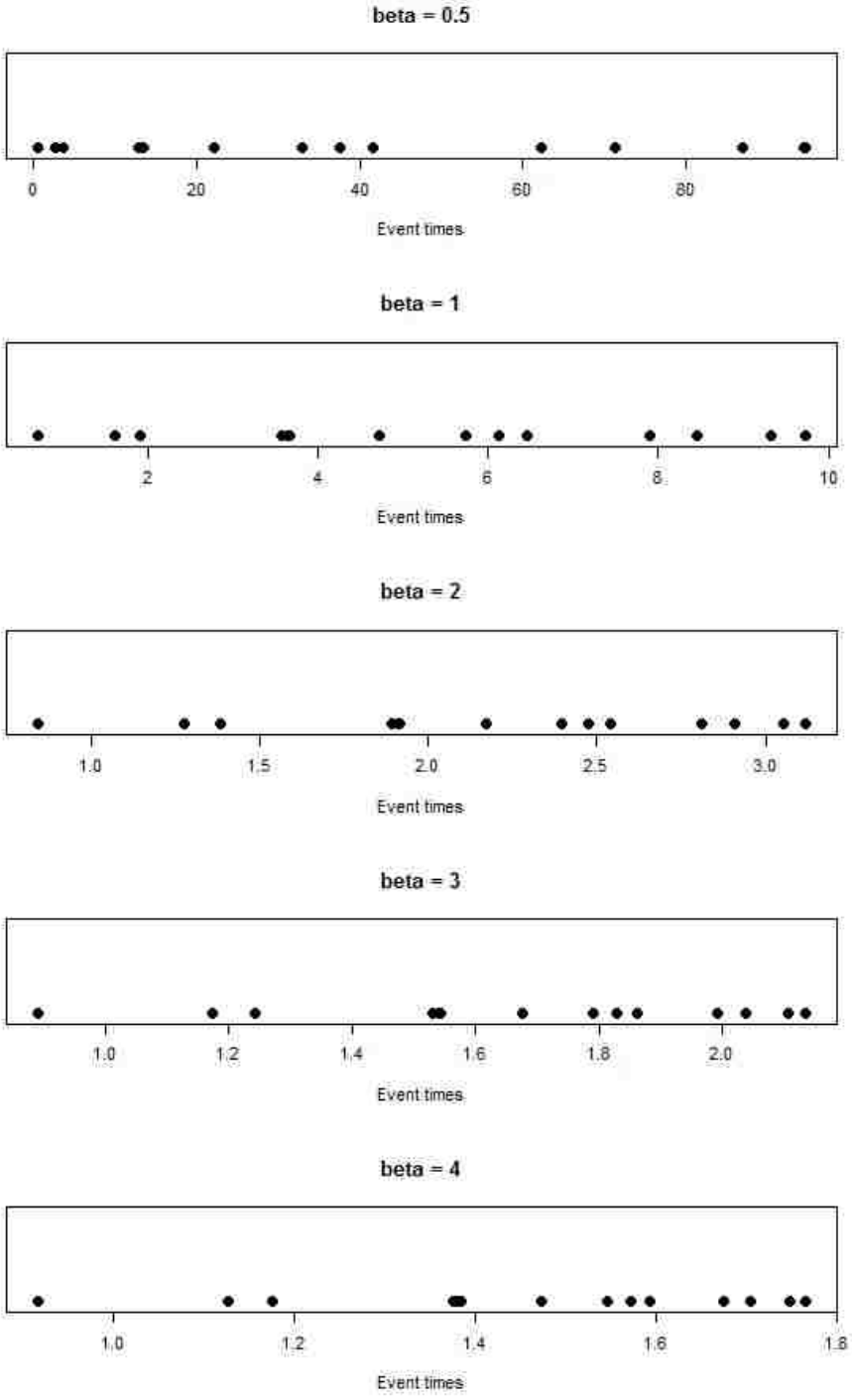


Figure 2.7: Simulation from NHPP's with different choices of β.

To be specific, let us set $\theta_0 = 1, \sigma = 2$ and $n = 25$ and to adequately define the critical region, we first simulate 25 vectors each of size 10000 from $N(1, 4)$, take their mean to

get their approximate distribution and calculate the 95th percentile out of it. This point should be serving as the (approximate) cutoff defining the right-tailed critical region. To get an estimate of the power at point $\theta = \theta$, we repeat the same exercise with a sample generated from a $N(\theta, 4)$ and find the proportion of means exceeding the cutoff. Figure (2.8) below compares the theoretical function to this empirical estimate, and we can see that it is reasonably well approximated. Similar accuracy should be carried over to the more complicated real cases at hand, involving PtPs, and non-smooth intensities.

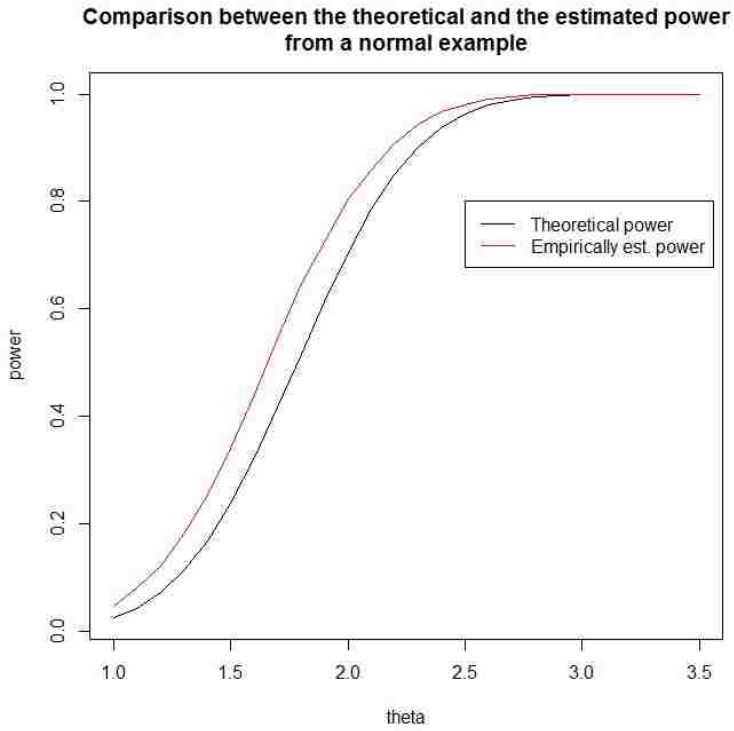


Figure 2.8: Comparison between theoretical and estimated power for a greater than type alternative under Normal sampling

2.4 Power comparisons under different intensities

Intensity functions, as described previously, influence failure patterns in a significant way, and statistics performing well in detecting a process's increased restlessness under a given intensity might lose efficiency when confronted with another. A previous section has shown the UMP-ness of the forward Z-test under power law intensities of the form (2.77). Cox (1955) [29] studies the Laplace's test statistic

$$L = \sum_{i=1}^{n} \frac{T_i}{T_n} \tag{2.86}$$

and shows how its standardized version, distributed asymptotically normally under the null assumption of homogeneity, turns out to be UMPU for testing process deterioration under intensities of the form

$$\lambda(t) = \alpha \exp(\beta t) \tag{2.87}$$

This section takes a tour of the types of intensities that will be explored in this dissertation and discovers the framework under which switching time could prove beneficial. Unless explicitly mentioned, we shall henceforth be concerned with the failure truncated scheme only. Excellent texts such as Bain and Engelhardt (1991) [10], Rigdon and Basu (2000) [122] describe ways to extract similar results under the time truncated framework.

2.4.1 Weibull (Power law) intensity

Bain and Engelhardt (1980) [7], Bassin (1969) [13], Crow (1974, 1982) [32] [34], Finkelstein (1976) [49] and Lee and Lee (1978) [89] have worked extensively with this form of the intensity function (graphed in Fig (2.9)), providing, as described previously, the maximum likelihood

estimates of β and θ for the intensity

$$\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1}, \ t > 0, \tag{2.88}$$

their confidence intervals, and goodness of fit tests. Crow (1974) [32] tests β treating θ as a nuisance parameter. To create each of the tables to follow, we have used $\alpha = 0.05$ to generate 10^4 NHPPs with the specified β using the time-scale transformation described in section 2.3, and checked how many of these are indeed classified as non-stationary by the forward and backward test. θ was held fixed at unity, and estimate for β for each case has been collected. For the forward test, the theoretical powers have been stored in the parentheses.

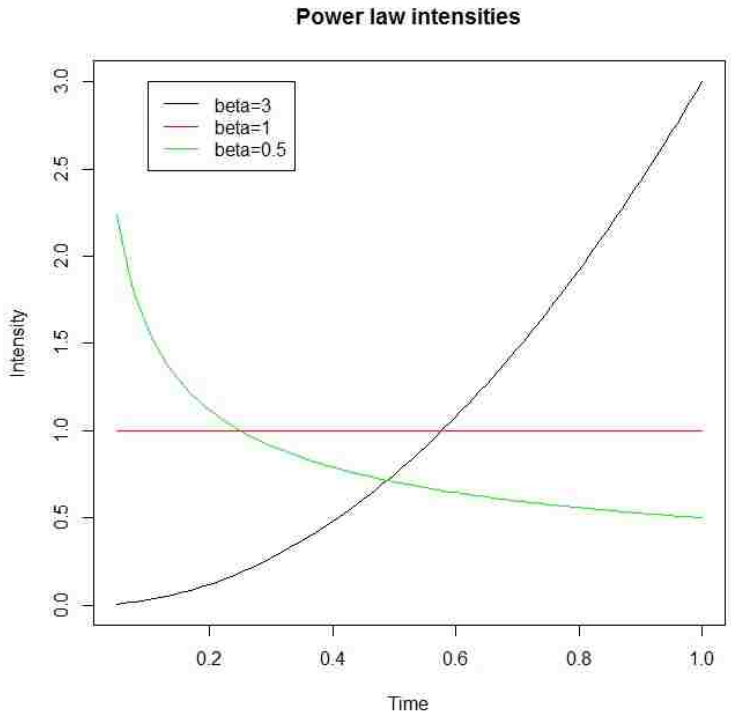


Figure 2.9: Power law (Weibull) intensities for different choices of β

Table 2.5: Power comparison between the forward and backward test under PLP assumption with $\beta = 0.6, \hat{\beta} = 0.591$

Sample size	Forward test	Backward test
$n = 15$	0.6402 (0.6386)	0.3769
$n = 25$	0.8188 (0.8165)	0.4969
$n = 35$	0.9052 (0.9102)	0.6017

Table 2.6: Power comparison between the forward and backward test under PLP assumption with $\beta = 0.8, \hat{\beta} = 0.768$

Sample size	Forward test	Backward test
$n = 15$	0.2281 (0.2332)	0.1482
$n = 25$	0.3203 (0.3162)	0.1917
$n = 35$	0.3929 (0.3908)	0.2319

Table 2.7: Power comparison between the forward and backward test under PLP assumption with $\beta = 1, \hat{\beta} = 1.007$

Sample size	Forward test	Backward test
$n = 15$	0.0541 (0.05)	0.0542
$n = 25$	0.0482 (0.05)	0.0563
$n = 35$	0.0499 (0.05)	0.0502

Table 2.8: Power comparison between the forward and backward test under PLP assumption with $\beta = 1.2, \hat{\beta} = 1.214$

Sample size	Forward test	Backward test
$n = 15$	0.1447 (0.1472)	0.1104
$n = 25$	0.2014 (0.2032)	0.1326
$n = 35$	0.2563 (0.2563)	0.1567

Table 2.9: Power comparison between the forward and backward test under PLP assumption with $\beta = 1.5, \hat{\beta} = 1.550$

Sample size	Forward test	Backward test
$n = 15$	0.3984 (0.3936)	0.2321
$n = 25$	0.5918 (0.5926)	0.3416
$n = 35$	0.7434 (0.7389)	0.4356

The numbers in the parentheses are the theoretical powers for the forward test, expressions for which have been derived previously, and their close agreement with the empirical calculations, in a spirit similar to Fig (2.8), is noted. In addition to numerically demonstrating the expected UMP-ness of the established forward Z test, these tables document Z_B's power loss corresponding to different sample sizes.

2.4.2 Compound power law intensity

One may next investigate the power performance of the backward test under other non-standard intensities, which are still smooth in nature. We found that Engelhardt and Bain (1987) [9] puts a $Gamma(x, \gamma)$ "prior" on $\lambda = \theta^{-\beta}$ as

$$\pi(\lambda) = \frac{1}{\gamma^x \Gamma(x)} \lambda^{x-1} \exp(-\lambda/\gamma), \quad \lambda > 0 \tag{2.89}$$

Using transformation of variables, a logical "prior" for θ would be:

$$\pi(\theta) = \frac{\beta}{\gamma^x \Gamma(x)} \theta^{-\beta x - 1} \exp(-1/\gamma\theta^{\beta}), \quad \theta > 0 \tag{2.90}$$

(2.90) is a special case of the generalized inverse gamma distribution introduced by Mead (2015) [105].

Thus, a compound PLP may be simulated using Bayesian smoothing with

- The "likelihood":

$$\lambda(t|\theta) = \theta^{-\beta}\beta t^{\beta-1}, \quad t > 0 \tag{2.91}$$

- The "prior"

$$\pi(\theta) = \frac{\beta}{\gamma^x \Gamma(x)} \theta^{-\beta x - 1} \exp(-1/\gamma\theta^{\beta}), \quad \theta > 0 \tag{2.92}$$

- The "pseudo" unconditional intensity:

$$\lambda(t) = \int_0^{\infty} \lambda(t|\theta)\pi(\theta)d\theta = x\gamma\beta t^{\beta-1}, \quad t > 0 \tag{2.93}$$

The mean function will thus be

$$\Lambda(t) = \int_0^t x\gamma\beta y^{\beta-1} dy = x\gamma t^{\beta}, \tag{2.94}$$

and the non-homogeneous times can be obtained from the homogeneous times using

$$NH.Times = \left(\frac{H.Times}{x\gamma}\right)^{\frac{1}{\beta}}. \tag{2.95}$$

With 10^4 as the simulation strength and an α value of 0.05, we then have the following set of power results.

Table 2.10: Power comparison between the forward and backward test under compound PLP assumption (GIG prior) with $x = 0.5, \gamma = 1$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5349	0.2745	0.2702	0.1428	0.0516	0.0501
$n = 25$	0.7352	0.3964	0.4376	0.2186	0.0474	0.0499
$n = 35$	0.8583	0.4971	0.6115	0.3137	0.0491	0.0496

Table 2.11: Power comparison between the forward and backward test under compound PLP assumption (GIG prior) with $x = 1, \gamma = 2$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5348	0.2753	0.2641	0.1391	0.0511	0.0506
$n = 25$	0.7324	0.3887	0.4579	0.2301	0.0476	0.0458
$n = 35$	0.8556	0.4941	0.6059	0.3082	0.0527	0.0486

Table 2.12: Power comparison between the forward and backward test under compound PLP assumption (GIG prior) with $x = 2, \gamma = 2$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5306	0.2779	0.2684	0.1436	0.0500	0.0527
$n = 25$	0.7326	0.3855	0.4408	0.2162	0.0516	0.0491
$n = 35$	0.8612	0.5025	0.6091	0.3195	0.0483	0.0518

Table 2.13: Power comparison between the forward and backward under compound PLP assumption (GIG prior) test with $x = 3, \gamma = 2$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5403	0.2794	0.2561	0.1394	0.0458	0.0463
$n = 25$	0.7310	0.3962	0.4466	0.2273	0.0502	0.0492
$n = 35$	0.8571	0.5010	0.6060	0.3161	0.0510	0.0531

Table 2.14: Power comparison between the forward and backward test under compound PLP assumption (GIG prior) with $x = 5, \gamma = 1$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5355	0.2751	0.2592	0.1406	0.0469	0.0476
$n = 25$	0.7418	0.3948	0.4537	0.2296	0.0502	0.0472
$n = 35$	0.8628	0.4971	0.6107	0.3154	0.0508	0.0498

Table 2.15: Power comparison between the forward and backward test under compound PLP assumption (GIG prior) with $x = 7.5, \gamma = 1$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5267	0.2674	0.2637	0.1434	0.0524	0.0487
$n = 25$	0.7331	0.3921	0.4528	0.2352	0.0490	0.0498
$n = 35$	0.8632	0.5023	0.6101	0.3112	0.0476	0.0461

Another way of generating a compound PLP is by using a recent three parameter Amorosro prior (2.97) introduced by Crooks (2015) [30] where

- The "likelihood":

$$\lambda(t|\theta) = \theta^{-\beta}\beta t^{\beta-1}, \quad t > 0 \tag{2.96}$$

- The "prior"

$$\pi(\theta) = \frac{1}{\Gamma(c)}|\frac{d}{b}|(\frac{\theta}{b})^{cd-1}\exp(-(\frac{\theta}{b})^d), \quad \theta > 0 \tag{2.97}$$

- The "pseudo" unconditional intensity:

$$\lambda(t) = \int_0^\infty \lambda(t|\theta)\pi(\theta)d\theta = \beta t^{\beta-1}\frac{1}{\Gamma(c)}(-(-\frac{1}{b})^d)^{-c+\frac{\beta}{d}}b^{-cd}\Gamma(c-\frac{\beta}{d}), \quad t > 0 \tag{2.98}$$

The mean function thus becomes

$$\Lambda(t) = \int_0^t \lambda(y)dy = \frac{1}{\Gamma(c)}(-(-\frac{1}{b})^d)^{-c+\frac{\beta}{d}}b^{-cd}\Gamma(c-\frac{\beta}{d})t^\beta, \tag{2.99}$$

and the non-homogeneous times can be obtained from the homogeneous times using

$$NH.Times = \left(\frac{H.Times}{\frac{1}{\Gamma(c)}(-(-\frac{1}{b})^d)^{-c+\frac{\beta}{d}}b^{-cd}\Gamma(c-\frac{\beta}{d})t^\beta}\right)^{\frac{1}{\beta}} \tag{2.100}$$

The power comparisons, with 10^4 simulations and $\alpha = 0.05$, follow.

Table 2.16: Power comparison between the forward and backward test under compound PLP assumption (Amoroso prior) with $b = 1, c = 2, d = 3$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5379	0.2712	0.2656	0.1437	0.0534	0.0506
$n = 25$	0.7369	0.3923	0.4526	0.2370	0.0493	0.0513
$n = 35$	0.8581	0.4954	0.6094	0.3083	0.0547	0.0484

Table 2.17: Power comparison between the forward and backward test under compound PLP assumption (Amoroso prior) with $b = 5, c = 4, d = 5$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5357	0.2736	0.2638	0.1433	0.0516	0.0495
$n = 25$	0.7374	0.3994	0.4435	0.2228	0.0481	0.0482
$n = 35$	0.8563	0.4939	0.6097	0.3174	0.0490	0.0534

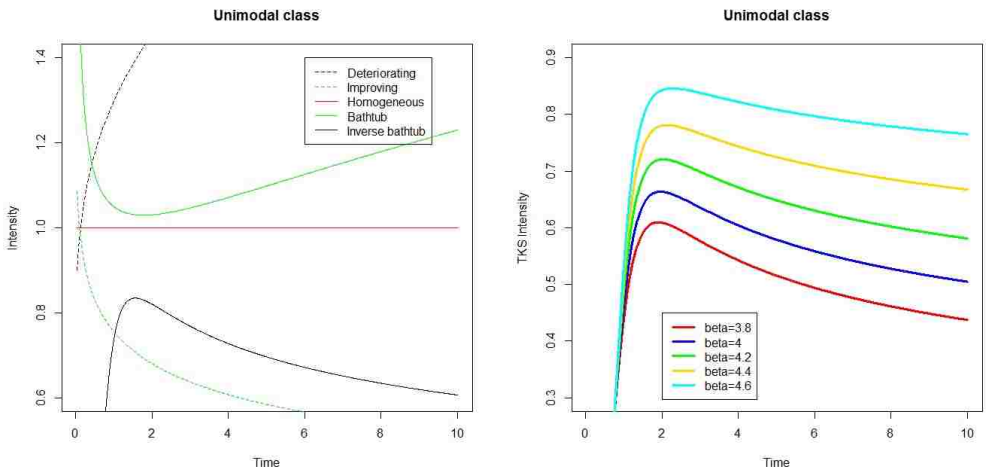
Table 2.18: Power comparison between the forward and backward test under compound PLP assumption (Amoroso prior) with $b = 15, c = 5, d = 7$

	True $\beta = 0.6$		*True* $\beta = 1.5$		*True* $\beta = 1$	
Sample size	Z power	Z_B power	Z power	Z_B power	Z power	Z_B power
$n = 15$	0.5366	0.2684	0.2659	0.1431	0.0509	0.0481
$n = 25$	0.7387	0.3952	0.4607	0.2334	0.0530	0.0471
$n = 35$	0.8638	0.5009	0.6024	0.3125	0.0517	0.0509

The hyperparameters in each of compound PLPs were carefully chosen to cover a wide type of intensity shapes. The superiority of the forward test Z is still persistent, especially with small sample sizes, when at times, it becomes twice as powerful as Z_B.

2.4.3 Unimodal smooth intensity

Power law or compound Power law intensities encountered in sections 2.4.1 and 2.4.2 are able to describe intensities which are either strictly increasing or decreasing. Dimitrakopoulou et



(a) Generalized smooth intensities (b) Unimodal intensities

Figure 2.10: Unimodal intensities introduced by Dimitrakopoulou et al. (2007)

al. (2007) [35] introduced unimodal intensities of the form

$$\lambda(t) = \alpha\beta\lambda t^{\beta-1}(1+\lambda t^{\beta})^{\alpha-1},\ t > 0 \tag{2.101}$$

which can handle bathtub or inverse bathtub shaped scenarios, graphed in Fig (2.10).

The following graph (Fig. (2.11)) compares the power functions of the forward, backward and Laplace tests, and identifies a range of β values over which the backward test outperforms the forward.

We shall not pursue this class of intensity any further in this dissertation, but will, however, note the possible existence of intensities (even within the smooth family) under which, the forward supremacy might be challenged. The merit in adding an extra parameter may be judged by deviance type tests.

2.4.4 Step intensity

Smooth intensities (unimodal or not) are adequate for representing changes in the failure pattern which are gradual, over a considerable period of time. If these changes are abrupt,

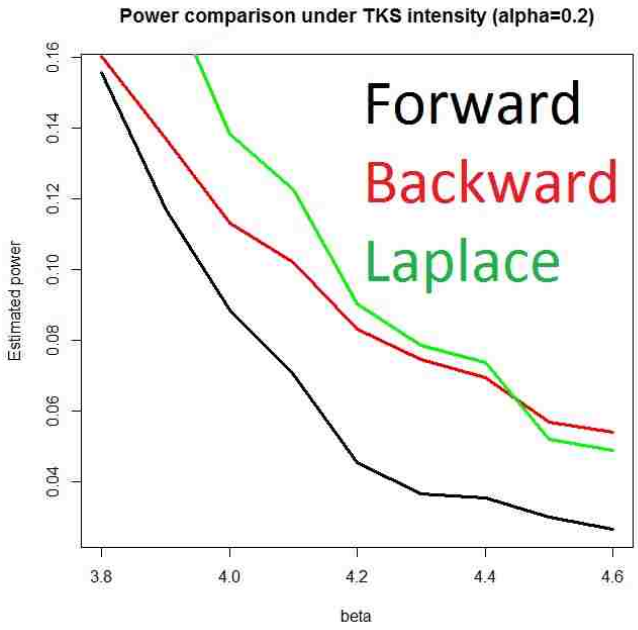


Figure 2.11: Power comparisons among forward, backward and Laplace tests under unimodal intensities

then *step intensities* of the type graphed in Fig (2.12), a special class of rough intensities of the form

$$\lambda(t) = \sum_{i=1}^{p} k_i I_{(\tau_{i-1},\tau_i]}(t),\ t > 0 \tag{2.102}$$

where $I_A(.)$ represents the usual indicator function on a set A, are often useful.

The two-step case ($p = 2$) and the power performance of the forward, backward, and Laplace tests have been analyzed in considerable detail by Ho (1993) [66]. To generalize, we shall look at an example when $k_1 = 1, k_2 = 2, k_3 = 1, n = 15$ and the sampling frequency is 1:1:1. This means within each interval $(0, \tau_1], (\tau_1, \tau_2], (\tau_2, \infty)$, we would expect 5 (=15/3) shocks. To elaborate

- On $(0, \tau_1]$, the process is a HPP with rate 1. $\Rightarrow$ the inter-event time is exponential with average $\frac{1}{1} = 1$. So the waiting time for 5 events $= 5 \times 1$ units. $\Rightarrow \tau_1 = 5$.

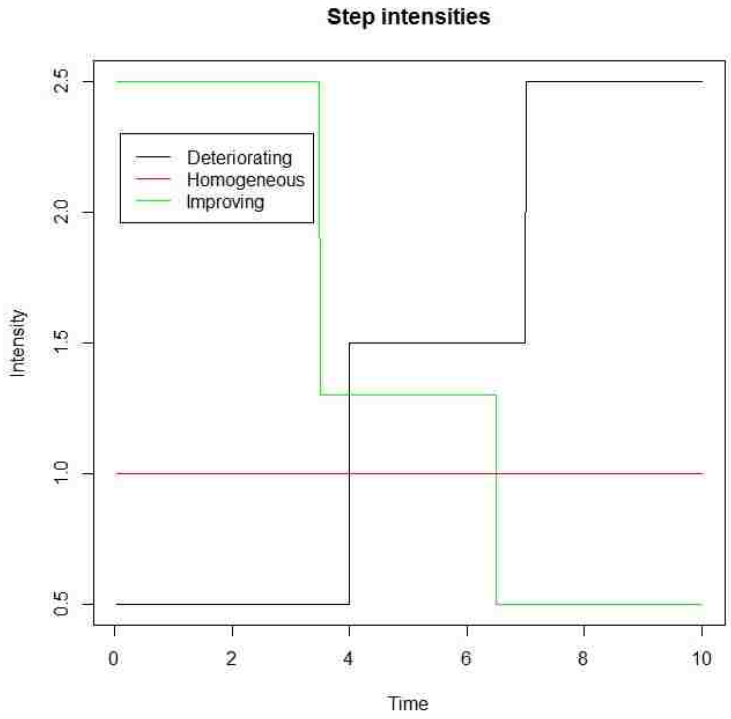


Figure 2.12: Step intensities with three steps

- On $(\tau_1, \tau_2]$, the process is a HPP with rate 2. $\Rightarrow$ the inter-event time is exponential with average $\frac{1}{2}$. So the waiting time for 5 events $= 5 \times \frac{1}{2}$ units. $\Rightarrow \tau_2 = 5 + \frac{5}{2} = 7.5$.

From the generic step intensity shown in (2.102), with $p = 2$, we have:

$$\Lambda(t) = \int_0^t \lambda(x)dx = \begin{cases} k_1 t & 0 \le t \le \tau_1 \\ k_1\tau_1 + k_2(t - \tau_1) & \tau_1 < t \le \tau_2 \\ k_1\tau_1 + k_2(\tau_2 - \tau_1) + k_3(t - \tau_2) & t > \tau_2 \end{cases}$$

Thus, from a simulated HPP of rate 1, say, $X = (x_1, x_2, ..., x_n)$, the NHPP, say Y, can be generated through the following inverse transformation:

$$y_i = \begin{cases} \frac{x_i}{k_1} & x_i \le k_1\tau_1 \\ \frac{x_i - k_1\tau_1}{k_2} + \tau_1 & k_1\tau_1 < x_i \le k_1\tau_1 + k_2(\tau_2 - \tau_1) \\ \frac{x_i - k_1\tau_1 - k_2(\tau_2 - \tau_1)}{k_3} + \tau_2 & x_i > k_1\tau_1 + k_2(\tau_2 - \tau_1) \end{cases}$$

Figures (2.13) and (2.14) below depict the conversion using unimodal and bathtub-shaped step intensities.

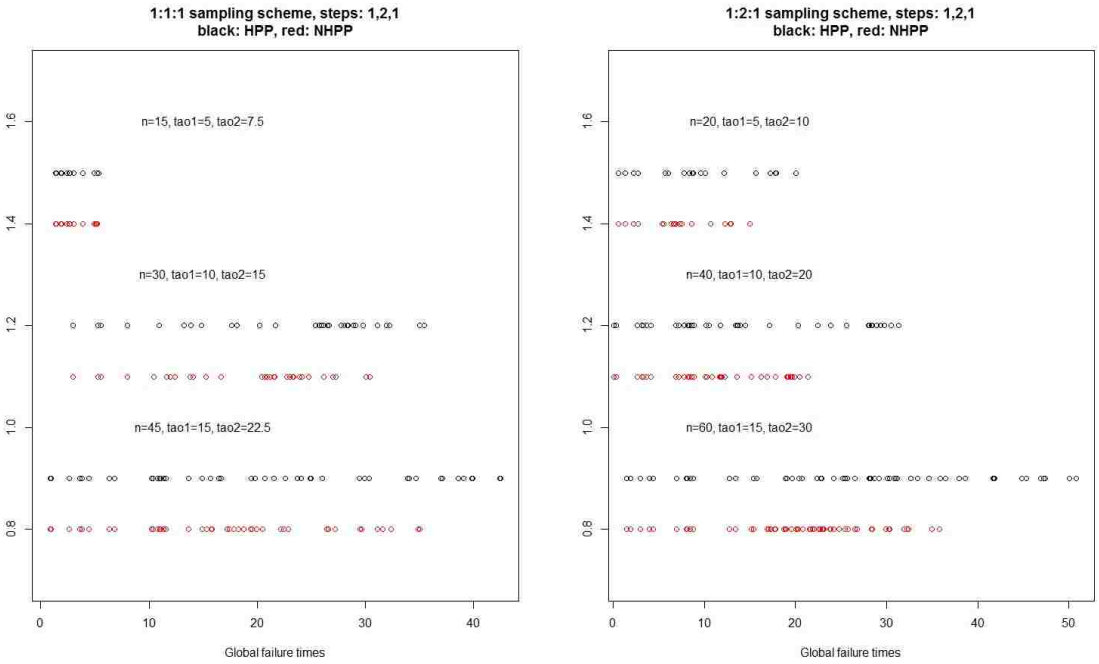


Figure 2.13: Time transformations under unimodal step intensities

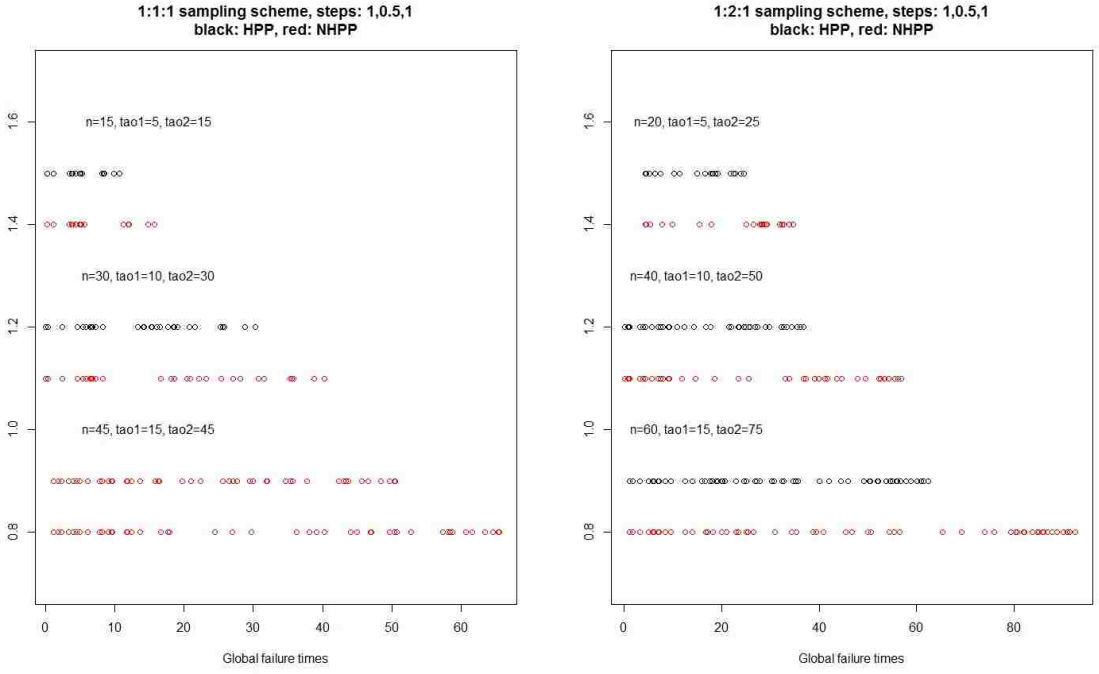


Figure 2.14: Time transformations under bathtub shaped step intensities

Bain et al. (1985) [8] and Engelhardt et al. (1990) [43] have studied the forward test Z and the Laplace test L under smooth intensities and increasing rough intensities with up to three steps, recommending L for instance, to guard against step intensities. Using deterministic two-step rough intensities, Ho (1993) [66] studied and compared the trend detection abilities among Z, Z_B and L, assuming failure truncation. Using similar parameters (the placement of knots, the height of the steps, the sampling frequency, the simulation strength, etc) and group frequencies (1:1, 1:2, and 2:1 for Groups I, II and III, respectively), we retraced the steps to get Table (2.19), the values in which are in close agreement with Ho (1993) [66], showing the backward test works best if the jump is placed late into the process.

Table 2.19: Power comparison between Z, Z_B and L with $k_1 = 1, k_2 = 3$

	Group I			Group II			Group III		
	n=10	n=20	n=40	n=10	n=20	n=40	n=10	n=20	n=40
Z	0.2682	0.3999	0.6066	0.2990	0.5052	0.7532	0.2172	0.2738	0.4192
Z_B	0.3064	0.5343	0.8213	0.2862	0.4880	0.7743	0.2711	0.4602	0.7374
L	0.3341	0.5530	0.8246	0.3628	0.6313	0.8935	0.2767	0.3987	0.6353

The purpose of this section is to generalize the $Z - Z_B$ comparisons to three steps, and to a random mixing of two steps. 10^4 iterations, the number used by Ho (1993) [66], was used.

Table 2.20: Power comparison between the forward and backward test with $k_1 = 1, k_2 = 2, k_3 = 1$, sampling frequency = 1:1:1

	Sampling freq = 1:1:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 15$	0.0623	0.0543	5.003	4.9071	5	7.5
$n = 30$	0.0635	0.0562	9.9586	9.9976	10	15
$n = 45$	0.0690	0.0519	14.9817	14.9953	15	22.5

Table 2.21: Power comparison between the forward and backward test with $k_1 = 1, k_2 = 2, k_3 = 1$, sampling frequency = 1:2:1

	Sampling freq = 1:2:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 20$	0.0796	0.0634	5.0445	9.7782	5	10
$n = 40$	0.0989	0.0704	9.9301	19.8930	10	20
$n = 60$	0.1116	0.0785	14.9382	29.8900	15	30

Table 2.22: Power comparison between the forward and backward test with $k_1 = 1, k_2 = 0.5, k_3 = 1$, sampling frequency = 1:1:1

	Sampling freq = 1:1:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 15$	0.0940	0.0984	5.0307	4.8726	5	15
$n = 30$	0.1111	0.0935	10.0045	9.9606	10	30
$n = 45$	0.1263	0.1059	14.9930	14.9970	15	45

Table 2.23: Power comparison between the forward and backward test with $k_1 = 1, k_2 = 0.5, k_3 = 1$, sampling frequency = 1:2:1

	Sampling freq = 1:2:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 20$	0.1233	0.1065	5.0200	9.8050	5	25
$n = 40$	0.1509	0.1356	10.0263	19.8596	10	50
$n = 60$	0.1800	0.1567	14.9873	29.9782	15	75

Table 2.24: Power comparison between the forward and backward test with $k_1 = 1, k_2 = 2, k_3 = 3$, sampling frequency = 1:1:1

	Sampling freq = 1:1:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 15$	0.2623	0.2352	5.0088	4.8811	5	7.5
$n = 30$	0.4378	0.4719	10.0435	9.9387	10	15
$n = 45$	0.5886	0.6693	15.0238	14.9823	15	22.5

Table 2.25: Power comparison between the forward and backward test with $k_1 = 1, k_2 = 2, k_3 = 3$, sampling frequency = 1:2:1

	Sampling freq = 1:2:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 20$	0.3047	0.2422	5.0010	9.7489	5	10
$n = 40$	0.5067	0.4670	9.9475	19.9217	10	20
$n = 60$	0.6774	0.6691	15.0038	29.9359	15	30

Table 2.26: Power comparison between the forward and backward test with $k_1 = 3, k_2 = 2, k_3 = 1$, sampling frequency = 1:1:1

	Sampling freq = 1:1:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 18$	0.2878	0.2845	6.0445	5.8903	2	5
$n = 36$	0.5398	0.4925	12.0228	11.9709	4	10
$n = 54$	0.7371	0.6651	18.0583	18.0416	6	15

Table 2.27: Power comparison between the forward and backward test with $k_1 = 3, k_2 = 2, k_3 = 1$, sampling frequency = 1:2:1

	Sampling freq = 1:2:1		*Average counts*		*Knot placements*	
Sample size	Z power	Z_B power	$E(0, \tau_1]$	$E(\tau_1, \tau_2]$	τ_1	τ_2
$n = 24$	0.2925	0.2963	5.9738	11.7766	2	8
$n = 48$	0.5434	0.5387	12.0093	23.8463	4	16
$n = 72$	0.7227	0.7238	18.0263	35.9533	6	24

In groups of two different frequencies $1:1:1$ and $1:2:1$, we have surveyed, in Tables (2.20) to (2.27), three-step intensities that are unimodal, bathtub shaped, monotonically increasing and decreasing. It is interesting to note that the two-step Z_B superiority conclusion seen in Table (2.19) extends to three steps too, evidenced by Table (2.24) and there exist other situations, such as the one shown in Table (2.27) where Z_B performs almost as good and at times, better than the forward test. Next, we investigate their performance against

a three-step alternative of the form:

$$H_a : \lambda(t) = 1I_t(0, \tau_1] + 3I_t(\tau_1, \tau_2] + 5I_t(\tau_2, \infty) \tag{2.103}$$

where the height of the steps are larger compared to the previous examples. Arguably, this leads to easier non-stationarity detection and the locations of the jump pairs are emphasized in the following comparison table.

Table 2.28: Power comparison between Z, Z_B and L with $k_1 = 1, k_2 = 3, k_3 = 5$

	Group I			Group II			Group III		
	n=10	n=20	n=40	n=10	n=20	n=40	n=10	n=20	n=40
	(3,7)	(6,14)	(12,18)	(2,5)	(5,11)	(12,24)	(5,8)	(12,18)	(26,38)
Z	0.6321	0.5264	0.7726	0.3511	0.5472	0.7677	0.2678	0.3211	0.4210
Z_B	0.4910	0.4536	0.7206	0.2388	0.4311	0.7360	0.3172	0.5027	0.7373
L	0.6721	0.6393	0.8903	0.3593	0.6315	0.8931	0.3437	0.4640	0.6299

To appreciate the increase in power, we choose the last category (Group III, n = 40) and change the step heights $k_2 = 3$ and $k_3 = 5$ in the intensity above to different combinations to get the following surface diagrams (Figs (2.15) and (2.16)). The backward test largely dominates the other two.

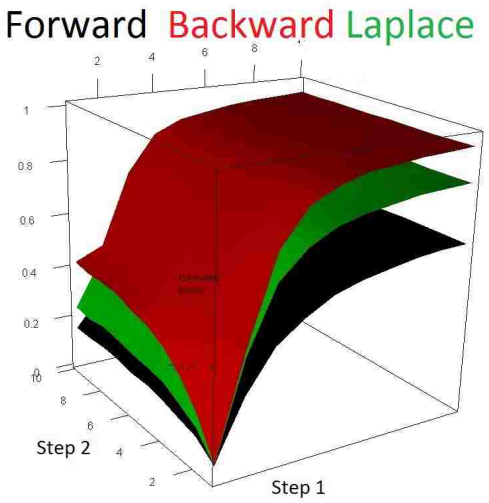


Figure 2.15: Power comparison among Z, Z_B and L with high values of k_2, k_3 (view 1)

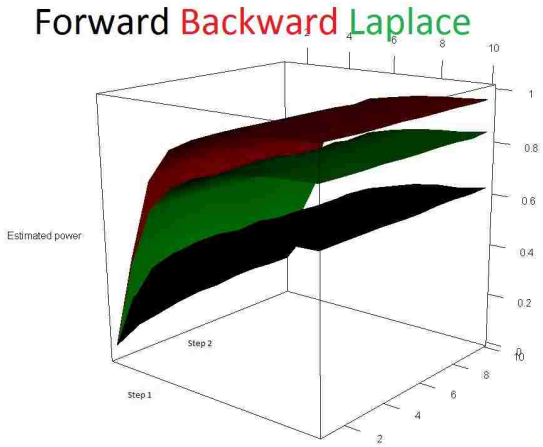


Figure 2.16: Power comparison among Z, Z_B and L with high values of k_2, k_3 (view 2)

Random mixing of steps

Next, with a sample of size 40, we consider a 50-50 mixture of an increasing step intensity ($k_1 = 1, k_2 = 3$) and a decreasing step intensity ($k_1 = 3, k_2 = 1$). The rationale behind these mixtures will be discussed in detail in the next chapter. Crudely, from a simulation viewpoint, it represents sampling the process from the increasing intensity 50% of the times and from the decreasing intensity, the remaining 50% of the times. Gleaning information from Table

(2.19) and (2.24), the jumps were placed late into the process at $\tau_{inc} = 27, \tau_{dec} = 27/3$ to accentuate the difference in the power performance. The backward test using Z_B shows better classification accuracy (Table(2.29)) under this hybrid framework too.

Table 2.29: Power comparison under failure truncation (50-50 step mixing) among the forward, backward and Laplace test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 27/3$, combination: $(2:1) \times (2:1)$

Test	Estimated power
Forward test Z	0.5441
Backward test Z_B	0.7218
Laplace test L	0.7069

Thus, in summary, this chapter suggests that under a large class of smooth intensities (such as the Power laws or compound Power laws), switching time will not be profitable from the viewpoint of achieving a power higher than Z, but under rough step intensities, it will.

Chapter 3

Bi-directional testing

Statistics employed to weed out non-stationary tendencies react differently to changes in the underlying intensity. We have witnessed how the forward test Z remains optimal under power laws and a host of smooth intensities, but loses power in the face of rough step intensities. The backward test Z_B renders better classification accuracy in those cases. Noting how non-stationarity (or non-homogeneity) embraces both process deterioration and improvement, one of the goals of this dissertation is to reap the benefits from both Z and Z_B through the creation of a single, implementable tool, fusing their optimal properties. What follows in this chapter is an adumbration of that development.

3.1 Prelude

In our quest to combine the forward and backward test, we explore a series of options, each paving the way for the next logical proposal. When the unidirectional arms Z and Z_B are used directly in the bidirectional test function, we emphasize it by placing an ordered superscript $\{X, Y\}$ to clarify that Z is "looking" in the X direction, while Z_B is "looking" in the Y direction. $\{X, Y\} \in \{L, R\}$, with L representing the left, and R, the right. This superscript will be dropped eventually to create a double bidirectional test $\phi(Z_{BD})$, when the

individual components will not be Z and Z_B anymore, but their maximum and minimum. Here are our initial proposals.

$$\phi(Z_{BD}^{LR}) = \begin{cases} 1 & \text{if } Z < \chi^2_{1-\frac{\alpha}{2},2n-2} \text{ or } Z_B > \chi^2_{\frac{\alpha}{2},2n-2}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

$$\phi(Z_{BD}^{LL}) = \begin{cases} 1 & \text{if } Z < \chi^2_{1-\frac{\alpha}{2},2n-2} \text{ or } Z_B < \chi^2_{1-\frac{\alpha}{2},2n-2}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

$$\phi(Z_{BD}^{RR}) = \begin{cases} 1 & \text{if } Z > \chi^2_{\frac{\alpha}{2},2n-2} \text{ or } Z_B > \chi^2_{\frac{\alpha}{2},2n-2}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

$$\phi(Z_{BD}^{RL}) = \begin{cases} 1 & \text{if } Z > \chi^2_{\frac{\alpha}{2},2n-2} \text{ or } Z_B < \chi^2_{1-\frac{\alpha}{2},2n-2}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

Thus, to detect general non-stationarity, in $\phi(Z_{BD}^{RR})$ for instance, we are employing the forward test Z to detect process improvement and the backward test Z_B to detect deterioration. The unidirectional tests are run at level $\alpha/2$ each to maintain the overall level condition. Arguably, thus, $\phi(Z_{BD}^{RR})$ and $\phi(Z_{BD}^{LL})$ are expected to be efficient in signaling non-stationarity in both directions, while $\phi(Z_{BD}^{LR})$ and $\phi(Z_{BD}^{RL})$ will be more efficient in sensing deterioration and improvement, respectively. Using a simulation strength of 10^4, samples of size 40 (i.e., 40 events within each generation), and $\alpha = 0.1$, we embark on a large scale power study, using step intensities with different trends (increasing/decreasing) and different placement of jumps (early, midway, late). To confirm intuitive expectations, for the bidirectional tests, we keep a running record of the proportion of cases rejected by Z only, by Z_B only, and by both. These are represented by a, b and c, respectively. The powers, given by $a + b - c$ are also stored. The established unidirectional tests' performances are also stored alongside. Each

table monitors the average number of events observed in both the "pre-change" interval and the "post-change" interval, to confirm our expected sampling frequencies.

3.1.1 Deterministic steps

The steps are first chosen deterministically, i.e., for an increasing step scenario $k_1 = 1, k_2 = 3$ with the knot placed at $\tau = 20$, each one of the 10^4 non-homogeneous cases had their parameters controlled that way while being simulated.

Increasing intensity

Table 3.1: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 1, k_2 = 3, \tau = 20$, frequency =1:1

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.6149	0.8163	a=	0.0351	0.6138	0.0010	0.0011
		b=	0.2376	0.0000	0.8162	0.0000
		c=	0.5787	0.0000	0.0001	0.0000
		pow=	0.8514	0.6138	0.8173	0.0011

$$E(N(0,\tau]) = 19.9316, E(N(\tau,\infty)) = 20.0684.$$

Table 3.2: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 1, k_2 = 3, \tau = 14$, frequency =1:2

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.7497	0.7723	a=	0.1006	0.7495	0.0002	0.0002
		b=	0.1234	0.0000	0.7723	0.0000
		c=	0.6489	0.0000	0.0000	0.0000
		pow=	0.8729	0.7495	0.7725	0.0002

$$E(N(0,\tau]) = 13.9456, E(N(\tau,\infty)) = 26.0544.$$

Table 3.3: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 1, k_2 = 3, \tau = 26$, frequency =2:1

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.4111	0.7273	a=	0.0162	0.4068	0.0043	0.0041
		b=	0.3362	0.0005	0.7268	0.0003
		c=	0.3906	0.0000	0.0000	0.0002
		pow=	0.7430	0.4073	0.7311	0.0046

$$E(N(0,\tau]) = 25.9863, E(N(\tau,\infty)) = 14.0137.$$

We find that $\phi(Z_{BD}^{LR})$ consistently outperforms the forward and backward tests and the other versions of the bidirectional test. This is expected since as mentioned previously, both arms have been employed to detect deterioration, which is the framework we have worked under. This is also evidenced quantitatively by $\phi(Z_{BD}^{LR})$, unlike the rest, deriving most of its strength from joint rejection, i.e., from the c component. In these cases, the height of the second step k_2 was held fixed at 3. To check whether the bidirectional supremacy will prevail over other heights, we let k_2 vary to get Fig (3.1)

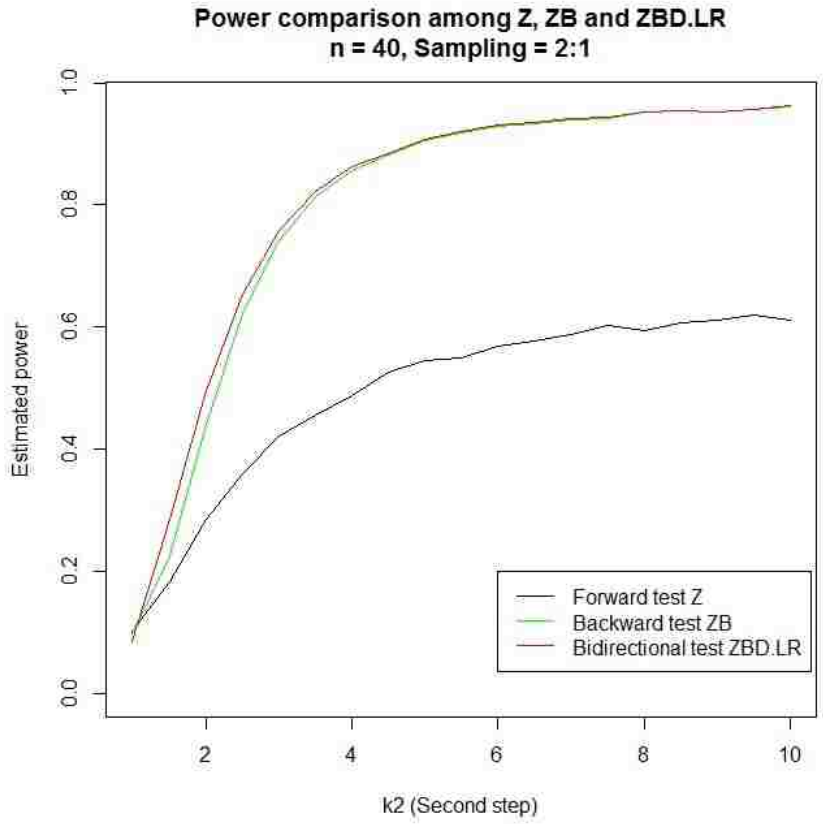


Figure 3.1: Generalized power comparison among Z, Z_B and Z_{BD}^{LR} with $n = 40$ and a $(2:1)$ sampling scheme

We note that as detection becomes easier, corresponding to higher values of k_2, the bidirectional test is almost as good as the unidirectional backward test, but for smaller values, corresponding to harder detection, it significantly outperforms the rest.

Decreasing intensity

Now, we move on to simulating improving systems from decreasing intensities of the form $k_1 = 3, k_2 = 1$. Due to reasons described earlier, ϕ_{BD}^{RL} is now expected to perform the best. The following tables corroborate.

Table 3.4: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 3, k_2 = 1, \tau = 20/3$, frequency =1:1

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.8169	0.6230	a=	0.0001	0.0001	0.8167	0.2348
		b=	0.0007	0.6223	0.0006	0.0403
		c=	0.0000	0.0000	0.0007	0.5820
		pow=	0.0008	0.6224	0.8174	0.8571

$$E(N(0,\tau]) = 19.9699, E(N(\tau,\infty)) = 20.0301.$$

Table 3.5: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 3, k_2 = 1, \tau = 16/3$, frequency =2:3

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.8090	0.5235	a=	0.0000	0.0000	0.8089	0.3113
		b=	0.0021	0.5214	0.0020	0.0237
		c=	0.0000	0.0000	0.0001	0.4977
		pow=	0.0021	0.5214	0.8110	0.8327

$$E(N(0,\tau]) = 16.0521, E(N(\tau,\infty)) = 23.9479.$$

Table 3.6: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 3, k_2 = 1, \tau = 24/3$, frequency =3:2

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.7643	0.7133	a=	0.0002	0.0002	0.7641	0.1437
		b=	0.0002	0.7131	0.0002	0.0927
		c=	0.0000	0.0000	0.0000	0.6204
		pow=	0.0004	0.7133	0.7643	0.8568

$$E(N(0,\tau]) = 24.0868, E(N(\tau,\infty)) = 15.9132.$$

Analogous to the increasing intensity scenario, we find that that $\phi(Z_{BD}^{RL})$ now is performing the best, since both arms have been employed to detect improvement, which $k_1 = 3, k_2 = 1$ represents. This is confirmed by a high c component. In these cases, the height of the first

step k_1 was held fixed at 1. To check whether we have similar results under other choices, we let k_1 vary to get Fig (3.2)

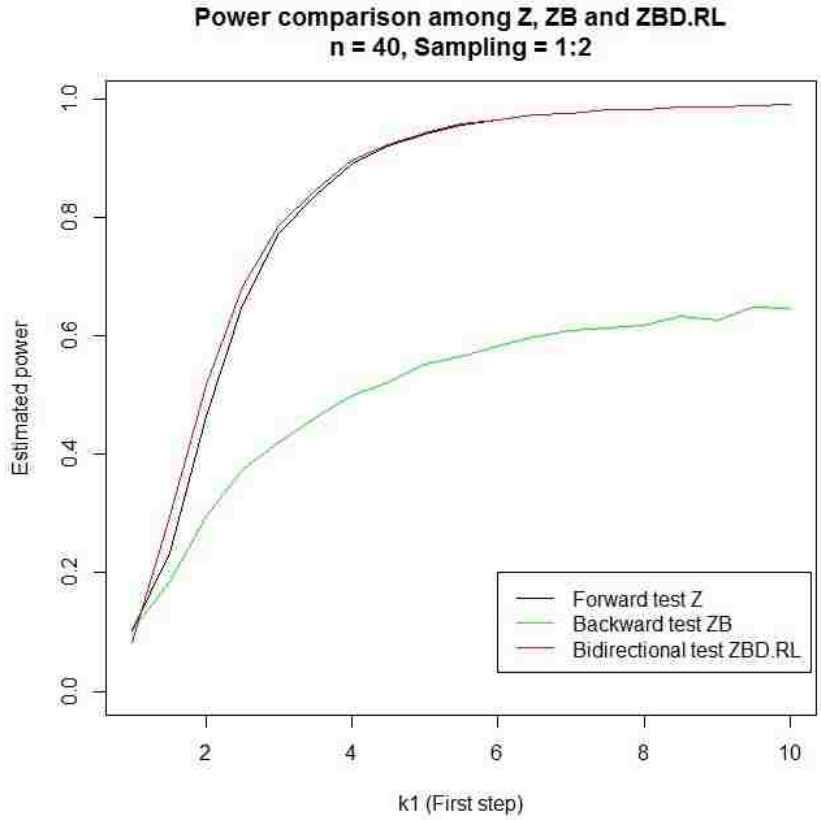


Figure 3.2: Generalized power comparison among Z, Z_B and Z_{BD}^{RL} with $n = 40$ and a $(1 : 2)$ sampling scheme

Once again, we find that easy detection invites similar performance from this bidirectional test and a unidirectional one (Z this time, unlike the previous case), while hard detection establishes the bidirectional relevance.

3.1.2 Random mixing of steps

Absolute knowledge about the trend of the step intensity (and the location of the knot) is a luxury one might not always afford. Thus, our next set of simulation exercise exploits a

50-50 random mix of an increasing $k_1 = 1, k_2 = 3$ step and a decreasing $k_1 = 3, k_2 = 1$ step intensity. Each table records the proportion of increasing and decreasing cases generated.

Table 3.7: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 20/3, combination : (1:1) \times (1:1)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.7120	0.7258	a=	0.0159	0.3000	0.4119	0.1118
		b=	0.1205	0.3212	0.4045	0.0210
		c=	0.2841	0.0000	0.0001	0.3002
		pow=	0.4205	0.6212	0.8165	0.4330

$$\hat{p_{inc}} = 0.4943, \hat{p_{dec}} = 0.5057.$$

Table 3.8: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 13/3, combination : (1:1) \times (1:2)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.6815	0.6217	a=	0.0169	0.2986	0.3827	0.1803
		b=	0.1290	0.2110	0.4105	0.0084
		c=	0.2817	0.0000	0.0002	0.2026
		pow=	0.4276	0.5096	0.7934	0.3913

$$\hat{p_{inc}} = 0.5002, \hat{p_{dec}} = 0.4998.$$

Table 3.9: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 27/3, combination : (1:1) \times (2:1)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.6457	0.7808	a=	0.0164	0.3034	0.3423	0.0493
		b=	0.1276	0.3662	0.4146	0.0732
		c=	0.2870	0.0000	0.0000	0.2930
		pow=	0.4310	0.6696	0.7569	0.4155

$$\hat{p_{inc}} = 0.5051, \hat{p_{dec}} = 0.4949.$$

Table 3.10: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 20/3, combination : (1:2) \times (1:1)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.7956	0.7001	a=	0.0563	0.3818	0.4137	0.1126
		b=	0.0532	0.3214	0.3786	0.0202
		c=	0.3255	0.0000	0.0001	0.3012
		pow=	0.4350	0.7032	0.7924	0.4340

$$\hat{p_{inc}} = 0.4956, \hat{p_{dec}} = 0.5044.$$

Table 3.11: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 13/3, combination : (1:2) \times (1:2)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.7628	0.5819	a=	0.0616	0.3894	0.3734	0.1827
		b=	0.0544	0.1997	0.3822	0.0090
		c=	0.3278	0.0000	0.0000	0.1907
		pow=	0.4438	0.5891	0.7556	0.3824

$$\hat{p_{inc}} = 0.5073, \hat{p_{dec}} = 0.4927.$$

Table 3.12: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 27/3, combination : (1:2) \times (2:1)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.7191	0.7386	a=	0.0585	0.3752	0.3439	0.0543
		b=	0.0547	0.3672	0.3714	0.0776
		c=	0.3167	0.0000	0.0000	0.2896
		pow=	0.4299	0.7424	0.7153	0.4215

$$\hat{p_{inc}} = 0.4913, \hat{p_{dec}} = 0.5069.$$

Table 3.13: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 20/3, combination : (2:1) \times (1:1)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.6081	0.6711	a=	0.0087	0.1953	0.4126	0.1217
		b=	0.1705	0.3140	0.3569	0.0229
		c=	0.1866	0.0000	0.0002	0.2911
		pow=	0.3658	0.5093	0.7697	0.4357

$$\hat{p_{inc}} = 0.5004, \hat{p_{dec}} = 0.4996.$$

Table 3.14: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 13/3, combination : (2:1) \times (1:2)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.5843	0.5678	a=	0.0088	0.1981	0.3862	0.1827
		b=	0.1679	0.2106	0.3572	0.0017
		c=	0.1893	0.0000	0.0000	0.2035
		pow=	0.3660	0.4087	0.7434	0.3933

$$\hat{p_{inc}} = 0.5014, \hat{p_{dec}} = 0.4986.$$

Table 3.15: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 27/3, combination : (2:1) \times (2:1)$

Z power	Z_B power		Z_{BD} powers			
			Z_{BD}^{LR}	Z_{BD}^{LL}	Z_{BD}^{RR}	Z_{BD}^{RL}
0.5307	0.7192	a=	0.0088	0.1876	0.3429	0.0555
		b=	0.1762	0.3642	0.3548	0.0766
		c=	0.1788	0.0000	0.0002	0.2876
		pow=	0.3638	0.5518	0.6979	0.4197

$$\hat{p_{inc}} = 0.5028, \hat{p_{dec}} = 0.4972.$$

We find that in the majority of cases, among the bidirectional proposals, $\phi(Z_{BD}^{RR})$ performs the best. The next set of surface diagrams (of one of the sampling frameworks, viewed from

different orientations) strengthens this conclusion by letting the second step of the increasing intensity and the first step of the decreasing intensity vary systematically.

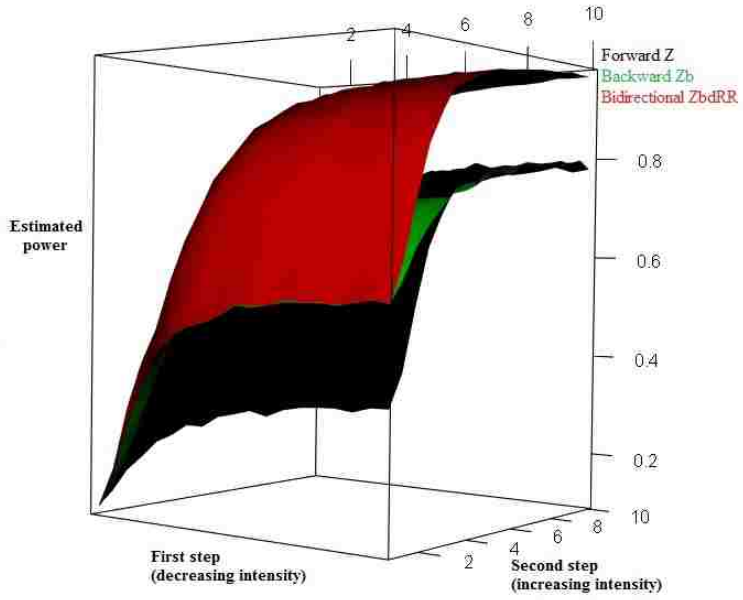


Figure 3.3: Generalized power comparison among Z, Z_B and Z_{BD}^{RR} with $n = 40$ and a $(2:1) \times (1:2)$ sampling scheme

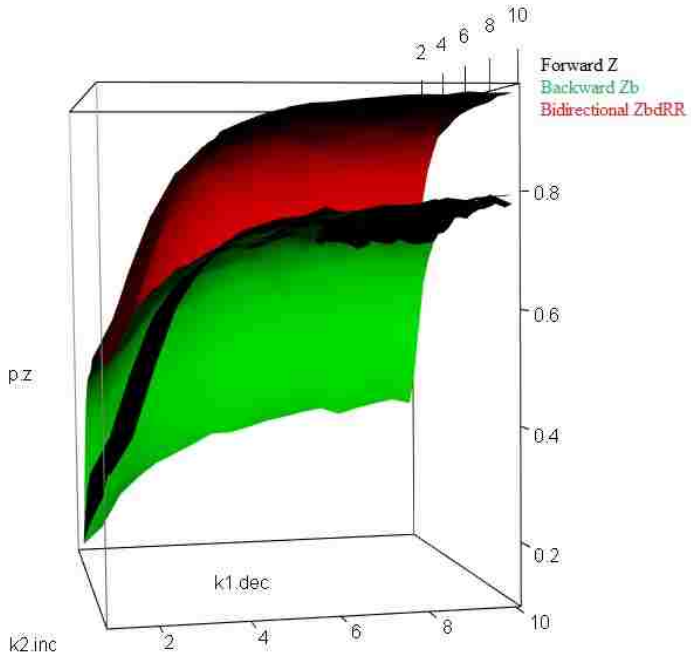


Figure 3.4: Generalized power comparison among Z, Z_B and Z_{BD}^{RR} with $n = 40$ and a $(2:1) \times (1:2)$ sampling scheme

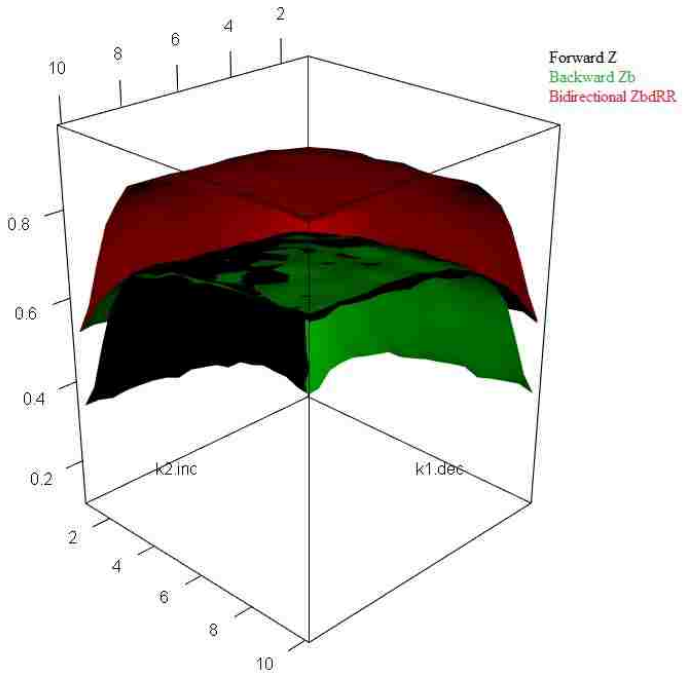


Figure 3.5: Generalized power comparison among Z, Z_B and Z_{BD}^{RR} with $n = 40$ and a $(2:1) \times (1:2)$ sampling scheme

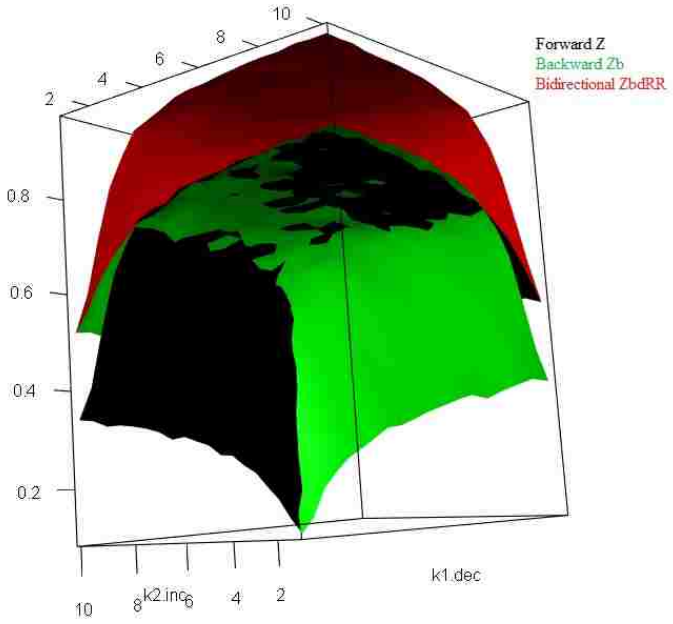


Figure 3.6: Generalized power comparison among Z, Z_B and Z_{BD}^{RR} with $n = 40$ and a $(2:1) \times (1:2)$ sampling scheme

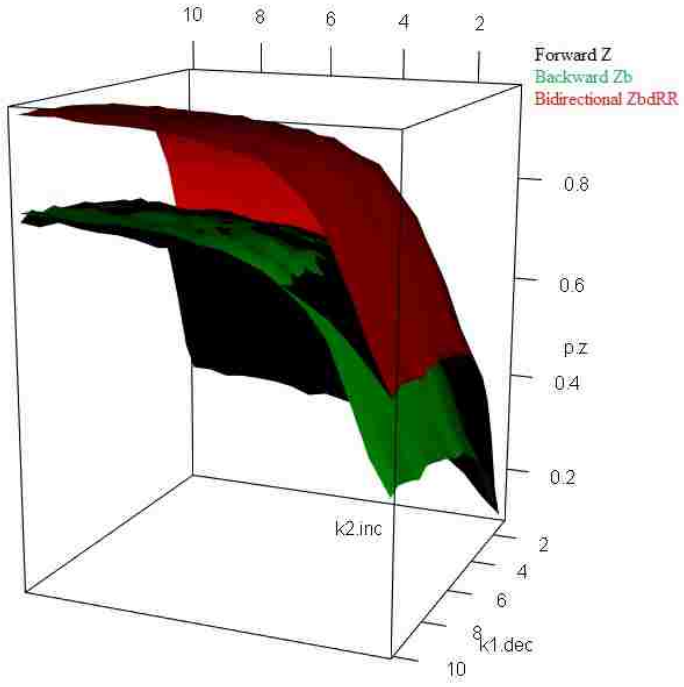


Figure 3.7: Generalized power comparison among Z, Z_B and Z_{BD}^{RR} with $n = 40$ and a $(2:1) \times (1:2)$ sampling scheme

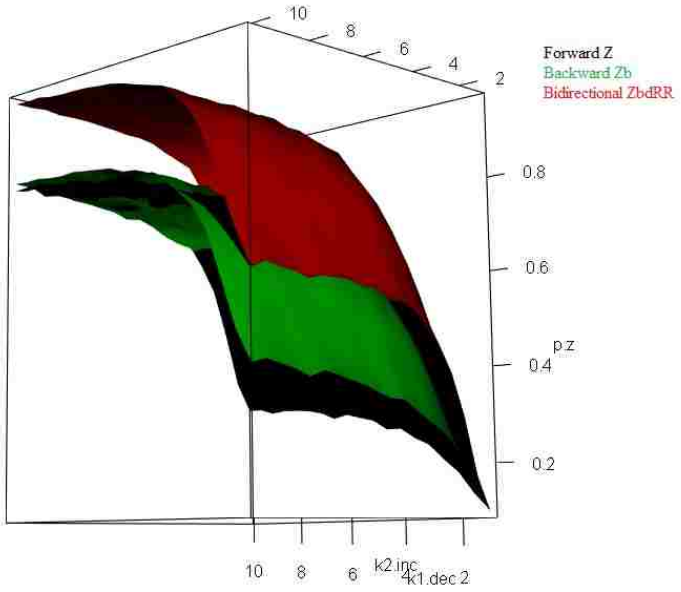


Figure 3.8: Generalized power comparison among Z, Z_B and Z_{BD}^{RR} with $n = 40$ and a $(2:1) \times (1:2)$ sampling scheme

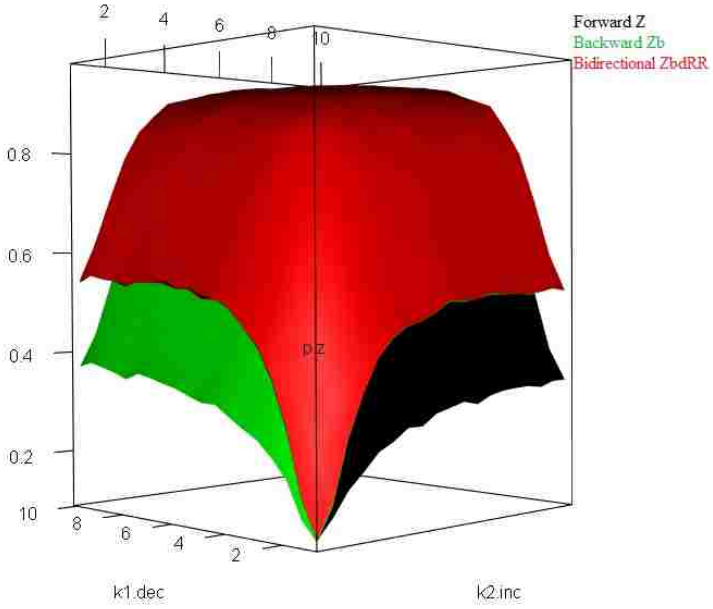


Figure 3.9: Generalized power comparison among Z, Z_B and Z_{BD}^{RR} with $n = 40$ and a $(2:1) \times (1:2)$ sampling scheme

The graphs above demonstrate the power efficiency of $\phi(Z_{BD}^{RR})$ under a late-early place-

ment of knots. It is interesting to note, however, that despite being the most regnant, this superiority is not persistent for every sampling scenario, even within the bidirectional options. There exist cases such as Table(3.12) where $\phi(Z_{BD}^{LL})$ outperforms $\phi(Z_{BD}^{RR})$ and Tables (3.9) and (3.15) where the unidirectional backward test emerges as the best candidate, even over the bidirectional options proposed in this section. Competitions between the unidirectional tests prove interesting too, where with the exception of the 1:2 sampling scheme on the decreasing intensity (i.e., early placement of knots), $\phi(Z_B)$ is outperforming $\phi(Z)$. Thus, with the intention of reducing the schematic complexity and striving towards a uniform winner, we modify the contributing arms of the bidirectional tests next, to propose refined definitions.

3.2 Bi-directional proposals

It is worth recalling that deterioration inflates Z_B (and hence, deflates Z) while improvement inflates Z (and hence deflates Z_B). Thus, under general non-stationarity, the maximum of these two statistics will be large, and the minimum will be small. This observation motivates two one-tailed tests based on $R = max(Z, Z_B)$ and $L = min(Z, Z_B)$ as potential replacements of the two-tailed tests introduced in the last section. In the context of testing non-homogeneity, with the null H_0 asserting stationarity, they are defined as follows.

Definition 3.1. *The maximum based R test is defined as*

$$\phi(R) = \begin{cases} 1 & if \quad R \geq c_\alpha^R \\ 0 & otherwise \end{cases} \tag{3.5}$$

Definition 3.2. *The minimum based L test is defined as*

$$\phi(L) = \begin{cases} 1 & if \quad L \leq c_\alpha^L \\ 0 & otherwise \end{cases} \tag{3.6}$$

where c_α^R and c_α^L are the usual upper and lower α points of the null densities of R and L respectively, summarized for different choices of α, with a maximum sample size of 50, in the Appendix. The prospect of two-tailed tests may still be entertained, with the arms modified as in

Definition 3.3. *The double bidirectional test is defined as*

$$\phi(Z_{DB}) = \begin{cases} 1 & if L \leq c_{\frac{\alpha}{2}}^L \quad or \quad R \geq c_{\frac{\alpha}{2}}^R \\ 0 & otherwise \end{cases} \tag{3.7}$$

Another way of combining contributions from the unidirectional pieces would be by checking the p-values from both the maximum based R test and the minimum based L test and reject the stationarity assumption for extremely low values of the minimum p-value. This option, termed the *dual bidirectional* test is thus, formally defined as

Definition 3.4. *With* $P = min\{P_{H_0}(L \leq l), P_{H_0}(R \geq r)\}$*, the dual bidirectional test is defined as*

$$\phi(P_{DB}) = \begin{cases} 1 & if P \leq p_\alpha \\ 0 & otherwise \end{cases} \tag{3.8}$$

where p_α is the lower α point of the null distribution of P, tabulated in the Appendix.

The null distributions for each of these bidirectional statistics R, L, and P have been derived using a stronger simulation strength of 10^5, ensuring numerical stability across different runs. Figure (3.10) below graphs the null distribution of P, as an example, with a sample of size 40.

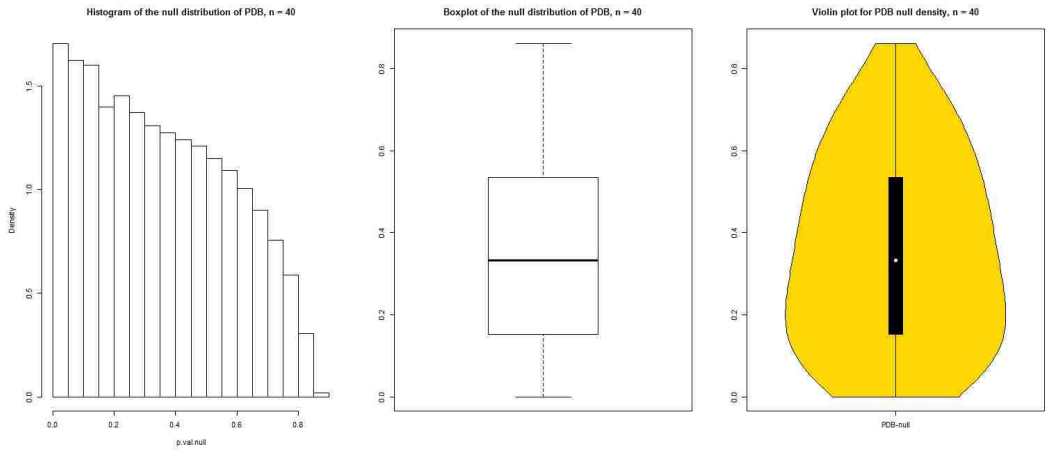


Figure 3.10: Null distribution of dual bidirectional statistic P with $n = 40$

Each of these newly proposed tests value contributions from both Z and Z_B, and offers a new way of combining them. An interesting fact about the two-tailed double bidirectional test $\phi(Z_{BD})$ is that it will almost never be able to achieve a predefined size α since

$$\begin{aligned} E_{H_0}\{\phi(Z_{DB})\} &= P_{H_0}\{L \leq c^L_{\frac{\alpha}{2}} \text{ or } R \geq c^R_{\frac{\alpha}{2}}\} && (3.9)\\ &= P_{H_0}\{L \leq c^L_{\frac{\alpha}{2}}\} + P_{H_0}\{R \geq c^R_{\frac{\alpha}{2}}\} - P_{H_0}\{L \leq c^L_{\frac{\alpha}{2}} \text{ and } R \geq c^R_{\frac{\alpha}{2}}\} && (3.10)\\ &= \frac{\alpha}{2} + \frac{\alpha}{2} - \text{ positive (a.s.)} && (3.11)\\ &< \alpha \text{ (a.s.)} && (3.12) \end{aligned}$$

The intersection term is almost always positive unless we have for instance, $R = L$. Chapter 2 has described *explosive* cases (not considered in this dissertation) under which this may happen. Other tests however, will be able to reach size α since

$$\begin{aligned} E_{H_0}\{\phi(P_{DB})\} &= P_{H_0}\{P \leq p_\alpha\} && (3.13)\\ &= \alpha \text{ (by definition)} && (3.14) \end{aligned}$$

Table (3.16) below, checking the true α value for these two tests, suggests this fact

numerically. This also helps to explain $\phi(P_{DB})$'s slightly better power performance compared to $\phi(Z_{DB})$ in the tables to follow.

Table 3.16: True level checking

Nominal α	Z_{DB} True α	P_{DB} True α
$\alpha = 0.01$	0.0083	0.0098
$\alpha = 0.05$	0.0442	0.0500
$\alpha = 0.1$	0.0905	0.1018

In the next section, we employ these tests by conducting power studies along veins introduced previously.

3.3 Power comparisons

The control parameters and the simulation schemes remain similar to the ones described in the previous section, and hence, will not be repeated here to avoid monotony.

3.3.1 50-50 mixing of two step intensities

Table 3.17: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 20/3, combination : (1:1) \times (1:1)$

Test	Estimated power
Forward test Z	0.7069
Backward test Z_B	0.7244
Right bidirectional test R	0.8124
Left bidirectional test L	0.6193
Double bidirectional test Z_{DB}	0.7685
$P-$dual bidirectional test P_{DB}	0.7858

Table 3.18: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 13/3, combination : (1:1) \times (1:2)$

Test	Estimated power
Forward test Z	0.6881
Backward test Z_B	0.6166
Right bidirectional test R	0.7981
Left bidirectional test L	0.5069
Double bidirectional test Z_{DB}	0.7279
$P-$dual bidirectional test P_{DB}	0.7482

Table 3.19: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 27/3, combination : (1:1) \times (2:1)$

Test	Estimated power
Forward test Z	0.6455
Backward test Z_B	0.7643
Right bidirectional test R	0.7461
Left bidirectional test L	0.6641
Double bidirectional test Z_{DB}	0.7507
$P-$dual bidirectional test P_{DB}	0.7698

Table 3.20: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 20/3, combination : (1:2) \times (1:1)$

Test	Estimated power
Forward test Z	0.7903
Backward test Z_B	0.6963
Right bidirectional test R	0.7845
Left bidirectional test L	0.7027
Double bidirectional test Z_{DB}	0.7825
$P-$dual bidirectional test P_{DB}	0.7968

Table 3.21: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 13/3, combination : (1:2) \times (1:2)$

Test	Estimated power
Forward test Z	0.7596
Backward test Z_B	0.5779
Right bidirectional test R	0.7543
Left bidirectional test L	0.5835
Double bidirectional test Z_{DB}	0.7295
$P-$dual bidirectional test P_{DB}	0.7486

Table 3.22: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 27/3, combination : (1:2) \times (2:1)$

Test	Estimated power
Forward test Z	0.7176
Backward test Z_B	0.7393
Right bidirectional test R	0.7178
Left bidirectional test L	0.7398
Double bidirectional test Z_{DB}	0.7526
$P-$dual bidirectional test P_{DB}	0.7713

Table 3.23: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 20/3, combination : (2:1) \times (1:1)$

Test	Estimated power
Forward test Z	0.6060
Backward test Z_B	0.6715
Right bidirectional test R	0.7671
Left bidirectional test L	0.5104
Double bidirectional test Z_{DB}	0.7008
$P-$dual bidirectional test P_{DB}	0.7231

Table 3.24: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 13/3, combination : (2:1) \times (1:2)$

Test	Estimated power
Forward test Z	0.5691
Backward test Z_B	0.5652
Right bidirectional test R	0.7363
Left bidirectional test L	0.3981
Double bidirectional test Z_{DB}	0.6541
$P-$dual bidirectional test P_{DB}	0.6761

Table 3.25: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 27/3, combination : (2:1) \times (2:1)$

Test	Estimated power
Forward test Z	0.5441
Backward test Z_B	0.7218
Right bidirectional test R	0.7083
Left bidirectional test L	0.5582
Double bidirectional test Z_{DB}	0.6858
$P-$dual bidirectional test P_{DB}	0.7047

3.3.2 Deterministic steps

Steps, either increasing or decreasing, are chosen here with absolute certainty.

Table 3.26: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 1, k_2 = 3, \tau = 20$

Test	Estimated power
Forward test Z	0.6039
Backward test Z_B	0.8185
Right bidirectional test R	0.8195
Left bidirectional test L	0.6032
Double bidirectional test Z_{DB}	0.7615
$P-$dual bidirectional test P_{DB}	0.7803

Table 3.27: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 1, k_2 = 3, \tau = 14$

Test	Estimated power
Forward test Z	0.7417
Backward test Z_B	0.7729
Right bidirectional test R	0.7732
Left bidirectional test L	0.7417
Double bidirectional test Z_{DB}	0.7840
$P-$dual bidirectional test P_{DB}	0.8027

Table 3.28: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 1, k_2 = 3, \tau = 26$

Test	Estimated power
Forward test Z	0.4242
Backward test Z_B	0.7313
Right bidirectional test R	0.7361
Left bidirectional test L	0.4192
Double bidirectional test Z_{DB}	0.6619
$P-$dual bidirectional test P_{DB}	0.6813

Table 3.29: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 3, k_2 = 1, \tau = 20/3$

Test	Estimated power
Forward test Z	0.8154
Backward test Z_B	0.6335
Right bidirectional test R	0.8166
Left bidirectional test L	0.6325
Double bidirectional test Z_{DB}	0.7690
$P-$dual bidirectional test P_{DB}	0.7874

Table 3.30: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 3, k_2 = 1, \tau = 14/3$

Test	Estimated power
Forward test Z	0.7879
Backward test Z_B	0.4514
Right bidirectional test R	0.7910
Left bidirectional test L	0.4489
Double bidirectional test Z_{DB}	0.7061
$P-$dual bidirectional test P_{DB}	0.7279

Table 3.31: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_1 = 3, k_2 = 1, \tau = 26/3$

Test	Estimated power
Forward test Z	0.7248
Backward test Z_B	0.7333
Right bidirectional test R	0.7257
Left bidirectional test L	0.7331
Double bidirectional test Z_{DB}	0.7623
$P-$dual bidirectional test P_{DB}	0.7801

In a spirit similar to Ho (1993) [66], we collect summarized verdicts in the recommendation Table (3.32).

Table 3.32: Prescription for step intensities.

Guarding against	Location	Recommendation
Increasing intensity	Early	$P-$dual bidirectional test P_{DB}
	Midway	Right bidirectional test R
	Late	Right bidirectional test R
Decreasing intensity	Early	Right bidirectional test R
	Midway	Right bidirectional test R
	Late	$P-$dual bidirectional test P_{DB}

Thus for instance, if one suspects an increasing nature of the underlying intensity and an early change, then the best test to confirm it would be the dual bidirectional P_{DB}.

3.3.3 50-50 mixing of two power laws

Competing with the forward Z test under a (deterministic) power law setting would be futile since as noted in Chapter 2, it is UMP under the condition. We found that superiority persists even under a 50-50 mix of an increasing and a decreasing power law.

Table 3.33: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.8678
Backward test Z_B	0.5808
Right bidirectional test R	0.7138
Left bidirectional test L	0.7351
Double bidirectional test Z_{DB}	0.8104
$P-$dual bidirectional test P_{DB}	0.8234

Table 3.34: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, \beta_L = 0.7, \beta_R = 1.3$

Test	Estimated power
Forward test Z	0.6047
Backward test Z_B	0.3479
Right bidirectional test R	0.5039
Left bidirectional test L	0.4493
Double bidirectional test Z_{DB}	0.5295
$P-$dual bidirectional test P_{DB}	0.5519

Table 3.35: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.3523
Backward test Z_B	0.2176
Right bidirectional test R	0.3007
Left bidirectional test L	0.2699
Double bidirectional test Z_{DB}	0.2919
$P-$dual bidirectional test P_{DB}	0.3134

Table 3.36: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, \beta_L = 0.9, \beta_R = 1.1$

Test	Estimated power
Forward test Z	0.1654
Backward test Z_B	0.1286
Right bidirectional test R	0.1513
Left bidirectional test L	0.1430
Double bidirectional test Z_{DB}	0.1373
$P-$dual bidirectional test P_{DB}	0.1523

Note that when detection gets "hard" (as in Table (3.36) with $\beta_L = 0.9, \beta_R = 1.1$ both being very close to 1), our dual bidirectional proposal gets extremely close to the forward test.

3.3.4 25-25-25-25 mixing

A new sampling framework, not considered previously, this represents, for each simulation, choosing an increasing power law, a decreasing power law, an increasing step intensity or a decreasing step intensity with probability 1/4 each. Arguably, this represents the most realistic assumption in the absence of useful knowledge on the governing intensity.

Table 3.37: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 20/3, combination : (1:1) \times (1:1), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.7887
Backward test Z_B	0.6453
Right bidirectional test R	0.7610
Left bidirectional test L	0.6733
Double bidirectional test Z_{DB}	0.7887
$P-$dual bidirectional test P_{DB}	0.8054

Table 3.38: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 20/3, combination : (1:1) \times (1:1), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.5347
Backward test Z_B	0.4676
Right bidirectional test R	0.5634
Left bidirectional test L	0.4396
Double bidirectional test Z_{DB}	0.5322
$P-$dual bidirectional test P_{DB}	0.5509

Table 3.39: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 13/3, combination : (1:1) \times (1:2), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.7738
Backward test Z_B	0.5948
Right bidirectional test R	0.7522
Left bidirectional test L	0.6168
Double bidirectional test Z_{DB}	0.7661
$P-$dual bidirectional test P_{DB}	0.7825

Table 3.40: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 13/3, combination : (1:1) \times (1:2), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.5252
Backward test Z_B	0.4153
Right bidirectional test R	0.5471
Left bidirectional test L	0.3937
Double bidirectional test Z_{DB}	0.5095
$P-$dual bidirectional test P_{DB}	0.5294

Table 3.41: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 27/3, combination : (1:1) \times (2:1), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.7594
Backward test Z_B	0.6666
Right bidirectional test R	0.7288
Left bidirectional test L	0.6971
Double bidirectional test Z_{DB}	0.7853
$P-$dual bidirectional test P_{DB}	0.8006

Table 3.42: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 20, \tau_{dec} = 27/3, combination : (1:1) \times (2:1), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.5055
Backward test Z_B	0.4918
Right bidirectional test R	0.5268
Left bidirectional test L	0.4704
Double bidirectional test Z_{DB}	0.5262
$P-$dual bidirectional test P_{DB}	0.5446

Table 3.43: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 20/3, combination : (1:2) \times (1:1), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.8243
Backward test Z_B	0.6338
Right bidirectional test R	0.7448
Left bidirectional test L	0.7141
Double bidirectional test Z_{DB}	0.7922
$P-$dual bidirectional test P_{DB}	0.8066

Table 3.44: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 20/3, combination : (1:2) \times (1:1), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.5723
Backward test Z_B	0.4476
Right bidirectional test R	0.5406
Left bidirectional test L	0.4794
Double bidirectional test Z_{DB}	0.5364
$P-$dual bidirectional test P_{DB}	0.5560

Table 3.45: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 13/3, combination : (1:2) \times (1:2), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.8107
Backward test Z_B	0.5821
Right bidirectional test R	0.7375
Left bidirectional test L	0.6553
Double bidirectional test Z_{DB}	0.7763
$P-$dual bidirectional test P_{DB}	0.7914

Table 3.46: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 13/3, combination : (1:2) \times (1:2), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.5618
Backward test Z_B	0.4066
Right bidirectional test R	0.5378
Left bidirectional test L	0.4309
Double bidirectional test Z_{DB}	0.5208
$P-$dual bidirectional test P_{DB}	0.5392

Table 3.47: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 27/3, combination : (1:2) \times (2:1), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.7963
Backward test Z_B	0.6576
Right bidirectional test R	0.7189
Left bidirectional test L	0.7358
Double bidirectional test Z_{DB}	0.7869
$P-$dual bidirectional test P_{DB}	0.8050

Table 3.48: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 13, \tau_{dec} = 27/3, combination : (1:2) \times (2:1), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.5369
Backward test Z_B	0.4696
Right bidirectional test R	0.5090
Left bidirectional test L	0.4982
Double bidirectional test Z_{DB}	0.5207
$P-$dual bidirectional test P_{DB}	0.5392

Table 3.49: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 20/3, combination : (2:1) \times (1:1), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.7347
Backward test Z_B	0.6249
Right bidirectional test R	0.7410
Left bidirectional test L	0.6185
Double bidirectional test Z_{DB}	0.7561
$P-$dual bidirectional test P_{DB}	0.7725

Table 3.50: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 20/3, combination : (2:1) \times (1:1), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.4727
Backward test Z_B	0.4381
Right bidirectional test R	0.5311
Left bidirectional test L	0.3797
Double bidirectional test Z_{DB}	0.4882
$P-$dual bidirectional test P_{DB}	0.5083

Table 3.51: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 13/3, combination : (2:1) \times (1:2), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.7180
Backward test Z_B	0.5696
Right bidirectional test R	0.7274
Left bidirectional test L	0.5605
Double bidirectional test Z_{DB}	0.7320
$P-$dual bidirectional test P_{DB}	0.7482

Table 3.52: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 13/3, combination : (2:1) \times (1:2), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.4589
Backward test Z_B	0.3780
Right bidirectional test R	0.5122
Left bidirectional test L	0.3256
Double bidirectional test Z_{DB}	0.4628
$P-$dual bidirectional test P_{DB}	0.4833

Table 3.53: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 27/3, combination : (2:1) \times (2:1), \beta_L = 0.6, \beta_R = 1.5$

Test	Estimated power
Forward test Z	0.6971
Backward test Z_B	0.6399
Right bidirectional test R	0.7025
Left bidirectional test L	0.6347
Double bidirectional test Z_{DB}	0.7422
$P-$dual bidirectional test P_{DB}	0.7576

Table 3.54: Power comparison under failure truncation among the forward, backward and bidirectional test with $n = 40, k_{1inc} = 1, k_{2inc} = 3, k_{1dec} = 3, k_{2dec} = 1, \tau_{inc} = 27, \tau_{dec} = 27/3, combination : (2:1) \times (2:1), \beta_L = 0.8, \beta_R = 1.2$

Test	Estimated power
Forward test Z	0.4361
Backward test Z_B	0.4607
Right bidirectional test R	0.4923
Left bidirectional test L	0.4047
Double bidirectional test Z_{DB}	0.4836
$P-$dual bidirectional test P_{DB}	0.5035

Table (3.55) below condenses these findings under the uniform mixture framework.

Table 3.55: Prescription for uniform mix

Inc- Dec	PLP	Use	Inc- Dec	PLP	Use	Inc- Dec	PLP	Use
Early-Early	Easy	Z	Mid-Early	Easy	P_{DB}	Late-Early	Easy	P_{DB}
	Hard	Z		Hard	R		Hard	P_{DB}
Early-Mid	Easy	Z	Mid-Mid	Easy	P_{DB}	Late-Mid	Easy	P_{DB}
	Hard	Z		Hard	R		Hard	R
Early-Late	Easy	P_{DB}	Mid-Late	Easy	P_{DB}	Late-Late	Easy	P_{DB}
	Hard	P_{DB}		Hard	P_{DB}		Hard	P_{DB}

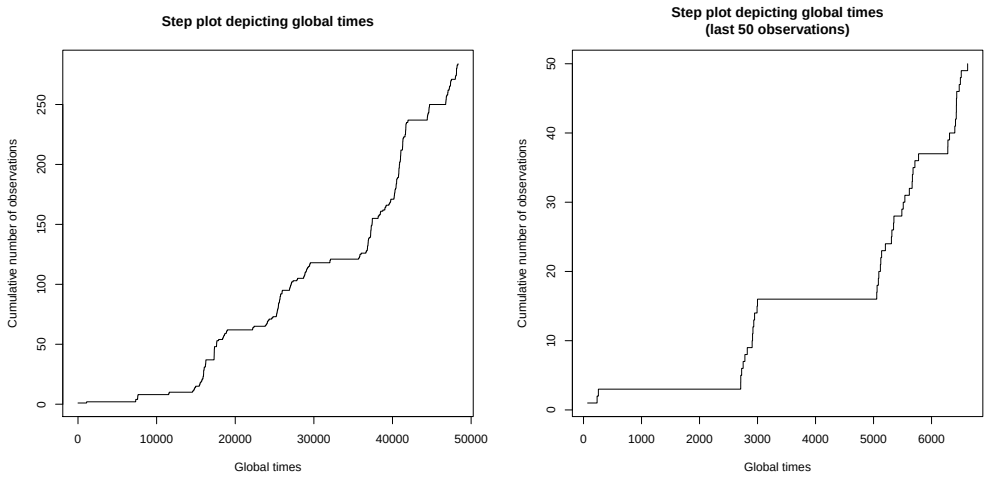
It is interesting to note that most frequently, some member of the bidirectional class triumphs over the only unidirectional test that feature here with comparable power. This holds especially if the suspected changes happen midway or late in the process, regardless

of how easy or hard it is to detect the smooth power-law contribution. Between the top two bidirectional candidates, $\phi(P_{DB})$ occupy the table more frequently than $\phi(R)$, and even at times when it gets beaten by $\phi(R)$, the difference in power is negligible. Thus, it may be argued that in the face of complete ignorance about the nature of the underlying intensity, $\phi(P_{DB})$ offers the best identification performance, and hence may be treated as an "all-purpose" test. It must be pointed out this recommendation is to reduce the categorization burden. An investigator extremely pedantic about the choice is always at liberty to ignore this prescription and refer to Tables (3.32) and (3.55) to choose the test most apt for the condition he is working under.

3.3.5 Case study: Dow Jones Industrial Average

The eruptive patterns of Mt. Etna, described in Chapter 1, was used by Ho (1992) [65] to demonstrate the possibility of creating control-chart type diagrams using χ^2 thresholds for the unidirectional tests. Multiple testing was not considered but was recommended for future work and adjusting the significance level was mentioned to control the overall Type-I error. Our proposed change detection algorithm described later, exploits this control through examining the Benjamini and Hochberg (1995) [14] 's proposed thresholds. We devote the current section, however, to generalize Ho (1992) [65] 's pointwise testing approach to generate control chart type figures, with the newly developed bidirectional tests. The data set we will analyze is the one on Dow Jones Industrial Averages, detailed in the introductory chapter.

The left panel Fig (3.11a) represents the entire data set, spanning 132 years worth of data, starting on Feb 16, 1885. Choosing this day as $t_0 = 0$, we have 283 observations, and



(a) DJIA data set (entire) (b) DJIA data set (last 50)

Figure 3.11: Step diagrams representing Dow Jones Industrial Average closing milestones

the right panel, Fig (3.11b) represents the last 50 of these. A spike represents the global time a closing milestone was achieved. Process deterioration is apparent both from the original and the truncated data set with high β estimates: $\hat{\beta} = 2.156$ (overall) and $\hat{\beta} = 1.859$ (last 50).

Figure (3.12) was generated using χ^2 cutoffs for Z and Z_B

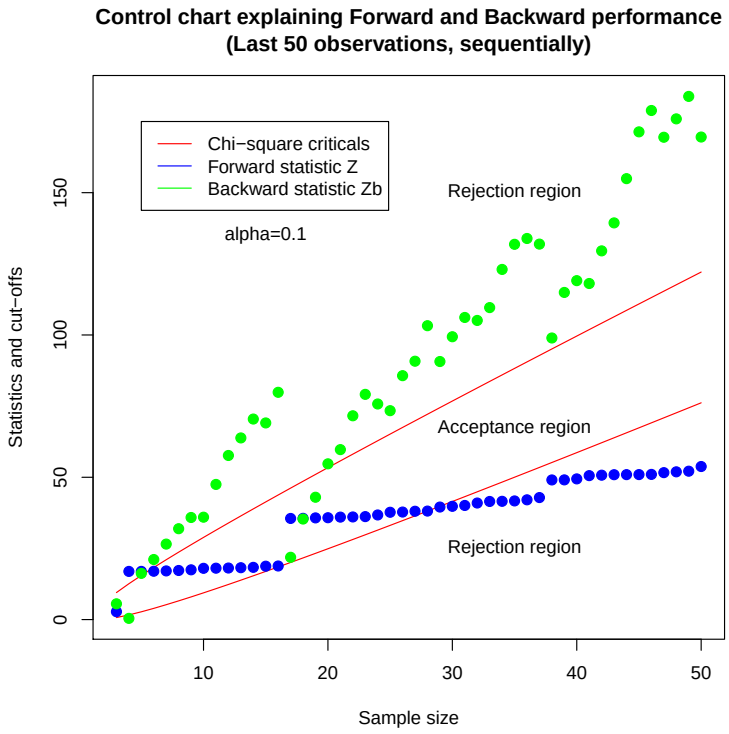


Figure 3.12: Control chart type behavior of Z and Z_B

while Figs (3.13) and (3.14) were created using thresholds from the empirically generated null densities from the bidirectional statistics stored in the Appendix.

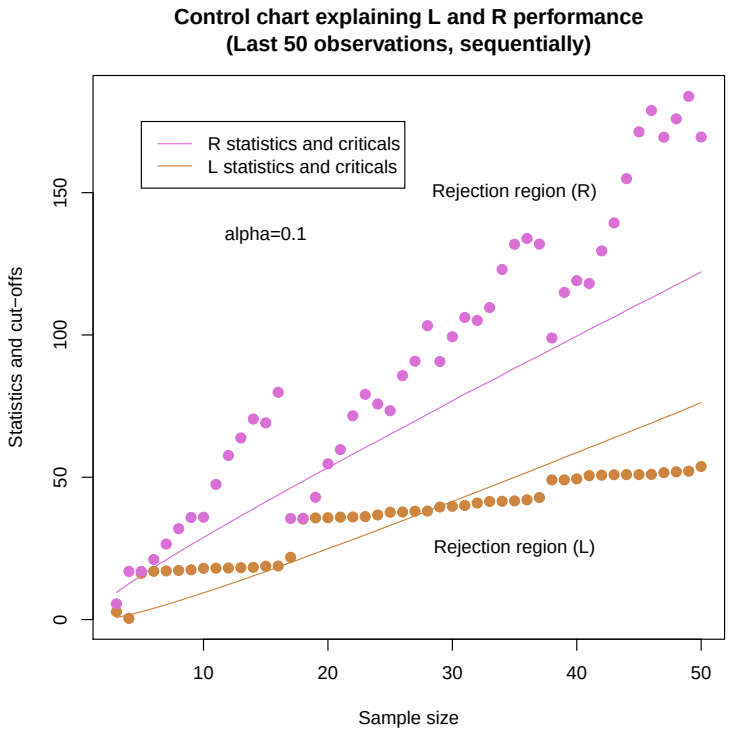


Figure 3.13: Control chart type behavior of L and R

Since these null densities run till samples of size 50, the truncated data set was set to include the 50 most recent observations.

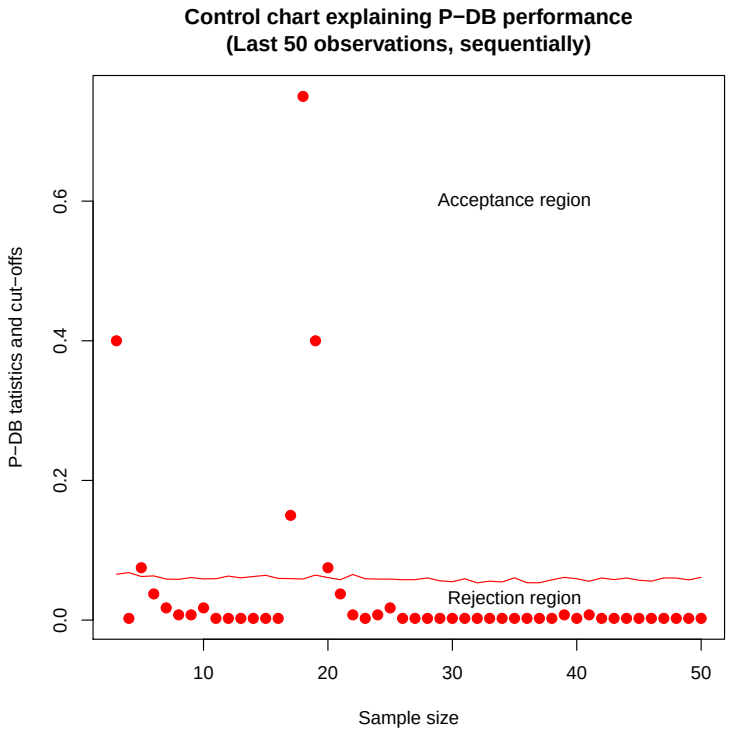


Figure 3.14: Control chart type behavior of P_{DB}

The p-value inspired dual bidirectional test $\phi(P_{DB})$ appears to be the strictest of all, with the highest proportion of rejections. Next, we collect the p-values from the competing tests and stack them alongside a running β estimate in Fig (3.15), to understand how difficult unearthing non-stationarity is around that sample size. A deep dashed line is added horizontally at level 1, suggesting stationarity, and traditional p-value thresholds at the 0.1 and 0.05 levels are also added as faint dashed lines.

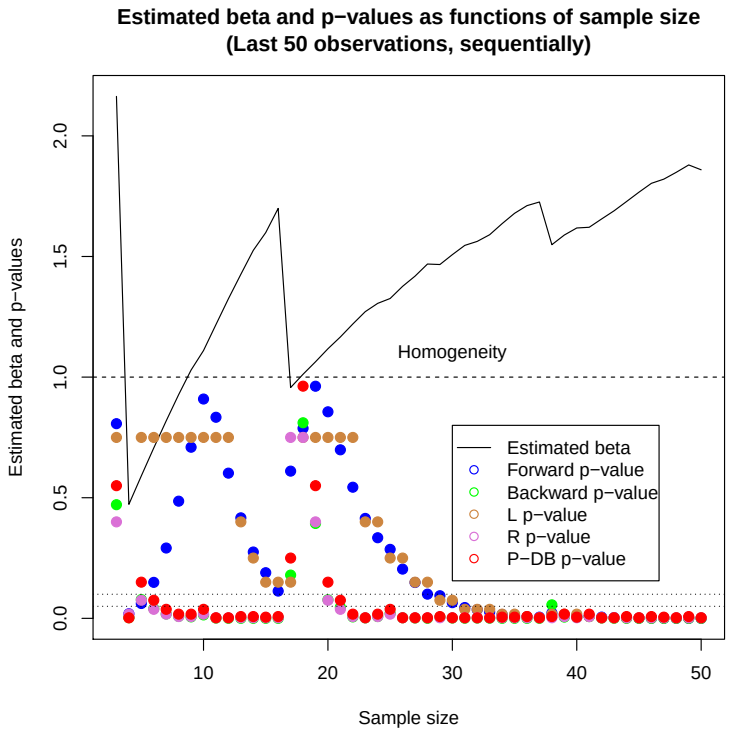


Figure 3.15: p-value comparison among unidirectional and bidirectional members (DJIA data set)

We observe several key facts: For large sample sizes, every test agrees on the non-homogeneity of the process. For smaller sample sizes, however, this unanimity is elusive. For sample sizes around 9 or 10, the β estimate renders the illusion of homogeneity and tests such as $\phi(Z)$ or $\phi(L)$ (with high p-values) are falling prey. Others like $\phi(Z_B), \phi(R)$, and $\phi(P_{DB})$ insist on non-homogeneity. Such inferiority of Z and L is not an artifact of small sample size - this reoccurs around sample size 25 - 27. A similar phenomenon occurs around sample sizes 17 and 18, but here, all of the tests incorrectly claim homogeneity. At sample size 20, with an estimated β of 1.1175, R and Z_B reject the assumption of homogeneity at the 0.1 level, while P_{DB}, L and Z, fail to reject it at that level. At sample size 21, with an estimated β of 1.1665, R and Z_B reject the assumption of homogeneity both at levels 0.1

and 0.05, while P_{DB} rejects it only at the 0.1 level. Z and L fail to reject it.

Next, we have scanned the process to locate instances of possible late and midway changes, to suggest the superiority of $\phi(P_{DB})$ in terms of real data. Our first data set consists of 16 observations, with one distinct jump occurring late in the process, described in Table (3.56).

Table 3.56: DJIA subset containing late jump

Dates	Global Time (T)	Interevent Time (X)
12-Jul-99	0	0
23-Dec-99	164	164
14-Jan-00	186	22
3-Oct-06	2640	2454
4-Oct-06	2641	1
19-Oct-06	2656	15
14-Nov-06	2682	26
14-Dec-06	2712	30
24-Jan-07	2753	41
18-Apr-07	2837	84
25-Apr-07	2844	7
2-May-07	2851	7
16-May-07	2865	14
30-May-07	2879	14
12-Jul-07	2922	43
19-Jul-07	2929	7
5-Mar-13	4985	2056

The shocks (i.e., reaching the closing milestones) can be seen to occur more frequently from the latter half of 2006. The dotplot and the step diagram attached in Fig (3.16) confirm this pattern too.

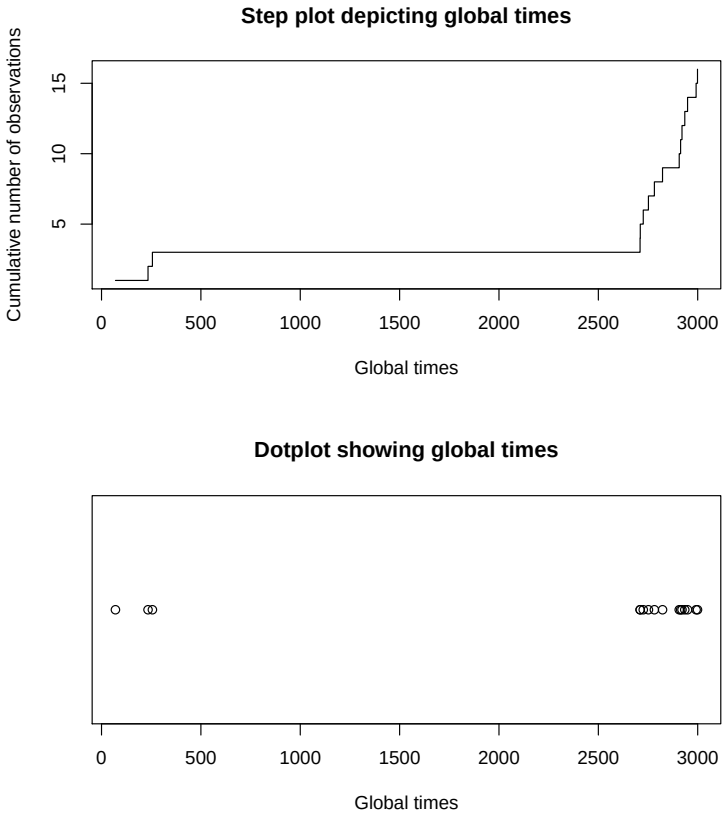


Figure 3.16: Representation of late shock (DJIA data set)

The β estimate here is 1.69935 and we summarize the test comparisons in Table (3.57):

Table 3.57: Performance among competing tests under real late change (DJIA data)

Test	p-value
Z	0.11304
Z_B	$4.14\text{x}10^{-6}$
L	0.15
R	0.0025
P_{DB}	0.0075

Here, the tests using Z and L fail to detect non-homogeneity while the rest are able. Our second data set consists of an earlier sequence of 15 observations with one significant jump occurring somewhat midway through the process. This set is contained in Table (3.58).

Table 3.58: DJIA subset containing midway jump

Dates	Global Time (T)	Interevent Time (X)
17-May-61	0	0
4-Aug-61	79	79
11-Sep-63	847	768
29-Oct-63	895	48
22-Jan-64	980	85
28-Feb-64	1017	37
18-Mar-64	1036	19
2-Jul-64	1142	106
11-Sep-64	1213	71
20-Oct-64	1252	39
28-Jan-65	1352	100
30-Apr-65	1444	92
11-Oct-65	1608	164
29-Oct-65	1626	18
5-Jan-66	1694	68
10-Nov-72	4195	2501

The β estimate here is 2.49823 and we summarize the new comparisons in Table (3.59):

Table 3.59: Performance among competing tests under real midway change (DJIA data)

Test	p-value
Z	0.0073
Z_B	0.07739
L	0.0075
R	0.075
P_{DB}	0.0175

Here, the tests using Z_B and R fail to detect non-homogeneity while the rest are able to (at the 0.05 level). Combining conclusions garnered from the two real subsets above, we might argue that $\phi(P_{DB})$ is a test apt for almost every condition. There exist times when each one of the rest fails to unearth the underlying deteriorating evolution.

3.4 Change point detection

With the efficiency of the bidirectional category and the best member vividly apparent, it is of natural interest to channel P_{DB}'s (or R's) classification prowess through productive avenues. Estimating the points in time around which a sudden cause corrupts an otherwise stable flow, is where, it is felt, these tests will have the most immediate impact.

3.4.1 Algorithm

Our prescription for continuous monitoring of a PtP is to perform a sequence of hypothesis tests and trust the earliest one that detects non-stationarity. The type-I error probability inflation resulting from multiple testing can be controlled in several ways, two of which are described below.

Controlling the Family Wise Error Rate

Given a series of m hypothesis tests, the traditional Bonferroni procedure (Holm 1979) [72] recommends conducting individual tests at a fraction of the overall type-I error probability. Let A_i denote the event that a type-I error has not been committed for the ith test ($i = 1, 2, .., m$) and α_0 be the common significance level for each of the m tests. The Bonferroni inequality gives

$$P(\cap_{i=1}^{m} A_i) \geq \sum_{i=1}^{m} P(A_i) - (m-1) = m(1-\alpha_0) - (m-1) \tag{3.15}$$

Now $P(\cap_{i=1}^{m} A_i) = 1 - P(FWER)$, using DeMorgan's law and noting that $FWER$ stands for making at least one type-I error. Thus to ensure $P(FWER) < \alpha$, we must have $\alpha_0 \leq \frac{\alpha}{m}$. An improved version is the sequential Bonferroni procedure (Holm (1979) [72]) where one

arranges the p-values in ascending order $p_{(i)}$ and requires

$$p_{(r)} \leq \frac{\alpha}{m+1-r} \tag{3.16}$$

to declare all $H_{(i)}$ positive for $i = 1, 2, .., r$. However, large values of m or dependence among tests may make declaring positives extremely unlikely. An improved version thus, is controlling the False Discovery Rates.

Controlling the False Discovery Rates

Introduced by Benjamini and Hochberg (1995) [14], this technique prevents the probability of type-II error inflation and the consequent power reduction, inevitable while controlling the FWER, by controlling the False Discovery Rate (FDR), taken as the average proportion of type-I errors among significant hypotheses. With a given α, one needs to find the largest index k with

$$p_{(k)} \leq \frac{k}{mc(m)}\alpha \tag{3.17}$$

where the $p_{(i)}$'s, as before, represent the ordered p-values. The corresponding hypotheses $H_{(i)}, 1 = 1, 2, .., k$ will be deemed significant. Under direct dependence among the tests, we take $c(m) = 1$, and under inverse dependence, $c(m) = \sum_{i=1}^{m} \frac{1}{i}$.

The change detection algorithm

Inspired by Chen (2010) [24], we propose a similar methodology to detect anomalous behavior, through a control of False Discovery Rates. In contrast to this work where only Z and Z_B were considered, we offer two versions of a more efficient algorithm, using the top bidirectional tests, due to their superior power performance, described previously. The algorithm runs as follows:

i) Fix α and the desirable sample size m. For the unidirectional tests Z and Z_B, m may be taken as n, the size of the entire process. For the bidirectional tests, however, m must be 50 at most, since the null distributions for these test statistics (contained in the Appendix) have been tabulated that far. The first global time is taken as the time origin, and the first test checks whether the flow from this origin to the second event is a stationary (or homogeneous) PtP.

ii) Next, with a chosen statistic (Z, Z_B, R, P_{DB} etc), arrange the hypotheses in order of ascending p-values

$$p_{(1)} < p_{(2)} < ... < p_{(m-1)} \tag{3.18}$$

and find the largest k such that

$$p_{(k)} \leq \frac{k}{m-1}\alpha \tag{3.19}$$

and declare the corresponding $H_{(i)}$s significant, $i = 1, 2, .., k$. Record the earliest time generated by the significant tests.

iii a) *(Weaker)* Let the time of the c th event be the earliest among the significant hypotheses from an "all-purpose" test like P_{DB} (or R, with minimal difference). Then the time of the $(c-1)$th event is the estimated time of change. This is supported by Table (3.55) which demonstrates how under a majority of cases, bidirectional tests may be treated as more powerful than unidirectional ones.

iii b) *(Stronger)* Let the time of the cth event be the earliest among the significant hypotheses from P_{DB} and let the time of the dth event be the earliest among the significant hypotheses from R. Then choosing the time of the $(c-1)$th event or the one of the $(d-1)$th event is governed by Table (3.32), where the estimates of the slope of the underlying intensity

are given by $\hat{\beta}$ defined in Chapter 2. Real data analyses conducted later will clarify both versions.

iv) The first change point identified (if any) will mark the termination of the first regime. This time point, in turn, will serve as the new time origin for another run of the algorithm to detect a (possible) second change point.

v) Repeat the process until no new change points are identified.

The present exercise tries to establish the newly developed tests, notably R and P_{DB}, as efficient change point detection instruments. We shall follow an algorithm studied for regime identification for Mt. Etna (Chen (2010) [24]) and modify it as follows:

a) *The size of the series:* Implementing the algorithm with the forward and backward versions of the Z test was straightforward, owing to the known null distribution of the statistics. For the recent statistics such as R and P_{DB}, tabulated null distribution values (contained in the Appendix) run till $n = 50$. So, we propose to detect change points using sequences of length at most 50. If for any $i \in \{1, 51, 101, ...\}$ the algorithm on the set $\{t_i, t_{i+1}, ..., t_{i+49}\}$ fails to generate a change point, we shall argue that these are part of a stationary phase. Otherwise, the earliest observation will be detected as the change point, and a new regime will start using this as the time origin.

b) *The discreteness of the null distribution:* Once again, unlike Z or Z_B, where probability calculations from the null density can be routinely carried out in softwares such as R, the null distributions of R and P_{DB} have been summarized in discrete tables. To get p-values for instance, we shall have to resort to approximations. We do this in several ways: by averaging the immediate neighbors, by re-simulating the null density with specific choices of the sample size and observation values or by beta regression (similar to a logistic regression).

We highlight the first two in the tables to follow. Generally, the neighbor averaging method proves advantageous when the block of significant tests is well defined. As an example, we will focus on Appendix C, where R's performance in detecting the first change point is stored. The p-value corresponding to the first test, the 41st for instance, in reality, is a number between 0 and 0.005. The null distribution contained in Appendix N (corresponding to a sample of size 42) claims

$$P_{H_0}(R > 122.42251) \approx 0.005 \tag{3.20}$$

while the statistic observed in this case is in excess of 122.42251. The p-value listed is thus, a simple average of 0 and 0.005. Note, however, that even if this number was close to 0.005 (< 0.0092), the tests in the "significant block" would simply undergo a permutation. For instance, tests 41 and 49 might interchange places or every test might move one step up, with 41 occupying 49's position. The earliest significant test, the 41st, would still remain unaltered. A significant test would not turn out to be insignificant or an insignificant test would not become significant under exact calculations. This is because the p-value of the 6th test (the first in the "insignificant block") 0.0175, an average of 0.025 and 0.01, would continue to be insignificant (w.r.t. the corresponding BH threshold 0.01) even if the exact value was extremely close to 0.01. This is what we mean by a "well defined significant block". On the other hand, one might come across cases like Appendix L where, to summarize P_{DB}'s performance, some fine tuning of the averages could be necessary. Tests 2 and 12, both with estimated p-values of 0.0375 sit extremely close to the boundary separating the significant tests from the insignificant ones. Unlike the previous case, the interval $(0.025, 0.05)$ generating this average covers the corresponding BH thresholds, thereby complicating the estimation

process. In such situations, we have re-simulated the null densities (for P_{DB} in this case) with the problematic sample sizes (3 and 13 in this case), and measured the proportion of extremes to get better estimates of these p-values, which are represented in the parentheses alongside. Test 2, with a more exact p-value of 0.043 (still contained in $(0.025, 0.05)$) turns out to be insignificant while test 12, with a more exact p-value of 0.0293 (still contained in $(0.025, 0.05)$) turns out to be significant.

Case study 1: Mount Etna

The first chapter introduced the Italian volcano Mt. Etna as a possible candidate for change point studies. Its eruptive patterns have been studied by Ho (1992) [65] and Chen (2010) [24], among others. The latter of these two conducts regime identification using the unidirectional tests $\phi(Z)$ and $\phi(Z_B)$. Using the same data set, we employ the new bidirectional candidates and the *weaker* version of the proposed algorithm, to achieve similar ends. The results have been summarized in the Appendix (A through H). The significant block has been separated from the non-significant block using horizontal separators.

These are exactly identical to the ones found by Chen (2010). Examining the β estimates we can argue that the Z_B test results are more reliable and conclude that Jan 30, 1974 (the date implied by the 41st test) marks the end of the first and the beginning of the second regime. Results from the newly developed R and P_{DB} tests (using 50 observations) are attached next.

We observe that both the new tests make the 42nd time instant (from the 41st test) Jan 30, 1974, as the first change point. An advantage of using the bidirectional tests with the weaker version of the detection algorithm is that one need not worry about the nature of

the underlying β, unlike a Z and Z_B analysis (Chen (2010) [24]) where its value dictates the reliability and hence, the choice of the more accurate test. This is because previously, it has been established that R or/and P_{DB} is/are more powerful than both of them regardless of the nature of steps (increasing or decreasing). Choosing this time point as the new origin, we start our analyses for the second regime and observe that the R and P_{DB} tests signal the presence of one more volcanic regime for Mt. Etna. The location of the end of the first regime is the same as the backward test. In this way, the technique may be used in batches of size 50, to weed out as many significant change points as there are.

Case study 2: Dow Jones Industrial Averages

We return once again to the DJIA data set, this time with the prospect of estimating the location(s) of the possible change(s). This is in contrast to the previous analysis in section 3.3.5, where we were concerned with checking whether the failure pattern is stable or stationary. Appendices I through L summarize our findings.

Since the potential change points (7, 15, 13th observations) occur early in the process and since the β estimates are more than 1 in this region ($\hat{\beta}_7 = 1.6228, \hat{\beta}_{15} = 2.0462, \hat{\beta}_{13} = 1.9124$), we use Table (3.32) to go with the P_{DB} test conclusion and claim the 13th observation as the possible location of the change point. This is October 20, 1925, and is two time instances ahead of the one identified by R. Note how here we have used the stronger form of the algorithm. Proceeding as in the first case study on Mt. Etna, the second change point identified is the 18th observation in the second block of 50 observations, which is July 2, 1929. This way of weeding changes out may be continued in batches of size 50 till one exhausts the entire data set. Some of the more recent changes detected by P_{DB} however,

merit special mention. Changes located on October 14, 1996 and July 16, 1998 are close to the time points October 21, 1996, and July 13, 1998, respectively, identified as change points by the E-divergence test (James (2014) [75]), one of the competitors studied later. Additionally, the change estimated at January 14, 2000 (identified as January 3, 2000, by the E-divergence method) is likely to be caused by the passing of the Gramm-Leach-Bliley Act, and the one at July 14, 2007 (identified as October 15, 2007, by the E-divergence method) is a possible consequence of the US financial meltdown, triggered by subprime mortgages.

3.4.2 Competitors

A vast array of change-detection algorithms, both from within mainstream statistics and beyond (notably, computer science, described in the concluding chapter), clamor for popularity, and this section, to lay foundations for the performance comparisons to follow, surveys a few relevant candidates. Chapter 1 has defined the identification problem at hand and has touched upon the differences between two distinct approaches: the batch and the sequential methods. What follows, is a necessary elaboration in a spirit similar to Ross (2015) [129].

Batch detection scenario

Change identification here amounts to choosing one of:

$$H_0 : X_i \sim F_0(x; \theta_0), i = 1, 2, .., n \tag{3.21}$$

$$H_1 : X_i \sim \begin{cases} F_0(x; \theta_0), & i = 1, 2, ..., k \\ F_1(x; \theta_1), & i = k+1, k+2, ..n \end{cases} \tag{3.22}$$

where the change (assumed at most one) in the properties of the study variables (such as the inter-event times) is assumed to occur immediately after the kth instance in a fixed sample of size n. Under a parametric setting, this is affected by a parameter update from θ_0 to θ_1.

Conditional on the change point, the variables are independently and identically distributed according to some F_i ($i = 0$ before the change, and $i = 1$ after it).

If the θs represent location or scale parameters, then two sample t or F tests are often used, under the assumption of normality. In the absence of such knowledge, Mann-Whitney or Mood tests can be employed to detect possible location or scale updates, and other non-parametric options such as Lepage, Kolmogorov-Smirnov or Cramer Von-Misses to unearth more intricate structural changes. In principle, a two sample statistic $D_{k,n}$ is agreed upon, extreme values of which signals dissimilarity between the pre-change sample (those before the kth observation) and the post-change sample (those after the kth observation), and consequently, a change in the distribution generating the values. Significant largeness or smallness is quantified through some tolerance level $h_{k,n}$. In practice, a working statistic D_n is created by choosing the largest of the $D_{k,n}$'s:

$$D_n = \max_{k=2,3,..,n-1} D_{k,n} = \max_{k=2,3,..,n-1} \left| \frac{\hat{D}_{k,n} - \mu_{\hat{D}_{k,n}}}{\sigma_{\hat{D}_{k,n}}} \right| \tag{3.23}$$

since the true value of k is unknown, an in fact, almost our target. A natural estimate for the true change point τ will thus, be

$$\tau = \operatorname*{argmax}_{k=2,3,..,n-1} D_{k,n} \tag{3.24}$$

Here $\mu_{\hat{D}_{k,n}}$ and $\sigma_{\hat{D}_{k,n}}$ represent the average and the standard deviations of $\hat{D}_{k,n}$'s. A general formal test runs thus:

$$\phi(D_n) = \begin{cases} 1 & \text{if } D_n > h_n \\ 0 & \text{otherwise} \end{cases} \tag{3.25}$$

where h_n is chosen to satisfy the level α condition, typically as the upper α point of the null density of D_n. Applying the technique, however, often presents problems: the most

notable one being a dearth of closed-form expressions for the null densities of several specific statistics. Asymptotic results for the t and Mann-Whitney choice of the $D_{k,n}$ statistics can be had from Hawkins (1977) [55] and Pettitt (1979) [114]. In addition, Worsley (1982) [146] offers asymptotic bounds for a category of other choices for D_n. The CPM framework mentioned in Chapter 1, and detailed later, exploits numerical simulations to estimate the null densities for small sample sizes.

Sequential detection scenario

In stark contrast to the previous setup, under the sequential setting, observations trickle in continuously in time, rendering the notion of an overall fixed sample size, inoperable. Batch methods, using the two sample statistics described previously, may however be extended if one agrees, as the tth observation x_t arrives, to treat $\{x_1, x_2, .., x_t\}$ as a t-length set, and compare D_t to h_t with $D_t > h_t$ signaling significant change. Hawkins et al. (2003) [57] and Ross et al. (2011) [125] describe how, for most common choices for D_t, the update to D_{t+1} is computationally efficient. Since a sequence of tests is being performed, caution must be exercised to find the thresholds h_t's. Traditionally, they are chosen to make the probability of Type-I error time-homogeneous, i.e.

$$P(D_1 > h_1) = \alpha \tag{3.26}$$

$$P(D_t > h_t | D_{t-1} \leq h_{t-1}, .., D_1 \leq h_1) = \alpha,\ t > 1 \tag{3.27}$$

for a fixed α. Under stationarity, the average run length (ARL_0) defined as the average number of instances scanned before sounding a false alarm is $1/\alpha$. The CPM package in R implements lookup tables condensing the above conditional distributions (created using Monte Carlo simulations) to get the threshold sequence $\{h_t\}$.

Change-Point Model (CPM) framework

While the previous section lays out the two main ways to conduct a change point analysis, the present section describes another source of variation, dictated by the amount of process knowledge. Hawkins et al. (2003) [57], while introducing this CPM framework with a normal choice for F, i.e., with

$$X_i \sim \begin{cases} N(\mu_1; \sigma_1^2), & i = 1, 2, ..., \tau \\ N(\mu_2; \sigma_2^2), & i = \tau + 1, \tau + 2, ..n \end{cases} \tag{3.28}$$

identifies the following scenarios:

i) Complete knowledge about process parameters

Here, the experimenter is supposed to be aware of μ_1 and μ_2 and $\sigma = \sigma_1 = \sigma_2$ completely and is only expected to estimate the change-point location τ. The Cumulative Sum (CUSUM) chart, constructed using:

$$S_0 = 0 \tag{3.29}$$

$$S_i = \max(0, S_{i-1} + X_i - k) \tag{3.30}$$

with $k = \frac{\mu_1 + \mu_2}{2}$ is an apt tool to signal anomalous behaviour. A shift in mean from μ_1 to μ_2 with $\mu_2 > \mu_1$ is indicated if $S_i > h$ where h (constructed using known parameters μ and σ) is created to fix ARL_0 at some predefined level. With the first step $k_1 = 1$ of a rough intensity and $k_2 = 3$, a typical CUSUM chart looks like the following.

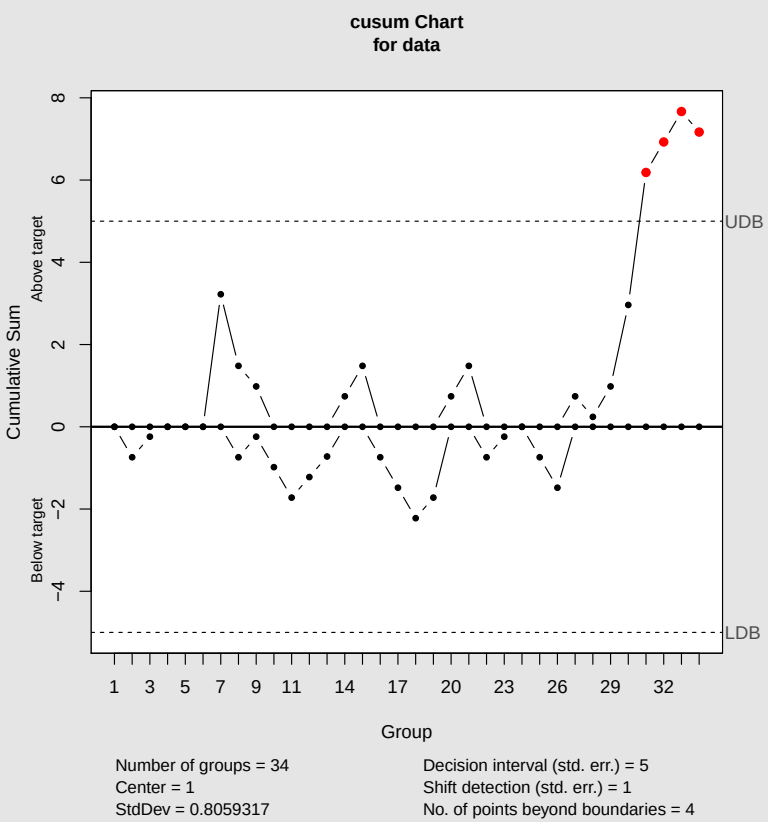


Figure 3.17: CUSUM chart with $k_1 = 1, k_2 = 3$

More about this technique and its theoretical attractiveness can be found in Lai (2001) [87] and Hawkins and Olwell (1998) [56]. The Exponentially Weighted Moving Average (EWMA) procedure is closely related and similarly depends on complete knowledge of process parameters.

ii) Partial knowledge about process parameters

Lai (2001) [87] relaxed the complete knowledge constraint partially by allowing the post-change process mean μ_2 be unknown. MLEs for μ_2 and τ are extracted, and the resulting generalized likelihood ratio test checks the assumption of a change point against the one of a "clean" data set. Pignatiello and Samuel (2001) [115] have worked with an identical framework and Gombay (2000) [52] generalized it even further by allowing the nuisance

parameter σ be unknown.

iii) Complete ignorance about process parameters

Undoubtedly the most realistic of all the assumptions, it is under this framework that Hawkins et al. (2003) [57] put forward their CPM formulation. Essentially, this concerns conducting another generalized likelihood ratio test as follows:

Defining $\bar{X_{jn}} = \frac{\sum_{i=1}^{j} X_i}{j}$ as the pre-change mean, $\bar{X_{jn}}^* = \frac{\sum_{i=j+1}^{n} X_i}{n-j}$ as the post-change mean, and $V_{JN} = \sum_{i=1}^{j}(X_i - \bar{X_{jn}})^2 + \sum_{i=j+1}^{n}(X_i - \bar{X_{jn}}^*)^2$ as the error sum of squares, a traditional two-sample t-statistic for comparing the two means would be

$$T_{jn} = \sqrt{\frac{j(n-j)}{n}} \frac{\bar{X_{jn}} - \bar{X_{jn}}^*}{\hat{\sigma_{jn}}} \tag{3.31}$$

with $\hat{\sigma^2_{jn}} = \frac{V_{jn}}{n-2}$. Under the null assumption of stationarity, $T_{jn} \sim t_{n-2}$. The MLE of the true change-point is thus

$$\hat{\tau} = \underset{1 \le j \le n-1}{\operatorname{argmax}} |T_{jn}| \tag{3.32}$$

and a change is signaled if

$$\max_{1 \le j \le n-1} |T_{jn}| > h_n \tag{3.33}$$

Bonferroni bounds are then employed to estimate the thresholds $\{h_n\}$'s. It is worth mentioning that this is merely a historical record of Hawkins et al. (2003)'s [57] original work. In creating the competitors to follow, the normality assumption (3.28) will be removed, and hence the t statistic (which, for this example can be taken as the D_n candidate) will be replaced by more exotic choices.

Competitors

Quite a handful of tests are available in statistical literature that check whether a Poisson process is non-stationary, or more specifically, whether a trend has crept in the intensity

function that drives the mechanism. In recent times, the problem has attracted attention both from the theoretic (Brodsky (2017) [19]) and pragmatic (Chen and Gupta (2011) [25]) viewpoints. In beautifully crafted review articles, both Antoch and Jaruskova (2007) [3], and Lindqvist (2006) [98] garner an opulent medley, along with their applicability. Tests (such as the CPM-based ones) that are not exclusively designed for Poisson processes may also be applied if the correct incoming variables are chosen (such as the inter-event times instead of the number of shocks observed in a given interval).

1. Laplace test: Encountered previously in Chapter 2, this test is used to check if the data follows a Poisson process with constant intensity against one with monotonic trends. The test statistic is

$$L = \frac{\sum_{i=1}^{n-1} T_i - ((n-1)/2)T_n}{\sqrt{(n-1){T_n}^2/12}}, \tag{3.34}$$

where the symbols have their usual meanings. The asymptotic null distribution of L is normal, and the appropriate quantiles may be treated as critical points.

2. Generalized Anderson Darling (GAD) test: To test the null assumption of a renewal process, Kvaly et al. (2001) [85] introduced the following statistic

$$GAD = \frac{(n-4)\bar{X}^2}{\hat{\sigma}^2} \sum_{i=1}^{n} \left\{ {q_i}^2 \log\left(\frac{i}{i-1}\right) + (q_i + r_i)^2 \log\left(\frac{n-i+1}{n-i}\right) - \frac{r_i^2}{n} \right\}, \tag{3.35}$$

where

$$q_i = \frac{T_i - iX_i}{T_n}, \quad r_i = \frac{nX_i}{T_n} - 1, \quad \hat{\sigma}^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2. \tag{3.36}$$

Kvaloy et al. (2001) [85] recommend using $\hat{\sigma}^2$ as an estimator of σ^2 instead of s^2 since the former is a better choice in the face of trend.

3. Mann test: Another way to check the renewal process assumption against the

existence of monotonic trend is through the Mann statistic

$$M = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} I(X_i < X_j) \tag{3.37}$$

where $I(.)$ is the usual indicator function.If $n < 10$ then one can use lookup tables to get the cutoffs, but for $n \geq 10$, M is rescaled to be standard normally distributed with expectation $\mu = n(n-1)/4$ and variance $\sigma^2 = (2n^3 + 3n^2 - 5n)/72$.

4. Parametric Poisson test: With $\{N_i\}$ representing a sequence of independent Poisson variables with rates $\{\lambda_i\}$s, $i = 1, 2, .., c$, the change point detection under this framework boils down to choosing one of:

$$H_0 : \lambda_1 = \lambda_2 = ... = \lambda_c = \lambda \tag{3.38}$$

$$H_1 : \lambda_1 = \lambda_2 = ... = \lambda_k = \lambda \neq \lambda_{k+1} = \lambda_{k+2} = ... = \lambda_c = \lambda^{'} \tag{3.39}$$

The null likelihood

$$L_0(\lambda) = \prod_{i=1}^{c} \frac{e^{-\lambda} \lambda^{n_i}}{n_i!} \tag{3.40}$$

and the alternate likelihood

$$L_1(\lambda, \lambda^{'}) = \prod_{i=1}^{k} \frac{e^{-\lambda} \lambda^{n_i}}{n_i!} \prod_{i=k+1}^{c} \frac{e^{-\lambda^{'}} \lambda^{'n_i}}{n_i!} \tag{3.41}$$

enable one to construct a likelihood ratio statistic

$$L_k = -2 \log \frac{L_0(\hat{\lambda})}{L_1(\hat{\lambda}, \hat{\lambda^{'}})} \tag{3.42}$$

where the hats represent the usual m.l.es. The optimum change-point position is given by the value of k that maximizes L_k, say $\hat{k}$, and the null assumption is rejected if $L_{\hat{k}} < C$ where C is appropriately chosen to satisfy the level condition. Information on the null

distribution of $\max_k L_k$ can be had from Chen and Gupta (2011) [25]. Generalized versions of these likelihood ratio based tests may be implemented using the `changepoint` package in R, created by Killick and Eckley (2014) [80], where one may control the type of change desired (mean, variance, both, etc), the number of change points, the penalty function, etc.

5. Pettitt's test: Similar to a test introduced later, this location test introduced by Pettitt (1979) [114] was widely used to monitor climatic and hydrological data. With

$$U_{t,T} = \sum_{i=1}^{t} \sum_{j=t+1}^{T} sgn(X_i - X_j) \tag{3.43}$$

where

$$sgn(X_i - X_j) = \begin{cases} 1 & \text{if } X_i > X_j \\ 0 & \text{if } X_i = X_j \\ -1 & \text{if } X_i < X_j \end{cases} \tag{3.44}$$

and

$$K_T = \max_{1 \le t \le T} |U_{t,T}| \tag{3.45}$$

the argument t generating K_T was chosen as the likely estimate of the change point provided K_T was significantly high.

6. Buishand's test: This is a test for a location shift, assuming one change point, using the statistic

$$U = \frac{1}{n(n+1)} \sum_{k=1}^{n-1} \left(\frac{S_k}{D_x} \right)^2 \tag{3.46}$$

with D_x as the standard deviation of the X variables, and $S_k = \sum_{i=1}^{k} (X_i - \bar{X})$ representing the cumulative deviations (note that $\bar{X}$ is calculated on the entire data set). The p-values are estimated using Monte Carlo simulations. Buishand (1982) [20] describes this test in detail, along with some of its real applications.

7. CPM-Exp test: Using $Exponential(\lambda_0)$ and $Exponential(\lambda_1)$ choices for F_0 and F_1 in (3.22), Ross (2014) [128] has constructed expressions for the generalized likelihood ratio

statistic

$$M_{k,n} = -2\log(\frac{L_0}{L_1}) \tag{3.47}$$

where L_0 and L_1 denote the maximized likelihoods under the null and the alternate hypothesis, respectively. $M_{k,n}$ can then be taken as $D_{k,n}$ under the general CPM framework.

8. CPM-Adjusted Exp test: As n explodes, it can be shown (Ross (2014) [128]) that the average of $M_{k,n}$ defined previously, approaches $-2k\{\psi(k) - \log(k)\}$, which need not be 1, the expectation of a chi-square variable with one degree of freedom. Here $\psi(k) = \frac{\Gamma(k)}{\Gamma'(k)}$ is the usual digamma function. To rectify this, Ross (2014) [128] scaled $M_{k,n}$ down as

$$M^c_{k,n} = \frac{M_{k,n}}{E(M_{k,n})} \tag{3.48}$$

which makes the mean hover around 1. $D_{k,n}$ in the original CPM framework may thus, now be played by $M^c_{k,n}$.

9. CPM-Mann-Whitney test: This relies on Pettitt's (1979) [114] proposal of a U statistic based on the Mann-Whitney two-sample test:

$$U_{k,n} = \sum_{i=1}^{k}\sum_{j=k+1}^{n} P_{ij} \quad 1 \leq k \leq n-1 \tag{3.49}$$

where

$$P_{ij} = sgn(X_i - X_j) = \begin{cases} 1 & \text{if } X_i > X_j \\ 0 & \text{if } X_i = X_j \\ -1 & \text{if } X_i < X_j \end{cases} \tag{3.50}$$

Conover (1999) [27] relates $U_{k,n}$ to the rank of X_i, i.e. R_i as

$$U_{k,n} = 2\sum_{i=1}^{k} R_i - k(n+1) \tag{3.51}$$

implying

$$E(U_{k,n}) = 0, \quad Var(U_{k,n}) = \frac{k(n-k)(n+1)}{3} \tag{3.52}$$

Thus, $D_{k,n}$ in the CPM framework may be taken as

$$D_{k,n} = \frac{U_{k,n}}{\sqrt{k(n-k)(n+1)/3}} \tag{3.53}$$

More on this technique can be found in Hawkins and Deng (2010) [59].

10. CPM-Mood test: The Mood test developed by Mood (1954) [106] is efficient in detecting scale parameter shifts, with reasonable power performance, observed by Duran (1976) [37]. Defining the rank of the ith observation as

$$r(X_i) = \sum_{i \neq j}^{n} I(X_i \geq X_j) \tag{3.54}$$

the statistic quantifies the amount of discrepancy between the rank of a point and its average

$$M^{'} = \sum_{X_i} \left(r(X_i) - \frac{n+1}{2} \right)^2. \tag{3.55}$$

A standardized version of $M^{'}$, namely M, can then be taken as D_n in the CPM framework. Details about the standardization can be had from Ross et al. (2011) [125].

11. CPM-Lepage test: With the Mann-Whitney test designed to detect location changes and the Mood test to detect scale shifts, a need is often felt to combine the two and create a test efficient for both aspects. Lepage-type tests (Lepage (1971) [92]) offers an alternative by using

$$L = U^2 + M^2 \tag{3.56}$$

with U and M defined previously. L can then be incorporated into the CPM framework. More on this test can be found in Ross et al. (2011) [125].

12. CPM-Kolmogorov-Smirnov test: This exploits the comparison between the empirical distribution functions of the pre-change and the post-change sample defined as:

$$\hat{F}_{S_1}(x) = \frac{1}{k} \sum_{i=1}^{k} I(X_i \leq x) \tag{3.57}$$

$$\hat{F}_{S_2}(x) = \frac{1}{n-k} \sum_{i=k+1}^{n} I(X_i \leq x) \tag{3.58}$$

and $D_{k,n}$ in the CPM framework is taken as:

$$D_{k,n} = \sup_x |\hat{F}_{S_1}(x) - \hat{F}_{S_2}(x)| \tag{3.59}$$

Techniques for standardization can be had from Ross and Adams (2012) [127].

13. CPM-Cramer-Von-Mises test: This uses the square of the average distance to quantify discrepancy between the two empirical functions. $D_{k,n}$ in the CPM framework is now:

$$D_{k,n} = \int_{-\infty}^{\infty} |\hat{F}_{S_1} - \hat{F}_{S_2}| dF_t(x) \tag{3.60}$$

with $F_t(.)$ standing for the empirical c.d.f. for the pooled sample. For implementation purposes, one may use:

$$D_{k,n} = \sum_{i=1}^{n} |\hat{F}_{S_1}(X_i) - \hat{F}_{S_2}(X_i)|^2 \tag{3.61}$$

Ross and Adams (2012) [127] may be consulted for the necessary standardization.

14. E-divergence test: A technique originally developed by Matteson and James (2013) [103] to detect any number of change points in multivariate time series observations, it is distribution free and is capable of detecting changes of several kinds. A priori knowledge on the number of change points is not required, however, the observations must be independent, and have finite αth absolute moments with $\alpha \in (0, 2]$. It uses Szekely and Rizzo (2005, 2010) [140] [141]'s divergence measure to check whether two vectors $X, Y \in R^d$ with characteristic functions $\phi_X(t)$ and $\phi_Y(t)$ are identically distributed. This is often viewed as an *energy* statistic, hence the name "E-divergence". Using Matteson and James (2013)'s [103] proposal

of a specific weight function, the measure takes the form

$$D(X,Y;\alpha) = \int_{R^d} |\phi_X(t) - \phi_Y(t)|^2 \left(\frac{2\pi^{d/2}\Gamma(1-\alpha/2)}{\alpha 2^\alpha \Gamma((d+\alpha)/2)} |t|^{d+\alpha} \right)^{-1} dt \tag{3.62}$$

The null assumption of similarity is rejected for exceedingly high values of this divergence. James and Matteson (2013, 2014) [103] [75] introduce a binary tree based bisection algorithm called "E-divisive" for hierarchical divisive change point estimation. The significance of an estimated change point and its corresponding p-value is found through permuting the observation collected thus far.

To choose the best competitors and motivate the error analyses to follow, we have done preliminary calculations such as the one in Table (3.60) to check the time of detection as a function of increasing step heights.

Table 3.60: Average time to detection comparisons among Laplace, Mann, GAD and the unidirectional and bidirectional tests with different step sizes, $n = 40, \tau = 4$

Test	$k_2 = 2$	$k_2 = 3$	$k_2 = 4$
LAP	17.73	17.32	16.60
GAD	10.09	9.76	9.59
MANN	17.95	17.62	17.27
Z	25.08	23.42	23.44
Z_B	13.96	11.7	11.62
L	25.21	25.2	25.32
R	10.62	9.76	7.67
Z_{DB}	4.04	4.1	3.09

As the step heights become larger, detection becomes easier (i.e., quicker), and under every situation, we find that the bi-directional proposals are outperforming the established tests in terms of quicker detection times. From now on, we thus ignore the first three competitors Laplace, Mann, and GAD tests and focus on the remaining (more recently proposed) change detection competitors described before. One instance was generated in

each of the cases described in Tables (3.61), (3.62), and (3.63) to check the effectiveness of the CPM class.

Table 3.61: Change detection comparison among the CPM class, the unidirectional and bidirectional tests with $n = 50, \tau = 12, \alpha = 0.05, k_1 = 1, k_2 = 3$

Test	Obs. closest to estimated change point	Time of change
CPM-Exp	None	None
CPM-Adjusted Exp	None	None
CPM-Mann-Whitney	None	None
CPM-Mood	None	None
CPM-Lepage	None	None
CPM-Kolmogorov-Smirnov	None	None
CPM-CramerVon-Mises	None	None
Z	None	None
Z_B	31	18.329
R	31	18.329
P_{DB}	31	18.329

Table 3.62: Change detection comparison among the CPM class, the unidirectional and bidirectional tests with $n = 50, \tau = 25, \alpha = 0.05, k_1 = 1, k_2 = 3$

Test	Obs. closest to estimated change point	Time of change
CPM-Exp	28	29.837
CPM-Adjusted Exp	28	29.837
CPM-Mann-Whitney	22	26.092
CPM-Mood	46	35.359
CPM-Lepage	46	35.359
CPM-Kolmogorov-Smirnov	28	29.837
CPM-CramerVon-Mises	28	29.837
Z	None	None
Z_B	31	31.743
R	31	31.743
P_{DB}	37	33.093

Table 3.63: Change detection comparison among the CPM class, the unidirectional and bidirectional tests with $n = 50, \tau = 36, \alpha = 0.05, k_1 = 1, k_2 = 3$

Test	Obs. closest to estimated change point	Time of change
CPM-Exp	36	36.731
CPM-Adjusted Exp	36	36.731
CPM-Mann-Whitney	36	36.731
CPM-Mood	None	None
CPM-Lepage	36	36.731
CPM-Kolmogorov-Smirnov	36	36.731
CPM-CramerVon-Mises	36	36.731
Z	None	None
Z_B	44	38.089
R	43	37.933
P_{DB}	44	38.089

These isolated instances suggest that the CPM class may produce better or worse results compared to our proposals depending on the location of the knot. The next section strengthens these facts through simulations and varying the step heights.

3.4.3 Estimation performance

The performance of change point techniques may be compared using a host of different metrics, depending on the nature of decisions generated. For those that exploit a hypothesis testing framework (similar to binary classifiers) detecting only the presence or absence of a change, one may use measures such as Accuracy, Sensitivity, Precision, etc. defined as

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{3.63}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.64}$$

$$Precision = \frac{TP}{TP + FP} \tag{3.65}$$

where TP, FP, TN, FN represent the true positives, false positives, true negatives, false negatives, respectively. The power analyses conducted previously in this chapter perform a similar role. Aminikhanghahi (2017) [2] and Cook (2015) [28] survey a class of other measures. But for those that transcend mere classification to estimating the possible change, the performance measures are slightly different. The most prevalent options (Aminikhanghahi (2017) [2]) to quantify the performance of offline or retrospective change point estimation algorithms are:

1) Mean Absolute Error/Difference (MAE/D) defined as:

$$MAE = \frac{\sum_i |\text{Actual change point}_i - \text{Estimated change point}_i|}{\text{number of change points}} \tag{3.66}$$

2) Root Mean Square Error (RMSE) defined as

$$RMSE = \sqrt{\frac{\sum_i (\text{Actual change point}_i - \text{Estimated change point}_i)^2}{\text{number of change points}}} \tag{3.67}$$

3) Mean Signed Difference (MSD) defined as

$$MSD = \frac{\sum_i (\text{Estimated change point}_i) - \text{Actual change point}_i}{\text{number of change points}} \tag{3.68}$$

Each serves a different purpose (e.g., MSD with an algebraic sign informs whether the algorithm signals non-stationarity before or after the actual change) and the competitors will now be evaluated in the light of these metrics. The estimated change point time distributions under different simulated environments (i.e., early, midway, and late placement of knots) have been graphed through violin plots with connectors joining the deciles for ready reference. A

violin plot is often treated as a hybrid of a traditional boxplot and a kernel density plot, and depict hidden concentrations around specific locations. The fact that some of the algorithms might generate multimodal change time densities led us to choose these plots over the usual boxplot type representations. The horizontal dashed line in each case represents the true time of change. We have carried out the analyses initially with a fixed height for the second step: $k_2 = 3$, under an increasing intensity framework, $k_2 = 1$ under a decreasing intensity framework, and have recorded the performance of the competitors using 200 simulations, samples of size 50, and an α value of 0.05. Such results have been summarized in Tables (3.64) - (3.66) (correspondingly Figs (3.18) - (3.20)) and Tables (3.67) - (3.69) (Figs (3.30) - (3.32)).

Next, we have systematically varied the height of the second step to check how the algorithms react to change detection scenarios more difficult than $k_2 = 3$. These results have been condensed in Figs 3.21 - 3.29. Several comments on them are in order:

i) Under an increasing step intensity framework, higher values of the second step correspond to easier change detection. Hence, one would expect the error amount to diminish with an increase in k_2. This explains the downward gradient in Figs 3.21 - 3.29. Moreover, a change early in the process is easier to detect which is why the slopes get less steep as we move from early to midway to late placement of knots.

ii) A high error value at $k_2 = 1$ (which represents the null environment of stationarity), and a rapid descent are signatures of an efficient algorithm. All our proposals behave accordingly.

iii) As evidenced by the MSD graphs, our proposals, like most of their competitors, have error curves lying above the horizontal line at the null error which represents detecting

non-stationarity *after* the true change has occurred. This is intuitive and desirable since the new environment must take some time to be detected.

iv) Our best proposal P_{DB} is most efficient when detection of non-stationarity is hard, especially in the region $k_2 \in (1, 1.5)$, when it outperforms its most worthy competitors, notably the parametric one and CPM-Exp or CPM-AdjustedExp, in terms of smaller average errors. Over other domains, it remains competitive. That the parametric ones (Parametric, CPM-Exp, CPM-AdjustedExp) will perform well is no accident since our simulation prescription fits them perfectly: the inter-event times from the two homogeneous pieces are exponentially distributed by construction. Our proposals, however, are free of such confines and perform just as well.

A note on the implementation aspects of these algorithms: All the CPM-based candidates have been run using the `cpm` package in R introduced by Ross (2015) [129]. The parametric Poisson based test, using the `changepoint` package in R introduced by Killick and Eckley (2014) [80], the Pettitt and Buish tests using the `trend` package in R introduced by Pohlert (2018) [116], and the E-divergence test using the `ecp` package in R introduced by James and Matteson (2014) [75]. In addition to the unidirectional forward Z and backward Z_B tests, we have included the top bidirectional candidates, the maximum based R test and the p-value inspired dual P_{DB} test. The null distributions for the last two, used to run our proposed algorithm, can be found as tables in the Appendix.

Increasing step intensity

Table 3.64: Change point detection comparison, $n = 50, \tau = 12, \alpha = 0.05, k_1 = 1, k_2 = 3$

Test	Q_1	Q_2	Q_3	MAE	MSE	MSD
Parametric	10.55	11.56	13.37	2.594	16.16	0.299
Pettitt	12.29	13.24	15.46	2.567	15.11	2.259
Buish	12.01	12.46	13.56	1.413	5.17	1.015
CPM-Exp	11.88	12.89	21.60	4.594	53.85	3.935
CPM-Adjusted Exp	11.88	12.74	18.14	4.302	49.88	3.666
CPM-Mann-Whitney	12.26	15.59	24.27	6.501	79.43	6.018
CPM-Mood	20.38	23.80	25.64	10.579	135.58	10.116
CPM-Lepage	12.14	14.16	24.30	6.539	81.34	5.838
CPM-Kolmogorov-Smirnov	12.53	21.65	25.15	8.125	103.67	7.779
CPM-CramerVon-Mises	0	21.68	24.57	12.160	150.85	1.840
E-Divergence	12.54	22.05	25.15	8.384	105.33	7.690
Z	15.18	20.41	24.32	8.536	96.79	7.609
Z_B	14.00	16.81	23.97	7.027	78.39	6.612
R	13.44	15.28	21.13	5.756	55.01	4.727
P_{DB}	12.32	14.08	19.13	5.597	51.77	2.948

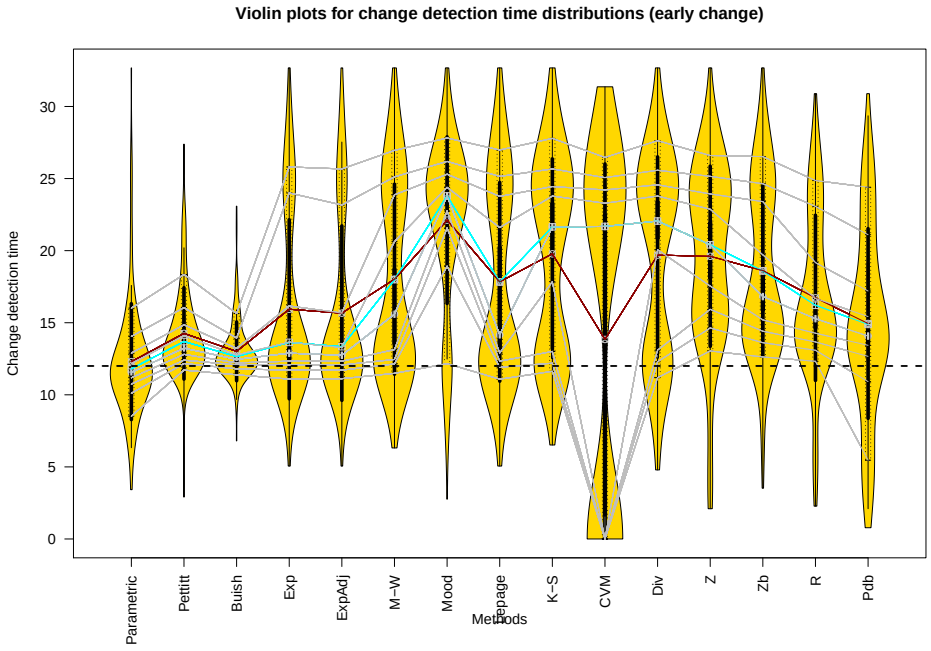


Figure 3.18: Estimated time of change distributions with $\tau = 12$ (early), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 3$

Table 3.65: Change point detection comparison, $n = 50, \tau = 25, \alpha = 0.05, k_1 = 1, k_2 = 3$

Test	Q_1	Q_2	Q_3	MAE	MSE	MSD
Parametric	22.87	24.35	25.85	2.673	17.86	-0.974
Pettitt	24.84	25.29	26.32	1.735	7.39	0.108
Buish	23.00	24.89	25.53	1.971	9.96	-0.949
CPM-Exp	24.72	25.49	28.26	3.066	23.63	1.562
CPM-Adjusted Exp	24.69	25.46	27.79	2.721	17.99	1.442
CPM-Mann-Whitney	25.06	26.40	31.75	4.179	32.74	2.649
CPM-Mood	31.19	32.68	34.31	8.243	75.73	6.606
CPM-Lepage	24.88	26.65	32.39	4.864	41.37	2.652
CPM-Kolmogorov-Smirnov	25.46	30.52	33.23	5.429	45.37	4.367
CPM-CramerVon-Mises	0	31.94	33.94	13.699	251.34	-2.300
E-Divergence	30.21	32.50	34.16	7.729	72.56	5.804
Z	30.70	32.37	34.19	7.932	73.06	6.624
Z_B	26.89	30.39	32.70	6.539	63.02	3.545
R	25.94	28.57	31.23	6.502	70.64	1.444
P_{DB}	18.60	27.81	30.86	7.468	91.82	-0.589

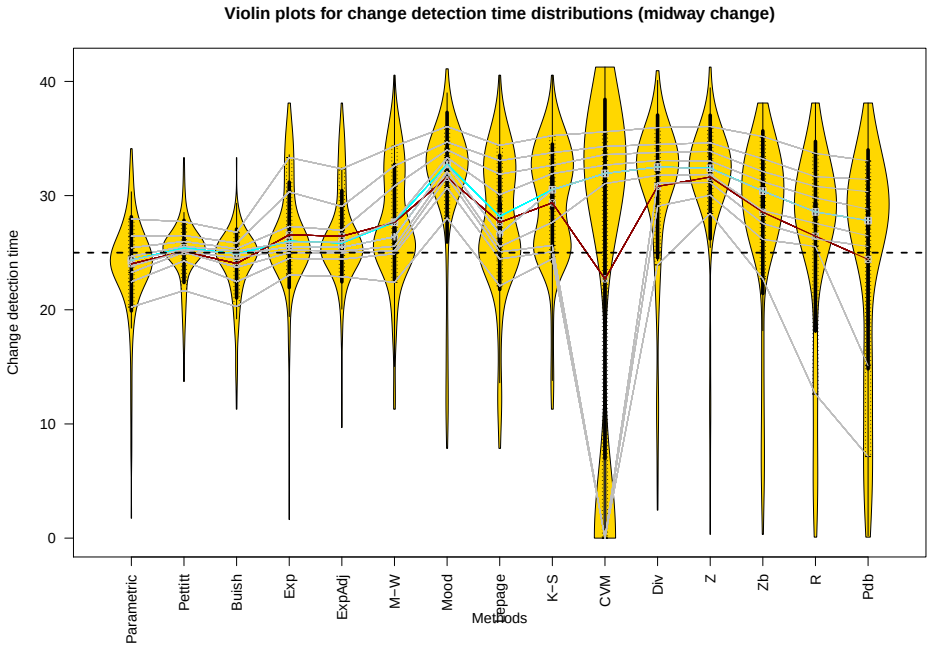


Figure 3.19: Estimated time of change distributions with $\tau = 25$ (midway), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 3$

Table 3.66: Change point detection comparison, $n = 50, \tau = 40, \alpha = 0.05, k_1 = 1, k_2 = 3$

Test	Q_1	Q_2	Q_3	MAE	MSE	MSD
Parametric	37.29	39.59	40.84	3.899	68.33	-2.239
Pettitt	24.17	35.61	39.94	9.201	193.36	-8.760
Buish	21.36	33.51	39.20	10.545	221.48	-10.351
CPM-Exp	39.94	40.93	42.60	3.214	38.21	0.136
CPM-Adjusted Exp	39.91	40.84	42.55	3.150	37.85	0.058
CPM-Mann-Whitney	39.91	41.20	43.36	3.517	41.66	0.369
CPM-Mood	41.28	43.13	44.85	4.139	31.84	2.498
CPM-Lepage	40.16	41.73	44.02	3.664	38.55	1.059
CPM-Kolmogorov-Smirnov	40.34	41.97	44.09	3.667	32.71	1.348
CPM-CramerVon-Mises	0	0	43.06	26.619	1009.15	-23.431
E-Divergence	40.87	42.57	44.64	4.639	46.29	1.284
Z	41.26	43.09	44.77	4.141	37.34	2.361
Z_B	40.99	42.31	44.09	4.179	50.86	1.313
R	40.79	42.04	43.60	4.966	83.16	-0.224
P_{DB}	40.27	41.89	43.44	6.058	113.85	-1.588

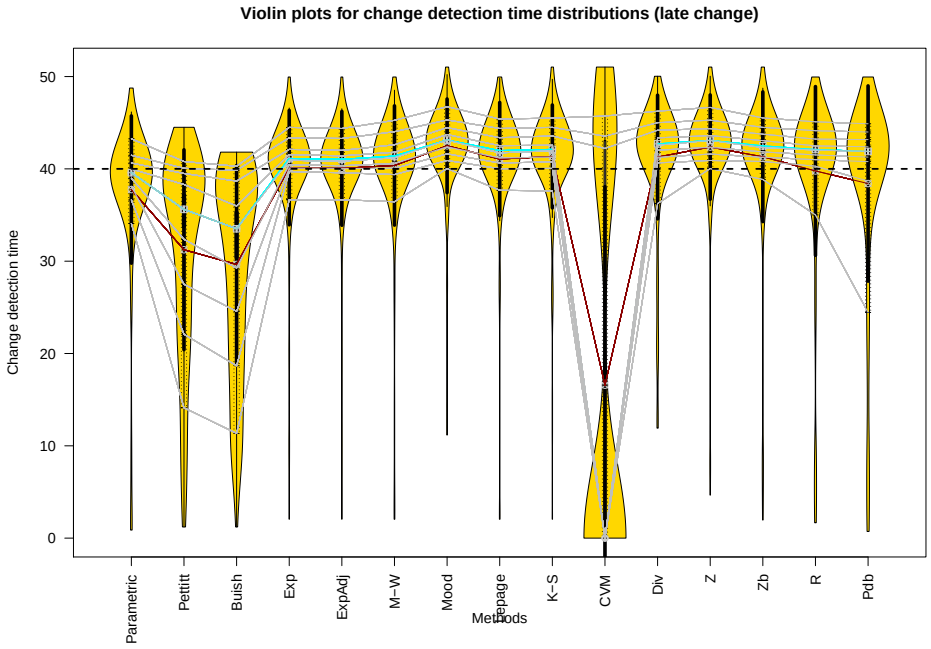


Figure 3.20: Estimated time of change distributions with $\tau = 40$ (late), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 3$

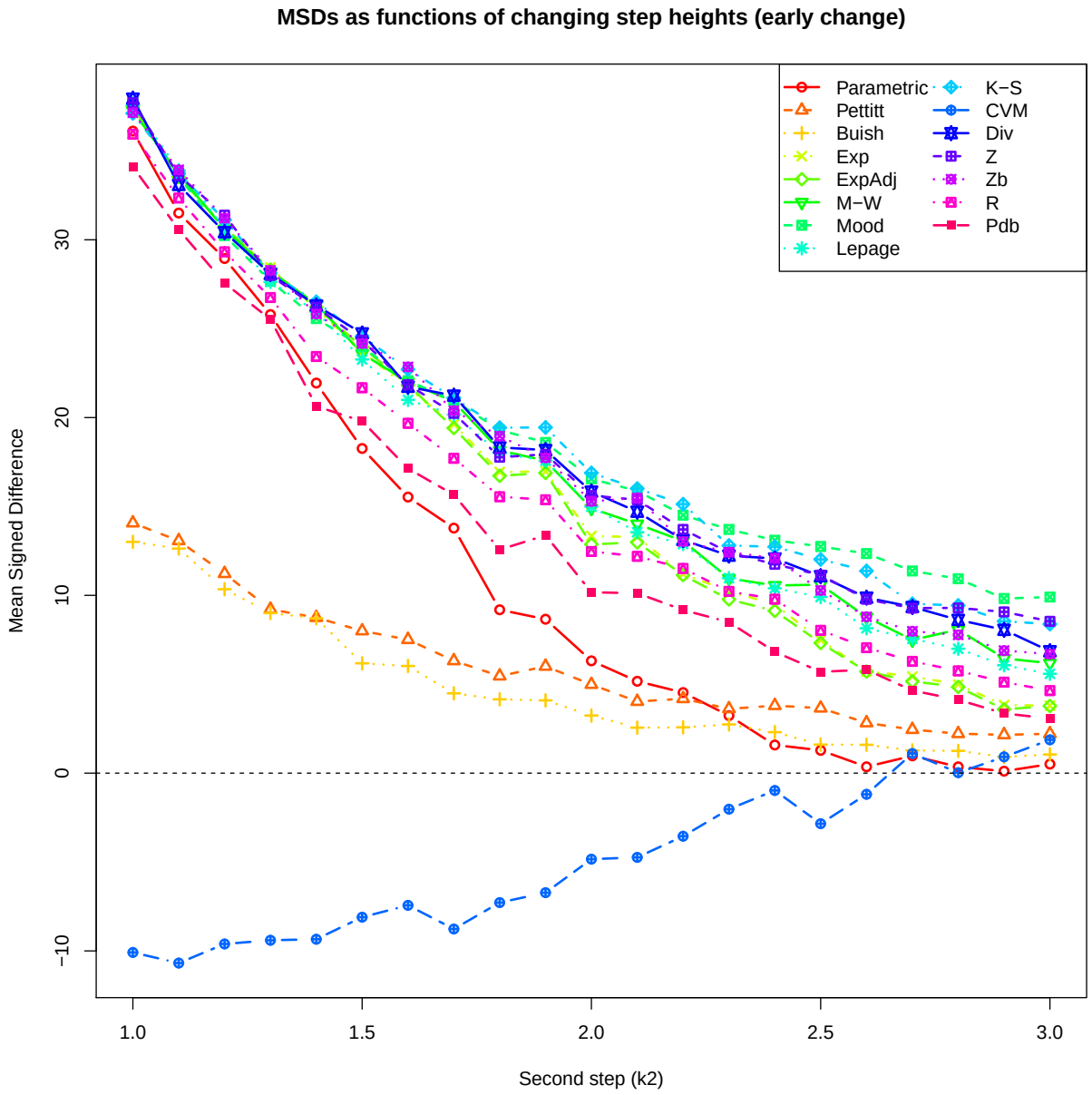


Figure 3.21: MSD comparisons with $\tau = 12$ (early), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

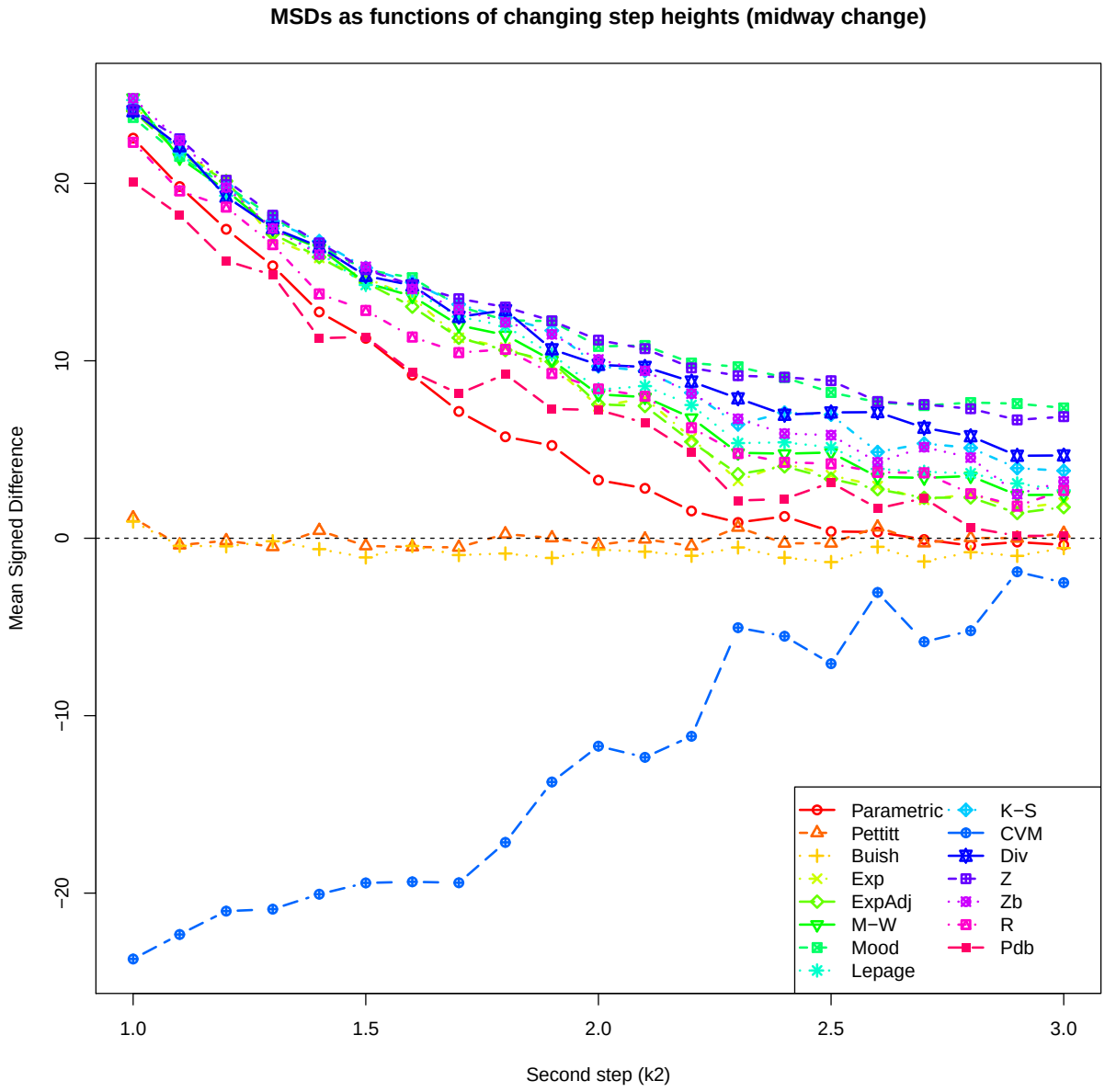


Figure 3.22: MSD comparisons with $\tau = 25$ (midway), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

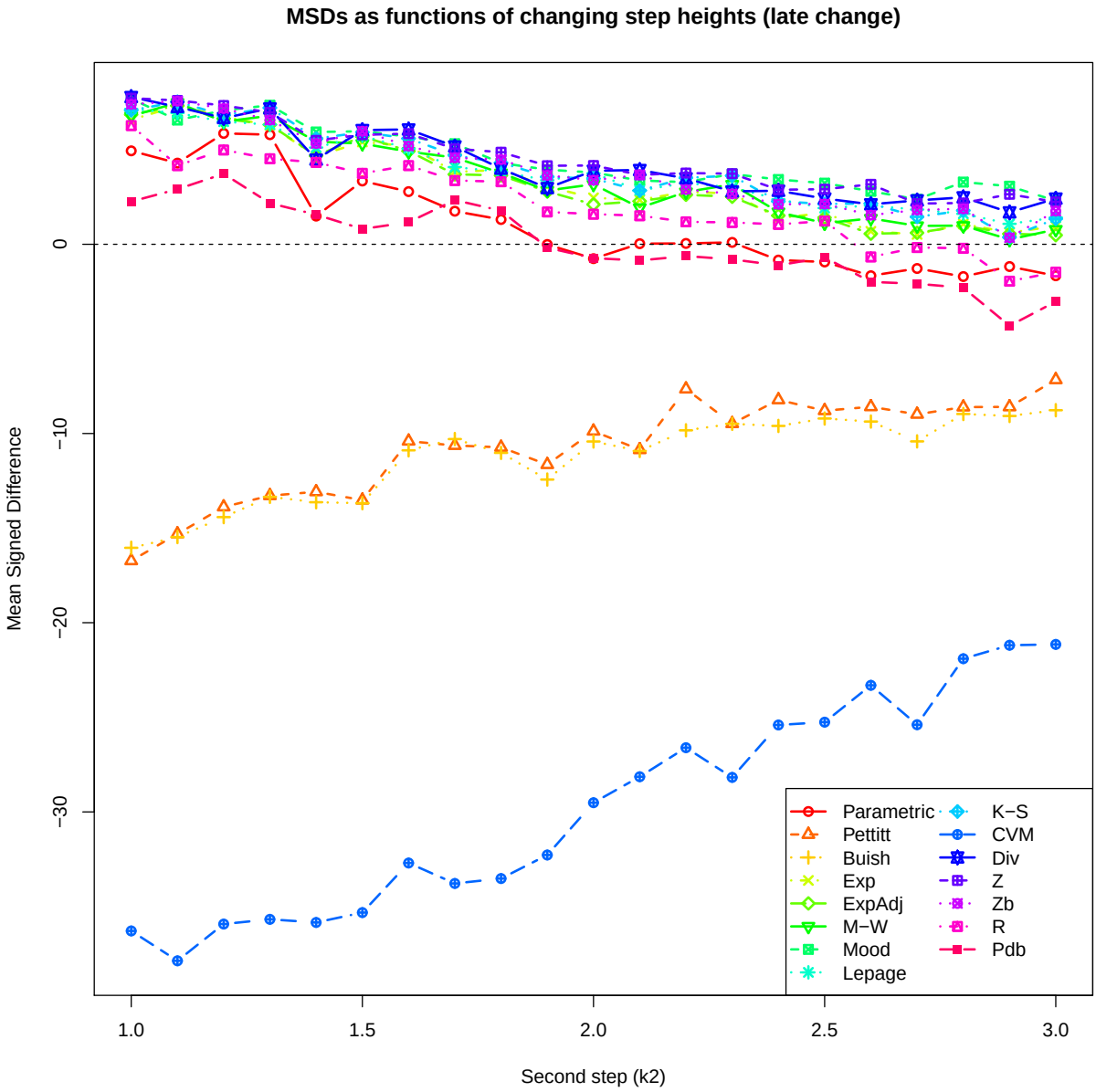


Figure 3.23: MSD comparisons with $\tau = 40$ (late), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

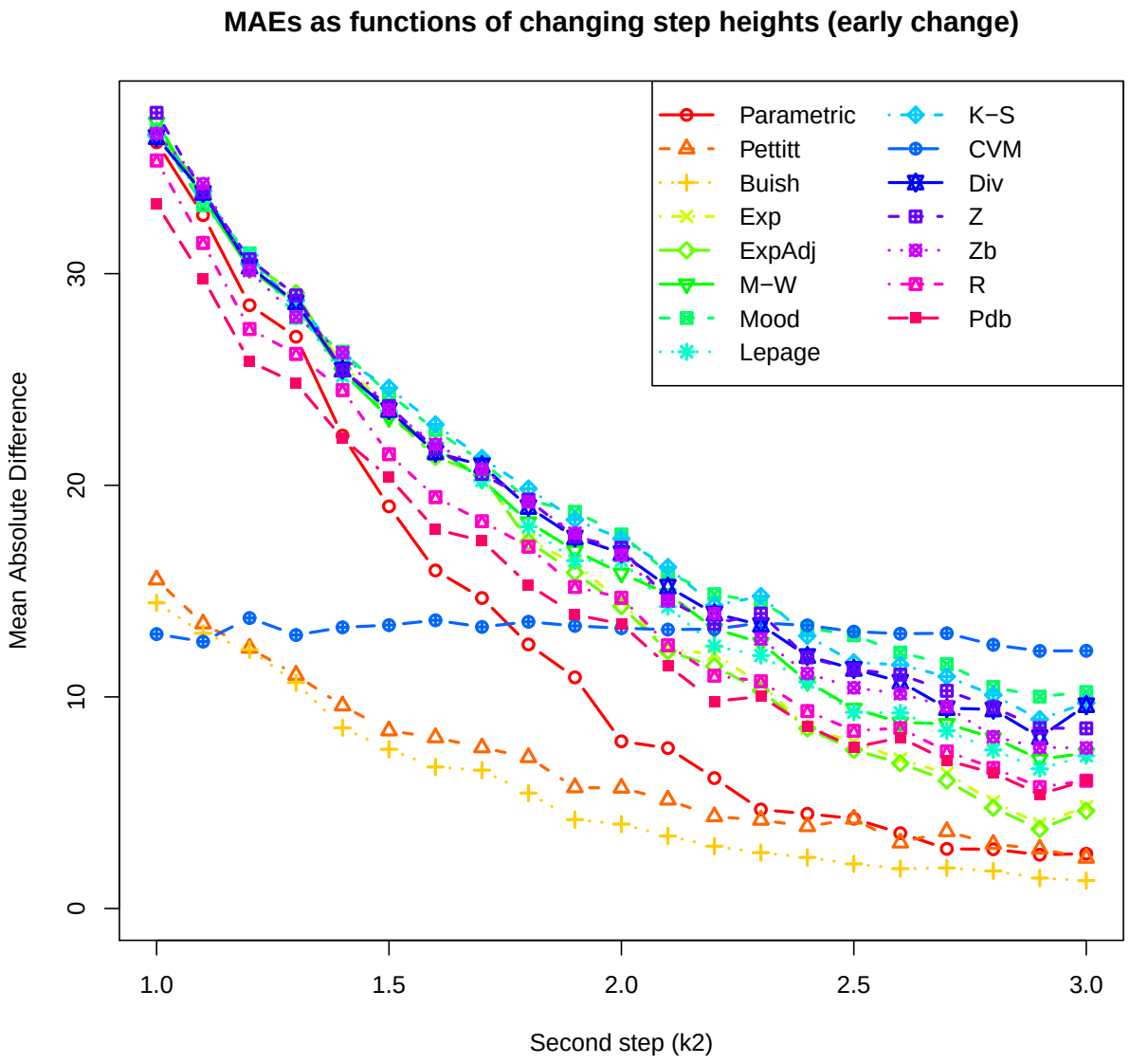


Figure 3.24: MAE comparisons with $\tau = 12$ (early), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

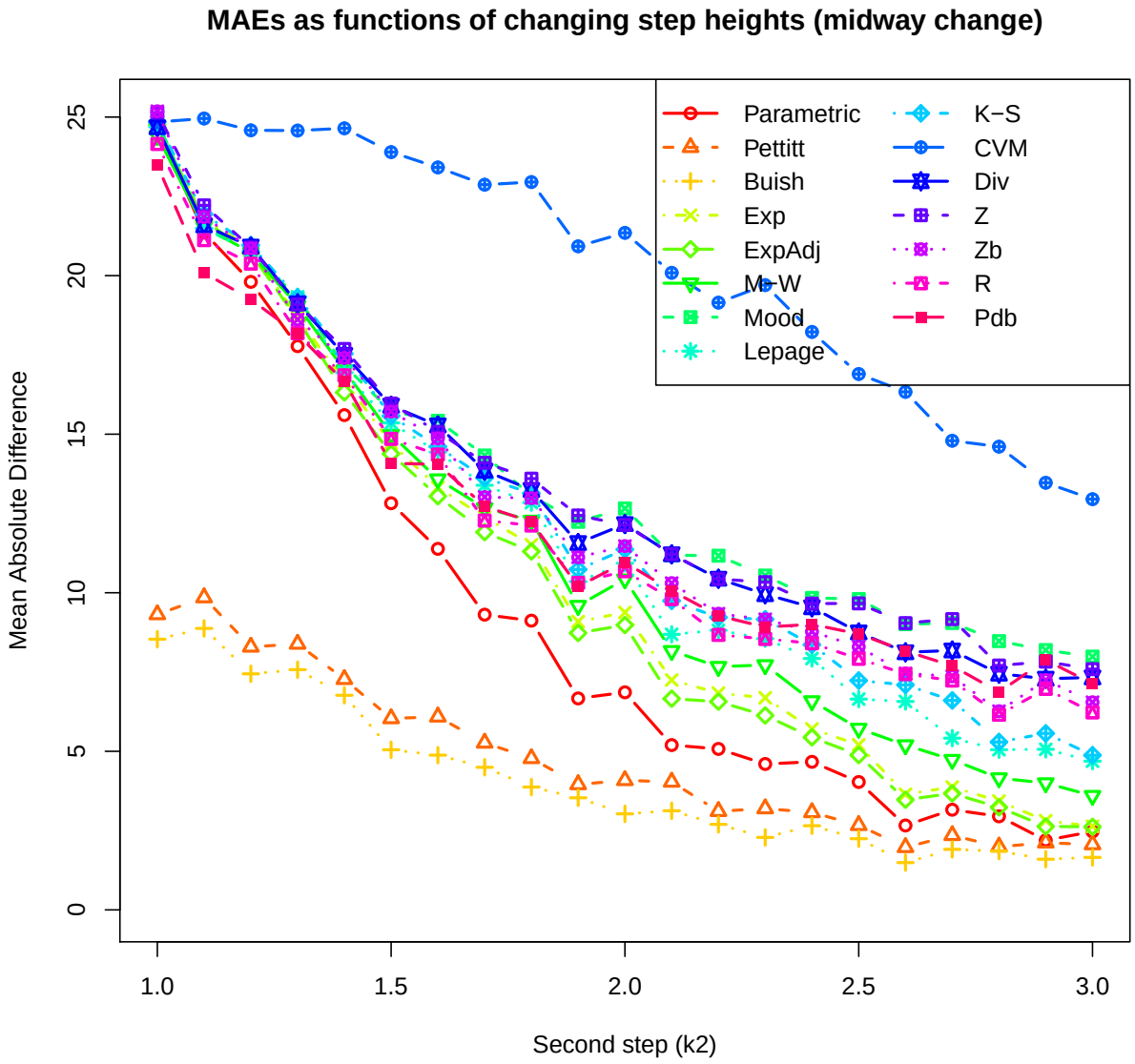


Figure 3.25: MAE comparisons with $\tau = 25$ (midway), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

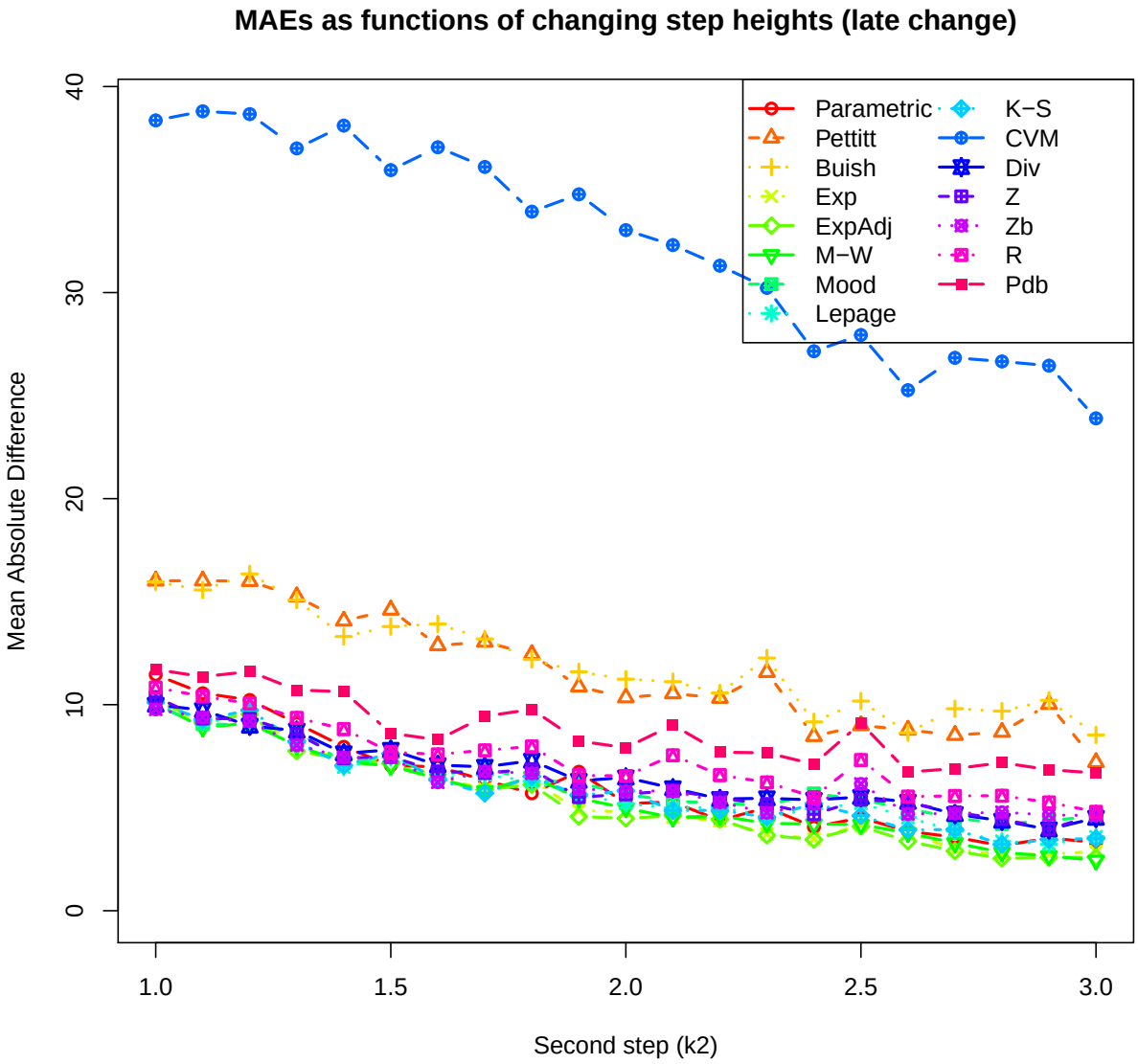


Figure 3.26: MAE comparisons with $\tau = 40$ (late), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

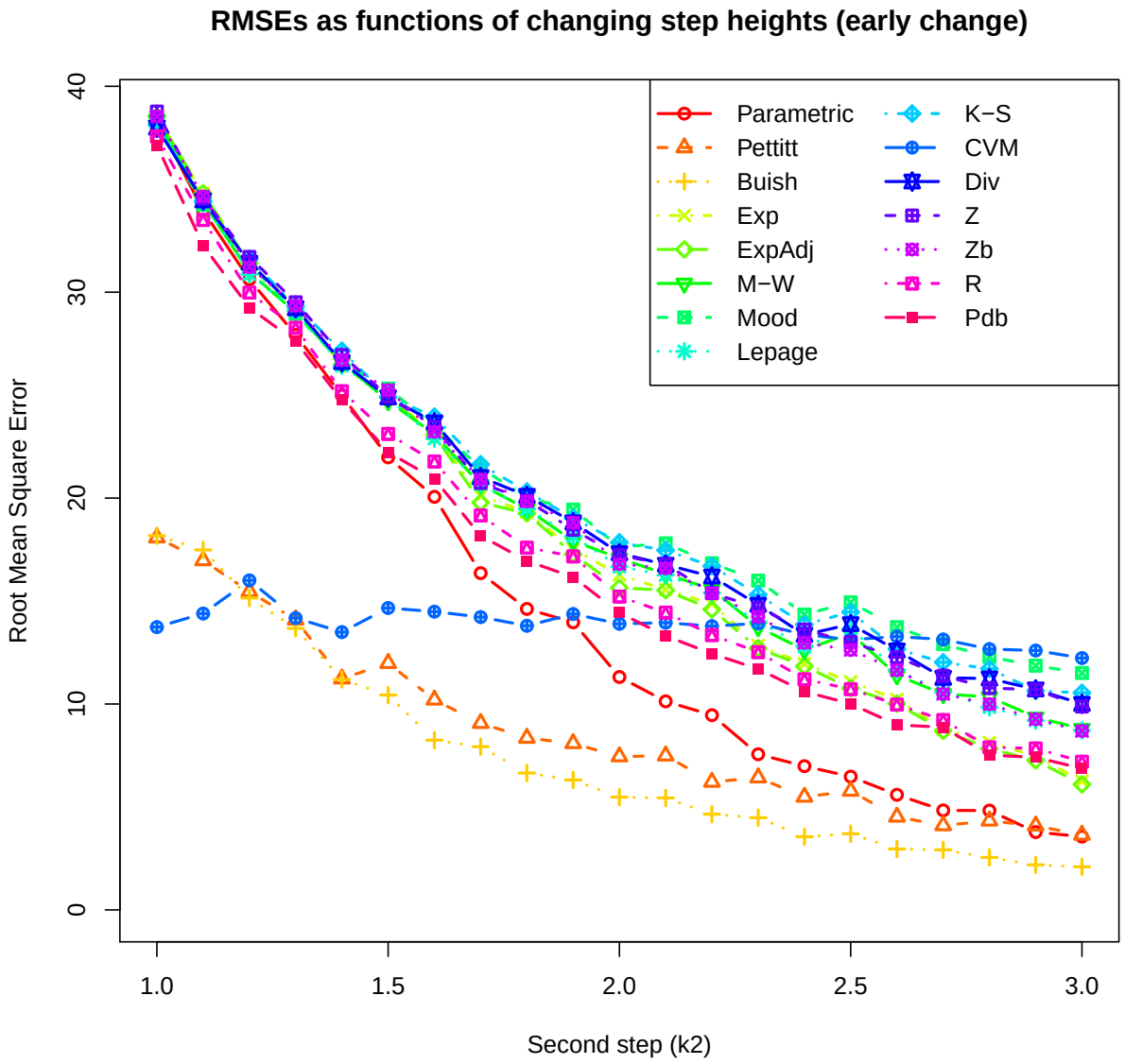


Figure 3.27: RMSE comparisons with $\tau = 12$ (early), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

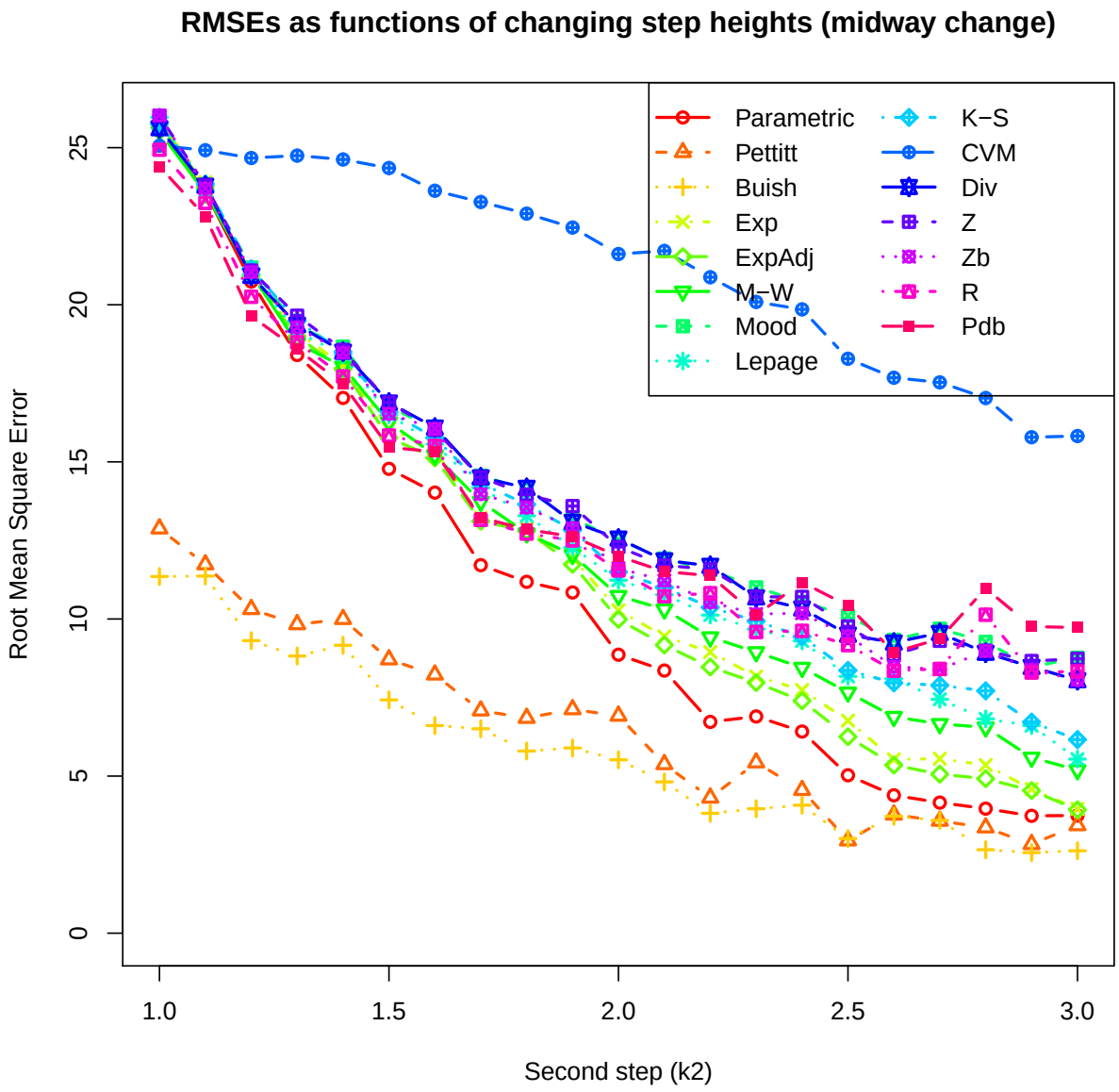


Figure 3.28: RMSE comparisons with $\tau = 25$ (midway), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

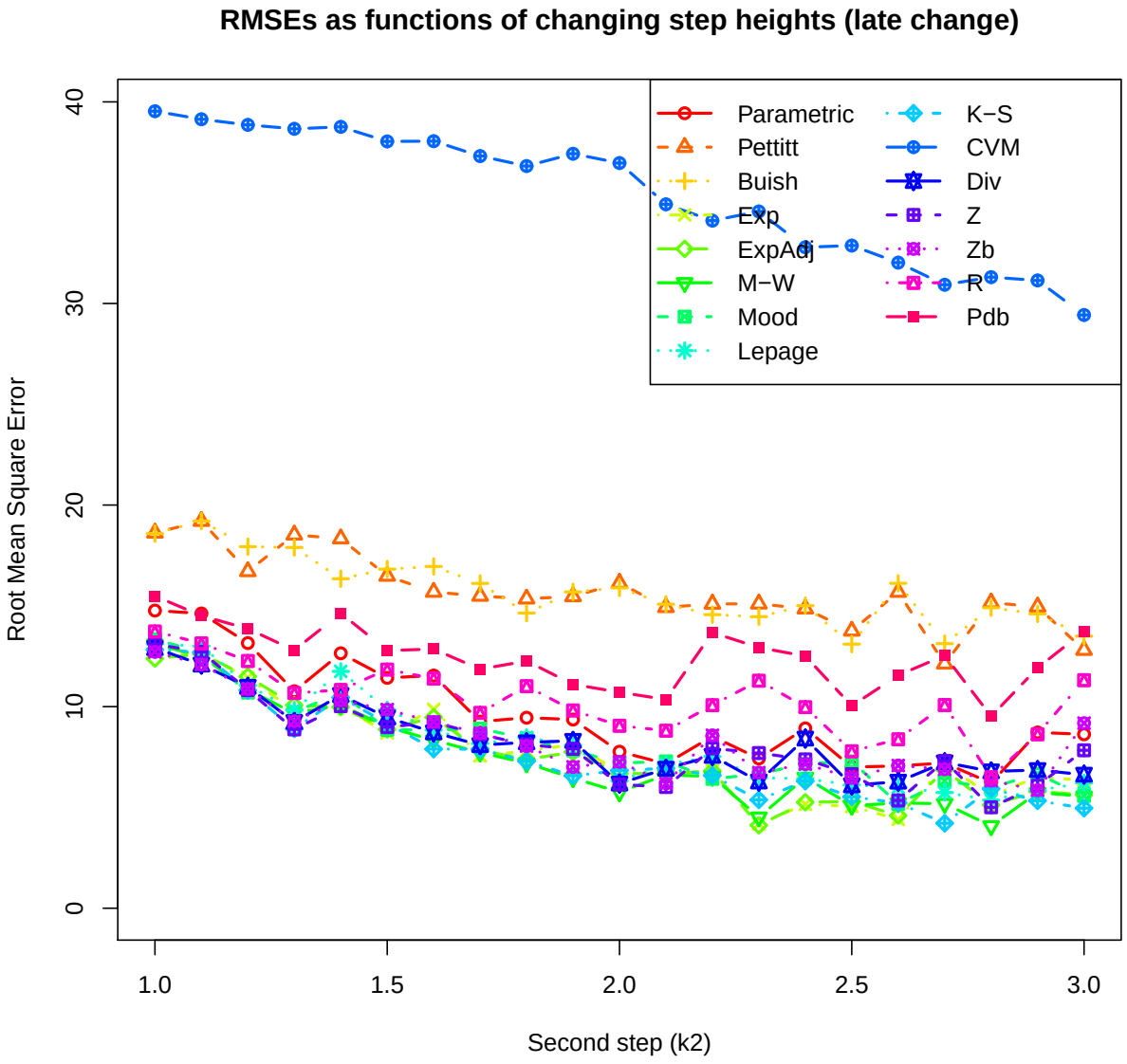


Figure 3.29: RMSE comparisons with $\tau = 40$ (late), $n = 50, \alpha = 0.05, k_1 = 1, k_2 = 1(0.5)3$

Decreasing step intensity

Table 3.67: Change point detection comparison, $n = 50, \tau = 12/3, \alpha = 0.05, k_1 = 3, k_2 = 1$

Test	Q_1	Q_2	Q_3	MAE	MSE	MSD
Parametric	2.923	3.868	6.128	8.724	326.09	7.459
Pettitt	3.613	4.507	9.825	4.893	87.12	4.273
Buish	4.128	6.722	12.050	5.804	88.69	5.528
CPM-Exp	3.147	4.402	38.020	14.944	572.45	14.073
CPM-Adjusted Exp	3.417	4.340	37.650	14.049	537.46	13.194
CPM-Mann-Whitney	3.674	15.040	41.660	19.481	752.16	18.824
CPM-Mood	36.080	41.450	46.730	37.605	1483.36	37.569
CPM-Lepage	4.753	37.210	44.450	26.703	1043.84	26.132
CPM-Kolmogorov-Smirnov	3.881	32.660	42.080	21.496	809.13	20.872
CPM-CramerVon-Mises	0	31.87	41.87	22.353	817.59	18.672
E-Divergence	32.05	38.88	45.61	31.769	1221.89	31.753
Z	10.30	31.25	42.53	23.634	842.84	23.512
Z_B	7.971	14.170	28.290	19.013	641.53	18.777
R	8.601	18.670	38.690	20.318	683.99	19.912
P_{DB}	4.682	8.602	23.010	12.428	376.45	11.591

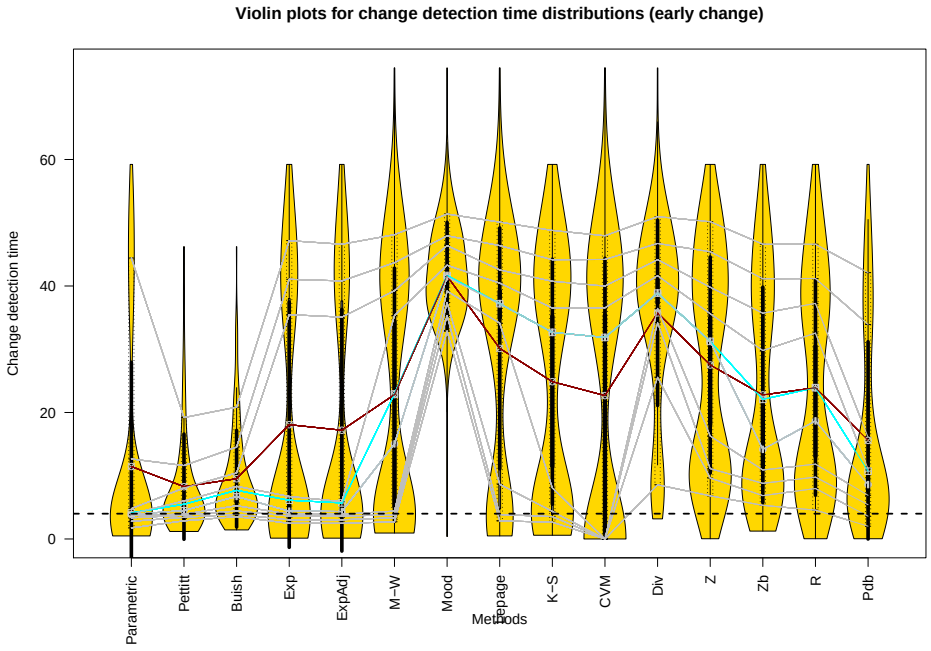


Figure 3.30: Estimated time of change distributions with $\tau = 12/3$ (early), $n = 50, \alpha = 0.05, k_1 = 3, k_2 = 1$

Table 3.68: Change point detection comparison, $n = 50, \tau = 25/3, \alpha = 0.05, k_1 = 3, k_2 = 1$

Test	Q_1	Q_2	Q_3	MAE	MSE	MSD
Parametric	6.818	7.898	8.825	2.705	37.60	0.402
Pettitt	7.311	7.974	8.536	1.675	11.43	0.189
Buish	7.907	8.366	9.834	1.554	6.73	0.792
CPM-Exp	7.579	8.222	9.898	5.278	118.70	3.856
CPM-Adjusted Exp	7.556	8.211	9.781	5.004	110.24	3.521
CPM-Mann-Whitney	7.609	8.546	24.210	8.585	205.64	7.479
CPM-Mood	25.860	32.370	37.570	22.647	604.18	22.526
CPM-Lepage	8.145	12.100	32.240	12.126	307.04	11.388
CPM-Kolmogorov-Smirnov	7.842	9.732	29.980	9.949	240.57	8.946
CPM-CramerVon-Mises	0	28.44	35.89	19.643	474.69	14.393
E-Divergence	12.040	25.060	34.680	16.352	415.75	16.308
Z	18.560	24.870	33.160	17.762	399.89	17.280
Z_B	11.990	15.980	23.090	11.407	219.58	10.278
R	15.750	21.790	31.200	15.446	320.37	14.239
P_{DB}	7.120	12.560	17.330	7.825	112.29	5.177

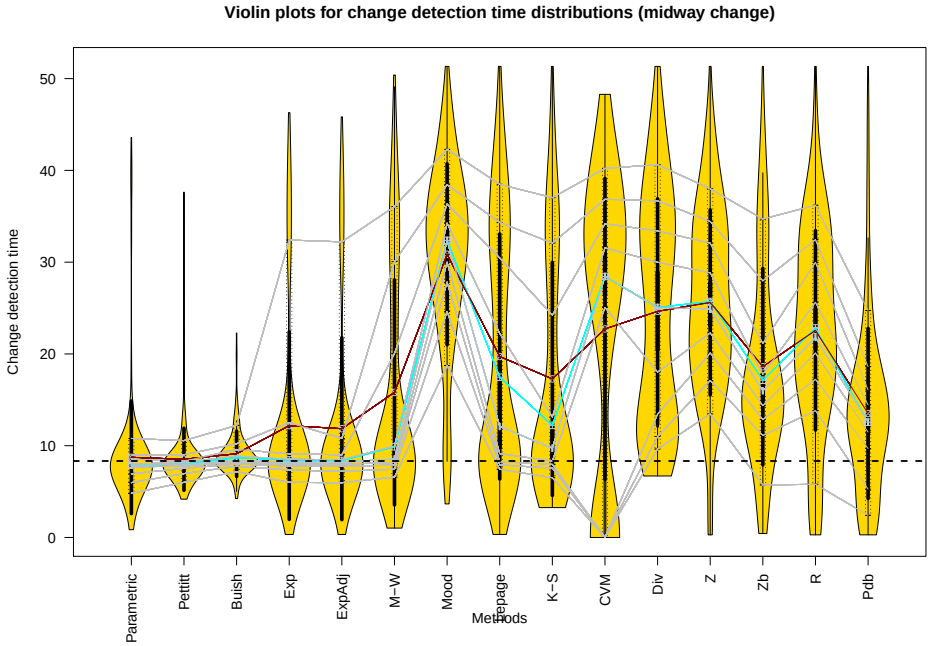


Figure 3.31: Estimated time of change distributions with $\tau = 25/3$ (midway), $n = 50, \alpha = 0.05, k_1 = 3, k_2 = 1$

Table 3.69: Change point detection comparison, $n = 50, \tau = 40/3, \alpha = 0.05, k_1 = 3, k_2 = 1$

Test	Q_1	Q_2	Q_3	MAE	MSE	MSD
Parametric	8.881	12.420	12.970	2.989	22.31	-2.428
Pettitt	6.861	10.500	12.510	3.828	25.99	-3.760
Buish	8.923	12.190	13.100	2.683	15.84	-2.458
CPM-Exp	12.390	13.260	14.870	3.191	28.05	0.812
CPM-Adjusted Exp	12.350	13.200	14.610	3.146	27.56	0.641
CPM-Mann-Whitney	12.280	13.750	20.460	4.986	55.56	2.995
CPM-Mood	14.200	20.900	26.780	8.691	128.51	7.933
CPM-Lepage	12.830	14.220	21.060	4.936	56.88	3.662
CPM-Kolmogorov-Smirnov	12.630	14.050	20.720	4.759	50.29	3.190
CPM-CramerVon-Mises	0	0	23.79	13.438	199.44	-2.295
E-Divergence	14.270	16.790	22.210	6.060	75.07	5.589
Z	16.560	21.690	26.300	9.099	122.84	8.431
Z_B	15.370	18.860	21.880	6.367	60.19	5.179
R	15.840	21.400	25.430	8.917	320.37	7.365
P_{DB}	12.690	17.030	20.030	6.107	54.21	2.386

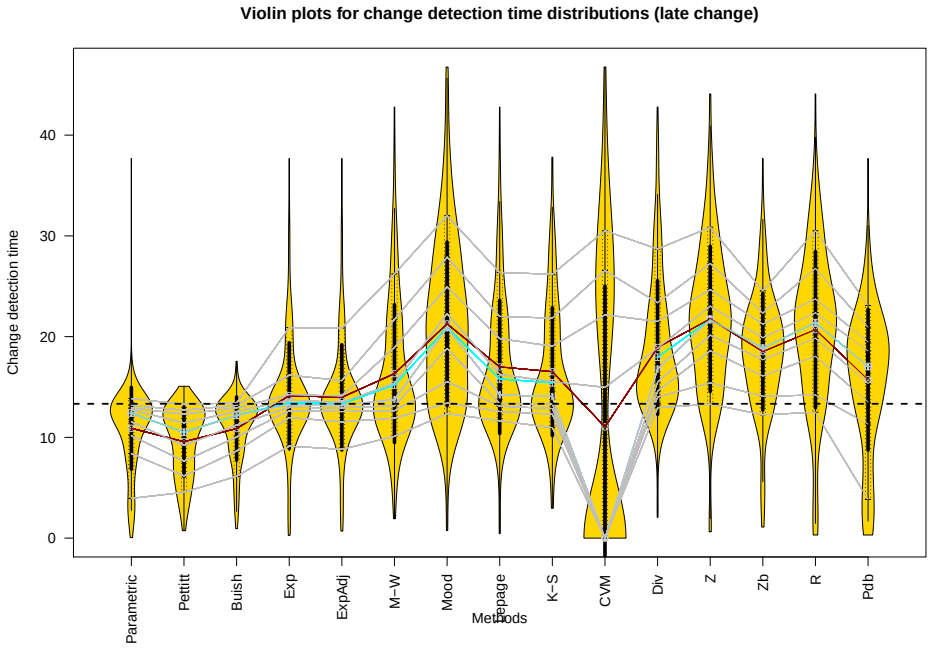


Figure 3.32: Estimated time of change distributions with $\tau = 40/3$ (late), $n = 50, \alpha = 0.05, k_1 = 3, k_2 = 1$

Chapter 4

Validation

The goal of this chapter is to substantiate change point estimation results obtained using the bidirectional tests with readily implementable graphic tools. Bridges such as

$$N(t) - N(t-1) := X_t \tag{4.1}$$

that connect a continuous time stochastic process (a Poisson process $\{N(t)\}_{t\geq 0}$ for instance) to a discrete time process (a time series $\{X_t\}_{t=1,2,\ldots}$ for instance) will be crucial in our analyses to follow. In particular, we will rid ourselves of the trouble of continuous monitoring, and instead will count the number of events observed over an interval defined discretely. The central theme that should bind the chapter is this close connection between a time series and a point process, coupled with our firm conviction that each area will be able to learn significantly from the other. A relation such as this has been sparsely conjectured in the literature: Brillinger (1994) [18] tries to find unifying characteristics embracing time series, point processes, marked point processes and hybrids through the introduction of "stationary increment process" and examination of 2nd and 3rd order autocovariances. This paper talks about a method of converting a linear point process to a 0-1 time series but refrains from forecasting and inferences. Rigas (1996) [121] creates a type of bivariate process, where one component is a time series, the other is a point process and then moves on to inferences.

Henschel et al. (2008) [61] make a casual remark about converting a point process to a time series but does not explicitly show how. Much of our work on this domain, especially those surveyed in the first four sections, have been published. The mood for this chapter will, therefore, be expository. We invite readers interested in details and intricate technicalities to consult Tan, Bhaduri, and Ho (2014) [142], Ho and Bhaduri (2015) [68], Ho et al. (2016) [69], and Ho and Bhaduri (2017) [70].

4.1 Empirical Recurrence Rates (ERRs)

Originally devised by Ho (2008) [67] the Empirical Recurrence Rate (ERR) statistic tracks the maximum likelihood estimate of the rate parameter of a homogeneous or stationary Poisson process. Corresponding to n event times $t_1, t_2, .., t_n \in [0, T]$ of a PtP, the ERR time series $\{Z_l\}$ may be generated at equidistant time points $h, 2h, 3h, .., Nh(= T)$ according to

$$Z_l = \frac{N(lh)}{lh} = \frac{\sum_{i=1}^{l} X_i}{lh}, \quad l = 1, 2, .., N \tag{4.2}$$

where the ith observation in the $\{X_i\}$ time series stores the number of observations in the interval $((i-1)h, ih]$, connected to the process $\{N(t)\}_{t\geq 0}$ according to (4.1). Chen (2010) [24] notes how ERR can act as a tool to confirm the patterns in the underlying intensity: a deteriorating process will make its ERR curve increase, while a stable process, governed by a constant intensity, should induce a horizontal ERR curve. Exploiting ERR's ability to generate "pseudo-observations" over barren data periods, Tan, Bhaduri, and Ho (2014) [142] and Ho and Bhaduri (2015) [68] have used this statistic to model sporadic events: the former deals with sandstorms, where seasonality is obvious, and the latter, with earthquakes, where it is not. Time series models (Box-Jenkins (1976) [17]) were fitted and forecasts were

extracted to predict future events.

4.2 Empirical Recurrence Rates Ratios (ERRRs)

The smoothing properties of ERRs may be extended to compare two time series $\{X_t\}$ and $\{Y_t\}$, often to be treated as two distinct "arms" of the same non-stationary PtP, one representing the pre-change sequence, and the other, the post-change sequence. Ho et al. (2016) [69] and Ho and Bhaduri (2017) [70] introduced and examined the Empirical Recurrence Rates Ratio ($ERRR$), defined as

$$R_{X,Y;l} = \frac{N_X(l)}{N_X(l) + N_Y(l)} = \frac{\sum_{k=1}^{t} X_k}{\sum_{k=1}^{t} X_k + \sum_{k=1}^{t} Y_k}, \quad l = 1, 2, .., N \tag{4.3}$$

where symbols have meanings similar to the ones in the previous section on ERR. The statistic, essentially storing ratios of cumulative counts, is a ratio of two ERRs, and is bounded by 0 and 1. A high value at a given instance implies that the first series $\{X_t\}$ is more active till then, while a value close to 0.5 indicates $\{X_t\}$ and $\{Y_t\}$ are equally active. Ho and Bhaduri (2017) [70] touch upon the statistic's inferential aspects and introduce two artificial time series to accentuate $ERRR$'s workings

Case 1: *(An equal size ratio series)* Assume that the discrete valued time series $\{X_t\}$ and $\{Y_t\}$ are given by

$$X_t = 1, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, 0, 0, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, 0, 0, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, ... \tag{4.4}$$

and

$$Y_t = 2, 4, 2, 1, 0, 0, 0, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, 0, 0, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, 0, 0, 0, 0, 2, 4, ... \tag{4.5}$$

It is apparent that when one is active, the other is relatively dormant, which induces a wave-like property into the $ERRR$ curve, graphed in Fig (4.1)

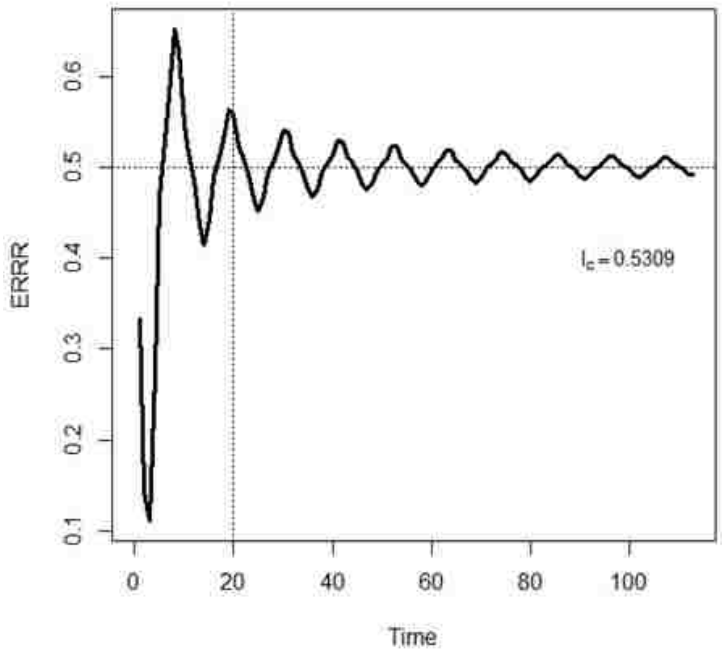


Figure 4.1: $ERRR$ plot and I_c index for inversely related equal size ratio series

and the fact that they are more or less equally intense places the curve around the 0.5 line.

Case 2: *(An unequal size ratio series)* Assume next that the discrete valued time series $\{X_t\}$ and $\{Y_t\}$ are given by

$$X_t = 1, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, 0, 0, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, 0, 0, 0, 0, 2, 4, 2, 2, 4, 2, 0, 0, ... \quad (4.6)$$

and

$$Y_t = 1, 2, 1, 0, 0, 0, 0, 0, 0, 1, 2, 1, 1, 2, 1, 0, 0, 0, 0, 0, 1, 2, 1, 1, 2, 1, 0, 0, 0, 0, 0, 1, 2, 1, 1, 2, 1, 0.... \quad (4.7)$$

The active-dormant inverse dependence still persists, keeping the curve graphed in Fig (4.2) wavy, but here, the first series is almost twice as intense as the second, positioning the

$ERRR$ curve away from the 0.5 line.

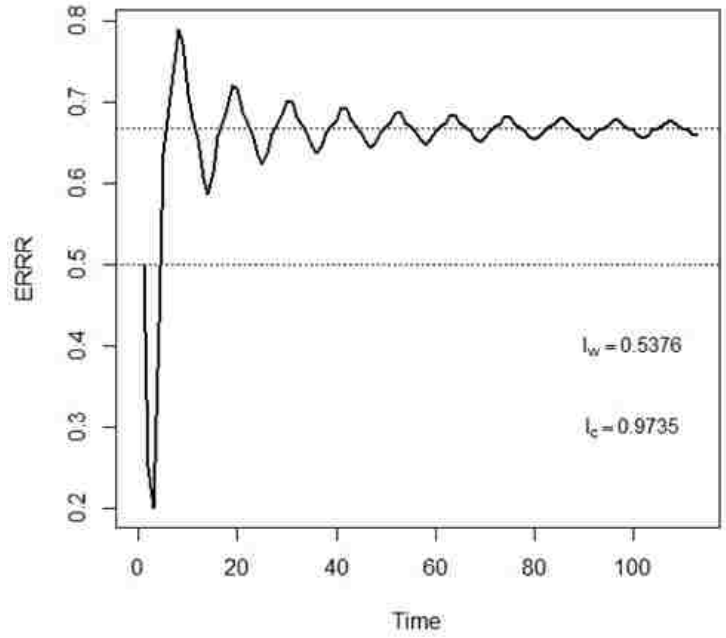


Figure 4.2: ERRR-plot of artificial data for unequal-size competing processes. The I_w is based on a threshold (= 0.6668) calculated by excluding the first 20 $ERRR$s (the burn-in period)

The indices contained in the figures will be explained in the next subsection. With the Hawaiian volcanoes described in the first chapter, treating the Kilauea eruption counts as the first series $\{X_t\}$ and the one for Mauna Loa as the second series $\{Y_t\}$, the $ERRR$ curve takes a sinusoidal form depicted in Fig (4.3)

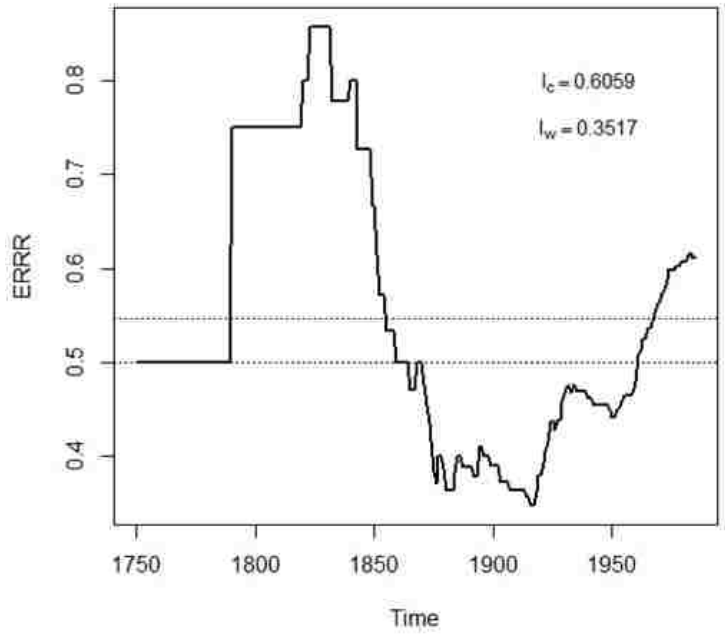


Figure 4.3: ERRR-plot of Kilauea vs Mauna Loa. The I_w is based on a threshold (= 0.5468) calculated by using the entire data set

The $\binom{3}{2} = 3$ pairwise comparisons from the West Atlantic hurricane basin, described in Chapter 1 generates the $ERRR$ curve shown in Fig (4.4) below.

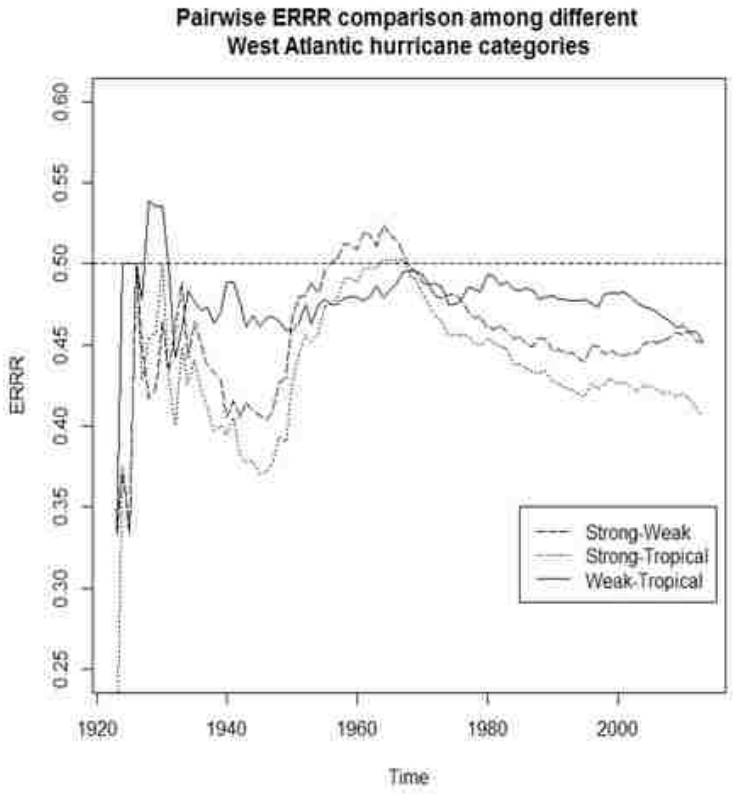


Figure 4.4: ERRR curves for hurricane counts

The noted weather scientist K. Emanuel, as pointed out in Chapter 1, feels that with a continually warming climate, it is increasingly difficult to start a devastating hurricane due to a hike in saturation deficit which works against its creation, but if it gets started somehow, it has the potential to become more intense. Thus, the total number of storms should decline globally, but the proportion of hurricanes which are intense should rise. This once again hints at a possible inverse dependence between the strong and weak categories and that suspicion is confirmed by the sinusoidal pattern of the ERRR curve generated in Fig (4.4). $ERRR$ analyses and time series fits to it have also been used to explain bank failures during the recent US economic meltdown. Ho et al. (2016) [69]documents that exercise.

4.2.1 Indices

While the appearance and the location of the ERRR curve provide preliminary notions about the nature of the interaction between the contributing series, measures quantifying such attributes are necessary for meaningful comparisons. Ho and Bhaduri (2017) [70] observes that in the spirit of the first hypothetical pair, if two series are equally competitive (and hence can be conveniently treated as parts of the same stationary PtP), the resulting ERRR curve should hover around the 0.5 line with faithful regularity. Thus

Definition 4.1. *For an ERRR series R of length n, the index of competitiveness I_c is defined as*

$$I_c = \frac{1}{n}\sum_{i=1}^{n} I\{r_i > 0.5\}, \tag{4.8}$$

which represents the time proportion of the curve's occupation of the space above the 0.5 reference line.

For the equally intense first hypothetical pair, $I_c = 0.5309$ while for the second pair with the first almost twice as intense as the second, this index shoots up to 0.9735. For the volcanic interaction series, the index clocks 0.6059. It is worth noting, however, that I_c merely renders an adequate measure of the curve's location, while remaining oblivious to any oscillating pattern that may be present. To quantify this property, we have proposed

Definition 4.2. *Definition: For an ERRR series R of length n, the index of waviness I_w is defined as*

$$I_w = \frac{1}{n}\sum_{i=1}^{n} I\{r_i > \bar{r}\}, \tag{4.9}$$

which represents the proportion of times the curve stays above the average level.

Ho and Bhaduri (2017) [70] spell out an algorithm of systematically deleting the first set of k ERRR observations to come up with a reliable, and sequentially modified $\bar{r}$ value, because of problems with burn-ins. The value around which I_w stabilizes may be taken as a reliable estimate of waviness. For the volcanic ERRR curve, the table below tracks this index, which converges to 0.4.

Table 4.1: Sequential history of the I_w index based on eruption counts

k	Modified mean	I_w	k	Modified mean	I_w
0	0.54679	0.3517	11	0.54908	0.3689
1	0.54699	0.3532	12	0.54929	0.3705
2	0.54719	0.3547	13	0.54952	0.3722
3	0.54739	0.3562	14	0.54974	0.3789
4	0.54759	0.3578	15	0.54997	0.3756
5	0.54781	0.3593	16	0.55019	0.3773
6	0.54801	0.3609	17	0.55042	0.3789
7	0.54822	0.3624	18	0.55065	0.3807
8	0.54843	0.364	19	0.55089	0.3825
9	0.54865	0.3656	20	0.55113	0.3843
10	0.54866	0.3673			

The first of these two measures will be used later to create a test and others serving more specialized ends may be formulated too. It can be observed that three successive points are sufficient to indicate whether or not a change in trend, indicating the presence of a small wave, has occurred in the ERRR series. So a good measure of "waviness" should take into account all possible triads of the form (r_{t-1}, r_t, r_{t+1}) and hence, the contribution from all the small waves. Bearing that in mind, we propose the following measures to quantify the nature and extent of a wave-like property:

Simple Static Wave Index (SSW Index)

For each $t = 2, 3, ..., (n-1)$, define an indicator-like function as follows:

$$I_t^{SS} = \begin{cases} 1 & \frac{r_t - r_{t-1}}{r_{t+1} - r_t} < 0, \\ \epsilon & r_{t+1} - r_t = 0, \\ 0 & \frac{r_t - r_{t-1}}{r_{t+1} - r_t} > 0. \end{cases} \tag{4.10}$$

Thus the value of I_t^{SS} indicates whether a wave exists at t: if it's 1, then a wave exists, if it is 0, then the ongoing local trend is preserved, indicating the absence of a wave. The choice of ϵ is subjective and should be preferably close to 0. It is designed to make the measure survive even in the face of pathological and artificially constructed situations, such as when the two original series are both exactly identical to each other. The Simple Static Wave Index (SSW Index) can now be defined as

$$SSW = \frac{1}{n-2} \sum_{t=2}^{n-1} I_t^{SS} \tag{4.11}$$

The numerator counts the total number of wavelets and to compare big and small data sets on the same scale, the normalization with the $n-2$ factor is essential. The SSW index is a sort of a simple average and hence, rather easy to interpret. Additionally, it enjoys the desirable property of being bounded: at most, it can be 1 (when the points alternate consistently) and at least it should be 0 (when the series is perfectly monotonic w.r.t trend and bereft of any variation). Otherwise, a large SSW index should in general, indicate a considerable degree of waviness. We do not feel obliged to prescribe a precise demarcation between waviness and non-waviness in terms of an SSW threshold. This is the signature of several other useful statistical measures as well: on the (0,1) range of the simple correlation coefficient, there doesnt exist any well-defined boundary between being weakly positively correlated and being strongly positively correlated.

Simple Dynamic Wave Index (SDW Index)

The SSW index, though extremely intuitive, doesnt pay attention to the height of the individual wavelets. Since it only worries about the number of times a locally ongoing trend-equilibrium is disturbed, it will not be able to differentiate between two series which are only scalar multiples of each other. Thus, to formulate a more dynamic measure, lets introduce, for each $t = 2, 3, , (n-1)$, another indicator-like operator as follows:

$$I_t^{SD} = \begin{cases} 1 + |r_t - r_{t-1}| + |r_{t+1} - r_t| & \frac{r_t - r_{t-1}}{r_{t+1} - r_t} < 0, \\ \epsilon & r_{t+1} - r_t = 0, \\ 0 & \frac{r_t - r_{t-1}}{r_{t+1} - r_t} > 0, \end{cases} \tag{4.12}$$

and based on it, define the Simple Dynamic Wave Index (SDW) as

$$SDW = \frac{1}{n-2} \sum_{t=2}^{n-1} I_t^{SD}. \tag{4.13}$$

This measure is unbounded, but whenever it detects the presence of a wavelet, it records the amount of change in trend as well. Using the definitions, it is not difficult to show that for any given series, the following inequality holds

$$SSW \leq SDW. \tag{4.14}$$

Weighted Dynamic Wave (WDW) Index

We observe that both the SSW and SDW indices are constructed using simple arithmetic means which lacks robustness in the presence of outliers. Additionally, in almost all applications of ERRR, we are able to identify an initial burn-in period – a period during which the series typically, behaves wildly, which might not necessarily be indicative of the pattern to follow. Furthermore, as a consequence of the strong laws of large numbers, the ERRR series is always expected to stabilize asymptotically around the process mean. These imply that

all the points should not exert equal influence in creating the measures and that ideally, a weighted mean seems to be more apt. If the choice of a reasonable weight function can be agreed upon, then a weighted version, WDW, can be constructed as follows:

$$WDW = \frac{\sum_{t=2}^{n-1} w_t I_t^{SD}}{\sum_{t=2}^{n-1} w_t}. \tag{4.15}$$

Realizing the burn-in issues and the fact that a big wave towards the latter half of the series (when it is expected to stabilize) is highly significant, one should choose $w_t \propto t$.

Wave Contribution Index (WC Index)

The numerator of SSW introduced above counts the total number of wavelets and does not differentiate between a peak and a valley. There might be times when such information is crucial, and the measure to be introduced now is an attempt to address this issue and also to understand how much does each of the original series contribute to the wave index. We introduce a pair of indicator variables I_t^P and I_t^V as follows:

$$I_t^P = \begin{cases} 1 & r_{t-1} < r_t \text{ and } r_t > r_{t+1}, \\ 0 & \text{otherwise.} \end{cases} \tag{4.16}$$

$$I_t^V = \begin{cases} 1 & r_{t-1} > r_t \text{ and } r_t < r_{t+1}, \\ 0 & \text{otherwise.} \end{cases} \tag{4.17}$$

Thus I_t^P and I_t^V record the occurrence of a peak and a valley at t. Summing up these indicators, we should have the total number of peaks and valleys and remembering that a valley is induced only by an increased activity of the first series and a peak is introduced by an increased activity of the second, we define the Wave Contribution Index of the first series as:

$$WC_X = \frac{\sum_{t=2}^{n-1} I_t^V}{\sum_{t=2}^{n-1} I_t^{SS}} = \frac{\sum_{t=2}^{n-1} I_t^V/(n-2)}{SSW}, \tag{4.18}$$

and of the second series as

$$WC_Y = \frac{\sum_{t=2}^{n-1} I_t^P}{\sum_{t=2}^{n-1} I_t^{SS}} = \frac{\sum_{t=2}^{n-1} I_t^P/(n-2)}{SSW}. \tag{4.19}$$

These are bounded by 0 and 1 too and can be interpreted as the separate contribution of the two series to the total amount of waviness present. Their relative importance can be judged by:

$$WC_{X,Y} = \frac{WC_X}{WC_Y}, \tag{4.20}$$

with a value in excess of 1 indicating the presence of more valleys than peaks.

4.2.2 Bootstrapping the ERRR

The feeling regarding the "waviness" of a generated ERRR curve, furnished by the I_w index, may be strengthened by demonstrating there exists no other permuted version of the curve which would have given a radically different index. To generate these rearranged ERRR curves, Ho and Bhaduri (2017) [70] resorted to block bootstrapping, and the exercise will be briefly touched upon in this section. The method of ordinary bootstrapping due to Efron (1979) [39] is ineffective here since isolated resampling will smother the inherent dependence, arguably one of the most crucial aspects of any time series. Hall (1985) [54] and Carlstein (1986) [22] 's generalization to block resampling was therefore, adopted.

In theory, the method requires a user-defined block size b and sampling chunks of the original time series $\{X_{I+1}, X_{I+2}, ..., X_{I+b}\}$ where $I \in \{0, 1, 2, .., n-b\}$ is randomly picked. Joining these blocks end to end will create a time series with properties similar to the parent. Higher values of b will imply the retention of a greater degree of dependence. We direct readers interested in more formal notational constructs and technicalities (such as the

choice of the block size) to Ho and Bhaduri (2017) [70]. Figure (4.5) depicts two resamples of block size 25 each generated from the ERRR series of volcanic interactions.

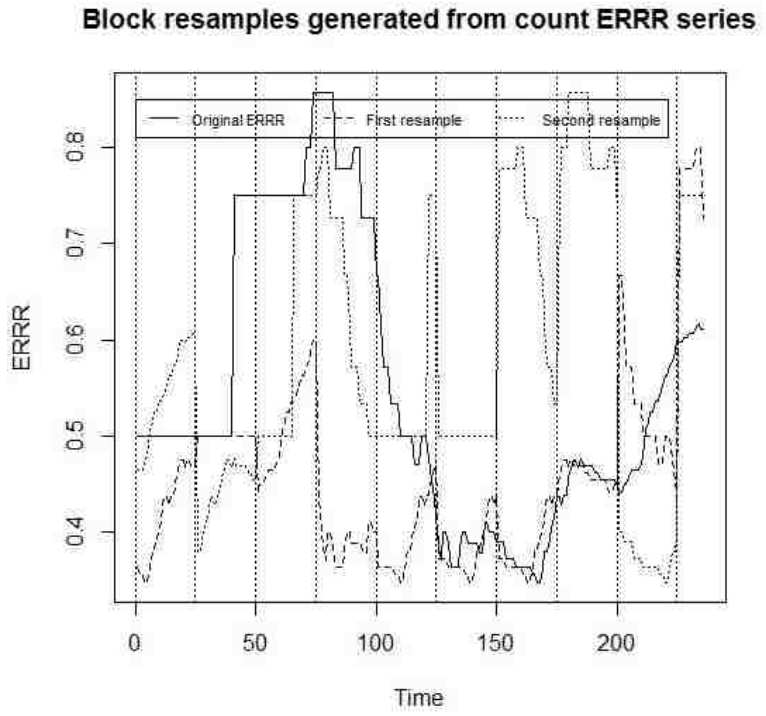


Figure 4.5: Generation of block resamples from the parent volcanic ERRR time series

To inquire whether the competitiveness or wave-like pattern in the parent series will be retained in other possible arrangements of the curve, we have gathered 1000 bootstrapped (or resampled) versions with block size 5, calculated I_c and I_w from each, and have summarized their distributions in Figures (4.6) and (4.7) and Table(4.2).

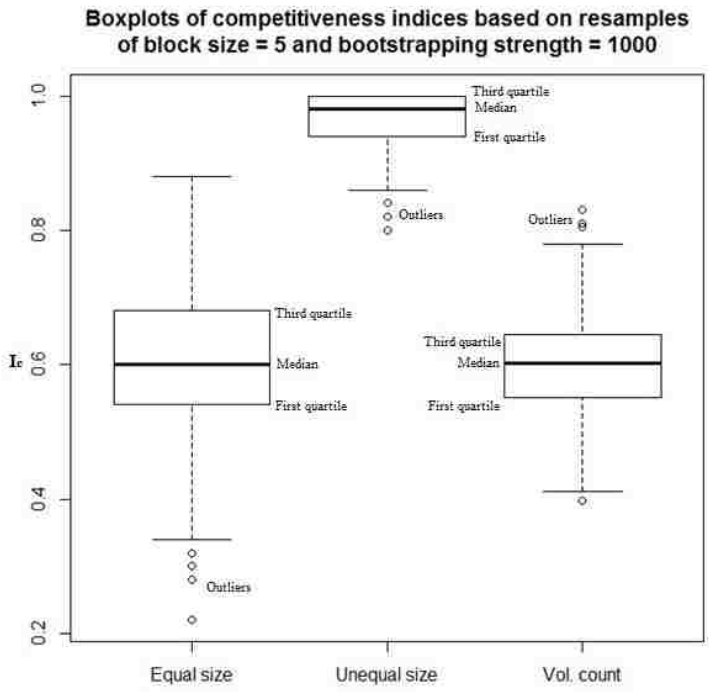


Figure 4.6: Descriptive summarization and comparison of the distribution of I_c indices calculated from several bootstrapped ERRR series

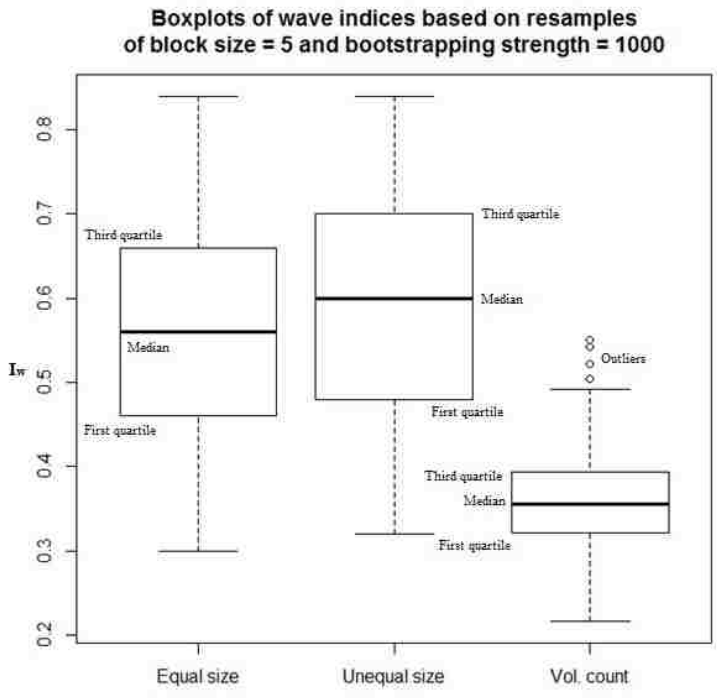


Figure 4.7: Descriptive summarization and comparison of the distribution of I_w indices calculated from several bootstrapped ERRR series

Table 4.2: Comparison between the point estimates and the bootstrap 95 % intervals of the I_c and I_w indices across different data sets

	Equal size	Unequal size	Volcanic count
I_c (from the original parent series)	0.5309	0.9735	0.6
I_c (the bootstrap 95 % interval)	[0.47, 0.79]	[0.93, 0.99]	[0.42, 0.71]
I_w (from the original parent series)	0.71	0.5376	0.4
I_w (the bootstrap 95 % interval)	[0.40, 0.72]	[0.41, 0.75]	[0.24, 0.49]

Those from the hypothetical equal and unequal size ratio series encountered previously, have been added alongside as ready references. It is interesting to note that the distribution of these indices from the bootstrapped samples tends to be centered around the indices calculated from the original parent. In particular, the 95% bootstrapped interval perfectly contain the point indices calculated previously, which enable one have stronger faith in them, the ERRR curve, and the general conclusion about the inverse nature of dependence between

these two volcanoes. Following Hesterberg (2015) [62] 's recommendation for reporting summary statistics, Ho and Bhaduri (2017) [70] records the standard error of the bootstrap interval and the aggregated bias. In that work, we have also described a way to circumnavigate (through differencing and modifying the definition of these indices) a technical problem regarding weak stationarity. Finally, a $SARIMA(2,0,1) \times (2,1,1)_{32}$ model was fitted to the ERRR series (we shall encounter the fit in the final section) and 500 years' worth of forecasts were extracted. Figure (4.8) summarizes the indices' distributions from the resampled versions of the pooled parent (i.e. including both the observed and the forecast values).

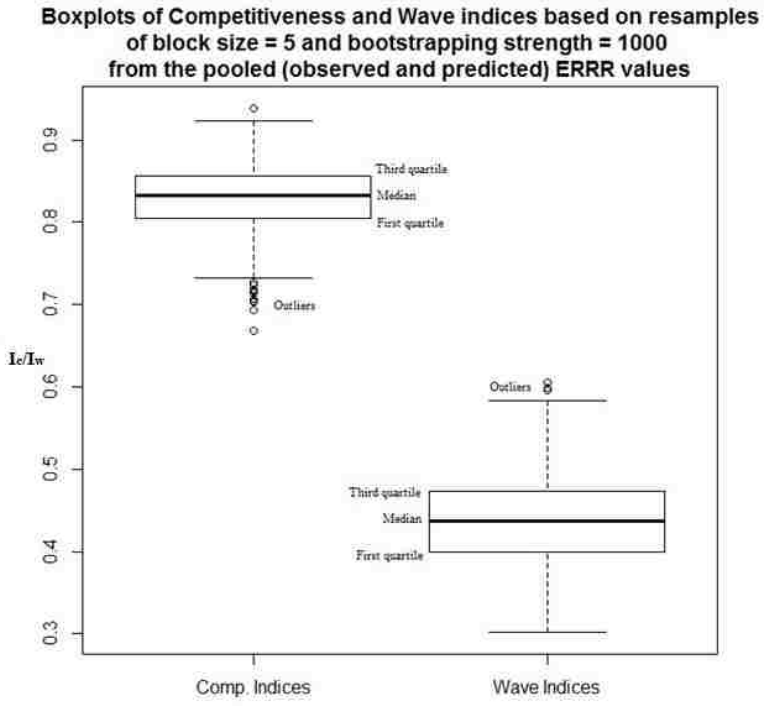


Figure 4.8: I_c and I_w indices from the bootstrapped version of the pooled ERRR volcanic time series

This was done partly to enlarge the size of the parent series and partly to make it weakly stationary without transformations or differencings. These distributions are, however, similar

to those from the observed series.

4.3 ERRR as a validation tool

The problem of comparing two independent Poisson rates has been well studied in literature, especially the ones involving clinical trials. In parallel arm trials, for instance, one often compares the incidence rates λ_1 and λ_2 between two study groups, one receiving a drug of interest, and another receiving a dummy pill (i.e., a placebo effect). Although not essential, in this section, we assume both groups are tracked over the same predetermined number of observations with t_i being the stopping time for the i-th group $i \;\; (i = 1, 2)$. This is to make the comparisons among the established tests detailed below, and our ERRR-based non-parametric proposals more natural. The framework we shall be working under will thus, be a failure truncated one.

4.3.1 A review of established tests

In order to check the over-prevalence of one group in comparison to the other through choosing one of

$$H_0 : \lambda_1 \geq \lambda_2 \;\; \text{vs} \;\; H_a : \lambda_1 < \lambda_2 \tag{4.21}$$

Shan (2015) [133] records a collection of procedures along with various small sample alternative proposals. In keeping with our usual premise, $N_i(t_i) \sim Pois(\lambda_i t_i) \;\; (i = 1, 2)$.

Wald test

Using the m.l.e. $\hat{\theta} = N_2(t)/t - N_1(t)/t$ of the rate difference $\theta = \lambda_2 - \lambda_1$, the Wald statistic is defined as:

$$T_{Wald} = \frac{N_2(t_2)/t_2 - N_1(t_1)/t_1}{\sqrt{N_2(t_2)/t_2^2 - N_1(t_1)/t_1^2}} \tag{4.22}$$

Ng and Tang (2005) show it is asymptotically normally distributed.

Score test

Modification of the estimated variance of $\hat{\theta}$ leads to the score test statistic

$$T_{Score} = \frac{N_2(t_2)/t_2 - N_1(t_1)/t_1}{\sqrt{(N_1(t_1) + N_2(t_2))/(t_1 t_2)}} \tag{4.23}$$

T_{Score} is asymptotically normal as well.

Signed root likelihood ratio test

The signed likelihood ratio statistic, expressed as:

$$T_{LR} = sign\left(\frac{N_2(t_2)}{t_2} - \frac{N_1(t_1)}{t_1}\right)\left\{N_1(t_1)\log\frac{N_1(t_1)}{t_1} + N_2(t_2)\log\frac{N_2(t_2)}{t_2} - (N_1(t_1) + N_2(t_2))\log\frac{N_1(t_1) + N_2(t_2)}{t_1 + t_2}\right\}^{1/2} \tag{4.24}$$

provides an alternative to the Wald-type tests described above.

Conditional test (C-Test)

This exploits the fact that conditional on $N_1(t_1)+N_2(t_2) = n$, the variable $N_2(t_2) \sim Bin(n, \kappa)$ with $\kappa = \frac{t_2}{t_1+t_2}$. Krishnamoorthy and Thomson (2004) [84] have proposed an alternative E-test based on estimated p-values.

4.3.2 Our ERRR-based non-parametric proposals

The location of the $ERRR$ curve introduced previously contains information about the intensities of the competing branches. Recall that by construction, the curve will remain under the 0.5 line if the second series $\{Y_t\}$ in (4.3) is more active or intense than the first. This corresponds to $\lambda_2 > \lambda_1$ where λ_1 and λ_2 represent the Poisson rates for the $\{X_t\}$ and $\{Y_t\}$ series, respectively. As an example, we may return to the eruptions of Mt. Etna considered in Chapters 1 and 3. Choosing 1974 as the year containing the change point (identified by

all the unidirectional and bidirectional tests) separating the first regime from the second, we can create two time series $\{X_t\}$ and $\{Y_t\}$ using (4.1), from the first and second regime, respectively. Table (4.3) below, elaborates (of course, a few values from either the $\{X_t\}$ or the $\{Y_t\}$ series may have to be removed to make their lengths match, which is essential for $ERRR$ creation)

Table 4.3: Time series from regime 1 and 2 for Mt. Etna

Regime 1	Counts	Regime 2	Counts
1941	0	1975	3
1942	1	1976	0
1943	0	1977	0
1944	0	1978	3
1945	0	1979	1
1946	0	1980	0
1947	1	1981	1
1948	0	1982	0
1949	1	1983	1
1950	1	1984	0
1951	0	1985	3
1952	0	1986	1
1953	0	1987	0
1954	0	1988	0
1955	0	1989	2
1956	1	1990	0
1957	0	1991	1
1958	0	1992	0
1959	0	1993	0
1960	0	1994	0
1961	0	1995	0
1962	0	1996	0
1963	0	1997	0
1964	1	1998	0
1965	0	1999	0
1966	0	2000	0
1967	0	2001	1
1968	2	2002	2
1969	0	2003	0
1970	0	2004	1
1971	1	2005	0
1972	0	2006	1
1973	0	2007	0
1974	1	2008	1

and the resulting curve is graphed in Fig (4.9).

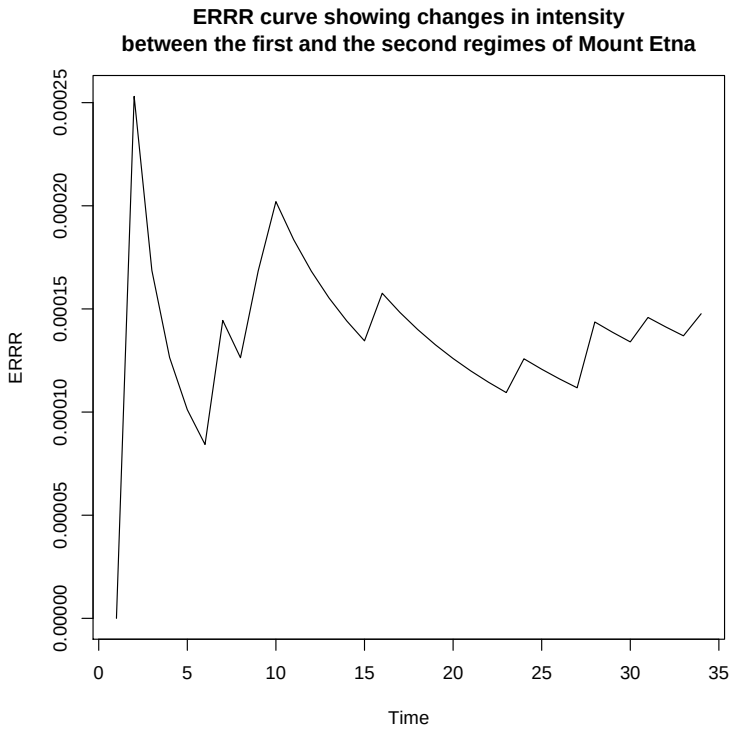


Figure 4.9: $ERRR$ plot suggesting an intense second regime for Mt. Etna

The curve lies entirely below the 0.5 threshold which suggests that the second regime is more intense than the first. All our $ERRR$-based proposals to compare Poisson rates rely on this positioning of the $ERRR$ curve.

ERRR-I_c test

The index of competitiveness I_c introduced by Ho and Bhaduri (2017) [70] and described previously, informs us of $ERRR$'s rough location and is used here to understand the nature of dependence. Low values of I_c (compared to the null threshold) will lead to the rejection of H_0. For instance, in the Mt. Etna example just considered, $I_c = 0$.

ERRR-Wilcoxon

Wilcoxon's signed rank test is a non-parametric test for location and using it on the $ERRR$ values will lead to checking

$$H_0 : \theta \geq 0.5 \quad \text{vs} \quad H_a : \theta < 0.5 \tag{4.25}$$

where θ is the true location parameter for the ERRR distribution.

ERRR-Signed median

This is a different non-parametric test for location) and we will use it on the $ERRR$ values, again with $\theta = 0.5$.

The last two tests are well known in nonparametric literature, and thus, we refrain from unnecessary elaboration. Readers interested in their construction may consult Hollander and Wolfe (1999) [71]. With a given data set and a hypothesized value of the location parameter, p-values are routinely calculated on softwares such as R.

4.3.3 Comparisons

Prior to embarking on a large scale simulation study, we work the steps out on one snapshot in detail, to accentuate the differences among the competing options, through an illustrative example. Employing failure truncation, we have sampled twenty observations from each of the two "arms" or "regimes" or "study groups", whichever interpretation seems to be relevant to the context at hand. The first is an HPP with rate 1, while the second is another HPP with rate 2. The observations are

$$\{N_1\} = 2.276816, 3.026426, 3.659351, 3.905895, 4.247297, 4.276704, 8.165672, .., 21.768739, 22.278512 \tag{4.26}$$

and

$$\{N_2\} = 1.345796, 1.477389, 1.826783, 2.193003, 2.492421, 2.682602, 3.163741, ..., 7.781710, 8.064482 \tag{4.27}$$

Thus, $T_{Wald} = \frac{20/8.064482 - 20/22.278512}{\sqrt{20/8.064482^2 + 20/22.278512^2}} = 2.6829$ with a one-sided p-value of 0.0036,

$T_{Score} = \frac{20/8.064482 - 20/22.278512}{\sqrt{(20+20)/(8.064482 \text{X} 22.278512)}} = 3.3534$ with a one-sided p-value of 0.0004,

$T_{LR} = sign(\frac{20}{8.064482} - \frac{20}{22.278512})\{20 \log \frac{20}{22.278512} + 20 \log \frac{20}{8.064482} - (20+20) \log \frac{20+20}{22.278512+8.064482}\}^{1/2}$

$= 2.22596$, with a one-sided p-value of 0.0130, while the p-value from the conditional test is given by:

$\sum_{i=20}^{40} \frac{40!}{i!(40-i)!} \kappa^i (1-\kappa)^{40-i}$ with $\kappa = \frac{8.064482}{22.278512+8.064482}$. This simplifies to 0.0013.

To implement our ERRR-based alternatives, we note that the smaller terminal time is 8.064482. Thus, we discretize the time axis into $\{1, 2, ..., 9\}$ and count the number of observations from each series falling into the intervals $(0, 1], (1, 2], ...(8, 9]$. With X representing the first sequence and Y, the second:

$$\{X\} = 0, 0, 1, 3, 2, 0, 0, 0, 2 \tag{4.28}$$

and

$$\{Y\} = 0, 3, 3, 4, 0, 5, 2, 2, 1 \tag{4.29}$$

The resulting ERRR values are

$$ERRR(X, Y) = 0.00, 0.00, 0.1428571, 0.2857143, 0.3750000, 0.2857143, 0.2608696, 0.2400000, 0.2857143 \tag{4.30}$$

where, in keeping with FDR convention, we have regarded 0/0 as 0. Since these ERRR values are all less than 0.5, the resulting competitiveness index is 0, and hence the I_c-based p-value is 0. Applying the non-parametric options on (4.30), we have 0.00701453 as the Wilcoxon p-value and 0.00390625 as the signed-median p-value. Table (4.4) below checks whether the different proposals are able to pronounce stable flow in case both arms are part of the same stationary process. With simulations of strength 10^4, the median p-values are stored for different sample sizes.

Table 4.4: Median p-value comparisons with $\lambda_1 = 1, \lambda_2 = 1$, and different sample sizes.

Test	$n = 20$	$n = 30$	$n = 40$
Wald	0.5471	0.4819	0.5107
Score	0.5471	0.4819	0.5107
LR	0.5333	0.4872	0.5076
Conditional C	0.6091	0.5332	0.5551
ERRR-I_c	0.5659	0.4372	0.5054
ERRR-Wilcoxon	0.7614	0.3599	0.5999
ERRR-signed-median	0.7529	0.3668	0.6508

All the tests, with high p-values, are able to detect the similarity in rates under the null assumption. Next, as the second regime gets more intense, we document similar results in Tables (4.5) and (4.6) below.

Table 4.5: Median p-value comparisons with $\lambda_1 = 1, \lambda_2 = 1.5$, and different sample sizes.

Test	$n = 20$	$n = 30$	$n = 40$
Wald	0.0754	0.0744	0.0408
Score	0.0647	0.0673	0.0351
LR	0.1451	0.1471	0.1024
Conditional C	0.0891	0.0869	0.0459
ERRR-I_c	0.1571	0.1863	0.1634
ERRR-Wilcoxon	0.0029	0.0005	$3.19\text{x}10^{-5}$
ERRR-signed-median	0.0009	$8.68\text{x}10^{-5}$	$9.54\text{x}10^{-7}$

Table 4.6: Median p-value comparisons with $\lambda_1 = 1, \lambda_2 = 2$, and different sample sizes.

Test	$n = 20$	$n = 30$	$n = 40$
Wald	0.0169	0.0075	0.0017
Score	0.0079	0.0033	0.0005
LR	0.0498	0.0311	0.0117
Conditional C	0.0145	0.0059	0.0009
ERRR-I_c	0.0000	0.0000	0.0000
ERRR-Wilcoxon	0.0071	0.0008	$7.1\text{x}10^{-5}$
ERRR-signed-median	0.0039	0.0002	$7.6\text{x}10^{-6}$

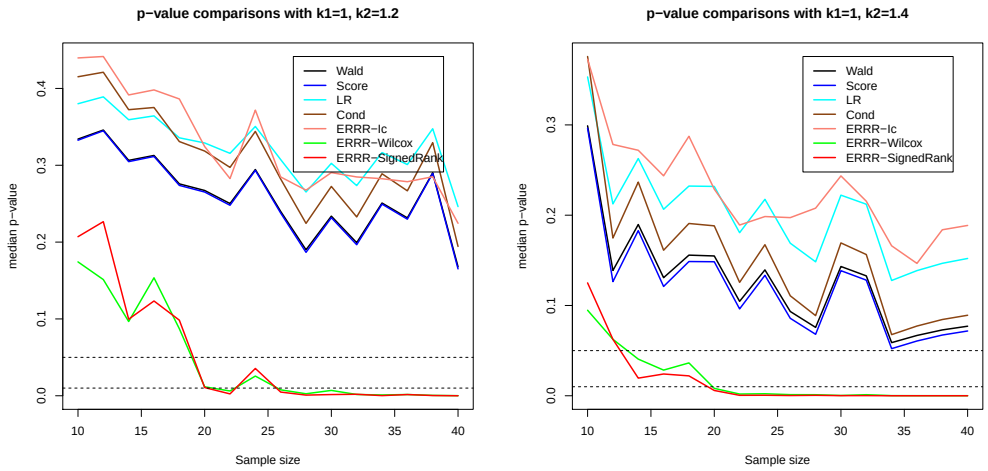
The last two tables demonstrate that when detection is hard, with λ_2 close to λ_1, our

ERRR-based proposals with small p-values, are still successful in detecting the changed rate when the traditional methods fail at the usual 5% or 1% level. Figure (4.10) provides another depiction.

The erratic movements of the p-value curve from the I_c based technique can be controlled by a proper trimming of the $ERRR$ values, especially by getting rid of the initial burn-in period, the definition of which, we believe, is better left at the hands of experts. For small values of λ_2 however, this measure performs almost as well as the established ones.

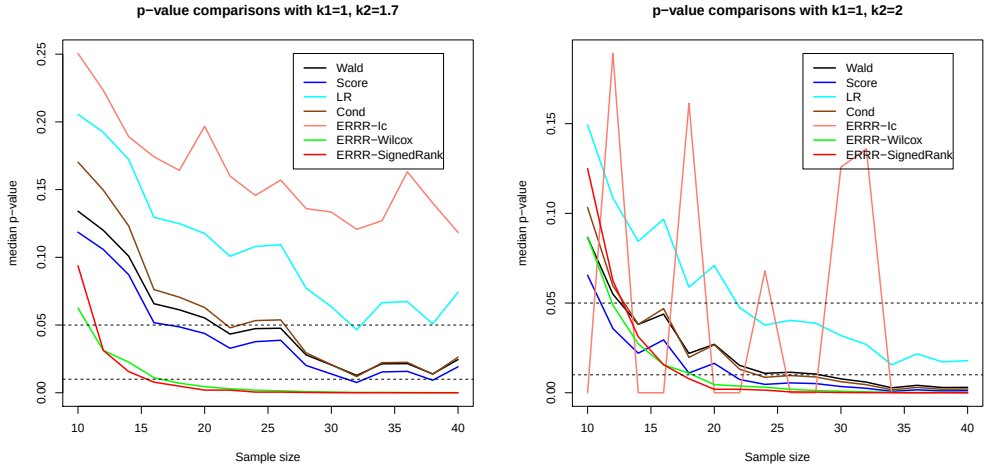
A technical caveat

Purists might correctly want to check the assumptions going through, prior to applying non-parametric tests (such as the Wilcoxon's or signed-rank) on $ERRR$s. One of these wants independence among the "data" values. In a majority of cases, $ERRR$s however, turn out to be dependent, as is evidenced by the first panel of Fig (4.11). The ACF function exhibits a slow decay, indicative of a long-memory process. Decreasing the sampling frequency offers a solution to this dependence problem. As panel (b) demonstrates, sampling every 10-th ERRR value reduces the dependence amount, with the lagged ACF resembling a white noise. The conclusions detailed in the previous sections remain valid under this lagged framework. Table (4.7) below demonstrates. The choice of the lag value should be left to experts in the concerned field.



(a) Median p-value comparison with $\lambda_2 = 1.2$

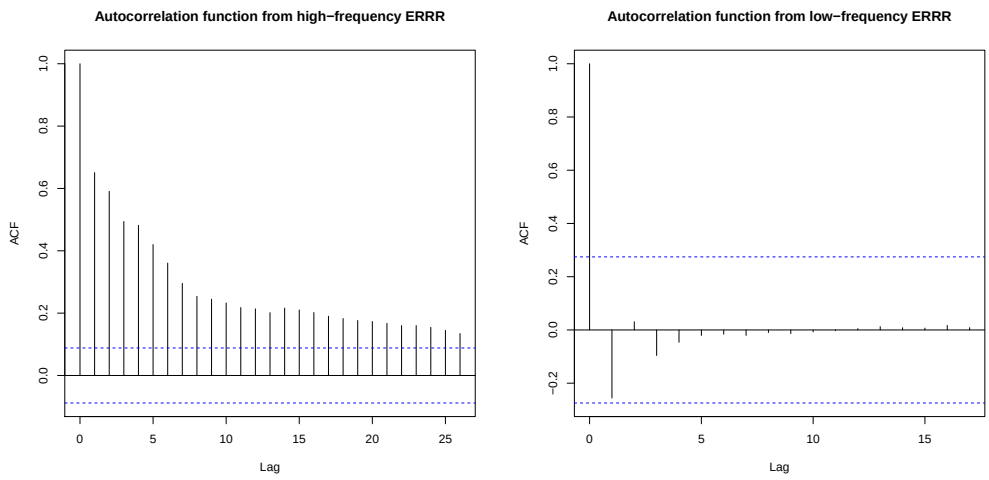
(b) Median p-value comparison with $\lambda_2 = 1.4$



(c) Median p-value comparison with $\lambda_2 = 1.7$

(d) Median p-value comparison with $\lambda_2 = 2$

Figure 4.10: Median p-value comparisons with changing k_2 and sample size.



(a) Slow decay of high frequency ERRR curve

(b) Rapid decay of low frequency ERRR curve

Figure 4.11: Effect on ACF due to a change in $ERRR$'s sampling frequency

Table 4.7: Median p-value comparisons with $\lambda_1 = 1, \lambda_2 = 2$, on lagged ERRR values.

Test	median-pvalue
Wald	0.00065
Score	0.00016
LR	0.00673
Conditional C	0.00033
ERRR-I_c	0.0743
ERRR-Wilcoxon	$2.813\text{x}10^{-10}$
ERRR-signed-median	$2.309\text{x}10^{-14}$

4.4 ERRRs as Hidden Markov chains

Through graphic illustrations and numerical calculations, this section will work on the interpretability of the $ERRR$ curve as a Markov chain (M.C.) governing one of the two PtPs involved. A discrete time, discrete space M.C., the type which the $ERRR$s will eventually be, is well studied in literature. Ross (2010) [132] among many others, is an excellent resource describing this widely used one-step dependence tool. Hidden Markov Models (HMMs), building on M.C.s, are however relatively less known. We shall thus, briefly describe HMMs

in section (4.4.1). The Kilauea-Mauna Loa interaction series and the "Strong-Weak" hurricane interaction series (both described in Chapter 1) will be the two examples surveyed.

We discretized the volcanic ERRR series into what eventually would turn out to be a Markov chain according to the scheme:

$$c(t) = \begin{cases} 1 & \text{if } r_t \in [0, 0.4) \\ 2 & \text{if } r_t \in [0.4, 0.6) \\ 3 & \text{if } r_t \in [0.6, 0.8) \\ 4 & \text{if } r_t \in [0.8, 1] \end{cases}$$

The reason for choosing four partitions and the above thresholds will be explained in due course. Formal tests for checking the Markov property exist in literature (e.g. Zucchini and MacDonald (2009) [150]) and since our ultimate goal would be to interpret $\{C_t : t = 0, 1, 2, ...\}$ as a hidden Markov chain, one of those rigorous tests was first performed. For testing:

$$H_0 : C_1, C_2, ..., C_T | c_0 \text{ are independent} \tag{4.31}$$

$$H_a : C_1, C_2, ..., C_T | c_0 \text{ is a Markov Chain with unknown t.p.m } \Gamma = ((\gamma_{ij})) \tag{4.32}$$

The test statistic $U = 2\sum_{i=1}^{m}\sum_{j=1}^{m} n_{ij}\frac{\hat{\gamma}_{ij}}{\hat{\gamma}_j}$ follows a chi-square distribution with $(m-1)^2$ d.f. under the null hypothesis where m is the number of states n_{ij} is the number of transitions from state i to j, $\hat{\gamma}_{ij} = \frac{n_{ij}}{n_{i+}}$. In volcanic context, $m = 4$ and $U = 206.822$ and since $\chi^2_{9,0.1} = 14.68, \chi^2_{9,0.05} = 16.92, \chi^2_{9,0.01} = 21.67$, we reject the assumption of independence with a reasonable degree of confidence and conclude that the sequence $\{C_t : t = 0, 1, 2, ...\}$ can be treated as a Markov chain.

Once the Markovian property has been objectively established, we can proceed to get an estimate of the underlying transition probability matrix $\Gamma = ((\gamma_{ij}))$ which contains useful information about the rate of flow from state to state. The usual non-parametric estimates

of the transition probabilities are given for $i, j = 1, 2, ..., m$ by:

$$\hat{\gamma}_{ij} = \frac{n_{ij}}{n_{i+}}. \tag{4.33}$$

This estimation leads to the following estimate for the volcanic eruption case:

$$\hat{\Gamma} = \begin{pmatrix} 0.9 & 0.1 & 0 & 0 \\ 0.032 & 0.952 & 0.016 & 0 \\ 0 & 0.0169 & 0.949 & 0.034 \\ 0 & 0 & 0.182 & 0.818 \end{pmatrix}$$

The chain generated from the "Strong-Weak" hurricane interaction will be discretized (once again, due to reasons elaborated later) according to:

$$c(t) = \begin{cases} 1 & \text{if } r_t \in [0, 0.45) \\ 2 & \text{if } r_t \in [0.45, 1] \end{cases}$$

and the estimated transition probability matrix will be

$$\hat{\Gamma} = \begin{pmatrix} 0.911 & 0.089 \\ 0.147 & 0.853 \end{pmatrix}$$

4.4.1 Models of interest

In this section, we will describe and explore the applicability of an established model termed as the Poisson Hidden Markov Model (Poisson-HMM) in the context of the available data and will propose a new alternative, termed ERRR Hidden Markov Model (ERRR-HMM). Through a series of steps covering different facets of statistical inference, we shall strive to establish how the proposed model works as good as, and sometimes even better than the one prevalent in literature with the added advantage of deterministically finding the chain that generates the observations - something that the Poisson-HMM lacks.

The Poisson-HMM

The use of Hidden Markov Models in statistical inference in general and in speech recognition or time series in particular is rather profuse (Ephraim and Merhav (2002) [44], Leroux and

Puterman (1992) [93]). Specifically, a Hidden Markov Model (HMM) $\{X_t : t \in N\}$ is a specific kind of dependent mixture. If $\overrightarrow{X}^{(t)}$ and $\overrightarrow{C}^{(t)}$ represent the histories of the observations and the undelying chain from time 1 to t, we can summarize the simplest model of this kind by:

$$P(C_t|\overrightarrow{C}^{t-1}) = P(C_t|C_{t-1}), t = 2, 3, ... \tag{4.34}$$

$$P(X_t|\overrightarrow{X}^{t-1}, \overrightarrow{C}^{t}) = P(X_t|C_t), t \in N \tag{4.35}$$

The model consists of an unobserved parameter process $\{C_t : t = 1, 2, ...\}$ satisfying the Markov property and a state dependent observable process $\{X_t : t = 1, 2, ...\}$ such that conditional on the knowledge of the current state C_t, the distribution of X_t is independent of the past states or observations. When the state-dependent distributions are chosen to be Poisson i.e. when $\forall x = 0, 1, 2, ..., i = 1, 2, .., m$, the number of states of the hidden chain,

$$p_i(x) = P(X_t = x|C_t = i) = e^{-\lambda_i}\frac{\lambda_i^x}{x!} \tag{4.36}$$

we have the m-state Poisson-HMM studied by Zucchini and MacDonald (2009) [150], among others. With the help of a time series storing earthquake counts, they have nicely demonstrated how the Poisson-HMM can capture serial dependence and can solve the problem of overdispersion. Intuitively, they suppose that each count in the time series is generated by one of m Poisson distributions (which is active at that specific time instant), with means $\lambda_1, \lambda_2, ..., \lambda_m$, where the choice of the mean is made by a second random mechanism, the parameter process. The mean λ_i is selected with probability δ_i where $\overrightarrow{\delta}$ is the stationary distribution of the transition probability matrix (t.p.m) of the underlying chain. If the initial distribution of the chain is $\overrightarrow{\delta}$ and the t.p.m is Γ, the likelihood of observing the sequence

$x_1, x_2, ..., x_t$ is given by:

$$L_T = \overrightarrow{\delta} P(x_1)\Gamma P(x_2)...\Gamma P(x_T)1^{'} \tag{4.37}$$

where P is a $m \times m$ diagonal matrix storing the state-dependent probabilities given in (4.36) above, formally defined by:

$$\begin{pmatrix} p_1(x) & 0 & ... & 0 \\ 0 & p_2(x) & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & p_m(x) \end{pmatrix}$$

If $\overrightarrow{\delta}$, the distribution of C_1, is the stationary distribution of the chain, then (4.37) simplifies to:

$$L_T = \overrightarrow{\delta}\Gamma P(x_1)\Gamma P(x_2)...\Gamma P(x_T)1^{'} \tag{4.38}$$

Since $\overrightarrow{\delta}$ can be solved out using:

$$\overrightarrow{\delta}(I_m - \Gamma + U) = 1^{'} \tag{4.39}$$

where $1^{'}$ is a vector of ones and U is an $m \times m$ matrix of ones, we observe that the real parameters that need to be estimated are the elements of the t.p.m Γ and $\overrightarrow{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_m)$. Zucchini and MacDonald (2009) [150] has elaborated on maximum likelihood estimation in this context, through direct maximization of the likelihood and through the E.M. algorithm. We shall agree to denote the parameters obtained this way by $\overrightarrow{\hat{\lambda}}$ and $\hat{\Gamma}_{P-HMM}$.

The ERRR-HMM

We observe that in the Poisson-HMM model detailed above, the total number of parameters that need to be estimated is $m+(m-1) \times m = m^2$ (assuming that the number of states m is fixed beforehand) and the process will only give us an estimate of the most likely states of the underlying hidden chain. In the presence of a companion series (like Mauna Loa eruption

counts) that is likely to influence the observation series of interest (the Kilauea eruption counts), we feel that the (properly discretized version) of the observed ERRR series can be treated as the underlying Markov chain generating the observations. The t.p.m in (4.37) or in (4.38) need not be estimated anymore, and we can use the non-parametric version of the ERRR-t.p.m. derived in the previous section in those equations. Critics might argue that the estimation issue of the matrix has not yet been fully circumnavigated, but we must realize that such estimation procedure is not directly involved in this specific context and hence the t.p.m can be safely assumed to be known or given. This would lead to a parsimonious model, requiring the estimation of only m of the λ_i's and would additionally provide a real (*not likely* as the Poisson-HMM does) realization of the Markov chain which, in the previous context, would forever, remain unobserved.

Our prescription currently is this: plug in the non-parametric estimate of the t.p.m $\hat{\Gamma}_{ERRR-HMM}$ in (4.37) or (4.38) and get m.l.e's of the λ_i's. Unfortunately, at present, we are having issues with the numerical maximization of the likelihood under this plug-in framework, and so, temporarily, we will go through the Poisson-HMM framework and at the very end, will replace $\hat{\Gamma}_{P-HMM}$ by $\hat{\Gamma}_{ERRR-HMM}$. The estimate $\overrightarrow{\hat{\lambda}}$ will remain unchanged. So the model momentarily will be "pseudo-parsimonious" and while we endeavor to resolve this issue, with the help of the two running examples we shall demonstrate how this even less-than-perfect model performs almost as good as the Poisson-HMM in some aspects and better than it in some others.

Prior to embarking on any detailed inferential analyses, we must pause to realize that the data sets we have explored fit nicely into the framework of a Poisson-HMM: we are fundamentally trying to model counts for which a convenient choice of the state dependent

densities could be Poisson. Marginal overdispersion is observed in the Kilauea eruption series, for instance (with a mean of 0.267 and a variance of 0.401) and for handling such situation, this type of model is rather apt (Zucchini and MacDonald (2009) [150]). Our observation of interest would be the Kilauea eruption count series, and the first hurdle that we stumble upon is an objective choice of the number of hidden states m.

4.4.2 Choice of the number of hidden states

We are aware of the fact that the number of states m can be treated as another parameter inherent in the model and can probably be estimated using a Bayesian framework by putting a non-informative prior on it and by comparing the posterior odds. We shall, however, not tread that path and will instead hope to choose m by exploiting the lowest AIC or BIC criteria while fitting models of different orders. Towards that, we have the following results for the volcanic interactions in Hawaii.

Table 4.8: Optimum no of states (Volcanic interaction)

m	-logLik.	AIC	BIC
2	133.38	274.76	288.62
3	133.24	284.47	315.65
4	126.08	284.17	339.59

Although strict adherence to the AIC criteria would lead us to choose the simple 2-state model, we are particularly encouraged by the significant drop in the negative of the log-likelihood value and hence opt for the slightly complex four-state model. Since we would like to compare the ERRR-HMM model to this Poisson-HMM, we choose to partition the observed continuous ERRR series (which should serve as the underlying chain) into four states as well. For the hurricane case, however, we have the following results:

Table 4.9: Optimum no of states (Hurricane interaction)

m	-logLik.	AIC	BIC
2	184.85	377.72	387.76
3	184.30	386.60	409.20
4	183.81	399.61	439.79

In the latter case, the decrease in the negative log-likelihood value is not appreciable, and so we decided to stick to the 2-state model. We observe in passing that as long as the underlying ERRR series is non-trivial (i.e., not consistently taking a constant value), we can always discretize it into two or more states.

4.4.3 The estimates: Maximum likelihood for Poisson-HMM and plug-in for ERRR-HMM

For numerical maximization of the Poisson-HMM likelihood shown above, we follow Zucchini and MacDonald's (2009) [150] advice of reparametrization and a choice of the observed quantiles as the seed values of the λ_i's. Off-diagonal seeds of 0.05 in Γ seemed to work rather well and the following are the maximum likelihood estimates of the parameters under the Poisson-HMM:

$$\hat{\lambda_1} = 3.49 \times 10^{-69}, \hat{\lambda_2} = 1.03 \times 10^{-32}, \hat{\lambda_3} = 0.8, \hat{\lambda_4} = 0.923 \tag{4.40}$$

$$\hat{\Gamma}_{P-HMM} = \begin{pmatrix} 0.87 & 0.11 & 7.79 \times 10^{-57} & 0.01 \\ 6.39 \times 10^{-7} & 0.19 & 0.81 & 1.92 \times 10^{-25} \\ 0.83 & 1.2 \times 10^{-81} & 0.166 & 4.24 \times 10^{-13} \\ 1.02 \times 10^{-111} & 0.003 & 1.94 \times 10^{-113} & 0.96 \end{pmatrix}$$

Following the prescription in (4.33), the estimates for the ERRR-HMM is given by (4.40) and:

$$\hat{\Gamma}_{ERRR-HMM} = \begin{pmatrix} 0.9 & 0.1 & 0 & 0 \\ 0.032 & 0.952 & 0.016 & 0 \\ 0 & 0.0169 & 0.949 & 0.034 \\ 0 & 0 & 0.182 & 0.818 \end{pmatrix}$$

The negative of the log-likelihood value for the Poisson-HMM was seen in the last section to be 126.08. The corresponding value for the ERRR-HMM turns out to be 150.057 - a difference that is tolerable, especially in the light of the benefits that the latter affords and the estimated matrices seem to agree on most of the 16 spots.

For the hurricane case, the corresponding estimates are:

$$\hat{\lambda_1} = 2.571, \hat{\lambda_2} = 4.874 \tag{4.41}$$

$$\hat{\Gamma}_{P-HMM} = \begin{pmatrix} 0.929 & 0.071 \\ 0.184 & 0.816 \end{pmatrix}$$

and the non-parametric estimate of the t.p.m from the discretized ERRR series turn out to be:

$$\hat{\Gamma}_{ERRR-HMM} = \begin{pmatrix} 0.911 & 0.089 \\ 0.147 & 0.853 \end{pmatrix}$$

4.4.4 Global decoding

Given the observed history, estimation of the most likely sequence of hidden states is often of interest, i.e. one wants $c_1, c_2, .., c_T$ such that the conditional probability:

$$P(\overrightarrow{C^{(T)}} = \overrightarrow{c^{(T)}} | \overrightarrow{X^{(T)}} = \overrightarrow{x^{(T)}}) \tag{4.42}$$

is maximized. The Viterbi algorithm provides a useful aid to the computations and proceeds by writing:

$$\epsilon_{1i} = P(C_1 = i, X_1 = x_1) = \delta_i p_i(x) \tag{4.43}$$

and for $t = 2, 3, ..., T$

$$\epsilon_{ti} = \max_{c_1,..c_{t-1}} P(\overrightarrow{C^{(t-1)}} = \overrightarrow{c^{(t-1)}}, C_t = i, \overrightarrow{X^{(T)}} = \overrightarrow{x^{(T)}}) \tag{4.44}$$

Zucchini and MacDonald (2009) [150] shows that for $t = 2, 3, .., T$ and $i = 1, 2, .., m$, the latter simplifies to:

$$\epsilon_{tj} = (\max_{i}(\epsilon_{t-1,i}\gamma_{ij}))p_j(x_t) \tag{4.45}$$

where γ_{ij} as usual, is the (i, j)th element of the t.p.m. The required maximizing sequence of states $i_i, i_2, .., i_T$ can be found using the recursion:

$$i_T = \underset{i=1,..,m}{\operatorname{argmax}}\, \epsilon_{Ti} \tag{4.46}$$

and for $t = T - 1, T - 2, .., 1$ from:

$$i_t = \underset{i=1,..,m}{\operatorname{argmax}}(\epsilon_{ti}\gamma_{i,i_{t+1}}) \tag{4.47}$$

An application of this algorithm using elements of the estimated t.p.m from the two models leads to Figs (4.12) and (4.13). The second panel describes how the underlying state space is likely to have evolved in the absence of information on the companion series; the third panel describes the corresponding situation in the presence of such knowledge while the top panel is observable, devoid of any estimation process and free of related uncertainties. The close agreement (mostly in terms of the ongoing trend) between the discretized states of ERRR and the estimated states from both models provides empirical evidence for and strengthens our belief in that ERRR can indeed be treated as the chain that generated the Kilauea eruption counts. It may be observed that the ERRR-HMM provides a better approximation to the discretized states of ERRR which is intuitively acceptable since it learns from the estimated t.p.m of this process. But remarkably, the pattern of the estimated states from the two competing models (i.e., the lower two panels) are pretty similar as well.

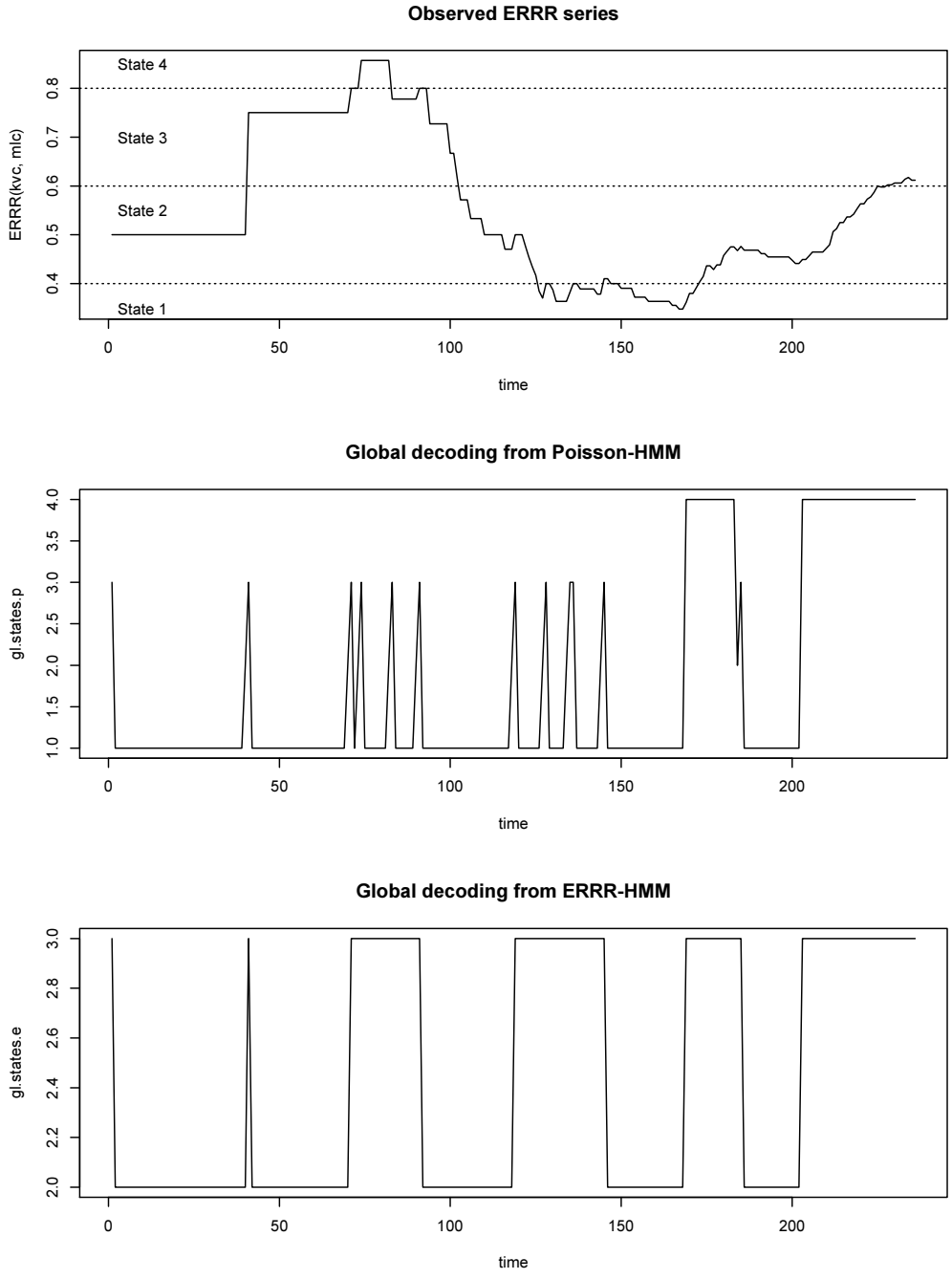


Figure 4.12: Global decoding for volcanic interaction

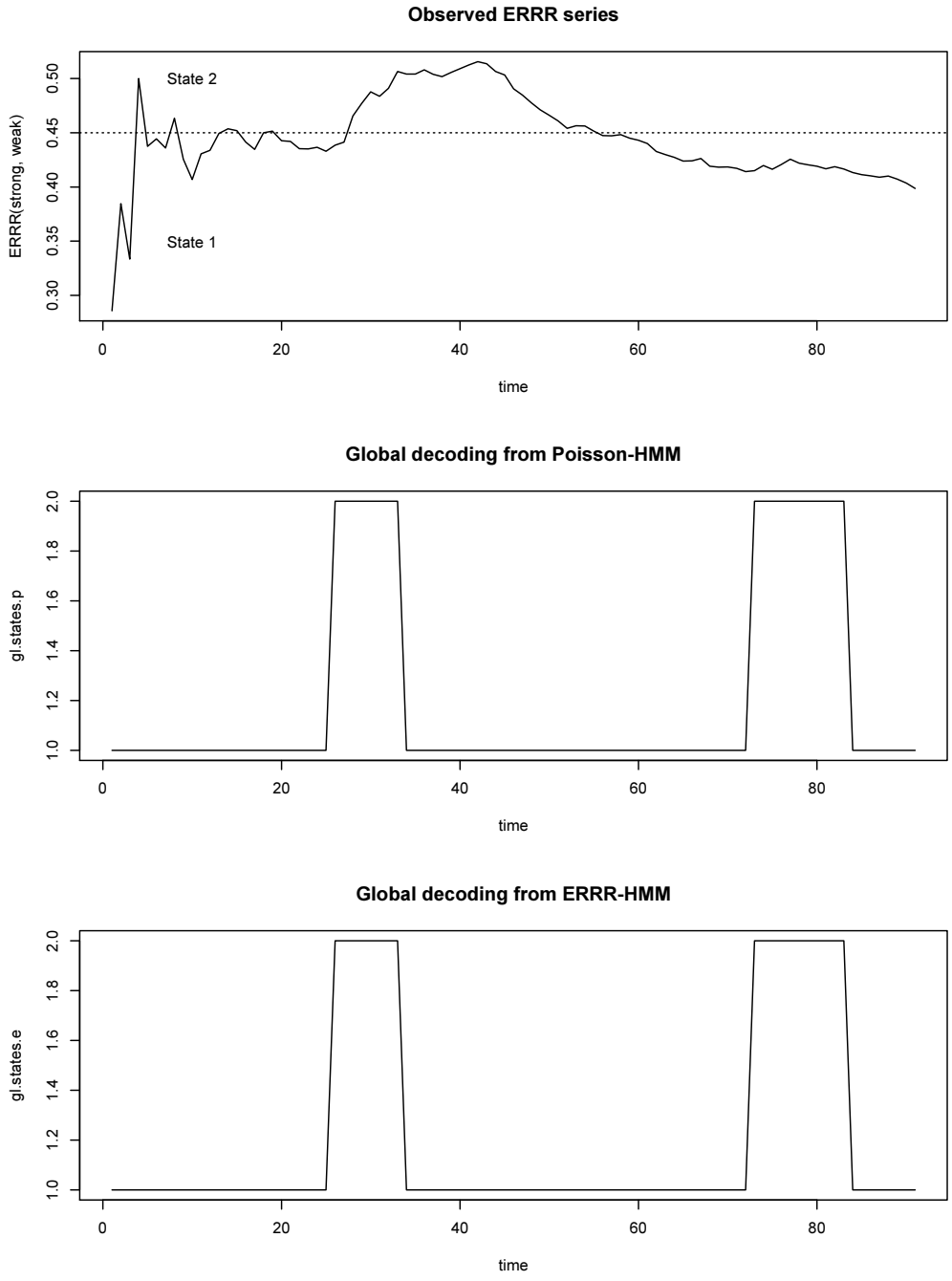


Figure 4.13: Global decoding for hurricane counts

4.4.5 State prediction

If need be, one can predict future states of the hidden chain using the observed history of the process through:

$$P(C_{T+h} = i|\overrightarrow{X^{(T)}} = \overrightarrow{x^{(T)}}) = \overrightarrow{\alpha}_T \Gamma^h(,i)/L_T \tag{4.48}$$

where h is the prediction horizon and $\overrightarrow{\alpha}_t = \delta P(x_1) \prod_{s=2}^{t} \Gamma P(x_s)$. This holds for all $i = 1, 2, ..m$ and thus, for each h, we should have a probability distribution on the state space. At horizon frames of 1, 100 and 500 for instance, i.e. at times $t = 236 + 1 = 237, 236 + 100 = 336, 236 + 500 = 736$, we have the following distributions from the two competing models for the Kilauea-Mauna Loa interaction.

Table 4.10: Probability distribution on state space (P-HMM), volcanic case

m	$h = 1$	$h = 100$	$h = 500$
1	0.137	0.574	0.577
2	0.057	0.091	0.091
3	0.100	0.088	0.088
4	0.705	0.247	0.243

Table 4.11: Probability distribution on state space (ERRR-HMM), volcanic case

m	$h = 1$	$h = 100$	$h = 500$
1	0.005	0.120	0.128
2	0.102	0.383	0.401
3	0.760	0.413	0.378
4	0.132	0.078	0.070

So at $h = 1$, the Poisson-HMM predicts the most likely state to be 4, while the ERRR-HMM predicts it to be 3 and similarly for the other values of h. We observe that as the prediction horizon increases, both the distributions are attracted to the "low" states, and

the variation in state forecasts from the ERRR-HMM is not extremely wild. We repeated this exercise for a sequence of h from 1 through 600, picked the most likely states from each of the two models and created the lower two panel of the adjoining graph (Fig (4.14)).

It is of natural curiosity to inquire how the ERRR sequence, viewed as a time series would flow in the near (or distant) future and whether it follows the pattern predicted by the two lower panels. Towards that, we have performed a traditional time series analysis on the ERRR series, found a SARIMA model of order 32 to work best and have extracted the 600 years' forecasts out of it. The exact fitting mechanism is similar to the ones shown in Tan (2014) [142], Ho and Bhaduri (2015) [68] and Ho et al. (2016) [69], and is not elaborated here, since it will be distracting to the main theme. The results are in close agreement with better performance exhibited by ERRR-HMM, which mostly predicts state 2. Poisson-HMM, which mostly predicts the nearby state of 1, is also rather close.

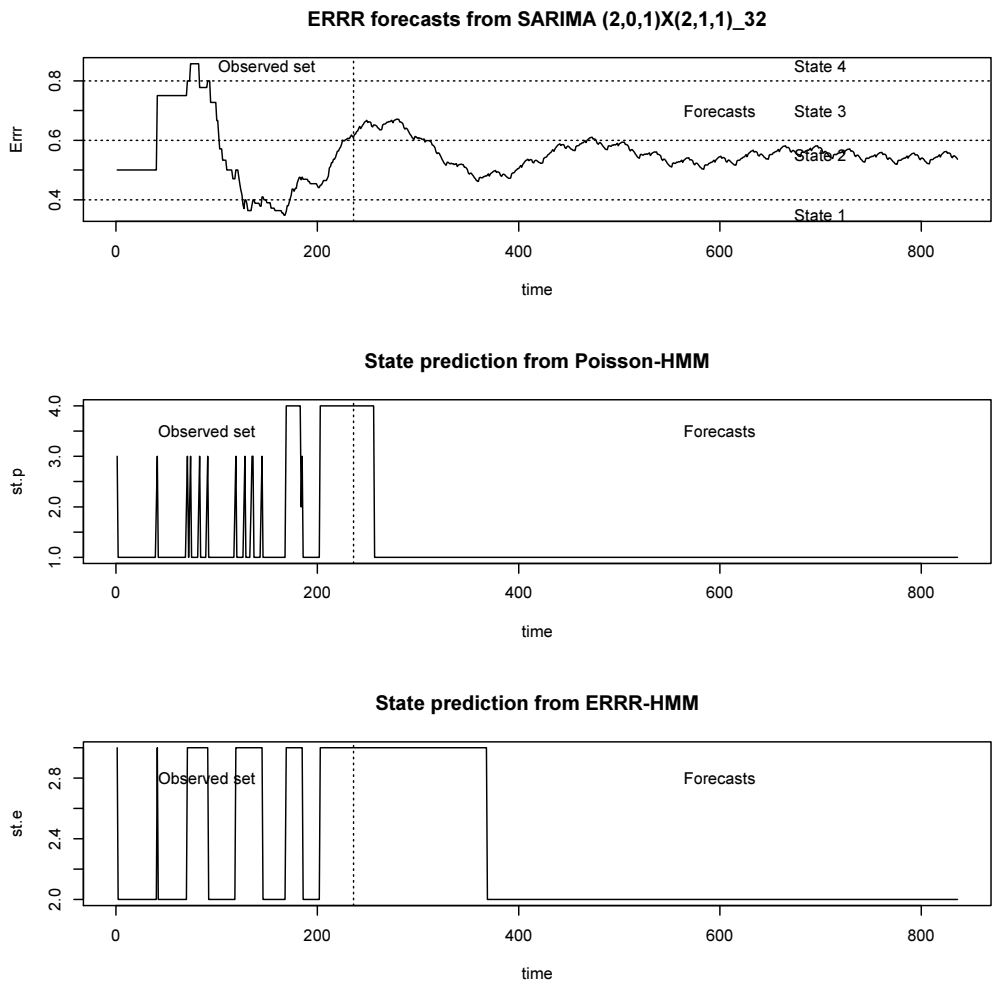


Figure 4.14: State prediction (volcanic case)

4.4.6 Forecast distributions and cross-validation

We can now turn our attention to the values that are actually observable with emphasis on their forecasts and put the two competing models to test in this regard. Zucchini and MacDonald (2009) [150] shows that the forecast distributions can be conveniently expressed as:

$$P(X_{T+h} = x|\overrightarrow{X^{(T)}} = \overrightarrow{x^{(T)}}) = \sum_{i=1}^{m} \epsilon_i(h)p_i(x) \tag{4.49}$$

where $\epsilon_i(h)$ is the i^{th} entry of the vector $\overrightarrow{\alpha}_T/(\overrightarrow{\alpha}_T 1')$. Just as in the previous section, for each value of h, we can expect a probability distribution, not on the state space now, but on the observation space, and with a view to check which model does better in the face of evidence observed already, we adopt a cross-validation type approach: we partition the 236 years' worth of data into a training set of 216 observations and a prediction set of the last 20 observations. We re-estimate the parameters (both the average and the t.p.m's) for each model with the training set and with different values of the forecast horizon, observe the following probability distributions:

Table 4.12: Probability distribution on observation space (P-HMM), volcanic case

x	$h = 1$	$h = 3$	$h = 5$
0	0.507	0.523	0.566
1	0.308	0.298	0.271
2	0.134	0.130	0.118
3	0.039	0.038	0.035
4	0.008	0.008	0.008
5	0.001	0.001	0.001

Although the number of counts, in theory is unbounded, in practice we have not calculated the probabilities beyond $x = 5$ because of their negligibility. Even a cursory glance at the

Table 4.13: Probability distribution on observation space (ERRR-HMM), volcanic case

x	$h=1$	$h=3$	$h=5$
0	0.408	0.422	0.437
1	0.361	0.352	0.344
2	0.165	0.159	0.155
3	0.050	0.049	0.047
4	0.012	0.011	0.011
5	0.002	0.002	0.002

prediction set (last 20 observations) is enough to convince one of the increased activity of Kilauea in the recent years. The probabilities listed above clearly indicate that the ERRR-HMM is a better than Poisson-HMM in picking up this fact which undoubtedly provides further support towards its superiority. One the hurricane side, similar analyses give

Table 4.14: Probability distribution on observation space (P-HMM), hurricane case

x	$h=1$	$h=3$	$h=5$
0	0.059	0.058	0.058
1	0.156	0.154	0.153
2	0.211	0.210	0.209
3	0.199	0.198	0.198
4	0.150	0.150	0.151
5	0.098	0.099	0.100
6	0.059	0.060	0.061
7	0.034	0.035	0.035
8	0.018	0.019	0.019
9	0.009	0.009	0.010
10	0.004	0.004	0.005
11	0.001	0.002	0.002
12	0.000	0.001	0.001

4.4.7 One-out conditional distributions

We now intend to investigate how the distribution of X_t conditioned on all the other observations of the HMM would react under the two different models and also, how do they

Table 4.15: Probability distribution on observation space (ERRR-HMM), hurricane case

x	$h=1$	$h=3$	$h=5$
0	0.050	0.047	0.046
1	0.134	0.129	0.126
2	0.190	0.184	0.181
3	0.190	0.187	0.186
4	0.155	0.156	0.157
5	0.112	0.115	0.117
6	0.074	0.078	0.080
7	0.046	0.048	0.049
8	0.026	0.028	0.029
9	0.013	0.014	0.015
10	0.006	0.007	0.007
11	0.003	0.003	0.003
12	0.001	0.001	0.001

compare with the actual observed counts. Using $\overrightarrow{X}^{(-t)} := (X_1, .., X_{t-1}, X_{t+1}, .., X_T)$, the required conditional density is given by (Zucchini and MacDonald (2009) [150])

$$P(X_t = x|\overrightarrow{X}^{(-t)} = \overrightarrow{x}^{(-t)}) = \sum_{i=1}^{m} w_i(t)p_i(x) \tag{4.50}$$

where the scales $w_i(t)$ are appropriate functions of the observations $\overrightarrow{x}^{(-t)}$ and the model parameters. For reasonable values of x once again, such distributions are calculated and compared in Figs (4.15) and (4.16) both between them and also with the actual observed counts. Remembering what this conditional distribution does, we can see that the ERRR-HMM can borrow strength more efficiently from the remaining observations as compared to the Poisson-HMM and is able to do far better in detecting large values of x. To clarify, the figure corresponding to $t = 212$ in the volcanic case (Fig(4.15)) demonstrates how using the remaining observations, the true number of Kilauea eruptions (i.e., 4) seems more plausible under the ERRR-HMM framework. Incidentally, in the context of rare and catastrophic events such as volcanic eruptions, such high values are more worrying and hence should be

estimated with better efficiency.

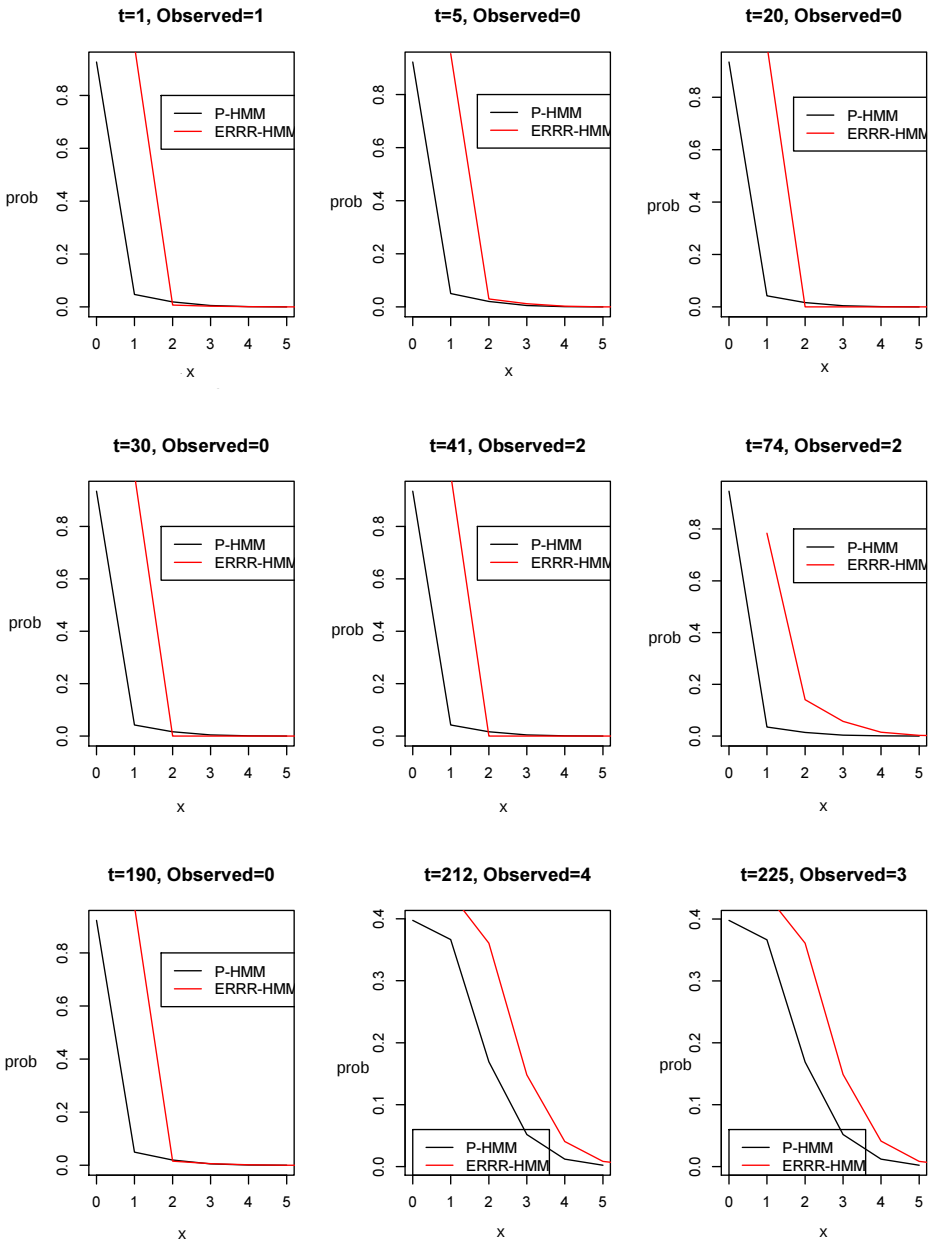


Figure 4.15: One-out conditionals, volcanic case

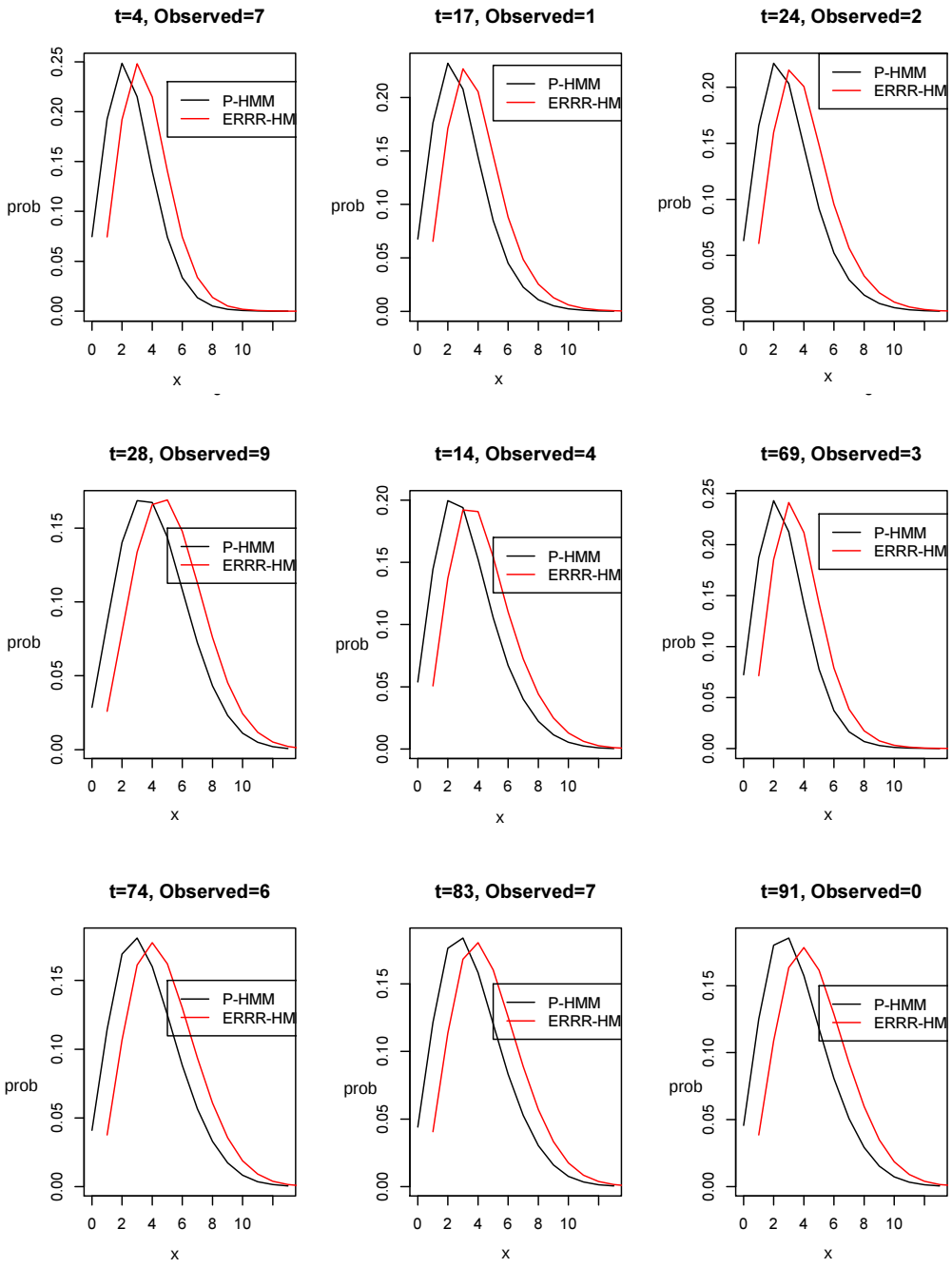


Figure 4.16: One-out conditionals, hurricane case

4.4.8 Analysis of pseudo-residuals

Zucchini and MacDonald (2009) [150] shows that if the underlying model is correct, then

$$z_t = \Phi^{-1}(P(X_t \leq x_t | \overrightarrow{X^{-t}} = \overrightarrow{x^{-t}})) \tag{4.51}$$

is a realization of a standard normal variable for the continuous case. For the discrete version, the normal pseudo-residual segment is $[z_t^-, z_t^+]$ where

$$z_t^- = \Phi^{-1}(P(X_t < x_t | \overrightarrow{X^{-t}} = \overrightarrow{x^{-t}})) \tag{4.52}$$

$$z_t^+ = \Phi^{-1}(P(X_t \leq x_t | \overrightarrow{X^{-t}} = \overrightarrow{x^{-t}})) \tag{4.53}$$

Figures (4.17) and (4.18) below compare the competing models with respect to this diagnostic tool, and find them to be more or less equally effective.

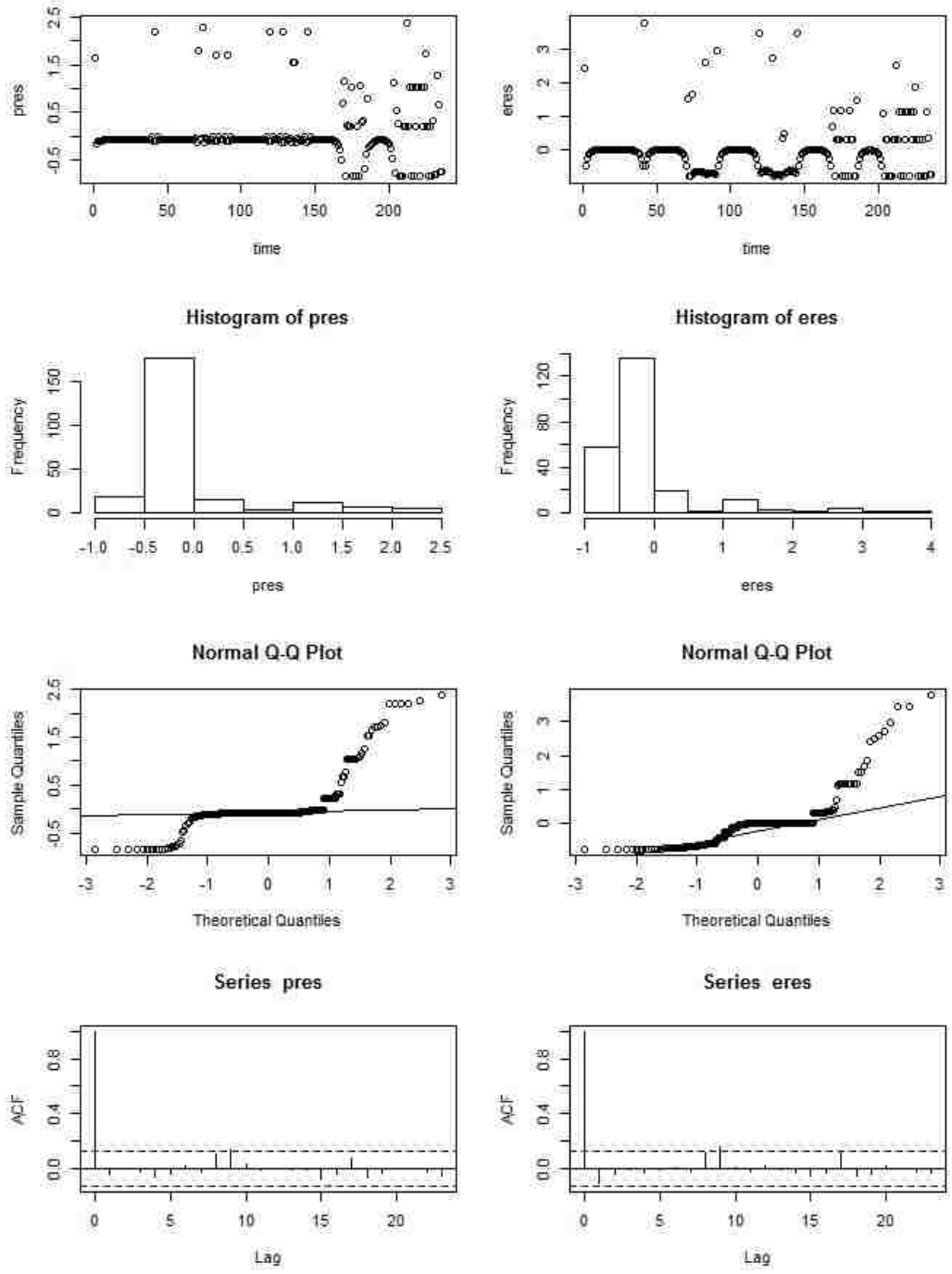


Figure 4.17: Pseudo-residuals (volcanic case). Left panel: $P-HMM$, right panel: $ERRR-HMM$

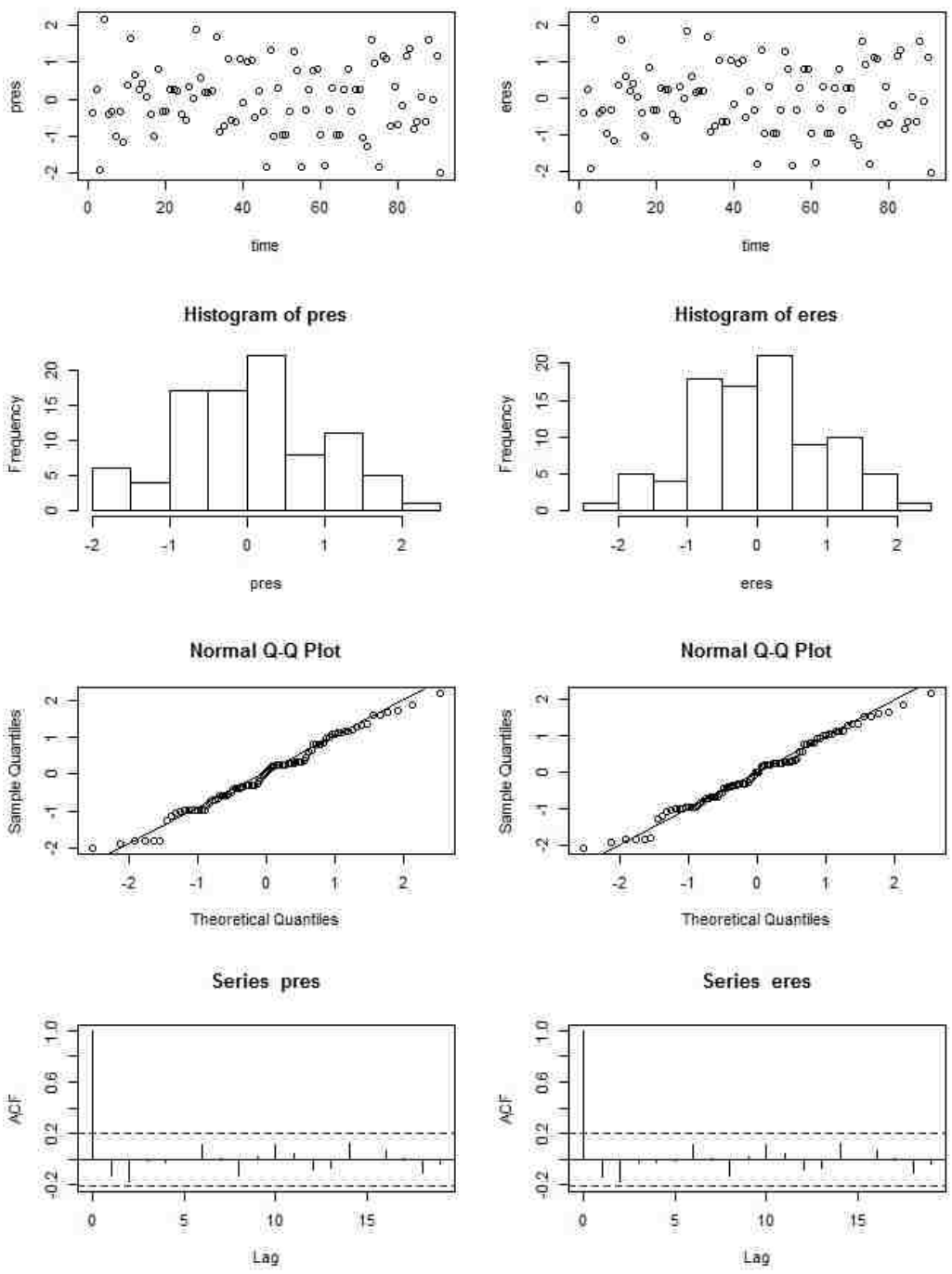


Figure 4.18: Pseudo-residuals (hurricane case). Left panel: $P-HMM$, right panel: $ERRR-HMM$

In summary, through the final section, we have tried to establish that the statistic ERRR can be conveniently discretized and converted to a Markov chain which can be thought to generate one of the PtPs. Traditional HMM analysis assumes the chain to be unobservable and utilizes computation heavy algorithms only to get the most likely chain. Bereft of tedious calculations, ERRR-HMM readily provides an observable version of it, and the inferences are better or at least in close agreement. In each of the inferential aspects explored, we have let the initial (stationary) distribution of the discretized ERRR chain to be dictated by its t.p.m. Had we forced this initial distribution to coincide with the one generated by the Poisson-HMM, we would possibly have had better conclusions. Zucchini and MacDonald (2009) [150] claims that if the Markov chain underlying a stationary HMM is time reversible, the HMM is so, too. For two state chains such as the hurricane case, the chain should certainly be time reversible. We know that an irreducible (homogeneous, discrete time, finite space) Markov chain has a unique, strictly positive stationary distribution. As long as our ERRR series is not monotonic, we can always find a partition that ensures the discretized chain will be irreducible. Stationary distributions can thus be found and the method detailed above can be carried out. Even after an objective choice of the number of hidden states, subjectivity still remains regarding the placement of the partitions. Towards that, we can say that a detailed sensitivity analysis with regards to this aspect on the two examples explored did not generate drastically different conclusions.

Chapter 5

Conclusions

Change point detection is arguably one of the pivotal problems confronting modern inference, embracing notions from control theory, estimation, and hypothesis testing. Using predominantly frequentist techniques, this dissertation, anchored to the Poissonian framework of event generation, addressed this problem with the intention of forging a nexus between a powerful test and a change detection algorithm. This is in keeping with current practice in this domain (for instance, with the CPM framework studied) where a testing-estimation routine seems unavoidable. Ingenuity was demonstrated, however, in conceiving the two protagonists.

The prospect of flipping time was noted to be proficuous by Ho (1993) [66] in creating a backward version of an already existing statistic. As its first exercise, this study unearthed new instances of intensities, both smooth (Dimitrakopoulou et al. (2007) [35]) and rough (intensities with two steps and more, and their mixtures) where the test using this backward statistic, retains its superiority, with regards to detecting non-stationarity. Piquant schemes of borrowing strength from both versions (through taking their maximum, choosing their minimum p-value, etc.) were then proposed to create a unified category of tools called "bidirectional tests". The equable thread that binds every member of this family is an effort

to preserve the anomaly identification properties, optimum under the working intensity, regardless of the version (the established forward Z or the backward Z_B) that provides it, and yet achieve higher power. Two palmary members of this family were then identified to carry out the estimation task, to be undertaken next.

The proposed algorithm exploits ideas from testing multiple hypotheses and False Discovery Rate (FDR) control. When run with the best bidirectional member, it proves to be successful in promptly detecting and estimating even minor deviations from stationarity. We have offered two versions of the prescription, and sensitivity expectations or available computational resources should dictate the choice. In addition to enjoying the classification accuracy of a test more powerful than the prevalent unidirectional ones, the weaker version relaxes the need to conduct two or more tests simultaneously while the algorithm is being run. The latter approach was advised by Chen (2010) [24]. Stifling assumptions on such parameters as the number of change points possible, have not been imposed, and extensive simulations and real examples demonstrate the applicability in diverse fields. Attention has been paid to ease the computational burden, and graphic tools, often rendering useful corollaries (such as efficient prediction, interpretability as Hidden Markov chains, etc.), have been offered as validation instruments.

The research that this dissertation initiates is, however, far from over. Change point detection problems in time series are currently being handled by computer scientists through relative density ratio estimation (Liu et al. (2013) [100]) and random forest techniques (Auret and Aldrich (2010) [5]). Bhaduri, Zhan, and Chiu (2017) [15] proposed a class of stochastic *weak learning estimators* in the context of dynamically evolving systems. Under relatable frameworks, it will be interesting to examine how the bidirectional class competes

against these alternatives. FDR control and ordered hypothesis testing are gaining prominence through *accumulation tests* introduced recently by Li and Barber (2017) [94] and our algorithm can be embellished upon by considering their insights, or by controlling positive-FDR (Storey (2003) [138]), for instance. Under a rough intensity scenario, tests may be constructed to guess the most likely number of steps and the location of knots. Similarity between PtPs has been studied by several authors, most notably, Kalzanov (1970) [77], who introduced a distance-type metric between the two distribution functions of the inter-event times, and Rand (1971) [117], who examined similarity through segment membership of different pairs of observations. Finding ways of incorporating these ideas into the detection algorithm, merits effort. An R package may be developed for ready implementation of our proposed methodology.

With complications and intricacies originating from myriad sources, the research climate for change expiscation is growing procellous with each passing day. Beguiled by the fecund foundation laid, we vow to stalk, with renewed verve and anticipation, the bidirectional class's blossoming, and remain confident of its efficient response to the exigencies of a tumultuous non-stationary future.

Appendix A: Ranked p-values from the Z-test on the whole process (Mt. Etna)

Test id	Ranked p-value	BH
61	0.000004943087942	0.0008196721311
60	0.000007102322927	0.001639344262
56	0.0000093608746	0.002459016393
59	0.00001018716082	0.003278688525
55	0.00001280563653	0.004098360656
58	0.00001459336562	0.004918032787
57	0.0000229692616	0.005737704918
54	0.0000251782327	0.006557377049
53	0.00003105257094	0.00737704918
52	0.00005289834801	0.008196721311
51	0.0001007964124	0.009016393443
50	0.0001696307746	0.009836065574
49	0.0002426596414	0.0106557377
48	0.0003440195984	0.01147540984
47	0.0005130903299	0.01229508197
46	0.0008496615905	0.0131147541
45	0.001464533111	0.01393442623
44	0.002459381927	0.01475409836
43	0.003314336522	0.01557377049
42	0.005215017362	0.01639344262
41	0.007962090384	0.01721311475
40	0.01286469193	0.01803278689
39	0.0165993583	0.01885245902
38	0.02141540717	0.01967213115
37	0.03257389687	0.02049180328
35	0.03580081964	0.02131147541
36	0.0388791885	0.02213114754
34	0.03934395714	0.02295081967
33	0.05709060267	0.0237704918
30	0.05812311929	0.02459016393
29	0.06468381119	0.02540983607
32	0.07362151428	0.0262295082

28	0.07614297032	0.02704918033
27	0.07649347305	0.02786885246
31	0.08524508684	0.02868852459
24	0.0954898563	0.02950819672
23	0.1042650064	0.03032786885
26	0.1084977748	0.03114754098
22	0.1361354441	0.03196721311
25	0.1482536128	0.03278688525
21	0.1711449883	0.03360655738
14	0.1809190262	0.03442622951
13	0.1920789062	0.03524590164
20	0.2056377189	0.03606557377
15	0.2360564393	0.0368852459
19	0.2456975735	0.03770491803
17	0.2575871762	0.03852459016
16	0.261195852	0.0393442623
12	0.2769413146	0.04016393443
18	0.2978102764	0.04098360656
10	0.3171459564	0.04180327869
11	0.3272108335	0.04262295082
3	0.4024184082	0.04344262295
9	0.4131393761	0.04426229508
7	0.4308539278	0.04508196721
8	0.4917445048	0.04590163934
4	0.6443505328	0.04672131148
1	0.6529830323	0.04754098361
6	0.6544987963	0.04836065574
2	0.8130724892	0.04918032787
5	0.9888456031	0.05

Appendix B: Ranked p-values from the Z_B-test on the whole process (Mt. Etna)

Test id	Ranked p-value	BH
55	0.00000001942997829	0.0008196721311
52	0.00000005039281525	0.001639344262
53	0.0000001596420469	0.002459016393
56	0.0000004625586953	0.003278688525
54	0.000001689198267	0.004098360656
51	0.00000430953076	0.004918032787
47	0.000005386461944	0.005737704918
48	0.000009717672124	0.006557377049
46	0.00001049194415	0.00737704918
61	0.00001455600076	0.008196721311
49	0.00001568116004	0.009016393443
50	0.00001710726971	0.009836065574
60	0.00002086860746	0.0106557377
59	0.00003085407754	0.01147540984
58	0.0000465430868	0.01229508197
45	0.0000966166795	0.0131147541
57	0.000143757168	0.01393442623
43	0.00043487751	0.01475409836
44	0.0007938378494	0.01557377049
42	0.001426623997	0.01639344262
41	0.002329017367	0.01721311475
6	0.02120557212	0.01803278689
40	0.02523164687	0.01885245902
38	0.03022646284	0.01967213115
39	0.03265153509	0.02049180328
7	0.03565259439	0.02131147541
34	0.065719317	0.02213114754
35	0.1107571925	0.02295081967
27	0.1144994673	0.0237704918
5	0.1226855514	0.02459016393
3	0.1321114403	0.02540983607
37	0.1415490663	0.0262295082

23	0.1569500487	0.02704918033
30	0.1858172486	0.02786885246
29	0.1858826209	0.02868852459
13	0.1921243749	0.02950819672
36	0.1941121472	0.03032786885
33	0.2002182532	0.03114754098
24	0.2047482687	0.03196721311
28	0.2273046814	0.03278688525
22	0.2473751119	0.03360655738
26	0.2789840633	0.03442622951
10	0.2988885683	0.03524590164
14	0.3017283711	0.03606557377
32	0.3487205709	0.0368852459
21	0.3573262561	0.03770491803
12	0.4151648096	0.03852459016
9	0.4209662152	0.0393442623
20	0.4583324904	0.04016393443
8	0.4602265212	0.04098360656
31	0.4776843603	0.04180327869
11	0.4964767453	0.04262295082
19	0.5998942672	0.04344262295
25	0.6104584302	0.04426229508
15	0.6158370696	0.04508196721
1	0.6529830323	0.04590163934
17	0.7068332469	0.04672131148
16	0.7277255267	0.04754098361
4	0.8396791684	0.04836065574
2	0.8559219995	0.04918032787
18	0.8822767596	0.05

Appendix C: Ranked p-values from the R-test on the first 50 observations (Mt. Etna)

Test id	Ranked p-value	BH
41	0.0025	0.001020408163
42	0.0025	0.002040816327
43	0.0025	0.00306122449
44	0.0025	0.004081632653
45	0.0025	0.005102040816
46	0.0025	0.00612244898
47	0.0025	0.007142857143
48	0.0025	0.008163265306
49	0.0025	0.009183673469
6	0.0175	0.01020408163
40	0.0175	0.0112244898
7	0.0375	0.01224489796
38	0.0375	0.01326530612
39	0.0375	0.01428571429
34	0.075	0.01530612245
5	0.15	0.01632653061
13	0.15	0.01734693878
23	0.15	0.01836734694
27	0.15	0.0193877551
29	0.15	0.02040816327
30	0.15	0.02142857143
33	0.15	0.02244897959
35	0.15	0.02346938776
36	0.15	0.02448979592
37	0.15	0.02551020408
10	0.25	0.02653061224
22	0.25	0.02755102041
24	0.25	0.02857142857
26	0.25	0.02959183673
28	0.25	0.0306122449
3	0.4	0.03163265306
8	0.4	0.03265306122

9	0.4	0.03367346939
11	0.4	0.03469387755
12	0.4	0.03571428571
14	0.4	0.03673469388
20	0.4	0.03775510204
21	0.4	0.0387755102
31	0.4	0.03979591837
32	0.4	0.04081632653
1	0.75	0.04183673469
2	0.75	0.04285714286
4	0.75	0.04387755102
15	0.75	0.04489795918
16	0.75	0.04591836735
17	0.75	0.04693877551
18	0.75	0.04795918367
19	0.75	0.04897959184
25	0.75	0.05

Appendix D: Ranked p-values from the P_{DB}-test on the first 50 observations (Mt. Etna)

Test id	Ranked p-value	BH
47	0	0.001020408163
41	0.0025	0.002040816327
42	0.0025	0.00306122449
43	0.0025	0.004081632653
44	0.0025	0.005102040816
45	0.0025	0.00612244898
46	0.0025	0.007142857143
48	0.0025	0.008163265306
49	0.0075	0.009183673469
6	0.0375	0.01020408163
38	0.0375	0.0112244898
39	0.0375	0.01224489796
40	0.0375	0.01326530612
7	0.075	0.01428571429
34	0.075	0.01530612245
35	0.075	0.01632653061
36	0.075	0.01734693878
37	0.075	0.01836734694
24	0.15	0.0193877551
27	0.15	0.02040816327
28	0.15	0.02142857143
29	0.15	0.02244897959
30	0.15	0.02346938776
31	0.15	0.02448979592
32	0.15	0.02551020408
33	0.15	0.02653061224
3	0.25	0.02755102041
5	0.25	0.02857142857
13	0.25	0.02959183673
14	0.25	0.0306122449
21	0.25	0.03163265306
22	0.25	0.03265306122

23	0.25	0.03367346939
25	0.25	0.03469387755
26	0.25	0.03571428571
10	0.35	0.03673469388
12	0.35	0.03775510204
15	0.35	0.0387755102
16	0.35	0.03979591837
17	0.35	0.04081632653
18	0.35	0.04183673469
19	0.35	0.04285714286
20	0.35	0.04387755102
8	0.55	0.04489795918
9	0.55	0.04591836735
11	0.55	0.04693877551
1	0.75	0.04795918367
2	0.85	0.04897959184
4	0.925	0.05

Appendix E: Ranked p-values from the Z-test for the second regime (Mt. Etna)

Test id	Ranked p-value	BH
16	0.03053105253	0.0025
17	0.03803814244	0.005
18	0.04001599977	0.0075
19	0.04218615555	0.01
20	0.04457540001	0.0125
3	0.1352083669	0.015
1	0.2051282051	0.0175
9	0.2185917881	0.02
10	0.2722428017	0.0225
13	0.2726137302	0.025
4	0.2733656995	0.0275
8	0.2793013439	0.03
15	0.303185794	0.0325
2	0.3044004016	0.035
14	0.4041262415	0.0375
7	0.4042843554	0.04
11	0.4236103038	0.0425
5	0.46520855	0.045
12	0.4698107875	0.0475
6	0.54100298	0.05

Appendix F: Ranked p-values from the Z_B-test for the second regime (Mt. Etna)

Test id	Ranked p-value	BH
16	0.002717885183	0.0025
17	0.02977388948	0.005
18	0.05371973481	0.0075
19	0.08372560997	0.01
20	0.1158422112	0.0125
3	0.1191436052	0.015
9	0.180707555	0.0175
1	0.2051282051	0.02
8	0.2346468783	0.0225
13	0.2389519006	0.025
11	0.268618233	0.0275
15	0.3281177433	0.03
10	0.430398062	0.0325
14	0.4747686349	0.035
7	0.4763004965	0.0375
2	0.4832001146	0.04
5	0.5904182423	0.0425
4	0.9091774219	0.045
6	0.9134659	0.0475
12	0.9899680316	0.05

Appendix G: Ranked p-values from the R-test for the second regime (Mt. Etna)

Test id	Ranked p-value	BH
19	0	0.0025
20	0	0.005
16	0.075	0.0075
17	0.075	0.01
18	0.075	0.0125
3	0.4	0.015
8	0.4	0.0175
9	0.4	0.02
10	0.4	0.0225
11	0.4	0.025
13	0.4	0.0275
15	0.4	0.03
1	0.75	0.0325
2	0.75	0.035
4	0.75	0.0375
5	0.75	0.04
6	0.75	0.0425
7	0.75	0.045
12	0.75	0.0475
14	0.75	0.05

Appendix H: Ranked p-values from the P_{DB}-test for the second regime (Mt. Etna)

Test id	Ranked p-value	BH
20	0	0.0025
1	0.0025	0.005
17	0.0025	0.0075
2	0.0375	0.01
18	0.0375	0.0125
4	0.075	0.015
19	0.075	0.0175
10	0.15	0.02
3	0.25	0.0225
9	0.25	0.025
14	0.25	0.0275
8	0.35	0.03
11	0.35	0.0325
16	0.35	0.035
12	0.55	0.0375
5	0.925	0.04
6	0.925	0.0425
7	0.925	0.045
13	0.9625	0.0475
15	0.9625	0.05

Appendix I: Ranked p-values from the Z-test for the first regime (DJIA)

Test id	Ranked p-value	BH
49	2.18E-15	0.001020408163
48	8.71E-15	0.002040816327
46	1.59E-14	0.00306122449
47	3.08E-14	0.004081632653
45	6.16E-14	0.005102040816
44	2.23E-13	0.00612244898
43	8.59E-13	0.007142857143
42	3.27E-12	0.008163265306
41	1.18E-11	0.009183673469
40	4.30E-11	0.01020408163
35	6.40E-11	0.0112244898
39	1.51E-10	0.01224489796
34	2.43E-10	0.01326530612
38	5.09E-10	0.01428571429
33	8.98E-10	0.01530612245
37	1.73E-09	0.01632653061
32	3.10E-09	0.01734693878
36	5.72E-09	0.01836734694
31	9.74E-09	0.0193877551
30	3.41E-08	0.02040816327
29	6.22E-08	0.02142857143
28	2.03E-07	0.02244897959
27	6.19E-07	0.02346938776
26	1.74E-06	0.02448979592
25	5.44E-06	0.02551020408
24	1.63E-05	0.02653061224
23	4.64E-05	0.02755102041
22	0.0001235767114	0.02857142857
21	0.0003007180559	0.02959183673
20	0.0007718935122	0.0306122449
19	0.00158901732	0.03163265306
18	0.003581549277	0.03265306122
17	0.007218956194	0.03367346939

16	0.01398218526	0.03469387755
15	0.02880627938	0.03571428571
6	0.03561792393	0.03673469388
14	0.05480955458	0.03775510204
13	0.07289474711	0.0387755102
5	0.1069806187	0.03979591837
12	0.1285694159	0.04081632653
11	0.222618907	0.04183673469
8	0.2594836933	0.04285714286
4	0.2874266603	0.04387755102
1	0.2989100817	0.04489795918
10	0.3634506048	0.04591836735
7	0.4552340209	0.04693877551
9	0.5454772057	0.04795918367
3	0.6448744181	0.04897959184
2	0.8659836365	0.05

Appendix J: Ranked p-values from the Z_B-test for the first regime (DJIA)

Test id	Ranked p-value	BH
35	0	0.001020408163
43	0	0.002040816327
44	0	0.00306122449
45	0	0.004081632653
46	0	0.005102040816
34	6.66E-16	0.00612244898
49	2.00E-15	0.007142857143
42	8.22E-15	0.008163265306
33	2.20E-14	0.009183673469
41	4.75E-14	0.01020408163
29	6.46E-14	0.0112244898
48	1.37E-13	0.01224489796
26	1.99E-13	0.01326530612
32	2.84E-13	0.01428571429
31	4.97E-13	0.01530612245
28	8.63E-13	0.01632653061
47	2.29E-12	0.01734693878
40	3.17E-12	0.01836734694
27	3.80E-12	0.0193877551
25	1.95E-11	0.02040816327
30	3.54E-11	0.02142857143
39	1.08E-10	0.02244897959
38	8.56E-10	0.02346938776
24	1.10E-09	0.02448979592
23	3.05E-08	0.02551020408
37	4.98E-08	0.02653061224
22	3.57E-07	0.02755102041
21	6.47E-07	0.02857142857
36	2.47E-06	0.02959183673
6	2.11E-05	0.0306122449
20	2.51E-05	0.03163265306
19	2.88E-05	0.03265306122
5	4.16E-05	0.03367346939

18	0.0002081008264	0.03469387755
17	0.0005063089311	0.03571428571
16	0.0006936279382	0.03673469388
4	0.00180598794	0.03775510204
15	0.006652678286	0.0387755102
13	0.008909579235	0.03979591837
2	0.01315918152	0.04081632653
12	0.02169336306	0.04183673469
14	0.03904152939	0.04285714286
3	0.05567207285	0.04387755102
11	0.1047896358	0.04489795918
8	0.1975646226	0.04591836735
1	0.2989100817	0.04693877551
10	0.4809663822	0.04795918367
9	0.7040870037	0.04897959184
7	0.9214588344	0.05

Appendix K: Ranked p-values from the R-test for the first regime (DJIA)

Test id	Ranked p-value	BH
4	0.0025	0.001020408163
5	0.0025	0.002040816327
6	0.0025	0.00306122449
16	0.0025	0.004081632653
17	0.0025	0.005102040816
18	0.0025	0.00612244898
19	0.0025	0.007142857143
20	0.0025	0.008163265306
21	0.0025	0.009183673469
22	0.0025	0.01020408163
23	0.0025	0.0112244898
24	0.0025	0.01224489796
25	0.0025	0.01326530612
26	0.0025	0.01428571429
27	0.0025	0.01530612245
28	0.0025	0.01632653061
29	0.0025	0.01734693878
30	0.0025	0.01836734694
31	0.0025	0.0193877551
32	0.0025	0.02040816327
33	0.0025	0.02142857143
34	0.0025	0.02244897959
35	0.0025	0.02346938776
36	0.0025	0.02448979592
37	0.0025	0.02551020408
38	0.0025	0.02653061224
39	0.0025	0.02755102041
40	0.0025	0.02857142857
41	0.0025	0.02959183673
42	0.0025	0.0306122449
43	0.0025	0.03163265306
44	0.0025	0.03265306122
45	0.0025	0.03367346939

46	0.0025	0.03469387755
47	0.0025	0.03571428571
48	0.0025	0.03673469388
49	0.0025	0.03775510204
13	0.0075	0.0387755102
15	0.0075	0.03979591837
2	0.0175	0.04081632653
12	0.0175	0.04183673469
14	0.0375 (0.03894)	0.04285714286
3	0.075	0.04387755102
8	0.15	0.04489795918
11	0.15	0.04591836735
1	0.25	0.04693877551
10	0.4	0.04795918367
7	0.75	0.04897959184
9	0.75	0.05

Appendix L: Ranked p-values from the P_{DB}-test for the first regime (DJIA)

Test id	Ranked p-value	BH
27	0	0.001020408163
47	0	0.002040816327
5	0.0025 (0.0037)	0.00306122449
6	0.0025 (0.0042)	0.004081632653
17	0.0025	0.005102040816
18	0.0025	0.00612244898
19	0.0025	0.007142857143
20	0.0025	0.008163265306
21	0.0025	0.009183673469
22	0.0025	0.01020408163
23	0.0025	0.0112244898
24	0.0025	0.01224489796
26	0.0025	0.01326530612
28	0.0025	0.01428571429
29	0.0025	0.01530612245
30	0.0025	0.01632653061
32	0.0025	0.01734693878
33	0.0025	0.01836734694
34	0.0025	0.0193877551
35	0.0025	0.02040816327
36	0.0025	0.02142857143
37	0.0025	0.02244897959
38	0.0025	0.02346938776
39	0.0025	0.02448979592
40	0.0025	0.02551020408
41	0.0025	0.02653061224
42	0.0025	0.02755102041
43	0.0025	0.02857142857
44	0.0025	0.02959183673
45	0.0025	0.0306122449
46	0.0025	0.03163265306
48	0.0025	0.03265306122

4	0.0075	0.03367346939
16	0.0075	0.03469387755
25	0.0075	0.03571428571
31	0.0075	0.03673469388
49	0.0075	0.03775510204
13	0.0175	0.0387755102
15	0.0175	0.03979591837
2	0.0375 (0.043)	0.04081632653
12	0.0375 (0.0293)	0.04183673469
14	0.075 (0.066)	0.04285714286
3	0.15	0.04387755102
1	0.25	0.04489795918
8	0.25	0.04591836735
11	0.25	0.04693877551
7	0.55	0.04795918367
10	0.55	0.04897959184
9	0.925	0.05

Appendix M: Empirical null distribution of L (lower α points)

	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.025$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.5$
$n = 2$	0.005342245	0.0100358	0.02500824	0.04987752	0.1027124	0.2115563	0.3267497	0.5746217
$n = 3$	0.141976053	0.2043901	0.32870548	0.48700641	0.7124095	1.0631895	1.3687415	1.9227077
$n = 4$	0.523628684	0.6593564	0.95825341	1.23343330	1.6374296	2.2064713	2.6661593	3.4631339
$n = 5$	1.089314500	1.3636549	1.75278920	2.18488940	2.7345944	3.4839738	4.0772423	5.0844331
$n = 6$	1.833420466	2.1610725	2.68873919	3.21334757	3.9481603	4.8636473	5.5578203	6.7255033
$n = 7$	2.685081315	3.0845979	3.72252674	4.40535801	5.2161091	6.3045719	7.1193572	8.4261628
$n = 8$	3.597272957	4.1073490	4.85766811	5.64304801	6.5749604	7.7929018	8.6952330	10.1707963
$n = 9$	4.562319611	5.1753124	6.03681592	6.94001869	7.9843505	9.3104161	10.3089203	11.9140336
$n = 10$	5.569752580	6.2352027	7.29071030	8.24023853	9.4129051	10.8680716	11.9528652	13.6964902
$n = 11$	6.758062790	7.4011449	8.57794098	9.60997610	10.8639633	12.4497371	13.6447642	15.4632712
$n = 12$	7.863731364	8.6572326	9.82588267	11.02532093	12.3185485	14.0180799	15.2651146	17.2590039
$n = 13$	9.084100158	9.9462277	11.23756206	12.39891836	13.8058329	15.6458095	16.9791641	19.0476480
$n = 14$	10.297832843	11.1201283	12.56335636	13.83385703	15.3873649	17.2946663	18.6637605	20.8725935
$n = 15$	11.474710051	12.3957852	13.91514005	15.34561529	16.8760636	18.9363083	20.3408053	22.6836592
$n = 16$	12.756572238	13.7717425	15.39553392	16.76678463	18.4950954	20.6067213	22.0988016	24.5158470
$n = 17$	14.055966586	15.1360480	16.78949470	18.30423942	20.0426151	22.2563400	23.8511865	26.3310460
$n = 18$	15.474528100	16.4287413	18.25880569	19.75092270	21.6493279	23.9451092	25.5884176	28.1751235
$n = 19$	16.767443658	17.7781953	19.68708759	21.35217412	23.2612252	25.6592464	27.3669993	30.0104381
$n = 20$	17.958095609	19.3079792	21.19729880	22.91769329	24.9505653	27.3343553	29.0955890	31.8859533
$n = 21$	19.483600017	20.6629021	22.64242750	24.38606345	26.5151265	29.0672264	30.8741481	33.6749618
$n = 22$	20.850332860	22.0349841	24.16042419	25.99914869	28.1056636	30.7172403	32.5778501	35.5446394
$n = 23$	22.133302263	23.5977894	25.69931418	27.56765639	29.7466516	32.4673600	34.3972715	37.3906334
$n = 24$	23.799195241	25.0380515	27.12134034	29.18794761	31.4021429	34.1687555	36.1953612	39.2296275
$n = 25$	25.072788709	26.5553100	28.74744747	30.75630652	33.1050623	35.9098239	37.9579031	41.1587186
$n = 26$	26.462182056	28.0280348	30.27839304	32.31292615	34.7378258	37.6564874	39.7640703	43.0016917
$n = 27$	27.849363133	29.5520078	31.82855645	34.05971027	36.4527245	39.4222599	41.5292967	44.9063343
$n = 28$	29.331437496	31.0008064	33.47346273	35.54630750	38.1141159	41.1981746	43.3280709	46.7345636
$n = 29$	30.730734495	32.5269221	34.96803181	37.19390252	39.8755715	42.9091039	45.1382074	48.6215126
$n = 30$	32.234975312	34.0700712	36.65181957	38.80503022	41.4772691	44.7519491	46.9857719	50.4922630
$n = 31$	33.708172251	35.5534691	38.09313729	40.40079118	43.1734072	46.4234445	48.7085168	52.3523932
$n = 32$	35.244672209	37.1658219	39.75441330	42.10678584	44.8997852	48.2893707	50.5517619	54.2174933
$n = 33$	36.777018563	38.6888885	41.27131732	43.69313106	46.5638311	50.0343661	52.3536855	56.0865580
$n = 34$	38.348183185	39.9491583	43.04393930	45.43268063	48.2928330	51.7815589	54.2071997	58.0209004
$n = 35$	39.816324554	41.6144949	44.52772657	47.17357197	50.0067832	53.5222504	56.0135859	59.8968109
$n = 36$	41.276049279	43.1624638	46.23088375	48.73887878	51.6739904	55.3969858	57.8751186	61.7508283
$n = 37$	42.901115607	44.8033417	47.73729772	50.47316040	53.4451702	57.1285176	59.7095016	63.7104367
$n = 38$	44.281003491	46.3191629	49.29791711	52.16321175	55.1457748	58.8786033	61.5165750	65.5753271
$n = 39$	46.053415349	48.0244277	51.17484245	53.80053554	57.0112770	60.7412850	63.3011303	67.4990098
$n = 40$	47.436455752	49.3832757	52.66801079	55.54662587	58.6571617	62.4825106	65.1849778	69.3442466
$n = 41$	48.746692106	51.0554295	54.46792142	57.22123744	60.4575842	64.2642134	67.0079775	71.1846340
$n = 42$	50.815360681	52.8000542	56.25597833	58.83610602	62.0870262	66.0697314	68.8301803	73.2086183
$n = 43$	52.312225668	54.3818203	57.79002117	60.48881068	63.9377521	67.8471136	70.6566811	75.0359549
$n = 44$	54.007011034	55.9265049	59.38657809	62.18832222	65.5763192	69.6779034	72.5537893	76.9233017
$n = 45$	55.273209340	57.7055638	60.89177975	63.93025185	67.3733428	71.4653683	74.3533448	78.8536660
$n = 46$	56.981149969	59.2709983	62.56791970	65.55567786	69.0319233	73.2840537	76.2085185	80.6849747
$n = 47$	58.518336365	60.6717924	64.48207364	67.28052322	70.8953967	75.1069846	78.0167268	82.6427556
$n = 48$	59.899163307	62.3672742	65.99213571	68.99497944	72.5457160	76.8882785	79.8979782	84.4790092
$n = 49$	61.543289290	63.9134021	67.57583430	70.75218472	74.4049951	78.7221092	81.7086905	86.4243621
$n = 50$	63.209286066	65.6515265	69.41948116	72.57376112	76.2788323	80.5359501	83.5870453	88.3294849

Appendix N: Empirical null distribution of R (upper α points)

	$\alpha = 0.5$	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
$n = 2$	2.778707	3.795736	4.607012	5.985454	7.361888	8.721125	10.58674	12.05585
$n = 3$	5.326644	6.717709	7.775147	9.452466	11.159827	12.729777	14.83075	16.43215
$n = 4$	7.754247	9.425155	10.647987	12.595892	14.439616	16.231088	18.52799	20.22054
$n = 5$	10.121040	12.011445	13.345203	15.486466	17.558034	19.500944	21.97255	23.45275
$n = 6$	12.452044	14.505021	16.001035	18.331968	20.500876	22.414891	25.22073	26.98483
$n = 7$	14.740375	16.958165	18.528387	21.021563	23.302349	25.389576	28.31222	30.14186
$n = 8$	17.012620	19.326650	21.044965	23.706163	26.051486	28.421446	31.28769	33.45931
$n = 9$	19.270595	21.738646	23.510008	26.342530	28.804060	31.152186	34.42222	36.47165
$n = 10$	21.484098	24.122356	25.896832	28.834812	31.592201	33.994215	37.21263	39.38467
$n = 11$	23.708728	26.494918	28.382059	31.397587	34.132858	36.769686	40.20616	42.19131
$n = 12$	25.887247	28.750796	30.790455	33.896401	36.795509	39.465996	42.89707	45.15359
$n = 13$	28.083132	31.082163	33.114634	36.444391	39.348538	42.124783	45.60122	48.17036
$n = 14$	30.280477	33.362499	35.518762	38.828410	41.875601	44.597173	48.30348	50.67796
$n = 15$	32.446608	35.701476	37.839896	41.418493	44.389070	47.400607	50.98134	53.82622
$n = 16$	34.637844	37.931809	40.268665	43.753318	47.063167	50.041080	53.58107	56.23565
$n = 17$	36.801737	40.217390	42.536455	46.235617	49.436647	52.490561	56.24384	59.11452
$n = 18$	38.976774	42.479856	44.808442	48.576310	51.892524	54.948014	58.92686	62.06399
$n = 19$	41.102505	44.711508	47.132952	51.059574	54.372164	57.739784	61.51741	64.50460
$n = 20$	43.263955	46.979450	49.542937	53.379899	56.927102	60.080653	64.23637	67.04411
$n = 21$	45.419612	49.150641	51.718119	55.740282	59.280871	62.759605	66.90614	69.33512
$n = 22$	47.594243	51.467368	54.124702	58.154416	61.779243	65.057663	68.84567	72.35462
$n = 23$	49.751032	53.612520	56.402079	60.458619	64.194240	67.630423	71.78741	75.00798
$n = 24$	51.857216	55.873525	58.634805	62.707431	66.557426	70.074571	74.50173	77.53545
$n = 25$	54.031759	58.068203	60.869541	65.115807	69.012325	72.516204	77.12052	80.30358
$n = 26$	56.094905	60.253118	63.108013	67.386863	71.445163	75.060820	79.29284	83.07089
$n = 27$	58.262371	62.500493	65.412009	69.762066	73.889342	77.484116	82.01144	85.26922
$n = 28$	60.423879	64.717988	67.689205	72.139176	76.225500	79.863829	84.49671	87.57447
$n = 29$	62.503099	66.818176	69.916316	74.427050	78.533650	82.458443	86.92732	90.18114
$n = 30$	64.550165	69.052519	72.150626	76.855642	80.911925	84.773578	89.38088	93.00390
$n = 31$	66.749111	71.290654	74.318337	79.265213	83.359979	87.225856	92.05437	95.27042
$n = 32$	68.918406	73.451908	76.587054	81.423640	85.585271	89.627545	94.24308	97.62325
$n = 33$	70.948355	75.660914	78.844894	83.649291	88.054484	92.025624	96.88809	100.25143
$n = 34$	73.070413	77.846357	81.062688	85.911091	90.352141	94.409356	98.92714	102.51973
$n = 35$	75.271250	80.068881	83.331552	88.367438	92.598172	96.856468	101.68246	105.12280
$n = 36$	77.350051	82.139683	85.544096	90.534519	95.051308	99.288474	103.95091	107.47745
$n = 37$	79.477195	84.354283	87.703481	92.721227	97.394176	101.685745	106.40725	110.86389
$n = 38$	81.566715	86.518806	89.932478	95.026760	99.687849	103.778782	109.19197	112.13101
$n = 39$	83.635003	88.770515	92.301753	97.330339	101.931128	106.211497	111.21844	115.04304
$n = 40$	85.783429	90.899807	94.361173	99.598144	104.256093	108.601252	113.94892	117.49656
$n = 41$	87.857371	93.080034	96.489382	101.901928	106.654949	110.804001	116.51365	120.46131
$n = 42$	89.984960	95.210355	98.761611	104.134740	108.873303	113.410375	118.81017	122.42251
$n = 43$	92.088890	97.301804	100.959648	106.251957	111.243619	115.700721	120.81188	124.97037
$n = 44$	94.266732	99.544029	103.123737	108.666148	113.546470	117.999050	123.76389	127.33136
$n = 45$	96.291074	101.661930	105.323961	110.926857	115.920390	120.596727	125.71277	129.86508
$n = 46$	98.399294	103.845852	107.554974	113.013258	118.234741	122.675341	128.61859	132.32189
$n = 47$	100.496617	106.047411	109.788978	115.304991	120.414385	125.206301	130.93504	134.59212
$n = 48$	102.570952	108.078494	111.928884	117.638503	122.668839	127.511232	133.17462	137.11794
$n = 49$	104.658983	110.315043	114.018283	119.829051	124.987648	129.678749	135.45010	139.74275
$n = 50$	106.774750	112.414306	116.371446	122.186279	127.529187	131.922635	137.87874	141.666

Appendix O: Empirical null distribution of P (lower α points)

	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.025$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$
$n = 2$	0.004799	0.009797	0.028100	0.051895	0.096900	0.190960	0.298640	0.394660
$n = 3$	0.004500	0.007299	0.018800	0.034200	0.065680	0.139600	0.212570	0.291000
$n = 4$	0.003600	0.006999	0.016198	0.034200	0.067990	0.137180	0.212110	0.282600
$n = 5$	0.003400	0.006300	0.015598	0.030300	0.062400	0.131300	0.206900	0.280660
$n = 6$	0.002700	0.006099	0.015400	0.030100	0.063300	0.135600	0.206800	0.282500
$n = 7$	0.002900	0.005698	0.014900	0.028800	0.058690	0.125900	0.195000	0.270100
$n = 8$	0.001900	0.005600	0.014200	0.028900	0.058400	0.128380	0.196900	0.263700
$n = 9$	0.004100	0.007000	0.014600	0.030095	0.060900	0.127480	0.190700	0.261000
$n = 10$	0.002800	0.005200	0.014700	0.028600	0.059000	0.124680	0.194800	0.265660
$n = 11$	0.002800	0.005399	0.013700	0.028390	0.059290	0.127800	0.195870	0.271260
$n = 12$	0.003300	0.006400	0.015800	0.030395	0.062990	0.127560	0.197700	0.273300
$n = 13$	0.002100	0.005499	0.013300	0.028700	0.060500	0.127200	0.191970	0.261660
$n = 14$	0.002000	0.004500	0.014700	0.029200	0.062400	0.127780	0.197600	0.266100
$n = 15$	0.002500	0.005499	0.016598	0.030800	0.064090	0.126800	0.192980	0.262100
$n = 16$	0.001900	0.005200	0.014598	0.030700	0.059700	0.125080	0.195340	0.266260
$n = 17$	0.002700	0.005998	0.014995	0.030395	0.059380	0.121980	0.190000	0.267620
$n = 18$	0.002400	0.005199	0.013998	0.027195	0.058900	0.126180	0.195900	0.265820
$n = 19$	0.003100	0.006300	0.015095	0.029400	0.064490	0.126180	0.194770	0.259500
$n = 20$	0.003100	0.006400	0.015198	0.027700	0.060890	0.129960	0.197900	0.265400
$n = 21$	0.002300	0.005199	0.014400	0.027195	0.057890	0.123800	0.195670	0.269560
$n = 22$	0.003400	0.006099	0.014895	0.031400	0.065290	0.128480	0.193700	0.264600
$n = 23$	0.003300	0.005597	0.015598	0.030300	0.059200	0.124900	0.186850	0.261400
$n = 24$	0.003000	0.005499	0.013700	0.030190	0.058700	0.121400	0.186270	0.256320
$n = 25$	0.002800	0.005100	0.012300	0.027990	0.058700	0.126000	0.200640	0.269400
$n = 26$	0.003400	0.005699	0.014498	0.026895	0.057800	0.121800	0.189400	0.262600
$n = 27$	0.003200	0.005199	0.013495	0.027900	0.057900	0.124480	0.189540	0.261560
$n = 28$	0.003100	0.006699	0.013798	0.029595	0.060400	0.123780	0.189800	0.262560
$n = 29$	0.002300	0.005200	0.015000	0.027300	0.056290	0.120780	0.186000	0.252860
$n = 30$	0.002900	0.005699	0.013100	0.028695	0.054990	0.125880	0.191670	0.260200
$n = 31$	0.002700	0.005398	0.012700	0.028795	0.059190	0.129480	0.193370	0.265800
$n = 32$	0.003200	0.006199	0.014095	0.026300	0.053300	0.122000	0.190470	0.262660
$n = 33$	0.003000	0.005900	0.012200	0.027000	0.055800	0.120900	0.187070	0.257400
$n = 34$	0.002499	0.005098	0.013400	0.026900	0.054580	0.120600	0.190500	0.265120
$n = 35$	0.002500	0.005197	0.016998	0.032995	0.060500	0.124380	0.195400	0.270060
$n = 36$	0.001700	0.003499	0.010798	0.024900	0.053500	0.119000	0.189240	0.258900
$n = 37$	0.002700	0.005199	0.013098	0.025895	0.053480	0.117380	0.182670	0.252160
$n = 38$	0.002200	0.004600	0.013898	0.031195	0.057690	0.123360	0.190140	0.256200
$n = 39$	0.003300	0.005399	0.012900	0.027900	0.061190	0.123080	0.189800	0.261800
$n = 40$	0.002500	0.004499	0.012698	0.028295	0.059180	0.119940	0.189200	0.261760
$n = 41$	0.001700	0.004400	0.011100	0.025700	0.055600	0.118980	0.188400	0.264500
$n = 42$	0.003000	0.006894	0.015700	0.030800	0.060290	0.122500	0.190100	0.260200
$n = 43$	0.003500	0.005700	0.014098	0.028100	0.057990	0.123000	0.195000	0.262060
$n = 44$	0.002100	0.004699	0.013200	0.028600	0.060290	0.122400	0.190100	0.262500
$n = 45$	0.003300	0.006300	0.015295	0.027495	0.057100	0.122160	0.189000	0.258700
$n = 46$	0.002000	0.004499	0.012898	0.028195	0.055700	0.115880	0.185000	0.262460
$n = 47$	0.002500	0.005099	0.013498	0.029400	0.060400	0.125800	0.196700	0.266100
$n = 48$	0.002900	0.006300	0.013900	0.027600	0.060190	0.127500	0.197140	0.263100
$n = 49$	0.002400	0.004900	0.014298	0.027500	0.057600	0.121500	0.190940	0.259400
$n = 50$	0.001900	0.004598	0.014300	0.029095	0.061190	0.130380	0.194640	0.257100

	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 0.95$	$\alpha = 0.975$	$\alpha = 0.99$	$\alpha = 0.995$
$n = 2$	0.511000	0.605280	0.709120	0.804120	0.902700	0.955505	0.977200	0.992602	0.995701
$n = 3$	0.384250	0.464720	0.550460	0.649900	0.753300	0.809400	0.842300	0.868101	0.882004
$n = 4$	0.357250	0.437400	0.529130	0.620100	0.722320	0.785400	0.821500	0.849601	0.859301
$n = 5$	0.349250	0.429840	0.510800	0.604900	0.697820	0.758600	0.798702	0.832702	0.848100
$n = 6$	0.350050	0.431640	0.515300	0.603520	0.700510	0.758505	0.797800	0.832100	0.846902
$n = 7$	0.332600	0.410200	0.499230	0.593240	0.693110	0.755200	0.793400	0.821801	0.839801
$n = 8$	0.355100	0.433400	0.519600	0.607740	0.705520	0.762800	0.803402	0.832303	0.845700
$n = 9$	0.342750	0.425380	0.509400	0.598220	0.698210	0.761300	0.795800	0.826002	0.842201
$n = 10$	0.333500	0.413600	0.500450	0.586200	0.693310	0.758515	0.795010	0.824901	0.840701
$n = 11$	0.338400	0.415400	0.494760	0.588300	0.689100	0.751105	0.789202	0.821102	0.832602
$n = 12$	0.349400	0.430080	0.510600	0.591120	0.687320	0.747305	0.786700	0.819005	0.833700
$n = 13$	0.336750	0.420380	0.501130	0.588320	0.686100	0.749120	0.788400	0.824101	0.841000
$n = 14$	0.343850	0.425700	0.506600	0.589420	0.688300	0.746400	0.786300	0.814800	0.830802
$n = 15$	0.342200	0.423700	0.498630	0.584900	0.688030	0.752705	0.789402	0.819200	0.831601
$n = 16$	0.342950	0.421500	0.503900	0.599500	0.693300	0.753405	0.797000	0.824600	0.836501
$n = 17$	0.338000	0.411100	0.492100	0.579300	0.682310	0.742100	0.783000	0.818502	0.838400
$n = 18$	0.339500	0.418840	0.499230	0.588060	0.686110	0.746910	0.785500	0.808304	0.824302
$n = 19$	0.340350	0.419280	0.502400	0.587800	0.693100	0.753500	0.790108	0.823410	0.836302
$n = 20$	0.328900	0.404400	0.488500	0.579420	0.677400	0.740920	0.779100	0.810701	0.826201
$n = 21$	0.338500	0.414440	0.495220	0.586400	0.688710	0.753305	0.791812	0.821203	0.833907
$n = 22$	0.348350	0.426340	0.506530	0.590900	0.694810	0.753805	0.792903	0.822701	0.838801
$n = 23$	0.332550	0.407920	0.493630	0.580500	0.680120	0.740910	0.780307	0.814301	0.827200
$n = 24$	0.340600	0.414920	0.497500	0.584100	0.679410	0.743000	0.780502	0.813703	0.831800
$n = 25$	0.333450	0.409700	0.493900	0.588220	0.691710	0.749405	0.789000	0.821307	0.837802
$n = 26$	0.337350	0.417940	0.501100	0.581200	0.680520	0.744020	0.778502	0.812000	0.832601
$n = 27$	0.336900	0.411280	0.495500	0.581600	0.680600	0.742205	0.778305	0.809106	0.827602
$n = 28$	0.338100	0.416140	0.494300	0.588500	0.683600	0.741400	0.778000	0.810900	0.825504
$n = 29$	0.331250	0.405700	0.495830	0.584700	0.684900	0.752705	0.789407	0.823305	0.836000
$n = 30$	0.344450	0.418880	0.504060	0.591840	0.692900	0.751400	0.784012	0.816501	0.828904
$n = 31$	0.330450	0.407580	0.489560	0.580200	0.681510	0.744510	0.784705	0.814303	0.829400
$n = 32$	0.333700	0.410900	0.490300	0.583740	0.679310	0.747215	0.786205	0.820000	0.835300
$n = 33$	0.331600	0.411900	0.498550	0.583460	0.684100	0.744620	0.787302	0.817500	0.834500
$n = 34$	0.337550	0.411640	0.494530	0.579100	0.680310	0.745600	0.787100	0.817802	0.832302
$n = 35$	0.342750	0.420540	0.501600	0.586820	0.686610	0.746450	0.786902	0.820802	0.837100
$n = 36$	0.331050	0.408900	0.494990	0.582800	0.685900	0.745905	0.783808	0.812310	0.827000
$n = 37$	0.328600	0.408100	0.491500	0.576220	0.678230	0.740705	0.779915	0.816900	0.832601
$n = 38$	0.339700	0.420040	0.496490	0.585400	0.684800	0.743600	0.780617	0.812211	0.826003
$n = 39$	0.334650	0.410500	0.490190	0.581800	0.682610	0.746100	0.786805	0.814807	0.832403
$n = 40$	0.335000	0.411040	0.491800	0.581240	0.682400	0.745205	0.784302	0.812400	0.827105
$n = 41$	0.338300	0.412200	0.489730	0.582520	0.679800	0.744800	0.782302	0.811507	0.823800
$n = 42$	0.333000	0.410240	0.491400	0.582300	0.678400	0.740005	0.780810	0.811103	0.828801
$n = 43$	0.333000	0.411700	0.487530	0.576500	0.678430	0.744500	0.784105	0.816500	0.829702
$n = 44$	0.338600	0.415400	0.504600	0.596720	0.689710	0.747200	0.785405	0.817700	0.830700
$n = 45$	0.333000	0.414400	0.497860	0.581700	0.683200	0.750700	0.784615	0.816203	0.829800
$n = 46$	0.332800	0.409500	0.490860	0.582100	0.678210	0.740220	0.782502	0.815705	0.832702
$n = 47$	0.336800	0.411040	0.490000	0.586800	0.685630	0.741325	0.778902	0.810403	0.829303
$n = 48$	0.333800	0.414200	0.493630	0.578100	0.681720	0.746410	0.784100	0.816201	0.830400
$n = 49$	0.341300	0.418040	0.495900	0.580800	0.680300	0.748300	0.785412	0.818206	0.832400
$n = 50$	0.331550	0.408040	0.488960	0.579040	0.684910	0.744110	0.786402	0.818003	0.832303

Bibliography

[1] Alzaatreh, A., Lee, C., and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron* 71:63-79.

[2] Aminikhanghahi, S., and Cook, D.J. (2017). A survey of methods for time series change point detection. *Knowl Inf Syst*, 51(2): 339-367.

[3] Antoch, J., and Jaruskova, D. (2007). Testing a homogeneity of stochastic processes. *Kybernetika*, 43(4): 415-430.

[4] Andersen, P.K., Borgan, O., Gill, R.D., and Keiding, N. (1993). *Statistical Models based on Counting Processes*, Springer Series in Statistics, Springer-Verlag, New York.

[5] Auret, L., and Aldrich, C. (2010). Change point detection in time series data with random forests. *Control Engineering Practice*, 18 (8):990-1002.

[6] Bain, L.J., and Engelhardt, M. (1980). Inferences on the parameters and current system reliability for a time truncated Weibull process. *Technometrics*, 22(3):421 - 426.

[7] Bain, L.J., and Engelhardt, M. (1980). Inferences on the parameters and current system reliability for a time truncated Weibull process. *Technometrics*, 22(3):421 - 426.

[8] Bain, L. J., Engelhardt, M., and Wright, F. T. (1985). Tests for an increasing trend in the intensity of a Poisson process: A power study. *Journal of the American Statistical Association* 80:419-422.

[9] Engelhardt, M., and Bain, L.J. (1987). Statistical analysis of a Compound Power-Law model for repairable systems. *IEEE Transactions on Reliability*, 36 (4):392-396.

[10] Bain, L. J., and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life-Testing Models Theory and Methods*, (2nd ed.), New York: Marcel Dekker.

[11] Bain, L. J., and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life Testing models, Theory and Methods*. Second Edition, Marcel Dekker, New York.

[12] Bakun, W.H., Aagaard, B., Dost, B., Ellsworth, W.L., Hardebeck, J.L., Harris, R.A., Ji, C., Johnston, M.J.S., Langbein, J., Lienkaemper, J.J., Michael, A.J., Murray, J.R., Nadeau, R.M., Reasenberg, P.A., Reichle, M.S., Roeloffs, E.A., Shakal, A., Simpson, R.W., and Waldhauser, F. (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature* 437:969974.

[13] Bassin, W. M. (1969). Increasing hazard functions and overhaul policy. *Proceedings of the 1969 Annual Symposium on Reliability*, Chicago: IEEE 173-178.

[14] Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A practical and poowerful approach to multiple testing. *Journal of Royal Statistical Society, Series B*, 57(1):289-300.

[15] Bhaduri, M., Justin Zhan, J., and Chiu, C. (2017). A novel weak estimator for dynamic systems. *IEEE Access* 5:27354-27365.

[16] Bhat, U.N. (1984). *Elements of Applied Stochastic Processes*, Second Edition, Wiley-Interscience, John Wiley and Sons.

[17] Box, G.E.P., and Jenkins, G.M. (1976). *Time series analysis: forecasting and control*. Holden-Day, San Francisco.

[18] Brillinger, D.R. (1994). Time series, point processes, and hybrids. *The Canadian Journal of Statistics*, 22(2):177-206.

[19] Brodsky, B. (2017). *Change-point analysis in nonstationary stochastic models*. CRP Press.

[20] Buishand, T.A. (1982). Some Methods for Testing the Homogeneity of Rainfall Records. *Journal of Hydrology*, 58, 1127.

[21] Cameron, A.C., and Trivedi, P.K. (1998). *Regression Analysis of Count Data*, Cambridge University Press, USA.

[22] Carlstein, E. (1986). The use of subseries methods for estimating the variance of a general statistic from a stationary time series. *Ann Stat* 14:11711179.

[23] Casella, G., and Berger, R.L. (2002). *Statistical Inference*, 2nd Edition, Cencage Learning.

[24] Chen, Q (2010). Poisson Process Monitoring, Test and Comparison. *UNLV Theses, Dissertations, Professional Papers, and Capstones. 1059.*

[25] Chen, J., and Gupta, A. K. (2011). *Parametric statistical change point analysis: With applications to Genetics, Medicine and Finance*. 2nd ed. Birkhauser.

[26] Chib, S. (1998). Estimation and Comparison of Multiple Change-Point Models. *Journal of Econometrics*, 86(2):221241.

[27] Conover, W. J. (1999). *Practical Nonparametric Statistics*. New York, NY: Wiley.

[28] Cook, D.J., and Krishnan, N.C. (2015). *Activity learning: discovering, recognizing, and predicting human behavior from sensor data*. Wiley.

[29] Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society Ser. B*, 17:129-164.

[30] Crooks, G.E. (2015). The Amoroso distribution. *arXiv:1005.3274 [math.ST].*

[31] Crow, L.H. (1974). Reliability analysis for complex repairable systems. *Reliability and Biometry*, SIAM, 379-410.

[32] Crow, L. H. (1974). Reliability analysis for complex repairable systems, reliability and biometry. *eds. F. Proschan and R. J. Serfling, Philadelphia:* SIAM:379-410.

[33] Crow, L.H. (1982). Confidence interval procedures for the Weibull process with applications to reliability growth. *Technometrics*, 24(1):67-72.

[34] Crow, L. H. (1982). Confidence interval procedures for the Weibull process with applications to reliability growth. *Technometrics* 24:67-72.

[35] Dimitrakopoulou, T., Adamidis, K., and Loukas, S. (2007). A Lifetime Distribution With an Upside-Down Bathtub-Shaped Hazard Function, *IEEE Transactions on Reliability*, 56 (2):308-311.

[36] Duane, J.T. (1964). Learning curve approach to reliability monitoring. *IEEE Transactions on Aerospace*, 2:563-566.

[37] Duran, B. (1976). Survey of Nonparametric Tests for Scale, *Communications in StatisticsTheory and Methods*, 5:12871312.

[38] Efron, B., and Zhang, N.R. (2011). False Discovery Rates and Copy Number Variation. *Biometrika*, 98(2):251271.

[39] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann Stat* 7:126.

[40] Emanuel, K. (2003). Tropical cyclones. *Annual Review of Earth and Planetary Sciences*, 31:75-104.

[41] Emanuel, K. (2006). Hurricanes: Tempests in a greenhouse. *Physics Today*, 59:74-75.

[42] Emanuel, K. (2007). Environmental factors affecting tropical cyclone power dissipation. *Journal of Climate*, 20, 54975509.

[43] Engelhardt, M., Guffey, J. M., and Wright, F. T. (1990). Tests for positive jumps in the intensity of a Poisson process: A power study. *IEEE Transactions on Reliability* 39:356-360.

[44] Ephraim, Y., and Merhav, N. (2002). Hidden Markov Processes *IEEE Transactions on Information Theory*, 48(6):1518 - 1569.

[45] Erdman, C., and Emerson, J.W. (2007). bcp: An R Package for Performing a Bayesian Analysis of Change Point Problems. *Journal of Statistical Software*, 23(3), 113.

[46] Fearnhead, P., and Liu, Z, (2007). On-line Inference for Multiple Changepoint Problems. *Journal of the Royal Statistical Society B*, 69(4):589605.

[47] Feller, W. (2005). *An Introduction to Probability Theory and its Applications*, Second Edition, John Wiley and Sons.

[48] Finkelstein, J.M. (1976). Confidence bounds on the parameters of the Weibull process. *Technometrics*, 18(1):115-117.

[49] Finkelstein, J. M. (1976). Confidence bounds on the parameters of the Weibull process. *Technometrics*, 18(1):115-117.

[50] Gamiz, M.L., Kulasekara, K.B., Limnios, N., and Lindqvist, B.H. (2011). *Applied Nonparametric Statistics in Reliability*, Springer Series in Reliability Engineering, Springer-Verlag, London.

[51] Gertsbakh, I. (2005). *Reliability Theory With Applications to Preventive Maintenance*, Springer-Verlag.

[52] Gombay, E. (2000). Sequential Change-point Detection with Likelihood Ratios. *Statistics and Probability Letters* 49:195204.

[53] Gustafsson, F. (2000). *Adaptive Filtering and Change Detection*. John Wiley and Sons.

[54] Hall, P. (1985). Resampling a coverage pattern. *Stoch Process Appl* 20:231246.

[55] Hawkins, D.M. (1977). Testing a Sequence of Observations for a Shift in Location. *Journal of the American Statistical Association*, 72(357):180186.

[56] Hawkins, D. M., and Olwell, D. H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*. Springer Verlag, New York, NY.

[57] Hawkins, D.M., Qiu, P.H., and Kang, C.W. (2003). The Changepoint Model for Statistical Process Control. *Journal of Quality Technology*, 35(4):355366.

[58] Hawkins, D.M., and Zamba, K.D. (2005a). A Change-Point Model for a Shift in Variance. *Journal of Quality Technology*, 37(1):2131.

[59] Hawkins, D.M., and Deng, Q. (2010). A Nonparametric Change-Point Control Chart. *Journal of Quality Technology*, 42(2):165173.

[60] Hawkins, D.M. (2001). Fitting Multiple Change-Point Models to Data. *Computational Statistics and Data Analysis*, 37(3):323341.

[61] Henschel, K., Hellwig, B., Amtage, F., Vesper, J., Jachan, M., Lucking, C.H., Timmer, J., and Schelter, B. (2008). Multivariate analysis of dynamical processes. *The European Physical Journal Special Topics*, 165:25-34.

[62] Hesterberg, T.C. (2015). What teachers should know about the bootstrap: resampling in the undergraduate statistics curriculum. *Am Stat*, 69(4):371386.

[63] Hilbe, J.M. (2014). *Modeling Count Data*, Cambridge University Press.

[64] Hinkley, D.V., Hinkley, E.A. (1970). Inference about Change-Point in a Sequence of Binomial Variables. *Biometrika*, 57(3):477488.

[65] Ho, C.-H. (1992). Statistical Control Chart for Regime Identification in Volcanic Time Series. *Mathematical Geology*, 24:775-787.

[66] Ho, C.-H. (1993). Forward and backward tests for an abrupt change in the intensity of a Poisson process. *Journal of Statistical Computation and Simulation*, 48:245-252.

[67] Ho, C.-H. (2008). Empirical recurrence rate time series for volcanism: application to Avachinsky volcano, Russia. *J Volcanol Geotherm Res* 173:1525.

[68] Ho, C.-H., and Bhaduri, M. (2015). On a novel approach to forecast sparse rare events : applications to Parkfield earthquake prediction. *Natural Hazards*, 78(1):669-679.

[69] Ho, C.-H., Zhong, G., Cui, F., and Bhaduri, M. (2016). Modeling interaction between bank failure and size. *Journal of Finance and Bank Management*, 4(1), 15-33.

[70] Ho, C.-H., and Bhaduri, M. (2017). A Quantitative Insight into the Dependence Dynamics of the Kilauea and Mauna Loa Volcanoes, Hawaii. *Mathematical Geosciences*, 49: 893-911.

[71] Hollander, M., Wolfe, D.A. (1999). *Nonparametric Statistical Methods (2nd edition)*, Wiley Intescience.

[72] Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2): 65-70.

[73] Inclan, C., and Tiao, G.C. (1994). Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. *Journal of the American Statistical Association*, 89(427): 913923.

[74] Jacobsen, M. (2006) *Point Process Thoery and Applications*, Probability and its Applications, Birkhauser.

[75] James, N.A., and Matteson, S.D. (2014). ecp: An R Package for Nonparametric Multiple Change Point Analysis of Multivariate Data. *Journal of Statistical Software*, 62(7): 1-25.

[76] Jelinski, Z., and Moranda, P. (1972). Software reliability research. *Statistical Computer Performance Evaluation*, 465484.

[77] Kalzanov, K. (1970) (In Russian). The closeness of the Palm and Poisson flows. *Izv. Akad. Nauk UzSSR Sere Fiz-Mat. Nauk* 14 (3): 69-71.

[78] Karr, A.F. *Point Processes and their Statistical Inference*, Probability: Pure and Applied, Dekker, New York, 1986.

[79] Killick, R., and Eckley, I.A. (2014). changepoint: An R Package for Changepoint Analysis, *Journal of Statistical Software* 58 (3).

[80] Killick, R., and Eckley, I.A. (2014). changepoint: An R Package for Changepoint Analysis, *Journal of Statistical Software* 58 (3).

[81] Klein, F.W. (1982). Patterns of historical eruptions at Hawaiian volcanoes. *Journal of Volcanology and Geothermal Research*, 12:1- 35.

[82] Knill, O. (2009). *Probability Theory and Stochastic Processes with Applications*, Overseas Press.

[83] Kovalenko, I.N., Kuznetsov, N.Y., and Pegg, P.A. (1997) *Mathematical Theory of Reliability of Time Dependent Systems with Practical Applications*, John Wiley and Sons.

[84] Krishnamoorthy, K., and Thomson, J. (2004). A more powerful test for comparing two Poisson means. *Journal of Statistical Planning and Inference*, 119: 23-25.

[85] Kvaloy, J. T., Lindqvist, B.H., and Malmedal, H. (2001). A statistical test for monotonic and non-monotonic trend in repairable systems. *Conference Proceedings ESREL 2001, Torino, 16-20 September*, 1563-1570.

[86] Lai, T.L. (1995). Sequential Changepoint Detection in Quality Control and Dynamical Systems. *Journal of the Royal Statistical Society B*, 57(4): 613658.

[87] Lai, T. L. (2001). Sequential Analysis: Some Classical Problems and New Challenges. *Statistica Sinica* 11: 303350.

[88] Lee, L., and Lee, S.K. (1978). Some results on inferences for the Weibull process. *Technometrics*, 20 :41-45.

[89] Lee, L., and Lee, S. K. (1978). Some results on inferences for the Weibull process. *Technometrics*, 20, 41-45.

[90] Lehmann, E.H., and Romano, J.P. (2005). *Testing Statistical Hypotheses*, 3rd Edition, Springer.

[91] Lemonte, A.J., Barreto-Souza, W., Cordeiro, G.M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics* 27: 31-53.

[92] Lepage, Y. (1971), Combination of Wilcoxians and AnsariBradley Statistics. *Biometrika*, 58: 213217.

[93] Leroux B.G., and Puterman, M.L. (1992). Maximum-penalized likelihood estimation for independent and Markov-dependent mixture models. *Biometrics*, 48:545558.

[94] Li, A., and Barber, R.F. (2017) Accumulation Tests for FDR Control in Ordered Hypothesis Testing. *Journal of the American Statistical Association*, 112(518):837-849.

[95] Limnios, N., and Nikulin M. (eds) (2000) *Recent Advances in Reliability Theory, Methodology, Practice, and Inference*, Birkhauser, Boston.

[96] Lindsey, J.K (1995). *Modelling Frequency and Count Data*, Oxford Statistical Science Series, Oxford University Press.

[97] Lindsey, J.K. (2004). *Statistical Analysis of Stochastic Processes in Time*, Cambridge University Press.

[98] Lindqvist, B.H. (2006). On the Statistical Modeling and Analysis of Repairable Systems. *Statistical Science*, 21 (4):532551.

[99] Lipman, P.W. (1980). The southeast rift zone of Mauna Loa: implications for structural evolution of Hawaiian volcanoes. *Am J Sci 280-A*: 752-776

[100] Liu, S., Yamad, M., Collier, N., and Sugiyama, M. (2013). Change-point detection in time-series data by relative density-ratio estimation. *Neural Networks*, 43, 72-83.

[101] Lowen, S. B., and Teich, M. C. (2005). *Fractal-Based Point Processes*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons.

[102] Maguire, B.A., Pearson, E.S., and Wynn, A.H.A. (1952). The time intervals between industrial accidents. *Biometrika*, 39:168180.

[103] Matteson, D.S., and James, N.A. (2013). A Nonparametric Approach for Multiple Change Point Analysis of Multivariate Data. *ArXiv e-prints*. To appear in the *Journal of the American Statistical Association*, 1306.4933.

[104] Matthes, K., Kerstan, J., and Mecke, J. (1978) *Infinitely Divisible Point Processes*, John Wiley and Sons.

[105] Mead, M.E. (2015). Generalized Inverse Gamma Distribution and its Application in Reliability. *Communications in Statistics, Theory and Methods*, 44 (7):1426-1435.

[106] Mood, A. (1954). On the Asymptotic Efficiency of Certain Nonparametric Two-Sample Tests. *Annals of Mathematical Statistics*, 25:514533.

[107] Mooley, D. A. (1981). Applicability of the Poisson Probability Model to the Severe Cyclonic Storms Striking the Coast around the Bay of Bengal. *Sankhya: The Indian Journal of Statistics, Series B (1960-2002)*, 43(2).

[108] Mulargia, F., Tinti, S., and Boschi, E. (1985). A Statistical Analysis of Flank Eruptions on Etna Volcano. *J. Volcano. Geotherm. Res.*,23:263272.

[109] Mulargia, F., Gasperini, P., and Tinti, S. (1987). Identifying Different Regimes in Eruptive Activity: An Application to Etna Volcano. *J. Volcano. Geotherm. Res.* 34: 89106.

[110] Nelson, B.L. (1995). *Stochastic Modeling: Analysis and Simulation*, Industrial Engineering Series, McGraw-Hill International Edition.

[111] Page, E.S. (1954). Continuous Inspection Schemes. *Biometrika*, 41(1/2): 100115.

[112] de Pascoa, M.A.R., Ortega, E.M.M., Cordeiro, G.M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, 8: 411433.

[113] Pettitt, A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Applied Statistics* 28(2):126135.

[114] Pettitt , A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Journal of the Royal Statistical Society C*, 28(2):126135.

[115] Pignatiello, J. J., Jr., and Samuel, T. R. (2001). Estimation of the Change Point of a Normal Process Mean in SPC Applications. *Journal of Quality Technology* 33:8295.

[116] Pohlert, T. (2018), Non-Parametric Trend Tests and Change-Point Detection. *CRAN repository.*

[117] Rand, W.M, (1971). Objective Criteria for the Evaluation of Clustering Methods. *Journal of the American Statistical Association*, 66:846 850.

[118] Reiss, R.-D. (1993). *A Course on Point Processes*, Springer Series in Statistics, Springer-Verlag, New York.

[119] Resnick, S. I. (2002). *Adventures in Stochastic Processes*, Birkhauser, Boston.

[120] Resnick, S.I. (2005). *A Probability Path*, Birkhauser, Boston.

[121] Rigas, A.G. (1996). Estimation of certain parameters of a stationary hybrid process involving a time series and a point process. *Mathematical Biosciences*, 133:197-218.

[122] Rigdon, S.E., and Basu, A.P. (2000). *Statistical Methods for the Reliability of Repairable Systems*, Wiley series in Probability and Statistics, John Wiley and Sons.

[123] Roberts, S.W. (1959). Control Chart Tests Based on Geometric Moving Averages. *Technometrics*, 42(1):97101.

[124] Ross, G.J. (2012). Modelling Financial Volatility in the Presence of Abrupt Changes. *Physica A: Statistical Mechanics and Its Applications*, 392(2):350360.

[125] Ross, G.J., Tasoulis, D.K., and Adams, N.M. (2011). Nonparametric Monitoring of Data Streams for Changes in Location and Scale. *Technometrics*, 53(4):379389.

[126] Ross, G.J., and Adams, N.M. (2011). Sequential Monitoring of a Bernoulli Sequence When the Pre-Change Parameter is Unknown. *Computational Statistics*, 28(2):463479.

[127] Ross, G.J., and Adams, N.M. (2012). Two Nonparametric Control Charts for Detecting Arbitrary Distribution Changes. *Journal of Quality Technology*, 44(12):102116.

[128] Ross, G.J. (2014). Sequential Change Detection in the Presence of Unknown Parameters. *Statistics and Computing*, 24(6):10171030.

[129] Ross, G.J.(2015). Parametric and Nonparametric Sequential Change Detection in R: The cpm Package. *Journal of Statistical Software*, 66 (3).

[130] Ross, S.M. (1990) *A Course in Simulation*, Macmillan Publishing Company, USA.

[131] Ross, S.M. (1996) *Stochastic Processes*, Second Edition, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons.

[132] Ross, S.M. (2010) *Introduction to Probability Models*, Tenth Edition, Academic Press, Elsevier.

[133] Shan, G. (2015). Exact unconditional testing procedures for comparing two independent Poisson rates, *Journal of Statistical Computation and Simulation*, 85(5):947-955.

[134] Shorack, G.R., and Wellner, J.A. (1986). *Empicial Processes with Applications to Statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons.

[135] Smethurst, L., James, M.R., Pinkerton, H., and Tawn, J.A. (2009). A statistical analysis of eruptive activity on Mount Etna, Sicily. *Geophys.J.Int.* 179:655-666.

[136] Snyder, D.L. (1975). *Random Point Processes*, Wiley Interscience, John Wiley and Sons.

[137] Stephens, D.A. (1994). Bayesian Retrospective Multiple-Changepoint Identification. *Journal of the Royal Statistical Society C*, 43(1):159178.

[138] Storey, J.D. (2003). The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics* 31(6):20132035.

[139] Stoyan, D., Kendall, W.S., and Mecke, J. (1987) *Stochastic Geometry and its Applications*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons.

[140] Szekely, G.J., and Rizzo, M.L. (2005). Hierarchical Clustering Via Joint Between-Within Distances: Extending Wards Minimum Variance Method. *Journal of Classification*, 22(2):151 183.

[141] Szekely, G.J., and Rizzo, M.L. (2010). Disco Analysis: A Nonparametric Extension of Analysis of Variance. *The Annals of Applied Statistics*, 4(2):10341055.

[142] Tan,S., Bhaduri,M., and Ho, C.-H. (2014). A statistical model for long term forecasts of strong dust sand storms, *Journal of Geoscience and Environment Protection*, 2: 16-26.

[143] Tartakovsky, A.G., Rozovskii, B.L., Blazek, R.B., and Kim. H. (2006). A Novel Approach to Detection of Intrusions in Computer Networks via Adaptive Sequential and Batch-Sequential ChangePoint Detection Methods. *IEEE Transactions on Signal Processing*, 54(9):33723382.

[144] Unser, M., and Tafti, P.D. (2014).*An Introduction to Sparse Stochastic Processes*, Cambridge University Press, UK.

[145] Wickman, F. E. (1966). Repose-Period Pattems of Volcanoes *Ark. Mineral. Geol.*, 4: 291-367.

[146] Worsley, K.J. (1982). An Improved Bonferroni Inequality and Applications. *Biometrika*, 69(2):297302.

[147] Zeileis, A., Leisch, F., Hornik, K., and Kleiber, C. (2002). strucchange: An R Package for Testing for Structural Change in Linear Regression Models. *Journal of Statistical Software*, 7(2):138.

[148] Zhou, C., Zou, C., Zhang, Y., and Wang, Z. (2009). Nonparametric Control Chart Based on ChangePoint Model. *Statistical Papers* 50(1):1328.

[149] Zou, C., and Tsung, F. (2010). Likelihood Ratio-Based Distribution-Free EWMA Control Charts. *Journal of Quality Technology*, 42(2):174196.

[150] Zucchini, W., and MacDonald, I.L. (2009) *Hidden Markov Models for time series*, CRC Press.

Curriculum Vitae

Graduate College
University of Nevada, Las Vegas

Moinak Bhaduri

Degrees:

Bachelor of Science - Statistics (Hons), 2010
St. Xavier's College, Kolkata, India

Special Honors and Awards:

1) Recipient of the *UNLV Outstanding Graduate Student Teaching Award (2nd place for the 2014-2015 academic year):* a $2,000 stipend awarded by the Graduate College, UNLV and the College of Sciences, UNLV for demonstrated excellence in teaching undergraduates.

2) Recipient of the *Wolzinger Family Research Scholarship (for Fall 2016 and Spring 2017)*: a $10,000 stipend awarded by the College of Sciences, UNLV for demonstrated excellence in mathematical research.

3) Recipient of the *Nevada Regents' Graduate Scholar Award (for 2018)*: a $5,000 stipend awarded by the Nevada Board of Regents, for excellence in academics, leadership, contributions to the university and the community, and demonstrated potential for continued success.

4) Recipient of the *GPSA Merit Award (for the 2015 academic year)*: a $300 stipend awarded by the Graduate and Professional Students' Association (GPSA), UNLV for demonstrated outstanding scholarship in the field of Statistics.

5) Recipient of the *Graduate Statistician* status: awarded by the Royal Statistical Society based in London, UK, in January 2016 in recognition of demonstrated statistical expertise.

6) Recipient of the *2015 Marcelo Cruz Internship Award* : worked as a quantitative analyst (intern) for risk management and financial modeling at Precsio Consulting during May - August 2015 in Central Arizona. Specific projects related to Operational Risk (with emphasis on Bayesian approach to loss distribution approximation) and CCAR Stress Testing (with emphasis on the development of ARIMAX models to forecast bank variables).

7) Recipient of *First place, Platform presentation (Science, Session B)* at the 18th Annual Graduate and Professional Students' Research Forum, UNLV on March 12, 2016: a $200 stipend awarded by the GPSA, UNLV.

8) Recipient of *University of Nevada, Las Vegas's GPSA funding and research grant* for Summer 2015 ($725) and Summer 2016 ($500) cycles.

9) Recipient of *Honorable mention, Platform presentation (Science and Health Science, Session B)* at the 19th Annual Graduate and Professional Students' Research Forum, UNLV on April 8, 2017.

10) Recipient of UNLV Graduate College's *Grad. Medallion*, Spring 2018.

11) Recipient of UNLV Department of Mathematical Science's *Graduate Teaching Assistantship*, Fall 2012 - Spring 2018.

12) Recipient of UNLV Department of Computer Science's *Graduate Research Assistantship*, Summer 2017.

13) Recipient of UNLV's *Top Tier Doctoral Graduate Research Assistantship* (Department of Computer Science), Spring 2018.

14) Recipient of UNLV's *Access Grant*, Fall 2013 - Spring 2018: a $2000 - $3000 yearly grant for continually excellent academic performance.

15) Awarded the *Visiting Student Research Fellowship* for the year 2012 by the Centre of Applicable Mathematics, Tata Institute of Fundamental Research, Bangalore, India: Worked at the TIFR-CAM during April - May 2012 under the supervision of Dr. Amit Apte and Dr. Sreekar V. on the principle of free energy minimization applied to nonlinear filtering.

16) *Elected member of Pi Mu Epsilon* (a national honorary mathematics society), Nevada Beta Chapter.

Peer-reviewed journal publications:

1) Moinak Bhaduri, Justin Zhan and Carter Chiu, "A novel weak estimator for dynamic systems", *IEEE Access, 2017, (5), 27354-27365.*

This work was supported in part by the United States Department of Defense under Grant W911NF-17-1-0088 and Grant W911NF-16-1-0416, in part by the AEOP/REAP Programs Support, in part by the National Science Foundation under Grant 1625677 and Grant 1710716, and in part by the United Healthcare Foundation under Grant UHF1592.

2) Moinak Bhaduri, Justin Zhan, Carter Chiu and Felix Zhan, "A novel online and non-parametric approach for drift detection in big data", *IEEE Access, 2017, (5), 15883-15892.*

This work was supported in part by the United States National Science Foundation under Grant 1625677.

3) Chih-Hsiang Ho and Moinak Bhaduri, "A quantitative insight into the dependence dynamics of the Kilauea and Mauna Loa volcanoes, Hawaii", *Mathematical Geosciences, 2017, 49(7), 893-911.*

4) Chih-Hsiang Ho, Guancun Zhong, Fangjin Cui and Moinak Bhaduri, "Modeling interaction between bank failure and size", *Journal of Finance and Bank Management, 2016, 4(1), 15-33.*

5) Chih-Hsiang Ho and Moinak Bhaduri, "On a novel approach to forecast sparse rare events: applications to Parkfield earthquake prediction", *Natural Hazards, 2015, 78(1), 669-679.*

6) Siqi Tan, Moinak Bhaduri, Chih-Hsiang Ho, "A statistical model for long term forecasts

of strong dust sand storms", *Journal of Geoscience and Environment Protection, 2014, 2, 16-26.*

Papers under review:

1) Moinak Bhaduri and Chih-Hsiang Ho, "On a Temporal Investigation of Hurricane Strength and Frequency", *Environmental Modeling and Assessment, (under review).*

2) Moinak Bhaduri, Justin Zhan, Carter Chiu and Felix Zhan, "A Novel Distribution-Free Approach For Efficient Online Drift Identification", *IEEE Transactions on Knowledge and Data Engineering, (under review).*

3) Moinak Bhaduri and Justin Zhan, "Investigating similarity among time series through Empirical Recurrence Rates Ratio", *IEEE Access, (under review).*

In preparation:

1) Moinak Bhaduri and Chih-Hsiang Ho, "Power functions of the forward and backward tests for a power-law process", to be turned in to *Communications in Statistics - Simulation and Computation.*

2) Moinak Bhaduri and Chih-Hsiang Ho, "Modeling adventures with atypical point processes", a research monograph to be turned in to one of the following publishers: *Springer (SpringerBriefs), John Wiley and Sons, Taylor and Francis (CRC Press), Elsevier and Cambridge University Press.*

Conference presentations:

1) 2017 Joint Statistical Meetings, Baltimore, July 30, 2017, MD, USA: Presented the work titled "On a Time Reversal Approach for Testing a Poisson Intensity Against Smooth and Rough Alternatives".

2) UNLV Graduate and Professional students' Research Forum, April 8, 2017, Las Vegas, Nevada, USA: Presented the work titled "On Consequences of Time Reversal for Repairable Systems".

3) 2016 Joint Statistical Meetings, Chicago, August 4, 2016, IL, USA: Presented the work titled "On the efficient interpretability of Empirical Recurrence Rates Ratio as a hidden markov chain".

4) UNLV Graduate and Professional students' Research Forum, March 12, 2016, Las Vegas, Nevada, USA: Presented the work titled "An Investigation into the Forecasting Power of Empirical Recurrence Rates and Ratios with Emphasis on Rare Event Modeling".

5) 2015 Joint Statistical Meetings, Seattle, August 12, 2015, Washington, USA: Presented the work titled "On a temporal investigation of hurricane strength and frequency".

6) UNLV Department of Mathematical Sciences' Graduate Students' Colloquium, April 30, 2015, Las Vegas, Nevada, USA: Presented the work titled "On an efficient class of smoothing statistics with applications to rare event modeling".

7) UNLV Graduate and Professional students' Research Forum, March 21, 2015, Las Vegas, Nevada, USA: Presented the work titled "On a statistical investigation of the dependence structure between two related time series: Application to hurricane frequency modeling".

8) IISA conference on Research Innovation in Statistics for Health, Education, Technology and Society, July 2014, Riverside, CA, USA : Presented the work titled "On the comparison of two inversely related non-homogeneous Poisson processes.".

9) 19th Industrial Mathematical and Statistical Modeling Workshop, July 2013, NCSU, Raleigh, NC, USA : "Photoresponsive polymer design for Solar Concentrator self steering Heliostats". This paper was co-authored.

10) 99th Indian Science Congress, January 3, 2012 - January 7, 2012, Orissa, India : Presented the work titled "On the characterization of MaxEnt Density Functions under Non Shannon Entropies and some Related Problems."

11) 98th Indian Science Congress, January 3, 2011 - January 7, 2011, Chennai, India : Presented the work titled "On a Multistage Randomized Statistical test structure".

12) 97th Indian Science Congress, January 3, 2010 - January 7, 2010, Kerala, India : Presented the work titled "Derivation of Optimal Order Quantity under constraints in Classical Newsboy Setup". This paper was co-authored.

Leadership and service contributions:

1) Served on the *Top Tier Graduate Policies, Procedures and Mentorship Committee* (Fall 2016 and Spring 2017): primary duties included offering advice on diverse issues pertaining to the university's goal of achieving Top Tier status.

2) Served as a graduate student mentor on the *Rebel Research and Mentorship Program* (Fall 2016 and Spring 2017): primary duties included proposing a research question and motivating and supervising an undergraduate student's work towards meeting the intended target.

3) Served as the *Department of Mathematical Science's representative to the GPSA, UNLV* (Fall 2016 and Spring 2017): primary duties included serving as an effective contributor to and communicator of campus wide events and debatable issues such as the introduction of "Technology Fee".

4) Acted as an *invited panelist on the workshop titled "Designing Effective Research Assignments"* organized by the Graduate College, UNLV and GPSA, October 9, 2015: offered advice on how to formulate, administer and evaluate efficient research exercises to be given to undergraduate and graduate students.

5) Co-organized UNLV Department of Mathematical Sciences' weekly *Graduate Students' Colloquium* over Fall 2014 and Spring 2015.

Professional certifications:

1) Completed the *UNLV Graduate College Mentorship Certification Program*, demon-

strating a commitment to mentoring and academic excellence. This year long, professional development program provides graduate students with the skills and knowledge necessary to effectively be mentored and serve as a mentor in higher education settings.

2) Completed the *UNLV Graduate College Rebel Research and Mentorship Program*. Supervised and guided an undergraduate student over two semesters in efficiently modeling death rates due to leukemia through the use of novel statistics and forecasting tools.

Membership in professional organizations:

Student member of American Statistical Association, Bernoulli Society and Institute of Mathematical Statistics.

Fellow of Royal Statistical Society.

Teaching history at UNLV:

Fall 2017: Applied Statistics for Biological Sciences (STAT 391), Discussion.

Spring 2017: Applied Statistics for Biological Sciences (STAT 391), Discussion and Statistical Methods I (STAT 411), Discussion.

Fall 2016: Applied Statistics for Biological Sciences (STAT 391), Discussion and Applied Statistics for Engineers (STAT 463-663), Discussion.

Summer 2016: Applied Statistics for Biological Sciences (STAT 391), Lecture.

Spring 2016: Applied Statistics for Biological Sciences (STAT 391), Discussion and Statistical Methods I (STAT 411), Discussion.

Fall 2015: Introduction to Statistics (STAT 152), Lecture and Applied Statistics for Biological Sciences (STAT 391), Discussion.

Spring 2015: Applied Statistics for Engineers (STAT 463-663), Discussion.

Fall 2014: Statistical Methods I (STAT 411), Discussion.

Summer 2014: Finite Mathematics (MATH 132), Lecture.

Spring 2014: Precalculus II (MATH 127), Lecture and Applied Statistics for Engineers (STA 463-663), Discussion.

Fall 2013: Applied Statistics for Engineers (STA 463-663), Discussion.

Summer 2013: Precalculus I (MATH 126), Lecture.

Spring 2013: Finite Mathematics (MATH 132), Lecture and Calculus I (MATH 181), Discussion.

Fall 2012: Calculus I (MATH 181), Discussion.

Emails:
moinak.bhaduri@unlv.edu, bhadurim@unlv.nevada.edu.

Dissertation Title:

Bi-directional Testing For Change Point Detection In Poisson Processes

Dissertation Examination Committee:

Chairperson, Chih-Hsiang Ho, Ph.D.
Committee Member, Amei Amei, Ph.D.
Committee Member, Kaushik Ghosh, Ph.D.
Committee Member, Malwane Ananda, Ph.D.
Graduate Faculty Representative, Guogen Shan, Ph.D.