


August 2022

Bidirectional Testing for Repairable Systems Reliability: Power Asymmetries, Panel of Control Charts, and Reliability Graphics

Sung Keun Koo
University of Nevada, Las Vegas

Follow this and additional works at: <https://digitalscholarship.unlv.edu/thesesdissertations>

 Part of the [Statistics and Probability Commons](#)

Repository Citation

Koo, Sung Keun, "Bidirectional Testing for Repairable Systems Reliability: Power Asymmetries, Panel of Control Charts, and Reliability Graphics" (2022). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 4837.

<http://dx.doi.org/10.34917/36948187>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Dissertation has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

BIDIRECTIONAL TESTING FOR REPAIRABLE SYSTEMS RELIABILITY: POWER
ASYMMETRIES, PANEL OF CONTROL CHARTS, AND RELIABILITY GRAPHICS

By

Sung Keun Koo

Bachelor of Arts - Hospitality Business
Michigan State University
2007

Master of Science - Applied Statistics
DePaul University
2016

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
December 2022

Copyright © 2023 by Sung Keun Koo
All Rights Reserved

Dissertation Approval

The Graduate College
The University of Nevada, Las Vegas

November 2, 2022

This dissertation prepared by

Sung Keun Koo

entitled

Bidirectional Testing for Repairable Systems Reliability: Power Asymmetries, Panel of Control Charts, and Reliability Graphics

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

Amei Amei, Ph.D.
Examination Committee Member

Malwane Ananda, Ph.D.
Examination Committee Member

Jee Woong Park, Ph.D.
Graduate College Faculty Representative

Alyssa Crittenden, Ph.D.
*Vice Provost for Graduate Education &
Dean of the Graduate College*

ABSTRACT

BIDIRECTIONAL TESTING FOR REPAIRABLE SYSTEMS RELIABILITY: POWER ASYMMETRIES, PANEL OF CONTROL CHARTS, AND RELIABILITY GRAPHICS

by

Sung Keun Koo

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

To a practitioner who chooses to be on the safe side, we offer an option between a cocktail of tests and a solo one-size-fits-all test as a needed antidote to power asymmetries. Two bidirectional tests are first established to empower a basic pair of tests as asymmetrical performances. Unsurprisingly, we've seen either bidirectional device championing in one alternative setting, but also being turned against the very setting altered with just one of the composed elements. Progressing by filtering out the bad and enhancing the good, we assemble a hybrid from the empowered pair to restore power symmetries that are thought to be automatic and required for quality assurance. Data elaborated by the maximum power-deficits certify that the hybrid, termed as the dual bidirectional test, is robust with essential quality assurance in addition to a user-friendly version for the practitioners conducting the work. To further substantiate the applicability of a p -value induced dual bidirectional test, we challenge the legendary Laplace test to a duel. The outcomes reveal some elaborate concerns associated with the defending champion while supporting the challenger as an all-purpose test and a safe bet for practitioners doing the applicable real-world case studies. We extend the statistical process control architecture to an all-in-one panel of coherent bidirectional control charts fit for group sequential testing and multisystem. The charting tools start with a basic pair of tests that are more complementary than competitive, and are bound together with a common set of control limits. Another unified control limits, partners of an empowered pair of tests, form a second bidirectional

control chart. And, a series of novel group sequential control charts, centering on a hybrid of the empowered pair, characterize the test as the only tool possessing the unique feature of having proxies that are free of the system sample size and the sampling schemes. It is, therefore, capable of producing a one-size-fits-all control chart, to be considered in equipment performance assessments of repairable systems and, by extension, other applications. Moreover, we start the development by showing a simple dot-plot that cumulates all the dot-jumps up to each time-point of occurrences and leaves us an often mysterious but meaningful footprint to shed any light on. Its mystery lies in the sequence of the slopes and/or curvature associated with each corresponding hill edging up. We then launch elaborated graphics panels to gain a broad perspective on the data analysis to exchange information, support one another, share findings of reliability trends and the change-point(s) from both directions, and even interconnected with power asymmetries. A cocktail of bidirectional tests is a social lubricant in the saga of absorbing graphics to make it easier to collaborate on new and/or improved technologies in reliability engineering and, by extension, other applications.

ACKNOWLEDGEMENTS

Words cannot express my gratitude to Dr. Ho who is my professor and chair of my committee, for his invaluable patience and feedback. I also could not have undertaken this journey without my defense committee members, who are Dr. Amei, Dr. Ananda, and Dr. Park with generously provided knowledge and expertise. Additionally, this endeavor would not have been possible without the generous support from the Department of Mathematical Sciences and Graduate College at the University of Nevada, Las Vegas. I am also grateful to my classmates and cohort members, especially my office mates, for their editing help and moral support. Lastly, I would be remiss in not mentioning my family, especially my parents and my younger brother. Their belief in me has kept my spirits and motivation high during this process. I would also like to thank my dog, Hodu, for all the entertainment and emotional support.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER 1 INTRODUCTION	1
1.1 Weibull Distribution	1
1.2 Power Law Process	6
1.2.1 Point Process	6
1.2.2 Poisson Process	9
1.2.3 Forward Test	12
1.2.4 Backward Test	22
1.3 Bidirectional Testing	29
1.3.1 Right-Tailed Bidirectional Test	30
1.3.2 Left-Tailed Bidirectional Test	30
1.3.3 Dual Bidirectional Test	31
1.3.4 Empirical Null Distributions	31
CHAPTER 2 AN INTELLIGIBLE VERSION	36
2.1 Motivation	36
2.2 Propositions	36
2.3 The p -value Proxies	37
2.3.1 For the R -Test	37
2.3.2 For the L -Test	38
2.3.3 For the P_{DB} -Test	39
2.4 Performance Assessments	43
CHAPTER 3 POWER COMPARISONS	46
3.1 Motivation	46
3.2 Simulation Techniques	47
3.2.1 Power Law Process	47
3.2.2 Step-Function Intensities	49
3.3 Power Comparisons for All Five Tests	51
3.3.1 Failure-Truncated Sampling	52
3.3.2 Time-Truncated Sampling	59
3.4 An Overview	62
3.4.1 Road Map	63
3.4.2 Diagnosis of Power Asymmetries	65
3.5 Conclusions	68

CHAPTER 4 THE P_{DB}-TEST VERSUS THE LAPLACE TEST	71
4.1 Motivation	71
4.1.1 Theoretical Results	72
4.1.2 Simulation Studies	72
4.2 Power Comparisons	75
4.2.1 Power Law Process	76
4.2.2 Two-Step-Intensity	78
4.2.3 Nonmonotonic Three-Step-Intensity	81
4.3 The Verdict	84
4.4 Applications	87
4.5 Conclusions – An All-Purpose Test	88
CHAPTER 5 PANEL OF BIDIRECTIONAL CONTROL CHARTS	91
5.1 Motivation	91
5.2 The Preliminaries	93
5.3 Panel of Fully Sequential Control Charts	96
5.3.1 Single System	97
5.3.2 Multisystem	100
5.4 Panel of Group Sequential Control Charts	102
5.4.1 GS- P_{DB} -Chart and Multi-GS- P_{DB} -Chart	102
5.4.2 GS- (P_{DB}, β_a) -chart and Multi-GS- (P_{DB}, β_a) -chart	105
5.5 An Overview	111
5.6 Conclusions – An All-in-One Panel of Bidirectional Charts	112
CHAPTER 6 RELIABILITY GRAPHICS FOR QUALITY ASSURANCE	114
6.1 Introduction	114
6.2 Fingerprinting Graphics	115
6.2.1 The IET-plot	115
6.2.2 The CT-plot	118
6.3 Motivating Examples – Power Asymmetries	121
6.4 Panel of OR -Charts	126
6.4.1 Simulated Data	126
6.4.2 Real Data	133
6.5 Conclusions	137
BIBLIOGRAPHY	140
CURRICULUM VITAE	148

LIST OF TABLES

1.1	Critical regions for Z under the different alternative hypothesis.	17
1.2	Critical regions for Z_B under the different alternative hypothesis.	28
1.3	Empirical null distribution of the R -test.	33
1.4	Empirical null distribution of the L -test.	34
1.5	Empirical null distribution of the P_{DB} -test.	35
2.1	Functions for test statistics (Z and Z_B) and the p -value proxies (R , L , and P_{DB}), failure-truncated and time-truncated sampling with size n	44
2.2	Mean absolute deviation (variance) between tabulated empirical cumulative probabilities and their proxies for various sample sizes, failure-truncated sampling.	44
3.1	Estimated powers for testing H_0 versus H_A : a power law process with $\beta = (0.2 - 2.0)$, respectively, and $n = 40$ (failure-truncated sampling).	53
3.2	Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 1$, $\lambda_2 = 3$, and change point at τ (failure-truncated sampling).	55
3.3	Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 3$, $\lambda_2 = 1$, and change point at τ (failure-truncated sampling).	56
3.4	Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$ (failure-truncated sampling).	58
3.5	Estimated powers for testing H_0 versus H_A : a power law process with $\beta = (0.2 - 2.0)$, respectively, and $n = 40$ (time-truncated sampling).	59
3.6	Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 1$, $\lambda_2 = 3$, and change point at τ (time-truncated sampling).	59
3.7	Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 3$, $\lambda_2 = 1$, and change point at τ (time-truncated sampling).	60
3.8	Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$ (time-truncated sampling).	61
3.9	Maximum power-deficit overview.	63
4.1	Functions for the test statistics (L_p , Z , and Z_B) and the p -value proxies (R , L , and P_{DB}), failure-truncated and time-truncated sampling with size n . The cumulative failure times (t_i 's and t) are recognized as the main series of events.	72
4.2	Mean absolute deviation (variance) between the empirical cumulative probabilities ($p = 0.005 - 0.995$) of P_{DB} and their proxies for sample sizes $n = 6$ (10) 46, failure-truncated sampling.	74
4.3	Estimated powers for testing H_0 versus H_A : a power law process with $\beta = (0.1 - 3.0)$ for $n = 5, 15, 25$, and 40	76
4.4	Estimated powers for testing H_0 versus H_A : a two-step-intensity with change point at τ and $\lambda_1 = 1$, $\lambda_2 = 3$	79
4.5	Estimated powers for testing H_0 versus H_A : a two-step-intensity with change point at τ and $\lambda_1 = 3$, $\lambda_2 = 1$	80
4.6	Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$	82

4.7	Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$	83
4.8	Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$, time-truncated sampling.	83
4.9	Power-deficit overview.	86
5.1	Functions for the test statistics (Z and Z_B), the p -value proxies (R , L , and P_{DB}), $(1 - \alpha)100\%$ control limits (Z , Z_B , R , L , and P_{DB}) with R code, and m.l.e. of β for both sampling schemes of size n . The cumulative failure times (t_i 's and t) are recognized as the main series of events.	96
5.2	Top three data sets in terms of the number of failure times for air-conditioning equipment, planes 3, 6, and 7.	108

LIST OF FIGURES

1.1	Weibull distribution's p.d.f. for various values of shape parameter β	2
1.2	Bathtub diagram for different values of the shape parameter β	5
1.3	Distributions of the forward statistic Z corresponding to different parameter values .	16
1.4	Power curves under different types of alternatives	20
2.1	Pairwise comparison of the empirical null distribution ($n = 21$, failure-truncated sampling) and the corresponding order statistic for (a) R and X_{max} , (b) L and X_{min} , and (c) P_{DB} and U_{min}	45
3.1	Simulation from NHPP's with different choices of β with $n = 15$	48
3.2	Power Law (Weibull) intensities for different choices of β	49
3.3	Three types of graph in three-step-intensity	50
3.4	Simulation flowchart: (Sampling) schemes \rightarrow {(F)ailure-truncated sampling, (T)ime-truncated sampling} \rightarrow {PLP, (2-step) intensity, (3-step) intensity}, with simulation results from Tables 3.1 to 3.8.	52
3.5	Structure overview – simulation: failure-truncated sampling (Table 3.1 – 3.4); time-truncated sampling (Table 3.5 – 3.8).	65
3.6	Maximum power-deficit-based performance comparisons for tests of R and L for (a) failure-truncated sampling and (b) time-truncated sampling.	66
3.7	Maximum power-deficit-based performance comparisons for tests of Z and Z_B for (a) failure-truncated sampling and (b) time-truncated sampling.	66
3.8	Maximum power-deficit-based performance comparisons for tests of R , L , and P_{DB} for (a) failure-truncated sampling and (b) time-truncated sampling.	67
3.9	Box-plots for all 60 maximum power-deficits.	68
4.1	Simulation flowchart: (Sampling) schemes \rightarrow {(F)ailure-truncated sampling, (T)ime-truncated sampling} \rightarrow {PLP, (2-step) intensity, (3-step) intensity}, with simulation results from Tables 4.3 to 4.8.	75
4.2	Power-deficit-based performance comparisons for testing H_0 versus H_A : a PLP with $\beta \neq 1$ for (a) $n = 5$, (b) $n = 15$, (c) $n = 25$, and (d) $n = 40$	78
4.3	Power-deficit-based performance comparing tests of P_{DB} and L_P for (a) two-step-intensity, (b) three-step-intensity, and (c) all simulations (failure-truncated sampling, FTS; time-truncated sampling, TTS).	81
4.4	Structure overview – simulation: failure-truncated sampling (F: Tables 4.3 – 4.7); time-truncated sampling (T: Table 4.8).	85
4.5	Box-plots for power-deficits: (a) all the power-deficits and (b) power-deficits excluding zeros.	87
4.6	Panel of (a) P_{DB} -chart and (b) L_P -chart for failure times of an aircraft generator. .	88
5.1	Panel of (a) (Z, Z_B)-chart, (b) (R, L)-chart, and (c) P_{DB} -chart for a PLP with $\beta = 0.2$ for $n = 10$	97
5.2	Panel of (a) (Z, Z_B)-chart, (b) (R, L)-chart, and (c) P_{DB} -chart for a PLP with $\beta = 2.0$ for $n = 10$	98

5.3	Panel of (a) (Z, Z_B) -chart, (b) (R, L) -chart, and (c) P_{DB} -chart for an NHPP with $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 3)$ and the combination of [sample size, location of change point] is [20, 15], where $t_{15} = 0.8971$ and $t_{20} = 0.9836$	99
5.4	Panel of (a) (Z, Z_B) -chart, (b) (R, L) -chart, and (c) P_{DB} -chart for an NHPP with $\lambda(t)$, where $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ and the combination of [sample size, locations of change point] is [144, 49, 96], where $t_{49} = 15.1941$, $t_{96} = 63.3500$, and $t_{144} = 79.1885$. 100	100
5.5	Panel of (a) Multi- Z -chart and (b) Multi- P_{DB} -chart for PLPs with $\beta = 0.2$ and $\beta = 2.0$	101
5.6	Panel of control charts for the milestones of DJIA and NASDAQ from January 1, 1982 to December 31, 2021: (a) Z -chart for (1) DJIA and (2) NASDAQ; (b) Z_B -chart for (1) DJIA and (2) NASDAQ; (c) Multi-GS- P_{DB} -chart for both with (1) 40 groups (per year) and (2) 20 groups (every 2 years).	104
5.7	Panel of GS- (P_{DB}, β_a) -chart, 40 groups, for (a) DJIA and (b) NASDAQ from January 1, 1982 to December 31, 2021, eliminating the first set of all the plots at $g = 1$ (burn-in period).	106
5.8	Multi-GS- (P_{DB}, β_a) -chart, 40 groups, for DJIA and NASDAQ from January 1, 1982 to December 31, 2021, eliminating the first set of the plots at $g = 1$ (burn-in period). 107	107
5.9	Panel of Multi-GS- (P_{DB}, β_a) -charts for (a) All 3 planes together, and three pairwise charting of: (b) 3 versus 6, (c) 3 versus 7, and (d) 6 versus 7. All control charts have the same group size, $g = 10$	109
5.10	Panel of Multi-GS- (P_{DB}, β_a) -charts for (a) All 3 planes together, and three pairwise charting of: (b) 3 versus 6, (c) 3 versus 7, and (d) 6 versus 7. All control charts have the same group size, $g = 20$, eliminating the first set of all the plots at $g = 1$ (burn-in period).	110
5.11	Structure overview for panels of control charts.	112
6.1	IET-plot for a total of 59 major earthquakes, occurred in Japan during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale. . .	116
6.2	IET-plot for a total of 53 major earthquakes, occurred in China during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale. . .	117
6.3	Panel of 12 IET-plots for the failure times of air-conditioning equipment on 12 aircraft. 118	118
6.4	CT-plot for a total of 59 major earthquakes, occurred in Japan during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale. . .	119
6.5	CT-plot for a total of 53 major earthquakes, occurred in China during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale. . .	120
6.6	Panel of 12 CT-plots for the failure times of air-conditioning equipment on 12 aircraft. 121	121
6.7	Panel of (a) IET-plot and (b) CT-plot for a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$	122
6.8	Panel of (a) Z -chart, (b) R -chart, (c) L -chart, and (d) P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$	123
6.9	Panel of (a) IET-plot and (b) CT-plot for the reversal of a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$	124

6.10	Panel of (a) Z -chart, (b) R -chart, (c) L -chart, and (d) P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for the reversal of a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$	125
6.11	Panel of (a) OR - Z -chart, (b) OR - R -chart, (c) OR - L -chart, and (d) OR - P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$	127
6.12	Panel of (a) OR - Z -chart, (b) OR - R -chart, (c) OR - L -chart, and (d) OR - P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for air-conditioning equipment on aircraft 1 to 4.	130
6.13	Panel of (a) OR - Z -chart, (b) OR - R -chart, (c) OR - L -chart, and (d) OR - P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for air-conditioning equipment on aircraft 5 to 8.	131
6.14	Panel of (a) OR - Z -chart, (b) OR - R -chart, (c) OR - L -chart, and (d) OR - P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for air-conditioning equipment on aircraft 9, 10, 12, and 13.	132
6.15	Graphics panel of (a) IET-plot, (b) CT-plot, (c) OR - Z -chart, (d) OR - R -chart, (e) OR - L -chart, and (f) OR - P_{DB} -chart for air-conditioning equipment on Aircraft 13.	133
6.16	Multi-CT-plot for comparisons of intensity functions of three simulated PLPs with $n = 20$, at $\beta = 0.5, 1.0$, and 2.0 : time-truncated sampling, all set at $T = 1$	135
6.17	Graphics panel of (a) IET-plot, (b) CT-plot, (c) OR - Z -chart, (d) OR - R -chart, (e) OR - L -chart, and (f) OR - P_{DB} -chart for air-conditioning equipment on Aircraft 8.	136

CHAPTER 1

INTRODUCTION

1.1 Weibull Distribution

In a technological world, nearly all the applications of practical sciences to industry or commerce depend upon the continued functioning of a wide array of complex machinery and equipment for their health, safety, mobility, and economic welfare. Continual assessment of new product reliability and ongoing control of reliability are crucial necessities in today's competitive arena. This phenomenon brings about several reasons to discuss the Weibull distribution in the dissertation at first. Firstly, the Weibull distribution with a scale parameter $\theta > 0$ and a shape parameter $\beta > 0$ is widely used in the field of statistics, engineering, medicine, social science, finance, insurance, biology, and elsewhere related to the context of a lifetime not only to consider the reliability also to repair the failures. Since it is essential to know if the repairable system's reliability has grown, decreased, or remained stable during the period, the cumulative failure times are recognized as the main events. Second, the Weibull distribution is related to the power law process (PLP), which is known as a commonly used model for the repairable systems. Third, if repairs bring a system back to a good as a new state, the times between failures are independent identically distributed (iid) Weibull random variables (Rigdon and Basu (2000) [81]).

The shape parameter β is the key to understand the behavior of Weibull distribution's probability density function (p.d.f.), which is:

$$f(t) = (\beta/\theta)(t/\theta)^{\beta-1}e^{-(t/\theta)^\beta}, t > 0, \quad (1.1)$$

where $\beta > 0$, and $\theta > 0$. Three different regimes depend on the value of shape parameter β . For $\beta < 1$, the p.d.f. tends to infinity as the time approaches zero. The function is unbounded, and

there is no turning point. This case is well suited to the time series data with heavy tails. When $\beta = 1$, the p.d.f. is finite at the starting point and has no turning point. Finally, if $\beta > 1$, the p.d.f. looks like the bell-shaped curve of the normal distribution with a hump, but it's asymmetric. As the value of β increases, the function resembles a normal distribution. The Figure 1.1 displays the Weibull distribution's p.d.f. for various values of shape parameter β for a fixed $\theta = 1$.

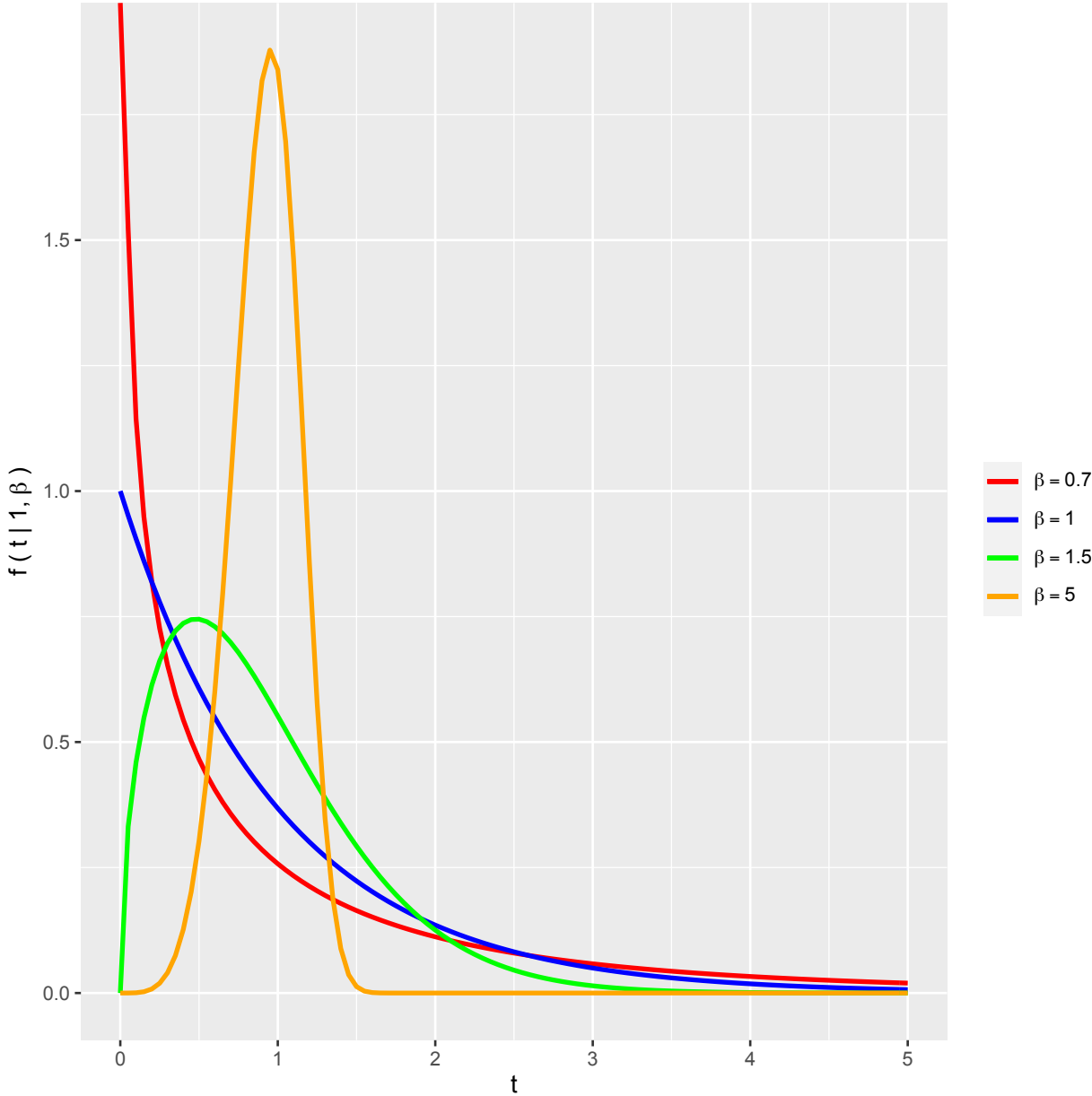


Figure 1.1. Weibull distribution's p.d.f. for various values of shape parameter β .

The hazard function well explains the properties of the Weibull distribution. This function describes how likely something will fail given that it has survived so far. In other words, this function is the conditional probability given that the failure we are concerned about has not occurred yet.

Let's consider the probability that the failure occurs between time t and $t + dt$. This is the probability that it has survived until time t , and it will fail in the next dt , and it may be expressed as:

$$f(t)dt = S(t)h(t)dt, t > 0 \quad (1.2)$$

where $f(t)$ is the p.d.f. for the failure at time t , $S(t)$ is the survival function describing the probability that the failure has not occurred until time t , and $h(t)$ is the hazard function.

The cumulative distribution function (c.d.f.) is the integral of the p.d.f. and it represents the probability that the failure happened before some time t such as:

$$F(t) = \int_0^t f(x)dx = 1 - e^{-(t/\theta)^\beta}, t > 0. \quad (1.3)$$

The survival function is the probability that there is no failure until time t , which is

$$S(t) = 1 - F(t) = e^{-(t/\theta)^\beta}, t > 0. \quad (1.4)$$

From (1.2), (1.3), and (1.4) equations, we can determine that the hazard function is the negative rate of change of \ln of the survival function:

$$h(t) = -\frac{d \ln(S(t))}{dt} = -\frac{d \ln(1 - F(t))}{dt} = \frac{f(t)}{S(t)}, t > 0. \quad (1.5)$$

It turns out that the hazard function can be described as:

$$h(t) = (\beta/\theta)(t/\theta)^{\beta-1}, t > 0. \quad (1.6)$$

This hazard function can distinguish three different behaviors: (1) when $\beta < 1$ known as the decreasing failure rate or improving, the times between the failures tend to increase, (2) if $\beta = 1$,

the rate of failure is fairly constant with advancing age, and (3) for $\beta > 1$ known as the increasing failure rate or deteriorating, the times between the failures tend to get shorter as time goes on. The Figure 1.2 called the "bathtub" curve is helpful for visualizing the distinctions among different values of the shape parameter β . This curve consists of three periods: an infant mortality period with a decreasing failure rate followed by a standard life period with a relatively constant failure rate and concluding with a wear-out period that displays an increasing failure rate.

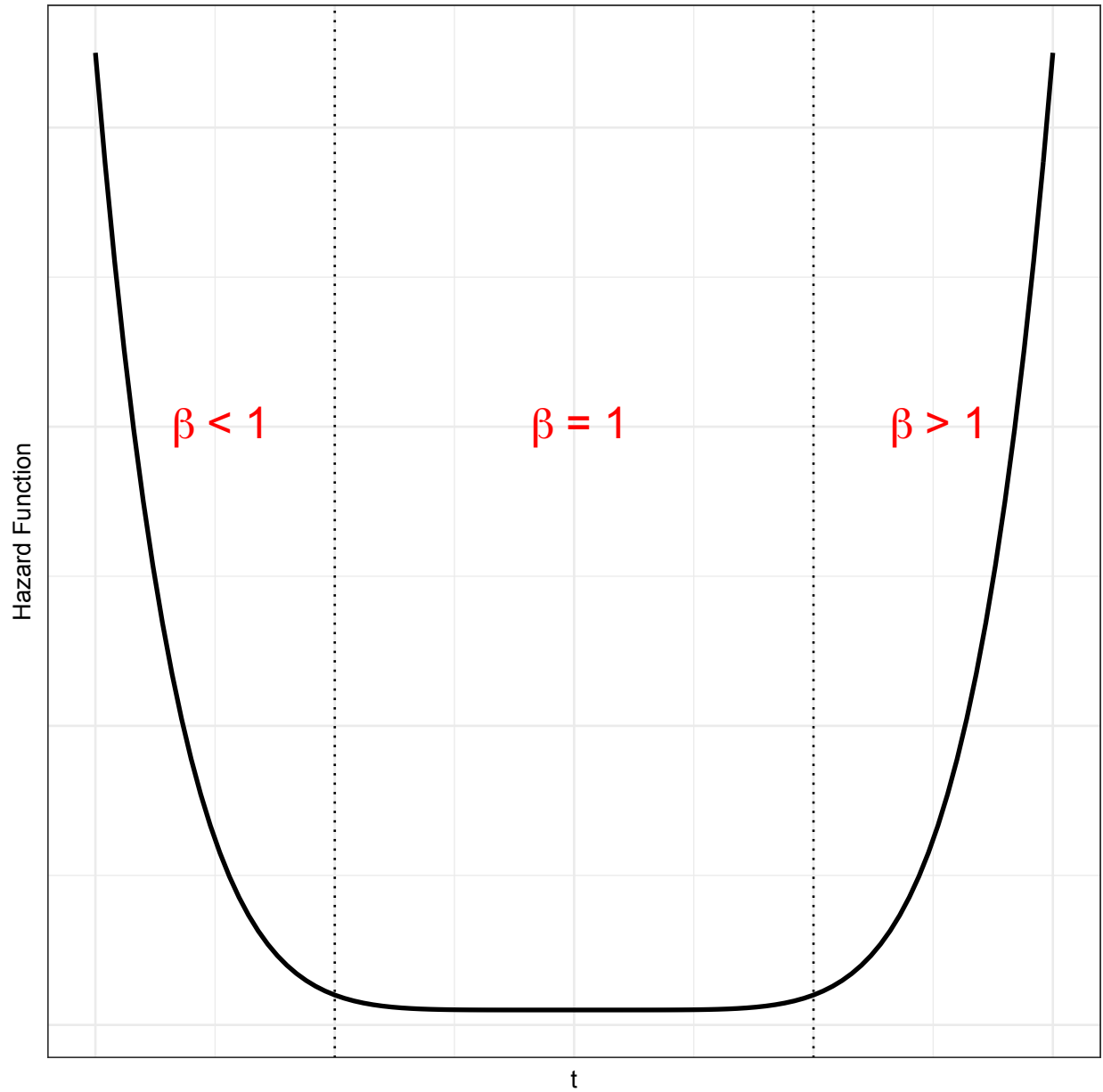


Figure 1.2. Bathtub diagram for different values of the shape parameter β .

As mentioned at the beginning of this chapter, the Weibull distribution has many applications ranging in the areas mainly in the context of a lifetime to consider the reliability, since nearly most depend upon the continued functioning of a wide array of complex machinery and equipment for their health, safety, mobility, and economic welfare. Moreover, as the lifetime of events such as most environmental phenomena with the inter-occurrences times between the successive events

(e.g., earthquakes, volcanoes, floods, wildfires, etc) have skewed shape and heavy tail behavior of the underlying distribution, the Weibull distribution can be the perfect distribution to analyze it. Johnson et al. (1994) [62] provided extensive works on the Weibull distribution and its applications. Rinne (2008) [82] also researched an extensive list of applications of Weibull distribution.

1.2 Power Law Process

This chapter provides the fundamental tests on detecting the non-stationarity of the cumulative failure time against stationarity. Primarily, it recalls the forward test Z -test known as the optimal test under the PLP and its time-reversed counterpart known as the Z_B -test under the time-truncated sampling.

1.2.1 Point Process

The point process, frequently called a counting process, is a stochastic model for describing events occurring at a random time. This process tracks the number of occurrences of events, which are recorded on the time axis. Throughout the dissertation, the cumulative failure times of repairable systems are recognized as the main events. Thus, the models for the repairable systems must be able to describe the cumulative failure times in time by a point process, which tracks the number of occurrences of events.

Definition 1.1. *Let $N(t)$ be the random variable that denotes the number of failures in the interval $[0, t]$. When N has as its argument an interval, we can write the number of failures in the interval $(a, b]$ with $b > a > 0$ as:*

$$N(a, b] = N(b) - N(a). \tag{1.7}$$

Definition 1.2. *The mean function of a point process is defined as:*

$$\mu(t) = E(N(t)) \tag{1.8}$$

which represents the expected number of events to time t .

Since $N(t)$ is a non-decreasing step function, there is no doubt that the mean function $\mu(t)$ must be non-decreasing. Thus, monotonicity of $N(t)$ leads to monotonicity of $\mu(t)$, and the mean function of a point process is right continuous (Rigdon and Basu (2000) [81]).

Alternative methods of specifying the point process are to find the joint density of the failure times $0 < t_1 < t_2 < \dots < t_n$ or the inter-event times such as

$$x_i = t_i - t_{i-1}, \quad i = 1, 2, \dots, n. \quad (1.9)$$

Theorem 1.1. *For any point process,*

1. $t_i > v \iff N(v) < i$
2. $t_i \leq w \iff N(w) \geq i$
3. $v < t_i \leq w \iff N(v) < i \leq N(w)$

Using the results from Theorem, the joint density of any of the following sets determine the joint density of others (Rigdon and Basu (2000) [81]).

1. $N(u_1), N(u_2), \dots, N(u_n)$ for any n and for any u_i 's
2. t_1, t_2, \dots, t_n for any n
3. x_1, x_2, \dots, x_n for any n

Definition 1.3. *A point process has independent increments if for all n and for all $r_1 < s_1 \leq r_2 < s_2 \leq \dots < r_n \leq s_n$,*

$$P(N(r_1, s_1] = k_1, \dots, N(r_n, s_n] = k_n) = \prod_{i=1}^n P(N(r_i, s_i] = k_i). \quad (1.10)$$

This implies that the random variables $N(r_1, s_1], N(r_2, s_2], \dots, N(r_n, s_n]$ are independent.

If we investigate the repairable systems, our question could be whether the pattern of failures is constant with advancing age, or whether the times between two sequences of consecutive failures tend to increase for an improving system or decrease for a deteriorating system over the time t .

Definition 1.4. *A point process has stationary increments if for all $k > 0$*

$$P(N(t, t + s] = k) \tag{1.11}$$

is independent of t .

This implies that the pattern of failures is stable or stationary over time t . The resulting point process is called a homogeneous Poisson process (HPP).

Definition 1.5. *The intensity function of a point process is*

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(N(t, t + \Delta t] \geq 1)}{\Delta t}. \tag{1.12}$$

This intensity function is the unconditional probability of failure in a short time interval divided by the length of the interval. It gives the instantaneous probability of occurring at least one failure in a small time interval. Thus, getting higher intensity gives a higher chance of failures over intervals, and fewer failures over intervals on which $\lambda(t)$ is small. For some models, it is more necessary to consider the conditional probability given the failure history of the process.

Definition 1.6. *The complete intensity function is*

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(N(t, t + \Delta t] \geq 1 | H_t)}{\Delta t} \tag{1.13}$$

where H_t denotes the entire history of the failure process through time t and represents the set of failure times $\{t_i : i = 1, 2, \dots, N(t)\}$.

1.2.2 Poisson Process

The Poisson process is one of the most widely-used point process. This process is a model for a series of discrete events where the average time between events is known, but the exact timing of events is random. In other words, this process is usually used in scenarios where we are counting the occurrences of the event that occur at a specific rate, but completely at random.

A point process $N(t)$ is said to be a Poisson process if:

1. $N(0) = 0$.
2. For any $a < b \leq c < d$, the random variables $N(a, b]$ and $N(c, d]$ are independent. This property is called the independent increments.
3. There is a function λ such that:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(N(t, t + \Delta t] = 1)}{\Delta t}. \quad (1.14)$$

This function is called the intensity function of the Poisson process.

4. The last property preclude the possibility of simultaneous failures.

$$\lim_{\Delta t \rightarrow 0} \frac{P(N(t, t + \Delta t] \geq 2)}{\Delta t} = 0. \quad (1.15)$$

Properties (1) through (4) of the Poisson process imply (Rigdon and Basu (2000) [81]) that

$$P(N(t) = n) = \frac{1}{n!} \left(\int_0^t \lambda(x) dx \right)^n \exp\left(- \int_0^t \lambda(x) dx\right), \quad n = 0, 1, 2, \dots \quad (1.16)$$

$$N(t) \sim \text{Poisson}\left(\int_0^t \lambda(x) dx\right). \quad (1.17)$$

The expected number of failures as events through time t , known as the the mean function, can be calculated as:

$$\mu(t) = E(N(t)) = \int_0^t \lambda(x) dx. \quad (1.18)$$

The most common form of the intensity functions of the Poisson process is

$$\lambda(t|\Theta) = \frac{d}{dt}\mu(t|\Theta) = (\beta/\theta)(t/\theta)^{\beta-1} \quad (1.19)$$

where Θ is a vector of parameters, $\beta > 0$, and $\theta > 0$. This form, termed the PLP, has found applications in reliability analysis due to its flexibility and the fact that the distribution of the time to first arrival in the process is Weibull (Finkelstein 1976 [37]; Lee and Lee1978 [65]; Bain and Engelhardt 1980 [8]; Crow 1982 [24]; Bain and Engelhardt 1991 [9]; Ho 1993 [49], 1998 [51]; Rigdon and Basu 2000 [81]; Altun and Comert 2016 [3]). Depending on the sampling schemes, the maximum likelihood estimates of parameters from the PLP of $\lambda(\cdot)$ can be calculated slightly differently.

A goodness-of-fit test, maximum likelihood (ML) estimates of β and θ , confidence intervals, inference procedures, and sequential testing for this process are presented in Bassin (1969) [12], Crow (1974 [23], 1982 [24]), Finkelstein (1976) [37], Lee and Lee (1978) [65], Bain and Engelhardt (1980) [8], Crow (1982) [24], and Ho (1993 [49], 1998 [51]). Extended applications, for example, change-point detections are demonstrated in Bhaduri 2018 [14].

- **Time-truncated sampling**

If a nonhomogeneous Poisson process (NHPP) with intensity function is observed until time t , and if the cumulative failure times are $t_1 < t_2 < \dots < t_n < t$ where n is a random number of failures in the interval $(0, t]$, the failures are said to be the time-truncated sampling with a predetermined time t (Rigdon and Basu (2000) [81]). The likelihood equation for the time-truncated case can be derived as follows: The joint density of (N, T_1, \dots, T_N) is

$$f(n, t_1, \dots, t_n) = \begin{cases} f_N(n)f(t_1, \dots, t_n|n), & n \geq 1 \\ f_N(0), & n = 0. \end{cases} \quad (1.20)$$

The random variable N has a Poisson distribution with mean $(t/\theta)^\beta$, so

$$f_N(n) = \frac{[(t/\theta)^\beta]^n \exp[-(t/\theta)^\beta]}{n!}, \quad n = 0, 1, \dots \quad (1.21)$$

The random variables $T_1 < T_2 < \dots < T_N < T$ given $N = n$ are distributed as n order statistics from the distribution with c.d.f.:

$$G(y) = \begin{cases} 0 & y < 0 \\ \mu(y)/\mu(t) & 0 \leq y \leq t \\ 1 & y > t. \end{cases} \quad (1.22)$$

For the PLP, this is:

$$G(y) = \frac{(y/\theta)^\beta}{(t/\theta)^\beta} = \left(\frac{y}{t}\right)^\beta, 0 \leq y \leq t. \quad (1.23)$$

The density corresponding to G would be

$$g(y) = \frac{\beta}{t} \left(\frac{y}{t}\right)^{\beta-1}, 0 \leq y \leq t. \quad (1.24)$$

Thus, the joint density of $t_1 < t_2 < \dots < t_n < t$ given n with the PLP of $\lambda(\cdot)$ is

$$\begin{aligned} f(t_1, t_2, \dots, t_n | n) &= n! \prod_{i=1}^n G'(t_i) \\ &= n! \prod_{i=1}^n \frac{\beta}{t} \left(\frac{t_i}{t}\right)^{\beta-1} \\ &0 < t_1 < t_2 < \dots < t_n < t. \end{aligned} \quad (1.25)$$

Thus, the joint density of n and $t_1 < t_2 < \dots < t_n < t$ is

$$\begin{aligned} f(n, t_1, t_2, \dots, t_n) &= \frac{\beta^n}{\theta n \beta} \left(\prod_{i=1}^n t_i\right)^{\beta-1} \exp[-(t/\theta)^\beta], \\ n \geq 1, 0 < t_1 < t_2 < \dots < t_n < t. \end{aligned} \quad (1.26)$$

To get the maximum likelihood estimation (MLE) of parameters, it is required to take the natural logarithm of this joint density and set the first partial derivatives with respect to θ and β equal to zero. The log-likelihood function is

$$l(\theta, \beta | n, t) = n \ln \beta - n \ln \theta + (\beta - 1) \sum_{i=1}^n \ln t_i - n(\beta - 1) \ln \theta - \left(\frac{t}{\theta}\right)^\beta. \quad (1.27)$$

If at least one failure occurs before time t then MLE's exist and are equal to:

$$\hat{\beta} = n / \sum_{i=1}^n \ln(t/t_i), \hat{\theta} = t/n^{1/\hat{\beta}}. \quad (1.28)$$

- **Failure-truncated sampling**

If the cumulative failure times of an NHPP are $t_1 < t_2 < \dots < t_n$, the failures are said to be the failure-truncated sampling with a predetermined number of failures n (Rigdon and Basu (2000) [81]).

Under this sampling scheme, the Poisson process is continually monitored until a predetermined number of failures n , which is deterministic, and the time of the last occurrence t_n is uncertain. After observing the cumulative failure times until a predetermined number of failures n , the maximum likelihood estimates of the parameters can be obtained as

$$\hat{\beta} = n / \sum_{i=1}^{n-1} \ln(t_n/t_i), \hat{\theta} = t_n/n^{1/\hat{\beta}}. \quad (1.29)$$

1.2.3 Forward Test

Rigdon and Basu (2000) [81] show that the quantity $Z = 2n\beta/\hat{\beta}$ is pivotal, having a $\chi^2(2n)$ for the time-truncated sampling. In keeping on working on a time-reversed version of Z formulated by Ho (1993) [49], this test based on the Z is called as the forward test, which is

$$Z = 2n/\hat{\beta} = 2 \sum_{i=1}^{n-1} \ln(t_n/t_i), \text{ or} \quad (1.30)$$

$$Z = 2n/\hat{\beta} = 2 \sum_{i=1}^n \ln(t/t_i), \quad (1.31)$$

with a time-truncated sampling and under the null hypothesis $H_0 : \beta = 1$, where conditional on $N > 0$ and for an observed number of failures, $n > 0$, the failure times are denoted by $0 < t_1 < t_2 < \dots < t_n < t$.

1.2.3.1 The Null and the Alternative Distributions of Z

Rigdon and Basu (2000) [81] focus on testing that the failure process is homogeneous or not. But we design the test to consider that the failure process is homogeneous (i.e. $\beta = 1$), improving (i.e. $\beta < 1$), or deteriorating (i.e. $\beta > 1$) such as:

$$H_0 : \beta = 1 \text{ vs } H_A : \beta > 1 \quad (1.32)$$

$$H_0 : \beta = 1 \text{ vs } H_A : \beta < 1 \quad (1.33)$$

$$H_0 : \beta = 1 \text{ vs } H_A : \beta \neq 1. \quad (1.34)$$

• **Time-Truncated Sampling**

Rigdon and Basu (2000) [81] shows how to form the test statistic called the forward test Z , as well as the distributions of null hypothesis and alternative hypothesis under the time-truncated sampling.

Theorem 1.2. *For the time-truncated sampling, $2n\beta/\hat{\beta} \sim \chi^2(2n)$.*

Proof. Conditioned on the number of occurrences $N = n$, the failure times $t_1 < t_2 < \dots < t_n < t$ are distributed as n order statistics from the distribution with c.d.f.:

$$G(y) = \begin{cases} 0 & y \leq 0 \\ \mu(y)/\mu(t) & 0 < y < t \\ 1 & y \geq t. \end{cases} \quad (1.35)$$

For the PLP, the c.d.f. are:

$$G(y) = \begin{cases} 0 & y \leq 0 \\ (\frac{y}{t})^\beta & 0 < y < t \\ 1 & y \geq t. \end{cases} \quad (1.36)$$

Let Y be a random variable with c.d.f. G , then

$$(\frac{y}{t})^\beta = G(y) = P(Y \leq y) = P(Y/t \leq y/t) = P((Y/t)^\beta \leq (y/t)^\beta), \quad 0 < y < t. \quad (1.37)$$

This shows that $(Y/t)^\beta$ is uniformly distributed on the interval $(0, 1)$. Thus, $(t_i/t)^\beta, i = 1, 2, \dots, n$ are distributed as n order statistics from $U(0, 1)$. Applying the fact that if U has a uniform distribution on the interval $(0, 1)$, then $X = -\theta \ln U$ has an exponential distribution with mean θ , we can claim that

$$\sum_{i=1}^n -\ln(t_i/t)^\beta = -\beta \sum_{i=1}^n \ln(t_i/t) \quad (1.38)$$

is a gamma $(n, 1)$ distribution. After multiplying 2 to a gamma $(n, 1)$ density, we conclude that

$$-2\beta \sum_{i=1}^n \ln(t_i/t) = 2\beta \sum_{i=1}^n \ln(t/t_i) = 2n\beta \left(\frac{n}{\sum_{i=1}^n \ln(t/t_i)} \right)^{-1} = 2n\beta/\hat{\beta} \quad (1.39)$$

has a chi-square distribution with $2n$ degrees of freedom.

□

After proving that $2n\beta/\hat{\beta} \sim \chi^2(2n)$, we can show that the general form of the test statistic called the forward test Z , which is

$$Z = 2n\beta/\hat{\beta}. \quad (1.40)$$

It is required to find the p.d.f. of Z under the null hypothesis and the alternative hypothesis to calculate both errors and the power. Under the null hypothesis $H_0 : \beta = 1$, $Z = 2n\beta_0/\hat{\beta} = 2n/\hat{\beta}$ follows $\chi^2(2n)$. In the same way, under the alternative hypothesis, $2n\beta/\hat{\beta} \sim \chi^2(2n)$ and the technique for a transformation of a random variable is used to derive the alternative distribution of Z as follows:

Let $Z = 2n\beta_0/\hat{\beta} = \frac{\beta_0}{\beta} \frac{2n\beta}{\hat{\beta}} = \frac{\beta_0}{\beta} X$, where $X \sim \chi^2(2n)$, then

$$Z = \frac{\beta_0}{\beta} X \implies X = \frac{\beta}{\beta_0} Z = g^{-1}(Z) \implies \frac{d}{dz} g^{-1}(z) = \frac{\beta}{\beta_0} \quad (1.41)$$

$$f_X(x) = \frac{1}{\Gamma(n)2^n} x^{n-1} e^{-\frac{x}{2}}, \quad x > 0 \quad (1.42)$$

$$\begin{aligned} f_Z(z) &= f_X(g^{-1}(z)) \left| \frac{d}{dz} g^{-1}(z) \right| \\ &= \frac{1}{\Gamma(n)2^n} \left(\frac{\beta}{\beta_0} z \right)^{n-1} e^{-\frac{\beta z}{2\beta_0}} \beta \beta_0 \\ &= (\beta/\beta_0)^n \frac{1}{\Gamma(n)2^n} z^{n-1} e^{-\frac{\beta z}{2\beta_0}}, \quad z > 0. \end{aligned} \quad (1.43)$$

The same technique is used to obtain the probability density of $\hat{\beta}$ by setting that $X = \frac{2n\beta}{\hat{\beta}}$, $Y = \hat{\beta} = \frac{2n\beta}{X} = g(X)$, and $\frac{d}{dy} g^{-1}(y) = -2n\beta \frac{1}{y^2}$. Using the p.d.f. of X above, we can derive

$$f_{\hat{\beta}}(\hat{\beta}) = \frac{1}{\Gamma(n)} \frac{(n\beta)^n}{\hat{\beta}^{n+1}} e^{-\frac{n\beta}{\hat{\beta}}}, \quad \hat{\beta} > 0. \quad (1.44)$$

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

Setting $\frac{\beta}{t} = x$ and $dt = -\frac{\beta}{x^2}dx$, the expectation of $\hat{\beta}$ can be calculated as follows:

$$\begin{aligned}
 E(\hat{\beta}) &= \int_{\infty}^0 \frac{1}{\Gamma(n)} n^n x^n e^{-nx} \left(-\frac{\beta}{x^2}\right) dx \\
 &= \frac{n^n \beta}{\Gamma(n)} \int_0^{\infty} e^{-nx} x^{n-2} dx \\
 &= \frac{n^n \beta}{\Gamma(n)} \frac{\Gamma(n-1)}{n^{n-1}} \\
 &= \frac{n\beta}{(n-1)}.
 \end{aligned} \tag{1.45}$$

This result draws a conclusion that $\frac{n-1}{n}\hat{\beta}$ is unbiased estimator for β (Rigdon and Basu (2000) [81]).

The null and alternate distributions of the forward statistic Z for the time-truncated sampling with $n = 10$ are compared in Figure 1.3.

Comparison of the null and alternative distribution of Z statistic (n = 10)

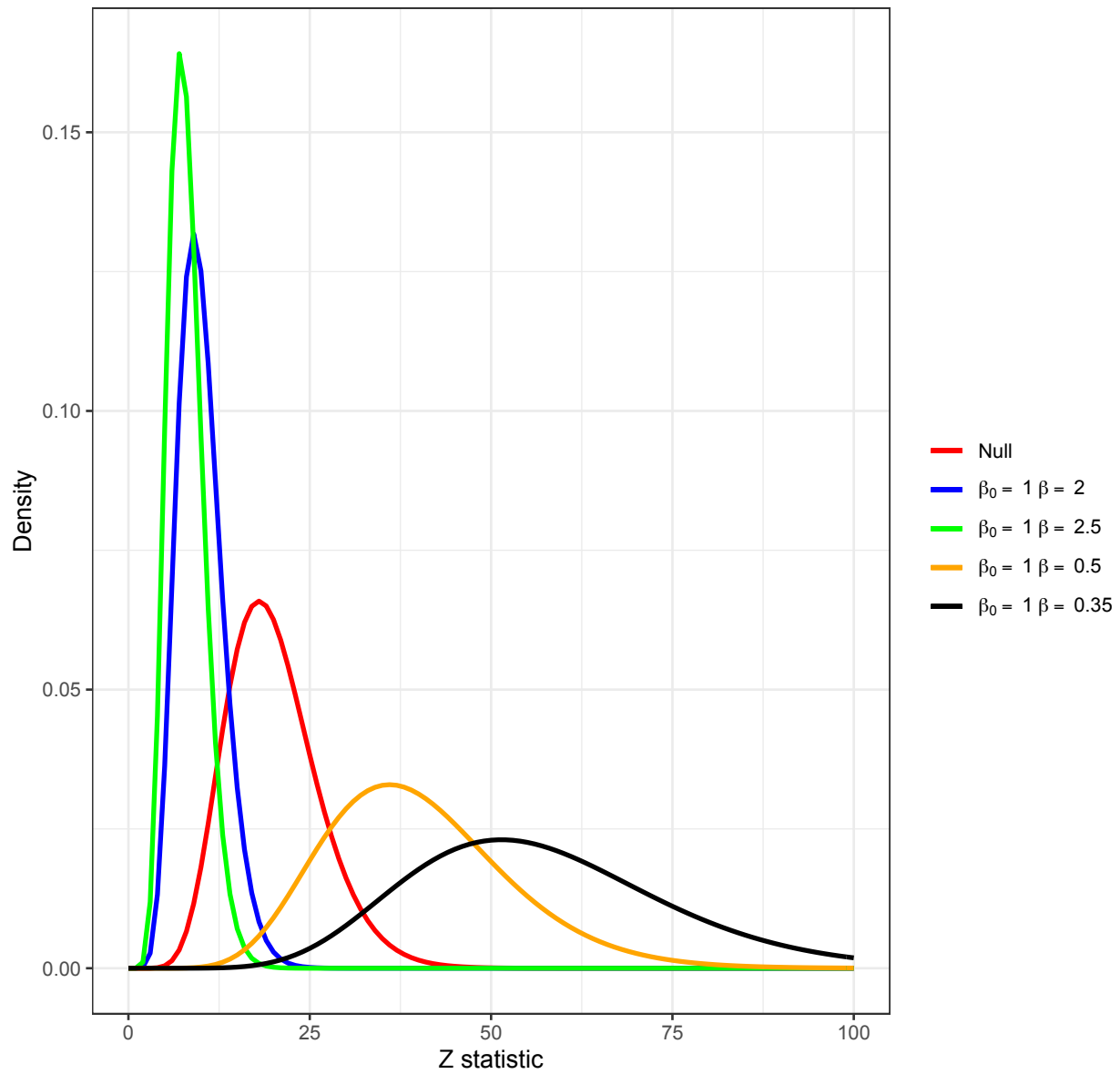


Figure 1.3. Distributions of the forward statistic Z corresponding to different parameter values

- **Failure-truncated sampling**

Same technique is employed to get the analogous results for the failure-truncated sampling by noting that $2n\beta/\hat{\beta}$ follows $\chi^2(2n-2)$ (Rigdon and Basu (2000) [81]). The distribution of $Z = 2n/\hat{\beta}$

under the null hypothesis $H_0 : \beta = 1$ is $\chi^2(2n - 2)$, and the alternative distribution is such as:

$$f_Z(z) = (\beta/\beta_0)^{n-1} \frac{1}{\Gamma(n-1)2^{n-1}} z^{n-2} e^{-\frac{\beta z}{2\beta_0}}, \quad z > 0. \quad (1.46)$$

1.2.3.2 Critical Region

The critical region is the set of outcomes of a statistical test for which the null hypothesis is to be rejected at the level of significance α . The statistical test will be one-tailed or two-tailed, depending on the nature of the null hypothesis and the alternative hypothesis. Under the one-tailed “greater than” type of alternative, large values of $\hat{\beta}$ tend to conclude to reject the null hypothesis. Also, a small value of statistic $2n\beta_0/\hat{\beta} = 2n/\hat{\beta}$ under $H_0 : \beta = 1$ leads to the rejection of the null hypothesis. Thus, the rejection region of the “greater than” type of alternative under the time-truncated sampling at the level of significance (α) is

$$\{z : z = 2n/\hat{\beta} < \chi_{1-\alpha}^2(2n)\} \quad (1.47)$$

where $\chi_{1-\alpha}^2(2n)$ is the lower α point of a chi-square distribution with $2n$ degrees of freedom. Equivalently, after solving for $\hat{\beta}$, the above set of critical region is same as:

$$\{\hat{\beta} : \hat{\beta} > 2n/\chi_{1-\alpha}^2(2n)\}. \quad (1.48)$$

The Table 1.1 shows the critical regions in Z depending on the different alternative hypothesis under the different sampling schemes.

Table 1.1. Critical regions for Z under the different alternative hypothesis.

	$H_A : \beta > 1$	$H_A : \beta < 1$	$H_A : \beta \neq 1$
Time-truncated	$Z < \chi_{1-\alpha}^2(2n)$	$Z > \chi_{\alpha}^2(2n)$	$Z < \chi_{1-\frac{\alpha}{2}}^2(2n)$ or $Z > \chi_{\frac{\alpha}{2}}^2(2n)$
Failure-truncated	$Z < \chi_{1-\alpha}^2(2n-2)$	$Z > \chi_{\alpha}^2(2n-2)$	$Z < \chi_{1-\frac{\alpha}{2}}^2(2n-2)$ or $Z > \chi_{\frac{\alpha}{2}}^2(2n-2)$

1.2.3.3 Analytical Power Function of Z-test

The power of a hypothesis testing is defined as the probability of making a correct decision, such as rejecting the null hypothesis when the alternative hypothesis is true. Simply put, the power is the probability of not making a Type II error which is the non-rejection of a false null hypothesis. The power should be between 0 and 1; if the power is close to 1, the hypothesis testing is good at detecting a false null hypothesis. If all other things are held constant and α increases, the power will increase. This is because a larger α means a larger rejection region and thus a greater probability of rejecting the null hypothesis. In the dissertation, the power is calculated by the Monte Carlo technique. The basic idea of the Monte Carlo technique is 1) to generate a data set assuming the alternative hypothesis is true, 2) to test the null hypothesis using the data set, 3) to save the results of the test, such as rejecting or failing to reject the null hypothesis, and 4) to repeat steps 1–3 10^4 times. The proportion of rejecting the null hypothesis estimates statistical power by the Monte Carlo technique.

Since the alternative distribution of Z is derived in the previous section, we can express the power functions under different alternatives and sampling schemes. Under the one-tailed “greater than” type of alternative, the small value of Z leads to the rejection of the null hypothesis, indicating that the critical region should be the left-tailed. Let’s define $\pi_{X,Y}(\beta)$ as the power function under the X th type alternative and the Y th type sampling scheme. The power function under the “greater than” type of alternative for the time-truncated sampling should be such as:

$$\begin{aligned}
 \pi_{G,T}(\beta) &= P_{H_A}(Z < \chi_{1-\alpha}^2(2n)) \\
 &= \int_0^{\chi_{1-\alpha}^2(2n)} (\beta/\beta_0)^n \frac{1}{\Gamma(n)2^n} z^{n-1} e^{-\frac{\beta z}{2\beta_0}} dz \\
 &= \int_0^{\chi_{1-\alpha}^2(2n)} \beta^n \frac{1}{\Gamma(n)2^n} z^{n-1} e^{-\beta z/2} dz.
 \end{aligned} \tag{1.49}$$

Setting $\beta z = t \implies dt = \beta dz$, we simplify the power function such as:

$$\begin{aligned}\pi_{G,T}(\beta) &= \int_0^{\beta\chi_{1-\alpha}^2(2n)} \frac{1}{\Gamma(n)2^n} t^{n-1} e^{-\frac{t}{2}} dt \\ &= \Psi_{2n}(\beta\chi_{1-\alpha}^2(2n))\end{aligned}\tag{1.50}$$

where $\Psi_m(\cdot)$ is the c.d.f. of a chi-square distribution with m degrees of freedom.

Using the pivotal property of the quantity $2n\beta_0/\hat{\beta} = 2n/\hat{\beta} \sim \chi^2(2n)$ with $H_0 : \beta = 1$, we get the same power function as follows:

$$\begin{aligned}\pi_{G,T}(\beta) &= P_{H_A}(Z < \chi_{1-\alpha}^2(2n)) \\ &= P_{H_A}(2n/\hat{\beta} < \chi_{1-\alpha}^2(2n)) \\ &= P_{H_A}(2n\beta/\hat{\beta} < \beta\chi_{1-\alpha}^2(2n)) \\ &= \Psi_{2n}(\beta\chi_{1-\alpha}^2(2n)).\end{aligned}\tag{1.51}$$

The power functions under the “less than” type and the “two-tailed” alternatives for the time-truncated case take the forms as:

$$\pi_{L,T}(\beta) = 1 - \Psi_{2n}(\beta\chi_{\alpha}^2(2n)),\tag{1.52}$$

$$\pi_{T,T}(\beta) = \Psi_{2n}(\beta\chi_{1-\alpha/2}^2(2n)) + 1 - \Psi_{2n}(\beta\chi_{\alpha/2}^2(2n)).\tag{1.53}$$

Analogous power functions can be expressed for the failure-truncated case:

$$\pi_{G,F}(\beta) = \Psi_{2n-2}(\beta\chi_{1-\alpha}^2(2n-2)),\tag{1.54}$$

$$\pi_{L,F}(\beta) = 1 - \Psi_{2n-2}(\beta\chi_{\alpha}^2(2n-2)),\tag{1.55}$$

$$\pi_{T,F}(\beta) = \Psi_{2n-2}(\beta\chi_{1-\alpha/2}^2(2n-2)) + 1 - \Psi_{2n-2}(\beta\chi_{\alpha/2}^2(2n-2)).\tag{1.56}$$

The power curves under different types of alternatives are graphed for the different choices of the sample size n in the Figure 1.4.

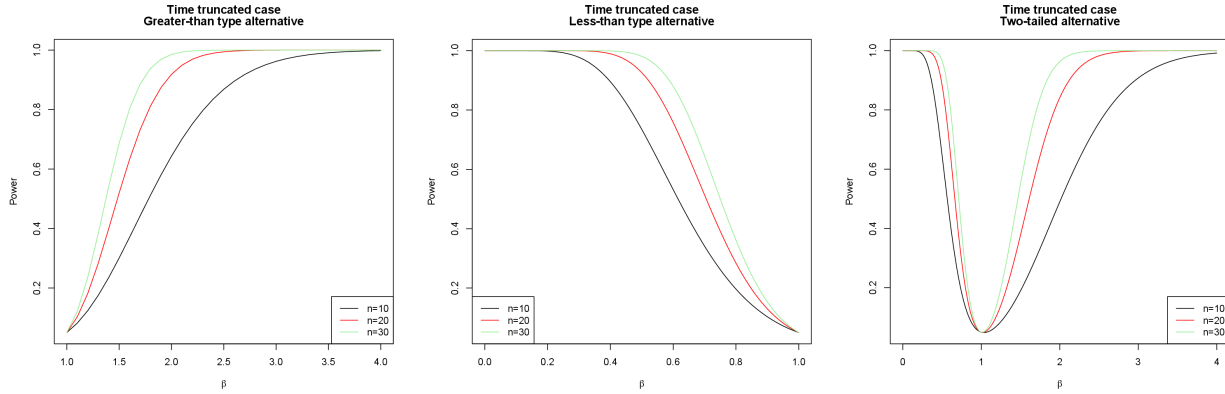


Figure 1.4. Power curves under different types of alternatives

1.2.3.4 UMP of the Forward Test Z

In this section, we prove the UMP (Uniformly Most Powerful) of the Z for the time-truncated sampling, which shows that the test gives greater power with the right alternative hypothesis than any others. There are two crucial notions, which are Monotone-likelihood ratio and Karlin-Rubin Theorem, to prove:

- Monotone-likelihood ratio (MLR): A family of p.d.f.'s or p.m.f.'s $\{g(t|\theta) : \theta \in \Theta\}$ for a univariate random variable T with real-valued parameter θ has a MLR if, $\forall \theta_2 > \theta_1, g(t|\theta_2)/g(t|\theta_1)$ is a monotone (nonincreasing or nondecreasing) function of t on $\{t : g(t|\theta_1) > 0 \cup g(t|\theta_2) > 0\}$.
-

Theorem 1.3. (Karlin-Rubin). Consider testing $H_0 : \theta \leq \theta_0$ versus $H_A : \theta > \theta_0$. Suppose that T is a sufficient statistic for θ and the family of p.d.f.'s or p.m.f.'s $\{g(t|\theta) : \theta \in \Theta\}$ 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 $\alpha = P_{\theta_0}(T > t_0)$.

Theorem 1.4. (UMP of the Z). For testing $H_0 : \beta = 1$ vs $H_A : \beta > 1$, the forward test using Z statistic is conditionally UMP.

Proof. From the forward statistic Z , we can formulate that:

$$Z = 2n\beta/\hat{\beta} = -2\beta \sum_{i=1}^n \ln(t_i/t) = -2 \sum_{i=1}^n \ln(t_i/t)^\beta \quad (1.57)$$

$$\implies -Z/2 = \sum_{i=1}^n \ln(t_i/t)^\beta \quad (1.58)$$

$$\implies e^{-Z/2} = \prod_{i=1}^n (t_i/t)^\beta. \quad (1.59)$$

To prove that the the Z in the time-truncated sampling is UMP, there are two steps. First, we need to prove that $\hat{\beta}$ is sufficient for β . The joint density of $t_1 < t_2 < \dots < t_n < t$ given $N = n$ is as follows:

$$f(t_1, t_2, \dots, t_n | n) = n! \prod_{i=1}^n \frac{\lambda(t_i)}{\mu(t)}, \quad 0 < t_1 < t_2 < \dots < t_n < t. \quad (1.60)$$

With the PLP of $\lambda(\cdot)$, the joint density is:

$$f(t_1, t_2, \dots, t_n | n) = n! \prod_{i=1}^n \frac{\beta}{t} \left(\frac{t_i}{t}\right)^{\beta-1}, \quad 0 < t_1 < t_2 < \dots < t_n < t. \quad (1.61)$$

Using the relation from (1.59), the joint density can be:

$$\begin{aligned} f(t_1, t_2, \dots, t_n | n) &= n! \prod_{i=1}^n \frac{\beta}{t} \left(\frac{t_i}{t}\right)^{\beta-1} \\ &= n! \prod_{i=1}^n \beta \left(\frac{t_i}{t}\right)^\beta \frac{1}{t_i} \\ &= n! \beta^n \prod_{i=1}^n \left(\frac{t_i}{t}\right)^\beta \prod_{i=1}^n \frac{1}{t_i} \\ &= n! \beta^n \exp[-z/2] \prod_{i=1}^n \frac{1}{t_i} \\ &= \{\beta^n \exp[-\frac{n\beta}{\hat{\beta}}]\} \{n! \prod_{i=1}^n \frac{1}{t_i}\}. \end{aligned} \quad (1.62)$$

After simplifying the above joint density, $\hat{\beta}$ is sufficient for β with the factorization $h(t_i) = n! \prod_{i=1}^n \frac{1}{t_i}$ and $g(\beta, \hat{\beta}) = \beta^n \exp[-\frac{n\beta}{\hat{\beta}}]$.

Next, in the second part, it is required to show that $\hat{\beta}$ has a monotone likelihood ratio with the real-valued parameter β . Using one to one transformation technique, the p.d.f. of $\hat{\beta}$ is from inverse

gamma $(n, n\beta)$ such as:

$$f_{\hat{\beta}}(\hat{\beta}) = \frac{1}{\Gamma(n)} \frac{(n\beta)^n}{\hat{\beta}^{n+1}} \exp\left[-\frac{n\beta}{\hat{\beta}}\right], \hat{\beta} > 0. \quad (1.63)$$

Choosing $\beta_2 > \beta_1$, the ratio could be:

$$\frac{f(\hat{\beta}|\beta_2)}{f(\hat{\beta}|\beta_1)} = \frac{\frac{1}{\Gamma(n)} \frac{(n\beta_2)^n}{\hat{\beta}^{n+1}} \exp\left[-\frac{n\beta_2}{\hat{\beta}}\right]}{\frac{1}{\Gamma(n)} \frac{(n\beta_1)^n}{\hat{\beta}^{n+1}} \exp\left[-\frac{n\beta_1}{\hat{\beta}}\right]} = \left(\frac{\beta_2}{\beta_1}\right)^n \exp\left[-\frac{n}{\hat{\beta}}(\beta_2 - \beta_1)\right] > 0. \quad (1.64)$$

Since $\beta_2 > \beta_1$, the ratio of $\frac{f(\hat{\beta}|\beta_2)}{f(\hat{\beta}|\beta_1)}$ is a monotone function of $\hat{\beta}$. This result also confirms that the density of $\hat{\beta}$ belongs to the exponential family. Thus, we can coincide with Bain and Engelhardt's result (1980) [8] that Z is UMP under the time-truncated sampling. Similar conclusions can be proved for the failure-truncated sampling (Bain and Engelhardt (1991) [9]). \square

1.2.4 Backward Test

In the previous section, we proved the forward statistic Z and showed its properties under the time-truncated sampling.

$$Z = 2n\beta/\hat{\beta} = -2\beta \sum_{i=1}^n \ln(t_i/t) = 2\beta \sum_{i=1}^n \ln(t/t_i). \quad (1.65)$$

This statistic Z describes in terms of the trend of failure in the time-truncated sampling. If the repairable system is deteriorating (i.e. $\beta > 1$) with increasing intensity, most of the t_i values tend to cluster around a predetermined time t . Consequently, $\frac{t_i}{t}$ values get closer to 1, leading to small values of the forward statistic Z . Under this deteriorating process, $1 - \frac{t_i}{t}$ values are close to 0. Thus, unlike the forward statistic Z , the following value

$$Z_B = -2\beta \sum_{i=1}^n \ln(1 - t_i/t) = 2\beta \sum_{i=1}^n \ln(t/(t - t_i)) \quad (1.66)$$

is inflated. A backward test statistic Z_B , which is a reversed version of Z , is called the backward test by Ho (1993) [49], where

$$Z_B = 2 \sum_{i=1}^{n-1} \ln(t_n/(t_n - t_i)), \text{ or} \quad (1.67)$$

$$Z_B = 2 \sum_{i=1}^n \ln(t/(t - t_i)), \quad (1.68)$$

respectively, for both sampling schemes under the null hypothesis $H_0 : \beta = 1$. This test was originally introduced by Ho (1993) [49] to provide the test which are more powerful than the forward test Z or the Laplace test L_P under the assumption of rough alternate intensities.

1.2.4.1 The Relationship between Two Versions of Tests

Theorem 1.5. *The forward and the backward test statistics Z and Z_B are related through:*

$$Z_B = -2\beta \sum_{i=1}^n \ln(t/t_i - 1) + Z. \quad (1.69)$$

Proof. (Method 1). The first method of proof is furnished through the following observations:

$$t_i/t = (t_i/t - t_i)(1 - t_i/t) = (1/(t/t_i - 1))(1 - t_i/t). \quad (1.70)$$

After taking the natural logarithm on the both sides, we get

$$\ln(t_i/t) = -\ln(t/t_i - 1) + \ln(1 - t_i/t). \quad (1.71)$$

Summing both sides and multiplying both sides by -2β :

$$-2\beta \sum_{i=1}^n \ln(t_i/t) = 2\beta \sum_{i=1}^n \ln(t/t_i - 1) - 2\beta \sum_{i=1}^n \ln(1 - t_i/t) \quad (1.72)$$

$$\implies -2\beta \sum_{i=1}^n \ln(1 - t_i/t) = -2\beta \sum_{i=1}^n \ln(t/t_i - 1) - 2\beta \sum_{i=1}^n \ln(t_i/t) \quad (1.73)$$

$$\implies Z_B = -2\beta \sum_{i=1}^n \ln(t/t_i - 1) + Z. \quad (1.74)$$

(Method 2). The second version of proof is done by following observations:

$$-\frac{Z}{2\beta} = \sum_{i=1}^n \ln(t_i/t) = \ln \prod_{i=1}^n (t_i/t) \quad (1.75)$$

$$\implies \prod_{i=1}^n (t_i/t) = e^{-\frac{Z}{2\beta}}, \quad (1.76)$$

and similarly

$$-\frac{Z_B}{2\beta} = \sum_{i=1}^n \ln(1 - t_i/t) = \ln \prod_{i=1}^n (1 - t_i/t) \quad (1.77)$$

$$\implies \prod_{i=1}^n (1 - t_i/t) = e^{-\frac{Z_B}{2\beta}}. \quad (1.78)$$

By dividing the above two, we obtain

$$\frac{e^{-\frac{Z}{2\beta}}}{e^{-\frac{Z_B}{2\beta}}} = \frac{\prod_{i=1}^n t_i/t}{\prod_{i=1}^n (1 - t_i/t)} \quad (1.79)$$

$$\implies e^{-\frac{1}{2\beta}(Z - Z_B)} = \prod_{i=1}^n \frac{t_i/t}{1 - t_i/t} = \prod_{i=1}^n t_i/(t - t_i) = \prod_{i=1}^n 1/(t/t_i - 1) \quad (1.80)$$

$$\implies -\frac{1}{2\beta}(Z - Z_B) = \ln\left(\prod_{i=1}^n 1/(t/t_i - 1)\right) = \sum_{i=1}^n \ln(1/(t/t_i - 1)) = -\sum_{i=1}^n \ln(t/t_i - 1) \quad (1.81)$$

$$\implies Z - Z_B = 2\beta \sum_{i=1}^n \ln(t/t_i - 1) \quad (1.82)$$

$$\implies Z_B = -2\beta \sum_{i=1}^n \ln(t/t_i - 1) + Z. \quad (1.83)$$

□

1.2.4.2 The Null and the Alternative Distributions of Z_B

Lemma 1.1. *If $X \sim \text{Beta}(m, n)$, then $E(X^{-\frac{1}{\beta}}) = \frac{B(n, m - \frac{1}{\beta})}{B(m, n)}$.*

Proof. The technique for a transformation of a random variable is used for the proof. Let's define

$Y = X^{-\frac{1}{\beta}} = g(X)$, then

$$X = Y^{-\beta} = g^{-1}(Y) \implies \frac{d}{dy}g^{-1}(Y) = -\beta Y^{-\beta-1} \quad (1.84)$$

$$f_X(x) = \frac{1}{B(m, n)} x^{m-1} (1-x)^{n-1}, \quad x \in (0, 1) \quad (1.85)$$

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \left| \frac{d}{dy}g^{-1}(y) \right| \\ &= \frac{1}{B(m, n)} (y^{-\beta})^{m-1} (1 - y^{-\beta})^{n-1} \beta y^{-\beta-1}, \quad y \in (1, \infty) \end{aligned} \quad (1.86)$$

$$\begin{aligned}
E(Y) &= \int_1^\infty y \frac{1}{B(m, n)} (y^{-\beta})^{m-1} (1 - y^{-\beta})^{n-1} \beta y^{-\beta-1} dy \\
&= \frac{\beta}{B(m, n)} \int_1^\infty y^{-\beta m} (1 - y^{-\beta})^{n-1} dy \\
&= \frac{B(n, m - \frac{1}{\beta})}{B(m, n)} = E(X^{-\frac{1}{\beta}}).
\end{aligned} \tag{1.87}$$

□

Theorem 1.6. *If $X^\beta \sim U(0, 1)$, then the density of $Y = (1 - X)^\beta$ is such as:*

$$f_Y(y) = y^{\frac{1}{\beta}-1} (1 - y^{\frac{1}{\beta}})^{\beta-1}, \quad y \in [0, 1]. \tag{1.88}$$

Proof. The proof is done by the technique for a transformation of a random variable with $X^\beta = (1 - Y^{\frac{1}{\beta}})^\beta = g^{-1}(Y)$.

$$\begin{aligned}
f_Y(y) &= f_{X^\beta}(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| \\
&= y^{\frac{1}{\beta}-1} (1 - y^{\frac{1}{\beta}})^{\beta-1}, \quad y \in (0, 1).
\end{aligned} \tag{1.89}$$

With the density of Y , the c.d.f. can be defined as:

$$F_Y(y) = P(Y \leq y) = \int_0^y u^{\frac{1}{\beta}-1} (1 - u^{\frac{1}{\beta}})^{\beta-1} du = 1 - (1 - y^{\frac{1}{\beta}})^{\beta-1}, \quad y \in (0, 1) \tag{1.90}$$

which is the c.d.f. of a Kumaraswamy-Generalized distribution by Pascoa et al. (2011) [27]. □

Theorem 1.7. *If $X^\beta \sim U(0, 1)$, then the density of $Z = -2 \ln(1 - X)^\beta$ is:*

$$f_Z(z) = \frac{1}{2} e^{-\frac{z}{2\beta}} (1 - e^{\frac{z}{2\beta}})^{\beta-1}, \quad z \in [0, \infty). \tag{1.91}$$

Proof. From the previous theorem, the density of $Y = (1 - X)^\beta$ is followed as:

$$f_Y(y) = y^{\frac{1}{\beta}-1} (1 - y^{\frac{1}{\beta}})^{\beta-1}, \quad y \in [0, 1]. \tag{1.92}$$

Applying the same technique with $Z = -2 \ln(1 - X)^\beta = -2 \ln(Y) \implies Y = e^{-\frac{z}{2}} = g^{-1}(Z)$, then the density of Z is given by:

$$\begin{aligned} f_Z(z) &= f_Y(g^{-1}(z)) \left| \frac{d}{dz} g^{-1}(z) \right| \\ &= \frac{1}{2} e^{-\frac{z}{2\beta}} (1 - e^{\frac{z}{2\beta}})^{\beta-1}, \quad z \in [0, \infty). \end{aligned} \quad (1.93)$$

□

Using $X = Y/t$, we can conclude that $-2 \ln(1 - Y/t)^\beta$ has the density given by (1.91). Thus, $-2 \ln(1 - t_i/t)^\beta$ are order statistics from a distribution of $-2 \ln(1 - Y/t)^\beta$. Applying this fact, then

$$Z_B = -2\beta \sum_{i=1}^n \ln(1 - t_i/t) = -2 \sum_{i=1}^n \ln(1 - t_i/t)^\beta \stackrel{d}{=} -2 \sum_{i=1}^n \ln(1 - Y_i/t)^\beta \quad (1.94)$$

where $\stackrel{d}{=}$ means that equal in distribution. Even it is difficult to find the density of Z_B , it is enough to find the distribution of the sum of the iid variables of $-2 \ln(1 - Y_i/t)^\beta$.

Theorem 1.8. *The backward statistic Z_B is non-pivotal.*

Proof. From the (1.94),

$$Z_B \stackrel{d}{=} -2 \sum_{i=1}^n \ln(1 - Y_i/t)^\beta = \sum_{i=1}^n Z_i \quad (1.95)$$

where $\{Z_i = -2 \ln(1 - Y_i/t)^\beta\}$'s are iid and each density is given by (1.91). The moment generating function (mgf) of Z_i is as follows:

$$M_{Z_i}(t) = E(e^{tZ_i}) = \frac{1}{2} \int_0^\infty e^{z(t-\frac{1}{2\beta})} (1 - e^{\frac{z}{2\beta}})^{\beta-1} dz. \quad (1.96)$$

Setting $u = 1 - e^{\frac{z}{2\beta}}$,

$$M_{Z_i}(t) = \beta \int_0^1 u^{\beta-1} (1 - u)^{-2t\beta} du = \beta B(\beta, 1 - 2t\beta) \quad (1.97)$$

which is only defined on $\{t : t < \frac{1}{2\beta}\}$. Using the iid of the $\{Z_i\}$'s, the mgf of Z_B is:

$$M_{Z_B}(t) = [\beta B(\beta, 1 - 2t\beta)]^n = \beta^n \left[\frac{\Gamma(\beta)\Gamma(1 - 2t\beta)}{\Gamma(1 + (1 - 2t)\beta)} \right]^n. \quad (1.98)$$

□

Since the mgf of Z_B contains β for the other choice of values not 1, the distribution of Z_B also contains β , thus the backward statistic is non-pivotal. Even this mgf is not a form to recognize, there is a method, called inverse Fourier transformation, to find the true density of Z_B .

Given the characteristic function $\phi_X(\cdot)$ of a random variable X such as:

$$\phi_X(t) = E(e^{itx}) = \int_{-\infty}^{\infty} e^{itx} f(x) dx, \quad (1.99)$$

we can find $f(\cdot)$ by an inverse Fourier transformation as:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t) dt. \quad (1.100)$$

Since the characteristic function and the mgf and are connected through:

$$\phi_X(t) = E(e^{itx}) = M_X(it), \quad (1.101)$$

the characterise function for Z_B should be:

$$\phi_{Z_B}(t) = M_{Z_B}(it) = \beta^n [B(\beta, 1 - 2it\beta)]^n. \quad (1.102)$$

Using an inverse Fourier transformation, the true density of Z_B is such as:

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \beta^n [B(\beta, 1 - 2it\beta)]^n dt. \quad (1.103)$$

- **The Special Case $\beta = 1$**

The mgf of Z_B for the time-truncated sampling has been derived in the previous section:

$$M_{Z_B}(t) = \beta^n [B(\beta, 1 - 2t\beta)]^n, \quad t < \frac{1}{2\beta}. \quad (1.104)$$

If $\beta = 1$ known as the constant or homogeneous failure pattern over time, then the mgf should be:

$$\begin{aligned} M_{Z_B}(t) &= [B(1, 1 - 2t)]^n \\ &= \left[\frac{\Gamma(1)\Gamma(1 - 2t)}{\Gamma(2 - 2t)} \right]^n \\ &= \left[\frac{\Gamma(1)\Gamma(1 - 2t)}{(1 - 2t)\Gamma(1 - 2t)} \right]^n \\ &= \left[\frac{1}{1 - 2t} \right]^n, \quad t < \frac{1}{2} \end{aligned} \quad (1.105)$$

which is the mgf of the chi-square distribution with $2n$ degrees of freedom. Thus, the distribution of backward statistic Z_B with $\beta = 1$, known as the homogeneity, is a chi-square distribution with $2n$ degrees of freedom for the time-truncated sampling.

Also, for the failure-truncated sampling of size n and under the null hypothesis $H_0: \beta = 1$, Z and Z_B , have the same distribution, and under this hypothesis, $Z_B \sim \chi^2(2n - 2)$. However, this property does not hold for other choices of β for Z_B . In other words, even though for the forward version, $\beta Z \sim \chi^2(2n - 2)$ for any $\beta > 0$, this pivotal is not shared by the backward version, and the chi-square distribution only holds for the backward version when $\beta = 1$ (Ho 1993 [49]; Bhaduri 2018 [14]). The null distribution yields the same result under either sampling scheme, if the sample is shifted by one size: $n = m - 1$, where $n =$ time-truncated sample size and $m =$ failure-truncated sample size.

1.2.4.3 Critical Regions

If the repairable system is deteriorating ($\beta > 1$), most of the t_i values tend to cluster around the t , so $(1 - t_i/t)$ values are close to 0. We already found that we reject the null hypothesis with small value of Z under the “greater than” type of alternative. We reject the null hypothesis for large value of the backward statistic Z_B under the “greater than” type of alternative. Since we proved that Z_B follows a chi-square distribution only when the null hypothesis $H_0 : \beta = 1$, the Table 1.2 shows the critical regions in the Z_B depending on the different alternative hypothesis under the different sampling schemes when $H_0 : \beta = 1$.

Table 1.2. Critical regions for Z_B under the different alternative hypothesis.

	$H_A : \beta > 1$	$H_A : \beta < 1$	$H_A : \beta \neq 1$
Time-truncated	$Z_B > \chi_{1-\alpha}^2(2n)$	$Z_B < \chi_{\alpha}^2(2n)$	$Z_B < \chi_{1-\frac{\alpha}{2}}^2(2n)$ or $Z_B > \chi_{\frac{\alpha}{2}}^2(2n)$
Failure-truncated	$Z_B > \chi_{1-\alpha}^2(2n - 2)$	$Z_B < \chi_{\alpha}^2(2n - 2)$	$Z_B < \chi_{1-\frac{\alpha}{2}}^2(2n - 2)$ or $Z_B > \chi_{\frac{\alpha}{2}}^2(2n - 2)$

For instance, one is interested in testing for the process: $H_0 : \beta = 1$ vs $H_a : \beta > 1$ based on a sample size 10. Since Z_B follows the chi-square distribution with 20 degrees of freedom for the time-truncated sampling under the null hypothesis $\beta_0 = 1$, the critical region should be $[28.4119, \infty]$ by exact chi-square calculations. However, this method is not useful for checking non-homogeneity values such as $\beta_0 = 2, 3, \dots$, since Z_B follows the chi-square distribution only when $H_0 : \beta = 1$. However, the critical region can be found by measuring the proportion ($\alpha = .1$) in excess of z^* . This method is the inversion method with the true density of Z_B to find z^* under $H_0 : \beta = 1$ by solving as follows:

$$\frac{1}{2\pi} \int_{z^*}^{\infty} \int_{-\infty}^{\infty} e^{-itz} 1^{10} B(1, 1 - 2it)^{10} dt dz = 0.1. \quad (1.106)$$

1.2.4.4 Power Function of Z_B -test

The critical region is found by the chi-square calculations or inversion method for a given null hypothesis. Under the “greater than” type of alternative, the power is the portion of observations generated under the alternative choice of β that exceeds the threshold. After deciding z^* for the critical region, the inversion method with the true density of Z_B under the alternative choice of β is used to calculate the power function such as:

$$\pi_{G,B}(\beta) = \frac{1}{2\pi} \int_{z^*}^{\infty} \int_{-\infty}^{\infty} e^{-itz} \beta^n B(\beta, 1 - 2it\beta)^n dt dz. \quad (1.107)$$

1.3 Bidirectional Testing

Although both Z and Z_B share a common chi-square distributions as their null distribution under $H_0 : \beta = 1$, the critical region is the left-tail for Z and right tail for Z_B in testing for a one-sided increasing trend, and vice versa for the decreasing trend. Additionally, even we’ve proved the Z ’s superiority under the PLP, there are asymmetrical performances between Z and Z_B (Ho 1993 [49]). For example, Z_B is more powerful than Z in detecting (a) an increasing two-step-intensity alternative

for the failure-truncated sampling, and (b) a decreasing two-step-intensity alternative for the time-truncated sampling. Since the deterioration inflates Z_B (or deflates Z) while the improvement inflates Z (or deflates Z_B), we introduce new tests with the optimal properties from both Z and Z_B to detect the non-stationarity (or non-homogeneity).

1.3.1 Right-Tailed Bidirectional Test

In the repairable systems, deterioration with the increasing intensity inflates Z_B , while improvement with the decreasing intensity inflates Z . Under the non-stationary systems, the maximum of these statistics is getting larger. For the sake of maintaining a broad two-sided alternative hypothesis, we pair each right leg of both tests to form a right-tailed bidirectional test, characterized by a right-tailed critical region. Since R takes the maximum value from the combination of Z and Z_B , the large value of R tends to reject the null hypothesis claiming homogeneity. In the context of testing non-homogeneity, with H_0 asserting otherwise, a maximum-based test is defined as:

Definition 1.7. *With $R = \text{Max}(Z, Z_B)$, the right-tailed bidirectional test is defined as:*

$$\phi(R) = \begin{cases} 1, & \text{if } R \geq c_\alpha^R \\ 0, & \text{otherwise} \end{cases} \quad (1.108)$$

where c_α^R is the upper $100\alpha^{th}$ percentile of the null distribution of R . This test is referred to as the R -test, or R throughout the remaining study.

1.3.2 Left-Tailed Bidirectional Test

In the same fashion, since L takes the minimum value from the combination of Z and Z_B , the small value of L leads to the rejection of the null hypothesis with a left-tailed critical region. In the context of testing non-homogeneity, with H_0 asserting otherwise, a minimum-based test is defined as:

Definition 1.8. *With $L = \text{Min}(Z, Z_B)$, the left-tailed bidirectional test is defined as:*

$$\phi(L) = \begin{cases} 1 & \text{if } L \leq c_\alpha^L \\ 0 & \text{otherwise} \end{cases} \quad (1.109)$$

where c_α^L is the lower $100\alpha^{th}$ percentile of the null distribution of L . This test is referred to as the L , or L throughout the remaining study. With the same principle of the R , the small value of L leads to reject the null hypothesis with a left-tailed critical region.

1.3.3 Dual Bidirectional Test

Another test could be comparing the p -values based on the bidirectional tests R and L . Choosing the smaller p -value between the maximum based on R and the minimum based on L , the H_0 asserting homogeneity is rejected with the extremely low p -value. This option, termed the dual bidirectional test is thus, formally defined as:

Definition 1.9. *With $P_{DB} = \text{Min}\{P_{H_0}(L \leq l), P_{H_0}(R \geq r)\}$, the dual bidirectional test is defined as:*

$$\phi(P_{DB}) = \begin{cases} 1 & \text{if } P_{DB} \leq p_\alpha \\ 0 & \text{otherwise} \end{cases} \quad (1.110)$$

where l and r are the observed test statistics, and $c_\alpha^{P_{DB}}$ is the lower $100\alpha^{th}$ percentile of the null distribution of P_{DB} . Since P_{DB} test takes the minimum value from the combination of p -values from both R and L (1.108) and L (1.109), the small value of P_{DB} leads to the rejection of the null hypothesis with a left-tailed critical region. This test is referred to as the P_{DB} -test or P_{DB} , because it has a p -value induced test statistic.

1.3.4 Empirical Null Distributions

The critical values (originally produced for the failure-truncated sampling in Bhaduri 2018 [14]) are extended to include that of the time-truncated sampling. The method of using order statistics is employed to simulate an HPP under time-truncated sampling. After generating an order statistics, $u_{(1)} < u_{(2)} < \dots < u_{(n)}$, from the uniform distribution $(0,1)$, the null distribution of Z and Z_B are

straightforward. The tabulated null bidirectional distributions furnish the critical values estimated by the simulation with size $N = 10^4$, ensuring numerical stability across different runs for a different choice of α and sample sizes up to $n = 50$. They are listed in Tables from 1.3 to 1.5 for R , L , and P_{DB} , respectively. Those critical values from the empirical null distributions are used to calculate the power which is the portion of observations generated under the alternative with the structure of the Monte Carlo simulations with 10^4 iterations. The critical values of R and L by simulation are close to the χ^2 with the right degrees of freedom, using of it for the approximation as well. Both sampling schemes yield the same results if the sample is shifted by one size, as now noted in the footnote of each table. Extensions to critical values with sample sizes over 50 are straightforward with today's computing technologies.

To check the accuracy of the results from the simulation studies, we also estimate the Monte Carlo error (MCE), defined as the standard deviation of the Monte Carlo estimator taken across hypothetical N repetitions of the simulation. Here, the critical values of P_{DB} at $\alpha = 0.1$ and $n = 40$ with different number of simulations are performed with the results summarized as (N , Mean, Median, MCE): (100, 0.0590, 0.0591, 0.0006), (500, 0.0591, 0.0591, 0.0006), and (1000, 0.0591, 0.0591, 0.0006), respectively, which confirm the coherences of the Monte Carlo method (MCM). Practices resembling the above accuracy checking will be repeated to guard the integrity of the MCM, should sharp dissonances arise.

Table 1.3. Empirical null distribution of the R -test.

n (or $m - 1$)	$\alpha = 0.5$	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
1	2.76259	3.79368	4.61742	6.01252	7.43771	8.82506	10.66636	12.00176
2	5.32610	6.73754	7.77333	9.49008	11.19715	12.79352	14.96972	16.58467
3	7.76303	9.43076	10.66238	12.62517	14.48170	16.27738	18.58517	20.27207
4	10.10980	11.98017	13.32651	15.47724	17.49580	19.43530	21.85535	23.59617
5	12.42576	14.48732	15.94396	18.27325	20.49064	22.50523	25.13755	27.03184
6	14.71607	16.93867	18.49171	20.96555	23.31582	25.50324	28.24446	30.18008
7	16.96605	19.34730	21.02629	23.63173	26.04287	28.29317	31.31948	33.46228
8	19.24758	21.76299	23.54114	26.32911	28.88441	31.28316	34.31901	36.31103
9	21.45358	24.10742	25.96012	28.85463	31.48720	34.01073	37.04510	39.45562
10	23.69674	26.50450	28.44189	31.45434	34.28740	36.86675	40.17500	42.55875
11	25.90015	28.81487	30.79092	33.91083	36.78415	39.50690	42.91206	45.49608
12	28.08460	31.10241	33.17450	36.39437	39.40271	42.17813	45.66175	47.96833
13	30.24182	33.38780	35.53193	38.80962	41.83210	44.58839	48.23849	51.13652
14	32.43075	35.67317	37.89541	41.31345	44.46291	47.32479	50.90457	53.54462
15	34.60049	37.93522	40.24877	43.79044	46.99693	50.00544	53.89229	56.48313
16	36.80468	40.22437	42.60411	46.20453	49.45632	52.59392	56.22777	58.90075
17	38.93818	42.47041	44.88308	48.59943	52.04321	55.22288	58.99215	61.81646
18	41.10030	44.70812	47.19624	51.07271	54.53403	57.76255	61.92054	64.68762
19	43.28436	46.97789	49.49510	53.39487	56.95178	60.38342	64.31317	67.23934
20	45.45902	49.24856	51.83492	55.75066	59.20902	62.56931	66.64971	69.64516
21	47.56828	51.42327	54.07548	58.20524	61.79274	65.12872	69.31541	72.22356
22	49.70588	53.66163	56.34347	60.44500	64.20394	67.66134	71.88456	74.93144
23	51.80661	55.82266	58.55526	62.79905	66.58829	70.03232	74.26518	77.18792
24	53.97184	58.05699	60.88109	65.18466	68.99439	72.51321	76.81869	79.93586
25	56.11324	60.25292	63.10763	67.48075	71.32257	74.98080	79.22916	82.22116
26	58.25655	62.52827	65.43137	69.82024	73.73753	77.39443	81.95604	85.30411
27	60.37751	64.70543	67.69596	72.17025	76.15421	79.88920	84.43537	87.84940
28	62.46744	66.85732	69.82941	74.34704	78.42661	82.23217	86.90191	90.14095
29	64.64753	69.12219	72.16090	76.81343	81.00318	84.84853	89.68073	93.10866
30	66.78896	71.38293	74.51746	79.21807	83.40181	87.18431	92.10368	95.58734
31	68.89859	73.51126	76.61667	81.44399	85.76024	89.49425	94.37618	97.70716
32	70.97500	75.66018	78.84264	83.67877	87.97682	91.95972	96.71393	99.98882
33	73.10887	77.82006	81.05650	85.92275	90.40613	94.36789	99.15488	102.95130
34	75.16808	79.97505	83.33152	88.24318	92.69073	96.89902	102.03886	105.73834
35	77.35576	82.18788	85.54640	90.64357	95.12778	99.27702	104.21867	107.83860
36	79.43211	84.41624	87.74260	92.81780	97.26202	101.64589	106.81243	110.48551
37	81.54294	86.51411	89.86794	95.01258	99.69488	104.16198	109.44412	113.13416
38	83.64238	88.76837	92.17712	97.31172	102.03225	106.28557	111.46263	115.14064
39	85.81631	90.90753	94.36576	99.55392	104.16692	108.48713	113.82813	117.84133
40	87.91300	93.13927	96.60852	101.95897	106.62420	110.92530	116.35325	120.22401
41	89.93419	95.15864	98.68305	104.14069	108.87683	113.30276	118.99165	122.82949
42	92.10702	97.40866	101.01474	106.43982	111.35621	115.78230	121.15330	125.36994
43	94.20144	99.58255	103.17588	108.61722	113.51957	117.87329	123.25148	127.05246
44	96.29968	101.67572	105.33692	110.85339	115.83471	120.36238	126.07313	130.35565
45	98.36778	103.77828	107.45827	113.11453	118.14202	122.77036	128.44119	132.56211
46	100.47670	105.96660	109.69420	115.38890	120.55440	125.13160	130.88120	134.89820
47	102.60970	108.12190	111.91870	117.64990	122.70920	127.44060	133.24370	137.46020
48	104.64490	110.25200	114.09700	119.82870	124.95200	129.78250	135.66940	139.92260
49	106.82860	112.42820	116.31220	122.14890	127.28050	132.04100	137.72580	141.51270
50	108.86510	114.60230	118.48180	124.29890	129.53450	134.17700	139.95830	144.02770

n = Time-truncated sample size
 m = Failure-truncated sample size

Table 1.4. Empirical null distribution of the L -test.

n (or $m - 1$)	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.025$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.5$
1	0.00496	0.00968	0.02440	0.04912	0.10148	0.20937	0.32514	0.57871
2	0.14113	0.20553	0.32787	0.48013	0.70458	1.05673	1.36202	1.91759
3	0.52924	0.67660	0.94625	1.23677	1.63307	2.20575	2.66604	3.45845
4	1.08364	1.34365	1.76883	2.19033	2.74536	3.49416	4.08466	5.06973
5	1.84232	2.17549	2.69929	3.26488	3.94926	4.86997	5.56764	6.74274
6	2.69344	3.09186	3.76884	4.42238	5.24687	6.31889	7.12695	8.43950
7	3.57607	4.07729	4.88337	5.63821	6.58139	7.78416	8.69264	10.15679
8	4.58167	5.15628	6.07308	6.92219	7.98760	9.31569	10.31460	11.91371
9	5.64919	6.31048	7.28921	8.21018	9.37776	10.87094	11.95484	13.69039
10	6.67236	7.39876	8.48837	9.57427	10.86845	12.43528	13.59541	15.45398
11	7.88192	8.62764	9.84200	11.00875	12.33747	14.04595	15.27930	17.24390
12	8.96052	9.81878	11.15279	12.40290	13.86817	15.68820	16.99069	19.05882
13	10.34489	11.24661	12.59560	13.87275	15.40217	17.31803	18.68749	20.87337
14	11.47277	12.41510	13.93401	15.32600	16.96005	18.99670	20.42902	22.68435
15	12.74278	13.77007	15.36507	16.81694	18.50790	20.60748	22.10096	24.48719
16	13.99977	15.09067	16.82381	18.34486	20.12679	22.30810	23.84769	26.36333
17	15.37219	16.51118	18.22243	19.74212	21.64856	23.94395	25.57070	28.12884
18	16.68265	17.83106	19.66802	21.28818	23.23193	25.63209	27.31804	30.01377
19	18.10879	19.34236	21.19211	22.88618	24.89359	27.33956	29.10000	31.85378
20	19.49000	20.74359	22.76034	24.52446	26.59117	29.08694	30.86245	33.70737
21	20.79040	22.08443	24.16476	25.98771	28.12703	30.73597	32.58790	35.52336
22	22.16999	23.63075	25.74368	27.57779	29.76956	32.46485	34.39625	37.38189
23	23.75310	25.17462	27.30697	29.22074	31.49282	34.27596	36.21837	39.27462
24	25.06207	26.48464	28.79441	30.80348	33.13472	35.98709	38.00466	41.16940
25	26.40665	27.84508	30.31999	32.35567	34.78848	37.69247	39.76566	42.99167
26	28.09580	29.59211	31.96157	33.99234	36.42509	39.37967	41.50782	44.84832
27	29.53220	31.12188	33.51560	35.55661	38.08687	41.17439	43.33050	46.72264
28	30.75435	32.44268	35.00445	37.21620	39.78666	42.97774	45.16957	48.60607
29	32.39192	33.97122	36.55021	38.78450	41.42280	44.65543	46.88708	50.44333
30	33.68649	35.49459	38.16219	40.44531	43.18725	46.43378	48.73811	52.35432
31	35.27916	37.06908	39.73772	42.12037	44.84337	48.18767	50.53781	54.21078
32	36.84649	38.59552	41.41324	43.84722	46.68289	50.01705	52.40115	56.14410
33	38.24502	40.05474	42.90363	45.41105	48.28918	51.77920	54.21380	58.02875
34	39.80278	41.71766	44.50803	47.03421	50.02725	53.62885	56.06333	59.92897
35	41.14301	43.12522	46.19119	48.74126	51.75320	55.38316	57.89626	61.83402
36	42.86949	44.86493	47.83983	50.46096	53.48904	57.12711	59.68710	63.66223
37	44.32329	46.38906	49.47411	52.11902	55.20546	58.92955	61.53197	65.56219
38	45.93260	47.96664	51.03559	53.77772	56.93415	60.70973	63.34384	67.44238
39	47.40184	49.67057	52.72153	55.44111	58.64378	62.48056	65.16311	69.33439
40	49.25394	51.28861	54.26215	57.13216	60.43461	64.29103	67.03257	71.24293
41	50.79032	52.80223	56.06614	58.90344	62.16971	66.06346	68.86692	73.13810
42	52.47493	54.36625	57.58985	60.53676	63.88144	67.79632	70.60799	74.98391
43	53.78684	56.04082	59.30166	62.29918	65.63380	69.72141	72.52166	76.92106
44	55.16547	57.40073	60.88929	63.90559	67.38559	71.49733	74.37956	78.82325
45	57.03045	59.37039	62.67875	65.72370	69.20876	73.33371	76.23899	80.77581
46	58.62665	60.84379	64.48524	67.47125	70.95868	75.09224	78.03343	82.65800
47	60.06928	62.38479	65.86486	68.96245	72.55007	76.88860	79.88194	84.54005
48	61.71044	64.14565	67.68884	70.79906	74.45834	78.74278	81.75321	86.43494
49	63.22715	65.67830	69.26954	72.49601	76.23223	80.55290	83.55064	88.31438
50	65.01923	67.31776	70.93106	74.18551	77.93719	82.39122	85.49748	90.25809

n = Time-truncated sample size
 m = Failure-truncated sample size

Table 1.5. Empirical null distribution of the P_{DB} -test.

n (or $m - 1$)	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.025$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$
1	0.00501	0.01001	0.02501	0.05001	0.10000	0.20000	0.30000	0.40000
2	0.00312	0.00639	0.01627	0.03318	0.06890	0.14239	0.21959	0.30046
3	0.00289	0.00595	0.01522	0.03124	0.06428	0.13385	0.20698	0.28242
4	0.00292	0.00589	0.01489	0.03048	0.06284	0.13021	0.20144	0.27504
5	0.00283	0.00561	0.01458	0.03003	0.06195	0.12827	0.19856	0.27215
6	0.00273	0.00555	0.01435	0.02934	0.06074	0.12705	0.19689	0.26944
7	0.00278	0.00562	0.01421	0.02912	0.06047	0.12660	0.19575	0.26771
8	0.00274	0.00559	0.01420	0.02917	0.06009	0.12538	0.19467	0.26645
9	0.00279	0.00562	0.01433	0.02917	0.05997	0.12551	0.19405	0.26577
10	0.00280	0.00553	0.01410	0.02910	0.06008	0.12524	0.19384	0.26556
11	0.00272	0.00554	0.01413	0.02899	0.05996	0.12471	0.19311	0.26519
12	0.00274	0.00555	0.01431	0.02908	0.05978	0.12432	0.19247	0.26391
13	0.00274	0.00554	0.01414	0.02903	0.05972	0.12459	0.19252	0.26413
14	0.00272	0.00548	0.01395	0.02860	0.05922	0.12415	0.19207	0.26266
15	0.00275	0.00552	0.01416	0.02891	0.05935	0.12441	0.19258	0.26426
16	0.00273	0.00552	0.01410	0.02905	0.05947	0.12424	0.19144	0.26274
17	0.00272	0.00551	0.01402	0.02870	0.05943	0.12390	0.19161	0.26315
18	0.00271	0.00556	0.01414	0.02897	0.05959	0.12374	0.19108	0.26241
19	0.00267	0.00542	0.01392	0.02874	0.05924	0.12344	0.19075	0.26153
20	0.00265	0.00543	0.01389	0.02855	0.05889	0.12332	0.19074	0.26166
21	0.00270	0.00549	0.01403	0.02877	0.05924	0.12369	0.19102	0.26138
22	0.00268	0.00540	0.01406	0.02875	0.05928	0.12327	0.19110	0.26228
23	0.00265	0.00552	0.01403	0.02868	0.05914	0.12321	0.19055	0.26145
24	0.00276	0.00558	0.01398	0.02872	0.05916	0.12316	0.19107	0.26169
25	0.00269	0.00548	0.01398	0.02885	0.05932	0.12344	0.19104	0.26169
26	0.00269	0.00541	0.01382	0.02832	0.05911	0.12315	0.19072	0.26143
27	0.00267	0.00543	0.01383	0.02850	0.05917	0.12368	0.19134	0.26165
28	0.00275	0.00555	0.01410	0.02873	0.05926	0.12343	0.19088	0.26122
29	0.00265	0.00544	0.01403	0.02864	0.05915	0.12367	0.19186	0.26222
30	0.00264	0.00546	0.01409	0.02875	0.05930	0.12326	0.19089	0.26059
31	0.00271	0.00549	0.01419	0.02889	0.05952	0.12332	0.19075	0.26194
32	0.00275	0.00548	0.01406	0.02866	0.05919	0.12288	0.19049	0.26073
33	0.00271	0.00550	0.01398	0.02859	0.05918	0.12334	0.19099	0.26148
34	0.00270	0.00547	0.01389	0.02864	0.05911	0.12341	0.19060	0.26182
35	0.00267	0.00544	0.01390	0.02874	0.05926	0.12382	0.19129	0.26192
36	0.00268	0.00546	0.01394	0.02854	0.05913	0.12281	0.19001	0.26022
37	0.00269	0.00547	0.01395	0.02858	0.05926	0.12361	0.19102	0.26098
38	0.00269	0.00546	0.01389	0.02839	0.05886	0.12303	0.19035	0.26103
39	0.00267	0.00539	0.01390	0.02851	0.05885	0.12266	0.19003	0.26093
40	0.00270	0.00546	0.01402	0.02873	0.05923	0.12362	0.19086	0.26106
41	0.00268	0.00549	0.01404	0.02866	0.05905	0.12293	0.19033	0.26081
42	0.00272	0.00545	0.01387	0.02844	0.05905	0.12347	0.19072	0.26096
43	0.00267	0.00543	0.01403	0.02880	0.05905	0.12292	0.19051	0.26076
44	0.00271	0.00551	0.01405	0.02868	0.05879	0.12286	0.19043	0.26107
45	0.00274	0.00550	0.01405	0.02877	0.05919	0.12298	0.19054	0.26130
46	0.00273	0.00551	0.01405	0.02864	0.05897	0.12328	0.19041	0.26074
47	0.00271	0.00546	0.01398	0.02859	0.05872	0.12303	0.19037	0.26075
48	0.00268	0.00546	0.01398	0.02859	0.05875	0.12261	0.19029	0.26083
49	0.00267	0.00546	0.01396	0.02837	0.05853	0.12269	0.19011	0.26014
50	0.00270	0.00549	0.01395	0.02869	0.05915	0.12297	0.19087	0.26103

n = Time-truncated sample size
 m = Failure-truncated sample size

CHAPTER 2

AN INTELLIGIBLE VERSION

2.1 Motivation

Simulation studies are an important tool for statistical research, aimed better to understand the behavior and performance of the calculated statistics. So far, it would seem that the main deliverable is tabulated empirical null distributions for test statistics R , L , and P_{DB} (Bhaduri 2018 [14]), which are estimated by simulations with size $N = 10^4$, ensuring numerical stability across different runs depending on the value of α and sample size n . However, it is old-fashioned and limited by the range of n -values up to 50 for both sampling schemes. Folks don't use those tables to calculate p -values and powers anymore. Rather, they expect a distribution function like `pnorm` in R (e.g., Kerns 2010 [63]), or similar. This Chapter reveals that theory would showcase the way to calculate p -values and powers with the bidirectional tests not only to overcome the hurdles but also to supercharge the desired transformation throughout with the power: simplicity and practicality.

2.2 Propositions

The theoretical functional forms quantifying the dependence between the tests and that of the complete analytical approach are yet to be explored in bidirectional tests, R , L , and P_{DB} . We set out to explore the possibility of the option for the practitioners, because the distributions of the order statistics and p -values are known. Additional steps in their development for testing H_0 : The intensity $\lambda(\cdot)$ is constant versus H_A : $\lambda(\cdot)$ is not constant (increasing or decreasing) are taken as follows:

Proposition 2.1. *(a) The null distribution of $R = \text{Max}(Z, Z_B) \approx X_{(2)}$, and (b) the null distribution of $L = \text{Min}(Z, Z_B) \approx X_{(1)}$, where $X_{(\cdot)}$ are the order statistics of a random sample of size 2 from*

a $\chi^2(k)$ with $k = 2(m - 1)$ or $2n$, respectively, for a failure-truncated sampling of size m and n for that of a time-truncated sampling.

Proposition 2.2. *The null distribution of $P_{DB} = \text{Min}(P_{H_0}(L \leq l), P_{H_0}(R \geq r)) \approx U_{(1)}$, where $U_{(1)}$ is the first order statistic of a random sample of size 2 from $U(0, 1)$.*

2.3 The p -value Proxies

2.3.1 For the R -Test

Both Z and Z_B share a common chi-square distributions as their null distribution under the $H_0 : \beta = 1$. The null distribution of $R = \text{Max}(Z, Z_B) \approx X_{(2)}$ is the maximum order statistic of a random sample of size 2 from a $\chi^2(k)$ with $k = 2(m - 1)$ or $2n$, respectively, for a failure-truncated sampling of size m and n for that of a time-truncated sampling. If one is interested in finding the critical value for $R = \text{Max}(Z, Z_B)$ with $\alpha = .1$ by this method,

$$\begin{aligned}
 1 - F_R(r) &= 1 - P(R \leq r) \\
 &= 1 - P(\text{Max}(Z, Z_B) \leq r) \\
 &\approx 1 - P(Z \leq r)P(Z_B \leq r) \\
 &= 1 - P(\chi^2 \leq r)P(\chi^2 \leq r) \\
 &= 1 - (P(\chi^2 \leq r))^2 \\
 &= 1 - (F_{\chi^2(k)}(r))^2 = 0.1 \\
 \implies (F_{\chi^2(k)}(r))^2 &= 0.9 \\
 \implies F_{\chi^2(k)}(r) &= \sqrt{0.9}.
 \end{aligned} \tag{2.1}$$

From the last equation, we can solve the equation to find the critical value r by the chi-square table or the software R. Similar way is employed to calculate the p -value for $R = \text{Max}(Z, Z_B)$. The

p -value of $R = r$ is

$$\begin{aligned}
P_{H_0}(R \geq r) &= 1 - P_{H_0}(R < r) \\
&= 1 - P_{H_0}(\text{Max}(Z, Z_B) < r) \\
&\approx 1 - [F_{\chi^2(k)}(r)]^2,
\end{aligned} \tag{2.2}$$

with $k = 2(m - 1)$ or $2n$, respectively, for a failure-truncated sampling of size m and n for that of a time-truncated sampling.

2.3.2 For the L -Test

Both Z and Z_B share a common chi-square distributions as their null distribution under the $H_0 : \beta = 1$. The null distribution of $L = \text{Min}(Z, Z_B) \approx X_{(1)}$ is the minimum order statistic of a random sample of size 2 from a $\chi^2(k)$ with $k = 2(m - 1)$ or $2n$, respectively, for a failure-truncated sampling of size m and n for that of a time-truncated sampling. If one is interested in finding the critical value for $L = \text{Min}(Z, Z_B)$ with $\alpha = .1$ by this method,

$$\begin{aligned}
F_L(l) &= P(L \leq l) \\
&= 1 - P(L > l) \\
&= 1 - P(\text{Min}(Z, Z_B) > l) \\
&\approx 1 - P(Z > l)P(Z_B > l) \\
&= 1 - P(\chi^2 > l)P(\chi^2 > l) \\
&= 1 - (P(\chi^2 > l))^2 \\
&= 1 - (1 - P(\chi^2 \leq l))^2 \\
&= 1 - (1 - F_{\chi^2(k)}(l))^2 = 0.1 \\
\implies (1 - F_{\chi^2(k)}(l))^2 &= 1 - 0.1 = 0.9 \\
\implies 1 - F_{\chi^2(k)}(l) &= \sqrt{0.9} \\
\implies F_{\chi^2(k)}(l) &= 1 - \sqrt{0.9}
\end{aligned} \tag{2.3}$$

From the last equation, we can solve the equation to find the critical value l by the chi-square table or the software R. Same technique is used to calculate the p -value for $L = \text{Min}(Z, Z_B)$. The p -value of $L = l$ is

$$\begin{aligned}
P_{H_0}(L \leq l) &= 1 - P_{H_0}(L > l) \\
&= 1 - P_{H_0}(\text{Min}(Z, Z_B) > l) \\
&\approx 1 - [1 - F_{\chi^2(k)}(l)]^2,
\end{aligned} \tag{2.4}$$

with $k = 2(m - 1)$ or $2n$, respectively, for a failure-truncated sampling of size m and n for that of a time-truncated sampling.

2.3.3 For the P_{DB} -Test

Under any circumstances, neither of the bidirectional tests is free from having a pair of two undesirable low power (bad) legs, as per the symmetry/asymmetry properties. Of course, in any setting, the odds are in favor of having at least one high power (good) leg in each pair. For R and L , the existence of having the best or worst combination of the legs in detecting the same alternative would definitely form a striking contrast in power comparison, where the extent is yet to be confirmed. An intelligible approach is to collectively treat R and L as a transit to a test that can simultaneously filter out the bad and enhance the good reflected by each of the bidirectional critical regions. This leads us to compare the p -values from both the maximum-based R and the minimum-based L , and to reject the homogeneity assumption for extremely low values of the minimum p -value.

Definition 2.1. *The p -value is the probability of obtaining results as extreme as the observed results of a statistical hypothesis test, assuming that the null hypothesis is correct.*

Theorem 2.1. *The p -value follows the uniform distribution $(0, 1)$.*

Proof. Consider a left-sided one-tailed hypothesis test. Then, the p -value is a function of the test statistic $P = F_X(X)$ and $p = F_X(x)$, where x is the observed test statistic and $F_X(x)$ is the c.d.f. of the test statistic under the null hypothesis.

Then, we can obtain the c.d.f. of the p -value as

$$\begin{aligned}
 F_P(p) &= P(P \leq p) \\
 &= P(F_X(X) \leq p) \\
 &= P(X \leq F_X^{-1}(p)) \\
 &= F_X(F_X^{-1}(p)) \\
 &= p
 \end{aligned} \tag{2.5}$$

Consider a right-sided one-tailed hypothesis test. Then, the p -value is a function of the test statistic $P = 1 - F_X(X)$ and $p = 1 - F_X(x)$, where x is the observed test statistic and $F_X(x)$ is the c.d.f. of the test statistic under the null hypothesis.

Then, we can obtain the c.d.f. of the p -value as

$$\begin{aligned}
 F_P(p) &= P(P \leq p) \\
 &= P(1 - F_X(X) \leq p) \\
 &= P(-F_X(X) \leq p - 1) \\
 &= P(F_X(X) > 1 - p) \\
 &= 1 - P(F_X(X) \leq 1 - p) \\
 &= 1 - P(X \leq F_X^{-1}(1 - p)) \\
 &= 1 - F_X(F_X^{-1}(1 - p)) \\
 &= 1 - (1 - p) \\
 &= 1 - 1 + p \\
 &= p
 \end{aligned} \tag{2.6}$$

Consider a two-sided hypothesis test. Then, the p -value is $2\text{Min}\{P(X \leq x), P(X \geq x)\}$. If $\text{Min} = P(X \leq x)$, the p -value is a function of the test statistic $P = 2F_X(X)$ and $p = 2F_X(x)$, where x is the observed test statistic and $F_X(x)$ is the c.d.f. of the test statistic under the null hypothesis.

Then, we can obtain the c.d.f. of the p -value as

$$\begin{aligned}
2F_P(p) &= 2P(P \leq p) \\
&= 2P(2F_X(X) \leq p) \\
&= 2P(F_X(X) \leq p/2) \\
&= 2P(X \leq F_X^{-1}(p/2)) \\
&= 2F_X(F_X^{-1}(p/2)) \\
&= 2p/2 \\
&= p
\end{aligned} \tag{2.7}$$

If $Min = P(X \geq x)$, then, the p -value is a function of the test statistic $P = 2(1 - F_X(X))$ and $p = 2(1 - F_X(x))$, where x is the observed test statistic and $F_X(x)$ is the c.d.f. of the test statistic under the null hypothesis.

Then, we can obtain the c.d.f. of the p -value as

$$\begin{aligned}
2F_P(p) &= 2P(P \leq p) \\
&= 2P(2(1 - F_X(X)) \leq p) \\
&= 2P(-F_X(X) \leq p/2 - 1) \\
&= 2P(F_X(X) > 1 - p/2) \\
&= 2(1 - P(F_X(X) \leq 1 - p/2)) \\
&= 2(1 - P(X \leq F_X^{-1}(1 - p/2))) \\
&= 2(1 - F_X(F_X^{-1}(1 - p/2))) \\
&= 2(1 - (1 - p/2)) \\
&= p
\end{aligned} \tag{2.8}$$

Thus, the p -value follows the uniform distribution $(0, 1)$. □

Theorem 2.2. *If two independent random variables X and Y follow the uniform distribution $(0, 1)$, the minimum of these two variables follows the beta distribution $(1, 2)$.*

Proof. Let's define $U_{(1)} = \text{Min}(U_1, U_2)$.

$$\begin{aligned}
P(U_{(1)} \geq x) &= P(\text{Min}(U_1, U_2) \geq x) \\
&= P(U_1 \geq x, U_2 \geq x) \\
&\approx P(U_1 \geq x)P(U_2 \geq x) \\
&= (1-x)(1-x) = (1-x)^2 \\
\implies F_{U_{(1)}}(x) &= P(U_{(1)} \leq x) = 1 - (1-x)^2 = 2x - x^2 \\
\implies f_{U_{(1)}}(x) &= \frac{d}{dx}F_{U_{(1)}}(x) = \frac{d}{dx}(2x - x^2) = 2 - 2x, \quad x \in (0, 1)
\end{aligned} \tag{2.9}$$

We can confirm that the minimum of two independent variables from the uniform distribution $(0, 1)$ follows the beta distribution $(1, 2)$ proved by the c.d.f. and it's p.d.f. \square

Also, using the order statistics, the c.d.f. of $U_{(1)} = \text{Min}(U_1, U_2)$, where $U_{(1)}$ is the minimum order statistic of a random sample of size 2 from $U(0, 1)$, is

$$\begin{aligned}
F_{U_{(1)}}(x) &= P(U_{(1)} \leq x) \\
&= 1 - P(U_{(1)} > x) \\
&\approx 1 - P(U_1 > x)P(U_2 > x) \\
&= 1 - \prod_{i=1}^2 P(U_i > x) \\
&= 1 - \prod_{i=1}^2 (1 - P(U_i \leq x)) \\
&= 1 - \prod_{i=1}^2 (1 - x) \\
&= 1 - (1-x)^2 = 2x - x^2
\end{aligned} \tag{2.10}$$

Proving that the minimum of two independent variables from the uniform distribution $(0, 1)$ is beta distribution $(1, 2)$ in the Theorem (2.2), we can find the critical value in P_{DB} test by solving the quadratic equation of the theoretical c.d.f. of beta distribution $(1, 2)$ or using the order statistics. After solving it, the correct critical value should be between 0 and 1 due to the domain of beta

distribution. For example, if one wants to find the critical value for P_{DB} test with $\alpha = .1$,

$$\begin{aligned}
 F_{P_{DB}}(p) &= P(P_{DB} \leq p) = 2p - p^2 = 0.1 \\
 &\implies 2p - p^2 - 0.1 = 0 \\
 &\implies p = 0.0513, 1.9486.
 \end{aligned}$$

Due to the proper domain of the beta distribution, the critical value should be 0.02532 with $\alpha = .1$.

Same technique is used to calculate the p -value for $P_{DB} = \text{Min}(P_{H_0}(L \leq l), P_{H_0}(R \geq r))$. The p -value of $P_{DB} = p$ is

$$\begin{aligned}
 P_{H_0}(P_{DB} \leq p) &= 1 - P_{H_0}(P_{DB} > p) \\
 &= 1 - P_{H_0}(\text{Min}(P_{H_0}(L \leq l), P_{H_0}(R \geq r)) > p) \\
 &= 1 - P_{H_0}(\text{Min}(P_{H_0}(L \leq l), P_{H_0}(R \geq r)) > p) \tag{2.11} \\
 &\approx 1 - [1 - F_{U(0,1)}(p)]^2 \\
 &= 1 - (1 - p)^2.
 \end{aligned}$$

2.4 Performance Assessments

The realization of the above propositions becomes Table 2.1, where the cumulative failure times, t_i 's and t are recognized as the main events, and $F(\cdot)$ represents the lower-tail probability of the distribution. In essence, Table 2.1 has so many simple yet profound implications. The functions are not abstract – rather, grounded in practicality. For instance, an observed test statistic, p , of P_{DB} yields an associated p -value proxy $\approx 1 - (1 - p)^2$ with reasonable precision, which, in turn, can be inverted for obtaining a critical value $\approx 1 - \sqrt{(1 - \alpha)}$ for a size α hypothesis testing. Recall that $P_{DB} = \text{Min}(P_{H_0}(L \leq l), P_{H_0}(R \geq r))$, a p -value induced dual bidirectional test statistic, apparently possesses the unique feature of having proxies that are free of either the sample size or the sampling schemes, which will be factored into the development as the study unfolds.

Table 2.1. Functions for test statistics (Z and Z_B) and the p -value proxies (R , L , and P_{DB}), failure-truncated and time-truncated sampling with size n .

	Failure-truncated	Time-truncated
$Z = z$	$z = 2 \sum_{i=1}^{n-1} \ln(t_n/t_i)$	$z = 2 \sum_{i=1}^n \ln(t/t_i)$
$Z_B = z_B$	$z_B = 2 \sum_{i=1}^{n-1} \ln(t_n/(t_n - t_i))$	$z_B = 2 \sum_{i=1}^n \ln(t/(t - t_i))$
p -value for $R = r$	$1 - [F_{\chi^2(2n-2)}(r)]^2$	$1 - [F_{\chi^2(2n)}(r)]^2$
p -value for $L = l$	$1 - [1 - F_{\chi^2(2n-2)}(l)]^2$	$1 - [1 - F_{\chi^2(2n)}(l)]^2$
p -value for $P_{DB} = p$	$1 - (1 - p)^2$	$1 - (1 - p)^2$

Graphical comparisons between the empirical cumulative probabilities of the null distribution ($n = 21$, failure-truncated sampling) tabulated in the Tables from 1.3 to 1.5 and their proxies using the order statistics are presented in Figure 2.1. To further assess the utilities of the proxies, samples of performance data are summarized in Table 2.2, measured by the mean absolute deviations. Surprisingly, results surrogated by the order statistics reveal that the degree of dependence between the test statistics hindering a complete analytical approach for obtaining the p -values is generally minor, especially for moderate to small p -values where a statistical significance threshold is normally set. Consequently, this user-friendly version of the test makes the p -values much more attainable.

Table 2.2. Mean absolute deviation (variance) between tabulated empirical cumulative probabilities and their proxies for various sample sizes, failure-truncated sampling.

	$n = 3$	$n = 10$	$n = 21$	$n = 60$
$F_{P_{DB}}(x) \leq 0.100$	0.0072 (9.17e-05)	0.0042 (3.37e-05)	0.0033 (2.57e-05)	0.0043 (3.21e-05)
Overall	0.0398 (2.03e-03)	0.0249 (8.26e-04)	0.0263 (9.27e-04)	0.0240 (8.00e-04)
$F_R(x) \geq 0.900$	0.0005 (3.97e-07)	0.0007 (4.63e-04)	0.0006 (9.94e-07)	0.0007 (1.21e-06)
Overall	0.0156 (3.93e-04)	0.0155 (3.80e-04)	0.0153 (3.73e-04)	0.0156 (3.84e-04)
$F_L(x) \leq 0.100$	0.0006 (7.15e-07)	0.0003 (2.61e-07)	0.0008 (1.02e-06)	0.0006 (8.45e-07)
Overall	0.0173 (4.95e-04)	0.0171 (4.72e-04)	0.0172 (4.85e-04)	0.0166 (4.41e-04)

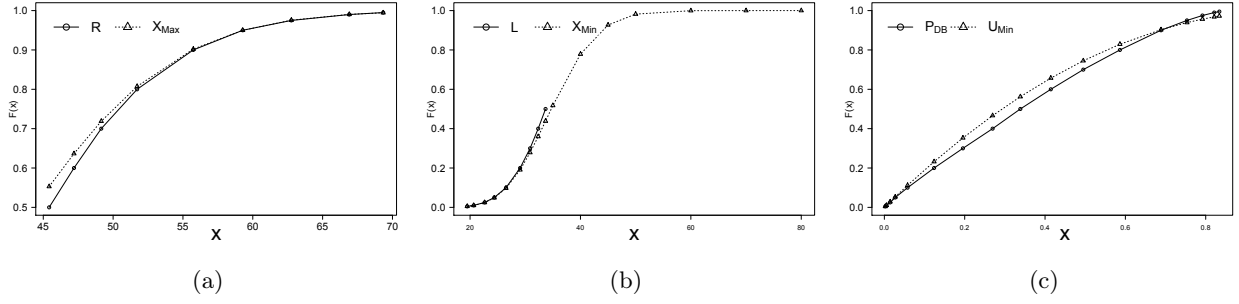


Figure 2.1. Pairwise comparison of the empirical null distribution ($n = 21$, failure-truncated sampling) and the corresponding order statistic for (a) R and X_{max} , (b) L and X_{min} , and (c) P_{DB} and U_{min} .

CHAPTER 3

POWER COMPARISONS

3.1 Motivation

Series of events come in different forms – one of the most sought-after and analyzed data imaginable. Continual assessment of new product reliability and ongoing control of reliability are crucial necessities in today’s competitive arena. Thus, the models for the repairable systems must be able to describe the occurrence of events in time by a point process, which tracks the number of occurrences of events (Cox 1955 [21]; Cox and Lewis 1966 [22]; Ascher 1984 [5]; Engelhardt and Bain 1987 [32]; Ho 1998 [51]; Limnios and Nikulin 2000 [66]; Rigdon and Basu 2000 [81]; Gertsbakh 2005 [39]; Epstein 2008 [34]; Tobias 2011 [90]; Modarres and Kaminskiy 2016 [70]; Cha and Finkelstein 2018 [16]; McPherson 2019 [69]; Rai 2020 [79]; Signoret and Leory 2021 [86]). A critical aspect of engineering applications is knowing if the system’s reliability has grown, decreased, or remained stable during the period. Therefore, the cumulative failure times are recognized as the main series of events.

Along these lines, the Poisson process is a classical point process, supporting a growing body of theories, methods, and applications. It may be reasonable to assume that the intensity, $\lambda(\cdot)$, is constant, so tests of $H_0: \lambda(\cdot)$ is constant versus $H_A: \lambda(\cdot)$ is not constant (increasing or decreasing) are of interest. The results of such tests could indicate whether the simple HPP may be adequate or whether a more general NHPP is required in modeling the occurrences of natural phenomena (Ho 1990 [46], 1991 [47], 1992 [48], 1996 [50]; Ho and Bhaduri 2017 [53]; Amei et al. 2012 [4]; Tan et al. 2014 [88]). Outcomes based on risk and hazard assessment studies, often modeled by an NHPP, may have long-lasting effects on humankind in one aspect or another (Ho 1991 [47], 1992 [48], 2010 [52]; Ho and Smith 1997 [54], 1998 [55]; Ho et al. 1991 [56], Ho et al. 2006 [57]; Ho et al. 2016 [58]).

By the simulating the events, as per Ho (1993) [49], there are asymmetrical and symmetrical

performances between Z and Z_B . First, for the asymmetry, Z_B is more powerful than Z in detecting (a) an increasing two-step-intensity alternative for the failure-truncated sampling, and (b) a decreasing two-step-intensity alternative for the time-truncated sampling. In other words, both legs (right and left, say) from Z and its Z_B counterpart set to do the same job for a specific alternative setting differ in power. As for the symmetrical property, it is easy to show that these two tests are equal in performance by properly allocating pairs of symmetrical change-points and trends (decreasing versus increasing) of the step-function intensities, because Z_B is a reversed version of Z . Consequently, both legs of either test are valuable, which rationalize the urge to compare the power with bidirectional tests.

3.2 Simulation Techniques

3.2.1 Power Law Process

This section explains how to simulate the data from a PLP. There are several ways to simulate events from a PLP under time-truncated sampling. Order statistics from the uniform distribution $(0, 1)$ are employed to simulate an NHPP with a power law intensity. After generating the order statistics, $u_{(1)} < u_{(2)} < \dots < u_{(n)}$, from a random sample of size n from the uniform distribution $(0, 1)$, the NHPP can be obtained through the solving $\mu(t_i)/\mu(t) = u_{(i)}$ where $\mu(t_i) = \int_0^{t_i} \lambda(x)dx = \int_0^{t_i} \frac{\beta}{\theta} (\frac{x}{\theta})^{\beta-1} dx = (\frac{t_i}{\theta})^\beta$ and $\mu(t) = (\frac{t}{\theta})^\beta$. Since the distribution of t_i/t , given n , does not depend on t , we set $t = 1$ and get the NHPP:

$$t_1 = u_{(1)}^{\frac{1}{\beta}}, t_2 = u_{(2)}^{\frac{1}{\beta}}, \dots, t_n = u_{(n)}^{\frac{1}{\beta}}. \quad (3.1)$$

After simulating the events, we can check that this process is flexible enough to model the repairable systems which are improving (i.e. $\beta < 1$), deteriorating (i.e. $\beta > 1$), or remaining homogeneous (i.e. $\beta = 1$) with respect to time t . For example, if $\beta > 1$, the derivative of the

intensity function $\frac{d\lambda(t|\Theta)}{dt}$ is always positive,

$$\frac{d}{dt}(\beta/\theta)(t/\theta)^{\beta-1} = (\beta/\theta^2)(\beta-1)(t/\theta)^{\beta-2} > 0. \quad (3.2)$$

This indicates that the failure process gradually gets more intense over time, known as the deteriorating since $\lambda(t|\Theta)$ keeps increasing. If $\beta < 1$, failures are getting less prevalent with the advent of time, indicating a hallmark of an improving system. Lastly, when $\beta = 1$, it is an HPP, implying that the system is neither improving nor wearing out over the time, but is more or less uniformly spread out with no apparent clustering tendency. The Figure 3.1 and Figure 3.2 display the different patterns depending on the values of β .

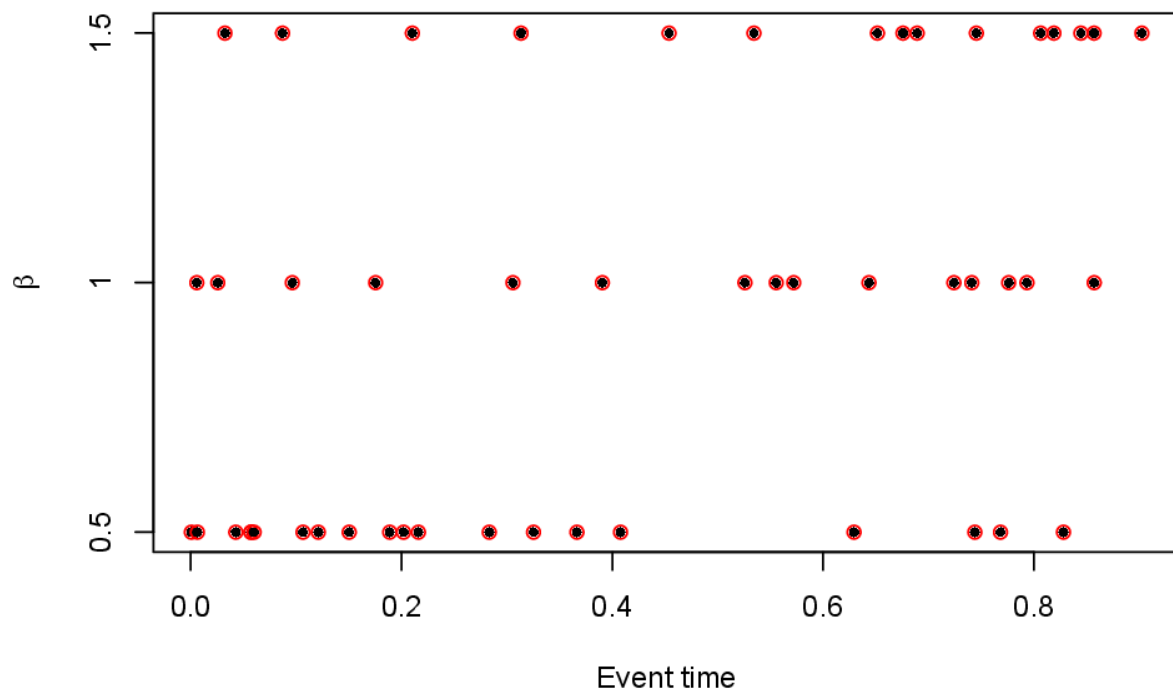


Figure 3.1. Simulation from NHPP's with different choices of β with $n = 15$

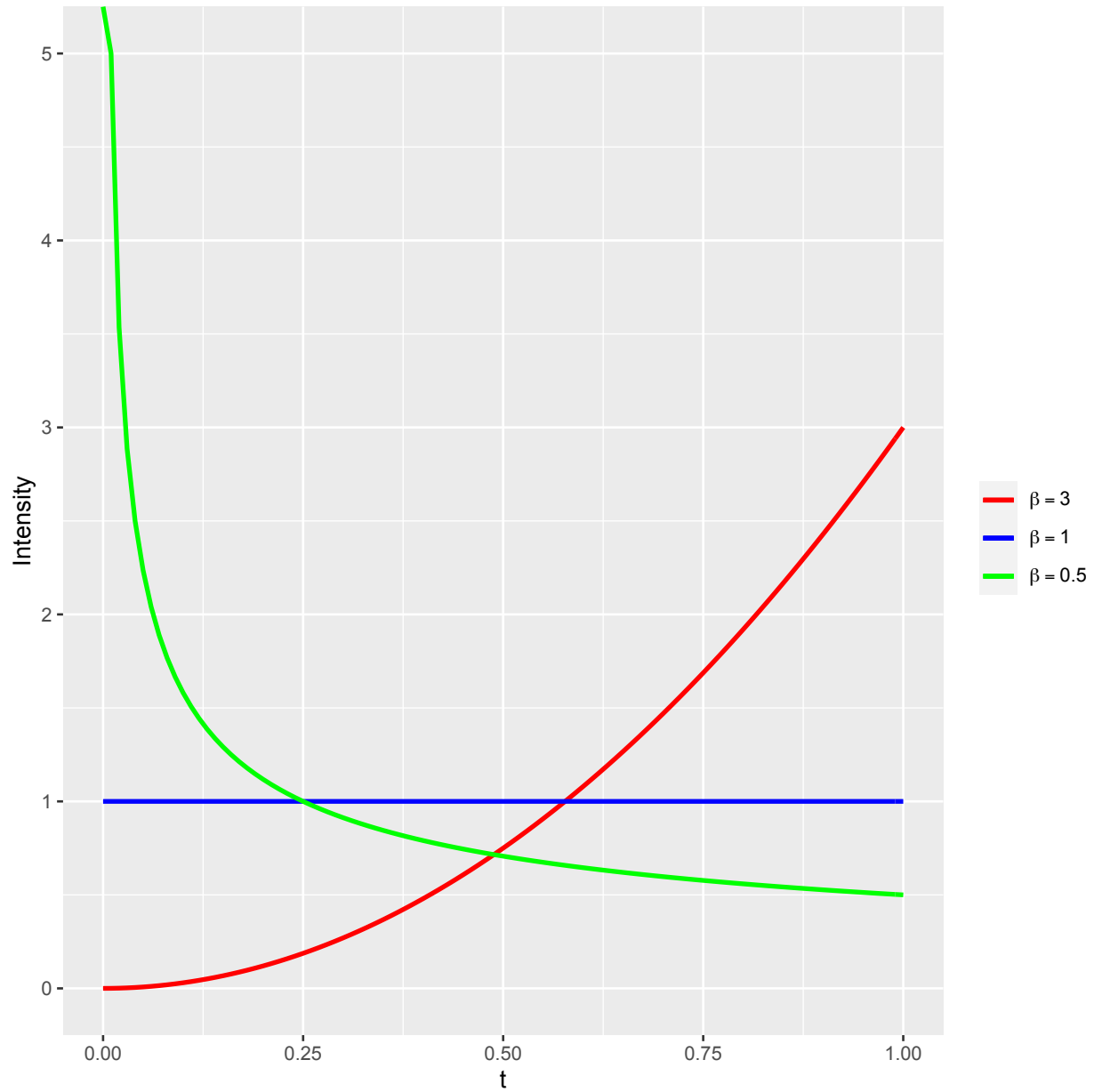


Figure 3.2. Power Law (Weibull) intensities for different choices of β

3.2.2 Step-Function Intensities

Smooth intensities are adequate for representing changes in the failure pattern, which are gradual over a considerable period. But if changes are abrupt, the failure rates are different over a while, meaning that each interval has its own value of intensity. This type of intensity graphed in the

Figure 3.3, called step-intensity, has the form:

$$\lambda(t) = \sum_{i=1}^p \lambda_i I_{(\tau_{i-1}, \tau_i]}(t), \quad t > 0 \quad (3.3)$$

where $I_A(\cdot)$ is the indicator function on set A .

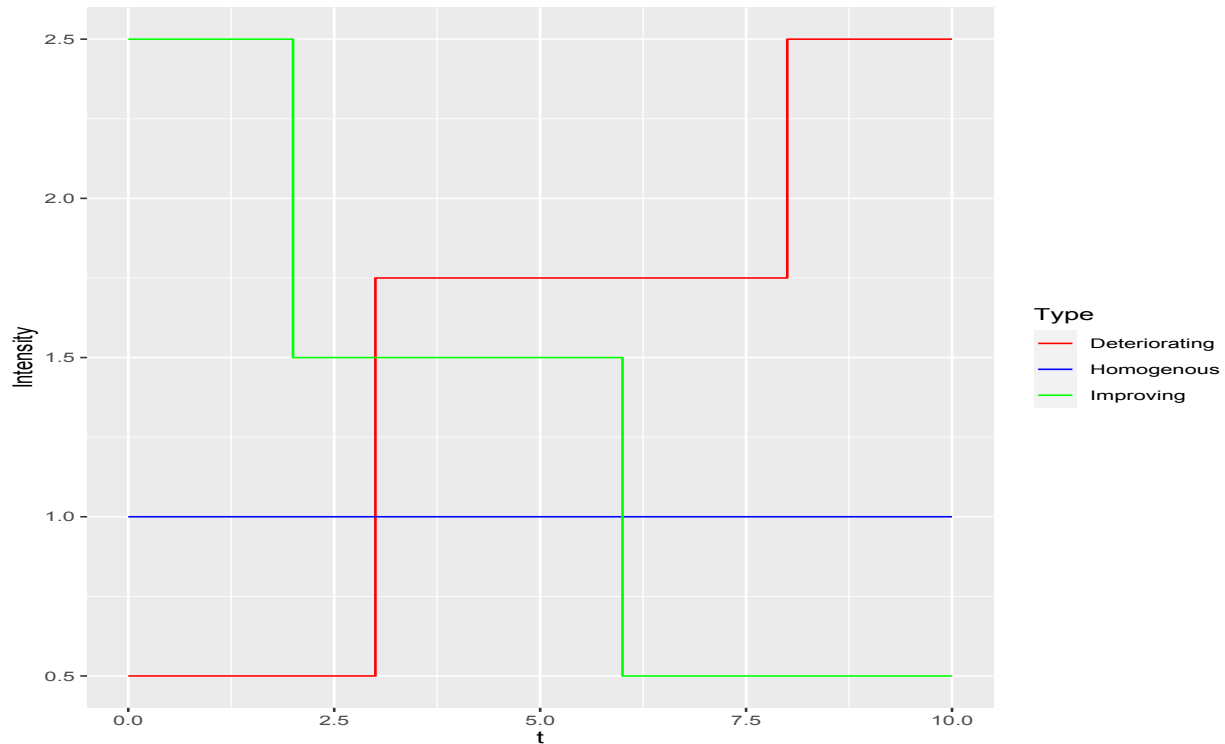


Figure 3.3. Three types of graph in three-step-intensity

The two-step intensity ($p = 2$) of the form can be defined as $\lambda(t) = \lambda_1$ for $0 < t \leq \tau$, and $\lambda(t) = \lambda_2$ for $\tau < t$. Moreover, we can calculate mean function in each interval such as:

$$\mu(t) = \int_0^t \lambda(x) dx = \begin{cases} \lambda_1 t & 0 < t \leq \tau \\ \lambda_1 \tau + \lambda_2 (t - \tau) & \tau < t \end{cases} \quad (3.4)$$

Generating an HPP of rate 1 (X), the NHPP (Y) can be obtained through the following inverse transformation:

$$y_i = \begin{cases} \frac{x_i}{\lambda_1} & x_i \leq \lambda_1 \tau \\ \frac{x_i - \lambda_1 \tau}{\lambda_2} + \tau & \lambda_1 \tau < x_i \end{cases} \quad (3.5)$$

For example, when $\lambda_1 = 5$, $\lambda_2 = 1$, $n = 30$ and sampling frequency is 1 : 1, this means within each interval $(0, \tau]$ and (τ, ∞) has 15 events ($= 30/2$). On $(0, \tau]$, as the process is an HPP with rate 5,

the inter-event time is exponential with average $1/5$. Thus, the waiting time for 15 events is $15 \times 1/5 = 3$ ($= \tau$).

For an observed number of failures, $n > 0$, the failure times are denoted by $0 < t_1 < t_2 < \dots < t_n < t$. The three-step intensity of the form under the time-truncated sampling can be defined as $\lambda(t) = k_1$, for $0 < t \leq \tau_1 T$, $\lambda(t) = k_2$, for $\tau_1 T < t \leq \tau_2 T$, and $\lambda(t) = k_3$, for $\tau_2 T < t \leq T$. Moreover, we can calculate mean function in each interval such as:

$$\mu(t) = \int_0^t \lambda(x) dx = \begin{cases} \lambda_1 t & 0 < t \leq \tau_1 T \\ \lambda_1 \tau_1 + \lambda_2 (t - \tau_1) & \tau_1 T < t \leq \tau_2 T \\ \lambda_1 \tau_1 + \lambda_2 (\tau_2 - \tau_1) + \lambda_3 (t - \tau_2) & \tau_2 T < t \leq T \end{cases} \quad (3.6)$$

Generating the order statistics, $u_{(1)} < u_{(2)} < \dots < u_{(n)}$, from a random sample of size n from the uniform distribution $(0, 1)$, the NHPP (Y) can be obtained through the solving $\mu(t_i)/\mu(t) = u_{(i)}$:

$$y_i = \begin{cases} \frac{u_{(i)} \mu(t)}{\lambda_1} & u_{(i)} \leq \frac{\lambda_1 \tau_1 t}{\mu(t)} \\ \frac{u_{(i)} \mu(t) - \lambda_1 \tau_1}{\lambda_2} + \tau_1 & \frac{\lambda_1 \tau_1 t}{\mu(t)} < u_{(i)} \leq \frac{\lambda_1 \tau_1 + \lambda_2 (\tau_2 t - \tau_1)}{\mu(t)} \\ \frac{u_{(i)} \mu(t) - \lambda_1 \tau_1 - \lambda_2 (\tau_2 - \tau_1)}{\lambda_3} + \tau_2 & \frac{\lambda_1 \tau_1 + \lambda_2 (\tau_2 t - \tau_1)}{\mu(t)} < u_{(i)} \end{cases} \quad (3.7)$$

Since the distribution of t_i/t , given $N = n$, does not depend on t , we set $t = 1$.

3.3 Power Comparisons for All Five Tests

The following power study is consistently for testing $H_0: \beta = 1$ versus H_A with a two-sided nominal level of $\alpha = 0.1$, where the distribution under H_A will be a known NHPP. Proxies developed above are not used for any of the simulations in this Chapter. The structure of the Monte Carlo simulations starts with a failure-truncated sampling scheme with 10^4 iterations for each alternative distribution. Tables from 3.1 to 3.4 summarize the results of this part. We then extend the same structure to a time-truncated sampling scheme to complete the power comparisons. Tables from 3.5 to 3.8 summarize the results of the second part. Specifically, a total of 60 settings (excluding the reference set, $\beta = 1$) of simulation will be considered. Every setting composes an element from each of the following components: (a) An intensity differentiated by type, trend, and the location/number of

change point(s), (b) sample size, and (c) sampling schemes as organized in the simulation flowchart (Figure 3.4).

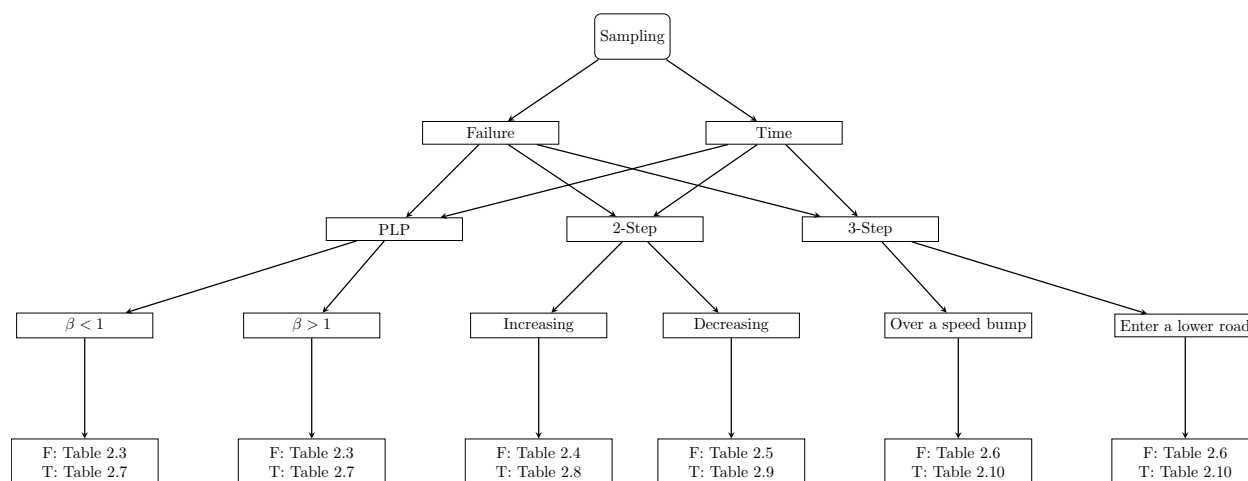


Figure 3.4. Simulation flowchart: (Sampling) schemes \rightarrow {(F)ailure-truncated sampling, (T)ime-truncated sampling} \rightarrow {PLP, (2-step) intensity, (3-step) intensity}, with simulation results from Tables 3.1 to 3.8.

3.3.1 Failure-Truncated Sampling

Three types of alternative distribution to be evaluated are: a PLP with $\beta \neq 1$ and an NHPP with a two/three-step-intensity function. Composed elements associated with each setting will be furnished as a footnote wherever applicable.

3.3.1.1 Power Law Process

For the first simulation set, we set $n = 40$, the total number of failures simulated/observed. In each case, H_A is a PLP with a β (1) listed in Table 3.1, which summarizes the estimated powers. The corresponding powers of the Z -test (1.30) are also computed analytically based on its theoretical power functions.

Table 3.1. Estimated powers for testing H_0 versus H_A : a power law process with $\beta = (0.2 - 2.0)$, respectively, and $n = 40$ (failure-truncated sampling).

Simulation	1	2	3	β				Overall Average	Maximum Deficit
Test	0.2	0.6	0.8	1.0	1.2	1.5	2.0		
Z (Analytical ^a)	1.0000	0.9378	0.4275	0.1000	0.2837	0.7939	0.9973		
Z	1.0000	0.9403	0.4299	0.1045	0.2869	0.7951	0.9976	0.7416 (1) ^b	0
Z_B	0.9998	0.6594	0.2541	0.1054	0.1784	0.4935	0.9011	0.5810	0.1606
R	1.0000	0.9404	0.4363	0.1071	0.1749	0.4938	0.9015	0.6578	0.0838
L	0.9998	0.6593	0.2482	0.1035	0.2909	0.7953	0.9976	0.6651 (3)	0.0765
P_{DB}	1.0000	0.9200	0.4016	0.0982	0.2536	0.7345	0.9943	0.7173 (2)	0.0243

^a Computed analytically based on the power function of the Z -test.

^b Top three tests per the averaged power.

Table 3.1 is meant to provide a baseline for all other simulation studies with abrupt change(s). First, values at $\beta = 1$ are coherent among tests along with that of analytical results of Z , which will be excluded from further power comparisons. An initial set of critical observations addresses the existence of power asymmetry under a PLP alternative and, consequently, the rationale of this study is summarized as follows:

- For a decreasing PLP ($\beta < 1$), R gets its full power from the right leg of Z which is a UMP for the PLP alternative. On the other hand, L is the worst in this category, because the leg from Z is useless for testing a decreasing PLP and the other one from Z_B cannot compete with that of the UMP.
- If the alternative is an increasing PLP ($\beta > 1$), the tables are turned and now L gets its full power from the left leg of the UMP and R takes the place of L as the worst. Note that R and L are tied at the first place with Z , respectively, in each trend accounted for simulation errors.
- Taking no notice of asymmetry, the averaged powers of R and L are close (Table 3.1). The

swing of the powers as the first sign of power asymmetries between R and L motivates us to consider a competitive indicator, termed as the maximum power-deficit, where individual deviation from the group maximum is quantified for performance assessments.

- In both cases, the P_{DB} -test maintains a steadfast second place with an averaged overall power of 71.73%, which is a mere loss of 2.43% versus that of the champion (= 74.16% for the Z), equivalently reflected by its averaged maximum power-deficit for a quick reference (Table 3.1). An elaborate set of maximum power-deficits, made for asymmetry enhancement and performance assessment, will appear as the study unfolds.

Both tests, the empowered pair, share the load in preserving the optimality of the Z -test, which is a sign of power asymmetry right at the baseline. The dual bidirectional P_{DB} -test, built into the structure, appears to be working to safeguard power symmetry as per the baseline analysis, which provides a platform for a straightforward extension to an alternative with pieces of HPPs, divided by one or two change points – a step-function intensity. Optimality and power asymmetry of the backward test are to appear in the settings that follow.

3.3.1.2 Two-Step-Intensity

We consider $n = 10, 20$, and 40 for all three groups here. In each group, a two-step-intensity of the form $\lambda(t) = 1$ for $0 < t \leq \tau$, and $\lambda(t) = 3$ for $\tau < t < \infty$ is considered with various values of τ for each sample size. For Group 1, the location of change occurs at $\tau = 5, 10$, and 20 for $n = 10, 20$, and 40 , respectively (i.e., the midpoint with respect to each n). For Group 2, jumps are set approximately one-third into the process at $\tau = 4, 7$, and 14 . Likewise, they are set at $\tau = 6, 13$, and 26 , approximately two-thirds into the process, for Group 3. Table 3.2 summarizes the results.

Table 3.2. Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 1$, $\lambda_2 = 3$, and change point at τ (failure-truncated sampling).

Simulation	7	8	9	10	11	12	13	14	15		
	Group ^a										
	1			2			3				
Test	$n = 10$ (4.9540) ^b	$n = 20$ (10.0085)	$n = 40$ (19.9476)	$n = 10$ (4.0297)	$n = 20$ (7.0081)	$n = 40$ (13.9469)	$n = 10$ (5.9691)	$n = 20$ (12.9367)	$n = 40$ (25.9515)	Overall Average	Maximum Deficit
Z	0.2650	0.4010	0.5997	0.3057	0.5030	0.7495	0.2106	0.2781	0.4116	0.4138	0.1164
Z_B	0.3008	0.5274	0.8218	0.2949	0.4866	0.7833	0.2796	0.4665	0.7359	0.5218 (2)	0.0084
R	0.3154	0.5314	0.8234	0.3072	0.4886	0.7841	0.3015	0.4796	0.7407	0.5302 (1)	0
L	0.2722	0.4060	0.5998	0.3110	0.5075	0.7495	0.2219	0.2859	0.4123	0.4184	0.1118
P_{DB}	0.3249	0.5091	0.7837	0.3256	0.5464	0.8039	0.2794	0.4237	0.6883	0.5205 (3)	0.0097

^a Combinations of [Group (sample size, location of τ)] are as follows: [1(10, 5)(20, 10)(40, 20)], [2(10, 4)(20, 7)(40, 14)], and [3(10, 6)(20, 13)(40, 26)].

^b The estimated mean number of occurrences in $[0, \tau]$.

For this simulation study, we include an abrupt change in each simulated process. Results in Table 3.2 enhance the utility of the proposed bidirectional tests a major step forward, because the Z -test has the lowest estimated power in nearly all the circumstances. The position of inferiority in power for Z is as high as a net of 32% as compared with that of the backward test Z_B (Table 3.2). Additional highlights are listed below:

- A small sample size ($n = 10$) or early location of change, set approximately one-third into the process, gives the P_{DB} -test an edge.
- The Z_B -test outperforms Z for both the middle and late τ settings.
- The R -test is consistently the winner for almost all the cases conducted, followed closely by Z_B , and then the P_{DB} -test.
- As per an averaged power-based ranking, we consider the result as a convincing three-way tie where the powers are (0.5302, 0.5218, 0.5205) for (R, Z_B, P_{DB}) , respectively. The corresponding overall averaged power is only 0.4138 for Z .

Therefore, for testing $H_0 : \beta = 1$ versus H_A , where the distribution under H_A is a two-step-

intensity of the form $\lambda(t) = 1$ for $0 < t \leq \tau$, and $\lambda(t) = 3$ for $\tau < t < \infty$ with various values of τ , the dual bidirectional P_{DB} -test, built into the structure, is equally competitive for detecting an abrupt change in a Poisson process. Of course, the conventional Z -test clearly is not set for these situations, which frequently arise in real-world applications.

In Table 3.3, we reverse the trend of the intensity function to the following: $\lambda(t) = 3$ for $0 < t \leq \tau$, and $\lambda(t) = 1$, for $\tau < t < \infty$. For Group 1, the location of change occurs at $\tau = 5/3, 10/3$, and $20/3$ for $n = 10, 20$, and 40 , respectively. For Group 2, jump points are set at $\tau = 4/3, 7/3$, and $14/3$. They are set at $\tau = 6/3, 13/3$, and $26/3$, again, approximately two-thirds into the process for Group 3.

Table 3.3. Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 3$, $\lambda_2 = 1$, and change point at τ (failure-truncated sampling).

Simulation	16	17	18	19	20	21	22	23	24		
	Group ^a										
	1			2			3				
Test	$n = 10$ (4.9900) ^b	$n = 20$ (10.0543)	$n = 40$ (20.0057)	$n = 10$ (3.9686)	$n = 20$ (7.0276)	$n = 40$ (14.0181)	$n = 10$ (5.9243)	$n = 20$ (12.9471)	$n = 40$ (25.9747)	Overall Average	Maximum Deficit
Z	0.2917	0.5379	0.8177	0.3140	0.5162	0.7809	0.2614	0.4398	0.7105	0.5189 (3)	0.0057
Z_B	0.2859	0.4392	0.6372	0.2565	0.3202	0.4545	0.2876	0.4694	0.7225	0.4303	0.0943
R	0.2999	0.5411	0.8189	0.3231	0.5220	0.7836	0.2673	0.4399	0.7106	0.5229 (2)	0.0017
L	0.2819	0.4400	0.6364	0.2501	0.3175	0.4522	0.2849	0.4731	0.7228	0.4287	0.0959
P_{DB}	0.3071	0.5317	0.8018	0.3065	0.4799	0.7226	0.2960	0.4995	0.7764	0.5246 (1)	0

^a Combinations of [Group (sample size, location of τ)] are as follows: [1 (10, 5/3) (20, 10/3) (40, 20/3)], [2 (10, 4/3) (20, 7/3) (40, 14/3)], and [3 (10, 6/3) (20, 13/3) (40, 26/3)].

^b The estimated mean number of occurrences in $[0, \tau]$.

A straightforward reversal of a (1, 3) trend to (3, 1) for (λ_1, λ_2) revives the power of Z to a close third place (Table 3.3):

- If the location of change is set approximately two-thirds into the process, the P_{DB} -test has an edge for this decreasing trend, although in general, the advantage of a small sample size at 10 diminishes. Interestingly enough, Z ranks the last in this setting of change.
- If the location of change is set approximately one-third or halfway into the process, the ranking is R , Z , and P_{DB} . The L -test happens to be the worst.

- In this study, we consider the result as a convincing three-way tie where the powers are (0.5246, 0.5229, 0.5189) for (P_{DB}, R, Z) , respectively. The corresponding overall averaged power is 0.4287 for L , the lowest.

So far, the demonstrated asymmetry has been more than offset by a dual bidirectional testing concept, delivered by P_{DB} .

3.3.1.3 Nonmonotonic Three-Step-Intensity

This simulation study (Table 3.4) involves two sets with three groups under each set. All the groups have a common sample size of 36. In Set A, a three-step-intensity of the form $\lambda(t) = 1$ for $0 < t \leq \tau_1$, $\lambda(t) = 3$, for $\tau_1 < t \leq \tau_2$, and $\lambda(t) = 1$, for $\tau_2 < t < \infty$ is considered with various values of τ_1 and τ_2 . Combinations of [Group (location of τ_1, τ_2)] are as follows: [A1 (12, 16)], [A2 (9, 12)], and [A3 (18, 21)]. In short, for Set B, the corresponding $\lambda(t)$ is $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ with various values of τ_1 and τ_2 . Along these lines, combinations of [Group (location of τ_1, τ_2)] are as follows: [B1 (4, 16)], [B2 (3, 12)], and [B3 (6, 15)].

Table 3.4. Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$ (failure-truncated sampling).

Simulation	25	26	27	28	29	30		
	Group ^a			Group ^b				
	$\lambda_1 = 1, \lambda_2 = 3, \text{ and } \lambda_3 = 1$			$\lambda_1 = 3, \lambda_2 = 1, \text{ and } \lambda_3 = 3$				
	1	2	3	1	2	3		
	$n = 36$	$n = 36$	$n = 36$	$n = 36$	$n = 36$	$n = 36$		
Test	(11.9792) ^c	(8.9939)	(17.9402)	(12.0539)	(8.9952)	(17.9703)	Overall	Maximum
	(11.9541) ^d	(8.9872)	(8.9373)	(11.9918)	(8.9619)	(8.9219)	Average	Deficit
Z	0.1476	0.0831	0.1873	0.3239	0.1866	0.3685	0.2161 (3)	0.0543
Z_B	0.1235	0.1500	0.1425	0.2724	0.3216	0.2335	0.2072	0.0632
R	0.0584	0.0569	0.1267	0.5234	0.4317	0.4256	0.2704 (1)	0
L	0.2150	0.1740	0.2006	0.0654	0.0734	0.1468	0.1458	0.1246
P_{DB}	0.1393	0.1111	0.1703	0.3862	0.3216	0.3485	0.2461 (2)	0.0243

^a Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (12, 16)], [2 (9, 12)], and [3 (18, 21)].

^b Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (4, 16)], [2 (3, 12)], and [3 (6, 15)].

^c The estimated mean number of occurrences in $[0, \tau_1]$.

^d The estimated mean number of occurrences in $(\tau_1, \tau_2]$.

This nonmonotonic intensity function with three steps provides additional insight into the power study. In the case where $(\lambda_1, \lambda_2, \lambda_3) = (1, 3, 1)$, the second segment has a mean occurrence of about 12, 9, and 9, respectively, for each pair of τ_1 and τ_2 (Table 3.4). The unit rate of the middle segment ($\lambda_2 = 3$) is three times that of its adjacent segments ($\lambda_1 = \lambda_3 = 1$). The total number of expected occurrences for its neighbors are set as either the same or twice as many, which would take the middle segment a relatively small timescale to run its course. Therefore, testing a hypothesis such as this may be perceived as driving over a speed bump. Along these lines, the other case may be perceived as entering a lower road ready for a new asphalt pavement installation.

As expected, most of the estimated powers are low, and the asymmetrical performances continue. In these settings, L takes the first place in Set A, but the last for Set B, and vice versa for the R -test. In the end, P_{DB} stands firm between R and Z , because the losses of L in Set B are quite significant.

3.3.2 Time-Truncated Sampling

The following types of alternative distribution are evaluated again in a time-truncated setting: a PLP with $\beta \neq 1$ and an NHPP with a two/three-step-intensity function. Tables from 3.5 to 3.8 below represent the counterparts of tables from 3.1 to 3.4.

Results summarized in Table 3.5 are essentially the same as that in Table 3.1. Again, values at $\beta = 1$ are coherent among tests and the analytical result of Z (1.31), which will be excluded from further power comparisons.

Table 3.5. Estimated powers for testing H_0 versus H_A : a power law process with $\beta = (0.2 - 2.0)$, respectively, and $n = 40$ (time-truncated sampling).

Simulation	1	2	3	β				Overall Average	Maximum Deficit
Test	0.2	0.6	0.8	1.0	1.2	1.5	2.0		
Z (Analytical*)	1.0000	0.9423	0.4341	0.1000	0.2887	0.8036	0.9970		
Z	1.0000	0.9421	0.4333	0.0956	0.2933	0.8061	0.9976	0.7453 (1)	0
Z_B	0.9999	0.6621	0.2521	0.0974	0.1826	0.5008	0.9126	0.5850	0.1603
R	1.0000	0.9421	0.4367	0.0959	0.1767	0.4986	0.9112	0.6608	0.0845
L	0.9999	0.6626	0.2482	0.0972	0.2992	0.8069	0.9977	0.6690 (3)	0.0763
P_{DB}	1.0000	0.9265	0.3969	0.1023	0.2564	0.7522	0.9947	0.7211 (2)	0.0242

^a Computed analytically based on the power function of the Z -test.

^b Top three tests per the averaged power.

Table 3.6. Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 1$, $\lambda_2 = 3$, and change point at τ (time-truncated sampling).

Simulation	7	8	9	10	11	12	13	14	15	Overall Average	Maximum Deficit
Test	τ										
	1/2			1/3			2/3				
	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$	$n = 10$	$n = 20$	$n = 40$		
Z	(2.4974) ^a	(5.0104)	(10.0414)	(1.4367)	(2.8547)	(5.7344)	(3.9709)	(8.0146)	(16.0109)	0.5428 (2)	0.0019
Z_B	0.2383	0.3937	0.6556	0.1504	0.2308	0.3731	0.3354	0.5453	0.8192	0.4157	0.1290
R	0.2395	0.3946	0.6539	0.1515	0.2305	0.3704	0.3368	0.5477	0.8176	0.4158	0.1289
L	0.3881	0.5787	0.7924	0.3036	0.4926	0.7351	0.3579	0.5222	0.7325	0.5447 (1)	0
P_{DB}	0.3505	0.5493	0.7950	0.2363	0.4121	0.6832	0.3712	0.5709	0.8365	0.5338 (3)	0.0109

^a The estimated mean number of occurrences in $[0, \tau]$.

Results summarized in Table 3.6 are mostly opposite to that in Table 3.2:

- Late location of change, set approximately two-thirds into the process, gives the P_{DB} -test an edge opposite to that of the failure-truncated setting.
- The Z -test outperforms Z_B (1.68) for both the middle and early τ settings – the tables are turned at a reversed setting of τ , if the sampling scheme is different.
- The L -test is the winner for the cases where the early or middle location of change is placed.
- Overall, we consider the result as a convincing three-way tie where the powers are (0.5447, 0.5428, 0.5338) for (L, Z, P_{DB}) , respectively. The corresponding overall averaged power is 0.4157 for Z_B . The top three, (R, Z_B, P_{DB}) , get replaced in this time-truncated setting.

Table 3.7. Estimated powers for testing H_0 versus H_A : a two-step-intensity with $\lambda_1 = 3$, $\lambda_2 = 1$, and change point at τ (time-truncated sampling).

Simulation	16	17	18	19	20	21	22	23	24	Overall Average	Maximum Deficit
	τ										
	1/2			1/3			2/3				
Test	$n = 10$ (7.4821) ^a	$n = 20$ (15.0265)	$n = 40$ (29.9529)	$n = 10$ (5.9942)	$n = 20$ (12.0198)	$n = 40$ (24.0673)	$n = 10$ (8.5675)	$n = 20$ (17.1495)	$n = 40$ (34.2712)		
Z	0.2325	0.3963	0.6590	0.3423	0.5503	0.8244	0.1479	0.2281	0.3754	0.4173	0.1254
Z_B	0.3768	0.5640	0.7951	0.3536	0.5151	0.7350	0.2948	0.4892	0.7298	0.5392 (2)	0.0035
R	0.2338	0.3970	0.6561	0.3463	0.5529	0.8240	0.1505	0.2287	0.3733	0.4180	0.1247
L	0.3801	0.5685	0.7963	0.3563	0.5198	0.7362	0.2992	0.4966	0.7319	0.5427 (1)	0
P_{DB}	0.3500	0.5493	0.7881	0.3791	0.5705	0.8317	0.2395	0.4325	0.6708	0.5346 (3)	0.0081

^a The estimated mean number of occurrences in $[0, \tau]$.

A straightforward reversal of a (1, 3) trend to (3, 1) for (λ_1, λ_2) for the time-truncated sampling pushes the power of Z to the last place (Table 3.7). The asymmetrical patterns repeat themselves:

- If the location of change is set approximately one-third into the process, the P_{DB} -test has an edge for this decreasing trend.

- If the location of change is set approximately two-thirds or halfway into the process, the ranking is L, Z_B , and P_{DB} . The R -test happens to be the worst.
- In this study, we consider the result as a convincing three-way tie where the powers are (0.5427, 0.5392, 0.5346) for (L, Z_B, P_{DB}) , respectively. The corresponding overall averaged power is 0.4173 for Z , the lowest.

Again, the demonstrated asymmetry suffered here for the conventional Z -test is equally dramatic as that based on the failure-truncated sampling. Results summarized in Table 3.8 are mostly the same as that of Table 3.4 with one exception: the final rankings are P_{DB}, L , and Z_B .

Table 3.8. Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3; $j = 1, 2$ (time-truncated sampling).

Simulation	25	26	27	28	29	30		
	Group ^a			Group ^b				
	$\lambda_1 = 1, \lambda_2 = 3, \text{ and } \lambda_3 = 1$			$\lambda_1 = 3, \lambda_2 = 1, \text{ and } \lambda_3 = 3$				
	1	2	3	1	2	3		
	$n = 40$	$n = 40$	$n = 40$	$n = 40$	$n = 40$	$n = 40$		
Test	(8.0170) ^c (24.0011) ^d	(6.6579) (20.0126)	(13.3302) (20.0044)	(17.1474) (5.7180)	(11.9949) (4.0055)	(24.0037) (3.9910)	Overall Average	Maximum Deficit
Z	0.1543	0.0442	0.3235	0.1648	0.1343	0.1626	0.1639	0.0377
Z_B	0.1572	0.3258	0.0424	0.1635	0.1710	0.1326	0.1654 (3)	0.0362
R	0.0082	0.0279	0.0270	0.2779	0.2198	0.2126	0.1289	0.0727
L	0.3054	0.3438	0.3405	0.0473	0.0838	0.0810	0.2003 (2)	0.0013
P_{DB}	0.1760	0.2476	0.2608	0.1902	0.1662	0.1691	0.2016 (1)	0

^a Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (1/3, 2/3)], [2 (1/4, 1/2)], and [3 (1/2, 3/4)].

^b Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (1/3, 2/3)], [2 (1/4, 1/2)], and [3 (1/2, 3/4)].

^c The estimated mean number of occurrences in $[0, \tau_1]$.

^d The estimated mean number of occurrences in $(\tau_1, \tau_2]$.

Time-truncated sampling studies have given us added awareness of what stands for performance in hypothesis testing when its symmetry is threatened with respect to the composed elements in the settings for an alternative distribution. We, therefore, see merit in both of the simulation schemes.

Anatomy based on a body of performance data, consolidating the above eight tables (from 3.1 to

3.8), provides additional insights into the behavior of all the tests – an overview.

3.4 An Overview

It is to be mentioned that the test with the lowest score of maximum power-deficit is the superlative one, which is provided for each of the tests under scrutiny. Firstly, pattern recognition of power asymmetries is bested by Figure 3.5, which also serves as a structure and performance overview of all the simulations conducted in this study. Furthermore, all the maximum power-deficits, made for performance assessment and asymmetry confirmation, are presented in Table 3.9.

Table 3.9. Maximum power-deficit overview.

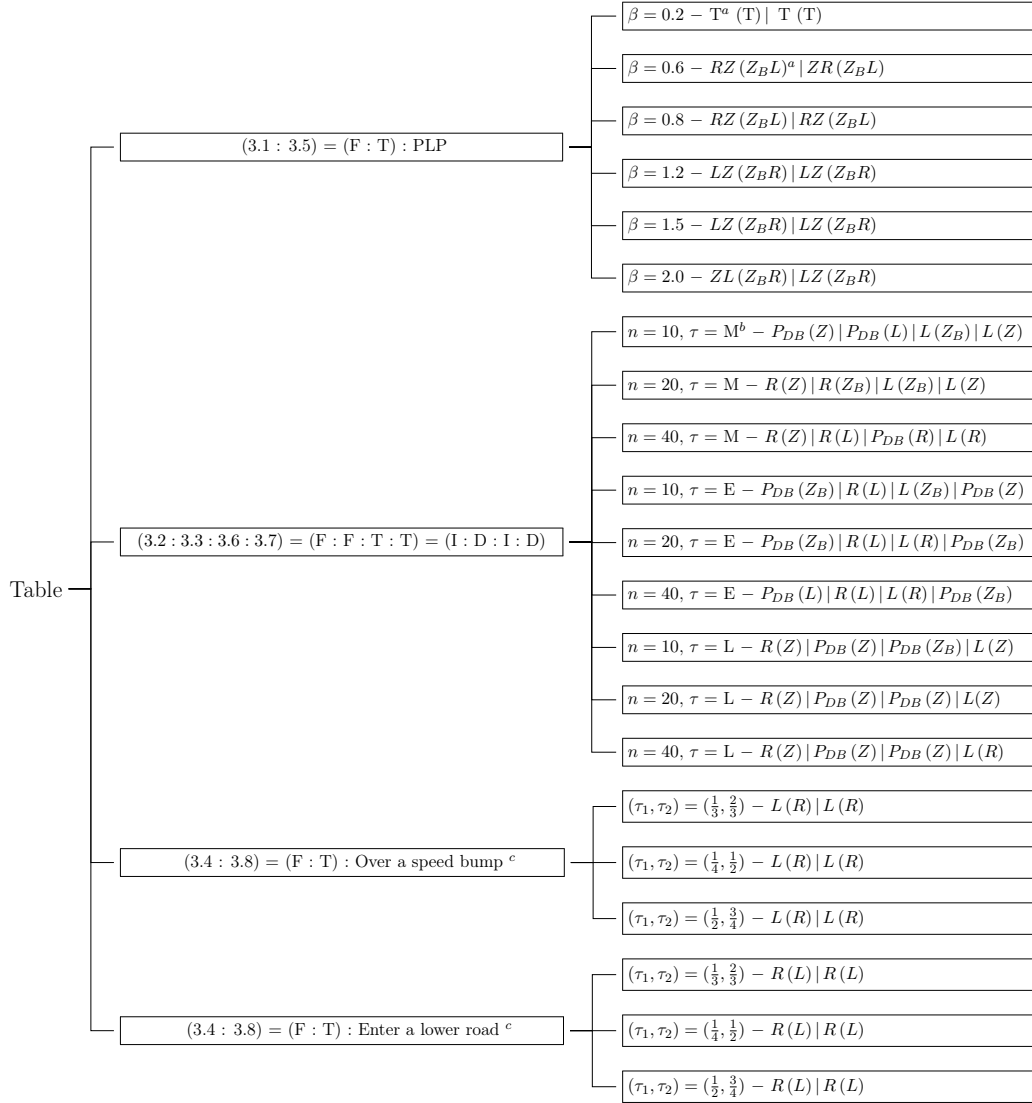
Simulation	Failure-truncated					Time-truncated				
	Z	Z_B	R	L	P_{DB}	Z	Z_B	R	L	P_{DB}
	Table 3.1					Table 3.5				
1	0	0.0002	0	0.0002	0	0	0.0001	0	0.0001	0
2	0.0001	0.2810	0	0.2811	0.0204	0	0.2800	0	0.2795	0.0156
3	0.0064	0.1822	0	0.1881	0.0347	0.0034	0.1846	0	0.1885	0.0398
4	0.0040	0.1125	0.1160	0	0.0373	0.0059	0.1166	0.1225	0	0.0428
5	0.0002	0.3018	0.3015	0	0.0608	0.0008	0.3061	0.3083	0	0.0547
6	0	0.0965	0.0961	0	0.0033	0.0001	0.0851	0.0865	0	0.0030
	Table 3.2					Table 3.6				
7	0.0599	0.0241	0.0095	0.0527	0	0.0036	0.1498	0.1486	0	0.0376
8	0.1304	0.0040	0	0.1254	0.0223	0.0042	0.1850	0.1841	0	0.0294
9	0.2237	0.0016	0	0.2236	0.0397	0.0037	0.1394	0.1411	0.0026	0
10	0.0199	0.0307	0.0184	0.0146	0	0.0007	0.1532	0.1521	0	0.0673
11	0.0434	0.0598	0.0578	0.0389	0	0.0054	0.2618	0.2621	0	0.0805
12	0.0544	0.0206	0.0198	0.0544	0	0.0018	0.3620	0.3647	0	0.0519
13	0.0909	0.0219	0	0.0796	0.0221	0.0106	0.0358	0.0344	0.0133	0
14	0.2015	0.0131	0	0.1937	0.0559	0.0510	0.0256	0.0232	0.0487	0
15	0.3291	0.0048	0	0.3284	0.0524	0.1052	0.0173	0.0189	0.1040	0
	Table 3.3					Table 3.7				
16	0.0154	0.0212	0.0072	0.0252	0	0.1476	0.0033	0.1463	0	0.0301
17	0.0032	0.1019	0	0.1011	0.0094	0.1722	0.0045	0.1715	0	0.0192
18	0.0012	0.1817	0	0.1825	0.0171	0.1373	0.0012	0.1402	0	0.0082
19	0.0091	0.0666	0	0.0730	0.0166	0.0368	0.0255	0.0328	0.0228	0
20	0.0058	0.2018	0	0.2045	0.0421	0.0202	0.0554	0.0176	0.0507	0
21	0.0027	0.3291	0	0.3314	0.0610	0.0073	0.0967	0.0077	0.0955	0
22	0.0346	0.0084	0.0287	0.0111	0	0.1513	0.0044	0.1487	0	0.0597
23	0.0597	0.0301	0.0596	0.0264	0	0.2685	0.0074	0.2679	0	0.0641
24	0.0659	0.0539	0.0658	0.0536	0	0.3565	0.0021	0.3586	0	0.0611
	Table 3.4					Table 3.8				
25	0.0674	0.0915	0.1566	0	0.0757	0.1511	0.1482	0.2972	0	0.1294
26	0.0909	0.0240	0.1171	0	0.0629	0.2996	0.0180	0.3159	0	0.0962
27	0.0133	0.0581	0.0739	0	0.0303	0.0170	0.2981	0.3135	0	0.0797
28	0.1995	0.2510	0	0.4580	0.1372	0.1131	0.1144	0	0.2306	0.0877
29	0.2451	0.1101	0	0.3583	0.1101	0.0855	0.0488	0	0.1360	0.0536
30	0.0571	0.1921	0	0.2788	0.0771	0.0500	0.0800	0	0.1316	0.0435
Average ^a	0.0678 (3)	0.0958	0.0376 (2)	0.1228	0.0329 (1)	0.0736 (3)	0.1070	0.1354	0.0434 (2)	0.0385 (1)

^a Averaged maximum power-deficit under all the simulations.

3.4.1 Road Map

Figure 3.5 provides a detailed study plan and key results, outlined underneath:

- Summarizing tables with like settings described in the adjacent twigs are stacked in four large branches. An ordered pair, (3.1 : 3.5), stands for Table 3.1 first then Table 3.5, and so forth.
- (3.1 : 3.5) = (F : T) indicates the corresponding sampling scheme, failure-truncated (F) or time-truncated (T) sampling, and so forth for more than two tables. Likewise, in the second branch, (I : D : I : D) symbolizes the respective trend order (a) associated with a step-function-intensity: monotonically increasing (I), or decreasing (D); (b) adopted by an orderly array of tables: 3.2, 3.3, 3.6, and 3.7.
- Consequently, each twig also stacks executed simulations by the same order as structured, where the total number of simulations accrued are affixed with divided cells.
- Each cell houses the best and worst test of each simulation, denoted as, say, $R(L)$ – the worst test is in the open parenthesis. Additionally, $ZR(L)$ means that Z and R are tied at the first place; “T” represents a case with more than two ties.



^a T = a tie for at least three; $RZ (Z_B L)$ = best (worst).

^b Location of change: (a) E = early into a process, (b) M = middle, and (c) L = late.

^c Over a speed bump = $(\lambda_1, \lambda_2, \lambda_3) = (1, 3, 1)$; Enter a lower road = $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$.

Figure 3.5. Structure overview – simulation: failure-truncated sampling (Table 3.1 – 3.4); time-truncated sampling (Table 3.5 – 3.8).

3.4.2 Diagnosis of Power Asymmetries

The performance board posted on Figure 3.5 is only for depicting that both R and L conquer the entire road map, evidenced by numerous pairs of $R(L)$ and $L(R)$ – a sign of power asymmetries. In practice, if a practitioner did find themselves in one of those scenarios, we think they would struggle to choose among these alternatives – especially if they led to disparate outcomes.

We've seen R and L championing in one alternative setting, but also being turned against the very setting altered with just one of the composed elements (Figure 3.5). Figure 3.6 reinforces this claim, where the sign of one set of the maximum power-deficits has been reversed to enhance the existences of power asymmetries. Recall that both tests own two parallel components outsourced from Z and Z_B , embedded with power asymmetries themselves (Figure 3.7), which have been discussed.

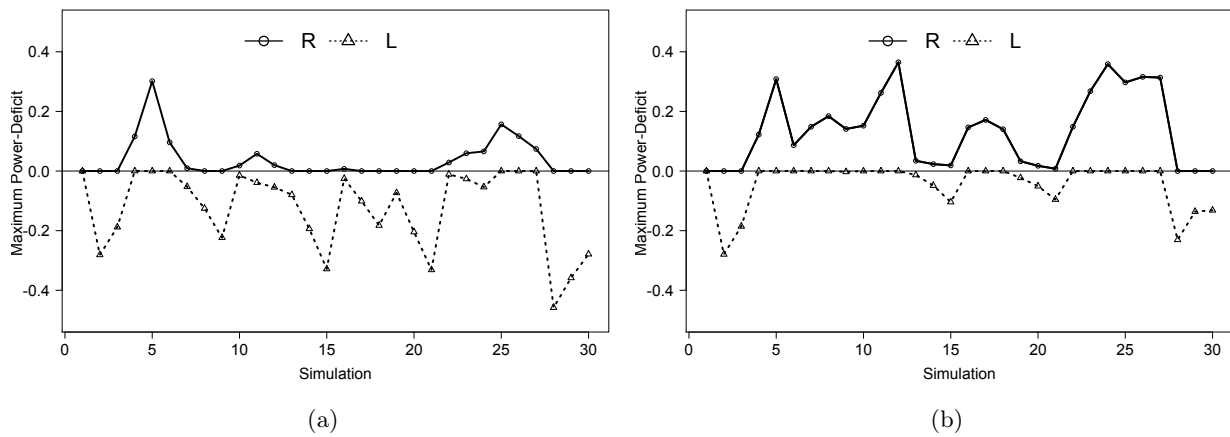


Figure 3.6. Maximum power-deficit-based performance comparisons for tests of R and L for (a) failure-truncated sampling and (b) time-truncated sampling.

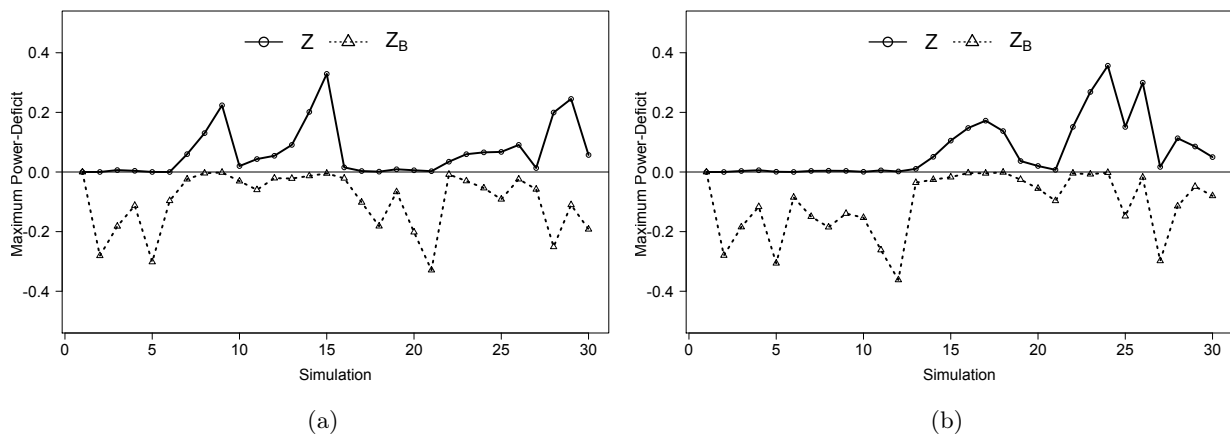


Figure 3.7. Maximum power-deficit-based performance comparisons for tests of Z and Z_B for (a) failure-truncated sampling and (b) time-truncated sampling.

As the first line of defense against power asymmetries, we consolidate R (1.108) and L (1.109) to assemble a hybrid. Performance data justify that the strategy works: Table 3.9, Figures 3.8 and 3.9. Most importantly, patterns presented in Figure 3.8 indicate that the power asymmetries diminish to the utmost for P_{DB} (1.110), and meanwhile Figure 3.9, a graphic version of Table 3.9, is simple and yet powerful in defending a claim that P_{DB} is competitive.

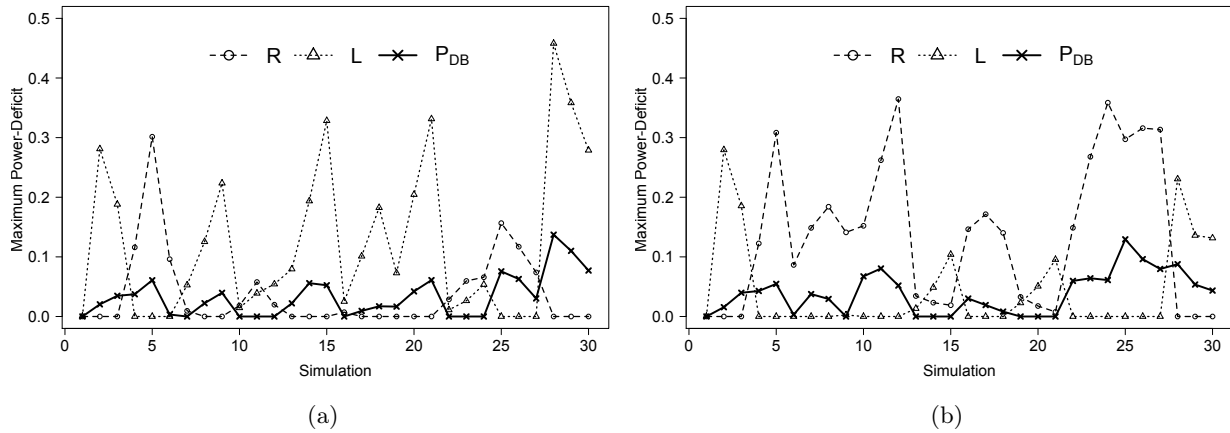


Figure 3.8. Maximum power-deficit-based performance comparisons for tests of R , L , and P_{DB} for (a) failure-truncated sampling and (b) time-truncated sampling.

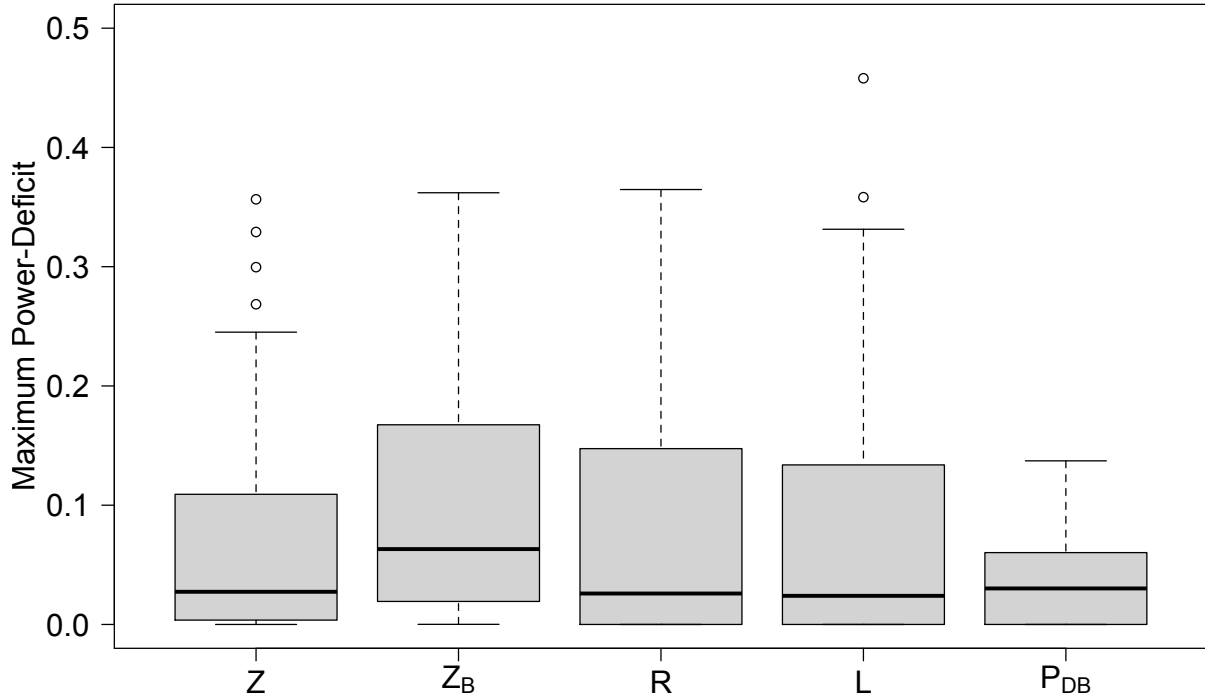


Figure 3.9. Box-plots for all 60 maximum power-deficits.

3.5 Conclusions

Lastly, we are ready to sum up the proposed cocktail of coherent bidirectional tests, $(Z, Z_B) \rightarrow [R, L] \rightarrow \{P_{DB}\}$, visualized via reliability graphics and justified by the following, for power symmetry and repairable systems reliability:

- There are a lot of power asymmetries going on as per the simulation results. Every setting composes an element from each of the following components: (a) An intensity differentiated by the type of function, trend, and location/number of change point(s), (b) sample size, and (c) sampling schemes.
- The asymmetrical characteristics of the basic pair (Z, Z_B) carry over into the empowered

bidirectional version $[R, L]$. We've seen either test championing in one alternative setting, but also being turned against the very setting altered with just one of the composed elements.

- We assemble a hybrid from the empowered pair to restore power symmetry, which is p -value induced with the dual purpose of filtering out the bad and enhancing the good. Namely, it is coherent with no tests expended on the journey from Z to P_{DB} .
- A user-friendly version of the test, adapted for modern practitioners, makes the p -values much more attainable. Henceforth, there is a so-called “robust” test that may possess power symmetry – essential quality assurance certified first in this study.
- In this study we focus on repairable systems, but many, if not most, of the concepts we present apply more broadly.

In closing, a cocktail of coherent tests, which is rooted in a basic pair of tests, (Z, Z_B) , that are more complementary than competitive, and are bound together with a common set of critical values. Both tests are exact tests. Another unified partners of an empowered pair of tests, $[R, L]$, form a second bidirectional test. A hybrid, $\{P_{DB}\}$, of the empowered pair, characterizes the test as the only tool possessing the unique feature of having an intelligible version, grounded in practicality that is free of either the system sample size or the sampling schemes. It is, therefore, capable of producing a one-size-fits-all test, proven to be a safe bet for practitioners doing the repairable systems reliability work. Cox (1955) [21] discusses the use of a test statistic attributed to Laplace, which yields a UMPU test of constant versus increasing intensity based on the time-truncated framework for moderate or large sample sizes (Rigdon and Basu 2000 [81]). To further substantiate the applicability of P_{DB} , next Chapter is to address the competitiveness between these tests, developed generations apart. The results, building upon the performance of the legendary Laplace test, would be of practical significance to practitioners or statistical consultants helping with the projects through to their

completions. We then will extend the statistical process control architecture to a panel of coherent control charts fit for group sequential testing and multisystem.

CHAPTER 4

THE P_{DB} -TEST VERSUS THE LAPLACE TEST

4.1 Motivation

Cox (1955) [21] discusses using a test statistic attributed to Laplace, L , which yields a UMPU test of constant versus increasing intensity based on the time-truncated framework for moderate or large sample sizes (Rigdon and Basu 2000 [81]). The notation L and the name “Laplace test” result from the fact that it is suggested much earlier by Laplace as a test of randomness (Bain and Engelhardt 1991 [9]). A slightly different form of the argument is required if the observed time period is taken up to a preassigned number of events – a failure-truncated case (Cox and Lewis 1966 [22]). Extended research and applications of the Laplace test are presented in, for example, Ascher and Feingold (1978) [6], Bastos et al. (1990) [13], Gaudoin (1992) [38], Kim et al. (1992) [64], Ho (1993 [49], 1998 [51]), and Roche-Carrier et al. (2020) [83]. Here, the Laplace test will be referred to as the L_P -test, or L_P , because it coincides with a different test introduced underneath.

According to the Chapter 3, we open the door to an option between a cocktail of tests, $(Z, Z_B) \rightarrow [R, L] \rightarrow \{P_{DB}\}$, and a solo one-size-fits-all test, certified first as robust with essential quality assurance in addition to a user-friendly version for the practitioners conducting the work. Two bidirectional tests $[R, L]$ are first established to empower a basic pair of tests (Z, Z_B) , documented as tests with asymmetrical performances (Ho 1993 [49]). Progressing by filtering out the bad and enhancing the good, they assemble a hybrid, $\{P_{DB}\}$, from the empowered pair to restore power symmetries that are thought to be automatic and required for quality assurance. To further substantiate the applicability of a p -value induced dual bidirectional test, we challenge the legendary Laplace test, the mainstream and a classic for the repairable systems reliability and more, to a duel.

4.1.1 Theoretical Results

Consider an NHPP with intensity function of the form $\lambda(t|\Theta) = \alpha \exp(\beta t)$, L_P yields a UMPU test of $H_0: \beta = 0$ versus $H_A: \beta > 0$ (Cox 1955 [21]). Under H_0 , conditional on $N = n$ for the time-truncated sampling, and by the Central Limit Theorem, the distribution of L_P is approximately normal. Along these lines, the test based on Z is UMP test for $H_0: \beta = 1$ versus $H_A: \beta > 1$ in the PLP setting, where the intensity function is $\lambda(t|\Theta) = (\beta/\theta)(t/\theta)^{\beta-1}$, and $\beta > 0$ and $\theta > 0$ (Bassin 1969 [12]; Crow 1974 [23], 1982 [24]; Finkelstein 1976 [37]; Lee and Lee 1978 [65]; Bain and Engelhardt 1980 [8]). Another statistic Z_B , which is a reversed version of Z , is called the backward test by Ho (1993) [49]. Theoretical results of the functions of the test statistics for the above hypotheses and the relevant distributions of the order statistics are summarized in Table 4.1.

Table 4.1. Functions for the test statistics (L_p , Z , and Z_B) and the p -value proxies (R , L , and P_{DB}), failure-truncated and time-truncated sampling with size n . The cumulative failure times (t_i 's and t) are recognized as the main series of events.

Statistic	Failure-truncated	Time-truncated
L_p	$l_p = (\sum_{i=1}^{n-1} t_i / (n-1) - t_n/2) / (t_n / \sqrt{12(n-1)})$	$l_p = (\sum_{i=1}^n t_i / n - t/2) / (t / \sqrt{12n})$
Z	$z = 2 \sum_{i=1}^{n-1} \ln(t_n/t_i)$	$z = 2 \sum_{i=1}^n \ln(t/t_i)$
Z_B	$z_B = 2 \sum_{i=1}^{n-1} \ln(t_n/(t_n - t_i))$	$z_B = 2 \sum_{i=1}^n \ln(t/(t - t_i))$
<hr/>		
P -value		
$R = r$	$1 - [F_{\chi^2(2n-2)}(r)]^2$	$1 - [F_{\chi^2(2n)}(r)]^2$
$L = l$	$1 - [1 - F_{\chi^2(2n-2)}(l)]^2$	$1 - [1 - F_{\chi^2(2n)}(l)]^2$
$P_{DB} = p$	$1 - (1 - p)^2$	$1 - (1 - p)^2$

4.1.2 Simulation Studies

The performance of both L_P and Z -test are discussed for smooth alternatives and step functions with one to three (regular or irregular) jumps in articles of Bain et al. (1985) [10] and Engelhardt et al. (1990) [33]. The above authors present their results for the power of tests against increasing intensity alternatives for the time-truncated sampling. A power study for both the time-truncated

and the failure-truncated cases for L_P , Z , and Z_B indicates that their results cannot be interpreted symmetrically to alternatives that are decreasing step-function intensities (Ho 1993). As per another simulation study of a cocktail of tests in Chapter 3, $(Z, Z_B) \rightarrow [R, L] \rightarrow \{P_{DB}\}$, the asymmetrical characteristics of the basic pair (Z, Z_B) carry over into the empowered bidirectional version $[R, L]$, where either test championing in one alternative setting, but also being turned against the very setting altered with just one of the composed elements set in the study. Hence, an alternative, $\{P_{DB}\}$, accommodating essential quality assurance for the practitioners conducting the work, becomes available, which rationalizes our urge for to conduct a power contest between tests of P_{DB} and L_P .

Simulation studies are an essential tool for statistical research, aimed better to understand the behavior and performance of the calculated statistics, evidenced by the above studies delivering results such as power asymmetries appeared to be not answerable beforehand by theories or real data alone. The distributions of the order statistics and p -values are known, however, the theoretical functional forms quantifying the dependence between the tests of the basic pair and that of the empowered pair hindering a complete analytical approach for obtaining the exact p -values of $[R, L]$ and $\{P_{DB}\}$, are yet to be explored. For the sake of self-containedness, theoretical results for the order statistics of a random sample of size 2 from a $\chi^2(k)$ and from $U(0, 1)$, respectively, are reproduced from Chapter 2 and are listed in Table 4.1. To further assess the utilities of the p -value proxies, we generate samples of performance data, measured by the mean absolute deviation (MAD) and their corresponding variance, between the empirical cumulative probabilities ($p = 0.005 - 0.995$) of P_{DB} and their proxies for sample sizes $n = 6(10)46$, failure-truncated sampling, and summarized the results in Table 4.2 for quality assurance before the implementations. Results surrogated by the order statistics reveal that the degree of dependence between the test statistics is generally minor, especially for moderate to small p -values. In particular, for $p \leq 0.1$ where a statistical significance

threshold is normally set, the MADs are confirmed to be $\approx 0.5\%$ for all the cases, which hopefully would be acceptable for most of the modern practitioners. Consequently, this user-friendly test version can be used for sample sizes beyond 50 for the anticipated power comparisons between P_{DB} and L_P . Again, the empirical null distributions of the test statistics estimated by the simulations and the theories of the order statistics jointly make this option possible.

Table 4.2. Mean absolute deviation (variance) between the empirical cumulative probabilities ($p = 0.005 - 0.995$) of P_{DB} and their proxies for sample sizes $n = 6(10)46$, failure-truncated sampling.

p	n					
	6	16	26	36	46	
0.0050	0.0055	0.0055	0.0054	0.0054	0.0055	
0.0100	0.0111	0.0110	0.0108	0.0109	0.0110	
0.0250	0.0285	0.0280	0.0274	0.0277	0.0279	
0.0500	0.0578	0.0573	0.0558	0.0563	0.0565	
0.1000	0.1178	0.1154	0.1147	0.1148	0.1145	
0.2000	0.2380	0.2330	0.2311	0.2305	0.2314	
0.3000	0.3550	0.3462	0.3451	0.3439	0.3446	
0.4000	0.4663	0.4564	0.4545	0.4527	0.4535	
0.5000	0.5715	0.5613	0.5579	0.5552	0.5574	
0.6000	0.6697	0.6570	0.6552	0.6529	0.6548	
0.7000	0.7583	0.7461	0.7449	0.7424	0.7445	
0.8000	0.8385	0.8276	0.8267	0.8243	0.8260	
0.9000	0.9087	0.9003	0.8999	0.8991	0.9000	
0.9500	0.9415	0.9357	0.9357	0.9346	0.9357	
0.9750	0.9587	0.9544	0.9536	0.9532	0.9533	
0.9900	0.9705	0.9668	0.9661	0.9659	0.9657	
0.9950	0.9752	0.9719	0.9713	0.9713	0.9710	
MAD	(All p)	0.0287 (1.07e-03)	0.0253 (8.67e-04)	0.0247 (8.30e-04)	0.0239 (7.85e-04)	0.0245 (8.20e-04)
	$(p \geq 0.9)$	0.0090 (1.43e-04)	0.0073 (9.79e-05)	0.0076 (1.01e-04)	0.0072 (9.49e-05)	0.0078 (1.05e-04)
		0.0053 (5.07e-05)	0.0047 (3.81e-05)	0.0044 (3.50e-05)	0.0044 (3.52e-05)	0.0043 (3.32e-05)
	$(p \leq 0.1)$					

4.2 Power Comparisons

The following power study is consistently for testing H_0 : a simple HPP versus H_A : a more general NHPP with a two-sided nominal level of $\alpha = 0.1$, where the distribution under H_A will be a known NHPP. Proxies developed above are used only for simulations of sizes above 50. The structure of the Monte Carlo simulations starts with a failure-truncated sampling scheme with 10^4 iterations for each alternative distribution. Tables from 4.3 to 4.7 summarize the results of this part. As the study unfolds, we then extend some of the structures to a time-truncated sampling scheme to complete the power comparisons. To bridge the gap by following in the footsteps of Bain et al. (1985) [10], and Engelhardt et al. (1990) [33], every setting in the simulation composes an element from each of the components: (a) an intensity differentiated by type, trend, and the location/number of change point(s), (b) sample size, and (c) sampling schemes. Three types of alternative distribution to be evaluated are: a PLP with $\beta \neq 1$ and an NHPP with a two/three-step-intensity function as organized in the simulation flowchart (Figure 4.1). Composed elements associated with each setting will be furnished as a footnote wherever applicable.

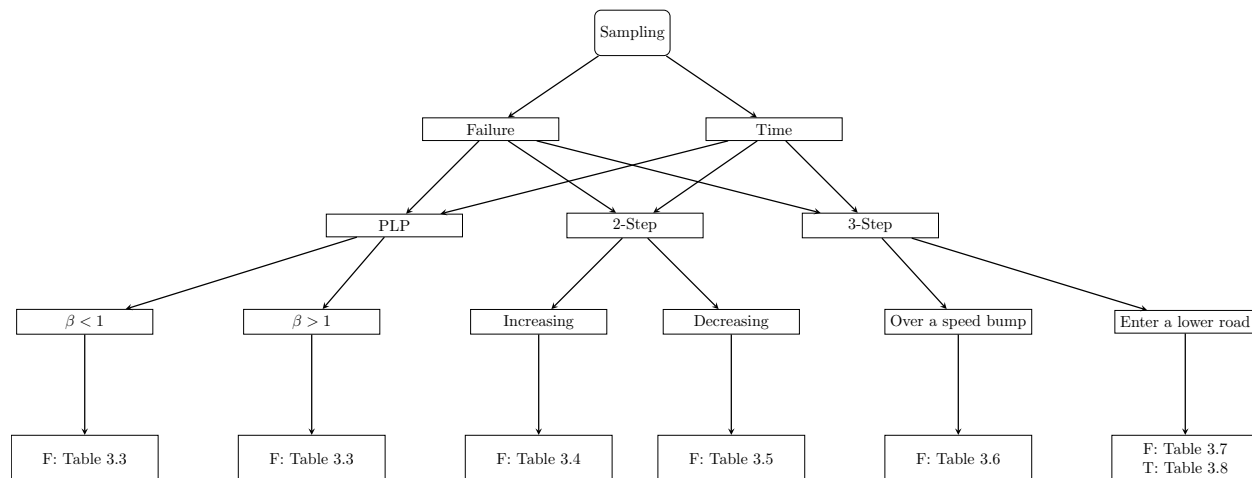


Figure 4.1. Simulation flowchart: (Sampling) schemes \rightarrow {(F)ailure-truncated sampling, (T)ime-truncated sampling} \rightarrow {PLP, (2-step) intensity, (3-step) intensity}, with simulation results from Tables 4.3 to 4.8.

4.2.1 Power Law Process

In each case of the first simulation set, H_A is a PLP with a $\beta = (0.1 - 3.0)$, respectively, and the total number of failures simulated/observed are $n = 5, 15, 25$, and 40 , respectively, listed in Table 4.3 which summarizes the estimated powers. In addition to a measurement of the averaged power, a competitive indicator, termed as power-deficit, where the absolute difference between two estimated powers is quantified for additional performance assessments.

Table 4.3. Estimated powers for testing H_0 versus H_A : a power law process with $\beta = (0.1 - 3.0)$ for $n = 5, 15, 25$, and 40 .

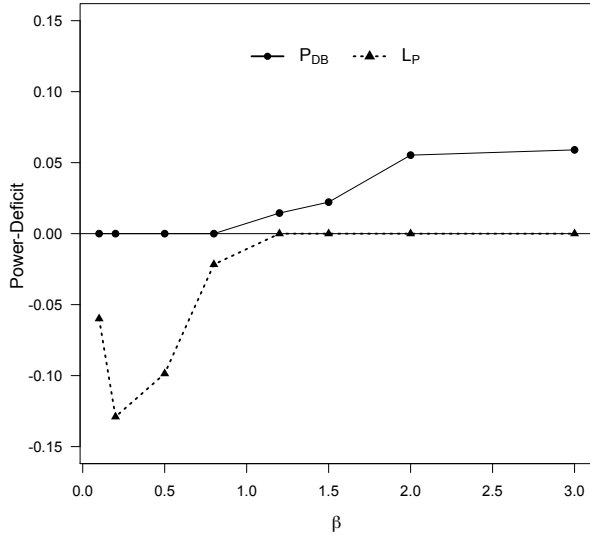
n	Test	β									Overall Average	Power Deficit
		0.1	0.2	0.5	0.8	1.0	1.2	1.5	2.0	3.0		
	Simulation	1	2	3	4		5	6	7	8		
5	P_{DB}	0.9918	0.9188	0.4498	0.1627	0.1011	0.0994	0.1375	0.2368	0.5164	0.4391 (1) ^a	0
	L_P	0.9319	0.7898	0.3512	0.1410	0.1020	0.1139	0.1597	0.2921	0.5754	0.4193 (2)	0.0198
	Simulation	9	10	11	12		13	14	15	16		
15	P_{DB}	1.0000	0.9999	0.8147	0.2268	0.0990	0.1335	0.3275	0.7213	0.9861	0.6512 (1)	0
	L_P	1.0000	0.9963	0.6951	0.1980	0.1008	0.1486	0.3504	0.7325	0.9845	0.6381 (2)	0.0131
	Simulation	17	18	19	20		21	22	23	24		
25	P_{DB}	1.0000	1.0000	0.9420	0.3031	0.1003	0.1843	0.5303	0.9277	1.0000	0.7359 (1)	0
	L_P	1.0000	1.0000	0.8737	0.2602	0.1019	0.1906	0.5277	0.9251	1.0000	0.7221 (2)	0.0138
	Simulation	25	26	27	28		29	30	31	32		
40	P_{DB}	1.0000	1.0000	0.9910	0.3875	0.0988	0.2394	0.7266	0.9937	1.0000	0.7922 (1)	0
	L_P	1.0000	1.0000	0.9674	0.3349	0.1013	0.2447	0.7145	0.9904	1.0000	0.7814 (2)	0.0108

^a Rankings per the averaged power.

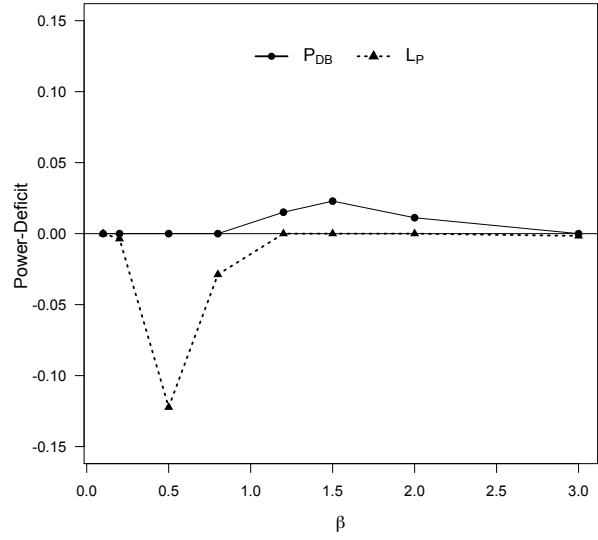
First, values at $\beta = 1$ are coherently close to 0.1 for all cases, which will be excluded from further power comparisons. Power-deficit-based performance comparisons depicted in Figure 4.2 conclude that P_{DB} is a clear winner in round one, where the sign of one set of the power-deficits has been reversed to facilitate the comparisons. In the contests with PLP settings, the maximum power-deficit for P_{DB} is about 0.06 and about 0.12 for L_P on multiple occasions. Interestingly enough, L_P is consistently a laggard regardless of sample sizes for a decreasing PLP ($\beta < 1$) alternative – the first sign of minor power asymmetries shared by both tests of the empowered version, $[R, L]$, at

the same PLP settings in Chapter 3.

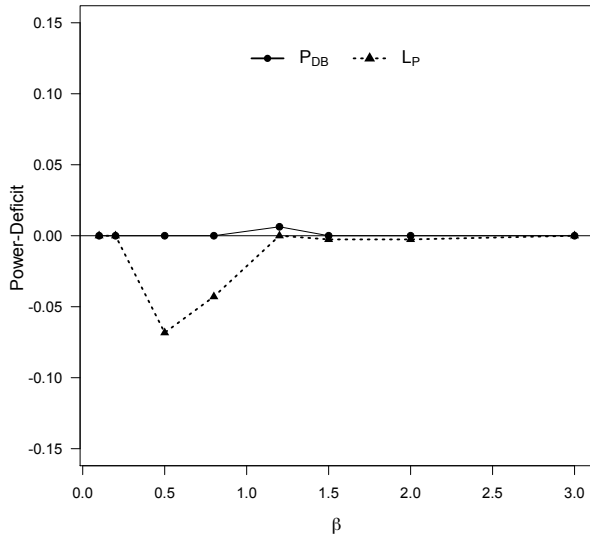
It is encouraging that at $\beta = 0.2$, a speedy improvement for the system, it only takes 5 samples for the P_{DB} to confirm the development with an estimated power of 0.9188 versus 0.7898 for L_P . On the other hand, at $\beta = 1.5$ as the system is deteriorating, the sample size at 40 lifts the power to about 72% for both from the low thirties, where n is only 15 (Table 4.3). Recalled that the PLP is a widely used point process for repairable systems reliability, and the Z is proven as a UMP test. Therefore, P_{DB} has an edge in this setting, and the race continues for an alternative with pieces of HPPs, divided by one or two change points – a step-function intensity.



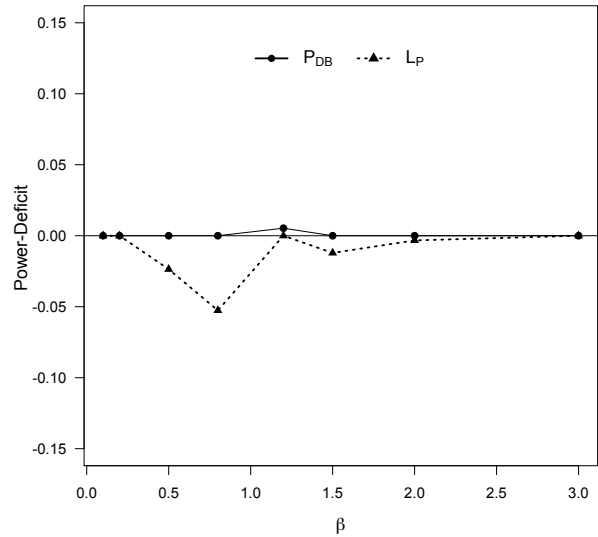
(a) $n = 5$



(b) $n = 15$



(c) $n = 25$



(d) $n = 40$

Figure 4.2. Power-deficit-based performance comparisons for testing H_0 versus H_A : a PLP with $\beta \neq 1$ for (a) $n = 5$, (b) $n = 15$, (c) $n = 25$, and (d) $n = 40$.

4.2.2 Two-Step-Intensity

First, we consider two triplets of samples (10, 20, 40) and (60, 80, 100). For each triplet, a two-step-intensity of the form $\lambda(t) = 1$ for $0 < t \leq \tau$, and $\lambda(t) = 3$ for $\tau < t < \infty$ is considered with various

values of τ to form three groups per triplet. Combinations of [Group (sample size, location of τ)] are as follows: [1 (10, 5) (20, 10) (40, 20)], [2 (10, 4) (20, 7) (40, 14)], and [3 (10, 6) (20, 13) (40, 26)]; [1 (60, 30) (80, 40) (100, 50)], [2 (60, 24) (80, 28) (100, 35)], and [3 (60, 36) (80, 52) (100, 65)]. Table 4.4 summarizes the estimated powers for alternatives of all the increasing two-step-intensity functions. In this settings, the maximum power-deficit for P_{DB} is about 0.09 and about 0.07 for L_P with a random pattern. If the location of change is set approximately two-thirds into the process, the P_{DB} -test has an edge on this increasing trend. L_P , however, shows higher powers for jumps set earlier in the process. Power gains for either are not impressive.

Table 4.4. Estimated powers for testing H_0 versus H_A : a two-step-intensity with change point at τ and $\lambda_1 = 1$, $\lambda_2 = 3$.

Simulation	33	34	35	36	37	38	39	40	41		
	Group ^a										
	1			2			3				
Test	$n = 10$ (4.9540) ^c	$n = 20$ (10.0085)	$n = 40$ (19.9476)	$n = 10$ (4.0297)	$n = 20$ (7.0081)	$n = 40$ (13.9469)	$n = 10$ (5.9691)	$n = 20$ (12.9367)	$n = 40$ (25.9515)	Overall Average	Power Deficit
P_{DB}	0.3249	0.5091	0.7837	0.3256	0.5464	0.8039	0.2794	0.4237	0.6883	0.5205 (2)	0.0293
L_P	0.3403	0.5544	0.8223	0.3716	0.6403	0.8965	0.2893	0.4136	0.6204	0.5498 (1)	0
Simulation	42	43	44	45	46	47	48	49	50		
	Group ^b										
	1			2			3				
Test	$n = 60$ (29.9955)	$n = 80$ (39.9916)	$n = 100$ (49.9304)	$n = 60$ (24.0712)	$n = 80$ (28.0569)	$n = 100$ (35.0610)	$n = 60$ (36.0166)	$n = 80$ (52.0049)	$n = 100$ (64.9710)	Overall Average	Power Deficit
P_{DB}	0.9124	0.9646	0.9869	0.9175	0.9706	0.9901	0.8635	0.9102	0.9583	0.9415 (2)	0.0012
L_P	0.9351	0.9752	0.9914	0.9661	0.9944	0.9987	0.8449	0.8570	0.9217	0.9427 (1)	0

^a Combinations of [Group (sample size, location of τ)] are as follows: [1 (10, 5) (20, 10) (40, 20)], [2 (10, 4) (20, 7) (40, 14)], and [3 (10, 6) (20, 13) (40, 26)].

^b Combinations of [Group (sample size, location of τ)] are as follows: [1 (60, 30) (80, 40) (100, 50)], [2 (60, 24) (80, 28) (100, 35)], and [3 (60, 36) (80, 52) (100, 65)].

^c The estimated mean number of occurrences in $[0, \tau]$.

Table 4.5. Estimated powers for testing H_0 versus H_A : a two-step-intensity with change point at τ and $\lambda_1 = 3, \lambda_2 = 1$.

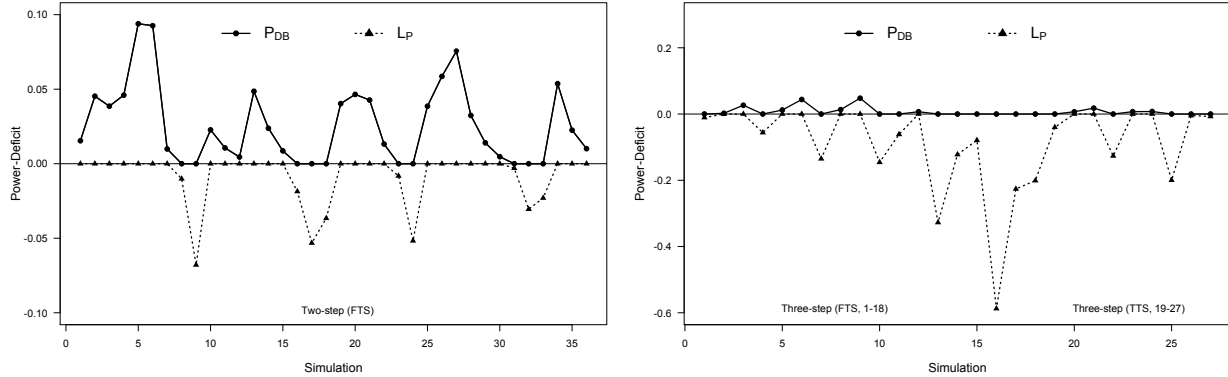
Simulation	51	52	53	54	55	56	57	58	59		
	Group ^a										
	1			2			3				
Test	$n = 10$ (4.9779) ^c	$n = 20$ (9.9276)	$n = 40$ (20.0133)	$n = 10$ (4.0009)	$n = 20$ (7.0411)	$n = 40$ (14.0155)	$n = 10$ (5.9026)	$n = 20$ (12.9583)	$n = 40$ (25.9786)	Overall Average	Power Deficit
P_{DB}	0.3047	0.5260	0.7923	0.3156	0.4759	0.7376	0.2999	0.5027	0.7782	0.5258 (2)	0.0284
L_P	0.3450	0.5725	0.8350	0.3288	0.4677	0.6859	0.3385	0.5613	0.8537	0.5542 (1)	0
Simulation	60	61	62	63	64	65	66	67	68		
	Group ^b										
	1			2			3				
Test	$n = 60$ (29.9727)	$n = 80$ (40.0649)	$n = 100$ (49.9598)	$n = 60$ (23.9800)	$n = 80$ (27.9779)	$n = 100$ (34.9649)	$n = 60$ (36.0462)	$n = 80$ (51.9412)	$n = 100$ (65.0260)	Overall Average	Power Deficit
P_{DB}	0.9068	0.9656	0.9878	0.8847	0.9337	0.9717	0.9137	0.9651	0.9873	0.9462 (2)	0.0090
L_P	0.9392	0.9796	0.9925	0.8817	0.9032	0.9487	0.9674	0.9876	0.9974	0.9552 (1)	0

^a Combinations of [Group (sample size, location of τ)] are as follows: [1 (10, 5/3) (20, 10/3) (40, 20/3)], [2 (10, 4/3) (20, 7/3) (40, 14/3)], and [3 (10, 6/3) (20, 13/3) (40, 26/3)].

^b Combinations of [Group (sample size, location of τ)] are as follows: [1 (60, 30/3) (80, 40/3) (100, 50/3)], [2 (60, 24/3) (80, 28/3) (100, 35/3)], and [3 (60, 36/3) (80, 52/3) (100, 65/3)].

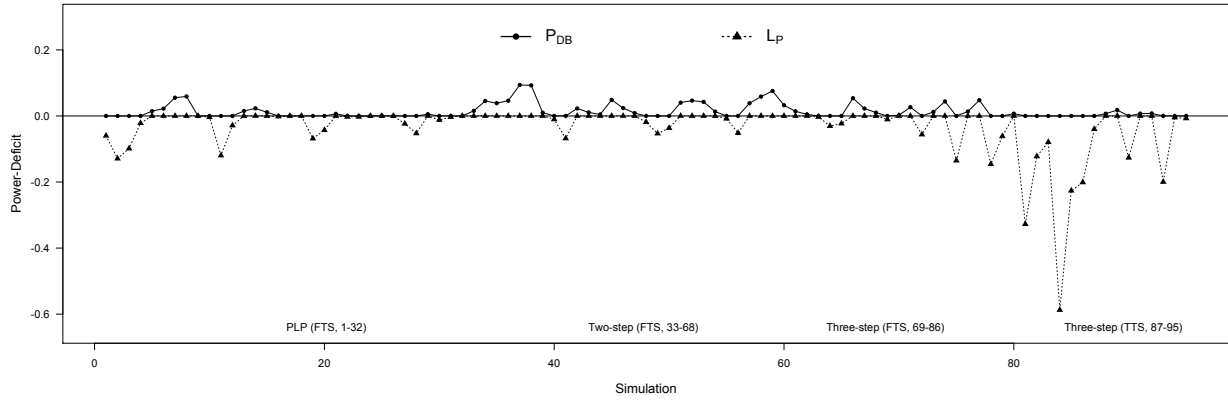
^c The estimated mean number of occurrences in $[0, \tau]$.

A straightforward reversal of a (1, 3) trend to (3, 1) for (λ_1, λ_2) yields Table 4.5. Again, combinations of [Group (sample size, location of τ)] are: [1 (10, 5/3) (20, 10/3) (40, 20/3)], [2 (10, 4/3) (20, 7/3) (40, 14/3)], and [3 (10, 6/3) (20, 13/3) (40, 26/3)]; [1 (60, 30/3) (80, 40/3) (100, 50/3)], [2 (60, 24/3) (80, 28/3) (100, 35/3)], and [3 (60, 36/3) (80, 52/3) (100, 65/3)]. In these settings, the maximum power-deficit for P_{DB} is about 0.08, and about 0.05 for L_P with a random pattern as well. In contrast to the previous observations, if the location of change is set approximately two-thirds into the process, L_P has a slight edge on this decreasing trend – a so-called symmetrical performance justified by P_{DB} showing higher powers for jumps set earlier in the process. Figure 4.3(a) presents the results of the power-deficit-based performance comparisons for alternatives, including both the increasing and decreasing two-step-intensity functions. L_P is slightly better per the overall power averages and power deficits presented in both tables and Figure 4.3(a) – we consider L_P wins by a narrow margin.



(a) Two-step-intensity

(b) Three-step-intensity



(c) All simulations

Figure 4.3. Power-deficit-based performance comparing tests of P_{DB} and L_P for (a) two-step-intensity, (b) three-step-intensity, and (c) all simulations (failure-truncated sampling, FTS; time-truncated sampling, TTS).

4.2.3 Nonmonotonic Three-Step-Intensity

Here, in Table 4.6, we consider a triplet of samples (36, 72, 144), where a three-step-intensity of the form $\lambda(t) = 1$ for $0 < t \leq \tau_1$, $\lambda(t) = 3$, for $\tau_1 < t \leq \tau_2$, and $\lambda(t) = 1$, for $\tau_2 < t < \infty$ is considered with various values of τ_1 and τ_2 to form three groups for each sample size in the triplet. Combinations of [Group (location of τ_1, τ_2)] are labeled as follows: for $n = 36$, [1 (12, 16)], [2 (9, 12)], and [3 (18, 21)]; for $n = 72$, [1 (24, 32)], [2 (18, 24)], and [3 (36, 42)]; for $n = 144$, [1 (48, 64)], [2 (36, 48)], and [3 (72, 84)]. The maximum power-deficit for P_{DB} is 0.0478, and 0.1351 for

L_P (Table 4.6).

Table 4.6. Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$.

Simulation	69	70	71	72	73	74	75	76	77		
	$\lambda_1 = 1, \lambda_2 = 3, \text{ and } \lambda_3 = 1$										
	Group ^a			Group ^b			Group ^c				
	1	2	3	1	2	3	1	2	3		
	$n = 36$ (11.9792) ^d (11.9541) ^e	$n = 36$ (8.9939) (8.9872)	$n = 36$ (17.9402) (8.9373)	$n = 72$ (23.9276) (24.0007)	$n = 72$ (18.0683) (18.0437)	$n = 72$ (35.9773) (18.0044)	$n = 144$ (48.0819) (48.0964)	$n = 144$ (35.9723) (36.0197)	$n = 144$ (72.1291) (35.9708)	Overall Average	Power Deficit
P_{DB}	0.1393	0.1111	0.1703	0.1720	0.1500	0.2008	0.2551	0.2362	0.2646	0.1888 (1)	0
L_P	0.1287	0.1133	0.1968	0.1159	0.1642	0.2447	0.1200	0.2495	0.3124	0.1828 (2)	0.0060

^a Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (12, 16)], [2 (9, 12)], and [3 (18, 21)].

^b Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (24, 32)], [2 (18, 24)], and [3 (36, 42)].

^c Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (48, 64)], [2 (36, 48)], and [3 (72, 84)].

^d The estimated mean number of occurrences in $[0, \tau_1]$.

^e The estimated mean number of occurrences in $(\tau_1, \tau_2]$.

Similarly, in Table 4.7, the corresponding $\lambda(t)$ is $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ with various values of τ_1 and τ_2 as well to form three groups for each sample size in the same triplet, (36, 72, 144). Combinations of [Group (location of τ_1, τ_2)] are as follows: for $n = 36$, [1 (4, 16)], [2 (3, 12)], and [3 (6, 15)]; for $n = 72$, [1 (8, 32)], [2 (6, 24)], and [3 (12, 30)]; for $n = 144$, [1 (16, 64)], [2 (12, 48)], and [3 (24, 60)]. The maximum power-deficit for P_{DB} is 0.0069, and 0.5873 for L_P – a technical knockout! Figure 4.3(b) is a graphical version of Tables 4.6 and 4.7 combined, which depicts that the dual bidirectional P_{DB} -test is competitive for detecting abrupt change(s) in a series of events, that frequently arise in real-world applications.

Table 4.7. Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$.

Simulation	78	79	80	81	82	83	84	85	86		
	$\lambda_1 = 3, \lambda_2 = 1$, and $\lambda_3 = 3$										
	Group ^a			Group ^b			Group ^c				
	1	2	3	1	2	3	1	2	3		
	$n = 36$ (12.0539) ^d (11.9918) ^e	$n = 36$ (8.9952) (8.9619)	$n = 36$ (17.9703) (8.9219)	$n = 72$ (23.9902) (23.9291)	$n = 72$ (17.9669) (18.0270)	$n = 72$ (35.9775) (18.0046)	$n = 144$ (48.1045) (48.0327)	$n = 144$ (35.8948) (35.9661)	$n = 144$ (71.8981) (35.9918)	Overall Average	Power Deficit
P_{DB}	0.3862	0.3216	0.3485	0.5659	0.4769	0.5023	0.8269	0.7480	0.7467	0.5470 (1)	0
L_P	0.2405	0.2603	0.3554	0.2385	0.3546	0.4231	0.2396	0.5219	0.5457	0.3532 (2)	0.1938

^a Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (4, 16)], [2 (3, 12)], and [3 (6, 15)].

^b Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (8, 32)], [2 (6, 24)], and [3 (12, 30)].

^c Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (16, 64)], [2 (12, 48)], and [3 (24, 60)].

^d The estimated mean number of occurrences in $[0, \tau_1]$.

^e The estimated mean number of occurrences in $(\tau_1, \tau_2]$.

A serious setback for L_P happens at the following setting: $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ for $n = 144$ with values of τ_1 and τ_2 as [1 (16, 64)]. This nonmonotonic intensity function with three steps provides additional insight into the power study. Surprisingly, for a large sample size of 144, L_P is not prepared for these situations, while P_{DB} enjoys a relatively high power of 0.8269 (Table 4.7). These pieces of HPPs would certainly be declared as an HPP by the L_P approximately 76% of the time with unexpected consequences, should the data be real, say, reliability data. Moreover, for the same setting and by increasing the sample size from 36 to 144 (Table 4.7), the estimated power for P_{DB} jumps from 0.3862 to 0.8269 while L_P stalls at about 0.24.

Table 4.8. Estimated powers for testing H_0 versus H_A : a three-step-intensity with λ_i and change points at τ_j , where $i = 1, 2$, and 3 ; $j = 1, 2$, time-truncated sampling.

Simulation	87	88	89	90	91	92	93	94	95		
	$\lambda_1 = 3, \lambda_2 = 1$, and $\lambda_3 = 3$										
	Group ^a										
	1	2	3	1	2	3	1	2	3		
	$n = 40$ (17.1474) ^b (5.7180) ^c	$n = 40$ (11.9949) (4.0055)	$n = 40$ (24.0037) (3.9910)	$n = 80$ (34.3850) (11.3578)	$n = 80$ (24.0245) (7.9981)	$n = 80$ (47.9409) (8.0061)	$n = 120$ (51.3958) (17.1903)	$n = 120$ (35.9425) (12.0553)	$n = 120$ (72.0252) (11.9601)	Overall Average	Power Deficit
P_{DB}	0.1902	0.1662	0.1691	0.2630	0.2179	0.2132	0.3456	0.2679	0.2776	0.2345 (1)	0
L_P	0.1503	0.1730	0.1869	0.1366	0.2250	0.2208	0.1459	0.2633	0.2703	0.1969 (2)	0.0376

^a Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (1/3, 2/3)], [2 (1/4, 1/2)], and [3 (1/2, 3/4)].

^b The estimated mean number of occurrences in $[0, \tau_1]$.

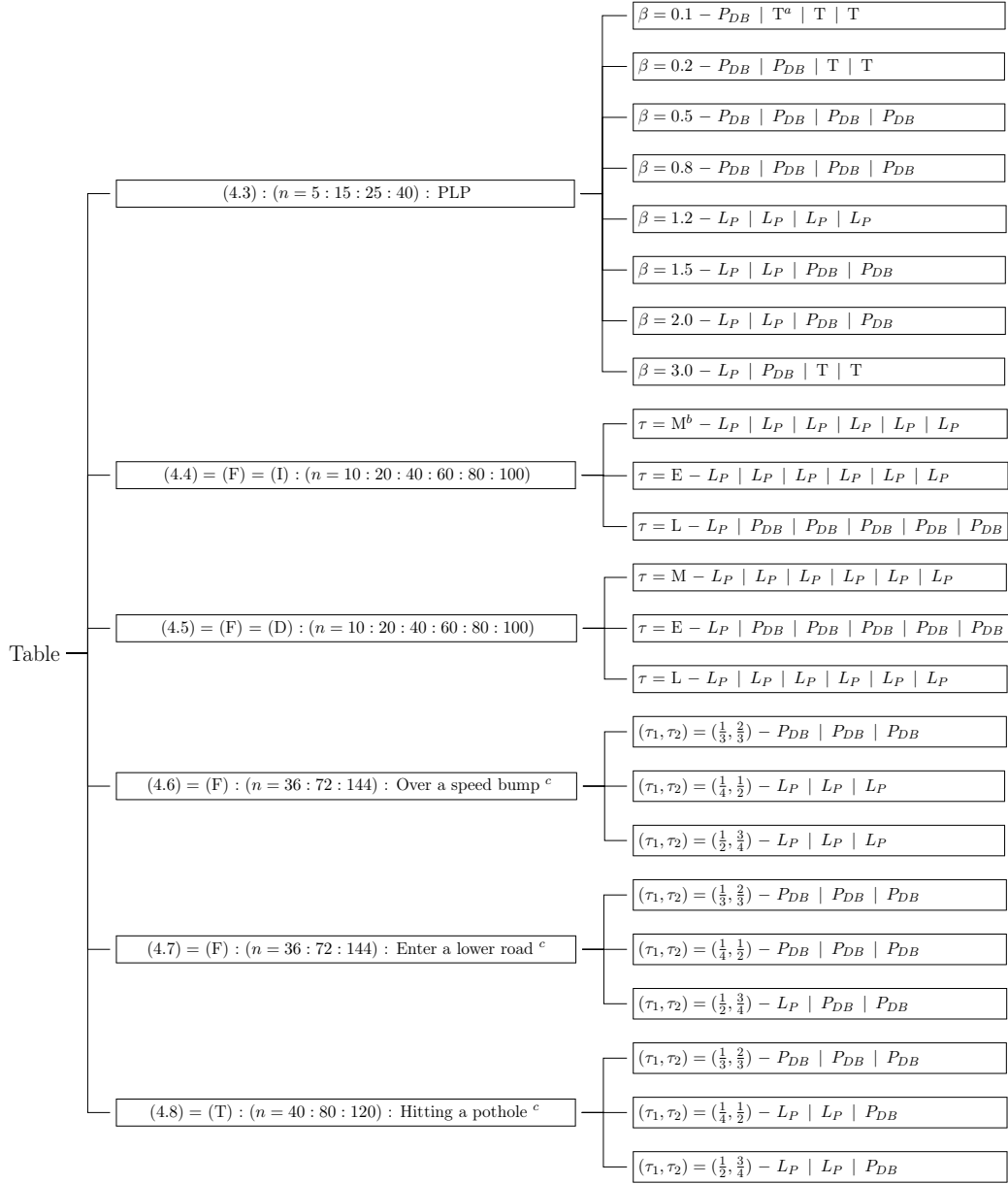
^c The estimated mean number of occurrences in $(\tau_1, \tau_2]$.

Filled with curiosity, we extend the above settings by considering a triplet of samples (40, 80, 120) for the time-truncated sampling. Similarly, in Table 4.8, the corresponding $\lambda(t)$ is $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ with various values of τ_1 and τ_2 also to form three groups for each sample size in the triplet. Combinations of [Group (location of τ_1, τ_2)] are as follows: [1 (1/3, 2/3)], [2 (1/4, 1/2)], and [3 (1/2, 3/4)]. In contrast to the above results, the maximum power-deficit for P_{DB} is 0.0178 and 0.1997 for L_P , respectively, which catch our attention. For example, for $n = 120$ with values of τ_1 and τ_2 as [1 (1/3, 2/3)]. In this case, where $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$, the change points to partition the entire process into three pieces of equal length on a timescale. The unit rate of the middle segment ($\lambda_2 = 1$) is a third of its adjacent segments ($\lambda_1 = \lambda_3 = 3$). Thus, the total number of the expected occurrences for its neighboring segments is three times as many. Consequently, the middle piece takes about 17 out of 120 (Table 4.8). Therefore, testing a hypothesis like this may be perceived as hitting a pothole, a minor setback for both tests, where the estimated power is only 0.3456 and 0.1459 for P_{DB} and L_P , respectively. Moreover, for the same setting and by tripling the sample size from 40 to 120 (Table 4.8), the estimated power for P_{DB} jumps from 0.1902 to 0.3456 while L_P stalls at about 0.15 for all three simulations.

Thus, time-truncated sampling studies have given us added awareness of what stands for performance in hypothesis testing for a series of events. We, therefore, see merit in both of the simulation schemes. Anatomy based on a body of performance data, consolidating the above six Tables (from 4.3 to 4.8), provides additional insights into the behavior of both tests.

4.3 The Verdict

Firstly, Figure 4.4 serves as a structure and performance overview of all the simulations conducted in this study. Summarizing tables with like settings described in the adjacent twigs are stacked in six large branches: (a) an ordered pair, “(4.3) : ($n = 5 : 15 : 25 : 40$) : PLP,” symbolizes the respective



^a T = a tie.

^b Location of change: (a) E = early into a process, (b) M = middle, and (c) L = late.

^c Over a speed bump = $(\lambda_1, \lambda_2, \lambda_3) = (1, 3, 1)$; Enter a lower road (F) = $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ = Hitting a pothole (T).

Figure 4.4. Structure overview – simulation: failure-truncated sampling (F: Tables 4.3 – 4.7); time-truncated sampling (T: Table 4.8).

sample size conducted in Table 4.3 for a PLP alternative, and so forth; (b) “(4.4) = (F) = (I)” indicates the corresponding sampling scheme conducted in Table 4.4 is a failure-truncated sampling scheme (F). (Likewise, use T otherwise.) And, “(I)” notes that the trend adopted in the simulation

for the intensity is monotonically increasing (or otherwise, use “D” for the decreasing trend); (c) Consequently, each twig also stacks executed simulations by the same order as structured, where the total number of simulations accrued are affixed with divided cells; (d) each cell houses the winner of each simulation, denoted as, say, L_P . Additionally, “T” represents a tie.

Table 4.9. Power-deficit overview.

Table 4.3											
Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P
1	0	0.0599	9	0	0	17	0	0	25	0	0
2	0	0.1290	10	0	0.0036	18	0	0	26	0	0
3	0	0.0986	11	0	0.1196	19	0	0.0683	27	0	0.0236
4	0	0.0217	12	0	0.0288	20	0	0.0429	28	0	0.0526
5	0.0145	0	13	0.0151	0	21	0.0063	0	29	0.0053	0
6	0.0222	0	14	0.0229	0	22	0	0.0026	30	0	0.0121
7	0.0553	0	15	0.0112	0	23	0	0.0026	31	0	0.0033
8	0.0590	0	16	0	0.0016	24	0	0	32	0	0

Table 4.4						Table 4.5					
Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P
33	0.0154	0	42	0.0227	0	51	0.0403	0	60	0.0324	0
34	0.0453	0	43	0.0106	0	52	0.0465	0	61	0.0140	0
35	0.0386	0	44	0.0045	0	53	0.0427	0	62	0.0047	0
36	0.0460	0	45	0.0486	0	54	0.0132	0	63	0	0.0030
37	0.0939	0	46	0.0238	0	55	0	0.0082	64	0	0.0305
38	0.0926	0	47	0.0086	0	56	0	0.0517	65	0	0.0230
39	0.0099	0	48	0	0.0186	57	0.0386	0	66	0.0537	0
40	0	0.0101	49	0	0.0532	58	0.0586	0	67	0.0225	0
41	0	0.0679	50	0	0.0366	59	0.0755	0	68	0.0101	0

Table 4.6			Table 4.7			Table 4.8		
Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P	Simulation	P_{DB}	L_P
69	0	0.0106	78	0	0.1457	87	0	0.0399
70	0.0022	0	79	0	0.0613	88	0.0068	0
71	0.0265	0	80	0.0069	0	89	0.0178	0
72	0	0.0561	81	0	0.3274	90	0	0.1264
73	0.0120	0	82	0	0.1223	91	0.0071	0
74	0.0439	0	83	0	0.0792	92	0.0076	0
75	0	0.1351	84	0	0.5873	93	0	0.1997
76	0.0133	0	85	0	0.2261	94	0	0.0046
77	0.0478	0	86	0	0.2010	95	0	0.0073

Average ^a	P_{DB}	L_P
	0.0138 (1)	0.0347 (2)

^aAveraged power-deficit under all the simulations.

Furthermore, Table 4.9 accrues all the power-deficits made for performance assessment. Graphic versions of Table 4.9 are presented in Figure 4.3(c) and Figure 4.5, which undoubtedly claim that P_{DB} is robust and competitive.

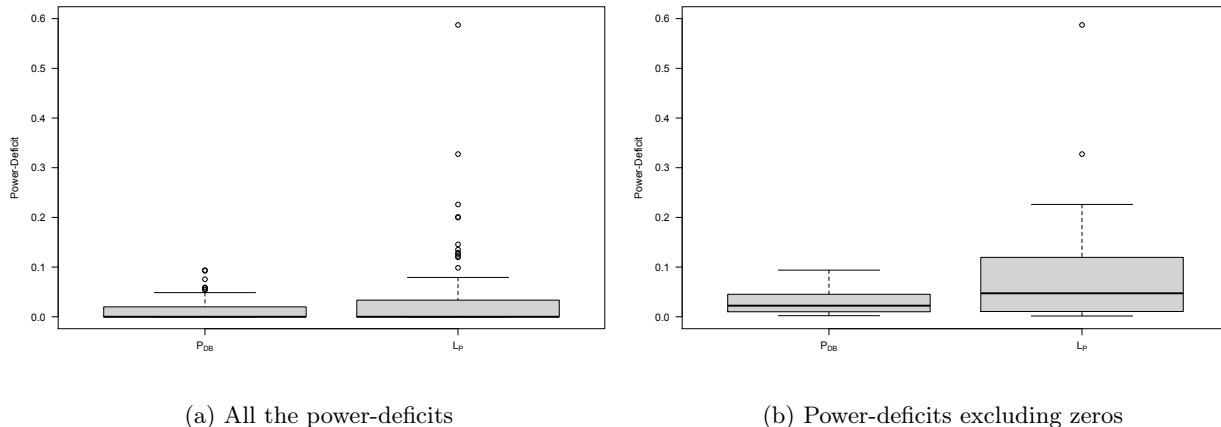


Figure 4.5. Box-plots for power-deficits: (a) all the power-deficits and (b) power-deficits excluding zeros.

4.4 Applications

The motivation and goal of this study is to extend the existing tools and present an alternative accommodating essential quality assurance for the practitioners conducting the work. A straightforward application of the proposed cocktail of tests is to extend a single bidirectional control chart (Ho 1998) to a panel of control charts, because good graphics are vital in data analysis. They are essential in communicating the analysis to others.

For illustration, consider the failure times of an aircraft generator shown in Table 4.8 of Rigdon and Basu (2000) [81]. These failure times, t_i ($= 10, 55, 166, 205, 341, 488, 567, 731, 1308, 2050, 2453, 3115, 4017, \text{ and } 4596$) were read from a figure in Duane (1964) [31]. A panel of two control charts, P_{DB} -chart and L_p -chart for the above data set, are presented in Figure 4.6. The control charts are both for testing $H_0 : \beta = 1$ (the system's reliability remains stable) versus H_A , with an applicable 90% control limit proxies. In general, for a real data set, the distribution under H_A would be an unknown NHPP indicating that the system's reliability has grown or decreased during the observed period.

In this application, the data are real, one of the tests is an approximated normal test while the other is a proxy. First, both control charts produce the same key result, indicating the first out-of-control signal at t_9 – an indication of the techniques’ applicability, reliability, and quality assurance. Additionally, the L_p -chart confirms a decreasing (improving) trend for the reliability data. This is one of the best ways to get the message over clearly without misunderstanding.

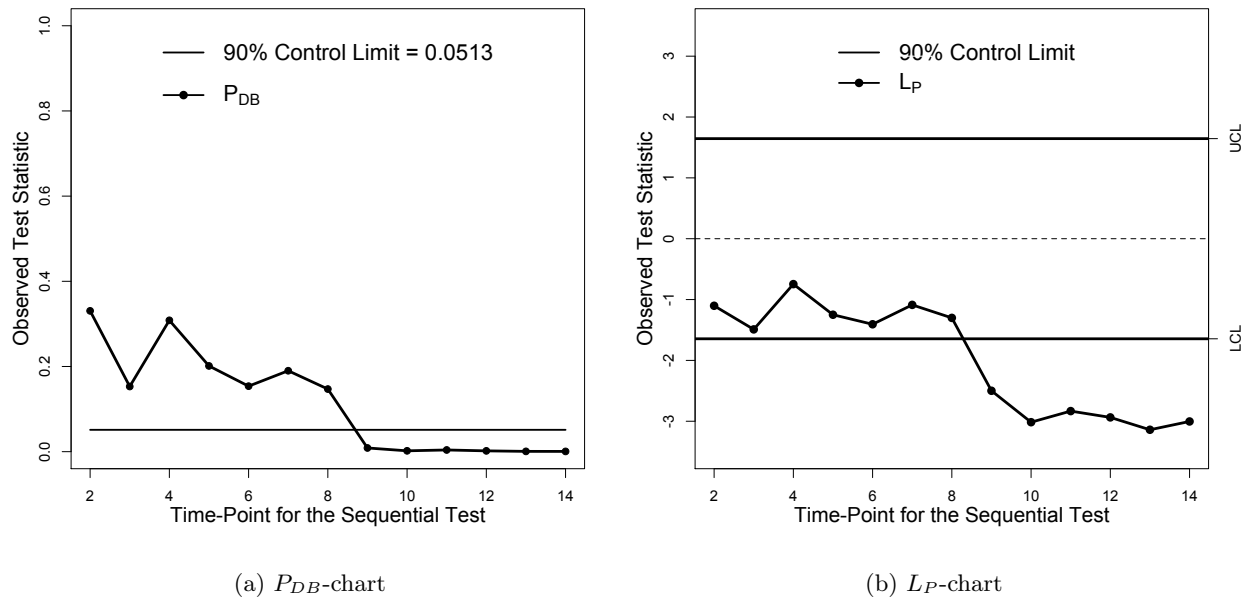


Figure 4.6. Panel of (a) P_{DB} -chart and (b) L_P -chart for failure times of an aircraft generator.

4.5 Conclusions – An All-Purpose Test

We have already witnessed the tug-of-war between the opponents for powers through 95 simulations which, on the contrary, pose some practical concerns for the practitioners conducting the Laplace test. Every scenario in the study involves a known underlying alternative distribution, 10^4 data sets with fixed sample size, and a test. The associated outcome is recorded as significant or otherwise. Unlike in a real-data application, the scenario is reproduced 10^4 times in the same fashion for a simulation-based study, and a relative frequency of significant outcomes is reported as an estimated

test power. In our study, the sample size is extended for each scenario to include a plausible range for power trend analyses. In addition to presenting a significant result building upon the performance of the legendary Laplace test, what is outlined below would be of practical significance to practitioners doing the repairable systems reliability work or to statistical consultants helping with the projects through to their completions.

- Three key factors are essential to point out: data distribution, sample size, and the test performed. For instance, at one point, the L_P -test bears a deficit in power at a magnitude of 0.5873. At a relatively large sample size of 144, a process having three pieces of HPPs would indeed be declared as an HPP by the L_P approximately 76% of the time, should the data be real, say, reliability data. Recall that, for the same scenario and by increasing the sample size from 36 to 144 (Table 4.7), the estimated power for P_{DB} jumps from 0.3862 to 0.8269 while L_P stalls at about 0.24. The implication is that increasing sample size typically increases power to a degree, if there are no mismatches between the adopted test and the true underlying distribution of the data. However, the underlying models are mostly unknown beforehand but usually determined by the outcomes of the tests, p -values, for instance. Therefore practitioners practicing multiple tests would find themselves struggling with the scenarios leading to disparate outcomes, should the data be real. Thus, this study provides additional insights and leads off a future work galvanizing the issues.
- On a more positive note, it is encouraging that at $\beta = 0.2$ for a PLP alternative distribution, a speedy improvement for the system, it only takes 5 samples for the P_{DB} to confirm the statistical significance of the development of a new system with an estimated power of 0.9188 versus 0.7898 for L_P . Similarly, at $\beta = 1.5$ as the system is deteriorating, increasing the sample size to 40 lifts the power to about 72% for both tests from the low thirties, where n is only 15. Of course, both scenarios would appear circumstantial, should the data be real

and solo. Here, they are backed by the implementations of 10^4 independent and identical processes, which obviously reinforce the importance, the interconnectedness, and the essential reliability and quality assurance of the roles played by the distribution of the data, sample size, and the test performed.

- The gap is now bridged. A cocktail of coherent tests, $(Z, Z_B) \rightarrow [R, L] \rightarrow \{P_{DB}\}$, which is rooted in a basic pair of tests, (Z, Z_B) , that are more complementary than competitive and are bound together with a common set of critical values. Both tests are the exact tests. Another unified and coherent p -value proxies, partners of an empowered pair of tests, $[R, L]$, form a second bidirectional test. A hybrid, $\{P_{DB}\}$, of the empowered pair, characterize the test as the only tool possessing the unique feature of having an intelligible version, grounded in practicality, that is free of either the system sample size or the sampling schemes. Now, we declare that P_{DB} is a so-called all-purposes test, supported by 95 simulations. It is a safe bet for practitioners doing applicable real-world case studies.

Furthermore, functions for p -value proxies are capable of producing an all-in-one control chart, to be considered in the aforementioned areas of applications by extending the statistical process control architecture to a panel of coherent control charts fit for group sequential testing and multisystem.

CHAPTER 5

PANEL OF BIDIRECTIONAL CONTROL CHARTS

5.1 Motivation

Data used to track the reliability growth trend of repairable systems occur naturally in a sequential fashion. It should be obvious that the chronological order in which the events occur is an extremely important aspect of a stochastic process, if a significant time trend is to be objectively identified. It is thus desirable to study some basic properties of repeated significance tests on series of events which come in different forms. Statistical process/quality control is a sophisticated concept, because it recognizes that variability will be present and requires only that the pattern of variability remain the same. A variable (or process) that continues to be described by the same distribution when observed over time is said to be in statistical control, or simply in-control. We are already quite advanced in the art of thinking statistically when we describe a variable as stable or in-control if its distribution does not change with time. The control chart was invented by Walter A. Shewhart working for Bell Labs in the 1920s. Books by Montgomery (1985) [71], Ryan (1989) [84], James (1991) [60], Grant and Leavenworth (1996 [40]), Xie et al. (2002) [97], Smith (2003) [87], DeVor et al. (2006) [28], Wheeler (2010) [94], Luko (2018) [67], Oakland and Oakland (2018) [74], Hardwick (2019) [41], Chandra (2020) [17], and Tran (2021) [91] review much of the work in this area.

Control charts, first developed in the 1920s and 1930s, provide a mechanism for recognizing whether the process is in control. A control chart will be effective if it shows a point outside the control limits almost as soon as the process goes out of control. A basic element of control charting is that data have been collected from the process of interest at a sequence of time points. Depending on the aspect of the process under investigation, some statistic is chosen. The value of this statistic is then calculated for each sample in turn. A traditional control chart then results from plotting

these calculated values over time. If the points on the chart all lie between the two control limits, the process is deemed to be in control. That is, the process is believed to be operating in a stable fashion reflecting only natural random variations. An out-of-control “signal” occurs whenever a plotted point falls outside the limits. This is assumed to be attributable to a significant trend (increasing or decreasing) in the intensity of the process. Designs of control limits so that an in-control process generates very few false alarms, whereas a process not in control quickly gives rise to a point outside the limits are discussed in, but not limited to, Macgregor (1988) [68], Saniga (1989) [85], Montgomery and Mastrangelo (1991) [73], Box and Kramer (1992) [13], Vander et al. (1992) [92], Montgomery et al. (1994) [72], Tatum(1996) [89], Willemain and Runger (1996) [95] Crowder et al. (1997) [25], Palm et al. (1997) [75], Ho (1998) [51], Woodall and Montgomery (1999) [96], Zou et al. (2006) [101], Riaz and Does (2009) [80], Qiu and Zou (2010) [78], Ashton et al. (2015) [7], Colin and Vanhoucke (2015) [20], and Qiu (2020) [77]. Additional statistical process control charting techniques designed for specific purposes are presented in, for example, Ali (2020 [1], 2021 [2]), Centofanti et al. (2021) [15], Zhao and Castillo (2021) [99], Yang and Qiu (2021) [98], and references therein.

Also, there is a strong analogy between the logic of control charting and hypothesis testing. The null hypothesis (H_0) here is that the process is in control. When an in-control process yields a point outside the control limits (an out-of-control signal), a type I error has occurred. Appropriate choice of control limits (corresponding to specifying a rejection region in hypothesis testing) will set this error probability at a desired level. For example, Ho (1998) [51] presents a dual process control chart combining graphical and numerical descriptions of data to monitor the repairable systems reliability, claimed as an all-purposes mechanism for monitoring stochastic processes to identify instability and unusual circumstances. We extend the statistical process control architecture to an all-in-one panel of coherent bidirectional control charts fit for group sequential testing and multisystem.

5.2 The Preliminaries

Suppose we assume that the successive occurrence times of a specific process follow a PLP, and let t_1, \dots, t_n be the first n successive times of events. These times are measured from the beginning of the observation period (cumulative length of time over which the events occur), so $t_1 < t_2 < \dots < t_n$. The following theoretical results (Bassin 1969 [12]; Crow 1974 [23], 1982 [24]; Finkelstein 1976 [37]; Lee and Lee 1978 [65]; Bain and Engelhardt 1980 [8]; Bain and Engelhardt 1991 [9]; Ho 1998 [51]) are useful for constructing the control limits:

$$\hat{\beta}_n = n / \sum_{i=1}^{n-1} \ln(t_n/t_i). \quad (5.1)$$

And, if the process is an HPP, the observed test statistic, $2n/\hat{\beta}_n$, should continue to be described by a chi-square distribution with $2n - 2$ degrees of freedom by drawing the $(1 - \alpha)100\%$ control limits at

$$LCL_\alpha = \text{Lower control limit} = \chi_{\alpha/2}^2(2n - 2), \text{ and} \quad (5.2)$$

$$UCL_\alpha = \text{Upper control limit} = \chi_{1-\alpha/2}^2(2n - 2). \quad (5.3)$$

The next step in examining the process is to plot the statistic [= $2n/\hat{\beta}_n$ or $2 \sum_{i=1}^{n-1} \ln(t_n/t_i)$] against the time order in which the measurements were recorded. Since it requires at least two repose times for the statistical process control at each stage, cumulative sums of log ratios (CSLR, Ho 1998 [51]) can be defined by:

$$S_2 = 2 \ln(t_2/t_1) \quad (5.4)$$

$$S_3 = 2[\ln(t_3/t_1) + \ln(t_3/t_2)] = 2 \sum_{i=1}^2 \ln(t_3/t_i) \quad (5.5)$$

⋮

$$S_l = 2[\ln(t_l/t_1) + \dots + \ln(t_l/t_{l-1})] = 2 \sum_{i=1}^{l-1} \ln(t_l/t_i) = S_{l-1} + 2(l-1) \ln(t_l/t_{l-1}). \quad (5.6)$$

These cumulative sums are plotted over time. That is, at time l of the i th stage, plot a point at height S_l . At the current time point r in the current stage i , the plotted points are $(2, S_2)_i, (3, S_3)_i, \dots, (r, S_r)_i$, if $i > 1$. If at current time r , either $S_r \leq \chi_{\alpha/2}^2(2r - 2)$ or $S_r \geq \chi_{1-\alpha/2}^2(2r - 2)$, the process is judged to be out of control. The first inequality suggests the process has shifted to an increasing time trend at time $r - 1$. Similarly, the second inequality suggests the process has shifted to a decreasing time trend.

Ho (1993) [49] recommends that parallel tests, termed as Z and Z_B (forward and backward, respectively), be performed on the same data set to guard against the step-function intensities of both kinds (increasing and decreasing). Based on this idea, Ho (1998) [51] adds the backward cumulative sums of log ratio (BCSLR) at each time point l ,

$$S_l^B = 2 \sum_{i=1}^{l-1} \ln[t_l / (t_l - t_i)] \quad (5.7)$$

to guard against the change-point Poisson process with an increasing intensity function. Again, these backward test statistics designed for the backward processes are also plotted over time along with the previously defined S_l 's in the same control chart. Therefore, at the current time point r , the plotted points are $\{(l, S_l)\}_1^r$ and $\{(l, S_l^B)\}_1^r$ for the forward process and the backward process, respectively. If at current time, the first signal of either $S_r^{(\cdot)} \leq \chi_{\alpha/2}^2(2r - 2)$ or $S_r^{(\cdot)} \geq \chi_{1-\alpha/2}^2(2r - 2)$, indicates the process is out of control at the r th time point. Of course, the out-of-control signal sent by S_r^B in a so-called dual process control chart should be interpreted in a reversed order.

The dual process control, built into the structure, serves as a springboard for a straightforward extension to a panel of coherent bidirectional control charts. The baseline control chart reviewed above will be referred to as the (Z, Z_B) -chart, consistently with that of Ho (1993 [49], 1998 [51]), and Chapter 3, where a cocktail of tests are studied. The potential benefits of this work are options among a pool of control charts labeled with the corresponding test statistic(s) such as a Z -chart, (R, L) -chart, P_{DB} -chart, and so forth.

We certified P_{DB} , a solo one-size-fits-all test, as robust with essential quality assurance in addition to a user friendly version, grounded in practicality, for the practitioners conducting the work. Therefore, this user-friendly version of the test can be used for the control limits. Per Table 5.1, functions for p -value proxies are inverted and added for the corresponding control limit(s) along with the R code (e.g., Kerns 2010 [63]). Table 5.1 has so many simple yet profound implications. The functions are not abstract - rather, grounded in practicality. Consequently, these user-friendly versions of the tests make the p -values much more attainable, which will be factored into the design of the control charts as the study unfolds.

Table 5.1. Functions for the test statistics (Z and Z_B), the p -value proxies (R , L , and P_{DB}), $(1 - \alpha)100\%$ control limits (Z , Z_B , R , L , and P_{DB}) with R code, and m.l.e. of β for both sampling schemes of size n . The cumulative failure times (t_i 's and t) are recognized as the main series of events.

Statistic	Failure-truncated	Time-truncated
Z	$z = 2 \sum_{i=1}^{n-1} \ln(t_n/t_i)$	$z = 2 \sum_{i=1}^n \ln(t/t_i)$
Z_B	$z_B = 2 \sum_{i=1}^{n-1} \ln(t_n/(t_n - t_i))$	$z_B = 2 \sum_{i=1}^n \ln(t/(t - t_i))$
P -value		
$R = r$	$1 - [F_{\chi^2(2n-2)}(r)]^2$	$1 - [F_{\chi^2(2n)}(r)]^2$
$L = l$	$1 - [1 - F_{\chi^2(2n-2)}(l)]^2$	$1 - [1 - F_{\chi^2(2n)}(l)]^2$
$P_{DB} = p$	$1 - (1 - p)^2$	$1 - (1 - p)^2$
Control Limits		
Z	$\chi_{\alpha/2}^2(2n - 2), \chi_{1-\alpha/2}^2(2n - 2)$	$\chi_{\alpha/2}^2(2n), \chi_{1-\alpha/2}^2(2n)$
Z_B	$\chi_{\alpha/2}^2(2n - 2), \chi_{1-\alpha/2}^2(2n - 2)$	$\chi_{\alpha/2}^2(2n), \chi_{1-\alpha/2}^2(2n)$
R	$\chi_{\sqrt{1-\alpha}}^2(2n - 2)$	$\chi_{\sqrt{1-\alpha}}^2(2n)$
L	$\chi_{1-\sqrt{1-\alpha}}^2(2n - 2)$	$\chi_{1-\sqrt{1-\alpha}}^2(2n)$
P_{DB}	$1 - \sqrt{1 - \alpha}$	$1 - \sqrt{1 - \alpha}$
R Code		
Z	$qchisq(\alpha/2, 2n - 2), qchisq(1 - \alpha/2, 2n - 2)$	$qchisq(\alpha/2, 2n), qchisq(1 - \alpha/2, 2n)$
Z_B	$qchisq(\alpha/2, 2n - 2), qchisq(1 - \alpha/2, 2n - 2)$	$qchisq(\alpha/2, 2n), qchisq(1 - \alpha/2, 2n)$
R	$qchisq(sqrt(1 - \alpha), 2n - 2)$	$qchisq(sqrt(1 - \alpha), 2n)$
L	$qchisq(1 - sqrt(1 - \alpha), 2n - 2)$	$qchisq(1 - sqrt(1 - \alpha), 2n)$
P_{DB}	$1 - sqrt(1 - \alpha)$	$1 - sqrt(1 - \alpha)$
β (m.l.e.)	$n / \sum_{i=1}^{n-1} \ln(t_n/t_i)$	$n / \sum_{i=1}^n \ln(t/t_i)$

5.3 Panel of Fully Sequential Control Charts

In this section, the control charts are consistently for testing $H_0 : \beta = 1$ (the system's reliability remains stable) versus H_A , with applicable 90% control limit proxies, where the distribution will be a known NHPP indicating that the system's reliability has grown or decreased during the observed period. Here, fully sequential control charts are about stability: complete visualization, continuous attention, quick action for a significant change, about charting an essential path forward, a very large category of control charts.

5.3.1 Single System

The test based on Z is UMP for the PLP setting. First, a panel of (Z, Z_B) -chart, (R, L) -chart, and P_{DB} -chart for a simulated PLP with $\beta = 0.2$ and $n = 10$ is presented in Figure 5.1. It is encouraging that at $\beta = 0.2$, a speedy improvement for the system, it only takes 5 samples for all the bidirectional charts to confirm the development of the system. On the other hand, at $\beta = 2.0$ as the system is deteriorating, both Z_B and R stay in-control for the entire period of 10 observations, while Z and L go out of control at the eighth sample and at the tenth for P_{DB} (Figure 5.2).

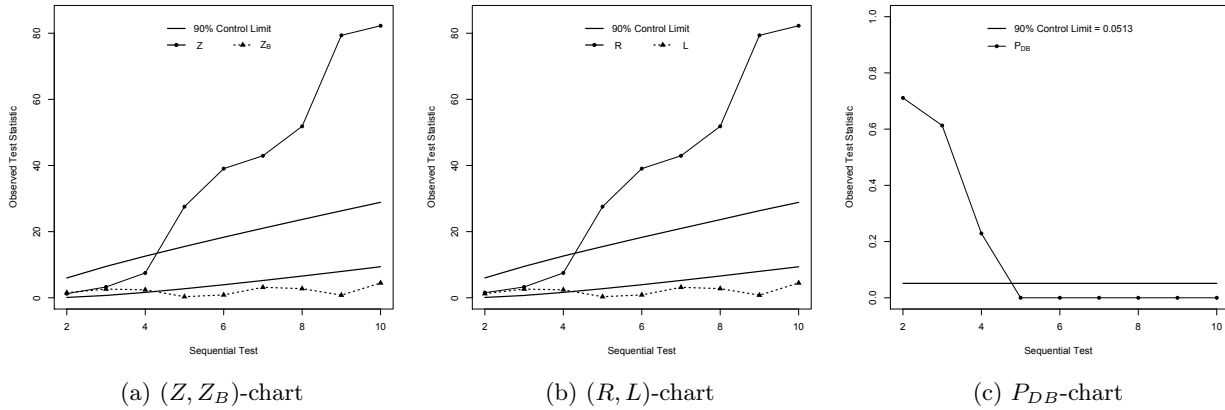


Figure 5.1. Panel of (a) (Z, Z_B) -chart, (b) (R, L) -chart, and (c) P_{DB} -chart for a PLP with $\beta = 0.2$ for $n = 10$.

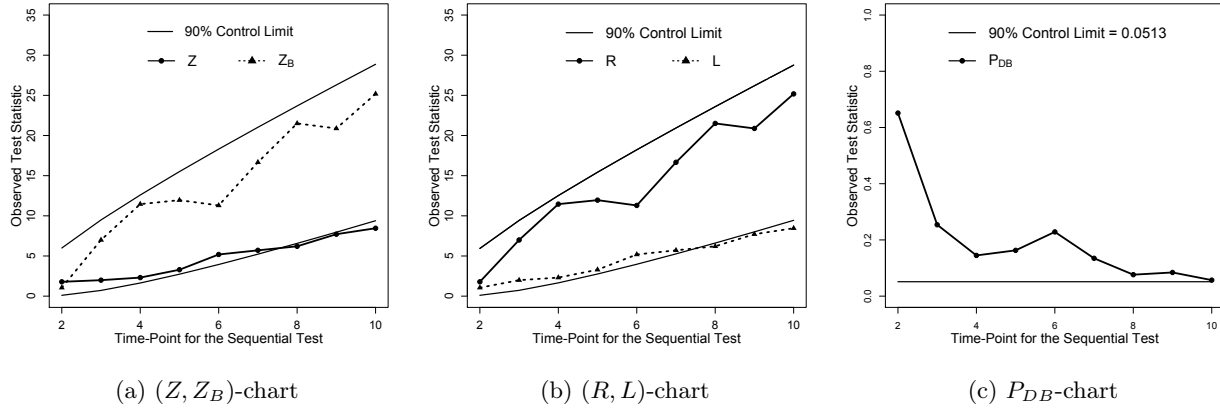


Figure 5.2. Panel of (a) (Z, Z_B) -chart, (b) (R, L) -chart, and (c) P_{DB} -chart for a PLP with $\beta = 2.0$ for $n = 10$.

We claim that the transformation from a cocktail of tests to a panel of coherent control charts is rooted in simplicity and practicality, yet is profound enough to modernize the landscape and infrastructure of the control charts. So, the charting continues for an alternative with pieces of HPPs, divided by one or two change points – a step-function intensity.

Again, a panel of (Z, Z_B) -chart, (R, L) -chart, and P_{DB} -chart for a simulated NHPP with $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 3)$ and the combination of [sample size, location of change point] is $[20, 15]$. Figure 5.3 appears interesting: (a) the base pair (Z, Z_B) and its empowered pair $[R, L]$ display almost the same pattern, and (b) Z and L stay in-control while all the others are able to detect an out-of-control before the process ends.

Consequently, a conventional Z -chart alone would fail to do the job, which justifies the utility and merits for a panel of control charts safeguarding power symmetry as per the work of Chapter 3.

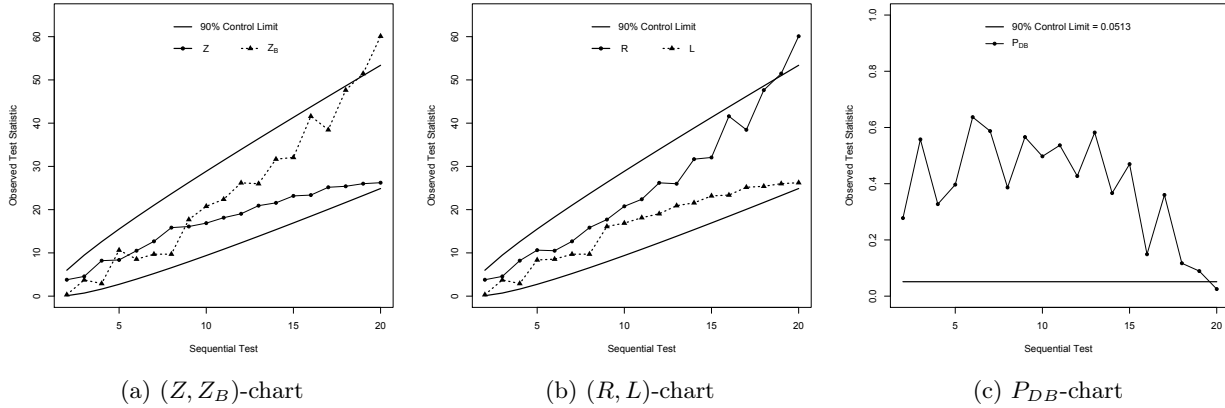


Figure 5.3. Panel of (a) (Z, Z_B) -chart, (b) (R, L) -chart, and (c) P_{DB} -chart for an NHPP with $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 3)$ and the combination of [sample size, location of change point] is [20, 15], where $t_{15} = 0.8971$ and $t_{20} = 0.9836$.

Lastly, a panel of (Z, Z_B) -chart, (R, L) -chart, and P_{DB} -chart for a simulated NHPP with $\lambda(t)$, where $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ and the combination of [sample size, locations of change point] is [144, 49, 96], where $t_{49} = 15.1941$, $t_{96} = 63.3500$, and $t_{144} = 79.1885$. In this case, all three HPP pieces have approximately the same number of occurrences of about 48, as per the locations of the change points. The unit rate of the middle segment ($\lambda_2 = 1$) is a third of that of its adjacent segments ($\lambda_1 = \lambda_3 = 3$). The total number of the occurrences for its neighboring segments is set exactly the same, which would take the middle segment approximately three times as long to run its course. Therefore, testing a hypothesis such as this may be perceived as entering a relatively lower and longer road ready for a new asphalt pavement installation.

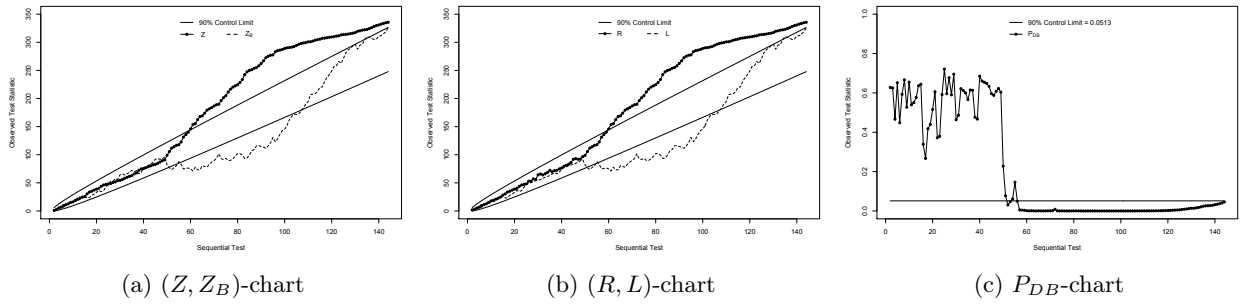


Figure 5.4. Panel of (a) (Z, Z_B) -chart, (b) (R, L) -chart, and (c) P_{DB} -chart for an NHPP with $\lambda(t)$, where $(\lambda_1, \lambda_2, \lambda_3) = (3, 1, 3)$ and the combination of [sample size, locations of change point] is [144, 49, 96], where $t_{49} = 15.1941$, $t_{96} = 63.3500$, and $t_{144} = 79.1885$.

Figure 5.4 summarizes the results of a fully sequential testing for all five tests with $n = 144$: (a) all charts consistently send out-of-control signals around the events at the mid-fifties without false alarms during the first piece of HPP for a fixed 90% control limits; (b) again, the base pair (Z, Z_B) and its empowered pair $[R, L]$ share the same pattern; (c) Z_B and L become in-control again after about 110 events; (d) in contrast, proxy of the control limit of the P_{DB} -chart is a horizontal line, reflecting patterns of the observed test statistics by bouncing up and down above the limit throughout the course of the first HPP, and leveling off shortly after absorbing the first process change. This added visual effect applies to the patterns of the p -values, which is just a slight upward translation of Figure 5.4(c), as per the proxy function.

5.3.2 Multisystem

Recall that the PLP is a widely used point process for repairable systems reliability, and the Z is proven as a UMP test. Therefore, a solo Z -chart can be used for multiple systems following a PLP with $\beta = 2.0$ and 0.2 with $n = 10$ (Figure 5.5a), so is the P_{DB} -chart (Figure 5.5b). Charts for multisystem in the panel (Figure 5.5) are labeled accordingly as Multi- Z -chart, and so forth. For Multi- Z -chart, the improving system at $\beta = 0.2$ inflates Z , and vice versa for at $\beta = 2.0$. At $\beta = 0.2$, Z goes out of control at the fifth time-point, and at the eighth at $\beta = 2.0$. For Multi- P_{DB} -chart,

the first out of control happens at the fifth time-point at $\beta = 0.2$ and the tenth at $\beta = 2.0$.

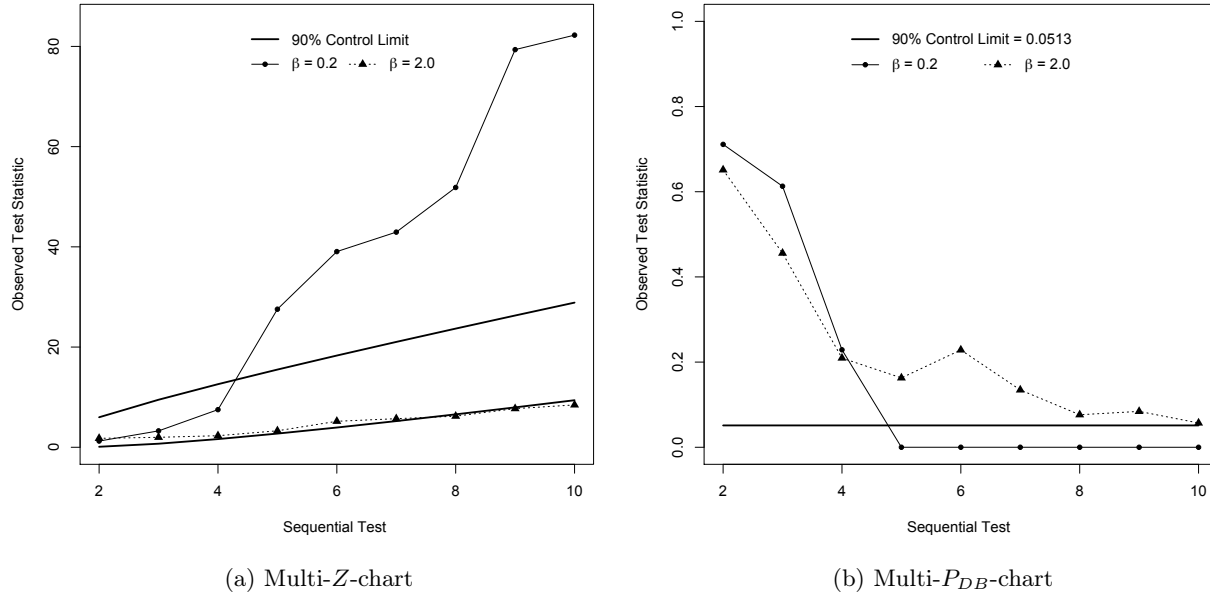


Figure 5.5. Panel of (a) Multi- Z -chart and (b) Multi- P_{DB} -chart for PLPs with $\beta = 0.2$ and $\beta = 2.0$.

In general, Figures from 5.1 to 5.5 conclude the following: (a) for a PLP, all tests provide consistent results; (b) for the two-step intensity, Z and L are at odds with the majority failing to reject an HPP at $\alpha = 0.1$, should only a fixed sample test be performed; (c) for the three-step intensity, Z_B and L stay out-of-control during only a mid-portion of the process; (d) R and P_{DB} present the findings consistently for these four data sets; (e) only the basic (Z, Z_B) -chart is able to show the directions of the trend.

In short, results of a panel of control charts presented so far are comprehensive and agreeable to the known underlying distributions, and by extension, real applications. Recall that, fully sequential control charts are about stability: complete visualization, continuous attention, quick action for a significant change, about charting an essential path forward, and a very large category of control charts. However, one may face the dilemma of performing a fixed-sample-size hypothesis testing or

a fully sequential approach (Figure 5.4). Therefore, the urge for additional options continues.

5.4 Panel of Group Sequential Control Charts

It is often more efficient and reliable to make a special effort to analyze interim results at periodic intervals, say every few time points (or months). If each analysis coincides with a meeting of program organizers then prompt action can follow. This well-known group sequential approach has been the state-of-the-art in designing clinical trials because of ethical as well as practical considerations.

The general method of performing repeated significance tests at different stages of time during the accumulation of a body of data can be implemented in a straightforward fashion, because P_{DB} is a p -value induced dual bidirectional test statistic, possessing the unique feature of having proxies that are free of either the sample size or the sampling schemes and is able to restore power symmetries that are thought to be automatic and required for quality assurance. In other words, the proxy of any level of control limit is a simple straight line, universally applicable under all kinds of mode of sampling scheme, sequential testing, number of systems/components, and alternative/underlying distribution. Group sequential control charts that follow are about multiple comparisons: relationship, multisystem, multi-components of a single system, about interim analysis of a single system with big data.

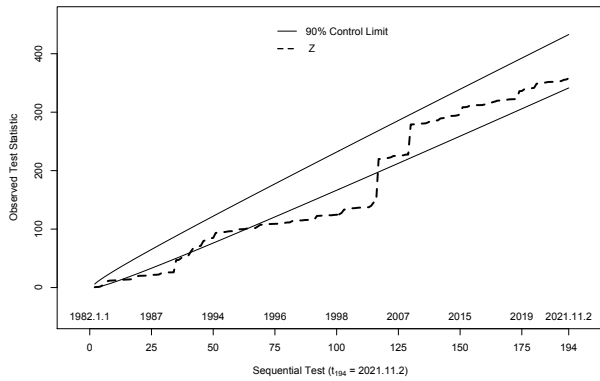
5.4.1 GS- P_{DB} -Chart and Multi-GS- P_{DB} -Chart

Panels for a series of novel group sequential control charts, centering on the hybrid, $\{P_{DB}\}$, will be first presented for the following data sets, closing milestones of DJIA and NASDAQ, 1/1/1982 - 12/31/2021:

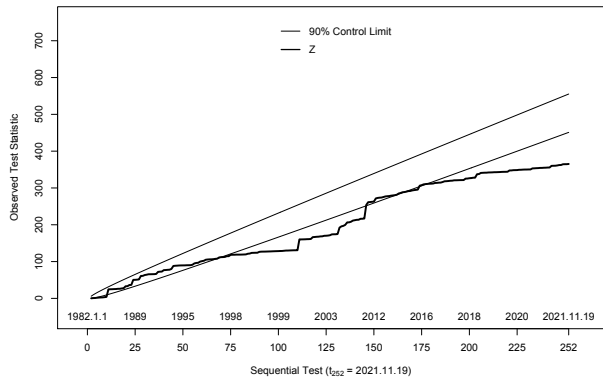
(1)https://en.wikipedia.org/wiki/Closing_milestones_of_the_Dow_Jones_Industrial_Average

(2)https://en.wikipedia.org/wiki/Closing_milestones_of_the_Nasdaq_Composite

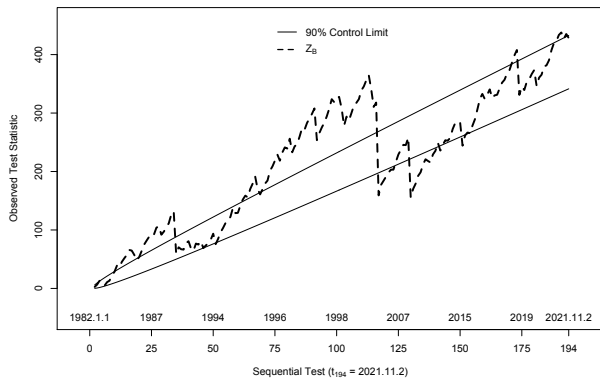
First, a panel of control charts is presented with a pair of Z -charts in Figure 5.6a: (1) DJIA and (2) NASDAQ, respectively. Here, during a span of 40 years, DJIA accomplishes 194 milestones while NASDAQ has 252, equally spaced by the number of sequential tests marked in the x-axis of the Z -chart. Likewise, the Z_B -charts in Figure 5.6b: (1) DJIA and (2) NASDAQ, respectively, point to a reversed direction of exits for the same trend of an out-of-control signal. The corresponding years are added to show that at any fixed period of observation time, the number of events of interest varies among the systems considered. Consequently, the only control chart, capable of accommodating multiple systems without additional adjustments of the control limits, is the P_{DB} -chart. They are presented in Figure 5.6c: (1) 40 groups (per year) and (2) 20 groups (every 2 years) of sequential testing, respectively, for both the DJIA and NASDAQ. The new chart is to be referred to as a GS- P_{DB} -chart, where straightforward pattern recognitions and comparisons among systems become effortless.



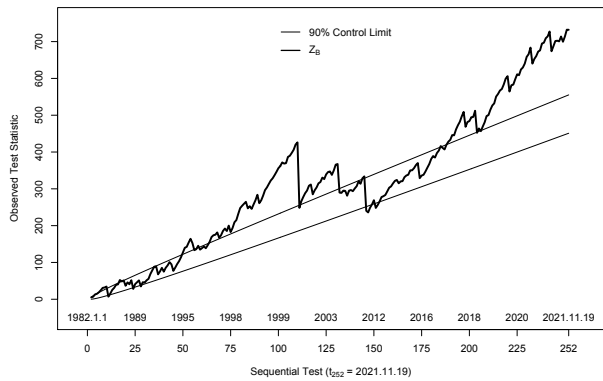
(a) (1) Z-chart for DJIA



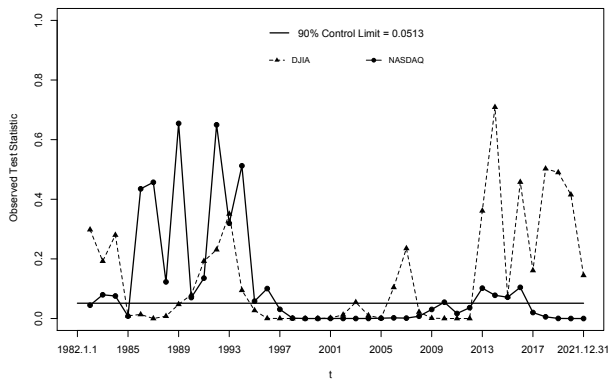
(a) (2) Z-chart for NASDAQ



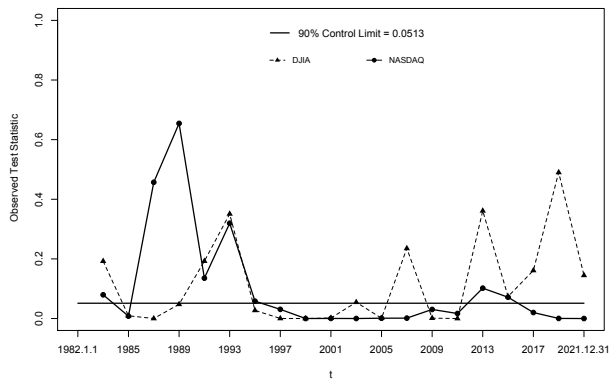
(b) (1) Z_B -chart for DJIA



(b) (2) Z_B -chart for NASDAQ



(c) (1) Multi-GS- P_{DB} -chart with 40 groups (per year)



(c) (2) Multi-GS- P_{DB} -chart with 20 groups (every 2 years)

Figure 5.6. Panel of control charts for the milestones of DJIA and NASDAQ from January 1, 1982 to December 31, 2021: (a) Z-chart for (1) DJIA and (2) NASDAQ; (b) Z_B -chart for (1) DJIA and (2) NASDAQ; (c) Multi-GS- P_{DB} -chart for both with (1) 40 groups (per year) and (2) 20 groups (every 2 years).

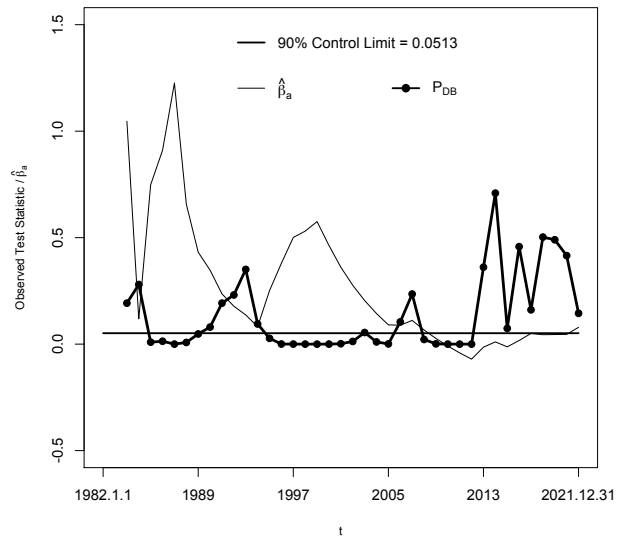
5.4.2 GS- (P_{DB}, β_a) -chart and Multi-GS- (P_{DB}, β_a) -chart

5.4.2.1 The Basics

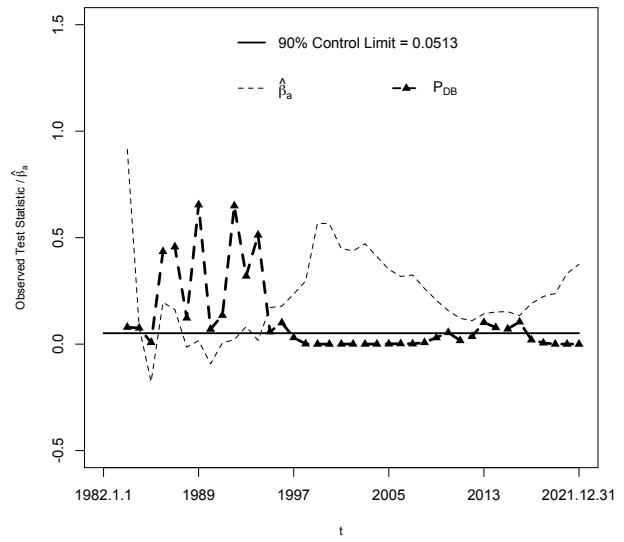
It is to be mentioned that we open the door to an option between a cocktail of tests, $(Z, Z_B) \rightarrow [R, L] \rightarrow \{P_{DB}\}$, and a solo one-size-fits-all test, which is here to facilitate the development of a panel of coherent bidirectional control charts capable of charting multiple systems. Recall that, for the Z -chart, we plot the statistic $[= 2n/\hat{\beta}_n$ or $2 \sum_{i=1}^{n-1} \ln(t_n/t_i)]$ against the time order in which the measurements were recorded for the failure-truncated sampling. Thus, the trend information presented in a Z -chart can be added to a P_{DB} -chart and a GS- P_{DB} -chart for both sampling schemes by:

- Adding the corresponding β curve, adjusted by $\beta_a = \beta - 1$ to a P_{DB} -chart, to be referred as a (P_{DB}, β_a) -chart. Thus, the reference line becomes $\beta_a = 0$ for all the (P_{DB}, β_a) -chart with the minimum value setting at -1 (or -0.5) for the y-axis. The time-trend becomes: increasing ($\beta_a > 0$), decreasing ($\beta_a < 0$), or random ($\beta_a = 0$), where the reference line is the x-axis.
- Repeating Step (a) for a GS- P_{DB} -chart, and likewise, the final chart is to be referred to as a GS- (P_{DB}, β_a) -chart.

Consequently, a panel of GS- (P_{DB}, β_a) -charts, 40 groups, for (a) DJIA and (b) NASDAQ from January 1, 1982 to December 31, 2021 is presented in Figure 5.7. In general, values of $\hat{\beta}_n$ obtained in the burn-in period may be discarded without jeopardizing the clarity of the overall trend of the system, and the pair can be pooled into an all-in-one: (a) control chart, termed as a Multi-GS- (P_{DB}, β_a) -chart (Figure 5.8); (b) panel of coherent control charts that follows.



(a) DJIA



(b) NASDAQ

Figure 5.7. Panel of GS- $(P_{DB}, \hat{\beta}_a)$ -chart, 40 groups, for (a) DJIA and (b) NASDAQ from January 1, 1982 to December 31, 2021, eliminating the first set of all the plots at $g = 1$ (burn-in period).

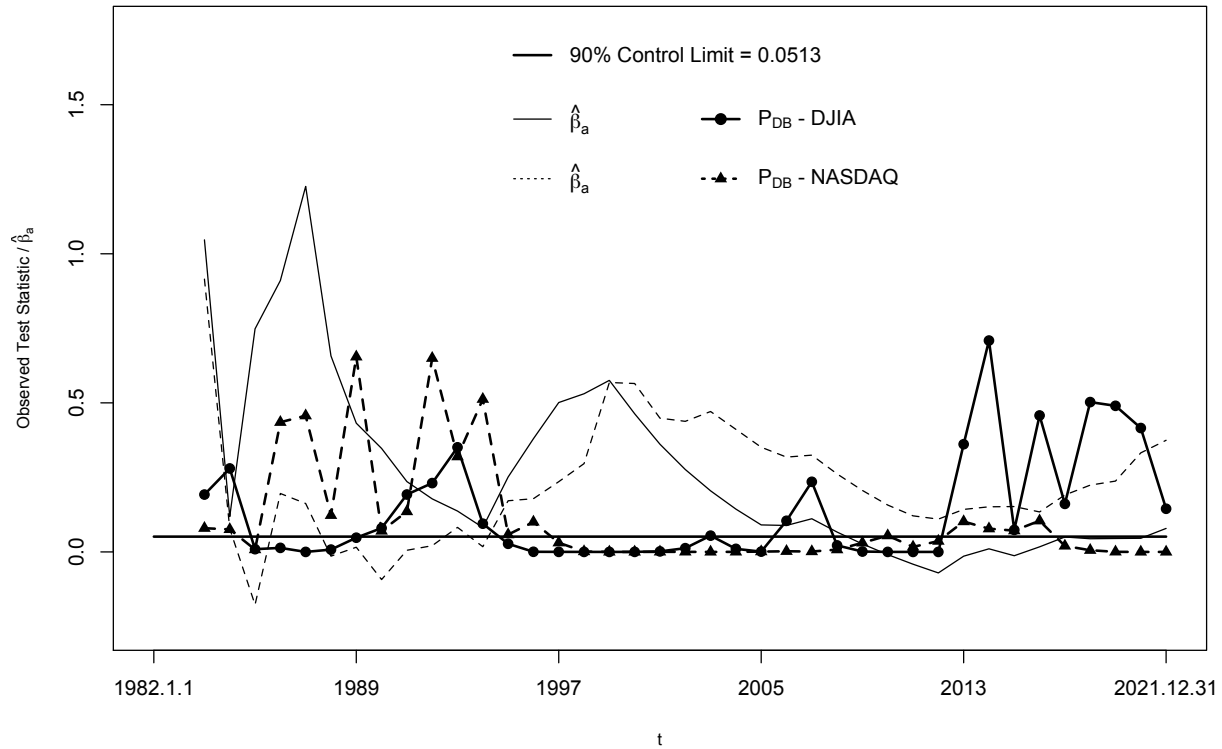


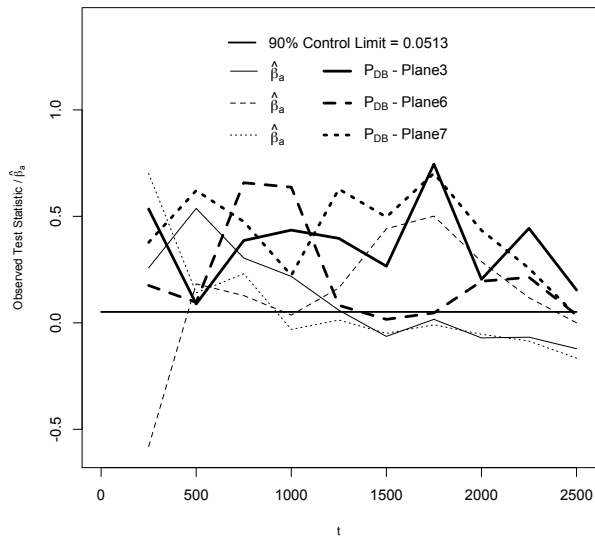
Figure 5.8. Multi-GS- (P_{DB}, β_a) -chart, 40 groups, for DJIA and NASDAQ from January 1, 1982 to December 31, 2021, eliminating the first set of the plots at $g = 1$ (burn-in period).

5.4.2.2 Application with Pairwise Comparisons

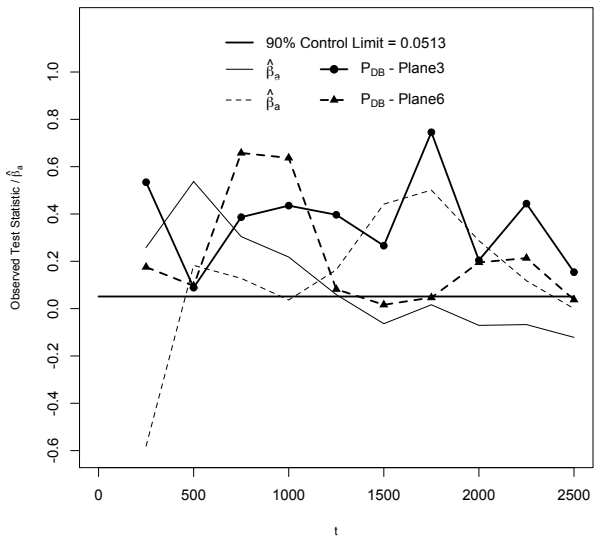
For general multisystem applications, we consider the failure times for air-conditioning equipment on 13 aircraft that were given by Proschen (1963) [76], shown in Table 5.3 of Rigdon and Basu (2000) [81]. Without loss of generality, the top three data sets in terms of the number of failure times (planes 3, 6, and 7) are listed in Table 5.2 for a panel of Multi-GS- (P_{DB}, β_a) -charts with $t = 2500$, set as the present time of observation. Figure 5.9 summarizes the results with a panel of multisystem control charts of (a) All 3 planes together, and three pairwise charting of: (b) 3 versus 6, (c) 3 versus 7, and (d) 6 versus 7. All control charts have the same group size, $g = 10$.

Table 5.2. Top three data sets in terms of the number of failure times for air-conditioning equipment, planes 3, 6, and 7.

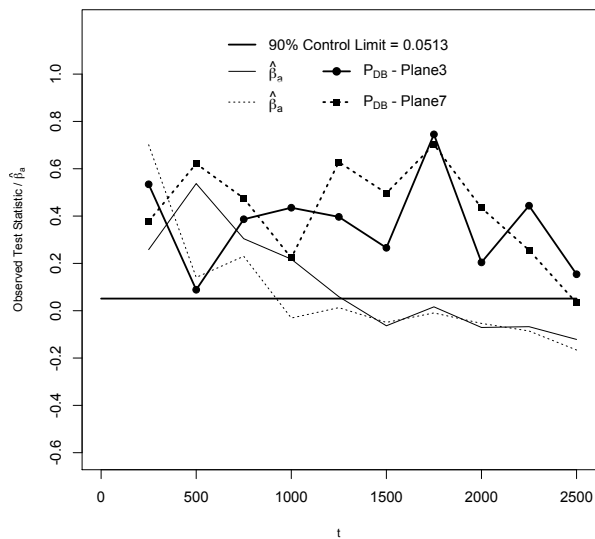
Plane	Failure Times								
3	90	100	160	346	407	456	470	494	550
	570	649	733	777	836	965	983	1008	1164
	1474	1550	1576	1620	1643	1705	1835	2043	2113
	2214	2422							
6	23	284	371	378	498	512	574	621	846
	917	1163	1184	1226	1246	1251	1263	1364	1383
	1397	1411	1482	1493	1507	1518	1534	1624	1625
	1641	1693	1788						
7	97	148	159	163	304	322	464	532	609
	689	690	706	812	1018	1100	1154	1185	1401
	1447	1558	1597	1660	1678	1869	1887	2055	2079



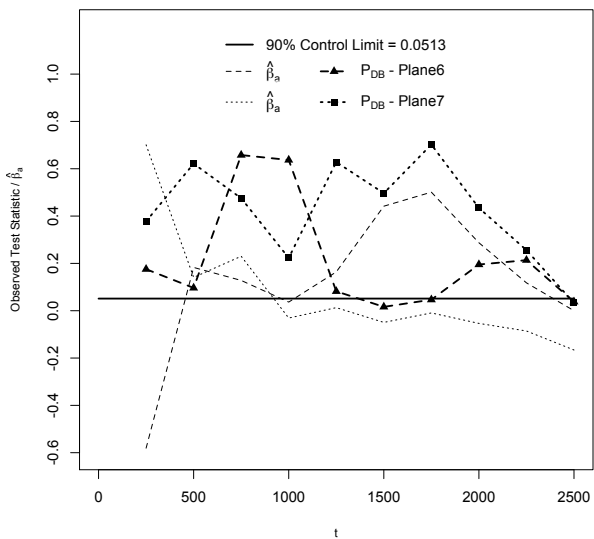
(a) All 3 planes



(b) 3 versus 6

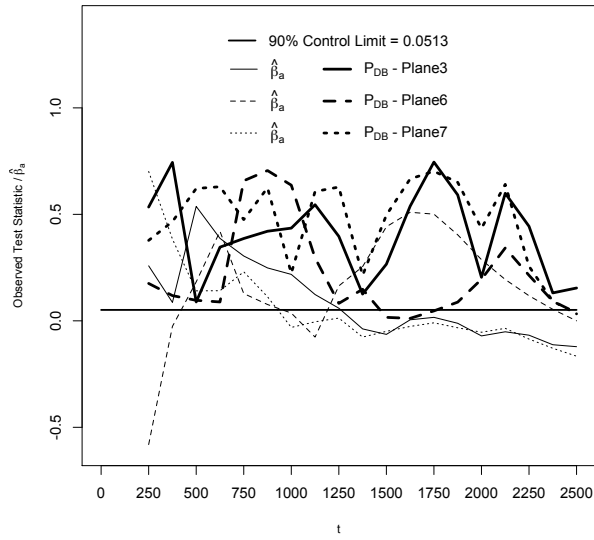


(c) 3 versus 7

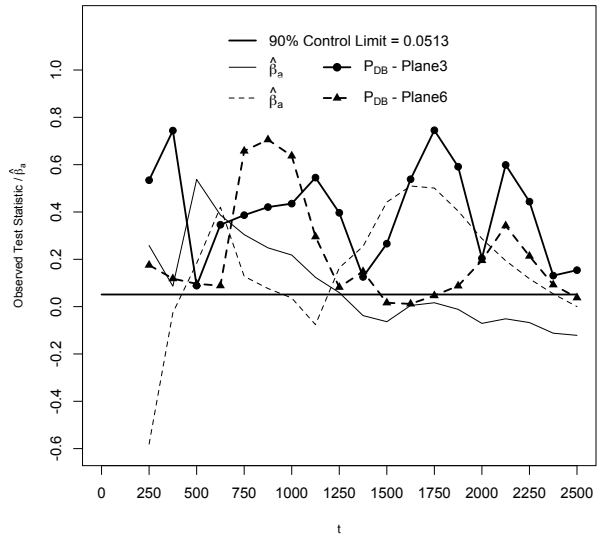


(d) 6 versus 7

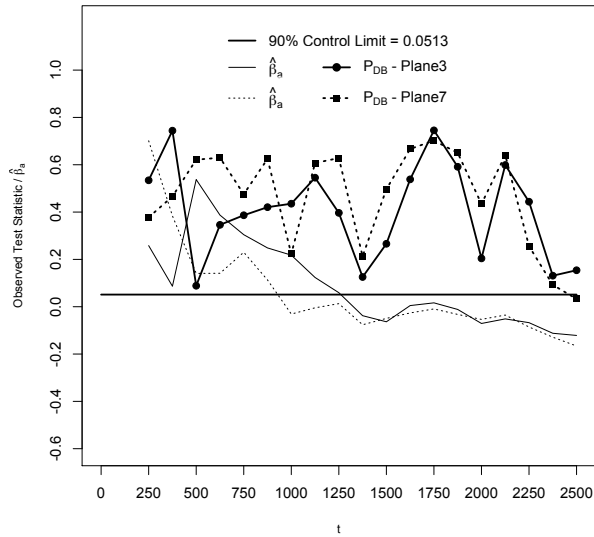
Figure 5.9. Panel of Multi-GS- (P_{DB}, β_a) -charts for (a) All 3 planes together, and three pairwise charting of: (b) 3 versus 6, (c) 3 versus 7, and (d) 6 versus 7. All control charts have the same group size, $g = 10$.



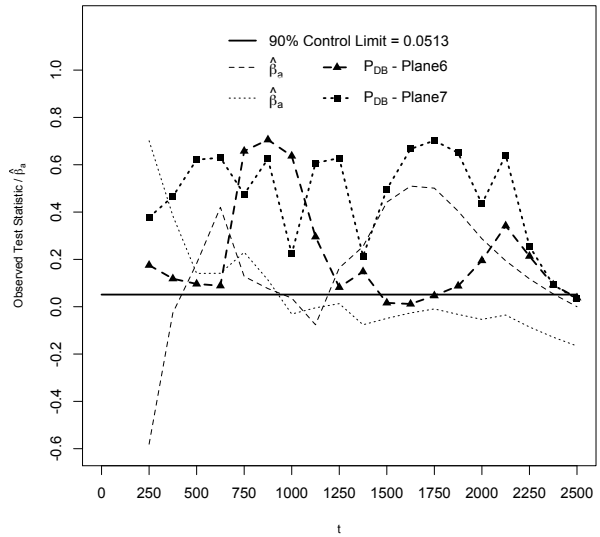
(a) All 3 planes



(b) 3 versus 6



(c) 3 versus 7



(d) 6 versus 7

Figure 5.10. Panel of Multi-GS- (P_{DB}, β_a) -charts for (a) All 3 planes together, and three pairwise charting of: (b) 3 versus 6, (c) 3 versus 7, and (d) 6 versus 7. All control charts have the same group size, $g = 20$, eliminating the first set of all the plots at $g = 1$ (burn-in period).

Again, information collected for the burn-in period may be discarded without jeopardizing the authenticity of the graphs as demonstrated in Figure 5.10, where the group size has been doubled

to $g = 20$, eliminating the first set of all the plots at $g = 1$. Figures 5.9 and 5.10 have a fixed group length ($= t/g$) throughout the entire observation period, which is not required in practice for the proxy of the P_{DB} – Another selling point for the proposed all-in-one panel of bidirectional control charts!

5.5 An Overview

In Figure 5.11, we write about two core categories of control charts: fully sequential charts and group sequential charts. Group sequential charts are about multiple comparisons: relationship, multisystem, multi-components of a single system, about interim analysis of a single system with big data. Fully sequential charts are about stability: complete visualization, continuous attention, quick action for a significant change, about charting an essential path forward, a very large category control charts. Control charts based on the P_{DB} -test are expansive; they are practical, simple, flexible, and run across both categories: P_{DB} -chart, Multi- P_{DB} -chart, and Multi- (P_{DB}, β_a) -chart in the fully sequential category; GS- P_{DB} -chart, GS- (P_{DB}, β_a) -chart, Multi-GS- P_{DB} -chart, and Multi-GS- (P_{DB}, β_a) -chart in the other category. Thus, control charts of the following fashion: multi- (Z, Z_B) , or $-(R, L)$ -chart; GS- Z, Z_B, R , or L -chart are possible but may become counterproductive, and are recommended for references only unless the associated tests are known to be the most powerful for the underlying distribution. In general, control charts having a natural affinity for each other, mapped in Figure 5.11, form a panel, facilitating timely and handy diagnostic features. The structure overview also simplifies the sharing of new resources – a kind of one-stop shop for bidirectional control charts.

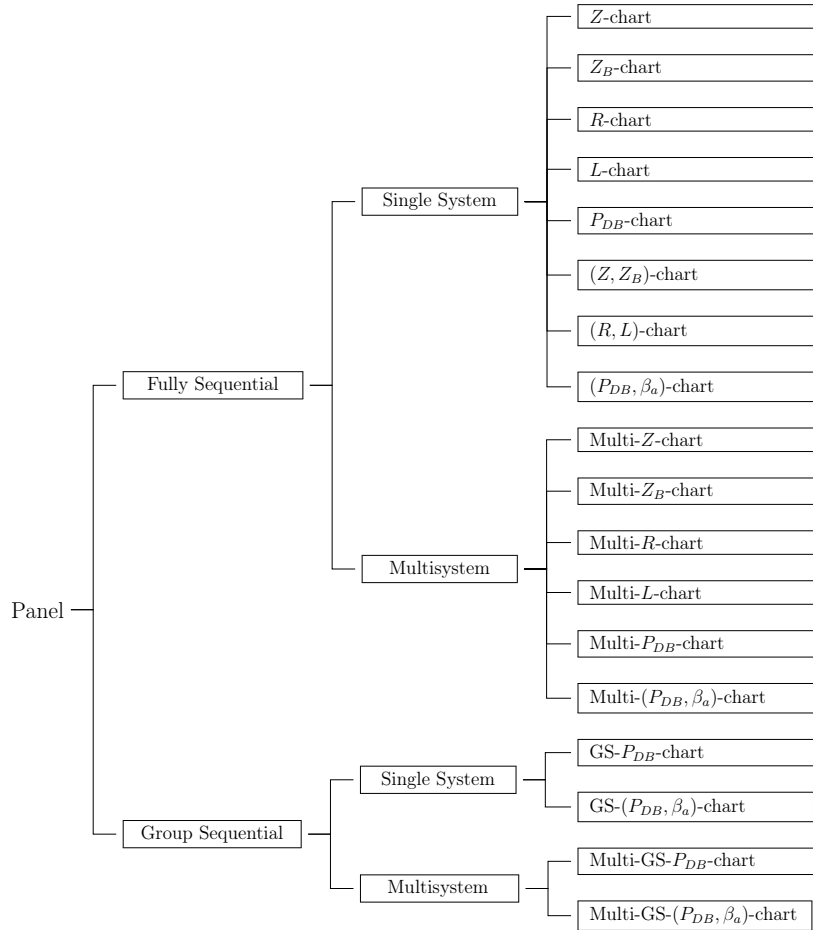


Figure 5.11. Structure overview for panels of control charts.

5.6 Conclusions – An All-in-One Panel of Bidirectional Charts

This work, in essence, represents a transformation and modernization of the fundamentals and classics in statistical process control. The coherence, utility, and merit of the entire structure development are elaborated with panels of relevant control charts based on a variety of data sets possessing a combination of: simulated or real, small or large, with known or unknown underlying distribution. Also, a supporting tool box, adapted to the simplified version, includes a structure overview and all the functions and R code for the test statistics (Z and Z_B), the β_a curve, the p -value proxies (R , L , and P_{DB}), and the $(1 - \alpha)100\%$ control limits (Z , Z_B , R , L , and P_{DB}) for

both sampling schemes with size n .

The backbone of the transformation is a cocktail of coherent bidirectional tests, $(Z, Z_B) \rightarrow [R, L] \rightarrow \{P_{DB}\}$. Each all-in-one panel consists of a group of affiliated control charts, fit for group sequential testing and multisystem. The testing tools start with a basic pair of tests, (Z, Z_B) , that are more complementary than competitive, and are bound together with a common set of control limits. Both tests are exact test. Another unified and coherent control limit proxies, partners of an empowered pair of tests, $[R, L]$, form a second bidirectional control chart. And, a series of novel group sequential control charts, centering on a hybrid, $\{P_{DB}\}$, of the empowered pair, characterize the test as the only tool possessing the unique feature of having an intelligible version that is free of the system sample size and the sampling schemes, and is able to restore power symmetries that are thought to be automatic and required for quality assurance.

These control charts are rooted in the p -value proxies, which are grounded in practicality and simplicity, and are inverted for the corresponding $(1 - \alpha)100\%$ control limit – a simple straight line, universally applicable under all kinds of mode of sampling scheme, sequential testing with flexible group length, number of systems/components, and alternative/underlying distribution. In particular, Multi-GS- (P_{DB}, β_a) -chart, the flagship, is a dual bidirectional and a one-size-fits-all control chart, to be considered in equipment performance assessments of repairable systems, and by extension, other applications – A big scope for our future work.

CHAPTER 6

RELIABILITY GRAPHICS FOR QUALITY ASSURANCE

6.1 Introduction

Suppose we assume that the successive occurrence times of a specific process follow an NHPP, and let $T_1 < T_2 < \dots < T_n$ be the first n successive times of events. These times are measured from the beginning of the observation period (cumulative length of time over which the events occur), so $T_1 < T_2 < \dots < T_n$, denoted by $\{T_i\}$. The cumulative times are recognized as the main series of events. The corresponding series of inter-event times is referred to as $\{t_i\}$, where $t_i = T_i - T_{i-1}, i = 1, 2, \dots, n$, and $T_0 = 0$. The reliability data associated with any repairable systems and all the NHPP data with similar nature can be abbreviated as $\{T_i\}$ or $\{t_i\}$, referred to as the reliability graphics in general, provide firsthand information on the reliability data concerning the functions of intensity, mean, and more.

This work is about creating self-evident and perceivable graphs, or what we call graphics, for the repairable systems and, by extension, other applications. In the words of Faraway (2021) [36], good graphics are vital in data analysis. They help us avoid mistakes and suggest the form of the modeling to come. They are also crucial in communicating our analysis to others. Many in our audience or readership will focus on the graphs. This is our best opportunity to convey our message without misunderstanding. In some cases, the graphics can be so convincing that the formal analysis becomes just a confirmation of what has already been seen. In this chapter, we start the development by showing a simple dot-plot. A complementary graph that follows cumulates all the dot-jumps up to each time-point of occurrences and leaves us an often mysterious but meaningful footprint to shed any light on. Its mystery lies in the sequence of the slopes and/or curvature associated with each corresponding hill edging up. Both plots, full of diagnostic elements, are equivalent to the

so-called “prior” in the Bayesian data analysis.

6.2 Fingerprinting Graphics

Reliability is a crucial concern for all products. Rapid technological advances, the development of highly sophisticated products, intense global competition, and greater customer expectations have continually placed increasing pressure on manufacturers to design and build higher reliability products. Informally, reliability may be defined as (Doganaksoy et al. 2021 [29]) “quality over time.” Most systems and some parts are repaired when a failure occurs. For repairable systems, this may involve replacing a failed component or subsystem, which typically generates a sequence of failure times, $\{T_i\}$ or $\{t_i\}$, on the same unit. Most importantly, a careful study of field reliability issues, such as those described in Dalal et al. (1989) [26], usually suggests the correct data, graphics, and analyses at the right time, coupled with prompt action, could have avoided the severity of the catastrophe.

For pattern recognition, data size for the intended graph matters. The NHPP is a classical point process supporting a growing body of theories, methods, and applications for the occurrences of natural phenomena. Restricted by the systems’ lives, most reliability data sizes are inevitably and relatively small in contrast to the occurrences of the natural phenomena. To better address the utility and power of the graphing techniques presented hereafter, we first adopt the seismological data of Japan from https://en.wikipedia.org/wiki/List_of_earthquakes_in_Japan. We then use multiple sets of real reliability data of various sizes, small to moderate, to further elaborate on the merits of all the developed reliability graphics.

6.2.1 The IET-plot

To produce an inter-event time plot, termed as the IET-plot, for $\{t_i\}$ of size n , label the x-axis as Time-Point, equally spaced from 0 to n , and y-axis as Inter-Event Time with scaling values starting

close to the minimum of t_i 's for a better eye-ball identification of the pattern and potential change-point(s), reflected by the heights of the inter-event times. Figure 6.1 is the IET-plot for a total of 59 main shocks that occurred in Japan between January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale. The graph maps the sequence of the inter-event times (in years), which may be interpreted as the so-called genetic code of seismicity in Japan. The plot sends a clear concern to the seismologists with an increasing intensity trend – half of the significant earthquakes occurred during the last 20 years!

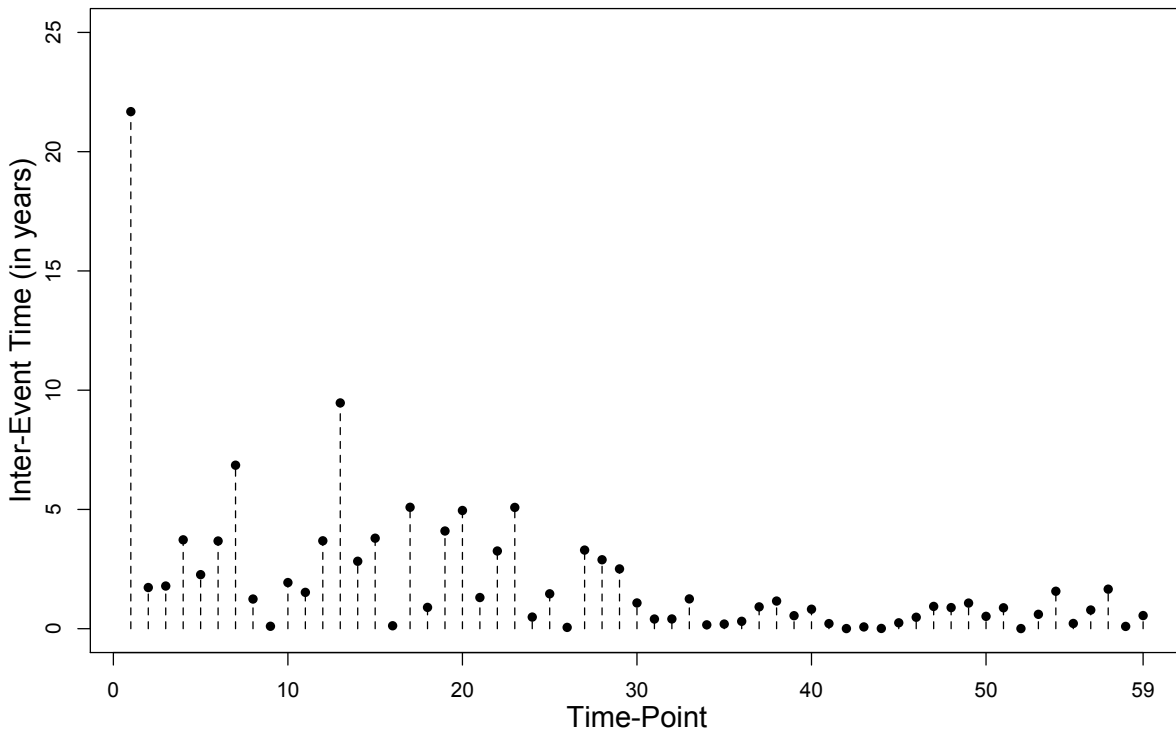


Figure 6.1. IET-plot for a total of 59 major earthquakes, occurred in Japan during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale.

A back-to-back comparison of the perceived alarming trend is again presented in an IET-plot of 53 main shocks that occurred in China during the same observation period and range of magnitudes

(Figure 6.2). These graphics are vital in data analysis. They would help the experts with the form of the modeling to come.

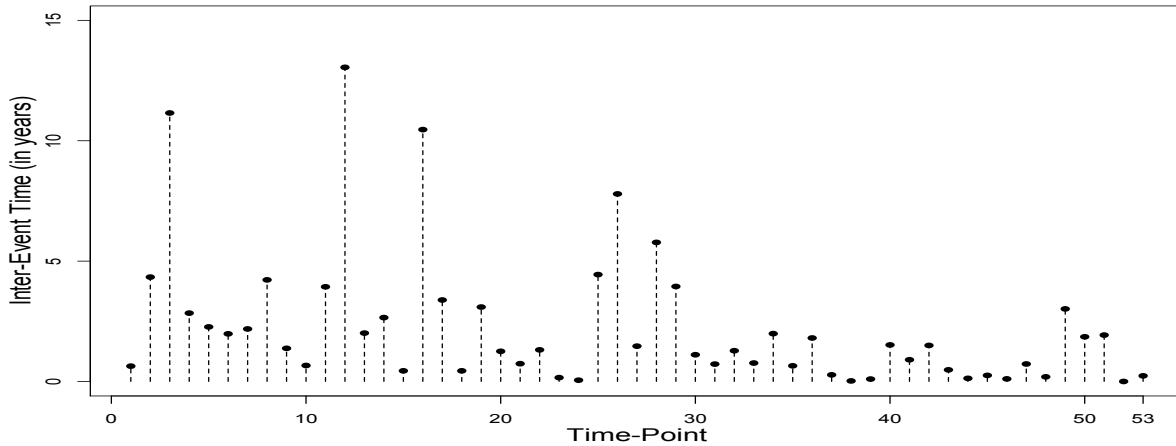


Figure 6.2. IET-plot for a total of 53 major earthquakes, occurred in China during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale.

For general multisystem reliability graphics, we consider the failure times for the air-conditioning equipment on 12 of the 13 aircraft given by Proschen (1963) [76], shown in Table 5.3 of Rigdon and Basu (2000) [81]. We exclude Aircraft 11, where only one data point was recorded. In Figure 6.3, a single panel summarizes the corresponding 12 IET-plots. The panel facilitates an opportunity to get initial individual and/or aggregated messages without misunderstanding. Of course, some reliability data require additional diagnostic graphics to confirm the findings.

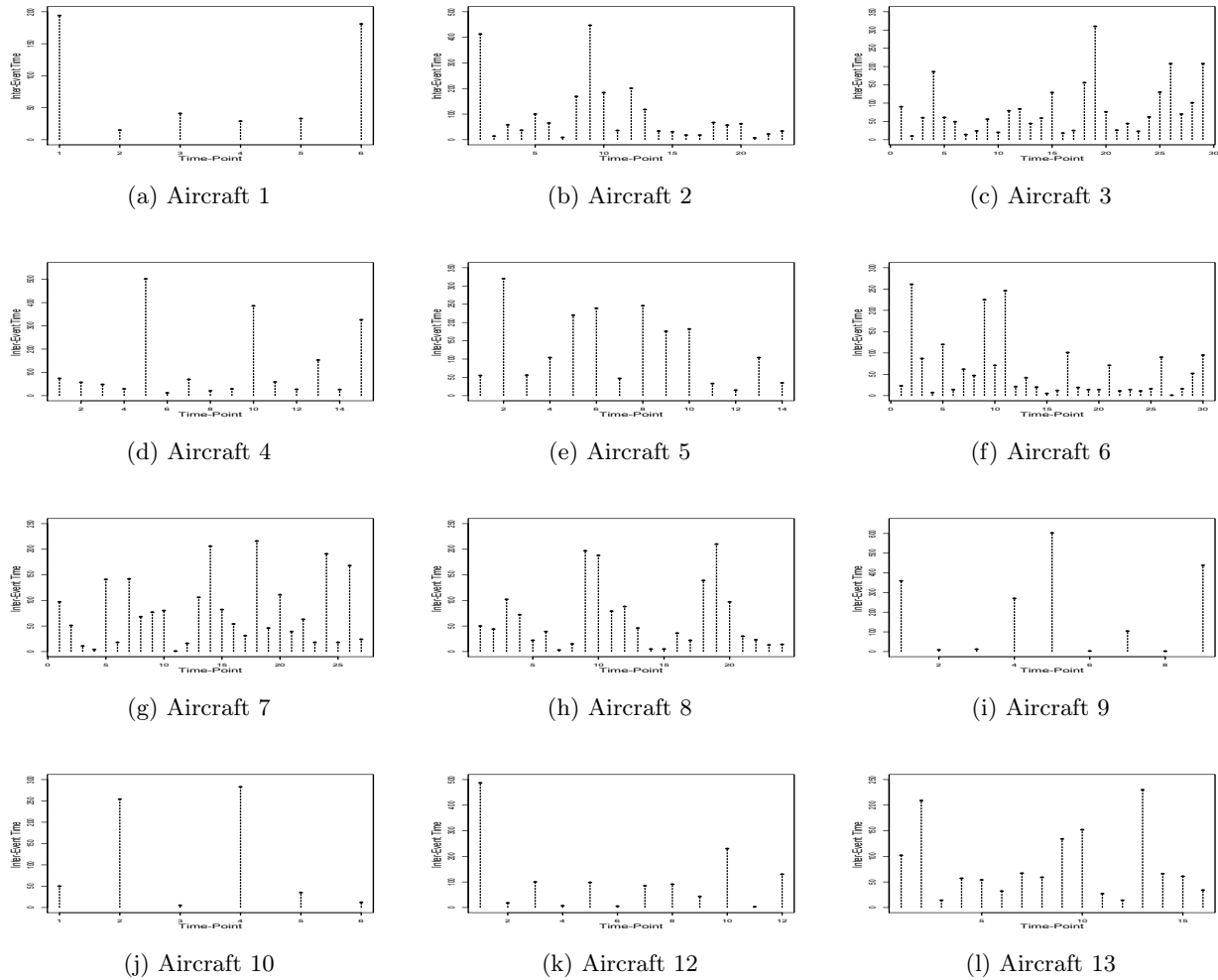


Figure 6.3. Panel of 12 IET-plots for the failure times of air-conditioning equipment on 12 aircraft.

6.2.2 The CT-plot

An alternative is to produce a cumulative time plot, termed as the CT-plot, for $\{T_i\}$ of size n , by labeling the x-axis as Time-Point, equally spaced from 0 to n , and the y-axis as Cumulative Time. In an IET-plot, we see a sequence of bouncing musical notes, while the counterpart, a CT-plot, transforms all of them into uphill stairs needed to climb to reach the top of a mountain. The CT-plot cumulates all the jumps up to each time-point and leaves us with a mysterious but meaningful footprint to shed light on. The CT-plots corresponding to the IET-plots above are shown in Figures from 6.4 to 6.6. These two fingerprinting graphics are views or summaries of the raw data rather

than the data themselves. Data used to track the reliability growth trend of repairable systems occur naturally in a sequential fashion. It should be evident that the chronological order in which the events occur is a vital aspect of a stochastic process, if a significant time trend is to be objectively identified. We are, therefore, motivated to drill down, brush, and investigate values with other dynamic graphics.

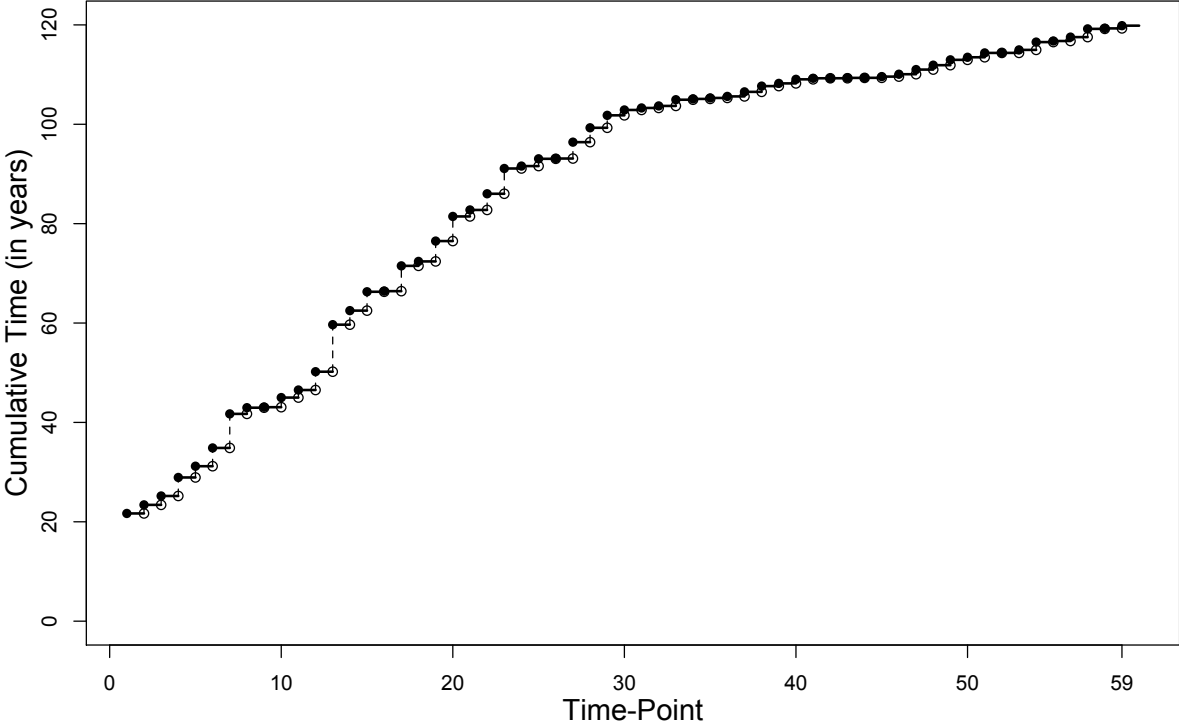


Figure 6.4. CT-plot for a total of 59 major earthquakes, occurred in Japan during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale.

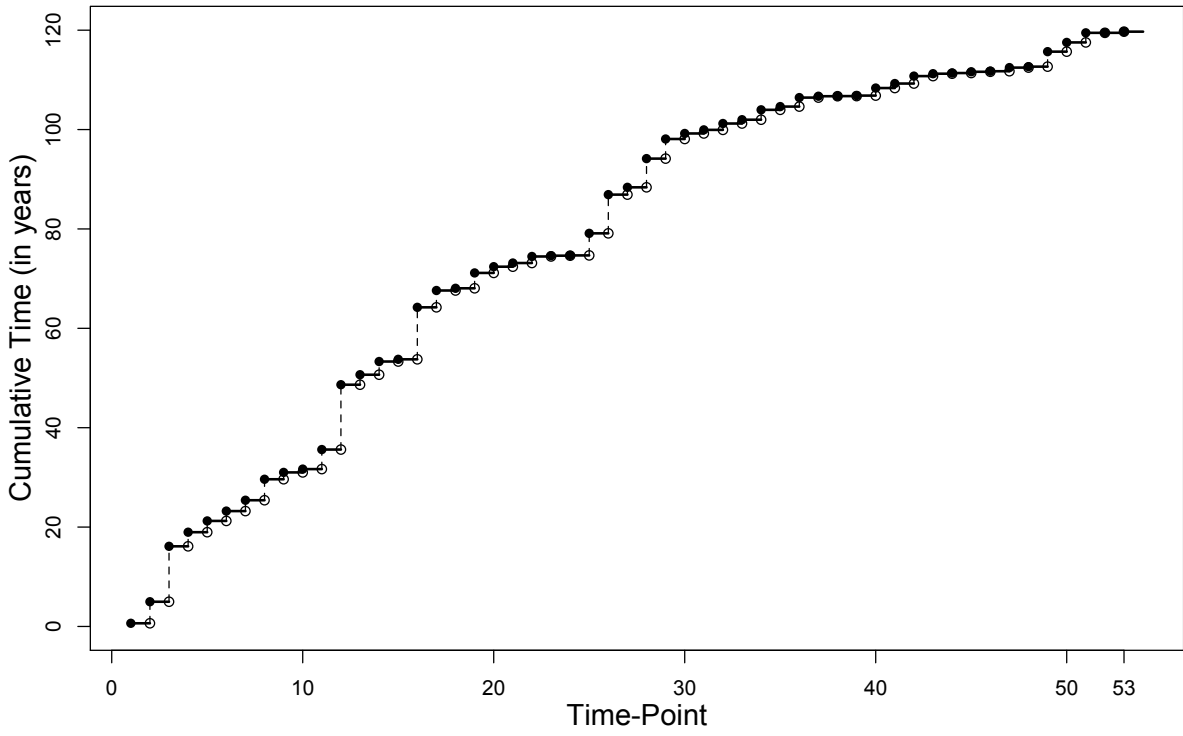


Figure 6.5. CT-plot for a total of 53 major earthquakes, occurred in China during January 1, 1902 and December 31, 2021 with magnitudes of 6 or higher on the Richter scale.

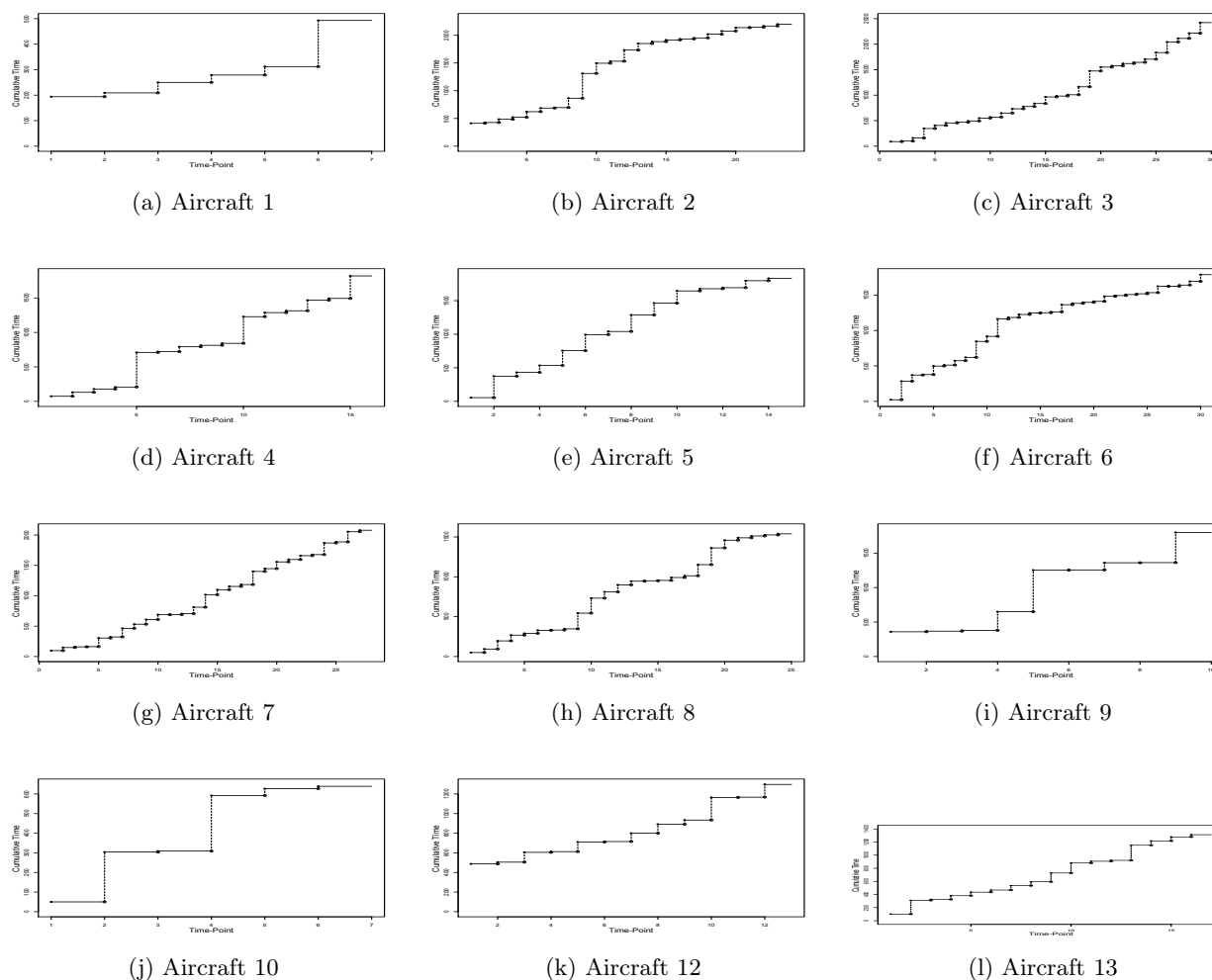
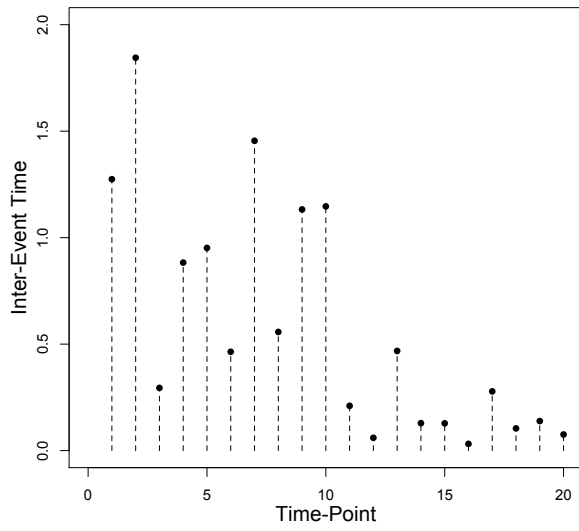


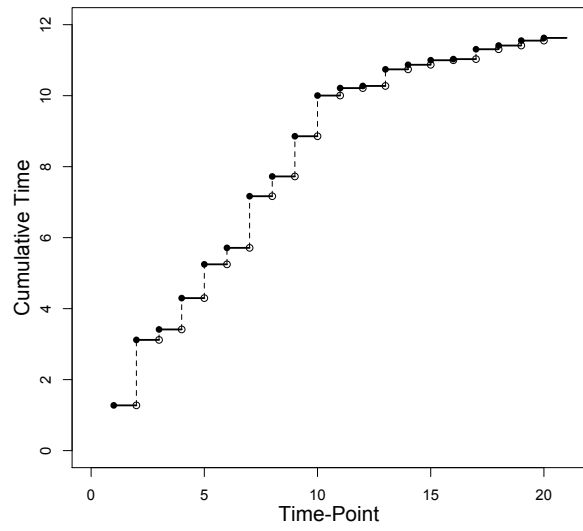
Figure 6.6. Panel of 12 CT-plots for the failure times of air-conditioning equipment on 12 aircraft.

6.3 Motivating Examples – Power Asymmetries

The asymmetrical characteristics of the basic pair (Z, Z_B) carry over into the empowered bidirectional version (R, L) . We have seen either test of the empowered pair, $[R, L]$, championing in one alternative setting, but also being turned against the very setting altered with just one of the composed elements in Chapter 3. They help us avoid mistakes and suggest the form of the graphics to come.

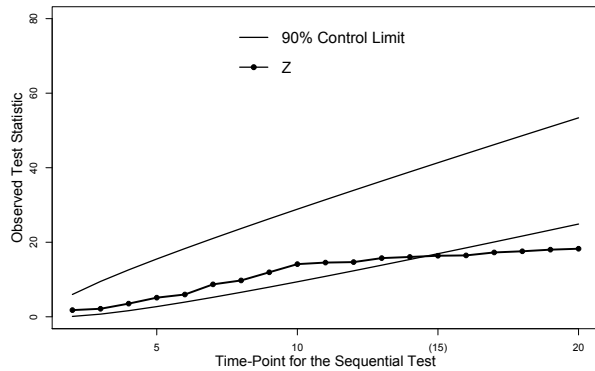


(a)

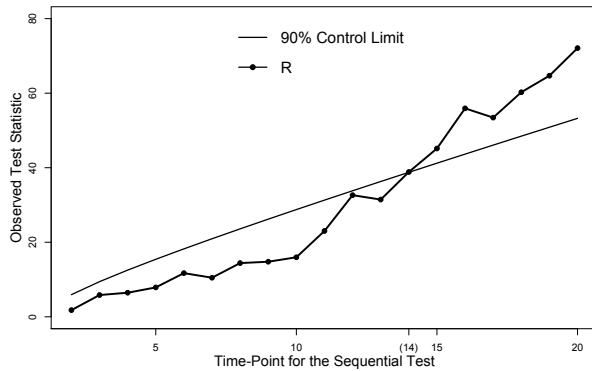


(b)

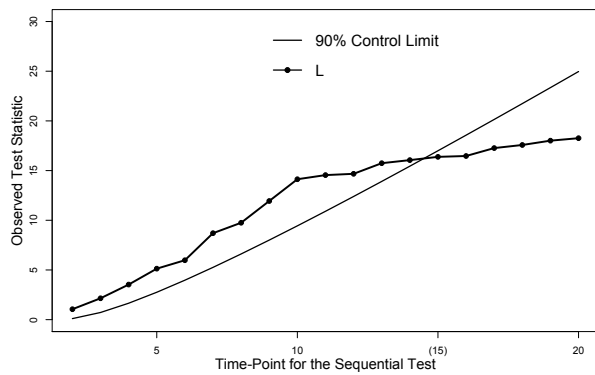
Figure 6.7. Panel of (a) IET-plot and (b) CT-plot for a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$.



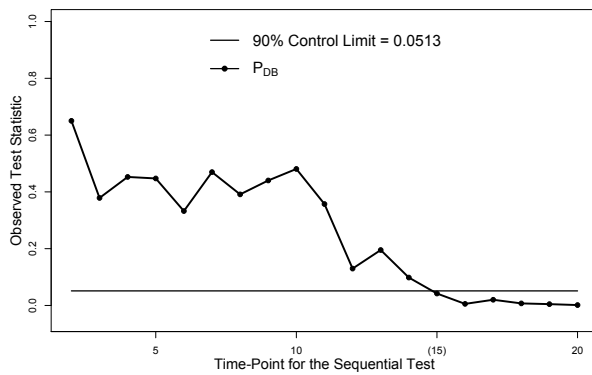
(a) Z -chart



(b) R -chart



(c) L -chart



(d) P_{DB} -chart

Figure 6.8. Panel of (a) Z -chart, (b) R -chart, (c) L -chart, and (d) P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$.

First, a panel of IET-plot and CT-plot for a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$, is presented in Figure 6.7. As expected, two slopes shown in Figure 6.7(b) follow a piecewise linear function of two, where the first appears to be steeper than the second for an increasing two-step-intensity function. Also, the location of the change-point is where the piecewise line bends.

Implementations of the cocktail of tests and the proxies outlined in Table 5.1 for the simulated data yield a panel of coherent Z -chart, R -chart, L -chart, and P_{DB} -chart (Figure 6.8). The control

charts are all for testing $H_0: \beta = 1$ (the system’s reliability remains stable) versus H_A , with an applicable 90% control limit proxies. In general, the distribution under H_A would be an unknown NHPP indicating that the system’s reliability has grown or decreased during the observed period. The R -chart stays in-control until the 14th time-point, while all the others are consistently at the 15th, marked with (14) and (15), respectively, in the graphics. In this case, the IET- and CT-plots in Figure 6.7, are so convincing that the analysis becomes just a confirmation of what has already been seen. Therefore, these fingerprinting graphics may be perceived as the so-called “prior” in the Bayesian data analysis and as a crucial diagnostic tool for the complete data analysis.

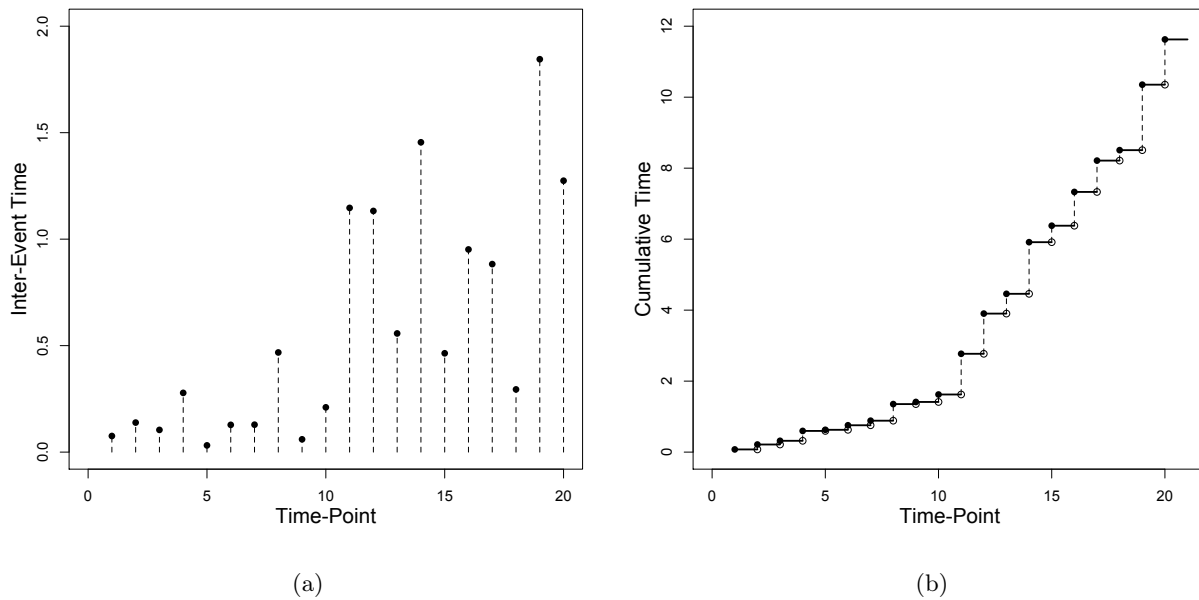
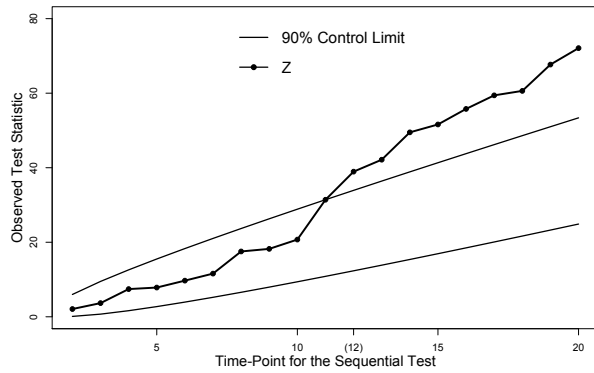
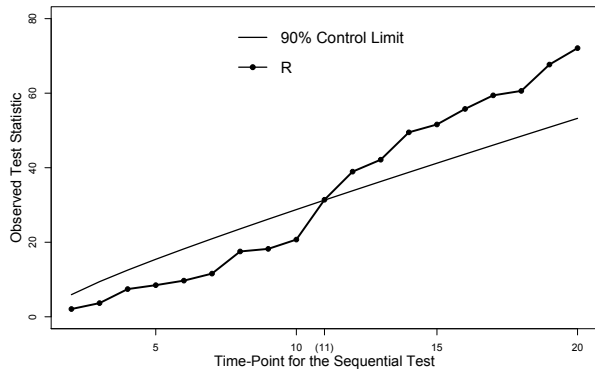


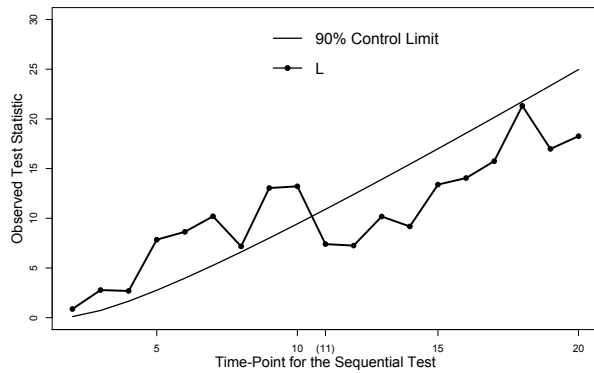
Figure 6.9. Panel of (a) IET-plot and (b) CT-plot for the reversal of a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$.



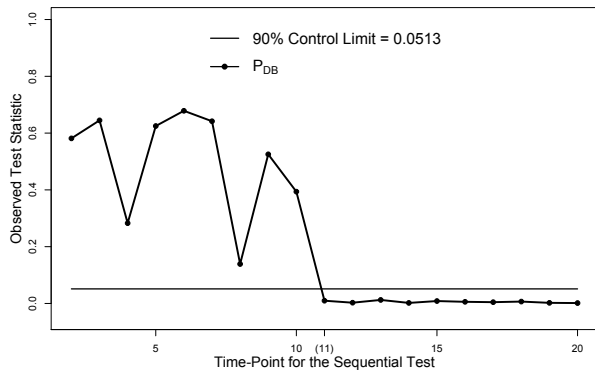
(a) Z -chart



(b) R -chart



(c) L -chart



(d) P_{DB} -chart

Figure 6.10. Panel of (a) Z -chart, (b) R -chart, (c) L -chart, and (d) P_{DB} -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for the reversal of a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$.

To substantiate the existence of power asymmetries, we reproduce Figure 6.7 based on the reversed data, $\{T_n - T_{n-i}\}$, for $i = 1, \dots, n$, and $T_0 = 0$ (Figure 6.9). Again, the two slopes shown in Figure 6.9(b) follow a piecewise linear function of two, where the second is steeper than the first for a decreasing two-step-intensity function. For the reversed process, the Z -chart stays in-control until the 12th time-point, while all the others are consistently at the 11th, per the ordering of the reversed sequence presented in Figures 6.9 and 6.10. The change-point is set at the middle point for both data sets with reversed trends. Clearly, it justifies that we need a single frame to

consolidate all the information accrued, including asymmetrical performances caught by the cocktail of bidirectional tests.

6.4 Panel of *OR*-Charts

6.4.1 Simulated Data

We are ready to propose a control chart, including both the original and its reversed version in a single frame, termed as an *OR-Z*-chart, and so forth by replacing *Z* with the test adopted. We recommend that a panel of the *OR-Z*-, *R*-, *L*-, and *P_{DB}*-charts be presented for unified and comprehensive analysis. The panel in Figure 6.11 summarizes the results of the simulated data based on the proposed graphing technique. Note that, for the reversal, the first out-of-control signals for (*Z*, *R*, *L*, *P_{DB}*) are (12, 11, 11, 11). They are converted back to (8, 9, 9, 9), time-points in line with the original sequence. In contrast, we recall that the original sequence's first out-of-control signals for (*Z*, *R*, *L*, *P_{DB}*) are (15, 14, 15, 15).

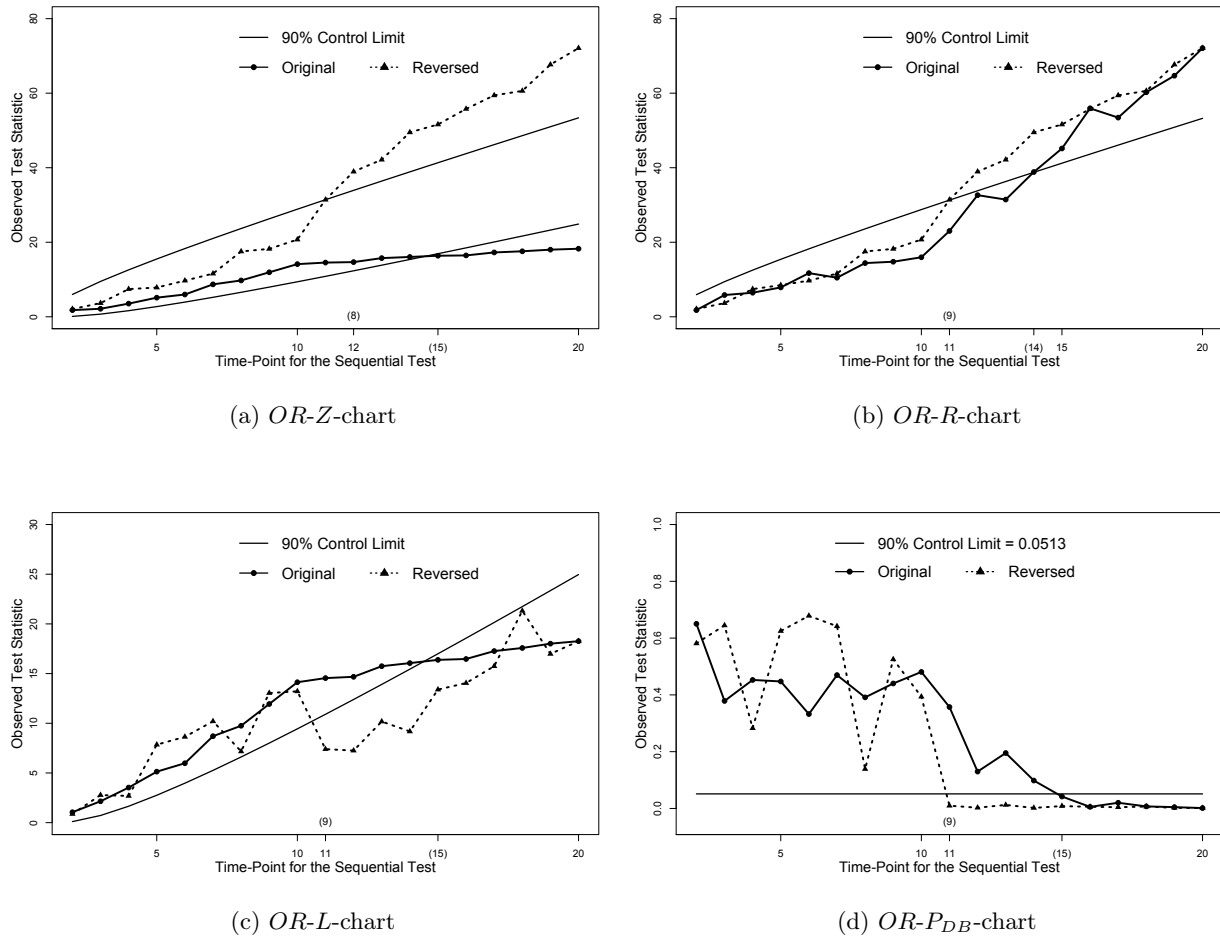


Figure 6.11. Panel of (a) $OR-Z$ -chart, (b) $OR-R$ -chart, (c) $OR-L$ -chart, and (d) $OR-P_{DB}$ -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for a simulated data, $\{T_i\}$, of an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$, and a two-step-intensity $\lambda(t)$, where $(\lambda_1, \lambda_2) = (1, 5)$.

To best elaborate the outcomes, we recall the raw data, $\{t_i\} = \{1.2745, 1.8448, 0.2944, 0.8828, 0.9517, 0.4641, 1.4548, 0.5572, 1.1324, 1.1467, \mathbf{0.2103}, \mathbf{0.0600}, \mathbf{0.4679}, \mathbf{0.1288}, \mathbf{0.1280}, \mathbf{0.0316}, \mathbf{0.2782}, \mathbf{0.1043}, \mathbf{0.1387}, \mathbf{0.0757}\}$: an NHPP with $n = 20$, change-point at $\tau = T_{10} = 10.00$ on a continuous time scale, and a two-step-intensity $\lambda(t)$, where $\lambda_1 = 1$ and $\lambda_2 = 5$. In other words, two pieces of HPPs form a single NHPP, divided at the change-point, $T_{10} = 10.00$. Clearly, the inter-event times in the second HPP (data highlighted in bold), or the so-called regime, are considerably smaller than those of the first one. And, the $OR-R$ -chart stays in-control until the

14th time-point for the original process, indicating that the changes are not significant until after the t_{14} is included in the test statistic. Consequently, per the R -test and a 90% control limit, the first HPP ends at T_{13} , one-step-back from the 14th time point, which is the detected change-point location, three steps away from the true $\tau = T_{10} = 10.00$. All the other tests are one step slower than R .

As for the reversed version, depicted in Figure 6.11, the first out-of-control signals for (Z, R, L, P_{DB}) are (8, 9, 9, 9), marked above the corresponding x-axis, in line with the ordering of the original sequence. Again, for instance, the OR - R -chart stays in-control until the 9th time-point for the reversed process, indicating that the changes are not significant until after the t_9 is included in the test statistic. Analogously, per the R -test and a 90% control limit, the second (or last) HPP begins at T_{10} , one-step forward from the 9th, which is the detected change-point location, exactly at the true $\tau = T_{10} = 10.00$. L and P_{DB} join R for the same outstanding performance!

To sum up:

- What is to be done? Look reality in the eye from the graphics and act upon it.
- It takes a longer time to detect a change-point for an NHPP with an increasing two-step-intensity function than a decreasing one – Beware of power asymmetries! Perceived that the test statistic tracks the balance of a checking account, which takes more number of small deposits to feel the significance of the increase as the balance is already high. On the other hand, a single deposit of \$1,000 to an account with a balance of \$99 would be a fascinating eye-opener! – Look at data from a different angle.
- A first forwardly detected out-of-control signal at the k th time-point indicates the HPP ends at T_{k-1} , the location of a change-point.
- A first backwardly detected out-of-control signal at the k th time-point indicates the HPP

begins at T_{k+1} .

- The Z_B -test, a member of the bidirectional tests, is excluded from the panel, because it is a backward version of Z (Ho 1993) at every time-point of the sequential test.

Change detection is a classic research area in statistical process control (Montgomery 2007). Numerous approaches have been developed for change detection (Hinkley 1971 [45]; Barry and Hartigan 1993 [11]; Wardell et al. 1994 [93]; Hawkins 2001 [42]; Hawkins et al. 2003 [44]; Zou et al. 2009 [100]; Hawkins and Deng 2010 [43]; Holland and Hawkins 2014 [59]; Cia et al. 2019 [19]; Dong et al. 2019 [30]; Eryilmaz and Kan 2019 [35]; Qiu 2020 [77]; Chen et al. 2021 [18]; Jin et al. 2022 [61]). The graphics discussed above apply the concepts to address the issues of power asymmetries based on simulated data with a known distribution. A glance at the complexities of real reliability data analysis, where the underlying parametric distributions are unknown, is forthcoming.

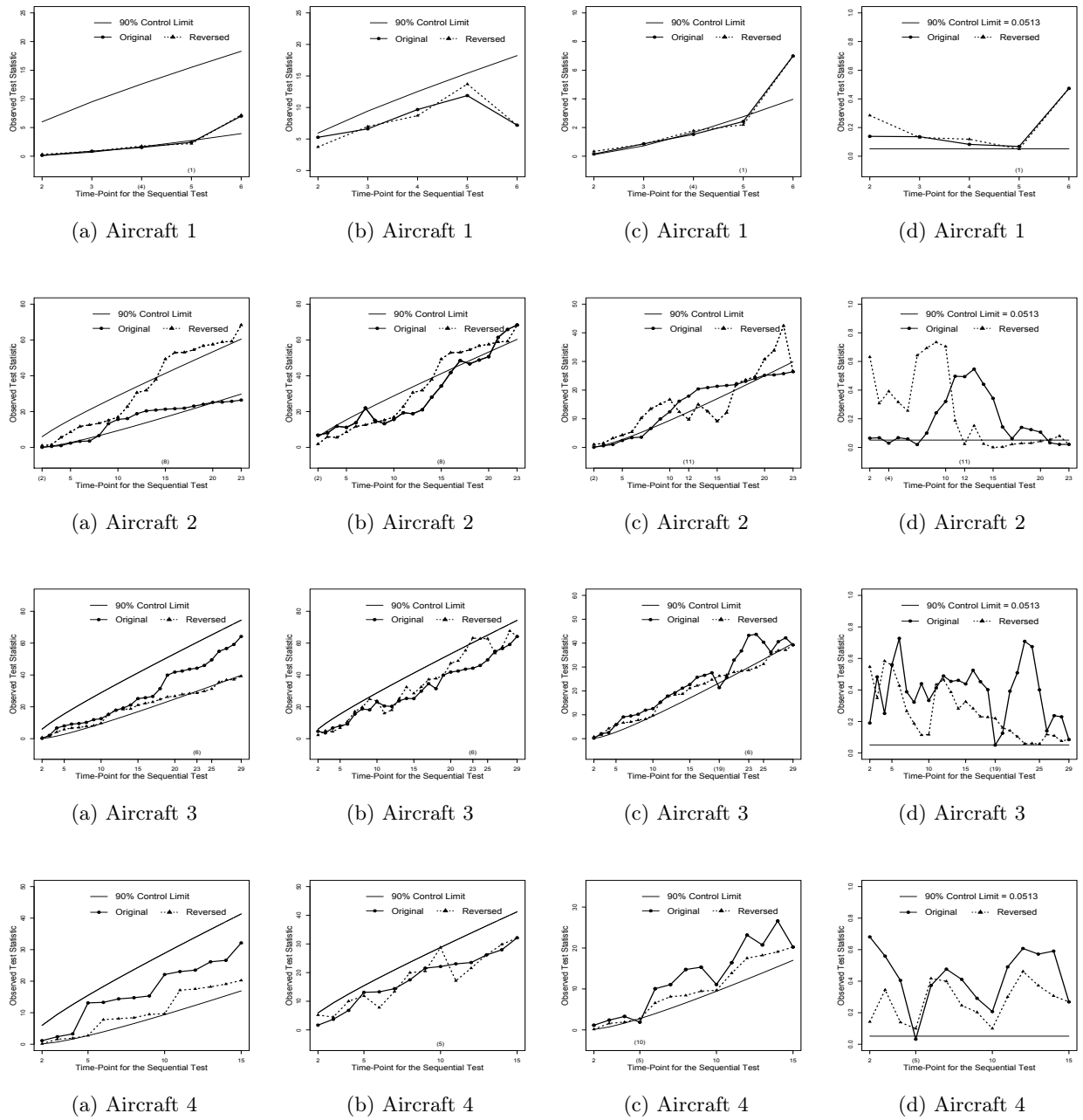


Figure 6.12. Panel of (a) $OR-Z$ -chart, (b) $OR-R$ -chart, (c) $OR-L$ -chart, and (d) $OR-P_{DB}$ -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for air-conditioning equipment on aircraft 1 to 4.

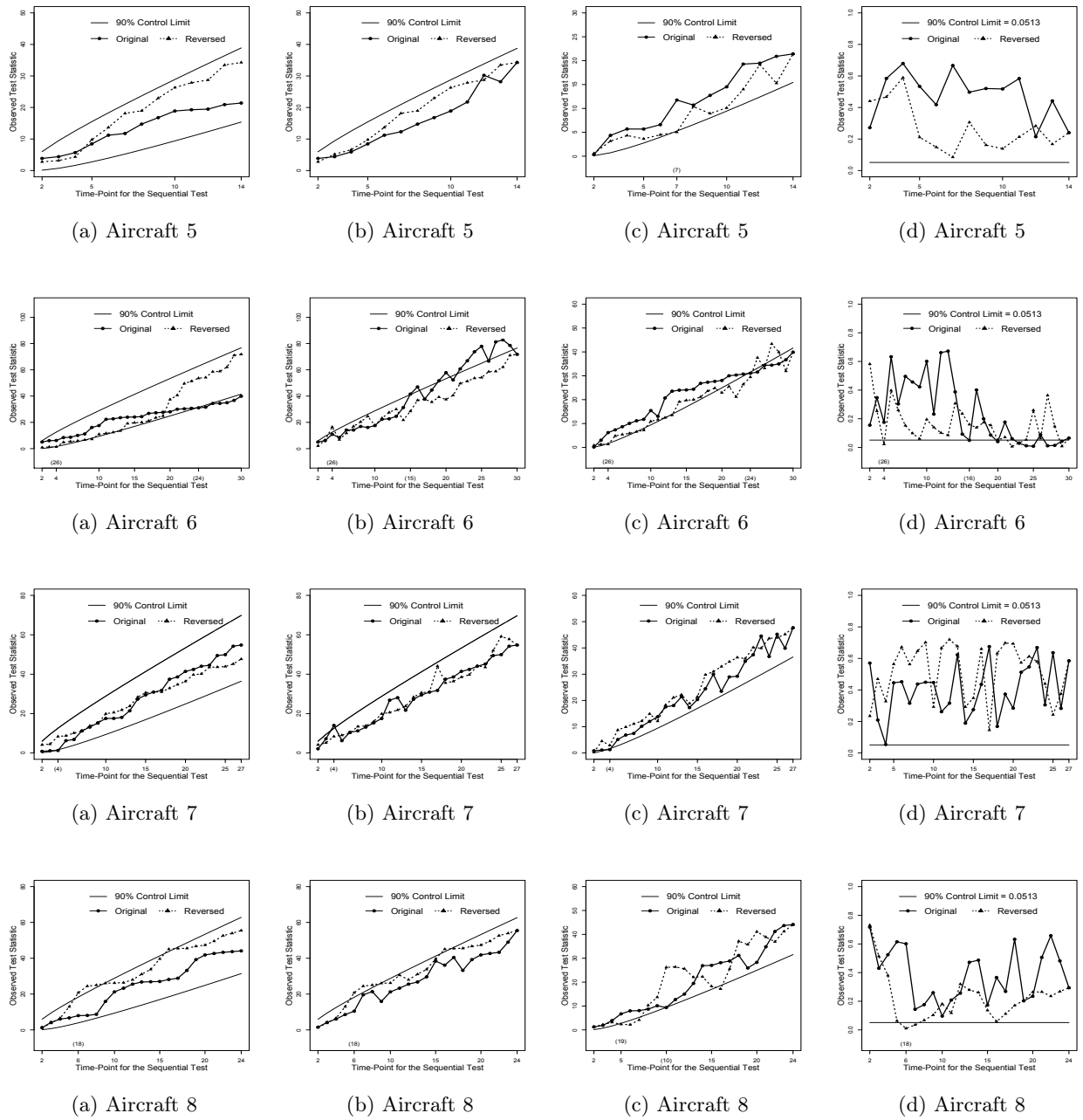


Figure 6.13. Panel of (a) $OR-Z$ -chart, (b) $OR-R$ -chart, (c) $OR-L$ -chart, and (d) $OR-P_{DB}$ -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for air-conditioning equipment on aircraft 5 to 8.

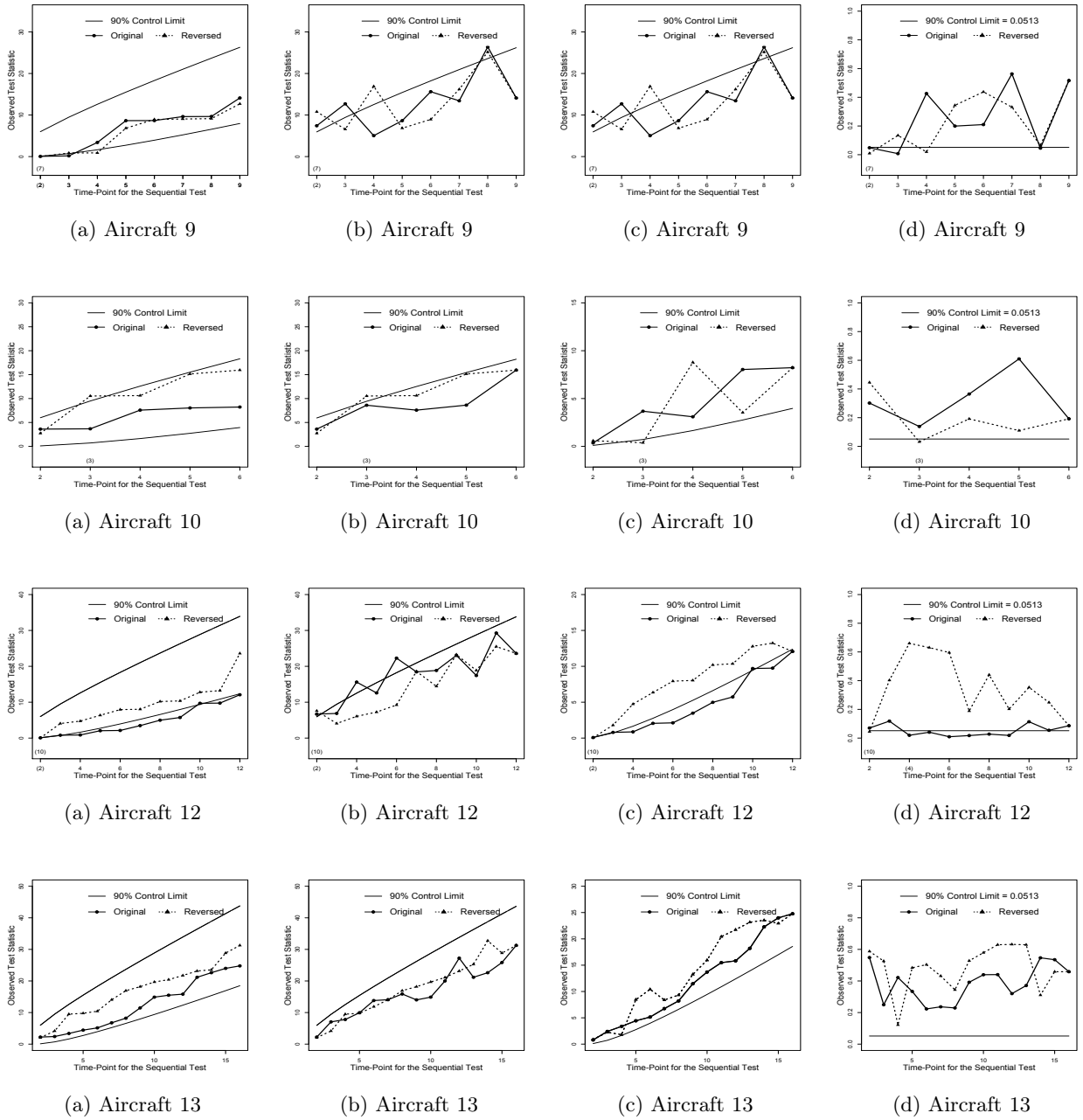


Figure 6.14. Panel of (a) $OR-Z$ -chart, (b) $OR-R$ -chart, (c) $OR-L$ -chart, and (d) $OR-P_{DB}$ -chart, all are at a 90% control limit for testing $H_0 : \beta = 1$, for air-conditioning equipment on aircraft 9, 10, 12, and 13.

6.4.2 Real Data

6.4.2.1 Aircraft 13

For general multisystem reliability graphics, we extend Figures 6.3 and 6.6 to figures from 6.12 to 6.14, which summarize the corresponding 12 $OR-Z$ -, R -, L -, and P_{DB} -charts for air-conditioning equipment on the 12 aircraft. Scattered results associated with Aircraft 13 seem straightforward and are used first as the main focus of the application, designed to optimize the utility of the proposed graphics in this study. We reproduce a panel of graphics by consolidating the following in Figure 6.15: The IET-plot, CT-plot, and all the $OR-Z$ -, R -, L -, and P_{DB} -charts for air-conditioning equipment on Aircraft 13, termed as a complete graphics panel. The graphics feature the following:

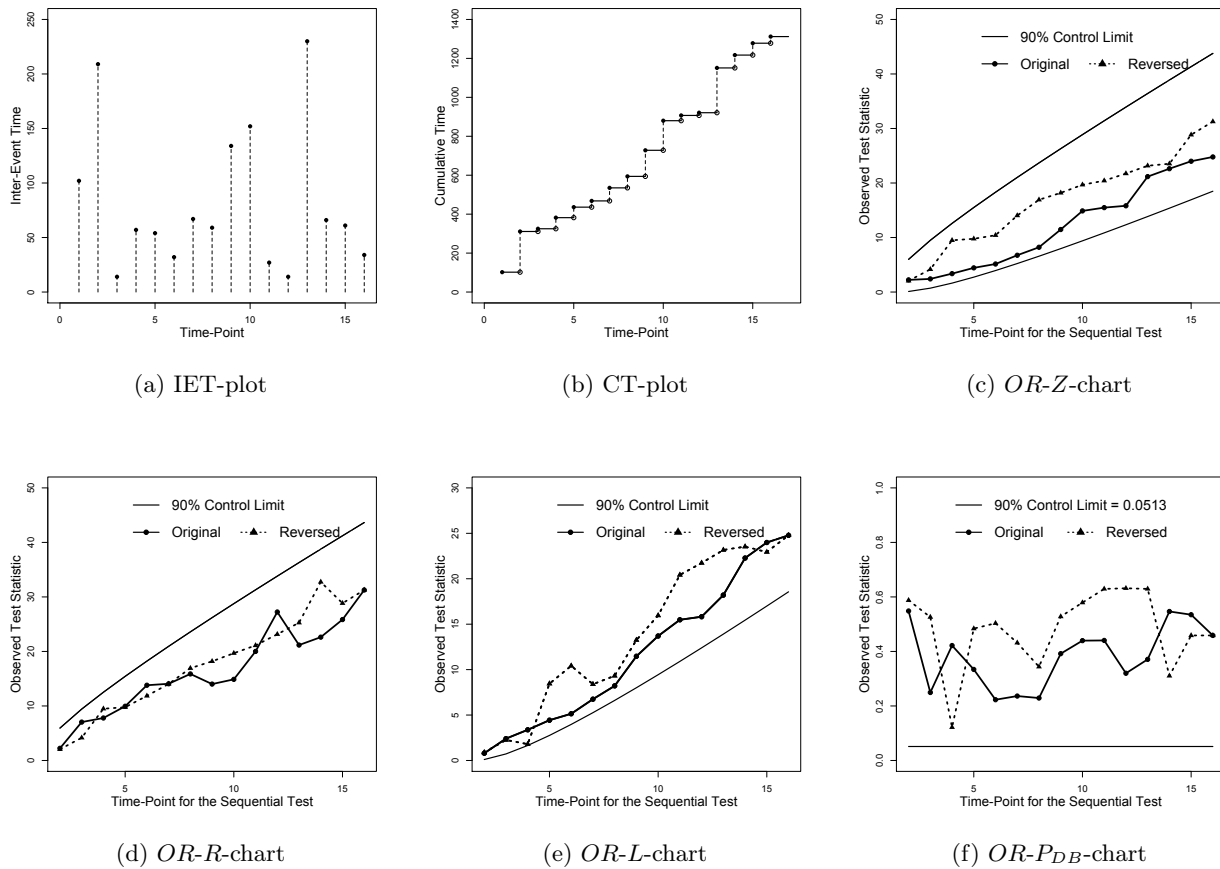


Figure 6.15. Graphics panel of (a) IET-plot, (b) CT-plot, (c) $OR-Z$ -chart, (d) $OR-R$ -chart, (e) $OR-L$ -chart, and (f) $OR-P_{DB}$ -chart for air-conditioning equipment on Aircraft 13.

- The IET-plot shows that the pattern of the inter-event times appears to be random.
- The upward slope of the aggregated stairs in the CT-plot appears to be a straight line, reinforcing the message sent by the IET-plot.
- With a 90% control limit, all the dual-bidirectional control charts show no out-of-control signals, indicating that the reliability data follow an HPP.

The above characterization provides critical baseline information for analyzing all the other aircraft.

6.4.2.2 Aircraft 8

With the urge for additional pattern recognition, we recall the PLP, which has a monotonic intensity function: constant (Figure 6.15), increasing, or decreasing; and add a CT-plot of multisystem, termed Multi-CT-plot. Figure 6.16 is a Multi-CT-plot for comparisons of three simulated PLPs with $n = 20$, at $\beta = 0.5, 1$, and 2 : time-truncated sampling, all set at $T = 1$. Three monotonic patterns are observed: a straight line segment for $\beta = 1$; a curve concave upward for $\beta < 1$ (decreasing); concave downward otherwise. Again, the mystery of a non-constant monotonic intensity lies in the concavities and/or curvature of the observed curves presented in Figures 6.4 and 6.5 for large seismological data sets. The accrued shapes of the CT-plots for monotonic intensity functions will facilitate the initial assessments of a reliability data analysis.

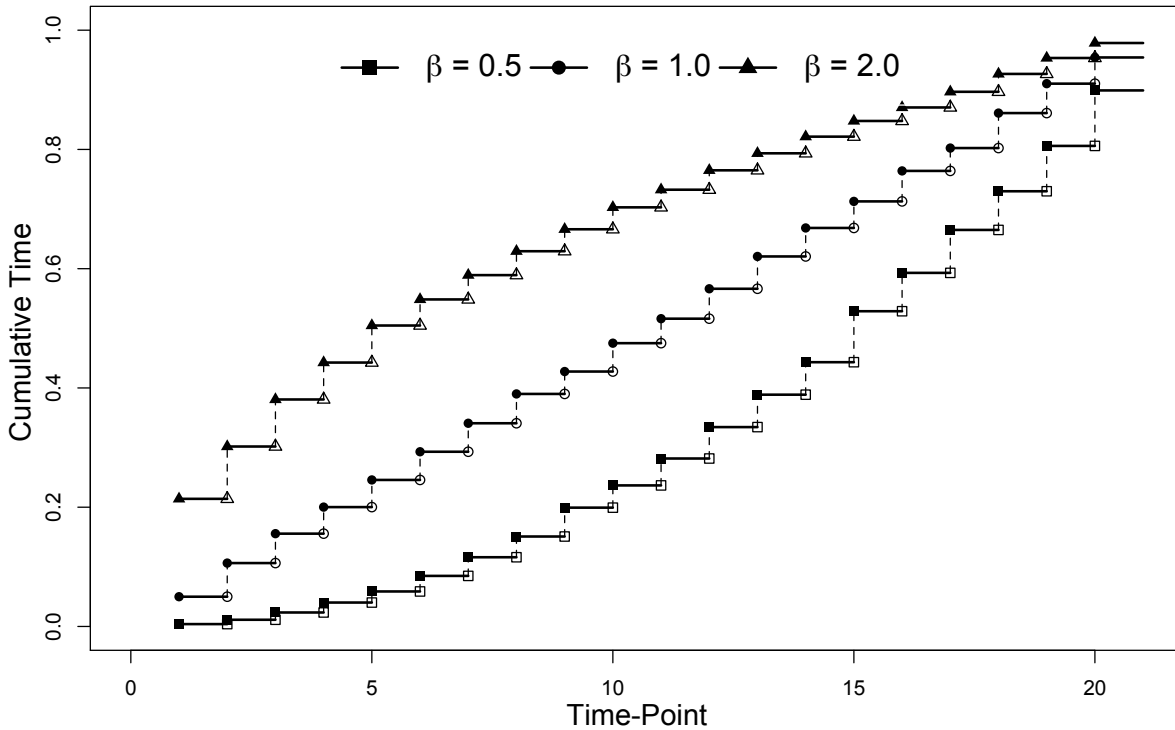


Figure 6.16. Multi-CT-plot for comparisons of intensity functions of three simulated PLPs with $n = 20$, at $\beta = 0.5, 1.0$, and 2.0 : time-truncated sampling, all set at $T = 1$.

We are ready to present a complete graphics panel of Aircraft 8 (Figure 6.17), facilitating analyses of all the other aircraft:

- The IET-plot depicts an increasing pattern cycle: from t_3 to t_8 ; t_9 to t_{17} ; t_{18} to t_{24} . The CT-plot reflects three connected segments from curves that concave downward – Both plots can exchange information and support each other.
- In the *OR-Z*-chart, the sequence of the observed test statistics follows a similar pattern mimicking that of the CT-plot for the original process. In contrast, traces of the reversed process are hollowed out, forming upward concavities (a decreasing trend) – They share findings of reliability trends from both directions in a single chart.

- L signals its only out-of-control at the 10th time-point, while all the others stay in-control based on the original process. For the reversal, the first out-of-control signals are: L at the 19th time-point, and all the others are at the 18th time-point – They are interconnected with power asymmetries and, consequently, offer a broad perspective on the data analysis.

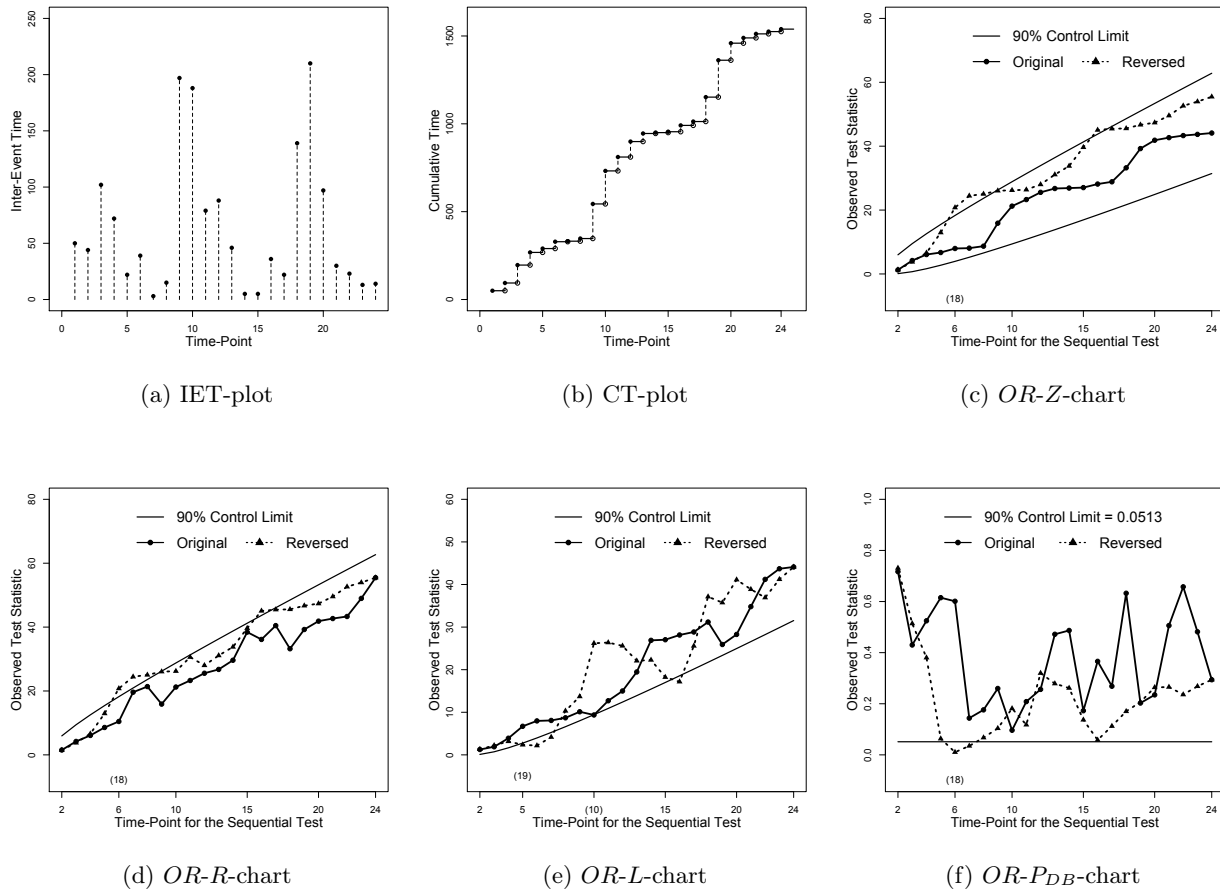


Figure 6.17. Graphics panel of (a) IET-plot, (b) CT-plot, (c) OR -Z-chart, (d) OR -R-chart, (e) OR -L-chart, and (f) OR - P_{DB} -chart for air-conditioning equipment on Aircraft 8.

The above analyses of Aircraft 8 provide a case study of the proposed graphics panel beyond a simple HPP (Aircraft 13). All the observations listed above indicate that every graph is a strong advocate for at least one of the others in a complete graphics panel. For instance, per the last point above, L detects multiple change-points for Aircraft 8, supported by both plots (Figure 6.17 (a)

and (b)), which are more complementary than competitive.

Data of all the remaining aircraft can be analyzed in the same fashion, except that towering dots/bars in an IET-plot often get our first attention, because they dwarf all of the others in a data set, and consequently, would squeeze and mask the curvature of the curves in the corresponding CT-plot due to the scaling of the y-values, especially if the data are small. For example, the IET-plot of Aircraft 2 shows two dominating failure times, and one of them occurs at the first time-point (Figure 6.3 (b)), which suppresses the appearance of the curvature anticipated in the CT-plot (Figure 6.6(b)).

6.5 Conclusions

There is always at least one story to tell in a graph, not to mention a panel of graphics:

- First, beware of power asymmetries! Look at data from a different angle, because system safety and reliability are at risk.
- We start the development by showing a simple IET-plot that shows that approximately 50% of the 120 years of large earthquakes occurred in Japan during the most recent 20 years. Dot-plots appear in various forms, and an IET-plot is indeed one of them. Towering dots/bars quickly get our first attention and/or concern, which allows us to get the first line of defense set in time. If the reliability data analysis were a film in the making, the IET-plot would be a film score, which encompasses an enormous variety of styles of music, depending on the nature of the films they accompany.
- In an IET-plot, we see a sequence of bouncing musical notes, while the counterpart, a CT-plot, transforms all of them into stairs edging up, needed to climb to reach the top of a mountain. A CT-plot cumulates all the dot-jumps up to each time-point and leaves us an often mysterious but meaningful footprint to shed any light on. The mystery of valuable

diagnostic elements is buried in the sequence of the slopes and/or curvature (Figures 6.4 and 6.5) associated with each corresponding hill. For instance, two slopes shown in Figure 6.7(b) follow a piecewise linear function of two, where the first appears to be steeper than the second for an increasing two-step-intensity function and vice versa (Figure 6.9(b)). Consequently, the above algorithm can be used to identify a nonmonotonic multi-step-intensity function, coupled with a Multi-CT-plot for comparisons of multisystem. Moreover, in a CT-plot, change-points of the process correspond to locations where the slopes change, if the patterns follow what has been described. The CT-plot has the edge over the IET-plot for plotting large data sets and multisystem.

- The degree of uncertainties and ambiguities projected by the characteristics and nature of the data delivered by the fingerprinting graphics could invite subjective expert opinions. Reliability engineering and technologies in these areas are well-advanced enough to incorporate the concepts and methodologies of statistical process control charts to efficiently consolidate all the valuable information in the graphics, built with the support of sound theories and methods. Along this line of argument, we then launch elaborated graphics panels of the *OR*-charts, recording the behaviors of the air conditioning equipment on 12 aircraft. Each all-in-one panel consists of a group of affiliated dual-bidirectional *OR*-charts, fit to safeguard power symmetry. The P_{DB} , declared as a so called all-purpose test, is a safe bet for practitioners doing applicable real-world case studies. The *OR*-charts, incorporating both processes of the original and its reversal in a single frame marking the first out-of-control signals from both directions in the charts, are a significant leap for modern reliability data science.
- In short, the graphics panel starts with an assortment of appetizers, including a set of two complimentary plots, {IET-plot and CT-plot}, followed by a cocktail of dual-bidirectional

control charts: $\{OR-Z\text{-chart}, OR-R\text{-chart}, OR-L\text{-chart}, \text{ and } OR-P_{DB}\text{-chart}\}$, designed to gain a broad perspective on the data analysis. Each of these graphics can coherently exchange information, support one another, share findings of reliability trends and the change-point(s) from both directions, and even interconnected with power asymmetries.

Graphics play a vital role in modern data science. In essence, this work represents a transformation and modernization of the fundamentals and classics in reliability data analysis. The coherence, utility, and merit of the entire structure development are elaborated with panels of relevant graphics based on a variety of data sets possessing a combination of: simulated or real, small or large, with known or unknown underlying distribution, and are supported by two fingerprinting plots considered as the so-called “prior” in the Bayesian data analysis. In summary, a cocktail of bidirectional tests is a social lubricant in the saga of absorbing reliability graphics to make it easier to collaborate on future technologies, projects, and ideas in reliability engineering and, by extension, other real-world case studies.

BIBLIOGRAPHY

- [1] S. Ali. A predictive bayesian approach to sequential time-between-events monitoring. *Quality and Reliability Engineering International*, 36:365–387, 2020.
- [2] S. Ali. Time-between-events monitoring using nonhomogeneous poisson process with power law intensity. *Quality and Reliability Engineering International*, 37:3157–3178, 2021.
- [3] M. Altun and S.V. Comert. A change-point based reliability prediction model using field return data. *Reliability Engineering and System Safety*, 156:175–184, 2016.
- [4] A. Amei., W. Fu, and C-H. Ho. Time series analysis for predicting the occurrences of large scale earthquakes. *International Journal of Applied Science and Technology*, 2(7):64–75, 2012.
- [5] H. Ascher. *Repairable Systems Reliability: Modeling, Inference, Misconceptions and Their Causes*. Marcel Dekker, 1984.
- [6] H.E. Ascher and H. Feingold. Application of laplace’s test to repairable system reliability. in *Proceedings of the International Conference on Reliability and Maintainability (Societe Pour la Diffusion des Sciences et des Arts, France)*, pages 219–225, 1978.
- [7] T. Ashton, N. Evangelopoulos, and V. Prybutok. Quantitative quality control from qualitative data: control charts with latent semantic analysis. *Qual. Quant.*, 49:1081–1099, 2015.
- [8] L.J. Bain and M. Engelhardt. Inferences on the parameters and current system reliability for a time truncated weibull process. *Technometrics*, 22(3):421–426, 1980.
- [9] L.J. Bain and M. Engelhardt. *Statistical Analysis of Reliability and Life-Testing Models Theory and Methods*. 2nd ed. New York, NY, USA: Marcel Dekker, 1991.
- [10] L.J. Bain, M. Engelhardt, and F.T. Wright. Tests for an increasing trend in the intensity of a poisson process: A power study. *Journal of the American Statistical Association*, 80(390):419–422, 1985.
- [11] D. Barry and J.A. Hartigan. A bayesian analysis for change point problems. *Journal of the American Statistical Association*, 88:309–319, 1993.

- [12] W.M. Bassin. Increasing hazard functions and overhaul policy. *Proceedings of the 1969 Annual Symposium on Reliability*, 48:173–178, 1969.
- [13] M.R.d. Bastos Martini, K. Kanoun, and J. Moreira de Souza. Software-reliability evaluation of the tropico-r switching system. *IEEE Trans. Reliability*, 39:369–379, 1990.
- [14] M. Bhaduri. Bi-directional testing for change point detection in poisson processes. *Ph.D. dissertation, Department of Mathematical Sciences, Univ. Nevada, Las Vegas, NV, USA*, 2018.
- [15] F. Centofani, A. Lepore, A. Menafoglio, and B. Palumbo. Functional regression control chart. *Technometrics*, 63:281–294, 2021.
- [16] J.H. Cha and M. Finkelstein. *Point Processes for Reliability Analysis: Shocks and Repairable Systems*. Springer, 2018.
- [17] M.J. Chandra. *Statistical Quality Control*. CRC Press, 2020.
- [18] Z. Chen, Y. Li, D. Zhou, T. Xia, and E. Pan. Two-phase degradation data analysis with change-point detection based on gaussian process degradation model. *Reliability Engineering and System Safety*, 216:107916, 2021.
- [19] X. Cia, Y. Tian, and W. Ning. Change-point analysis of the failure mechanisms based on accelerated life tests. *Reliability Engineering and System Safety*, 188:515–522, 2019.
- [20] J. Colin and M. Vanhoucke. Developing a framework for statistical process control approaches in project management. *International Journal of Project Management*, 33:1289–1300, 2015.
- [21] D.R. Cox. Some statistical methods connected with series of events. *Journal of the Royal Statistical Society Ser. B*, 17(2):129–164, 1955.
- [22] D.R. Cox and P.A.W. Lewis. *The Statistical Analysis of Series of Events*. London: Chapman and Hall, 1966.
- [23] L.H. Crow. Reliability analysis for complex repairable systems, reliability and biometry. *eds. F. Proschan and R. J. Serfling, Philadelphia, PA, USA: SIAM*., pages 379–410, 1974.
- [24] L.H. Crow. Confidence interval procedures for the weibull process with applications to reliability growth. *Technometrics*, 24(1):67–72, 1982.
- [25] S.V. Crowder, D.M. Hawkins, M.R. Reynolds, and E. Yashchin. Process control and statistical inference. *Journal of Quality Technology*, 29:134–139, 1997.

- [26] S. Dalal, E.B. Fowlkes, and B. Hoadley. Risk analysis of the space shuttle: Pre-challenger prediction of failure. *Journal of the American Statistical Association*, 84:945–957, 1989.
- [27] M.A.R. de Pascoa, E.M.M. Ortega, and G.M. Cordeiro. The kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, 8:411–433, 2011.
- [28] R. DeVor, T-H. Chang, and J. Sutherland. *Statistical Quality Design and Control: Contemporary Concepts and Methods*. 2nd ed. Pearson, 2006.
- [29] N. Doganaksoy, W.Q. Meeker, and G.G. Hahn. *Achieving Product Reliability*. Routledge, 2021.
- [30] H. Dong, N. Chen, and K. Wang. Modeling and change detection for count-weighted multi-layer networks. *Technometrics*, 62:184–195, 2019.
- [31] J.T. Duane. Learning curve approach to reliability monitoring. *IEEE Transactions on Aerospace*, 2:563–566, 1964.
- [32] M. Engelhardt and L.J. Bain. Statistical analysis of a compound power-law model for repairable systems. *IEEE Transactions on Reliability*, 36(4):392–396, 1987.
- [33] M. Engelhardt, J.M. Guffey, and F.T. Wright. Tests for positive jumps in the intensity of a poisson process: A power study. *IEEE Transactions on Reliability*, 39(3):356–360, 1990.
- [34] B. Epstein. *Mathematical Models for Systems Reliability*. Chapman and Hall, 2008.
- [35] S. Eryilmaz and C. Kan. Reliability and optimal replacement policy for an extreme shock model with a change point. *Reliability Engineering and System Safety*, 190:106513, 2019.
- [36] J.J. Faraway. *Linear Models with Python*. Chapman and Hall, 2021.
- [37] J.M. Finkelstein. Confidence bounds on the parameters of the weibull process. *Technometrics*, 18(1):115–117, 1976.
- [38] O. Gaudoin. Optimal properties of the laplace trend test for software-reliability models. *IEEE Transactions on Reliability*, 41:525–532, 1992.
- [39] I. Gertsbakh. *Reliability Theory with Applications to Preventive Maintenance*. Springer, 2005.
- [40] E. Grant and R. Leavenworth. *Statistical Quality Control*. 8rd ed. McGraw-Hill Science, 1996.
- [41] C. Hardwick. *Practical Control Charts: Control Charts Made Easy!* Liberation Media Ltd, 2019.

- [42] D.M. Hawkins. Fitting multiple change-point models to data. *Computational Statistics & Data Analysis*, 37:323–341, 2001.
- [43] D.M. Hawkins and Q. Deng. A nonparametric change-point control chart. *Journal of Quality Technology*, 42:165–173, 2010.
- [44] D.M. Hawkins, P. Qiu, and C.W. Kang. The change point model for statistical process control. *Journal of Quality Technology*, 35:355–366, 2003.
- [45] D.V. Hinkley. Inference about the change-point from cumulative sum tests. *Biometrika*, 58:509–523, 1971.
- [46] C-H. Ho. Bayesian analysis of volcanic eruptions. *Journal of Volcanology and Geothermal Research*, 43:91–98, 1990.
- [47] C-H. Ho. Nonhomogeneous poisson model for volcanic eruptions. *Mathematical Geology*, 23:167–173, 1991.
- [48] C-H. Ho. Statistical control chart for regime identification in volcanic time series. *Mathematical Geology*, 24:775–787, 1992.
- [49] C-H. Ho. Forward and backward tests for an abrupt change in the intensity of a poisson process. *Journal of Statistical Computation and Simulation*, 48:245–252, 1993.
- [50] C-H. Ho. Volcanic time trend analysis. *Journal of Volcanology and Geothermal Research*, 74:171–177, 1996.
- [51] C-H. Ho. Repeated significance tests on accumulating data of repairable systems. *Communications in Statistics - Theory and Methods*, 27:1181–1200, 1998.
- [52] C-H. Ho. Hazard area and recurrence rate time series for determining the probability of volcanic disruption of the proposed high-level radioactive waste repository at yucca mountain, nevada, usa. *Bulletin of Volcanology*, 72:205–219, 2010.
- [53] C-H. Ho and Bhaduri. M. A quantitative insight into the dependence dynamics of kilauea and mauna loa volcanoes, hawaii. *Mathematical Geosciences*, 49(7):893–911, 2017.
- [54] C-H. Ho and E.I. Smith. Volcanic hazard assessment incorporating expert knowledge: Application to the yucca mountain region, nevada, u.s.a. *Mathematical Geology*, 29:615–627, 1997.
- [55] C-H. Ho and E.I. Smith. A spatial-temporal/3-d model for volcanic hazard assessment: Application to the yucca mountain region, nevada. *Mathematical Geology*, 30:497–510, 1998.

- [56] C-H. Ho, E.I. Smith, D.L. Feuerbach, and T.R. Naumann. Eruptive probability calculation for the yucca mountain site, u.s.a.: Statistical estimation of recurrence rates. *Bulletin of Volcanology*, 54:50–56, 1991.
- [57] C-H. Ho, E.I. Smith, and D. Keenan. Hazard area and probability of volcanic disruption of the proposed high-level radioactive waste repository at yucca mountain, nevada, usa. *Bulletin of Volcanology*, 69:117–123, 2006.
- [58] C-H. Ho, G. Zhong, F. Cui, and M. Bhaduri. Modeling interaction between bank failure and size. *Journal of Finance and Bank Management*, 4(1):01–10, 2016.
- [59] M.D. Holland and D.M. Hawkins. A control chart based on a nonparametric multivariate change-point model. *Journal of Quality Technology*, 46:63–77, 2014.
- [60] E. James. *Statistical Process Control For Quality Improvement: A Training Guide To Learning SPC*. Pearson, 1991.
- [61] H. Jin, G. Yin, Yuan B., and F. Jiang. Bayesian hierarchical model for change point detection in multivariate sequences. *Technometrics*, 64:177–186, 2022.
- [62] N.L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*. Wiley-Interscience, 1994.
- [63] G.J. Kerns. *Introduction to Probability and Statistics Using R*. CRC Press LLC, 2010.
- [64] C.J. Kim, S.J. Lee, and S.H. Kang. Evaluation of feeder monitoring parameters for incipient fault detection using laplace trend statistic. *IEEE Transactions on Industry Applications*, 40:1718–1724, 1992.
- [65] L. Lee and S.K. Lee. Some results on inferences for the weibull process. *Technometrics*, 20:21–45, 1978.
- [66] N. Limnios and M. Nikulin. *Recent Advances in Reliability Theory, Methodology, Practice, and Inference*. Boston: Birkhauser, 2000.
- [67] C. Luko. *Presentation of Data and Control Chart Analysis*. 9th ed. ASTM International, 2018.
- [68] J.F. Macgregor. On-line statistical process control. *Chemical Engineering Progress*, 84:21–31, 1988.
- [69] J.W. McPherson. *Reliability Physics and Engineering*. Springer, 2019.

- [70] M. Modarres and M.P. Kaminskiy. *Reliability Engineering and Risk Analysis: A Practical Guide*. CRC Press, 2016.
- [71] D.C. Montgomery. *Introduction to Statistical Quality Control*. John Wiley, 1985.
- [72] D.C. Montgomery, J.B. Keats, G.C. Runger, and W.S. Messina. Integrating statistical process control and engineering process control. *Journal of Quality Technology*, 26:79–87, 1994.
- [73] D.C. Montgomery and C.M. Mastrangelo. Some statistical process control methods for auto-correlated data. *Journal of Quality Technology*, 23:179–193, 1991.
- [74] J. Oakland and R.J. Oakland. *Statistical Process Control*. 7th ed. Routledge, 2018.
- [75] A.C. Palm, R.N. Rodriguez, F.A. Spring, and D.J. Wheeler. Some perspectives and challenges for control chart methods. *Journal of Quality Technology*, 29:122–127, 1997.
- [76] F. Proschan. Theoretical explanation of observed decreasing failure rate. *Technometrics*, 18:375–383, 1963.
- [77] P. Qiu. Big data? statistical process control can help! *Statistica Sinica*, 74:329–344, 2020.
- [78] P. Qiu and C. Zou. Control chart for monitoring nonparametric profiles with arbitrary design. *Statistica Sinica*, 20:1655–1682, 2010.
- [79] R.N. Rai. *Repairable Systems Reliability Analysis (Performability Engineering Series)*. Wiley-Scrivener, 2020.
- [80] M. Riaz and R.J.M.M. Does. A process variability control chart. *Computational Statistics*, 24:345–368, 2009.
- [81] S.E. Rigdon and A.P. Basu. *Statistical Methods for the Reliability of Repairable Systems*. John Wiley and Sons, 2000.
- [82] H. Rinne. *The Weibull Distribution*. Chapman and Hall/CRC, 2008.
- [83] N.L. Roche-Carrier, G.D. Ngoma, Y. Kocaefe, and F. Erchiqui. Reliability analysis of underground rock bolters using the renewal process, the non-homogeneous poisson process and the bayesian approach. *IEEE International Journal of Quality & Reliability Management*, 37:223–242, 2020.
- [84] T.P. Ryan. *Statistical Methods for Quality Improvement*. John Wiley and Sons, 1989.

- [85] E. Saniga. Economic statistical control chart designs with an application to \bar{x} and r charts. *Technometrics*, 31:313–320, 1989.
- [86] J.P. Signoret and A. Leory. *Reliability Assessment of Safety and Production Systems: Analysis, Modelling, Calculations and Case Studies*. Springer, 2021.
- [87] G.M. Smith. *Statistical Process Control and Quality Improvement*. 5th ed. Pearson, 2003.
- [88] S. Tan, M. Bhaduri, and C-H. Ho. A statistical model for long-term forecasts of strong sand storms. *Journal of Geoscience and Environment Protection*, 2:16–24, 2014.
- [89] L.G. Tatum. Control charts for the detection of a periodic component. *Technometrics*, 38:152–160, 1996.
- [90] P.A. Tobias. *Applied Reliability*. Chapman and Hall, 2011.
- [91] K.P. Tran. *Control Charts and Machine Learning for Anomaly Detection in Manufacturing*. Springer, 2021.
- [92] S. Vander Weil, W.T. Tucker, F.W. Faltin, and N. Doganaksoy. Algorithmic statistical process control: Concepts and an application. *Technometrics*, 34:286–297, 1992.
- [93] D.G. Wardell, H. Moskowitz, and R.D. Plante. Run length distributions of special-cause control charts for correlated processes. *Technometrics*, 36:3–17, 1994.
- [94] D.J. Wheeler. *Understanding Statistical Process Control*. 3rd ed. SPC PRESS, 2010.
- [95] T.R. Willemain and G.C. Runger. Designing control charts using an empirical reference distribution. *Journal of Quality Technology*, 28:31–38, 1996.
- [96] W.H. Woodall and D.C. Montgomery. Research issues and ideas in statistical process control. *Journal of Quality Technology*, 31:376–386, 1999.
- [97] M. Xie, T.N. Goh, and V. Kuralmani. *Statistical Models and Control Charts for High-Quality Processes*. Springer, 2002.
- [98] K. Yang and P. Qiu. Adaptive process monitoring using covariant information. *Technometrics*, 63:313–328, 2021.
- [99] X. Zhao and E.D. Castillo. An intrinsic geometrical approach for statistical process control of surface and manifold data. *Technometrics*, 63:295–312, 2021.

- [100] C. Zou, P. Qiu, and D. Hawkins. Nonparametric control chart for monitoring profiles using change point formulation and adaptive smoothing. *Statistica Sinica*, 19:1337–1357, 2009.
- [101] C. Zou, Y. Zhang, and Z. Wang. A control chart based on a change-point model for monitoring linear profiles. *IIE Trans.*, 38:1093–1103, 2006.

CURRICULUM VITAE

Graduate College
University of Nevada, Las Vegas

Sung Keun Koo

skkoo0412@gmail.com

Degrees:

Bachelor of Arts - Hospitality Business, 2007
Michigan State University

Master of Science - Applied Statistics, 2016
DePaul University

Dissertation Title:

Bidirectional Testing for Repairable Systems Reliability: Power Asymmetries,
Panel of Control Charts, and Reliability Graphics

Dissertation Examination Committee:

Chairperson, Chih-Hsiang Ho, Ph.D.
Committee Member, Amei Amei, Ph.D.
Committee Member, Malwane Ananda, Ph.D.
Graduate Faculty Representative, Jee Woong Park, Ph.D.