



HAL
open science

Dynamic regression models and their applications in survival and reliability analysis

Xuan Quang Tran

► **To cite this version:**

Xuan Quang Tran. Dynamic regression models and their applications in survival and reliability analysis. Statistics [math.ST]. Université de Bordeaux, 2014. English. NNT : 2014BORD0147 . tel-01201910

HAL Id: tel-01201910

<https://theses.hal.science/tel-01201910>

Submitted on 18 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITE DE BORDEAUX

ÉCOLE DOCTORALE DE MATHÉMATIQUE & INFORMATIQUE

SPÉCIALITÉ : Mathématiques Appliquées et Calcul Scientifique

TRAN XUAN QUANG

**Title: Dynamic regression models and their
applications in Survival and Reliability analysis**

**Titre: Les modèles de régression dynamique et leurs
applications en analyse de survie et fiabilité**

Directeur de thèse : **M. Mikhail Nikulin, Professeur Emérite,
Professeur des Universités**

Soutenue le 26/09/2014

Membres du jury

M. Mikhail NIKULIN	Prof. Emérite, Université de Bordeaux	Directeur
Mme. Catherine HUBER-CAROL	Prof. Emérite, Université René Descartes	Président
M. Henning LÄUTER	Prof. Emérite, Université de Potsdam, Germany	Rapporteur
M. Mounir MESBAH	Prof. Université de Paris 6	Rapporteur
M. Nikolaos LIMNIOS	Prof. Université de Technologie de Compiègne	Examinateur
M. Léo GERVILLE-REACHE	MdC, HDR, Université de Bordeaux	Examinateur
M. Vincent COUALLIER	MdC, Université de Bordeaux	Examinateur

Acknowledgements

During my Ph.D, I have received a lot of help, encouragement from my teachers, colleagues, family and friends. I would like to take this opportunity to express my sincere gratitude to them.

I shall begin with the most important person for this thesis, my advisor, Emeritus Professor. M. Mikhail NIKULIN for his patience, motivation, enthusiasm and immense knowledge. He has been more than just a guide to me. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study. Also I would like to thank Mme. Voila NIKULINA. Since the day I landed here in France, she has helped me a lot.

I would like to thank my sincere gratitude to M. Mounir MESBAH, M. Henning LÄUTER for their valuable reports, to the president of the jury Mme. Catherine HUBER-CAROL for her kind suggestions and examiners of the defense, M. Nikolaos LIMNIOS , M. Vincent COUALLIER, M. Léo GERVILLE-REACHE for their precious remarks. Their works have guided me a lot during this thesis.

I sincerely thank my Lab for providing me with a great environment for the research. I extend my gratitude to all the team members of the department and also to the secretaries for their administrative material.

Many thanks to the Ministry of Education and Training of Vietnam and Campus France for providing me the opportunity for higher studies. Also I say thanks to my colleagues at Thai Binh University, Vietnam for their assistance in the past three years.

I would like to thank my gratitude to Muhammad Ramzan TAHIR who has collaborated and helped me a lot. Also thanks to Noureddine SAAIDIA for his collaborations during my Ph.D.

I'm more grateful to NGUYEN Nhu Ha, NGUYEN Thuy Linh for all the help you have given me in my life. I'm thankful to association of Vietnamese students in Bordeaux (UEVBx), especially to my friends, BUI Thi Xuan Hong, PHAM Hai Long, NGUYEN

Huy Chung, for making my life easy at Bordeaux. I hope we can all keep in touch.

I would like to specially thank my parents for giving birth to me at the first place and supporting me spiritually throughout my life, also my thanks go to my family-in-law, my brother, my big sister for their encouragements for me.

In the end I would like to thank my wife and my sons for all the extraordinary support, encouragement and love that I have received from them all along my life.

Thank you so much to everyone!

Abstract

This thesis was designed to explore the dynamic regression models, assessing the statistical inference for the survival and reliability data analysis. These dynamic regression models that we have been considered including the parametric proportional hazards and accelerated failure time models contain the possibly time-dependent covariates. We discussed the following problems in this thesis.

At first, we presented a generalized chi-squared test statistics \mathbf{Y}_n^2 that is a convenient to fit the survival and reliability data analysis in presence of three cases: complete, censored and censored with covariates. We described in detail the theory and the mechanism to used of \mathbf{Y}_n^2 test statistic in the survival and reliability data analysis. Next, we considered the flexible parametric models, evaluating the statistical significance of them by using \mathbf{Y}_n^2 and log-likelihood test statistics. These parametric models include the accelerated failure time (AFT) and a proportional hazards (PH) models based on the Hypertabastic distribution. These two models are proposed to investigate the distribution of the survival and reliability data in comparison with some other parametric models. The simulation studies were designed, to demonstrate the asymptotically normally distributed of the maximum likelihood estimators of Hypertabastic's parameter, to validate of the asymptotically property of \mathbf{Y}_n^2 test statistic for Hypertabastic distribution when the right censoring probability equal 0% and 20%.

In the last chapter, we applied those two parametric models above to three scenes of the real-life data. The first one was done the data set given by Freireich et al. [46] on the comparison of two treatment groups with additional information about log white blood cell count, to test the ability of a therapy to prolong the remission times of the acute leukemia patients. It showed that Hypertabastic AFT model is an accurate model for this dataset. The second one was done on the brain tumour study with malignant glioma patients, given by Sauerbrei & Schumacher [105]. It showed that the best model is Hypertabastic PH on adding five significance covariates. The third application was done on the data set given by Semenova & Bitukov [106] on the survival times of the multiple myeloma patients. We did not propose an exactly model for this dataset. Because of that was an existing one intersection of survival times. We, therefore, suggest fitting other dynamic model as Simple Cross-Effect model for this dataset.

Keywords: Accelerated Failure Time model, Covariates, Cox model, Dynamic regression models, Flexible parametric models, Generalized Chi-squared statistic, Goodness-of-fit tests, Hypertabastic survival models, Leukemia cancer, Malignant glioma cancer, Multiple Myeloma cancer, Proportional Hazards model, Survival data analysis.

Résumé

Cette thèse a été conçue pour explorer les modèles dynamiques de régression, d'évaluer les inférences statistiques pour l'analyse des données de survie et de fiabilité. Ces modèles de régression dynamiques que nous avons considérés, y compris le modèle des hasards proportionnels paramétriques et celui de la vie accélérée avec les variables qui peut-être dépendent du temps. Nous avons discuté des problèmes suivants dans cette thèse.

Nous avons présenté tout d'abord une statistique de test du chi-deux généralisée \mathbf{Y}_n^2 qui est adaptative pour les données de survie et fiabilité en présence de trois cas, complètes, censurées à droite et censurées à gauche avec les covariables. Nous avons présenté en détail la forme pratique de \mathbf{Y}_n^2 statistique en analyse des données de survie. Ensuite, nous avons considéré deux modèles paramétriques très flexibles, d'évaluer les significations statistiques pour ces modèles proposées en utilisant \mathbf{Y}_n^2 statistique. Ces modèles incluent du modèle de vie accélérés (AFT) et celui de hasards proportionnels (PH) basés sur la distribution de Hypertabastic. Ces deux modèles sont proposés pour étudier la distribution de l'analyse de la durée de survie en comparaison avec d'autres modèles paramétriques. Nous avons validé ces modèles paramétriques en utilisant \mathbf{Y}_n^2 . Les études de simulation ont été conçues.

Dans le dernier chapitre, nous avons proposé les applications de ces modèles paramétriques à trois données de bio-médicale. Le premier a été fait les données étendues des temps de rémission des patients de leucémie aiguë qui ont été proposées par Freireich et al. [46] sur la comparaison de deux groupes de traitement avec des informations supplémentaires sur les log du blanc du nombre de globules. Elle a montré que le modèle Hypertabastic AFT est un modèle précis pour ces données. Le second a été fait sur l'étude de tumeur cérébrale avec les patients de gliome malin, ont été proposées par Sauerbrei & Schumacher [105]. Elle a montré que le meilleur modèle est Hypertabastic PH à l'ajout de cinq variables de signification. La troisième demande a été faite sur les données de Semenova & Bitukov [106], à concernant les patients de myélome multiple. Nous n'avons pas proposé un modèle exactement pour ces données. En raison de cela était les intersections de temps de survie. Par conséquent, nous vous conseillons d'utiliser un autre modèle dynamique que le modèle de la Simple Cross-Effect à installer ces données.

Mots clés: Analyse de survie, Covariables, Modèle de Cox, Modèle des Hasards Proportionnels, Modèles de survie de Hypertabastic, Modèles des paramétriques flexibles, Modèle de régression dynamique, Modèle de vie accélérée, Cancer de leucémie, Cancer de gliome malin, Cancer de myélome multiple, Test d'ajustement, Statistique du test du Chi-deux généralisée.

Contents

Acknowledgements	i
Abstract	iii
Contents	v
List of Figures	viii
List of Tables	ix
General Introduction	1
1 Goals	1
2 Outline of Thesis	6
3 The data	7
3.1 Complete data	8
3.2 Censored failure time data	8
3.3 Censored failure time data with covariates	10
4 Some real-life data examples for survival and reliability analysis	11
5 Some basic notations	15
1 Generalized Chi-squared type tests	17
1 Introduction	17
2 A brief review of Pearson chi-squared tests	18
3 Generalized Chi-squared tests for complete data	25
3.1 Idea of test construction	25
3.2 Generalized Chi-squared test statistic	27
3.3 Remark on the choice of grouping cells	28
4 Some modification of chi-squared tests	29
5 Application to testing normality	30
5.1 Chi-squared test for normality	30
5.2 Power considerations	35
5.3 Testing normality for Millikan's measurements	40
6 Generalized chi-squared tests for right censored data	41

6.1	The right censored failure-time data	41
6.2	Maximum likelihood estimators for right censored data	43
6.3	Chi-squared test construction	45
6.4	Chi-squared test statistic	47
2	Parametric survival models	49
1	Introduction	49
2	The consideration models	50
2.1	Proportional hazards (PH) model	52
2.2	Accelerated failure time (AFT) model	54
3	Right censored failure-time data with covariates	58
4	Estimates of parameters	60
4.1	Parametric estimations	60
4.2	Asymptotic properties of the maximum likelihood estimators	63
4.3	Approximate confidence estimate of survival function	65
5	Generalized chi-squared tests for parametric survival models	66
5.1	Idea of test construction	67
5.2	Asymptotically properties of test statistic	68
5.3	Test statistic	70
5.4	On the choice of random cells boundaries	71
3	Validity of Hypertabastic models	73
1	Introduction	73
2	Hypertabastic distribution $H(a, b)$	74
3	Validity of $H(a, b)$ using generalized chi-squared tests	76
3.1	Complete data	76
3.2	Right censored data	78
4	Some other GOF tests for $H(a, b)$ distribution	80
5	Validity of Hypertabastic AFT model	82
5.1	Hypertabastic AFT model	82
5.2	MLE for the regression parameters of AFT_H model	84
5.3	Estimators of the reliability characteristic	86
5.4	Asymptotic distribution of the regression parameters estimators	86
5.5	Approximate CI for the reliability characteristic	88
5.6	Validity of Hypertabastic AFT model using \mathbf{Y}_n^2	88
6	Validity of Hypertabastic PH model	90

6.1	Hypertabastic PH model	90
6.2	MLE for the regression parameters	91
6.3	Estimators of the reliability characteristic	92
6.4	Asymptotic distribution of the parameters estimators	93
6.5	Approximate CI for the reliability characteristic	94
6.6	Validity of Hypertabastic PH model using \mathbf{Y}_n^2	95
7	Simulation study	97
7.1	Hypertabastic's parameters Estimation	97
7.2	Simulated distribution of \mathbf{Y}_n^2 statistic for $H(a, b)$	98
7.3	Power considerations	99
4	Some numerical applications	105
1	Introduction	105
2	Statistical analysis and inferences	106
2.1	Analysis of the quantities of the oil field	106
2.2	Analysis of remission duration from a clinical trial for acute leukemia	110
2.3	Analysis of brain tumour study	113
2.4	Analysis of multiple myeloma patients	119
	Conclusion and Perspectives	123
	Appendix	125
1	Publications During PhD	125
2	The data	126
2.1	Remission durations for acute leukemia study	126
2.2	Brain tumour study with malignant glioma patients	127
2.3	Multiple myeloma data	140
	Bibliography	143

List of Figures

1	KM estimates plots for remission duration of acute leukemia patients . . .	12
2	KM estimates for malignant glioma patients receiving different age groups	14
3	KM estimates of multiple myeloma patients receiving different treatment .	15
1.1	P.D.F. of $N(1, 0.2^2)$, $Log(1, 0.2)$, $L(1, 0.25)$ and $W(1, 5)$ distributions . . .	36
3.1	Hypertabastic hazard plot with difference of parameter (a, b)	75
3.2	Simulated ω_n^2 statistic and JB distributions	81
3.3	Simulated \mathbf{A}_n statistic and JB distributions	82
3.4	Simulated \mathbf{Y}_n^2 statistical distribution for $H(a, b)$ with $n = 200$, $r = 13$. . .	99
3.5	Power of tests for $H(a, b)$ against $LN(2, 0.5)$ alternative	100
3.6	Power of tests for $H(a, b)$ against $LL(2, 3)$ alternative	101
3.7	Power of tests for $H(a, b)$ against $BS(0.5, 2.5)$ alternative	101
3.8	Power of tests for $H(a, b)$ against $IG(0.5, 3.5)$ alternative	102
4.1	Density curves of consideration distributions for quantities of oil data . . .	108
4.2	E.C.D.F. and estimate C.D.F. curves for quantities of oil data	109
4.3	KM and Hypertabastic survival estimates plots for acute leukemia patients	111
4.4	KM, AFT_H survival estimates of acute leukemia receiving different groups	113
4.5	Hypertabastic rate of hazard curve for malignant glioma.	118
4.6	KM, PH_H estimates and 95% CI for malignant glioma patients	119
4.7	KM and fitted models without covariates of multiple myeloma patients . .	121

List of Tables

1	The quantities of 49 locations of the same area of an oil field.	11
1.1	The approximated critical values of Shapiro-Wilk, Shapiro-Francia statistics.	38
1.2	Power of test for normality against Logistic(1, 0.2) distribution.	39
1.3	Power of test for normality against Laplace(1, 0.25) distribution.	39
1.4	Power of test for normality against Weibull(1, 5) distribution.	39
1.5	The values of the charge of the electron, data of Millikan.	40
1.6	Values of generalized chi-squared and its modified tests for Millikan's dataset	41
3.1	Upper α percentage points of ω_n^2 and \mathbf{A}_n statistical distributions	82
3.2	Simulated average absolute errors of $\hat{\theta}_n$ versus their true value for $H(a, b)$.	98
3.3	Power of \mathbf{Y}_n^2 test when the data are right censored with $n = 200, k = 8$. .	104
4.1	Statistical inference for the quantities of oil data under the null hypotheses	107
4.2	Statistical inference for quantities of oil data when $2NPG$ used	109
4.3	Comparison models for remission duration of acute leukemia patients ($k = 7$).	112
4.4	95% CI for the remission duration of acute leukemia patients	112
4.5	The variables for malignant glioma trial	114
4.6	Values of -2ℓ and \mathbf{Y}_n^2 for PH_H stragety model	115
4.7	Comparison models for malignant glioma data ($k = 15$).	117
4.8	MLE of the parameters of PH_H model and 95% CI.	118
4.9	Fitted models for the multiple myeloma patients ($k = 7$).	120
4.10	Values of \mathbf{Y}_n^2 with different intervals for multiply myeloma data	121

General Introduction

1 Goals

Survival models have been studied and applied in several fields, such as medicine, public, health, epidemiology, engineering, social and behavioral sciences. During the 2nd half of 20th century, a lot of development is made and played a vital role in survival and reliability analysis. Some frequently used survival models that can be mentioned here include the Proportional Hazards (PH) (is so-called Cox) and Accelerated Failure Time (AFT) models.

The Cox model became famous and commonly used as soon after it was proposed by Cox [34] in 1972. This model is the most widely used in clinical analysis because of it has "the advantage of estimating the effect of covariates as the logarithm of hazard ratio without the need to estimate the baseline hazard" (Nelson et al. [81]). There are some references, for example, Cox [35], Andersen & Gill [5], Andersen et al. [6], Tsiatis [117], Hosmer et al. [56] which we recommend as the good discussions on the PH model and its applications. However, this model has also some limitations like the hazard ratio at any moment does not depend on the values of the variables before this moment, so that, the use of Cox model can lead to lost or omit the important clinical information. Furthermore, the behavior of the hazard function that can be more interesting to the medical research because it is directly related to the survival times of a type of disease. That is why many extensions have been made for the Cox model. We can cite here some recent significant contributions that have the good discussions, for instances, Qing Chen [30], Royston [99, 100], Royston & Parmar [102], Lin & Ying [73]. Qing Chen [30] proposed an accelerated hazards regression model to study the relationship between survival times and covariates through a scale change between hazard functions. Royston & Parmar [102] proposed a flexible parametric model using restricted cubic splines for censored survival data.

An alternate to Cox regression and perhaps a better way to evaluate the risks in cohorts however exists in terms of the AFT models, which are efficiently used in the time to event data. AFT models postulate a direct relationship between the logarithm of survival

time and covariates. Therefore, we can directly appraise the impact of the covariates on the survival time instead of the hazard ratio as in the Cox model. This characteristic allows for an easier interpretation of the results since the parameters measure the effect of the correspondent covariate on the mean survival time. A semi-parametric AFT model is much less used in survival analysis than the Cox model because of complicated estimation procedures. A non-parametric AFT model also can be used. However, a frequently used is a parametric AFT model. Different from the Cox model, the parametric AFT model is mostly applied to the failure time regression analysis and accelerated life testing. Some key references, for instance, Miller [78], Bagdonavičius [9], Meeker & Escobar [77], Bagdonavičius & Nikulin [14], Martinussen & Scheike [75], Kalbfleisch & Prentice [60].

On the other hand, in recent years, the investigation of the parametric models as the baseline functions for the PH or AFT models is also interested. In this regard, Hjort [54] remarked that "The success of Cox regression has perhaps had the unintended side-effect that practitioners too seldomly invest efforts in studying the baseline hazard... A parametric version ... if found to be adequate, would lead to more precise estimation of survival probabilities and related quantities and concurrently contribute to a better understanding of the phenomenon under study...". A basic assumption of the parametric models, one often used as the Exponential, Weibull, Gamma, Log-normal, Log-logistic to fit the data. We can outline some key references, such as, Royston & Lambert [101], Kleinbaum & Klein [65], Lee & Wang [69], Collett [33], Lawless [67]. Following these above contributions we see that using the parametric models, we can receive some advantages as

- The parametric models can be brought the smooth estimates of the hazard and survival functions for any combination of covariates. In addition, the parametric models can completely adapt to the data influenced by time-dependent covariates.
- Unlike the PH regression model, with the parametric models we can consider to fit the data with more time scales of interest.

These traditional parametric models commonly used in literature. However, these models may have some disadvantages, such as

- For the main parametric models, the Exponential, Weibull or Gamma have a hazard functions that always goes in the same direction with time-up, time-down or constant. This means that these models cannot fit the data well.
- For the Log-normal, Log-logistic distributions, although their hazard functions have one turning-point, their posture graphs sometimes shown the flexibility poorly. Specifically, the hazard rate of Log-normal has uniquely non-monotonic and Log-logistic

hazard rate has monotone decreasing or increases from 0 to its maximum value and then decreases to 0.

This indicate that the traditional parametric models are not enough flexible to represent real data adequately. So one often discovers the new parametric models with the rich properties of hazard functions. One of them is a Hypertabastic distribution which was introduced by Tabatabai, Bursac, Willams and Singh [113] in 2007. Hypertabastic distribution seems quite complicated to calculate because it is built from the class of Hyperbolic functions. As compared with traditional distributions, the Hypertabastic hazard rate function also can be monotone increasing, monotone decreasing or \cap -shaped in independence of parameter values. Apart from that, Hypertabastic distribution can be brought some following benefits.

- Hypertabastic can be used as the alternative for traditional distributions. This distribution also completely adapts to the data when the hazard ratio has monotone. For example, Tabatabai et al. [113] concluded that the hazard rate for glioma brain cancer patients was increasing function for both groups of misonidazole and without misonidazole treatments.
- When the Hypertabastic hazard function has \cap -shaped, this distribution may be played a role in competition with the Birnbaum-Saunders and Inverse Gaussian distributions. In this document, a part of simulation study carried out to demonstrate this advantage.
- The Hypertabastic hazard function sometimes has the decline rapidly at some period of time after the peak. This property allows this distribution becomes more flexible and compatible with the real-life datasets.

Based on the log-likelihood and Akaike Information Criterion (AIC) statistics, Tabatabai [113–115] also suggested using Hypertabastic distribution as the baseline function for parametric Cox and AFT models to fit the clinical data, such as, one data set from multiple myeloma, the second taken from a randomized study of glioma patients who underwent radiotherapy treatment with and without radiosensitizer misonidazole and other set from survival times of breast cancer patients. Although the log-likelihood ratio and AIC statistics are really good when they are used to select the significance variables or to evaluate the correlation of the regression parameters in such model, but they will not give any warning of that when the candidate models fit poorly. Therefore, these statistics are no more useful to fit for the real-life datasets influenced by many covariates.

We believe that the Hypertabastic survival models may be lead to more promise of statistical inferences if we use a good assessing adequacy of fit (or so-called a goodness-of-fit (GOF) test) for real-life data influenced by the covariates. In fact, there are many ways to test GOF, some frequently used are described in section 4-chapter 1 and section 4-chapter 3 of this document, though our primary focus is the generalized chi-squared statistic (\mathbf{Y}_n^2) that was proposed by Nikulin [83, 85] in 1973. Let us cite some good references that studied the tests are those of Pearson [95], Kolmogorov [66], Smirnov [110], Cramer [37], von Mises [124], Birnbaum [26], Anderson & Darling [7] and many other. There is no general theory that tells us how to choose a good test. Because that a selected test depends many factors including the quality of the data, properly designed and applied of test that one use. This often refers to the estimation and then finding the limit of distribution of the test statistic. However, for almost of tests, the only method currently available to get a limit distribution of test statistic is by simulation. **Then why we choose \mathbf{Y}_n^2 statistic?** Before answering this question, we give bellow a brief review of a historical development of \mathbf{Y}_n^2 .

The first to initiate the GOF test is Pearson in 1900 with the introduction of the chi-squared test statistic that is well-known as a Pearson's statistic. A century later, at the international conference *Goodness-of-Fit Tests and Model Validity* to celebrate the 100th anniversary of the landmark paper by Karl Pearson on chi-square GOF test, Cox [36] said "For perhaps 70 years following its introduction the chi-squared test was one of the most widely used tools of formal statistical analysis...". However, the limiting distribution of Pearson's statistic has χ^2 distribution only for simple null hypothesis. So, Pearson's statistic is not available for composite hypotheses, especially, when the data are not being always fully observation and more influenced by the possible time-dependent covariates. As a consequence, many investigations of modified Pearson's statistic are considered. The majority of these considerations are discussed in the direction of replacing the unknown parameters of the null hypotheses by their estimates based on the groups or original sample. For example, Fisher [44] shown that the distribution of Pearson's statistic has in limit χ^2 with the degrees of freedom that is reduced by the number of parameters estimated from the sample. Fisher's result is only correct when the parameters are estimated from the frequency vector by the method of minimum chi-squared Pearson's statistic. Later, Cramér [37] also indicated that Fisher's result remains valid if the parameters of the null hypothesis are replaced by the multinomial maximum likelihood estimator based on grouped data. It is evident that the frequency vector generally are not the sufficient statistics. This means that we taken the censored data if we used the grouped data to estimate the parameters of the null hypothesis. Thereafter, in 1954, Chernoff & Lehmann [31] clearly shown that

the limiting distribution of Pearson's statistic does not follow a χ^2 distribution if the parameters of the null hypotheses are estimated by the method of maximum likelihood estimators (MLE) based on the original sample.

Until 1973, Nikulin [83, 85] is the first who proposed a modified chi-squared test statistic for shape and scale distribution. The essence of this statistic is used the property of multivariate normal distribution of the vector which is based on the differences between two estimators of the probabilities to fall into grouping intervals: one estimator is based on the empirical distribution function, one other is based on the MLE of parameters of the null hypothesis using trivial sufficient statistic. Next, Nikulin, by using the property "best asymptotically normal" of MLE and Wald's method (see, Moore [80]), solved the problem completely for any null hypotheses. Apart from that, with Nikulin's statistic the cell boundaries are grouped as the random continuous functions using the predetermined probabilities for the cells. This approach does not change the characteristic of the chi-squared of the limiting distribution of Nikulin's statistic and it usually use in survival and reliability analysis. A year later, Rao & Robson [97] obtained the same result for the exponential family of distribution by using the method of moment to estimate the parameters. Since then this statistic refers as Nikulin-Rao-Robson (NRR) statistic.

After its introduction the NRR statistic was the most widely utilized in literature. Some good key references, have been published, we recommend, for instance, LeCam et al. [68], Drost [38, 39], Aguirre & Nikulin [3], Van der Vaart [118], Voinov [123], Voinov et al. [120], Bagdonavičius et al. [19], Voinov et al. [121] and many other. Apart from that, a lot of investigations done to recommend the application of chi-squared NRR test statistic in analysis of survival and reliability. For example, Habib & Thomas [51], Hollander & Peña [55], Zhang [126] considered natural modifications of the NRR statistic to the case of censored data.

In 2011, Bagdonavičius & Nikulin [16, 21], by developing the idea of Akritas [4] of comparing observed and expected numbers of failures in time intervals, introduced the similar test which we called generalized chi-squared test statistic (\mathbf{Y}_n^2) for the censored failure-time data. Later, the most of considerations of \mathbf{Y}_n^2 statistic and its applications are considered in recent years, such as, Voinov et al. [122], Gerville-Réache et al. [47], Haghighi & Nikulin [52], Bagdonavičius et al. [23, 24], Bagdonavičius & Nikulin [18], Nikulin & Tran [88], Goual & Seddik-Ameur [48], Dupuy [41], Nikulin et al. [92]. In the thesis of Tahir [116] at University Bordeaux I, he has discussed \mathbf{Y}_n^2 on the validation of some parametric models and its application to fit the Birnbaum-Saunders AFT model. Goual [48] also proposed the Inverse Generalized Weibull AFT model and applied it to the survival and reliability analysis using \mathbf{Y}_n^2 .

We see that the use of \mathbf{Y}_n^2 may be led some addition points as follows.

First, this statistic was taken the full information of the original data which are a trivial sufficient statistic to estimate the parameters of the null hypothesis. This means that all observation of the data are used. This statistic make sure overcome the obstacles of Pearson's statistic that still provide sufficient basic information of the sample.

Secondly, based on the MLE, this statistic used completely the Fisher information that was to measure the information relative to a parameter contained in a model selection. This gives a more objective assessments of the selected model of fit.

Further, this statistic is good adapted to fit the statistical models used in survival and reliability analysis, when the data are censored and we take into account the influence of covariates.

From these above reasons, *we choose a Hypertabastis baseline distribution as an object of our study and applying \mathbf{Y}_n^2 statistic to the assessing adequacy of fit for the parametric survival models based on this baseline distribution.*

Our goals, foremost, are to explain in detail the mechanism used of \mathbf{Y}_n^2 test statistic for the complete or right censored or right censored with time varying covariates in the survival and reliability analysis. Later, we present the statistical inferences of Hypertabastic survival models including the estimated parameters and the goodness-of-fit test. The tool which we use in the goodness-of-fit procedure to be \mathbf{Y}_n^2 test statistic. Finally, we discuss on the validation of these models above applying to the real-life datasets with the development of R-language.

2 Outline of Thesis

In addition the general introduction, conclusion, bibliography and appendix. This thesis includes four chapters.

In chapter 1, we first give a brief review of Pearson's statistic and some estimated methods for parameters of the null hypotheses. Next, we focus describe in detail our approach on the NRR statistic to fit the complete and non-complete data. We also illustrate the power of NRR test statistic in comparison with some other GOF tests for the testing of normality by simulation study. In this chapter, we also take the Millikan's measurements data (see, Linnik [74], Voinov et al. [121]), that described the values of the charge of the electron to validate the testing of normality using NRR statistic.

Chapter 2 addresses the question: how do we use the parametric survival models for the right-censored data influenced by possible time-dependent covariates? So, our primary

focus, in this chapter, is to consider some parametric survival models, including parametric PH and parametric AFT models. A theoretical global of the properties of these models as the estimations of the regression parameters and reliability characteristic are given explicit. Later, we consider the problem of testing for these models using \mathbf{Y}_n^2 .

Chapter 3 considers the validity of Hypertabastic survival models by using the theoretical global of \mathbf{Y}_n^2 which we discussed in the second chapter after presenting the properties of Hypertabastic distribution. Next, the simulation studies were carried out, the first is to illustrate the \sqrt{n} -consistent property of Hypertabastic parameters, the second is to prove the limiting distribution of \mathbf{Y}_n^2 do not change for the testing of Hypertabastic distribution and other is to demonstrative the flexibility of Hypertabastic distribution in comparison with Log-normal (LN), Log-logistic (LL), Birnbaum-Saunders (BS) and Inverse Gaussian (IG) distributions when the data are the right censored. We also demonstrate the performance of \mathbf{Y}_n^2 in comparison its value with Anderson-Darling, Cramér-von Mises-Simirnov statistics.

In the last chapter, we focus discuss the applications and the benefits of Hypertabastic survival models to adapt for the three cases of real-life datasets including: completely observation, right censored and right censored under the covariates. A complete dataset of the quantities of the oil field which we are taken to validate the Hypertabastic distribution using \mathbf{Y}_n^2 statistic. Next, three real-life datasets from three motivating studies are explored, the first is to discuss on the remission duration from a clinical trial for acute leukemia study by applying to the parametric Hypertabastic AFT model; The second data set from the brain tumour study of a randomized clinical trial comparing two chemotherapy regimens in patients with malignant glioma, is to consider for the Hypertabastic PH model; The third dataset was designed by the Hematology Center, in the Main Military Clinical Hospital to compare the time of responses in two groups of treatment, one treated the drug together Bortezomibe, other one received the drug without Bortezomibe. There is exiting an intersection of survival times of two patient groups. So, this dataset is to explore for several parametric survival models. Although those three datasets are already studied but we would like here to propose a more accurate model for each of datasets by using \mathbf{Y}_n^2 and log-likelihood statistics.

3 The data

We offer as follows the failure time data which the response of interest is the time from a well defined time origin to until some events happen or to the end-point time of

study τ . If the moment τ is stochastic, it is really necessary to consider the distribution of τ . Because, in this circumstance, the distribution of τ will affect to the reliability characteristics estimation. We can cite here some key references, for example, Martinussen & Scheike [75], Huber et al. [59], Shen [109]. Following the purpose of our works, we assume that the end-point τ is constant and suit the aim of each study.

Returning the survival analysis, suppose that the data $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ are available on the failure times of n individuals usually taken to be independent during an experiment. Follows that, the failure time data \mathbf{X} can be classified by the way: complete or censoring observations or censored with the time-dependent covariates (stresses, explanatory variables) or censored with covariates and observed degradation. For the our aims, we only describe here three scenes of the data, as follows.

3.1 Complete data

If the failure time data are complete observations, that is $X_i = T_i$ for all $i = 1, 2, \dots, n$ where

$$\mathbf{T} = (T_1, T_2, \dots, T_n)^T, \quad (1)$$

are independent identically distributed (i.i.d.) random variables, then the statistical analysis of parametric models is doing by classical way.

In this situation, if the parameter of the models are known, we can use any test statistic, on the contrary, if the parameters of models are unknown we can use the standard maximum likelihood method to estimate the unknown parameters and we propose using the \mathbf{Y}_n^2 statistic for the assessing adequacy of fit the completely data.

3.2 Censored failure time data

Today, the commonly failure time data \mathbf{X} are non-complete with the most common example being the *right censoring mechanisms*. Censoring occurs when the value of observation is only partially known, i.e. the time T is observed if $T \leq C$, $C > 0$ is called the right-censoring time. Otherwise, we only know that $T > C$. On the contrary, left-censoring means that the time T is observed if $T \geq C$. The right censoring mechanisms can be various the type I, type II or type III.

If n subjects are observed at fixed study time t then censoring is called Type I censoring, in this case, $C = t$ for all subjects.

If the study is terminated whenever a specified number r ($r < n$) of failures have occurred, it is called type II censoring. The time of the r^{th} failure is then defined as the

censoring time C for all subjects.

However, the commonly data which we observed in survival analysis are right censoring, therefore, our focus is to discuss on the right censoring data. Let us C_1, C_2, \dots, C_n are the random censoring times and if the failure times T_i, C_i are mutually independent positive random variables for all $i = 1, 2, \dots, n$. Let

$$X_i = T_i \wedge C_i, \quad \delta_i = \mathbf{1}_{\{T_i \leq C_i\}}, \quad i = 1, 2, \dots, n. \quad (2)$$

where $a \wedge b = \min\{a, b\}$ and $\mathbf{1}_A$ is an indicator of the event A . We obtain the following data

$$(X_1, \delta_1), (X_2, \delta_2), \dots, (X_n, \delta_n). \quad (3)$$

If $\delta_i = 1$, it is known that a failure occurs at the time $T_i = X_i$, else, if $\delta_i = 0$, then the failure occurs after the time X_i , i.e. the subject is censored by $C_i = X_i$. We note that the distribution of bi-variate element (X_i, δ_i) is not trivial, since the first component is continuous and the second element is discrete.

Right censored samples are often presented by different way, which is well adapted to the problems considered in survival and reliability analysis. In particular, we shall use the following counting processes for construction the chi-squared type tests for parametric survival models. Set

$$(N_1(t), Y_1(t), t \geq 0), (N_2(t), Y_2(t), t \geq 0), \dots, (N_n(t), Y_n(t), t \geq 0), \quad (4)$$

where $N_i(t) = \mathbf{1}_{\{X_i \leq t, \delta_i=1\}}$ is the number of failures of the i^{th} unit in the interval $[0, t]$, $Y_i(t) = \mathbf{1}_{\{X_i \geq t\}}$ indicate the number "at risk" of i^{th} unit just prior the moment t .

The advantage of using data presentation (4) is the values of counting processes

$$\{N_i(u), Y_i(u), 0 \leq u \leq t, i = 1, 2, \dots, n\},$$

is showed the history of failures and censorings up to time t . Indeed, the two presented ways of data presentation in (3) and (4) are equivalent. We shall describe in detail the presentation of data in terms of counting processes in section 6 - chapter 1 and section 3 - chapter 2 of this document.

When the data are the right censored failure time, one can use again the maximum likelihood method for parametric situation, and the Kaplan-Meier (KM) estimator in non-parametric case (see, Kaplan & Meier [61]) to estimate the survival functions. The tool which one can be used to the assessing adequacy of fit for the right censored data, one can

be cited here, some test statistics proposed by Habib & Thomas [51], Akritas [4], Hjort [53], Hollander & Peña [55], Kim [63] and Bagdonavičius et al. [16]. Of course, we suggest again using the statistic \mathbf{Y}_n^2 of Bagdonavičius et al. for this situation, since this statistic have the nice advantages as the presented above in section 1.

3.3 Censored failure time data with covariates

Besides observing the censored observations, the failure time data \mathbf{X} may be governed by individual characteristics of units, that are expressed as the covariates. We suppose that any covariate is given in terms of deterministic time function

$$z(\cdot) = (z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot))^T : [0, \infty) \longrightarrow \mathbf{R}^m, \quad z(\cdot) \in \mathbf{E}, \quad (5)$$

where \mathbf{E} is a set of all possible covariates and $z_i(\cdot)$ is a scalar function. Note that the covariate $z(\cdot)$ can be endogenous or exogenous. They can be categorized into two classes: the constant covariates over the time (and we note \mathbf{E}_1 is the set of all constant covariates) and the time-varying covariates (and we note \mathbf{E} in this situation). We also can classify the covariates as the constant, or step-stresses, or degradation stresses, or cyclic stresses, or random stresses (commonly used in reliability and degradation models). More about the type of covariates, one can see, Meeker & Escobar [77], Dupuy [41].

It is evident that the changes in the values of covariates imply the changes in the life conditions of patient. These changes can increase or decrease the risk of failure, and hence all evolution processes of population go more quickly or more slowly. For example, the survival time of cancer patients may be affected by the area of residence, age, gender, groups of treatment and some other characteristics. A shabby area of residence, a high age can make the failure time of cancer patient happen faster and contrary...

As a consequence, the information obtained from covariates data may be used in clinical trial and reliability in order to quantify the prediction of survival, to control and to improve the efficiency of different technology and procedures. So, to construct the dynamic regression models at first we must to determine the impact of covariates, and after to understand by which way they influence on survival. We present as follows the presentation of the right censored data with covariates.

Suppose that n items are observed and the i^{th} item is affected by covariates $z_i(\cdot) = (z_{i1}(\cdot), \dots, z_{im}(\cdot))^T \in \mathbf{E}$. The data are assumed to be independently right censored which

we present as follows

$$(X_1, \delta_1, z_1(s)), (X_2, \delta_2, z_2(s)), \dots, (X_n, \delta_n, z_n(s)), \quad 0 \leq s \leq \tau, \quad (6)$$

where X_i and δ_i , for all $i = 1, 2, \dots, n$, are defined in (2).

In this situation, one need the special methods for the special models. A test statistic \mathbf{Y}_n^2 which we propose to fit this kind of data, is the most consistent. In addition our motivation is to work on the parametric survival models and to apply them to each kind of data. These models is considered based on the several failure-time baseline function as Weibull, Log-normal, Log-logistic, Birnbaum-Saunders, Inverse Gaussian and Hypertabastic distributions. And the our discussed focus only include the dynamic parametric PH and AFT models using Hypertabastic distribution as the baseline function. So, before considering those two dynamic models, we give for illustration some well-known real-life data which we analyze in this document.

4 Some real-life data examples for survival and reliability analysis

Example 1 (The quantities of the oil field).

The data in Table 1 are taken from the book of Bagdonavičius et al. [19] to present the quantities T_i (in conditional units) of 49 locations of the same area of an oil field.

Table 1: The quantities of 49 locations of the same area of an oil field.

8.7	6.6	10.0	24.3	7.9	1.3	26.2	8.3	0.9	7.1	5.9	16.8	6.0	13.4
31.7	8.3	28.3	17.1	16.7	19.7	5.2	18.9	1.0	3.5	2.7	12.0	8.3	14.8
6.3	39.3	4.3	19.4	6.5	7.4	3.4	7.6	8.3	1.9	10.3	3.2	0.7	19.0
26.2	10.0	17.7	14.1	44.8	3.4	3.5							

Bagdonavičius et al. [19] suggested that this dataset is suitable for Log-normal distribution. We take again this dataset in Chapter 4(subsection 2.1) to illustrate the flexibility of Hypertabastic distribution using a chi-squared \mathbf{Y}_n^2 statistic.

Example 2 (Remission durations for acute leukemia study).

The times of remission from clinical trial for acute leukemia of children study were first introduced by Freireich et al. [46] in 1963. Since then, this dataset frequently studied in

literature, such as, Lawless [67] disused then confidence intervals (CI) of remission times using the KM estimator and Nelson-Aalen estimator, and fitted the Weibull distribution for this dataset containing the treatment groups variable. Klein & Moeschberger [64] also considered the same dataset to estimate the probability of survival times using the KM and Nelson-Aalen estimates. They also suggested that the relapse rates of acute leukemia are different in two treatment groups (6-MP and placebo) by using the stratified log rank test. Later, Lee & Wang [69] also concluded that the rates of relapse from remission are strongly different in the two groups.

Kleinbaum & Klein [65] concluded that there are different in two groups treatment and the treatment effect is confounded by the effect of $\log(\text{WBC})$. They also used the Cox model to fit this dataset. In 2014, using generalized chi-squared statistic, Nikulin & Tran [88] used the same data given non acceptable fit for Exponential AFT model. Figure 1 shows the KM estimates of survival function in two groups for each type of treatment.

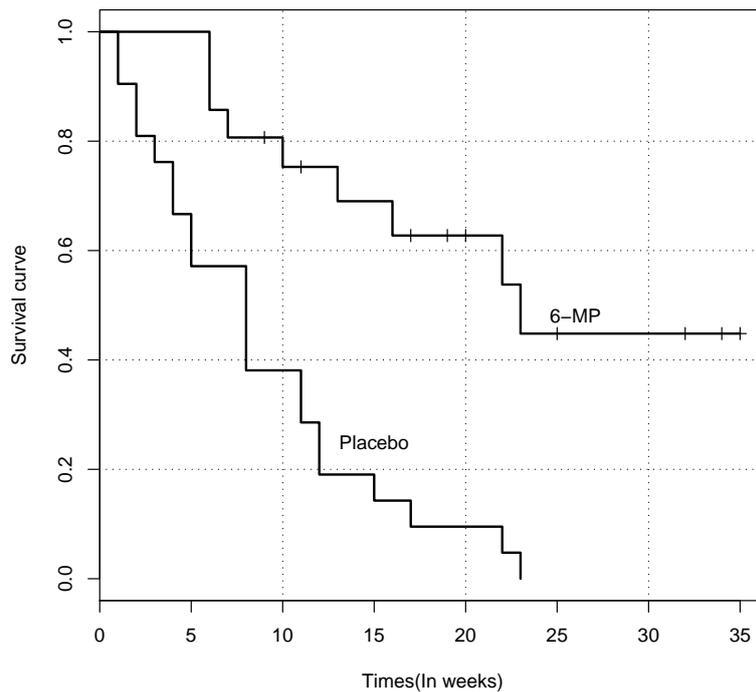


Figure 1: KM estimates plots for remission duration of acute leukemia patients receiving different groups.

We taken, in Chapter 4(subsection 2.2) of this document, the extended data including two effect variables: treatment variable and the index of log while blood cell ($\log(\text{WBC})$).

This dataset reproduced in subsection 2.1 of Appendix of this document.

Let us assume that we do not have any information of the model that is good to fit the remission duration of acute leukemia influenced by two groups of patients with additional the $\log(\text{WBC})$ variable for each patient. So, we need to propose an accurate model that can be fitted this dataset best. To do this task, we first take this dataset without covariates to fit for Hypertabastic distribution. Later, let us consider the parametric AFT models by taking the Hypertabastic distribution as the baseline hazard function to adapt the remission duration of acute leukemia patient under two consideration covariates. A GOF test procedure which we used to be \mathbf{Y}_n^2 test statistic. We also consider the role of applying the Hypertabastic AFT model in comparison with other parametric AFT models. Since then, we given 95% CI for the remission duration of the acute leukemia patients.

Example 3 (Brain tumour study with malignant glioma patients).

A randomized clinical trial comparing two chemotherapy regimens in 447 patients with malignant glioma that are studied by Sauerbrei & Schumacher [105]. During an accrual period of five years, 293 patients have died and the median survival time from the randomization was about 11 months.

The data contains 12 explanatory covariates including age, three ordinal and eight binary covariates. Three covariates are measured on an ordinal scale that are the index of Karnofsky, the type of surgical resection and the grade of malignancy. One of these covariates is represented by two dummy covariates resulting 16 covariates in all. Royston & Sauerbrei [103] have discussed the strategy selection of covariates analyzing with the PH model. Figure 2 illustrates the survival times of malignant glioma patients receiving different age groups. It shown that the age variable has the effect to the survival times of the patients.

This data set is used in Chapter 4 (subsection 2.3) to explorer the distribution of the survival times of the patients with malignant glioma under the significance covariates. For the overall model, we suggest using the parametric Hypertabastic PH model to fit this dataset, in comparison it with other parametric baseline distribution function using the \mathbf{Y}_n^2 statistic. We also describe the stratified Hypertabastic PH model for the selection of covariates as follows the approach of Collett [33].

Example 4 (Multiple myeloma study).

We consider the time of response of multiple myeloma patients that were carried out in The Hematology Center, in the Main Military Clinical Hospital named after N. N. Burdenko, to compare the response times in two patient groups of treatment. The difference

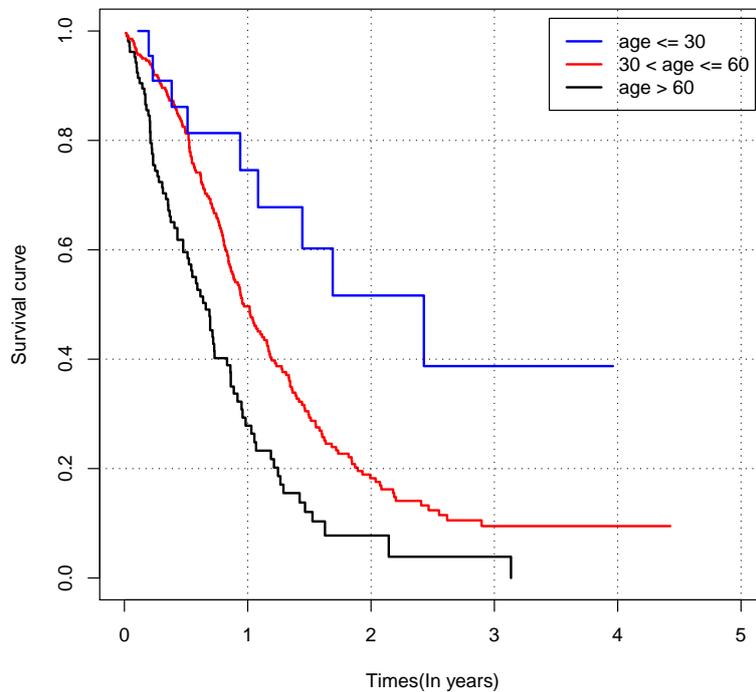


Figure 2: KM survival estimates for malignant glioma patients receiving different age groups.

in these groups is in the fact that the first group received the chemotherapy together Bortezomibe, which is marketed as Velcade by Millennium Pharmaceuticals, other received drugs without Bortezomibe. The data include 60 patients with 56 observations and 4 censored patients, influenced by four covariates: type of chemotherapy, type of response; sex; and age in years. Detail of the data are presented in Semenova & Bitukov [106], and reproduced in Appendix of this document. Figure 3 below illustrate the KM estimates of survival function for multiple myeloma patients receiving two groups of treatment.

From Figure 3, we see that the Cox model can be inappropriate for this dataset. Semenova & Bitukov [106] suggest using the parametric Log-normal Hsieh model included a combination of covariates using three statistics: Kolmogorov, Cramér von Mises-Smirnov and Anderson-Darling statistics. However, those three statistics are not consistent for the incomplete data with covariates. Therefore, we shall explore this dataset on adding a combination of all covariates to select a more accurate model.

We, in Chapter 4 (subsection 2.4), propose using the parametric survival models for relating the distribution of response time to the scheme of all covariates. The class of

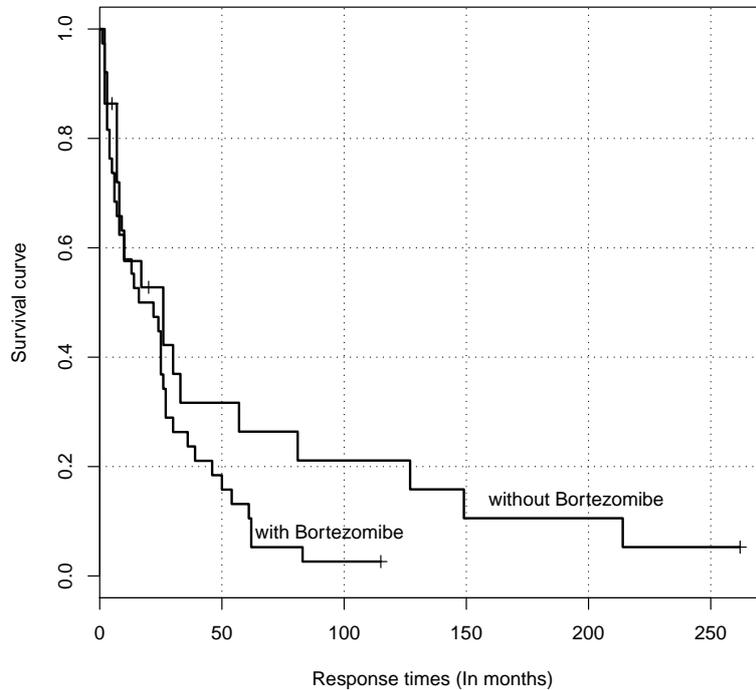


Figure 3: KM estimates of multiple myeloma patients receiving different treatment.

baseline parametric which we use to fit this dataset, including Hypertabastic, Weibull, Log-normal, Log-logistic, Birnbaum-Saunders and Inverse Gaussian distributions.

5 Some basic notations

Let us the positive random variable T denotes the failure time of a subject. The probability of a patient surviving up to time t is given by survivor function or reliability function $S(\cdot)$ as follows

$$S(t) = \mathbf{P}\{T > t\}, \quad t > 0.$$

The function

$$F(t) = 1 - S(t), \quad t > 0,$$

is called the cumulative distribution function (C.D.F.) of the failure time T . Further, the function

$$f(t) = \frac{dF(t)}{dt} = -S'(t), \quad t > 0,$$

is called the probability density function (P.D.F.) of T . However, mostly failure time data is expressed in terms of hazard rate function which can be written as

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{1}{h} \mathbf{P}\{t \leq T < t + h \mid T \geq t\} = -\frac{d[\ln S(t)]}{dt}, \quad t > 0.$$

The hazard rate specifies the instantaneous rate of mortality or failure at time t . It follows that $S(\cdot)$ can be written as

$$S(t) = e^{-\Lambda(t)}, \quad \text{where} \quad \Lambda(t) = \int_0^t \lambda(u) du, \quad t > 0,$$

is the cumulative hazard function of the failure time T . It is clear that the cumulative hazard function is an increasing function from $\Lambda(0) = 0$ to $\Lambda(\infty) = \infty$.

Chapter 1

Generalized Chi-squared type tests

1 Introduction

It is well-known that the goodness-of-fit (GOF) test procedure is an important tool to select the probability distribution suitable the data. In the development of statistical literature, there are many type of tests that have been developed. One of them is necessary to mention to be a commonly test statistic named a chi-squared measure that is proposed by Pearson [95] in 1900 and it, nowadays, is well-known as Pearson statistic. The distribution of the Pearson statistic in the limit has only chi-square when we take the test for simple hypothesis. Because this limiting distribution depends totally on the parameters of the fitted model. This means that it would not be a chi-squared if we fit the data for the composite hypothesis. Since then, many modifications and extensions of Pearson's statistic have been developed, such as, Fisher [44], Cramer [37], Chernoff & Lehmann [31], Roy [98], Plackett [96], Stigler [111]. However, there exist some disadvantages in the use of them for the validation. For example, Fisher showed that the limiting distribution of Pearson's statistic has a chi-squared with the degrees of freedom is reduced by the number of estimated parameters if and only if the minimum or maximum chi-squared based on the vector of frequencies, are used to estimate the parameters of fitted models. In reality, however, the frequency vector is not the unique vector that can be found from the initial data, more it in general is not the sufficient statistic. We cannot use this statistic to replace for original data. Furthermore, the estimation based on the vector of frequencies, are not best asymptotically efficient. Thus, it is necessary to find a better GOF test to make sure overcome the above obstacles that still provide sufficient basic information about the sample.

Note that it exists one famous modified classical chi-squared test that was introduced

by Nikulin [83, 85], is well adapted. The initially non-grouped data are used to estimate parameters of the fitted models, the endpoints of grouping intervals are taken as a random function depend on the parameters, based on the asymptotic properties of the multivariate normal distribution, a test statistic proposed by Nikulin good overcame the disadvantages of Pearson's test statistic as well as the test statistics proposed earlier. A year later, Rao and Robson [97] found the same result for the exponential family. That is why this statistic nowadays is well-known as Nikulin-Rao-Robson (NRR) statistic, for example, one can see, Drost [38, 39], Van de Vaart [118], Voinov et al. [119]. The direction of modified Pearson's statistic was continued develop by Dzhaparidze & Nikulin [42], Hsuan & Robson [58], Bolshev & Mirvaliev [27], Dudley [40], LeCam et al. [68], Nikulin & Mirvaliev [87], Nikulin et al. [90] and many other.

In addition the section of introduction, the consideration of the brief view about Pearson's statistic, this chapter include four section. We present our approach to give explicitly a chi-squared test based on NRR statistic for two cases of data: complete and right censoring observations in sections 3 and 6. Of course, we focus discussion our approach to build \mathbf{Y}_n^2 statistic, consider the selection of cell boundaries as the Neyman-Pearson (2NPG) classes to improve the power of \mathbf{Y}_n^2 statistic. In two sections 4 and 5, we consider some modifications of this statistic and employ on the simulated power of these tests for normality. An example of Millikan's data is also considered for the testing of normality.

2 A brief review of Pearson chi-squared tests

Let $\mathbf{T} = (T_1, T_2, \dots, T_n)^T$ be a simple sample, i.e. T_1, T_2, \dots, T_n are i.i.d. random variables. Suppose we have a family of distribution $F(t, \boldsymbol{\theta})$ in \mathbf{R} , $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)^T \in \Theta$, Θ is open in \mathbf{R}^m .

We wish to test the null composite hypothesis H_0 , according to which

$$H_0 : \mathbf{P}_{\boldsymbol{\theta}}(T_1 \leq t) = F(t, \boldsymbol{\theta}), \quad t \in \mathbf{R}^1, \quad \boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)^T \in \Theta. \quad (1.1)$$

Assume that for each $\boldsymbol{\theta} \in \Theta$ the measure $\mathbf{P}_{\boldsymbol{\theta}}$ is absolute continuous with respect to certain σ -finite measure μ , given on Borelian σ -algebra \mathcal{B} . Let us

$$f(t, \boldsymbol{\theta}) = \frac{d\mathbf{P}_{\boldsymbol{\theta}}}{d\mu}(t), \quad |t| < \infty,$$

is the density of the probability distribution $\mathbf{P}_{\boldsymbol{\theta}}$ with respect to the measure μ .

We define a measurable partition J_1, J_2, \dots, J_r of \mathbf{R} by the boundary points

$$-\infty = a_0 < a_1 < a_2 < \dots < a_r = +\infty,$$

such that

$$p_i(\boldsymbol{\theta}) = \mathbf{P}_{\boldsymbol{\theta}}(T_1 \in J_i) = \int_{a_{i-1}}^{a_i} dF(t, \boldsymbol{\theta}) =: p_i > 0, \quad i = 1, 2, \dots, r. \quad (1.2)$$

From (1.2), we see that $\sum_{i=1}^r p_i = 1$. Let $\boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_r)^T$ be a frequency vector obtained by grouping the initial random variables $\mathbf{T} = (T_1, T_2, \dots, T_n)^T$ over J_i . It is evident that

$$\nu_i = \int_{a_{i-1}}^{a_i} dF_n(t), \quad i = 1, 2, \dots, r,$$

where

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{T_i \leq t\}},$$

is the empirical cumulative distribution function (E.C.D.F.) obtained from the data T_1, T_2, \dots, T_n .

So, for each $\boldsymbol{\theta} \in \Theta$, we have the random vector

$$\mathbf{X}_n(\boldsymbol{\theta}) = \left(\frac{\nu_1 - np_1(\boldsymbol{\theta})}{\sqrt{np_1(\boldsymbol{\theta})}}, \frac{\nu_2 - np_2(\boldsymbol{\theta})}{\sqrt{np_2(\boldsymbol{\theta})}}, \dots, \frac{\nu_r - np_r(\boldsymbol{\theta})}{\sqrt{np_r(\boldsymbol{\theta})}} \right)^T, \quad (1.3)$$

and the random variable of Pearson's statistic

$$\mathbf{X}_n^2(\boldsymbol{\theta}) = \mathbf{X}_n^T(\boldsymbol{\theta})\mathbf{X}_n(\boldsymbol{\theta}) = \sum_{i=1}^r \frac{(\nu_i - np_i(\boldsymbol{\theta}))^2}{np_i(\boldsymbol{\theta})}. \quad (1.4)$$

If $\boldsymbol{\theta}$ is known equal $\boldsymbol{\theta}_0$, we have the next

Theorem 1.1 (Pearson, 1990). *If H_0 is true, then for any x positive*

$$\lim_{n \rightarrow \infty} \mathbf{P}_{\boldsymbol{\theta}_0} \{ \mathbf{X}_n^2(\boldsymbol{\theta}_0) \leq x \} = \mathbf{P} \{ \chi_{r-1}^2 \leq x \}, \quad (1.5)$$

where χ_{r-1}^2 has the chi-squared distribution with $r - 1$ degrees of freedom.

Theoretically, the Pearson's theorem is only valid when the null hypothesis H_0 is a simple and the data are grouped, i.e. parameters $\boldsymbol{\theta}$ is known and the endpoints a_i are defined in advance of the experiment. In fact, however, one do not known all of them. So, to use the Pearson's statistic, in general, one must overcome two above problems.

To take the first problem, one can choose one distribution $F(t, \boldsymbol{\theta}^*)$ from the family $\{F(t, \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\}$ in the best possible way using the data. Obviously, it is necessary to consider the distribution $\mathbf{P}_{\boldsymbol{\theta}_0}\{X_n^2(\boldsymbol{\theta}^*) \leq x\}$, where $\mathbf{P}_{\boldsymbol{\theta}_0} = \mathbf{P}_0$ is the true unknown distribution and $\boldsymbol{\theta}^*$ is our point estimators, replace for $\boldsymbol{\theta}$ in (1.4). In this regard, if we use the grouped data $\boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_r)^T$, we can obtain the minimum chi-squared $\{\tilde{\boldsymbol{\theta}}_n\}$, multinomial maximum likelihood $\{\check{\boldsymbol{\theta}}_n\}$ or modified minimum chi-squared $\{\bar{\boldsymbol{\theta}}_n\}$ estimators of $\boldsymbol{\theta}$ as follows the approach of Fisher [44] and Cramér [37]. On the contrary, when we use the simple sample $\mathbf{T} = (T_1, T_2, \dots, T_n)^T$ to estimate the parameters $\boldsymbol{\theta}$, we can obtain the MLE $\{\hat{\boldsymbol{\theta}}_n\}$. Besides giving some method of finding estimators $\boldsymbol{\theta}^*$ and their asymptotically properties, we consider below the effects of the methods estimators of parameters on the limiting distribution of $\mathbf{X}_n^2(\boldsymbol{\theta}^*)$ following the works of Greenwood and Nikulin [50], Cramér [37], Fisher [44], Chernoff and Lehmann [31], Hsuan and Robson [58], Chibisov [32].

At first, suppose that the following conditions of Cramér [37] hold

1. $0 < c < p_i(\boldsymbol{\theta})$, and $\sum_{i=1}^r p_i(\boldsymbol{\theta}) = 1$, for all $i = 1, 2, \dots, r$ and $\boldsymbol{\theta} \in \Theta$.

2. $\frac{\partial^2 p_i(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}$ are continuous functions on Θ .

3. The rank of the matrix $\mathbf{C} = \mathbf{C}(\boldsymbol{\theta}) = \|\frac{\partial p_i(\boldsymbol{\theta})}{\partial \theta_j}\|$ is m . So, the rank of the matrix $\mathbf{B} = \mathbf{B}(\boldsymbol{\theta})$ is m , where

$$\mathbf{B}(\boldsymbol{\theta}) = \left[\frac{1}{\sqrt{p_{ji}(\boldsymbol{\theta})}} \frac{\partial p_j(\boldsymbol{\theta})}{\partial \theta_i} \right]_{r \times m}.$$

The matrix $\mathbf{J}(\boldsymbol{\theta}) = \mathbf{B}^T(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})$ is called the Fisher information matrix of multinational distribution with parameters $\mathbf{p}(\boldsymbol{\theta})$ and $n\mathbf{J}(\boldsymbol{\theta})$ is called the Fisher information matrix in the vector of frequencies $\boldsymbol{\nu}$.

Minimum chi-squared estimator: Let $\tilde{\boldsymbol{\theta}}_n$ is an estimate of $\boldsymbol{\theta}$, such that

$$\mathbf{X}_n^2(\tilde{\boldsymbol{\theta}}_n) = \min_{\boldsymbol{\theta} \in \Theta} \mathbf{X}_n^2(\boldsymbol{\theta}) \iff \tilde{\boldsymbol{\theta}}_n = \operatorname{argmin}_{\boldsymbol{\theta} \in \Theta} \mathbf{X}_n^2(\boldsymbol{\theta}),$$

this is called the minimum chi-squared method of estimating $\boldsymbol{\theta}$, and $\{\tilde{\boldsymbol{\theta}}_n\}$ is called the sequence of minimum chi-squared estimators. It is so-called Fisher's estimators. Under some regularity conditions, the estimate $\{\tilde{\boldsymbol{\theta}}_n\}$ can be found by solving the system equations below

$$\sum_{i=1}^r \frac{\nu_i - np_i(\boldsymbol{\theta})}{p_i(\boldsymbol{\theta})} \frac{\partial p_i(\boldsymbol{\theta})}{\partial \theta_j} = 0, \quad j = 1, 2, \dots, m.$$

Multinomial maximum likelihood estimator: We see that the vector frequency $\boldsymbol{\nu}$

has the multinomial distribution $M_r(n, \mathbf{p})$, for $0 \leq n_i \leq n$, $\sum_{i=1}^r n_i = n$

$$\mathbf{P}_{\boldsymbol{\theta}}(\nu_1 = n_1, \nu_2 = n_2, \dots, \nu_r = n_r) = \frac{n!}{n_1! \dots n_r!} p_1^{n_1}(\boldsymbol{\theta}) \dots p_r^{n_r}(\boldsymbol{\theta}).$$

The likelihood function of $\boldsymbol{\nu}$ is

$$\ell(\boldsymbol{\theta}) = \frac{n!}{\nu_1! \dots \nu_r!} \prod_{i=1}^r p_i^{\nu_i}(\boldsymbol{\theta}) \dots p_r^{\nu_r}(\boldsymbol{\theta}).$$

The method of maximum likelihood is to use any estimator $\check{\boldsymbol{\theta}}_n$ which is maximize the likelihood function $\ell(\boldsymbol{\theta})$, that is

$$\check{\boldsymbol{\theta}}_n = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}} \ell(\boldsymbol{\theta}),$$

called the multinomial maximum likelihood estimator. In 1946, Cramér [37] showed that one can use the multinomial maximum likelihood estimator $\check{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}$ to replace for the minimum chi-squared estimator $\tilde{\boldsymbol{\theta}}_n$.

Modified minimum chi-squared estimator: Third method that also was proposed by Fisher [44], is so-called modified minimum chi-squared estimator $\bar{\boldsymbol{\theta}}_n$ minimizing the quadratic form

$$\bar{\mathbf{X}}_n^2(\boldsymbol{\theta}) = \sum_{i=1}^r \frac{(\nu_i - np_i(\boldsymbol{\theta}))^2}{\nu_i},$$

with respect to $\boldsymbol{\theta}$. Note that here we use the estimator $p_i = \frac{\nu_i}{n}$ ($i = 1, 2, \dots, r$) in the denominator of $\bar{\mathbf{X}}_n^2(\boldsymbol{\theta})$.

Using grouped data $\boldsymbol{\nu}$, we have three methods of finding estimators of $\boldsymbol{\theta}$, that are a minimum chi-squared, multinomial maximum likelihood and modified minimum chi-squared estimators, denote by $\{\tilde{\boldsymbol{\theta}}_n\}$, $\{\check{\boldsymbol{\theta}}_n\}$ and $\{\bar{\boldsymbol{\theta}}_n\}$, respectively. We use the notation $\bar{\bar{\boldsymbol{\theta}}}_n$ to denote any of them. One see that such sequences of estimates are asymptotically equivalent and have similar asymptotic properties, namely, the statistic $\bar{\mathbf{X}}_n(\bar{\bar{\boldsymbol{\theta}}}_n)$ have the same limit of distribution for $\bar{\bar{\boldsymbol{\theta}}}_n$ equal $\tilde{\boldsymbol{\theta}}_n$, $\check{\boldsymbol{\theta}}_n$ or $\bar{\boldsymbol{\theta}}_n$ in (1.4). Furthermore, such sequences of estimates $\bar{\bar{\boldsymbol{\theta}}}_n$ is consistent, that is

$$\mathbf{P}_{\boldsymbol{\theta}}\{|\bar{\bar{\boldsymbol{\theta}}}_n - \boldsymbol{\theta}| > \epsilon\} \rightarrow 0, \quad \text{and} \quad \sqrt{n}(\bar{\bar{\boldsymbol{\theta}}}_n - \boldsymbol{\theta}) \simeq AN(\mathbf{0}_s, \mathbf{J}^{-1}(\boldsymbol{\theta})), \quad \text{as } n \rightarrow \infty.$$

We have the next theorem

Theorem 1.2 (Fisher, 1928). *If H_0 is true and $\bar{\bar{\boldsymbol{\theta}}}_n$ is a consistent sequence estimators of*

$\boldsymbol{\theta}$ then

$$\lim_{n \rightarrow \infty} \mathbf{P}\{\mathbf{X}_n^2(\bar{\boldsymbol{\theta}}_n) > x\} = \mathbf{P}\{\chi_{r-m-1}^2 > x\}. \quad (1.6)$$

Fisher's theorem seems to solve the first problem by using the vector of frequencies $\boldsymbol{\nu}$. In fact that the data are often non-grouped, so that we present follow the method of Chernoff and Lehmann [31] to estimate the parameters $\boldsymbol{\theta}$ using non-grouped data.

Maximum likelihood estimator (MLE): In addition to Cramér's conditions 1.-3., we suppose that the following regularity conditions which lead up to the Chernoff and Lehmann [31] result, are hold

4. The second derivatives

$$\frac{\partial^2 f(t, \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}, \quad i, j = 1, 2, \dots, m,$$

of the density probability $f(t, \boldsymbol{\theta}) = \frac{d}{dt} F(t, \boldsymbol{\theta})$ exist and are continuous on Θ .

5. The Fisher information matrix of one observation T_i

$$\mathbf{i}(\boldsymbol{\theta}) = \mathbf{E}_{\boldsymbol{\theta}} \dot{\ell}_i(\boldsymbol{\theta}) \dot{\ell}_i^T(\boldsymbol{\theta}), \quad i = 1, 2, \dots, n, \quad (1.7)$$

is finite and positive definite for each $\boldsymbol{\theta} \in \Theta$, where

$$\dot{\ell}_i(\boldsymbol{\theta}) = \mathbf{grad} \log f(T_i, \boldsymbol{\theta}) = (\dot{\ell}_{i1}(\boldsymbol{\theta}), \dot{\ell}_{i2}(\boldsymbol{\theta}), \dots, \dot{\ell}_{im}(\boldsymbol{\theta}))^T,$$

is the vector-informant of the observation T_i , where

$$\dot{\ell}_{ij}(\boldsymbol{\theta}) = \frac{\partial \log f(T_i, \boldsymbol{\theta})}{\partial \theta_j}, \quad j = 1, 2, \dots, m.$$

6. Differentiation with respect to parameters under the integral sign in

$$\int_{-\infty}^{+\infty} f(t, \boldsymbol{\theta}) dt = 1,$$

is acceptable, that is

$$\frac{\partial}{\partial \theta_j} \int_{-\infty}^{+\infty} f(t, \boldsymbol{\theta}) dt = \int_{-\infty}^{+\infty} \frac{\partial}{\partial \theta_j} f(t, \boldsymbol{\theta}) dt = 0.$$

The identify

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f(T_i, \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \Theta, \quad (1.8)$$

is called the likelihood function of the simple sample $\mathbf{T} = (T_1, T_2, \dots, T_n)^T$. Note that \mathbf{T} is a trivial sufficient statistic that is only observed. So, if we use this sufficient statistic to estimate the parameters of the null hypothesis, we always ensure to keep and to use all information of the subjects.

We denote $\{\hat{\boldsymbol{\theta}}_n\}$ be a sequence of MLE of the parameter $\boldsymbol{\theta}$ based on the initial data \mathbf{T} maximizing the likelihood function $L(\boldsymbol{\theta})$ in (1.8), that is, for each n then

$$\dot{\ell}(\hat{\boldsymbol{\theta}}_n) = \mathbf{0}_m,$$

where $\dot{\ell}(\boldsymbol{\theta})$ is the vector-informant of sample $\mathbf{T} = (T_1, T_2, \dots, T_n)^T$, is given as

$$\dot{\ell}(\boldsymbol{\theta}) = \mathbf{grad} \log L(\boldsymbol{\theta}) = \sum_{i=1}^n \dot{\ell}_i(\boldsymbol{\theta}).$$

Under conditions 1.-6., the sequence of estimators $\hat{\boldsymbol{\theta}}_n$ satisfies that for any $\epsilon > 0$, exist C_ϵ, n_ϵ , such that

$$\mathbf{P}_{\boldsymbol{\theta}}\{\sqrt{n}|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}| > C_\epsilon\} < \epsilon, \quad \text{for all } n > n_\epsilon.$$

A sequence $\{\hat{\boldsymbol{\theta}}_n\}$ is so-called \sqrt{n} -consistent for parameter $\boldsymbol{\theta}$. Suppose that $\boldsymbol{\theta}_0$ is true value of $\boldsymbol{\theta}$. Under the conditions 1.-6., the asymptotic properties of $\{\hat{\boldsymbol{\theta}}_n\}$ can be given as

1. $\dot{\ell}(\hat{\boldsymbol{\theta}}_n) \xrightarrow{P} 0, \quad \hat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0,$
2. $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}}\mathbf{i}^{-1}(\boldsymbol{\theta})\dot{\ell}(\boldsymbol{\theta}) + O(1_m),$
3. $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} N_m(\mathbf{0}_m, \mathbf{i}^{-1}(\boldsymbol{\theta}_0)),$
4. $\frac{1}{\sqrt{n}}\dot{\ell}(\boldsymbol{\theta}_0) \xrightarrow{d} N_m(\mathbf{0}_m, \mathbf{i}(\boldsymbol{\theta}_0)),$
5. $-\frac{1}{n}\ddot{\ell}(\boldsymbol{\theta}_0) \xrightarrow{d} \mathbf{i}(\boldsymbol{\theta}_0), \quad -\frac{1}{n}\ddot{\ell}(\boldsymbol{\theta}_n) \xrightarrow{d} \mathbf{i}(\boldsymbol{\theta}_0),$

so, $\hat{\boldsymbol{\theta}}_n$ is "best asymptotically normal" (B.A.N.) for $\boldsymbol{\theta}$. From the above properties of $\{\hat{\boldsymbol{\theta}}_n\}$, we offer below a theorem of Chernoff & Lehmann [31], Greenwood & Nikulin [50].

Theorem 1.3 (Chernoff and Lehmann, 1954). *If the condition 1.-6. hold, based on the B.A.N. properties of $\hat{\boldsymbol{\theta}}_n$, then*

$$\lim_{n \rightarrow \infty} \mathbf{P}\{\mathbf{X}_n^2(\hat{\boldsymbol{\theta}}_n) < x \mid H_0\} = \mathbf{P}\{\chi_{r-m-1}^2 + \sum_{j=1}^m \lambda_j \xi_j^2 < x\}, \quad (1.9)$$

where, $\chi_{r-m-1}^2, \xi_1, \xi_2, \dots, \xi_m$ are independent random variables, $\xi_i \simeq N(0, 1)$ and $\lambda_i = \lambda_i(\boldsymbol{\theta})$ are the roots of the equation

$$|(1 - \lambda)\mathbf{i}(\boldsymbol{\theta}) - \mathbf{J}(\boldsymbol{\theta})| = 0.$$

So, in all cases if we use only grouped data to estimate the parameters of the fitted models, then only 3 cases the Pearson statistic $\mathbf{X}_n^2(\boldsymbol{\theta})$ has in limit a chi-squared distribution with $r - 1$ (in the case of Pearson), and $r - m - 1$ (in the cases of Fisher and Cramér) degrees of freedom. If we use initial data to estimate the parameters, then as follows the Chernoff & Lehmann's theorem, the limiting distribution of $\mathbf{X}_n^2(\hat{\boldsymbol{\theta}}_n)$ is totally changed and depends on $\boldsymbol{\theta}$. This means that the classical Pearson's test statistic is not easy to apply, particularly, in survival and reliability analysis.

Returning to the test of the composite hypothesis H_0 in (1.1), we assume that the boundaries are chosen, at least to some extent, in accordance with observations. In fact that $a_j = a_j(T_1, T_2, \dots, T_n)$ can be falsified the limiting distribution of $\mathbf{X}_n^2(\bar{\boldsymbol{\theta}}_n)$, since the vector of frequencies $\boldsymbol{\nu}^* = (\nu_1^*, \nu_2^*, \dots, \nu_r^*)^T$ based on $a_j = a_j(T_1, T_2, \dots, T_n)$ cannot be suitable a multinomial distribution. Furthermore, during to collect the data into grouping, we may be eliminated some information of observed realization of the random variables T_1, T_2, \dots, T_n . We have a remark below

Remark 1.1. *Note that the statistic $\boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_r)^T$ is not sufficient. Since*

$$\nu_j = \sum_{i=1}^n \mathbf{1}_{(a_{j-1}, a_j]}(T_i), \quad j = 1, 2, \dots, r,$$

*this means that some information is lost if we use only grouped data. More we say that the statistic $\boldsymbol{\nu}$ is obtained using the so-called **censoring mechanism of grouping data**. So this statistic cannot replace the full informative data $\mathbf{T} = (T_1, \dots, T_n)^T$. For example, the Fisher chi-squared test, based on the minimum chi-square estimator $\tilde{\boldsymbol{\theta}}_n$, is not the best if we have information that the statistic $\boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_r)^T$ is not sufficient.*

As follows remark 1.1 then the extensions of Pearson's statistic clearly bring out misconceptions and misunderstanding in the use of this method. Only the NRR statistic which is proposed by Nikulin ([83–86]) in 1973, Rao & Robson [97] in 1974, is the best restricted two problems above. Because that this statistic is constructed based on the MLE $\boldsymbol{\theta}$ using the trivial sufficient statistic (so it is the BAN), the cell boundaries are chosen as a random function depend the parameters and the vector of probabilities $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$. On the other hand, remark 1.1 also indicated that why we recommend using statistic of Nikulin to test the statistical models used in survival analysis and reliability, when the data are censored influenced by covariates. This problem will discuss detail in the next chapter. We refer here to the book of Greenwood & Nikulin [49, 50], where this statistic is described and explained in detailed.

3 Generalized Chi-squared tests for complete data

3.1 Idea of test construction

Suppose that the regularity conditions 1.-6. hold and the asymptotically properties of MLE's $\hat{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}$ are satisfied. We fix the vector of positive probabilities $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$, such that $\mathbf{p}^T \mathbf{1}_r = 1$ and $0 < p_i < 1$.

Let us $\boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_r)^T$ be the vector of frequencies obtained by grouping the initial data over the r intervals of the partition

$$(a_0, a_1(\hat{\boldsymbol{\theta}}_n)]; (a_1(\hat{\boldsymbol{\theta}}_n), a_2(\hat{\boldsymbol{\theta}}_n)]; \dots; [a_{r-1}(\hat{\boldsymbol{\theta}}_n), a_r), \quad (1.10)$$

$$\text{where } a_j(\boldsymbol{\theta}) = F^{-1}\left(\sum_{i=1}^j p_i, \boldsymbol{\theta}\right), a_0 = -\infty, a_r = +\infty, \quad (j = 1, 2, \dots, r-1), \quad (1.11)$$

and

$$F^{-1}(u, \boldsymbol{\theta}) = \inf_{0 < u < 1} \{t : F(t, \boldsymbol{\theta}) \geq u, \boldsymbol{\theta} \in \boldsymbol{\Theta}\},$$

is the inverse of the cumulative distribution function $F(t, \boldsymbol{\theta})$.

We know that the statistic $\boldsymbol{\nu}$ follows the multinomial distribution $M_r(\mathbf{p}, n)$ with parameters r, n, \mathbf{p} under H_0 . From the properties of the multinomial distribution it follows the next (see also, Nikulin et al. [90])

Theorem 1.4. *If $n \rightarrow \infty$, $r = \text{const} \geq 2$, then from the central limit theorem it follows under the H_0 the statistic $\boldsymbol{\nu}$ is asymptotically normal with parameters*

$$\mathbf{E}\boldsymbol{\nu} = n\mathbf{p} + O(\mathbf{1}_r) \quad \text{and} \quad \mathbf{E}(\boldsymbol{\nu} - n\mathbf{p})(\boldsymbol{\nu} - n\mathbf{p})^T = n\mathbf{D} + O(\mathbf{1}_{r \times r}), \quad (1.12)$$

where the matrix $\mathbf{D} = \mathbf{P} - \mathbf{p}\mathbf{p}^T$ is degenerated (since $\mathbf{1}_r^T \boldsymbol{\nu} = n$), hence $\text{rank}(\mathbf{D}) = r - 1$ and \mathbf{P} is a non-degenerated diagonal matrix with p_1, \dots, p_r on the main diagonal.

To proof this theorem one can apply the methods and techniques of Chernoff & Lehmann [31], Watson [125], Moore [79], Nikulin [83], Rao & Robson [97].

We similarly obtain the vector statistic $\mathbf{X}_n(\boldsymbol{\theta})$ and Pearson's sum $\mathbf{X}_n^2(\boldsymbol{\theta})$ in (1.3) and (1.4), respectively. We shall use one trivial theorem stating one important property of multivariate normal distribution $N_r(\mathbf{a}, \boldsymbol{\Sigma})$ according to which if we have a r -dimensional normal vector $\mathbf{Z} \sim N_r(\mathbf{a}, \boldsymbol{\Sigma})$ then

$$(\mathbf{Z} - \mathbf{a})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Z} - \mathbf{a}) \sim \chi_s^2,$$

where Σ^- is a general inverse of the matrix Σ , s is the rank of the matrix Σ , $s \leq r$. For example, in the case of simple hypothesis, the covariance matrix \mathbf{D} in (1.12) can be written as

$$\mathbf{D} = \mathbf{P} - \mathbf{p}\mathbf{p}^T, \quad \text{and} \quad \mathbf{D}^- = \mathbf{P}^{-1} + \frac{1}{p_r}\mathbf{1}_r\mathbf{1}_r^T,$$

since $\mathbf{1}_r^T\mathbf{P} = \mathbf{p}^T$, $\mathbf{1}_r^T\mathbf{p} = 1$, $\mathbf{P}\mathbf{1}_r = \mathbf{p}$. Note that by definition a general inverse matrix \mathbf{D}^- verifies the next relation

$$\mathbf{D}\mathbf{D}^-\mathbf{D} = \mathbf{D},$$

So the Pearson's vector $\mathbf{X}_n(\boldsymbol{\theta})$ in (1.3) can be written as

$$\mathbf{X}_n(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}}\mathbf{P}^{-1/2}(\boldsymbol{\nu} - n\mathbf{p}(\boldsymbol{\theta})), \quad (1.13)$$

Pearson's statistic $\mathbf{X}_n^2(\boldsymbol{\theta})$ therefore can be given as follows

$$\mathbf{X}_n^2(\boldsymbol{\theta}) = \frac{1}{n}(\boldsymbol{\nu} - n\mathbf{p}(\boldsymbol{\theta}))^T\mathbf{P}^{-1}(\boldsymbol{\nu} - n\mathbf{p}(\boldsymbol{\theta})). \quad (1.14)$$

From the central limit theorem, then the probability distribution of the statistic $\boldsymbol{\nu}$ can be approximated by a chi-squared distribution with $r - 1$ degrees of freedom. Thus, the probability distribution of $\mathbf{X}_n(\boldsymbol{\theta})$ can be adjoined the multivariate normal distribution, that is

$$\mathbf{X}_n(\boldsymbol{\theta}) \xrightarrow{d} \mathbf{X} \sim N_r(\mathbf{0}, \Sigma(\boldsymbol{\theta})),$$

where $\Sigma(\boldsymbol{\theta})$ is the covariance matrix of $\mathbf{X}_n(\boldsymbol{\theta})$, $\text{rank}(\Sigma) = r - 1$ and

$$\Sigma(\boldsymbol{\theta}) = \mathbf{P} - \mathbf{p}\mathbf{p}^T - \mathbf{W}(\boldsymbol{\theta})\mathbf{i}^{-1}(\boldsymbol{\theta})\mathbf{W}^T(\boldsymbol{\theta}), \quad (1.15)$$

where the matrix $\mathbf{W}(\boldsymbol{\theta}) = [w_{ij}]_{r \times m}$ include the following elements

$$w_{ij} := w_{ij}(\boldsymbol{\theta}) = \int_{a_{i-1}(\boldsymbol{\theta})}^{a_i(\boldsymbol{\theta})} \frac{\partial}{\partial \theta_j} f(t, \boldsymbol{\theta}) dt, \quad (i = 1, 2, \dots, r, j = 1, 2, \dots, m). \quad (1.16)$$

We obtained

$$\mathbf{Y}_n^2(\boldsymbol{\theta}) = \mathbf{X}_n^T(\boldsymbol{\theta})\Sigma^-(\boldsymbol{\theta})\mathbf{X}_n(\boldsymbol{\theta}) \xrightarrow{d} \mathbf{Y}^2 \sim \chi_{r-1}^2. \quad (1.17)$$

3.2 Generalized Chi-squared test statistic

Under the hypothesis H_0 , the generalized chi-squared test can be written as

$$\mathbf{Y}_n^2(\hat{\boldsymbol{\theta}}_n) = \mathbf{X}_n^T(\hat{\boldsymbol{\theta}}_n) \boldsymbol{\Sigma}^-(\hat{\boldsymbol{\theta}}_n) \mathbf{X}_n(\hat{\boldsymbol{\theta}}_n), \quad (1.18)$$

where $\boldsymbol{\theta}$ is replaced by the MLE $\hat{\boldsymbol{\theta}}_n$, and

$$\boldsymbol{\Sigma}^-(\hat{\boldsymbol{\theta}}_n) = \mathbf{P}^{-1} + \mathbf{P}^{-1} \mathbf{W}(\hat{\boldsymbol{\theta}}_n) [\mathbf{i}(\hat{\boldsymbol{\theta}}_n) - \mathbf{J}(\hat{\boldsymbol{\theta}}_n)]^{-1} \mathbf{W}^T(\hat{\boldsymbol{\theta}}_n) \mathbf{P}^{-1}. \quad (1.19)$$

If the regularity conditions hold, using the class-integrals with $a_j(\boldsymbol{\theta})$ equal a_i do not depend on $\boldsymbol{\theta}$ then the expression (1.16) can be written as

$$w_{ij} := \int_{a_{i-1}}^{a_i} \frac{\partial}{\partial \theta_j} f(t, \boldsymbol{\theta}) dt = \frac{\partial}{\partial \theta_j} \int_{a_{i-1}}^{a_i} f(t, \boldsymbol{\theta}) dt =: \frac{\partial p_i(\boldsymbol{\theta})}{\partial \theta_j}. \quad (1.20)$$

The statistic \mathbf{Y}_n^2 has a particularly convenient form which is the sum of the Pearson statistic \mathbf{X}_n^2 and a non-negative quadratic form \mathbf{Q} , and it can be presented as follows.

$$\begin{aligned} \mathbf{Y}_n^2 &= \mathbf{X}_n^2 + \mathbf{Q}, \quad (1.21) \\ \mathbf{X}_n^2 &= \sum_{i=1}^r \frac{\nu_i^2}{np_i} - n, \quad \mathbf{Q} = \frac{1}{n} \mathbf{v}^T \mathbf{G}^{-1} \mathbf{v}, \quad \mathbf{v} = (v_1, v_2, \dots, v_m)^T, \quad v_j = \sum_{i=1}^r \frac{c_{ij} \nu_i}{p_i}, \\ \mathbf{G} &= [g_{ll'}]_{m \times m}, \quad g_{ll'} = i_{ll'} - \sum_{i=1}^r \frac{c_{il} c_{il'}}{p_i}, \\ c_{ij} &= \frac{\partial p_i(\hat{\boldsymbol{\theta}}_n)}{\partial \theta_j} = f(a_i(\hat{\boldsymbol{\theta}}_n), \hat{\boldsymbol{\theta}}_n) \frac{\partial a_i(\hat{\boldsymbol{\theta}}_n)}{\partial \theta_j} - f(a_{i-1}(\hat{\boldsymbol{\theta}}_n), \hat{\boldsymbol{\theta}}_n) \frac{\partial a_{i-1}(\hat{\boldsymbol{\theta}}_n)}{\partial \theta_j}, \end{aligned}$$

and $i_{ll'}$ is the elements of the Fisher information matrix $\mathbf{i}(\hat{\boldsymbol{\theta}}_n)$ are defined in (1.7).

Theorem 1.5 (Nikulin, 1973). *Under the regularity conditions 1.-6., if the hypothesis H_0 is true then for any fixed positive x*

$$\lim_{n \rightarrow \infty} \mathbf{P}_{\boldsymbol{\theta}} \{ \mathbf{Y}_n^2(\hat{\boldsymbol{\theta}}_n) \leq x \} = \mathbf{P} \{ \chi_{r-1}^2 \leq x \}, \quad (1.22)$$

where χ_{r-1}^2 has a chi-squared distribution with $r - 1$ degrees of freedom.

3.3 Remark on the choice of grouping cells

We wish to test that the data are in concordance with the hypothesis H_0 at the significance level α . If the available data are non-grouped, then it is necessary to divide the real line into a finite number of intervals $I_j = (a_{j-1}, a_j]$ so the \mathbf{Y}_n^2 test statistic is more powerful. From the partition J_j in (1.10), we see that the endpoint a_i depends on the choice of vector $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$.

If there is no alternative, as follows Greenwood & Nikulin [50], we recommend using the partition J_i as *the equiprobable intervals (EQPG)* on the real line, that is

$$p_1 = p_2 = \dots = p_r = \frac{1}{r}, \quad \text{with } r \leq \min\left(\frac{1}{\alpha}, \ln n\right),$$

with this choice of r , the expected number of observation in each class is not small and the variance of the statistic \mathbf{Y}_n^2 is smallest, so that in this case the statistic \mathbf{Y}_n^2 is easier to calculate.

If there is an alternative, suppose that we wish to test for the null composite hypothesis

$$H_0 : \mathbf{P}_{\boldsymbol{\theta}}(T_1 \leq t) = F(t, \boldsymbol{\theta}), \quad t \in \mathbf{R}^1, \quad \boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)^T \in \boldsymbol{\Theta},$$

against the alternative

$$H_1 : \mathbf{P}_{\boldsymbol{\gamma}}\{T_1 \leq t\} = G(t, \boldsymbol{\gamma}), \quad \boldsymbol{\gamma} \in \mathbf{G} \subset \mathbf{R}^s.$$

We could choose \mathbf{p} to maximize the power of the statistic \mathbf{Y}_n^2 . Greenwood & Nikulin defined *the class of Neyman-Pearson (2NPG)* to construct the partition I_j as follows.

Let us $f(t, \boldsymbol{\theta})$ and $g(t, \boldsymbol{\gamma})$ are the density probability function of null and alternative hypotheses, respectively. To construct grouping cells I_j , we find the points a_1, a_2, \dots, a_r of intersections of two densities by solving equation $f(t, \hat{\boldsymbol{\theta}}_n) = g(t, \hat{\boldsymbol{\gamma}}_n)$, where $\hat{\boldsymbol{\theta}}_n$ and $\hat{\boldsymbol{\gamma}}_n$ are the MLE of parameters $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$, respectively. We obtain r intervals

$$I_1 = (-\infty, a_1] \cup [a_r, +\infty), \quad I_2 = (a_1, a_2], \quad \dots, \quad I_r = (a_{r-1}, a_r], \quad (1.23)$$

and two 2NPG classic given as

$$\mathbf{I}_1 = \{t : f(t, \hat{\boldsymbol{\theta}}_n) \geq g(t, \hat{\boldsymbol{\gamma}}_n)\}, \quad \mathbf{I}_2 = \{t : f(t, \hat{\boldsymbol{\theta}}_n) < g(t, \hat{\boldsymbol{\gamma}}_n)\}. \quad (1.24)$$

Based on these 2NPG classic, we obtain the frequency vector $\boldsymbol{\nu} = (\nu_1, \nu_2)^T$ and positive

probability vector $\mathbf{P} = (P_1, P_2)^T$ that

$$\begin{aligned} \nu_1 &= n \int_{\mathbf{I}_1} dF_n(t), & \nu_2 &= n \int_{\mathbf{I}_2} dF_n(t), \\ P_1(\boldsymbol{\theta}) &= \int_{\mathbf{I}_1} f(u, \hat{\boldsymbol{\theta}}_n) du =: P_1, & P_2(\boldsymbol{\theta}) &= \int_{\mathbf{I}_2} f(u, \hat{\boldsymbol{\theta}}_n) du =: P_2, \end{aligned}$$

where $F_n(\cdot)$ is the E.C.D.F. based on our data. We have, for $j = 1, 2, \dots, m$

$$c_{1j} = \frac{\partial}{\partial \theta_j} P_1(\hat{\boldsymbol{\theta}}_n), \quad c_{2j} = \frac{\partial}{\partial \theta_j} P_2(\hat{\boldsymbol{\theta}}_n), \quad \nu_1 + \nu_2 = n, \quad P_1 + P_2 = 1.$$

As in consequence, to test for H_0 against H_1 alternative, we can use a chi-squared test based on statistic

$$\mathbf{Y}_n^2 = \frac{\nu_1^2}{nP_1} + \frac{\nu_2^2}{nP_2} - n + \frac{1}{n} \mathbf{V}^T \mathbf{G}^{-1} \mathbf{V},$$

where

$$\begin{aligned} \mathbf{V} &= (V_1, V_2, \dots, V_m)^T, & V_j &= \frac{c_{1j}\nu_1}{P_1} + \frac{c_{2j}\nu_2}{P_2}, \quad j = 1, 2, \dots, m, \\ \mathbf{G} &= [g_{ll'}]_{m \times m}, & g_{ll'} &= i_{ll'} - \left(\frac{c_{1l}c_{1l'}}{P_1} + \frac{c_{2l}c_{2l'}}{P_2} \right). \end{aligned}$$

In this circumstance, the distribution of \mathbf{Y}_n^2 statistic has in limit a chi-squared with 1 degrees of freedom as n to ∞ . Thus, the hypothesis H_0 is rejected if $\mathbf{Y}_n^2 > \chi_\alpha^2(1)$.

Note that to take the 2NPG class with respect to the alternative close to each other, we selected parameters of alternative in such a manner that graphs of the alternative would be visually as close to each other as possible. This procedure is particularly useful when n is not large because the numbers of observations in just two grouping intervals is larger and the relevant central limit theorem is the Moivre-Laplace version. We discuss, in section 5 of this document, the use the 2NPG for testing alternative. A simulation study was carried out to demonstrate the power of \mathbf{Y}_n^2 test when the 2NPG classic is used to test normality against the Logistic, Laplace and Weibull alternatives.

4 Some modification of chi-squared tests

If the regularity conditions hold, the matrices $\mathbf{B}(\boldsymbol{\theta})$ in condition 3. and $\mathbf{W}(\boldsymbol{\theta})$ in (1.16) exist a simple connection that defined by expression (1.20). Using the expression

$$\mathbf{J}(\boldsymbol{\theta}) = \mathbf{W}^T(\boldsymbol{\theta}) \mathbf{P}^{-1} \mathbf{W}(\boldsymbol{\theta}),$$

and the identity

$$(\mathbf{I} - \mathbf{B}\mathbf{J}^{-1}\mathbf{B}^T)^{-1} = \mathbf{I} - \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T + \mathbf{B}\{(\mathbf{J} - \mathbf{B}\mathbf{B})^{-1} + (\mathbf{B}^T\mathbf{B})^{-1}\}\mathbf{B}^T,$$

McCulloch [76] presented (1.18) as

$$\mathbf{Y}_n^2(\hat{\boldsymbol{\theta}}_n) = \frac{1}{n}(\boldsymbol{\nu} - n\mathbf{p})^T \boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}}_n) \frac{1}{n}(\boldsymbol{\nu} - n\mathbf{p}) = \mathbf{D}_n^2(\hat{\boldsymbol{\theta}}_n) + \mathbf{S}_n^2(\hat{\boldsymbol{\theta}}_n),$$

where

$$\mathbf{D}_n^2(\hat{\boldsymbol{\theta}}_n) = (\boldsymbol{\nu} - n\mathbf{p})^T \left\{ \mathbf{P}^{-1} - \mathbf{P}^{-1/2} \mathbf{W}(\hat{\boldsymbol{\theta}}_n) [\mathbf{W}^T(\hat{\boldsymbol{\theta}}_n) \mathbf{W}(\hat{\boldsymbol{\theta}}_n)]^{-1} \mathbf{W}^T(\hat{\boldsymbol{\theta}}_n) \mathbf{P}^{-1/2} \right\} (\boldsymbol{\nu} - n\mathbf{p}),$$

is the Dzhaparidze-Nikulin [42] statistic that we note by DN statistic and

$$\mathbf{S}_n^2(\hat{\boldsymbol{\theta}}_n) = \frac{1}{n}(\boldsymbol{\nu} - n\mathbf{p})^T \mathbf{P}^{-1} \mathbf{W}(\hat{\boldsymbol{\theta}}_n) \left\{ [\mathbf{i}(\hat{\boldsymbol{\theta}}_n) - \mathbf{J}(\hat{\boldsymbol{\theta}}_n)]^{-1} + \mathbf{J}^{-1}(\hat{\boldsymbol{\theta}}_n) \right\} \mathbf{W}^T(\hat{\boldsymbol{\theta}}_n) \mathbf{P}^{-1} (\boldsymbol{\nu} - n\mathbf{p}),$$

the statistic $\mathbf{S}_n^2(\boldsymbol{\theta})$ is so-called McCulloch statistic (McCu).

McCulloch [76] has been proved that the statistics $\mathbf{D}_n^2(\hat{\boldsymbol{\theta}}_n)$ and $\mathbf{S}_n^2(\hat{\boldsymbol{\theta}}_n)$ are asymptotically independent and DN statistic $\mathbf{D}_n^2(\hat{\boldsymbol{\theta}}_n)$ has in limit distribution as χ_{r-m-1}^2 , while McCu statistic $\mathbf{S}_n^2(\hat{\boldsymbol{\theta}}_n)$ is distributed asymptotically as χ_m^2 . More about of those modifications, one can see Voinov et al.[121]. In the following section, we discuss the test of normality using the generalized chi-squared statistic \mathbf{Y}_n^2 following the contributions of Nikulin [85], Nikulin et al. [91].

5 Application to testing normality

5.1 Chi-squared test for normality

Let T_1, T_2, \dots, T_n are i.i.d. random variables, we consider to test the hypothesis H_0 , according to which

$$H_0 : T_1 \simeq \Phi\left(\frac{t - \mu}{\sigma}\right), \quad \boldsymbol{\theta} = (\mu, \sigma)^T \in \Theta = \{|\mu| < \infty, \sigma > 0\}, \quad t \in \mathbf{R}^1,$$

where $\Phi(\cdot)$ is the distribution function of the standard normal law.

The probability density function of T_i is

$$\frac{1}{\sigma} \varphi\left(\frac{t - \mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right), \quad t \in \mathbf{R}^1.$$

We see that there many results of test for normality have been contributed in the statistical literature. We consider here generalized chi-squared test statistic \mathbf{Y}_n^2 that are described in (1.21).

Returning to the hypothesis H_0 for normality, it is clear that

$$\mathbf{E}_\theta T_i = \mu, \quad \text{and} \quad \mathbf{Var}_\theta T_i = \sigma^2, \quad i = 1, 2, \dots, n.$$

If parameter $\theta = (\mu, \sigma^2)^T$ is unknown, then we can estimate it using the data by the MLE's $\hat{\theta}_n$. If the H_0 is true then the likelihood function $\ell(\cdot)$ on Θ can be given as

$$\ell(\theta) = \frac{(2\pi)^{-n/2}}{\sigma^n} \exp\left\{-\frac{1}{2\sigma^2} \left[\sum_{i=1}^n T_i^2 - 2\mu \sum_{i=1}^n T_i + n\mu^2 \right]\right\},$$

from where one can see that a minimal sufficient statistic for θ is

$$\mathbf{T} = \left\{ \sum_{i=1}^n T_i, \quad \sum_{i=1}^n T_i^2 \right\},$$

and the statistics

$$\bar{T}_n = \frac{1}{n} \sum_{i=1}^n T_i, \quad \bar{s}_n^2 = \frac{1}{n} \sum_{i=1}^n (T_i - \mu)^2, \quad s_n^2 = \frac{1}{n} \sum_{i=1}^n (T_i - \bar{T}_n)^2,$$

are the MLE's for the components μ and σ^2 of the parameter θ . It is well known that for the regular family of distributions there exists a sequence of MLE's $\hat{\theta}_n$ such that $\hat{\theta}_n = \arg \max_{\theta \in \Theta} \ell(\theta)$, where $\ell(\theta)$ is the likelihood functions of T_1, T_2, \dots, T_n , and

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N_2(\mathbf{0}, \mathbf{i}^{-1}(\theta)),$$

where $\mathbf{i}(\theta)$ is the Fisher's information matrix of one observation X_1 .

$$\mathbf{i}(\theta) = \mathbf{i}(\mu, \sigma^2) = \frac{1}{\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \theta = (\mu, \sigma^2)^T.$$

For $i = 1, 2, \dots, n$. Set

$$\eta_i = \begin{cases} \frac{1}{\sigma}(T_i - \mu), & \text{if } \mu \text{ and } \sigma \text{ are known,} \\ \frac{1}{\sigma} \sqrt{\frac{n}{n-1}}(T_i - \bar{T}_n), & \text{if } \mu \text{ is unknown, } \sigma \text{ is known,} \\ \frac{1}{\bar{s}_n}(T_i - \mu), & \text{if } \mu \text{ is known, } \sigma \text{ is unknown,} \\ \frac{1}{s_n}(T_i - \bar{T}_n), & \text{if } \mu \text{ and } \sigma \text{ are known.} \end{cases}$$

Let $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$ be the vector of positive probabilities, $0 < p_i < 1$, such that $\mathbf{p}^T \mathbf{1}_r = 1$ and let define the quantiles of the standard normal distribution

$$-\infty = x_0 < x_1 < \dots < x_{r-1} < x_r = +\infty, \quad \text{with} \quad x_j = \Phi^{-1}\left(\sum_{i=1}^j p_i\right), \quad j = 1, 2, \dots, r-1.$$

Suppose that $\nu = (\nu_1, \nu_2, \dots, \nu_r)^T$ is a frequency vector obtained by grouping the statistics η_1, \dots, η_n over the r intervals of the partition $(x_0, x_1], (x_1, x_2], \dots, (x_{r-1}, x_r)$. It is evident that the same statistic is obtained by grouping the initial random variables T_1, T_2, \dots, T_n over the random intervals J_i , ($i = 1, \dots, r$)

$$J_i = \begin{cases} (x_{i-1}\sigma + \mu; x_i\sigma + \mu], & \text{if } \mu \text{ and } \sigma \text{ are known,} \\ (x_{i-1}\sigma + \bar{T}_n; x_i\sigma + \bar{T}_n], & \text{if } \mu \text{ is unknown, } \sigma \text{ is known,} \\ (x_{i-1}\tilde{s}_n + \mu; x_i\tilde{s}_n + \mu], & \text{if } \mu \text{ is known, } \sigma \text{ is unknown,} \\ (x_{i-1}s_n + \bar{T}_n; x_i s_n + \bar{T}_n], & \text{if } \mu \text{ and } \sigma \text{ are unknown.} \end{cases}$$

It is clear that if H_0 is true then for any $i = 1, \dots, n$

$$\mathbf{P}\{\eta_i \in (x_{j-1}, x_j]\} = \mathbf{P}\{T_i \in J_j\} = p_j, \quad j = 1, 2, \dots, r.$$

As follows the theorem 1.4, under the H_0 the statistic ν is asymptotically normal with parameters

$$\mathbf{E}\nu = n\mathbf{p} + O(\mathbf{1}_r) \quad \text{and} \quad \mathbf{E}(\nu - n\mathbf{p})(\nu - n\mathbf{p})^T = n\mathbf{B} + O(\mathbf{1}_{r \times r}),$$

where

$$\mathbf{B} = \begin{cases} \mathbf{P} - \mathbf{p}\mathbf{p}^T, & \text{if } \mu \text{ and } \sigma \text{ are known,} \\ \mathbf{P} - \mathbf{p}\mathbf{p}^T - \mathbf{a}\mathbf{a}^T, & \text{if } \mu \text{ is unknown, } \sigma \text{ is known,} \\ \mathbf{P} - \mathbf{p}\mathbf{p}^T - \frac{1}{2}\mathbf{b}\mathbf{b}^T, & \text{if } \mu \text{ is known, } \sigma \text{ is unknown,} \\ \mathbf{P} - \mathbf{p}\mathbf{p}^T - \mathbf{a}\mathbf{a}^T - \frac{1}{2}\mathbf{b}\mathbf{b}^T, & \text{if } \mu \text{ and } \sigma \text{ are unknown,} \end{cases}$$

the matrix \mathbf{B} is degenerated, $\text{rank}(\mathbf{B}) = r - 1$ in all considered 4 cases and

$$\mathbf{a} = (a_1, \dots, a_r)^T, \quad a_j = \varphi(x_j) - \varphi(x_{j-1}), \quad (1.25)$$

$$\mathbf{b} = (b_1, \dots, b_r)^T, \quad b_j = -x_j\varphi(x_j) + x_{j-1}\varphi(x_{j-1}). \quad (1.26)$$

Set

$$X_n(\theta) = \left(\frac{\nu_1 - np_1}{np_1}, \frac{\nu_2 - np_2}{np_2}, \dots, \frac{\nu_r - np_r}{np_r} \right)^T,$$

the limiting distribution of $X_n(\theta)$ can be approximated by the multivariate normal law, that is

$$X_n(\theta) \xrightarrow{d} \mathbf{X} \sim N_r(\mathbf{0}, \mathbf{\Sigma}), \quad \text{rank}(\mathbf{\Sigma}) = r - 1,$$

where $\mathbf{\Sigma}$ is the covariance matrix of $X_n(\theta)$

$$\mathbf{\Sigma} = \mathbf{P}^{-1/2} \mathbf{B} \mathbf{P}^{-1/2} = \begin{cases} \mathbf{E}_r - \mathbf{q} \mathbf{q}^T, & \text{if } \mu \text{ and } \sigma \text{ are known,} \\ \mathbf{E}_r - \mathbf{q} \mathbf{q}^T - \mathbf{P}^{-1/2} \mathbf{a} \mathbf{a}^T \mathbf{P}^{-1/2}, & \text{if } \mu \text{ is unknown, } \sigma \text{ is known,} \\ \mathbf{E}_r - \mathbf{q} \mathbf{q}^T - \frac{1}{2} \mathbf{P}^{-1/2} \mathbf{b} \mathbf{b}^T \mathbf{P}^{-1/2}, & \text{if } \mu \text{ is known, } \sigma \text{ is unknown,} \\ \mathbf{E}_r - \mathbf{q} \mathbf{q}^T - \mathbf{P}^{-1/2} [\mathbf{a} \mathbf{a}^T + \frac{1}{2} \mathbf{b} \mathbf{b}^T] \mathbf{P}^{-1/2}, & \text{if } \mu \text{ and } \sigma \text{ are unknown,} \end{cases}$$

where \mathbf{E}_r is $r \times r$ unit matrix, $\mathbf{q} = (\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_r})^T$, \mathbf{a} and \mathbf{b} are defined in (1.25) - (1.26). We obtain (see also, Nikulin et al. [90])

$$\mathbf{Y}_n^2 = X_n^T(\hat{\theta}_n) \mathbf{\Sigma}^{-1} X_n(\hat{\theta}_n) \xrightarrow{d} \mathbf{Y}^2 \sim \chi_{r-1}^2.$$

So, the test statistics \mathbf{Y}_n^2 is

$$\mathbf{Y}_n^2 = X_n^T(\hat{\theta}_n) \mathbf{\Sigma}^{-1} X_n(\hat{\theta}_n) = \begin{cases} \mathbf{X}_n^2, & \text{if } \mu \text{ and } \sigma \text{ are known,} \\ \mathbf{X}_n^2 + \frac{1}{n\lambda_1} \alpha^2, & \text{if } \mu \text{ is unknown, } \sigma \text{ is known,} \\ \mathbf{X}_n^2 + \frac{1}{n\lambda_2} \beta^2, & \text{if } \mu \text{ is known, } \sigma \text{ is unknown,} \\ \mathbf{X}_n^2 + \frac{\lambda_2 \alpha^2 - 2\lambda_3 \alpha \beta + \lambda_1 \beta^2}{n(\lambda_1 \lambda_2 - \lambda_3^2)}, & \text{if } \mu \text{ and } \sigma \text{ are unknown,} \end{cases}$$

where

$$\begin{aligned} \mathbf{X}_n^2 &= X_n^T X_n = \frac{1}{n} \sum_{j=1}^r \frac{\nu_j^2}{p_j} - n, \quad \alpha = \sum_{j=1}^r \frac{a_j \nu_j}{p_j} = \mathbf{a}^T \mathbf{P}^{-1} \boldsymbol{\nu}, \\ \beta &= \sum_{j=1}^r \frac{b_j \nu_j}{p_j} = \mathbf{b}^T \mathbf{P}^{-1} \boldsymbol{\nu}, \quad \lambda_1 = 1 - \sum_{j=1}^r \frac{a_j^2}{p_j} = 1 - \mathbf{a}^T \mathbf{P}^{-1} \mathbf{a}, \\ \lambda_2 &= 2 - \sum_{j=1}^r \frac{b_j^2}{p_j} = 2 - \mathbf{b}^T \mathbf{P}^{-1} \mathbf{b}, \quad \lambda_3 = - \sum_{j=1}^r \frac{a_j b_j}{p_j} = -\mathbf{a}^T \mathbf{P}^{-1} \mathbf{b}, \end{aligned}$$

and Σ^- is the generalized inverse matrix of Σ are given as

$$\Sigma^- = \begin{cases} \mathbf{E}_r + \mathbf{q}\mathbf{q}^T, & \text{if } \mu \text{ and } \sigma \text{ are known,} \\ \mathbf{E}_r + \mathbf{q}\mathbf{q}^T + \frac{1}{\lambda_1} \mathbf{a}\mathbf{a}^T, & \text{if } \mu \text{ is unknown, } \sigma \text{ is known,} \\ \mathbf{E}_r + \mathbf{q}\mathbf{q}^T + \frac{1}{\lambda_2} \mathbf{b}\mathbf{b}^T, & \text{if } \mu \text{ is known, } \sigma \text{ is unknown,} \\ \mathbf{E}_r + \mathbf{q}\mathbf{q}^T + \frac{(\lambda_2 \mathbf{a}\mathbf{a}^T - \lambda_3 (\mathbf{a}\mathbf{b}^T + \mathbf{b}\mathbf{a}^T) + \lambda_1 \mathbf{b}\mathbf{b}^T)}{\lambda_1 \lambda_2 - \lambda_3^2}, & \text{if } \mu \text{ and } \sigma \text{ are unknown.} \end{cases}$$

Under the hypothesis H_0 , the statistics \mathbf{Y}_n^2 has an asymptotically chi-squared distribution with $(r - 1)$ degrees of freedom. So, the hypothesis H_0 is rejected with approximate significance level α if $\mathbf{Y}_n^2 > \chi_\alpha^2(r - 1)$.

In the selection of random grouping cells a_j , which depends on the choice of vector $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$ such that $\mathbf{p}^T \mathbf{1}_r = 1$ and all $p_i > 0$. So it is necessary to select p maximizing power the value of test statistic. We, therefore, present follows some remarks on the choice of the intervals for grouping data for testing normality.

Remark 1.2. *If there is no alternative, we can choose the partition J_i as equiprobable intervals (EQPG), that is $p_1 = p_2 = \dots = p_r = \frac{1}{r}$, then $\lambda_3 = 0$ and we have*

$$\mathbf{Y}_n^2 = \frac{r}{n} \sum_{j=1}^r \nu_j^2 - n + \frac{r^2}{n\lambda_1} \left(\sum_{j=1}^r a_j \nu_j \right)^2 + \frac{r^2}{n\lambda_2} \left(\sum_{j=1}^r b_j \nu_j \right)^2.$$

Remark 1.3. *If the vector $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$ is chosen symmetric probability (SPG), so that $p_1 = p_r$, $p_2 = p_{r-1}$, $p_3 = p_{r-2}$ and so on, then $\lambda_3 = 0$. So that, we obtain the following expression of \mathbf{Y}_n^2 statistic*

$$\mathbf{Y}_n^2 = \mathbf{X}_n^2 + \frac{\alpha^2}{n\lambda_1} + \frac{\beta^2}{n\lambda_2}.$$

Furthermore, Ogawa [93] found that this symmetric choice is optimal for testing normality.

Remark 1.4. *If there is alternative, to optimize the choice of p , we may be used the class of Neyman-Pearson (2NPG) as follows.*

Suppose that we take the alternative hypothesis

$$H_1 : \mathbf{P}_\gamma \{T_1 \leq x\} = G(x, \gamma), \gamma \in \mathbf{G} \subset \mathbf{R}^m.$$

Let us $\varphi(x, \theta) = \frac{1}{\sigma} \varphi\left(\frac{x-\mu}{\sigma}\right)$ and $g(x, \gamma)$ are the density probability function of null and alternative hypotheses, respectively. It is necessary to find the intersection points

a_1, a_2, \dots, a_r of two density probability by solving following equation

$$\frac{1}{\hat{\sigma}} \varphi \left(\frac{x - \hat{\mu}}{\hat{\sigma}} \right) = g(x, \hat{\gamma}_n),$$

where $\hat{\theta}_n = (\hat{\mu}, \hat{\sigma}^2)^T$ and $\hat{\gamma}_n$ are the MLE's of parameters $\theta = (\mu, \sigma^2)^T$ and γ , respectively.

We obtain two Neyman-Pearson type classes (2NPG) give as follows

$$\mathbf{I}_1 = \{t : \varphi(t, \hat{\theta}_n) \geq g(t, \hat{\gamma}_n)\}, \quad \mathbf{I}_2 = \{t : \varphi(t, \hat{\theta}_n) < g(t, \hat{\gamma}_n)\}, \quad (1.27)$$

and the frequency vector $\nu = (\nu_1, \nu_2)^T$, the positive probability vector $\mathbf{p} = (p_1, p_2)^T$ are given as

$$\begin{aligned} \nu_j &= n \int_{\mathbf{I}_j} dF_n(t), \quad j = 1, 2, \\ p_1 &:= p_1(\theta) = \int_{\mathbf{I}_1} \varphi(u, \hat{\theta}_n) du, \quad p_2 := p_2(\theta) = \int_{\mathbf{I}_2} \varphi(u, \hat{\theta}_n) du, \quad \theta = (\mu, \sigma)^T, \end{aligned}$$

where $F_n(\cdot)$ is the EDF based on our data. Note that

$$\nu_1 + \nu_2 = n, \quad p_1 + p_2 = 1,$$

and the first derivative of probability vector

$$a_1 = \frac{\partial p_1(\theta)}{\partial \mu}, \quad a_2 = \frac{\partial p_2(\theta)}{\partial \mu}, \quad b_1 = \frac{\partial p_1(\theta)}{\partial \sigma}, \quad b_2 = \frac{\partial p_2(\theta)}{\partial \sigma},$$

satisfy $a_1 + a_2 = b_1 + b_2 = 0$. We construct the statistic \mathbf{Y}_n^2 as defined

$$\begin{aligned} \mathbf{Y}_n^2 &= \mathbf{X}_n^2 + \frac{\lambda_2 \alpha^2 - 2\lambda_3 \alpha \beta + \lambda_1 \beta^2}{n(\lambda_1 \lambda_2 - \lambda_3^2)}, \\ \text{where } \mathbf{X}_n^2 &= \frac{1}{n} \left(\frac{\nu_1^2}{p_1} + \frac{\nu_2^2}{p_2} \right), \quad \lambda_1 = 1 - \frac{a_1^2}{p_1 p_2}, \quad \lambda_2 = 2 - \frac{b_1^2}{p_1 p_2}, \quad \lambda_3 = -\frac{a_1 b_1}{p_1 p_2}, \\ \alpha &= \frac{a_1 \nu_1}{p_1} + \frac{a_2 \nu_2}{p_2}, \quad \beta = \frac{b_1 \nu_1}{p_1} + \frac{b_2 \nu_2}{p_2}. \end{aligned}$$

5.2 Power considerations

We investigate here the power of \mathbf{Y}_n^2 in comparison with other test statistics as Dzhaparidze-Nikulin (\mathbf{D}_n^2), McCulloch (\mathbf{S}_n^2), Shapiro-Wilk (\mathbf{W}), Shapiro-Francia (\mathbf{W}'), Cramér von Mises-Smirnov (ω_n^2) and Anderson-Darling (\mathbf{A}_n^2) statistics for the normal distribution as

null hypothesis against some alternative hypotheses including Logistic, Laplace and Weibull distributions. The \mathbf{Y}_n^2 is constructed based two methods of selection grouping cells, one is the equiprobable intervals (*EQPG*) and one other is the class of Neyman-Pearson (2NPG).

1. For \mathbf{Y}_n^2 statistic using 2NPG class to choose endpoints a_j , from Figure 1.1, we see that there are four points of intersection among Normal and alternative references. Thus, it is reasonable to construct 2NPG class from these points.

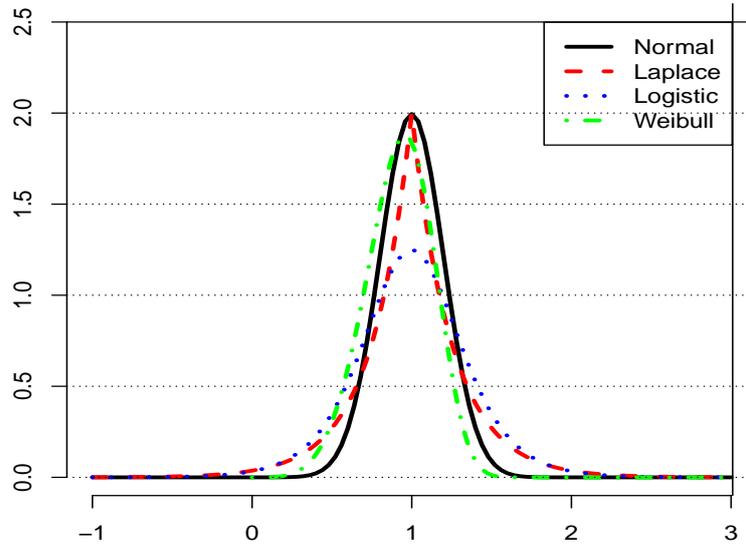


Figure 1.1: Probability density functions of $N(1, 0.2^2)$, $Log(1, 0.2)$, $L(1, 0.25)$ and $W(1, 5)$ distributions.

For example, under Logistic distribution alternative $Log(1, 0.2)$, solving of two density functions, we obtained four point of intersections

$$a_1 = 0.1686993, a_2 = 0.7501097, a_3 = 1.237212, a_4 = 1.8538186,$$

and the following 2NPG class can be considered as (see, Figure 1.1)

$$\mathbf{I}_1 = (0.1686993, 0.7501097] \cup (1.237212, 1.8538186],$$

and $\mathbf{I}_2 = (-\infty, 0.1686993] \cup (0.7501097, 1.237212] \cup [1.8538186, +\infty).$

Under Laplace $L(1, 0.25)$ alternative, we obtain the points of intersection of two den-

sities function

$$a_1 = 0.1761564, \quad a_2 = 0.8260756, \quad a_3 = 1.1723921, \quad a_4 = 1.8313972,$$

so 2NPG class give by (see, Figure 1.1)

$$\mathbf{I}_1 = (0.1761564, 0.8260756] \cup (1.1723921, 1.8313972],$$

and $\mathbf{I}_2 = (-\infty, 0.1761564] \cup (0.8260756, 1.1723921] \cup (1.8313972, +\infty).$

We obtain similarly the 2NPG class for Normal against Weibull $W(1, 5)$ distribution

$$\mathbf{I}_1 = (-\infty, 0.14823] \cup (0.58350, 0.94616] \cup [1.28586, \infty),$$

and $\mathbf{I}_2 = (0.14823, 0.58350] \cup (0.94616, 1.28586).$

2. For the Cramer-von Mises-Smirnov (ω_n^2) and Anderson-Darling (\mathbf{A}_n) statistics, we take the results of Lemeshko et al. [70–72] to compute the power of these tests at $\alpha = 0.05$ significance level.

3. The Shapiro-Wilk [108] statistic is defined as

$$\mathbf{W} = \frac{\left(\sum_{i=1}^n a_i X_{(i)}\right)^2}{\sum_{i=1}^n (X_i - \bar{X}_n)^2}, \quad (1.28)$$

where $\mathbf{T}^{(\cdot)} = (T_{(1)}, T_{(2)}, \dots, T_{(n)})^T$ is the vector of order statistics associated with the simple sample T_1, T_2, \dots, T_n and $\mathbf{a} = (a_1, a_2, \dots, a_n)^T$ are the coefficients having the forms

$$\mathbf{a} = \frac{\mathbf{m}^T \Sigma^{-1}}{(\mathbf{m}^T \Sigma^{-1} \Sigma^{-1} \mathbf{m})^{\frac{1}{2}}},$$

$$\mathbf{m} = (m_1, m_2, \dots, m_n)^T; \quad \Sigma = [\sigma_{ij}]_{n \times n},$$

the expected values of normal order statistics and the covariance matrix of order statistics from the standard normal distribution, which do not depend on unknown parameters μ and σ under the normality hypothesis.

In 1972, Shapiro-Francia [107] proposed a modification of Shapiro - Wilk statistic which

can be used for large sample and defined as

$$\mathbf{W}' = \frac{\left(\sum_{i=1}^n \mathbf{b}_i X_{(i)}\right)^2}{\sum_{i=1}^n (X_i - \bar{X}_n)^2}, \quad (1.29)$$

where, $\mathbf{b} = (b_1, b_2, \dots, b_n)^T$ are the coefficients has the form

$$\mathbf{b} = \frac{\mathbf{m}^T}{(\mathbf{m}^T \mathbf{m})^{\frac{1}{2}}}.$$

with \mathbf{m} is the expected values of normal order statistics $\mathbf{T}^{(\cdot)}$.

The Shapiro - Wilk and Shapiro-Francia statistics are a widely used method for evaluating the condition of normality. However, the distribution of these test statistics are unknown under the null hypothesis and we must be approximated them.

Using the Monte-Carlo method for 50000 simulated samples from a standard normal distribution with sample size n , we obtained its estimated critical values corresponding at two level of significance $\alpha = 5\%$ and 10% . Following Table 1.1, we presented the approximated critical values for each test which is used for reference.

Table 1.1: The approximated critical values of Shapiro - Wilk (\mathbf{W}) and Shapiro-Francia (\mathbf{W}') statistics for testing of normality.

α	Tests	Sample size n									
		40	50	80	100	150	200	250	300	400	500
0.05	\mathbf{W}	0.944	0.953	0.969	0.974	0.982	0.986	0.988	0.990	0.992	0.994
	\mathbf{W}'	0.943	0.953	0.968	0.974	0.982	0.986	0.988	0.990	0.992	0.994
0.10	\mathbf{W}	0.953	0.961	0.973	0.978	0.985	0.988	0.990	0.991	0.993	0.994
	\mathbf{W}'	0.954	0.961	0.974	0.978	0.985	0.988	0.990	0.992	0.993	0.995

Each simulated sample under the alternative hypothesis is repeated $N = 10000$ times for the sample size n . For each operation, we compute the value of each test, then we calculate the power of test that is the difference between N and the number of times where each statistic is less than its critical value divided by N by subtraction, the significance level $\alpha = 0.05$ is used. The result of power tests for each alternative are given in Table 1.2-1.4.

From the practical point of view, the coefficients \mathbf{a} and \mathbf{b} of Shapiro-Wilk and Shapiro-Francia tests are given only for $20 < n \leq 50$ by Shapiro, Wilk and Francia. These are not available for $n > 50$. However, in the case of the sample size is small, these tests then may not have enough power to detect non-normality in the population. Furthermore, the results of the simulations shown that the Shapiro-Wilk and Shapiro-Francia tests get less

Table 1.2: Power of test for normality against Logistic(1, 0.2) distribution.

Tests	Sample size n									
	40	50	80	100	150	200	250	300	400	500
W	0.830	0.802	0.736	0.697	0.591	0.514	0.428	0.354	0.249	0.170
W'	0.791	0.755	0.681	0.632	0.524	0.444	0.366	0.291	0.201	0.131
ω_n^2	0.125	0.145	0.179	0.210	0.280	0.340	0.412	0.478	0.591	0.692
A_n	0.134	0.160	0.200	0.235	0.318	0.390	0.468	0.538	0.657	0.753
D_n²	0.058	0.063	0.064	0.067	0.077	0.088	0.095	0.112	0.129	0.157
S_n²	0.113	0.128	0.213	0.268	0.373	0.462	0.536	0.610	0.692	0.763
Y_n²(EQPG)	0.103	0.124	0.160	0.180	0.244	0.296	0.339	0.390	0.468	0.550
Y_n²(2NPG)	0.083	0.100	0.144	0.174	0.250	0.324	0.393	0.465	0.580	0.678
r	4	5	8	10	15	20	25	30	35	40

Table 1.3: Power of test for normality against Laplace(1, 0.25) distribution.

Tests	Sample size n									
	40	50	80	100	150	200	250	300	400	500
W	0.561	0.486	0.297	0.203	0.07	0.024	0.007	0.001	0	0
W'	0.491	0.416	0.240	0.158	0.055	0.017	0.005	9.10^{-4}	0	0
ω_n^2	0.449	0.536	0.737	0.821	0.941	0.984	0.996	0.998	0.999	1.000
A_n	0.454	0.539	0.738	0.824	0.941	0.983	0.996	0.999	0.999	1.000
D_n²	0.125	0.180	0.299	0.364	0.514	0.642	0.745	0.822	0.929	0.974
S_n²	0.339	0.416	0.604	0.682	0.830	0.899	0.936	0.963	0.985	0.995
Y_n²(EQPG)	0.341	0.423	0.614	0.701	0.845	0.922	0.958	0.979	0.996	0.999
Y_n²(2NPG)	0.390	0.474	0.680	0.781	0.922	0.976	0.992	0.997	0.999	1
r	4	5	8	10	15	20	25	30	35	40

Table 1.4: Power of test for normality against Weibull(1, 5) distribution.

Tests	Sample size n									
	40	50	80	100	150	200	250	300	400	500
W	0.926	0.916	0.887	0.861	0.799	0.740	0.663	0.610	0.476	0.374
W'	0.935	0.929	0.909	0.886	0.841	0.789	0.730	0.680	0.552	0.443
ω_n^2	0.068	0.077	0.102	0.122	0.165	0.213	0.267	0.309	0.411	0.507
A_n	0.071	0.078	0.107	0.131	0.184	0.239	0.298	0.349	0.469	0.571
D_n²	0.048	0.053	0.065	0.074	0.082	0.093	0.107	0.117	0.141	0.161
S_n²	0.066	0.059	0.067	0.069	0.070	0.061	0.063	0.059	0.059	0.063
Y_n²(EQPG)	0.059	0.065	0.067	0.075	0.088	0.097	0.106	0.111	0.145	0.166
Y_n²(2NPG)	0.073	0.083	0.118	0.142	0.191	0.244	0.296	0.337	0.445	0.536
r	4	5	8	10	15	20	25	30	35	40

powerful while the sample size increases. Thus, Shapiro-Wilk and Shapiro-Francia have disadvantages.

As can see Table 1.2 - 1.4, the power of Anderson-Darling (\mathbf{A}_n), \mathbf{Y}_n^2 with 2NPG class

and Cramer-von Mises-Smirnov (ω_n^2) statistics are the best power, in which, the first powers are \mathbf{A}_n and $\mathbf{Y}_n^2(2NPG)$ in all of combination hypothesis considerations. In addition the power of $\mathbf{Y}_n^2(2NPG)$ test statistic shows that this statistic is the best stability among all when we use the 2NPG class as the optimal grouping cells.

5.3 Testing normality for Millikan's measurements

We take in this section the data of Millikan from the book of Linnik [74] to demonstrate the practicability, the performance of the considered above statistics for the test of normality. The Millikan's data are given in Table 1.5 below.

Table 1.5: The values of the charge of the electron, data of Millikan.

4.781	4.764	4.777	4.809	4.761	4.769	4.795	4.776	4.765	4.790	4.792	4.806
4.769	4.771	4.785	4.779	4.758	4.779	4.792	4.789	4.805	4.788	4.764	4.785
4.779	4.772	4.768	4.772	4.810	4.790	4.775	4.789	4.801	4.791	4.799	4.777
4.772	4.764	4.785	4.788	4.799	4.749	4.791	4.774	4.783	4.783	4.797	4.781
4.782	4.778	4.808	4.740	4.790	4.767	4.791	4.771	4.775	4.747		

Voinov et al. [121] suggested that the test does not reject the hypothesis of the normal distribution for Millikan's determinations of the charge of the electron by using the NRR and McCulloch tests. We analyse here the testing of Millikan's data for normality based on the above considered statistics. Under the normal distribution, we obtain the MLE's $\hat{\theta}_n = (4.78081, 0.000233^2)^T$.

Using result of Lemeshko [72] and from the equation (1.28), the value of Shapiro-Wilk, Shapiro-Francia, Cramer-von Mises- Smirnov and Anderson-Darling test statistics are given

$$\mathbf{W} = 0.984, \quad \mathbf{W}' = 0.985, \quad \omega_n^2 = 0.024, \quad \mathbf{A}_n = 0.207,$$

and the p -value at 0.05 significance level of those test statistics are 0.643, 0.634, 0.916, 0.861, respectively. We realize that the null hypothesis that the Millikan's determinations of the charge of the electron is generated by a normal distribution is clearly no rejected. The values of \mathbf{Y}_n^2 and \mathbf{S}_n^2 statistics when the function of the number of equipropable random ($EQPG$) and symmetric probability (SPG) intervals are given in Table 1.6.

Therefore, we concluded a statistical inference that the hypothesis of the normal distribution do not rejected for Millikan's determinations of the charge of the electron in all used considered test statistics.

Table 1.6: Values of \mathbf{Y}_n^2 and \mathbf{S}_n^2 test statistics for Millikan's determinations receiving from the difference of the choice of grouping cells: the class of Neyman-Pearson (2NPG) and the symmetric probability choice (SPG).

r	<i>EQPG</i>		<i>SPG</i>		<i>p</i> -value			
	\mathbf{Y}_n^2	\mathbf{S}_n^2	\mathbf{Y}_n^2	\mathbf{S}_n^2	<i>EQPG</i>		<i>SPG</i>	
	\mathbf{Y}_n^2	\mathbf{S}_n^2	\mathbf{Y}_n^2	\mathbf{S}_n^2	\mathbf{Y}_n^2	\mathbf{S}_n^2	\mathbf{Y}_n^2	\mathbf{S}_n^2
4	2.100	1.757	1.668	1.583	0.551	0.415	0.643	0.452
5	2.454	1.640	2.103	1.701	0.652	0.440	0.716	0.427
6	2.355	0.873	1.859	1.466	0.798	0.646	0.868	0.480
8	6.199	3.203	4.524	1.052	0.516	0.202	0.717	0.590
10	7.872	2.583	4.970	0.762	0.547	0.274	0.836	0.683

6 Generalized chi-squared tests for right censored data

In section 3 we described clearly a generalized chi-squared test statistic for composite hypothesis when the data are complete. In the survival and reliability analysis, however, the data are frequently incomplete observed that usually well-known as censoring times. In these circumstances, the exact survival times of these subjects are unknown. So, it is necessary to consider the test for the censored failure-time data.

In the investigations of the modified \mathbf{Y}_n^2 statistic adapted for the censored failure time data, Habib & Thomas [51], Hollander & Peña [55] considered natural modifications of the \mathbf{Y}_n^2 statistic on the case of censored data without covariates. These tests are based on the differences between two estimators of the probabilities to fall into grouping intervals. The first estimator is based on the estimated KM of the cumulative distribution functions, the estimated second is based on the MLE's of unknown parameters of the fitted model using initial non-grouped censored data.

In 2011, developing an idea of Akritas [4] (also was developed by Hjort [53]) that is to compare the different of observed and expected numbers of failures in time intervals and the choice of random grouping intervals used as data functions, Bagdonavičius et al. [16] were introduced a modified generalized chi-squared test statistic \mathbf{Y}_n^2 that is well adapted for the censored failure time data no covariates. We discuss here this statistic \mathbf{Y}_n^2 as follows the contributions of Bagdonavičius & Nikulin [16], Bagdonavičius et al.[20].

6.1 The right censored failure-time data

Suppose that we observed the right censored failure time data

$$(X_1, \delta_1); (X_2, \delta_2); \dots; (X_n, \delta_n), \quad (1.30)$$

where

$$X_i = T_i \wedge C_i, \quad \delta_i = \mathbf{1}_{\{T_i \leq C_i\}}, \quad (i = 1, 2, \dots, n),$$

where T_1, T_2, \dots, T_n are absolutely continuous i.i.d. random variables, C_1, C_2, \dots, C_n are censoring times. Suppose that T_1, T_2, \dots, T_n and C_1, C_2, \dots, C_n are *mutually independent*. If $\delta_i = 1$ then it means that the failure happens at the moment $T_i = X_i$, on the contrary the failure occurs after the moment X_i . We note here that in classical statistic, we observe the data $(T_1, T_2, \dots, T_n)^T$ and is present of censoring that we have in (1.30).

Assume that the P.D.F. of the random variable T_i belongs to a parametric family

$$\{f(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta} := (\theta_1, \theta_2, \dots, \theta_m)^T \in \Theta \subset \mathbf{R}^m\},$$

where $f(\cdot, \boldsymbol{\theta})$ is the density of T_i with respect to σ -finite measure μ and Θ is open. Denote

$$\begin{aligned} S(t, \boldsymbol{\theta}) &= \mathbf{P}_{\boldsymbol{\theta}}\{T_i > t\}, & \Lambda(t, \boldsymbol{\theta}) &= -\ln S(t, \boldsymbol{\theta}), \\ \lambda(t, \boldsymbol{\theta}) &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbf{P}_{\boldsymbol{\theta}}\{t \leq T_i < t+h \mid T_i \geq t\}, & t > 0, \end{aligned}$$

are the survival, the cumulative hazard and the hazard rate functions of T_i , respectively.

Note that there is another way to describe right censored data (1.30) by using the counting processes . Indeed, for any $t \geq 0$, set

$$\begin{aligned} N_i(t) &= \mathbf{1}_{\{X_i \leq t, \delta_i = 1\}}, & Y_i(t) &= \mathbf{1}_{\{X_i \geq t\}}, \\ N(t) &= \sum_{i=1}^n N_i(t) \quad , & Y(t) &= \sum_{i=1}^n Y_i(t), \end{aligned}$$

$N_i(t)$ is the number of failures of the i^{th} unit in the interval $[0, t]$, $Y_i(t)$ indicate the number "at risk" of i^{th} unit just prior the moment t . The process $N(t)$ therefore shows the number of observed failures in the interval $[0, t]$ ans the process $Y(t)$ shows the number of objects which are "at risk" just prior to time t . The data (1.30) is equivalent to the sample

$$(N_1(t), Y_1(t), t \geq 0), (N_2(t), Y_2(t), t \geq 0), \dots, (N_n(t), Y_n(t), t \geq 0). \quad (1.31)$$

The stochastic processes $N_i(\cdot)$ and $Y_i(\cdot)$ shown dynamics of failures and censoring over time. If the values of

$$\{N_i(s), Y_i(s), 0 \leq s \leq t, i = 1, 2, \dots, n\},$$

are known then history of failures and censoring in the interval $[0, t]$ are known. Thus, the data (1.31) give all history of failures and censoring during the experiment. More about of the stochastic processes, one can see, Borgan [29], Hjort [53], Andersen et al. [6], Bagdonavičius & Nikulin [14].

6.2 Maximum likelihood estimators for right censored data

Let $G_i(\cdot)$ and $g_i(\cdot)$ be the survival and density functions of the censoring time C_i , respectively. We suppose that the right censoring is non-informative which means that the function G_i does not depend on $\boldsymbol{\theta}$. So, The likelihood function $L(\boldsymbol{\theta})$ of the data can be written as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f^{\delta_i}(X_i, \boldsymbol{\theta}) S^{1-\delta_i}(X_i, \boldsymbol{\theta}) g^{1-\delta_i}(C_i) G^{\delta_i}(C_i),$$

Because the G_i and g_i do not contain $\boldsymbol{\theta}$, so they can be rejected. The likelihood function of our data can be written as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f^{\delta_i}(X_i, \boldsymbol{\theta}) S^{1-\delta_i}(X_i, \boldsymbol{\theta}) = \prod_{i=1}^n \lambda^{\delta_i}(X_i, \boldsymbol{\theta}) S(X_i, \boldsymbol{\theta}). \quad (1.32)$$

The log-likelihood function is

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \{\delta_i \ln \lambda(X_i, \boldsymbol{\theta}) + \ln S(X_i, \boldsymbol{\theta})\} = \sum_{i=1}^n \{\delta_i \ln \lambda(X_i, \boldsymbol{\theta}) - \Lambda(X_i, \boldsymbol{\theta})\}, \quad \boldsymbol{\theta} \in \Theta. \quad (1.33)$$

If $\lambda(u, \boldsymbol{\theta})$ is a sufficiently smooth function of the parameter $\boldsymbol{\theta}$ then the MLEs' $\hat{\boldsymbol{\theta}}_n$ satisfies the equation $\dot{\ell}(\hat{\boldsymbol{\theta}}_n) = \mathbf{0}_m$, where $\dot{\ell}(\boldsymbol{\theta})$ is the score vector

$$\dot{\ell}(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}) = \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_1}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_2}, \dots, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_m} \right)^T.$$

The Fisher information matrix is $\mathbf{I}(\boldsymbol{\theta}) = -\mathbf{E}_{\boldsymbol{\theta}} \ddot{\ell}(\boldsymbol{\theta})$, where

$$\ddot{\ell}(\boldsymbol{\theta}) = \sum_{i=1}^n \delta_i \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \lambda(X_i, \boldsymbol{\theta}) - \sum_{i=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \Lambda(X_i, \boldsymbol{\theta}).$$

Under the data (1.31), from the following identifies

$$\int_0^\infty Y_i(u) \lambda(u, \boldsymbol{\theta}) du = \int_0^{X_i} \lambda(u, \boldsymbol{\theta}) du = -\ln S(X_i, \boldsymbol{\theta}),$$

$$\int_0^\infty \ln \lambda(u, \boldsymbol{\theta}) dN_i(u) = \begin{cases} \ln \lambda(X_i, \boldsymbol{\theta}), & \text{if } \delta_i = 1 \\ 0, & \text{if } \delta_i = 0 \end{cases} = \delta_i \ln \lambda(X_i, \boldsymbol{\theta}),$$

then the log-likelihood function (1.33) can be written as

$$\ell(\boldsymbol{\theta}) = \int_0^\infty \{\ln \lambda(u, \boldsymbol{\theta}) dN(u) - Y(u)\lambda(u, \boldsymbol{\theta})\} du. \quad (1.34)$$

The score function in this case is

$$\dot{\ell}(\boldsymbol{\theta}) = \int_0^\infty \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}) \{dN(u) - Y(u)\lambda(u, \boldsymbol{\theta})\} = \int_0^\infty \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}) dM(u, \boldsymbol{\theta}),$$

where

$$M(t, \boldsymbol{\theta}) = N(t) - \int_0^t Y(u)\lambda(u, \boldsymbol{\theta}) du,$$

is the zero mean martingale with respect to the filtration generated by the data.

So, the Fisher information matrix in this case is

$$\mathbf{I}(\boldsymbol{\theta}) = -\mathbf{E}_{\boldsymbol{\theta}} \ddot{\ell}(\boldsymbol{\theta}) = \mathbf{E}_{\boldsymbol{\theta}} \sum_{i=1}^n \int_0^\infty \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}) \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}) \right)^T \lambda(u, \boldsymbol{\theta}) Y_i(u) du.$$

Let us $\hat{\boldsymbol{\theta}}_n$ be the MLE of the parameters $\boldsymbol{\theta}$ which is maximized the likelihood function $L(\boldsymbol{\theta})$ hold. The most results of asymptotic properties of $\hat{\boldsymbol{\theta}}_n$ can be obtained using the stochastic processes $M(t, \boldsymbol{\theta})$. Assume that the following conditions (Borgan [29], Hjort [53]) hold. Suppose that the processed N_i, Y_i are observed during finite time τ .

Condition 1.1.

1) There exists a neighbourhood Θ_0 of $\boldsymbol{\theta}_0$ such that for all n and $\boldsymbol{\theta} \in \Theta_0$, and almost all $t \in [0, \tau]$, the partial derivatives of $\lambda(t, \boldsymbol{\theta})$ of the first, second and third order with respect to $\boldsymbol{\theta}$ exist and are continuous in $\boldsymbol{\theta}$ for $\boldsymbol{\theta} \in \Theta_0$. Moreover, they are bounded in $[0, \tau] \times \Theta_0$ and the log-likelihood function may be differentiated three times with respect to $\boldsymbol{\theta} \in \Theta_0$ by interchanging the order of integration and differentiation.

2) $\lambda(t, \boldsymbol{\theta})$ is bounded away from zero in $[0, \tau] \times \Theta_0$.

3) A positive deterministic function $y(t)$ exists such that $\sup_{t \in [0, \tau]} |Y(t)/n - y(t)| \rightarrow 0$.

4) Under condition 1)-3), the following matrix

$$\mathbf{i}(\boldsymbol{\theta}_0) = \lim_{n \rightarrow \infty} \frac{\mathbf{I}(\boldsymbol{\theta}_0)}{n} = \int_0^\infty \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}_0) \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}_0) \right)^T \lambda(u, \boldsymbol{\theta}_0) Y_i(u) du,$$

is positive definite.

Under the conditions 1.1, we obtain the following the asymptotic properties of $\hat{\boldsymbol{\theta}}_n$

1. $\hat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0$,
2. $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = \frac{1}{\sqrt{n}} \mathbf{i}^{-1}(\boldsymbol{\theta}_0) \dot{\ell}(\boldsymbol{\theta}_0) + O_P(1)$,
3. $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} N_m(\mathbf{0}_m, \mathbf{i}^{-1}(\boldsymbol{\theta}_0))$,
4. $\frac{1}{\sqrt{n}} \dot{\ell}(\boldsymbol{\theta}_0) \xrightarrow{d} N_m(\mathbf{0}_m, \mathbf{i}(\boldsymbol{\theta}_0))$,
5. $\frac{-1}{n} \ddot{\ell}(\boldsymbol{\theta}_0) \xrightarrow{P} \mathbf{i}(\boldsymbol{\theta}_0)$ and $\frac{-1}{n} \ddot{\ell}(\hat{\boldsymbol{\theta}}_n) \xrightarrow{P} \mathbf{i}(\boldsymbol{\theta}_0)$.

6.3 Chi-squared test construction

We consider the problem of testing of the null composite hypothesis H_0

$$H_0 : F(t) \in \mathcal{F}_0 = \{F(t, \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta \subset \mathbf{R}^m, t > 0\}, \quad (1.35)$$

this means that the C.D.F. $F(\cdot)$ belongs to the class \mathcal{F}_0 of C.D.F. of the form $F(t, \boldsymbol{\theta})$, here $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)^T \in \Theta \subset \mathbf{R}^m$ is unknown m -dimensional parameter and F_0 is a known distribution function.

To construct a chi-squared test statistic for the composite hypothesis (1.35), we divide $(0, \tau)$ into k intervals by the points

$$0 = a_0 < a_1 < \dots < a_{k-1} < a_k = \tau, \quad \text{we note } I_j = (a_{j-1}, a_j].$$

Let us

$$U_j = N(a_j) - N(a_{j-1}) = \sum_{i: X_i \in I_j} \delta_i, \quad j = 1, 2, \dots, k,$$

are the number of observed failures in the j -th interval. We consider the equality

$$\mathbf{E}N(t) = \mathbf{E} \int_0^t \lambda(u, \boldsymbol{\theta}_0) Y(u) du,$$

where $\hat{\boldsymbol{\theta}}_n$ is the MLE's of $\boldsymbol{\theta}$, we denote by $\lambda(t, \boldsymbol{\theta})$ is the hazard function.

We consider the following vector \mathbf{Z}

$$\mathbf{Z} = (Z_1, Z_2, \dots, Z_k)^T; \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j), \quad j = 1, 2, \dots, k, \quad (1.36)$$

of the differences between the numbers of observed and "expected" failures

$$e_j = \int_{I_j} \lambda(u, \hat{\boldsymbol{\theta}}_n) Y(u) du, \quad (1.37)$$

in the intervals I_1, I_2, \dots, I_k , which is considered by Akritas [4].

To investigate the properties of the statistic \mathbf{Z} , it is necessary to use the properties of the stochastic properties

$$H_n(t) = \frac{1}{\sqrt{n}} \left(N(t) - \int_0^t \lambda(u, \hat{\boldsymbol{\theta}}_n) Y(u) du \right), \quad t > 0.$$

Using the paper of Hjort [53], we have that the following lemma is hold

Lemma 1.1. *If the condition 1.1 hold, then the stochastic $H_n(t)$ is the weak convergence in the space $D[0, \tau]$ of cad-lag functions with Skorokhod metric, that is*

$$H_n(t) \xrightarrow{d} V, \quad \text{on } D[0, \tau],$$

exists, where V is the zero mean Gaussian martingale such that for all $0 \leq s \leq t$,

$$\text{Cov}(V(s), V(t)) = A(s) - C^T(s) \mathbf{i}^{-1}(\boldsymbol{\theta}_0) C(t),$$

and

$$A(t) = \int_0^t \lambda(u, \boldsymbol{\theta}_0) Y(u) du, \quad C(t) = \int_0^t \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}_0) Y(u) du,$$

where $\boldsymbol{\theta}_0$ is the true value of $\boldsymbol{\theta}$.

For $j, j' = 1, 2, \dots, k$. Set

$$\begin{aligned} V_j &= V(a_j) - V(a_{j-1}), \quad \sigma_{jj'} = \text{Cov}(V_j, V_{j'}), \\ A_j &= A(a_j) - A(a_{j-1}), \quad \mathbf{C}_j = (C_{1j}, C_{2j}, \dots, C_{mj})^T = C(a_j) - C(a_{j-1}), \\ \boldsymbol{\Sigma} &= [\sigma_{jj'}]_{m \times m}, \quad \mathbf{C} = (\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_k)_{m \times k}, \end{aligned}$$

and denote by \mathcal{A} is $k \times k$ diagonal matrix with diagonal elements A_1, A_2, \dots, A_k . We have the next

Theorem 1.6 (Bagdonavičius and Nikulin, 2011). *Under conditions 1.1 then*

$$\mathbf{Z} \xrightarrow{d} Y \simeq N_k(\mathbf{0}_k, \boldsymbol{\Sigma}), \quad \text{as } n \rightarrow \infty. \quad (1.38)$$

where the matrix $\boldsymbol{\Sigma} = \mathcal{A} - C^T \mathbf{i}^{-1}(\boldsymbol{\theta}_0) C$ has the rank equal r .

From the theorem 1.6, we obtain the statistic \mathbf{Y}_n^2 given as follows

$$\mathbf{Y}_n^2 = \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} \xrightarrow{d} \mathbf{Y}^2 \sim \chi_r^2, \quad \text{where } r = \text{rank}(\boldsymbol{\Sigma}), \quad (1.39)$$

where Σ^- is the generalized inverse matrix of Σ . Denote $\mathbf{G} = \mathbf{i} - \mathcal{C}\mathcal{A}^{-1}\mathcal{C}^T$, then one can direct calculate that

$$\Sigma^- = \mathcal{A}^{-1} + \mathcal{A}^{-1}\mathcal{C}^T\mathbf{G}^{-1}\mathcal{C}\mathcal{A}^{-1}. \quad (1.40)$$

As in consequence, we need to find the inverse \mathcal{A}^{-1} of the matrix \mathcal{A} and the general inverse \mathbf{G}^{-1} of the $m \times m$ matrix \mathbf{G} . We take the next theorem

Theorem 1.7 (Bagdonavičius and Nikulin, 2011). *Under condition 1.1, the following estimators of $A_j, \mathbf{C}_j, \mathbf{i}(\theta_0)$ and Σ are consistent*

$$\begin{aligned} \hat{A}_j &= \frac{U_j}{n}, \quad \hat{\mathbf{C}}_j = \frac{1}{n} \int_{I_j} \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \hat{\boldsymbol{\theta}}_n) dN(u), \\ \hat{\mathbf{i}} &= \frac{1}{n} \int_0^\tau \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \hat{\boldsymbol{\theta}}_n) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \hat{\boldsymbol{\theta}}_n) \right)^T dN(u), \quad \text{and} \quad \hat{\Sigma} = \hat{\mathcal{A}} - \hat{\mathcal{C}}^T \hat{\mathbf{i}}^{-1} \hat{\mathcal{C}}. \end{aligned}$$

6.4 Chi-squared test statistic

Theorem 1.6 and 1.7 imply that, for random intervals of grouping cells, a test for the hypothesis H_0 can be based on the statistic

$$\hat{\mathbf{Y}}_n^2 = \mathbf{Z}^T \hat{\Sigma}^- \mathbf{Z}, \quad (1.41)$$

where, $\hat{\Sigma}^-$ that is the general inverse of the covariance matrix $\hat{\Sigma}$, can be written as

$$\hat{\Sigma}^- = \hat{\mathcal{A}}^{-1} + \hat{\mathcal{A}}^{-1} \hat{\mathcal{C}}^T \hat{\mathbf{G}}^{-1} \hat{\mathcal{C}} \hat{\mathcal{A}}^{-1}, \quad \hat{\mathbf{G}} = \hat{\mathbf{i}} - \hat{\mathcal{C}} \hat{\mathcal{A}}^{-1} \hat{\mathcal{C}}^T.$$

The test statistic (1.41) in practically, can be written in the form

$$\hat{\mathbf{Y}}_n^2 = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j} + \hat{\mathbf{Q}}, \quad (1.42)$$

where

$$U_j = \sum_{i: X_i \in I_j} \delta_i, \quad \hat{\mathbf{Q}} = \hat{\mathbf{W}}^T \hat{\mathbf{G}}^{-1} \hat{\mathbf{W}}, \quad \hat{\mathbf{W}} = \hat{\mathcal{C}} \hat{\mathcal{A}}^{-1} \mathbf{Z}, \quad (1.43)$$

$$\hat{\mathbf{G}} = [\hat{\mathbf{g}}_{ll'}]_{m \times m}, \quad \hat{\mathbf{g}}_{ll'} = \hat{i}_{ll'} - \sum_{j=1}^k \hat{C}_{lj} \hat{C}_{l'j} \hat{A}_j^{-1}, \quad (1.44)$$

$$\hat{i}_{ll'} = \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\partial \ln \lambda(X_i, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_l} \frac{\partial \ln \lambda(X_i, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_{l'}}, \quad \hat{C}_{lj} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \frac{\partial \ln \lambda(X_i, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_l}. \quad (1.45)$$

Follows (1.39), under the hypothesis H_0 , the limiting distribution of the statistic $\hat{\mathbf{Y}}_n^2$ is a chi-squared with r degrees of freedom, where

$$r = \text{rank}(\boldsymbol{\Sigma}^-) = \text{Tr}(\boldsymbol{\Sigma}^- \boldsymbol{\Sigma}).$$

So, the hypothesis H_0 rejected with an approximate significance level α if $\hat{\mathbf{Y}}_n^2 > \chi_\alpha^2(r)$.

Remark 1.5 (On the choice of random grouping cells a_j). *As follows Bagdonavičius & Nikulin [16], we describe here the choice of grouping cells a_j as the random data functions. The idea is to divide the interval $[0, \tau]$ into k small intervals with equal expected numbers of failures. Define*

$$E_k = \sum_{i=1}^n \Lambda(X_i, \hat{\boldsymbol{\theta}}_n), \quad E_j = \frac{j}{k} E_k.$$

Let us $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ are the ordered sample from X_1, X_2, \dots, X_n . Set

$$b_i = (n - i) \Lambda(X_{(i)}, \hat{\boldsymbol{\theta}}_n) + \sum_{l=1}^i \Lambda(X_{(l)}, \hat{\boldsymbol{\theta}}_n),$$

if i is the smallest natural number verifying $E_j \in [b_{i-1}, b_i]$ then \hat{a}_j verifying the equality

$$(n - i + 1) \Lambda(\hat{a}_j, \hat{\boldsymbol{\theta}}_n) + \sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\boldsymbol{\theta}}_n) = E_j,$$

$$\text{so } \hat{a}_j = \Lambda^{-1} \left(\frac{E_j - \sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\boldsymbol{\theta}}_n)}{n - i + 1}, \hat{\boldsymbol{\theta}}_n \right), \quad \hat{a}_k = \max \{X_{(n)}, \tau\}, \quad (1.46)$$

where Λ^{-1} is the inverse of the function Λ . With this choice of intervals, we have

$$0 < \hat{a}_1 < \hat{a}_2 < \dots < \hat{a}_k, \quad \text{and } e_j = \frac{E_k}{k}, \quad \text{for all } j.$$

More about the construction of the chi-squared \mathbf{Y}_n^2 statistic and the several applications of this statistic for failure time distributions, one can see, Bagdonavičius et al. [19, 21], Voinov et al. [121].

Note that the approach of construction of a chi-squared statistic based on the difference of observed and expected numbers of failures in time intervals, the grouping cells are taken as the random function, will continue to be developed by us in section 5 of chapter 2 when the data are the right-censored with covariates. We also consider the equations (1.46) in section 3, chapter 3 of this document to estimate the endpoints a_j constructing the chi-squared \mathbf{Y}_n^2 statistic for Hypertabastic distribution when the data are right censored.

Chapter 2

Parametric survival models

1 Introduction

Returning to the survival data that was designed from the response of the time of units functioning under various values covariates. The failure processes of population happen quickly or slowly depending on the effect of observed covariates data. For example, to investigation the brain tumour study, a randomized clinical trial was designed to compare two chemotherapy regimes of the patients with malignant glioma cancer together with 16 covariates are recorded for this study (see, Sauerbrei & Schumacher [105]). The analytically results shown that there are five out of sixteen observed variables had an impact on the survival time of each of patients (see, subsection 2.3, chapter 4). In these circumstances, to construct the good accelerated life models at first we have to evaluate the impact of covariates, after to understand by which way they influence on survival. Follows this, in addition to identifying the most appropriate model for the observed, then a problem posed is to study the effect of covariates on the survival, as well as to consider the estimation of survival under the covariates. In literature development of the survival analysis, the non-parametric and semi-parametric survival models that have been usually utilized as Cox model (see, Cox [34]), Additive Hazards model (see, Aalen [2]), AFT model (see, Bagdonavičius [9]), Hsieh model (see, Hsieh [57]), Additive-multiplicative hazard model (see, Lin and Ying [73]), or the model with the crossing of survival functions which is introduced by Bagdonavičius & Nikulin [14] and more other.

Along with non-parametric, semi-parametric models, many researchers developed parametric survival models such as Billingsley [25], Collett [33], Lee & Wang [69], Kim & Balakrishnan [62], Kleinbaum & Klein [65], Bagdonavičius et al. [20].

Since our purpose is to employ the parametric survival models, to quantify the estimated

survival, to predict the influence of the covariates on survival time, to find a suitable model for the failure times, from which to control or improve the efficiency of the technology and procedures different if covariate information can be used well. Therefore, our focus in this chapter is to discuss of the parametric survival models. In addition to a section of introduction, this chapter include four sections: Section 2, we present a global of accelerated life models given covariates. Section 3, we focus to present the right-censored failure time under covariates which are often obtained in the survival and reliability analysis. Section 4 and 5, we discuss the statistical inference for parametric survival models including the problem of estimated parameters using the MLE's when the data are right censored possibly time-dependence covariates. In this chapter, we also focus to study the generalized chi-squared test procedures for parametric survival models. Our discussions, in this chapter, are refer to the important references of Bagdonavičius et al. ([17, 20–22]), Martinussen & Scheike [75], Nikulin & Tran [88], Nikulin & Wu [89] and other.

2 The consideration models

Let $z(\cdot)$ be a m -dimensional covariate

$$z(\cdot) = (z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot))^T : [0, \infty) \longrightarrow \mathbf{R}^m, \quad z(\cdot) \in \mathbf{E}, \quad (2.1)$$

where $z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot)$ are the univariate time-dependent covariates and \mathbf{E} is a set of all possible m -dimensional covariates. For instance, if one wishes to examine the effect of the chemotherapy together with Bortezomibe treatment versus chemotherapy without Bortezomibe on the survival of multiply myeloma patients studied by Hematology Center, in the Main Military Clinical Hospital, then the patients in the study were randomly assigned to one of two treatment groups. In addition to treatment, several other covariates, for example, type of response, sex, and age of the patients were also observed to analyses the effect of them on the survival times of the multiple myeloma patients. In this study, the covariates that are affected on the survival of multiple myeloma patients, are those two treatment groups, the sex, the age and the type of response.

Let us consider a non-negative random variable $T_{z(\cdot)}$ that is to describe the failure time of an item under the vector of covariate $z(\cdot)$. It can be expected that the survival and hazard rate functions depend on the covariate $z(\cdot)$ as well as on the data history. If the history of one item is described by the covariate $z(\cdot)$, $z(\cdot) \in \mathbf{E}$ and the failure-time $T_{z(\cdot)}$, then the survival function of $T_{z(\cdot)}$ under covariate $z(\cdot)$ is given as

$$S_{z(\cdot)}(t) = \mathbf{P}\{T_{z(\cdot)} \geq t \mid z(u) : 0 \leq u \leq t\}, \quad z(\cdot) \in \mathbf{E}.$$

The hazard rate function of $T_{z(\cdot)}$ under covariate $z(\cdot)$ can be written as

$$\lambda_{z(\cdot)}(t) = \lim_{h \rightarrow 0} \frac{1}{h} \mathbf{P}\{t \leq T_{z(\cdot)} < t + h \mid T_{z(\cdot)} \geq t\} = -\frac{S'_{z(\cdot)}(t)}{S_{z(\cdot)}(t)}, \quad z(\cdot) \in \mathbf{E}.$$

The cumulative hazard and probability density functions given covariate $z(\cdot)$, respectively, are

$$\Lambda_{z(\cdot)}(t) = -\ln \{S_{z(\cdot)}(t)\}, \quad f_{z(\cdot)}(t) = -S'_{z(\cdot)}(t), \quad z(\cdot) \in \mathbf{E}.$$

The failure time $T_{z(\cdot)}$ could be called the resource of the item, but the notion of the resource it should not depend on $z(\cdot) \in \mathbf{E}$.

We say that a covariate z_2 is higher than a covariate z_1 writing $z_2 > z_1$, if for any $t \geq 0$ the inequality $S_{z_1}(t) \geq S_{z_2}(t)$ holds and exists $t_0 > 0$ such that $S_{z_1}(t) > S_{z_2}(t)$. We also say that the covariate z_2 is *accelerated* with respect to the covariate z_1 .

We denote $\lambda_0(\cdot) = \lambda_{z_0(\cdot)}(\cdot)$ as a baseline hazards rate function, where $z_0(\cdot)$ is a usual (so-called as an ideal, normal) covariate. We also use the notation $S_0(\cdot)$ and $\Lambda_0(\cdot)$ instead of the baseline survival $S_{z_0(\cdot)}(\cdot)$ and cumulative hazard $\Lambda_{z_0(\cdot)}(\cdot)$ functions, respectively. Note that usual covariate $z_0(\cdot)$ can be included in \mathbf{E} or not. If $z_0(\cdot)$ doesn't include in \mathbf{E} , then it often describe some ideal or virtual conditions. If $z_0(\cdot)$ include in \mathbf{E} , then $S_0(\cdot)$ include in the class $\{S_{z(\cdot)}(\cdot), z(\cdot) \in \mathbf{E}\}$ also. Then there exists a function denoting the relation between two survival functions $S_{z(\cdot)}(\cdot)$ and $S_0(\cdot)$, that are known as the named of transferable function (so-called the linking functions or the equipercentile equation). More detail on the transferable function, one can see in the publication of Bagdonavičius & Nikulin [14], especially, one can see in thesis of El Fassi. K. [43] at Pierre et Marie Curie-Paris University.

Each specified accelerated life model relates the hazard function (or other function) to the covariate in some particular way. As follows the significant contributions of Bagdonavičius [9, 10], Bagdonavičius & Nikulin [11, 12, 15], Nelson [82], Cox [34], Meeker & Escobar [77], Chen [30], Martinussen & Scheke [75], Nikulin & Wu [89], we briefly consider some well-known survival models which are useful utilized in the survival and reliability data analysis.

2.1 Proportional hazards (PH) model

The proportional hazards model was proposed by Cox [34] in 1972 in order to evaluate the effects of different covariates influencing to the life-time of a system. Since then, this model has been well-known used in survival and reliability analysis.

Definition 2.1. *The Proportional hazards (PH) (or Cox) model holds on a set of covariates \mathbf{E} if for all $z(\cdot) \in \mathbf{E}$ then*

$$\lambda_{z(\cdot)}(t) = r\{z(t)\}\lambda_0(t), \quad z(\cdot) \in \mathbf{E}, \quad (2.2)$$

where $r(\cdot)$ is a positive function on \mathbf{E} and $\lambda_0(\cdot)$ (it could $\lambda_0(\cdot) = \lambda_{z_0(\cdot)}(\cdot)$, where $z_0(\cdot)$ is the standard usual covariate) is an unspecified baseline hazard rate function. The equality (2.2) implies that

$$\Lambda_{z(\cdot)}(t) = \int_0^t r\{z(u)\}d\Lambda_0(u), \quad z(\cdot) \in \mathbf{E}, \quad (2.3)$$

where

$$\Lambda_0(t) = \int_0^t \lambda_0(u)du.$$

The survival function of the PH model is given by

$$S_{z(\cdot)}(t) = \exp\left\{-\int_0^t r\{z(u)\}d\Lambda_0(u)\right\}, \quad z(\cdot) \in \mathbf{E}. \quad (2.4)$$

If $r(\cdot)$ is unknown then we obtain the non-parametric model. If $r(\cdot)$ is parametrized in the form

$$r(z) = e^{\beta^T z}, \quad z(\cdot) \in \mathbf{E}, \quad (2.5)$$

where $\beta = (\beta_1, \beta_2, \dots, \beta_m)^T$ is a vector of finite-dimensional unknown regression parameters. Under the parametrization (2.5), we obtain the semi-parametric PH model on \mathbf{E} with time-dependent covariates

$$\lambda_{z(\cdot)}(t) = e^{\beta^T z(t)}\lambda_0(t), \quad z(\cdot) \in \mathbf{E}. \quad (2.6)$$

If the covariates are constant over time, the hazard rate, the survival and the cumulative functions of PH model given constant covariate $z \in \mathbf{E}_1$, can be written as follows

$$\lambda_z(t) = e^{\beta^T z}\lambda_0(t), \quad S_z(t) = \exp\left\{-e^{\beta^T z}\Lambda_0(t)\right\}, \quad \Lambda_z(t) = e^{\beta^T z}\Lambda_0(t), \quad z \in \mathbf{E}_1,$$

respectively.

Returning to consider the PH model in (2.2), the hazard ratio $HR(t, z_1, z_2)$ of hazard rates given different fixed constant covariates z_1 and z_2 is

$$HR(t, z_1, z_2) = \frac{\lambda_{z_1}(t)}{\lambda_{z_2}(t)} = e^{\beta^T(z_1 - z_2)}, \quad z_1, z_2 \in \mathbf{E}_1.$$

This ratio is constant over time, and more, for any t the conditional probability to fail in a time interval $(t, t + s)$ given that a unit is functioning at the moment t depends only on the value of the covariate $z(\cdot)$ in that interval but does not depend on the values of the covariate until the moment t , that is

$$\mathbf{P}\{T_{z(\cdot)} \leq t + s \mid T_{z(\cdot)} > t\} = 1 - \exp\left\{\int_t^{t+s} e^{\beta^T z(u)} d\Lambda_0(u)\right\}, \quad z(\cdot) \in \mathbf{E},$$

where the baseline hazard rate function $\lambda_0(\cdot)$ does not depend on the covariate. This explains why we can say that the PH model has the absence of memory property and the PH model is not useful utilized analyzing the failure time regression data in reliability.

Although, the PH regression is almost always utilized in survival analysis. However, if the PH assumption does not hold, the PH model may be lead to the mistakes in estimating or evaluating the statistical inference about the effect of a given prognostic factor on lifetime. So that the parametric models that where the assumption of a specified probability distribution for the data is valid, may be brought out some advantages. On the other hand, inferences based on such a parametric proportional hazard assumption will be more precise. Therefore, in some circumstances, the parametric PH model can be used to instead of the classical semi-parametric PH model.

The parametric PH model is obtained when the baseline hazard function $\lambda_0(t)$ is taken from some parametric family of hazard rates

$$\lambda_0(t) \in \lambda_0(t, \gamma), \quad \gamma \in G \subset \mathbf{R}^q.$$

So that, the parametric PH model under the vector of covariate $z(\cdot)$ gives

$$\lambda_{z(\cdot)}(t) = e^{\beta^T z(t)} \lambda_0(t, \gamma), \quad \gamma \in G \subset \mathbf{R}^q. \quad (2.7)$$

The primary difference between of two type of models is that the baseline hazard function is assumed to follow a specific distribution when the parametric PH model is fitted to the data, whereas the PH model does not needed the assumption. The coefficients are estimated by partial likelihood in the PH model but maximum likelihood in the parametric

PH model. In literature, the most commonly used baseline hazard functions are Exponential, Weibull, Gamma and Gompertz distributions. Note that the parametric PH model was rarely used, because the parametric AFT model that will outline in the next section, is also simple for analysis and more natural.

Let us consider below a specific PH model on a set of covariates \mathbf{E} when the baseline hazard rate is the constant

$$\lambda_0(t) \simeq \lambda_0 = \text{const.}$$

If the covariate z is constant over the time, then the PH model gives

$$\lambda_{z(\cdot)}(t) = e^{\beta^T z(t)},$$

it means that the distribution of the failure time T_z is exponential under the constant covariate $z(t) = z$.

On the other hand, if the covariate $z(t)$ depends on the time t as the formula

$$z(t) = \alpha \ln t, \quad \alpha > 0,$$

it follows from model (2.7) that the hazard rate of the lifetime $T_{z(\cdot)}$ is

$$\lambda_{z(\cdot)}(t) = e^{\beta z(t)} \lambda_0(t) = \lambda_0 e^{\alpha \beta \ln t} = \lambda_0 t^\gamma,$$

where $\gamma = \alpha \beta$. It is evident that the failure time $T_{z(\cdot)}$, in this case, follows the Weibull distribution.

As the same way, one obtain the distribution of failure time $T_{z(\cdot)}$ that belongs the Gompertz distribution

$$\lambda_{z(\cdot)}(t) = e^{\beta z(t)} \lambda_0(t) = \lambda_0 e^{\alpha \beta t},$$

if the covariate is formulated as below $z(t) = \alpha t$.

2.2 Accelerated failure time (AFT) model

Definition 2.2. *The accelerated failure time (AFT) model holds on the set of all covariates \mathbf{E} if there exists on \mathbf{E} a positive function $r(\cdot)$ and a survival function S_0 (S_0 does not depend on $z(\cdot)$) such that for any $z(\cdot) \in \mathbf{E}$ the following equality holds on \mathbf{E} .*

$$S_{z(\cdot)}(t) = S_0 \left(\int_0^t r\{z(u)\} du \right), \quad z(\cdot) \in \mathbf{E}, \quad (2.8)$$

if the covariates are constant over time then the AFT model (2.8) is written as

$$S_z(t) = S_0(r(z)t), \quad z \in \mathbf{E}_1, \quad (2.9)$$

where \mathbf{E}_1 is a set of all constant covariates.

The AFT model (2.8) may be considered as non-parametric model (if the function $S_0(t)$ and $r(z)$ are unknown), semi-parametric model (if the function S_0 is unknown, the function $r(z)$ has the parametrization form) or parametric model (if the baseline function $S_0(t)$ belongs to a class of parametric distribution functions). Some key reference, we recommend, for example, Bagdonavičius [9], Lin & Ying [73], Bagdonavičius & Nikulin [13] and many other.

The AFT model (2.8) can be written in terms of hazard rate function as below

$$\lambda_{z(\cdot)}(t) = r\{z(t)\} q\{\Lambda_{z(\cdot)}(t)\}, \quad z(\cdot) \in \mathbf{E}, \quad (2.10)$$

where the function $q(\cdot)$ is given by the following equation

$$H(u) = \int_0^{-\ln u} \frac{dv}{q(v)}, \quad 0 < u < 1, \quad H = S_0^{-1}.$$

As can be seen from (2.10) that different from the PH model, the hazard rate $\lambda_{z(\cdot)}(t)$, under AFT model, at the moment t depends not only on the covariate applied at this moment but also on the covariate applied in the past, that is, in the interval $[0, t)$.

Suppose that we consider the AFT model in (2.8), if $S_0 = S_{z_0(\cdot)}$, where $z_0(\cdot)$ is a usual covariate, then

$$\begin{aligned} S_{z(\cdot)}(t) &= S_{z_0(\cdot)}\left(\int_0^t r\{z(u)\} du\right) \\ &\iff \int_0^t r\{z(u)\} du = S_{z_0(\cdot)}^{-1}\left(S_{z(\cdot)}(t)\right), \quad z(\cdot) \in \mathbf{E}. \end{aligned} \quad (2.11)$$

From (2.8) and (2.11), one can see that in the case of the AFT model on \mathbf{E} the function $r(\cdot)$ changes locally the scale of time to failure distribution.

Let us set

$$f_{z(\cdot)}(t) = \int_0^t r\{z(u)\} du,$$

then the random variable

$$R_{z(\cdot)} = f_{z(\cdot)}(T_{z(t)}) = \int_0^{T_{z(t)}} r\{z(u)\} du = S_{z_0(\cdot)}^{-1}\left(S_{z(\cdot)}(T_{z(t)})\right), \quad z(\cdot) \in \mathbf{E},$$

is called the resource under failure time $T_{z(\cdot)}$ and the survival function of $R_{z(\cdot)}$ is $S_{z_0(\cdot)}$, where

$$S_{z_0(\cdot)}^{-1}(t) = \inf \{u : S_{z_0(\cdot)}(u) \geq t\},$$

and more, the derivative of $f_{z(\cdot)}(t)$ at the moment t

$$\frac{\partial f_{z(\cdot)}(t)}{\partial t} = r(z(t)),$$

at the continuity points of $r\{z(\cdot)\}$. Note that the functional

$$f_{z(\cdot)}(\cdot) : \mathbf{E} \times [0, \infty) \rightarrow [0, \infty),$$

is called the transferable function or linking function (see, El Fassi. K. [43]).

On the other hand, under the constant covariates, from (2.8), it follows that for any $x_1, x_2 \in \mathbf{E}_1$, the survival functions S_{x_1} and S_{x_2} are related in the following way

$$S_{x_2}(t) = S_0(r(x_2)t) = S_{x_1}\left(\frac{r(x_2)}{r(x_1)}t\right) = S_{x_1}(\rho(x_2, x_1)t), \quad x_1, x_2 \in \mathbf{E}_1, \quad (2.12)$$

where $\rho(x_2, x_1) = \frac{r(x_2)}{r(x_1)}$ shows the degree of scale variation. It is evident that $\rho(x, x) = 1$. The equality (2.12) show that the AFT model on \mathbf{E}_1 coincide with the Sedyakin's model (see, Bagdonavičius [10]) on \mathbf{E}_1 .

In practically, the function $r(\cdot)$ is often written in the following parametrization

$$r(z) = e^{-\beta^T \mathbf{z}}, \quad (2.13)$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$ is a vector of unknown regression parameters and

$$\mathbf{z} = (1, z_1, z_2, \dots, z_m)^T = (1, \varphi_1(z), \varphi_2(z), \dots, \varphi_m(z))^T,$$

is a vector of specified functions φ_i of covariates z . In this circumstance, the AFT model is given by the next formula

$$S_{z(\cdot)}(t) = S_0\left(\int_0^t e^{-\beta^T z(u)} du\right), \quad z(\cdot) \in \mathbf{E}, \quad (2.14)$$

and the hazard rate is

$$\lambda_{z(\cdot)}(t) = e^{-\beta^T z(t)} \lambda_0\left(\int_0^t e^{-\beta^T z(u)} du\right), \quad z(\cdot) \in \mathbf{E}. \quad (2.15)$$

For constant covariates then

$$S_z(t) = S_0\left(e^{-\beta^T z t}\right), \quad z \in \mathbf{E}_1. \quad (2.16)$$

We see that the AFT model on \mathbf{E}_1 can also be given as a log-linear model or linear transformation model, since the logarithm of the failure time T_z given z can be written as

$$\ln T_z = \beta^T z + \varepsilon, \quad \varepsilon \sim S(t) = S_0(\ln t), \quad z \in \mathbf{E}_1,$$

it is evident that if the distribution of ε is normal in the case of log-normal failure time distribution than we have the standard linear regression model.

We note here that the choice of the functions $\varphi_i(z)$ is very important in accelerated life testing because the usual covariate used in the experiment, and the bad choice of the model may give bad predict of reliability characteristics given the usual covariate. We cite here some specified choice of φ_i given univariate z .

1. *Log-linear model*: $z_1 = z$, then $e^{-\beta^T z} = e^{-\beta_0 - \beta_1 z}$.
2. *Power rule model*: $z_1 = \ln z$, then $e^{-\beta^T z} = e^{-\beta_0 - \beta_1 \ln z} = \alpha_1 z^{-\beta_1}$.
3. *Arrhenius model*: $z_1 = \frac{1}{z}$, then $e^{-\beta^T z} = e^{-\beta_0 - \frac{\beta_1}{z}} = \alpha_1 e^{-\frac{\beta_1}{z}}$.
4. *Meeker-Luvalle model*: $z_1 = \ln \frac{z}{1-z}$, then $e^{-\beta^T z} = e^{-\beta_0 - \beta_1 \ln \frac{z}{1-z}} = \alpha_1 \left(\frac{z}{1-z}\right)^{-\beta_1}$.

Some times a better choice is taking $m = 2$, then $z = (1, z_1, z_2)^T$. In this case, we often use the Eyring and generalized Eyring models.

5. *Eyring model*: $z_1 = \ln z$, $z_2 = \frac{1}{z}$, then $e^{-\beta^T z} = e^{-\beta_0 - \beta_1 \ln z - \frac{\beta_2}{z}} = \alpha_1 z e^{-\frac{\beta_2}{z}}$.
6. *Generalized Eyring model*: $z_1 = \ln z$, $z_i = \frac{1}{z^i}$, $i = 2, \dots, m$, then

$$e^{-\beta^T z} = \exp \left\{ -\beta_0 - \beta_1 \ln z - \sum_{i=2}^m \frac{\beta_i}{z^i} \right\}.$$

Unlikely to the PH model, the semi-parametric and non-parametric AFT models are rarely used. This is because of the function S_0 , in those cases, are supposed completely unknown. So the modified variants of likelihood functions are not differentiable and even not continuous functions, the limiting of covariance matrices of the regression parameters depend on the derivatives of the probability density functions. This leads to the estimation procedure becomes more complicated than the PH model. On the contrary, a parametric AFT model is mostly utilized in survival, in failure time regression analysis and accelerated life testing in particular. From (2.16), one can see that if the baseline survival function S_0

is taken from Exponential or Weibull distribution functions, that is,

$$S_0(t) = \exp\left\{-\frac{t}{\theta}\right\} \quad \text{or} \quad S_0(t) = \exp\left\{-\left(\frac{t}{\theta}\right)^\nu\right\},$$

respectively, then the AFT models are coincided with PH model on \mathbf{E}_1 .

Nevertheless, it is mention that one disadvantage of parametric AFT model is usually the estimation of these models are carried out by assuming a distribution for the failure time, which in most cases is unknown. As in consequence, Orbe et al. [94] were proposed the methodology which is based on the method of Stute [112] to estimate survival regression models with censored observations without assuming the distribution of the lifetime variable.

The AFT model is the parametric if the baseline survival function S_0 is taken from a class of parametric distributions

$$S_0(t, \gamma), \quad \gamma = (\gamma_1, \gamma_2, \dots, \gamma_q)^T \in G \subset \mathbf{R}^q.$$

So, the parametric AFT model can be written by the following form

$$S_{z(\cdot)}(t) = S_0\left(\int_0^t e^{-\beta^T z(u)} du, \gamma\right), \quad z(\cdot) \in \mathbf{E}, \quad (2.17)$$

if the covariates are constant over time, then

$$S_z(t) = S_0\left(e^{-\beta^T z} t, \gamma\right), \quad z \in \mathbf{E}_1.$$

In literature of parametric AFT model analysis, the most commonly used specified survival functions are Exponential, Weibull, Gamma, Log-normal, Log-logistic, Birnbaum-Saunders, Inverse Gaussian, Inverse Weibull, Generalized Weibull and Inverse Generalized Weibull, v.v. families of distributions.

3 Right censored failure-time data with covariates

Let us consider n items observed and the i^{th} - item is tested under the covariate

$$z_i(\cdot) = (z_{i1}(\cdot), z_{i2}(\cdot), \dots, z_{im}(\cdot))^T \in \mathbf{E}. \quad (2.18)$$

Suppose that we consider the right censored survival data

$$(X_1, \delta_1, z_1(s)), (X_2, \delta_2, z_2(s)), \dots, (X_n, \delta_n, z_n(s)), \quad 0 \leq s \leq \tau. \quad (2.19)$$

where

$$X_i = T_i \wedge C_i, \quad \text{and} \quad \delta_i = \mathbf{1}_{\{T_i \leq C_i\}}, \quad (i = 1, 2, \dots, n), \quad (2.20)$$

T_i is being failure time, C_i is being censored times. The random variable δ_i is the indicator of the event $\{T_i \leq C_i\}$ and τ is the finite time of the experiment.

If the covariates are constant over the times, we obtain the right censored failure time under the constant covariates as follow

$$(X_1, \delta_1, z_1), (X_2, \delta_2, z_2), \dots, (X_n, \delta_n, z_n). \quad (2.21)$$

In the following we describe the right-censored failure time data (2.19) in terms of counting processes keeping the significance contributions of Aalen [1], Andersen et al. [6], Fleming & Harrington [45], Bagdonavičius & Nikulin [14].

Set

$$N_i(t) = \mathbf{1}_{\{t \geq X_i, \delta_i=1\}} = \begin{cases} 1, & \text{if } t \geq X_i \text{ and } X_i = T_i, \\ 0, & \text{if } t < X_i \text{ and } X_i = C_i, \end{cases} \quad (2.22)$$

be the number of failures of the i^{th} item in the interval $[0, t]$. It is equal to 1 if failure is observed in this interval, otherwise it is equal to 0. Also set

$$Y_i(t) = \mathbf{1}_{\{t \leq X_i\}} = \begin{cases} 1, & \text{if } t \leq X_i, \\ 0, & \text{if } t > X_i, \end{cases} \quad (2.23)$$

then $Y_i(t)$ is the at-risk process of the i^{th} item, it's mean that it is not censored and not failed just prior the moment t . And then

$$N(t) = \sum_{i=1}^n N_i(t), \quad Y(t) = \sum_{i=1}^n Y_i(t), \quad t \geq 0, \quad (2.24)$$

are the total number of failures observed in the interval $[0, t]$ and the number of subjects at risk for failure just prior to the moment t , respectively. We also say that the value $Y(t)$, for any t , gives the number of the patients who are at risk for failure during a small time interval $(t - \epsilon, t]$ for an arbitrarily small positive ϵ , since any unit that fails exactly at time t must be both in the risk set at the failure time and known to be at risk before the failure

occurred.

The stochastic processes $N_i(t), Y_i(t)$ is called the counting processes. We say that if (X_i, δ_i) ($i = 1, 2, \dots, n$) are given then the stochastic processes $N_i(t), Y_i(t)$ can be found using their definition. Conversely, suppose that $N_i(t), Y_i(t)$ are defined follow (2.22) and (2.23), if $N_i(t)$ has the jump at X_i then $X_i = T_i$ and $\delta_i = 1$, this means that we obtain an observation $T_i = X_i$. If $N_i(t) = 0$ for any $t \geq 0$ then $X_i = C_i$ and $\delta_i = 0$, it is censoring time. On the other hand, the moment X_i is the moment of the jump of $Y_i(t)$ from 1 to 0. For this reason, the data (2.19) can be written in terms of counting processes $N_i(t), Y_i(t)$ as follows

$$(N_1(t), Y_1(t), t \geq 0), (N_2(t), Y_2(t), t \geq 0), \dots, (N_n(t), Y_n(t), t \geq 0). \quad (2.25)$$

We note that the representation (2.25) give us the dynamics of the stories of failures and censoring up to time t . The notion of the history is formalized by the notion of the filtration. The very concept history is well formalized in terms of the concept filtration of a random process. One very important advantage of data presentation in the form (2.25) is the following: The processes $N_i(\cdot)$ and $Y_i(\cdot)$ show dynamics of failure and censoring mechanism over time. If the values of

$$\{N_i(u), Y_i(u), 0 \leq u \leq t, i = 1, 2, \dots, n\}, \quad (2.26)$$

are known then the history of failures and censoring up to the moment t is known.

We want to cite here some important references that have been used the counting processed for survival analysis, for example, Aven & Jensen [8], Huber et al. [59], Klein et al. [64], Lawless et al.[67], Bagdonavičius et al. [21].

Following, we consider the data described in formulas (2.19) (also (2.21) or (2.25)) for constructing the parameter estimation and the assessing adequacy of fit for the parametric survival models starting from section 4 to the end of chapter 2. We also validate the Hypertabastic survival models to fit the same data in the chapter 3 of this document.

4 Estimates of parameters

4.1 Parametric estimations

Assumption that distributions of all n units that are tested under different explanatory variables or stresses generally are different, are absolutely continuous with the sur-

vival functions $S_i(t, \boldsymbol{\theta})$, the probability densities $f_i(t, \boldsymbol{\theta})$, the hazard rates $\lambda_i(t, \boldsymbol{\theta})$ and the cumulative hazard $\Lambda_i(t, \boldsymbol{\theta})$, specified by a common possibly multidimensional parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_s)^T \in \Theta \subset \mathbf{R}^s$.

Let us consider the right censored failure-time data (2.19) under covariates $z_i(\cdot) \in \mathbf{E}$. Denote by $\bar{G}_i(t)$ and $g_i(t)$ are the survival and the density functions of of the censoring time C_i , respectively. In this case for non-informative and independent censoring we may give the following expressions for the likelihood function

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^n f_i^{\delta_i}(X_i, \boldsymbol{\theta}) S_i^{1-\delta_i}(X_i, \boldsymbol{\theta}) \cdot \bar{G}_i^{\delta_i}(X_i) g_i^{1-\delta_i}(X_i), \quad \boldsymbol{\theta} \in \Theta.$$

Since the problem is to estimate the parameter $\boldsymbol{\theta}$, we can skip the multipliers which do not depend on this parameter. So under non-informative censoring the likelihood function has the next form

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^n f_i^{\delta_i}(X_i, \boldsymbol{\theta}) S_i^{1-\delta_i}(X_i, \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \Theta. \quad (2.27)$$

Using the relation $f_i(t, \boldsymbol{\theta}) = \lambda_i(t, \boldsymbol{\theta}) S_i(t, \boldsymbol{\theta})$ the likelihood function (2.27) can be written

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^n \lambda_i^{\delta_i}(X_i, \boldsymbol{\theta}) S_i(X_i, \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \Theta. \quad (2.28)$$

The log-likelihood function of (2.28) has the form

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \sum_{i=1}^n \{\delta_i \ln \lambda_i(X_i, \boldsymbol{\theta}) + \ln S_i(X_i, \boldsymbol{\theta})\} \\ &= \sum_{i=1}^n \{\delta_i \ln \lambda_i(X_i, \boldsymbol{\theta}) - \Lambda_i(X_i, \boldsymbol{\theta})\}, \quad \boldsymbol{\theta} \in \Theta. \end{aligned} \quad (2.29)$$

The log-likelihood function is maximized at the same point as the likelihood function. As before the estimator $\hat{\boldsymbol{\theta}}_n$, maximizing the likelihood function $L_n(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$, is called the MLE (see, Hjort [54]). Denote by $\hat{\boldsymbol{\theta}}_n = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_s)^T$, are the MLE of $\boldsymbol{\theta}$ under H_0 . If $\lambda_i(u, \boldsymbol{\theta})$ is sufficiently smooth function of the parameter $\boldsymbol{\theta}$ then the MLEs' $\hat{\boldsymbol{\theta}}_n$ satisfies the equation

$$\dot{\ell}(\hat{\boldsymbol{\theta}}_n) = \mathbf{0}_s, \quad (2.30)$$

where $\dot{\ell}(\boldsymbol{\theta})$ is the score vector

$$\dot{\ell}(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}) = (\dot{\ell}_1(\boldsymbol{\theta}), \dot{\ell}_2(\boldsymbol{\theta}), \dots, \dot{\ell}_s(\boldsymbol{\theta}))^T,$$

where the score function $\dot{\ell}_j(\boldsymbol{\theta})$, ($j = 1, 2, \dots, s$) given as

$$\dot{\ell}_j(\boldsymbol{\theta}) = \frac{\partial}{\partial \theta_j} \ell(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \delta_i \frac{\partial}{\partial \theta_j} \ln \lambda_i(X_i, \boldsymbol{\theta}) - \frac{\partial}{\partial \theta_j} \Lambda_i(X_i, \boldsymbol{\theta}) \right\}, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}. \quad (2.31)$$

As before, the Fisher information matrix is $\mathbf{I}(\boldsymbol{\theta}) = -\mathbf{E}_{\boldsymbol{\theta}} \ddot{\ell}(\boldsymbol{\theta})$, where

$$\ddot{\ell}(\boldsymbol{\theta}) = \sum_{i=1}^n \delta_i \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \lambda_i(X_i, \boldsymbol{\theta}) - \sum_{i=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \Lambda_i(X_i, \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}. \quad (2.32)$$

Note here that the data (2.19) can be presented in the form (2.25). So, we can write the log-likelihood $\ell(\boldsymbol{\theta})$ and the score $\dot{\ell}(\boldsymbol{\theta})$ functions in terms of the processes $N_i(\cdot)$ and $Y_i(\cdot)$ as follows. We see that

$$\int_0^{+\infty} \ln \lambda_{z_i(\cdot)}(u, \boldsymbol{\theta}) dN_i(u) = \begin{cases} \ln \lambda_{z_i(\cdot)}(X_i, \boldsymbol{\theta}), & \text{if } \delta_i = 1, \\ 0, & \text{if } \delta_i = 0, \end{cases} = \delta_i \ln \lambda_{z_i(\cdot)}(X_i, \boldsymbol{\theta}),$$

and

$$\begin{aligned} \int_0^{+\infty} Y_i(u) \lambda_{z_i(\cdot)}(u, \boldsymbol{\theta}) du &= \begin{cases} \int_0^t \lambda_{z_i(\cdot)}(u, \boldsymbol{\theta}) du, & \text{if } t \leq X_i, \\ 0, & \text{if } t > X_i, \end{cases} \\ &= \int_0^{X_i} \lambda_{z_i(\cdot)}(u, \boldsymbol{\theta}) du = -\ln S_{z_i(\cdot)}(X_i, \boldsymbol{\theta}). \end{aligned}$$

So, those two expressions (2.29), (2.31) imply that

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \int_0^{+\infty} \{ \ln \lambda_i(u, \boldsymbol{\theta}) dN_i(u) - Y_i(u) \lambda_i(u, \boldsymbol{\theta}) \} du, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}, \quad (2.33)$$

$$\dot{\ell}_j(\boldsymbol{\theta}) = \int_0^{+\infty} \frac{\partial}{\partial \theta_j} \ln \lambda_i(u, \boldsymbol{\theta}) dM_i(u, \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}, \quad (2.34)$$

where

$$M_i(t, \boldsymbol{\theta}) = N_i(t) - \int_0^t Y_i(u) \lambda_i(u, \boldsymbol{\theta}) du,$$

is the zero mean martingale with respect to the filtration generated by the data.

The Fisher's information matrix is given by formula

$$\begin{aligned} \mathbf{I}(\boldsymbol{\theta}) &= -\mathbf{E}_{\boldsymbol{\theta}} \ddot{\ell}(\boldsymbol{\theta}) \\ &= \mathbf{E}_{\boldsymbol{\theta}} \sum_{i=1}^n \int_0^{\infty} \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda_i(u, \boldsymbol{\theta}) \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda_i(u, \boldsymbol{\theta}) \right)^T \lambda_i(u, \boldsymbol{\theta}) Y_i(u) du. \end{aligned} \quad (2.35)$$

4.2 Asymptotic properties of the maximum likelihood estimators

Suppose that the sufficient conditions of Borgan [29] (one can also be found in Bagdonavičius & Nikulin [14]) hold. These conditions are

Condition 2.1.

1. There exists a neighbourhood Θ_0 of the true value $\boldsymbol{\theta}_0$ of $\boldsymbol{\theta}$ such that for all i , $\boldsymbol{\theta} \in \Theta_0$ the derivatives of $\lambda_i(u, \boldsymbol{\theta})$ of the first, the second and the third order with respect to $\boldsymbol{\theta}$ exist and are continuous in $\boldsymbol{\theta}$ for $\boldsymbol{\theta} \in \Theta_0$. Moreover, they are bounded in $[0, \tau] \times \Theta_0$ and the log-likelihood function may be differentiated three times with respect to $\boldsymbol{\theta} \in \Theta_0$ by interchanging the order of integration and differentiation.

2. There exists a positive definite matrix $\mathbf{i}(\boldsymbol{\theta}) = [\sigma_{jj'}(\boldsymbol{\theta})]_{s \times s}$ such that

$$\frac{1}{n} \sum_{i=1}^n \int_0^{\tau} \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda_i(u, \boldsymbol{\theta}_0) \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda_i(u, \boldsymbol{\theta}_0) \right)^T \lambda_i(u, \boldsymbol{\theta}_0) Y_i(u) du \xrightarrow{P} \mathbf{i}(\boldsymbol{\theta}_0), \quad \text{as } n \rightarrow \infty.$$

3. For all $j > 0$ and all $\epsilon > 0$ then

$$\frac{1}{n} \sum_{i=1}^n \int_0^{\tau} \left\{ \frac{\partial}{\partial \theta_j} \ln \lambda_i(u, \boldsymbol{\theta}_0) \right\}^2 \mathbf{1}_{\left\{ \left| n^{-1/2} \frac{\partial}{\partial \theta_j} \ln \lambda_i(u, \boldsymbol{\theta}_0) \right| > \epsilon \right\}} Y_i(u) du \xrightarrow{P} 0.$$

4. For any n and i , there exist measurable functions g_{in} and h_{in} , not dependent on $\boldsymbol{\theta}$, such that for all $t \geq 0$

$$\sup_{\boldsymbol{\theta} \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_{j'} \partial \theta_{j''}} \lambda_i(u, \boldsymbol{\theta}_0) \right| \leq g_{in}(t),$$

and

$$\sup_{\boldsymbol{\theta} \in \Theta_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_{j'} \partial \theta_{j''}} \lambda_i(u, \boldsymbol{\theta}_0) \right| \leq h_{in}(t),$$

for all j, j', j'' . Moreover, for all j, j'

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau g_{in}(t) Y_i(t) dt, \quad \frac{1}{n} \sum_{i=1}^n \int_0^\tau h_{in}(t) \lambda_i(t, \boldsymbol{\theta}_0) Y_i(t) dt,$$

and

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau \left\{ \frac{\partial}{\partial \theta_j} \ln \lambda_i(u, \boldsymbol{\theta}_0) \right\}^2 \lambda_i(t, \boldsymbol{\theta}_0) Y_i(u) du,$$

all converge in probability to finite quantities as $n \rightarrow +\infty$, and, for all $\epsilon > 0$ then

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau h_{in}(t) \mathbf{1}_{\{n^{-1/2} h_{in}(t) > \epsilon\}} Y_i(t) \lambda_i(t, \boldsymbol{\theta}_0) dt \xrightarrow{P} 0.$$

As follows Bogran, we have the next

Theorem 2.3 (Bogran, 1984). *Under independent right censoring and the sufficient conditions 2.1 hold, suppose that $\hat{\boldsymbol{\theta}}_n$ is the MLE of $\boldsymbol{\theta}$ then*

1. $\hat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0$.
2. $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \frac{1}{\sqrt{n}} \mathbf{i}^{-1}(\boldsymbol{\theta}_0) \dot{\ell}(\boldsymbol{\theta}_0) + O_P(1)$.
3. $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{d} N_s(0, \mathbf{i}^{-1}(\boldsymbol{\theta}_0))$.
4. $\frac{1}{\sqrt{n}} \dot{\ell}(\boldsymbol{\theta}_0) \xrightarrow{d} N_s(0, \mathbf{i}(\boldsymbol{\theta}_0))$.
5. $-\frac{1}{n} \ddot{\ell}(\boldsymbol{\theta}_0) \xrightarrow{P} \mathbf{i}(\boldsymbol{\theta}_0)$, and $-\frac{1}{n} \ddot{\ell}(\hat{\boldsymbol{\theta}}_n) \xrightarrow{P} \mathbf{i}(\boldsymbol{\theta}_0)$.

From theorem 2.3, we obtain the Wald statistic for $\boldsymbol{\theta}$ that follows

$$(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})^T \mathbf{i}(\hat{\boldsymbol{\theta}}_n) (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{d} \chi_s^2,$$

where χ_s^2 is the chi-squared distribution with s degrees of freedom.

The asymptotic distribution of $\hat{\boldsymbol{\theta}}_n$ when n as ∞ is approximately normally distributed, that is

$$\hat{\boldsymbol{\theta}}_n \simeq N(\hat{\boldsymbol{\theta}}, n\Sigma^{-1}(\boldsymbol{\theta}_0)), \quad (2.36)$$

in practically, the covariance matrix $n\Sigma(\boldsymbol{\theta})$ can be estimated by $\mathbf{i}(\boldsymbol{\theta})$. Follow (2.36) then the approximated $(1 - \alpha)$ confidence intervals (CI) for the parameter θ_j , ($j = 1, 2, \dots, s$)

$$\theta_j \in \left\{ \hat{\theta}_j - w_{1-\alpha/2} \sigma(\hat{\theta}_j), \hat{\theta}_j + w_{1-\alpha/2} \sigma(\hat{\theta}_j) \right\}, \quad (2.37)$$

where $w_{1-\alpha/2}$ is the $(1 - \frac{\alpha}{2})$ quantile of the normal distribution and

$$\sigma^2(\hat{\theta}_j) = e_j^T \mathbf{i}^{-1}(\hat{\theta}_n) e_j,$$

e_j is the unit vector with the i^{th} -element equal 1, other elements equal 0.

4.3 Approximate confidence estimate of survival function

Suppose that $\hat{\theta}_n$ is a MLE of θ . Suppose that $\{a_n\}$ is a sequence of real numbers. Based on the functional delta method (see, Andersen et al. [6]), we have

$$a_n(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \Sigma^{-1}(\theta)) \quad \text{as } a_n \rightarrow \infty. \quad (2.38)$$

Suppose that $g : \mathbf{R}^s \rightarrow \mathbf{R}$ is a function verifying the conditions of Andersen et al. [6]. This implies that

$$a_n(g(\theta) - g(\hat{\theta}_n)) \xrightarrow{d} N(0, \mathbf{J}_g(\theta) \Sigma^{-1}(\theta) \mathbf{J}_g^T(\theta)) \quad \text{as } a_n \rightarrow \infty, \quad (2.39)$$

where

$$\mathbf{J}_g(\theta) = \left(\frac{\partial g(\theta)}{\partial \theta_1}, \frac{\partial g(\theta)}{\partial \theta_2}, \dots, \frac{\partial g(\theta)}{\partial \theta_s} \right)^T.$$

If n is large then (2.38) implies that

$$\hat{\theta}_n \simeq N_s(\theta, a_n^2 \Sigma^{-1}(\theta)).$$

The convergence (2.39) implies

$$g(\hat{\theta}_n) \simeq N_s(g(\theta), \mathbf{J}_g(\theta) a_n^2 \Sigma^{-1}(\theta) \mathbf{J}_g^T(\theta)).$$

Suppose that $a_n^{-2} \mathbf{i}^{-1}(\hat{\theta}_n)$ is a consistent estimator of $\Sigma^{-1}(\theta)$. Thus, we have

$$\frac{g(\hat{\theta}_n) - g(\theta)}{\sigma_g(\hat{\theta}_n)} \simeq N(0, 1), \quad (2.40)$$

where $\sigma_g^2(\hat{\theta}_n) = \mathbf{J}_g(\hat{\theta}_n) \mathbf{i}^{-1}(\hat{\theta}_n) \mathbf{J}_g^T(\hat{\theta}_n)$.

In the following we construct the CI for survival function of i^{th} -item: $g(\theta) = S_i(t, \theta)$. Because the survival function take values in the interval $(0, 1]$, so the approximation by

the normal law is improved by using the transformations

$$Q_i(t, \boldsymbol{\theta}) = \ln \frac{S_i(t, \boldsymbol{\theta})}{1 - S_i(t, \boldsymbol{\theta})}, \quad \text{for } i = 1, 2, \dots, n.$$

Taking into account that

$$\left(\ln \frac{u}{1-u} \right)' = \frac{1}{u(1-u)},$$

and using (2.38), we obtained

$$\frac{Q_i(t, \hat{\boldsymbol{\theta}}_n) - Q_i(t, \boldsymbol{\theta})}{\sigma_{Q_i}(\hat{\boldsymbol{\theta}}_n)} \simeq N(0, 1),$$

where

$$\sigma_{Q_i}^2(\hat{\boldsymbol{\theta}}_n) = \frac{\mathbf{J}_{S_i}^T(\hat{\boldsymbol{\theta}}_n) \mathbf{i}^{-1}(\hat{\boldsymbol{\theta}}_n) \mathbf{J}_{S_i}(\hat{\boldsymbol{\theta}}_n)}{S_i^2(t, \hat{\boldsymbol{\theta}}_n) [1 - S_i(t, \hat{\boldsymbol{\theta}}_n)]^2},$$

and

$$\mathbf{J}_{S_i}(\hat{\boldsymbol{\theta}}_n) = \left(\frac{\partial S_i(t, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_1}, \frac{\partial S_i(t, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_2}, \dots, \frac{\partial S_i(t, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_s} \right)^T.$$

As in consequence, the approximated $(1 - \alpha)$ confidence intervals (CI) for the survival function $S_i(t, \boldsymbol{\theta})$ is

$$\left\{ \left[1 + \frac{1 - S_i(t, \hat{\boldsymbol{\theta}}_n)}{S_i(t, \hat{\boldsymbol{\theta}}_n)} e^{\sigma_{Q_i}(\hat{\boldsymbol{\theta}}_n) w_{1-\alpha/2}} \right]^{-1}; \left[1 + \frac{1 - S_i(t, \hat{\boldsymbol{\theta}}_n)}{S_i(t, \hat{\boldsymbol{\theta}}_n)} e^{-\sigma_{Q_i}(\hat{\boldsymbol{\theta}}_n) w_{1-\alpha/2}} \right]^{-1} \right\}, \quad (2.41)$$

where $w_{1-\alpha/2}$ is the α quantile of the standard normal law.

5 Generalized chi-squared tests for parametric survival models

Suppose that n independent failure time variables are observed. Let us consider the hypothesis H_0 stating that the survival function given the vector of covariates $z(\cdot) \in \mathbf{E}$

$$H_0 : S(t, \boldsymbol{\theta}) = S_0(t, \boldsymbol{\theta}, z(\cdot)), \quad \boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_s)^T \in \Theta \subset \mathbf{R}^s, \quad z(\cdot) \in \mathbf{E}, \quad (2.42)$$

where $S_0(\cdot)$ is a specified function of time t . The hypothesis H_0 in (2.42) can be also formulated in terms of the hazard function $\lambda(t, \boldsymbol{\theta}) = -\frac{S'(t, \boldsymbol{\theta})}{S(t, \boldsymbol{\theta})}$ or the cumulative hazard function $\Lambda(t, \boldsymbol{\theta}) = -\ln \{S(t, \boldsymbol{\theta})\}$.

About the hypotheses (2.42), we can construct a test for the parametric AFT model (2.17); parametric PH model (2.7); the parametric generalized proportional hazards models; or the class of models with cross effects of survival functions (for detail see, Bagdonavičius et al. [21]).

Additional information about the baseline function $S_0(\cdot)$, we have to cite here the families of the Weibull, Log-logistic, Log-normal, Generalized Weibull, Exponentiated Weibull, Birnbaum-Saunders, Inverse Gaussian, Generalized Inverse Weibull distributions. In research of Bagdonavičius et al. [20], Tahir [116], Saaidia [104], Goual [48], they have been discussed on them.

We shall present in detail a test for parametric survival models include parametric AFT and PH models using Hypertabastic distribution as the baseline hazard functions, in chapter 3 of this document. Therefore, we consider the foremost the generalized chi-squared test for general hypothesis H_0 with possibly time dependent covariates following Bagdonavičius et al. [21, 22], Nikulin & Tran [88]. Using the same approach like the section 6 of chapter 1, the choice of random grouping intervals as data functions is also considered.

5.1 Idea of test construction

We shall give chi-squared test for the hypothesis H_0 from right censored failure time regression data that are described in (2.21) or in counting processes (2.25), respectively.

Let us $\hat{\boldsymbol{\theta}}_n$ be the MLE's of $\boldsymbol{\theta}$ under H_0 . The estimated parameters $\hat{\boldsymbol{\theta}}_n$ satisfied the score equations (2.30) assuming that the consistency and asymptotic normality of the $\hat{\boldsymbol{\theta}}_n$ (theorem 2.3) hold under H_0 .

Suppose that the following condition hold

Condition 2.2.

1. The processes N_i, Y_i, z_i are observed up to the finite time of the experiment τ ;
2. Survival distributions of all n objects given z_i are absolutely continuous with the survival functions $S_i(t, \boldsymbol{\theta})$ with the hazard rates $\lambda_i(t, \boldsymbol{\theta})$ and the cumulative hazard rates $\Lambda_i(t, \boldsymbol{\theta})$ of X_i under z_i ;
3. Censoring is non-informative and independent, supposing that the multiplicative intensity model holds. The compensators of the counting processes N_i with respect to the history of the observed processes are $\int_0^t Y_i(u) \lambda_i(u, \boldsymbol{\theta}) du$.

As the same the approach of the construction of the test statistic for right censored failure time data which we discussed in the first chapter, here, we divide the interval $[0, \tau]$

into k smaller intervals $I_j = (a_{j-1}, a_j]$, $a_0 = 0$, $a_k = \tau$, and denote by

$$U_j = N(a_j) - N(a_{j-1}), \quad (2.43)$$

is the number of observed failures in the j -th interval, $j = 1, 2, \dots, k$.

Under regularity conditions, suppose that the equality

$$\mathbf{E}N_i(t) = \mathbf{E} \int_0^t \lambda_i(u, \boldsymbol{\theta}) Y_i(u) du, \quad \boldsymbol{\theta} \in \Theta, \quad (2.44)$$

holds. It suggests that we can expect to observe

$$e_j = \sum_{i=1}^n \int_{a_{j-1}}^{a_j} \lambda_i(u, \hat{\boldsymbol{\theta}}_n) Y_i(u) du, \quad \boldsymbol{\theta} \in \Theta, \quad (2.45)$$

failures in the interval I_j .

Let us consider the stochastic process

$$H_n(t) = \frac{1}{\sqrt{n}} \left(N(t) - \sum_{i=1}^n \int_0^t \lambda_i(u, \hat{\boldsymbol{\theta}}_n) Y_i(u) du \right), \quad (2.46)$$

which characterizes the difference between observed and expected numbers of failures. So we obtain the vector $\mathbf{Z} = (Z_1, Z_2, \dots, Z_k)^T$ of the difference between observed and expected numbers of failures, where

$$Z_j = H_n(a_j) - H_n(a_{j-1}) = \frac{1}{\sqrt{n}} (U_j - e_j), \quad (j = 1, 2, \dots, k). \quad (2.47)$$

It is very reasonable that we can construct the test for H_0 based on the vector \mathbf{Z} .

5.2 Asymptotically properties of test statistic

It is evident that the properties of the statistic \mathbf{Z} depends the properties of the stochastic process $H_n(t)$, $t \geq 0$. To investigate the properties of $H_n(t)$, we need to use the asymptotic properties of the MLE in section 4. Set

$$\begin{aligned} S^{(0)}(t, \boldsymbol{\theta}) &= \sum_{i=1}^n Y_i(t) \lambda_i(t, \boldsymbol{\theta}), & S^{(1)}(t, \boldsymbol{\theta}) &= \sum_{i=1}^n Y_i(t) \frac{\partial \ln \lambda_i(t, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \lambda_i(t, \boldsymbol{\theta}), \\ S^{(2)}(t, \boldsymbol{\theta}) &= \sum_{i=1}^n Y_i(t) \frac{\partial^2 \ln \lambda_i(t, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \lambda_i(t, \boldsymbol{\theta}). \end{aligned}$$

We also suppose that the Conditions VI.1.1 given in Andersen et al. [6] and the following condition hold.

Condition 2.3. *There exist a neighborhood Θ_0 of θ_0 and continuous bounded on $\Theta_0 \times [0, \tau]$, the functions*

$$s^{(0)}(t, \boldsymbol{\theta}), \quad s^{(1)}(t, \boldsymbol{\theta}) = \frac{\partial s^{(0)}(t, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \quad s^{(2)}(t, \boldsymbol{\theta}) = \frac{\partial^2 s^{(0)}(t, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2},$$

such that for $j = 0, 1, 2$

$$\sup_{t \in [0, \tau], \boldsymbol{\theta} \in \Theta_0} \left\| \frac{1}{n} S^{(j)}(t, \boldsymbol{\theta}) - s^{(j)}(t, \boldsymbol{\theta}) \right\| \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty.$$

The conditions 2.3 imply that uniformly for $t \in [0, \tau]$

$$\frac{1}{n} \sum_{i=1}^n \int_0^t \lambda_i(u, \boldsymbol{\theta}_0) Y_i(u) du \xrightarrow{P} A(t), \quad \frac{1}{n} \sum_{i=1}^n \int_0^t \frac{\partial}{\partial \boldsymbol{\theta}} \lambda_i(u, \boldsymbol{\theta}_0) Y_i(u) du \xrightarrow{P} C(t),$$

where $A(\cdot)$ and $C(\cdot) = (C_1(\cdot), C_2(\cdot), \dots, C_s(\cdot))^T$ are finite functions. As in consequence, we have the next follow

Lemma 2.1 (Bagdonavičius et al., 2011). *If the conditions 2.1, 2.3 and Conditions VI.1.1 (see, Andersen et al. [6]) are hold, then the following convergence exists*

$$H_n \xrightarrow{d} V \quad \text{on } D[0, \tau],$$

here $D[0, \tau]$ is space of cadlag functions with Skorokhod metric, V is zero mean Gaussian martingale such that, for all $0 \leq u \leq v \leq \tau$

$$\text{cov}(V(u), V(v)) = A(u) - C^T(u) \mathbf{i}^{-1}(\boldsymbol{\theta}_0) C(v).$$

For $l = 1, 2, \dots, s$, $j, j' = 1, 2, \dots, k$, set

$$\begin{aligned} V_j &= V(a_j) - V(a_{j-1}), \quad \sigma_{jj'} = \text{cov}(V_j, V_{j'}), \quad \Sigma = [\sigma_{jj'}]_{k \times k}, \\ A_j &= A(a_j) - A(a_{j-1}), \quad C_{lj} = C_l(a_j) - C_l(a_{j-1}), \quad \mathcal{C} = [C_{ij}]_{s \times k}, \end{aligned}$$

and we denote by \mathcal{A} a $k \times k$ diagonal matrix with the diagonal elements A_1, \dots, A_k . From the lemme 2.1 one can obtain the theorem below

Theorem 2.4 (Bagdonavičius et al., 2011). *Under conditions 2.1, 2.3 and condition VI.1.1 Andersen et al. [6] then*

$$\mathbf{Z} \xrightarrow{d} Y \sim N_k(\mathbf{0}_s, \Sigma) \quad \text{as } n \rightarrow \infty, \quad (2.48)$$

where

$$\Sigma = \mathcal{A} - \mathcal{C}^T \hat{\mathbf{i}}^{-1}(\boldsymbol{\theta}_0) \mathcal{C}, \quad \text{rank}(\Sigma) = r. \quad (2.49)$$

As follows Theorem 2.4, we obtain a statistic

$$\mathbf{Y}_n^2 := \mathbf{Z}^T \Sigma^{-} \mathbf{Z} \xrightarrow{d} \chi_r^2, \quad (2.50)$$

where χ_r^2 is a chi-squared distribution with r degrees of freedom.

5.3 Test statistic

The formula (2.50) shows that a test for the hypothesis H_0 can be based on the statistic

$$\mathbf{Y}_n^2 = \mathbf{Z}^T \Sigma^{-} \mathbf{Z}, \quad (2.51)$$

where, the matrix Σ^{-} that is the general inverse of the covariance matrix Σ . And then, we have

$$\lim_{n \rightarrow \infty} \mathbf{P}\{\mathbf{Y}_n^2 < x \mid H_0\} = \mathbf{P}\{\chi_r^2 < x\}, \quad \text{for any } x > 0.$$

Note that if we set $\mathbf{G} = \mathbf{i} - \mathcal{C} \mathcal{A}^{-1} \mathcal{C}^T$. The formula

$$\Sigma^{-} = \mathcal{A}^{-1} + \mathcal{A}^{-1} \mathcal{C}^T \mathbf{G}^{-1} \mathcal{C} \mathcal{A}^{-1}, \quad (2.52)$$

implies that we need to invert only diagonal $k \times k$ matrix \mathcal{A} and find the general inverse of the $s \times s$ matrix \mathbf{G} .

From Theorem 2.4, it follows that under conditions 2.1 and 2.3, the following estimators of A_j , \mathcal{C}_j , Σ and $\mathbf{i}(\boldsymbol{\theta}_0)$ are consistent

$$\hat{A}_j = \frac{U_j}{n}, \quad \hat{\mathcal{C}}_j = \frac{1}{n} \sum_{i=1}^n \int_{I_j} \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda_i(u, \hat{\boldsymbol{\theta}}_n) dN_i(u), \quad \hat{\Sigma} = \hat{\mathcal{A}} - \hat{\mathcal{C}}^T \hat{\mathbf{i}}^{-1} \hat{\mathcal{C}}, \quad (2.53)$$

$$\hat{\mathbf{i}} = \frac{1}{n} \sum_{i=1}^n \int_0^\tau \frac{\partial \ln \lambda_i(u, \hat{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} \left(\frac{\partial \ln \lambda_i(u, \hat{\boldsymbol{\theta}}_n)}{\partial \boldsymbol{\theta}} \right)^T dN_i(u). \quad (2.54)$$

The statistic \mathbf{Y}_n^2 in fact has a convenient form which is the sum of the Pearson statistic

\mathbf{Y}_n^2 and a non-negative quadratic form \mathbf{Q} , and it can be presented as

$$\mathbf{Y}_n^2 := \mathbf{Z}^T \hat{\Sigma}^- \mathbf{Z} = \mathbf{X}_n^2 + \mathbf{Q}, \quad (2.55)$$

where $\hat{\Sigma}^-$ is the general inverse of the estimate covariance matrix $\hat{\Sigma}$ and

$$\mathbf{X}_n^2 = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j}, \quad \mathbf{Q} = W^T \mathbf{G}^- W, \quad \mathbf{G} = \mathbf{i} - \hat{\mathcal{C}} \hat{\mathbf{A}}^{-1} \hat{\mathcal{C}}^T, \quad (2.56)$$

$$W = (W_1, \dots, W_s)^T, \quad W_l = \sum_{j=1}^k \hat{C}_{lj} \hat{A}_j^{-1} Z_j, \quad (l = 1, 2, \dots, s), \quad (2.57)$$

and the elements of the matrix $\hat{\mathcal{C}} = [\hat{C}_{lj}]_{s \times k}$ and Fisher information matrix $\hat{\mathbf{i}} = [\hat{i}_{ll'}]_{s \times s}$ estimated by the formulas below

$$\hat{C}_{lj} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \frac{\partial \ln \lambda_i(X_i, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_l}, \quad (2.58)$$

$$\hat{i}_{ll'} = \frac{1}{n} \sum_{i=1}^n \delta_i \frac{\partial \ln \lambda_i(X_i, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_l} \frac{\partial \ln \lambda_i(X_i, \hat{\boldsymbol{\theta}}_n)}{\partial \theta_{l'}}. \quad (2.59)$$

The distribution of the statistic \mathbf{Y}_n^2 has in limit a chi-squared distribution with

$$r = \text{rank}(\Sigma^-) = \text{Tr}(\Sigma^- \Sigma),$$

degrees of freedom. If the matrix \mathbf{G} is non-degenerate then $r = k$.

Statistic inference: The hypothesis H_0 is rejected with approximate significance level α if $\mathbf{Y}_n^2 > \chi_\alpha^2(r)$.

5.4 On the choice of random cells boundaries

An usual experiment plan in accelerated life testing is to test several groups of units under different higher stress conditions. In such experiment it is possible that the failures of units from different groups are mostly concentrated in different non-intersecting intervals. So using common idea of constructing chi-squared test by division of the interval $[0, \tau)$ into smaller intervals and comparing observed and expected numbers of failures the choice of the ends of the intervals is very important because dividing into intervals of equal length may give intervals where the numbers of observed failures are zero or very small.

Let us consider the choice of the limits of grouping intervals as random data functions.

Define

$$E_k = \sum_{i=1}^n \int_0^{\tau} \lambda_i(u, \hat{\boldsymbol{\theta}}_n) Y_i(u) du = \sum_{i=1}^n \Lambda_i(X_i, \hat{\boldsymbol{\theta}}_n), \quad E_j = \frac{j}{k} E_k, \quad (j = 1, 2, \dots, k). \quad (2.60)$$

So we seek \hat{a}_j to have equal numbers of expected failures (not necessary an integer) in all intervals. So \hat{a}_j satisfy the equality

$$g(\hat{a}_j) = E_j, \quad g(a) = \sum_{i=1}^n \int_0^a \lambda_i(t, \hat{\boldsymbol{\theta}}_n) Y_i(t) dt.$$

Denote by $X_{(1)} \leq \dots \leq X_{(n)}$ the ordered sample from X_1, \dots, X_n . Note that the function

$$g(a) = \sum_{i=1}^n \Lambda_i(X_i \wedge a, \hat{\boldsymbol{\theta}}_n) = \sum_{i=1}^n \left[\sum_{l=i}^n \Lambda_{(l)}(a, \hat{\boldsymbol{\theta}}_n) + \sum_{l=1}^{i-1} \Lambda_{(l)}(X_{(l)}, \hat{\boldsymbol{\theta}}_n) \right] \mathbf{1}_{[X_{(i-1)}, X_{(i)}]}(a),$$

is continuous and increasing on $[0, \tau]$; here $X_{(0)} = 0$ and we denote by $\sum_{l=1}^0 c_l = 0$. Set

$$b_i = \sum_{l=i+1}^n \Lambda_{(l)}(X_{(i)}, \hat{\boldsymbol{\theta}}_n) + \sum_{l=1}^i \Lambda_{(l)}(X_{(l)}, \hat{\boldsymbol{\theta}}_n). \quad (2.61)$$

If $b_{i-1} \leq E_j \leq b_i$ then \hat{a}_j is the unique solution of the equation

$$\sum_{l=i}^n \Lambda_{(l)}(\hat{a}_j, \hat{\boldsymbol{\theta}}_n) + \sum_{l=1}^{i-1} \Lambda_{(l)}(X_{(l)}, \hat{\boldsymbol{\theta}}_n) = E_j. \quad (2.62)$$

For this choice, we have $e_j = E_k/k$ for any j and

$$0 < \hat{a}_1 < \hat{a}_2 < \dots < \hat{a}_k = \tau.$$

So the hypothesis H_0 is rejected with approximate significance level α if $\mathbf{Y}_n^2 > \chi_\alpha^2(r)$, the statistic \mathbf{Y}_n^2 computed using the formulas from (2.55) to (2.59), replacing a_j by \hat{a}_j in all formulas and taking

$$e_1 = \dots = e_k = \frac{1}{k} \sum_{i=1}^n \Lambda(X_i, \hat{\boldsymbol{\theta}}_n, z_i). \quad (2.63)$$

More about the construction of \mathbf{Y}_n^2 test statistic and its applications for the parametric survival models, one can see, Bagdonavičius et al [21, 22], Nikulin & Tran [88] and many other. Using the \mathbf{Y}_n^2 test statistic, in the next chapter, we shall validate the Hypertabastic survival models.

Chapter 3

Validity of Hypertabastic models

1 Introduction

On the application of generalized chi-squared test statistic \mathbf{Y}_n^2 , we look at some significant references as Aguirre & Nikulin [3], Bagdonavičius et al. [19, 21, 23], Zhang [126], Nikulin et al. [92], Tahir [116] and Saaidia [104] in their Bordeaux Ph.D thesis. There are several failure-time distributions, such as, Exponential, Weibull, Log-normal, Log-logistic, Birnbaum-Saunders, Inverse Gaussian, Generalized Weibull, Generalized Exponentiated Weibull distributions have been validated using the chi-squared test and applying to the real data. We interest a new probability distribution that has the nice properties with uni-modal shape of hazard function, was recently proposed by Tabatabai et al. [113], it is called Hypertabastic distribution. In addition to being used as the baseline function in the parametric models for applying to the clinical trial data as the myeloma cancer (Tabatabai et al. [113]) or the breast cancer (see, Tabatabai et al. [114, 115]), Tabatabai and his colleagues demonstrated the performance of this distribution in comparison with Weibull, Log-normal, Log-logistic and Gamma distributions based on the log-likelihood ratio and AIC statistics. But those two statistics have not useful to use for assessing adequacy of fit for the survival and reliability data analysis.

As follows our methodology and generalized chi-squared test that have been described in the first and second chapters. We wish give here to explicit the formula of our consideration tests for Hypertabastic survival models in three scenes of the data: Complete, right censored and right censored with covariates. We take, firstly, the Hypertabastic distribution as follow the text of Tabatabai and his colleagues [113] in section 2. Later, we focus to discuss the properties of the two parametric models included Hypertabastic AFT and Hypertabastic PH models. We also consider the testing of those two models using \mathbf{Y}_n^2

statistic. The simulation studies were carried out in section 4 to illustrate the power of \mathbf{Y}_n^2 test in comparison with other statistics. Using Monte Carlo procedure, the more precise models of statistic distribution of considered tests are also analyzed for this distribution when unknowns parameters are estimated by the MLE.

2 Hypertabastic distribution $H(a, b)$

We say that the random variable T follows Hypertabastic distribution noting $H(a, b)$ if the C.D.F. of this distribution is given as

$$F(t, \theta) = 1 - \text{Sech}\left\{\frac{a}{b}\left(1 - t^b \text{Coth}(t^b)\right)\right\}, \quad t > 0, \theta = (a, b)^T, a > 0, b > 0, \quad (3.1)$$

where, $\text{Sech}(\cdot)$ and $\text{Coth}(\cdot)$ are hyperbolic secant and hyperbolic cotangent, respectively. Its survival function is

$$S(t, \theta) = \text{Sech}\left\{\frac{a}{b}\left(1 - t^b \text{Coth}(t^b)\right)\right\}, \quad t > 0, a > 0, b > 0. \quad (3.2)$$

P.D.F. and hazard functions of $H(a, b)$ distribution are given the expressions below

$$f(t, \theta) = at^{b-1}\{t^b \text{Csch}^2(t^b) - \text{Coth}(t^b)\} \text{Sech}(W) \text{Tanh}(W), \quad (3.3)$$

and

$$\lambda(t, \theta) = at^{b-1}\{t^b \text{Csch}^2(t^b) - \text{Coth}(t^b)\} \text{Tanh}(W), \quad (3.4)$$

respectively, where

$$W = \frac{a}{b}\left(1 - t^b \text{Coth}(t^b)\right).$$

The consideration of the different hazard shapes bring out the different biological mechanisms of disease progression which can be assist clinicians, researches and pharmacologists to keep track of the disease status over time. The Hypertabastic hazard function has interesting properties (for detail, see Tabatabai et al. [113]) as follows.

- If $0 < b \leq 0.25$, then its hazard rate is decreasing from ∞ to 0.
- If $0.25 < b \leq 1$, then its hazard rate is \cap -shaped: The first increases with time until it reaches its maximum and then decreases.
- If $1 < b \leq 2$, then its hazard rate is increasing with upward concavity until it reaches its inflection point and then it continues to increase with downward concavity thereafter.
- If $b > 2$, then its hazard rate is increasing with upward concavity.

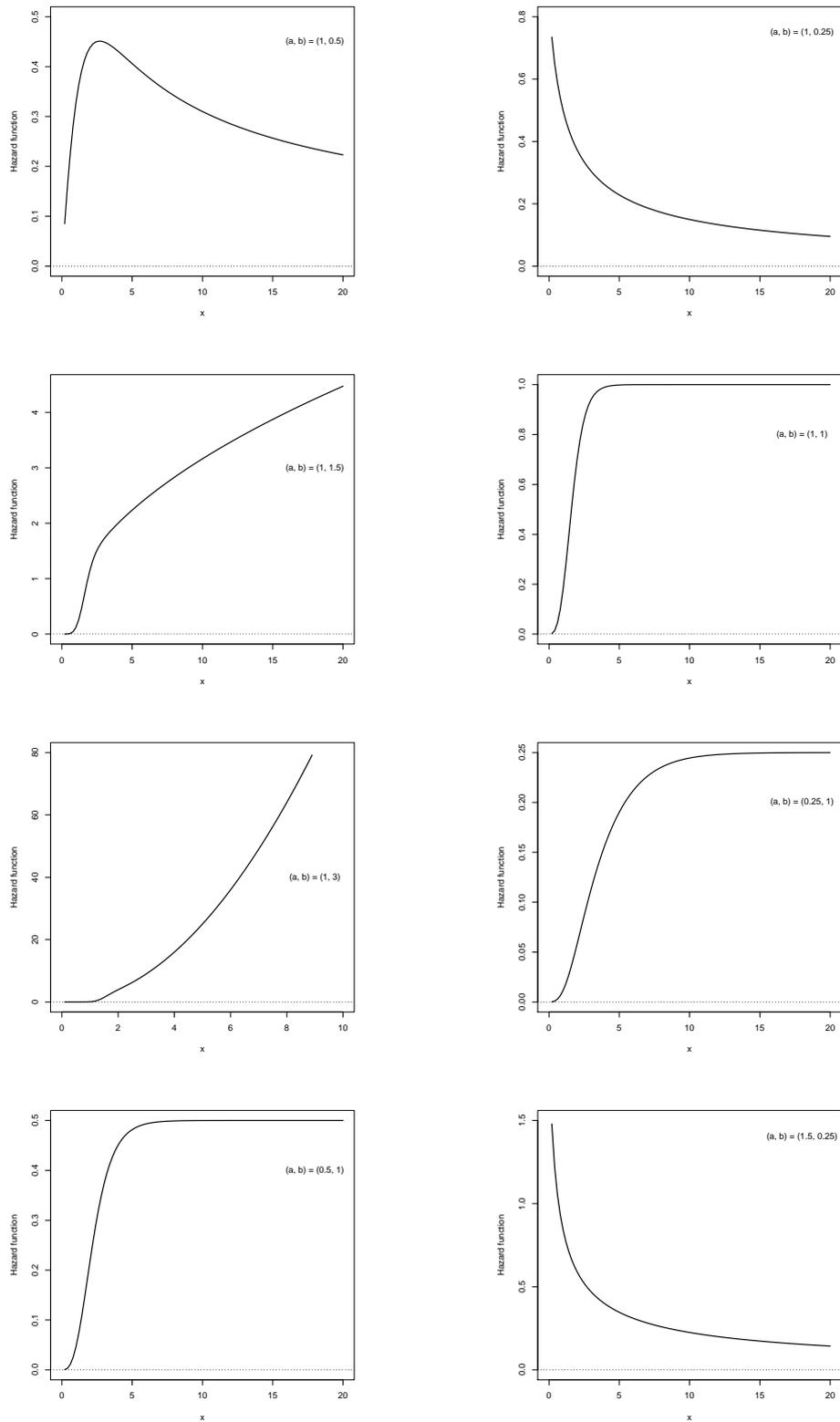


Figure 3.1: Hypertabastic hazard plot with difference of parameter (a, b) .

In addition to being used a Hypertabastic distribution as the baseline function of parametric survival models, this distribution can be alternated for other as Exponential, Weibull, Gamma, Log-normal, Log-logistic, Inverse Gaussian and Birnbaum - Saunders distributions in survival and reliability analysis. In the next chapter, our statistical results will show that Hypertabastic survival models be a more accurate than other models for the real-life dataset based on \mathbf{Y}_n^2 statistic. However, if Hypertabastic survival model is rejected for the data, we need to select other model to fit best. So, it is really necessary to consider in the next the goodness-of-fit test for Hypertabastic survival models as the null hypotheses.

3 Validity of $H(a, b)$ using generalized chi-squared tests

We have the foremost to test the composite hypothesis H_0 , according to which

$$H_0 : F(t, \theta) = 1 - \text{Sech}\left\{\frac{a}{b}\left(1 - t^b \text{Coth}(t^b)\right)\right\}, \quad t > 0, a > 0, b > 0, \quad (3.5)$$

when the data can be the complete or right censored observations.

3.1 Complete data

Let $\mathbf{T} = (T_1, T_2, \dots, T_n)^T$ be a complete sample, i.e. T_1, T_2, \dots, T_n are i.i.d. random variables. We consider to test for the composite hypothesis H_0 in (3.5) from the sample \mathbf{T} .

The log-likelihood function of \mathbf{T} is

$$\ell(\theta) = n \ln a + (b - 1) \sum_{i=1}^n \ln T_i + \sum_{i=1}^n \ln \{U_i \text{Tanh}(W_i)\} + \sum_{i=1}^n \ln \{\text{Sech}(W_i)\}, \quad (3.6)$$

and the score functions are

$$\begin{aligned} \dot{\ell}_a(\theta) &= \frac{n}{a} - \frac{1}{a} \sum_{i=1}^n \frac{W_i(2 \text{Tanh}^2(W_i) - 1)}{\text{Tanh}(W_i)}, \\ \dot{\ell}_b(\theta) &= \sum_{i=1}^n \ln T_i + \frac{2b}{a} \sum_{i=1}^n \frac{N_i}{U_i} + \frac{1}{b} \sum_{i=1}^n \frac{K_i(2 \text{Tanh}^2(W_i) - 1)}{\text{Tanh}(W_i)}, \end{aligned}$$

where

$$\begin{aligned} W_i &= \frac{a}{b} \left(1 - T_i^b \text{Coth}(T_i^b)\right), \quad U_i = T_i^b \text{Csch}^2(T_i^b) - \text{Coth}(T_i^b), \\ K_i &= W_i - a T_i^b U_i \ln T_i, \quad \text{and} \quad N_i = T_i^b W_i \text{Csch}^2(T_i^b) \ln T_i, \quad (i = 1, 2, \dots, n). \end{aligned}$$

The MLE's $\hat{\theta}_n = (\hat{a}_n, \hat{b}_n)^T$ of $\theta = (a, b)^T$ can be obtained by solving the non-linear system of scores functions $(\dot{\ell}_a(\theta), \dot{\ell}_b(\theta))^T = \mathbf{0}_2$.

Using the formulas that are described in section 3 of chapter 1, we present below the mechanism of \mathbf{Y}_n^2 test statistic for the hypothesis (3.5) when the data are complete and the MLE's are usage.

Let us $\mathbf{p} = (p_1, p_2, \dots, p_r)^T$ is a vector of positive probability, $0 < p_i < 1$, such that $\mathbf{p}^T \mathbf{1}_r = 1$ and defining the endpoints a_j of the grouping intervals that are solution of the equation

$$1 - \text{Sech}\left\{\frac{\hat{a}_n}{\hat{b}_n}\left(1 - a_j^{\hat{b}_n} \text{Coth}(a_j^{\hat{b}_n})\right)\right\} = \sum_{i=1}^j p_i, \quad (j = 1, 2, \dots, r-1), \quad a_0 = 0, a_r = +\infty. \quad (3.7)$$

Suppose that $\boldsymbol{\nu} = (\nu_1, \nu_2, \dots, \nu_r)^T$ is a frequency vector arising from grouping T_1, \dots, T_n over the partition $I_j = (a_{j-1}, a_j]$ and the probability $\mathbf{p}(\theta) = (p_1(\theta), p_2(\theta), \dots, p_r(\theta))^T$ to fall into each cell is

$$p_j(\theta) = \int_{a_{j-1}}^{a_j} f(x, \theta) dx = F(a_j, \theta) - F(a_{j-1}, \theta), \quad (j = 1, 2, \dots, r).$$

The Fisher's information matrix $\mathbf{I}(\theta)$ for one observation T_i can be replaced by

$$\mathbf{I}(\hat{\theta}_n) = -\frac{1}{n} \begin{pmatrix} \frac{\partial^2 \ell(\hat{\theta}_n)}{\partial a^2} & \frac{\partial^2 \ell(\hat{\theta}_n)}{\partial ab} \\ \frac{\partial^2 \ell(\hat{\theta}_n)}{\partial ba} & \frac{\partial^2 \ell(\hat{\theta}_n)}{\partial b^2} \end{pmatrix},$$

where

$$\begin{aligned} \frac{\partial^2 \ell(\hat{\theta}_n)}{\partial a^2} &= -\frac{1}{\hat{a}_n^2} \left\{ n + \sum_{i=1}^n \frac{W_i^2 \text{Sech}^2(W_i)(1 + 2 \text{Tanh}^2(W_i))}{\text{Tanh}^2(W_i)} \right\}, \\ \frac{\partial^2 \ell(\hat{\theta}_n)}{\partial ba} &= \frac{1}{\hat{a}_n \hat{b}_n} \sum_{i=1}^n \frac{K_i}{\text{Tanh}(W_i)} \left\{ 2 \text{Tanh}^2(W_i) - 1 + \frac{W_i \text{Sech}^2(W_i)(2 \text{Tanh}^2(W_i) + 1)}{\text{Tanh}(W_i)} \right\}, \\ \frac{\partial^2 \ell(\hat{\theta}_n)}{\partial b^2} &= \frac{2}{\hat{a}_n} \sum_{i=1}^n \frac{N_i \ln T_i}{K_i} - \frac{2}{\hat{b}_n^2} \sum_{i=1}^n \frac{K_i(2 \text{Tanh}^2(W_i) - 1)}{\text{Tanh}(W_i)} + \frac{2 \hat{b}_n}{\hat{a}_n} \sum_{i=1}^n \frac{A_i \ln T_i}{U_i^2} - \\ &\frac{1}{\hat{b}_n} \sum_{i=1}^n \frac{T_i^{\hat{b}_n} (\hat{a}_n U_i + 2 \hat{b}_n N_i)(2 \text{Tanh}^2(W_i) - 1) \ln^2 T_i}{\text{Tanh}(W_i)} - \frac{1}{\hat{b}_n^2} \sum_{i=1}^n \frac{K_i^2 \text{Sech}^2(W_i)(2 \text{Tanh}^2(W_i) + 1)}{\text{Tanh}^2(W_i)}, \end{aligned}$$

where

$$W_i = \frac{\hat{a}_n}{\hat{b}_n} \left(1 - T_i^{\hat{b}_n} \text{Coth}(T_i^{\hat{b}_n})\right), \quad U_i = T_i^{\hat{b}_n} \text{Csch}^2(T_i^{\hat{b}_n}) - \text{Coth}(T_i^{\hat{b}_n}),$$

$$K_i = W_i - \hat{a}_n T_i^{\hat{b}_n} U_i \ln T_i, \quad N_i = T_i^{\hat{b}_n} W_i \operatorname{Csch}^2(T_i^{\hat{b}_n}),$$

$$A_i = T_i^{\hat{b}_n} U_i \operatorname{Csch}^2(T_i^{\hat{b}_n}) \left[W_i \ln T_i - \frac{K_i}{\hat{b}_n} - 2T_i^{\hat{b}_n} W_i \operatorname{Coth}(T_i^{\hat{b}_n}) \ln T_i \right] - \frac{2\hat{b}_n}{\hat{a}_n} N_i^2 \ln T_i.$$

Let $\mathbf{J}(\hat{\theta}_n) = \mathbf{B}^T(\hat{\theta}_n)\mathbf{B}(\hat{\theta}_n)$ is the Fisher's information matrix of the vector frequencies $\boldsymbol{\nu}$, where the matrix $\mathbf{B} = [b_{ij}]_{r \times 2}$ include the elements

$$b_{i1} = \frac{1}{\sqrt{p_i(\hat{\theta}_n)}} \frac{\partial p_i(\hat{\theta}_n)}{\partial a}, \quad b_{i2} = \frac{1}{\sqrt{p_i(\hat{\theta}_n)}} \frac{\partial p_i(\hat{\theta}_n)}{\partial b}, \quad (i = 1, 2, \dots, r).$$

From (1.18) - (1.21) in chapter 1, a chi-squared test statistic for H_0 can be written as

$$\mathbf{Y}_n^2(\hat{\theta}_n) = \mathbf{X}_n^2(\hat{\theta}_n) + \mathbf{Q}, \quad (3.8)$$

where $\mathbf{X}_n^2(\hat{\theta}_n) = \mathbf{X}_n^T(\hat{\theta}_n)\mathbf{X}_n(\hat{\theta}_n)$, $\mathbf{X}_n(\hat{\theta}_n)$ is the Pearson's vector

$$\mathbf{X}_n(\hat{\theta}_n) = \left(\frac{\nu_1 - np_1(\hat{\theta}_n)}{\sqrt{np_1(\hat{\theta}_n)}}, \frac{\nu_2 - np_2(\hat{\theta}_n)}{\sqrt{np_2(\hat{\theta}_n)}}, \dots, \frac{\nu_r - np_r(\hat{\theta}_n)}{\sqrt{np_r(\hat{\theta}_n)}} \right)^T,$$

and the quadratic form \mathbf{Q} give as the following expression

$$\mathbf{Q} = \mathbf{X}_n(\hat{\theta}_n)\mathbf{B}(\hat{\theta}_n)(\mathbf{I}(\hat{\theta}_n) - \mathbf{J}(\hat{\theta}_n))^{-1}\mathbf{B}^T(\hat{\theta}_n)\mathbf{X}_n(\hat{\theta}_n).$$

Following the theorem 1.5 in chapter 1, the hypothesis H_0 is rejected with approximate significance level α if $\mathbf{Y}_n^2(\hat{\theta}_n) > \chi_\alpha^2(r-1)$, where $\chi_\alpha^2(r-1)$ is the upper α percentage points of the chi-squared distribution with $r-1$ degrees of freedom.

3.2 Right censored data

Let us consider the composite hypothesis H_0 in (3.5) from the right censored failure time data that are given in (1.30).

In this situation, the log-likelihood function can be written as

$$\ell(\theta) = \sum_{i=1}^n \{ \delta_i [\ln a + (b-1) \ln X_i + \ln(U_i \operatorname{Tanh}(W_i))] + \ln \operatorname{Sech}(W_i) \}, \quad (3.9)$$

where

$$W_i = \frac{a}{b} \left(1 - X_i^b \operatorname{Coth}(X_i^b) \right), \quad U_i = X_i^b \operatorname{Csch}^2(X_i^b) - \operatorname{Coth}(X_i^b).$$

The MLE's $\hat{\theta}_n = (\hat{a}, \hat{b})^T$ of $\theta = (a, b)^T$ can be obtained by solving the non-linear system of scores functions: $(\dot{\ell}_a(\theta), \dot{\ell}_b(\theta))^T = \mathbf{0}_2$. Set

$$\hat{W}_i = \frac{\hat{a}}{\hat{b}} \left(1 - X_i^{\hat{b}} \text{Coth}(X_i^{\hat{b}})\right), \quad \hat{U}_i = X_i^{\hat{b}} \text{Csch}^2(X_i^{\hat{b}}) - \text{Coth}(X_i^{\hat{b}}),$$

we obtain the expression of the element of the matrix $\mathcal{C} = [\hat{C}_{lj}]_{2 \times k}$ and the Fisher information matrix $\hat{\mathbf{I}} = [\hat{i}_{ll}]_{2 \times 2}$ by using the formula (1.45), are given as follows

$$\begin{aligned} \hat{i}_{11} &= \frac{1}{\hat{a}^2} \sum_{i=1}^n \delta_i \left(1 + \frac{\hat{W}_i \hat{U}_i \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)}\right)^2, \\ \hat{i}_{22} &= \frac{1}{\hat{b}^2} \sum_{i=1}^n \delta_i \left(\hat{b} \ln X_i + \frac{2\hat{b}^2 X_i^{\hat{b}} \hat{W}_i \text{Csch}^2(X_i^{\hat{b}}) \ln X_i}{\hat{a} \hat{U}_i} - \frac{(\hat{W}_i - \hat{a} X_i^{\hat{b}} \hat{U}_i \ln X_i) \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)}\right)^2, \\ \hat{i}_{12} &= \frac{1}{\hat{a}\hat{b}} \sum_{i=1}^n \delta_i \left(1 + \frac{\hat{W}_i \hat{U}_i \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)}\right) \times \\ &\quad \left(\hat{b} \ln X_i + \frac{2\hat{b}^2 X_i^{\hat{b}} \hat{W}_i \text{Csch}^2(X_i^{\hat{b}}) \ln X_i}{\hat{a} \hat{U}_i} - \frac{(\hat{W}_i - \hat{a} X_i^{\hat{b}} \hat{U}_i \ln X_i) \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)}\right), \\ \hat{C}_{1j} &= \frac{1}{\hat{a}} \sum_{i: X_i \in I_j} \delta_i \left(1 + \frac{\hat{W}_i \hat{U}_i \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)}\right)^2, \\ \hat{C}_{2j} &= \frac{1}{\hat{b}} \sum_{i: X_i \in I_j} \delta_i \left(\hat{b} \ln X_i + \frac{2\hat{b}^2 X_i^{\hat{b}} \hat{W}_i \text{Csch}^2(X_i^{\hat{b}}) \ln X_i}{\hat{a} \hat{U}_i} - \frac{(\hat{W}_i - \hat{a} X_i^{\hat{b}} \hat{U}_i \ln X_i) \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)}\right). \end{aligned}$$

Choice \hat{a}_j : Let us $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ are the ordered sample from X_1, X_2, \dots, X_n . Set

$$\Lambda(X_{(i)}, \hat{\theta}_n) = -\ln \left\{ \text{Sech} \left\{ \frac{\hat{a}}{\hat{b}} \left(1 - X_{(i)}^{\hat{b}} \text{Coth}(X_{(i)}^{\hat{b}})\right) \right\} \right\}, \quad E_j = \frac{j}{k} \sum_{i=1}^n \Lambda(X_{(i)}, \hat{\theta}_n), \quad (3.10)$$

and

$$b_i = (n - i) \Lambda(X_{(i)}, \hat{\theta}_n) + \sum_{l=1}^i \Lambda(X_{(l)}, \hat{\theta}_n), \quad i = 1, 2, \dots, n. \quad (3.11)$$

If i is the smallest natural number verifying $E_j \in [b_{i-1}; b_i]$ then \hat{a}_j verifying the equality

$$(n - i + 1) \Lambda(\hat{a}_j, \hat{\theta}_n) + \sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\theta}_n) = E_j. \quad (3.12)$$

Because the inverse of the cumulative hazard function of $H(a, b)$ distribution can not be written in an explicit form. So, the equations in (3.12) are solved based on the numerical method.

4 Some other GOF tests for $H(a, b)$ distribution

We consider, in this section, some test statistics including the modified "omega-squared" and Anderson-Darling statistics for $H(a, b)$ distribution (3.5) for the complete data. Let $T_{(1)}, T_{(2)}, \dots, T_{(n)}$ are the ordered sample from T_1, T_2, \dots, T_n .

"Omega-squared" statistic. This statistic based on the integral of squared difference between of C.D.F. and E.C.D.F. and is given as

$$\mathbf{W}_n^2 = n \int_0^1 [F_n(u) - u]^2 du, \quad \text{where} \quad F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{T_i \leq t\}},$$

\mathbf{W}_n^2 is so-called Cramer-von Mises-Smirnov statistic. This statistic in fact has the more convenient following formula

$$\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^n \left[F(T_{(i)}, \theta) - \frac{2i-1}{2n} \right]^2. \quad (3.13)$$

Anderson-Darling statistic. The test based on Anderson-Darling statistic can be written as

$$\mathbf{A}_n = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left[\ln F(T_{(i)}, \theta) + \ln (1 - F(T_{(n-i+1)}, \theta)) \right]. \quad (3.14)$$

ω_n^2 and \mathbf{A}_n tests: The hypothesis H_0 is rejected if ω_n^2, \mathbf{A}_n exceed their critical values.

Remark 3.1 (Remark on the limit distribution of ω_n^2 and \mathbf{A}_n statistics).

1. If θ is known, the critical values of ω_n^2 and \mathbf{A}_n computed by statistical software when the sample n is small. If n is large then the critical values are found using following asymptotic results

$$n\omega_n^2 \xrightarrow{d} C = \int_0^1 B^2(t) dt, \quad n\mathbf{A}_n \xrightarrow{d} A = \int_0^1 \frac{B^2(t)}{t(1-t)} dt.$$

For detail the C.D.F. of the random variables C and A , one can see Bolshev [28].

2. If θ is unknown, we use the MLE of Hypertabastic's parameters $\hat{\theta}_n = (\hat{a}, \hat{b})^T$ instead of $\theta = (a, b)^T$ in the expressions (3.13) and (3.14). In this case, the distributions of ω_n^2 and \mathbf{A}_n statistics depend on the estimated parameters $\hat{\theta}_n = (\hat{a}, \hat{b})^T$. However, under $H(a, b)$ hypothesis (3.5), using the same approach of Lemeshko et al. [72], we suggest that the limiting distributions of ω_n^2, \mathbf{A}_n statistics are best approximated by the Johnson - SB

family of distributions (J_{SB}) with the cumulative function

$$J_{SB}(x, \theta) = \Phi\left\{\theta_0 + \theta_1 \ln\left(\frac{x - \theta_3}{\theta_2 + \theta_3 - x}\right)\right\}, \quad \theta = (\theta_0, \theta_1, \theta_2, \theta_3)^T,$$

where $\Phi(x)$ is the distribution function of the standard normal law. In fact, by simulating for the sample size $n = 200$ with $N = 10000$ of repeat on the difference combination Hypertabastic's parameters $(a, b) = (1, 0.2), (1, 0.5), (1, 1.5), (1.5, 0.75), (1, 1)$ and $(2, 3)$, then the limiting distributions of ω_n^2 , \mathbf{A}_n statistics when MLE are used, are given as follows

$$\begin{aligned} \omega_n^2 &\rightarrow J_{SB}(4.00256, 1.51431, 0.63650, 0.00652), \quad \text{as } n \rightarrow \infty, \\ \mathbf{A}_n &\rightarrow J_{SB}(4.55893, 1.69143, 4.41847, 0.05498), \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Figure 3.2 shows the curve of estimated distribution of ω_n^2 statistic using maximum likelihood method to estimate two unknown Hypertabastic's parameters. In this Figure, a black line shows the curve of the estimated C.D.F. of J_{SB} family with the estimated parameters $\hat{\theta} = (4.00256, 1.51431, 0.63650, 0.00652)^T$.

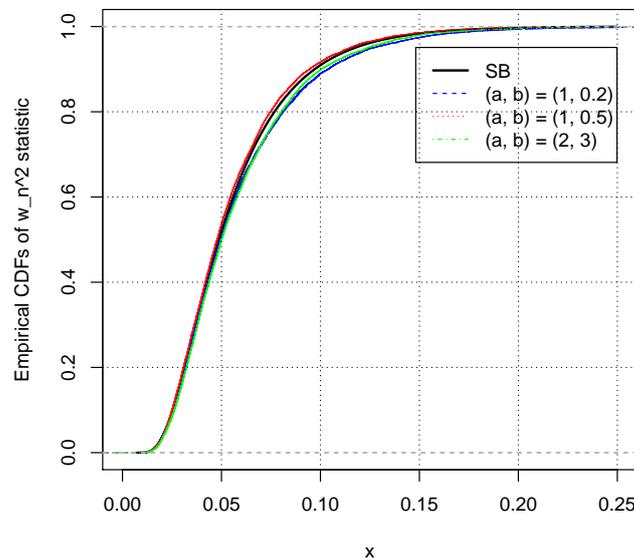


Figure 3.2: Simulated Cramer-von Mises-Smirnov statistic distribution for testing of Hypertabastic and estimate of JB distribution when MLE's used.

Similarly, Figure 3.3 shows the curve of estimated distribution of \mathbf{A}_n statistic for calculating difference MLE's under H_0 , a black line indicate the curve of the estimated C.D.F.

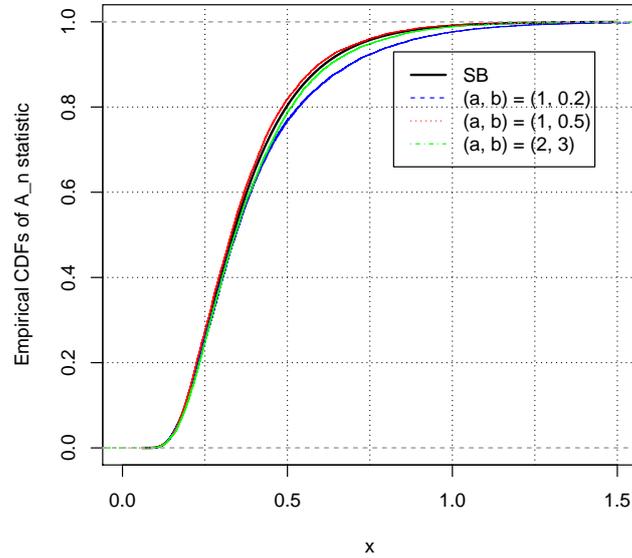


Figure 3.3: Simulated Anderson-Darling statistic distribution for testing of Hypertabastic and estimate of JB distribution when MLE's used.

of J_{SB} family with estimated parameters $\hat{\theta} = (4.55893, 1.69143, 4.41847, 0.05498)^T$.

Table 3.1: Upper α percentage points of ω_n^2 and \mathbf{A}_n statistical distributions for $H(a, b)$ in the case of MLE usage.

Test statistics	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
Cramer-von Mises-Smirnov's ω_n^2	0.1646562	0.1173248	0.0970515
Anderson-Darling's \mathbf{A}_n	0.9865201	0.724399	0.611299

Table 3.1 shows the upper α percentage points of statistic distributions of the ω_n^2 and \mathbf{A}_n tests that were constructed by modeled pseudo-random $H(a, b)$ sample size $n = 200$ for $N = 10000$ of repeat, when two unknown parameters estimated by their MLE's.

5 Validity of Hypertabastic AFT model

5.1 Hypertabastic AFT model

Suppose that

$$z(\cdot) = (z_0(\cdot), z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot))^T \in \mathbf{E}, z_0(\cdot) \equiv 1, \quad (3.15)$$

is the possible time-varying m -dimensional vector of covariate.

The Hypertabastic accelerated failure time (AFT_H) model under $z(\cdot) \in \mathbf{E}$ can be written in the form

$$S_{z(\cdot)}(t) = G_0 \left\{ \left(\int_0^t e^{-\beta^T z(u)} du \right)^b \right\}, \quad z(\cdot) \in \mathbf{E}, \quad (3.16)$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$ is the unknown regression parameter and $G_0(\cdot)$ is the baseline survival function, given as

$$G_0(u) = \text{Sech} \left\{ \frac{a}{b} [1 - u \text{Coth}(u)] \right\}, \quad a, b > 0, \quad (3.17)$$

here $a > 0$, $b > 0$ are the unknown parameter of $H(a, b)$ distribution. The scale parameter a do not change the posture and characteristics of the hazard rates and may be include in the coefficient β_0 . So, we obtain

$$S_{z(\cdot)}(t, \beta, \sigma) = G_0 \left\{ \left(\int_0^t e^{-\beta^T z(u)} du \right)^{\frac{1}{\sigma}}, \sigma \right\}, \quad \sigma = \frac{1}{b}, \quad z(\cdot) \in \mathbf{E}, \quad (3.18)$$

where

$$G_0(u, \sigma) = \text{Sech} \{ \sigma [1 - u \text{Coth}(u)] \}. \quad (3.19)$$

Set

$$G(u, \sigma) = G_0(e^u, \sigma), \quad \sigma > 0, \quad (3.20)$$

then the survival function of AFT_H model can be written as the form

$$S_{z(\cdot)}(t, \beta, \sigma) = G \left\{ \frac{1}{\sigma} \ln \left(\int_0^t e^{-\beta^T z(u)} du \right), \sigma \right\}, \quad \sigma > 0, \quad z(\cdot) \in \mathbf{E}, \quad (3.21)$$

if the covariate $z(\cdot) \equiv z$ is constant that

$$S_z(t, \beta, \sigma) = G \left\{ \frac{\ln t - \beta^T z}{\sigma}, \sigma \right\}, \quad \sigma > 0, \quad z \in \mathbf{E}_1, \quad (3.22)$$

where $G(\cdot, \sigma)$ and $G_0(\cdot, \sigma)$ are defined by the formulas (3.19) and (3.20), respectively. Note that, the formula (3.22) seem to us that the AFT_H model belongs the log-location scale parametric AFT classical models, such as AFT Weibull, AFT Log-normal or AFT Log-logistic models. However, the AFT_H is more interesting other. Since the Hypertabastic baseline function is more flexibility than shape-scale function distributions.

Following (3.21), we set $\boldsymbol{\theta} = (\beta^T, \sigma)^T$ and

$$f_{z(\cdot)}(\beta) = \int_0^t e^{-\beta^T z(u)} du, \quad v_{z(\cdot)}(\boldsymbol{\theta}) = \frac{1}{\sigma} \ln \left(\int_0^t e^{-\beta^T z(u)} du \right), \quad z(\cdot) \in \mathbf{E}, \quad (3.23)$$

under constant covariate then

$$f_z(\beta) = e^{-\beta^T z t}, \quad v_z(\boldsymbol{\theta}) = \frac{\ln t - \beta^T z}{\sigma}, \quad z \in \mathbf{E}_1. \quad (3.24)$$

We obtain the hazard rate function of AFT_H model

$$\lambda_{z(\cdot)}(t, \boldsymbol{\theta}) = e^{-\beta^T z(t)} e^{v_{z(\cdot)}(\boldsymbol{\theta})} f_{z(\cdot)}^{-1}(\beta) U_{z(\cdot)}(\boldsymbol{\theta}) \text{Tanh}(W_{z(\cdot)}(\boldsymbol{\theta})), \quad z(\cdot) \in \mathbf{E}, \quad (3.25)$$

where

$$W_{z(\cdot)}(\boldsymbol{\theta}) = \sigma \left[1 - e^{v_{z(\cdot)}(\boldsymbol{\theta})} \text{Coth}(e^{v_{z(\cdot)}(\boldsymbol{\theta})}) \right], \quad (3.26)$$

$$U_{z(\cdot)}(\boldsymbol{\theta}) = e^{v_{z(\cdot)}(\boldsymbol{\theta})} \text{Csch}^2(e^{v_{z(\cdot)}(\boldsymbol{\theta})}) - \text{Coth}(e^{v_{z(\cdot)}(\boldsymbol{\theta})}). \quad (3.27)$$

5.2 MLE for the regression parameters of AFT_H model

Returning to consider the right-censored data influenced by covaraites that are given in (2.19). Follows that, the likelihood function of AFT_H model is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \left\{ e^{-\beta^T z_i(X_i)} e^{V_i(\boldsymbol{\theta})} f_i^{-1}(\beta) U_i(\boldsymbol{\theta}) \text{Tanh}(W_i(\boldsymbol{\theta})) \right\}^{\delta_i} \text{Sech}W_i(\boldsymbol{\theta}), \quad (3.28)$$

$$\text{where } V_i(\boldsymbol{\theta}) = \frac{1}{\sigma} \ln \left(\int_0^{X_i} e^{-\beta^T z_i(u)} du \right), \quad f_i(\beta) = \int_0^{X_i} e^{-\beta^T z_i(u)} du, \quad z_i(\cdot) \in \mathbf{E}, \quad (3.29)$$

$$W_i(\boldsymbol{\theta}) = \sigma \left[1 - e^{V_i(\boldsymbol{\theta})} \text{Coth}(e^{V_i(\boldsymbol{\theta})}) \right], \quad (3.30)$$

$$U_i(\boldsymbol{\theta}) = e^{V_i(\boldsymbol{\theta})} \text{Csch}^2(e^{V_i(\boldsymbol{\theta})}) - \text{Coth}(e^{V_i(\boldsymbol{\theta})}), \quad (3.31)$$

if the covariate $z_i(\cdot)$ is constant then the likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \left\{ \frac{1}{t} e^{V_i(\boldsymbol{\theta})} U_i(\boldsymbol{\theta}) \text{Tanh}(W_i(\boldsymbol{\theta})) \right\}^{\delta_i} \text{Sech}W_i(\boldsymbol{\theta}), \quad z_i \in \mathbf{E}. \quad (3.32)$$

The log-likelihood function is

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \delta_i \left[-\beta^T z_i(X_i) - \ln f_i(\beta) + V_i(\boldsymbol{\theta}) + \ln \{U_i(\boldsymbol{\theta}) \operatorname{Tanh}(W_i(\boldsymbol{\theta}))\} \right] + \ln \operatorname{Sech}(W_i(\boldsymbol{\theta})) \right\}, \quad (3.33)$$

under the constant covariate, the log-likelihood function can be written as

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \delta_i \left[V_i(\boldsymbol{\theta}) + \ln (U_i(\boldsymbol{\theta}) \operatorname{Tanh}(W_i(\boldsymbol{\theta}))) - \ln X_i \right] + \ln (\operatorname{Sech}(W_i(\boldsymbol{\theta}))) \right\}, \quad (3.34)$$

$$\text{where } V_i(\boldsymbol{\theta}) = \frac{\ln X_i - \beta^T z_i}{\sigma}, \quad W_i(\boldsymbol{\theta}) = \sigma \left(1 - e^{V_i(\boldsymbol{\theta})} \operatorname{Coth}(e^{V_i(\boldsymbol{\theta})}) \right), \quad (3.35)$$

$$U_i(\boldsymbol{\theta}) = e^{V_i(\boldsymbol{\theta})} \operatorname{Csch}^2(e^{V_i(\boldsymbol{\theta})}) - \operatorname{Coth}(e^{V_i(\boldsymbol{\theta})}), \quad i = 1, 2, \dots, n. \quad (3.36)$$

The score functions are

$$\dot{\ell}_{\beta_l}(\boldsymbol{\theta}) := \frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta_l} = \sum_{i=1}^n \delta_i \{ f_i^l(\beta) - z_{il}(X_i) \} - \frac{1}{\sigma} \sum_{i=1}^n f_i^l(\beta) D_i(\boldsymbol{\theta}), \quad (3.37)$$

$$\dot{\ell}_{\sigma}(\boldsymbol{\theta}) := \frac{\partial \ell(\boldsymbol{\theta})}{\partial \sigma} = -\frac{1}{\sigma} \sum_{i=1}^n W_i(\boldsymbol{\theta}) B_i(\boldsymbol{\theta}) - \frac{1}{\sigma} \sum_{i=1}^n V_i(\boldsymbol{\theta}) D_i(\boldsymbol{\theta}), \quad (3.38)$$

where $\delta = \sum_{i=1}^n \delta_i$ and for all $l = 0, 1, \dots, m$, $i = 1, 2, \dots, n$,

$$f_i^l(\beta) = \frac{\int_0^{X_i} z_{il}(u) e^{-\beta^T z_i(u)} du}{\int_0^{X_i} e^{-\beta^T z_i(u)} du}, \quad B_i(\boldsymbol{\theta}) = \operatorname{Tanh}(W_i(\boldsymbol{\theta})) - \delta_i \frac{\operatorname{Sech}^2(W_i(\boldsymbol{\theta}))}{\operatorname{Tanh}(W_i(\boldsymbol{\theta}))}, \quad (3.39)$$

$$D_i(\boldsymbol{\theta}) = \delta_i \left(1 + \frac{2}{\sigma} A_i(\boldsymbol{\theta}) \right) - \sigma e^{V_i(\boldsymbol{\theta})} U_i(\boldsymbol{\theta}) B_i(\boldsymbol{\theta}), \quad (3.40)$$

$$\text{where } A_i(\boldsymbol{\theta}) = \frac{e^{V_i(\boldsymbol{\theta})} W_i(\boldsymbol{\theta}) \operatorname{Csch}^2(e^{V_i(\boldsymbol{\theta})})}{U_i(\boldsymbol{\theta})}, \quad (3.41)$$

and $f_i(\beta)$, $V_i(\boldsymbol{\theta})$, $W_i(\boldsymbol{\theta})$ and $U_i(\boldsymbol{\theta})$ are defined in (3.29)-(3.31), respectively.

Under the constant covariate $z_i(\cdot)$ the score functions (3.37), (3.38) can be written as

$$\dot{\ell}_{\beta_l}(\boldsymbol{\theta}) = -\frac{1}{\sigma} \sum_{i=1}^n z_{il} D_i(\boldsymbol{\theta}), \quad (l = 0, 1, \dots, m), \quad (3.42)$$

$$\dot{\ell}_{\sigma}(\boldsymbol{\theta}) = -\frac{1}{\sigma} \sum_{i=1}^n W_i(\boldsymbol{\theta}) B_i(\boldsymbol{\theta}) - \frac{1}{\sigma} \sum_{i=1}^n V_i(\boldsymbol{\theta}) D_i(\boldsymbol{\theta}), \quad (3.43)$$

where $B_i(\boldsymbol{\theta})$, $D_i(\boldsymbol{\theta})$ are defined in (3.39) - (3.41) and $V_i(\boldsymbol{\theta})$, $W_i(\boldsymbol{\theta})$ and $U_i(\boldsymbol{\theta})$ are defined in (3.29)-(3.31), respectively. The MLE's $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}^T, \hat{\sigma})^T$ are obtained by solving the system of equations

$$\dot{\ell}_{\beta_l}(\boldsymbol{\theta}) = \mathbf{0}_{m+1}, \quad \dot{\ell}_{\sigma}(\boldsymbol{\theta}) = 0, \quad (l = 0, 1, \dots, m).$$

5.3 Estimators of the reliability characteristic

Let us suppose that $z(\cdot) = (z_0(\cdot), z_1(\cdot), \dots, z_m(\cdot))^T \in \mathbf{E}$ is a possible time-varying covariate which may be different from $z_i(\cdot)$. Suppose that $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}^T, \hat{\sigma})^T$ are the MLE's of $\boldsymbol{\theta} = (\beta^T, \sigma)^T$. As follows the section 4 in chapter 2, the estimator of the survival function $S_{z(\cdot)}(t)$ of AFT_H model under covariate $z(\cdot)$ is

$$\hat{S}_{z(\cdot)}(t) = \text{Sech}\left\{\hat{\sigma}\left[1 - e^{\hat{V}_{z(\cdot)}} \text{Coth}(e^{\hat{V}_{z(\cdot)}})\right]\right\}, \quad \text{where} \quad \hat{V}_{z(\cdot)} = \frac{1}{\hat{\sigma}} \ln\left(\int_0^t e^{-\hat{\beta}^T z(u)} du\right). \quad (3.44)$$

Under the constant covariate $z(\cdot) = z$, we obtained

$$\hat{S}_z(t) = \text{Sech}\left\{\hat{\sigma}\left[1 - e^{\hat{V}_z} \text{Coth}\{e^{\hat{V}_z}\}\right]\right\}, \quad \text{where} \quad \hat{V}_z = \frac{\ln t - \hat{\beta}^T z}{\hat{\sigma}}. \quad (3.45)$$

5.4 Asymptotic distribution of the regression parameters estimators

Under the regularity conditions 2.1 in chapter 2, the limiting distribution of $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}^T, \hat{\sigma})^T$ has the normal law, that is

$$(\hat{\beta}, \hat{\sigma})^T \simeq N_{m+2}((\beta, \sigma)^T, \Sigma^{-1}(\hat{\beta}, \hat{\sigma})),$$

the covariance matrix $\Sigma^{-1}(\hat{\beta}, \hat{\sigma})$ is estimated by the inverse

$$\mathbf{I}^{-1}(\hat{\beta}, \hat{\sigma}) = [\mathbf{i}^{ls}(\hat{\beta}, \hat{\sigma})]_{(m+2) \times (m+2)},$$

of the Fisher information matrix $\mathbf{I}(\hat{\beta}, \hat{\sigma}) = \frac{-1}{n}[\mathbf{i}_{ls}(\hat{\beta}, \hat{\sigma})]_{(m+2) \times (m+2)}$ containing the elements below

$$\mathbf{i}_{ls}(\hat{\beta}, \hat{\sigma}) = \sum_{i=1}^n \ddot{f}_i^{ls} \left[\delta_i - \frac{1}{\hat{\sigma}} \hat{D}_i \right] - \frac{1}{\hat{\sigma}} \sum_{i=1}^n \dot{f}_i^l \dot{f}_i^s \{ \hat{K}_i + \hat{\sigma} e^{\hat{V}_i} \hat{U}_i^2 \hat{H}_i \}, \quad (3.46)$$

$$\mathbf{i}_{l(m+2)}(\hat{\beta}, \hat{\sigma}) = \frac{1}{\hat{\sigma}} \sum_{i=1}^n \dot{f}_i^l \left\{ \frac{1}{\hat{\sigma}} \hat{D}_i + \hat{P}_i - \hat{V}_i \hat{K}_i \right\}, \quad (3.47)$$

$$\mathbf{i}_{(m+2)(m+2)}(\hat{\beta}, \hat{\sigma}) = \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (\hat{W}_i \hat{B}_i + 2\hat{V}_i \hat{D}_i) - \frac{1}{\hat{\sigma}} \sum_{i=1}^n \hat{V}_i \{\hat{V}_i \hat{K}_i - \hat{P}_i\} - \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (\hat{W}_i - \hat{\sigma} \hat{V}_i e^{\hat{V}_i} \hat{U}_i) (\hat{B}_i + \hat{W}_i \hat{H}_i) + \frac{\delta}{\hat{\sigma}^2}, \quad (3.48)$$

$$\begin{aligned} \text{where } \hat{f}_i^l &= \frac{\int_0^{X_i} z_{il}(u) e^{-\hat{\beta}^T z_i(u)} du}{\int_0^{X_i} e^{-\hat{\beta}^T z_i(u)} du}, \quad (l = 0, 1, \dots, m), \\ \hat{f}_i^{ls} &= \frac{\int_0^{X_i} z_{il}(u) e^{-\hat{\beta}^T z_i(u)} du \int_0^{X_i} z_{is}(u) e^{-\hat{\beta}^T z_i(u)} du}{\left(\int_0^{X_i} e^{-\hat{\beta}^T z_i(u)} du\right)^2} - \frac{\int_0^{X_i} z_{il}(u) z_{is}(u) e^{-\hat{\beta}^T z_i(u)} du}{\int_0^{X_i} e^{-\hat{\beta}^T z_i(u)} du}, \\ \hat{V}_i &= \frac{1}{\hat{\sigma}} \ln \left(\int_0^{X_i} e^{-\hat{\beta}^T z_i(u)} du \right), \quad \hat{W}_i = \hat{\sigma} [1 - e^{\hat{V}_i} \text{Coth}(e^{\hat{V}_i})], \\ \hat{U}_i &= e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i}) - \text{Coth}(e^{\hat{V}_i}), \quad \hat{A}_i = \frac{e^{\hat{V}_i} \hat{W}_i \text{Csch}^2(e^{\hat{V}_i})}{\hat{U}_i}, \\ \hat{B}_i &= \text{Tanh}(\hat{W}_i) - \delta_i \frac{\text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)}, \quad \hat{D}_i = \delta_i \left(1 + \frac{2}{\hat{\sigma}} \hat{A}_i\right) - \hat{\sigma} e^{\hat{V}_i} \hat{U}_i \hat{B}_i, \\ \hat{H}_i &= \text{Sech}^2(\hat{W}_i) \left\{1 + 2\delta_i + \delta_i \frac{\text{Sech}^2(\hat{W}_i)}{\text{Tanh}^2(\hat{W}_i)}\right\}, \quad \hat{P}_i = e^{\hat{V}_i} \hat{U}_i \left\{\hat{B}_i - (\hat{W}_i - \hat{\sigma} \hat{V}_i e^{\hat{V}_i} \hat{U}_i) \hat{H}_i\right\}, \\ \hat{K}_i &= -\frac{2\delta_i}{\hat{\sigma}^2} \hat{Q}_i + e^{\hat{V}_i} \left\{\hat{U}_i \hat{B}_i + \frac{2}{\hat{\sigma}} e^{\hat{V}_i} \hat{W}_i \hat{B}_i \text{Csch}^2(e^{\hat{V}_i})\right\}, \\ \text{and } \hat{Q}_i &= \hat{A}_i \left\{1 - \frac{2}{\hat{\sigma}} \hat{A}_i - 2e^{\hat{V}_i} \text{Coth}(e^{\hat{V}_i})\right\} + \hat{\sigma} e^{2\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i}). \end{aligned}$$

If the covariate $z_i(\cdot)$ is constant over the time, we have

$$\mathbf{i}_{ls}(\hat{\beta}, \hat{\sigma}) = -\frac{1}{\hat{\sigma}} \sum_{i=1}^n z_{il} z_{is} \left\{ \hat{K}_i + \hat{\sigma} e^{\hat{V}_i} \hat{U}_i^2 \hat{H}_i \right\}, \quad (3.49)$$

$$\mathbf{i}_{l(m+2)}(\hat{\beta}, \hat{\sigma}) = \frac{1}{\hat{\sigma}} \sum_{i=1}^n z_{il} \left\{ \frac{1}{\hat{\sigma}} \hat{D}_i + \hat{P}_i - \hat{V}_i \hat{K}_i \right\}, \quad (3.50)$$

$$\begin{aligned} \mathbf{i}_{(m+2)(m+2)}(\hat{\beta}, \hat{\sigma}) &= \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (\hat{W}_i \hat{B}_i + 2\hat{V}_i \hat{D}_i) - \frac{1}{\hat{\sigma}} \sum_{i=1}^n \hat{V}_i \{\hat{V}_i \hat{K}_i - \hat{P}_i\} - \\ &\quad \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (\hat{W}_i - \hat{\sigma} \hat{V}_i e^{\hat{V}_i} \hat{U}_i) (\hat{B}_i + \hat{W}_i \hat{H}_i). \end{aligned} \quad (3.51)$$

In particular, the CI approximation at $(1 - \alpha)$ level for the parameter β_l

$$\beta_l \in \left\{ \hat{\beta}_l - w_{1-\alpha/2} \sigma(\hat{\beta}_l), \hat{\beta}_l + w_{1-\alpha/2} \sigma(\hat{\beta}_l) \right\}, \quad (l = 0, 1, \dots, m), \quad (3.52)$$

where $w_{1-\alpha/2}$ is the $(1 - \frac{\alpha}{2})$ quantile of the normal law and the variance $\sigma^2(\hat{\beta}_l)$ of the estimator $\hat{\beta}_l$ is estimated by

$$\sigma^2(\hat{\beta}_l) = \mathbf{i}^l(\hat{\beta}, \hat{\sigma}), \quad (l = 0, 1, \dots, m).$$

5.5 Approximate CI for the reliability characteristic

For any covariate $z(\cdot) = (z_0(\cdot), z_1(\cdot), \dots, z_m(\cdot))^T \in \mathbf{E}$, using (2.41) in chapter 2, we obtain the approximate $(1 - \alpha)$ CI for survival function $S_{z(\cdot)}$ that is given as follows.

$$\left\{ \left[1 + \frac{1 - \hat{S}_{z(\cdot)}(t)}{\hat{S}_{z(\cdot)}(t)} e^{\sigma_{Q_i}(\hat{\beta}, \hat{\sigma}) w_{1-\alpha/2}} \right]^{-1}; \left[1 + \frac{1 - \hat{S}_{z(\cdot)}(t)}{\hat{S}_{z(\cdot)}(t)} e^{-\sigma_{Q_i}(\hat{\beta}, \hat{\sigma}) w_{1-\alpha/2}} \right]^{-1} \right\}, \quad (3.53)$$

where $w_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the normal distribution, $\hat{S}_{z(\cdot)}(t)$ is given in (3.38) and

$$\sigma_{Q_i}^2(\hat{\beta}, \hat{\sigma}) = \frac{\mathbf{J}_{S_{z(\cdot)}}^T(\hat{\beta}, \hat{\sigma}) \mathbf{I}^{-1}(\hat{\beta}, \hat{\sigma}) \mathbf{J}_{S_{z(\cdot)}}(\hat{\beta}, \hat{\sigma})}{\hat{S}_{z(\cdot)}^2(t) [1 - \hat{S}_{z(\cdot)}(t)]^2}, \quad (3.54)$$

where

$$\begin{aligned} \mathbf{J}_{S_{z(\cdot)}}(\hat{\beta}, \hat{\sigma}) &= \left(\frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_0}, \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_1}, \dots, \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_m}, \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \sigma} \right)^T, \\ \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_l} &= \hat{f}_z^l e^{\hat{V}_z} \hat{U}_z \text{Sech}(\hat{W}_z) \text{Tanh}(\hat{W}_z), \quad l = 0, 1, \dots, m, \\ \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \sigma} &= -\frac{1}{\hat{\sigma}} (\hat{W}_z - \hat{\sigma} \hat{V}_z \hat{U}_z e^{\hat{V}_z}) \text{Sech}(\hat{W}_z) \text{Tanh}(\hat{W}_z), \\ \text{here } \hat{f}_z^l &= \frac{\int_0^t z_l(u) e^{-\hat{\beta}^T z(u)} du}{\int_0^t e^{-\hat{\beta}^T z(u)} du}, \quad \hat{V}_z = \frac{1}{\hat{\sigma}} \ln \left(\int_0^t e^{-\hat{\beta}^T z(u)} du \right), \\ \hat{W}_z &= \hat{\sigma} [1 - e^{\hat{V}_z} \text{Coth}(e^{\hat{V}_z})], \quad \hat{U}_z = e^{\hat{V}_z} \text{Csch}^2(e^{\hat{V}_z}) - \text{Coth}(e^{\hat{V}_z}). \end{aligned}$$

5.6 Validity of Hypertabastic AFT model using \mathbf{Y}_n^2

Under the right censored failure-time data under the constant covariates that are defined in (2.21), we consider the hypothesis H_0 stating AFT_H model as follows

$$H_0 : S_i(t) = \text{Sech}\{\sigma [1 - e^{V_i} \text{Coth}(e^{V_i})]\}, \quad \text{where } V_i = \frac{\ln t - \beta^T z_i}{\sigma}, \quad t > 0, \quad (3.55)$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$ is a vector of unknown regression parameters, $\sigma > 0$. The hypothesis H_0 in (3.55) can be written as the following expressions of the hazard rate or cumulative hazard functions, respectively

$$\lambda_i(t) = \frac{1}{t} e^{V_i} [e^{V_i} \text{Csch}^2(e^{V_i}) - \text{Coth}(e^{V_i})] \text{Tanh}\{\sigma [1 - e^{V_i} \text{Coth}(e^{V_i})]\},$$

$$\Lambda_i(t) = -\ln \{ \text{Sech}\{\hat{\sigma} [1 - e^{\hat{V}_i} \text{Coth}(e^{\hat{V}_i})]\} \}.$$

Suppose that $\hat{\theta}_n = (\hat{\beta}^T, \hat{\sigma})^T$ are the MLE's of $\theta = (\beta^T, \sigma)^T$. The statistic \mathbf{Y}_n^2 and its components are given from (2.55) to (2.59). We obtain

$$\mathbf{Y}_n^2 = \mathbf{X}_n^2 + \mathbf{Q}, \quad \mathbf{X}_n^2 = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j}, \quad \mathbf{Q} = \mathbf{W}^T \mathbf{G}^{-1} \mathbf{W}, \quad U_j = \sum_{i: X_i \in I_j} \delta_i,$$

$$\mathbf{G} = \mathbf{i} - \hat{\mathcal{C}} \hat{\mathcal{A}}^{-1} \hat{\mathcal{C}}^T, \quad \mathbf{W} = \hat{\mathcal{C}} \hat{\mathcal{A}}^{-1} \mathbf{Z}, \quad \mathbf{Z} = (Z_1, Z_2, \dots, Z_k)^T, \quad Z_j = \frac{1}{\sqrt{n}} (U_j - e_j).$$

The diagonal matrix $\hat{\mathcal{A}}$ includes the diagonal elements $\hat{A}_j = \frac{U_j}{n}$, the elements of the matrix $\mathbf{i} = [\hat{i}_{ll'}]_{(m+2) \times (m+2)}$ and $\mathcal{C} = [\hat{\mathcal{C}}_{lj}]_{(m+2) \times k}$ are calculated by taking the formulas (2.58) and (2.59) as follows

$$\hat{i}_{ll'} = \frac{1}{n \hat{\sigma}^2} \sum_{i=1}^n \delta_i z_{il} z_{il'} \left\{ 1 + \frac{2 \hat{W}_i e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i})}{\hat{\sigma} \hat{U}_i} + \frac{\hat{\sigma} \hat{U}_i e^{\hat{V}_i} \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \right\}^2,$$

$$\hat{i}_{l(m+1)} = \frac{1}{n \hat{\sigma}^2} \sum_{i=1}^n \delta_i z_{il} \left\{ 1 + \frac{2 \hat{W}_i e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i})}{\hat{\sigma} \hat{U}_i} + \frac{\hat{\sigma} \hat{U}_i e^{\hat{V}_i} \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \right\} \times$$

$$\left\{ 1 + \hat{V}_i + \frac{2 \hat{V}_i \hat{W}_i e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i})}{\hat{\sigma} \hat{U}_i} - \frac{(\hat{W}_i - \hat{\sigma} \hat{V}_i \hat{U}_i e^{\hat{V}_i}) \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \right\},$$

$$\hat{i}_{(m+1)(m+1)} = \frac{1}{n \hat{\sigma}^2} \sum_{i=1}^n \delta_i \left\{ 1 + \hat{V}_i + \frac{2 \hat{V}_i \hat{W}_i e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i})}{\hat{\sigma} \hat{U}_i} - \frac{(\hat{W}_i - \hat{\sigma} \hat{V}_i \hat{U}_i e^{\hat{V}_i}) \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \right\}^2,$$

$$\hat{\mathcal{C}}_{lj} = \frac{-1}{n \hat{\sigma}} \sum_{i: X_i \in I_j} \delta_i z_{il} z_{il'} \left\{ 1 + \frac{2 \hat{W}_i e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i})}{\hat{\sigma} \hat{U}_i} + \frac{\hat{\sigma} \hat{U}_i e^{\hat{V}_i} \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \right\},$$

$$\hat{\mathcal{C}}_{(m+1)j} = \frac{-1}{n \hat{\sigma}} \sum_{i: X_i \in I_j} \delta_i \left\{ 1 + \hat{V}_i + \frac{2 \hat{V}_i \hat{W}_i e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i})}{\hat{\sigma} \hat{U}_i} - \frac{(\hat{W}_i - \hat{\sigma} \hat{V}_i \hat{U}_i e^{\hat{V}_i}) \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \right\},$$

respectively, e_j is given in (2.43) and

$$\begin{aligned}\hat{V}_i &= \frac{\ln X_i - \hat{\beta}^T z_i}{\hat{\sigma}}, & \hat{W}_i &= \hat{\sigma} [1 - e^{\hat{V}_i} \text{Coth}(e^{\hat{V}_i})], \\ \hat{U}_i &= e^{\hat{V}_i} \text{Csch}^2(e^{\hat{V}_i}) - \text{Coth}(e^{\hat{V}_i}), & i &= 1, 2, \dots, n.\end{aligned}$$

For testing AFT_H model based on \mathbf{Y}_n^2 statistic, the endpoints a_j are found from the equations (2.61) - (2.63) based on the numerical method. Since the matrix \mathbf{G} is non-degenerate, so $\text{rank}(G) = k$. As follows Theorem 2.4, the limiting distribution of statistic \mathbf{Y}_n^2 is a chi-squared with k degrees of freedom. So, the hypothesis H_0 is rejected with an approximate significance level α if $\mathbf{Y}_n^2 > \chi_\alpha^2(k)$.

6 Validity of Hypertabastic PH model

6.1 Hypertabastic PH model

Suppose that each individual is taken with associated the possible time-varying m -dimensional vector of covariate

$$z(\cdot) = (z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot))^T \in \mathbf{E}.$$

We consider the parametric Hypertabastic PH (PH_H) model under the covariate $z(\cdot) \in \mathbf{E}$, are given in the form

$$\lambda_{z(\cdot)}(t, \boldsymbol{\theta}) = at^{b-1} [t^b \text{Csch}^2(t^b) - \text{Coth}(t^b)] \text{Tanh}\left(\frac{a}{b} [1 - t^b \text{Coth}(t^b)]\right) e^{\beta^T z(t)}, \quad z(\cdot) \in \mathbf{E}, \quad (3.56)$$

where $\beta = (\beta_1, \beta_2, \dots, \beta_m)^T$ is a vector of unknown regression parameter, a, b are the unknown parameters of the Hypertabastic baseline, we set $\boldsymbol{\theta} = (\beta^T, a, b)^T$.

The PH_H model can be written as the following form of survival and cumulative functions, respectively

$$S_{z(\cdot)}(t, \boldsymbol{\theta}) = \exp \left\{ - \int_0^t e^{\beta^T z(u)} d\Lambda_0(u, a, b) \right\}, \quad \Lambda_{z(\cdot)}(t, \boldsymbol{\theta}) = \int_0^t e^{\beta^T z(u)} d\Lambda_0(u, a, b), \quad (3.57)$$

where $\Lambda_0(t, a, b)$ is the baseline cumulative hazard of the $H(a, b)$ distribution given as

$$\Lambda_0(t, a, b) = - \ln \left\{ \text{Sech} \left(\frac{a}{b} [1 - t^b \text{Coth}(t^b)] \right) \right\}, \quad a, b > 0. \quad (3.58)$$

If the covariate $z(\cdot)$ is a constant over the time, we obtained follow the expressions of the survival, hazard rate and the cumulative hazard functions of PH_H model

$$\lambda_z(t, \boldsymbol{\theta}) = at^{b-1} [t^b \text{Csch}^2(t^b) - \text{Coth}(t^b)] \text{Tanh}\left(\frac{a}{b} [1 - t^b \text{Coth}(t^b)]\right) e^{\beta^T z}, \quad z \in \mathbf{E}_1, \quad (3.59)$$

$$S_z(t, \boldsymbol{\theta}) = \left\{ \text{Sech}\left(\frac{a}{b} [1 - t^b \text{Coth}(t^b)]\right) \right\}^{e^{\beta^T z}}, \quad \Lambda_z(t, \boldsymbol{\theta}) = e^{\beta^T z} \Lambda_0(t, a, b), \quad (3.60)$$

respectively.

6.2 MLE for the regression parameters

Returning to consider the right-censored data influenced by covaraites that are given in (2.19). Follows that, the likelihood function of PH_H model can be written as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \left\{ [aX_i^{b-1} U_i \text{Tanh}(W_i) e^{\beta^T z_i(X_i)}]^{\delta_i} \exp \left\{ - \int_0^{X_i} e^{\beta^T z_i(u)} \lambda_0(u, a, b) du \right\} \right\}, \quad (3.61)$$

where

$$W_i = \frac{a}{b} [1 - X_i^b \text{Coth}(X_i^b)], \quad U_i = X_i^b \text{Csch}^2(X_i^b) - \text{Coth}(X_i^b), \quad (3.62)$$

and $\lambda_0(u, a, b)$ is the hazard rate baseline function is given as (3.4).

The log-likelihood function of PH_H model is given as

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \delta_i [\ln a + (b-1) \ln X_i + \beta^T z_i(X_i) + \ln \{U_i \text{Tanh}(W_i)\}] - \int_0^{X_i} e^{\beta^T z_i(u)} \lambda_0(u, a, b) du \right\}, \quad \boldsymbol{\theta} = (\beta^T, a, b)^T. \quad (3.63)$$

If $z_i(\cdot)$ is constant covariate then

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \delta_i [\ln a + \beta^T z_i + (b-1) \ln X_i + \ln (U_i \text{Tanh}(W_i))] + e^{\beta^T z_i} \ln (\text{Sech}(W_i)) \right\}. \quad (3.64)$$

The score functions are

$$\begin{aligned} \dot{\ell}_{\beta_i}(\boldsymbol{\theta}) &= \sum_{i=1}^n \left\{ \delta_i z_{il}(X_i) - \int_0^{X_i} z_{il}(u) e^{\beta^T z_i(u)} \lambda_0(u, a, b) du \right\}, \\ \dot{\ell}_a(\boldsymbol{\theta}) &= \frac{1}{a} \sum_{i=1}^n \delta_i \left(1 + \frac{W_i \text{Sech}^2(W_i)}{\text{Tanh}(W_i)} \right) - \sum_{i=1}^n \int_0^{X_i} e^{\beta^T z_i(u)} \frac{\partial \lambda_0(u, a, b)}{\partial a} du, \end{aligned}$$

$$\dot{\ell}_b(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \delta_i \left(\ln X_i + \frac{2b N_i}{a U_i} - \frac{1}{b} \frac{K_i \operatorname{Sech}^2(W_i)}{\operatorname{Tanh}(W_i)} \right) - \int_0^{X_i} e^{\beta^T z_i(u)} \frac{\partial \lambda_0(u, a, b)}{\partial b} du \right\},$$

where W_i, U_i are given in (3.62) and

$$N_i = X_i^b W_i \operatorname{Csch}^2(X_i^b) \ln X_i, \quad K_i = (W_i - a X_i^b U_i \ln X_i). \quad (3.65)$$

Under constant covariate $z_i(\cdot) \equiv z_i \in \mathbf{E}_1$, the score function can be written as follows

$$\begin{aligned} \dot{\ell}_{\beta_l}(\boldsymbol{\theta}) &= \sum_{i=1}^n z_{il} \left\{ \delta_i + e^{\beta^T z_i} \ln(\operatorname{Sech}(W_i)) \right\}, \\ \dot{\ell}_a(\boldsymbol{\theta}) &= \frac{1}{a} \sum_{i=1}^n \left(\delta_i + \frac{W_i (\delta_i \operatorname{Sech}^2(W_i) - e^{\beta^T z_i} \operatorname{Tanh}^2(W_i))}{\operatorname{Tanh}(W_i)} \right), \\ \dot{\ell}_b(\boldsymbol{\theta}) &= \sum_{i=1}^n \delta_i \left(\ln X_i + \frac{2b N_i}{a U_i} \right) - \frac{1}{b} \sum_{i=1}^n \frac{K_i (\delta_i \operatorname{Sech}^2(W_i) - e^{\beta^T z_i} \operatorname{Tanh}^2(W_i))}{\operatorname{Tanh}(W_i)}, \quad z_i \in \mathbf{E}_1. \end{aligned}$$

The parameters $\boldsymbol{\theta} = (\beta^T, a, b)^T$ can be estimated by maximizing the log-likelihood function $\ell(\boldsymbol{\theta})$. Suppose that $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}^T, \hat{a}, \hat{b})^T$ is the MLE's of $\boldsymbol{\theta}$ which is obtained by solving the system equation

$$\dot{\ell}_{\beta_l}(\boldsymbol{\theta}) = \mathbf{0}_m, \quad \dot{\ell}_a(\boldsymbol{\theta}) = 0, \quad \dot{\ell}_b(\boldsymbol{\theta}) = 0, \quad (l = 1, 2, \dots, m).$$

6.3 Estimators of the reliability characteristic

Let us consider any possible time-varying covariate $z(\cdot) = (z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot)) \in \mathbf{E}$, $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}^T, \hat{a}, \hat{b})^T$ is the MLE's of $\boldsymbol{\theta}$. The estimator $\hat{S}_{z(\cdot)}(t)$ of survival function $S_{z(\cdot)}(t)$ of PH_H model under covariate $z(\cdot)$ has

$$\hat{S}_{z(\cdot)}(t) = \exp \left\{ -\hat{a} \int_0^t e^{\hat{\beta}^T z(u)} u^{\hat{b}-1} \left[u^{\hat{b}} \operatorname{Csch}^2(u^{\hat{b}}) - \operatorname{Coth}(u^{\hat{b}}) \right] \operatorname{Tanh} \left(\frac{\hat{a}}{\hat{b}} \left[1 - u^{\hat{b}} \operatorname{Coth}(u^{\hat{b}}) \right] \right) du \right\}, \quad (3.66)$$

if the covariate $z(\cdot)$ is constant over the time, we obtained below the estimator of the PH_H survival function

$$\hat{S}_z(t) = \left\{ \operatorname{Sech} \left(\frac{\hat{a}}{\hat{b}} \left[1 - t^{\hat{b}} \operatorname{Coth}(t^{\hat{b}}) \right] \right) \right\}^{e^{\hat{\beta}^T z}}, \quad z \in \mathbf{E}_1. \quad (3.67)$$

6.4 Asymptotic distribution of the parameters estimators

Suppose that the regularity condition 2.1 hold. As follows (2.36), we obtain the asymptotic distribution of the parameters estimators $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}^T, \hat{a}, \hat{b})^T$ has

$$(\hat{\beta}, \hat{a}, \hat{b})^T \simeq N_{m+2}((\beta, a, b)^T, \Sigma^{-1}(\beta, a, b)).$$

In particular, the estimator of the covariance matrix $\Sigma^{-1}(\beta, a, b)$ can be estimated by

$$\mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}_n) = [\mathbf{i}^{ls}(\hat{\boldsymbol{\theta}}_n)]_{(m+2) \times (m+2)},$$

where $\mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}_n)$ is inverse of the information matrix $\mathbf{I}(\hat{\boldsymbol{\theta}}_n) = \frac{-1}{n} [\mathbf{i}^{ls}(\hat{\boldsymbol{\theta}}_n)]_{(m+2) \times (m+2)}$ with the following elements

$$\begin{aligned} \mathbf{i}_{ls}(\hat{\boldsymbol{\theta}}_n) &= - \sum_{i=1}^n \int_0^{X_i} z_{il}(u) z_{is}(u) e^{\hat{\beta}^T z_i(u)} \lambda_0(u, \hat{a}, \hat{b}) du, \\ \mathbf{i}_{l(m+1)}(\hat{\boldsymbol{\theta}}_n) &= - \sum_{i=1}^n \int_0^{X_i} z_{il}(u) e^{\hat{\beta}^T z_i(u)} \frac{\partial \lambda_0(u, \hat{a}, \hat{b})}{\partial a} du, \\ \mathbf{i}_{l(m+2)}(\hat{\boldsymbol{\theta}}_n) &= - \sum_{i=1}^n \int_0^{X_i} z_{il}(u) e^{\hat{\beta}^T z_i(u)} \frac{\partial \lambda_0(u, \hat{a}, \hat{b})}{\partial b} du, \\ \mathbf{i}_{(m+1)(m+1)}(\hat{\boldsymbol{\theta}}_n) &= - \frac{1}{\hat{a}^2} \sum_{i=1}^n \delta_i \left\{ 1 + \frac{\hat{W}_i^2 [1 - \text{Tanh}^4(\hat{W}_i)]}{\text{Tanh}^2(\hat{W}_i)} \right\} - \sum_{i=1}^n \int_0^{X_i} e^{\hat{\beta}^T z_i(u)} \frac{\partial^2 \lambda_0(u, \hat{a}, \hat{b})}{\partial a^2} du, \\ \mathbf{i}_{(m+1)(m+2)}(\hat{\boldsymbol{\theta}}_n) &= - \frac{1}{\hat{a}\hat{b}} \sum_{i=1}^n \delta_i \frac{\hat{K}_i \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \left\{ 1 - \frac{\hat{W}_i [1 + \text{Tanh}^2(\hat{W}_i)]}{\text{Tanh}(\hat{W}_i)} \right\} - \\ &\quad \sum_{i=1}^n \int_0^{X_i} e^{\hat{\beta}^T z_i(u)} \frac{\partial^2 \lambda_0(u, \hat{a}, \hat{b})}{\partial a \partial b} du, \\ \mathbf{i}_{(m+2)(m+2)}(\hat{\boldsymbol{\theta}}_n) &= \frac{2\hat{b}}{\hat{a}} \sum_{i=1}^n \delta_i \left\{ \frac{\hat{N}_i}{\hat{b}\hat{U}_i} + \frac{\hat{N}_i^b}{\hat{U}_i} - \frac{2\hat{b}}{\hat{a}} \left(\frac{\hat{N}_i}{\hat{U}_i} \right)^2 \right\} - \sum_{i=1}^n \int_0^{X_i} e^{\hat{\beta}^T z_i(u)} \frac{\partial^2 \lambda_0(u, \hat{a}, \hat{b})}{\partial b^2} du + \\ &\quad \frac{1}{\hat{b}^2} \sum_{i=1}^n \delta_i \frac{\text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \left\{ (\hat{K}_i - \hat{b}\hat{K}_i^b) - \frac{\hat{K}_i^2 (1 + \text{Tanh}^2(\hat{W}_i))}{\text{Tanh}(\hat{W}_i)} \right\}, \end{aligned}$$

$$\text{where } \hat{W}_i = \frac{\hat{a}}{\hat{b}} [1 - X_i^{\hat{b}} \text{Coth}(X_i^{\hat{b}})], \quad \hat{U}_i = X_i^{\hat{b}} \text{Csch}^2(X_i^{\hat{b}}) - \text{Coth}(X_i^{\hat{b}}), \quad (3.68)$$

$$\hat{N}_i = X_i^{\hat{b}} W_i \text{Csch}^2(X_i^{\hat{b}}) \ln X_i, \quad \hat{N}_i^b = \left\{ -\frac{1}{\hat{b}} \hat{K}_i X_i^{\hat{b}} \text{Csch}^2(X_i^{\hat{b}}) + \hat{N}_i - 2\hat{N}_i X_i^{\hat{b}} \text{Coth}(X_i^{\hat{b}}) \right\} \ln X_i,$$

$$\hat{K}_i = (\hat{W}_i - \hat{a} X_i^{\hat{b}} \hat{U}_i \ln X_i), \quad \hat{K}_i^b = -\frac{1}{\hat{b}} \hat{K}_i - (2\hat{b}\hat{N}_i + \hat{a} X_i^{\hat{b}} \hat{U}_i \ln X_i) \ln X_i. \quad (3.69)$$

Under the constant covariates z_i , we obtain

$$\begin{aligned}
\mathbf{i}_{ls}(\hat{\boldsymbol{\theta}}_n) &= \sum_{i=1}^n z_{il} z_{is} e^{\hat{\beta}^T z_i} \ln \text{Sech}(\hat{W}_i), \\
\mathbf{i}_{l(m+1)}(\hat{\boldsymbol{\theta}}_n) &= -\frac{1}{\hat{a}} \sum_{i=1}^n z_{il} e^{\hat{\beta}^T z_i} \hat{W}_i \text{Tanh}(\hat{W}_i), \quad \mathbf{i}_{l(m+2)}(\hat{\boldsymbol{\theta}}_n) = \frac{1}{\hat{b}} \sum_{i=1}^n z_{il} e^{\hat{\beta}^T z_i} \hat{K}_i \text{Tanh}(\hat{W}_i), \\
\mathbf{i}_{(m+1)(m+1)}(\hat{\boldsymbol{\theta}}_n) &= -\frac{1}{\hat{a}^2} \sum_{i=1}^n \left\{ \delta_i + \frac{\hat{W}_i^2 \text{Sech}^2(\hat{W}_i) [\delta_i (1 + \text{Tanh}^2(\hat{W}_i)) + e^{\hat{\beta}^T z_i} \text{Tanh}^2(\hat{W}_i)]}{\text{Tanh}^2(\hat{W}_i)} \right\}, \\
\mathbf{i}_{(m+1)(m+2)}(\hat{\boldsymbol{\theta}}_n) &= \frac{1}{\hat{a}\hat{b}} \sum_{i=1}^n \left\{ \frac{\delta_i \hat{K}_i \text{Sech}^2(\hat{W}_i)}{\text{Tanh}(\hat{W}_i)} \left(-1 + \frac{\hat{W}_i (1 + \text{Tanh}^2(\hat{W}_i))}{\text{Tanh}(\hat{W}_i)} \right) + \right. \\
&\quad \left. e^{\hat{\beta}^T z_i} \hat{K}_i [\text{Tanh}(\hat{W}_i) + \hat{W}_i \text{Sech}^2(\hat{W}_i)] \right\}, \\
\mathbf{i}_{(m+2)(m+2)}(\hat{\boldsymbol{\theta}}_n) &= \frac{1}{\hat{b}^2} \sum_{i=1}^n \frac{(\hat{K}_i) - \hat{b} \hat{K}_i^b}{\text{Tanh}(\hat{W}_i)} (\delta_i \text{Sech}^2(\hat{W}_i) - e^{\hat{\beta}^T z_i} \text{Tanh}^2(\hat{W}_i)) - \\
&\quad \frac{1}{\hat{b}^2} \sum_{i=1}^n \frac{\hat{K}_i^2 \text{Sech}^2(\hat{W}_i) [\delta_i (1 + \text{Tanh}^2(\hat{W}_i)) + e^{\hat{\beta}^T z_i} \text{Tanh}^2(\hat{W}_i)]}{\text{Tanh}^2(\hat{W}_i)} + \\
&\quad \frac{2\hat{b}}{\hat{a}} \sum_{i=1}^n \delta_i \left\{ \frac{\hat{N}_i}{\hat{b}\hat{U}_i} + \frac{\hat{N}_i^b}{\hat{U}_i} - \frac{2\hat{b}}{\hat{a}} \left(\frac{\hat{N}_i}{\hat{U}_i} \right)^2 \right\},
\end{aligned}$$

where $\hat{W}_i, \hat{U}_i, \hat{N}_i, \hat{K}_i, \hat{N}_i^b$ and \hat{K}_i^b are given from (3.62) to (3.65), respectively.

6.5 Approximate CI for the reliability characteristic

Let us consider any covariate $z(\cdot) = (z_1(\cdot), z_2(\cdot), \dots, z_m(\cdot))^T \in \mathbf{E}$, using the formula (2.41) in chapter 2, we obtain the approximate $(1 - \alpha)$ confidence intervals for survival function $S_{z(\cdot)}$ that is given by the following formula

$$\left\{ \left[1 + \frac{1 - \hat{S}_{z(\cdot)}(t)}{\hat{S}_{z(\cdot)}(t)} e^{\sigma_{Q_i}(\hat{\boldsymbol{\theta}}_n) w_{1-\alpha/2}} \right]^{-1}; \left[1 + \frac{1 - \hat{S}_{z(\cdot)}(t)}{\hat{S}_{z(\cdot)}(t)} e^{-\sigma_{Q_i}(\hat{\boldsymbol{\theta}}_n) w_{1-\alpha/2}} \right]^{-1} \right\}, \quad (3.70)$$

where $w_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the normal distribution, $\hat{S}_{z(\cdot)}(t)$ is the estimator of $S_{z(\cdot)}(t)$ which is defined in the form (3.60) and

$$\sigma_{Q_i}^2(\hat{\boldsymbol{\theta}}_n) = \frac{\mathbf{J}_{S_{z(\cdot)}}^T(\hat{\boldsymbol{\theta}}_n) \mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}_n) \mathbf{J}_{S_{z(\cdot)}}(\hat{\boldsymbol{\theta}}_n)}{\hat{S}_{z(\cdot)}^2(t) [1 - \hat{S}_{z(\cdot)}(t)]^2}, \quad (3.71)$$

$$\text{where } \mathbf{J}_{S_{z(\cdot)}}(\hat{\beta}, \hat{\sigma}) = \left(\frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_1}, \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_2}, \dots, \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_m}, \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial a}, \frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial b} \right)^T,$$

$$\frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial \beta_l} = -\exp \left\{ -\int_0^t e^{\hat{\beta}^T z(u)} d\Lambda_0(u, \hat{a}, \hat{b}) \right\} \int_0^t z_l(u) e^{\hat{\beta}^T z(u)} d\Lambda_0(u, \hat{a}, \hat{b}),$$

$$\frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial a} = -\exp \left\{ -\int_0^t e^{\hat{\beta}^T z(u)} d\Lambda_0(u, \hat{a}, \hat{b}) \right\} \int_0^t e^{\hat{\beta}^T z(u)} \frac{\partial \lambda_0(u, \hat{a}, \hat{b})}{\partial a} d\Lambda_0(u, \hat{a}, \hat{b}),$$

$$\frac{\partial \hat{S}_{z(\cdot)}(t)}{\partial b} = -\exp \left\{ -\int_0^t e^{\hat{\beta}^T z(u)} d\Lambda_0(u, \hat{a}, \hat{b}) \right\} \int_0^t e^{\hat{\beta}^T z(u)} \frac{\partial \lambda_0(u, \hat{a}, \hat{b})}{\partial b} d\Lambda_0(u, \hat{a}, \hat{b}),$$

$\lambda_0(t, a, b)$ and $\Lambda_0(t, a, b)$ are the hazard rate and cumulative hazard functions of baseline Hypertabastic distribution and

$$\frac{\partial \lambda_0(t, \hat{a}, \hat{b})}{\partial a} = \frac{1}{\hat{a}} \left\{ 1 + \frac{\hat{W} \hat{U} \operatorname{Sech}^2(\hat{W})}{\operatorname{Tanh}(\hat{W})} \right\} \lambda_0(t, \hat{a}, \hat{b}),$$

$$\frac{\partial \lambda_0(t, \hat{a}, \hat{b})}{\partial b} = \frac{1}{\hat{b}} \left\{ \hat{b} \ln t + \frac{2\hat{b}^2 t^{\hat{b}} \hat{W} \operatorname{Csch}^2(t^{\hat{b}}) \ln t}{\hat{U}} - \frac{(\hat{W} - \hat{a} t^{\hat{b}} \hat{U} \ln t) \operatorname{Sech}^2(\hat{W})}{\operatorname{Tanh}(\hat{W})} \right\} \lambda_0(t, \hat{a}, \hat{b}),$$

$$\text{where } \hat{W} = \frac{a}{b} [1 - t^{\hat{b}} \operatorname{Coth}(t^{\hat{b}})], \quad \hat{U} = t^{\hat{b}} \operatorname{Csch}^2(t^{\hat{b}}) - \operatorname{Coth}(t^{\hat{b}}).$$

If the covariates $z(\cdot)$ is constant over the time, we have

$$\frac{\partial \hat{S}_z(t)}{\partial \beta_l} = z_l e^{\hat{\beta}^T z} \ln \{ \operatorname{Sech}(\hat{W}) \} \{ \operatorname{Sech}(\hat{W}) \}^{e^{\hat{\beta}^T z}},$$

$$\frac{\partial \hat{S}_z(t)}{\partial a} = -\frac{1}{\hat{a}} \hat{W} e^{\hat{\beta}^T z} \operatorname{Tanh}(\hat{W}) \{ \operatorname{Sech}(\hat{W}) \}^{e^{\hat{\beta}^T z}},$$

$$\frac{\partial \hat{S}_z(t)}{\partial b} = \frac{1}{\hat{b}} (\hat{W} - \hat{a} t^{\hat{b}} \hat{U} \ln t) e^{\hat{\beta}^T z} \operatorname{Tanh}(\hat{W}) \{ \operatorname{Sech}(\hat{W}) \}^{e^{\hat{\beta}^T z}}.$$

6.6 Validity of Hypertabastic PH model using \mathbf{Y}_n^2

Let us consider to test for the hypothesis H_0 that the right censored failure time regression (X_i, δ_i, z_i) under the constant covariates $z_i = (z_{i1}, z_{i2}, \dots, z_{im})^T$ follow the parametrization PH_H model, that is

$$H_0 : \lambda_{z_i}(t) = e^{\beta^T z_i} a t^{b-1} \{ t^b \operatorname{Csch}^2(t^b) - \operatorname{Coth}(t^b) \} \operatorname{Tanh}(W), \quad (3.72)$$

where $\beta = (\beta_1, \beta_2, \dots, \beta_m)^T \in R^m$ is the m -dimensional unknown regression parameters and

$$W = \frac{a}{b} (1 - t^b \operatorname{Coth}(t^b)), \quad a > 0, b > 0.$$

Suppose that $\hat{\boldsymbol{\theta}}_n = (\hat{\beta}^T, \hat{a}, \hat{b})^T$ is the MLE's of $\boldsymbol{\theta} = (\beta^T, a, b)^T$. We give follow the explicit formulas of generalized chi-squared test for H_0 in (3.67). Following the formulas (2.55) to (2.59), the components of \mathbf{Y}_n^2 statistic give as

$$\mathbf{Y}_n^2 = \mathbf{X}_n^2 + \mathbf{Q}, \quad \mathbf{X}_n^2 = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j}, \quad \mathbf{Q} = W^T \mathbf{G}^{-1} W, \quad U_j = \sum_{i: X_i \in I_j} \delta_i,$$

$$\mathbf{G} = \mathbf{I} - \hat{\mathcal{C}} \hat{\mathcal{A}}^{-1} \hat{\mathcal{C}}^T, \quad W = \hat{\mathcal{C}} \hat{\mathcal{A}}^{-1} \mathbf{Z}, \quad \mathbf{Z} = (Z_1, Z_2, \dots, Z_k)^T, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j),$$

where the diagonal matrix $\hat{\mathcal{A}}$ include the diagonal elements $\hat{A}_j = \frac{U_j}{n}$ and e_j is given in (2.63). Using the formulas (2.58) - (2.59), the expressed elements of the matrix $\hat{\mathbf{i}} = [\hat{i}_{ll'}]_{(m+2) \times (m+2)}$ and $\hat{\mathcal{C}} = [\hat{C}_{lj}]_{(m+2) \times k}$ are given as follows

$$\hat{i}_{ll'} = \frac{1}{n} \sum_{i=1}^n \delta_i z_{il} z_{il'},$$

$$\hat{i}_{l(m+1)} = \frac{1}{na} \sum_{i=1}^n \delta_i z_{il} \left\{ 1 + W_i \frac{\text{Sech}^2(W_i)}{\text{Tanh}(W_i)} \right\},$$

$$\hat{i}_{l(m+2)} = \frac{1}{n} \sum_{i=1}^n \delta_i z_{il} \left\{ \ln X_i + \frac{2b X_i^b W_i \text{Csch}^2(X_i^b) \ln X_i}{a U_i} - \frac{(W_i - a X_i^b U_i \ln X_i) \text{Sech}^2(W_i)}{b \text{Tanh}(W_i)} \right\},$$

$$\hat{i}_{(m+1)(m+1)} = \frac{1}{na^2} \sum_{i=1}^n \delta_i \left\{ 1 + W_i \frac{\text{Sech}^2(W_i)}{\text{Tanh}(W_i)} \right\}^2,$$

$$\hat{i}_{(m+2)(m+2)} = \frac{1}{n} \sum_{i=1}^n \delta_i \left\{ \ln X_i + \frac{2b X_i^b W_i \text{Csch}^2(X_i^b) \ln X_i}{a U_i} - \frac{(W_i - a X_i^b U_i \ln X_i) \text{Sech}^2(W_i)}{b \text{Tanh}(W_i)} \right\}^2,$$

$$\hat{i}_{(m+1)(m+2)} = \frac{1}{na} \sum_{i=1}^n \delta_i \left\{ 1 + W_i \frac{\text{Sech}^2(W_i)}{\text{Tanh}(W_i)} \right\} \times$$

$$\left\{ \ln X_i + \frac{2b X_i^b W_i \text{Csch}^2(X_i^b) \ln X_i}{a U_i} - \frac{(W_i - a X_i^b U_i \ln X_i) \text{Sech}^2(W_i)}{b \text{Tanh}(W_i)} \right\},$$

and $\hat{C}_{lj} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i z_{il}, \quad \hat{C}_{(m+1)j} = \frac{1}{na} \sum_{i: X_i \in I_j} \delta_i \left\{ 1 + W_i \frac{\text{Sech}^2(W_i)}{\text{Tanh}(W_i)} \right\},$

$$\hat{C}_{(m+2)j} = \frac{1}{n} \sum_{i: X_i \in I_j} \delta_i \left\{ \ln X_i + \frac{2b X_i^b W_i \text{Csch}^2(X_i^b) \ln X_i}{a U_i} - \frac{(W_i - a X_i^b U_i \ln X_i) \text{Sech}^2(W_i)}{b \text{Tanh}(W_i)} \right\},$$

The choice of grouping cells a_j : For the choice of grouping cell for testing of Hypertabastic PH model, the equations from (2.60) to (2.63) is to estimate the endpoints a_j by using the numerical method because that the inverse of the cumulative hazard function

of the Hypertbastic distribution can't write in an explicit form.

Note here that for testing of the PH_H model, the matrix \mathbf{G} is non-degenerate, so $rank(G) = k$. The limiting distribution of test statistic \mathbf{Y}_n^2 , therefore, is a chi-squared with k degrees of freedom. Thus, the hypothesis H_0 is rejected with an approximate significance level α if $\mathbf{Y}_n^2 > \chi_\alpha^2(k)$.

7 Simulation study

This our part include three sub-part, the first we demonstrate that the MLE of Hypertabastic distribution are \sqrt{n} -consistent in two cases of the data, the second we illustrate the limiting distribution of \mathbf{Y}_n^2 statistic for testing of Hypertabastic distribution in two cased of the data, and other we shall consider the powerful of \mathbf{Y}_n^2 statistic in comparison with other GOF test which we described in sub-section 4.

Follows this, in subsections 7.1 and 7.2, the observations T_i are generated by the numerical solution of the following equation

$$1 - Sech \left\{ \frac{a_0}{b_0} \left(1 - t^{b_0} Coth(t^{b_0}) \right) \right\} = \xi_i, \quad (3.73)$$

where ξ_i are the pseudo-random numbers of uniformly over the interval $[0, 1]$.

If the data are right censored, then the censoring times C_i are taken independently of T_i , belongs uniformly distributed random variables $U(0, c)$, where c is the positive constant which can be chosen dependent on p as follows: If we fix the level of censoring p ($0 < p < 1$) as well

$$p = \mathbf{P}\{C_i \leq T_i\} = \mathbf{P}\{C_i \leq c \leq T_i\} \quad \text{then} \quad p = \frac{1}{c} \int_0^c S(u, \boldsymbol{\theta}_0) du, \quad (3.74)$$

where $S(\cdot, \boldsymbol{\theta}_0)$ is the Hypertabastic survival function of T_i .

7.1 Hypertabastic's parameters Estimation

Since the problem of finding the maximization of $\ell(\theta)$ are usually difficult, so we investigated MLE's $\hat{\theta}_n$ of parameters θ by Monte Carlo simulation. To do this task, corresponding with the true values of the parameters $\theta_0 = (a_0, b_0)^T$, the samples of pseudo-random T_i from the Hypertabastic distribution (3.1).

An example of simulated MLE for true values of the parameters $\theta_0 = (1.0, 0.5)^T$. The simulated average absolute errors of MLE $\hat{\theta}_n = (\hat{a}_n, \hat{b}_n)^T$ versus their true value can

Table 3.2: Simulated average absolute errors of MLE $\hat{\theta}_n$ versus their true value $\theta_0 = (1.0, 0.5)^T$ when the right censoring probability p take values 0 (complete data) and 20%.

Sample size	$\simeq n^{-0.5}$	\hat{a}_n		\hat{b}_n	
		$p = 0\%$	$p = 20\%$	$p = 0\%$	$p = 20\%$
100	0.100	0.055	0.061	0.035	0.039
150	0.081	0.045	0.050	0.028	0.032
200	0.070	0.039	0.042	0.024	0.027
250	0.063	0.034	0.039	0.022	0.024
300	0.057	0.032	0.035	0.019	0.022
350	0.053	0.029	0.032	0.018	0.020
400	0.050	0.027	0.030	0.017	0.019
450	0.047	0.026	0.028	0.016	0.017
500	0.044	0.024	0.027	0.015	0.017

be find by using the *maxLik* package in R statistical software with simulated samples $n = 100, 150, 200, 250, 300, 350, 400, 450$ and 500.

The number of runs for every n was $N = 10000$. For each run, we obtain the corresponding MLE of each parameter. An each simulate sample n , the average absolute of errors are the total of absolute different of the MLE's against the true value on the number of runs N .

Table 3.2 shown that all simulated MLE converge faster than $n^{-0.5}$ and, hence confirm the theorem established by Fisher [44].

7.2 Simulated distribution of \mathbf{Y}_n^2 statistic for $H(a, b)$

Figure 3.4 illustrate the simulated distribution of generalized chi-squared \mathbf{Y}_n^2 statistic of the null hypothesis (3.1) recovering from difference initial value of parameters (a, b) for the sample size $n = 200$ and $r = 13$ intervals, versus the chi-squared distribution with $r - 1 = 12$ degree of freedom, the number of runs for every n was $N = 5000$.

As can see from Figure 3.4, one can observe that the statistical distribution of \mathbf{Y}_n^2 in the limit follows a chi-squared with $r - 1$ degrees of freedom within the statistical errors of simulation. We obtain the same results for different value of parameters and different number of equiprobable grouping intervals. It is means that the limiting distribution of the generalized chi-squared \mathbf{Y}_n^2 statistic is distribution free. This result also established the Theorem 1.5 of Nikulin.

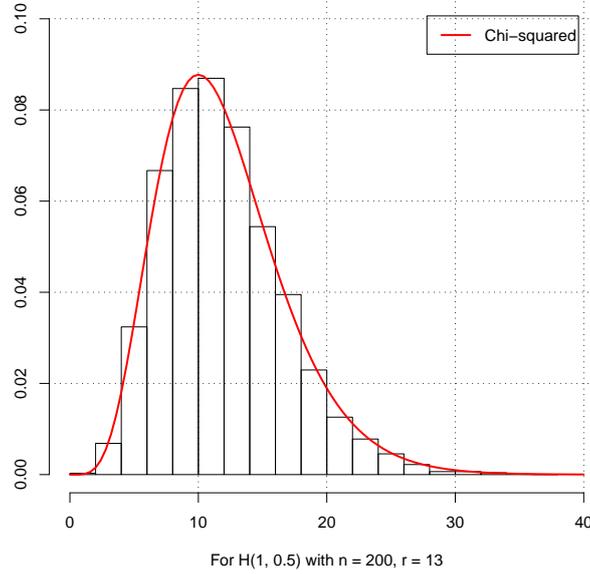


Figure 3.4: Simulated distribution of the \mathbf{Y}_n^2 statistic ($n = 200, r = 13$) under the null hypothesis $H(a, b)$ recovering difference of parameters versus the chi-squared distribution with 12 degree of freedom.

7.3 Power considerations

In this sub-section, we illustrate the flexibility of Hypertabastic distribution and the performance of \mathbf{Y}_n^2 statistic for considering two cases of the data included the completely and right censoring observations. The comparison distributions which we use in this sub-section include the following distributions.

1. Log-normal distribution ($LN(\theta, \nu)$)

$$H_1 : T_i \simeq F_{LN}(t, \theta, \nu) = \Phi\left(\frac{\ln t - \nu}{\theta}\right), \quad t > 0, -\infty < \nu < +\infty, \theta > 0, \quad (3.75)$$

where $\Phi(\cdot)$ is the distribution function of a normal random variable.

2. Log-logistic distribution ($LL(\theta, \nu)$)

$$H_1 : T_i \simeq F_{LL}(t, \theta, \nu) = 1 - \left\{1 + \left(\frac{t}{\theta}\right)^\nu\right\}^{-1}, \quad t > 0, \theta > 0, \nu > 0. \quad (3.76)$$

3. Birnbaum-Saunders distribution ($BS(\theta, \nu)$)

$$H_1 : T_i \simeq F_{BS}(t, \theta, \nu) = \Phi\left\{\frac{1}{\nu}\left[\left(\frac{t}{\theta}\right)^{\frac{1}{2}} - \left(\frac{\theta}{t}\right)^{\frac{1}{2}}\right]\right\}, \quad t > 0, \theta > 0, \nu > 0. \quad (3.77)$$

4. Inverse Gaussian distribution ($IG(\theta, \nu)$), set $\theta > 0$, $\nu > 0$, with $t > 0$ then

$$H_1 : T_i \simeq F_{IG}(t, \theta, \nu) = \Phi\left\{\sqrt{\nu}\left[\left(\frac{t}{\theta}\right)^{\frac{1}{2}} - \left(\frac{\theta}{t}\right)^{\frac{1}{2}}\right]\right\} + e^{2\nu}\Phi\left\{-\sqrt{\nu}\left[\left(\frac{t}{\theta}\right)^{\frac{1}{2}} + \left(\frac{\theta}{t}\right)^{\frac{1}{2}}\right]\right\}. \quad (3.78)$$

Complete data

For evaluating the power of the considered test statistics, we analyze a simulation result in which estimate the powerful of \mathbf{Y}_n^2 , ω_n^2 and \mathbf{A}_n tests for Hypertabastic ($H(a, b)$) against $LN(2, 0.5)$, $LL(2, 3)$, $BS(0.5, 2.5)$ and $IG(0.5, 3.5)$ distributions.

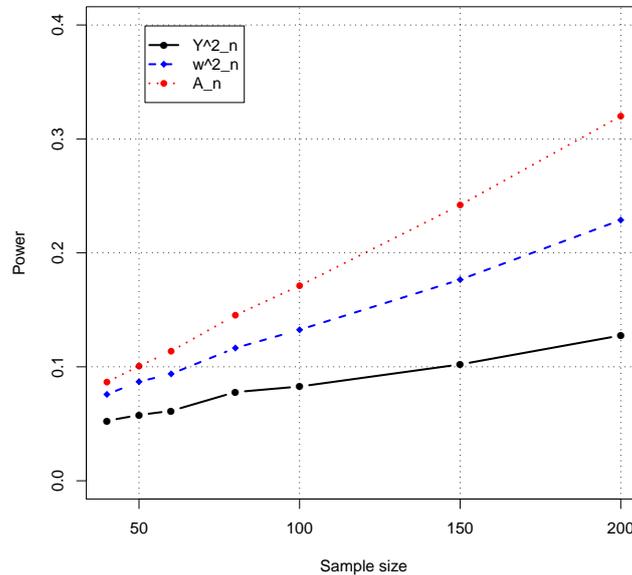


Figure 3.5: Power of tests for $H(a, b)$ against $LN(2, 0.5)$ alternative.

A Monte Carlo procedure is employed with $N = 10000$ times of run for each sample size $n = 40, 50, 60, 80, 100, 150$ and 200 at significance level $\alpha = 0.05$. Note that the test based on \mathbf{Y}_n^2 statistic is constructed by using the equi-probable intervals with $r = 4, 4, 5, 6, 8, 11$ and 13 corresponding with the n sample size.

Figure 3.5 - 3.8 illustrate the power of three consideration tests for $H(a, b)$ against $LN(2, 0.5)$, $LL(2, 3)$, $BS(0.5, 2.5)$ and $IG(0.5, 3.5)$, respectively.

The power of three tests considered indicate that the modified Anderson-Darling GOF test bring out the highest power among other modified in all of alternative distributions. On the contrary, the \mathbf{Y}_n^2 test for equiprobable grouping cells is generally lower powerful. However, one note that the power of \mathbf{Y}_n^2 test will be improved using the class of Neyman-

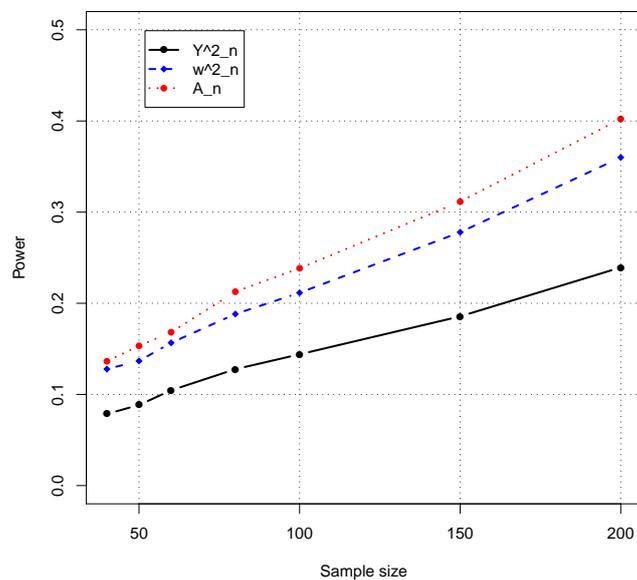


Figure 3.6: Power of tests for $H(a, b)$ against $LL(2, 3)$ alternative.

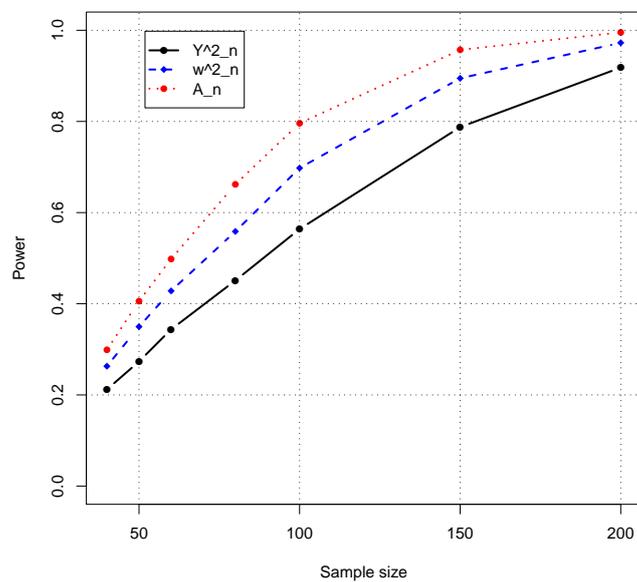


Figure 3.7: Power of tests for $H(a, b)$ against $BS(0.5, 2.5)$ alternative.

Pearson to group intervals. This is given the references of Voinov et al. [120], Nikulin et al. [90] on testing of normality against composite alternative hypotheses. Power studies also

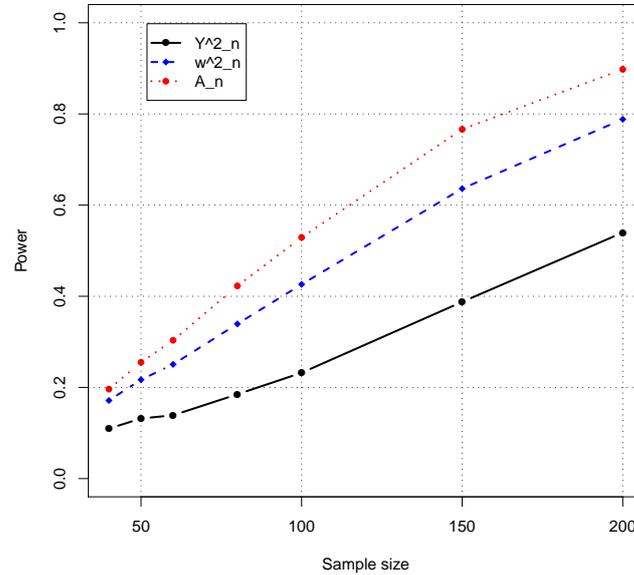


Figure 3.8: Power of tests for $H(a, b)$ against $IG(0.5, 3.5)$ alternative.

shown that the Hypertabastic distribution is the best flexibility, can be suitable with the sample belong to the Log-logistic and Log-normal distributions, this distribution also can be use as a hypothesis to oppose the Birnbaum-Saunders and Inverse Gaussian distributions in statistical analysis.

Although the powerful of Anderson-Darling (\mathbf{A}_n) and Cramer-von Mises-Smirnov (ω_n^2) statistics are higher than \mathbf{Y}_n^2 and these statistics may be used for Hypertabastic distribution when the data are censoring times. Follows that, the KM survival estimates can be used instead of the E.C.D.F. $F_n(t)$ in the formulas (3.13) and (3.14). In this circumstance, we need to estimate the asymptotic of the limiting distributions for each of these tests. However, by simulation study, we realize that the limiting distributions of these test statistics depend totally the MLE of Hypertabastic distribution and we cannot fit the observed statistics for any other distributions. Therefore, we do not recommend using the \mathbf{A}_n and ω_n^2 statistics to fit the Hypertabastic distribution for the right censored data.

Right censored data

We illustrate here the flexibility of Hypertabastic in comparison with other failure time distributions using \mathbf{Y}_n^2 statistic. We consider the case of independently censored samples with censoring proportion approximately $p = 20\%$. Using the Monte-Carlo simulation with

the number of runs $N = 1000$ for each considered hypothesis, the failure time T which are generated from 7 different combination parameters $\boldsymbol{\theta}_0$, are given by the formula

$$T = F^{-1}(\xi, \boldsymbol{\theta}_0), \quad (3.79)$$

where $F^{-1}(\cdot, \boldsymbol{\theta}_0)$ is the inverse function of the H , LN , LL , BS and IG alternative distributions, ξ are the pseudo-random numbers of uniformly over the interval $[0, 1]$ for the sample size of $n = 200$.

The censoring times C are taken independently of T , belongs uniformly distributed random variables $U(0, c)$, where c is the positive constant which can be chosen dependent on p as follows: If we fix the level of censoring p ($0 < p < 1$) as well

$$p = \mathbf{P}\{C \leq T\} = \mathbf{P}\{C \leq c \leq T\} \quad \text{then} \quad p = \frac{1}{c} \int_0^c S(u, \boldsymbol{\theta}_0) du, \quad (3.80)$$

where $S(\cdot, \boldsymbol{\theta}_0)$ is the survival function of T .

The number of grouping intervals is $k = 8$, we obtain the value of \mathbf{Y}_n^2 test for the each hypothesis comprising it with the critical value $\chi_8^2(0.05) = 15.50731$. So, the simulated powerful of \mathbf{Y}_n^2 which counts the number of times where $\mathbf{Y}_n^2 > \chi_8^2(0.05)$ divided by N , are given in following Table 3.3.

From simulated results, we see that the powerful of \mathbf{Y}_n^2 is lowest for Hypertabastic when the failure time data generate from Log-logistic in four out of seven combination parameters. The performance of \mathbf{Y}_n^2 for Birnbaum-Saunders and Inverse Gaussian are highest. So that, two distributions does not fit the data which are generated form Log-logistic distribution.

When the data generate from Log-normal alternative distribution, the Hypertabastic fit the data with highest precision, surpass Log-logistic in performance in all instances. In addition, the Birnbaum-Saunders and Inverse Gaussian distributions again present to offer a opposition with Log-normal alternative distribution.

As follows Table 3.3, we also see that the Hypertabastic again expressed higher opposition when the data are generated from the Birnbaum-Saunders and Inverse Gaussian alternative hypotheses. That is because the powerful of \mathbf{Y}_n^2 for this distribution are highest performance than other distributions. This was reaffirmed when the data are generated from the Hypertabastic alternative, the powerful of \mathbf{Y}_n^2 for Birnbaum-Saunders and Inverse Gaussian distributions are higher than other two distribution. So, we can be give a statistical inference that the Hypertabastic distribution can be in concord with the failure time data which belong to the Log-logistic and Log-normal distributions, Hypertabastic

Table 3.3: Power of \mathbf{Y}_n^2 test for LL , LN , BS , IG and H distributions against alternative hypothesis when the data are the right censored and MLE used ($n = 200$, $k = 8$).

Log-logistic alternative $LL(\theta, \nu)$							
	(2, 3)	(3, 2)	(2, 2)	(3, 5)	(0.25, 1.5)	(5, 3)	(3, 4)
H	0.201	0.137	0.169	0.380	0.223	0.228	0.224
LN	0.228	0.194	0.195	0.270	0.171	0.225	0.243
BS	0.809	0.944	0.930	0.684	0.976	0.809	0.734
IG	0.813	0.959	0.965	0.493	0.999	0.798	0.571

Log-normal alternative $LN(\theta, \nu)$							
	(0.75, 0.5)	(0.5, 0.75)	(0.75, 1.5)	(1, 1)	(2, 1)	(1, 0.5)	(0.5, 2)
H	0.160	0.170	0.147	0.178	0.153	0.184	0.139
LL	0.216	0.260	0.254	0.246	0.271	0.251	0.289
BS	0.466	1.000	0.962	0.802	0.824	0.451	1.000
IG	0.268	1.000	1.000	0.942	0.946	0.259	1.000

Birnbaum-Saunders alternative $BS(\theta, \nu)$							
	(2.5, 0.5)	(2, 3)	(0.5, 3)	(3, 2)	(1, 1)	(0.5, 1)	(2, 2.5)
H	0.496	0.999	0.993	0.994	0.815	0.770	0.998
LL	0.322	0.998	1.000	0.978	0.595	0.582	0.998
LN	0.155	0.998	0.993	0.901	0.350	0.312	0.982
IG	0.318	1.000	1.000	0.996	0.418	0.410	1.000

Inverse Gaussian alternative $IG(\theta, \nu)$							
	(0.5, 3.5)	(1, 2.5)	(3, 0.5)	(3, 2)	(1, 1)	(0.5, 0.5)	(2.5, 1)
H	0.894	0.769	0.827	0.586	0.908	0.963	0.774
LL	0.894	0.659	0.382	0.384	0.694	0.884	0.405
LN	0.737	0.447	0.204	0.191	0.457	0.724	0.208
BS	0.508	0.297	0.312	0.303	0.316	0.466	0.306

Hypertabastic alternative $H(a, b)$							
	(0.5, 0.5)	(0.25, 1)	(0.5, 1.5)	(1, 1)	(0.5, 1)	(1, 0.5)	(2, 0.75)
LL	0.103	0.164	0.166	0.129	0.142	0.130	0.180
LN	0.163	0.141	0.117	0.160	0.129	0.230	0.279
BS	0.917	0.697	0.627	0.790	0.688	0.939	0.936
IG	0.941	0.719	0.622	0.793	0.699	0.951	0.940

distribution also can be use as a hypothesis to oppose the Birnbaum-Saunders and Inverse Gaussian distributions in statistical analysis.

Chapter 4

Some numerical applications

1 Introduction

Parametric survival models are frequently being used in the analysis of survival and reliability and they are also consistent with many different real-life dataset. However, sometime the accepted model for some specific data is not the best model because some other models can be very close. In these circumstances, it is better to compare it with other models having similar properties.

In chapter 3, we have been described our approach to validate the Hypertabastic survival models using generalized chi-squared test statistic \mathbf{Y}_n^2 for three scenes of data: complete, right censored and right censored with covariates. We also designed the simulation studies to evaluate the flexibility of the Hypertabastic distribution showing the performance of the generalized chi-squared test statistic \mathbf{Y}_n^2 .

We therefore consider in this chapter the application of the Hypertabastic survival models for the some real-life dataset that are described in section 4 of General Introduction by using \mathbf{Y}_n^2 statistic. These datasets are already studied but we recommend here the different models from previously proposed.

For the datasets do not affected by the covariates, using \mathbf{Y}_n^2 , we demonstrate again the flexibility and the validity of the Hypertabastic distribution that can be use as the alternative distribution (see subsection 2.1).

For the data governed by one or more covariates, we focus explore the profits of Hypertabastic survival models in comparison with other parametric survival models by applying them to these datasets. The comparison models which we use here include Weibull (W), Log-normal (LN), Log-logistic (LL), Birnbaum-Saunders (BS) and Inverse Gaussian (IG) models. Specifically, the first dataset we focus on evaluating the role of AFT_H model by

applying it to the data set from the remission duration of acute leukemia patients with one influence covariate, see subsection 2.2. The second dataset, we present a methodology for covariate selection as follow the approach of Collett [33]. Later, we consider the PH_H model by fitting for the survival times of malignant glioma patients in the brain tumour study and focus on exploring the role of the most significance covariate in combination with remaining covariates, see subsection 2.3. We also give the estimation of the confidence interval (CI) for the regression parameters and the CI for the reliability characteristic for each of datasets besides making a most accurate model. The final dataset, we take an analysis to suggest that the parametric models sometime can be inappropriate for the data existing the intersection of survival functions, see subsection 2.4. So that we proposed using other model with cross-effect as the simple cross-effect (SCE) model to analyze this dataset.

2 Statistical analysis and inferences

2.1 Analysis of the quantities of the oil field

We take the data that are described in Example 1-section 4 about the quantities of the oil field. Bagdonavičius et al. [19] proposed using the Log-normal for the quantities of the oil. For this example, our aim is looking for accurate distribution for this dataset based on the \mathbf{Y}_n^2 statistic. The class of Neyman-Pearson (2NPG) are used here to help us get a best decision for selected distribution.

Indeed, we first obtain the MLE's of Hypertabastic's parameters: $\hat{a}_n = 0.3259$ and $\hat{b}_n = 0.5400$. Using the equi-probable grouping cell for $r = 6$ intervals, the values of endpoints a_i and the vector frequency ν_i are presented in the following Table.

i	0	1	2	3	4	5	6
a_i	0	3.445	5.716	8.470	12.530	20.459	$+\infty$
ν_i		9	4	13	5	11	7

The Fisher's information matrix \mathbf{I} of T_i and the information matrix \mathbf{J} of ν_i are

$$\mathbf{I} = \begin{pmatrix} 19.493 & 9.509 \\ 9.509 & 7.087 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 12.585 & 4.269 \\ 4.269 & 2.501 \end{pmatrix}.$$

The statistic $\mathbf{X}_n^2 = 7.449$, the quadratic form $\mathbf{Q} = 1.149$. So, the value of the statistic $\mathbf{Y}_n^2 = 8.598$. The asymptotic p -value of the test based on \mathbf{Y}_n^2 and the critical value $C_{0.05}$

at significance level $\alpha = 0.05$ are

$$p_{0.05} = \mathbf{P}\{\chi_5^2(0.05) > 8.598\} = 0.1261792, \quad C_{0.05} = 11.07050.$$

We see that the value of statistic \mathbf{Y}_n^2 is less than the critical value C_α . On the other hand, using the Cramer-von Mises- Smirnov or Anderson-Darling statistic, we also obtained the value of statistics are less than their critical value: $\omega_n^2 = 0.054987 < C_{0.05}^{cvm} = 0.11732$ and $\mathbf{A}_n = 0.41141 < C_{0.05}^{ad} = 0.72439$, respectively. So, all consideration statistics shown that we are not reasonable to refuse the hypothesis H_0 that the distribution of the quantities T_i is appropriate to the Hypertabastic distribution.

Furthermore, we see that this dataset also is in concord for all of consideration hypotheses, including LN, LL, BS and IG distributions, based on the value of \mathbf{Y}_n^2 . Table 4.1 shows the statistical inferences for $LN(\theta, \nu)$, $LL(\theta, \nu)$, $BS(\theta, \nu)$ and $IG(\theta, \nu)$ distributions.

Table 4.1: Statistical inference for null hypotheses that the quantities of oil data belong LN , LL , BS , IG distributions with $r = 6$ intervals.

	LN		LL		BS		IG	
	MLE	Std.err	MLE	Std.err	MLE	Std.err	MLE	Std.err
Shape (ν)	2.102	0.138	8.728	1.187	1.086	0.109	0.654	0.180
Scale (θ)	0.967	0.097	1.830	0.218	7.406	0.989	12.018	2.246
-2ℓ	341.898		342.230		184.695		347.079	
\mathbf{Y}_n^2	5.630		10.681		7.143		7.212	
p -value	0.343		0.058		0.210		0.205	

So, we continue analyse and compute the value of the statistic \mathbf{Y}_n^2 for Hypertabastic distribution against alternative hypotheses. Figure 4.1 shows the curve of P.D.F. of $H(a, b)$, $LN(\theta, \nu)$, $LL(\theta, \nu)$, $BS(\theta, \nu)$ and $IG(\theta, \nu)$ when the MLE usage. It exists the intersection of $H(a, b)$ density with other \cap -shape densities distribution functions.

So, we have a reason to use the 2NPG class for testing the hypothesis $H(a, b)$ distribution against $LN(\theta, \nu)$, $LL(\theta, \nu)$, $BS(\theta, \nu)$ and $IG(\theta, \nu)$ alternatives for this dataset. We proceed the \mathbf{Y}_n^2 using the 2NPG class for each pair of hypotheses as follows.

1. For a pair of hypotheses: $H(a, b)$ in (3.4) versus $LN(\theta, \nu)$ alternative (3.69): Solving the equation of two densities, we obtain three points of solutions: $a_1 = 0.79467$, $a_2 = 4.91036$, $a_3 = 31.52838$. Consider the following 2NPG class (see, Figure 4.1)

$$\mathbf{I}_1 = (0, 0.79467] \cup (4.91036, 31.52838],$$

$$\mathbf{I}_2 = (0.79467, 4.91036] \cup (31.52838, +\infty).$$

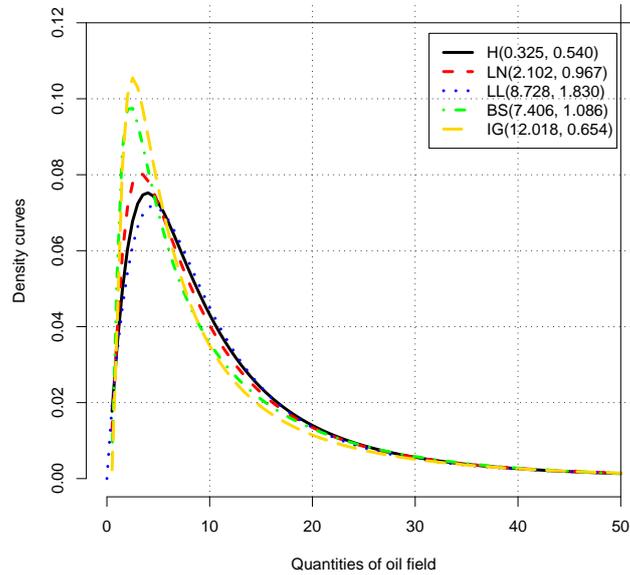


Figure 4.1: Density curves of consideration distributions for quantities of oil when MLE usage.

2. For the test of H_0 in (3.4) against $LL(\theta, \nu)$ alternative in (3.70). We obtained four points of intersection of densities: $a_1 = 0.73541$, $a_2 = 5.65267$, $a_3 = 15.81745$, $a_4 = 46.20444$ (see, Figure 4.1) and using the 2NPG class as

$$\begin{aligned} \mathbf{I}_1 &= (0.73541, 5.65267] \cup (15.81745, 46.20444], \\ \mathbf{I}_2 &= (0, 0.73541] \cup (5.65267, 15.81745] \cup (46.20444, +\infty), \end{aligned}$$

3. Similarly, for the test of the null hypothesis $H_0 : H(a, b)$ versus alternative $H_1 : BS(\theta, \nu)$. We consider the following 2NPG class

$$\begin{aligned} \mathbf{I}_1 &= (0, 0.720452] \cup (28.49408, 54.15587], \\ \mathbf{I}_2 &= (0.720452, 28.49408] \cup (54.15587, +\infty), \end{aligned}$$

where $a_1 = 0.720452$, $a_2 = 28.49408$, $a_3 = 54.15587$ are the points intersections of two densities of $H(a, b)$ and $BS(\theta, \nu)$ (see, Figure 4.1).

4. For testing of the null hypothesis $H_0 : H(a, b)$ versus alternative $H_1 : IG(\theta, \nu)$.

We consider the following 2NPG class (see, Figure 4.1)

$$\mathbf{I}_1 = (0, 0.913635] \cup (5.51039, 40.69827] \cup (236.79819, +\infty),$$

$$\mathbf{I}_2 = (0.913635, 5.51039] \cup (40.69827, 236.79819].$$

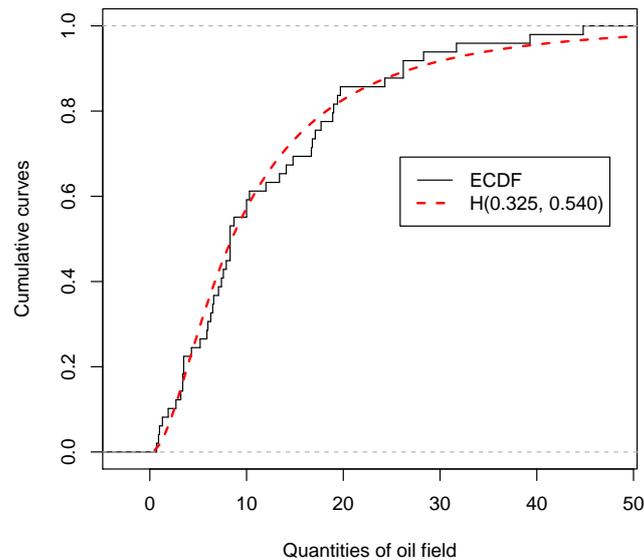


Figure 4.2: Empirical and $H(0.325, 0.540)$ cumulative curves of quantities of oil field data.

The statistical references which we obtain when the 2NPG class used to test the $H(a, b)$ against alternative hypotheses, are given in the Table 4.2.

Table 4.2: Statistical inference for quantities of oil data for testing of $H(a, b)$ against alternative hypotheses when 2NPG used.

	Alternative hypotheses H_1			
	$LN(2.102, 0.967)$	$LL(1.830, 8.728)$	$BS(7.406, 1.086)$	$IG(12.018, 0.654)$
$\nu = (\nu_1, \nu_2)^T$	$(27, 22)^T$	$(35, 14)^T$	$(4, 45)^T$	$(37, 12)^T$
\mathbf{Y}_n^2	0.05022	1.11645	0.00188	3.89851
p -value	0.82268	0.29068	0.96535	0.04832
Inference	Accepte H_0	Accepte H_0	Accepte H_0	Refuse H_0

From Table 4.2, all of statistical inferences give the acceptable for the $H(a, b)$. This indicates that the $H(a, b)$ distribution fits the data much better than all other considered

distributions. Figure 4.2 shown the plot of E.C.D.F. and $H(0.325, 0.540)$ estimates C.D.F. of the quantities of the oil field data.

2.2 Analysis of remission duration from a clinical trial for acute leukemia

Returning the time of remission of acute leukemia patients from a clinical trial that are described in Example 2 - section 3. We see that this dataset consist two covariates: groups of patient and $\log(\text{WBC})$ and we use $z = 1$ if the patient was given 6-MP and $z = 0$ if a placebo.

For this dataset, using likelihood ratio test, we see that the Cox model does not fit for this dataset. This is because that the value of likelihood ratio test is $\ell = 16.4$ with the p -value at 0.05 approximate 5.10^{-5} . Therefore, here, we focus consider a parametric AFT model where the Hypertabastic distribution is used as the baseline hazard function in comparison with other failure time baseline distribution functions. Besides indicating that AFT_H is an accurate model, we shall evaluate the effect of $\log(\text{WBC})$ on the remission of survival of acute leukemia patients. The MLE's of the parameters $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}$ of AFT_H model are

$$\hat{\beta}_0 = 3.428, \hat{\beta}_1 = 0.751, \hat{\beta}_2 = -0.741, \hat{\sigma} = 1.311.$$

Choose $k = 7$ intervals, intermediate results are given in following Table

j	1	2	3	4	5	6	7
a_j	2.979	4.787	6.736	9.084	12.247	17.173	$+\infty$
e_j	4.079	4.079	4.079	4.079	4.079	4.079	4.079
U_j	4	3	5	5	5	4	4
Z_j	-0.012	-0.166	0.142	0.142	0.142	-0.012	-0.012

The statistical inferences for testing of hypothesis H_0 that the acute leukemia patients belong the AFT_H model, are given as follows

$$\mathbf{I} = \begin{pmatrix} 2.413 & 0.889 & 7.327 & -0.071 \\ 0.889 & 0.889 & 2.635 & -0.060 \\ 7.327 & 2.635 & 24.313 & 0.185 \\ -0.071 & -0.060 & 0.185 & 0.842 \end{pmatrix}, \quad \mathbf{W} = (-0.367, -0.293, -0.949, -0.110)^T,$$

$$\mathbf{X}_n^2 = 0.901, \quad \mathbf{Q} = 2.140, \quad \mathbf{Y}_n^2 = 3.041, \quad \text{and } p_{0.05} = \mathbf{P}\{\chi_7^2(0.05) > 3.041\} = 0.881.$$

We see that the values of \mathbf{X}_n^2 and \mathbf{Q} are smaller. So, the value of \mathbf{Y}_n^2 statistic is also smaller. This means that the hypothesis H_0 of AFT_H is well accepted at the significance level 0.05. Figure 4.3 gives the KM and Hypertabastic remission survival estimates of acute Leukemia patients without the covariates.

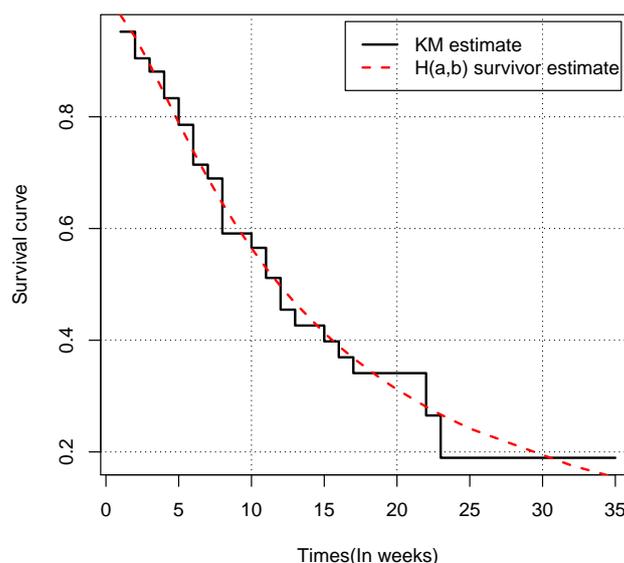


Figure 4.3: KM and Hypertabastic survival estimates plots for remission duration of acute leukemia patients.

Continuously, we make a comparison of the AFT_H model with other parametric AFT models based on the \cap -shape baseline functions, including Log-normal (LN), Log-logistic (LL), Birnbaum-Saunders (BS) and Inverse Gaussian (IG) distributions. We also consider the models without covariates included. Table 4.3 shows the statistical results for each of these comparison models to fit this dataset.

Note that the distribution of \mathbf{Y}_n^2 for all considered models has in limit a chi-squared distribution with $k = 7$ degrees of freedom. So that the critical value of \mathbf{Y}_n^2 test is $C_{0.05} = 14.067$ at significance level 0.05.

From Table 4.3, comparing the value of \mathbf{Y}_n^2 with its critical value $C_{0.05} = 14.067$, we see that this dataset without the covariates are consistent for all consideration models and the Log-normal model without the covariates is fitted this dataset best.

For the considered models contain each of two covariates: $\log(\text{WBC})$ or patient groups, there are three models: AFT_H , AFT_{LN} and AFT_{LL} , are well adapted for this data set. However, the value of \mathbf{Y}_n^2 for AFT_H contain any each of two covariates is lowest for all of

Table 4.3: Comparison models for remission duration of acute leukemia patients ($k = 7$).

Models	With overall variables		With log(WBC) variable		With patient groups		Without variables	
	-2ℓ	\mathbf{Y}_n^2	-2ℓ	\mathbf{Y}_n^2	-2ℓ	\mathbf{Y}_n^2	-2ℓ	\mathbf{Y}_n^2
AFT_H	183.682	3.041	196.240	3.627	213.951	2.408	230.782	4.395
PH_H	181.610	4.327	194.469	4.271	214.118	16.485	230.782	4.395
AFT_{LN}	132.746	7.934	145.796	10.723	158.272	3.063	175.649	1.767
AFT_{LL}	184.662	6.085	196.991	3.975	215.322	3.958	230.702	3.983
AFT_{BS}	93.742	17.195	108.713	14.136	115.133	14.197	135.357	8.669
AFT_{IG}	136.679	27.957	152.343	20.223	158.191	15.724	178.162	8.483

models considered. We also see that the AFT_{BS} and AFT_{IG} models containing log(WBC) variable are fitted badly for this dataset. Because of the values of \mathbf{Y}_n^2 for two these models are higher than its critical value, i.e. the hypothesis that this dataset belong one of two models AFT_{BS} or AFT_{IG} on adding log(WBC) variable is rejected, although the value of -2ℓ for two models are lowest. This indicate that the log(WBC) has the opposite effect in the AFT_{BS} and AFT_{IG} models. From Table 4.3, we suggest using AFT_H containing one out of two variables, log(WBC) or patient groups variables to fit this dataset.

Table 4.4: 95% CI for the remission duration of acute leukemia patients based on AFT_H model.

	Mean \hat{S}	Mean \hat{S}_{AFT_H}	Std.err	lower 95% CI	upper 95% CI
All of patients	0.579	0.603	0.034	0.567	0.637
Group 1: 6-MP	0.692	0.716	0.034	0.680	0.750
Group 2: Placebo	0.464	0.490	0.034	0.455	0.524

For overall model, we consider this dataset on adding a combination of two variables: groups of patient and log(WBC). In this circumstance, we see that the -2ℓ values of AFT_{BS} and AFT_{IG} models are lowest but on the contrary their values of \mathbf{Y}_n^2 are highest for all of consideration models. So, we cannot use two models AFT_{BS} and AFT_{IG} to fit this dataset on adding two variables: group of patients and log(WBC). Apart from that, the AFT_H model with a combination of two covariates continue play in optimal role in the assessing adequacy of fit the remission survival of acute leukemia patients, although we see that the parametric PH_H , AFT_{LL} and AFT_{LN} models also fit this dataset well. This is because that the value of \mathbf{Y}_n^2 for AFT_H model contain two variables is lowest for all of consideration models in Table 4.3.

Figure 4.4 shows the KM and AFT_H estimates of survival distribution function in two

groups for each type of treatment. Table 4.4 shows the 95% CI of the time of remissions of acute leukemia patients constructed on AFT_H model for each of two groups.

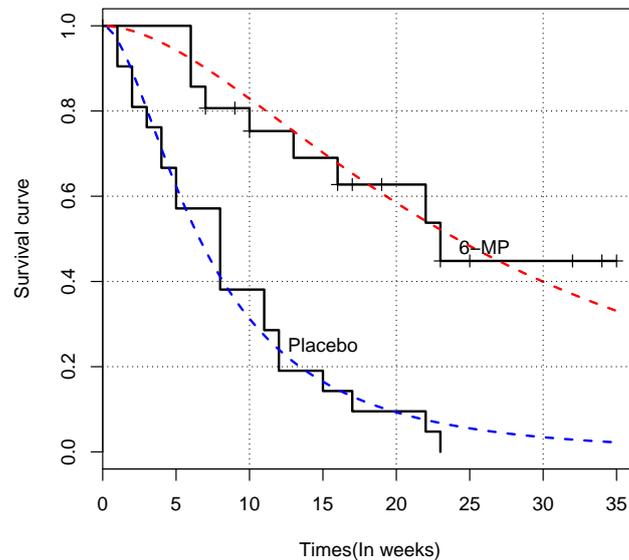


Figure 4.4: KM and AFT_H survival estimates plots for remission duration of acute leukemia patients receiving different groups.

We conclude that the AFT_H is an accurate model for the remission survival of acute leukemia patients from clinical trial study.

2.3 Analysis of brain tumour study

We continue considering the data set from the brain tumour study which discussed in Example 3 section 3. A randomized clinical trial comparing two chemotherapy regimens in 447 patients with malignant glioma cancer are presented by Sauerbrei & Schumacher [105]. The data contains 12 explanatory variables including age, three ordinal and eight binary variables. Three variable are measured on an ordinal scale that are the index of Karnofsky, the type of surgical resection and the grade of malignancy. One of these variables is represented by two dummy variables resulting 16 variables in all. Details for these variables are presented in Table 4.6 below.

We use here the data include 411 patients (274 events, 137 censored times) that can be see in the publications of Sauerbrei & Schumacher [105], Royston & Sauerbrei [103].

Table 4.5: The variables for malignant glioma trial.

Name	Variable	Descriptions	Code/Values
SEX	z_1	Sex	Male = 1, Female = 0
TIME	z_2	Time from first symptoms to clinical confirmation	Short = 1, Long = 2
GRADD1	z_3	Grade of malignancy ¹	Yes = 1, No = 0
GRADD2	z_4		
AGE	z_5	Age	Years
KARD1	z_6	Karnofsky index ¹	Yes = 1, No = 0
KARD2	z_7		
SURGD1	z_8	Type of surgical resection ¹	Yes = 1, No = 0
SURGD2	z_9		
CONVUL	z_{10}	Convulsions	Yes = 1, No = 0
CORT	z_{11}	Cortisone	Yes = 1, No = 0
EPI	z_{12}	Epilepsy	Yes = 1, No = 0
AMNESIA	z_{13}	Amnesia	Yes = 1, No = 0
OPS	z_{14}	Organic psycho-syndrome	Yes = 1, No = 0
APH	z_{15}	Aphasia	Yes = 1, No = 0
TRT	z_{16}	Treatment	Std = 0, New = 1

¹Categorical predictor represented by 2 dummy variables.

Using a bootstrap re-sampling procedure for cox regression, Sauerbrei & Schumacher [105] found five variables (z_5, z_3, z_8, z_6 and z_{12}) which might influence the outcome variable. Royston & Sauerbrei [103] analyzed the survival time of the malignant glioma patients for Cox model. We focus to discuss the validity of PH_H for this dataset using the \mathbf{Y}_n^2 statistic.

As follows the general variable selection strategy outlined by Collett [33], we consider foremost the PH_H model in chapter 3 to find the significant variables using the value of $-2\log$ -likelihood (-2ℓ) and \mathbf{Y}_n^2 statistics on the choice of grouping interval $k = 15$ for all considered models. For PH_H model, the parametrization function is $e^{\beta^T z}$. To take this task, there six steps were carried out to make the choice of the significant variables, are given as follows.

The first step, we fit this dataset for null PH_H model without all of variables. We see that the survival time of the malignant glioma patients are closed to a Hypertabastic distribution. This is because that the value of -2ℓ equal 683.597, the value of NRR statistic $\mathbf{Y}_n^2 = 16.963$ which is lower than the critical value $\chi_{0.05}^2 = 24.995$ at 5% level of significance. We thus expect that the addition of the variables to the model will minimize the value of the considered statistics.

The second step, we continue to fit this dataset for the models that contain each of

Table 4.6: Values of $-2\log$ -likelihood (-2ℓ) and \mathbf{Y}_n^2 for PH_H models fit to the malignant glioma patients.

Steps	Variables in models	-2ℓ	\mathbf{Y}_n^2	$p_{0.05}$ value
1	Non	683.597	16.963	0.321
2	z_1	683.500	18.735	0.224
	z_2	676.834	19.979	0.172
	z_3	649.035	16.012	0.381
	z_4	651.691	14.496	0.488
	z_5	620.059	31.475	0.007
	z_6	672.111	22.804	0.088
	z_7	674.708	17.312	0.300
	z_8	672.156	24.771	0.053
	z_9	681.814	18.248	0.249
	z_{10}	682.645	17.954	0.265
	z_{11}	671.384	19.529	0.190
	z_{12}	670.682	20.498	0.153
	z_{13}	679.508	18.782	0.223
	z_{14}	680.192	16.467	0.351
	z_{15}	680.044	17.237	0.304
	z_{16}	683.500	16.735	0.224
3	$z_5 + z_3$	599.651	21.067	0.134
	$z_5 + z_8$	605.190	25.850	0.039
	$z_3 + z_4$	644.187	15.010	0.450
	$z_6 + z_7$	668.582	22.314	0.099
	$z_8 + z_9$	671.875	25.625	0.042
	$z_5 + z_{14}$	618.860	33.776	0.003
	$z_5 + z_{15}$	620.054	31.872	0.006
4	$z_5 + z_3 + z_8$	571.461	19.111	0.208
	$z_5 + z_3 + z_2$	598.353	19.850	0.177
	$z_5 + z_3 + z_{13}$	598.489	24.320	0.059
	$z_5 + z_6 + z_7$	609.860	24.227	0.061
5	$z_5 + z_3 + z_8 + z_6$	562.535	16.797	0.331
	$z_5 + z_3 + z_8 + z_4$	568.362	17.469	0.291
	$z_5 + z_3 + z_8 + z_{11}$	568.588	21.988	0.108
	$z_5 + z_3 + z_8 + z_{12}$	567.032	21.256	0.128
6	$z_5 + z_3 + z_8 + z_6 + z_{12}$	558.606	17.458	0.292

sixteen variables. Of these variables, the -2ℓ statistic is largest diminished from 683.597 to 620.059 for the model that contain only AGE (z_5) variable. The next variables belong to KARD1 (z_3) and KARD2 (z_4). Besides, there are three variables that be to doubt its significance in the relation with the survival times of the malignant glioma patients. That

are SURGD2 (z_9), OPS (z_{14}) and APH (z_{15}). Since the reduction in -2ℓ on adding z_9 or z_{14} or z_{15} to the null model are 1.783 (nearly 20% at significance level 5%), 3.405 (nearly 6% at significance level 5%) and 3.553 (nearly 5.5% at significance level 5%), respectively. However, we will keep three variables under consideration for inclusion in the model. Apart from that, we eliminated three variable that are SEX (z_1), CONVUL (z_{10}) and TRT (z_{16}) outside the PH_H model. This is because that the reductions of -2ℓ on adding one of three variable are not significantly minimized. So, three variables (z_1 , z_{10} and z_{16}) will not be further considered for inclusion. Also in the second step, we see that the model given two variables: z_5 and z_3 carried out the maximal reduction for -2ℓ when we fit this dataset. Because of the reduction of -2ℓ from 620.059 to 599.651. Therefore, it is necessary to add two variables z_5 and z_3 in the models for inclusion.

In the third step, we fit the data for the models that include z_5 variable in addition with one of three doubt variables (z_9 , z_{14} and z_{15}). We also consider the relation of three double dummy variables that are z_3 with z_4 , z_6 with z_7 and z_8 with z_9 in the effect with the survival of malignant glioma patients. There are three variables (z_9 , z_{14} and z_{15}) that were eliminated in step 2. Because the values of -2ℓ do not well reduced when we fit for the PH_H model on adding each of two variable (z_{14} , z_{15}) with z_5 , and fit for the PH_H model on adding z_9 with z_8 . Specifically, when we fit this dataset for PH_H model on adding variable z_5 with z_{14} or z_{15} then the reduction of -2ℓ are corresponding 1.199 and 0.005 units. Similarly, if we taken the PH_H model contain two variable z_8 and z_9 then the reduction of -2ℓ is 0.281 unit. This is mean that the variable z_8 is contained z_9 . Therefore, it is not necessary to retain three variables (z_9 , z_{14} and z_{15}) in the PH_H model.

The next step, 4th step, is to fit the model that contains z_5 , z_3 and each of seven remaining variables. We see that the effect of losing TIME (z_2) variable did not significantly changed the value of statistical consideration. Indeed, the value of -2ℓ reduce a smaller units when we fit the model that contains z_5 , z_3 and z_2 , on adding the z_2 in the model then the decrease is 1.298 for -2ℓ statistic. So z_2 does not appear to be needed in the model. As the same statistical point of view, two variables, KARD2 (z_7) and AMNESIA (z_{13}) does not keep in the model. Since the reduction of -2ℓ is not significance minimized, it reduced from 611.271 to 609.860 units on adding z_7 in the model contain z_5 and z_6 and from 599.651 to 598.063 units on adding z_{13} in the model contain z_3 and z_5 . End of 4th step, we retained seven variables that are GRADD1 (z_3), GRADD2 (z_4), AGE (z_5), KARD1 (z_6), SURGD1 (z_8), CORT (z_{11}) and EPI (z_{12}) of which there are one variable that on its has some explanatory power is z_4 , which leads to a reduction in -2ℓ that is nearly significant at the 8.9%. The effect of the variables for the survival time of malignant glioma patients from PH_H model are shown in Table 4.6.

The 5th step is to fit the PH_H model includes four of the seven variables that we retained after the third step. We see that the model given combination of four variables (z_3, z_5, z_8 and s_6) leads to the largest reduction in -2ℓ , reducing the value of the statistic from 571.461 to 562.535. This reduction of 8.926 is significant at the 5% ($p = 0.0022$). The reduction in -2ℓ on adding z_4 in the PH_H model that contain a combination z_3, z_5, z_8 variables is 3.098, which is also significant at the 5% level $p = 0.089$. Similarly, if we fit the PH_H model include a combination z_3, z_5, z_8 and z_{11} then the reduction in -2ℓ is 2.873 (nearly 9.3% at significance level 5%). This indicates that we can be eliminated two variable (z_4, z_{11}) get out the model.

The selection variables strategy will continue in the 6th step that we can be eliminated the AMNESIA (z_{13}) variable. Because the value of -2ℓ is not significantly minimized, its reduction is 1.590 when we fit this dataset for PH_H model given by four variables: z_5, z_3, z_8 and z_{13} . In this step, we see that the reduction of -2ℓ is lowest when we fit this dataset for the model on adding five variables: z_5, z_3, z_8, z_6 and z_{12} .

We therefore conclude that the optimal PH_H model that is containing five variables: z_5, z_3, z_8, z_6 and z_{12} for the malignant glioma patients in which the AGE (z_5) is the most important variable.

Table 4.7: Comparison models for malignant glioma data ($k = 15$).

Models	With combination 5 variables				Without variables			
	\mathbf{X}_n^2	\mathbf{Q}	\mathbf{Y}_n^2	$p_{0.05}$	\mathbf{X}_n^2	\mathbf{Q}	\mathbf{Y}_n^2	$p_{0.05}$
PH_H	10.719	6.739	17.458	0.292	16.729	0.234	16.963	0.321
PH_{LN}	30.035	9.890	39.925	$4.6 \cdot 10^{-4}$	47.610	3.292	50.903	$8.10 \cdot 10^{-6}$
PH_{LL}	10.311	7.957	18.269	0.248	19.300	2.814	22.115	0.104
PH_W	17.135	1.498	18.634	0.179	29.714	0.013	29.728	0.008

Using \mathbf{Y}_n^2 statistic with the grouping intervals $k = 15$ for all of consideration models, we make a comparison of the Hypertabastic model with several other parametric models, including Log-Normal, Log-logistic and Weibull models. We note that the limiting distribution of \mathbf{Y}_n^2 statistic for PH_W model including a combination 5 variables has the chi-squared distribution with $k - 1$ degrees of freedom. Because of the rank of matrix \mathbf{G} in formula (2.49) is degenerate for PH_W model. Therefore, p value at 0.05% significance level for PH_W given by

$$\mathbf{P}\{\chi_{14}^2(0.05) > 18.634\} = 0.179.$$

Table 4.7 shows the statistical results of \mathbf{Y}_n^2 statistic for each of these models. In order to see the role of the covariates, we also include the values of these models considered

without variables. Hypertabastic is the most exact of these models, although we see that Log-logistic and Weibull also fit this dataset well.

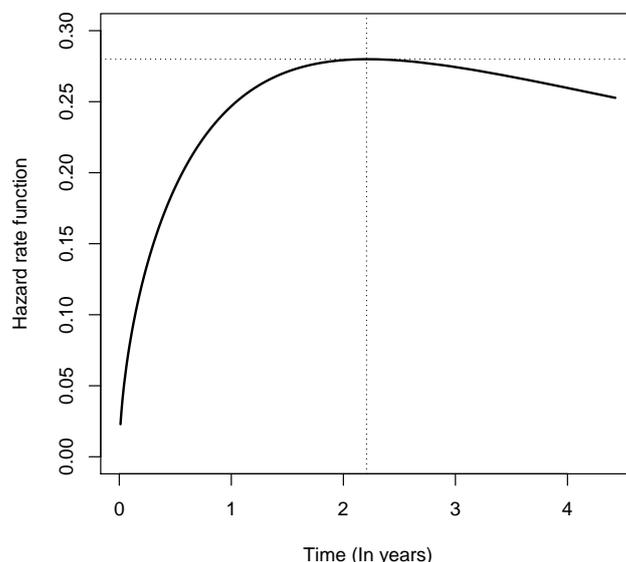


Figure 4.5: Hypertabastic rate of hazard curve for malignant glioma.

As also from Table 4.7, we see that the PH_H model contain 5 variables is an accurate model for this dataset. Using the method of MLE, we obtained in following Table 4.8 the MLE of the parameters of PH_H model.

Table 4.8: MLE of the parameters of PH_H model and 95% CI.

Variables	Parameter	Estimates	Std.err	lower 95% CI	upper 95% CI
AGE (z_5)	β_5	0.038	0.006	0.030	0.045
GRADE1 (z_3)	β_3	0.993	0.193	0.747	1.239
SURGD1 (z_8)	β_8	-1.096	0.193	-1.350	-0.842
KARD1 (z_6)	β_6	-0.377	0.127	-0.540	-0.215
EPI (z_{12})	β_{12}	-0.275	0.142	-0.456	-0.094
	a	0.772	0.168	0.545	0.999
	b	0.397	0.019	0.365	0.429

It is clear that the graph of Hypertabastic hazard rate is \cap -shape for malignant glioma patients. Since the estimates of shape parameter of Hypertabastic distribution when this distribution play the role of baseline hazard function in the PH model is $\hat{b} = 0.397$. Figure 4.5 shows the graph of Hypertabastic hazard for malignant glioma data. The Hypertabastic

hazard function shows that the failure rate reaches its maximum velocity in about 2.21 years.

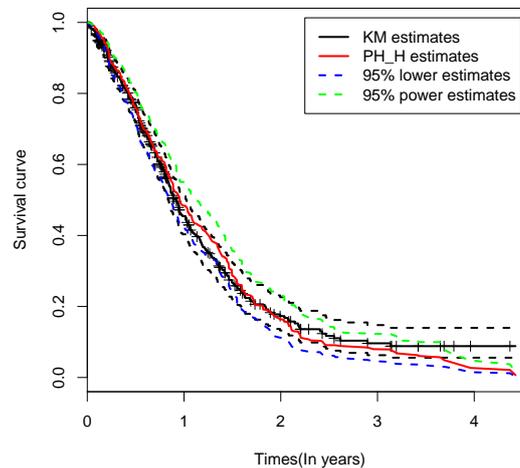


Figure 4.6: KM, PH_H estimates and 95% CI for survival times of malignant glioma patients.

Figure 4.6 shows the graph of KM estimates of survival time of malignant glioma patient in comparison with the PH_H and the 95% CI estimates survival function. We see that the CI of the survival of malignant glioma patients are very closed at the significant level 95%.

2.4 Analysis of multiple myeloma patients

The patients of multiply myeloma that were carried out by The Haematology Center, in the Main Military Clinical Hospital named after N.N.Burdenko, to compare the response time to the treatment in two groups of patients. The difference in these groups is in the fact that the first group received the chemotherapy together Bortezomibe, which is marketed as Velcade by Millennium Pharmaceuticals, other received drugs without Bortezomibe.

The data include 60 patients with 56 observations and 4 censored patients, of two treatment groups: chemotherapy without Bortezomibe ($z_1 = 0$) and chemotherapy together with Bortezomibe ($z_1 = 1$). In addition to variable of treatment, three such variables are given in the data that consider in their work, namely: Type of response (the value $z_2 = 0$ corresponds to the progression of the disease, $z_2 = 1$ - the general response); sex ($z_3 = 1$ - male, $z_3 = 0$ -female); and age in years (z_4). Detail of the data are presented in Semenova and Bitukov [106].

In Example 4 section 4, we illustrated the KM estimates of survival function for multiple myeloma patients receiving two treatment by Figure 3. As seen that, we believe that the Cox model can be inappropriate for this dataset.

We therefore propose considering the parametric survival models for relating the distribution of response time to the scheme of chemotherapy, type of the response, age, sex and age. The class of baseline parametric consider included Hypertabastic, Weibull, Log-normal, Log-logistic, Birnbaum-Saunders and Inverse Gaussian distributions.

For $k = 7$ interval, using \mathbf{Y}_n^2 and $-2\log$ -likelihood (-2ℓ) statistics for all fitted models, we first consider the GOF test for the models without covariates. Later, to evaluate the effect of type of chemotherapy covariate on the response times of multiple myeloma patients, we consider the models on adding type of chemotherapy covariate. For overall models, the models on adding a combination of all covariates are fitted for this dataset. Table 4.9 illustrates the statistical results of \mathbf{Y}_n^2 and -2ℓ for all fitted models.

Table 4.9: Fitted models for the multiple myeloma patients ($k = 7$).

Models	With		With		Without	
	all covariates		z_1 covariate		covariates	
	-2ℓ	\mathbf{Y}_n^2	-2ℓ	\mathbf{Y}_n^2	-2ℓ	\mathbf{Y}_n^2
AFT_W	191.183	13.296	204.292	15.295	508.635	26.514
AFT_H	481.397	36.208	501.072	24.719	501.839	7.373
AFT_{LN}	375.485	18.669	395.964	10.468	397.361	12.413
AFT_{LL}	477.052	16.904	501.761	10.354	502.606	13.709
AFT_{BS}	302.643	18.671	315.557	8.493	318.489	11.443
AFT_{IG}	377.045	19.481	396.755	21.110	397.680	14.186

From Table 4.9, for all of considered models without covariates, this dataset is consistent for the Hypertabastic, Log-normal, Log-logistic and Birnbaum-Saunders distributions at significance level $\alpha = 0.05$. Because their values of \mathbf{Y}_n^2 statistic are lower than their critical value $C_\alpha = \chi_{0.05}^2(7) = 14.067$. This dataset also is clearly rejected for Weibull distribution without covariates, because of its value of $\mathbf{Y}_n^2 = 26.514$ is higher than its critical value $C'_\alpha = \chi_{0.05}^2(6) = 12.591$. Figure 4.7 shown the KM and good fitted models for the response times of multiply myeloma patients without covariates.

Continuously, for the models contain the type of chemotherapy covariate (z_1), we see that three models, including AFT_W , AFT_H and AFT_{IG} are rejected. And this dataset responded to be the acceptable for AFT_{LN} , AFT_{LL} and AFT_{BS} on adding type of response covaraite. So, if we need find an accurate model contain only one covariate to fit this dataset on adding type of chemotherapy (z_1), then we proposed using AFT_{BS} model.

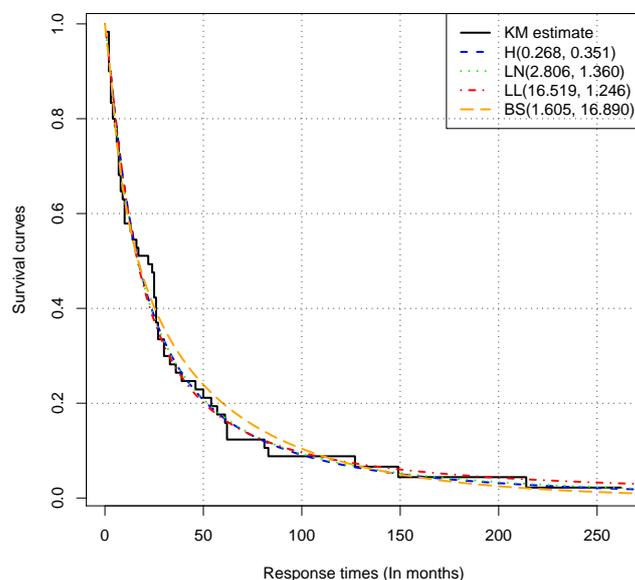


Figure 4.7: KM and fitted models without covariates for reponse times of multiple myeloma.

However, a model which we really need to be a prognostic model for overall that can be fitted best this dataset with a combination of all covariates. Table 4.10 shows the values of \mathbf{Y}_n^2 and their $p_{0.05}$ value at the significance level 0.05 for the overall considered models with the different grouping cells k . From Table 4.10, the value of \mathbf{Y}_n^2 for overall fitted models are greater than their critical values, it follows that all considered models are rejected.

Table 4.10: Values of \mathbf{Y}_n^2 with $k = 8, 9, 10, 11$ intervals when we fit the multiply myeloma data for overall models that contain all covariates.

Overall models	Number of intervals							
	$k = 8$		$k = 9$		$k = 10$		$k = 11$	
	\mathbf{Y}_n^2	$p_{0.05}$	\mathbf{Y}_n^2	$p_{0.05}$	\mathbf{Y}_n^2	$p_{0.05}$	\mathbf{Y}_n^2	$p_{0.05}$
AFT_W	8.662	0.277	21.501	0.005	38.890	1.10^{-5}	40.663	1.10^{-5}
AFT_H	45.851	2.10^{-7}	38.312	1.10^{-5}	82.481	1.10^{-13}	72.496	4.10^{-11}
AFT_{LN}	40.999	2.10^{-6}	23.548	5.10^{-3}	48.878	4.10^{-7}	43.664	8.10^{-6}
AFT_{LL}	45.658	2.10^{-7}	21.199	0.011	31.353	5.10^{-4}	42.666	1.10^{-5}
AFT_{BS}	36.208	2.10^{-5}	47.513	3.10^{-7}	58.178	9.10^{-9}	75.963	1.10^{-11}
AFT_{IG}	38.728	5.10^{-6}	47.567	3.10^{-7}	49.183	3.10^{-7}	50.498	5.10^{-7}

Note that we received almost same of the statistical inferences for overall fitted models when we fit this dataset using \mathbf{Y}_n^2 with the different of the number of intervals. Only the

value of \mathbf{Y}_n^2 with $k = 8$ intervals is lower its critical value $C_{0.05} = \chi_{0.05}^2(7) = 14.067$ when we fit this dataset for AFT_W on adding all covariates. So, we can temporarily conclude that the hypothesis is accepted for the AFT_W on adding all covariates. However, we recognize that we will continue to propose an another overall model that fit this dataset better than AFT_W model, such as the simple cross-effect model which was proposed by Bagdonavičius & Nikulin [14]. Because the survival times of two patient groups have the cross-effect. So, we hope that we, in the future, will to cooperate with our colleagues to realize this proposition.

Conclusion and Perspectives

Although a type of chi-squared statistic \mathbf{Y}_n^2 that became famous and commonly used as soon after it was proposed by Nikulin-Rao-Robson in 1973-1974. Since then there also have been many discussed on this statistic in literature. However, the most of the textbooks that are more technical in nature to use it as a tool of goodness-of-fit test procedures. In our works, besides proving that the use of this statistic are perfectly reasonable for the almost kind of survival data analysis: complete, censored failure time and censored failure time with covariates, we also presented in detail the mechanism used of \mathbf{Y}_n^2 test statistic so that we can apply it an optimal way.

Apart from that we have been validated a new probability distribution that is named a Hypertbastic distribution by using \mathbf{Y}_n^2 statistic. The power of \mathbf{Y}_n^2 test statistic for this distribution, against the parametric \cap -shape distributions, is simulated to demonstrate the flexibility of the considered distribution in survival and reliability analysis. Some real-life data was analyzed to confirm the advantages of this distribution in literature.

We also continued to study, apply and develop the idea of Tabatabai and his colleagues whom proposed using the Hypertabastic distribution as the baseline hazard rate function in the parametric survival models, applying to the clinical data. The classes of parametric survival models which we taken here include the parametric Proportional Hazards and Accelerated Failure Time models.

However, unlike the approaches only used the values of log-likelihood ratio and AIC statistics to solve the problem of Tabatabai and his colleagues, we used here an other approach based on the generalized chi-squared test statistic following the textbooks of Bagdonavičius and Nikulin to choose a good model for each of dataset. A chi-squared test statistic have been firstly constructed to fit the parametric Hypertabastic survival models for the data influenced by many covariates. We afterwards indicate that the Hypertabastic survival models is the accurate model to fit for two dataset, one set from the remission duration from a Clinical Trial for acute leukemia and one-other set from a randomized study comparing two chemotherapy regimes in patients with malignant glioma. Since

then, we make the prognosis of the confidence estimates of regression parameters and survival functions. These results may help the clinicians predict the development tendency of patients, thereby making the decisions treatment suitable. In addition to fitting the Hypertabastic AFT and Hypertabastic PH for those two datasets above, we also give a suggestion on the disadvantages of these models when we fit them for the dataset contain intersection of survival functions. For kind of dataset, the parametric survival models can not be consistent. Therefore, it is necessary to select other dynamic model that can be well adapted this kind of dataset, to alternate for the parametric models.

We also constructed in our works the programs for all of considered models based on R-languages.

We found that the Hypertabastic distribution is closed and plummeted after achieving a maximum speed if the shape parameter $0.5 < b < 1$ besides its uni-modal properties. This means that we have the reasons studying in the future the redundant system that's where use this distribution as baseline distribution function of the "hot" or "warm" conditions. We also wish to be together with our colleagues will continue considering the dynamic model for the multiple myeloma patients in sub-section 2.4 of chapter 4. Simultaneously, in the future, we wish to be together with our colleagues will develop a generally package of the generalized chi-squared test statistic to fit the parametric models for the censoring data possible influenced by covariates based on R-languages.

Appendix

1 Publications During PhD

1. V. Bagdonavičius and R. Levulienė and M. S. Nikulin and **X. Q. Tran**. (2012). On chi-squared type tests and their applications in survival analysis and reliability. *Zapiski Nauchnykh Seminarov (POMI)*, Vol. 42 (#15), 43–61.
2. M. S. Nikulin and **X. Q. Tran**. (2013). Chi-squared goodness-of-fit tests for Generalized Birnbaum-Saunders models for right censored data and its reliability applications. *Journal of Reliability: Theory & application, Electronic Journal of International Groups on Reliability*, Vol. 8 (#2), 7–20.
3. M. S. Nikulin and L. Gerville-Réache and **X. Q. Tran**. (2013). On Chi-Squared Goodness-of-Fit Test for Normality, chapter in: *Statistical Models and Methods for Reliability and Survival Analysis*, (Editors: Couallier, V. and Gerville-Réache, L. and Huber-Carol, H. and Linnios, N. and Mesbah, M.), Wiley-ISTE, New York, 213–227.
4. M. S. Nikulin and **X. Q. Tran**. (2014). On chi-squared testing in accelerated trials. *International Journal of Performability Engineering*, Vol. 10(#1), 53–62.
5. M. S. Nikulin and **X. Q. Tran** and M. R. Tahir. (2014). Goodness-of-fit test for Hypertabastic survival models with applications. *Zapiski Nauchnykh Seminarov (POMI)*, (Accepted).
6. N. Saaidia. and M. R. Tahir and **X. Q. Tran**. (2013). A Modified Chi-squared Goodness-of-Fit Test for the Kumaraswamy-Log-Logistic Distributions and its Applications in Reliability 45th Journées de Statistique, 21-25, May 2013, Toulouse, France. (Accepted).

7. **X. Q. Tran** and M. R. Tahir. (2014). Powerful evaluation of some modified goodness-of-fit tests for the Hypertabastic distribution. 5rd International conference on Accelerated Life Testing and Degradation Models, 11–13/June, Pau, France. (Oral presentation).
8. M. R. Tahir and **X. Q. Tran** and M. S. Nikulin. (2014). Comparison of survival models: A methodology and test for biomedical data. (In progress).
9. M. S. Nikulin and **X. Q. Tran** and M. R. Tahir. (2014). Use of Generalized Birnbaum-Saunders models to improve the reliability of Redundant systems. (In progress).

2 The data

2.1 Remission durations for acute leukemia study

The investigation of acute leukemia of children was carried out by Freireich et al. [46] in 1963. This study was designed to test the ability of a therapy to prolong the remission times. These patients were selected who had a remission induction (complete or partial) of their leukemia by treatment with the drug prednisone that was continued as the steroid treatment until complete bone marrow remission achievement or for a maximum of 28 days (PHASE I). The patients with marrow remission before 28 days or marrow remission on the 28th day were continued to treat randomly to 6-MP or placebo for the remission maintenance (PHASE II). The patients with marrow improvement after 28 days of steroid treatment or received placebo maintenance were treated with 6-MP for remission induction and maintenance (PHASE III). Freireich and his colleagues designed a treatment starting from 97 acute leukemia patients aged from 0 to 19 and conducted at 11 participating institutions, in the period from April 1959 to April 1960, 92 patients were considered acceptable for analysis. Patients were followed until their leukemia returned or until the end of the study (in weeks). This study was terminated when a report analyzing the results of 21 pair of patients are given - this number resulting in the sample path crossing a boundary line of the restricted sequential procedure.

Because of the log while blood cell ($\log(\text{WBC})$) count is often considered important predictor in the analysis of the remission duration of leukemia patient. Therefore, we consider in this document the data including the remission survival times for the two treatment groups with additional information about white blood cell count for each patient

studied. Table 1 give the 21 pair of patients as follow the context of Freireich as well as the textbook of Kleinbaum & Klein [65].

Table 1: Remission duration of 6-MP versus placebo in acute leukemia patients.

No	Drug	Length of Remission	log(WBC)	No	Drug	Length of Remission	log(WBC)
1	Placebo	1	2.80	22	6-MP	10	2.96
2	Placebo	5	3.49	23	6-MP	20 ⁺	2.01
3	Placebo	22	2.73	24	6-MP	7	4.43
4	Placebo	4	4.36	25	6-MP	19 ⁺	2.05
5	Placebo	3	4.01	26	6-MP	32 ⁺	2.20
6	Placebo	15	2.30	27	6-MP	6	4.06
7	Placebo	12	1.50	28	6-MP	23	2.57
8	Placebo	8	2.32	29	6-MP	17 ⁺	2.16
9	Placebo	8	3.52	30	6-MP	22	2.32
10	Placebo	23	1.97	31	6-MP	35 ⁺	1.45
11	Placebo	17	2.95	32	6-MP	6	2.31
12	Placebo	5	3.97	33	6-MP	6	3.28
13	Placebo	2	4.91	34	6-MP	16	3.60
14	Placebo	11	2.12	35	6-MP	13	2.88
15	Placebo	11	3.49	36	6-MP	34 ⁺	1.47
16	Placebo	4	2.42	37	6-MP	9 ⁺	2.80
17	Placebo	8	3.05	38	6-MP	32 ⁺	2.53
18	Placebo	1	5.00	39	6-MP	6 ⁺	3.20
19	Placebo	12	3.06	40	6-MP	25 ⁺	1.78
20	Placebo	8	3.26	41	6-MP	10 ⁺	2.70
21	Placebo	2	4.48	42	6-MP	11 ⁺	2.60

⁺ Censored observations.

2.2 Brain tumour study with malignant glioma patients

A randomized clinical trial comparing two chemotherapy regimens in 447 patients with malignant glioma that are studied by Sauerbrei & Schumacher [105]. During an accrual period of five years, 293 patients have died and the median survival time from the randomization was about 11 months. The data contains 12 explanatory variables including age, three ordinal and eight binary variables. Three variable are measured on an ordinal scale that are the index of Karnofsky, the type of surgical resection and the grade of malignancy. One of these variables is represented by two dummy variables resulting 16 variables in all. Details for these variables are presented in Table 4.5 in chapter 4

Table 2: Brain tumour study of the patients with malignant glioma

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
1	1	2	0	0	64	1	0	0	0	1	0	0	0	1	0	1	157	1
2	0	2	0	0	59	0	0	1	0	1	1	1	0	0	1	0	300	1
3	1	2	0	0	60	1	0	1	0	1	1	0	0	0	1	0	238	1
4	0	1	1	1	54	0	0	1	0	1	1	0	0	0	0	1	196	1
5	0	2	0	0	63	0	0	0	0	1	1	0	0	0	1	0	443	1
6	0	1	0	0	58	1	0	0	0	1	1	1	0	0	0	1	372	1
7	0	1	1	1	53	0	0	0	0	1	1	0	0	0	0	0	131	1
8	0	2	1	0	61	1	1	1	1	1	1	1	0	0	0	1	254	1
9	0	1	1	1	57	1	1	1	0	0	1	0	0	0	0	0	233	1
10	0	1	0	0	62	1	0	1	0	1	1	0	0	1	1	1	434	1
11	1	2	0	0	50	1	0	0	0	0	1	0	0	0	0	1	673	1
12	1	1	1	1	64	1	1	1	0	0	1	0	0	0	0	0	324	1
13	0	1	1	1	63	1	1	1	1	1	1	0	0	0	0	0	385	1
14	0	2	1	0	67	1	0	1	0	1	1	0	1	0	1	1	37	1
15	0	1	1	1	45	0	0	1	0	1	1	1	0	0	0	1	200	1
16	0	2	1	1	52	1	0	1	0	1	0	0	0	1	0	1	169	1
17	1	2	1	0	68	1	0	0	0	1	1	1	1	1	0	0	53	1
18	1	2	1	0	62	1	0	0	0	0	1	0	0	0	0	1	225	1
19	1	2	1	1	57	1	0	1	0	1	1	0	1	0	0	0	30	1
20	0	1	0	0	44	1	1	1	0	0	1	0	0	0	0	0	1249	0
21	0	1	1	1	63	0	0	1	1	0	1	0	0	0	1	1	134	1
22	1	1	0	0	58	1	0	0	0	0	1	0	0	0	1	0	491	1
23	1	1	0	0	52	0	0	1	0	0	1	0	0	1	1	0	422	1
24	0	2	1	0	20	1	0	1	0	1	1	0	0	0	0	1	1447	0
25	0	2	1	0	43	1	0	1	1	1	0	0	0	0	0	1	614	1
26	0	1	1	1	23	1	0	1	0	1	0	0	0	0	0	0	887	1
27	0	2	1	1	49	1	1	1	1	1	1	1	0	0	0	0	239	1
28	1	2	1	1	23	0	0	1	0	1	1	0	0	1	0	1	72	1
29	0	2	0	0	45	1	1	0	0	1	1	1	0	0	0	1	1250	0
30	1	1	1	0	53	0	0	1	1	1	1	0	0	0	0	0	728	1
31	0	2	1	0	33	1	1	1	0	1	1	0	0	0	0	1	956	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
32	1	1	1	1	41	0	0	0	0	1	1	0	0	0	1	0	124	1
33	0	2	1	1	26	1	0	1	0	1	1	1	0	0	1	1	527	1
34	1	1	1	0	62	1	0	1	1	1	1	0	1	1	1	1	1145	1
35	0	2	1	1	45	1	0	1	0	1	1	0	0	0	0	1	507	1
36	1	1	1	1	59	0	0	1	1	0	1	0	0	0	0	0	533	1
37	0	2	1	1	57	1	0	1	0	1	1	0	0	0	1	1	395	1
38	0	2	0	0	46	1	1	1	0	0	1	1	0	0	0	0	1167	0
39	0	1	1	1	68	1	0	1	0	0	1	1	0	0	0	0	85	1
40	0	2	0	0	45	0	0	0	0	1	1	0	1	1	1	0	664	1
41	1	1	1	1	61	1	0	1	0	1	1	0	1	0	0	1	77	1
42	0	1	1	1	68	1	0	1	0	0	1	0	0	0	0	0	112	1
43	0	1	1	0	27	1	1	1	0	1	1	1	0	0	1	1	343	1
44	0	1	1	1	60	0	0	0	0	1	1	0	0	0	1	1	38	1
45	0	2	1	1	63	0	0	1	0	1	1	0	0	1	0	1	315	1
46	1	1	1	1	59	1	0	0	0	0	1	0	0	0	1	1	207	1
47	1	1	1	1	43	0	0	1	0	1	1	0	0	0	0	0	936	0
48	1	2	1	0	32	0	0	0	0	1	1	0	0	0	0	0	343	1
49	1	2	1	1	48	1	1	1	1	1	0	1	0	0	1	0	468	1
50	0	1	1	1	45	0	0	0	0	0	1	1	0	0	0	0	109	1
51	1	2	1	0	28	1	0	1	0	0	0	0	1	1	0	1	140	1
52	1	1	1	1	58	0	0	1	1	0	1	0	1	0	1	1	155	1
53	0	1	1	1	56	0	0	0	0	0	1	0	0	0	0	0	23	1
54	1	2	1	1	50	1	0	1	0	1	0	0	0	0	1	1	757	1
55	0	2	0	0	37	1	1	1	0	1	0	0	0	0	0	1	796	1
56	0	2	1	1	44	1	0	1	0	1	1	1	0	0	0	1	390	1
57	0	1	1	1	52	1	1	1	1	0	1	0	0	0	0	1	510	1
58	0	2	0	0	57	1	0	0	0	1	1	0	0	0	1	0	145	1
59	1	2	1	1	67	0	0	1	0	0	1	0	0	0	0	1	174	1
60	1	2	1	1	62	1	0	0	0	0	1	0	0	0	0	0	254	1
61	0	1	1	1	58	1	0	1	0	0	1	0	0	1	0	1	190	1
62	0	1	1	1	73	1	0	1	0	1	1	0	0	0	1	1	350	1
63	0	1	1	1	58	0	0	0	0	0	1	0	0	1	0	1	111	1
64	0	1	1	1	53	0	0	1	0	1	1	0	0	0	0	0	623	0

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
65	1	1	1	1	63	1	0	1	0	1	1	0	0	0	1	0	347	1
66	0	2	1	1	58	1	1	1	1	1	1	0	1	0	0	0	193	1
67	0	1	0	0	56	0	0	0	0	0	1	0	0	0	0	0	212	1
68	0	1	1	1	54	1	1	1	0	0	1	0	0	0	0	0	237	0
69	1	2	1	0	53	1	0	1	0	0	1	0	0	0	0	0	149	1
70	0	2	1	1	56	1	0	1	0	0	1	0	0	0	0	0	547	1
71	0	2	1	0	57	1	1	1	0	0	1	0	0	0	0	1	712	0
72	0	2	0	0	55	0	0	1	0	1	1	1	1	0	1	1	16	0
73	0	1	1	1	61	1	1	1	0	0	1	0	0	0	1	1	519	1
74	0	1	1	1	52	1	0	1	0	0	1	0	1	0	0	1	87	1
75	1	1	1	1	61	0	0	1	1	0	1	0	0	0	1	1	241	1
76	0	2	1	1	61	0	0	0	0	1	1	0	0	1	0	0	91	1
77	0	1	1	1	43	1	0	1	1	1	1	1	0	0	1	1	403	1
78	0	2	1	1	64	1	1	1	1	0	1	0	0	0	1	1	589	0
79	0	1	1	1	70	1	0	1	0	1	1	0	0	0	1	1	157	1
80	0	1	1	1	54	1	0	1	1	1	1	1	0	0	1	1	446	1
81	1	1	0	0	58	1	0	1	0	1	1	0	1	1	1	1	101	1
82	0	1	1	1	33	1	0	1	1	1	1	1	0	0	0	0	382	1
83	0	1	1	1	61	1	0	1	0	1	1	0	0	1	1	0	45	1
84	1	2	1	0	51	1	0	0	0	1	1	0	0	1	0	0	146	1
85	0	2	1	1	62	1	0	0	0	1	1	1	0	0	0	0	82	1
86	1	1	1	1	51	1	1	1	1	1	1	1	0	0	0	0	432	1
87	1	1	1	1	54	1	1	1	0	1	1	1	0	0	0	1	183	0
88	0	1	1	1	63	1	1	1	1	0	1	0	0	0	1	0	314	1
89	1	1	1	1	47	1	0	1	1	1	1	0	1	0	0	0	259	1
90	0	2	0	0	35	1	1	0	0	1	1	1	0	0	0	0	205	1
91	1	2	1	1	67	1	0	1	0	0	1	0	0	1	0	0	129	1
92	1	2	0	0	57	1	0	1	1	1	1	0	0	1	1	1	465	0
93	1	1	1	1	63	0	0	1	0	1	1	0	0	0	0	1	233	1
94	1	2	0	0	30	1	1	1	0	1	1	0	0	0	0	1	221	0
95	0	1	0	0	64	1	0	1	0	1	1	0	0	1	0	0	264	0
96	1	1	1	1	55	1	0	1	0	1	1	1	0	0	0	0	503	0
97	1	1	1	1	53	0	0	1	1	0	1	0	0	1	0	1	190	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
98	1	1	1	1	56	1	0	1	0	1	1	1	0	1	1	1	281	1
99	0	1	1	1	60	0	0	1	0	1	1	1	1	1	0	0	126	1
100	1	2	0	0	64	0	0	0	0	0	1	0	0	1	0	1	71	1
101	1	2	0	0	34	1	0	0	0	1	1	1	0	0	0	0	240	0
102	0	1	1	1	28	1	1	1	1	1	1	1	0	0	0	0	85	0
103	1	1	1	1	57	0	0	1	1	0	0	0	1	1	0	1	260	1
104	1	1	1	1	62	1	1	1	0	1	1	0	0	0	1	1	201	0
105	1	1	1	1	62	1	0	1	0	1	1	1	0	0	0	0	197	0
106	1	1	1	1	64	1	0	1	0	1	1	0	0	0	1	1	173	0
107	1	1	1	1	34	1	0	1	0	1	1	0	1	0	0	0	337	0
108	0	1	1	1	55	1	0	1	1	1	1	0	1	0	0	0	354	1
109	0	1	1	1	70	0	0	1	0	0	1	0	1	1	0	1	187	1
110	0	1	1	1	64	1	1	1	0	1	0	0	0	0	0	1	215	0
111	1	1	1	0	59	0	0	1	0	1	1	0	1	0	1	0	189	0
112	0	1	1	1	46	1	1	1	0	0	0	0	0	0	0	0	227	0
113	0	1	1	1	61	1	0	1	0	0	1	0	0	0	1	1	84	0
114	1	1	1	0	62	1	0	1	0	1	1	0	1	0	0	1	324	0
115	1	1	1	0	33	1	0	1	0	1	1	1	0	0	0	1	209	0
116	0	2	1	0	40	1	0	1	0	1	1	0	0	0	1	1	174	0
117	1	1	1	1	52	1	1	1	0	1	1	0	0	0	0	1	226	1
118	1	1	1	0	30	1	1	1	0	1	1	1	0	0	0	0	238	0
119	0	1	1	1	66	1	0	0	0	0	1	0	0	0	0	0	212	1
120	0	1	1	1	60	1	0	1	1	1	1	0	1	0	0	0	172	1
121	0	1	0	0	22	1	0	1	0	1	1	0	0	0	0	0	42	0
122	0	2	1	1	43	1	0	1	0	1	1	0	0	0	0	0	210	0
123	1	2	1	1	56	1	0	1	1	1	1	0	0	0	0	1	172	0
124	0	2	1	1	52	1	0	1	0	1	1	0	0	0	0	1	173	0
125	0	2	1	1	57	1	0	0	0	0	1	0	0	0	1	1	84	0
126	0	1	1	1	64	1	0	1	0	0	1	0	0	0	0	1	121	0
127	0	1	0	0	42	1	1	1	1	1	0	0	0	0	0	0	63	0
128	0	2	1	1	48	1	0	1	0	1	1	1	0	0	0	1	49	0
129	0	2	1	1	59	0	0	1	1	1	1	0	0	1	1	0	52	1
130	1	2	0	0	61	1	1	1	0	1	1	1	0	0	0	0	73	0

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
131	0	1	1	1	62	1	0	1	0	0	1	0	0	0	0	0	42	0
132	1	2	1	1	63	1	0	1	0	1	1	1	0	1	1	0	64	0
133	0	2	0	0	27	1	1	1	0	1	1	1	0	0	0	1	41	0
134	0	2	0	0	28	1	0	1	0	1	0	1	0	0	0	1	42	0
135	1	2	1	1	59	0	0	1	0	1	1	1	0	0	1	1	171	1
136	0	2	0	0	56	0	0	0	0	1	1	0	0	0	0	0	133	1
137	0	1	0	0	45	1	1	1	1	0	1	0	0	0	0	1	1617	0
138	1	1	1	1	52	1	0	1	1	1	1	0	0	0	0	1	10	1
139	0	1	1	1	31	1	1	1	1	1	0	0	0	0	0	0	292	1
140	1	1	1	1	59	1	0	1	1	1	1	0	0	0	0	1	309	1
141	0	1	1	1	52	1	1	1	1	1	0	0	0	0	0	0	535	1
142	0	1	1	1	56	1	0	1	1	1	0	0	0	0	0	1	328	1
143	1	1	1	0	48	0	0	1	1	1	1	0	0	0	0	0	674	1
144	0	1	0	0	52	0	0	1	0	1	1	0	0	0	0	1	81	1
145	0	2	1	1	68	1	0	1	0	1	1	0	0	0	0	0	77	1
146	0	1	1	1	57	1	0	1	0	1	1	0	0	0	0	1	152	1
147	1	1	1	1	71	0	0	0	0	1	1	0	0	0	0	1	62	1
148	0	1	1	1	47	1	1	1	1	1	0	1	0	0	0	0	517	1
149	0	1	1	1	63	1	0	1	0	1	0	0	0	0	1	0	36	1
150	0	1	0	0	58	1	1	1	0	1	1	0	0	0	0	0	85	1
151	0	1	1	1	50	1	0	1	0	1	1	0	1	0	0	1	521	0
152	0	2	1	1	22	1	1	1	0	1	0	1	0	0	0	0	187	1
153	0	1	0	0	50	1	1	1	0	1	0	1	0	0	1	1	879	1
154	0	1	0	0	32	1	1	1	1	1	1	1	0	0	0	0	843	0
155	0	1	1	0	50	0	0	1	0	1	1	0	0	0	0	1	29	1
156	0	1	1	1	43	0	0	1	1	1	1	0	0	0	1	1	691	1
157	1	1	1	1	60	1	1	1	0	1	0	1	0	0	0	1	545	1
158	1	2	0	0	45	0	0	1	1	1	0	0	0	0	0	0	836	0
159	1	1	1	1	20	1	0	1	1	1	0	1	0	0	0	0	1080	0
160	1	1	1	1	70	0	0	1	1	1	0	0	0	0	0	1	76	1
161	0	1	1	1	61	0	0	1	1	1	1	1	0	0	0	1	15	1
162	1	1	1	1	59	1	0	1	1	1	1	0	0	0	0	0	36	1
163	1	1	0	0	49	1	1	1	1	1	1	0	0	0	0	0	344	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
164	1	2	1	1	51	1	0	1	0	1	0	0	0	0	0	0	596	1
165	0	2	1	1	57	0	0	1	0	1	1	0	0	0	0	1	163	1
166	1	2	1	1	65	1	0	1	0	1	1	0	1	1	0	0	40	1
167	1	1	1	1	66	1	1	1	0	1	1	0	0	0	1	1	33	1
168	0	2	1	0	48	1	0	1	0	1	1	1	0	0	0	1	279	1
169	0	1	1	1	51	1	1	1	1	1	1	0	0	0	0	0	193	1
170	1	1	1	1	60	0	0	1	0	1	1	0	0	0	0	0	179	1
171	0	1	1	1	49	0	0	0	0	1	1	0	0	0	1	0	12	1
172	0	1	1	1	57	0	0	0	0	1	1	0	0	0	1	1	64	1
173	1	1	1	1	36	0	0	1	1	1	1	0	0	0	1	0	304	1
174	0	1	1	1	66	1	0	1	0	1	1	0	1	0	0	1	76	1
175	0	1	1	1	66	1	0	1	0	1	0	0	0	0	0	1	266	1
176	0	1	1	1	70	0	0	1	0	1	1	0	0	0	0	1	16	1
177	1	1	1	1	69	1	0	1	0	1	1	0	0	0	1	1	30	1
178	0	2	0	0	23	1	1	1	0	1	1	0	0	0	0	0	456	0
179	0	2	1	0	27	1	1	0	0	1	1	0	1	1	1	0	189	0
180	0	1	1	1	62	1	0	0	0	1	1	0	0	0	1	0	191	1
181	1	1	0	0	65	0	0	1	0	1	1	0	0	0	0	0	327	0
182	0	1	1	1	56	1	0	1	0	1	1	0	0	0	0	0	35	1
183	1	1	0	0	33	1	1	1	0	1	1	0	1	0	0	1	329	0
184	1	2	1	1	44	1	0	1	0	1	1	1	0	0	0	0	97	0
185	0	1	1	1	46	1	0	0	0	0	0	0	0	0	1	0	265	1
186	0	2	1	1	66	1	0	0	0	1	1	0	0	0	0	0	11	1
187	0	2	1	1	52	1	0	0	0	1	1	0	0	1	1	0	299	0
188	0	1	1	1	62	0	0	1	0	1	1	1	0	0	1	1	7	1
189	1	1	1	1	58	1	0	1	1	1	1	1	0	0	0	1	220	0
190	0	1	1	1	60	1	0	1	0	1	1	0	0	0	0	0	259	0
191	1	2	1	1	44	1	0	1	0	0	1	0	0	0	0	1	27	1
192	0	2	1	1	51	1	0	1	0	1	0	1	0	0	1	0	134	0
193	1	1	1	1	45	1	0	1	1	1	1	0	1	0	1	0	44	0
194	1	1	1	1	47	0	0	1	0	1	1	1	0	0	0	1	44	0
195	0	2	0	0	35	0	0	1	1	1	0	0	0	1	0	1	683	1
196	1	2	1	0	54	0	0	1	1	1	1	0	0	0	0	0	234	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
197	0	1	1	1	58	1	1	1	1	1	0	1	0	0	0	0	633	1
198	1	2	0	0	34	1	0	1	1	0	0	0	0	0	0	1	761	1
199	0	1	1	1	59	0	0	1	0	0	0	1	0	0	0	1	567	1
200	0	2	1	1	28	1	0	1	0	1	1	1	0	0	0	1	768	0
201	1	1	0	0	69	0	0	1	1	1	1	1	1	1	1	1	198	1
202	1	2	1	1	58	1	0	1	0	1	0	1	0	0	0	1	255	1
203	0	2	0	0	48	1	0	1	0	1	1	1	0	0	0	1	348	0
204	0	1	0	0	46	1	0	1	1	1	1	0	0	0	0	0	227	1
205	1	2	0	0	65	1	0	1	1	1	0	1	0	0	0	0	565	0
206	0	2	1	1	46	1	0	1	0	1	0	0	0	0	0	0	479	1
207	0	1	1	1	39	0	0	1	1	1	0	1	0	1	0	0	324	1
208	0	1	1	1	58	1	0	1	0	0	1	0	0	0	0	0	349	0
209	1	1	1	1	52	0	0	1	0	0	0	0	0	0	0	1	141	0
210	0	1	1	1	58	1	0	1	0	1	0	0	0	0	0	0	241	0
211	0	1	1	1	61	0	0	1	0	1	1	0	0	1	0	1	137	1
212	1	1	1	1	62	1	0	1	0	0	0	0	0	0	0	1	783	1
213	0	1	1	1	66	0	0	1	0	1	1	0	0	0	0	1	98	1
214	0	2	1	1	47	1	0	1	1	0	0	0	0	0	0	0	436	1
215	0	1	1	1	54	1	0	1	1	0	1	0	0	0	0	1	578	1
216	1	1	1	1	56	0	0	1	0	0	1	0	0	1	0	0	180	1
217	1	1	1	1	60	0	0	1	1	0	1	0	0	0	1	0	73	1
218	0	2	1	1	42	1	0	1	0	1	1	1	0	0	0	0	316	1
219	1	2	1	1	45	0	0	1	0	0	0	0	0	0	1	1	323	1
220	1	1	1	0	45	1	0	1	0	0	1	0	0	1	0	0	379	0
221	1	1	1	0	59	0	0	1	0	1	1	0	0	0	1	0	128	1
222	1	1	1	0	58	1	0	1	0	0	0	0	0	0	0	0	277	1
223	1	1	1	1	49	1	0	1	1	0	1	0	0	0	0	1	181	1
224	1	1	1	1	56	1	0	1	1	0	0	0	0	0	0	1	564	0
225	1	1	1	1	67	1	0	1	0	1	1	0	0	0	1	0	148	1
226	0	2	0	0	32	1	1	1	0	1	0	1	0	0	0	0	804	1
227	1	1	1	1	51	1	0	1	0	0	1	0	0	0	1	1	159	1
228	0	1	1	1	59	1	0	1	0	1	1	0	0	0	0	0	103	0
229	0	2	1	0	55	1	0	1	0	0	1	1	1	0	0	0	466	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
230	0	2	1	1	45	1	0	1	0	0	1	0	0	0	0	0	225	1
231	0	1	1	1	59	1	1	1	1	1	1	1	0	0	0	0	491	1
232	1	1	1	1	50	1	1	1	1	0	0	0	0	0	0	1	450	1
233	0	1	1	1	51	1	0	1	0	0	1	0	0	0	1	1	526	1
234	1	1	0	0	37	1	1	1	0	0	0	0	0	0	0	0	539	0
235	1	1	1	0	68	1	1	1	1	1	1	0	0	0	0	1	98	0
236	1	2	1	0	50	1	1	1	1	1	1	1	0	0	0	0	166	0
237	0	1	1	0	62	1	1	1	1	0	1	0	0	0	0	1	216	1
238	0	1	1	0	60	1	0	1	0	0	1	0	0	0	1	1	32	1
239	1	2	1	0	65	1	1	1	0	1	1	1	0	0	0	0	50	0
240	0	1	0	0	34	0	0	1	0	0	1	0	0	1	0	0	279	0
241	1	2	1	0	65	1	0	1	1	0	0	1	0	0	1	1	390	1
242	1	1	1	0	64	0	0	1	1	0	0	0	0	0	0	0	48	0
243	0	2	1	0	60	1	1	1	0	1	1	1	0	0	0	1	280	0
244	1	1	0	0	62	0	0	0	0	0	1	0	0	0	1	1	65	1
245	0	1	1	0	61	0	0	1	1	0	1	0	0	1	0	1	62	1
246	1	1	1	0	49	0	0	1	1	0	0	0	0	0	1	0	295	1
247	0	2	1	0	58	1	1	1	1	0	1	0	0	1	0	1	90	0
248	0	2	1	1	52	0	0	1	1	1	1	1	0	0	1	1	901	1
249	1	1	1	1	58	0	0	1	0	1	0	0	0	1	1	0	190	1
250	0	2	1	1	24	1	0	1	0	1	1	0	0	0	0	0	617	1
251	1	1	1	1	56	0	0	1	0	0	0	1	0	0	0	1	209	1
252	1	2	0	0	25	1	1	1	1	1	0	0	0	0	0	1	1384	0
253	0	1	1	1	40	0	0	1	1	1	1	0	1	1	0	1	103	1
254	0	2	0	0	59	0	0	1	1	0	0	1	0	0	0	0	371	1
255	0	1	1	1	50	1	0	1	1	0	1	0	1	1	0	1	346	1
256	0	1	1	1	41	0	0	1	1	0	1	0	1	0	0	0	932	1
257	0	2	1	1	55	1	0	1	1	1	0	0	0	1	0	0	1149	0
258	0	2	0	0	21	1	0	1	1	1	0	1	0	0	0	1	744	0
259	1	1	1	1	60	1	0	1	1	1	0	0	1	1	0	1	627	1
260	0	1	1	1	55	1	0	1	1	1	0	1	0	1	0	1	288	1
261	0	1	1	0	50	0	0	1	1	0	1	1	0	1	0	1	56	0
262	0	1	1	1	52	1	0	1	1	1	0	1	0	1	0	0	408	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
263	0	1	1	1	45	0	0	1	1	1	1	0	0	1	0	1	264	1
264	1	1	1	1	57	0	0	1	0	1	1	0	1	1	0	1	200	1
265	1	1	1	1	53	1	0	1	1	1	1	1	0	0	0	1	89	0
266	1	1	1	1	46	1	0	1	1	1	0	1	0	0	0	0	108	0
267	0	1	1	1	60	1	0	1	0	1	1	0	0	0	0	0	79	0
268	1	1	1	0	56	1	1	1	1	0	0	0	0	0	1	1	414	1
269	0	2	1	0	60	0	0	1	1	1	0	1	1	1	1	0	342	1
270	1	1	1	1	51	1	1	1	0	0	0	0	1	1	0	1	283	1
271	0	1	1	1	53	0	0	1	1	0	0	0	0	1	1	1	584	1
272	0	1	1	1	22	1	1	1	0	0	0	0	0	1	0	0	84	1
273	1	2	1	1	63	1	0	1	1	1	1	1	0	0	1	0	663	0
274	0	2	1	1	39	1	0	1	1	0	0	0	0	1	0	1	496	1
275	1	2	1	1	66	1	0	1	0	0	0	0	0	1	1	1	75	1
276	1	2	1	1	50	1	1	1	0	0	0	0	0	1	1	0	229	1
277	0	1	1	1	61	1	0	1	0	1	1	1	0	1	1	1	535	1
278	1	1	1	1	59	1	0	1	0	0	0	0	0	0	0	1	380	1
279	1	2	1	1	70	1	0	1	1	0	0	0	0	1	1	1	384	0
280	0	1	1	1	66	1	0	1	1	1	1	1	0	0	1	0	359	1
281	1	2	0	0	52	1	0	1	1	1	0	1	0	1	0	0	806	0
282	1	2	1	1	63	1	0	1	0	0	0	0	0	0	0	1	462	1
283	0	1	1	1	40	1	0	1	1	0	0	1	1	1	1	0	694	0
284	0	2	1	0	31	1	0	1	1	1	0	1	0	0	0	1	772	0
285	0	1	0	0	47	0	0	1	0	1	0	1	1	1	0	1	498	1
286	0	1	1	1	47	1	1	1	0	0	0	1	0	0	0	1	502	0
287	0	1	1	1	30	1	1	1	1	1	0	1	0	0	0	1	583	0
288	1	1	0	0	47	1	0	1	1	0	0	0	0	0	1	0	487	1
289	1	2	1	0	51	1	0	1	1	1	0	1	0	0	1	1	379	0
290	0	2	0	0	41	1	0	1	0	0	0	0	0	1	0	0	345	0
291	1	1	1	1	57	0	0	1	1	0	0	0	0	1	1	1	34	0
292	0	1	1	1	41	0	0	1	1	0	1	0	1	1	0	0	308	1
293	0	1	1	1	67	0	0	1	0	0	1	0	0	1	0	0	130	1
294	1	1	1	1	56	1	0	1	0	1	1	1	0	1	0	0	98	1
295	1	1	1	1	63	1	0	1	1	0	0	0	0	0	0	0	123	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
296	1	2	0	0	51	1	0	1	0	1	1	0	0	0	0	1	1596	0
297	1	1	0	0	55	1	0	1	0	1	0	0	0	0	0	1	426	1
298	0	2	0	0	45	1	0	1	0	1	1	1	0	0	1	0	1058	1
299	0	2	0	0	38	1	0	1	0	1	0	0	0	1	0	1	1333	0
300	1	1	1	0	64	1	0	1	0	1	1	0	0	0	1	0	594	1
301	0	1	0	0	42	1	0	1	0	1	0	0	0	0	0	1	349	1
302	0	1	0	0	62	1	0	1	0	1	0	0	0	0	1	0	263	1
303	0	1	1	1	65	0	0	1	0	1	0	0	1	1	0	1	304	1
304	0	1	1	0	38	1	0	1	0	1	0	0	0	0	0	0	192	1
305	0	2	1	1	60	1	0	1	0	1	1	1	0	0	0	1	385	1
306	1	2	1	0	49	0	0	1	0	1	1	0	0	0	1	0	196	1
307	0	2	1	0	35	0	0	1	0	1	0	1	0	0	1	1	1167	0
308	0	1	1	0	62	1	0	1	0	1	0	0	0	0	0	1	201	1
309	0	2	1	1	45	0	0	1	0	1	1	0	0	1	0	0	255	1
310	0	2	1	1	34	1	0	1	1	1	0	0	0	0	0	0	552	1
311	1	1	1	1	65	0	0	1	0	1	0	0	0	0	0	1	85	1
312	1	2	1	0	52	0	0	1	0	1	0	0	1	1	0	0	47	0
313	1	1	1	1	41	1	0	1	0	1	0	0	0	1	1	1	972	0
314	1	1	1	1	57	0	0	1	0	1	0	0	0	0	1	0	744	1
315	0	2	1	1	58	1	0	1	1	1	0	0	0	0	0	1	105	1
316	0	1	1	1	46	1	0	1	0	1	0	0	0	0	0	0	705	1
317	1	2	1	1	35	1	0	1	0	1	0	1	0	0	0	1	799	1
318	1	1	1	1	53	0	0	1	0	1	0	1	0	0	0	0	349	1
319	1	1	1	1	54	0	0	1	0	1	0	0	0	0	1	0	305	1
320	0	2	0	0	48	0	0	1	0	1	0	0	0	1	0	1	371	1
321	0	1	0	0	44	0	0	1	0	1	0	0	0	1	1	1	893	0
322	1	1	1	1	66	0	0	1	0	1	1	1	0	0	0	1	557	1
323	1	1	1	1	47	0	0	1	0	1	1	0	0	0	0	1	320	1
324	0	2	0	0	50	0	0	1	0	1	1	1	0	1	0	0	652	0
325	0	1	1	1	47	1	0	1	0	1	1	0	0	0	0	0	374	1
326	0	1	1	1	47	0	0	1	0	1	1	0	0	0	0	1	312	1
327	0	1	1	1	67	0	0	1	0	1	0	0	0	0	1	0	455	1
328	0	1	1	0	64	1	0	1	1	1	0	1	1	1	0	1	261	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
329	0	1	1	1	53	1	0	1	0	1	0	0	0	0	1	1	192	1
330	0	1	1	1	31	0	0	1	0	1	1	1	0	0	1	0	423	1
331	0	1	0	0	48	0	0	1	0	1	0	1	0	0	0	1	520	0
332	1	1	0	0	57	1	0	1	0	1	0	1	0	0	0	0	199	1
333	1	1	1	1	43	1	0	1	0	1	0	0	1	1	0	1	427	1
334	1	1	1	1	58	0	0	1	1	1	0	1	1	1	0	0	286	1
335	1	2	1	1	62	0	0	1	0	0	0	0	0	0	0	0	324	0
336	1	1	1	1	70	1	0	1	1	1	0	0	0	0	0	0	315	1
337	0	2	1	1	38	1	0	1	0	1	0	1	0	0	0	1	385	0
338	1	2	1	1	36	1	0	1	1	1	0	1	0	0	0	1	298	1
339	1	2	0	0	48	1	0	1	0	1	0	1	0	0	0	0	292	0
340	0	2	0	0	44	1	0	1	0	1	0	1	0	0	0	1	297	0
341	0	2	1	1	54	1	0	1	0	1	1	1	0	1	0	0	192	1
342	1	1	1	1	56	0	0	1	0	1	1	0	1	1	0	1	47	1
343	0	2	1	1	39	1	0	1	0	1	0	1	0	1	0	1	168	0
344	0	1	1	1	57	1	0	1	0	1	0	1	1	1	0	1	187	0
345	1	1	1	1	49	1	0	1	0	1	0	0	1	1	0	0	200	0
346	0	1	1	1	59	0	0	1	0	1	1	0	0	0	1	0	77	0
347	1	2	1	1	67	0	0	1	0	1	0	0	0	1	1	0	135	0
348	0	2	0	0	36	1	0	1	0	1	0	1	0	0	0	1	65	0
349	0	1	1	1	34	1	0	1	0	1	0	0	0	1	0	0	57	0
350	0	1	1	1	33	0	0	1	0	1	0	0	0	1	1	0	64	0
351	0	2	0	0	44	1	0	1	0	1	1	1	0	0	0	0	586	1
352	1	1	1	1	58	1	0	0	0	0	1	0	0	0	1	0	315	1
353	1	1	1	1	52	1	1	1	0	1	0	1	0	0	0	1	566	1
354	0	2	1	1	54	1	1	1	0	0	1	0	0	0	0	1	393	0
355	0	1	1	0	32	1	1	1	0	1	1	1	0	0	0	0	183	0
356	1	2	1	1	48	1	1	1	0	1	1	1	0	0	0	1	488	1
357	1	1	1	1	41	1	1	1	0	1	1	0	0	0	0	1	1347	0
358	0	2	1	1	62	1	0	1	1	1	1	1	0	0	0	1	252	1
359	0	1	1	0	16	1	1	1	0	0	0	0	0	0	0	0	396	1
360	1	1	1	1	46	1	1	1	1	1	1	0	0	0	0	0	290	1
361	0	1	1	1	60	1	1	1	1	1	1	0	0	0	1	1	6	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
362	1	2	1	1	70	1	0	1	1	1	1	1	0	1	0	0	102	1
363	1	1	1	1	59	1	0	1	1	1	1	0	1	1	0	1	173	0
364	1	1	1	1	56	1	1	1	0	1	1	0	0	1	1	0	295	1
365	0	1	1	1	60	1	1	1	1	1	1	1	0	0	0	0	430	1
366	0	1	1	1	48	1	0	1	1	0	1	0	0	0	0	0	274	0
367	0	2	1	1	54	1	0	1	1	0	1	1	0	0	0	0	417	0
368	1	1	1	0	53	1	0	1	0	1	0	0	0	0	0	1	301	0
369	1	1	1	1	54	0	0	1	0	1	1	0	0	1	1	1	87	1
370	0	1	1	0	58	1	0	1	1	0	0	0	0	0	0	1	308	1
371	1	2	1	0	66	0	0	1	1	1	0	1	0	0	0	0	497	0
372	1	1	1	1	55	1	0	1	0	1	0	0	0	0	0	1	273	1
373	0	1	1	1	54	1	0	1	0	1	0	0	0	0	1	0	245	1
374	1	1	1	1	58	1	1	1	0	0	0	0	0	0	0	0	272	1
375	0	2	1	0	36	1	0	1	0	1	0	1	0	0	0	0	227	1
376	1	1	1	1	62	0	0	1	1	1	1	0	0	0	0	1	471	1
377	1	2	1	0	67	0	0	1	1	0	1	0	0	1	1	0	376	1
378	0	2	1	1	46	0	0	1	0	1	1	0	0	0	1	0	166	1
379	0	2	1	1	50	0	0	1	0	1	1	0	0	1	1	1	121	1
380	0	2	1	1	40	1	0	1	1	1	1	1	0	1	0	1	45	0
381	0	2	1	0	49	1	1	1	1	0	0	0	1	1	0	0	336	1
382	1	2	1	1	51	1	1	1	1	1	1	0	1	0	0	1	592	1
383	1	2	1	1	54	1	1	1	1	1	1	1	0	0	0	1	340	1
384	0	2	1	1	49	1	1	1	0	0	1	0	0	0	1	0	293	1
385	0	2	1	1	61	1	1	1	1	0	1	0	0	0	0	0	267	1
386	0	1	1	1	61	1	1	1	1	0	1	0	0	0	0	1	114	1
387	1	1	1	1	45	1	1	1	0	1	1	1	0	0	1	0	250	1
388	1	1	0	0	57	1	1	1	0	0	1	1	0	0	0	1	295	0
389	0	1	1	1	66	1	1	1	0	0	1	0	0	0	1	1	335	1
390	1	1	1	1	65	1	1	1	0	0	1	0	0	1	1	1	154	0
391	1	1	1	1	64	1	1	1	0	0	1	0	0	1	1	0	174	1
392	0	2	1	1	41	1	0	1	0	0	1	1	0	0	1	1	60	0
393	1	1	0	0	47	1	1	1	1	0	1	0	0	0	0	1	155	0
394	1	2	1	1	48	1	1	1	1	0	1	0	0	0	0	0	76	1

To be continue

pt	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}	z_{16}	t	d
395	1	1	1	1	40	1	1	1	1	1	0	0	1	1	0	0	83	1
396	0	1	1	0	47	1	1	1	0	1	1	1	0	0	0	1	203	0
397	0	1	1	1	40	1	1	1	0	1	1	1	0	0	0	1	16	0
398	0	1	1	0	44	1	1	1	1	1	1	0	1	1	0	0	50	0
399	0	2	1	1	40	1	0	0	0	1	1	0	0	0	0	0	4	1
400	1	2	0	0	21	1	1	1	0	1	1	1	0	0	0	0	236	0
401	0	1	0	0	29	1	1	1	0	1	1	1	0	0	0	1	300	0
402	1	1	1	1	65	0	0	1	0	0	0	0	0	0	0	0	249	0
403	0	1	1	1	60	1	1	1	0	0	0	0	0	0	0	1	156	1
404	0	1	1	1	45	1	1	1	0	0	1	1	0	0	0	1	282	0
405	1	1	1	1	50	1	0	1	0	0	1	0	0	0	1	1	282	0
406	1	1	1	1	66	0	0	1	0	0	0	1	0	0	0	0	245	0
407	0	1	1	1	65	1	0	1	0	0	0	0	0	0	1	0	81	1
408	1	2	1	1	68	1	0	1	0	1	1	1	0	0	1	0	172	0
409	0	1	1	1	64	1	0	1	0	0	1	0	1	1	0	0	59	1
410	1	1	0	0	52	1	1	1	0	1	1	1	0	0	0	0	172	0
411	0	1	1	1	51	1	1	1	0	1	0	0	0	1	0	1	103	0

2.3 Multiple myeloma data

The patients of multiply myeloma study was carried out in The Hematology Center, in the Main Military Clinical Hospital named after N.N.Burdenko to compare the response times in two patient groups of treatment. The difference in these groups is in the fact that the first group received the chemotherapy together Bortezomibe, which is marketed as Velcade by Millennium Pharmaceuticals, other received drugs without Bortezomibe. The data include 60 patients with 56 observations and 4 censored patients, of two treatment groups: chemotherapy without Bortezomibe ($z_1 = 0$) and chemotherapy together with Bortezomibe ($z_1 = 1$). The first group consists 38 observations with one observed censoring and other include 22 observation with 3 observed censoring. In addition to variable of treatment, three such variables are given in the data that consider in their work, namely: Type of response (the value $z_2 = 0$ corresponds to the progression of the disease, $z_2 = 1$ - the general response); sex ($z_3 = 1$ -male, $z_3 = 0$ -female); and age in years (z_4). Detail of the data are presented in Semenova & Bitukov [106]. Note that 4 observation are independent randomly censored observations. The data are given in following Table.

Table 3: Multiply myeloma data

pt	z_1	z_2	z_3	z_4	survtime	censure	pt	z_1	z_2	z_3	z_4	survtime	censure
1	1	1	1	64	61	1	31	1	1	0	51	5	1
2	1	0	1	81	50	1	32	1	0	1	81	30	1
3	1	1	0	71	2	1	33	1	0	1	58	25	1
4	1	0	1	69	36	1	34	1	0	1	77	39	1
5	1	0	0	74	14	1	35	1	1	1	65	6	1
6	1	0	1	83	27	1	36	1	0	0	69	83	1
7	1	1	0	46	1	1	37	1	1	1	52	24	1
8	1	1	1	80	4	1	38	1	1	1	49	3	1
9	1	0	0	58	27	1	39	0	1	1	61	7	1
10	1	0	1	50	115	0	40	0	1	1	45	2	1
11	1	0	1	85	13	1	41	0	1	0	66	7	1
12	1	1	1	56	2	1	42	0	1	0	60	2	1
13	1	1	1	57	3	1	43	0	1	1	68	262	0
14	1	1	1	71	25	1	44	0	1	0	81	81	1
15	1	1	1	64	4	1	45	0	0	0	79	33	1
16	1	0	0	57	62	1	46	0	0	1	95	214	1
17	1	1	0	71	9	1	47	0	0	0	85	57	1
18	1	0	0	56	10	1	48	0	0	1	89	17	1
19	1	1	1	55	7	1	49	0	1	0	75	26	1
20	1	0	1	75	54	1	50	0	0	1	47	7	1
21	1	0	1	75	62	1	51	0	0	1	75	30	1
22	1	1	1	64	3	1	52	0	0	1	66	2	1
23	1	1	1	61	26	1	53	0	1	1	76	26	1
24	1	0	1	72	22	1	54	0	1	1	37	20	0
25	1	0	1	59	46	1	55	0	0	0	57	5	0
26	1	1	1	77	3	1	56	0	1	0	73	8	1
27	1	1	1	66	16	1	57	0	0	0	79	127	1
28	1	0	0	46	10	1	58	0	0	1	87	149	1
29	1	1	0	55	25	1	59	0	1	1	65	10	1
30	1	1	1	48	6	1	60	0	1	1	61	8	1

Bibliography

- [1] O. Aalen. Non-parametric inference for a family of counting processes. *The Annals of Statistics*, 6(4):701–726, 1978. [59](#)
- [2] O. Aalen. A model for nonparametric regression analysis of counting processes. In *Mathematical statistics and probability theory*, pages 1–25. Springer, 1980. [49](#)
- [3] N. Aguirre and M. S. Nikulin. Chi-squared goodness-of-fit test for the family of logistic distributions. *Kybernetika*, 30(3):214–222, 1994. [5](#), [73](#)
- [4] M. G. Akritas. Pearson-type goodness-of-fit tests: The univariate case. *Journal of the American Statistical Association*, 83(401):222–230, 1988. [5](#), [10](#), [41](#), [46](#)
- [5] P. K. Andersen and R. D. Gill. Cox’s regression model for counting processes: A large sample study. *The annals of statistics*, pages 1100–1120, 1982. [1](#)
- [6] P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding. *Statistical models based on counting processes*. Springer Verlag, New York, 1993. [1](#), [43](#), [59](#), [65](#), [69](#), [70](#)
- [7] T. W. Anderson and D. A. Darling. A test of goodness-of-fit. *Journal of the American Statistical Association*, 49(268):765–769, 1954. [4](#)
- [8] T. Aven and U. Jensen. *Stochastic models in reliability*. Springer Verlag, New York, 1999. [60](#)
- [9] V. Bagdonavičius. A statistical test of a model of additive accumulation of damage. *Theory of Probability & Its Applications*, 23(2):385–390, 1978. [2](#), [49](#), [51](#), [55](#)
- [10] V. Bagdonavičius. Accelerated life models when the stress is not constant. *Kybernetika*, 26(4):289–295, 1990. [51](#), [56](#)
- [11] V. Bagdonavičius and M. S. Nikulin. *Semiparametric models in accelerated life testing*. Queen’s University, Kingston, 1995. [51](#)
- [12] V. Bagdonavičius and M. S. Nikulin. Generalized Proportional Hazards model based on modified partial likelihood. *Lifetime Data Analysis*, 5(4):329–350, 1999. [51](#)
- [13] V. Bagdonavičius and M. S. Nikulin. Goodness-of-fit tests for accelerated life models. *Goodness-of-fit Tests and Model Validity*, pages 281–300, 2002. [55](#)

- [14] V. Bagdonavičius and M. S. Nikulin. *Accelerated life models: Modeling and statistical analysis*. Chapman & Hall/CRC, Boca Raton, 2002. [2](#), [43](#), [49](#), [51](#), [59](#), [63](#), [122](#)
- [15] V. Bagdonavičius and M. S. Nikulin. Statistical modeling in survival analysis and its influence on the duration analysis. In N. Balakrishnan and C. R. Rao, editors, *Handbook of statistics: Advances in survival analysis*, pages 411–429. Elsevier, New York, 2004. [51](#)
- [16] V. Bagdonavičius and M. S. Nikulin. Chi-squared goodness-of-fit test for right censored data. *International Journal of Applied Mathematics and Statistics*, 24(SI-11A):30–50, 2011. [5](#), [10](#), [41](#), [48](#)
- [17] V. Bagdonavičius and M. S. Nikulin. Chi-squared tests for general composite hypotheses from censored samples. *Comptes Rendus Mathematique*, 349(3):219–223, 2011. [50](#)
- [18] V. Bagdonavičius and M. S. Nikulin. Goodness-of-fit test for homogeneous markov processes. *Comptes Rendus Mathematique*, 351(3):149–154, 2013. [5](#)
- [19] V. Bagdonavičius, J. Kruopis, and M. S. Nikulin. *Non-parametric tests for complete data*. Wiley-ISTE, New York, 2010. [5](#), [11](#), [48](#), [73](#), [106](#)
- [20] V. Bagdonavičius, J. Kruopis, and M. S. Nikulin. *Non-parametric tests for censored Data*. John Wiley & Sons, New York, 2011. ISBN 9781118601983. [41](#), [49](#), [50](#), [67](#)
- [21] V. Bagdonavičius, R. J. Levulienė, and M. S. Nikulin. Chi-squared goodness-of-fit tests for parametric accelerated failure time models. *Communications in Statistics-Theory and Methods*, 42(15):2768–2785, 2011. [5](#), [48](#), [60](#), [67](#), [72](#), [73](#)
- [22] V. Bagdonavičius, R. Levulienė, M. S. Nikulin, and X. Q. Tran. On chi-squared type tests and their applications in survival analysis and reliability. *Zapiski Nauchnykh Seminarov*, 408:43–61, 2012. [50](#), [67](#), [72](#)
- [23] V. Bagdonavičius, R. J. Levulienė, and M. S. Nikulin. Exact goodness-of-fit tests for shape-scale families and type II censoring. *Lifetime data analysis*, 19(3):413–435, 2013. [5](#), [73](#)
- [24] V. Bagdonavičius, M. S. Nikulin, and R. Levulienė. On the validity of the Weibull-Gnedenko Model. *Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference*, page 259, 2013. [5](#)
- [25] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, New York, 1968. [49](#)
- [26] Z. W. Birnbaum. Distribution-free tests of fit for continuous distribution functions. *The Annals of Mathematical Statistics*, 24:1–8, 1953. [4](#)
- [27] L. N. Bol’shev and M. Mirvaliev. Chi square goodness-of-fit test for the Poisson, Binomial and Negative Binomial distributions. *Theory of Probability & Its Applications*, 23(3):461–474, 1978. [18](#)
- [28] L. N. Bol’shev and N. V. Smirnov. *Tables of mathematical statistics*, 1983. [80](#)

- [29] Ø. Borgan. Maximum likelihood estimation in parametric counting process models, with applications to censored failure time data. *Scandinavian Journal of Statistics*, pages 1–16, 1984. [43](#), [44](#), [63](#)
- [30] Y. Q. Chen. Accelerated hazards regression model and its adequacy for censored survival data. *Biometrics*, 57(3):853–860, 2001. [1](#), [51](#)
- [31] H. Chernoff and E. L. Lehmann. The use of maximum likelihood estimates in χ^2 tests for goodness of fit. *The Annals of Mathematical Statistics*, 25(3):579–586, 1954. [4](#), [17](#), [20](#), [22](#), [23](#), [25](#)
- [32] D. M. Chibisov. Certain chi-square type tests for continuous distributions. *Theory of Probability & Its Applications*, 16(1):1–22, 1971. [20](#)
- [33] D. Collett. *Modelling Survival Data in Medical Research*. CRC press, New York, 2003. [2](#), [13](#), [49](#), [106](#), [114](#)
- [34] D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 187–220, 1972. [1](#), [49](#), [51](#), [52](#)
- [35] D. R. Cox. Partial likelihood. *Biometrika*, 62(2):269–276, 1975. [1](#)
- [36] D. R. Cox. Karl Pearson and the Chi-squared test. In C. Huber-Carol, N. Balakrishnan, M. S. Nikulin, and M. Mesbah, editors, *Goodness-of-Fit Tests and Model Validity*, pages 3–8. Springer, New York, 2002. [4](#)
- [37] H. Cramér. *Mathematical Methods of Statistics*. Princeton University Press, Princeton, N J, 1946. [4](#), [17](#), [20](#), [21](#)
- [38] F. C. Drost. Asymptotics for generalized chi-square goodness-of-fit tests. *CWI Tracts*, 48: 1–104, 1988. [5](#), [18](#)
- [39] F. C. Drost. Generalized chi-square goodness-of-fit tests for location-scale models when the number of classes tends to infinity. *The Annals of Statistics*, pages 1285–1300, 1989. [5](#), [18](#)
- [40] R. M. Dudley. On χ^2 tests of composite hypotheses. *Theory of Probability & Its Applications, Banach Center Publications*, 5(3):75–87, 1979. [18](#)
- [41] J. F. Dupuy. Accelerated Failure Time Models: A Review. *International Journal of Performance Engineering*, 10(1):23–40, 2014. [5](#), [10](#)
- [42] K. O. Dzaparidze and M. S. Nikulin. On a modification of the standard statistics of Pearson. *Theory of Probability & Its Applications*, 19(4):851–853, 1974. [18](#), [30](#)
- [43] K. El Fassi. *Sur l'estimation non paramétrique de la fonction d'égalisation équipercetile. Application à la qualité de vie*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2009. [51](#), [56](#)

- [44] R. A. Fisher and others. On a property connecting the χ^2 measure of discrepancy with the method of maximum likelihood. *Journal of the Royal Statistical Society*, 87:442–449, 1928. [4](#), [17](#), [20](#), [21](#), [98](#)
- [45] T. R. Fleming and D. P. Harrington. *Counting processes and survival analysis*. John Wiley & Sons, New York, 1991. [59](#)
- [46] E. J. Freireich, E. Gehan, E. Frei, L. R. Schroeder, I. J. Wolman, R. Anbari, E. O. Burgert, S. D. Mills, D. Pinkel, O. S. Selawry, et al. The effect of 6-mercaptopurine on the duration of steroid-induced remissions in acute leukemia: A model for evaluation of other potentially useful therapy. *Blood*, 21(6):699–716, 1963. [iii](#), [iv](#), [11](#), [126](#), [152](#), [153](#)
- [47] L. Gerville-Réache, M. S. Nikulin, R. Tahir, et al. On reliability approach and statistical inference in demography. *Journal of Reliability and Statistical Studies*, 5:37–49, 2012. [5](#)
- [48] H. Goual and N. Seddik-Ameur. Chi-squared type test for the AFT-Generalized Inverse Weibull distribution. *Communications in Statistics-Theory and Methods*, 43(13):2605–2617, 2014. [5](#), [67](#)
- [49] P. Greenwood and M. S. Nikulin. Application of tests of chi-square type. *Journal of Mathematical Sciences*, 43(6):2776–2791, 1988. [24](#)
- [50] P. E. Greenwood and M. S. Nikulin. *A guide to chi-squared testing*. Wiley, New York, 1996. [20](#), [23](#), [24](#), [28](#)
- [51] M. G. Habib and D. R. Thomas. Chi-square goodness-of-fit tests for randomly censored data. *The Annals of Statistics*, 14(2):759–765, 1986. [5](#), [10](#), [41](#)
- [52] F. Haghighi and M. S. Nikulin. A chi-squared test for power Generalized Weibull family for the head-and-neck cancer censored data. *Journal of Mathematical Sciences*, 133(3):1333–1341, 2006. [5](#)
- [53] N. L. Hjort. Goodness-of-fit tests in models for life history data based on cumulative hazard rates. *The Annals of Statistics*, pages 1221–1258, 1990. [10](#), [41](#), [43](#), [44](#), [46](#)
- [54] N. L. Hjort. On inference in parametric survival data models. *International Statistical Review/Revue Internationale de Statistique*, pages 355–387, 1992. [2](#), [61](#)
- [55] M. Hollander and E. A. Peña. A chi-squared goodness-of-fit test for randomly censored data. *Journal of the American Statistical Association*, 87(418):458–463, 1992. [5](#), [10](#), [41](#)
- [56] D. W. Hosmer, S. Lemeshow, and S. May. *Applied survival analysis: Regression modeling of time to event data (second edition)*. Wiley-Interscience, New Jersey, 2008. [1](#)
- [57] F. Hsieh. On heteroscedastic hazards regression models: Theory and application. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(1):63–79, 2001. [49](#)
- [58] A. Hsuan and D. S. Robson. The χ^2 -goodness-of-fit tests with moment type estimators. *Communications in Statistics - Theory and Methods*, 5(15):1509–1519, 1976. [18](#), [20](#)

- [59] C. Huber, V. Solev, and F. Vonta. Interval censored and truncated data: Rate of convergence of NPMLE of the density. *Journal of Statistical Planning and Inference*, 139(5):1734–1749, 2009. 8, 60
- [60] J. D. Kalbfleisch and R. L. Prentice. *The statistical analysis of failure time data*. John Wiley & Sons, New York, 2011. 2
- [61] E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958. 9
- [62] J. A. Kim and N. Balakrishnan. Non-parametric tests for independence between lifetimes and covariates from censored bivariate normal samples. *Communications in Statistics & Simulation and Computation*, 34(3):685–710, 2005. 49
- [63] J. H. Kim. Chi-square goodness-of-fit tests for randomly censored data. *The Annals of Statistics*, 21(3):1621–1639, 1993. 10
- [64] J. Klein and M. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag, New York, 1997. 12, 60
- [65] D. G. Kleinbaum and M. Klein. *Survival Analysis: A Self-Learning Text*. Springer-Verlag, New York, 2005. 2, 12, 49, 127
- [66] A. N. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione (on the empirical determination of each law distribution). *Giornale dell’Istituto Italiano degli Attuari*, 4(1):83–91, 1933. 4
- [67] J. F. Lawless. *Statistical models and methods for lifetime data (Second edition)*. Wiley, New York, 2003. 2, 12, 60
- [68] L. LeCam, C. Mahan, and A. Singh. An extension of a theorem of H. Chernoff and E. L. Lehmann, chapter in: Recent advances in statistics. *Academic Press, Orlando*, 29:303–332, 1983. 5, 18
- [69] E. T. Lee and J. Wang. *Statistical methods for survival data analysis*. Wiley, New York, 2003. 2, 12, 49
- [70] B. Yu. Lemeshko, S. B. Lemeshko, and S. N. Postovalov. Comparative analysis of the power of goodness-of-fit tests for near competing hypotheses. I. The verification of simple hypotheses. *Journal of Applied and Industrial Mathematics*, 3(4):462–475, 2009. 37
- [71] B. Yu. Lemeshko, S. B. Lemeshko, and S. N. Postovalov. Comparative analysis of the power of goodness-of-fit tests for near competing hypotheses. II. Verification of complex hypotheses. *Journal of Applied and Industrial Mathematics*, 4(1):79–93, 2010.
- [72] B. Yu. Lemeshko, S. B. Lemeshko, and S. N. Postovalov. Statistic distribution models for some nonparametric goodness-of-fit tests in testing composite hypotheses. *Communications in Statistics—Theory and Methods*, 39(3):460–471, 2010. 37, 40, 80

- [73] D.Y. Lin and Z. Ying. Semiparametric analysis of general Additive-Multiplicative hazard models for counting processes. *The annals of Statistics*, pages 1712–1734, 1995. [1](#), [49](#), [55](#)
- [74] Yu. V. Linnik. Method of least squares and fundamentals of the theory of processing observations. *Fizmatgiz, Moscow*, 1958. [6](#), [40](#)
- [75] T. Martinussen and T. H. Scheike. *Dynamic regression models for survival data*. Springer, New York, 2006. [2](#), [8](#), [50](#), [51](#)
- [76] C. E. McCulloch. Relationships among some chi-square goodness of fit statistics. *Communications in Statistics-Theory and Methods*, 14(3):593–603, 1985. [30](#)
- [77] W. Q. Meeker and L. A. Escobar. *Statistical methods for reliability data*. Wiley-Interscience, New York, 1998. [2](#), [10](#), [51](#)
- [78] R. G. Miller. Least squares regression with censored data. *Biometrika*, 63(3):449–464, 1976. [2](#)
- [79] D. S. Moore. A chi-square statistic with random cell boundaries. *The Annals of Mathematical Statistics*, 42(1):147–156, 1971. [25](#)
- [80] D. S. Moore. Generalized inverses, Wald’s method, and the construction of chi-squared tests of fit. *Journal of the American Statistical Association*, 72(357):131–137, 1977. [5](#)
- [81] C. P. Nelson, P. C. Lambert, I. B. Squire, and D. R. Jones. Flexible parametric models for relative survival, with application in coronary heart disease. *Statistics in medicine*, 26(30):5486–5498, 2007. [1](#)
- [82] W. Nelson. *Accelerated life testing: statistical models, data analysis and test plans*. John Wiley and Sons, New York, 1990. [51](#)
- [83] M. S. Nikulin. Chi-square test for continuous distributions with shift and scale parameters. *Theory of Probability & Its Applications*, 18(3):559–568, 1973. [4](#), [5](#), [18](#), [24](#), [25](#)
- [84] M. S. Nikulin. About chi-square goodness-of-fit test for continuous distributions. *Teor. Veroyatnost. i Primenen*, 18(3):675–676, 1973.
- [85] M. S. Nikulin. Chi-square test for normality. In *Proceedings of International Vilnius Conference on Probability Theory and Mathematical Statistics*, volume 2, pages 119–122, 1973. [4](#), [5](#), [18](#), [30](#)
- [86] M. S. Nikulin. On a chi-square test for continuous distribution. *Theory of Probability and its Application*, 19:638–639, 1973. [24](#)
- [87] M. S. Nikulin and M. Mirvaliev. Goodness of fit chi-squared type criterions. *Industrial Laboratory*, pages 280–291, 1992. [18](#)
- [88] M. S. Nikulin and X. Q. Tran. On chi-squared testing in accelerated trials. *International Journal of Performability Engineering*, 10(1):53–62, 2014. [5](#), [12](#), [50](#), [67](#), [72](#)

- [89] M. S. Nikulin and H-D. I. Wu. *The Cox model and its applications*. Springer. To appear, New York, 2014. [50](#), [51](#)
- [90] M. S. Nikulin, L. Gerville-Réache, and X. Q. Tran. On Chi-squared goodness-of-fit test for Normality. In V. Couallier, L. Gerville-Réache, H. Huber-Carol, N. Limnios, and M. Mesbah, editors, *Statistical Models and Methods for Reliability and Survival Analysis*, pages 213–227. Wiley Online Library, New York, 2013. [18](#), [25](#), [33](#), [101](#)
- [91] M. S. Nikulin, L. Gerville-Réache, and X. Q. Tran. *Statistical Models and Methods for Reliability and Survival Analysis*, chapter On Chi-Squared Goodness-of-Fit Test for Normality, pages 213–227. Wiley Online Library, New York, 2013. [30](#)
- [92] M. S. Nikulin, X. Q. Tran, and M. R. Tahir. Goodness-of-fit test for Hypertabastic survival models with applications. *Zapiski Nauchnykh Seminarov (accepted)*, 2014. [5](#), [73](#)
- [93] J. Ogawa. *A Note on the Optimal Spacing of the Systematic Statistics Normal Distribution*. University of Calgary, Department of Mathematics, Statistics and Computing Science, 1974. [34](#)
- [94] J. Orbe, E. Ferreira, and V. Núñez-Antón. Comparing Proportional Hazards and Accelerated Failure Time models for survival analysis. *Statistics in medicine*, 21(22):3493–3510, 2002. [58](#)
- [95] K. Pearson. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175, 1900. [4](#), [17](#)
- [96] R. L. Plackett. Karl Pearson and the chi-squared test. *International statistical review*, 51(1):59–72, 1983. [17](#)
- [97] K. C. Rao and B. S. Robson. A chi-square statistic for goodness-of-fit tests within the Exponential family. *Communications in Statistics-Theory and Methods*, 3(12):1139–1153, 1974. [5](#), [18](#), [24](#), [25](#)
- [98] A. R. Roy. *On χ^2 -statistics with variable intervals*. PhD thesis, Department of Statistics, Stanford University., 1956. [17](#)
- [99] P. Royston. Flexible parametric alternatives to the Cox model, and more. *Stata Journal*, 1(1):1–28, 2001. [1](#)
- [100] P. Royston. Flexible parametric alternatives to the Cox model: Update. *The Stata Journal*, 4(1):98–101, 2004. [1](#)
- [101] P. Royston and P. C. Lambert. Flexible parametric survival analysis using stata: Beyond the Cox model. *Stata Press BOOKs*, 2006. [2](#)
- [102] P. Royston and M. K. Parmar. Flexible parametric Proportional Hazards and Proportional Odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in medicine*, 21(15):2175–2197, 2002. [1](#)

- [103] P. Royston and W. Sauerbrei. *Multivariable model-building: a pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables*. Wiley, New York, 2008. [13](#), [113](#), [114](#)
- [104] N. Saaidia. *Sur les familles des lois de fonction de hasard unimodale: Applications en fiabilité et analyse de survie*. PhD thesis, Bordeaux 1, 2013. [67](#), [73](#)
- [105] W. Sauerbrei and M. Schumacher. A bootstrap resampling procedure for model building: Application to the Cox regression model. *Statistics in medicine*, 11(16):2093–2109, 1992. [iii](#), [iv](#), [13](#), [49](#), [113](#), [114](#), [127](#), [152](#), [153](#)
- [106] M. Semenova and A. Bitukov. Parametric models in the analysis of patients with multiple myeloma. In B. Lemeshko, M. Nikulin, and N. Balakrishnan, editors, *Applied methods of statistical analysis. Applications in survival analysis, reliability and quality control*, pages 250–256. Novosibirsk, September 2013. [iii](#), [iv](#), [14](#), [119](#), [140](#), [152](#), [153](#)
- [107] S. S. Shapiro and R. S. Francia. An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67(337):215–216, 1972. [37](#)
- [108] S. S. Shapiro and M. B. Wilk. An analysis of variance test for the exponential distribution (complete samples). *Technometrics*, 14(2):355–370, 1972. [37](#)
- [109] P. S. Shen. Aalen’s additive risk model for left-truncated and right-censored data. *Communications in Statistics-Simulation and Computation*, 43(5):1006–1019, 2014. [8](#)
- [110] N. V. Smirnov. Sur la distribution de ω^2 . *C. R. Acad. Sci.*, 202:449–452, 1936. [4](#)
- [111] S. M. Stigler. Karl Pearson theoretical errors and the advances they inspired. *Statistical Science*, 23(2):261–271, 2008. [17](#)
- [112] W. Stute. Consistent estimation under random censorship when covariables are present. *Journal of Multivariate Analysis*, 45(1):89–103, 1993. [58](#)
- [113] M. A. Tabatabai, Z. Bursac, D. K. Williams, and K. P. Singh. Hypertabastic survival model. *Theor Biol Med Model*, 4:40–53, 2007. [3](#), [73](#), [74](#)
- [114] M. A. Tabatabai, W. M. Eby, N. Nimeh, H. Li, and K. P. Singh. Clinical and multiple gene expression variables in survival analysis of breast cancer: Analysis with the Hypertabastic survival model. *BMC medical genomics*, 5(1):63–76, 2012. [73](#)
- [115] M. A. Tabatabai, W. M. Eby, N. Nimeh, and K. P. Singh. Role of metastasis in Hypertabastic survival analysis of breast cancer: Interaction with clinical and gene expression variables. *Cancer Growth & Metastasis*, 5:1–17, 2012. [3](#), [73](#)
- [116] M. R. Tahir. *On validation of parametric models applied in survival analysis and reliability*. PhD thesis, Bordeaux 1, 2012. [5](#), [67](#), [73](#)
- [117] A. Tsiatis. *Semiparametric theory and missing data*. Springer, 2007. [1](#)

-
- [118] A. W. Van der Vaart. *Asymptotic statistics*. Cambridge university press, New York, 1998. [5](#), [18](#)
- [119] V. Voinov, R. Alloyarova, and N. Pya. Recent achievements in modified chi-squared goodness-of-fit testing. In *Statistical Models and Methods for Biomedical and Technical Systems*, pages 241–258. Springer, 2008. [18](#)
- [120] V. Voinov, N. Pya, and R. Alloyarova. A comparative study of some modified chi-squared tests. *Communications in Statistics - Simulation and Computation*, 38(2):355–367, 2009. [5](#), [101](#)
- [121] V. Voinov, M. S. Nikulin, and N. Balakrishnan. *Chi-Squared Goodness of Fit Tests with Applications*. Academic Press, New York, 2013. [5](#), [6](#), [30](#), [40](#), [48](#)
- [122] V. Voinov, N. Pya, N. Shapakov, and Y. Voinov. Goodness-of-fit tests for the Power Generalized Weibull probability distribution. *Communications in Statistics-Simulation and Computation*, 42(5):1003–1012, 2013. [5](#)
- [123] V. G. Voinov. On optimality of the Rao–Robson–Nikulin test. *Industrial Laboratory (in Russian)*, 72(3):65–70, 2006. [5](#)
- [124] R. von Mises and H. Geiringer. *Mathematical theory of probability and statistics*. Academic Press, New York, 1964. [4](#)
- [125] G. S. Watson. The χ^2 goodness-of-fit test for normal distributions. *Biometrika*, pages 336–348, 1957. [25](#)
- [126] B. Zhang. A chi-squared goodness-of-fit test for logistic regression models based on case-control data. *Biometrika*, 86(3):531–539, 1999. [5](#), [73](#)

Dynamic regression models and their applications in survival and reliability analysis

Thesis by
TRAN Xuan Quang

This thesis was designed to explore the dynamic regression models, assessing the statistical inference for the survival and reliability data analysis. These dynamic regression models that we have been considered including the parametric proportional hazards and accelerated failure time models contain the possibly time-dependent covariates. We discussed the following problems in this thesis.

At first, we presented a generalized chi-squared test statistics \mathbf{Y}_n^2 that is a convenient to fit the survival and reliability data analysis in presence of three cases: complete, censored and censored with covariates. We described in detail the theory and the mechanism to used of \mathbf{Y}_n^2 test statistic in the survival and reliability data analysis. Next, we considered the flexible parametric models, evaluating the statistical significance of them by using \mathbf{Y}_n^2 and log-likelihood test statistics. These parametric models include the accelerated failure time (AFT) and a proportional hazards (PH) models based on the Hypertabastic distribution. These two models are proposed to investigate the distribution of the survival and reliability data in comparison with some other parametric models. The simulation studies were designed, to demonstrate the asymptotically normally distributed of the maximum likelihood estimators of Hypertabastic's parameter, to validate of the asymptotically property of \mathbf{Y}_n^2 test statistic for Hypertabastic distribution when the right censoring probability equal 0% and 20%.

In the last chapter, we applied those two parametric models above to three scenes of the real-life data. The first one was done the data set given by Freireich et al. [46] on the comparison of two treatment groups with additional information about log white blood cell count, to test the ability of a therapy to prolong the remission times of the acute leukemia patients. It showed that Hypertabastic AFT model is an accurate model for this dataset. The second one was done on the brain tumour study with malignant glioma patients, given by Sauerbrei & Schumacher [105]. It showed that the best model is Hypertabastic PH on adding five significance covariates. The third application was done on the data set given by Semenova & Bitukov [106] on the survival times of the multiple myeloma patients. We did not propose an exactly model for this dataset. Because of that was an existing one intersection of survival times. We, therefore, suggest fitting other dynamic model as Simple Cross-Effect model for this dataset.

Keywords: Accelerated Failure Time model, Covariates, Cox model, Dynamic regression models, Flexible parametric models, Generalized Chi-squared statistic, Goodness-of-fit tests, Hypertabastic survival models, Leukemia cancer, Malignant glioma cancer, Multiple Myeloma cancer, Proportional Hazards model, Survival data analysis.

Les modèles de régression dynamique et leurs applications en analyse de survie et fiabilité

Thèse par
TRAN Xuan Quang

Cette thèse a été conçue pour explorer les modèles dynamiques de régression, d'évaluer les inférences statistiques pour l'analyse des données de survie et de fiabilité. Ces modèles de régression dynamiques que nous avons considérés, y compris le modèle des hasards proportionnels paramétriques et celui de la vie accélérée avec les variables qui peut-être dépendent du temps. Nous avons discuté des problèmes suivants dans cette thèse.

Nous avons présenté tout d'abord une statistique de test du chi-deux généralisée \mathbf{Y}_n^2 qui est adaptative pour les données de survie et fiabilité en présence de trois cas, complètes, censurées à droite et censurées à droite avec les covariables. Nous avons présenté en détail la forme pratique de \mathbf{Y}_n^2 statistique en analyse des données de survie. Ensuite, nous avons considéré deux modèles paramétriques très flexibles, d'évaluer les significations statistiques pour ces modèles proposées en utilisant \mathbf{Y}_n^2 statistique. Ces modèles incluent du modèle de vie accélérés (AFT) et celui de hasards proportionnels (PH) basés sur la distribution de Hypertabastic. Ces deux modèles sont proposés pour étudier la distribution de l'analyse de la durée de survie en comparaison avec d'autres modèles paramétriques. Nous avons validé ces modèles paramétriques en utilisant \mathbf{Y}_n^2 . Les études de simulation ont été conçues.

Dans le dernier chapitre, nous avons proposé les applications de ces modèles paramétriques à trois données de bio-médicale. Le premier a été fait les données étendues des temps de rémission des patients de leucémie aiguë qui ont été proposées par Freireich et al. [46] sur la comparaison de deux groupes de traitement avec des informations supplémentaires sur le log du blanc du nombre de globules. Elle a montré que le modèle Hypertabastic AFT est un modèle précis pour ces données. Le second a été fait sur l'étude de tumeur cérébrale avec les patients de gliome malin, ont été proposées par Sauerbrei & Schumacher [105]. Elle a montré que le meilleur modèle est Hypertabastic PH à l'ajout de cinq variables de signification. La troisième demande a été faite sur les données de Semenova & Bitukov [106], à concernant les patients de myélome multiple. Nous n'avons pas proposé un modèle exactement pour ces données. En raison de cela était les intersections de temps de survie. Par conséquent, nous vous conseillons d'utiliser un autre modèle dynamique que le modèle de la Simple Cross-Effect à installer ces données.

Mots clés: Analyse de survie, Covariables, Modèle de Cox, Modèle des Hasards Proportionnels, Modèles de survie de Hypertabastic, Modèles des paramétriques flexibles, Modèle de régression dynamique, Modèle de vie accélérée, Cancer de leucémie, Cancer de gliome malin, Cancer de myélome multiple, Test d'ajustement, Statistique du test du Chi-deux généralisée.