
UNIVERSITI PUTRA MALAYSIA

# NON-PARAMETRIC AND PARAMETRIC ESTIMATIONS OF CURE FRACTION USING RIGHT- AND INTERVAL-CENSORED DATA

BADER ALJAWDI

IPM 2011 7

# NON-PARAMETRIC AND PARAMETRIC ESTIMATIONS OF CURE FRACTION USING RIGHT- AND INTERVAL-CENSORED DATA

**BY**

**BADER ALJAWDI**

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Doctor of Philosophy**

**August 2011**

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirements for the degree of Doctor of Philosophy

**NON-PARAMETRIC AND PARAMETRIC ESTIMATIONS OF CURE FRACTION USING RIGHT- AND INTERVAL- CENSORED DATA**

By

**BADER ALJAWDI**

**August 2011**

**Chair: Associate Professor Mohd Rizam Abu Bakar, PhD**

**Faculty: Institute for Mathematical Research**

In cancer studies many patients will never experience the event of concern and are hence considered as cured. Therefore, the main interest of the cancer trials is in determining the fraction of cured patients so as to obtain the trends in the survival of cancer patients. Survival models incorporate the cure fraction in the analysis; namely cure rate models, are being widely used in cancer studies, where one of the most common cure models that can be employed in cancer clinical trials is the bounded cumulative hazard (BCH) model. In this thesis, we considered two methods via the expectation maximization (EM) algorithm for cure rate estimation based on the BCH model using the two censoring types common to cancer clinical trials; namely, right and interval censoring.

Then, a series of simulation studies was conducted to evaluate the performance of the proposed estimation approaches.

This research investigated the non-parametric maximum likelihood estimation method for cure rate estimation by considering two common estimators for the survival function: 1) The Kaplan Meier (KM) estimator, which is suitable for the right censoring case; and 2) The Turnbull Estimator, which is suitable for the interval type of data censoring. The parametric maximum likelihood estimation of the cure fraction was also investigated under the same circumstances considering two scenarios: 1) when covariates were excluded from the analysis. In this case, the estimation was developed based on the exponential and Weibull distributions using the right and interval censoring types; and 2) when covariates were incorporated into the analysis through the scale parameter of the exponential distribution only using the same types of data censoring.

The major research findings were as follows: 1) the non-parametric and parametric estimation methods using the right and interval censoring types produced highly efficient cure rate parameters when the censoring rate was decreased to the minimum possible; 2) Non-parametric estimation of the cure fraction using interval censored data based on Turnbull estimator resulted in more precise cure fraction than the

Kaplan Meier estimator considering the interval midpoint to represent the exact life time; 3) The parametric estimation of the cure fraction based on the exponential distribution and right and interval censoring types produced more consistent estimates than the Weibull distribution especially in case of heavy censoring; 4) Parametric estimation of the cure fraction was more efficient when some covariates had been involved in the analysis than when covariates had been excluded; and 5) the non-parametric estimation method is the viable alternative to the parametric one when the data set contains substantial censored observations while in the case of low censoring the parametric method is more attractive.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

**PENGANGGARAN TAK BERPARAMETER DAN BERPARAMETER BAGI PECAHAN SEMBUH MENGGUNAKAN DATA KANAN DAN SELANG TERTAPIS**

Oleh

**BADER ALJAWDI**

**Ogos 2011**

**Pengerusi: Profesor Madya Mohd Rizam Abu Bakar, PhD**

**Fakulti: Institut Penyelidikan Matematik**

Dalam kajian penyakit barah kebanyakan pesakit mungkin tidak mengalami peristiwa yang dikaji maka mereka dianggap sebagai sembuh. Oleh itu minat kajian bagi percubaan penyakit barah adalah untuk menentukan pecahan pesakit yang sembuh untuk memperlihatkan trend mandirian pesakit barah.

Didalam tesis ini, kami membina dua pendekatan analitik melalui algoritma pemaksimuman jangkaan (EM) untuk menganggar kadar sembuh berasaskan model bahaya terbatas melonggok (BCH) menggunakan dua jenis tapisan umum didalam percubaan klinikal penyakit barah iaitu tapisan kanan dan selang. Beberapa siri kajian

simulasi dijalankan untuk menilai kemampuan pendekatan penganggaran yang dicadangkan.

Kami bina kaedah penganggar kebolehjadian maksimum tak berparameter bagi penganggaran kadar sembuh melalui dua penganggar lazim bagi fungsi mandirian: 1) Penganggar Kaplan Meier (KM) yang sesuai untuk kes tapisan kanan, 2) Penganggar Turnbull yang sesuai untuk jenis tapisan selang.

Penganggaran kebolehjadian maksimum berparameter bagi pecahan sembuh turut dibina berdasarkan keadaan yang sama dengan mempertimbangkan dua scenario: 1) bila kovariat tidak disertakan semasa analisis. Dalam hal ini penganggaran dibina berdasarkan taburan eksponen dan Weibull menggunakan tapisan kanan dan tapisan selang dan 2) bila melibatkan kovariat melalui parameter skala taburan eksponen dan Weibull dengan menggunakan tapisan yang sama.

Hasil kajian adalah seperti berikut: 1) kaedah penganggaran tak berparameter dan berparameter menggunakan tapisan kanan dan selang menghasilkan penganggar kadar sembuh yang sangat cekap apabila kadar tapisan diperkecilkan kepada nilai yang sekecil mungkin; 2) Penganggaran tak berparameter bagi pecahan sembuh berasaskan

penganggar Turnbull meghasilkan pecahan sembuh yang lebih jitu berbanding penganggar Kaplan Meier apabila titik tengah selang mewakili masa hayat tepat; 3) Penganggaran berparameter bagi pecahan sembuh berdasarkan taburan eksponen dengan tapisan kanan dan selang menghasilkan penganggar yang konsisten jika dibandingkan dengan taburan Weibull terutamanya bagi kes tapisan yang banyak; 4) Penganggaran berparameter bagi pecahan sembuh lebih cekap apabila beberapa kovariat diambilkira dalam analisis berbanding dengan jika kovariat dikeluarkan; dan 5) Kaedah penganggaran tak berparameter adalah alternatif yang lebih berjaya jika dibandingkan dengan kaedah berparameter apabila data mengandungi cerapan tertapis yang cukup banyak manakala bagi kes bilangan tapisan yang rendah kaedah berparameter adalah lebih menarik.

**ACKNOWLEDGEMENTS**

*In The Name of ALLAH, the Most Merciful and Most Beneficent*

All praise do to Allah, Lord of the universe. Only by His grace and mercy this thesis can be completed.

This work was carried out with a hope to contribute towards the expansion of our currently limited knowledge on survival data analysis. The completion of this thesis would have been impossible if not for the assistance and direct involvement of so many kindhearted individuals. Thus, I am very much indebted to my previous mentors and I have no way of repaying such a debt except to express my sincerest gratitude.

First and foremost, I am very grateful to my supervisor Assoc. Prof. Dr. Hj. Mohd Rizam Abu Bakar, for his strong support, guidance, and patience for the very enriching and thought provoking discussions which helped to shape the thesis. I am also grateful to Prof. Dr. Noor Akma Ibrahim and Assoc. Prof. Dr. Habshah Midi in their capacities as members of committee. I am also indebted to the staff of the Institute for Mathematical Research, Universiti Putra Malaysia for their help and cooperation. Special thanks are extended to Dean and all committee members of School of Graduate Studies for providing financial support during the period of my study through Graduate Research Fellowship (GRF) Grant Scheme.

I wish to express my deepest gratitude to my parents, brothers and sisters for their prayers, continuous moral support and unending encouragement. Last but not least, I wish especially to acknowledge my beloved wife, Saba' Al-Atawnah, and my dearest son Abdullah and daughter Eman for their love, support, patience and understanding.

I certify that a Thesis Examination Committee has met on 10th of August 2011 to conduct the final examination of Bader Aljawdi on his thesis entitled "The non-parametric and parametric estimations of cure fraction using right- and interval-censored data" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

**Mohd Bakri bin Adam, PhD**
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Chairman)

**Jayanthi Arasan, PhD**
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Internal examiner)

**Ibrahim Bin Mohamed, PhD**
Associate Professor
Institute of Mathematical Sciences
Universiti Malaya
Malaysia
(External examiner)

**Muhammad Hanif Mian, PhD**
Professor
Departments of Mathematics
Lahore University of Management Sciences
Pakistan
(External examiner)

**NORITAH OMAR, PhD**
Associate Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 23 August 2011

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

**Mohd Rizam Abu Bakar, PhD**
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

**Noor Akma Ibrahim, PhD**
Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Member)

**Habsha Midi, PhD**
Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Member)

**HASANAH MOHD. GHAZALI, PhD**
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 23 August 2011

**DECLARATION**

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institution.

**BADER ALJAWDI**

Date: 10 August 2011

**TABLE OF CONTENTS**