NONPARAMETRIC COMPARISON OF REGRESSION FUNCTIONS IN TWO SAMPLES

By

Pimtong Sribera

A Dissertation Submitted in Partial
Fulfillment of the Requirements for the Degree of
Doctor of Philosophy (Statistics)
School of Applied Statistics
National Institute of Development Administration

ISBN 974-231-644-9

2004

NONPARAMETRIC COMPARISON OF REGRESSION FUNCTIONS IN TWO SAMPLES

Pimtong Sribera

School of Applied Statistics

The Examining Committee Approved This Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy (Statistics).

Associate Professor.....Committe Chairman

(Pachitjanut Siripanich, Ph.D.)

Professor.....Committee Co-Chairman

(Winfried Stute, Ph.D.)

Professor.....Committee

(Prachoom Suwattee, Ph.D.)

Associate Professor.....Committee

(Samruam Chongcharoen, Ph.D.)

.....Dean

(Lersan Bosuwan, Ph.D.)

Date...9-05-05...

ABSTRACT

Title of Dissertation : Nonparametric Comparison of Regression Functions in Two Samples

Author : Ms. Pimtong Srihera

Degree : Doctor of Philosophy (Statistics)

Year : 2004

In this work we provide a new methodology for comparing regression functions m_1 and m_2 in two samples. Since apart from smoothness no other (parametric) assumptions are required, our approach is based on a comparison of nonparametric estimators $\hat{m}_1$ and $\hat{m}_2$ of m_1 and m_2, respectively. Our test statistics are of the form

$$\hat{T} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} W\left(\frac{X_{1i} + X_{2j}}{2}\right)\left[\hat{m}_1\left(\frac{X_{1i} + X_{2j}}{2}\right) - \hat{m}_2\left(\frac{X_{1i} + X_{2j}}{2}\right)\right],$$

where X_{1i}, $1 \le i \le n_1$, and X_{2j}, $1 \le j \le n_2$, are the design variables in the two samples and W is a given weight function.

As our main results we obtain the limit distribution of $\hat{T}$ (properly standardized) under the null hypothesis $H_0 : m_1 = m_2$ and under local and global alternatives. We are also able to choose W so as to maximize power. Furthermore, the tests are asymptotically distribution-free under H_0 and shift- and scale-invariant.

In a simulation study we find out that our tests achieve the nominal level and have excellent power already for small to moderate sample sizes. As to proofs we heavily make use of results from empirical process theory, U-statistics and nonparametric curve estimation.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to Associate Professor Dr. Pachitjanut Siripanich, my advisor, for her precious advice, encouragement, kindness and confidence in me.

My deep gratitude and appreciation are extended to Professor Dr. Winfried Stute, my co-advisor, for his wonderful supervision, guidance, encouragement and patience. Thank you very much for making it all possible.

I also gratefully acknowledge the valuable comments and great effort in reading given to me by the committee members; Professor Dr. Prachoom Suwattee, Associate Professor Dr. Samruam Chongcharoen and Associate Professor Dr. Jirawan Jitthavech.

My hearty thanks go to Dr. Gerrit Eichner for providing me with support in the simulation study and to Mrs. Kerstin Lenk for making our preliminary work readable.

I take pleasure in expressing my thankfulness to the German Academic Exchange Service (DAAD) for giving me the chance to work with my co-advisor and supporting me throughout a year of study in Germany. I would also thank the National Institute of Development Administration (NIDA) and staff in the school of Applied Statistics for the generous support.

I would like to profoundly thank my two cheer-leader teams both in Thailand and in Germany, my friends and everybody around me. I was always recharged by your cheerfulness. My special thanks also go to Mr.Chainarong Phaiwitayasiritham for always being there.

Finally, I am greatly indebted to my parents. Everything has happened because you cherish me with your love and I would particularly like to thank my sisters and my brother for warming me with understanding, quiet but firm.

Pimtong Srihera

February 2005