Preface

Dedicated to Ronald Pyke

Empirical process theory has developed over the past 25 years into a very important and useful set of tools for mathematical statistics, and it continues to inspire interesting research on the boundary between probability and statistics. These notes aim to develop some of the key aspects of the modern theory, with a view toward statistical applications and uses. We will give a number of basic tools, inequalities, and some hints of the background weak convergence theory, and will then proceed to develop some of the basic Glivenko-Cantelli and Donsker theorems. Since the theoretical results are stated in terms of hypotheses on several kinds of bracketing numbers and covering numbers related to the "size" of the spaces and sets which index the processes, we will also develop and review some of the standard methods for controlling and bounding these covering numbers, including the combinatorial methods initiated by Vapnik and Chervonenkis, and the bracketing methods developed by Dudley and others. In Chapter 2 we develop some selected applications of empirical process theory in statistics. These applications include consistency of maximum likelihood estimators, methods for establishing rates of convergence, and limit theory for M- and Z- estimators. We have given a brief treatment of some of the bootstrap limit theory with indications of the connections to multiplier central limit theorems, and have also included a section on semiparametric mixture models. It seems to me that the applications of empirical process theory are just getting started. Many more are possible!

The subject of modern empircal processes developed rapidly in the 1980's and 1990's due to the contributions of A. Ya. Chervonenkis, Richard Dudley, Peter Gaenssler, Evarist Giné, Vladimir Koltchinskii, Mina Ossiander, David Pollard, Ronald Pyke, Michel Talagrand, V. N. Vapnik, A. W. van der Vaart, and Joel Zinn, among others. Development of the theory has been strongly affected by related developments in the areas of Gaussian process theory (with contributions by Richard Dudley, Xavier Fernique, Michel Ledoux, Michael Marcus, Michel Talagrand, and others), probability theory in Banach spaces (J. Hoffmann-Jørgensen, Jim Kuelbs, Michael Marcus), and (perhaps to a lesser degree) approximation theory (L. Birgé, E. M. Bronstein, M. S. Birman, A. N. Kolmogorov, P. Massart, M. Z. Solomjak, and others). These intertwining areas contribute to the fascination of the subject.

The present notes borrow considerably from the recent books by Richard Dudley (1999), de la Pena and Giné (1999), van der Vaart and Wellner (1996), and the lecture notes by Giné and Zinn (1986), and Giné (1997). Any mistakes are my own.

I sincerely thank the organizers of this Special Topics Course, Piet Groeneboom, and Delft University for the invitation to present this course. I owe thanks to Moulinath Banerjee and several students in Seattle, including Fadoua Balabdaoui, Ozgur Cetin, Lu Chen, Hao Liu, Marloes Maathuis, for suggestions and constructive criticism of preliminary versions of this material. I also owe thanks to the participants in the summer school *Corso Estivo di statistica e calcolo delle probabilità* held in Torgnon during July 2003, and the participants in the *Corso Estivo di Matematica Cortona* held in Cortona during August 2004 for numerous further corrections and comments. For further corrections in this version of the course given in Delft, I owe thanks to Willem Kruijer and Stefanie Donauer for spotting a considerable number of typographical and other errors.

Further comments, remarks, and corrections are very welcome and will be greatly appreciated.

Amsterdam, March 2005 Jon A. Wellner