Institute of Mathematical Statistics

LECTURE NOTES — MONOGRAPH SERIES

ESTIMATING FUNCTIONS: A SYNTHESIS OF LEAST SQUARES AND MAXIMUM LIKELIHOOD METHODS

V.P. Godambe
University of Waterloo

ABSTRACT

The development of the modern theory of estimating functions is traced from its inception. It is shown that this development has brought about a synthesis of the two historically important methodologies of estimation namely, the 'least squares' and the 'maximum likelihood'.

Key Words: Estimating functions; likelihood; score function.

1 Introduction

In common with most of the historical investigations, it is difficult to trace the origin of the subject of this conference: 'Estimating Functions'. However, in the last two centuries clearly there are three important precursors of the modern theory of estimating functions (EF): In the year 1805, Legendre introduced the least squares (LS) method. At the turn of the last century Pearson proposed the method of moments and in 1925 Fisher put forward the maximum likelihood (ML) equations. Of these three, the method of moments faded out in time because of its lack of any sound theoretical justification. However the other two methods namely the LS and the ML even at present play an important role in the statistical methodology. These two methods would also concern us in the following. The LS method was justified by what today is called the Gauss-Markoff (GM) theorem: The estimates obtained from LS equations are 'optimal' in the sense that they have minimum variance in the class of linear unbiased estimates. This was a *finite sample* justification. At about the same time Laplace provided a different 'asymptotic justification' for the method. Fisher justified the ML estimation, for it produced estimates which are *asymptotically* unbiased with smallest variance. This left open the question, is there a finite sample justification for the ML estimation corresponding to the GM theorem justification for the LS estimation?

The modern EF theory provided such a justification. According to the 'optimality criterion' of the EF theory, the score function (SF) is 'optimal'.