STATISTICAL INFERENCE IN FACTOR ANALYSIS

T. W. ANDERSON AND HERMAN RUBIN

COLUMBIA UNIVERSITY AND STANFORD UNIVERSITY

1. Introduction

In this paper we discuss some methods of factor analysis. The entire discussion is centered around one general probability model. We consider some mathematical problems of the model, such as whether certain kinds of observed data determine the model uniquely. We treat the statistical problems of estimation and tests of certain hypotheses. For these purposes the asymptotic distribution theory of some statistics is treated.

The primary aim of this paper is to give a unified exposition of this part of factor analysis from the viewpoint of the mathematical statistician. The literature on factor analysis is scattered; moreover, the many papers and books have been written from many different points of view. By confining ourselves to one model and by emphasizing statistical inferences for this model we hope to present a clear picture to the statistician.

The development given here is expected to point up features of model-building and statistical inference that occur in other areas where statistical theories are being developed. For example, nearly all of the problems met in factor analysis are met in latent structure analysis.

There are also some new results given in this paper. The proofs of these are mainly given in a technical Part II of the paper.

In confining ourselves to the mathematical and statistical aspects of one model, we are leaving out of consideration many important and interesting topics. We shall not consider how useful this model may be nor in what substantive areas one may expect to find data (and problems) that fit the model. We also do not consider methods based on other models. In doing this, we do not mean to imply that the model considered here is the most useful or important. It seems that this model has some usefulness and importance, it has been studied considerably, and one can give a fairly unified exposition of it.

Extensive discussion of the purposes and applications (as well as other developments) of factor analysis is given in books by psychologists (for example, Holzinger and Harmon [10], Thomson [23], Thurstone [24]). Some general discussion of statistical inference has been given in papers by Bartlett [9] and Kendall [12].

PART I. EXPOSITORY

2. The model

The model we consider is

$$(2.1) \qquad\qquad X = \Lambda f + U + \mu,$$