EMPIRICAL BAYES APPROACH

by /264

KO-CHUN WANG

B. A., National Taiwan University, 1962

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics and Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1970

Approved by:

Major Professor

TABLE OF CONTENTS

Page

1. INTRODUCTION 1
2. EMPIRICAL BAYES APPROACH TO STATISTIC ESTIMATION . . 2
   - 2.1 Bayes Estimator with Apriori Distribution . . 2
   - 2.2 Empirical Distribution Function 4
   - 2.3 Empirical Bayes Approach 6
3. EMPIRICAL BAYES APPROACH TO STATISTICAL DECISION PROBLEMS 11
   - 3.1 Notation 12
   - 3.2 Bayes Decision Function: Known G 12
   - 3.3 Empirical Bayes Decision Procedure 14
   - 3.4 Application--Poisson Case 19
   - 3.5 Remarks 22
4. EMPIRICAL BAYES APPROACH TO TESTING HYPOTHESIS . . . 22
   - 4.1 Testing Simple Hypotheses: Known p 22
   - 4.2 Testing Simple Hypotheses: Unknown p 25
   - 4.3 Testing Composite Hypotheses: Known G 30
   - 4.4 Testing Composite Hypotheses: Unknown G . . . 32
   - 4.5 Application--Normal Case 34
   - 4.6 Remarks 38
5. EXTENSION: SEQUENTIAL HYPOTHESIS TESTS FOR STATIONARY PROCESSES 39
   - 5.1 Definitions and Notation 39
   - 5.2 Bayes Solution: Known p 41
   - 5.3 Empirical Bayes Solution: Unknown p 43
6. CONCLUSION 45

ACKNOWLEDGMENT 48

REFERENCES 49

1. INTRODUCTION

Consider a sequence of observations generated by a statistical experiment which depends on an unknown parameter, and, in addition, consider as given that the parameter is a random variable but with an unknown distribution. It will usually be advantageous to treat the whole sequence of observations as a single entity rather than to treat each component separately. It was with this idea in mind that Robbins (1955) first discussed the use of apriori observations to approximate Bayes procedures, and hence to establish the "asymptotically optimal" statistical solutions. In that paper Robbins described such approximations as "empirical Bayes procedures".

It has been proved that whenever a statistical experiment comes to us with such a sequence of observations, the empirical Bayes approach offers certain advantages over any other approach which either ignores the fact that the unknown parameter is itself a random variable, or assumes a personal or subjective probability distribution of the parameter not subject to change with experience.

Since the time of Robbins' initial work--which Neyman (1962) admired as the first breakthrough in statistical theory during the past decade--on the utilization of previous experience for statistical inference, the theory of the empirical Bayes approach has been so developed that some of the theories and solutions are ready for practical applications.

This report considers some of the discussions by Robbins (1955, 1963, 1964), Johns (1957, 1961), Neyman (1962), Samuel (1963), and Tainiter (1965). The empirical Bayes approach to statistical estimation is described in Section 2, where for simplicity only discrete random variables are considered. In Section 3, the "asymptotically optimal" empirical Bayes rules and their applications for decision problems are considered. In Section 4, the asymptotically optimal empirical Bayes approach to the testing of hypotheses, together with some examples, is considered. Section 5 is an extension which deals with some recent works by Tainiter. The last section is for concluding remarks.

2. EMPIRICAL BAYES APPROACH TO STATISTIC ESTIMATION

2.1 Bayes Estimator with Apriori Distribution

Let X be a random variable with a known probability density function depending on an unknown real parameter Λ , namely,

$$p(x|\lambda) = Pr[X = x|\Lambda = \lambda].$$

Suppose Λ is itself a random variable with apriori distribution function

$$G(\lambda) = Pr[\Lambda \leq \lambda],$$

then the marginal p.d.f. of X is given by

$$P_G(x) = \Pr[X = x] = \int_\Lambda p(x|\lambda)\,dG(\lambda). \tag{2.1}$$

If mean square error is adopted as the measurement of accuracy, then the expected squared deviation of any estimator of Λ of the form $\Psi(x)$ is

$$\begin{aligned} E[\Psi(x) - \Lambda]^2 &= E\{E[(\Psi(x) - \Lambda)^2|\Lambda = \lambda]\} \\ &= E\{\textstyle\sum_x (\Psi(x) - \lambda)^2 p(x|\lambda)\} \\ &= \int_\Lambda \textstyle\sum_x p(x|\lambda)[\Psi(x) - \lambda]^2 dG(\lambda) \\ &= \textstyle\sum_x \int_\Lambda p(x|\lambda)[\Psi(x) - \lambda]^2 dG(\lambda). \end{aligned} \tag{2.2}$$

This quantity attains its minimum if, for each x) the estimator given by $\Psi_0 = \Psi_0(x)$ is such that

$$I(x) = \int_\Lambda p(x|\lambda)[\Psi_0 - \lambda]^2 dG(\lambda) = \text{minimum}.$$

But for any fixed x the quantity

$$\begin{aligned} I(x) &= \Psi_0^2 \int_\Lambda p\,dG - 2\Psi_0 \int_\Lambda p\lambda\,dG + \int_\Lambda p\lambda^2 dG \\ &= \left(\Psi_0 - \frac{\int_\Lambda p\lambda\,dG}{\int_\Lambda p\,dG}\right)^2 \int_\Lambda p\,dG + \left[\int p\lambda^2 dG - \frac{(\int_\Lambda p\lambda\,dG)^2}{\int_\Lambda p\,dG}\right] \end{aligned} \tag{2.3}$$

is a minimum w.r.t. Ψ_0 when

$$\Psi_0 = \frac{\int_\Lambda p\lambda\,dG}{\int_\Lambda p\,dG}.$$

The minimum value of I(x) is