

Sequential Inference of the Number of Multinomial Population

Hokwon Cho and Z. Govindarajulu

*Department of Mathematical Sciences
University of Nevada, Las Vegas, Las Vegas and
Department of Statistics
University of Kentucky, Lexington*

ABSTRACT

A sequential risk-efficient estimator with squared error loss is proposed for estimating the number of equally probable cells in a given multinomial distribution. It is assumed that the cost per observation is constant. Large-sample properties of the sequential estimator are investigated and a simulation study is carried out in order to examine its finite sample behavior.

CONTENTS

1. Introduction

1.1 Formulation

2. Asymptotic Behavior of the Procedure

2.1 First Order Asymptotic Results

3. Performance of the Procedure

3.1 Risk Efficiency

3.2 Regret

4. Simulation Study

4.1 Monte Carlo Experimentation

4.2 Approximation and Examples

BIBLIOGRAPHY

Boender, C. and Rinnooy Kan, A. (1983). A Bayesian analysis of the number of cells of a multinomial distribution. *The Statistician*, **32**, 240-248.

Boender, C. and Zielinski, R. (1985). A sequential Bayesian approach to estimating the dimension of a multinomial distribution. *Sequential Methods in Statistics*, Banach Center Publications, **16**, 37-41, PWN-Polish Scientific Publishers, Warsaw.

Bunge, J. and Fitzpatrick, M. (1993). Estimating the number of species: a review. *Journal of the American Statistical Association*, **88**, 364-373.

Chow, Y. S. and Martinsek, A. T. (1982). Bounded regret of a sequential procedure for estimation of the mean. *The Annals of Statistics*, **10**, 909-914.

Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean. *The Annals of Mathematical Statistics*, **36**, 457-462.

Chow, Y. S. and Yu, K. F. (1981). The performance of a sequential procedure for the estimation of the mean. *The Annals of Statistics*, **9**, 184-189.

Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd Ed. John Wiley, New York.

Finkelstein, M., Tucker, H. and Veeh, J. (1998). Confidence intervals for the number of unseen types. *Statistics and Probability Letters*, **37**, 423-430.

Goodman, L. (1949). On the estimation of the number of classes in a population. *The Annals of Mathematical Statistics*, **20**, 572-579.

Govindarajulu, Z. (1999). *The Elements of Sampling Theory and Methods*. Prentice-Hall, Inc. New Jersey.

Jordan, C. (1950). *Calculus of Finite Differences*, 2nd Ed., Chelsea Publishing Co., New York.

Lai, T. L. (1996). On uniform integrability and asymptotically risk-efficient sequential estimation. *Sequential Analysis*. **15**, 237-251.