

MATHEMATIK-ARBEITSPAPIERE

A: MATHEMATISCHE FORSCHUNGSPAPIERE

A FORMAL DERIVATION OF THE
CONDITIONAL LIKELIHOOD FOR
MATCHED CASE-CONTROL STUDIES

GERHARD OSIUS

No. 55

JANUARY 2001

FACHBEREICH MATHEMATIK UND INFORMATIK
UNIVERSITÄT BREMEN

Bibliothekstraße
D-28359 Bremen
Germany

A Formal Derivation Of The Conditional Likelihood For Matched Case-Control Studies ¹

Gerhard Osius ²

- 1 Introduction and Notation
- 2 Odds Ratios
- 3 Odds Ratio Models
- 4 Conditional Sampling Distribution and Likelihood
- 5 Case-Control Studies With 1:M Matching
- A Multivariate Noncentral Hypergeometric Distributions
- B Rank and Order Statistics

Abstract

Likelihood analysis for logistic regression models in matched case-control studies is based on a *conditional* likelihood. We provide a formal derivation for this conditional likelihood in a general setting by conditioning the sampling distribution of the risk factors (in each matching group) on its observed *order statistic* with respect to lexicographical ordering. The required joint distribution of the rank and order statistics for vectors is derived in a more general context. The conditional likelihood depends only upon the odds ratios of interest, i.e. for different values of the risk factor and same value of the matching variable. Hence the conditional analysis may be used in parametric (e.g., logistic regression) as well as in nonparametric models for the odds ratios. In the presence of only two disease categories (case and control) and 1: M matching the conditional likelihood is a product of multinomial probabilities (reducing to binomial probabilities for $M=1$). In this case the conditional analysis can be done within the framework of generalized linear models. In the presence of more than two disease categories the conditional distribution of the covariate in each matching group is a *multivariate noncentral hypergeometric distribution* (whose basic properties are also given) and the conditional analysis requires a higher computational effort.

1 This work was supported by the German National Science Foundation (DFG), grant OS-144/1 and is a revised version of the German report (Oct. 1997) for the DFG.

2 Institut für Statistik, Fachbereich 3, Universität Bremen, Postfach 330440, 28334 Bremen, Germany. E-mail: osius@math.uni-bremen.de